UNIVERSITY OF CALIFORNIA
SANTA CRUZ

Beyond Rule of 5 Drug Discovery:
Investigating the Drug-like Properties of
Cyclic Peptide Natural Products and PROTACs

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CHEMISTRY AND BIOCHEMISTRY

by

Victoria Gail Klein

December 2020

The Dissertation of Victoria Gail Klein
is approved:

_____
Professor R. Scott Lokey, Chair

_____
Professor Seth Rubin

_____
Professor Carrie Partch

_____
Quentin Williams
Acting Vice Provost and Dean of Graduate Studies

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Identifying the cellular target of a cordyheptapeptide A and synthetic derivatives

Chapter 2: Understanding and Improving the Membrane Permeability of VH032-Based PROTACs

Chapter 3: Amide-to-Ester Substitutions improve PROTAC permeability

LIST OF TABLES

Chapter 1: Identifying the cellular target of a cordyheptapeptide A and synthetic derivatives

Chapter 2: Understanding and Improving the Membrane Permeability of VH032-Based PROTACs

Chapter 3: Amide-to-Ester Substitutions improve PROTAC permeability

ABSTRACT

Beyond Rule of 5 Drug Discovery:

Investigating the Drug-like Properties of Cyclic Peptide Natural Products and PROTACs

Victoria Gail Klein

Drug discovery efforts have favored small molecules that can be described by Lipinski's "Rule of 5" (Ro5). Compounds that fit into the Ro5 are under 500 in molecular weight (MW), have an octanol-water partition coefficient of less than five, have fewer than five hydrogen bond donors (HBDs), and fewer than ten hydrogen bond acceptors (HBAs). While some have published variations on the Ro5 since its description by Lipinski in 1997 and though the Ro5 was simply a summary of orally active drugs at the time, the Ro5 has often been used as design parameters in drug discovery. Drugs that conform to the Ro5 are typically likely to be orally bioavailable and have favorable ADME properties (absorption, distribution, metabolism, and excretion). However, confining drug design to fit into this set of criteria restricts drug targets to proteins with well-defined active sites and leaves a large portion of the proteome "undruggable." To expand the druggable proteome, there has been a recent surge in antibody-based drugs. Antibodies are larger and able to target proteins without a well-defined active site. However, these expensive molecules are not orally bioavailable, and they are so large that they are only capable of targeting extracellular receptors. In this drug paradigm, intracellular protein disease targets either lacking

small molecule binding sites or participating in protein-protein interaction are left without treatment.

Further and critical expansion of the druggable proteome can be achieved with both natural product-inspired cyclic peptides and degradation-triggering proteolysis targeting chimeras (PROTACs). The intermediate size of these types of compounds lends the ability to affect intracellular disease targets while maintaining cellular permeability and oral bioavailability. Both molecule types can also bind to their protein targets at sites other than a deep pocket, overcoming this limitation common to typical small molecule drugs. The overall goal of my dissertation is to investigate the biological activity and physicochemical properties of these two "beyond Rule of 5" (bRo5) therapeutics capable of targeting previously "undruggable" protein targets.

Chapter one investigates the bioactivity and permeability of the cyclic peptide natural product cordyheptapeptide A and several synthetic derivatives. These natural product-inspired derivatives reveal that it is crucial to consider both bioactivity and permeability when optimizing a natural product. Additionally, using a combination of high-content screening and biochemical assay, I identify the intracellular target of cordyheptapeptide A as the eukaryotic translation elongation factor eEF1α. This is a critical disease target that is upregulated in many cancers and has yet to be successfully inhibited by a small molecule that adheres to the Ro5. Importantly, this work highlights that cyclic peptides dominated by aromatic and lipophilic sidechains, like cordyheptapeptide A, have the capacity to inhibit intracellular drug targets with sub-micromolar potencies.

In chapters two and three, I explore the physicochemical properties of PROTACs. These heterobifunctional molecules catalytically trigger the degradation of a protein target. PROTACs are typically more specific, more potent, and produce fewer off-target effects than typical Ro5 small molecule drugs. Our knowledge of PROTAC bioactivity is rapidly growing, but there is an urgent need to better understand these compounds' physicochemical properties. Chapter two uses two label-free mass spectrometry related assays to determine the passive permeability of several previously published PROTACs and examine how different structural features contribute to this permeability. I also demonstrate that amide-to-ester substitutions increase PROTAC permeability, PROTACs can form intramolecular hydrogen bonds that could be important for their permeability, and that PROTAC bioactivity is affected by permeability. However, strong target binding and ternary complex formation can overcome permeability deficits.

Chapter three expands on the findings from chapter two and investigates permeability improvements gained from an amide to ester substitution in compounds with a wide range of calculated lipophilicities and several different linkers. I describe how esters increase the permeability of PROTACs over a broad range of linkers and lipophilicities, but not for compounds that already have high lipophilicities (>4.5) that are moving into the range of lacking aqueous solubility. We also discovered that while amide-containing PROTACs are the most stable in plasma, ester-containing compounds in which the ester is near a larger drug-like side chain see only minimal hydrolysis in plasma. Combined, chapters two and three offer design guidelines for

developing permeable PROTACs, insights into how their structural features affect their permeability, and strategies to improve these compounds' permeabilities.

Overall, this dissertation emphasizes the need to 1) consider the relationship between bioactivity and permeability when optimizing new drug compounds and 2) expand our current drug discovery efforts to include bRo5 compounds to treat previously "undruggable" diseases.

DEDICATION

For my parents

ACKNOWLEDGEMENTS

Completing this Ph.D. would not have been possible without the support of so many around me. First, I would like to thank my adviser, Scott Lokey, for his continued support and guidance. Thank you for being a great source of creative ideas and helping me grow as a scientist. Thank you to my committee members, Carrie Partch and Seth Rubin, for the unwavering mentorship, advice, and encouragement through my time at UC Santa Cruz. I was fortunate to start this Ph.D. program with two amazing graduate students in my cohort, Lena Meyer and Terren Chang. I do not know how I would have made it through the first few years of graduate school without their friendship and collaboration. I also want to thank Karen Meece and Katie Cramton for doing all they do to keep the department running smoothly.

I have been lucky to have some amazing lab mates in my 5 years in the Lokey Lab and in the UCSC Screening Center. I would especially like to thank Josh Schwochert, Cameron Pye, Walter Bray, Chad Townsend, Matt Naylor, Andrew Ly, Alex Turmon, Quinn Edmonson. I would like to thank all of them for their friendship, encouragement, advice, and collaboration even after they left the lab. My research and quality of life been drastically improved thanks to their presence. I would like to say a special thank you to Jessie Ochoa for being a role model and source of inspiration. I felt so supported by you and cannot thank you enough for always being in my corner. I would also like to thank Professors Jack Taunton from UCSF and Alessio Ciulli from the University of Dundee and their research groups for their collaboration.

I would like to thank my parents for their unconditional love, inspiration, and encouragement though my life. Thank you always helping me to achieve my goals, supporting my education, and showing me the value of working hard. I am lucky to have been raised by such incredible parents. I would also like to thank my wonderful sisters, Cat, Liz, and Jenn, and their families, Jon, David, Liam, and Hayden, for a making my life so fun. Finally, I want to thank my amazing partner, Dr. Galen, for all that he has done for me throughout my Ph.D. Thank you for filling our life with love and joy especially through all the ups and downs that come with being a graduate student. I could not have done it without you.

The text of this dissertation includes reprint[s] of the following previously published material:

Klein, V. G., Townsend, C. E., Testa, A., Zengerle, M., Maniaci, C., Hughes, S. J., Chan, K. H., Ciulli, A., & Lokey, R. S. (2020). Understanding and Improving the Membrane Permeability of VH032-Based PROTACs. *ACS Medicinal Chemistry Letters*, 11(9), 1732–1738. https://doi.org/10.1021/acsmedchemlett.0c00265

The co-author listed in this publication directed and supervised the research which forms the basis for the dissertation.

Chapter One

Identifying the cellular target of a cordyheptapeptide A and synthetic derivatives

This chapter contains text and figures from the following manuscript: Victoria G Klein, Walter M Bray, Haoyuan Wang, Quinn Edmondson, Joshua Schwochert, Satoshi Ono, Matthew R. Naylor, Alexandra C Turmon, Justin H Faris, Okimasa Okada, Jack Taunton, R. Scott Lokey. *Identifying the cellular target of a cordyheptapeptide A and synthetic derivatives.* (Manuscript in preparation)

Abstract

Originating from the widely prized *Cordyceps* fungal genus, cordyheptapeptide A is a lipophilic cyclic peptide with significant therapeutic potential. While its target was previously unidentified, this natural product has been shown to be cytotoxic in multiple cancer cell lines. To better understand its bioactivity and physicochemical properties, we developed several synthetic cordyheptapeptide derivatives. Based on these derivatives, we identified a number of strategies to improve cyclic peptide permeability. Furthermore, we demonstrate that removing a backbone N-methyl can improve bioactivity without drastically affecting membrane permeability when a new intramolecular hydrogen bond is formed. Additionally, we identified the mechanism of action and major target of cordyheptapeptide A. This cyclic peptide diminishes protein and DNA synthesis through inhibition of the cancer-relevant eukaryotic translation elongation factor eEF1A. This work offers a strategy to study and improve cyclic peptide natural products while highlighting the ability of these lipophilic compounds to effectively inhibit intracellular disease targets.

1.1 Introduction

Macrocycles have long been pursued for their rich structural diversity and continue to provide a bountiful source of bioactive scaffolds.[1-3] Significant work has gone into understanding and improving both the biochemical and physiochemical properties of macrocycles with a recent emphasis on their uses as therapeutics.[4,5] Many cyclic peptide natural products are highly active in mammalian cells, prompting studies by our group and others into the factors that govern cell permeability in these large, non-"druglike" molecules.[6-9] In the course of our investigations into the relationship between molecular size and cell permeability,[7,8] we synthesized and investigated the properties of cordyheptapeptide A (**1**) (Figure 1-1). Originally isolated from *Cordyceps*, a fungal genus widely valued for its pharmaceutical potential,[10] the cordyheptapeptide family, including cordyheptapeptide B (**2**) and C, are reported to show toxicity toward bacteria, fungi, as well as a variety of cancer cell lines.[11-13] While a crystal structure[11] of **1** and its solution-phase total synthesis have been reported,[13] its biological target(s) and mechanism of action remain unknown.

During an image-based screening effort on a variety of natural and synthetic cyclic peptides, we identified an interesting and potent phenotypic activity of **1** in HeLa cells. This observation prompted us to develop an efficient solid-phase synthesis approach to **1** and several derivatives, enabling the investigation of structure-activity and structure-permeability relationships in the cordyheptapeptide family. Using NMR coupled with molecular dynamics simulations, we demonstrate that small modifications to the backbone of a cyclic peptide can drastically affect both the

permeability and bioactivity of a cyclic peptide. Additionally, by combining two high-content screening assays with targeted pull-down assays, we identified its cellular target as the translation elongation factor eEF1A.

1.2 Results

1.2.1 Improved total synthesis

The solution-phase total synthesis of cordyheptapeptide A (**1**, Figure 1-1 A) have been reported previously.[13] We developed a total synthesis for **1** using solid-phase peptide synthesis (SPPS), allowing for rapid, automated linear synthesis using Fmoc chemistry, N-methylated amino acid monomers, and in-sequence, on-resin N-methylation[14] (Supplementary Scheme 1). This higher-throughput synthesis was essential for generating a library of compounds to investigate structure-activity and structure-permeability relationships.

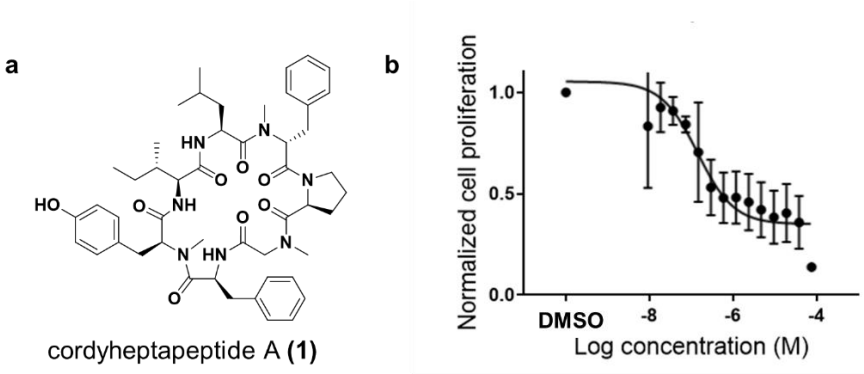


Figure 1-1: Cordyheptapeptide structure and activity *(A)* chemical structure of cordyheptapeptides A (**1**). *(B)* Effect of **1** on cell proliferation in HCT 116 cells after 72 h. Error bars are one SD. N =3

1.2.2 Bioactivity optimization

Since its isolation in 2006, **1** and its derivatives have shown cytotoxicity in a variety of cancer cell lines, with IC_{50} values in the mid- to low-μM range.[11] Using a 72-h resazurin-based cell proliferation assay, we confirmed literature results for **1** in HCT 116 cells (IC_{50} = 0.1 μM, Figure 1-1 B, Supplementary Figure 1-S1). The reported IC_{50} of cordyheptapeptide B (**2**, NMeTyr-2-NMePhe) also falls in the low-μM range in many cells lines and is often less potent than **1** when compared by cell line.[11-13] Our cell proliferation results in HCT 116 cells also showed **2** to be much less active than **1** (IC_{50} = 35 μM, Supplementary Figure 1-S1).

Using the modularity of our new cordyheptapeptide synthesis, we generated a variety of structural variants aimed at probing structure-permeability and structure-activity relationships (SAR). Though **1** and **2** differ by only one hydroxyl at the Tyr2 (**1**)/Phe2 (**2**) position, these two compounds have a 350-fold difference in antiproliferative activity in HCT 116 cells (Table 1-1). To further investigate this SAR and potentially improve the bioactivity of **1**, we generated a series of derivatives by varying each of its seven residues. Many noteworthy observations emerged from this SAR study (Table 1-1).

First, we performed an alanine scan[15] of the natural product (**1**), which revealed that all side chains are critical to its antiproliferative activity (**3** – **9**, IC_{50} > 75 μM, Table 1-1, Supplementary Figure 1-S1), suggesting that that each residue is essential to either target engagement, secondary structure, or cellular permeability. Next, since there are a wide variety of commercially available, non-proteinogenic L-Phe and D-

Phe derivatives including fluorine, chlorine, and heterocyclic substitutions, we synthesized a small SAR series based on variations at Phe3 and D-Phe6. All compounds from this initial SAR series (**19** – **26**) were less potent than **1**, ranging from moderately active (IC_{50} = 1 μM) to inactive (IC_{50} > 75 μM) (Table 1-1, Supplementary Figure 1-S1).

Table 1-1: Comparison of potency and properties between cordyheptapeptides A and B and synthetic derivatives. The compounds are compared by ALogP, IC_{50}, PAMPA permeability, and aqueous solubility. The features are colored on a log scale ranging from white (weakest potency, poorest permeability/solubility) to dark (highest potency, best permeability/solubility). Dark colors indicate the lowest values in the IC_{50} column and indicate the highest values in the permeability and solubility columns. The IC_{50} of a compound was determined by 72-h cell proliferation in HCT 116 cells. Notation "*" = data reported in Naylor *et al.* (2018)[7] "n.d." = none detected, and "--" indicated no measurement was taken. (See Supplementary Table 1-S1 for IC_{50} 95% confidence intervals)

	Compound	1	2	3	4	5	6	7	ALogP	IC50 (μM)	P_e x 10^{-6} (cm/s)	Solubility (μM)
	1	Ile	MeTyr	Phe	Sar	Pro	MeDPhe	Leu	4.2	0.1	1.2*	4
	2	Ile	MePhe	Phe	Sar	Pro	MeDPhe	Leu	4.5	35	2.3*	2
Alanine Scan	VK03 (**3**)	Ala	MeTyr	Phe	Sar	Pro	MeDPhe	Leu	2.9	> 75	0.2*	10
	VK04 (**4**)	Ile	MeAla	Phe	Sar	Pro	MeDPhe	Leu	2.9	> 75	5.9*	16
	VK05 (**5**)	Ile	MeTyr	Ala	Sar	Pro	MeDPhe	Leu	2.7	> 75	1.2*	17
	VK06 (**6**)	Ile	MeTyr	Phe	Sar	Ala	MeDPhe	Leu	3.9	> 75	2.5*	--
	VK07 (**7**)	Ile	MeTyr	Phe	Sar	MeAla	MeDPhe	Leu	4.1	> 75	7.8*	37
	VK08 (**8**)	Ile	MeTyr	Phe	Sar	Pro	MeDAla	Leu	2.7	> 75	1.3*	n.d.
	VK09 (**9**)	Ile	MeTyr	Phe	Sar	Pro	MeDPhe	Ala	3	> 75	2.2	9
N-methyl Scan	VK10 (**10**)	Ile	Tyr	Phe	Sar	Pro	MeDPhe	Leu	4	8	1.6	25
	VK11 (**11**)	Ile	MeTyr	Phe	Gly	Pro	MeDPhe	Leu	4	0.5	1.6	7
	VK12 (**12**)	Ile	MeTyr	Phe	Sar	Pro	DPhe	Leu	4	28	0.7	37
	VK13 (**13**)	Ile	Phe	Phe	Sar	Pro	MeDPhe	Leu	4.3	> 75	1.7	14
	VK14 (**14**)	Ile	MePhe	Phe	Gly	Pro	MeDPhe	Leu	4.3	17	0.9	2
	VK15 (**15**)	Ile	MePhe	Phe	Sar	Pro	DPhe	Leu	4.3	> 75	3.9	23
	VK16 (**16**)	MeAla	MeTyr	Phe	Sar	Pro	MeDPhe	Leu	4.4	37	6.7	56
	VK17 (**17**)	Ile	MeTyr	MePhe	Sar	Pro	MeDPhe	Leu	4.4	> 75	2.4	230
Halogen and Heterocyclic Substitutions	VK19 (**19**)	Ile	MeTyr	Phe(2-F)	Sar	Pro	MeDPhe	Leu	4.4	11	2.5	--
	VK20 (**20**)	Ile	MeTyr	Phe(3-Cl)	Sar	Pro	MeDPhe	Leu	4.9	34	1	3
	VK21 (**21**)	Ile	MeTyr	Phe(4-Cl)	Sar	Pro	MeDPhe	Leu	4.9	3.6	1.5	0.3
	VK22 (**22**)	Ile	MeTyr	Phe(4-F)	Sar	Pro	MeDPhe	Leu	4.4	1.8	1.9	8
	VK23 (**23**)	Ile	MeTyr	Ala(β-2-fur)	Sar	Pro	MeDPhe	Leu	3.3	> 75	3.4	58
	VK24 (**24**)	Ile	MeTyr	Ala(β-3-pyr)	Sar	Pro	MeDPhe	Leu	3.1	> 75	n.d.	516
	VK25 (**25**)	Ile	MeTyr	Phe	Sar	Pro	MeDPhe(4-Cl)	Leu	4.9	28	1	0.5
	VK26 (**26**)	Ile	MeTyr	Phe	Sar	Pro	MeDPhe(4-F)	Leu	4.4	3.4	3.4	4
	VK27 (**27**)	Ile	MeTyr	Phe(2-F)	Gly	Pro	MeDPhe	Leu	4.2	0.3	0.9	156
	VK28 (**28**)	Ile	MeTyr	Phe(3-Cl)	Gly	Pro	MeDPhe	Leu	4.7	0.2	1.3	0.8
	VK29 (**29**)	Ile	MeTyr	Phe(4-Cl)	Gly	Pro	MeDPhe	Leu	4.7	0.5	1.5	0.9
	VK30 (**30**)	Ile	MeTyr	Phe(4-F)	Gly	Pro	MeDPhe	Leu	4.2	0.1	0.9	2
	VK31 (**31**)	Ile	MeTyr	Ala(β-2-fur)	Gly	Pro	MeDPhe	Leu	3.1	> 75	1.6	20
	VK32 (**32**)	Ile	MeTyr	Ala(β-3-pyr)	Gly	Pro	MeDPhe	Leu	2.9	> 75	n.d.	5
	VK33 (**33**)	Ile	MeTyr	Phe	Gly	Pro	MeDPhe(4-Cl)	Leu	4.7	19	0.8	0.8
	VK34 (**34**)	Ile	MeTyr	Phe	Gly	Pro	MeDPhe(4-F)	Leu	4.2	0.3	1.3	10
	VK35 (**35**)	Ile	MeTyr	Phe(4-F)	Gly	Pro	MeDPhe(4-F)	Leu	4.4	28	0.7	19
Other Substitutions	VK36 (**36**)	Ile	MeTyr(OAc)	Phe	Sar	Pro	MeDPhe	Leu	4.2	2.9	0.4	5
	VK37 (**37**)	Ile	MeTyr(OMe)	Phe	Sar	Pro	MeDPhe	Leu	4.4	> 75	n.d.	14
	VK38 (**38**)	Ile	MeTyr(OMe)	Phe	Gly	Pro	MeDPhe	Leu	4.2	5.6	0.9	6
	VK39 (**39**)	Ile	MePhe(4-F)	Phe	Sar	Pro	MeDPhe	Leu	4.7	--	n.d.	11
	VK40 (**40**)	Ile	MePhe	Phe(4-F)	Gly	Pro	MeDPhe	Leu	4.5	1.5	0.9	23
	VK41 (**41**)	Ile	MeTyr	Phe(4-F)	β-Ala	Pro	MeDPhe	Leu	4.3	28	2.7	3
	VK42 (**42**)	Nva	MeTyr	Phe	Sar	Pro	MeDPhe	Leu	4.4	22	1.9	18

Since these initial SAR studies yielded no compounds with improved potency over the parent natural product, and because **1** showed only modest permeability (P_e = 1.2 x 10^{-6} cm/s) and low aqueous solubility (3.9 μM) (Table 1-1), we considered whether potency could be improved by increasing the compound's membrane permeability, aqueous solubility, or both. As backbone N-methylation causes a complex interplay of both local and global effects on permeability and bioactivity,[16-19] we synthesized and tested a variety of backbone N-methyl variants of **1**. Indeed, changes in backbone N-methylation had varying effects on both the bioactivity and the physiochemical properties of **1**. N-methylating a non-N-methylated position at Phe3 in **1,** while it produced a slight increase in permeability, dramatically decreased bioactivity (**17**, IC_{50} > 75 μM). Although the solid-state conformation does not necessarily represent its target-bound state, the crystal structure of **1** shows that the NH groups at positions 3 and 7 are both involved in intramolecular hydrogen bonds, demonstrating the potential importance of these NH groups in determining the overall conformation of the molecule.[7,11]

Removal of any individual N-methyl from **1** led to a decrease in potency, although the severity of this loss depended on the position. The Sar4 position (**11**) was most tolerant of N-Me removal (IC_{50} = 0.5 μM), followed by Tyr2 (**10**, IC_{50} = 8 μM), and then D-Phe6 (**12**, IC_{50} = 35 μM). However, whereas removal of the N-Me at MeDPhe6 also diminished permeability significantly, N-Me removal at MeTyr2 did not affect permeability. Interestingly, while removal of the N-Me at Sar4 (**11**) also had no

impact on permeability, this substitution saw only a 5-fold loss in potency (IC_{50} = 0.5 μM). We hypothesized that introducing amino acids with more conformational freedom, for example, substituting Pro5 for a MeAla, might allow for an alternative conformation with higher permeability. Indeed, the substitution of Pro5 for MeAla (**7**), which does not substantially change the calculated lipophilicity of the compound, resulted in a significant increase in permeability from 1.2 x 10^{-6} cm/s to 7.8 x 10^{-6} cm/s. This suggests that the Pro-containing natural product has a more restricted conformational landscape, thereby limiting access to nonpolar conformations in the membrane's low dielectric. Similarly, substituting Sar4 for a the more flexible β-alanine (**41**) slightly improved permeability (2.7 x 10^{-6} cm/s) over **1**. Though both the Pro5-to-MeAla and Sar4-to-β-alanine were more permeable, these modifications also eliminated antiproliferative activity in HCT 116 cells.

Since the Sar4-to-Gly4 substitution had no deleterious effect on permeability and only a modest decrease in potency (**11**, IC_{50} = 0.5 μM), we reasoned that the resulting increase in backbone flexibility could yield an SAR series with more opportunity for side chain optimization than the more rigid, natural backbone afforded. Thus, compounds **27** – **35** were synthesized, which had the same side chain modifications at positions 3 and 6 previously described for **19** – **26** (Table 1-1). Notably, except for the compounds that remained inactive in both sets, the Gly-containing compounds were consistently more active than their Sar matched pairs, with a 2- to 170-fold improvement in IC_{50} (Supplementary Figure 1-S2, Supplementary Table 1-S4). Considering that the Gly compounds had similar solubilities and similar

or slightly lower permeabilities compared to their Sar-containing matched pairs, the improved bioactivity may be derived from increased access to conformations that interact favorably with the target. Furthermore, **30** (Phe3-to-Phe[4-F] and Sar4-to-Gly) was equipotent to **1** and had comparable permeability and solubility. Applying the same modifications in **30** to cordyheptapeptide B (**2**) yielded a 23-fold increase potency (IC_{50} = 1.5 μM) and a 10-fold increase in solubility (**40**). Given that amide cis-trans isomerism in Pro and N-Me residues can dramatically change a cyclic peptide's overall conformation, the preservation (and in some cases, enhancement) of biological potency in the Gly4 series compared to their Sar4 congeners suggests that in the bioactive, target-bound state of **1**, the omega torsion at this position is likely *trans*.

1.2.3 Molecular dynamics simulations

To explore the hypothesis that conformational flexibility and/or cis-tran isomerism at Sar4 impacts properties as well as the observed SAR in this series, we studied the conformations of **1** and **11** in low- and high-dielectric media using multicanonical molecular dynamics (MCMD) simulations. MCMD[20-22] randomly samples a potential energy space, thereby overcoming large energy barriers such as cis-trans isomerization of Pro and N-methylated amino acids. After a production MCMD run, canonical ensembles between low- and high-temperature states are obtained by reweighting or resampling methods. In this study, canonical ensembles at T = 300 K were obtained for each solvent and compound. The AMBER ff03 force field was used

for amino acids,[23] and ForceField NCAA was used for N-methylated amino acids.[24] Solvents were treated explicitly.

The MD-derived free energy landscapes of **1** and **11** showed that **11** populates a wider variety of conformational states and was more flexible compared to the parent natural product, **1** (Figure 1-2 A-B, Supplementary Figure 1-S2 A). For example, in membrane-like chloroform, **1** adopted one major conformer with two intramolecular hydrogen bonds (IMHBs) for 47.4% of the MD simulation. On the other hand, the more flexible **11** had two major conformers each with three IMHBs for a combined total of 68.1% (Supplementary Figure 1-S2 B). Importantly, the newly exposed NH at position 4 partipated in this new third IMHB in **11**. This suggests that **11** can access multiple conformation in which the additional HBD is not solvent exposed, providing a potenial explanation as to why removal of this N-methyl did not cause a significant loss in permeability. Additionally, based on an average from 5,000 conformers, **1** and **11** have broadly compareable polar surface areas (PSA) in chloroform (159 ± 11 $Å^2$ and 164 ± 18 $Å^2$, respectively), supporting the conclusion that the Gly -NH is shielded from solvent. NMR temperature coefficient experiments in chloroform also showed significant changes in the chemical shift of this newly exposed amide NH suggesting a conformational change and increased access to multiple conformations (Supplementary Figures 1-S4 and 1-S5).

While **11** is more flexible in a hydrophobic environment and has access to greater conformational space according, the major MCMD-derived conformers of **11** and **1** are similar in water, with an average backbone RMSD of 0.5 Å (Figure 1-2 D).

Though these conformations do not necessarily reflect the target-bound state for either compound, the similarities of the major conformers for the two compounds suggest that the more flexible **11** could still interact with the same target binding site as **1**. Increased scaffold flexibility within the Gly series likely accounts for the generally increased potency compared to Sar matched pairs (Table 1-1, Supplementary Table 1-S4) as the more flexible Gly scaffolds may be capable of accessing conformations with improved binding interactions.

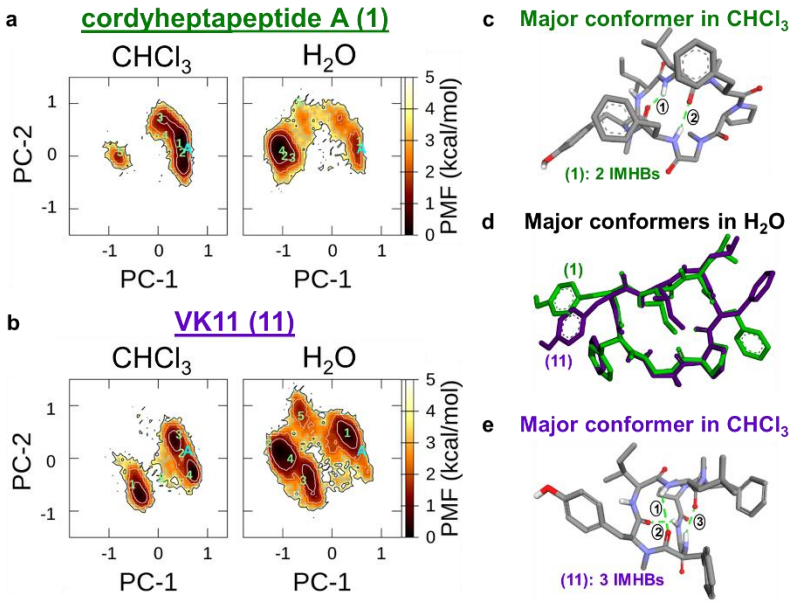


Figure 1-2: Molecular dynamics results comparing cordyheptapeptide A (1) and VK-11 *(A-B)* Principle component (PC) analysis of free energy landscapes in chloroform (CHCl3) and water (H_2O) for *(A)* cordyheptapeptide A (**1**) and *(B)* VK11 (**11**) that has a Sar4 to Gly substitution. "A" in cyan denotes the mirror-imaged X-ray structure (CCDC 287376). Green numbers within each plot indicate representative structures in each solvent. *(C)* Major MD simulated conformer of **1** in chloroform. *(D)* Overlay of the major conformer in water for **1** (in green) and VK-11 (in purple). *(E)* Major MD simulated conformer of VK-11 in chloroform.

To corroborate our MD results and the compared number of HBDs between **1** and **11**, we used a lipophilic permeability efficiency (LPE) metric. Developed by our group, LPE uses an experiemntal decadiene-water partition coefficient ($LogD_{dec/w}$) to assess passive permeability normalized to calculated ALogP. This is a conformation-independent lipophilicity metric that roughly reflects the maximum polarity achievable in aqueous solution.[7] Each exposed HBD is expected to reduce LPE by 1.5 – 2 units. For example, compounds **1** and **2** differ in only one hydroxyl HBD (Tyr2 vs. Phe2, respectively), causing an LPE drop of 1.4 in **1** (LPE **1 =** 2.3, LPE **2** = 3.7). A similar LPE difference of 1.9 value is observed between **11** and **14** for the same exposed Tyr (LPE **11 =** 2.0, LPE **14** = 3.9). However, the small LPE differences between **1** and **11** (LPE = 2.3 and 2.0, respectively) and between **2** and **14** (LPE = 3.7 and 3.9, respectively) reinforce the MD prediction that the new NH at position four is not exposed but is hidden within an IMHB, masking its polarity while passively crossing the lipophilic cell membrane (Supplementary Figure 1-S3, Supplementary Table 1-S3).

1.2.4 NCI COMPARE and Cytological Profiling: Mechanism of action determination

The pattern of cytotoxicity for the cordyheptapeptides has been well established, but its target and mechanism of action (MOA) have remained obscure. To investigate the MOA of **1**, we obtained an independent phenotypic profile of **1** using the NCI60 human tumor cell line assay from the DTP at the National Cancer Institute and its associated analysis algorithm, COMPARE.[25,26] The NCI60 assay has commonly

been used to identify mechanisms of action of bioactive compounds.[27] Briefly, The growth inhibitory activity and cytotoxicity of **1** was measured in the NCI60 panel of 59 cell lines. The results were compared against the NCI "Standard Agents" library of reference compounds using the COMPARE algorithm.[26,28] Compounds with highly correlated COMPARE signatures are suggested to have similar mechanisms of action. The NCI60 activity profile of **1** correlated most highly with that of **1** was phyllanthoside, a known eukaryotic protein synthesis inhibitor that binds directly to the 80S ribosome and blocks translation elongation.[28,29] Other compounds with similar, albeit lower, COMPARE correlations to **1** included microtubule poisons, DNA intercalators, and other eukaryotic protein synthesis inhibitors (Figure 1-3 A).

a

Compound	Mechanism of action	COMPARE correlation to **1**
phyllanthoside	Protein synthesis inhibitor	0.80
vinorelbine tartrate	Microtubule poison	0.73
bisantrene hydrochloride	Intercalating agent, DNA replication inhibitor	0.70
chromomycin A3	DNA binding, RNA synthesis inhibition	0.70
actinomycin D	Intercalating agent	0.68
bruceantin	Peptidyl transferase inhibition	0.66
vinblastine sulfate	Microtubule poison	0.62
didemnin B	Apoptosis, protein synthesis inhibition	0.61
paclitaxel	Apoptosis, cell division inhibitor, tubulin binder	0.61
echinomycin	Intercalating agent, DNA/RNA synthesis inhibitor	0.60

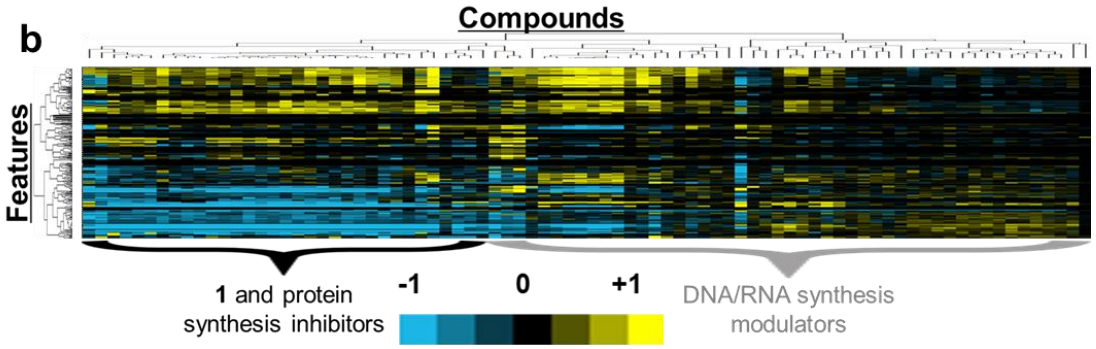


Figure 1-3: COMPARE results and cytological profiling dendrogram fingerprint (*A*) Top ten results from COMPARE analysis ordered by COMPARE correlation to **1** *(B)* Compound **1** was compared to DMSO controls across 254 features normalized with features represented on a scale from -1 (blue, below DMSO control) to +1 (yellow, above DMSO control) with black indicating no difference between compound and DMSO for that feature. Compounds were clustered by similarity to a reference library of 25 bioactive compounds with known mechanisms of action at three concentrations (50 µM, 10 µM, 2 µM).

To orthogonally augment the COMPARE data, we used cytological profiling (CP). This is a high-content screening platform that uses automated fluorescence microscopy and unbiased image analysis to correlate the cordyheptapeptide phenotype with that of compounds from a reference library of known drugs.[30-32] Briefly, HeLa cells were incubated with the compound of interest for 18 h and then stained with fluorescent probes for DNA (Hoechst), S-phase (EdU), and mitotic cells (anti-phospho-

histone H3), actin (phalloidin), and tubulin (anti-tubulin mAb). Computational analysis of the combined images yielded 254 cytological features, normalized to DMSO controls, to provide a fingerprint for each compound (Figure 1-3 B). Similar to COMPARE, compounds with a similar mechanism of action have similar fingerprints and cluster together. Our reference library contained 25 compounds with MOAs including protein synthesis inhibitors, DNA synthesis modulators, cell cycle modulators, and microtubule poisons (see Supplementary Figure 1-S8 for full list) at three concentrations (50 μM, 10 μM, and 2 μM). Across all three concentrations, compound **1** clustered most closely to a variety of eukaryotic protein synthesis inhibitors, including bouvardin, phyllanthoside, cycloheximide, didemnin B, ternatin, anisomycin, and ansatrienin A (Figure 1-3 B). Moreover, **1** was further distinguished from the microtubule poisons and DNA synthesis inhibitors which clustered more closely to **5**, a much less active cordyheptapeptide derivative (Figure 1-3 B, Supplementary Figure 1-S8). These corroborating results pointed toward protein synthesis inhibition as a likely MOA of **1**.

The fluorescent microscopy images from CP also allow us to directly eliminate microtubule poison (suggested by the COMPARE analysis) as the MOA of cordyheptapeptide. The COMPARE results for **1** showed a correlation with three microtubule poisons that are known to cause mitotic arrest (vinorelbine, vinblastine, and paclitaxel).[33-35] In contrast, **1** caused a *decrease* in the percentage of mitotic cells compared to a DMSO control as measured by the anti-phospho-histone H3 mitotic

marker (Supplementary Figure 1-S6), providing strong evidence that **1** does not act as a classic antimitotic drug.

NCI60 results from specific cell lines also allowed us to eliminate of poly(ADP-ribose) polymerase (PARP) inhibition as a potential mechanism of action. PARP inhibitors like olaparib have been shown to be synthetically lethal with BRCA1-negative breast cancer cell lines.[36] However, the NCI60 dose-response assay showed **1** to have similar IC_{50}s in both BRCA1 positive and negative breast cancer cell lines (Supplementary Figure 1-S7). Therefore, the NCI60 COMPARE data do not support PARP inhibition as an MOA for **1**. The fingerprint of **1** was also distinguished from the PARP inhibitor Olaparib in the CP clustering. Thus, the combination of CP and COMPARE allowed us to eliminate both microtubule cytoskeleton and PARP as potential targets.

1.2.5 Quantification of protein synthesis and DNA synthesis

Protein synthesis was measured using the bioorthogonal noncanonical amino acid tagging (BONCAT) method, a radiation-free alternative to the classic ^{35}S-methionine incorporation assay.[37] BONCAT quantifies protein synthesis based on the addition of a pulse of the methionine-mimetic L-homopropargyl glycine (L-HPG), which, when incorporated into proteins, provides a handle for covalent linkage to rhodamine-azide using Cu(I)-catalysis.[38] Cells are incubated with a compound of interest followed by a Met-depletion in Met-free cell culture medium and a 1-h pulse of L-HPG. The cells are fixed, stained, and imaged using fluorescent microscopy. A

24-h treatment with compound **1** caused a dose-dependent decrease in protein synthesis as measured by BONCAT in HeLa cells (IC_{50} = 0.1 µM, Figure 1-S4 A-C).

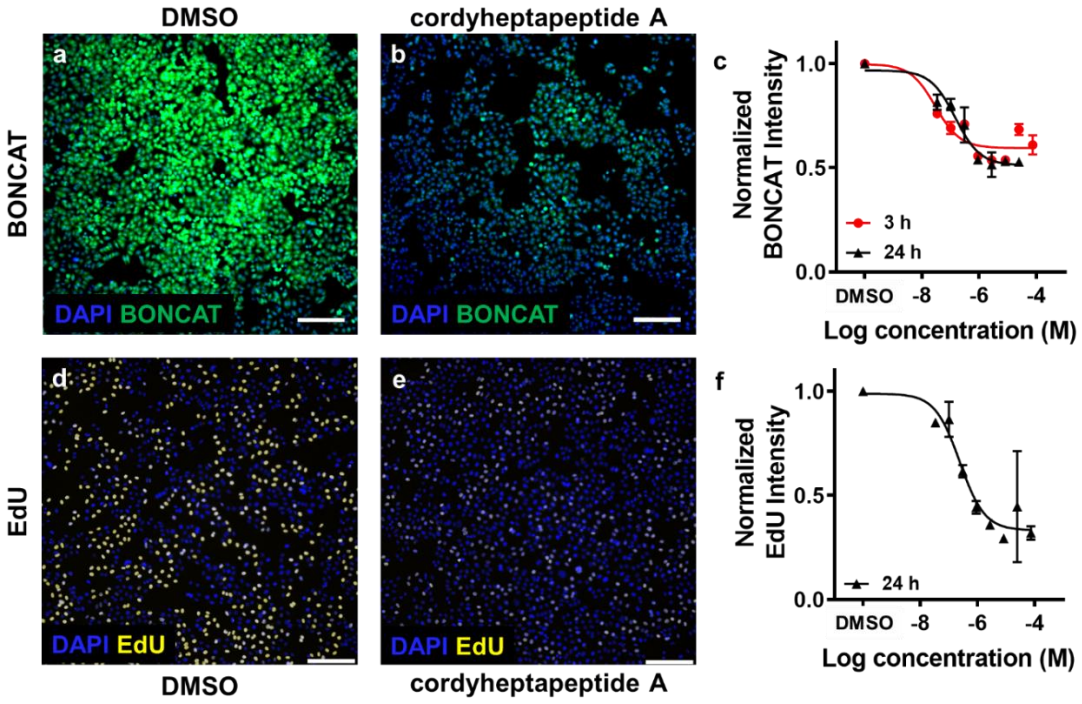


Figure 1-4: Effects of cordyheptapeptide A on protein synthesis and DNA synthesis *(A-B)* Representative images of HeLa cells stained for the following features: nuclei (Hoechst, blue), protein synthesis (BONCAT, green) after a 24-h incubation with *(A)* DMSO or *(B)* 100 nM **1** in 96-well plates. *(C)* Effect of **1** on normalized cell average BONCAT intensity in HeLa cells after a 3-h or 24-h incubation at 37 °C in 96-well plates. *(D-E)* Representative images of HeLa cells stained for the following features: nuclei (Hoechst, blue), and S-phase cells (EdU, yellow) after a 2-h incubation at 37 °C with *(D)* DMSO or *(E)* 0.6 µM **1** in 96-well plates. *(F)* Effect of **1** on normalized EdU intensity in HeLa cells after a 2-h or 24-h incubation at 37 °C in 96-well plates. Scale bars: 211 µm. Error bars are one SD.

Compound **1** also produced phenotypic changes in cellular DNA content in out CP assay. Similar to BONCAT, the rate of DNA synthesis is measured by treating cells with a 1-h pulse of ethynyl-deoxyuridine (EdU), a propargylated-thymidine mimic which is incorporated into DNA during S-phase and can subsequently be

quantified by reaction with rhodamine-azide using Cu(I)-catalysis.[39] Overall, **1** caused both a dose-dependent increase in the percentage of EdU-positive cells and a decrease in the average EdU intensity in the EdU-positive cells after a 24 h treatment in HeLa cells (IC_{50} = 0.2 μM, Figure 1-S4 D-F). This suggests a decrease in the rate of DNA synthesis, resulting in a build-up in the number of cells in S-phase.

Inhibition of protein synthesis has been shown to cause a decrease in the rate of DNA synthesis,[28] which could account for our observation that, in addition to its effect on protein synthesis, **1** caused a significant decrease in the rate of DNA synthesis as measured by EdU uptake. However, to confirm that the effect on DNA synthesis was secondary to its effect on protein synthesis, we performed the BONCAT assay after a 3-h treatment with **1**. Even at this shorter time point, **1** inhibited protein synthesis and had a more potent effect on protein synthesis than DNA synthesis. This result suggests that **1** likely acts as a direct protein synthesis inhibitor, which, in turn, has a secondary effect on DNA synthesis, consistent with both CP and NCI60 results (Figure 1-3 B).

1.2.6 Identification of eEF1A as the intracellular target of 1

To identify candidate targets related to the protein synthesis inhibition activity of **1**, we performed photo-affinity labeling experiments with a cordyheptapeptide derivative (**43**), which contains a propargyl glycine residue and a diazirine-proline derivative (Figure 1-5 A). Similar to photo-leucine,[40] the diazirine-containing "photo-proline" forms a carbene intermediate when irradiated with UV light and reacts with

an amino acid residue in close proximity. Briefly, **43** was incubated with HeLa cells then irradiated with UV with a filter with a max intensity at 365 nm to attach **43** to its cellular target covalently. After cell lysis, the propargyl was used to covalently connect rhodamine-biotin-azide to **43**, which effectively fluorescently tags the cellular target of **43**. Streptavidin-conjugated magnetic beads were used to concentrate the target protein from the cell lysate. The proteins were then separated by gel electrophoresis and imaged for rhodamine fluorescence. Shown in Figure 1-5 B, a fluorescent band just above 44 kDa is present and is darker in our experiment lane than in any of our negative control wells. This result suggested the molecular weight of our target is between 44 and 50 kDa.

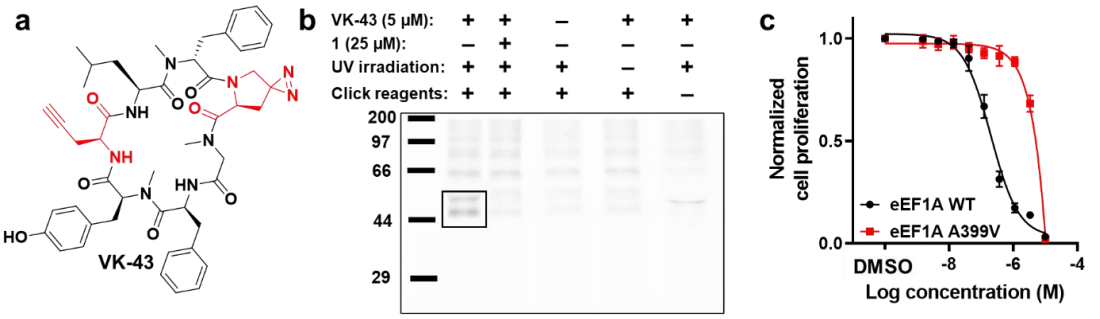


Figure 1-5: Molecular weight determination and target identification *(A)* Structure of cordyheptapeptide derivative (**43**) used for photo-crosslinking experiments. This derivative has a Pro-5 to photo-Pro substitution for photo-crosslinking and an Ile-1 to propargyl glycine substitution for click chemistry. *(B)* HeLa cells were incubated with **43** for 3 h at 37°C followed by UV irradiation (360 nm, 150 W, 10 minutes). Cells were then frozen overnight and lysed with RIPA buffer at 0 °C. Samples treated with TAMRA-biotin-azide, separated with magnetic streptavidin beads, eluted, separated by gel-electrophoresis, and imaged for fluorescence. *(C)* eEF1A WT HCT 116 cells and eEF1A A399V HCT 116 cells were incubated with **1** for 72 hours. Cell proliferation was measured with and Alamar blue assay. Error bars are one SD.

Based on these results, we postulated that the target could be eEF1A, the 50-kDa protein elongation factor that is involved in the recruitment of aminoacyl-tRNA to the A-site of the 80S ribosome. The protein eEF1A is essential for protein synthesis in eukaryotes, and among its known inhibitors are the cytotoxic cyclic peptide natural products didemnin B and ternatin (including its more potent analog ternatin-4). Like the cordyheptapeptides, ternatin is a head-to-tail cyclized heptapeptide with predominantly aliphatic side chains. Although the cordyheptapeptides and ternatin (and its more potent variant, ternatin-4) have different stereochemical and N-methylation patterns, their overall structural and physicochemical similarities, along with the results from the photoaffinity experiments, prompted us to test whether the inhibition of protein synthesis by **1** was also due to a direct interaction with eEF1A. In their confirmation of eEF1A as the cellular target of ternatin and its improved variant ternatin-4, Carelli *et al.*, showed that a point mutation in eEF1A (A399V) confers complete resistance.[41] A variety of structurally unrelated eEF1A inhibitors are also inactive against this mutant cell line.[41] While **1** caused a dose-dependent reduction in wildtype HCT 116 cell proliferation (IC_{50} ~ 0.2 μM), this effect was dramatically reduced in the A399V eEF1A mutant cells by more than 20-fold (IC_{50} ~ 7 μM, Figure 1-5 C). This "gold-standard" genetic evidence strongly suggests that eEF1A is the direct intracellular target of cordyheptapeptide A.

1.3 Discussion and Conclusion

Overexpression of eEF1A is associated with several cancer types, including prostate, breast, colon, and lung cancer.[42] While one traditional druglike molecule, gamendazole, has been shown to bind eEF1A, this is a non-specific inhibitor pursued more for its antispermatogenic and antifertility effects than its anticancer properties.[43] Conversely, cordyheptapeptide A is yet another macrocyclic natural product shown to bind to eEF1A and inhibit protein synthesis. This list of structurally diverse natural products includes ternatin (a cyclic heptapeptide),[41] didemnin B (a lariat depsipeptide),[44] plitidepsin (a natural didemnin B derivative),[45] cytotrienin A (a macrocyclic lactam), ansatrienin B (a macrocyclic lactam), nannocystin A (a hybrid peptide-polyketide macrocycle),[46] and ansatrienin B (a polyketide).[41] While there are major differences in their overall structures, the above compounds are all large, hydrophobic molecules. Our study adds to the mounting evidence that successful inhibition of this target requires a large, "beyond-Rule-of-5" (bRo5) molecule. The high potencies of these compounds support the continuation of work to investigate the design and synthesis of natural product-like drug molecules dominated by hydrophobic- and aromatic-rich side chains.[18,47,48] Additional investigation of these compounds could lead to a better understanding of the binding mechanisms of eEF1A and allow for faster MOA identification of novel natural products that fit into this class of protein synthesis inhibitors.

The synthetic and natural cordyheptapeptides in this study showed weak to moderate passive permeability coefficients in PAMPA ranging from 0.2×10^{-6} cm/s to

7.8 x 10^{-6} cm/s. The seven compounds with sub-micromolar cellular activities, including the parent natural product **1**, were not among the most permeable in this study, with P_e values ranging from 0.9 x 10^{-6} to 1.3 x 10^{-6} cm/s. Many of the backbone-modifying substitutions in this series led to enhanced permeability, including the substitution of Pro with Ala (**6**) or MeAla (**7**), addition of a backbone N-Me (**16**, **17**), *removal* of a backbone N-Me (**15**), or substitution of an α- with a β-amino acid (**41**). The impact of these substitutions on permeability is perhaps not surprising given the established sensitivity of passive permeability to backbone conformation. Unfortunately, each of these substitutions caused a dramatic decrease in cellular potency, highlighting the difficulty of optimizing permeability in cyclic peptide scaffolds without abrogating biochemical efficacy.[49,50]

The effect of backbone-altering substitutions on solubility also followed interesting trends. While removal of the N-Me from position six enhanced solubility (**12** vs. **1** and **15** vs. **2**), backbone N-methylation increased aqueous solubility at positions one (**16** vs. **3**) and three (**17** vs. **1**), N-methylation at position four caused a variable response in which either the Gly or Sar derivative was more slightly soluble depending on the compound. However, the solubilities of the matched pairs were broadly comparable. This paradoxical effect of N-methylation on solubility, as well as its positional dependence, has been observed in other cyclic peptide systems[7,51] and suggests that the effect of backbone modifications on properties is highly scaffold-dependent.

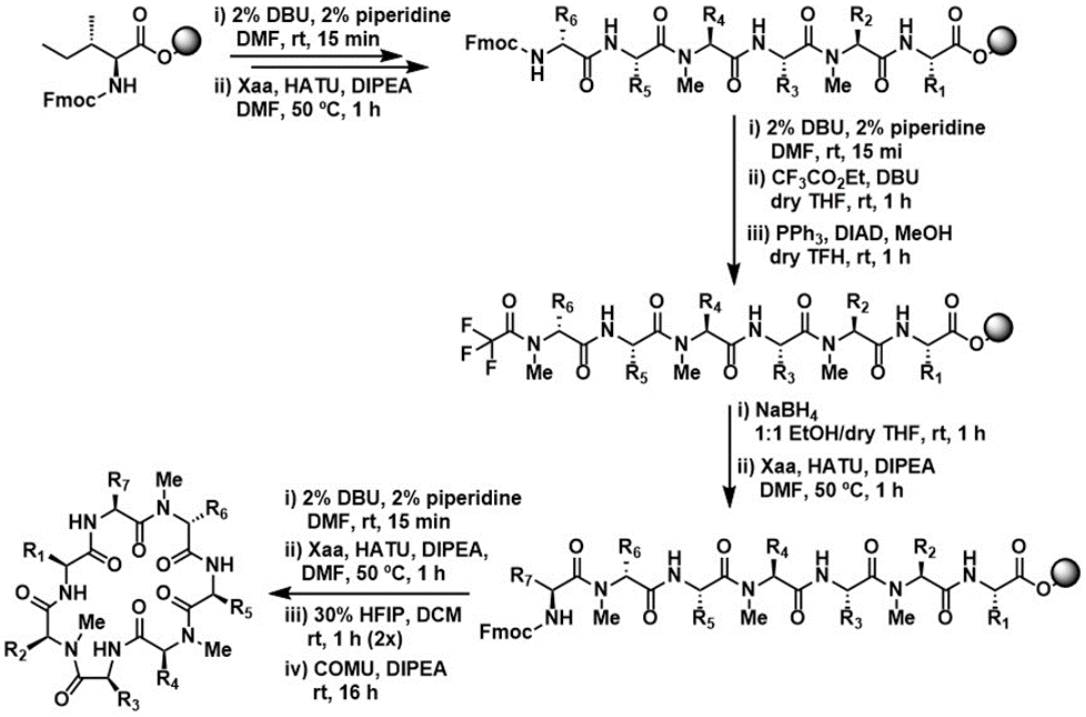
Changing the N-methylation pattern of the cyclic peptide backbone can also dramatically modulate its bioactivity.[49,50] This is most evident when the removal of the N-Me at position four was combined with halogenating a Phe (**19 – 26** vs. **27 – 34**) in which the Gly compounds (**27 – 34**) are active than the Sar derivatives despite generally having similar permeabilities and solubilities as their Sar-containing counterparts (**19 – 26**). The more flexible Gly-compounds could allow for conformations yielding increased access to favorable halogen-to-target binding. Alternatively, it is also possible that this increased bioactivity is derived from a new intermolecular hydrogen bond between the newly exposed -NH at position four and the protein target, eEF1A. This hypothesis is supported by recent reviews which have suggested that cyclic peptide backbone amides are responsible for a significant a significant portion of target binding.[52]

This study highlights the interwoven nature of macrocycle bioactivity and physicochemical properties, and a strategy to optimize these properties simultaneously via classical synthetic techniques such as alanine, N-methyl, and isostere scanning. In particular, a demethylation at position a yielded new scaffolds (**27**, **28**, and **30**) with maintained potency and permeability relative to cordyheptapeptide A (**1**), and a two-fold increase in bioactivity for another (**40**) relative to cordyheptapeptide B (**2**). Furthermore, this modification, in conjunction with a fluorine substitution, improved the activity of **2** to nearly that of **1**, the more active natural product. Consequently, modifications to the N-methylation pattern of a cyclic peptide scaffold offer effective strategies to improve the potency and oral bioavailability of bRo5 molecules in addition

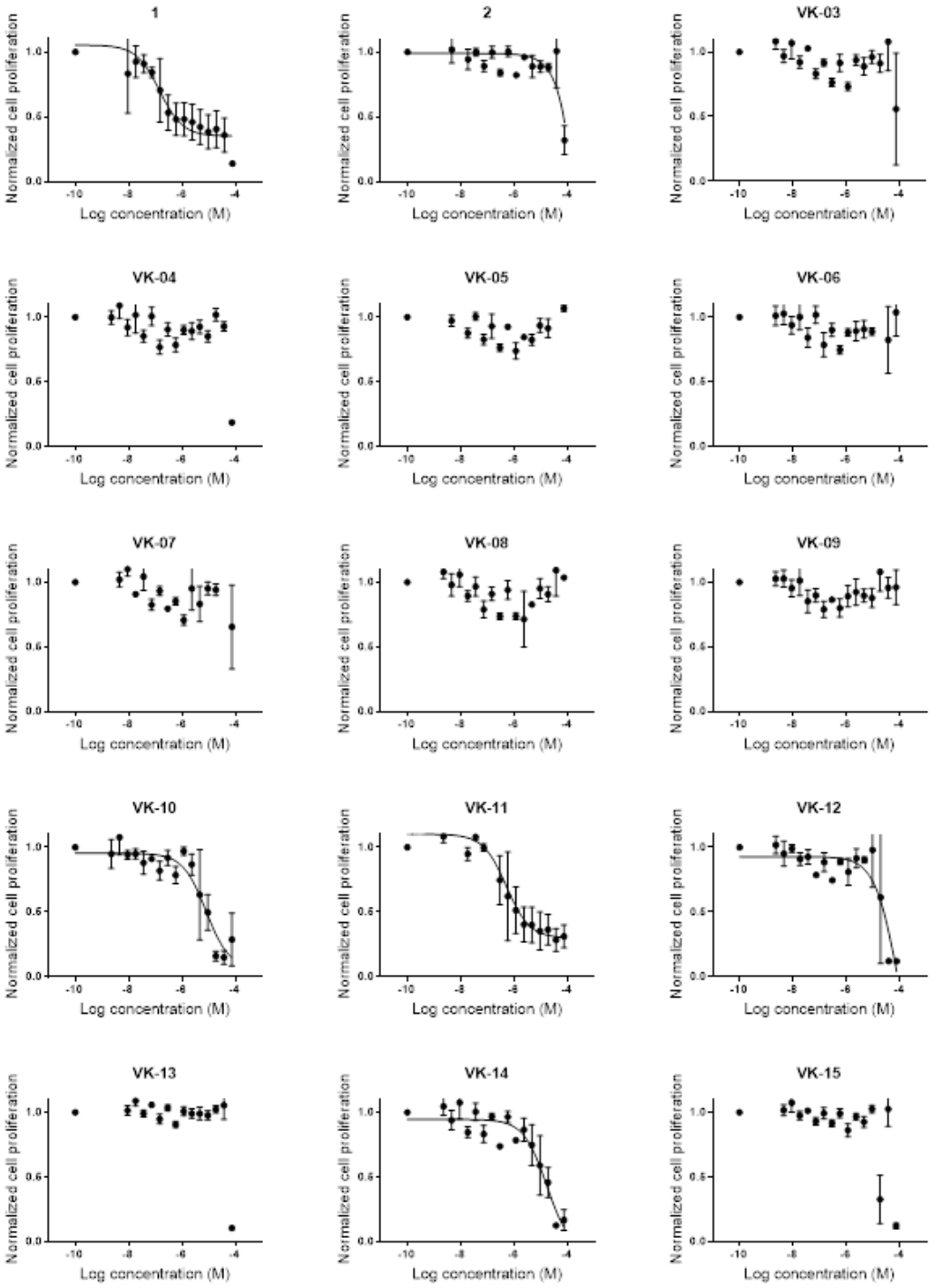
to the inclusion of more exotic, “druglike” side chains. This work improves our understanding of the physicochemical properties of cyclic peptide natural products. Additionally, it confirms that lipophilic macrocycles can inhibit intracellular proteins responsible for cellular growth. Therefore, our study motivates further investigation of cyclic peptides as potential cancer therapeutics, and modulators of protein targets once thought to be “undruggable”.

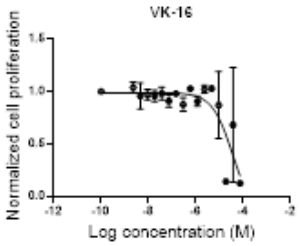
1.4 Supplementary Schemes, Figures, and Tables

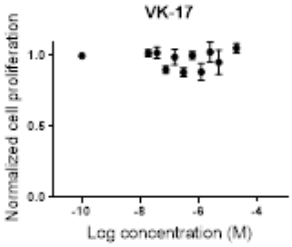
1.4.1 Supplementary Scheme 1-S1: Solid-phase synthetic pathway to cyclic, N-methylated heptapeptides

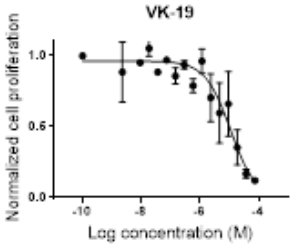


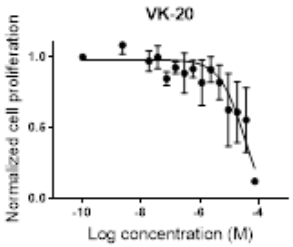
1.4.2 Supplementary Figure 1-S1: Cell proliferation IC50 dose-response curves

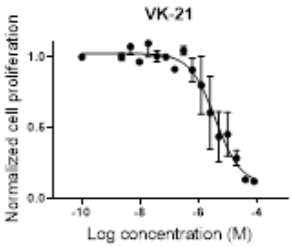


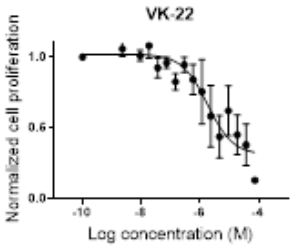

VK-16
Normalized cell proliferation
Log concentration (M)

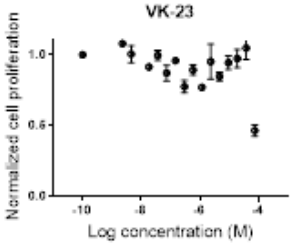

VK-17
Normalized cell proliferation
Log concentration (M)

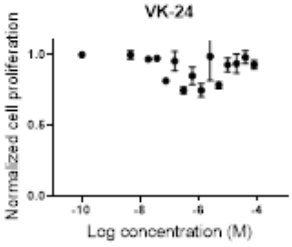

VK-19
Normalized cell proliferation
Log concentration (M)

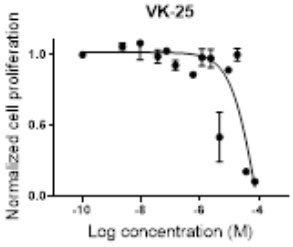

VK-20
Normalized cell proliferation
Log concentration (M)

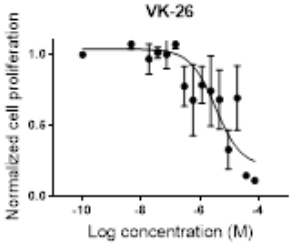

VK-21
Normalized cell proliferation
Log concentration (M)

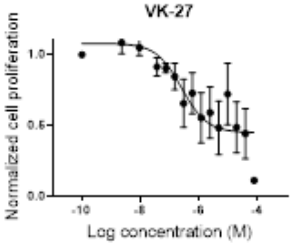

VK-22
Normalized cell proliferation
Log concentration (M)

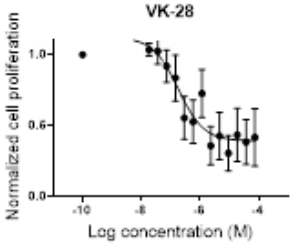

VK-23
Normalized cell proliferation
Log concentration (M)

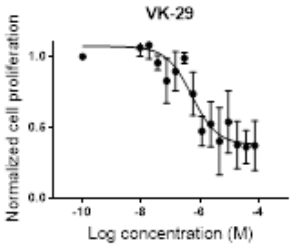

VK-24
Normalized cell proliferation
Log concentration (M)

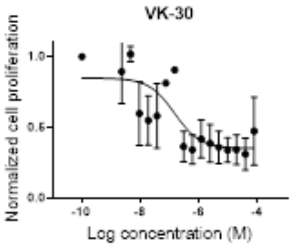

VK-25
Normalized cell proliferation
Log concentration (M)

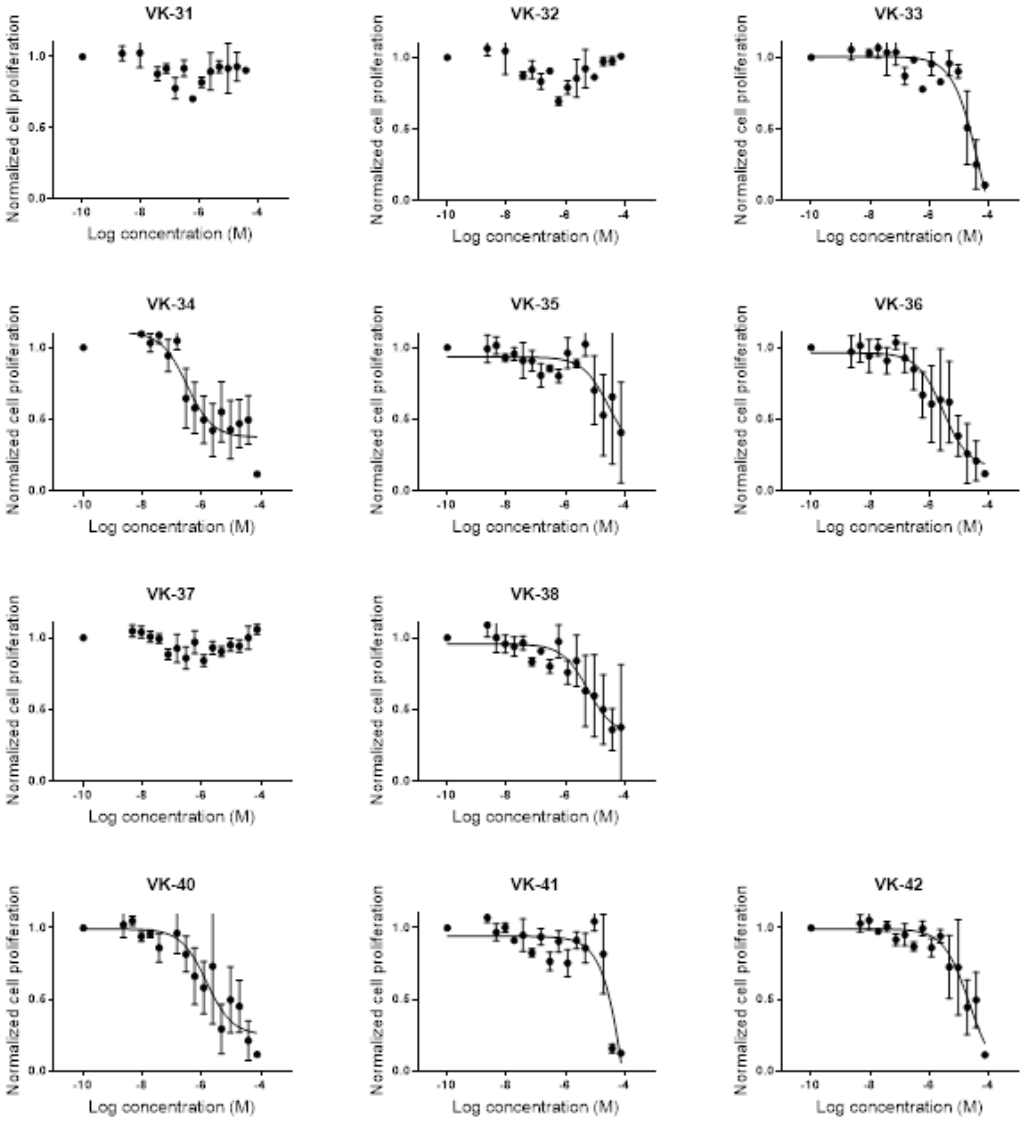

VK-26
Normalized cell proliferation
Log concentration (M)


VK-27
Normalized cell proliferation
Log concentration (M)


VK-28
Normalized cell proliferation
Log concentration (M)

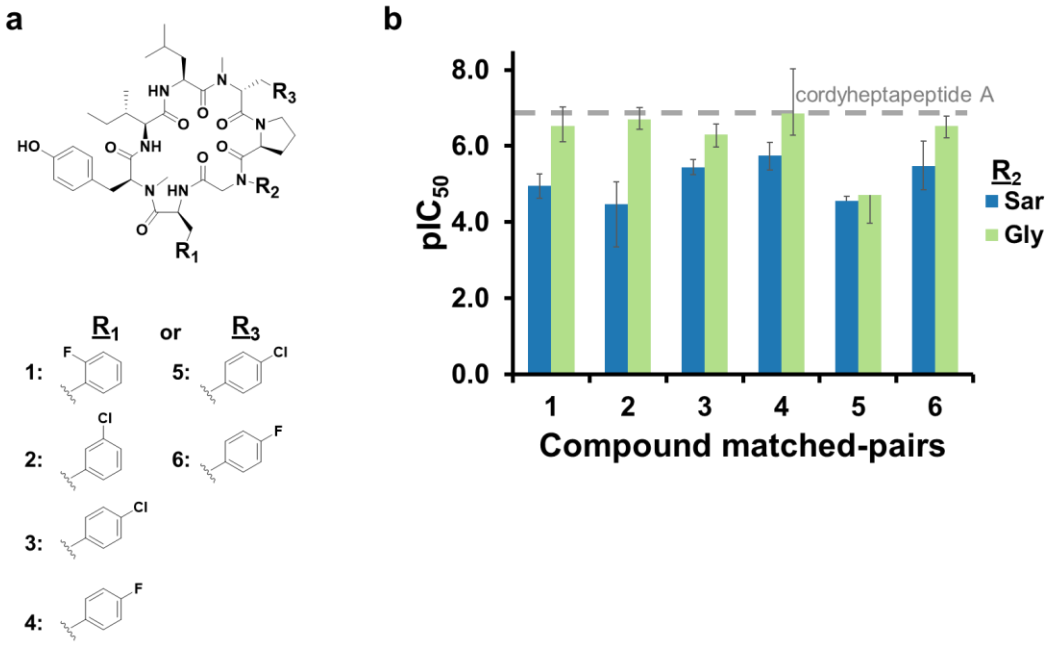

VK-29
Normalized cell proliferation
Log concentration (M)


VK-30
Normalized cell proliferation
Log concentration (M)



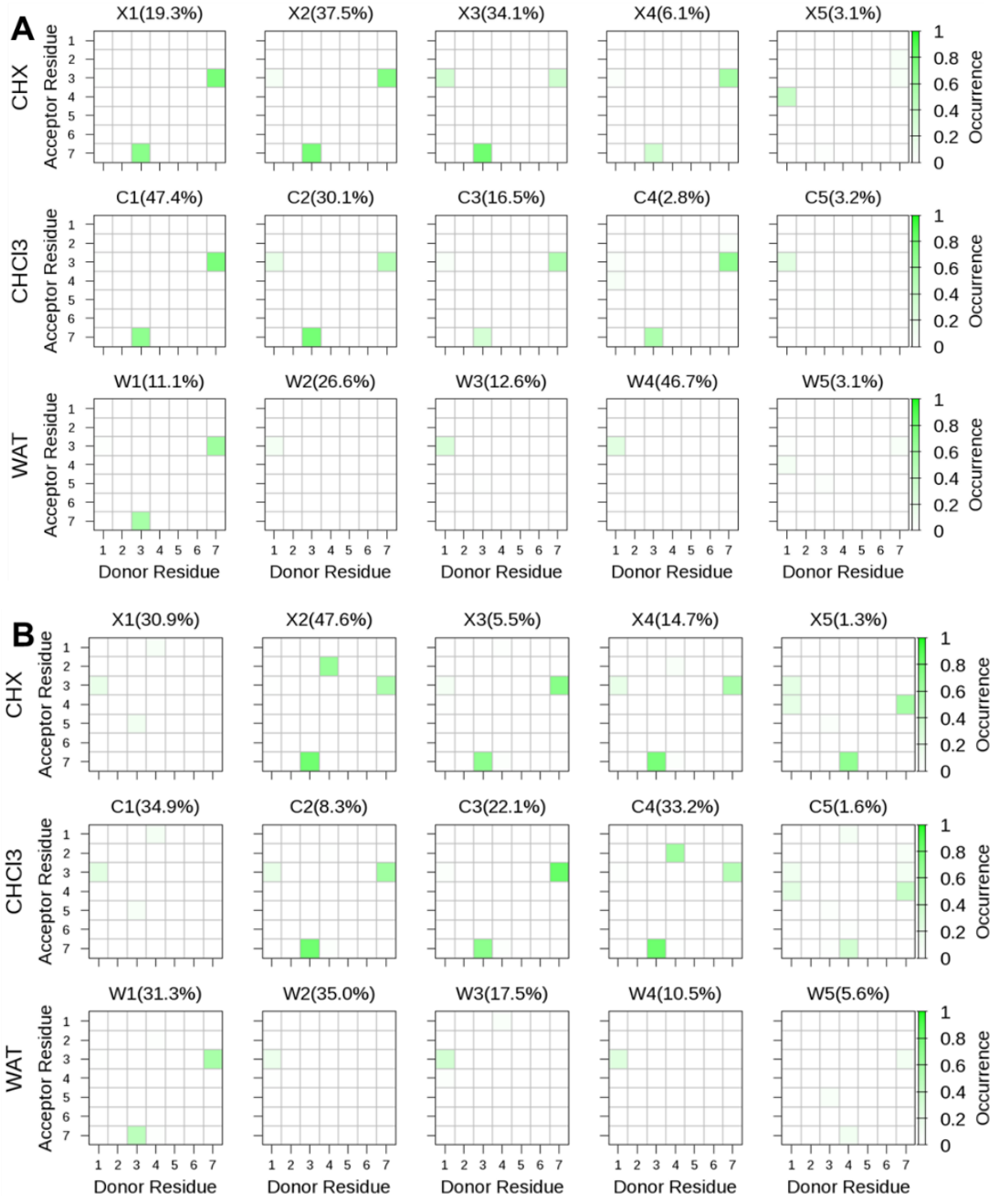
Cell proliferation IC_{50} dose-response curves Dose-response curves for compounds effects on cell proliferation in HCT 116 cells after a 72-h compound incubation time. Cell proliferation was measured using an Alamar blue assay. Error bars represent one standard deviation around the mean. Results were normalized to DMSO which is represented as a concentration of -10 on the x-axis. Error bars are one SD.

1.4.3 Supplementary Figure 1-S2: Gly and Sar matched pair analysis



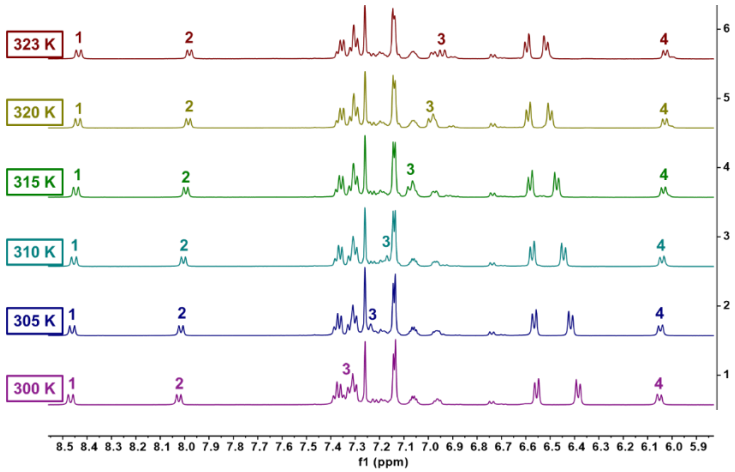
(a) Structure of matched compound pairs (b) Plot showing the pIC_{50} for the matched compound pairs. Gray dashed line represents the pIC_{50} for cordyheptapeptide A (**1**). Error bars are a 95% confidence interval. (n = 3)

1.4.4 Supplementary Figure 1-S3: Molecular dynamics simulations for VK11 (11)

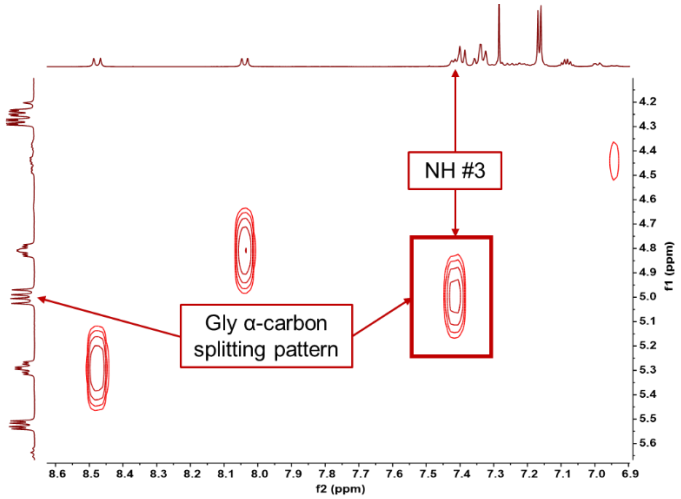


Intramolecular hydrogen bond pattern for the top five conformers in cyclohexane (CHX), chloroform ($CHCl_3$), and water (WAT) for *(A)* cordyheptapeptide A *(B)* **11**.

1.4.5 Supplementary Figure 1-S4: NMR temperature coefficient graphs for VK11 (11)

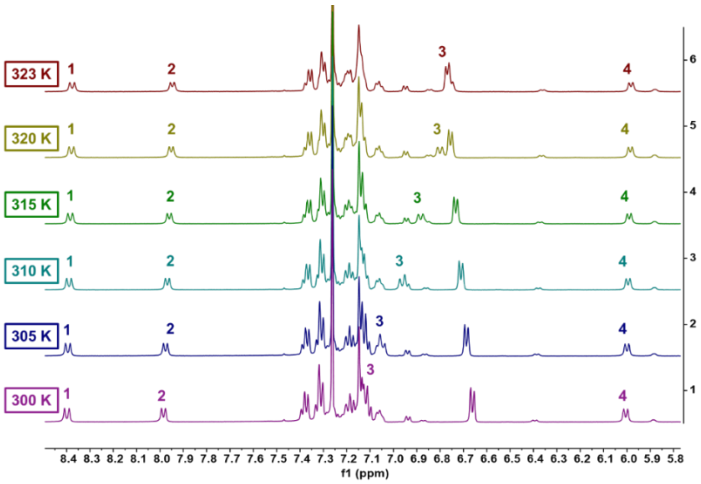


Amide[i]	Shift[ii]	*J* NH-Ha[iii]	Dppb/K[iv]
1	8.45	9.87	-1.56
2	8.00	8.42	-1.80
3	7.07	10.13	-16.25
4	6.04	8.44	-0.97

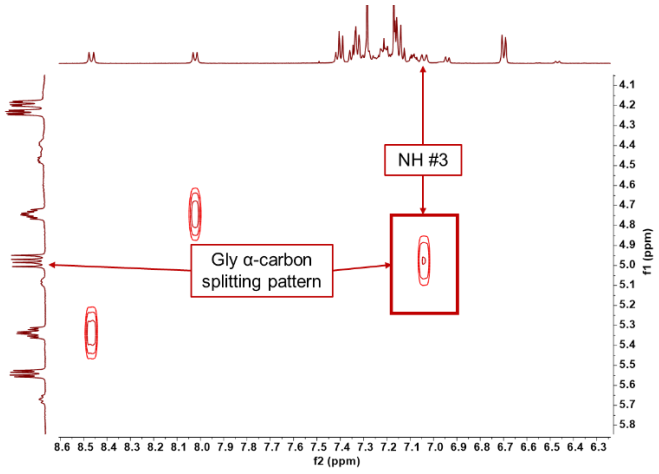


(Top) ^{1}H-NMR of **11** at six temperatures from 300 K to 323 K. Spectra show four amide bond peaks (labeled 1 – 4). The amide hydogen #3 moves to the right as the temperature increase 500 MHz in $CDCl_3$; *(Middle)* Temperature shift coefficients for (*i*) amides 1 – 4, (ii) chemical shift of amide peak -NH peak, (iii) coupling constant of amide -NH peak (iv) temperaure shift coeffieient in ppb; *(Bottom)* 2D-NMR COSY at 300 K confirming that hydrogen #3 is the Gly NH at postion four; rt, 500 MHz, $CDCl_3$.

1.4.6 Supplementary Figure 1-S5: NMR temperature coefficient graphs for VK14 (14)

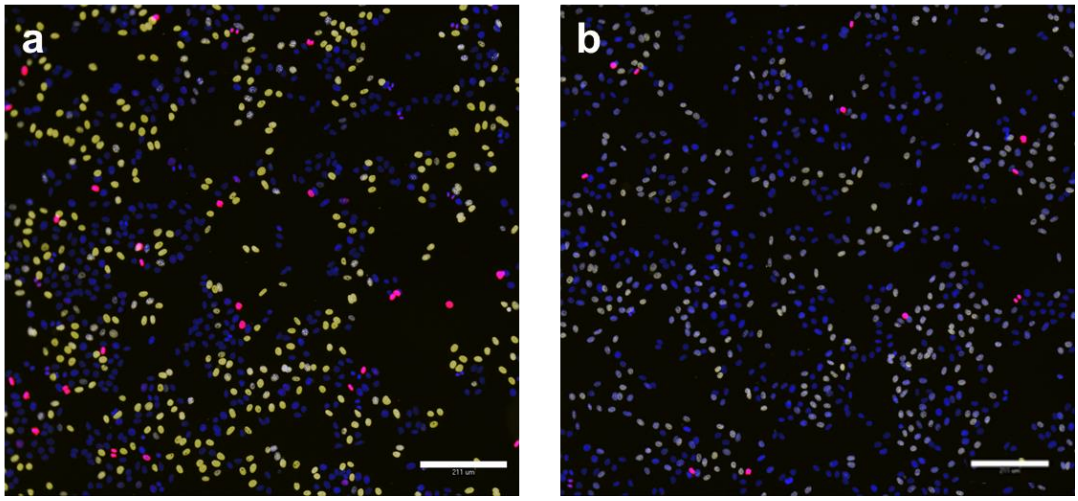


Amide[i]	Shift[ii]	*J* NH-Ha[iii]	Dppb/K[iv]
1	8.39	9.94	-0.86
2	7.97	8.38	-1.34
3	6.96	10.52	-16.00
4	5.99	8.31	-0.77



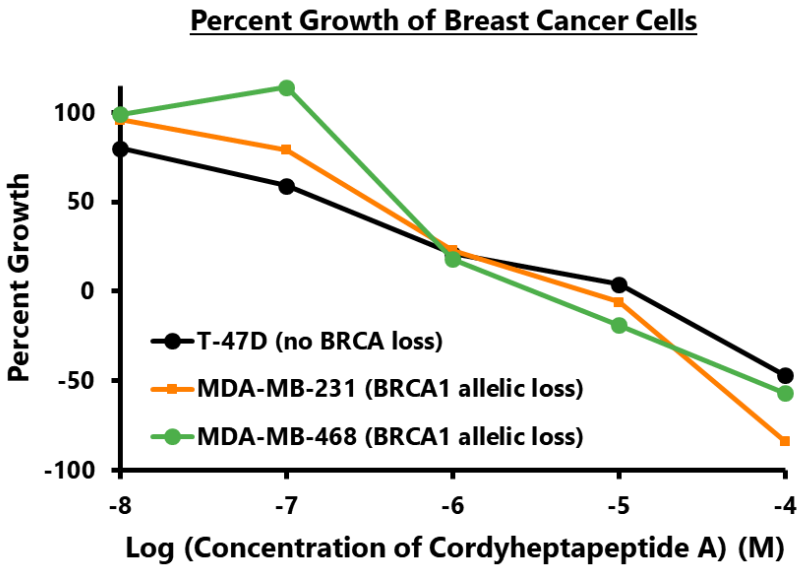
(Top) ^{1}H-NMR of **14** at six temperatures from 300 K to 323 K. Spectra show four amide bond peaks (labeled 1 – 4). The amide hydrogen #3 moves to the right as the temperature increase 500 MHz in $CDCl_3$; *(Middle)* Temperature shift coefficients for (*i*) amides 1 – 4, (ii) chemical shift of amide peak -NH peak, (iii) coupling constant of amide -NH peak (iv) temperaure shift coeffieient in ppb; *(Bottom)* 2D-NMR COSY at 300 K confirming that hydrogen #3 is the Gly NH at postion four; rt, 500 MHz, $CDCl_3$.

1.4.7 Supplementary Figure 1-S6: Effects of cordyheptapeptide A on DNA synthesis and mitosis



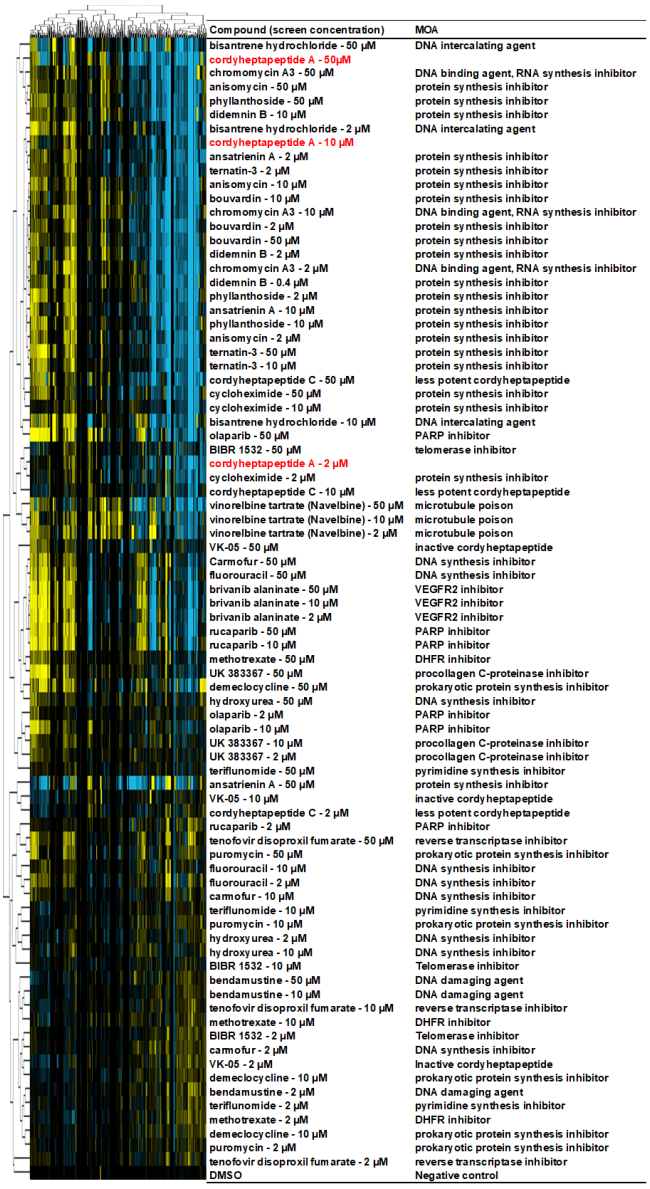
Representative images of HeLa cells stained for the following features: nuclei (Hoechst, blue), S-phase cells (EdU, yellow), and mitotic cells (anti-phospho-histone H3, pink) after an 18-hour incubation with *(A)* DMSO or *(B)* 9 nM **1** in 384-well plates. Scale bar = 211 µm.

1.4.8 Supplementary Figure 1-S7: NCI60 percent growth data for three breast cancer cell lines



Effect of cordyheptapeptide A on the percent growth of three breast cancer cell lines (one with no BRCA1 loss and two with BRCA1 allelic loss) from the NCI60 five-dose screen. The overlap of the data suggests that cordyheptapeptide A is not synthetically lethal with any of the cell lines in this figure.

1.4.9 Supplementary Figure 1-S8: Cytological profiling dendrogram fingerprint for follow-up CP library with full list of compounds



Compound (screen concentration)	MOA
bisantrene hydrochloride - 50 µM	DNA intercalating agent
cordyheptapeptide A - 50µM	
chromomycin A3 - 50 µM	DNA binding agent, RNA synthesis inhibitor
anisomycin - 50 µM	protein synthesis inhibitor
phyllanthoside - 50 µM	protein synthesis inhibitor
didemnin B - 10 µM	protein synthesis inhibitor
bisantrene hydrochloride - 2 µM	DNA intercalating agent
cordyheptapeptide A - 10 µM	
ansatrienin A - 2 µM	protein synthesis inhibitor
ternatin-3 - 2 µM	protein synthesis inhibitor
anisomycin - 10 µM	protein synthesis inhibitor
bouvardin - 10 µM	protein synthesis inhibitor
chromomycin A3 - 10 µM	DNA binding agent, RNA synthesis inhibitor
bouvardin - 2 µM	protein synthesis inhibitor
bouvardin - 50 µM	protein synthesis inhibitor
didemnin B - 2 µM	protein synthesis inhibitor
chromomycin A3 - 2 µM	DNA binding agent, RNA synthesis inhibitor
didemnin B - 0.4 µM	protein synthesis inhibitor
phyllanthoside - 2 µM	protein synthesis inhibitor
ansatrienin A - 10 µM	protein synthesis inhibitor
phyllanthoside - 10 µM	protein synthesis inhibitor
anisomycin - 2 µM	protein synthesis inhibitor
ternatin-3 - 50 µM	protein synthesis inhibitor
ternatin-3 - 10 µM	protein synthesis inhibitor
cordyheptapeptide C - 50 µM	less potent cordyheptapeptide
cycloheximide - 50 µM	protein synthesis inhibitor
cycloheximide - 10 µM	protein synthesis inhibitor
bisantrene hydrochloride - 10 µM	DNA intercalating agent
olaparib - 50 µM	PARP inhibitor
BIBR 1532 - 50 µM	telomerase inhibitor
cordyheptapeptide A - 2 µM	
cycloheximide - 2 µM	protein synthesis inhibitor
cordyheptapeptide C - 10 µM	less potent cordyheptapeptide
vinorelbine tartrate (Navelbine) - 50 µM	microtubule poison
vinorelbine tartrate (Navelbine) - 10 µM	microtubule poison
vinorelbine tartrate (Navelbine) - 2 µM	microtubule poison
VK-05 - 50 µM	inactive cordyheptapeptide
Carmofur - 50 µM	DNA synthesis inhibitor
fluorouracil - 50 µM	DNA synthesis inhibitor
brivanib alaninate - 50 µM	VEGFR2 inhibitor
brivanib alaninate - 10 µM	VEGFR2 inhibitor
brivanib alaninate - 2 µM	VEGFR2 inhibitor
rucaparib - 50 µM	PARP inhibitor
rucaparib - 10 µM	PARP inhibitor
methotrexate - 50 µM	DHFR inhibitor
UK 383367 - 50 µM	procollagen C-proteinase inhibitor
demeclocycline - 50 µM	prokaryotic protein synthesis inhibitor
hydroxyurea - 50 µM	DNA synthesis inhibitor
olaparib - 2 µM	PARP inhibitor
olaparib - 10 µM	PARP inhibitor
UK 383367 - 10 µM	procollagen C-proteinase inhibitor
UK 383367 - 2 µM	procollagen C-proteinase inhibitor
teriflunomide - 50 µM	pyrimidine synthesis inhibitor
ansatrienin A - 50 µM	protein synthesis inhibitor
VK-05 - 10 µM	inactive cordyheptapeptide
cordyheptapeptide C - 2 µM	less potent cordyheptapeptide
rucaparib - 2 µM	PARP inhibitor
tenofovir disoproxil fumarate - 50 µM	reverse transcriptase inhibitor
puromycin - 50 µM	prokaryotic protein synthesis inhibitor
fluorouracil - 10 µM	DNA synthesis inhibitor
fluorouracil - 2 µM	DNA synthesis inhibitor
carmofur - 10 µM	DNA synthesis inhibitor
teriflunomide - 10 µM	pyrimidine synthesis inhibitor
puromycin - 10 µM	prokaryotic protein synthesis inhibitor
hydroxyurea - 2 µM	DNA synthesis inhibitor
hydroxyurea - 10 µM	DNA synthesis inhibitor
BIBR 1532 - 10 µM	Telomerase inhibitor
bendamustine - 50 µM	DNA damaging agent
bendamustine - 10 µM	DNA damaging agent
tenofovir disoproxil fumarate - 10 µM	reverse transcriptase inhibitor
methotrexate - 10 µM	DHFR inhibitor
BIBR 1532 - 2 µM	Telomerase inhibitor
carmofur - 2 µM	DNA synthesis inhibitor
VK-05 - 2 µM	Inactive cordyheptapeptide
demeclocycline - 10 µM	prokaryotic protein synthesis inhibitor
bendamustine - 2 µM	DNA damaging agent
teriflunomide - 2 µM	pyrimidine synthesis inhibitor
methotrexate - 2 µM	DHFR inhibitor
demeclocycline - 10 µM	prokaryotic protein synthesis inhibitor
puromycin - 2 µM	prokaryotic protein synthesis inhibitor
tenofovir disoproxil fumarate - 2 µM	reverse transcriptase inhibitor
DMSO	Negative control

Compound **1** was compared DMSO controls across 254 features normalized with features represented on a scale from -1 (blue, below DMSO control) to +1 (yellow, above DMSO control) with black indicating no difference between compound and DMSO for that feature. (*A*) Compounds were clustered by similarity to a reference library of 2035 bioactive compounds with known mechanisms of action. (*B*) Cytological profiling was used to compare compound **1** to a smaller, follow-up set of 25 reference compounds at three concentrations with known mechanisms of action.

1.4.10 Supplementary Table 1-S1: IC_{50} confidence intervals

HCT 116 cell proliferation IC_{50} data 95% confidence intervals for data reported in Table 1 of the main text.

	Compound	IC50 (μM)	95% confidence interval (μM)
	1	0.1	0.07 - 0.3
	2	35*	N.D.
Alanine Scan	VK03 (**3**)	> 75	N/A
	VK04 (**4**)	> 75	N/A
	VK05 (**5**)	> 75	N/A
	VK06 (**6**)	> 75	N/A
	VK07 (**7**)	> 75	N/A
	VK08 (**8**)	> 75	N/A
	VK09 (**9**)	> 75	N/A
N-methyl Scan	VK10 (**10**)	8	4.5 - 14
	VK11 (**11**)	0.5	0.3 - 0.9
	VK12 (**12**)	28*	N.D.
	VK13 (**13**)	> 75	N/A
	VK14 (**14**)	17	9.0 - 33
	VK15 (**15**)	> 75	N/A
	VK16 (**16**)	37	10 - 118
	VK17 (**17**)	> 75	N/A
Halogen and Heterocyclic Substitutions	VK19 (**19**)	11	5.4 - 24
	VK20 (**20**)	34	8.6 - 445
	VK21 (**21**)	3.6	2.3 - 5.6
	VK22 (**22**)	1.8	0.8 - 4.3
	VK23 (**23**)	> 75	N/A
	VK24 (**24**)	> 75	N/A
	VK25 (**25**)	28*	20 - N.D.
	VK26 (**26**)	3.4	0.7 - 140
	VK27 (**27**)	0.3	0.09 - 0.8
	VK28 (**28**)	0.2	0.1 - 0.4
	VK29 (**29**)	0.5	0.3 - 1.0
	VK30 (**30**)	0.1	0.009 - 0.5
	VK31 (**31**)	> 75	N/A
	VK32 (**32**)	> 75	N/A
	VK33 (**33**)	19*	N.D.
	VK34 (**34**)	0.3	0.2 - 0.6
	VK35 (**35**)	28	5.8 - N.D.
Other Substitutions	VK36 (**36**)	2.9	1.0 - 7.5
	VK37 (**37**)	> 75	N/A
	VK38 (**38**)	5.6	1.7 - 18
	VK39 (**39**)	--	N/A
	VK40 (**40**)	1.5	0.5 - 4.2
	VK41 (**41**)	28*	N.D.
	VK42 (**42**)	22	9.0 - 66

***** estimated graphically
N/A = not appliable
N.D. = not determined

1.4.11 Supplementary Table 1-S2: NCI60 COMPARE results

The effect of cordyheptapeptide A (**1**) on 59 cancer cell lines was compared to NCI60 "Standard Agents" based on growth inhibition. The compounds with the top ten average correlations to cordyheptapeptide A are shown here.

Compound	Mechanism of action	COMPARE correlation to **1**
phyllanthoside	Protein synthesis inhibitor[28]	0.80
vinorelbine tartrate	Microtubule poison[53]	0.73
bisantrene hydrochloride	DNA intercalating agent[54] DNA binding agent,[55]	0.70
chromomycin A3	RNA synthesis inhibitor[56]	0.70
actinomycin D	DNA intercalating agent[57]	0.68
bruceantin	Protein synthesis inhibitor[58]	0.66
vinblastine sulfate	Microtubule poison[59]	0.62
didemnin B	Protein synthesis inhibitor[60]	0.61
paclitaxel	Microtubule poison[61]	0.61
echinomycin	DNA intercalating agent[62]	0.60

1.4.12 Supplementary Table 1-S3: Log $D_{(dec/w)}$ and LPE Results

Table showing Log $D_{(dec/w)}$ and LPE for the natural poducts **1** and **2** and their Sar4-to-Gly derivatives VK-11, and VK-14.

Compound	$LogD_{(dec/w)}$	ALogP	LPE
cordyheptapeptide A (**1**)	1.3	4.2	2.3
VK11 (**11**)	0.8	4.0	2.0
cordyheptapeptide B (**2**)	3.0	4.5	3.7
VK14 (**14**)	2.9	4.3	3.9

1.4.13 Supplementary Table 1-S4: Gly and Sar derivative bioactivity

Comparison of the between glycine-containing cordyheptapeptide derivatives and their sarcosine-substituted matched pairs

Matched pair	Gly-containing compound		Sar-containing compound		Fold improvement in IC_{50}
	Coumpound number	IC_{50} (μM)	Coumpound number	IC_{50} (μM)	
1	VK27 (**27**)	0.3	VK19 (**19**)	11	37
2	VK28 (**28**)	0.2	VK20 (**20**)	34	170
3	VK29 (**29**)	0.5	VK21 (**21**)	3.6	7
4	VK30 (**30**)	0.14	VK22 (**22**)	1.8	13
5	VK33 (**33**)	19	VK25 (**25**)	28	1
6	VK34 (**34**)	0.3	VK26 (**26**)	3.4	11

1.5 Methods

1.5.1 General Synthetic Information and Procedures

Commercially available chemicals were used without further modification unless otherwise specified. Reagents and solvents were purchased from Fisher Scientific. HATU was purchased from Combi-Blocks or Chem-Impex. Amino acids were purchased from Combi-Blocks, Chem-Impex, Sigma-Aldrich, or Oakwood unless otherwise stated. Piperidine and 1,9-decadiene were purchased from Spectrum Chemical and TCI Chemicals, respectively. Polystyrene 2-chlorotrityl resin was purchased from Rapp-Polymere (H10033). Antibodies were purchased from Thermo Scientific and Sigma Aldrich. LCMS compound purity analysis was performed on a Thermo Scientific Ultimate 3000 UPLC system with a Thermo Scientific OrbiTrap VelosPro mass spectrometer. This system used a Thermo Hypersil GOLD C18 (30 mm x 2.1 mm, 1.9 µm particle size) column eluting in water/acetonitrile with 0.1% formic acid. NMR spectra were recorded on a Bruker Advance III HD 500 MHz 5 mm BBO Smart Probe in $CDCl_3$.

1.5.2 Statistical Parameters

Unless otherwise noted, replicates were reported as averages with errors listed as standard deviations. Error for IC_{50} calculations are reported as a 95% confidence interval.

1.5.3 Cell Culture

Unless otherwise noted cell culture was performed using Dulbecco's Modified Eagle's medium (DMEM) (Corning 10013CVMP) supplemented with 10% fetal bovine serum (FBS) (Corning 35-015-CV) and penicillin (100 U/mL)/streptomycin (100 µg/mL) (Gibco, ThermoFisher Scientific 15070063). Cells were maintained with phosphate-buffer saline (PBS) (Gibco 14190235) and a solution of 0.25% trypsin with 2.21 mM EDTA (Corning 25-053-CI). All cells were cultured at 37°C in a 5% CO2 atmosphere. For mutant cell line proliferation used in target ID, HCT116 (ATCC, Manassas, VA) and HCT116-417 (mutant) cells were maintained in McCoy's 5A media (Gibco, Grand Island, NY) supplemented with 10% fetal bovine serum (Axenia Biologix, Dixon, CA), 100 units/mL penicillin, and 100 ug/mL streptomycin (Gibco).

1.5.4 Synthetic Methods

1.5.4.1 Solid phase peptide synthesis (SPPS)

Solid phase peptide synthesis (SPPS) was carried out using standard Fmoc chemistry on pre-loaded 2-chlorotrityl resin (loading values: 0.8 – 1.2 mmol/g). Manual SPPS was performed using a Boekel Scientific Shake 'N' Bake shaking incubator (model: 136400) at 50 °C. Automatic SPPS was performed using a Gyros Protein Technologies Prelude X Multiple Synthesizer with induction heating.

1.5.4.2 Manual SPPS amino acid coupling

Four eq of Xaa, 8 eq of DIPEA, and 4 eq of HATU were dissolved in DMF and added to the resin. The reaction was shaken for 60 minutes at 50 °C, then drained. The resin was washed with DMF (3 x 2 mL), 3:1 DCM:DMF (3 x 2 mL), and DCM (3 x 2 mL). Reaction completion was monitored by LCMS.

1.5.4.3 Manual SPPS Fmoc deprotection

A solution of 2% DBU and 2% piperidine in DMF was added to the resin, then drained. The solution was added to the resin again. The reaction was shaken for 15 minutes at 50 °C, then drained. The resin was washed with DMF (3 x 2 mL), 3:1 DCM:DMF (3 x 2 mL), and DCM (3 x 2 mL). Reaction completion was monitored by LCMS.

1.5.4.4 Automatic SPPS amino acid coupling

Four eq of Xaa, 6 eq of DIPEA, and 3.8 eq of COMU were added to the resin in DMF. The reaction was shaken for 10 minutes at 90 °C, then drained. The resin was washed with DMF (4 x 2 mL), DCM (3 x 2 mL), and DMF (2 x 2 mL).

1.5.4.5 Automatic SPPS N-terminal capping

The resin was capped between amino acids during automatic SPPS. A solution of 10% DIPEA and 10 % acetic acid in DMF was added to the resin. The resin was shaken for 5 minutes at room temperature, then drained. The resin was washed with DMF (6 x 2 mL), DCM (3 x 2 mL), and DMF (2 X 2 mL).

1.5.4.6 Automatic SPPS Fmoc deprotection

A solution of 2% DBU and 2% piperidine in DMF was added to the resin. The reaction was shaken for 2 minutes at 90 °C, then drained. This step was repeated twice. The resin was washed with DMF (4 x 2 mL), DCM (3 x 2 mL), and DMF (2 X 2 mL).

1.5.4.7 On-resin addition of an N-terminal TFA protecting group

The resin was rinsed with DCM (3 x 2 mL). A solution of 12 eq of DBU and 10 eq of ethyl trifluoroacetate in THF was added to the resin containing an n-terminal deprotected amino acid. The reaction was shaken at room temperature for 60 minutes, then drained. The resin was washed with DMF (3 x 2 mL), 3:1 DCM:DMF (3 x 2 mL), and DCM (3 x 2 mL). Reaction completion was monitored by LCMS.

1.5.4.8 N-alkylation using Mitsunobu conditions

The resin was rinsed with dry THF under inert conditions (3 x 3 mL). Five eq of triphenylphosphine were dissolved in dry THF and added to the resin. Ten eq of dry methanol was then added. Five eq of DIAD (stored over molecular sieves) was added dropwise to the resin mixture while vortexing the reaction tube. Reaction completion was monitored by LCMS and repeated till no more starting material was observed. The resin was washed with DMF (3 x 2 mL), 3:1 DCM:DMF (3 x 2 mL), and DCM (3 x 2 mL).

1.5.4.9 On-resin removal of an N-terminal TFA protecting group

A 1:1 solution of dry EtOH and dry THF was added to the resin. Twenty eq of sodium borohydride was added to the resin solution. The mixture was left open to the atmosphere for 60 minutes, stirring occasionally. Reaction completion was monitored by LCMS and repeated till no more starting material was observed. The resin was washed with methanol (3 x 3 mL), DMF (3 x 2 mL), 3:1 DCM:DMF (3 x 2 mL), and DCM (3 x 2 mL).

1.5.4.10 Peptide cleavage from resin

The finished linear peptide was cleaved off resin using three resin volumes of 30% HFIP in DCM, shaking for 60 minutes, after which the solution was collected. The resin was rinsed twice with 2 mL of DCM, collected in the same vial. The cleave was repeated allowing the reaction to shake for 30 minutes. Followed by another DCM rinse. The HFIP/DCM mixture was removed under N_2. The cleaved peptide was then dissolved in acetone and evaporated under reduced pressure (x2) to remove any residual HFIP.

1.5.4.11 Cyclization with COMU

The linear peptides were dissolved in 5 mL ACN, then 5 eq of DIPEA was added. This mixture was added dropwise to a solution of ACN with 1.2 eq of COMU. The final concentration of this reaction mixture was 1 mg of linear peptide per mL of ACN. The reaction was stirred at room temperature for 16 -24 hours until the reaction was

complete, monitored by LCMS. The solution was then evaporated under reduced pressure to prepare the crude cyclic peptide for purification.

1.5.4.12 Cyclization with PyBOP and HOAt

The linear peptides were dissolved in 5 mL ACN, then 10 eq of DIPEA was added. This mixture was added dropwise to a solution of ACN with 4 eq of PyBOP and 5 eq of HOAt. The final concentration of this reaction mixture was 1 mg of linear peptide per mL of ACN. The reaction was stirred at room temperature for 16 -24 hours until the reaction was complete, monitored by LCMS. The solution was then evaporated under reduced pressure to prepare the crude cyclic peptide for purification.

1.5.4.13 Purification of peptides

The cyclized peptides were purified on a Biotage Isolera Prime system using a KP-C18-Ultra 30g column eluting with ACN/H_2O with 0.1% TFA. Fractions were checked for purity by LCMS, then evaporated under reduced pressure.

1.5.4.14 Removal of acid-labile side chain protecting groups

The peptides were dissolved in a solution of 95:5 TFA:TES and allowed to react for 60 minutes at room temperature, stirring occasionally. The solution was removed under N_2. The cleaved peptide was then dissolved in acetone and evaporated under reduced pressure (x2) to remove any residual TFA.

1.5.5 Cell proliferation

Adherent HCT 116 cells were plated at a density of 1,200 cells/well at 30 μL per well in a 384-well plate (Greiner Bio-One 781090) and allowed to adhere overnight. The following day, compounds were added using a PerkinElmer Janus MDT pinning robot from a 384-well DMSO compound stock plate giving a range of concentrations in the assay plate from 4.5 nM to 75 μM in 1:2 dilutions. The cells were incubated for 72 hours at 37 °C and 5% CO_2. Five μL of a 0.7 mg/mL resazurin stock solution (sodium salt, in PBS) was added using an Finstruments Multidrop 384 automatic dispenser for a final resazurin concentration of 0.1 mg/mL per well. The cells were for an additional 90 minutes at 37 °C and 5% CO_2, after which a Perkin Elmer EnVision plate reader (2103 Multilabel Reader) was used to quantify fluorescence. Cell proliferation fluorescence was normalized to the average fluorescence of control cells incubated with DMSO in each assay plate. This assay was performed in biological and technical triplicate (n = 3) taken from distinct samples. IC_{50} values were calculated with GraphPad Prism (v 8.4.1) except for compounds **2**, **12**, **25**, **33**, and **41** which were estimated graphically.

1.5.6 Cytological profiling (CP)

Cytological profiling (CP)[30,31] was performed following the procedure described in Schulze *et* al[32] with the following modifications. Adherent HeLa cells were plated at a density of 3,600 cells/well at 100 μL per well in a 384-well plate and allowed to adhere overnight. The following day, compounds were added using a

PerkinElmer Janus MDT pinning robot from a compound stock plate giving a range of concentration in the 384-well plate (Greiner Bio-One 781090) from 4.5 nM to 75 μM. The cells were incubated with compound for 18 hours at 37 °C and 5% CO_2. Prior to fixing and staining, cells were incubated with 20 μM EdU in DMEM for 45 minutes at 37 °C. Cells were fixed in a 4% formaldehyde in PBS for 20 minutes at room temperature. Cells were washed three times with a BioTek automatic plate washer. Next, to permeabilize the cell membrane, cells were incubated in PBS with 0.5% Triton X for 10 minutes followed by another set of three washes on the plate washer.

To prepare cells for antibody treatment, they were blocked with a solution of 2% BSA in PBS for 20 minutes at room temperature then washed again. Click chemistry was performed by adding a master mix of click reagents to the cells (4mM $CuSO_4$, 2mg/mL sodium ascorbate, and 0.1 mM rhodamine azide in 100 mM Tris buffer pH 7.4). This master mix was made just before adding it to the cells which were then incubated for one hour at room temperature in the dark. The cells were washed, and the appropriate primary antibodies (rabbit anti-phospho histone-H3 (9H12L10, ThermoFisher Scientific 701258) or FITC-anti-tubulin (Sigma Aldrich F2043)) were added in PBS with 2% BSA and incubated at 4 °C for 18 hours. The following day, the cells were washed with PBS to remove excess primary antibody. Then the secondary antibody (chicken anti-rabbit Alexa Fluor 647, ThermoFisher scientific A-21443) was added along with rhodamine-phalloidin and Hoechst 33342 (Anaspec Inc 83218) stain in PBS with 2% BSA. The cells were incubated for 1 hour at room temperature in the dark. Cells were washed with PBS on the automatic plate washer. After which, the

plates were imaged using an ImageXpressMicro XLS epifluorescent microscope (Molecular Devices). The resulting images were analyzed using MetaXpress (v. 6.2.1.704, Molecular Devices) and built-in multi-wave cell scoring. These values were converted to feature scores and clustered and analyzed using Cytoscape (v 3.7.0, Pearson absolute correlation, maximum linkage). (n = 1)

1.5.7 Bioorthogonal Non-canonical Amino Acid Tagging (BONCAT) Assay

Adherent HeLa cells were plated at a density of 10,000 cells in 100 μL DMEM per well in 96-well plates and allowed to adhere overnight. Compounds were added as DMSO stocks dissolved in DMEM in 1:3 dilutions in concentrations from 34 nM to 75 μM for **1** and 9 nM to 22 μM for didemnin B, a positive control. DMSO was used as a negative control.

For the 24 hr time point, cells were incubated with compounds for 24 hours in DMEM at 37 °C and 5% CO_2. The media was removed and replaced with 100 μL of media without methionine ("DMEM –Met") (Thermo Fisher catalog #21013024, with added 1mM sodium pyruvate, 0.584 g/L glutamine, penicillin (100 U/mL)/streptomycin (100 μg/mL)) was added to each well. The plates were incubated for 1 hr at 37 °C and 5% CO_2. After this, 10 μL of 1 mM L-homopropargylglycine (L-HPG) in DMEM –Met was added to each well. Again, the plates were incubated for 1 hour 37 °C and 5% CO_2. For the 3 hr time point, cells were incubated with compound for 1.5 hours in "DMEM –Met" at 37 °C and 5% CO_2. After this, 10 μL of 1 mM of L-HPG in DMEM -Met was added to each well and incubated for another 1.5 hours 37

°C and 5% CO_2. This assay was performed in technical triplicate (n = 3) taken from distinct samples.

For both time points, cells were fixed in a 4% formaldehyde in PBS for 20 minutes at room temperature. Cells were washed three times with a BioTek automatic plate washer. Next, to permeabilize the cell membrane, cells were incubated in PBS with 0.5% Triton X for 10 minutes followed by another set of three washes on the plate washer. To prepare cells for antibody treatment, they were blocked with a solution of 2% BSA in PBS for 20 minutes at room temperature then washed again. Click chemistry was performed by adding a master mix of click reagents to the cells (4mM $CuSO_4$, 2mg/mL sodium ascorbate, and 0.1 mM rhodamine azide in 100 mM Tris buffer pH 7.4). This master mix was made just before adding it to the cells which were then incubated for one hour at room temperature in the dark, then washed. After which, Hoechst was added in PBS with 2% BSA. Cells were incubated for 20 minutes in the dark at room temperature, and then washed. After which, the plates were imaged using an ImageXpressMicro XLS epifluorescent microscope (Molecular Devices). The resulting images were analyzed using MetaXpress (v. 6.2.1.704, Moleular Devices) and built-in multi-wave cell scoring.

1.5.8 NCI COMPARE Analysis

Testing was performed by the Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute (http://dtp.cancer.gov). Cordyheptapeptide A (NSC number: S812201) was sent to the Developmental

Therapeutics Program (DTP) at the National Cancer Institute (NCI) to be analyzed in their human tumor cell line assay. Details of this assay protocol can be found in related publications on the DTP website.[25,26,63] Briefly, 59 human tumor cell lines were screened against **1** and analyzed for total growth inhibition (TGI), GI50 (concentration of **1** that causes 50% of growth inhibition), and LC50 (concentration of **1** at which half of the cells are killed). Using the NCI60 COMPARE algorithm, the GI50 of **1** in these cell lines was compared to the GI50 of the DTP's list of "Standard Agents" using the following parameters: minimum correlation (0.2), count results to return (50), minimum count common cell lines (40), and minimum standard deviation (0.05). The compounds were rank ordered by correlation from highest to lowest. If duplicate reference compounds were in the top 50 results, the average correlation is reported. The top 10 compounds are reported here with their COMPARE correlation value to **1** and a brief description of their mechanism of action. GI50 was chosen for follow-up studies because **1** was more highly correlated to other compounds from the "Standard Agent" set by this metric than either TGI or LC50.

1.5.9 Photo-crosslinking pull down assay

A cordyheptapeptide derivative (**43**) was synthesized using the above methods including an Fmoc-L-propargylglycine and an Fmoc-L-photo-proline (Iris Biotech GMBH) amino acids. (n = 1).

1.5.10 Cell proliferation in mutant eEF1a cell line

HCT116 or HCT116-417 cells were briefly trypsinized and repeatedly pipetted to produce a homogenous cell suspension. 2,500 cells were seeded in 100 uL complete growth media per well in 96-well clear-bottom plates. After allowing cells to grow/adhere overnight, cells were treated with 25 uL/well 5x drug stocks (0.1% DMSO final) and incubated for 72 hours. AlamarBlue (Life Technologies, Grand Island, NY) was used to assess cell viability per the manufacturer's instructions. Briefly, 12.5 uL alamarBlue reagent was added to each well, and plates were incubated at 37°C. Fluorescence intensity was measured every 30 min to determine the linear range for each assay (Ex 545 nm, Em 590 nm, SPARK, Tecan Austria GmbH, Austria). Proliferation curves were generated by first normalizing fluorescence intensity in each well to the DMSO-treated plate average. Normalized fluorescence intensity was plotted in GraphPad Prism (GraphPad, La Jolla, CA), and IC_{50} values were calculated from nonlinear regression curves. The reported IC_{50} values represent the average of at least three independent determinations (n = 3, ±SD).

1.5.11 Parallel artificial membrane permeability assay (PAMPA)

PAMPA[64,65] was used to determine the passive permeability of these compounds with modifications and calculations as described in Naylor *et al*.[7] and here. Replicates were taken from distinct samples (n = 4). PAMPA was carried out in a 96-well donor plate (Millipore MAIPNTR10) and a 96-well Teflon acceptor plate (Millipore MSSACCEPTOR). Analyte solutions were prepared from a 200 μM DMSO

compound stock solution, 200 μM propranolol in DMSO (control), and PBS pH 7.4 for a final concentration of 1 μM compound, 1 μM propranolol, and 5% DMSO by volume in PBS. The membrane of each donor well was prepared by adding 5 μL of 1% (w/v) lecithin (soybean, 90%) in n-dodecane to the underside of the membrane surface. Plate was then left to sit for 5 minutes to allow the lipids to adsorb to the membrane. The sample plates were prepared by adding 300 μL of 5% DMSO in PBS buffer (v/v) to the acceptor plate. Using a multichannel pipette, 150 μL of the 1 μM analyte solutions were added to the donor plate. The donor plate was then slowly lowered into the acceptor plate, avoiding bubbles. The assembly was placed in a dark, humid container and allowed to sit for 14 – 16 hours. The plates were separated, recording the time. Samples were prepared for LCMS by adding 50 μL of solution from the plate to 50 μL of methanol in a 96-well plate (Coring, COSTAR 3357). Relative concentrations of the samples were analyzed by LCMS. Integration and ion counts were calculated using automatic analysis of LCMS data using a lab-developed python program. PAMPA permeability was calculated using the following parameters and formulae:

Active surface area of membrane (mm^2): MSA = 240
Volume of acceptor well (μL): $V_A = 300$
Volume of donor well (μL): $V_D = 300$
Assay run time (s): T_S
Donor intensity: I_D
Acceptor intensity: I_A
Recovery intensity: I_R

[1] $Analyte\ Equlibrium$:

$$AnalyteEquil = \frac{(I_A * V_A) + (I_D * V_D)}{V_A + V_D}$$

[2] $Transmittane\ (discard\ data\ if\ T\ is > 0.95)$:

$$T = \frac{I_A}{AnalyteEquil}$$

[3] $\%\ Recovery$

$(low\ values\ represent\ material\ loss, ei.\ aggregation\ or\ adherence)$:

$$\%R = \frac{(I_A * V_A) + (I_D * V_D)}{I_R + V_D} * 100\%$$

[4] $Constant$:

$$C = \frac{(V_D * V_A)}{(V_D + V_A) * MSA * T_S}$$

[5] $Pe\ (x\ 10^{-6})\ after\ represented\ on\ log\ scale$:

$$Pe = (-C * ln\ ln\ (1 - T)\ ) * 10^{-6}$$

1.5.12 Thermodynamic solubility

From a 15 mM DMSO stock of the compound, 16.7 µL of compound was added to a 96 well plate. DMSO was evaporated under reduced pressure at 60 °C in a Genevac EZ-2 Plus centrifugal evaporator. To the evaporated film/solid, 125 µL of PBS, pH 7.4, was added to make a 2 mM solution. Plates were heat sealed, sonicated for 1 hour, and then shaken at 37 °C for 16 hours. The solution was filtered using a 0.7 µm glass fiber filter. Filtrate collected and diluted 1:10 in ACN. The solution was centrifuged at 1000x G for 10 minutes at room temperature and pressure. The supernatant was analyzed by LCMS, comparing the single ion chromatogram count to

a standard curve created using injections at known concentrations at 100uM, 10uM, 1uM, 0.1uM. Integrations were calculated using an in-lab developed peak integrating python program. (n = 1)

1.5.13 $LogD_{(dec/w)}$ shake flask partition coefficient assay

This assay was carried out following procedures described in Naylor *et al.*[7] with modification as described in Klein *et al*.[66] Replicates were taken from the same shake-flask experiment. Four samples were removed from each shake-flask experiment and four separate calculations were averaged (n = 4).

1.5.14 ALogP calculations

ALogP was calculated according to previously described methods.[67]

1.5.15 Lipophilic permeability efficiency (LPE) metric calculations

LPE was calculated following previously described procedures[7] using the following equation: LPE = LogD – 1.06(ALogP) + 5.47.

1.5.16 Temperature shift coefficient calculations

NMR temperature shift coefficient experiments were used to determine solvent exposure of amide bond -NH.[18,68-70] Peptide samples were dissolved in deuterated chloroform. A proton NMR was taken at temperatures of 300 K, 305 K, 310 K, 315 K, 320 K, and 323 K. Measurements were taken from the same sample following a period of thermal equilibration (n = 1). Samples were processed and analyzed using

MestReNova (v. 12.0.2). At each temperature, amide peaks were integrated, and the chemical shift and the coupling constant (*J* NH-HA) were calculated. Amide peaks were tracked by their coupling constant as the peaks shifted. The temperature shift coefficient was calculated using a linear regression of the change in chemical shift against the change in temperature. This value was then multiplied by 1000 to get values in ppb. Temperature shift coefficient < –4 ppb/K was considered solvent exposed.[18,68] In some cases, not at all amide peaks were located at each temperature due to overlap with other proton peaks in the NMR. The temperature shift coefficient for a given amide was only calculated if its proton peak couple be located in the proton NMR from at least two different temperatures.

1.5.17 Molecular dynamics simulations

The initial conformer was built using the Molecular Operating Environment (MOE)[1] protein builder, and LowModeMD[71] implemented in MOE was used to search for the most stable conformer in vacuo. PACKMOL[72] was used to solvate the system, where the numbers of water, chloroform, and cyclohexane molecules are 2000, 450, and 300, respectively. The AMBER ff03 force field was used for amino acids[23] and ForceField_NCAA was used for N-methylated amino acids.[24] The initially solvated system was minimized by steepest descent, followed by 50,000 steps of NVT calculation. The system was then equilibrated by NPT calculation at atmospheric pressure by applying the Berendsen barostat[73] for 500,000 steps. A positional restraint

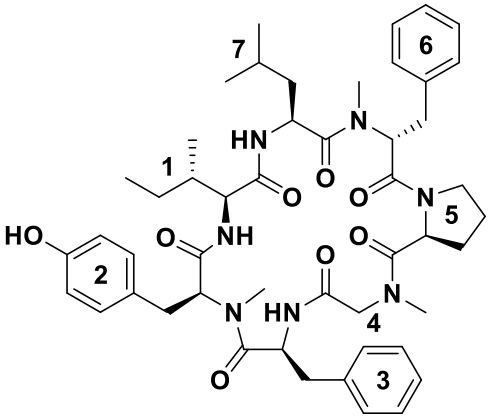
was imposed on Cα atoms during equilibration. The resultant system was used as an initial structure for the following simulations.

Virtual-system-coupled Trivial Trajectory Parallelization of Multicanonical Molecular Dynamics (TTP-V-McMD)[21,74] simulations were adopted to effectively sample conformations of cyclic peptides. A total of 336 pre-TTP-V-McMD runs were initiated with random velocities for each atom at T = 300 K. Then the systems were heated to T = 1503 K with 5,000 steps, followed by 495,000 steps at T = 1503 K to randomize the initial structures. Flat potential energy distributions between T=280 K and 1503 K were obtained by iterating TTP-V-McMD simulations for seven times with eight virtual states. The cut-off distance for Coulombic and van der Waals interactions was 1.0 nm, and PME was used to calculate the long-range electrostatic interaction. The NVT ensemble was used for all TTP-V-McMD simulations using the velocity rescaling method (Bussi thermostat).[75] The LINCS algorithm was used to constrain the bonds with a hydrogen atom, enabling a time step of 2.0 fs. For each solvent, 1.0×10^7 steps × 336 production runs (aggregating 6.72 μs) were performed. The structure and potential energy were stored every 2 ps. The virtual states were exchanged in every 5,000 steps. A resampling method was used to extract the canonical ensemble at T = 300 K, where structures were drawn from the simulated ensemble with relative probabilities according to their Boltzmann weights. Typically, ~ 20,000 conformers were obtained from the resampling method, and then 5,000 conformers were randomly selected to use for further cluster analysis. All 3.36 million conformers were taken into account to determine the free energy landscape by a potential of mean force (PMF)

calculation; $W = -k_B T \ln \rho$, where k_B is the Boltzmann constant and ρ is the density of state. An in-house implemented TTP-V-McMD using GROMACS version 5.1.4 was used for the simulations.

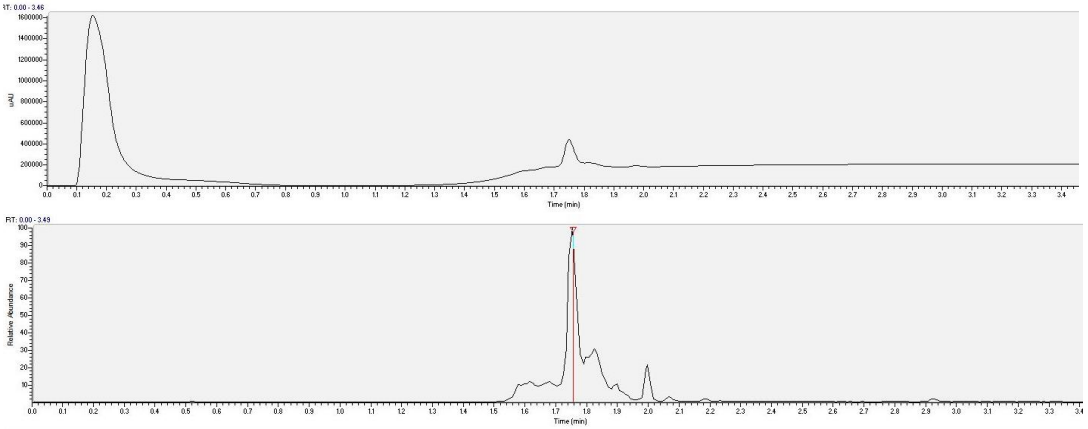
1.5.18 Spectra for synthesized compounds

1.5.18.1 NMR and LC/MS data for cordyheptapeptide A (**1**)

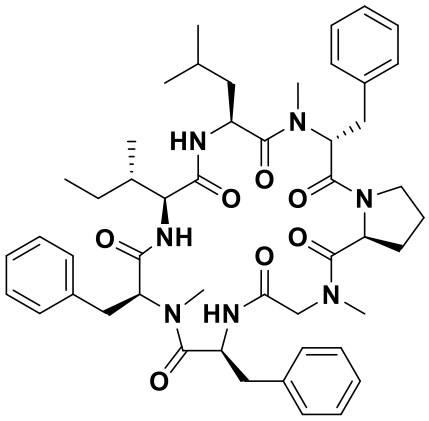


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.58 (d, *J* = 9.8 Hz, 1H), 8.19 (d, *J* = 9.0 Hz, 1H), 7.45 – 7.25 (m, 6H), 7.19 – 7.09 (m, 4H), 7.11 – 7.01 (m, 1H), 6.52 (d, *J* = 8.1 Hz, 2H), 6.24 (d, *J* = 8.2 Hz, 2H), 5.88 (d, *J* = 9.6 Hz, 1H), 5.56 (dd, *J* = 11.6, 4.8 Hz, 1H), 5.41 (d, *J* = 17.4 Hz, 1H), 5.36 (ddd, *J* = 11.6, 9.8, 3.9 Hz, 1H), 4.93 (ddd, *J* = 11.9, 9.1, 2.4 Hz, 1H), 4.45 (dd, *J* = 9.6, 3.1 Hz, 1H), 4.39 (dd, *J* = 9.3, 3.4 Hz, 1H), 3.77 (ddd, *J* = 12.0, 7.7, 4.3 Hz, 1H), 3.61 (dt, *J* = 11.8, 7.6 Hz, 1H), 3.43 – 3.25 (m, 3H), 3.17 (t, *J* = 12.0 Hz, 1H), 3.11 – 3.00 (m, 1H), 3.04 (s, 4H), 2.90 (s, 3H), 2.86 – 2.76 (m, 1H), 2.76 – 2.66 (m, 1H), 2.62 (s, 3H), 2.52 – 2.40 (m, 1H), 2.32 (dtt, *J* = 10.3, 6.8, 3.3 Hz, 1H), 2.08 – 2.00 (m, 1H), 1.95 – 1.85 (m, 1H), 1.85 – 1.77 (m, 1H), 1.56 (dtd, *J* = 10.4, 6.6, 3.5 Hz, 1H), 1.42 – 1.29 (m, 3H), 1.05 – 0.91 (m, 3H), 0.90 (dd, *J* = 6.8, 3.2 Hz, 7H), 0.84 (d, *J* = 6.5 Hz, 3H), 0.13 (ddd, *J* = 13.7, 10.7, 2.4 Hz, 1H); LCMS (*m/z*): calculated exact mass: 879.4895, experimental $[M+H]^+$: 880.4951.

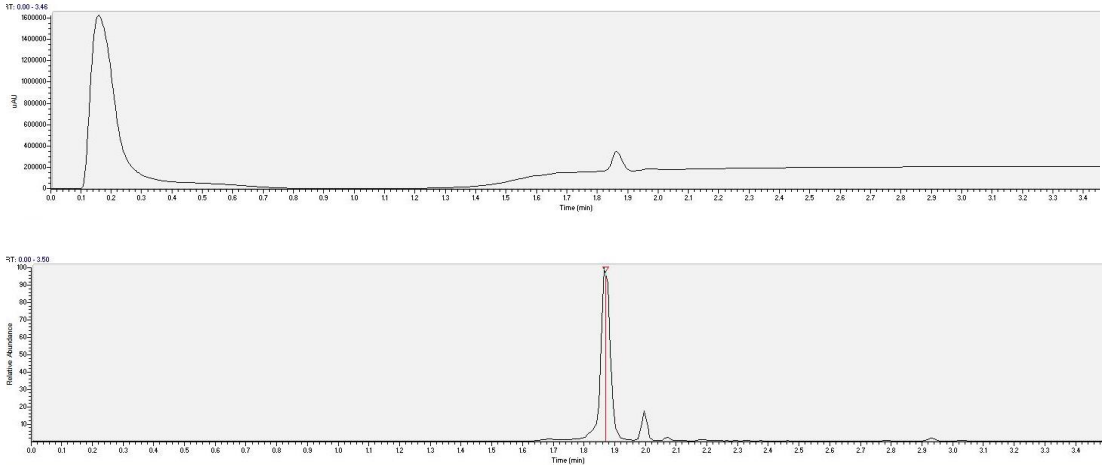


1.5.18.2 NMR and LC/MS data for cordyheptapeptide B (**2**)

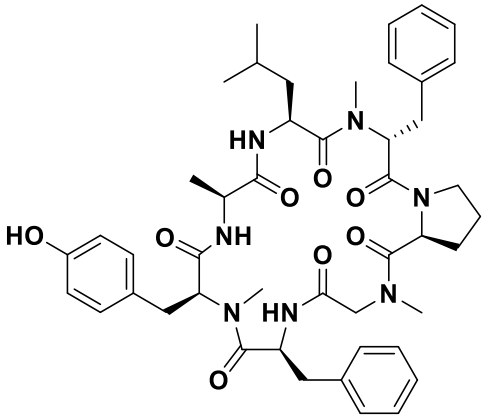


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe-MePhe-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.55 (d, *J* = 9.8 Hz, 1H), 8.19 (d, *J* = 9.1 Hz, 1H), 7.40 (t, *J* = 7.4 Hz, 2H), 7.32 (dq, *J* = 16.7, 9.2, 8.1 Hz, 4H), 7.20 (d, *J* = 29.7 Hz, 1H), 7.17 – 7.08 (m, 5H), 7.04 (t, *J* = 7.3 Hz, 3H), 6.43 (d, *J* = 7.4 Hz, 2H), 5.88 (d, *J* = 9.5 Hz, 1H), 5.55 (dd, *J* = 11.6, 4.8 Hz, 1H), 5.40 (d, *J* = 17.4 Hz, 1H), 5.33 (ddd, *J* = 12.9, 10.5, 3.8 Hz, 1H), 4.93 (td, *J* = 10.3, 9.4, 2.3 Hz, 1H), 4.45 (dd, *J* = 9.5, 3.0 Hz, 1H), 4.39 (dd, *J* = 9.3, 3.4 Hz, 1H), 3.77 (ddd, *J* = 12.1, 7.8, 4.3 Hz, 1H), 3.61 (dt, *J* = 11.9, 7.8 Hz, 1H), 3.49 (dd, *J* = 11.4, 4.4 Hz, 1H), 3.46 (s, 0H), 3.38 – 3.30 (m, 1H), 3.30 – 3.25 (m, 1H), 3.25 – 3.17 (m, 1H), 3.04 (s, 4H), 3.13 – 2.97 (m, 2H), 2.90 (s, 3H), 2.81 (ddd, *J* = 16.4, 12.3, 4.1 Hz, 2H), 2.65 (s, 3H), 2.47 (dq, *J* = 12.6, 9.5 Hz, 1H), 2.32 (tt, *J* = 7.1, 3.6 Hz, 1H), 2.17 (s, 1H), 2.04 (ddt, *J* = 11.8, 7.1, 4.0 Hz, 1H), 1.90 (qd, *J* = 7.1, 2.8 Hz, 1H), 1.80 (q, *J* = 10.8, 9.9 Hz, 1H), 1.57 (td, *J* = 6.7, 3.6 Hz, 1H), 1.44 – 1.24 (m, 3H), 0.98 (dt, *J* = 10.7, 6.7 Hz, 1H), 0.92 (dd, *J* = 9.8, 6.6 Hz, 6H), 0.88 (d, *J* = 6.8 Hz, 3H), 0.86 – 0.69 (m, 5H), 0.14 (dd, *J* = 14.2, 11.0 Hz, 1H); LCMS (*m/z*): calculated exact mass: 863.4945, experimental $[M+H]^+$: 864.5011.

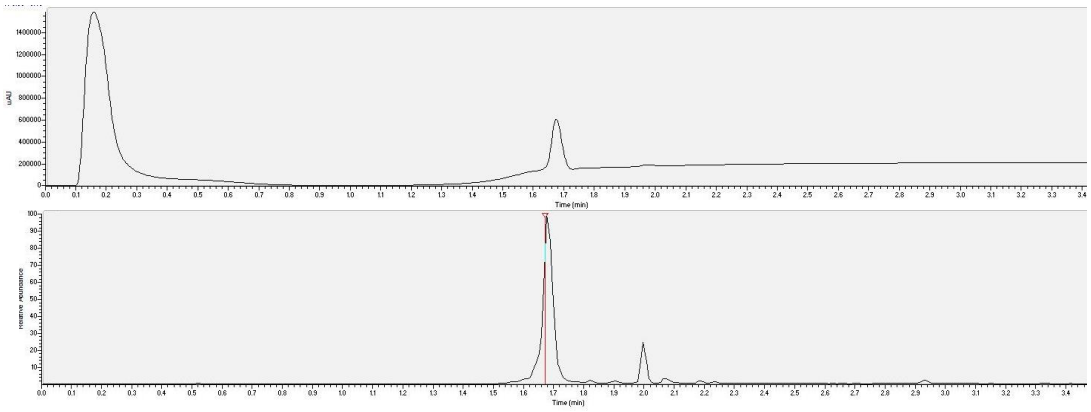


1.5.18.3 NMR and LC/MS data for VK-03 (**3**)

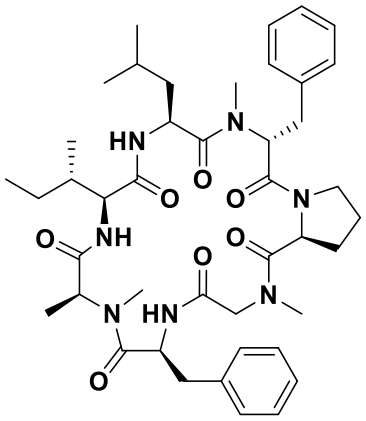


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe-MeTyr-Ala)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.61 (d, *J* = 9.8 Hz, 1H), 8.17 (d, *J* = 9.1 Hz, 1H), 7.46 – 7.31 (m, 6H), 7.21 – 7.13 (m, 4H), 7.07 (t, *J* = 7.0 Hz, 1H), 6.52 (d, *J* = 8.2 Hz, 2H), 6.21 (d, *J* = 8.3 Hz, 2H), 5.85 (d, *J* = 8.6 Hz, 1H), 5.57 (dd, *J* = 11.6, 4.8 Hz, 1H), 5.43 (d, *J* = 17.4 Hz, 1H), 5.39 – 5.31 (m, 1H), 4.92 (t, *J* = 10.1 Hz, 1H), 4.53 (p, *J* = 7.6 Hz, 1H), 4.42 (dd, *J* = 9.2, 3.4 Hz, 1H), 3.80 (ddd, *J* = 11.8, 7.6, 4.3 Hz, 1H), 3.63 (dt, *J* = 11.9, 7.8 Hz, 1H), 3.39 (d, *J* = 8.0 Hz, 1H), 3.36 (d, *J* = 3.0 Hz, 1H), 3.32 (dd, *J* = 15.1, 11.6 Hz, 1H), 3.18 (t, *J* = 12.0 Hz, 1H), 3.11 – 2.98 (m, 1H), 3.06 (s, 4H), 2.95 (s, 3H), 2.86 – 2.70 (m, 3H), 2.62 (s, 3H), 2.53 – 2.43 (m, 1H), 2.11 – 2.03 (m, 1H), 1.92 (s, 1H), 1.83 (dq, *J* = 18.7, 8.1 Hz, 1H), 1.47 (d, *J* = 7.3 Hz, 3H), 1.42 – 1.32 (m, 1H), 0.90 (dd, *J* = 21.9, 6.5 Hz, 6H), 0.25 – 0.16 (m, 1H); LCMS (*m/z*): calculated exact mass: 837.4425, experimental $[M+H]^{+}$: 838.4487.

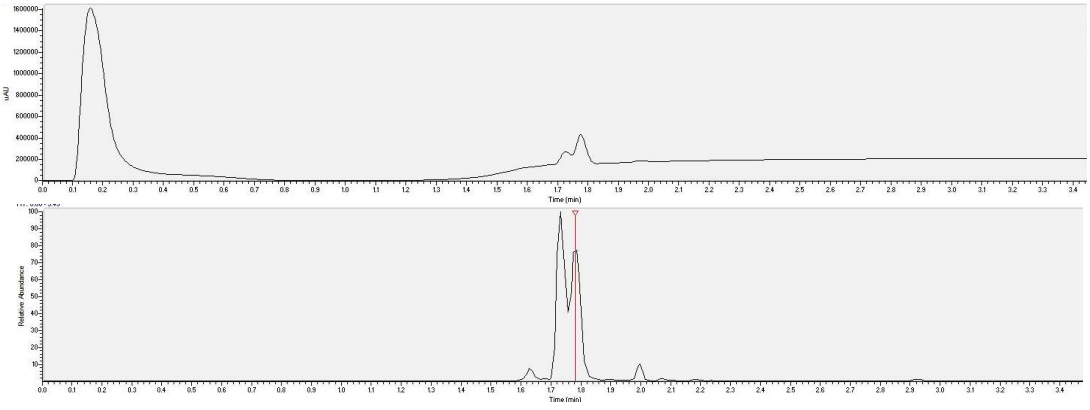


1.5.18.4 NMR and LC/MS data for VK-04 (**4**)

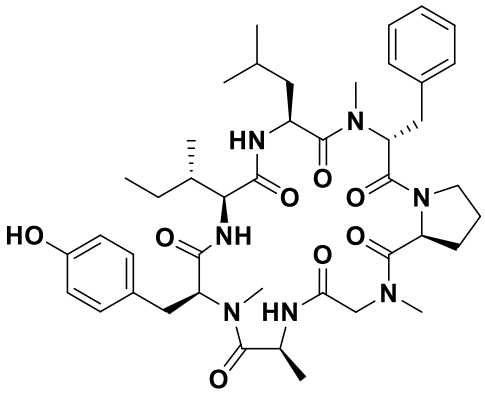


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe-MeAla-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.54 (d, *J* = 9.8 Hz, 1H), 8.34 (s, 0H), 8.27 (d, *J* = 9.1 Hz, 1H), 7.35 – 7.28 (m, 1H), 7.28 – 7.00 (m, 9H), 5.89 (d, *J* = 9.6 Hz, 1H), 5.59 (dd, *J* = 11.5, 4.9 Hz, 1H), 5.48 – 5.35 (m, 2H), 4.92 (ddd, *J* = 12.0, 9.1, 2.4 Hz, 1H), 4.43 (ddd, *J* = 12.6, 9.3, 3.2 Hz, 2H), 3.79 (ddd, *J* = 11.9, 7.7, 4.3 Hz, 1H), 3.63 (dt, *J* = 11.7, 7.7 Hz, 1H), 3.47 (q, *J* = 6.9 Hz, 1H), 3.40 (d, *J* = 17.3 Hz, 1H), 3.29 (dd, *J* = 15.1, 11.5 Hz, 1H), 3.19 (s, 2H), 3.11 – 2.98 (m, 5H), 2.94 (s, 3H), 2.83 – 2.74 (m, 1H), 2.70 – 2.57 (m, 1H), 2.48 (ddd, *J* = 12.6, 9.7, 7.0 Hz, 1H), 2.07 (dq, *J* = 12.7, 3.7 Hz, 1H), 1.91 (td, *J* = 7.6, 3.9 Hz, 1H), 1.88 – 1.78 (m, 1H), 1.50 (s, 1H), 1.38 (d, *J* = 7.0 Hz, 3H), 1.34 (s, 1H), 1.34 – 1.26 (m, 1H), 1.07 – 0.96 (m, 1H), 1.00 (s, 1H), 0.99 – 0.75 (m, 13H), 0.69 (s, 1H), 0.44 (dd, *J* = 14.9, 6.9 Hz, 1H), 0.10 (ddd, *J* = 14.0, 10.8, 2.5 Hz, 1H); LCMS (*m/z*): calculated exact mass: 787.4632, experimental $[M+H]^{+}$: 788.4703.

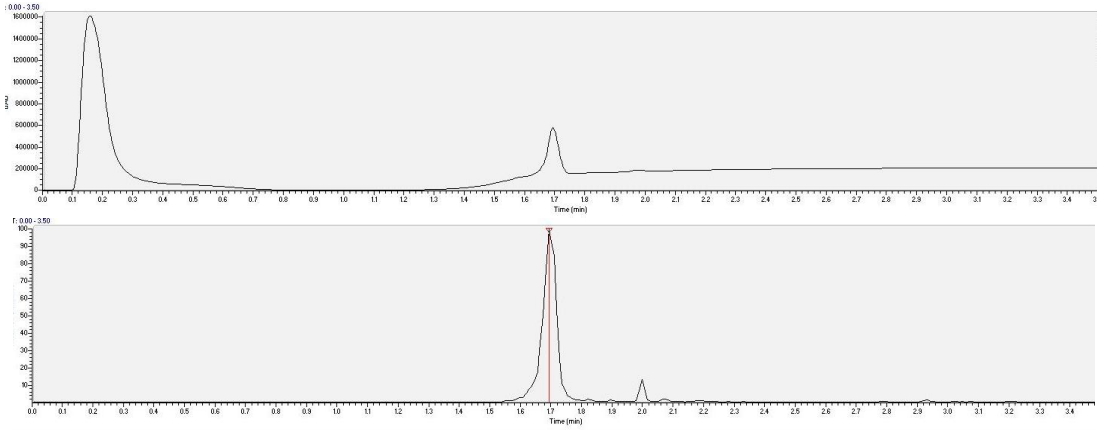


1.5.18.5 NMR and LC/MS data for VK-05 (**5**)

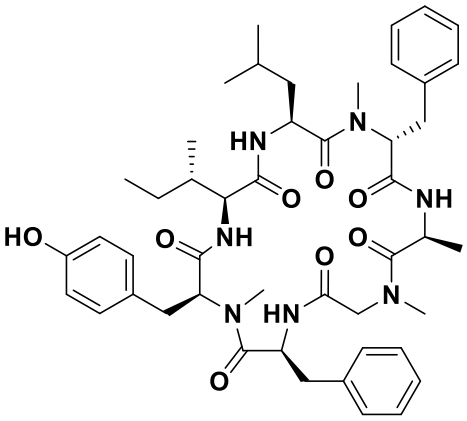


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Ala-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.36 (d, *J* = 9.9 Hz, 1H), 7.89 (d, *J* = 9.2 Hz, 1H), 7.24 – 7.11 (m, 4H), 7.05 (d, *J* = 8.3 Hz, 2H), 6.87 – 6.81 (m, 2H), 5.96 (d, *J* = 9.1 Hz, 1H), 5.49 (dd, *J* = 11.2, 5.0 Hz, 1H), 5.39 (d, *J* = 17.5 Hz, 1H), 5.25 – 5.14 (m, 1H), 4.94 – 4.86 (m, 1H), 4.76 (s, 1H), 4.40 (td, *J* = 8.6, 3.2 Hz, 2H), 3.77 (ddd, *J* = 12.0, 7.7, 4.5 Hz, 1H), 3.71 (dd, *J* = 9.7, 6.5 Hz, 1H), 3.61 (dt, *J* = 11.8, 7.7 Hz, 1H), 3.43 (dd, *J* = 13.3, 9.7 Hz, 1H), 3.35 (d, *J* = 17.5 Hz, 1H), 3.26 (dd, *J* = 15.0, 11.2 Hz, 1H), 3.21 – 3.11 (m, 1H), 3.13 – 3.04 (m, 1H), 3.03 (s, 3H), 2.98 (s, 3H), 2.94 (s, 3H), 2.71 (s, 1H), 2.52 – 2.43 (m, 1H), 1.30 (dd, *J* = 11.7, 7.3 Hz, 1H), 1.26 – 1.14 (m, 4H), 0.98 – 0.87 (m, 4H), 0.79 (dd, *J* = 18.5, 6.8 Hz, 8H), 0.09 (d, *J* = 12.4 Hz, 1H); LCMS (*m/z*): calculated exact mass: 803.4582, experimental $[M+H]^{+}$: 804.4645.

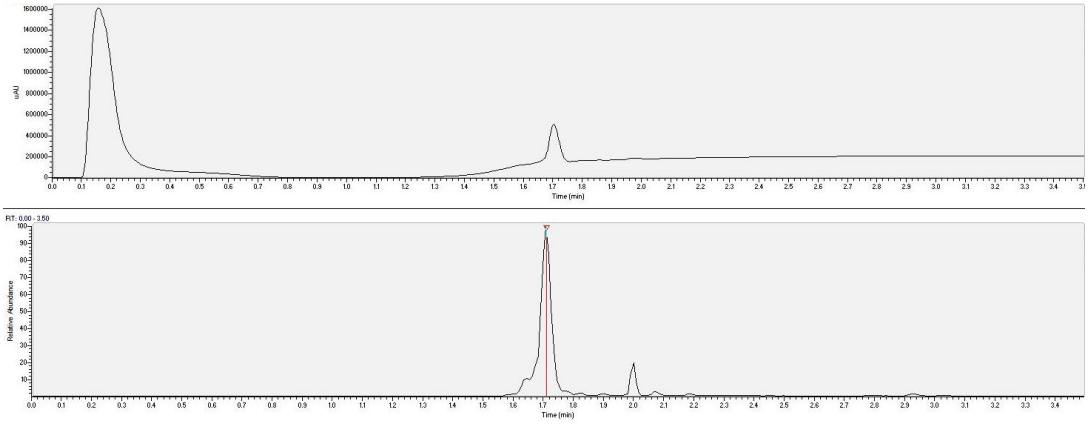


1.5.18.6 NMR and LC/MS data for VK-06 (**6**)

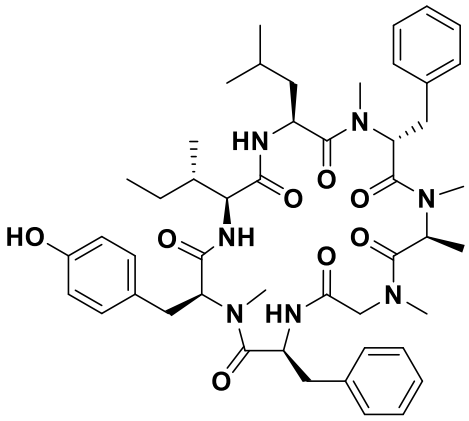


Sequence: Cyclo-(Leu-MeDPhe-Ala-Sar-Phe-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.52 (d, *J* = 9.9 Hz, 1H), 8.26 (d, *J* = 9.1 Hz, 1H), 7.43 (t, *J* = 7.4 Hz, 2H), 7.38 (d, *J* = 7.2 Hz, 1H), 7.32 (d, *J* = 7.5 Hz, 3H), 7.21 – 7.13 (m, 4H), 7.04 (t, *J* = 7.0 Hz, 1H), 6.69 (s, 1H), 6.55 – 6.50 (m, 2H), 6.31 (d, *J* = 8.1 Hz, 2H), 5.94 (d, *J* = 9.5 Hz, 1H), 5.55 (dd, *J* = 11.3, 4.8 Hz, 1H), 5.41 (d, *J* = 17.3 Hz, 1H), 5.40 – 5.31 (m, 1H), 5.11 – 5.02 (m, 1H), 4.44 (dd, *J* = 9.5, 3.1 Hz, 1H), 4.38 (dd, *J* = 9.1, 3.4 Hz, 1H), 3.79 (ddd, *J* = 12.0, 7.6, 4.3 Hz, 1H), 3.63 (dt, *J* = 11.8, 7.7 Hz, 1H), 3.45 (dd, *J* = 11.7, 4.1 Hz, 1H), 3.36 (d, *J* = 17.4 Hz, 1H), 3.32 – 3.13 (m, 3H), 3.09 (s, 3H), 3.12 – 3.05 (m, 1H), 3.00 (t, *J* = 12.2 Hz, 1H), 2.92 (s, 3H), 2.81 – 2.73 (m, 2H), 2.69 (s, 3H), 2.51 – 2.42 (m, 1H), 2.09 – 2.02 (m, 1H), 1.38 (ddd, *J* = 11.9, 7.4, 4.1 Hz, 1H), 1.28 (s, 1H), 1.07 – 1.00 (m, 1H), 0.97 (s, 1H), 0.95 (d, *J* = 7.3 Hz, 2H), 0.92 (d, *J* = 6.9 Hz, 3H), 0.86 (d, *J* = 7.1 Hz, 3H); LCMS (*m/z*): calculated exact mass: 853.4738, experimental $[M+H]^{+}$: 854.4800.



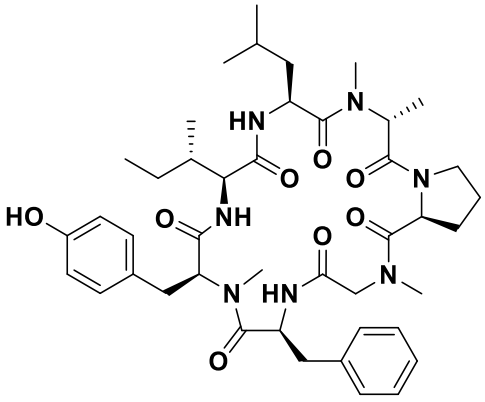
1.5.18.7 NMR and LC/MS data for VK-07 (**7**)



Sequence: Cyclo-(Leu-MeDPhe-MeAla-Sar-Phe-MeTyr-Ile)

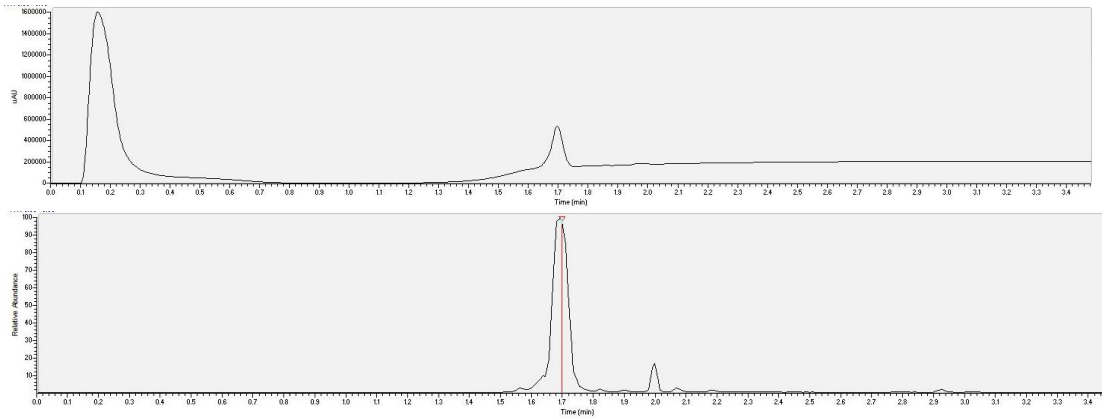
^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.77 (d, *J* = 9.0 Hz, 1H), 7.37 (d, *J* = 8.3 Hz, 1H), 7.30 – 7.21 (m, 4H), 7.24 – 7.14 (m, 3H), 7.14 – 7.04 (m, 2H), 6.97 – 6.92 (m, 2H), 6.73 – 6.68 (m, 2H), 6.64 (d, *J* = 7.8 Hz, 1H), 5.62 (dd, *J* = 9.4, 7.0 Hz, 1H), 5.35 (q, *J* = 6.6 Hz, 1H), 5.02 – 4.95 (m, 2H), 4.76 – 4.68 (m, 2H), 4.58 (s, 1H), 4.30 (dd, *J* = 8.9, 6.9 Hz, 1H), 3.21 (dd, *J* = 13.1, 9.4 Hz, 1H), 3.16 (s, 3H), 3.02 – 2.92 (m, 5H), 2.86 – 2.78 (m, 1H), 2.72 (s, 3H), 2.65 (s, 3H), 2.38 (dd, *J* = 13.5, 7.2 Hz, 1H), 2.09 (ddd, *J* = 28.7, 13.7, 7.1 Hz, 2H), 1.35 – 1.15 (m, 2H), 1.12 – 0.95 (m, 5H), 0.86 (dd, *J* = 16.4, 6.6 Hz, 6H), 0.79 (t, *J* = 7.3 Hz, 6H); LCMS (*m/z*): calculated exact mass: 867.4895, experimental $[M+H]^{+}$: .

1.5.18.8 NMR and LC/MS data for VK-08 (**8**)

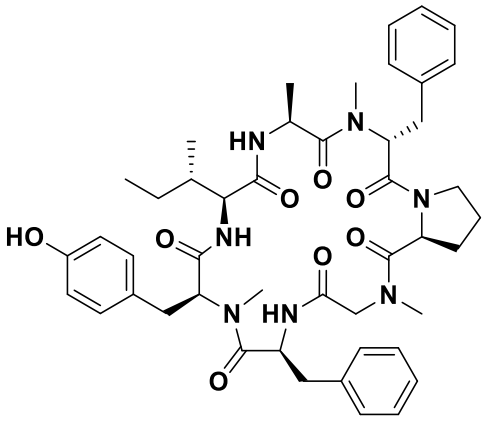


Sequence: Cyclo-(Leu-MeDAla-Pro-Sar-Phe-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.58 (d, *J* = 9.9 Hz, 1H), 8.30 (d, *J* = 9.2 Hz, 1H), 7.41 – 7.34 (m, 2H), 7.30 (d, *J* = 13.9 Hz, 6H), 6.60 (d, *J* = 8.2 Hz, 2H), 6.53 (d, *J* = 8.2 Hz, 2H), 6.02 (d, *J* = 9.3 Hz, 1H), 5.30 – 5.16 (m, 4H), 4.84 (s, 2H), 4.49 (ddd, *J* = 15.8, 9.2, 3.2 Hz, 2H), 3.78 (ddd, *J* = 12.0, 7.8, 4.4 Hz, 1H), 3.63 – 3.52 (m, 2H), 3.37 (t, *J* = 11.8 Hz, 1H), 3.31 (d, *J* = 17.5 Hz, 1H), 3.11 (dd, *J* = 12.9, 10.0 Hz, 1H), 3.02 (d, *J* = 20.0 Hz, 6H), 2.96 – 2.84 (m, 2H), 2.82 – 2.71 (m, 4H), 2.47 (dd, *J* = 11.5, 8.5 Hz, 1H), 2.40 (td, *J* = 6.8, 3.4 Hz, 1H), 2.13 – 2.02 (m, 1H), 1.96 – 1.88 (m, 1H), 1.84 (s, 2H), 1.43 (s, 1H), 1.37 (t, *J* = 12.2 Hz, 1H), 1.25 (d, *J* = 6.7 Hz, 3H), 1.13 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.5 Hz, 3H), 0.99 (t, *J* = 7.2 Hz, 3H), 0.93 (d, *J* = 6.9 Hz, 3H); LCMS (*m/z*): calculated exact mass: 803.4582, experimental $[M+H]^{+}$: 804.4653.

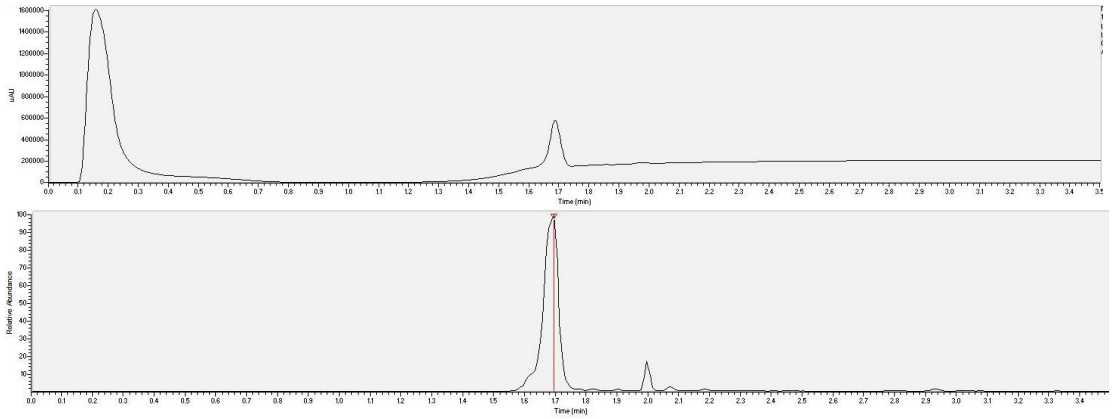


1.5.18.9 NMR and LC/MS data for VK-09 (**9**)

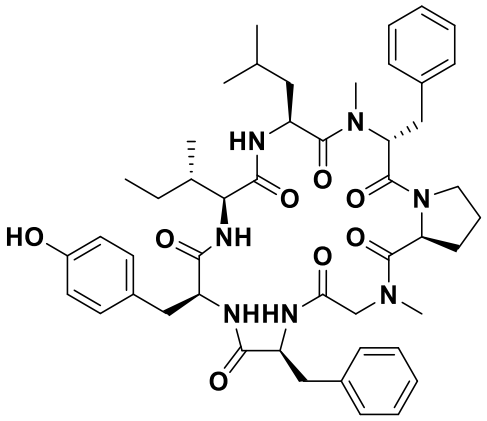


Sequence: Cyclo-(Ala-MeDPhe-Pro-Sar-Phe-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.53 (d, *J* = 9.1 Hz, 1H), 7.25 (d, *J* = 7.2 Hz, 5H), 7.22 – 7.14 (m, 7H), 7.01 (d, *J* = 8.3 Hz, 3H), 6.96 (d, *J* = 7.3 Hz, 2H), 6.84 (t, *J* = 6.7 Hz, 0H), 6.75 (t, *J* = 9.1 Hz, 3H), 6.12 (s, 1H), 6.01 (s, 1H), 5.79 (dd, *J* = 17.8, 7.8 Hz, 1H), 5.62 (dd, *J* = 11.4, 5.6 Hz, 1H), 5.00 – 4.89 (m, 1H), 4.78 (p, *J* = 6.5 Hz, 1H), 4.62 (d, *J* = 16.0 Hz, 1H), 4.38 (s, 1H), 4.38 – 4.32 (m, 1H), 4.16 (s, 1H), 3.52 – 3.39 (m, 1H), 3.36 – 3.22 (m, 2H), 3.12 (d, *J* = 10.5 Hz, 2H), 3.07 (s, 1H), 2.99 (dd, *J* = 23.0, 15.0 Hz, 10H), 2.76 (s, 3H), 2.77 – 2.64 (m, 4H), 2.00 (d, *J* = 14.4 Hz, 1H), 1.91 (s, 1H), 1.57 (d, *J* = 7.1 Hz, 1H), 1.32 (d, *J* = 6.7 Hz, 1H), 1.30 (s, 1H), 1.27 (d, *J* = 9.7 Hz, 3H), 1.21 (dd, *J* = 13.9, 7.2 Hz, 1H), 1.07 – 0.97 (m, 1H), 1.00 – 0.93 (m, 2H), 0.96 – 0.87 (m, 6H), 0.89 – 0.80 (m, 2H), 0.78 – 0.67 (m, 9H), 0.36 (s, 6H); LCMS (*m/z*): calculated exact mass: 837.4425, experimental $[M+H]^+$: 838.4492.

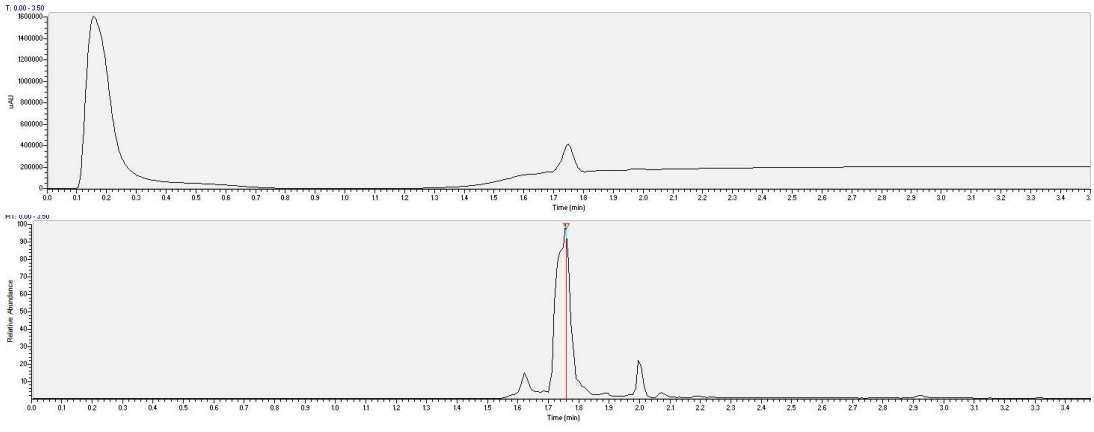


1.5.18.10 NMR and LC/MS data for VK-10 (**10**)

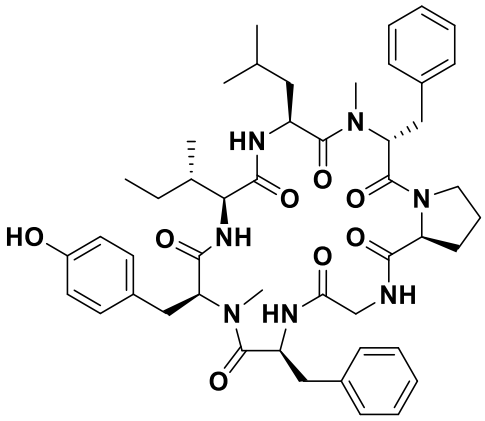


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe-Tyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.75 (d, *J* = 9.5 Hz, 1H), 7.63 (d, *J* = 8.9 Hz, 1H), 7.42 – 7.23 (m, 6H), 7.23 – 7.16 (m, 1H), 7.18 – 7.13 (m, 2H), 7.13 (d, *J* = 7.8 Hz, 2H), 7.11 – 7.04 (m, 1H), 6.78 (d, *J* = 7.3 Hz, 1H), 6.68 – 6.58 (m, 2H), 6.55 – 6.49 (m, 2H), 6.46 – 6.39 (m, 2H), 5.57 (dd, *J* = 11.4, 5.1 Hz, 1H), 5.34 (d, *J* = 16.4 Hz, 1H), 4.91 (ddd, *J* = 11.7, 8.8, 2.4 Hz, 1H), 4.82 (ddd, *J* = 11.6, 9.6, 3.9 Hz, 1H), 4.53 (dd, *J* = 10.4, 3.5 Hz, 1H), 4.28 (dd, *J* = 9.2, 3.0 Hz, 1H), 4.19 (s, 1H), 3.77 (ddd, *J* = 12.0, 7.9, 4.2 Hz, 1H), 3.70 – 3.58 (m, 1H), 3.57 (tt, *J* = 7.2, 3.6 Hz, 1H), 3.30 – 3.21 (m, 3H), 3.17 (d, *J* = 4.2 Hz, 1H), 3.12 (dd, *J* = 15.1, 5.1 Hz, 1H), 3.06 (dd, *J* = 13.1, 11.6 Hz, 1H), 3.02 (s, 3H), 2.81 (dd, *J* = 13.0, 4.6 Hz, 1H), 2.74 (dd, *J* = 13.2, 3.9 Hz, 1H), 2.67 (s, 3H), 2.48 – 2.33 (m, 1H), 2.19 (ddt, *J* = 10.2, 6.8, 3.4 Hz, 1H), 1.99 (ddd, *J* = 13.0, 6.9, 3.5 Hz, 1H), 1.90 – 1.76 (m, 1H), 1.38 – 1.21 (m, 2H), 1.25 (s, 1H), 1.07 – 0.80 (m, 16H), 0.80 (dd, *J* = 6.5, 3.0 Hz, 1H), 0.52 (q, *J* = 8.0 Hz, 2H), 0.19 (ddd, *J* = 13.9, 10.9, 2.4 Hz, 1H); LCMS (*m/z*): calculated exact mass: 865.4738, experimental $[M+H]^{+}$: 866.4789.

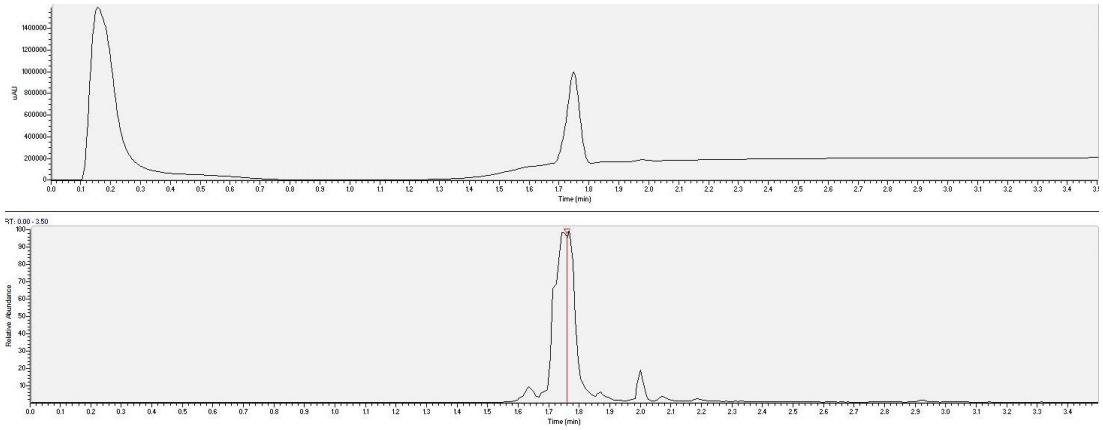


1.5.18.11 NMR and LC/MS data for VK-11 (**11**)

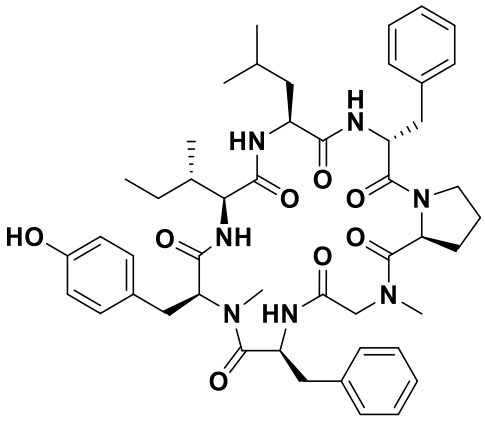


Sequence: Cyclo-(Leu-MeDPhe-Pro-Gly-Phe-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.60 (d, *J* = 9.7 Hz, 1H), 8.23 (d, *J* = 8.6 Hz, 1H), 7.86 (d, *J* = 9.8 Hz, 1H), 7.40 (t, *J* = 7.4 Hz, 2H), 7.37 – 7.22 (m, 4H), 7.15 (d, *J* = 5.6 Hz, 4H), 7.07 (td, *J* = 6.0, 2.9 Hz, 1H), 7.00 (ddd, *J* = 27.7, 19.1, 9.4 Hz, 1H), 6.62 (s, 12H), 6.58 (d, *J* = 7.9 Hz, 2H), 6.36 (d, *J* = 7.9 Hz, 2H), 6.13 (d, *J* = 8.2 Hz, 1H), 5.51 (dd, *J* = 11.6, 4.5 Hz, 1H), 5.27 (td, *J* = 10.6, 3.9 Hz, 1H), 4.99 (dd, *J* = 17.9, 9.9 Hz, 1H), 4.90 – 4.82 (m, 1H), 4.35 (ddd, *J* = 15.9, 8.6, 3.2 Hz, 2H), 4.21 (d, *J* = 2.3 Hz, 1H), 3.99 (s, 3H), 3.74 (ddd, *J* = 12.0, 7.6, 4.2 Hz, 1H), 3.67 (d, *J* = 17.8 Hz, 1H), 3.60 – 3.53 (m, 1H), 3.52 (s, 4H), 3.49 (dd, *J* = 11.1, 5.1 Hz, 1H), 3.27 (dd, *J* = 14.9, 11.6 Hz, 1H), 3.16 (t, *J* = 11.6 Hz, 1H), 3.10 (s, 3H), 3.08 (d, *J* = 7.2 Hz, 1H), 3.06 – 2.95 (m, 2H), 2.93 (s, 0H), 2.84 (dd, *J* = 12.2, 5.1 Hz, 1H), 2.78 (dd, *J* = 12.9, 3.9 Hz, 1H), 2.54 (s, 3H), 2.36 – 2.24 (m, 2H), 2.08 – 1.97 (m, 2H), 1.85 (ddt, *J* = 37.2, 19.8, 9.4 Hz, 3H), 1.51 – 1.41 (m, 2H), 1.37 (d, *J* = 18.0 Hz, 2H), 1.11 – 0.86 (m, 15H), 0.81 (dd, *J* = 17.9, 6.4 Hz, 5H), 0.54 (q, *J* = 7.9 Hz, 2H), 0.36 (td, *J* = 12.6, 11.5, 5.9 Hz, 1H); LCMS (*m/z*): calculated exact mass: 865.4738, experimental $[M+H]^+$: 866.4791.

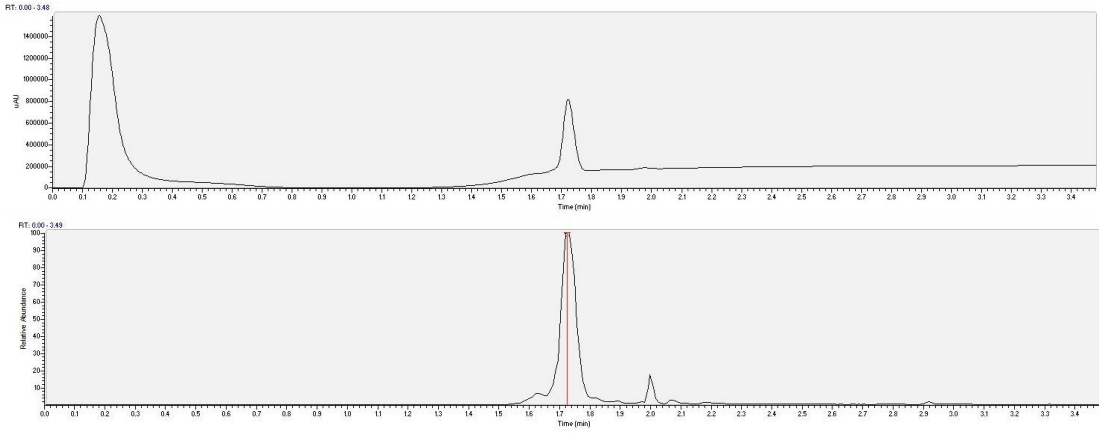


1.5.18.12 NMR and LC/MS data for VK-12 (**12**)

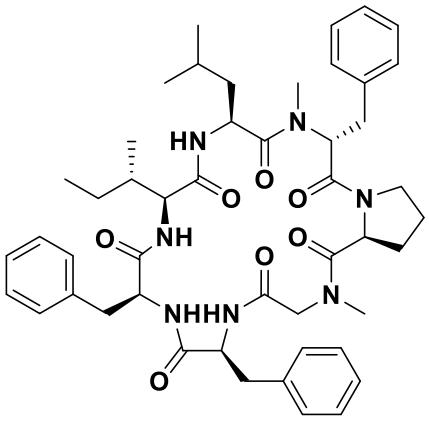


Sequence: Cyclo-(Leu-DPhe-Pro-Sar-Phe-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.12 (d, *J* = 7.5 Hz, 1H), 7.54 (d, *J* = 9.5 Hz, 1H), 7.38 – 7.14 (m, 12H), 7.16 – 7.05 (m, 1H), 7.03 (d, *J* = 8.6 Hz, 1H), 6.71 (d, *J* = 8.2 Hz, 2H), 6.63 (t, *J* = 6.6 Hz, 2H), 6.51 (dd, *J* = 16.6, 8.1 Hz, 1H), 6.02 (dd, *J* = 18.0, 6.9 Hz, 1H), 5.02 (dtd, *J* = 17.7, 9.1, 8.7, 5.2 Hz, 1H), 4.91 (ddd, *J* = 9.4, 7.4, 5.8 Hz, 1H), 4.78 (d, *J* = 16.9 Hz, 1H), 4.52 – 4.34 (m, 2H), 4.27 (dd, *J* = 7.4, 3.2 Hz, 1H), 3.72 (q, *J* = 7.1, 6.5 Hz, 1H), 3.54 (q, *J* = 7.3 Hz, 1H), 3.23 (s, 2H), 3.29 – 3.15 (m, 2H), 3.07 – 3.01 (m, 1H), 3.03 – 2.99 (m, 1H), 3.01 – 2.92 (m, 2H), 2.87 (dt, *J* = 10.0, 6.6 Hz, 1H), 2.69 (t, *J* = 7.0 Hz, 1H), 2.66 (s, 3H), 2.62 (d, *J* = 13.3 Hz, 1H), 2.56 (dt, *J* = 13.3, 6.3 Hz, 1H), 2.29 (ddt, *J* = 11.4, 7.7, 4.1 Hz, 1H), 2.03 (dp, *J* = 12.6, 6.6 Hz, 1H), 1.99 – 1.75 (m, 3H), 1.50 – 1.30 (m, 2H), 1.06 – 0.93 (m, 4H), 0.95 – 0.83 (m, 8H), 0.82 – 0.67 (m, 3H); LCMS (*m/z*): calculated exact mass: 865.4738, experimental $[M+H]^{+}$: 866.4794.

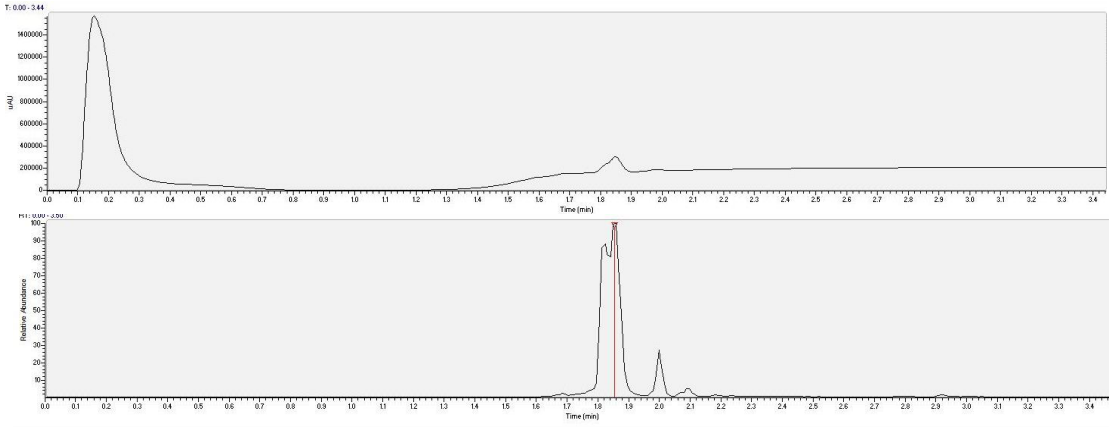


1.5.18.13 NMR and LC/MS data for VK-13 (**13**)

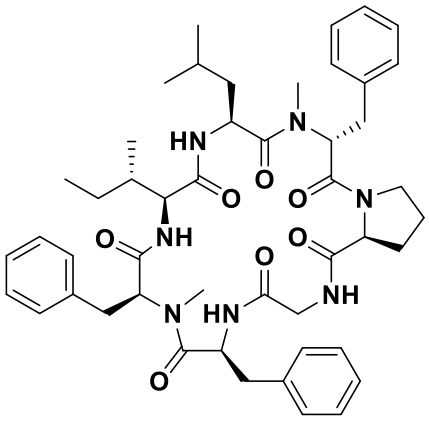


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe-Phe-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.72 (d, *J* = 9.5 Hz, 1H), 7.55 (d, *J* = 8.9 Hz, 1H), 7.43 – 7.36 (m, 2H), 7.39 – 7.32 (m, 1H), 7.35 – 7.26 (m, 2H), 7.27 (d, *J* = 1.7 Hz, 1H), 7.24 – 7.11 (m, 6H), 7.13 – 7.03 (m, 3H), 6.68 – 6.58 (m, 4H), 5.57 (dd, *J* = 11.4, 5.1 Hz, 1H), 5.40 (s, 1H), 5.34 (d, *J* = 16.3 Hz, 1H), 4.91 (ddd, *J* = 11.8, 8.9, 2.5 Hz, 1H), 4.82 (ddd, *J* = 11.5, 9.5, 3.9 Hz, 1H), 4.53 (dd, *J* = 10.4, 3.5 Hz, 1H), 4.49 – 4.40 (m, 0H), 4.27 (dd, *J* = 9.2, 3.0 Hz, 1H), 3.77 (ddd, *J* = 12.0, 7.9, 4.1 Hz, 1H), 3.68 – 3.55 (m, 2H), 3.34 (t, *J* = 12.4 Hz, 1H), 3.25 (dd, *J* = 15.5, 11.2 Hz, 2H), 3.21 – 3.04 (m, 3H), 3.02 (s, 3H), 2.94 – 2.83 (m, 2H), 2.75 (dd, *J* = 13.1, 3.8 Hz, 1H), 2.66 (s, 3H), 2.42 (dtd, *J* = 12.8, 9.8, 7.1 Hz, 1H), 2.26 – 2.13 (m, 1H), 1.99 (ddd, *J* = 12.8, 7.0, 3.6 Hz, 1H), 1.87 (s, 1H), 1.79 (s, 1H), 1.40 – 1.26 (m, 1H), 1.25 (s, 3H), 1.07 – 0.81 (m, 14H), 0.84 – 0.75 (m, 1H), 0.19 (ddd, *J* = 14.0, 11.0, 2.4 Hz, 1H); LCMS (*m/z*): calculated exact mass: 849.4789, experimental $[M+H]^+$: 850.4852.

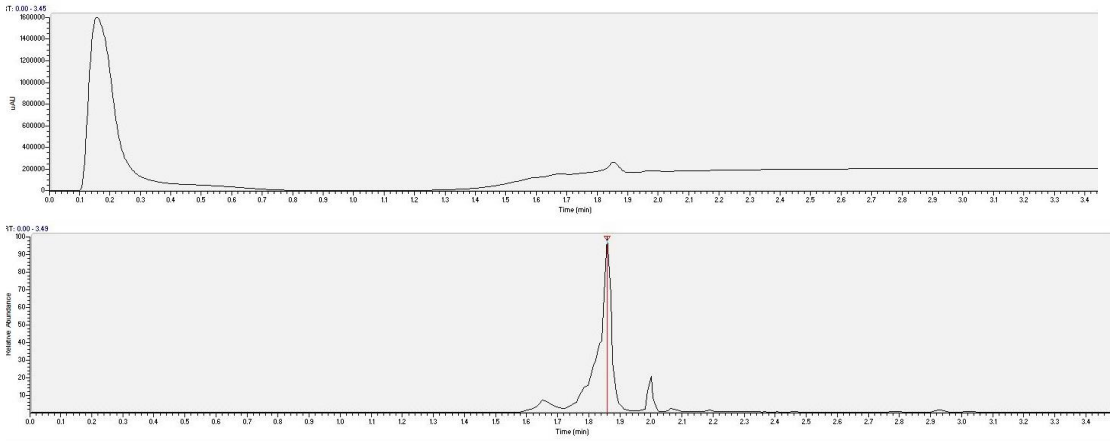


1.5.18.14 NMR and LC/MS data for VK-14 (**14**)

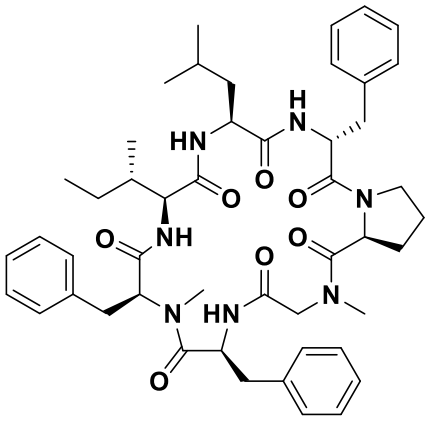


Sequence: Cyclo-(Leu-MeDPhe-Pro-Gly-Phe-MePhe-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.40 (d, *J* = 9.9 Hz, 1H), 7.98 (d, *J* = 8.5 Hz, 1H), 7.41 – 7.24 (m, 6H), 7.18 (q, *J* = 7.0, 5.5 Hz, 2H), 7.16 – 7.08 (m, 6H), 7.06 (d, *J* = 6.8 Hz, 1H), 6.66 (d, *J* = 7.5 Hz, 2H), 6.00 (d, *J* = 8.3 Hz, 1H), 5.52 (dd, *J* = 11.6, 4.7 Hz, 1H), 5.31 (td, *J* = 10.5, 4.4 Hz, 1H), 4.94 (dd, *J* = 17.8, 9.9 Hz, 1H), 4.73 (td, *J* = 9.4, 4.5 Hz, 1H), 4.42 (td, *J* = 17.1, 14.6, 10.3 Hz, 1H), 4.19 (ddd, *J* = 20.8, 8.8, 3.2 Hz, 2H), 3.74 (ddd, *J* = 12.1, 7.7, 4.5 Hz, 1H), 3.59 – 3.46 (m, 3H), 3.31 (t, *J* = 11.5 Hz, 1H), 3.23 (s, 1H), 3.21 (dd, *J* = 14.9, 11.4 Hz, 1H), 3.10 (d, *J* = 15.7 Hz, 4H), 3.08 – 2.95 (m, 2H), 2.89 – 2.74 (m, 2H), 2.63 (s, 3H), 2.37 – 2.28 (m, 1H), 2.24 (s, 1H), 2.17 (d, *J* = 1.4 Hz, 4H), 2.02 (d, *J* = 14.8 Hz, 2H), 1.89 – 1.81 (m, 2H), 1.53 (ddd, *J* = 14.8, 10.1, 5.5 Hz, 1H), 1.48 – 1.33 (m, 2H), 1.25 (s, 2H), 0.97 (s, 1H), 1.02 – 0.86 (m, 7H), 0.83 (d, *J* = 7.0 Hz, 4H), 0.78 (d, *J* = 6.4 Hz, 3H), 0.70 (ddd, *J* = 13.6, 8.5, 4.7 Hz, 1H); LCMS (*m/z*): calculated exact mass: 849.4789, experimental $[M+H]^+$: 850.4868.

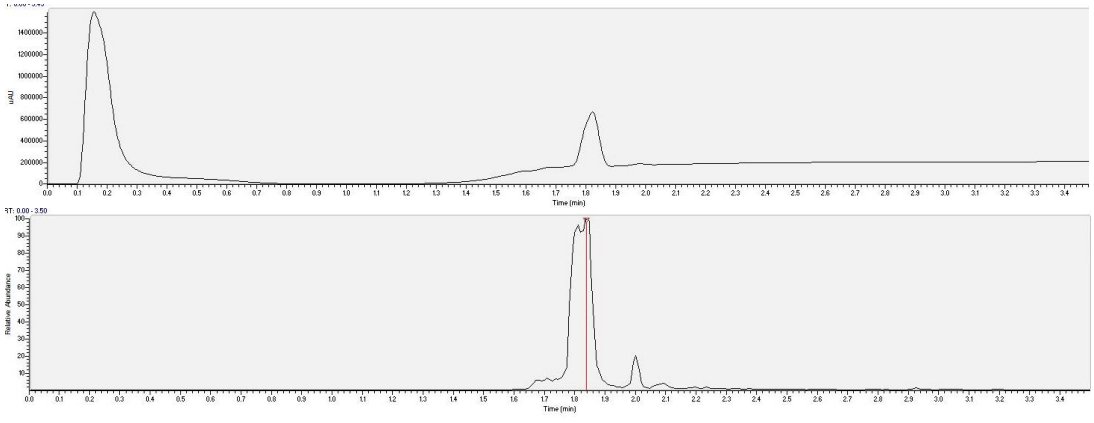


1.5.18.15 NMR and LC/MS data for VK-15 (**15**)

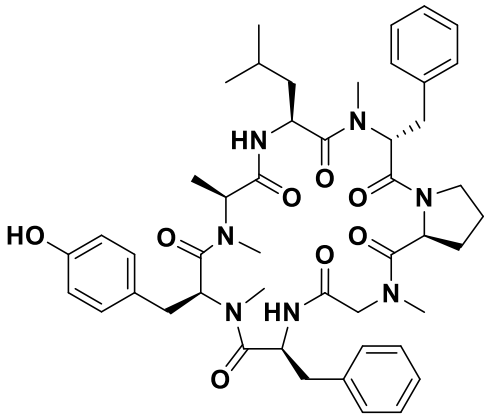


Sequence: Cyclo-(Leu-DPhe-Pro-Sar-Phe-MePhe-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.46 (d, *J* = 9.9 Hz, 0H), 8.27 (s, 1H), 8.07 (s, 1H), 7.98 (d, *J* = 9.5 Hz, 0H), 7.50 (s, 1H), 7.37 (dt, *J* = 8.3, 6.5 Hz, 3H), 7.34 – 7.27 (m, 8H), 7.30 – 7.22 (m, 10H), 7.24 – 7.16 (m, 7H), 7.14 (qd, *J* = 8.2, 2.4 Hz, 2H), 7.07 – 6.96 (m, 3H), 6.83 (s, 3H), 6.54 (dd, *J* = 20.8, 8.9 Hz, 1H), 6.45 – 6.40 (m, 1H), 5.93 (s, 1H), 5.81 (s, 1H), 5.28 (s, 1H), 5.06 – 4.99 (m, 3H), 4.92 (dt, *J* = 9.7, 6.8 Hz, 2H), 4.80 (d, *J* = 16.2 Hz, 2H), 4.56 (dd, *J* = 8.8, 3.9 Hz, 1H), 4.47 – 4.41 (m, 4H), 4.37 (td, *J* = 10.3, 9.8, 4.1 Hz, 1H), 4.29 (dd, *J* = 7.7, 3.5 Hz, 2H), 4.15 (d, *J* = 7.1 Hz, 1H), 3.82 (d, *J* = 16.5 Hz, 1H), 3.81 – 3.76 (m, 1H), 3.65 – 3.55 (m, 2H), 3.28 – 3.15 (m, 9H), 3.14 (s, 2H), 3.15 – 3.09 (m, 1H), 3.09 – 2.96 (m, 7H), 2.96 – 2.86 (m, 3H), 2.83 (ddd, *J* = 16.3, 7.4, 4.4 Hz, 1H), 2.74 – 2.68 (m, 1H), 2.67 (s, 4H), 2.66 (s, 1H), 2.64 – 2.52 (m, 3H), 2.40 – 2.28 (m, 1H), 2.22 (t, *J* = 7.6 Hz, 1H), 2.08 (tq, *J* = 13.0, 7.3, 6.9 Hz, 2H), 2.04 – 1.93 (m, 1H), 1.99 (s, 2H), 1.90 (s, 5H), 1.83 (d, *J* = 9.5 Hz, 1H), 1.62 (d, *J* = 7.6 Hz, 2H), 1.54 – 1.41 (m, 2H), 1.35 – 1.25 (m, 2H), 1.25 (s, 2H), 1.21 – 1.15 (m, 1H), 1.01 – 0.86 (m, 19H), 0.89 – 0.69 (m, 13H); LCMS (*m/z*): calculated exact mass: 849.4789, experimental $[M+H]^+$: 850.4853.



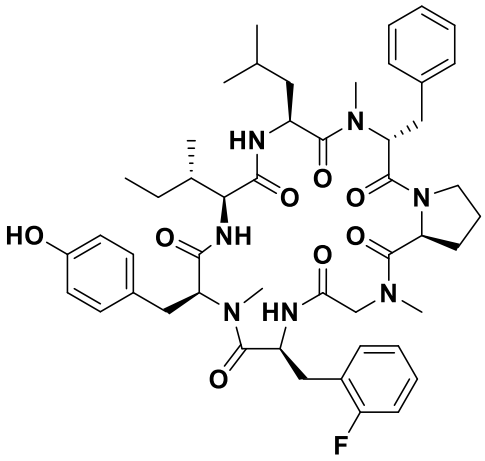
1.5.18.16 NMR and LC/MS data for VK-16 (**16**)



Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe-MeTyr-MeAla)

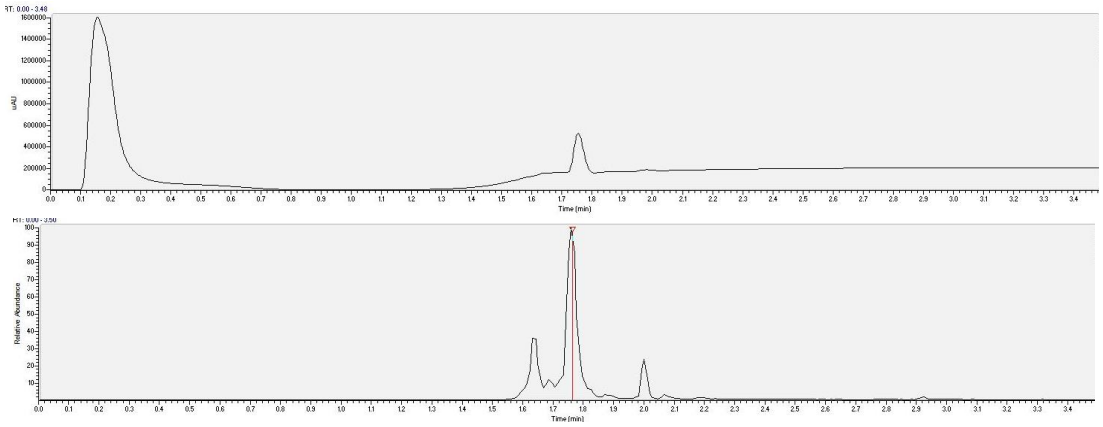
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.52 (d, *J* = 10.0 Hz, 1H), 7.38 – 7.21 (m, 7H), 7.12 (d, *J* = 7.1 Hz, 2H), 7.06 (t, *J* = 7.5 Hz, 2H), 6.98 (t, *J* = 7.9 Hz, 3H), 6.76 – 6.69 (m, 2H), 5.56 (dd, *J* = 11.5, 5.1 Hz, 1H), 5.45 – 5.34 (m, 2H), 5.26 (q, *J* = 7.2 Hz, 1H), 4.83 (ddd, *J* = 11.8, 9.2, 2.1 Hz, 1H), 4.69 (dd, *J* = 11.9, 4.7 Hz, 1H), 4.28 (dd, *J* = 9.3, 3.3 Hz, 1H), 3.76 (ddd, *J* = 12.0, 7.8, 4.3 Hz, 1H), 3.60 (dt, *J* = 11.8, 7.8 Hz, 1H), 3.38 (d, *J* = 17.4 Hz, 1H), 3.29 – 3.20 (m, 1H), 3.24 (s, 3H), 3.14 – 2.98 (m, 4H), 2.97 (s, 3H), 2.95 (s, 4H), 2.92 (dd, *J* = 12.5, 4.7 Hz, 1H), 2.88 (s, 3H), 2.75 (dd, *J* = 13.0, 4.3 Hz, 1H), 2.70 (s, 3H), 2.43 (dtd, *J* = 12.8, 9.6, 7.0 Hz, 1H), 2.24 (s, 3H), 2.08 – 1.99 (m, 1H), 1.92 – 1.78 (m, 1H), 1.78 (s, 1H), 1.42 – 1.30 (m, 1H), 1.32 – 1.24 (m, 1H), 1.11 (d, *J* = 7.3 Hz, 3H), 0.78 (dt, *J* = 10.4, 5.1 Hz, 8H); LCMS (*m/z*): calculated exact mass: 851.4582, experimental $[M+H]^{+}$

1.5.18.17 NMR and LC/MS data for VK-19 (**19**)

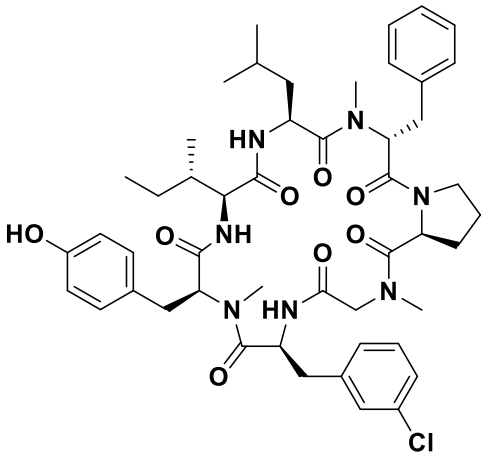


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe(2-F)-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.54 (d, *J* = 9.8 Hz, 1H), 8.14 (d, *J* = 9.1 Hz, 1H), 7.35 (d, *J* = 6.6 Hz, 1H), 7.22 (s, 1H), 7.20 – 7.09 (m, 7H), 7.06 – 6.97 (m, 2H), 6.68 (d, *J* = 7.9 Hz, 0H), 6.52 (d, *J* = 8.0 Hz, 2H), 6.32 (d, *J* = 8.0 Hz, 2H), 5.88 (d, *J* = 9.6 Hz, 1H), 5.55 (dd, *J* = 11.3, 4.8 Hz, 1H), 5.46 (td, *J* = 10.6, 3.8 Hz, 1H), 5.40 (d, *J* = 17.4 Hz, 1H), 5.29 (s, 1H), 4.99 – 4.91 (m, 1H), 4.45 (dd, *J* = 9.7, 3.0 Hz, 1H), 4.37 (dd, *J* = 9.0, 3.4 Hz, 1H), 4.25 (s, 4H), 4.18 (s, 1H), 3.77 (ddd, *J* = 11.9, 7.6, 4.3 Hz, 1H), 3.69 – 3.56 (m, 1H), 3.51 – 3.40 (m, 2H), 3.42 (s, 3H), 3.35 (d, *J* = 17.3 Hz, 1H), 3.31 – 3.16 (m, 2H), 3.20 (s, 2H), 3.15 – 3.06 (m, 1H), 3.05 (s, 4H), 3.03 – 2.95 (m, 1H), 2.89 (s, 3H), 2.86 – 2.73 (m, 2H), 2.70 (s, 3H), 2.50 – 2.41 (m, 1H), 2.36 – 2.29 (m, 1H), 2.08 – 2.01 (m, 1H), 1.92 – 1.84 (m, 1H), 1.81 (s, 1H), 1.55 (q, *J* = 7.0 Hz, 1H), 1.47 – 1.39 (m, 1H), 1.37 (dd, *J* = 22.3, 8.9 Hz, 1H), 1.26 (s, 4H), 0.99 – 0.85 (m, 16H), 0.84 (d, *J* = 6.5 Hz, 4H), 0.81 (s, 6H), 0.81 – 0.69 (m, 1H), 0.59 (t, *J* = 8.4 Hz, 1H), 0.52 (q, *J* = 8.0 Hz, 4H), 0.19 (t, *J* = 12.6 Hz, 1H); LCMS (*m/z*): calculated exact mass: 897.4800, experimental $[M+H]^+$: 898.4862.

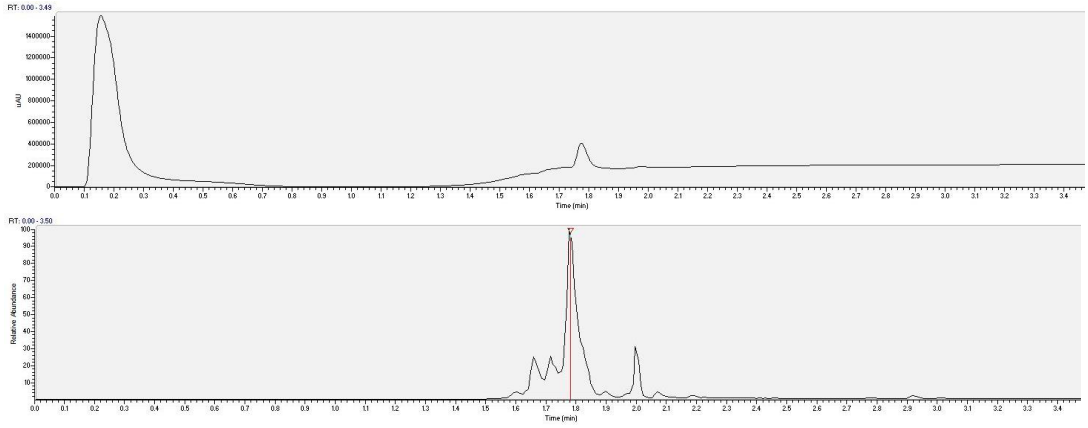


1.5.18.18 NMR and LC/MS data for VK-20 (**20**)

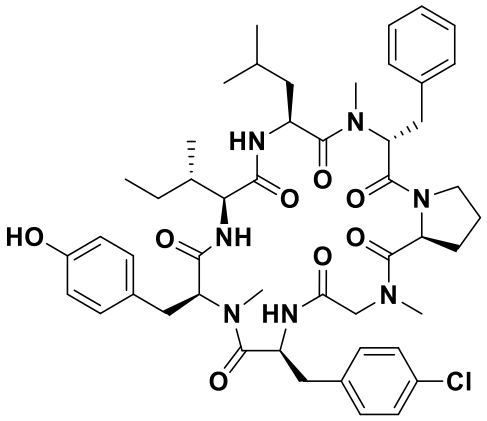


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe(3-Cl)-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.52 (d, *J* = 9.8 Hz, 1H), 8.09 (d, *J* = 9.1 Hz, 1H), 7.34 (d, *J* = 4.6 Hz, 2H), 7.30 (s, 1H), 7.24 – 7.08 (m, 5H), 7.04 (t, *J* = 7.0 Hz, 1H), 6.56 (s, 1H), 6.55 (s, 1H), 6.35 (d, *J* = 8.2 Hz, 2H), 5.88 (d, *J* = 9.5 Hz, 1H), 5.53 (dd, *J* = 11.6, 4.7 Hz, 1H), 5.39 (d, *J* = 17.4 Hz, 1H), 5.29 (td, *J* = 10.9, 9.9, 3.8 Hz, 1H), 4.94 (ddd, *J* = 11.9, 9.2, 2.4 Hz, 1H), 4.83 (s, 1H), 4.44 (dd, *J* = 9.5, 3.1 Hz, 1H), 4.38 (dd, *J* = 9.2, 3.4 Hz, 1H), 4.24 (s, 2H), 3.77 (ddd, *J* = 11.9, 7.7, 4.3 Hz, 1H), 3.69 – 3.56 (m, 1H), 3.46 (dd, *J* = 11.6, 4.3 Hz, 1H), 3.36 (s, 1H), 3.34 – 3.27 (m, 1H), 3.27 – 3.18 (m, 2H), 3.17 (s, 1H), 3.20 – 3.12 (m, 0H), 3.05 (s, 3H), 3.02 (dd, *J* = 15.9, 10.6 Hz, 2H), 2.96 (t, *J* = 12.3 Hz, 1H), 2.90 (s, 3H), 2.75 (ddd, *J* = 28.5, 12.4, 4.2 Hz, 2H), 2.69 (s, 3H), 2.51 – 2.41 (m, 1H), 2.38 – 2.29 (m, 1H), 2.07 – 2.01 (m, 1H), 1.90 (s, 1H), 1.89 – 1.80 (m, 1H), 1.80 (s, 1H), 1.49 – 1.38 (m, 0H), 1.38 – 1.28 (m, 3H), 1.26 (s, 2H), 0.94 (d, *J* = 8.2 Hz, 11H), 0.90 (d, *J* = 14.9 Hz, 7H), 0.89 – 0.80 (m, 4H), 0.83 – 0.71 (m, 1H), 0.63 – 0.56 (m, 1H), 0.52 (q, *J* = 7.9 Hz, 6H), 0.20 (t, *J* = 12.3 Hz, 1H); LCMS (*m/z*): calculated exact mass: 913.4505, experimental $[M+H]^{+}$: 914.4569.

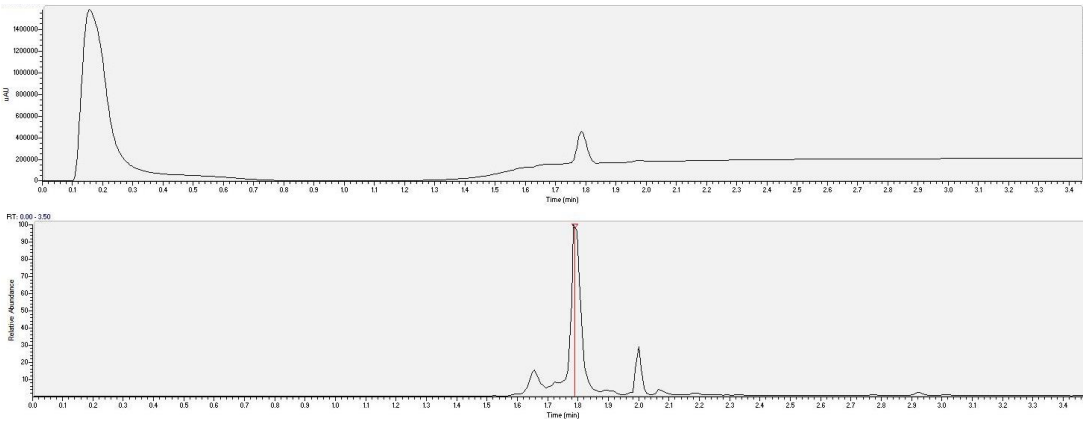


1.5.18.19 NMR and LC/MS data for VK-21 (**21**)

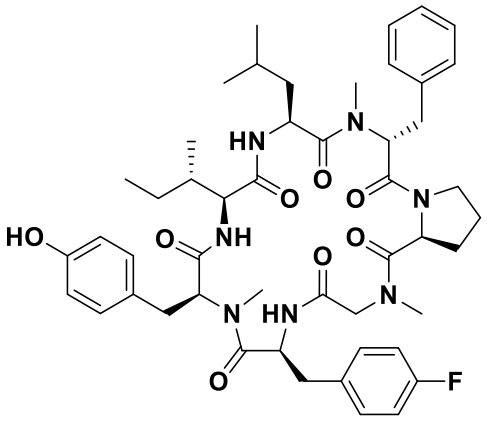


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe(4-Cl)-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.61 (d, *J* = 9.8 Hz, 1H), 8.21 (d, *J* = 9.0 Hz, 1H), 7.41 – 7.36 (m, 2H), 7.23 (d, *J* = 7.9 Hz, 2H), 7.21 – 7.09 (m, 4H), 7.03 (t, *J* = 7.2 Hz, 1H), 6.61 – 6.54 (m, 2H), 6.31 (d, *J* = 7.9 Hz, 2H), 5.89 (d, *J* = 9.3 Hz, 1H), 5.53 (dd, *J* = 11.6, 4.7 Hz, 1H), 5.40 (d, *J* = 17.5 Hz, 1H), 5.33 – 5.25 (m, 1H), 4.91 (q, *J* = 11.8, 11.1 Hz, 1H), 4.42 (ddd, *J* = 24.5, 9.5, 3.2 Hz, 3H), 4.23 (s, 1H), 3.77 (ddd, *J* = 12.0, 7.7, 4.1 Hz, 1H), 3.61 (dt, *J* = 11.8, 7.8 Hz, 1H), 3.49 (dd, *J* = 11.6, 4.3 Hz, 1H), 3.39 (d, *J* = 17.5 Hz, 1H), 3.29 (dd, *J* = 15.0, 11.7 Hz, 1H), 3.18 (q, *J* = 13.7, 12.0 Hz, 1H), 3.04 (dd, *J* = 15.2, 4.7 Hz, 1H), 3.04 (s, 3H), 2.96 (t, *J* = 12.2 Hz, 1H), 2.91 (s, 3H), 2.76 (ddd, *J* = 16.7, 12.5, 4.9 Hz, 2H), 2.66 (s, 3H), 2.51 – 2.41 (m, 1H), 2.31 (ddt, *J* = 10.7, 7.1, 3.6 Hz, 1H), 2.04 (ddd, *J* = 11.2, 6.9, 3.1 Hz, 1H), 1.90 (tt, *J* = 7.4, 4.1 Hz, 1H), 1.84 – 1.77 (m, 1H), 1.59 – 1.50 (m, 1H), 1.39 – 1.24 (m, 3H), 1.04 – 0.94 (m, 1H), 0.96 – 0.86 (m, 10H), 0.84 (d, *J* = 6.5 Hz, 3H), 0.52 (q, *J* = 7.9 Hz, 1H), 0.19 (t, *J* = 12.6 Hz, 1H); LCMS (*m/z*): calculated exact mass: 913.4505, experimental $[M+H]^{+}$: 914.4566.

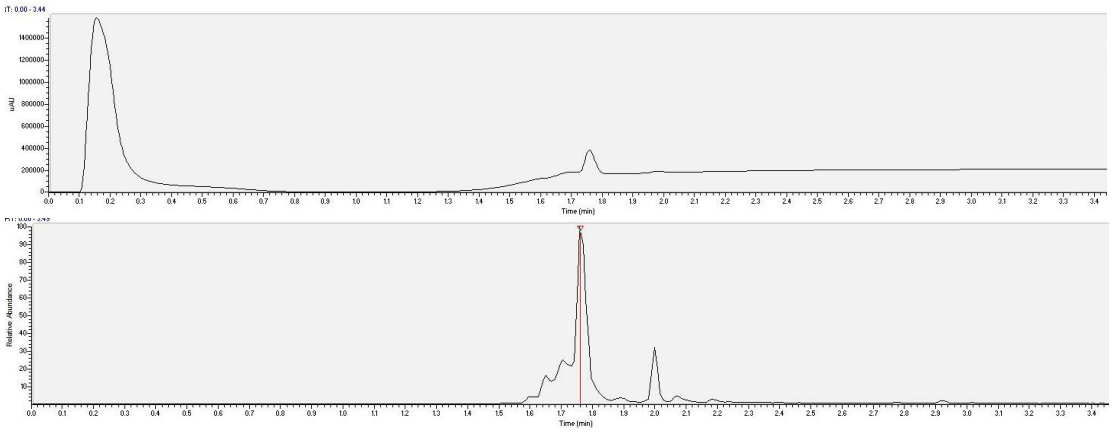


1.5.18.20 NMR and LC/MS data for VK-22 (**22**)

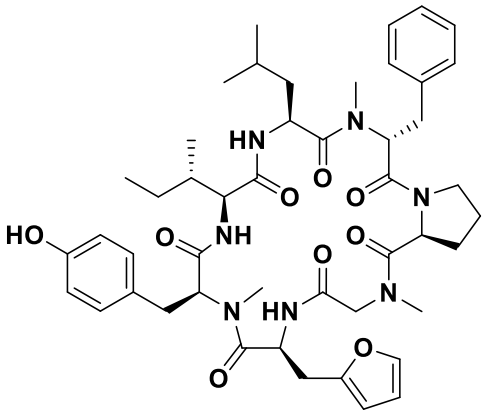


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe(4-F)-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.54 (d, *J* = 9.8 Hz, 1H), 8.12 (d, *J* = 9.1 Hz, 1H), 7.15 (d, *J* = 7.4 Hz, 2H), 7.10 (dt, *J* = 12.0, 7.8 Hz, 4H), 7.02 (t, *J* = 7.3 Hz, 1H), 6.54 (d, *J* = 8.1 Hz, 2H), 6.34 (d, *J* = 8.0 Hz, 2H), 5.88 (d, *J* = 9.5 Hz, 1H), 5.54 (dd, *J* = 11.5, 4.7 Hz, 1H), 5.39 (d, *J* = 17.4 Hz, 1H), 5.32 – 5.23 (m, 1H), 4.93 (t, *J* = 10.0 Hz, 1H), 4.44 (dd, *J* = 9.4, 3.1 Hz, 1H), 4.38 (dd, *J* = 9.1, 3.4 Hz, 1H), 4.25 (s, 3H), 3.77 (ddd, *J* = 11.9, 7.8, 4.4 Hz, 1H), 3.69 – 3.56 (m, 1H), 3.51 – 3.43 (m, 1H), 3.35 (d, *J* = 17.4 Hz, 1H), 3.29 (dd, *J* = 15.1, 11.6 Hz, 1H), 3.24 – 3.16 (m, 2H), 3.15 – 3.08 (m, 1H), 3.04 (s, 3H), 3.08 – 2.95 (m, 2H), 2.99 (s, 1H), 2.90 (s, 3H), 2.75 (ddd, *J* = 23.9, 12.6, 4.3 Hz, 2H), 2.69 (s, 3H), 2.49 – 2.43 (m, 1H), 2.39 – 2.29 (m, 3H), 1.90 (s, 1H), 1.85 – 1.77 (m, 1H), 1.47 – 1.38 (m, 1H), 1.32 (s, 2H), 1.26 (s, 3H), 1.05 – 0.86 (m, 15H), 0.89 – 0.78 (m, 3H), 0.60 (p, *J* = 7.6 Hz, 1H), 0.52 (q, *J* = 7.9 Hz, 3H), 0.16 (t, *J* = 12.6 Hz, 1H); LCMS (*m/z*): calculated exact mass: 897.4800, experimental $[M+H]^+$: 898.4861.

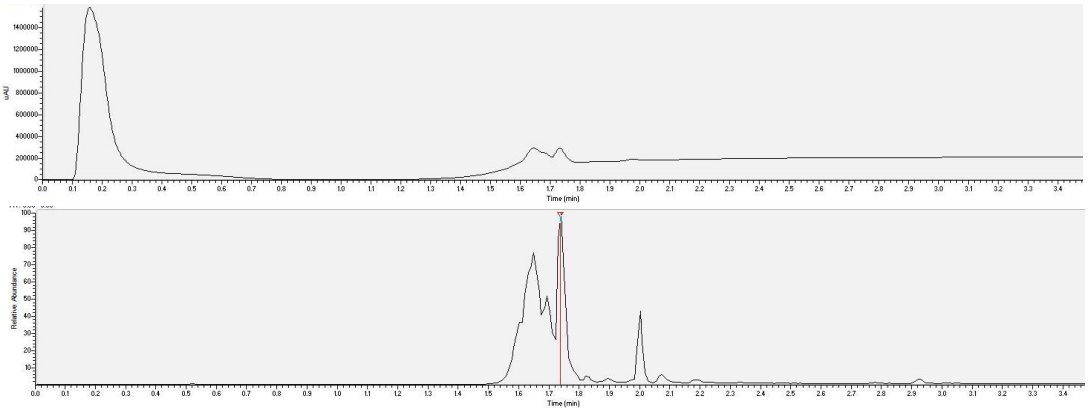


1.5.18.21 NMR and LC/MS data for VK-23 (**23**)

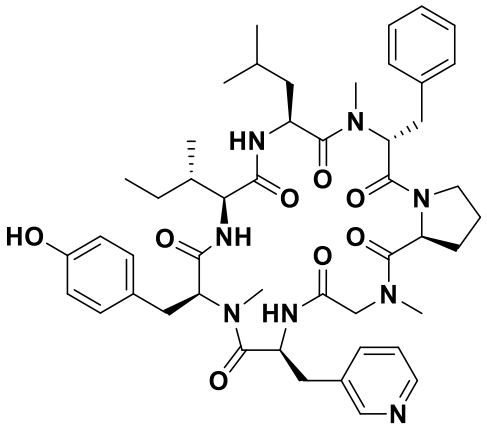


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Ala(β-2-fur)-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.54 (d, *J* = 9.9 Hz, 1H), 8.16 (d, *J* = 9.0 Hz, 1H), 7.44 (s, 1H), 7.26 (s, 3H), 7.15 (dq, *J* = 13.1, 6.9, 6.4 Hz, 5H), 6.70 (s, 3H), 6.40 (dt, *J* = 3.1, 1.5 Hz, 1H), 6.07 (d, *J* = 3.1 Hz, 1H), 5.92 (d, *J* = 9.4 Hz, 1H), 5.51 (dd, *J* = 11.5, 4.8 Hz, 1H), 5.48 – 5.35 (m, 2H), 4.88 (dd, *J* = 11.9, 8.7 Hz, 1H), 4.43 (dd, *J* = 9.3, 3.1 Hz, 1H), 4.39 (dd, *J* = 9.4, 3.4 Hz, 1H), 4.23 (s, 1H), 3.94 (s, 8H), 3.76 (ddd, *J* = 11.9, 7.6, 4.2 Hz, 1H), 3.65 – 3.53 (m, 2H), 3.41 – 3.32 (m, 1H), 3.29 (dd, *J* = 11.8, 8.1 Hz, 1H), 3.23 (dd, *J* = 20.7, 11.9 Hz, 1H), 3.17 – 3.09 (m, 1H), 3.06 (dd, *J* = 12.6, 7.7 Hz, 1H), 3.02 (s, 3H), 2.91 (s, 3H), 2.88 (dd, *J* = 13.1, 8.0 Hz, 1H), 2.82 – 2.71 (m, 4H), 2.50 – 2.43 (m, 1H), 2.32 – 2.24 (m, 1H), 2.11 – 1.99 (m, 1H), 1.89 (s, 1H), 1.84 – 1.77 (m, 1H), 1.27 (q, *J* = 11.7, 9.8 Hz, 3H), 1.01 – 0.93 (m, 1H), 0.88 (dt, *J* = 28.6, 6.7 Hz, 9H), 0.82 – 0.70 (m, 4H), 0.09 (t, *J* = 12.5 Hz, 1H); LCMS (*m/z*): calculated exact mass: 869.4687, experimental $[M+H]^{+}$: 870.4747.



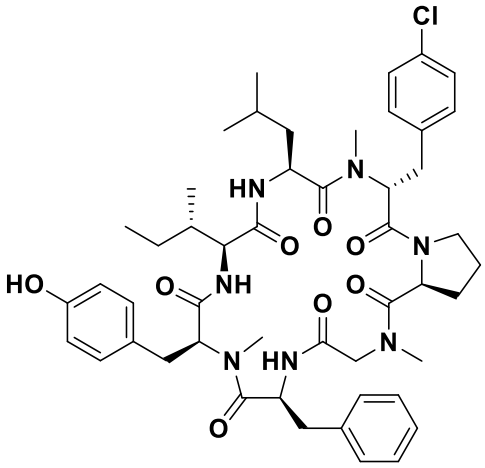
1.5.18.22 NMR and LC/MS data for VK-24 (**24**)



Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Ala(β-3-pyr)-MeTyr-Ile)

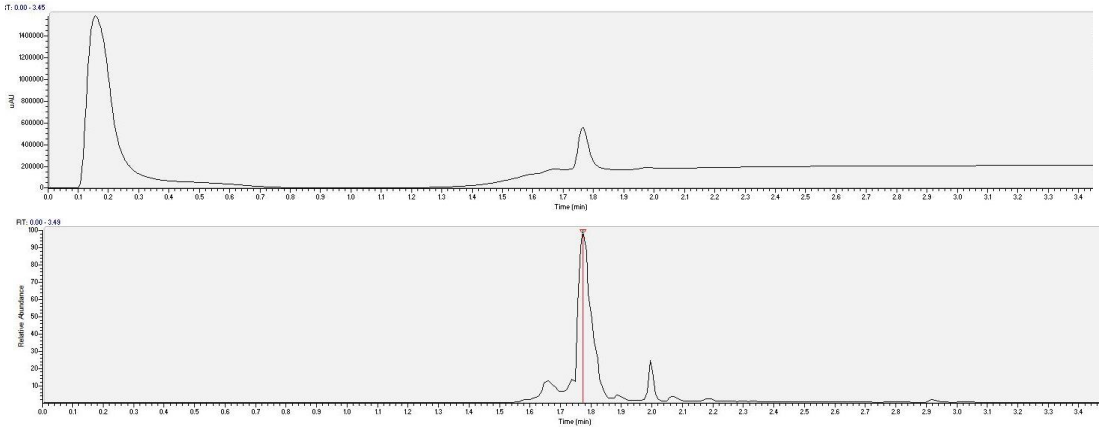
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.80 (d, *J* = 5.2 Hz, 1H), 8.74 (s, 1H), 8.52 (d, *J* = 9.8 Hz, 1H), 8.07 (d, *J* = 7.9 Hz, 1H), 7.80 (d, *J* = 6.9 Hz, 1H), 7.75 (d, *J* = 9.2 Hz, 1H), 7.26 (s, 16H), 7.23 (s, 2H), 7.15 (d, *J* = 7.5 Hz, 2H), 7.06 (t, *J* = 7.7 Hz, 2H), 6.95 (t, *J* = 7.5 Hz, 1H), 6.61 (d, *J* = 8.0 Hz, 2H), 6.33 (d, *J* = 7.8 Hz, 2H), 5.99 (d, *J* = 9.1 Hz, 1H), 5.47 (dd, *J* = 11.3, 4.6 Hz, 1H), 5.38 (d, *J* = 17.6 Hz, 1H), 5.25 – 5.18 (m, 1H), 5.01 (t, *J* = 10.5 Hz, 1H), 4.76 (s, 1H), 4.44 – 4.31 (m, 3H), 4.24 (s, 6H), 3.77 (ddd, *J* = 12.0, 7.7, 4.5 Hz, 1H), 3.66 – 3.54 (m, 2H), 3.49 (s, 1H), 3.43 (d, *J* = 17.0 Hz, 1H), 3.35 (d, *J* = 17.6 Hz, 1H), 3.29 – 3.15 (m, 3H), 3.12 – 3.01 (m, 4H), 2.99 (t, *J* = 11.7 Hz, 2H), 2.94 (s, 3H), 2.88 (dd, *J* = 12.2, 5.3 Hz, 1H), 2.84 – 2.72 (m, 1H), 2.71 (s, 3H), 2.50 (t, *J* = 10.4 Hz, 1H), 1.92 (s, 2H), 1.85 – 1.78 (m, 2H), 1.64 – 1.59 (m, 1H), 1.45 – 1.33 (m, 1H), 1.28 (s, 1H), 1.26 (s, 3H), 1.23 (d, *J* = 3.8 Hz, 1H), 1.09 – 0.98 (m, 1H), 0.98 – 0.80 (m, 14H), 0.40 (t, *J* = 12.4 Hz, 1H); LCMS (*m/z*): calculated exact mass: 880.4847, experimental $[M+H]^{+}$: .

1.5.18.23 NMR and LC/MS data for VK-25 (**25**)

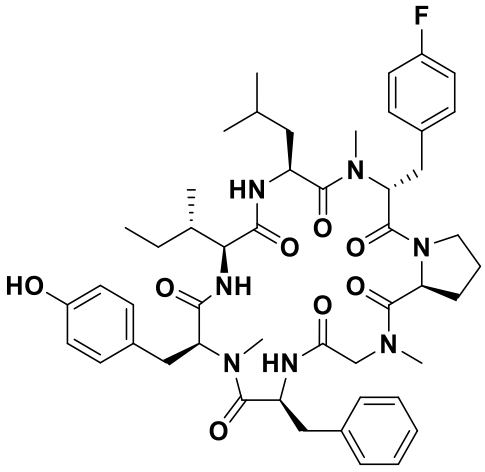


Sequence: Cyclo-(Leu-MeDPhe(4-Cl)-Pro-Sar-Phe-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.56 (d, *J* = 9.8 Hz, 1H), 8.24 (d, *J* = 9.0 Hz, 1H), 7.41 (t, *J* = 7.3 Hz, 2H), 7.35 (dd, *J* = 12.9, 7.1 Hz, 3H), 7.21 (s, 1H), 7.12 (s, 4H), 6.52 (d, *J* = 8.2 Hz, 2H), 6.25 (d, *J* = 8.0 Hz, 2H), 5.88 (d, *J* = 9.5 Hz, 1H), 5.55 (dd, *J* = 11.5, 5.0 Hz, 1H), 5.44 – 5.34 (m, 2H), 4.94 (ddd, *J* = 11.9, 9.0, 2.4 Hz, 1H), 4.45 (dd, *J* = 9.5, 3.1 Hz, 1H), 4.37 (dd, *J* = 9.3, 3.4 Hz, 1H), 3.76 (ddd, *J* = 12.1, 7.7, 4.3 Hz, 1H), 3.69 – 3.56 (m, 1H), 3.41 (dd, *J* = 11.7, 4.1 Hz, 1H), 3.35 (d, *J* = 17.4 Hz, 1H), 3.27 (dd, *J* = 15.1, 11.5 Hz, 1H), 3.18 (t, *J* = 12.0 Hz, 1H), 3.09 (t, *J* = 12.1 Hz, 1H), 3.05 – 2.97 (m, 1H), 3.01 (s, 3H), 2.90 (s, 3H), 2.84 (dd, *J* = 12.6, 3.9 Hz, 1H), 2.77 – 2.66 (m, 2H), 2.64 (s, 3H), 2.50 – 2.40 (m, 1H), 2.32 (ddt, *J* = 10.3, 7.0, 3.6 Hz, 1H), 2.07 – 2.00 (m, 1H), 1.88 (ddt, *J* = 11.3, 7.0, 4.4 Hz, 1H), 1.85 – 1.74 (m, 1H), 1.58 (ddd, *J* = 13.8, 6.7, 3.5 Hz, 1H), 1.48 – 1.29 (m, 2H), 1.28 (s, 4H), 1.25 (s, 2H), 1.04 – 0.87 (m, 9H), 0.85 (d, *J* = 6.5 Hz, 3H), 0.84 (s, 7H), 0.12 (ddd, *J* = 13.8, 10.6, 2.5 Hz, 1H); LCMS (*m/z*): calculated exact mass: 913.4505, experimental $[M+H]^+$: 914.4579.

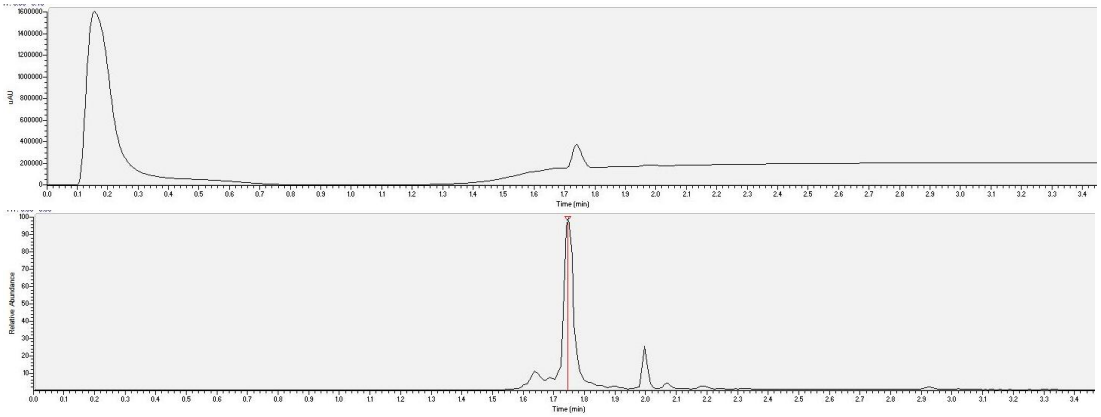


1.5.18.24 NMR and LC/MS data for VK-26 (**26**)

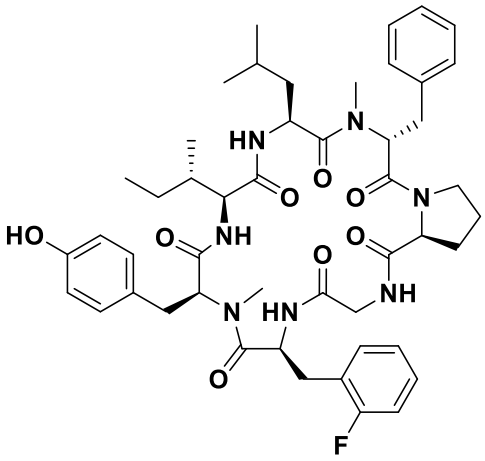


Sequence: Cyclo-(Leu-MeDPhe(4-F)-Pro-Sar-Phe-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.57 (d, *J* = 9.8 Hz, 1H), 8.24 (d, *J* = 9.0 Hz, 1H), 7.40 (t, *J* = 7.3 Hz, 2H), 7.36 (d, *J* = 7.0 Hz, 1H), 7.32 (d, *J* = 7.4 Hz, 2H), 7.13 (dd, *J* = 8.3, 5.3 Hz, 2H), 6.84 (t, *J* = 8.4 Hz, 2H), 6.53 (d, *J* = 8.1 Hz, 2H), 6.26 (d, *J* = 8.1 Hz, 2H), 5.89 (d, *J* = 9.5 Hz, 1H), 5.53 (dd, *J* = 11.4, 5.1 Hz, 1H), 5.41 – 5.33 (m, 2H), 4.94 (ddd, *J* = 11.9, 9.0, 2.5 Hz, 1H), 4.52 (s, 6H), 4.45 (dd, *J* = 9.5, 3.1 Hz, 1H), 4.38 (dd, *J* = 9.3, 3.4 Hz, 1H), 3.76 (ddd, *J* = 12.0, 7.8, 4.1 Hz, 1H), 3.69 – 3.56 (m, 1H), 3.42 (dd, *J* = 11.6, 4.1 Hz, 1H), 3.35 (d, *J* = 17.4 Hz, 1H), 3.26 (dd, *J* = 15.1, 11.4 Hz, 1H), 3.18 (t, *J* = 12.0 Hz, 1H), 3.14 – 2.97 (m, 2H), 3.02 (s, 3H), 2.90 (s, 3H), 2.82 (dd, *J* = 12.6, 3.9 Hz, 1H), 2.74 (dd, *J* = 12.3, 4.2 Hz, 1H), 2.64 (s, 3H), 2.51 – 2.40 (m, 1H), 2.33 (ddt, *J* = 10.4, 7.1, 3.7 Hz, 1H), 2.03 (ddt, *J* = 11.0, 7.1, 3.6 Hz, 1H), 1.94 – 1.73 (m, 2H), 1.66 – 1.54 (m, 1H), 1.50 – 1.29 (m, 2H), 1.26 (d, *J* = 8.6 Hz, 2H), 1.24 (s, 1H), 1.06 – 0.87 (m, 9H), 0.85 (d, *J* = 6.4 Hz, 3H), 0.15 (ddd, *J* = 13.8, 10.8, 2.5 Hz, 1H); LCMS (*m/z*): calculated exact mass: 897.4800, experimental $[M+H]^+$: 898.4867.

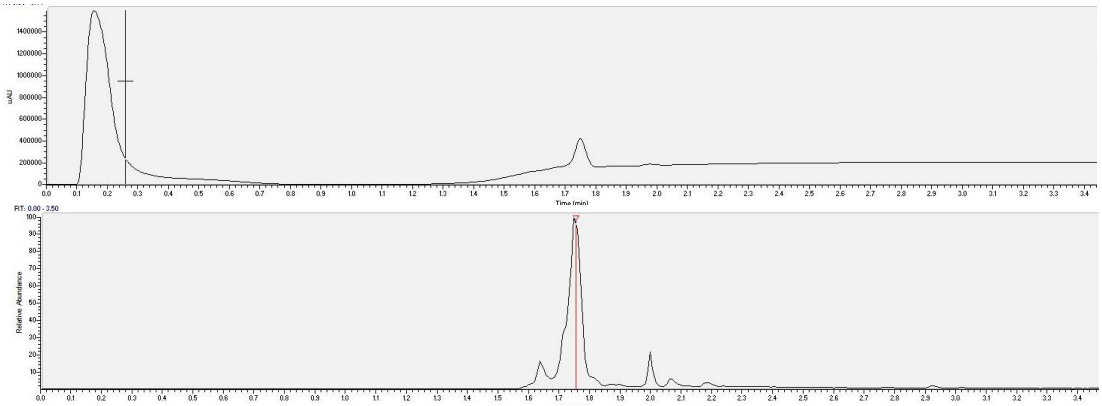


1.5.18.25 NMR and LC/MS data for VK-27 (**27**)

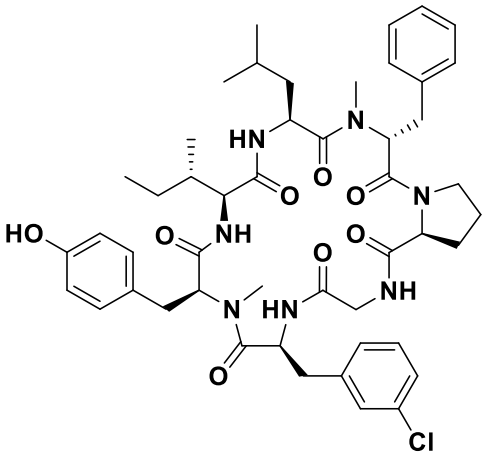


Sequence: Cyclo-(Leu-MeDPhe-Pro-Gly-Phe(2-F)-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.43 (d, *J* = 9.9 Hz, 1H), 7.96 (d, *J* = 8.4 Hz, 1H), 7.34 (dtd, *J* = 8.0, 6.0, 5.5, 2.6 Hz, 1H), 7.29 (d, *J* = 15.2 Hz, 1H), 7.26 – 7.19 (m, 1H), 7.19 – 7.12 (m, 6H), 7.12 – 6.95 (m, 3H), 6.75 (d, *J* = 8.4 Hz, 1H), 6.64 (d, *J* = 8.1 Hz, 2H), 6.58 (d, *J* = 8.1 Hz, 2H), 6.05 (d, *J* = 8.1 Hz, 1H), 5.53 (dd, *J* = 11.3, 4.8 Hz, 1H), 5.45 (td, *J* = 9.9, 5.0 Hz, 1H), 5.02 – 4.93 (m, 1H), 4.75 (td, *J* = 9.3, 4.6 Hz, 1H), 4.41 (s, 1H), 4.23 (dd, *J* = 8.1, 3.2 Hz, 1H), 4.17 (dd, *J* = 8.9, 3.2 Hz, 1H), 3.80 – 3.67 (m, 2H), 3.61 – 3.44 (m, 3H), 3.25 (ddd, *J* = 30.8, 13.8, 10.9 Hz, 2H), 3.12 (d, *J* = 21.6 Hz, 5H), 3.09 – 3.03 (m, 1H), 2.96 (dd, *J* = 12.5, 5.5 Hz, 1H), 2.92 – 2.85 (m, 2H), 2.84 – 2.73 (m, 1H), 2.67 (d, *J* = 12.3 Hz, 4H), 2.36 – 2.22 (m, 2H), 2.05 – 1.96 (m, 2H), 1.91 – 1.78 (m, 1H), 1.83 (s, 2H), 1.54 (ddd, *J* = 14.8, 10.0, 5.6 Hz, 1H), 1.50 – 1.34 (m, 2H), 1.30 (s, 1H), 1.28 (s, 4H), 1.28 (s, 4H), 1.23 (tt, *J* = 8.1, 4.0 Hz, 1H), 1.08 – 0.91 (m, 4H), 0.94 – 0.72 (m, 13H); LCMS (*m/z*): calculated exact mass: 883.4644, experimental $[M+H]^+$: 884.4713.

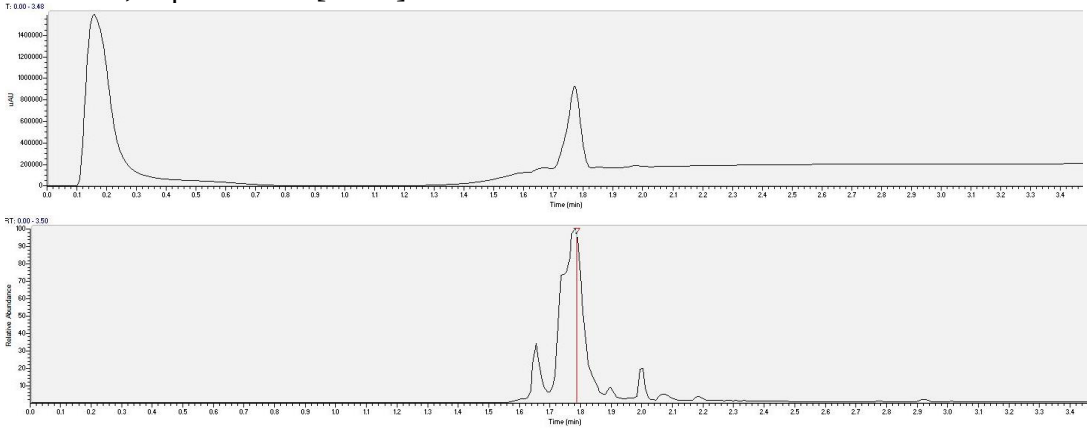


1.5.18.26 NMR and LC/MS data for VK-28 (**28**)

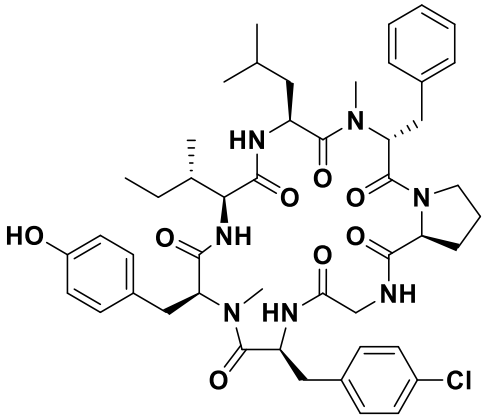


Sequence: Cyclo-(Leu-MeDPhe-Pro-Gly-Phe(3-Cl)-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.45 (d, *J* = 9.8 Hz, 1H), 7.93 (d, *J* = 8.5 Hz, 1H), 7.35 (s, 1H), 7.30 (t, *J* = 3.7 Hz, 4H), 7.19 (q, *J* = 5.9, 4.4 Hz, 2H), 7.14 (q, *J* = 3.2, 1.9 Hz, 5H), 7.07 (td, *J* = 5.8, 2.8 Hz, 1H), 6.91 (d, *J* = 5.9 Hz, 0H), 6.61 (d, *J* = 8.0 Hz, 2H), 6.52 (d, *J* = 8.1 Hz, 2H), 6.05 (d, *J* = 8.2 Hz, 1H), 5.48 (dd, *J* = 11.5, 4.6 Hz, 1H), 5.24 (td, *J* = 10.6, 4.1 Hz, 1H), 4.97 (dd, *J* = 17.9, 9.9 Hz, 1H), 4.75 (td, *J* = 9.7, 4.1 Hz, 1H), 4.24 (dd, *J* = 8.3, 3.1 Hz, 1H), 4.18 (dd, *J* = 8.9, 3.3 Hz, 1H), 3.73 (ddd, *J* = 12.1, 7.7, 4.5 Hz, 1H), 3.60 – 3.37 (m, 4H), 3.29 – 3.14 (m, 2H), 3.12 (s, 3H), 3.10 (s, 2H), 3.03 (ddd, *J* = 15.6, 12.0, 7.7 Hz, 3H), 2.94 – 2.84 (m, 2H), 2.73 (dd, *J* = 13.2, 4.2 Hz, 1H), 2.67 (s, 1H), 2.56 (s, 3H), 2.27 (dddd, *J* = 13.7, 10.9, 8.4, 6.0 Hz, 2H), 2.07 – 1.94 (m, 2H), 1.89 – 1.75 (m, 3H), 1.48 (ddt, *J* = 17.4, 11.9, 6.0 Hz, 1H), 1.45 – 1.29 (m, 1H), 1.31 (s, 3H), 1.25 (s, 1H), 1.19 (s, 1H), 1.02 (ddt, *J* = 13.0, 9.5, 6.7 Hz, 1H), 0.92 (dd, *J* = 22.2, 7.0 Hz, 10H), 0.85 (dd, *J* = 10.5, 6.5 Hz, 1H), 0.80 (d, *J* = 6.5 Hz, 4H), 0.62 (ddd, *J* = 13.6, 9.0, 4.2 Hz, 1H); LCMS (*m/z*): calculated exact mass: 899.4348, experimental $[M+H]^+$: 900.4423.

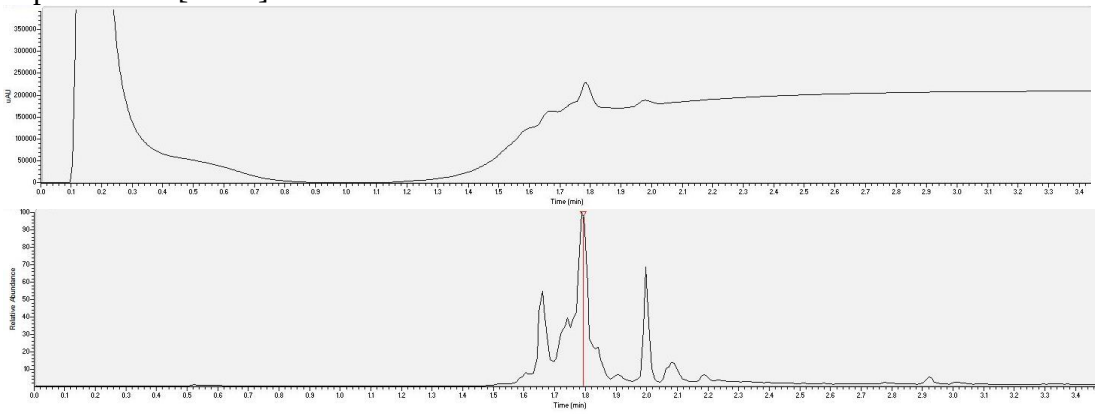


1.5.18.27 NMR and LC/MS data for VK-29 (**29**)

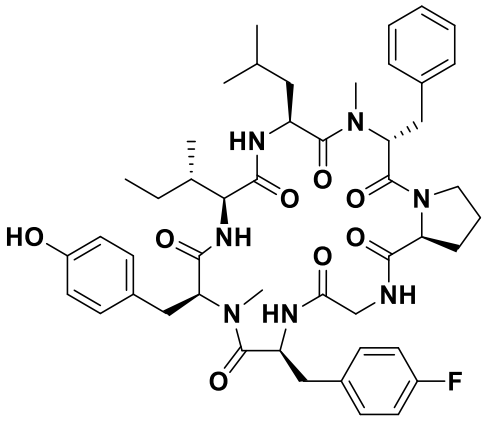


Sequence: Cyclo-(Leu-MeDPhe-Pro-Gly-Phe(4-Cl)-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.44 (d, *J* = 9.9 Hz, 1H), 7.93 (d, *J* = 8.4 Hz, 1H), 7.35 (d, *J* = 8.1 Hz, 2H), 7.31 – 7.18 (m, 3H), 7.20 – 7.11 (m, 1H), 7.14 (s, 2H), 7.12 (d, *J* = 7.7 Hz, 1H), 7.09 – 6.97 (m, 2H), 6.89 (d, *J* = 8.1 Hz, 1H), 6.60 (d, *J* = 8.1 Hz, 2H), 6.49 (d, *J* = 8.1 Hz, 2H), 6.03 (d, *J* = 8.3 Hz, 1H), 5.49 (dd, *J* = 11.5, 4.7 Hz, 1H), 5.24 (td, *J* = 10.6, 4.3 Hz, 1H), 5.01 – 4.92 (m, 1H), 4.73 (td, *J* = 9.4, 4.5 Hz, 1H), 4.39 (s, 1H), 4.23 (dd, *J* = 8.2, 3.1 Hz, 1H), 4.16 (dd, *J* = 8.8, 3.3 Hz, 1H), 3.73 (ddd, *J* = 12.2, 7.8, 4.5 Hz, 1H), 3.60 – 3.48 (m, 2H), 3.45 (dd, *J* = 10.7, 5.2 Hz, 1H), 3.41 (s, 1H), 3.28 – 3.14 (m, 2H), 3.12 (s, 3H), 3.04 (ddd, *J* = 17.1, 12.9, 7.8 Hz, 3H), 2.95 – 2.85 (m, 2H), 2.74 (dd, *J* = 13.1, 4.4 Hz, 1H), 2.57 (s, 3H), 2.32 – 2.21 (m, 2H), 2.04 – 1.95 (m, 2H), 1.83 (ddd, *J* = 14.9, 11.0, 7.6 Hz, 2H), 1.53 – 1.42 (m, 1H), 1.42 – 1.32 (m, 1H), 1.29 (dd, *J* = 11.0, 4.6 Hz, 1H), 1.26 (s, 3H), 1.01 (ddt, *J* = 12.4, 9.3, 6.6 Hz, 1H), 0.92 (dd, *J* = 15.8, 8.5 Hz, 5H), 0.90 – 0.79 (m, 7H), 0.78 (d, *J* = 6.6 Hz, 3H), 0.67 (ddd, *J* = 13.5, 8.7, 4.4 Hz, 1H); LCMS (*m/z*): calculated exact mass: 899.4348, experimental $[M+H]^{+}$: 900.4426.

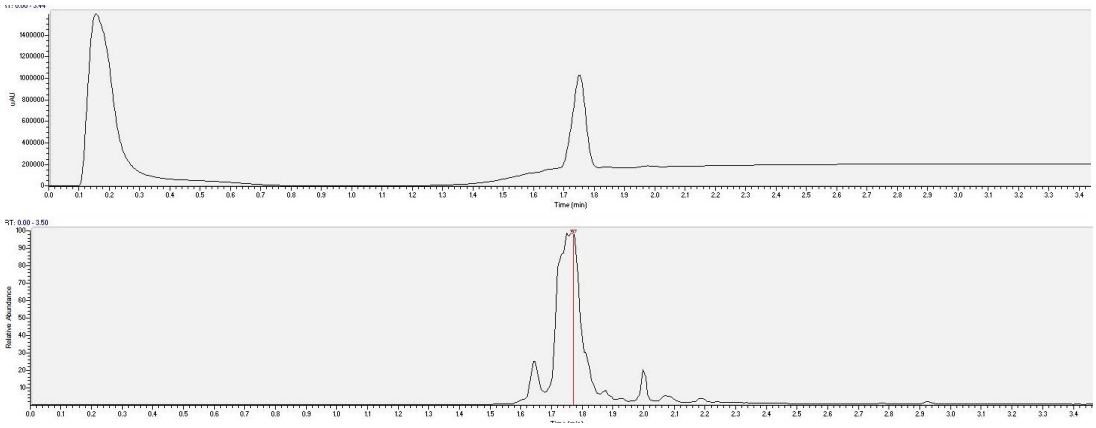


1.5.18.28 NMR and LC/MS data for VK-30 (**30**)

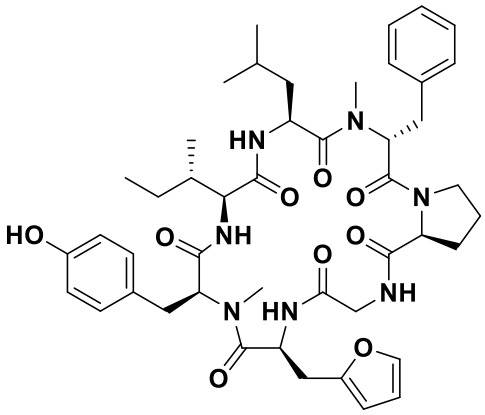


Sequence: Cyclo-(Leu-MeDPhe-Pro-Gly-Phe(4-F)-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.42 (d, *J* = 9.9 Hz, 1H), 7.94 (d, *J* = 8.5 Hz, 1H), 7.30 – 7.25 (m, 2H), 7.25 – 7.18 (m, 0H), 7.20 – 7.11 (m, 4H), 7.13 – 7.03 (m, 3H), 6.91 (d, *J* = 7.0 Hz, 1H), 6.58 (d, *J* = 8.0 Hz, 2H), 6.48 (d, *J* = 8.0 Hz, 2H), 6.04 (d, *J* = 8.3 Hz, 1H), 5.49 (dd, *J* = 11.5, 4.7 Hz, 1H), 5.22 (td, *J* = 10.5, 4.3 Hz, 1H), 4.97 (dd, *J* = 17.9, 9.8 Hz, 1H), 4.75 (td, *J* = 9.5, 4.4 Hz, 1H), 4.24 (dd, *J* = 8.3, 3.1 Hz, 1H), 4.17 (dd, *J* = 8.9, 3.3 Hz, 1H), 3.72 (ddd, *J* = 12.1, 7.7, 4.6 Hz, 1H), 3.59 – 3.46 (m, 2H), 3.43 (dd, *J* = 10.8, 5.1 Hz, 1H), 3.29 – 3.14 (m, 2H), 3.12 (s, 2H), 3.08 (s, 0H), 3.03 (dd, *J* = 15.1, 11.3 Hz, 2H), 2.93 – 2.84 (m, 1H), 2.73 (dt, *J* = 12.5, 6.1 Hz, 1H), 2.53 (s, 2H), 2.26 (ddd, *J* = 13.4, 9.6, 6.6 Hz, 2H), 2.02 – 1.95 (m, 1H), 1.90 – 1.76 (m, 2H), 1.53 – 1.35 (m, 1H), 1.34 – 1.24 (m, 2H), 1.02 (ddt, *J* = 13.0, 9.2, 6.6 Hz, 1H), 0.89 (dtd, *J* = 22.3, 15.7, 15.0, 6.6 Hz, 9H), 0.79 (t, *J* = 6.3 Hz, 3H), 0.65 (ddd, *J* = 13.6, 8.8, 4.5 Hz, 1H); LCMS (*m/z*): calculated exact mass: 883.4644, experimental $[M+H]^{+}$: 884.4712.



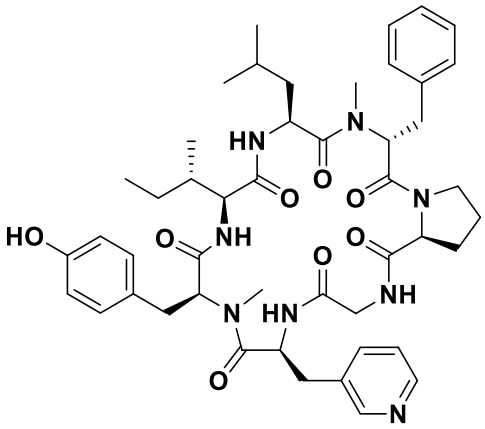
1.5.18.29 NMR and LC/MS data for VK-31 (**31**)



Sequence: Cyclo-(Leu-MeDPhe-Pro-Gly-Ala(β-2-fur)-MeTyr-Ile)

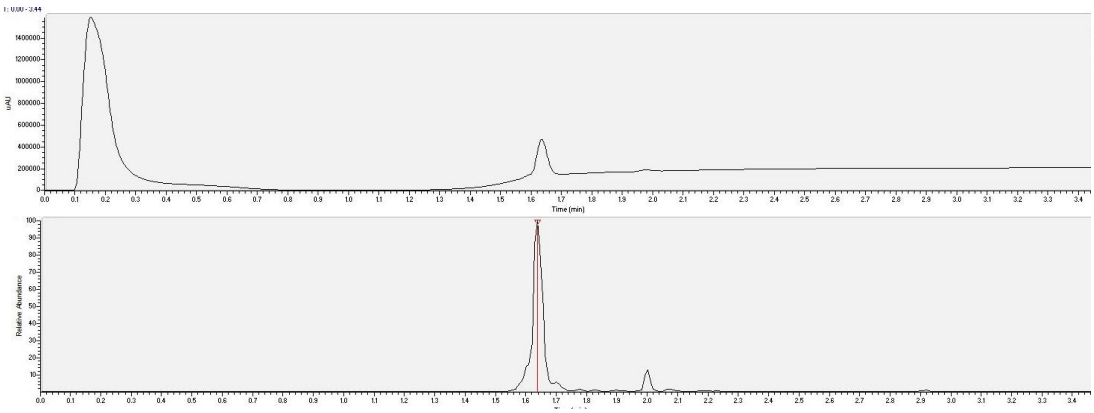
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.38 (d, *J* = 10.1 Hz, 1H), 7.89 (t, *J* = 7.4 Hz, 1H), 7.41 (s, 1H), 7.20 (s, 4H), 7.20 – 7.14 (m, 1H), 7.11 (d, *J* = 6.8 Hz, 6H), 7.00 (qt, *J* = 18.2, 8.7 Hz, 3H), 6.86 – 6.80 (m, 3H), 6.74 (t, *J* = 8.9 Hz, 4H), 6.37 (dd, *J* = 3.3, 1.8 Hz, 1H), 6.08 (d, *J* = 3.1 Hz, 1H), 6.01 (dd, *J* = 14.1, 8.4 Hz, 2H), 5.49 (dd, *J* = 12.1, 4.4 Hz, 1H), 5.41 (td, *J* = 10.0, 5.4 Hz, 1H), 4.99 – 4.87 (m, 1H), 4.57 (dt, *J* = 12.8, 6.5 Hz, 1H), 4.53 (s, 1H), 4.41 (s, 1H), 4.14 (ddd, *J* = 18.6, 8.5, 3.2 Hz, 1H), 4.09 – 4.01 (m, 1H), 3.96 (d, *J* = 5.9 Hz, 1H), 3.89 – 3.65 (m, 3H), 3.58 (d, *J* = 19.6 Hz, 1H), 3.56 – 3.52 (m, 1H), 3.52 (s, 1H), 3.50 (d, *J* = 9.6 Hz, 1H), 3.41 – 3.30 (m, 1H), 3.22 – 3.13 (m, 2H), 3.11 (d, *J* = 9.7 Hz, 6H), 3.07 (s, 7H), 3.05 – 2.95 (m, 2H), 2.92 (s, 1H), 2.86 (d, *J* = 11.4 Hz, 4H), 2.83 – 2.77 (m, 1H), 2.74 (s, 2H), 2.65 (s, 2H), 2.28 (dd, *J* = 12.2, 8.5 Hz, 1H), 2.06 – 1.99 (m, 1H), 1.99 (s, 4H), 1.92 (p, *J* = 7.1 Hz, 1H), 1.79 (s, 8H), 1.63 (q, *J* = 7.6 Hz, 1H), 1.46 – 1.27 (m, 3H), 1.25 (s, 4H), 1.24 (s, 6H), 1.20 (dd, *J* = 9.3, 6.3 Hz, 1H), 1.13 (dt, *J* = 13.4, 6.8 Hz, 1H), 0.89 (dt, *J* = 11.8, 6.7 Hz, 10H), 0.88 – 0.80 (m, 8H), 0.83 – 0.67 (m, 12H) ; LCMS (*m/z*): calculated exact mass: 855.4531, experimental $[M+H]^{+}$:

1.5.18.30 NMR and LC/MS data for VK-32 (**32**)

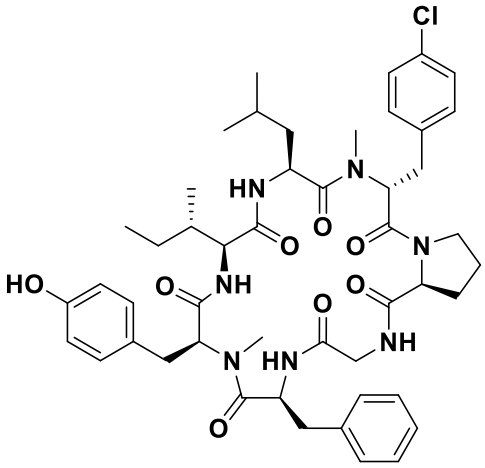


Sequence: Cyclo-(Leu-MeDPhe-Pro-Gly-Ala(β-3-pyr)-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.55 (d, *J* = 5.7 Hz, 1H), 8.40 – 8.31 (m, 2H), 7.80 (dd, *J* = 8.0, 5.7 Hz, 1H), 7.36 – 7.26 (m, 2H), 7.26 (s, 4H), 7.21 (t, *J* = 8.8 Hz, 4H), 6.93 (d, *J* = 8.1 Hz, 2H), 6.87 (d, *J* = 8.1 Hz, 2H), 6.46 (d, *J* = 9.6 Hz, 1H), 6.26 (d, *J* = 8.2 Hz, 1H), 5.90 (d, *J* = 3.9 Hz, 1H), 5.77 (s, 1H), 4.47 (ddt, *J* = 21.5, 13.4, 4.2 Hz, 3H), 4.25 – 4.15 (m, 4H), 4.01 – 3.90 (m, 2H), 3.88 (s, 1H), 3.53 – 3.35 (m, 3H), 3.19 – 3.02 (m, 4H), 2.95 (s, 3H), 2.90 (s, 3H), 2.83 (dd, *J* = 16.6, 6.9 Hz, 1H), 2.82 – 2.71 (m, 1H), 2.53 (dd, *J* = 14.1, 4.2 Hz, 1H), 2.28 – 2.18 (m, 1H), 2.01 (d, *J* = 16.8 Hz, 1H), 1.95 – 1.87 (m, 1H), 1.85 – 1.78 (m, 0H), 1.56 (t, *J* = 7.4 Hz, 1H), 1.49 – 1.34 (m, 3H), 1.27 – 1.10 (m, 2H), 1.05 (ddd, *J* = 13.7, 8.3, 5.0 Hz, 1H), 0.97 – 0.89 (m, 7H), 0.86 (d, *J* = 6.4 Hz, 4H), 0.80 (d, *J* = 6.4 Hz, 4H) ; LCMS (*m/z*): calculated exact mass: 866.4691, experimental $[M+H]^+$: 867.4749.

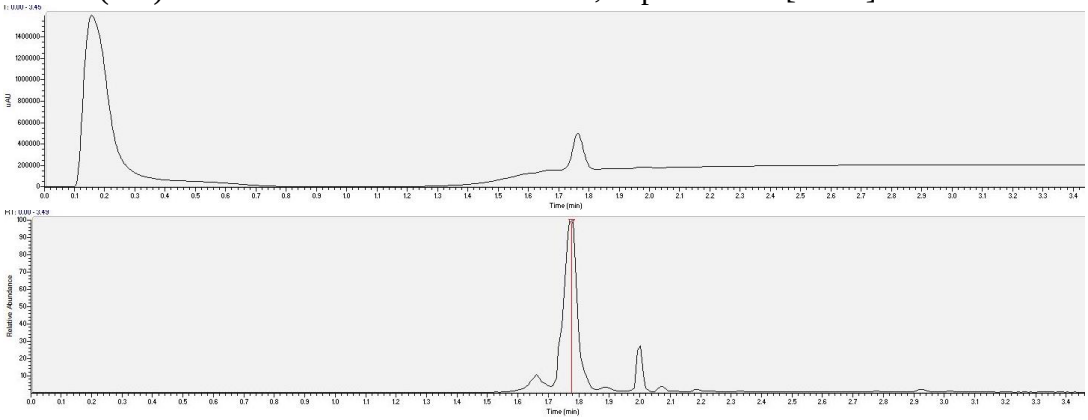


1.5.18.31 NMR and LC/MS data for VK-33 (**33**)

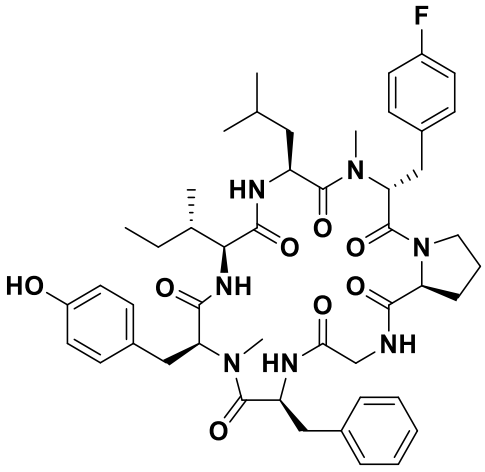


Sequence: Cyclo-(Leu-MeDPhe(4-Cl)-Pro-Gly-Phe-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.40 (d, *J* = 9.9 Hz, 1H), 8.04 (d, *J* = 8.4 Hz, 1H), 7.41 – 7.34 (m, 2H), 7.32 (q, *J* = 4.0, 3.4 Hz, 3H), 7.29 – 7.21 (m, 1H), 7.19 (s, 1H), 7.17 (d, *J* = 9.9 Hz, 1H), 7.11 (q, *J* = 8.2 Hz, 4H), 6.97 (dd, *J* = 15.9, 7.8 Hz, 1H), 6.58 (d, *J* = 8.0 Hz, 2H), 6.45 (d, *J* = 8.0 Hz, 2H), 6.04 (d, *J* = 8.2 Hz, 1H), 5.49 (dd, *J* = 11.3, 4.9 Hz, 1H), 5.34 (td, *J* = 10.4, 4.4 Hz, 1H), 4.95 (dd, *J* = 17.8, 9.8 Hz, 1H), 4.75 (td, *J* = 9.3, 4.7 Hz, 1H), 4.23 (dd, *J* = 8.3, 3.1 Hz, 1H), 4.15 (dd, *J* = 8.9, 3.4 Hz, 1H), 3.84 (s, 6H), 3.72 (ddd, *J* = 12.0, 7.7, 4.4 Hz, 1H), 3.57 (d, *J* = 17.3 Hz, 1H), 3.56 – 3.46 (m, 1H), 3.41 (dd, *J* = 10.8, 5.0 Hz, 1H), 3.29 – 3.11 (m, 2H), 3.10 (s, 3H), 3.09 – 2.97 (m, 2H), 2.86 (td, *J* = 15.9, 14.3, 4.8 Hz, 2H), 2.55 (s, 3H), 2.26 (dtd, *J* = 11.7, 8.8, 8.2, 5.8 Hz, 2H), 2.07 (dq, *J* = 14.1, 6.1 Hz, 0H), 2.02 – 1.94 (m, 1H), 1.82 (tq, *J* = 15.3, 5.9 Hz, 2H), 1.57 (ddd, *J* = 14.8, 10.1, 5.5 Hz, 1H), 1.46 – 1.33 (m, 2H), 1.35 – 1.21 (m, 3H), 1.01 (ddt, *J* = 12.8, 9.4, 6.7 Hz, 1H), 0.93 (dt, *J* = 7.4, 4.0 Hz, 6H), 0.89 (d, *J* = 7.0 Hz, 3H), 0.83 (t, *J* = 7.1 Hz, 4H), 0.73 (ddd, *J* = 13.7, 8.6, 4.8 Hz, 1H); LCMS (*m/z*): calculated exact mass: 899.4348, experimental $[M+H]^+$: 900.4406.

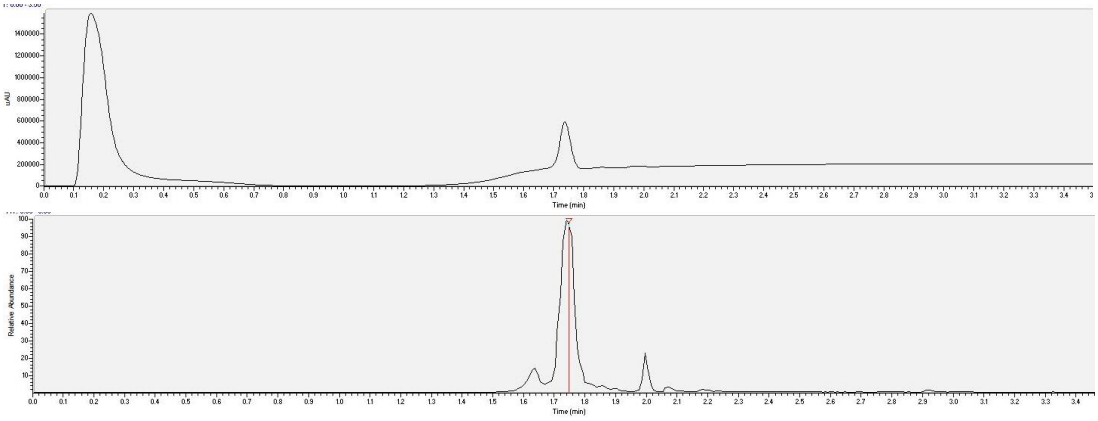


1.5.18.32 NMR and LC/MS data for VK-34 (**34**)

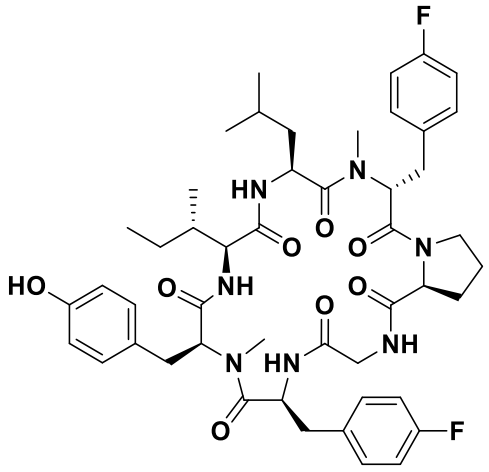


Sequence: Cyclo-(Leu-MeDPhe(4-F)-Pro-Gly-Phe-MeTyr-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.41 (d, *J* = 9.9 Hz, 1H), 8.04 (d, *J* = 8.5 Hz, 1H), 7.34 (dt, *J* = 28.3, 7.5 Hz, 4H), 7.26 – 7.20 (m, 1H), 7.11 (dd, *J* = 8.3, 5.3 Hz, 2H), 6.96 (q, *J* = 9.2, 8.5 Hz, 1H), 6.87 – 6.71 (m, 2H), 6.57 (d, *J* = 8.0 Hz, 2H), 6.45 (d, *J* = 8.0 Hz, 2H), 6.05 (d, *J* = 8.2 Hz, 1H), 5.48 (dd, *J* = 11.3, 5.0 Hz, 1H), 5.31 (td, *J* = 10.4, 4.5 Hz, 1H), 4.95 (dd, *J* = 17.9, 9.9 Hz, 1H), 4.76 (td, *J* = 9.3, 4.9 Hz, 1H), 4.23 (dd, *J* = 8.3, 3.1 Hz, 1H), 4.15 (dd, *J* = 8.9, 3.3 Hz, 1H), 3.71 (ddd, *J* = 12.0, 7.7, 4.5 Hz, 1H), 3.56 (d, *J* = 17.5 Hz, 1H), 3.56 – 3.45 (m, 1H), 3.40 (dd, *J* = 10.8, 5.1 Hz, 1H), 3.27 – 3.20 (m, 1H), 3.20 – 3.12 (m, 1H), 3.10 (s, 3H), 3.06 – 2.95 (m, 1H), 2.91 – 2.79 (m, 2H), 2.53 (s, 2H), 2.26 (dqd, *J* = 13.5, 9.4, 8.4, 5.4 Hz, 2H), 1.97 (ddt, *J* = 12.4, 7.8, 4.1 Hz, 1H), 1.89 – 1.72 (m, 2H), 1.57 (ddd, *J* = 14.7, 10.0, 5.5 Hz, 1H), 1.46 – 1.30 (m, 1H), 1.24 (h, *J* = 6.6 Hz, 1H), 1.02 (ddt, *J* = 12.7, 9.1, 6.6 Hz, 1H), 0.96 – 0.83 (m, 9H), 0.85 – 0.72 (m, 4H); LCMS (*m/z*): calculated exact mass: 883.4644, experimental $[M+H]^{+}$: 884.4707.



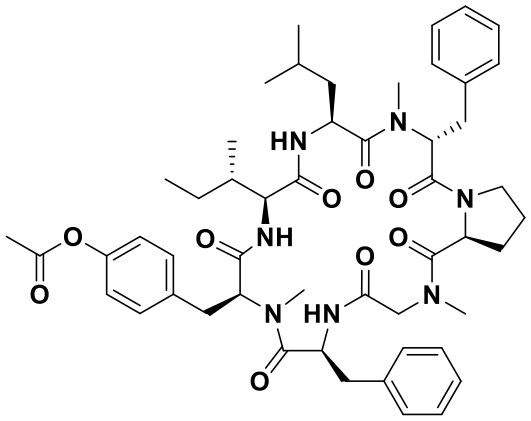
1.5.18.33 NMR and LC/MS data for VK-35 (**35**)



Sequence: Cyclo-(Leu-MeDPhe(4-F)-Pro-Gly-Phe(4-F)-MeTyr-Ile)

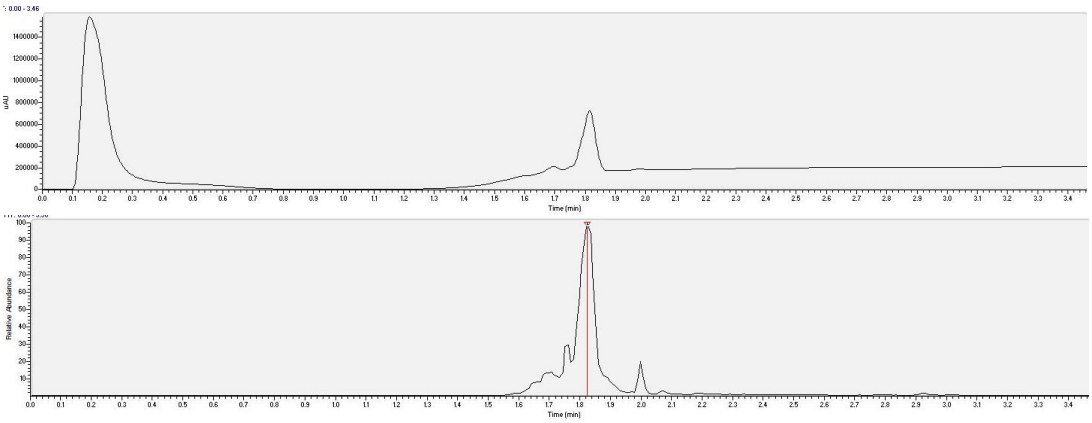
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.34 (dd, *J* = 10.0, 3.8 Hz, 1H), 8.00 – 7.95 (m, 1H), 7.26 (d, *J* = 13.8 Hz, 1H), 7.18 (dd, *J* = 20.6, 7.7 Hz, 1H), 7.12 (d, *J* = 5.4 Hz, 1H), 7.12 – 6.98 (m, 3H), 7.00 – 6.87 (m, 2H), 6.85 – 6.73 (m, 3H), 6.65 – 6.57 (m, 2H), 6.53 (d, *J* = 8.3 Hz, 1H), 6.02 (d, *J* = 8.2 Hz, 1H), 5.48 (dt, *J* = 11.4, 4.0 Hz, 1H), 5.26 (tt, *J* = 9.8, 4.7 Hz, 1H), 4.92 (dd, *J* = 17.9, 10.0 Hz, 1H), 4.75 (qd, *J* = 9.3, 5.0 Hz, 1H), 4.41 (q, *J* = 7.1 Hz, 1H), 4.40 (s, 6H), 4.20 (td, *J* = 8.8, 8.4, 3.2 Hz, 1H), 4.14 (td, *J* = 9.1, 3.2 Hz, 1H), 3.76 – 3.66 (m, 1H), 3.71 (s, 1H), 3.58 – 3.44 (m, 3H), 3.27 (dd, *J* = 12.6, 9.6 Hz, 1H), 3.24 – 3.13 (m, 1H), 3.11 (d, *J* = 1.3 Hz, 4H), 3.02 (ddd, *J* = 19.2, 10.6, 5.5 Hz, 2H), 2.98 – 2.90 (m, 1H), 2.90 – 2.74 (m, 2H), 2.71 (s, 1H), 2.65 (d, *J* = 14.5 Hz, 3H), 2.36 – 2.26 (m, 1H), 2.04 (s, 0H), 1.83 (ddd, *J* = 17.4, 11.1, 6.9 Hz, 1H), 1.59 – 1.48 (m, 0H), 1.48 – 1.39 (m, 1H), 1.42 – 1.36 (m, 1H), 1.39 – 1.29 (m, 7H), 1.32 (s, 6H), 1.32 – 1.27 (m, 1H), 1.26 (s, 3H), 1.24 (s, 1H), 1.09 – 0.95 (m, 1H), 0.98 – 0.88 (m, 6H), 0.88 (s, 1H), 0.88 – 0.81 (m, 4H), 0.80 (dd, *J* = 6.5, 1.2 Hz, 3H) ; LCMS (*m/z*): calculated exact mass: 901.4550, experimental $[M+H]^{+}$:

1.5.18.34 NMR and LC/MS data for VK-36 (**36**)

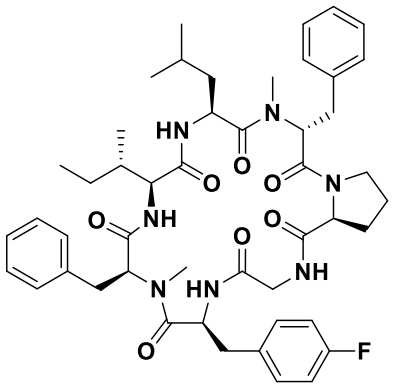


Sequence: Cyclo-(Leu-MeDPhe-Pro-Sar-Phe-MeTyr(OAc)-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.51 (d, *J* = 9.9 Hz, 1H), 8.11 (d, *J* = 9.1 Hz, 1H), 7.77 (d, *J* = 7.5 Hz, 1H), 7.54 (d, *J* = 7.5 Hz, 1H), 7.45 – 7.29 (m, 8H), 7.19 – 7.10 (m, 4H), 7.12 – 7.00 (m, 1H), 6.79 – 6.73 (m, 2H), 6.40 – 6.34 (m, 2H), 5.85 (d, *J* = 9.6 Hz, 1H), 5.55 (dd, *J* = 11.6, 4.8 Hz, 1H), 5.47 – 5.29 (m, 2H), 5.02 (d, *J* = 4.7 Hz, 1H), 4.94 (ddd, *J* = 11.9, 9.1, 2.4 Hz, 1H), 4.69 (s, 2H), 4.45 (dd, *J* = 9.6, 3.1 Hz, 1H), 4.38 (dd, *J* = 9.3, 3.5 Hz, 1H), 4.23 (d, *J* = 5.5 Hz, 1H), 3.77 (ddd, *J* = 12.0, 7.7, 4.3 Hz, 1H), 3.60 (dt, *J* = 11.8, 7.6 Hz, 1H), 3.44 (dd, *J* = 11.8, 3.9 Hz, 1H), 3.36 – 3.21 (m, 3H), 3.04 (s, 3H), 3.12 – 2.99 (m, 2H), 2.90 (s, 3H), 2.79 (dt, *J* = 12.1, 4.1 Hz, 2H), 2.67 (s, 3H), 2.47 (dtd, *J* = 12.6, 9.6, 7.1 Hz, 1H), 2.34 (tt, *J* = 6.8, 3.3 Hz, 1H), 2.30 (s, 3H), 2.10 – 2.00 (m, 1H), 1.89 (dq, *J* = 11.4, 3.5, 2.9 Hz, 1H), 1.81 (dt, *J* = 20.4, 10.0 Hz, 1H), 1.44 – 1.31 (m, 2H), 1.26 (s, 1H), 1.07 – 0.96 (m, 1H), 0.94 (d, *J* = 7.0 Hz, 2H), 0.91 (d, *J* = 7.0 Hz, 6H), 0.90 – 0.78 (m, 4H), 0.67 (s, 1H), 0.14 (ddd, *J* = 13.9, 10.8, 2.4 Hz, 1H) ; LCMS (*m/z*): calculated exact mass: 921.5000, experimental $[M+H]^+$: 922.5062.

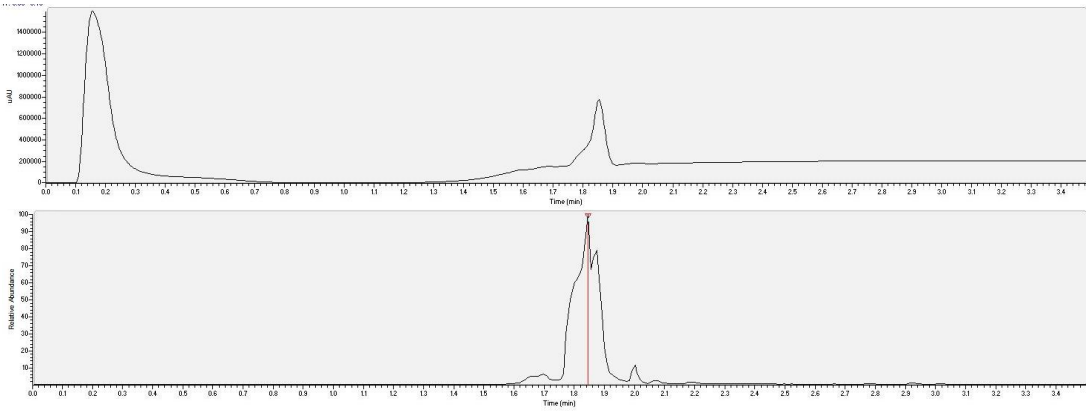


1.5.18.35 NMR and LC/MS data for VK-40 (**40**)

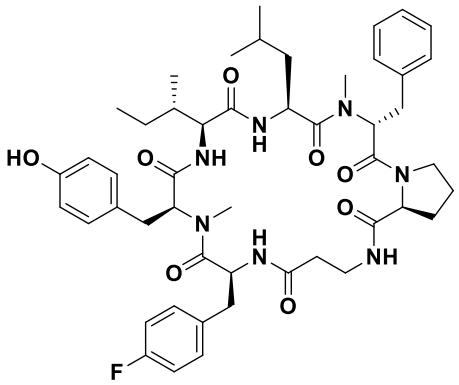


Sequence: Cyclo-(Leu-MeDPhe-Pro-Gly-Phe(4-F)-MePhe-Ile)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.41 (d, *J* = 9.9 Hz, 1H), 7.96 (d, *J* = 8.7 Hz, 1H), 7.65 (dd, *J* = 10.0, 1.8 Hz, 1H), 7.32 – 7.25 (m, 1H), 7.27 – 7.16 (m, 3H), 7.19 – 6.99 (m, 10H), 6.97 – 6.83 (m, 1H), 6.65 – 6.59 (m, 2H), 6.03 (d, *J* = 8.5 Hz, 1H), 5.48 (dd, *J* = 11.6, 4.7 Hz, 1H), 5.23 – 5.14 (m, 1H), 4.92 (dd, *J* = 17.9, 10.0 Hz, 1H), 4.80 (ddd, *J* = 10.9, 8.7, 3.8 Hz, 1H), 4.25 (dt, *J* = 8.5, 3.2 Hz, 2H), 4.07 (s, 4H), 3.71 (ddd, *J* = 12.1, 7.8, 4.4 Hz, 1H), 3.58 – 3.48 (m, 3H), 3.35 – 3.18 (m, 2H), 3.09 (s, 3H), 3.10 – 2.92 (m, 4H), 2.70 (dd, *J* = 13.1, 4.1 Hz, 1H), 2.61 (s, 3H), 2.41 – 2.28 (m, 1H), 2.24 (ddh, *J* = 10.5, 7.0, 3.5 Hz, 1H), 2.03 (ddt, *J* = 12.6, 6.5, 3.8 Hz, 1H), 1.82 (dddd, *J* = 29.3, 12.5, 9.9, 5.8 Hz, 2H), 1.49 – 1.28 (m, 3H), 1.04 – 0.91 (m, 1H), 0.92 (s, 1H), 0.94 – 0.84 (m, 6H), 0.86 – 0.76 (m, 7H), 0.48 (ddd, *J* = 13.4, 9.2, 3.7 Hz, 1H); LCMS (*m/z*): calculated exact mass: 867.4695, experimental $[M+H]^{+}$: 868.4769.

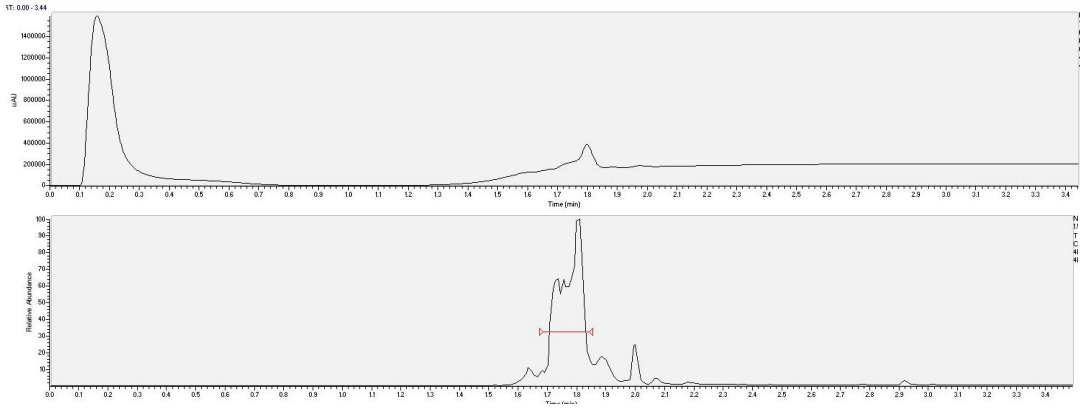


1.5.18.36 NMR and LC/MS data for VK-41 (**41**)



Sequence: *Cyclo-(Leu-MeDPhe-Pro-βAla-Phe(4-F)-MeTyr-Ile)*

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.15 (d, *J* = 9.9 Hz, 1H), 7.97 (d, *J* = 8.0 Hz, 1H), 7.59 (dd, *J* = 8.7, 4.2 Hz, 1H), 7.29 – 7.23 (m, 1H), 7.23 – 7.13 (m, 2H), 7.14 – 7.08 (m, 3H), 7.05 (t, *J* = 8.6 Hz, 2H), 6.66 – 6.60 (m, 2H), 6.57 – 6.52 (m, 2H), 6.03 (d, *J* = 7.3 Hz, 1H), 5.26 (td, *J* = 10.4, 5.0 Hz, 2H), 4.71 (q, *J* = 7.8 Hz, 1H), 4.19 (p, *J* = 7.6 Hz, 1H), 4.01 (dd, *J* = 7.3, 3.3 Hz, 1H), 3.94 (dd, *J* = 8.6, 3.1 Hz, 1H), 3.79 (s, 8H), 3.74 (ddd, *J* = 12.2, 7.5, 5.1 Hz, 1H), 3.55 – 3.44 (m, 2H), 3.23 (dd, *J* = 12.6, 10.5 Hz, 1H), 3.22 (s, 3H), 3.15 – 3.00 (m, 3H), 2.99 – 2.86 (m, 1H), 2.78 (ddd, *J* = 14.2, 10.2, 4.2 Hz, 1H), 2.73 – 2.67 (m, 1H), 2.67 (s, 3H), 2.55 (dd, *J* = 13.2, 6.9 Hz, 1H), 2.43 – 2.35 (m, 1H), 2.31 – 2.13 (m, 2H), 2.00 – 1.93 (m, 1H), 1.81 (dq, *J* = 11.1, 7.2, 6.2 Hz, 2H), 1.64 (dt, *J* = 13.5, 7.7 Hz, 1H), 1.53 – 1.37 (m, 2H), 1.25 (s, 1H), 1.17 – 1.00 (m, 2H), 1.00 – 0.91 (m, 7H), 0.94 – 0.83 (m, 6H), 0.80 (d, *J* = 6.5 Hz, 3H), 0.52 (q, *J* = 7.9 Hz, 2H) ; LCMS (*m/z*): calculated exact mass: 897.4800, experimental $[M+H]^+$: 898.4856.



1.5.18.37 NMR and LC/MS data for VK-42 (**42**)

1.5.18.38 NMR and LC/MS data for VK-43 (**43**)

Sequence: *Cyclo-(Leu-MeDPhe-PhotoPro-Sar-Phe-MeTyr-Pra)*

1.6 Abbreviations

ACN: acetonitrile; BONCAT: Bioorthogonal Non-canonical Amino Acid Tagging; BSA: Bovine serum albumin; COMU: (1-Cyano-2-ethoxy-2-oxoethylidenaminooxy)dimethylamino-morpholino-carbenium hexafluorophosphate; CP: cytological profiling; DBU: 1,8-Diazabicyclo[5.4.0]undec-7-ene; DCM: dichloromethane; DIAD: diisopropyl azodicarboxylate; DIPEA: N,N-diisopropylethylamine; DMEM: Dulbecco's Modified Eagle's medium; DMF: N,N-dimethylformamide; DMSO: dimethyl sulfoxide; DTP: Development Therapeutics Program (at the National Cancer Institute); EdU: 5-ethynyl-2´-deoxyuridine; eEF1A: eukaryotic translation elongation factor 1-alpha; EtOH: ethanol; FBS: fetal bovine serum; FITC: fluorescein isothiocyanate; Fmoc: fluorenylmethoxycarbonyl protecting group; GI50: concentration at which a compound causes 50% of growth inhibition; HATU: 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate; HFIP: hexafluoroisopropanol; LC50: concentration of compound at which half of the cells are killed (lethal dose); L-HPG: L-homopropargylglycine; $LogD_{(dec/w)}$: shake-flask partition coefficient between 1,9-decadiene and water; LPE: lipophilic permeability efficiency; MD: molecular dynamics; MOE: molecular operating environment; NCI: National Cancer Institute; PAMPA: parallel artificial membrane permeability assay; PBS: phosphate-buffered saline; SPPS: solid-phase peptide synthesis; TES: triethyl silane; TFA: trifluoroacetic acid; TGI: total growth inhibition; THF: tetrahydrofuran; TIC: total ion chromatogram; Xaa: unspecified amino acid

1.7 Acknowledgements

We thank Chad E. Townsend for creation of the Python code used to process mass directed assay.

1.8 Author Contributions

This project was designed by V.G.K, J.S., and R.S.L. with project initiation my J.S., Q.E. and M.R.N., and W.M.B and J.T. made significant contribution to guiding project direction. V.G.K., Q.E., A.T., and J.H.F. synthesized compounds. V.G.K., W.M.B, and J.S. collected and analyzed cytological profiling data. H.W. and J.T. performed and analyzed eEF1A A399V mutant cell line assays. S.O. and O.O. performed and analyzed molecular dynamics simulations. Manuscript was written by V.G.K and R.S.L. with the aid of M.R.N. and J.T. All authors edited the manuscript.

1.9 Funding Sources

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program (NSF DGE 1339067 to V.G.K.), Tobacco-Related Disease Research Program Postdoctoral Fellowship (28FT-0014 to H.-Y.W.), the National Institute of General Medicine Studies of the National Institutes of Health (NIH R01GM13113 to R.S.L), and Eli Lilly and Co. (Lilly Innovation Fellowship Award to M.R.N.). The UCSC Chemical Screening Center was supported by S10 RR022455. Any opinions, findings, and conclusions or recommendations expressed in

this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the National Institutes of Health. Testing was performed by the Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute. More information on this can be found on the Program's website: http://dtp.cancer.gov.

1.10 Notes

The data that support the findings of this study are available from the corresponding author upon reasonable request. Figures 1 – 5 and Table 1 from the main text have associated raw data, as do Supplementary Figures 1 – 8 and Supplementary Table 1 – 4. Supplementary Figure 7 and Supplementary Table 1 contain data by the Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute.

Chapter Two

Understanding and Improving the Membrane Permeability of VH032-Based PROTACs

This chapter contains text and figures from the following manuscript: Klein, V. G., Townsend, C. E., Testa, A., Zengerle, M., Maniaci, C., Hughes, S. J., Chan, K. H., Ciulli, A., & Lokey, R. S. (2020). Understanding and Improving the Membrane Permeability of VH032-Based PROTACs. *ACS Medicinal Chemistry Letters*, 11(9), 1732–1738. https://doi.org/10.1021/acsmedchemlett.0c00265

Abstract

Proteolysis targeting chimeras (PROTACs) are catalytic heterobifunctional molecules that can selectively degrade a protein of interest by recruiting a ubiquitin E3 ligase to the target, leading to its ubiquitylation and degradation by the proteasome. Most degraders lie outside the chemical space associated with most membrane-permeable drugs. Although many PROTACs have been described with potent activity in cells, our understanding of the relationship between structure and permeability in these compounds remains limited. Here, we describe a label-free method for assessing the permeability of several VH032-based PROTACs and their components by combining a parallel artificial membrane permeability assay (PAMPA) and a lipophilic permeability efficiency (LPE) metric. Our results show that the combination of these two cell-free membrane permeability assays provides new insight into PROTAC structure-permeability-relationships and offers a conceptual framework for predicting the physicochemical properties of PROTACs in order to better inform the design of more permeable and more effective degraders.

2.1 Introduction

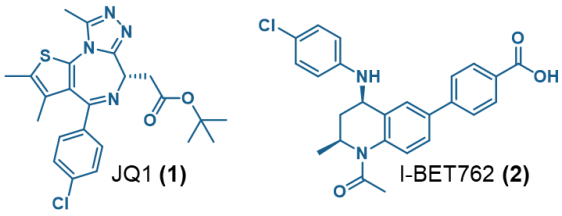
Proteolysis targeting chimeras (PROTACs) enhance our ability to drug biologically relevant targets through selective degradation.[76-78] These heterobifunctional compounds include an E3 ligase-binding ligand and a protein-targeting ligand connected by a linker. PROTACs facilitate proteasomal degradation by recruiting the target protein to an E3 ligase, leading to ubiquitination and subsequent degradation of the targeted protein.[79-81] Unlike traditional inhibitors, PROTACs are catalytic and have increased target-specificity derived largely from ternary complex protein-protein contacts.[82-84] While our understanding of the bioactivity of PROTACs is rapidly increasing, the physiochemical properties of these molecules have received relatively little attention.[85,86]

Due to the interest in PROTAC therapeutics, there is a clear need to better understand their physicochemical properties. Given their high molecular weight (MW > 800) and the presence of multiple hydrogen bond donors (HBDs) and acceptors (HBAs), PROTACs are expected to have low membrane permeability.[8,87-89] A recent study that used the label-based chloroalkane penetration assay (CAPA)[90] showed very low permeabilities for PROTACs relative to their individual components.[91] While this assay provides relative cell permeabilities across a large dynamic range, it does not provide permeability coefficients that can compared to other datasets. Also, CAPA requires a chloroalkane tag and therefore does not directly measure the permeability of the parent compound. Establishing a label-free method to quantify the permeability of PROTACs provides greater flexibility in compound design without needing to

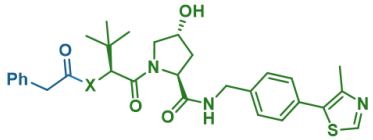
synthesize a second set of CAPA tag-containing molecules. While there are some mass spectrometry approaches to quantify the intracellular concentration of unlabeled compounds, these indirect studies do not inform on oral bioavailability and some do not differentiate between membrane-trapped compounds and those free for target binding.[91-94] The VHL-NanoLuc Fusion assay[95] offers label-free assessment of cell permeability, but results are confounded by their dependence on variable VHL-binding affinities. Here we report a label-free approach for studying the passive permeability of von Hippel-Lindau (VHL)-based PROTAC molecules using the parallel artificial membrane permeability assay (PAMPA) and lipophilic permeability efficiency (LPE).[7] These simple, high-throughput assays correlate strongly with cell-based permeabilities and oral bioavailability while being relatively inexpensive.[65] PAMPA quantifies orders-of-magnitude differences in PROTAC permeabilities with a low limit of quantitation. LPE provides insight as to how structural changes affect permeability.

2.2 Results and Discussion

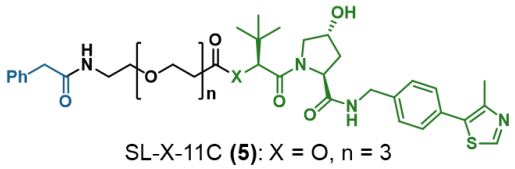
We tested the membrane permeabilities of JQ-1 (**1**), four model compounds (SL-X series) (**3** – **6**), and 11 previously published VHL-PROTACs.[96,97] These PROTACs include four series: MZ (**7** – **9**),[81,82,98] AT (**15** – **17**),[82] CM/CMP (**12** – **14**),[99] and MZP (**10** – **11**),[98] grouped according to the target-binding ligand and attachment to the VHL-recruiting ligand (Figures 2-2 and 2-5). Most previously published PROTACs have MWs ranging from 900 to 1200 and between four and six HBDs. Based on traditional criteria of drug-likeness, these compounds are expected to have low membrane permeability. This is indeed what we found. The highest PAMPA permeability measured for this set was $P_e = 0.6 \times 10^{-6}$ cm/s, slightly below the standard for "modest" permeability ($P_e = 1 \times 10^{-6}$ cm/s). Notably, we were able to quantify permeabilities for all our compounds with coefficients as low as 0.002×10^{-6} cm/s.



Small molecule bromodomain inhibitors



SL-X-9A (**3**): X = O
SL-X-9B (**4**): X = NH



SL-X-11C (**5**): X = O, n = 3
SL-X-11D (**6**): X = NH, n = 3

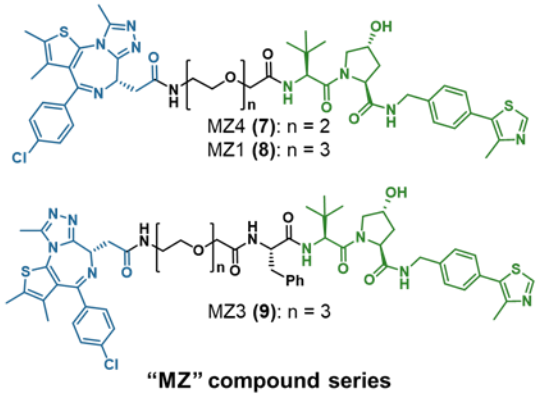
"SL-X" compound series

Cmpd	MW	ALogP	# of HBDs	# of HBAs	PAMPA	LogD (dec/w)	LPE
1	457	5.0	0	5	5.6	2.0	2.2
2	435	4.6	2	4	--	--	--
3	550	3.3	2	6	16	0.3	2.3
4	549	2.6	3	5	8.6	-0.7	2.0
5	753	2.2	3	10	0.3	-1.6	1.6
6	752	1.6	4	9	0.2	-2.2	1.6

Figure 2-1: Physiochemical properties of protein-targeting small molecules and model compounds Cmpd = compound; PAMPA units: x 10^{-6} cm/s; $LogD_{(dec/w)}$: 1,9-decadiene and PBS pH 7.4 shake flask partition coefficient; LPE = $LogD_{(dec/w)}$ - 1.06(ALogP) + 5.47; "--" = not determined

From our initial set of amide-containing compounds, the most permeable compound was **4** (P_e = 8.6 x 10^{-6} cm/s, 2.1), an N-terminally capped VH032 analog with a phenylacetamide acting as a simple protein-targeting model. Compound **4** was 43-fold more permeable than a similar compound, **6**, with a 3-unit PEG linker between the VH032 and the phenylacetamide. Strikingly, **4** was 4000-fold more permeable than the two least permeable compounds, **17** and **14** (Figures 2-2 and 2-5, respectively).

Furthermore, among all 11 PROTACs tested, there was a 300-fold difference between the most permeable compound, **7**, and least permeable compounds, **14** and **17**. In the MZ series alone (**7** – **9**, Figure 2-2), there was a 100-fold difference between the most (**7**, P_e = 0.6 x 10^{-6} cm/s) and least (**9**, P_e = 0.006 x 10^{-6} cm/s) permeable derivatives. Combined, these data demonstrate the large dynamic range of PAMPA and support its use for unlabeled, quantitative measurements.



Cmpd	MW	ALogP	# of HBDs	# of HBAs	PAMPA	LogD (dec/w)	LPE
7	959	3.7	4	11	0.6	-1.1	0.5
8	1003	3.6	4	12	0.03	-1.6	0.1
9	1150	4.7	5	13	0.006	-1.6	-1.1
15	961	3.7	4	11	0.005	-3.8	-2.3
16	1005	3.6	4	12	0.003	-4.3	-2.7
17	945	4.5	4	10	0.002	-3.8	-3.1

Figure 2-2: Physiochemical properties of "AT" and "MZ" PROTACs Cmpd = compound; PAMPA units: x 10^{-6} cm/s; $LogD_{(dec/w)}$: 1,9-decadiene and PBS pH 7.4 shake flask partition coefficient; LPE = $LogD_{(dec/w)}$ - 1.06(ALogP) + 5.47

MW and solvent-exposed HBDs can significantly affect membrane permeability.[87] Permeability generally decreases as MW increases,[100] leading to a significant reduction in permeability beyond MW = 1000.[8,101] All else being equal, the relatively high MWs (900-1200 Da) of the PROTACs represent a predicted size-dependent permeability cost of approximately one log unit compared to typical small molecules of the same lipophilicity (MW < 600).[7,8] However, because the PROTACs in this study are all in a similar MW range, comparisons between them reflect differences in their physical properties separate from the size penalty. Recent reviews argue that MW effects should not be considered alone because factors like hydrophobicity and HBDs affect permeability more prominently than MW.[102,103] Supporting this conclusion, our two least permeable PROTACs, **14** and **17**, had the highest and lowest MWs, respectively. Furthermore, **16** and **8** have nearly the same MW (1005 and 1003, respectively), the same calculated octanol-water partition coefficients (ALogP), and the same number of HBDs and HBAs, yet their permeabilities differ by 10-fold (Figure 2-2). Likewise, **15** and **7** are similar in terms of MW, ALogP, and HBAs/HBDs, but **7** is 120-fold more permeable than **15** (Figure 2-2). As expected, the compounds that had lower MW *and* fewer HBDs/HBAs, including **1**, **3**, and **4**, were significantly more permeable ($P_e \geq 5 \times 10^{-6}$ cm/s, Figure 2-1) than the PROTACs.

Permeability data alone provide little information on how structural features affect permeability. Therefore, we measured lipophilic permeability efficiency (LPE).[7]

Originated by our group, LPE quantifies the efficiency with which a compound achieves passive membrane permeability at a given lipophilicity based on the experimental hydrocarbon-water partition coefficient ($LogD_{(dec/w)}$) and ALogP. Combining PAMPA and LPE represents a powerful method for assessing how structural features contribute to compound permeability.

This is most evident when comparing the two matched pairs from the AT and MZ series: **15** vs. **7** and **16** vs. **8**. These compounds have the same ALogP, the same number of HBDs/HBAs, and MWs within 2 Da. Yet, the MZ compounds, **7** and **8**, are significantly more permeable than their counterparts from the AT series, **15** and **16**, respectively. These AT and MZ compounds differ only in the connection between their linker and VH032 ligand. In **7** and **8**, the VH032 ligand has an N-terminal *tert*-Leu connected to a linker through an amide bond. Alternatively, **15** and **16** have a penicillamine group in place of the *tert*-Leu which is attached to the linker through a thioether in place of the amide bond (Figure 2-2).

Clearly, the chemical environment surrounding HBDs affects the PAMPA permeability of these PROTACs, similar to the effects observed in other compounds in this MW range.[51,104] The LPE values of these compounds provide insight into the potential for these flexible molecules to adopt conformations capable of shielding HBDs. Typically, the addition of a solvent-exposed HBD reduces LPE by 1.8.[7] The *tert*-Leu-containing **7** has an LPE of 0.4, and its penicillamine counterpart, **15**, has an LPE of -2.3, suggesting that **15** has at least one additional exposed HBD compared to **7**. The same pattern is seen with **8** and **16** that have LPE values of 0.1 and -2.6,

respectively. These LPE data show that switching the *tert*-Leu for a penicillamine group exposes an -NH to solvent which likely contributes to the lower permeability of these AT compounds.

The crystal structure of **8** in a ternary complex with VHL and Brd4 further supports the presence of a shielded -NH in the MZ compound series.[82] Inspection of this structure shows that the *tert*-Leu amide -NH of **8** is in a position to be shielded from solvent by the *tert*-Leu side chain, and is within a short contact distance to the PEG oxygen, likely participating in an intramolecular hydrogen bond (IMHB) capable of shielding the -NH polarity from solvent (Figure 2-3). Co-crystal structures of binary complexes of VHL with bound ligands provide additional evidence for this phenomenon showing an oxygen (in a similar position to the PEG ether in **8**) that points in towards the *tert*-Leu -NH, potentially close enough to form an IMHB.[105] While the membrane permeating conformation is not necessarily the same as the target-bound conformation, these crystal structures provide a possible explanation for the difference in solvent-exposed HBDs between the MZ and AT compounds.

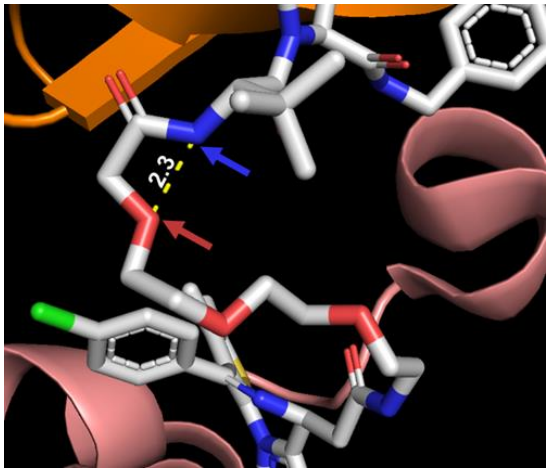


Figure 2-3: MZ1 ternary complex with VHL and Brd4 (PDB:5T35)[82] Crystal structure showing the ternary complex of MZ1 (colored by element) with Brd4 (pink) and VHL (orange). The VHL ligand *tert*-Leu -NH (blue arrow) is shielded by the *tert*-Leu side chain and is within hydrogen bonding distance of the VHL ligand PEG oxygen (red arrow).

This relationship between the MZ and AT compounds supports reducing the number of exposed HBDs to increase permeability. The extensive structural information on VHL ligand co-crystal structures have shown that the tert-Leu amide does not form a direct hydrogen bond with the VHL protein.[96,97,105] Hence, we hypothesized that removing an HBD by substituting an amide for an ester would lead to increased permeability, without detrimentally comprising VHL binding affinity. To test this, we synthesized **3** and **5**, ester derivatives of **4** and **6**, respectively, in which the N-terminal *tert*-Leu amide was replaced by an ester (Figure 2-1). As predicted, the ester derivatives were more permeable than their amide counterparts. Compound **3** was 2-fold more permeable than **4**, and **5** was 1.5-fold more permeable than **6**. Thus, the amide-to-ester substitution provides a viable option to increase the permeability of

these types of compounds, though with the caveat of the ester's potential susceptibility to intracellular esterase hydrolysis.

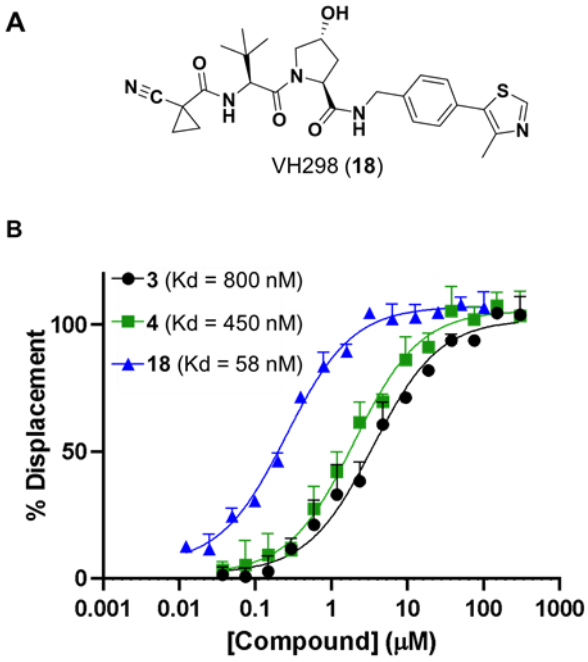
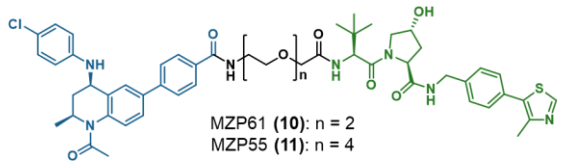


Figure 2-4: Fluorescence polarization (FP)-derived K_d of amide to ester substitution in SLX compounds *(A)* VH298, a small molecule inhibitor of the E3 ubiquitin ligase VHL, used as a positive control for high-affinity binding. *(B)* FP data for compounds **3**, **4**, and **18**.

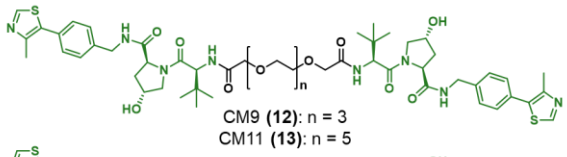
The LPE of the amide compounds (**4** and **6**) is nearly the same as the LPE of their ester compound counterparts (**3** and **5**, respectively), suggesting that the *tert*-Leu is likely shielding the polarity of the HBD in the amide-containing compounds as has be observed with beta-branched amino acids.[51,104] The relatively modest increase in permeability observed with these amide-to-ester substitutions reflects the unusually low desolvation penalty for the shielded amide NH – consistent with what observed in

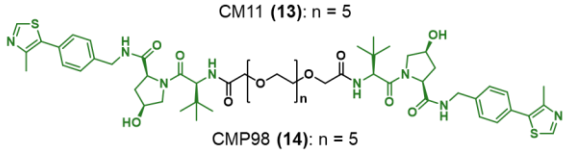
the MZ series. Therefore, it is possible that substituting a more exposed amide with an ester could lead to even greater improvement of membrane permeability. Using a competitive fluorescence polarization (FP) assay, we found that the ester-containing **3** was still capable of binding its target protein, VHL, with a K_d only 1.7-fold higher than that of the amide-containing **4** , albeit >10-fold higher than the potent VHL inhibitor VH298 (**18**, Figure 2-4).[105] The K_d increase in the ester compound further advocates for trying similar substitutions farther away from the VHL-binding ligand to maintain binding capacity while improving permeability.

Consistent with Foley *et al.*,[91] we found that permeability increased with decreasing linker length. This was expected, as increasing the length of the linker usually results in an increase in one or more of the MW, HBDs, or HBAs. For the AT and CM/CMP series, compound permeability was reduced by half with one or two additional PEG units in the linker, respectively (cf. **15** vs. **16**, and **12** vs. **13**, Figures 2-2 and 2-5). This effect was more prominent in the MZ series as **7** (2-unit PEG linker) was 20-fold more permeable than **8** (3-unit PEG linker). A 2-fold difference in permeability was also seen in the MZP compounds (**11**, 4-unit PEG linker, and **10**, 2-unit PEG linker, Figure 2-5). These results indicate that shorter linkers typically produce more permeable compounds. Moreover, for all but the MZP series, the compounds with shorter PEG linkers had higher LPE values, suggesting that compounds with shortened linkers were more efficient at permeating the membrane for their given lipophilicity.



"MZP" compound series





"CM"/"CMP" compound series

Cmpd	MW	ALogP	# of HBDs	# of HBAs	PAMPA	LogD (dec/w)	LPE
10	993	4.6	5	10	0.3	-0.7	-0.1
11	1081	4.3	5	12	0.16	-0.7	0.2
12	1091	0.9	6	14	0.009	-3.8	0.7
13	1179	0.6	6	16	0.005	-4.3	0.5
14	1179	0.6	6	16	0.002	-3.2	1.6

Figure 2-5: Physicochemical properties of "MZP" and "CM/CMP" PROTACs Cmpd = compound; PAMPA units: x 10^{-6} cm/s; $LogD_{(dec/w)}$: 1,9-decadiene and PBS pH 7.4 shake flask partition coefficient; LPE = $LogD_{(dec/w)}$ - 1.06(ALogP) + 5.47

Previous studies have advocated for the use of short alkyl linkers over PEG linkers to reduce total polar surface area to improve permeability.[91] Our results diverge in this respect, as we found that our only compound bearing an alkyl linker, **17**, was the least permeable (P_e = 0.002 x 10^{-6} cm/s). This compound was 2.5-fold less permeable than **15**, which has a 1-unit PEG linker. Compound **17** has one fewer HBA in this linker than **15** which could reduce solubility and therefore affect permeability. While PAMPA allows us to quantify the differences in permeabilities directly, analyzing LPE enable us to predict which structural features cause the permeability changes. Increasing the number of PEG units in the PROTAC linker reduces the LPE

of that compound (cf.**15** and **16**). If the HBAs in these PEG linkers were not contributing to IMHB, substituting the PEG linker in **15** with an alkyl linker as in **17** (removing HBAs) should have little effect on LPE. However, this is not what we observed. Instead, the LPE of alkyl-linked **17** is 0.8 lower than its PEG counterpart **15**, suggesting that the ether oxygen in the PEG linker of **15** is capable of shielding HBD, possibly the linker amide bond -NH (adjacent to JQ-1) in a manner similar to that observed for MZ1 (Figure 2-3). As the ΔLPE between **15** and **17** is less than the 1.8-unit difference expected for a fully exposed HBD, it is likely that the PEG ether provides only partial shielding by way of IMHB formation.

The same phenomenon is present in the SL-X series (**3-6**, Figure 2-1). Compounds **5** and **6** have an additional amide and 3-unit PEG linker compared to **3** and **4**, respectively. If no additional IMHBs were present in **5** and **6**, the inclusion of these additional HBAs and HBD should cause a decrease in LPE of at least 1.8, compared to **3** and **4**. Yet, the LPE values of **5** and **6** are only moderately lower than **3** (ΔLPE = 0.8) and **4** (ΔLPE = 0.4), respectively. Thus, the PEG linker is likely involved in IMHBs responsible for shielding some polarity. Moreover, using a linker capable of forming IMHB could shield the polarity of important HBDs responsible for target engagement, a feature that would not be possible with an alkyl linker. Therefore, the best linker type for a given PROTAC is likely scaffold dependent, further highlighting the need to examine the overall lipophilicity of the molecule when designing a PROTAC.

Comparing PAMPA and $LogD_{(dec/w)}$ to ALogP allows us to analyze permeability trends and predict permeability improvements. For compounds with

ALogPs up to ~4, there is a positive linear correlation between ALogP and permeability.[7] As lipophilicity increases beyond ALogP ~ 4-5, compounds become insoluble or membrane-retained, and their effective membrane permeabilities diminish (Figure 2-6A). Therefore, designing PROTACs to have an ALogP below 5.0 could bias these compounds towards higher permeabilities. The CM/CMP compounds have low permeabilities and lower ALogPs (<1) than the other PROTACs. As permeability typically increases with ALogP from 0 – 4, a lipophilicity increase , such as increasing the number of -CH_2- groups relative to oxygens in the linker, could greatly improve CM/CMP permeability.[106]

Plotting $LogD_{(dec/w)}$ vs. ALogP creates a visualization of the LPE metric which offers potential strategies to improve permeability (Figure 2-6B). For example, in the MZ series, **7** and **8** have low permeabilities (>0.6 x 10^{-6} cm/s) and moderately low LPE (>0.5). As **7** and **8** already have ALogP values close to 4.0, further increasing lipophilicity would likely push these compounds into the insoluble region and cause a further decrease in their membrane permeability (Figure 2-6A). Also, the addition of a Phe residue to **8** to generate **9** leads to a 1.2-unit decrease in LPE due to the addition of an amide NH (which is less than the 1.8-unit cost expected for the addition of an amide group, indicating partial IMHB). This decrease in LPE between **8** and **9** is partially offset by an increase in ALogP of 1.1 units, leading to a 5-fold decrease in permeability and putting **9** over the edge of the solubility cliff. This analysis suggests that the decreased degrader activity observed in cells for **9** compared to **8**,[81] could be, in part, due to these poor physiochemical properties. Both the decrease in LPE and a

significant increase in ALogP contribute to the very poor permeability of **9**. An alternative solution to improving the permeability of **7** and **8** would be to replace the amide linkage to the bromodomain warhead with a group (such as an ester) that does not contribute an HBD.

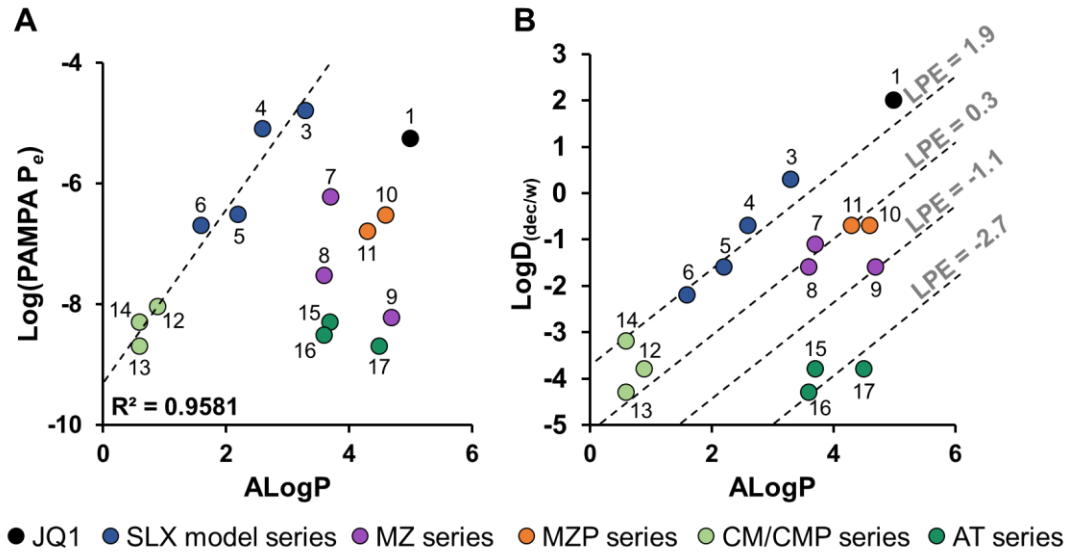


Figure 2-6: PROTAC permeability and LPE Graphs showing the *(A)* permeability vs. ALogP and the *(B)* $LogD_{(dec/w)}$ vs. ALogP for compounds **1-17**. Dashed line on *(A)* shows the linear correlation between PAMPA and ALogP for ALogP from 0 – 4 (R^2 = 0.9581). Dashed lines on *(B)* represent LPE classes, *m* = 1.06. LPE values (grey) are LPE averages for compounds that fall on or near the line.

The effect of structural features on permeability and bioactivity can be significant. Generally, the bromodomain-targeting compounds (MZ, AT, MZP), with extremely low permeabilities (≤0.006 cm/s), were less active in relevant cellular anti-proliferation assays than compounds with higher permeabilities (≥0.03 cm/s, SI table 2.2).[98] Specifically, **9** was both less permeable and less bioactive than **7** and **8**. This decreased bioactivity is likely attributed to the decreased permeability (Figure 2-2, SI

table 2-2), as binding affinities with the target proteins were broadly comparable.[81,96] Similarly, the AT compounds were the least active compounds tested, consistent with their much lower permeability (Figure 2-2, SI table 2-2). However, the related PROTAC, AT1, exhibited a 5-fold lower bind affinity for the VHL protein and formed less stable ternary complexes compared to **8**,[82,107] which could also contribute to the significant loss of cellular potency in the AT series. Conversely, the formation of a cooperative and stable ternary complex can override the impact of permeability.[107] For example, **8** forms a more stable complex with its targets, Brd4 and VHL, than compound **7**, leading **8** to be one log unit more active, despite being 20-fold less permeable than **7** (Figure 2-2, SI table 2-2).[98] Similarly, in the CM/CMP series, **13** is two log units more active than **12** in a cellular protein degradation assay despite being slightly less permeable (0.005 cm/s vs 0.009 cm/s, Figure 2-5).[99] This suggests that differences in efficacy between these two compounds are likely due to the relative stability of their respective ternary complexes[99] rather than differences in their extremely low permeabilities. These results suggest that efforts to improve the permeability should be monitored in conjunction with effects on ternary complex formation.

2.3 Conclusion

In this study, we have demonstrated that combining PAMPA and LPE provides insight into PROTAC structure-permeability relationships. These label-free assays model only passive permeability without the confounding effects of active transport. PAMPA and $LogD_{(dec/w)}$ are established methods, therefore comparisons can be made to data previously gathered using these methods. With this simple method for measuring the permeability of PROTACs in hand, a more systematic study on PROTAC permeability and pharmacokinetics is required. While this study provides some evidence, assessing the permeability of PROTACs over a complete range of ALogP values would allow us to develop a more detailed lipophilicity window to guide the design of PROTACs biased towards higher permeability. As esters are generally more prone to hydrolysis than amides, additional studies are required to assess the viability of amide-to-ester substitutions.

Finally, VH032-based PROTACs have a high number of HBDs and HBAs often present on both protein-binding domains of the molecule that are typically connected by a long flexible linker. This arrangement of HBDs and HBAs lends itself to the formation of IMHBs capable of shielding some of the PROTACs' polarity, enhancing permeability. The recently reported macrocyclization of PROTACs[108] could also prove beneficial in this regard by taking advantage of the IMHBs and HBD-shielding often achieved by cyclic peptides. Future studies on the permeability of these compounds, and expansion of these studies to include other PROTAC classes such as those based on cereblon-binding ligands, are warranted as they could create

opportunities to model and predict a network of IMHBs and fine-tune these interactions to produce more permeable and more bioactive PROTACs.

2.4 Supplementary Tables

2.4.1 Supplementary Table 2-S1: Compiled physicochemical data

Combined data for all compounds and conditionally formatted for PAMPA permeability, $LogD_{(dec/w)}$, and LPE. Dark colors represent better values for each column and the worst value is in white for each column.

Compound	Molecular weight	ALogP	Number of HBDs[a]	Number of HBAs[b]	PAMPA permeability ($\times 10^{-6}$ cm/s)	$LogD_{(dec/w)}$[c]	LPE[d]
JQ1 **(1)**	457	5.0	0	5	5.6 ± 1.7	2.0	2.2
I-BET726 **(2)**	435	4.6	2	4	--	--	--
SL-X-9A **(3)**	550	3.3	2	6	16.1 ± 1	0.3	2.3
SL-X-9B **(4)**	549	2.6	3	5	8.6 ± 0.7	-0.7	2.0
SL-X-11C **(5)**	753	2.2	3	10	0.3 ± 0.1	-1.6	1.6
SL-X-11D **(6)**	752	1.6	4	9	0.2 ± 0.03	-2.2	1.6
MZ4 **(7)**	959	3.7	4	11	0.6 ± 0.1	-1.1	0.5
MZ1 **(8)**	1003	3.6	4	12	0.03 ± 0.01	-1.6	0.1
MZ3 **(9)**	1150	4.7	5	13	0.006 ± 0.002	-1.6	-1.1
MZP61 **(10)**	993	4.6	5	10	0.3 ± 0.1	-0.7	-0.1
MZP55 **(11)**	1081	4.3	5	12	0.16 ± 0.14	-0.7	0.2
CM9 **(12)**	1091	0.9	6	14	0.009 ± 0.004	-3.8	0.8
CM11 **(13)**	1179	0.6	6	16	0.005 ± 0.001	-4.3	0.6
CMP98 **(14)**	1179	0.6	6	16	0.002 ± 0.0002	-3.2	1.7
AT3 **(15)**	961	3.7	4	11	0.005 ± 0.003	-3.8	-2.2
AT6 **(16)**	1005	3.6	4	12	$0.003 \pm 8.4 \times 10^{-5}$	-4.3	-2.6
AT2 **(17)**	945	4.5	4	10	0.002 ± 0.0003	-3.8	-3.1

a: HBD = hydrogen bond donor
b: HBA = hydrogen bond acceptor
c: 1,9-decadiene and PBS pH 7.4 shake flask partition coefficient
d: $LPE = LogD_{(dec/w)} - 1.06(ALogP) + 5.47$
"--" = value not determined

2.4.2 Supplementary Table 2-S2: PROTAC Bioactivity

Effects of MZ, AT, and MZP compounds on cell proliferation in two relevant BET-sensitive cell lines (MV4:11 and HL60) compared to the PAMPA permeabilities of these compounds

Compound	pEC_{50}[a] (MV4;11)	pEC_{50}[a] (HL60)	PAMPA permeability (x 10^{-6} cm/s)
MZ4 **(7)**	6.75*	5.84*	0.6
MZP61 **(10)**	6.24*	6.17*	0.3
MZP55 **(11)**	7.08*	6.37*	0.16
MZ1 **(8)**	7.57*	6.8*	0.03
MZ3 **(9)**	6.34	5.38	0.006
AT3 **(15)**	5.14	4.9	0.005
AT6 **(16)**	6.16	BLQ[b]	0.003
AT2 **(17)**	6.21	BLQ[b]	0.002

a: pEC_{50} measured after 48 h treatment
b: BLQ = below limit of quantitation
*: These data were previously reported in Chan *et al.* 2018 (ref[98])

2.5 Methods

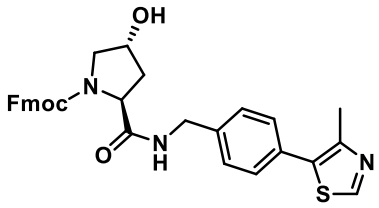
2.5.1 General Synthetic Information and Procedures

Reagents, compounds, and solvents were used without further modification unless otherwise stated. Reagents and solvents were purchased from Fisher Scientific. HATU was purchased from Combi-Blocks or Chem-Impex. Previously published PROTAC compounds were supplied by the Ciulli research group and used without further modification. I-BET726 (catalog #16872-1) and (+)-JQ1 (catalog #11187-1) were purchased from Cayman Chemical Company. Amino acids and linkers were purchased from Combi-Blocks or Oakwood. SynPhase polystyrene lanterns were purchased from Mimotopes. For PAMPA and $LogD_{(dec/w)}$ experiments, LC/MS samples were analyzed on a Thermo Scientific Ultimate 3000 UPLC system and Thermo Scientific Orbitrap VelosPro mass spectrometer with a Thermo Hypersil GOLD C18 30 x 2.1 (1.9 µm) column eluting with 5-95% ACN in H_2O with 0.1 % formic acid. Relative amount in each layer was quantified using an in-house python program to integrate the single ion chromatogram for the exact mass of each compound.

2.5.2 Synthetic Methods

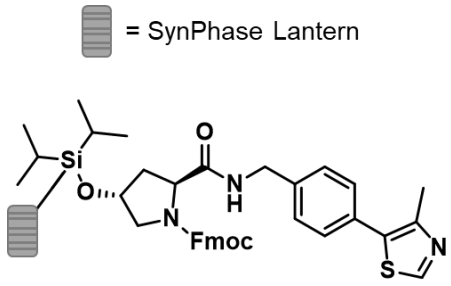
2.5.2.1 Loading SynPhase polystyrene L-series lanterns

The following compound (**19**) was synthesized according previously published procedures.[96]



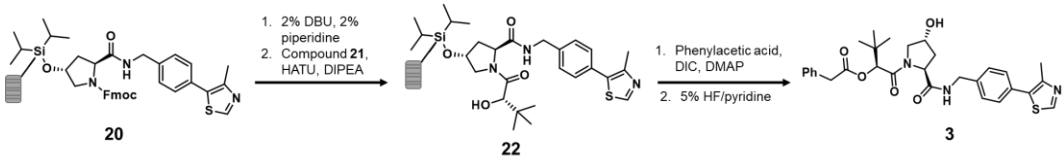
19

Compound **19** was conjugated to SynPhase polystyrene L-series lanterns (alkyl tethered diisopropylarylsilane linker, 22 µM, catalog # MIL10431000) that were previously treated with excess trifluoromethanesulfonic acid as previously described to make the following compound (**20**).[109] Note: Use caution while handling trifluoromethanesulfonic (aka triflic) acid. This is a strong acid that can cause severe burns and is strongly exothermic with polar solvents.



20

*2.5.2.2 Solid phase synthesis of compound **3***

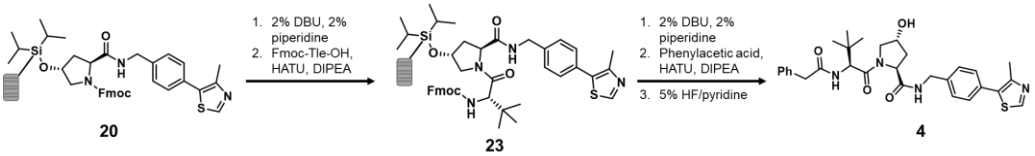


20 **22** **3**

A SynPhase lantern conjugated to compound **20** was treated with 2 mL of a solution of 2% DBU and 2% piperidine in DMF for 15 min at room temperature to remove the Fmoc protecting group. The lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL DCM. Next, a solution of a solution of (S)-2-hydroxy-3,3-

dimethylbutanoic acid (**21**) (19.8 mg, 0.15 mmol, 7 eq), HATU (57 mg, 0.15 mmol, 7 eq), and DIPEA (40 µL, 0.225 mmol, 10 eq) in DMF was added per lantern. Reaction was shaken on a linear shaker for 1 h at room temperature yielding compound **22**. The lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL DCM. Next, a solution of phenylacetic acid (50 mg, 0.37 mmol, 17 eq), DIC (48 µL, 0.44 mmol 20 eq), and DMAP (1 mg, 0.008 mmol, 0.4 eq) in 1 mL dry DCM was added to the lantern with compound **22** and shaken at room temperature for 1 h. Lantern was washed 3x with 2 mL of DMF and 3x with 2 mL DCM. Compound **3** was cleaved off the lantern with 5% HF/pyridine in THF and quenched with methoxytrimethylsilane according to literature procedures.[109] Compound **3** was purified on Waters mass-directed preparative HPLC system with an XBridge BEH130 5µm 19 x150 C18 column eluting with 10 – 100% ACN in water both with 0.1% formic acid. Sample identity was confirmed by LCMS. Note: Use *extreme* caution while handling HF/pyridine solution. Hydrogen fluoride (HF) forms hydrofluoric acid upon contact with water in the air or human tissue. Hydrofluoric acid is both highly corrosive and toxic. See hydrogen fluoride safety data sheet (SDS) before using HF/pyridine.

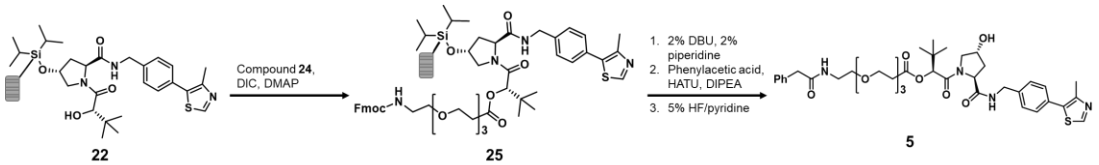
*2.5.2.3 Solid phase synthesis of compound **4***



A SynPhase lantern conjugated to compound **20** was treated with 2 mL of a solution of 2% DBU and 2% piperidine in DMF for 15 min at room temperature to

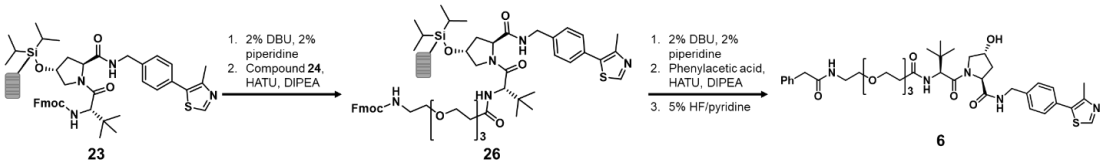
remove the Fmoc protecting group. Lanterns were rinsed 3x with 2 mL of DMF and 3x with 2 mL DCM. A solution of Fmoc-L-Tle-OH (53 mg, 0.15 mmol, 7 eq), HATU (57 mg, 0.15 mmol, 7 eq), and DIPEA (40 µL, 0.225 mmol, 10 eq) in DMF was added per lantern. Reactions vial containing the lanterns was mixed on a linear shaker for 4 h at room temperature to produce compound **23**. Lantern was washed 3x with 2 mL of DMF and 3x with 2 mL DCM. The Fmoc protecting group on **23** removed with 2 mL of a solution of 2% DBU and 2% piperidine in DMF for 15 min at room temperature. Lantern was washed 3x with 2 mL of DMF and 3x with 2 mL DCM. Next, a solution of phenylacetic acid (50 mg, 0.37 mmol, 17 eq), HATU (140 mg, 0.37 mmol, 17 eq), and DIPEA (96 µL, 0.44 mmol, 20 eq) in 1 mL DMF was added to the lantern and shaken at room temperature for 4 h. Lantern was washed 3x with 2 mL of DMF and 3x with 2 mL DCM. Compound **4** was cleaved off the lantern with 5% HF/pyridine in THF and quenched with methoxytrimethylsilane according to literature procedures.[109] Compound **4** was purified on Waters mass-directed preparative HPLC system with an XBridge BEH130 5µm 19 x150 C18 column eluting with 10 – 100% ACN in water both with 0.1% formic acid. Sample identity was confirmed by LCMS. Note: Use *extreme* caution while handling HF/pyridine solution. Hydrogen fluoride (HF) forms hydrofluoric acid upon contact with water in the air or human tissue. Hydrofluoric acid is both highly corrosive and toxic. See hydrogen fluoride safety data sheet (SDS) before using HF/pyridine.

*2.5.2.4 Solid phase synthesis of compound **5***



To a SynPhase lantern conjugated to compound **22**, a solution of Fmoc-12-amino-4,7,10-trioxadodecanoic acid (**24**) (67 mg, 0.15 mmol, 7 eq), DIC (48 µL, 0.44 mmol 20 eq), and DMAP (1 mg, 0.008 mmol, 0.4 eq) in 1 mL of dry DCM was added to a single lantern. Reactions vial containing the lanterns was mixed on a linear shaker for 4 h at room temperature to produce compound **25**. Lantern was washed 3x with 2 mL of DMF and 3x with 2 mL DCM. The Fmoc protecting group on **25** removed with 2 mL of a solution of 2% DBU and 2% piperidine in DMF for 15 min at room temperature. Lantern was washed 3x with 2 mL of DMF and 3x with 2 mL DCM. Next, a solution of phenylacetic acid (50 mg, 0.37 mmol, 17 eq), HATU (140 mg, 0.37 mmol, 17 eq), and DIPEA (96 µL, 0.44 mmol, 20 eq) in 1 mL DMF was added to the lantern and shaken at room temperature for 4 h. Lantern was washed 3x with 2 mL of DMF and 3x with 2 mL DCM. Compound **5** was cleaved off the lantern with 5% HF/pyridine in THF and quenched with methoxytrimethylsilane according to literature procedures.[109] Sample identity was confirmed by LCMS. Note: Use *extreme* caution while handling HF/pyridine solution. Hydrogen fluoride (HF) forms hydrofluoric acid upon contact with water in the air or human tissue. Hydrofluoric acid is both highly corrosive and toxic. See hydrogen fluoride safety data sheet (SDS) before using HF/pyridine.

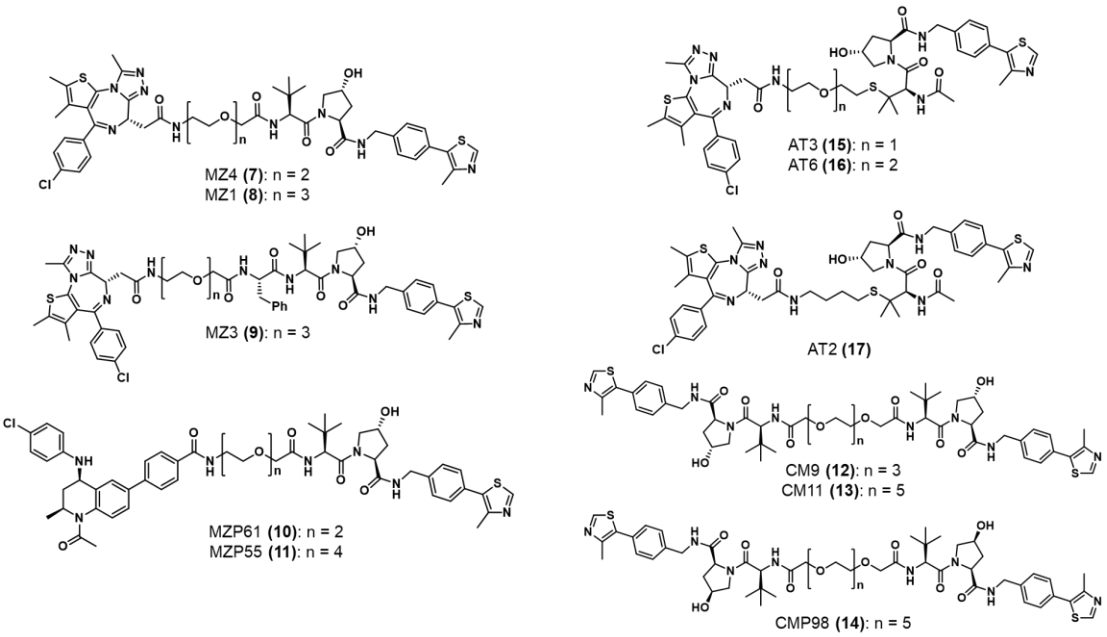
*2.5.2.5 Solid phase synthesis of compound **6***



A SynPhase lantern conjugated to compound **23** was treated with 2 mL of a solution of 2% DBU and 2% piperidine in DMF for 15 min at room temperature to remove the Fmoc protecting group. Lanterns were rinsed 3x with 2 mL of DMF and 3x with 2 mL DCM. A solution of Fmoc-12-amino-4,7,10-trioxadodecanoic acid (**24**) (67 mg, 0.15 mmol, 7 eq), HATU (57 mg, 0.15 mmol, 7 eq), and DIPEA (40 µL, 0.225 mmol, 10 eq) in DMF was added to a single lantern. Reactions vial containing the lanterns was mixed on a linear shaker for 4 h at room temperature to produce compound **26**. Lantern was washed 3x with 2 mL of DMF and 3x with 2 mL DCM. The Fmoc protecting group on **26** removed with 2 mL of a solution of 2% DBU and 2% piperidine in DMF for 15 min at room temperature. Lantern was washed 3x with 2 mL of DMF and 3x with 2 mL DCM. Next, a solution of phenylacetic acid (50 mg, 0.37 mmol, 17 eq), HATU (140 mg, 0.37 mmol, 17 eq), and DIPEA (96 µL, 0.44 mmol 20 eq) in 1 mL DMF was added to the lantern and shaken at room temperature for 4 h. Lantern was washed 3x with 2 mL of DMF and 3x with 2 mL DCM. Compound **6** was cleaved off the lantern with 5% HF/pyridine in THF and quenched with methoxytrimethylsilane according to literature procedures.[109] Sample identity was confirmed by LCMS. Note: Use *extreme* caution while handling HF/pyridine solution. Hydrogen fluoride (HF) forms hydrofluoric acid upon contact with water in the air or human tissue.

Hydrofluoric acid is both highly corrosive and toxic. See hydrogen fluoride safety data sheet (SDS) before using HF/pyridine.

*2.5.2.6 Synthesis of compound **7 – 17***



Compounds **7** – **17** were prepared as described in previously published literature, for compounds **7** – **9**,[81,82,98] for compounds **10** – **11**,[98] for compounds **12** – **14**,[99] and for compounds **15** – **17**.[82] Compound identity data (i.e. LC/MS and NMR) for these previously published compounds can be found in their associated citations as well.

2.5.3 Parallel artificial membrane permeability assay (PAMPA)

PAMPA[64,65] was used to determine the passive permeability of these compounds with modifications and calculations as described in Naylor *et al*. 2018[7] and here. Briefly, pure compound stocks (150 µL, 1 µM, 5% DMSO in phosphate-buffer

saline (PBS) pH 7.4) were added to a 96-well donor plate with 0.45 µm hydrophobic Immobilon-P membranes (Millipore MAIPNTR10) pre-soaked with a 1% solution of lecithin in n-dodecane. The donor plate was loaded in to a 96-well Teflon acceptor plate (Millipore MSSACCEPTOR) with 300 µL of PBS with 5% DMSO. After a ~15 h for compounds **7** – **17** or ~7 h for compounds **1** and **3** – **6**, the acceptor and donor plates were separated, and 50 µL was collect from each well. Samples were diluted with 50 µL methanol and analyzed by LC/MS as described above. Samples were run in quadruplicate with 1 µM propranolol as a control. Exact time in seconds was used in the calculations. Error was calculated as standard deviation.

2.5.4 $LogD_{(dec/w)}$ shake flask partition coefficient assay

Prior to performing this assay, 1,9-decadiene was saturated with H_2O by vortexing equal parts 1,9-decadiene and PBS (pH 7.4). The reagents are ready for use once the emulsion is fully separated. Three µL of a 200 µM DMSO stock was added to a 1.5 mL microcentrifuge tube. To this tube, 300 µL of H_2O-saturated 1,9-decadiene and 300 µL of PBS pH 7.4 were added. The final concentration of DMSO is 0.5%. The tubes were closed, sealed, and vortexed for 20 minutes. Tubes were then sonicated for 30 minutes, and centrifuged for 10 minutes at 14,000 x g. From each layer, 150 µL was removed and transferred to a costar 96-well plate (Coring, COSTAR 3357). To avoid contamination when samples from the water layer, a small air bubble was expelled from the pipette tip to displace any decadiene that accumulated in the pipette tip while moving from the organic layer to the lower aqueous layer. The solvent was evaporated

from the 96-well plate overnight using an EZ-2 Plus series Genevac centrifugal evaporator at 60 ºC. To each evaporated sample, 150 µL of DMSO was added. The 96-well plate was sealed and sonicated for 30 minutes. Samples were analyzed by LC/MS as described above. Assay was run in quadruplicate.

2.5.5 Lipophilic permeability efficiency (LPE) metric calculations

The experimental *1,9-decadiene*-water partition coefficient, $LogD_{(dec/w)}$, is highly sensitive to a molecule's net minimum hydrogen bond acidity in the membrane environment and reflects its ability to sequester HBDs. Whereas the calculated *octanol*-water partition coefficient ALogP is a knowledge-based 2-dimensional descriptor that reflects a compound's net overall lipophilicity in the aqueous environment (where polar groups are maximally exposed). By normalizing membrane partitioning by a compound's overall lipophilic character, LPE provides a measure of the permeability that is achievable at its maximum ALogP-defined lipophilicity. Beyond this, the addition of aliphatic character incurs solubility penalties, leading to, among other liabilities, diminished permeability. Because differences in LPE reflect free energy relationships derived from simple physical processes, group contributions are additive; any deviation from additivity can provide a sensitive indicator of scaffold effects on net polarity (i.e., HBD exposure) in the membrane environment. LPE was calculated following the procedure described by Naylor *et al.*[7] using the following equation:

$LPE = LogD_{(dec/w)} - 1.06(ALogP) + 5.47.$

2.5.6 Fluorescence Polarization (FP) assay

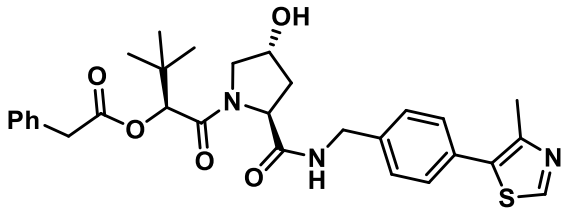
Fluorescence Polarization (FP) competitive binding assays were performed as described previously,[110] with all measurements taken using a PHERAstar FS (BMG LABTECH) with fluorescence excitation and emission wavelengths (λ) of 485 and 520 nm, respectively. Assays were run in triplicate using 384-well plates, with each well solution containing 15 nM VCB protein, 10 nM FAM-labeled HIF-1α peptide (FAM-DEALAHypYIPMDDDFQLRSF, "JC9") and decreasing concentrations of compound (14-point 2-fold serial dilution starting from 100 μM VH298 or 300 μM test compound). All components were dissolved from stock solutions using 100 mM Bis-Tris propane, 100 mM NaCl, 1 mM TCEP, pH 7.5, to yield a final assay volume of 15 μL. DMSO was added as appropriate to ensure a final concentration of 1% v/v. Control wells for zero displacement (VCB, JC9, no compound) and maximum displacement (JC9 and no protein) were also included. Percentage displacement values obtained by normalization to controls were plotted against Log[Compound], and IC50 values were determined for each titration using nonlinear regression analysis with Prism (v. 8.0.1, GraphPad). Ki values were back-calculated from the Kd for JC9 (3 nM, determined from direct binding) and fitted IC50 values, as described previously.[105,110]

2.5.7 Cell proliferation assay

Cell proliferation data were collected as described in Chan et al. 2018 pg. 511 titled "Cell Viability Assay".[98]

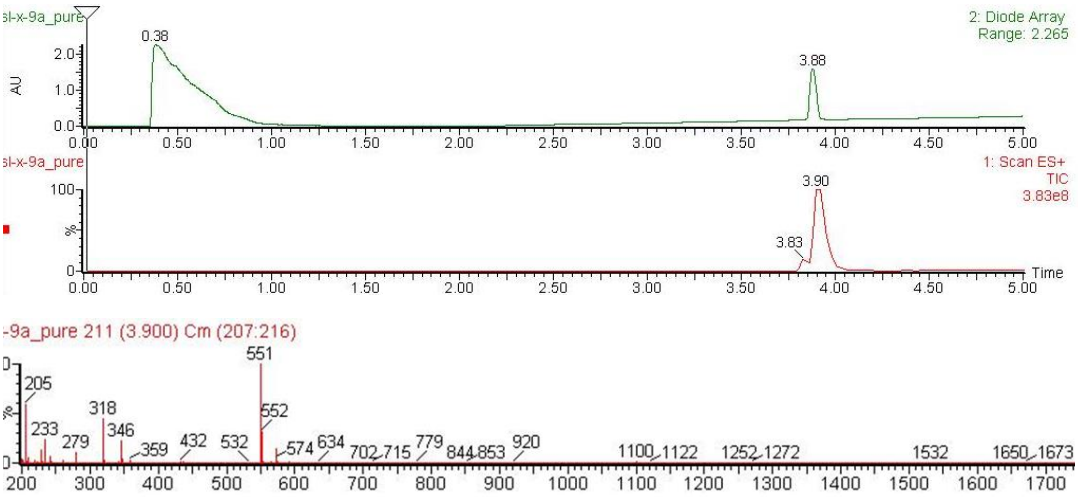
2.5.8 Spectra for synthesized compounds

*2.5.8.1 LC/MS trace for compound **3***

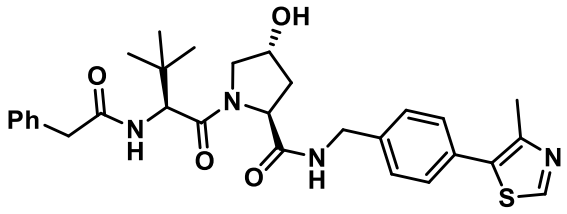


Molecular formula: $C_{30}H_{35}N_3O_5S$

Exact mass: 549, mass required (M+H): 550 mass found: 551, retention time 3.9 min

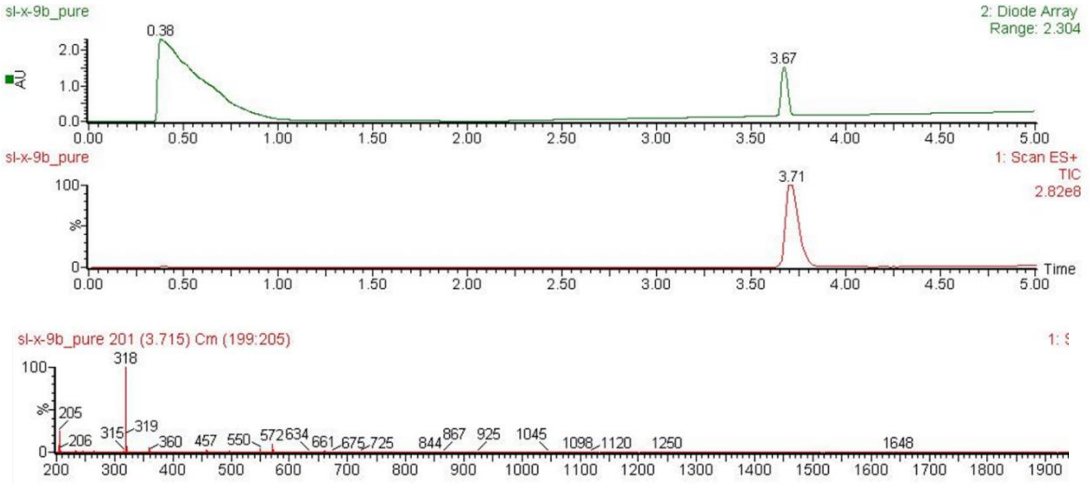


*2.5.8.2 LC/MS trace for compound **4***

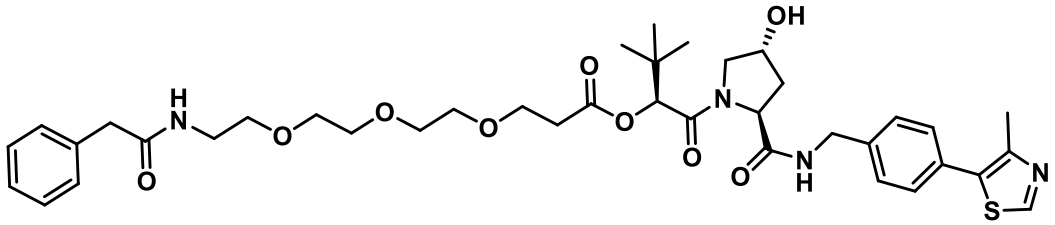


Molecular formula: $C_{30}H_{36}N_4O_4S$

Exact mass: 548, mass required (M+H): 549 mass found: 550, retention time 3.7 min

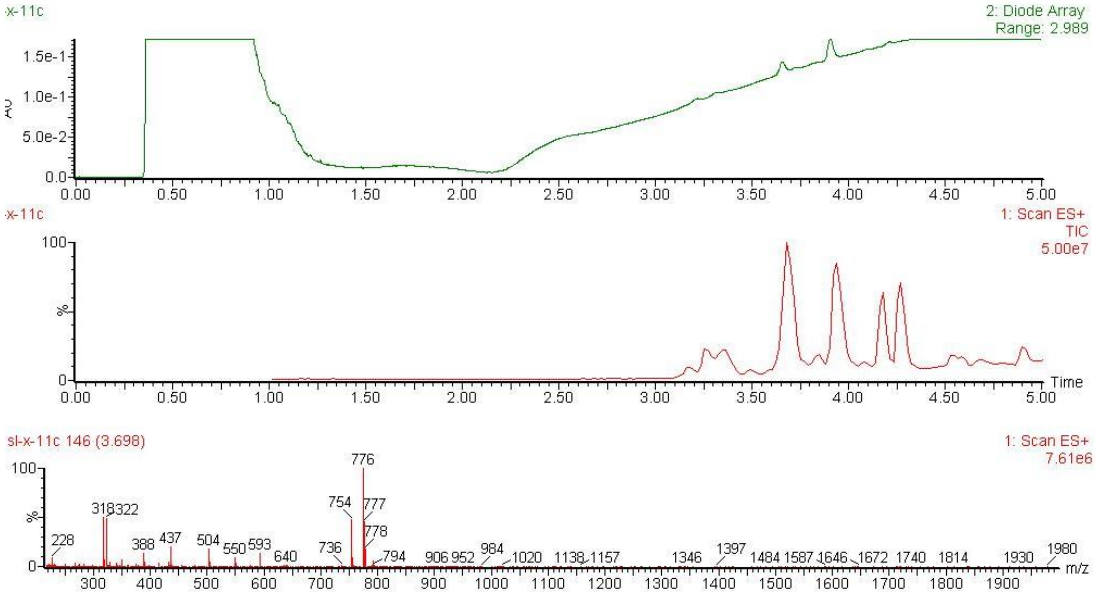


*2.5.8.3 LC/MS trace for compound **5***

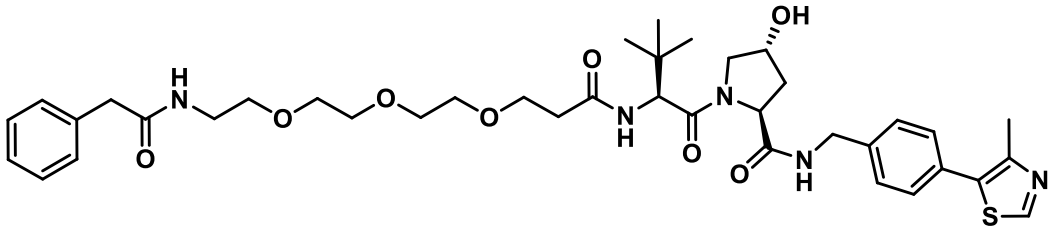


Molecular formula: $C_{39}H_{52}N_4O_9S$

Exact mass: 752, mass required (M+H): 753 mass found: 754, retention time 3.7 min

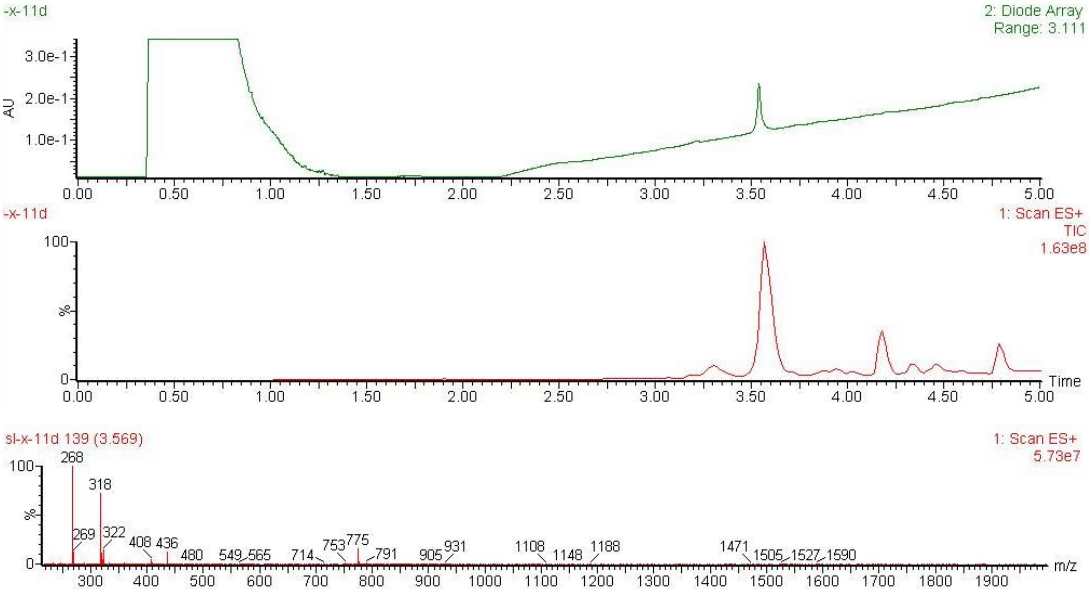


*2.5.8.4 LC/MS trace for compound **6***



Molecular formula: $C_{39}H_{53}N_5O_8S$

Exact mass: 751, mass required (M+H): 752 mass found: 753, retention time 3.6 min



2.6 Abbreviations

ACN: acetonitrile; CAPA: Chloroalkane Penetration Assay; DBU: 1,8-Diazabicyclo[5.4.0]undec-7-ene; DCM: Dichloromethane; DIC: N,N′-Diisopropylcarbodiimide; DIPEA: N,N-Diisopropylethylamine; DMAP: 4-Dimethylaminopyridine; DMF: N,N- Dimethylformamide; FP: Fluorescence polarization; HATU: 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate; HBA: hydrogen bond acceptor; HBD: hydrogen bond donor; HF/pyridine: Hydrogen fluoride pyridine; IMHB: intramolecular hydrogen bond; LPE: lipophilic permeability efficiency; MW: molecular weight; PAMPA: parallel artificial membrane permeability assay; PROTAC: Proteolysis targeting chimera; TFA: Trifluoroacetic acid; THF: Tetrahydrofuran; VHL: von Hippel-Lindau

2.7 Author Contributions

V.G.K., R.S.L, and A.C. designed the project and wrote the manuscript. Permeability experiments were performed by V.G.K. Mass spectrometry data was processed by C.E.T. FP data was collected and processed by S.J.H. Compounds were synthesized by R.S.L., A.T., M.Z., and C.M with design support from A.C. Cell proliferation data was collected by K.-H.C.) The manuscript was edited by all.

2.8 Funding Sources

The R.S.L. group was funded by the National Institute of General Medicine Studies of the National Institutes of Health (NIH R01GM131135). and the National Science

Foundation GRFP (NSF DGE 1339067 to V.G.K). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NIH or the NSF. The A.C. lab was funded by awards from the UK Biotechnology and Biological Sciences Research Council (BBSRC, grant BB/J001201/2), the European Research Council (ERC, Starting Grant ERC-2012-StG-311460 DrugE3CRLs), the Italian Ministry of Education, University and Research (Miur, Ph.D. Studentship to C.M.), and the European Commission (Marie Skłodowska-Curie Actions Individual Fellowship H2020-MSCA-IF-2014-655516 to K.-H.C.).

2.9 Notes

The authors declare the following competing financial interest(s): The A.C. laboratory receives or has received sponsored research support from Boehringer Ingelheim, Eisai Co., Nurix, Ono Pharmaceuticals, and Amphista Therapeutics. A.C. is a scientific founder, shareholder, non-executive director and consultant of Amphista Therapeutics, a company that is developing targeted protein degradation therapeutic platforms.

Chapter Three

Amide-to-Ester Substitutions improve PROTAC permeability

This chapter contains text and figures from the following manuscript: Klein, V. G., Sun, Y., Lokey, R.S. *Ester-to-amide substitutions improve PROTAC permeability* (Manuscript in preparation)

Abstract

Proteolysis targeting chimeras (PROTACs), also known as targeted degraders, are heterobifunctional compounds that harness the power of the cells E3 ubiquitin ligase proteasomal degradation pathway to degrade their target catalytically and selectively. Since their initial development in the early 2000s, PROTACs have been used to target a wide range of targets, but their development remains difficult and time consuming. As these types of compounds move towards the clinic, it is crucial to understand their physicochemical properties so that these hard to make degraders are not triaged for their poor pharmacokinetic properties. Here, we investigate the effects of an amide-to-ester substitution on PROTAC permeability, plasma stability, and lipophilic permeability efficiency (LPE) over a broad range of calculated lipophilicities and several linker types. We discovered that this amide-to-ester substitution improves permeability over the common PROTAC lipophilicity range and these ester containing-compounds remain stable in plasma when the ester is close to a bulky, drug-like side chain. This study offers design suggestions to bias PROTACs towards better physicochemical properties.

3.1 Introduction

Targeted degraders, also known as Proteolysis Targeting Chimeras (PROTACs), are becoming a widespread source of lead compounds aimed at expanding the "druggable" proteome and improving existing drugs.[79-81,98,111] PROTACs are heterobifunctional molecules typically contain a protein of interest (POI)-targeting ligand (or warhead) and a E3 ubiquitin ligase-targeting ligand connected by a flexible linker.[112,113] These degraders trigger catalytic and targeted degradation by recruiting both the POI and an E3 ubiquitin ligase, inducing ternary complex formation, ubiquitylation, and consequent proteasomal degradation of the POI.[114] Due their catalytic activity and creation of ternary complex formation, PROTACs overcome certain small molecule limitation and offer higher potency, higher selectivity, and fewer off-target effects compared to other drug compounds.[83,115] Furthermore, unlike traditional small molecule drugs, PROTAC activity is not dependent on the presence of a well-defined active site.[115,116] It is clear that PROTACs are an area of great interest in drug discovery. The number of publications is rapidly on the rise and there have been over 40 targets degraded thus far.[111] Notably, two PROTACs (ARV-110 and ARV-147) have recently moved into clinical trials.[117]

While numerous papers and reviews tout the benefits of the PROTAC strategy for targeted degradation, recent papers have also cited concerns over the potentially poor physicochemical properties of these "beyond rule of 5" compounds.[86,118,119] Similarly, others highlight the importance of investigating the physicochemical and pharmacokinetic properties of PROTACs as is key with other types of drug

molecules.[85,120] Recently, we and others have established that several highly potent PROTACs have very low permeabilities compared to other bioactive compounds and their individual degrader components.[66,91,121] Now that the low permeability of these compounds has been established, we were interested in improving the permeability of these compounds in an effort to improve their bioactivity.

In our previous work, we demonstrated that an amide-to-ester substitution at the *tert*-Leu of the von Hippel–Lindau (VHL)-recruiting ligand can increase membrane permeability.[66] While effective, these ester modification yielded only a modest increase in permeability over their amide counterparts due to the relatively high shielding at this position from the β-branched amino acid sidechain.[51] We suggested that substituting the amide connecting the linker to the POI warhead for and ester would cause a more dramatic increase in permeability. Therefore, to build on our proof-of-concept study, we developed a systematic set of compounds to test this hypothesis across a wide range of lipophilicities (ALogP) and linker lengths. In this study, we demonstrate that amide-to-ester substitutions increase PROTAC considerably at low to moderate lipophilicities. Furthermore, ester-containing compounds with drug-like side chains experienced only moderately reduced plasma stability compared to the amide parent compounds. These new physicochemical insights establish lipophilicity design parameters for the development of permeable PROTACs and offer strategies for improving PROTAC permeability.

3.2 Results and Discussion

3.2.1 PROTAC permeability increases with lipophilicity up to a threshold

Lipophilicity is key in the design of drugs with favorable ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties.[122] While suggested optimal lipophilicity ranges exist for typical small molecule drugs[123] and "beyond Rule of 5" compounds,[7] design parameters for PROTAC ideal lipophilicity remains unclear. In this paper, we performed a systematic investigation into the effect of lipophilicity on permeability for a set of seven model VHL-based PROTAC compounds (**1 – 7**, Figure 3-1). All the compounds had a VHL-based E3 ligase ligand (VH032) and a linker.

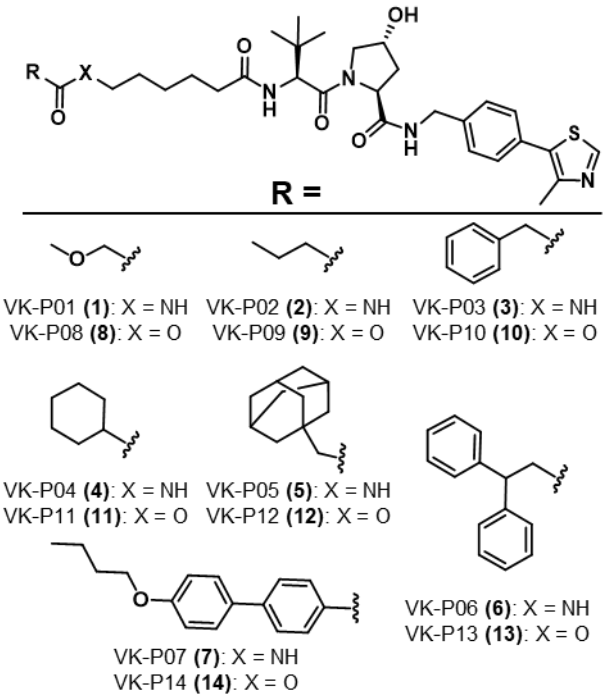


Figure 3-1: PROTAC liposcan compound structures Compounds are composed of a VH032 (VHL-targeting) ligand and a short alkyl linker which is connect to the warhead (POI-targeting) ligand (R) with either and amide (**1** – **7**) or and ester (**8** – **14**) (X)

We modulated the compounds' lipophilicities by using a variety of simple POI-targeting group mimics for a range of calculated lipophilicities (ALogP) from 1.2 – 6.0. As permeability can be affected by physicochemical properties like molecular weight (MW) and the number of hydrogen bond donors (HBDs) and acceptors (HBAs), we kept these values in a relatively small range (MW = 600 – 800, HBD = 3 – 4, HBA = 6 – 8, Table 3-1). Furthermore, we used a short alkyl linker for all of these compounds to eliminated permeability-affecting intramolecular hydrogen bonds (IMHBs) that can be formed between PEG linkers and amide -NHs in other parts of the molecule.[66]

Table 3-1: Physical properties of amide liposcan compounds Table showing the molecular weight (MW), calculated lipophilicity (ALogP), and the number of HBDs and HBAs of the first seven liposcan compounds that have an amide bond linking their POI-targeting mimic to their linker

	Cmpd[a]	MW[b]	ALogP	# of HBDs[c]	# of HBAs[d]
Amide	**VK-P01 (1)**	616	1.2	4	7
	VK-P02 (2)	614	2.6	4	6
	VK-P03 (3)	662	3.2	4	6
	VK-P04 (4)	654	3.6	4	6
	VK-P05 (5)	720	4.1	4	6
	VK-P06 (6)	752	4.9	4	6
	VK-P07 (7)	796	6.0	4	7

a: Cmpd = compound
b: MW = molecular weight
c: HBD = hydrogen bond donor
d: HBA = hydrogen bond acceptor

We first investigated the effects of PROTAC lipophilicity on permeability using PAMPA, a high-throughput permeability assay that is cheaper, faster, and well correlated to cell-based permeability assays.[65] Our group has also shown that PAMPA is beneficial for studying compounds with low expected permeabilities as the PAMPA

due to the assays low limit of detection.[66] Similar to other types of previously studied compounds,[7] PROTACs permeability increased with ALogP up to an ALogP of around 4.1 (cf. **1 – 5**, Figure 3-2, SI Table 3-S1). Above an ALogP of 4, permeability decreased as ALogP increased (cf. **6 – 7**) with no detectable permeability for **7** which had an ALogP of 6.0 (SI Table 3-S1). At these higher ALogP values (>4 – 5), compounds begin to lose aqueous solubility and become membrane retained, both of which can reduce passive membrane permeability.[124] This compound series suggests that PROTACs should be designed with an ALogP between 3 – 5 to bias them towards higher permeability. Moreover, relationship between lipophilicity and permeability offers a route to improving the permeability of PROTACs by making small modifications to increase ALogP up to ~4-5.

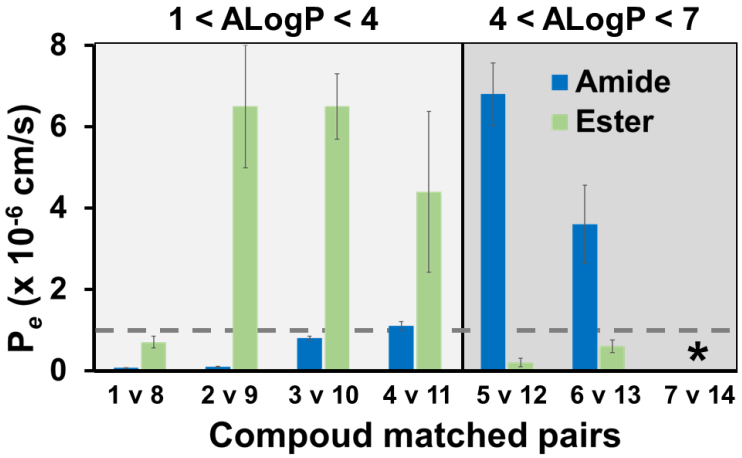


Figure 3-2: PAMPA permeability of liposcan matched pairs Comparison of amide- (blue) vs. ester-containing (green) liposcan compound PAMPA permeability (P_e) grouped by their lipophilicity-tuned warhead. Compounds are ordered by their calculated lipophilicity (ALogP). The start (*****) over compounds **7** and **14** (**7** v **14**) indicates that there was no detectable signal in the acceptor well. N = 4. Error bars are ±SD.

3.2.2 Amide-to-ester substitutions improve permeability over a broad ALogP range

In addition to lipophilicity, the number of HBDs in compounds is a crucial determinate of permeability.[103,125] Reducing the presence of solvent-exposed HBDs though N-methylation or occlusion from solvent by β-branching or other steric shielding are common methods to increase a compound's membrane permeability.[51,70,126,127] In our previous investigation on PROTAC permeability, we demonstrated that substituting the VH032 ligand's *tert*-Leu amide with an ester improved PROTAC permeability by about 2-fold.[66] Though a viable approach to improving permeability, substituting this amide for an ester only conferred a modest increase in permeability, likely due to the fact that the -NH was partial shielded from solvent by the adjacent β-branched α-carbon, limiting the permeability reducing effects of this HBD.[51,66] Additionally, substituting this amide (connecting the VHL ligand to the linker) for an ester reduced its affinity for the VHL protein.[66] Therefore, we created a new set of compounds with an amide to ester substitution at the other end of the linker in order to more strongly improve permeability while maintaining binding to the VHL E3 ligase.

We made a second set of liposcan compounds (**8 – 14**) to test if these ester-containing compounds were more permeable across a broad ALogP range of 1.9 – 6.6 than the amide derivatives (Figure 3-1, Table 3-2). These compounds were identical to the previous amide-containing compounds except for one amide to ester substitution at

the connection between the linker and the protein-targeting ligand model, creating 7 matched pairs for permeability analysis.

Table 3-2: Physical properties of ester liposcan compounds Table showing the molecular weight (MW), calculated lipophilicity (ALogP), and the number of HBDs and HBAs of the second seven liposcan compounds that have an ester bond linking their POI-targeting mimic to their linker

	Cmpd[a]	MW[b]	ALogP	# of HBDs[c]	# of HBAs[d]
Ester	VK-P08 (8)	617	1.9	3	8
	VK-P09 (9)	615	3.2	3	7
	VK-P10 (10)	663	3.8	3	7
	VK-P11 (11)	655	4.3	3	7
	VK-P12 (12)	721	4.7	3	7
	VK-P13 (13)	753	5.6	3	7
	VK-P14 (14)	797	6.6	3	8

a: Cmpd = compound
b: MW = molecular weight
c: HBD = hydrogen bond donor
d: HBA = hydrogen bond acceptor

Over an ALogP range of 1 – 4, the ester-containing compounds were 4- to 65-fold more permeable than their amide-containing counterparts (Figure 3-2, SI Table 3-S1). Substituting an amide for an ester not only removes an HBD, but also increases the ALogP by about 0.6. Both the reduction of HBDs and increase in ALogP are likely to lead to increased permeability at this ALogP range.[7,51,128] However, over an ALogP of ~4, the ester compounds were less permeable than their amide counterpart (Figure 3-2, SI Table 3-S1). This is likely due to the inverse relationship between permeability and lipophilicity as the ALogP increases over 4 due to a decreased aqueous solubility.[129] Furthermore, it is possible that the additional HBD present in the amide compounds could confer increased solubility over the ester derivatives. Similar to the amide-

containing **7** (ALogP = 6), its ester-containing counterpart, **14**, (ALogP = 6.6) had no detectable permeability (Figure 3-2, SI Table 3-S1). Overall, amide-to-ester substitution offers a potential mechanism to improve PROTAC permeability.

3.2.3 Amide-to-ester substitutions increase permeability across a range of linkers

Previous works have suggested that short alkyl linkers may be better for PROTAC permeability as these linkers help to minimize the already high TPSA and number of hydrogen bond acceptors present.[91,121] However, we have shown that effect of the linker on PROTAC permeability can be confounded by hydrogen bonding and the overall lipophilicity.[66] For this study, we designed a systematic set of four compounds to isolate the effects of the linker on permeability by reducing the protein targeting group to a simple phenyl group attached by an amide. The linkers included a short alkyl linker (**3**) and three PEG linkers ranging from one to three PEG units in length (**15** – **17**, respectively) (Figure 3-3 A). The alkyl-containing **3** had the highest permeability, and permeability decreased with increasing PEG chain linker length with **17** (3 PEG unit linker) showing no detectable permeability (Figure 3-3 B, SI Table 3-S2). This decrease in permeability is likely due to a combined effect of increased HBAs, increased MW, and decreased ALogP due to the increasing PEG chain length. As an amide-to-ester substitution improved the permeability of the liposcan model compounds across a broad range of lipophilicities, we made a second set of compounds with an amide to ester substitution between the linker and the protein-targeting ligand (Figure 3-3 A, **10**, **18** – **20**).

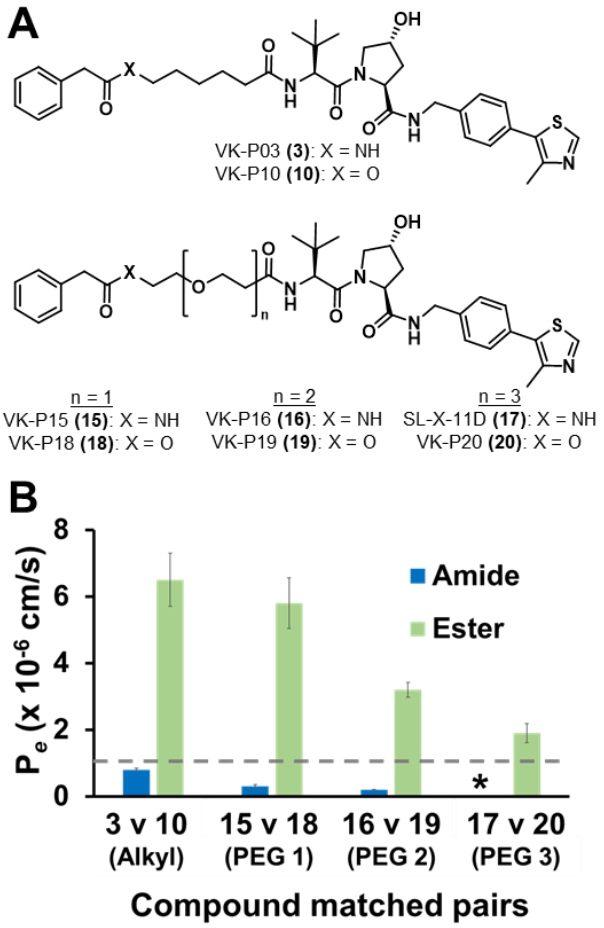


Figure 3-3: Structures and PAMPA permeabilities of PROTACs with varied linkers *(A)* Compounds are composed of a VH032 (VHL-targeting) ligand and phenyl ring acting as a warhead mimic which are connected by one of four linkers. The warhead (POI-targeting) ligand is connected to the linker with either and amide (**3**, **15** – 17) or and ester (**10**, **18** – **20**). *(B)* Comparison of amide- (blue) vs. ester-containing (green) compound PAMPA permeability (P_e) grouped by their linker type. The start (*****) over compound **17** indicates that there was no detectable signal in the acceptor well for this compound. N = 4. Error bars are ±SD.

As with the liposcan series of compounds, the ester-containing PROTACs were all had detectable permeabilities and were 8 – 19-fold more permeable than their amide-containing counterparts (Figure 3-3 B, SI Table 3-S2). Unlike the amide containing compounds which all had a permeability below P_e = 1 x 10^{-6} cm/s, all the ester

compounds had permeabilities in the modest to good range (P_e > 1 x 10^{-6} cm/s). Thus, an amide-to-ester substitution not only improves permeability, but also offers more flexibility in compound design as a wider range of ester linkers are more likely to be permeable compared to their amide counterparts. This design flexibility in crucial since small modifications to the linker significantly affect PROTAC bioactivity, ternary complex formation, and subsequent targeted degradation.[81,130,131]

3.2.4 PROTACs exhibit ligand-to-linker intramolecular hydrogen bonds

The basic structure of PROTACs, two small molecules connected by a flexible linker, lends itself to the formation of intramolecular hydrogen bonds (IMHBs). This important feature allows for polar atoms to be shuttled across the lipophilic cell membrane. It is difficult to determine the presence of IMHBs by looking at the 2D chemical structure alone. However, measuring the lipophilic permeability efficiency (LPE) of matched pairs indicates differences in the number of exposed HBDs.[7,66] LPE is a metric developed by our lab that balances aqueous solubility (calculated ALogP) and membrane partitioning (experimental $LogD_{(dec/w)}$) to determine the efficiency with which a compounds crosses a membrane at a given lipophilicity. This metric is particularly helpful in determining differences in solvent-exposed HBDs between compounds. Compounds with similar LPE values are likely to have the same number of solvent exposed HBDs. While a ΔLPE of 1.8 suggests the that the compound with a lower LPE has an additional exposed HBD compared to the compound with the higher LPE.

For the majority of the liposcan compound pairs (**2** – **6** vs. **9 – 13**, respectively), the ester compounds had higher LPEs than their counterpart amide compounds (Table 3-3). The ΔLPEs of between 1.1 – 1.8, suggests that the additional HBD in the amide compounds is partially to fully solvent exposed. Interestingly, for the majority of the amide compounds with a HBA four atoms away from the amide -NH, the LPEs of the matched pairs were more similar (**1** vs. **8**, **15** vs. **18**, and **17** vs. **20**, Table 3-1). Similar to previous work,[66] this suggests that the HBD in the amide compounds is making and IMHB with the oxygen in the PEG linker or warhead which we have shown in the past. Therefore, while ester bonds and alkyl linkers are better for permeability, when used in combination, a PEG linker and amide bond could be used to shield the polarity of important HBDs that are crucial to the bioactivity or solubility of the overall molecule.

Table 3-3: PROTAC lipophilic permeability efficiency (LPE) Table comparing amide and ester matched compound pairs. LPE = $LogD_{(dec/w)}$ - 1.06(ALogP) + 5.47.

	LPE			
	Amide		Ester	
Liposcan	VK-P01	2.0	1.8	VK-P08
	VK-P02	0.5	1.6	VK-P09
	VK-P03	0.6	2.3	VK-P10
	VK-P04	0.3	2.1	VK-P11
	VK-P05	1.0	2.8	VK-P12
	VK-P06	-0.1	1.5	VK-P13
	VK-P07	0.1	N/A	VK-P14
Linker scan	VK-P03	0.6	2.3	VK-P10
	VK-P15	2.2	2.7	VK-P18
	VK-P16	1.4	2.5	VK-P19
	VK-P17	2.0	2.4	VK-P20

3.2.5 Relative plasma stability of amide and ester compounds

While amide-to-ester substitutions offer increased permeability leading to increased flexibility in compound design, esters are also typically more susceptible to plasma-mediated hydrolysis which can lead to low *in vivo* efficiency.[132] For these ester substitutions to be a viable option in drug development, it is important to assess the differential stability of ester and amide compounds. To test this, we incubated **1 – 20** in human plasma at 37 ºC for 0, 15, 30, and 90 minutes. Overall, the amide-containing compounds were more stable in plasma than their ester-containing counterparts for the compounds that had small. This effect was more pronounced for the compound pairs with the smaller warheads with 10% or less compound loss of the amide compounds (**1 – 4**) compared the their ester counterparts (**8 – 11**) which had 60 – 90% compound loss at 90 minutes (SI Figure 3-S1). Ester compounds with larger warheads (**12 – 14**), and likely more steric shielding around the susceptible ester, had much less compound loss at 90 minutes (4 – 10%, SI Figure 3-S1). This reduced hydrolysis with bulkier sides that are more similar to typical protein-targeting groups present on PROTACS suggests that amide-to-ester substitutions could be used to increase PROTAC permeability with only mild reduction in the compound's *in vivo* activity.

3.3 Discussion and Conclusions

Composed to two small molecules and a linker, targeted degraders typically fall in "beyond Rule of 5" space.37 Drugs discovery has been shifting away from designing drugs in the traditional Ro5 model27,38 as more beyond Ro5 compounds are shown to be orally bioavailable.39,40 However, having some kind of guideline for physicochemical design parameters for beyond Ro5 compounds, like PROTACs, is critical to reduce high attrition rates of compounds that are lost to poor ADME properties in the clinic.41 In response to this pressing need, some have attempted to improve permeability, solubility, and to decrease efflux ratios through linker modifications,42 while other have attempted to reduce the number of remove amide bonds (HBDs) to improve the physicochemical properties of PROTACs.43 We and others have attempted to develop systematic studies of PROTAC physicochemical properties and new methods to studying these properties.16,18-20

Here, we used a systematic investigation of linker lengths and lipophilicity in combination with amide-to-ester substitutions to improve PROTAC permeability. We demonstrated the PROTACs achieve the highest permeability at moderate lipophilicities (2 – 4) and that, within this range, increasing the lipophilicity of a compound leads to increased permeability, as has been seen with other beyond Ro5 compounds.24 Recent medicinal chemistry reviews have cited lipophilicity as one of the most important features molecular properties, and that designing compounds in this range (that we have found to contain more permeable compounds) is also likely to reduce toxicity.41 We also demonstrate that amide-to-ester substitutions increase

PROTAC permeability in this ALogP range as well. Therefore, ester-containing compounds in this lipophilicity range are likely to have better overall pharmacokinetic properties than amide compounds or those with higher lipophilicities. Finally, though esters are more prone to hydrolysis and therefore tend to be less stable in plasma, we discovered that adding steric bulk to the chemical space surrounding the area (i.e. near the warhead) drastically reduces degradation in the plasma. Therefore, amide-to-ester substitutions remain a viable option for PROTAC pharmacokinetic improvement as long as the benefits in permeability outweigh the slight decrease in stability, leading to more compound reaching its protein target.

PROTACs have very low permeability but remain highly potent in part due to their catalytic activity and degradation of their target.18-20 We have previously determined that ROTACs with extremely low permeability were less active than similar PROTACs with better permeability.18 Therefore, a simple strategy to improve PROTAC bioactivity would be to improve PRTOAC permeability in subtle ways that do not affect complex PROTAC binding and ternary complex formation. The amide-to-ester substitution explored here is one possible avenue to achieve this goal. Further analysis and compounds should be developed to test this method on previously published PRTOACs to determine if improving permeability does in fact improve bioactivity.

3.4 Supplementary Figures and Tables

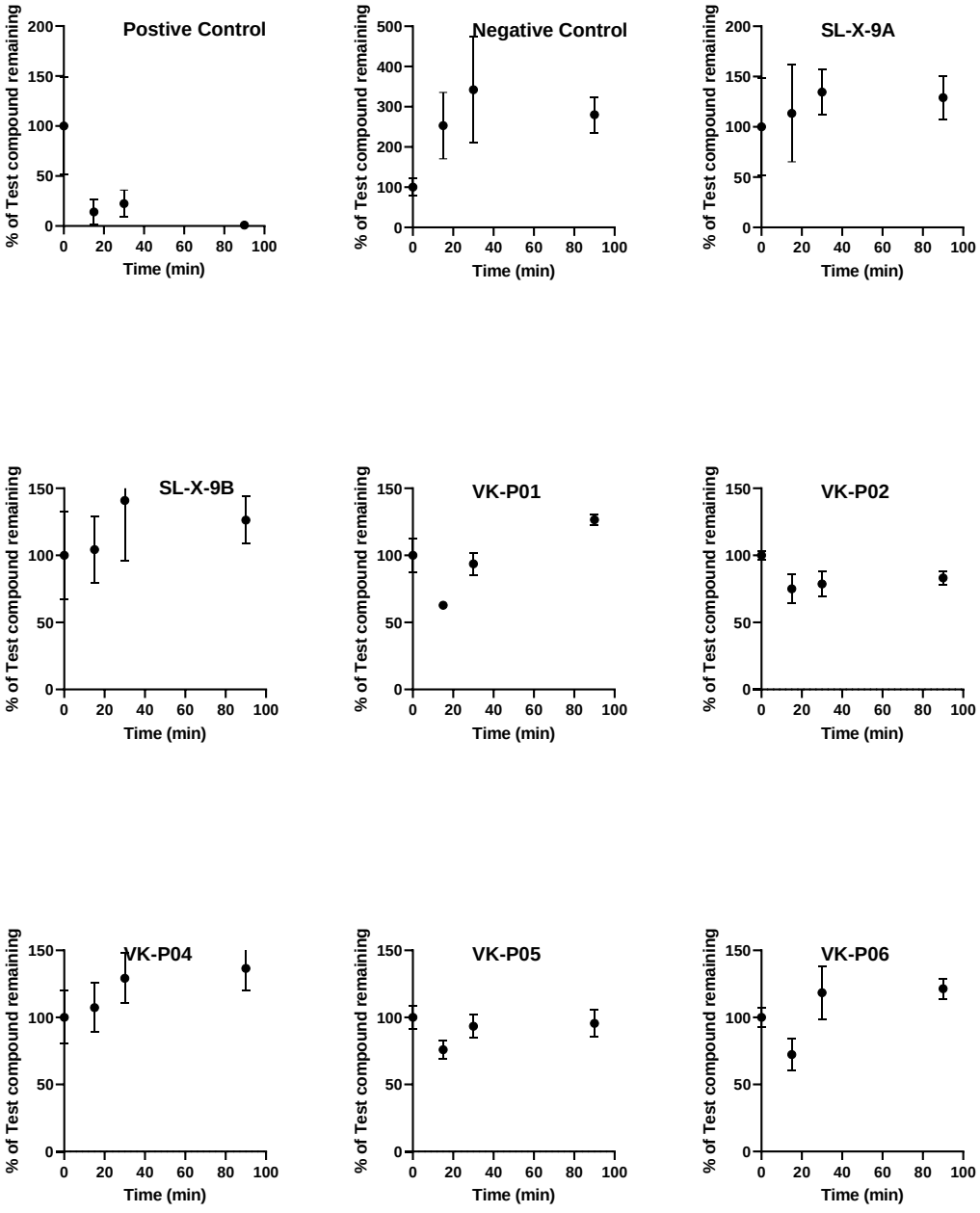
3.4.1 Supplementary Table 3-S1: Combine PROTAC physicochemical data

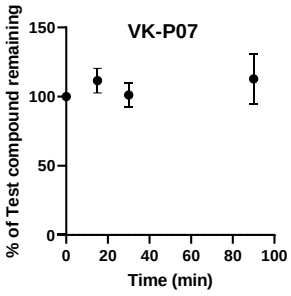
Table showing the physicochemical properties of all the compounds tested in this study. This includes calculated properties (MW, ALogP, HBD, and HBA), experimental results (PAMPA, LogD), and an efficiency metric (LPE).

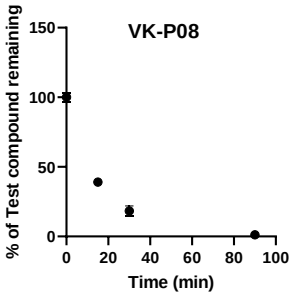
		Cmpd[a]	MW[b]	ALogP	# of HBDs[c]	# of HBAs[d]	PAMPA[e]	LogD (dec/w)[f]	LPE[g]
Liposcan	Amide	VK-P01 (1)	616	1.2	4	7	0.07	-2.2	2.0
		VK-P02 (2)	614	2.6	4	6	0.1	-2.2	0.5
		VK-P03 (3)	662	3.2	4	6	0.8	-1.5	0.6
		VK-P04 (4)	654	3.6	4	6	1.1	-1.4	0.3
		VK-P05 (5)	720	4.1	4	6	6.8	-0.2	1.0
		VK-P06 (6)	752	4.9	4	6	3.6	-0.3	-0.1
		VK-P07 (7)	796	6.0	4	7	0	1.0	0.1
	Ester	VK-P08 (8)	617	1.9	3	8	0.7	-1.7	1.8
		VK-P09 (9)	615	3.2	3	7	6.5	-0.5	1.6
		VK-P10 (10)	663	3.8	3	7	6.5	0.8	2.3
		VK-P11 (11)	655	4.3	3	7	4.4	1.1	2.1
		VK-P12 (12)	721	4.7	3	7	0.2	2.4	2.8
		VK-P13 (13)	753	5.6	3	7	0.6	1.9	1.5
		VK-P14 (14)	797	6.6	3	8	0	BLD	--
Linkers	Amide	VK-P15 (15)	664	1.8	4	7	0.3	-1.4	2.2
		VK-P16 (16)	708	1.7	4	8	0.2	-2.3	1.4
		VK-P17 (17)	752	1.6	4	9	0	-1.9	2.0
	Ester	VK-P18 (18)	665	2.5	3	8	5.8	-0.2	2.7
		VK-P19 (19)	709	2.3	3	9	3.2	-0.5	2.5
		VK-P20 (20)	753	2.2	3	10	1.9	-0.7	2.4

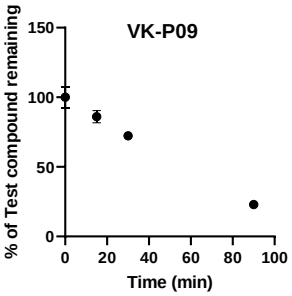
a: Cmpd = compound
b: MW = molecular weight
c: HBD = hydrogen bond donor
d: HBA = hydrogen bond acceptor
e: PAMPA = P_e x 10^{-6} cm/s
f: 1,9-decadiene and PBS 7.4 partition coefficients
g: LPE = $LogD_{(dec/w)}$ - 1.06(ALogP) + 5.47

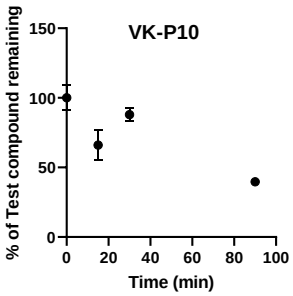
3.4.2 Supplementary Figure 3-S1: Plasma Stability Assay Graphs

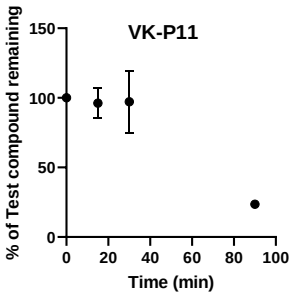


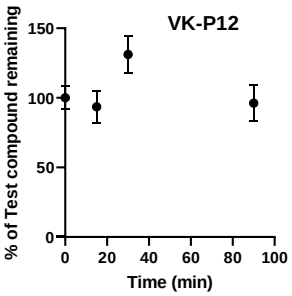

VK-P07
% of Test compound remaining
Time (min)
0
50
100
150
0
20
40
60
80
100

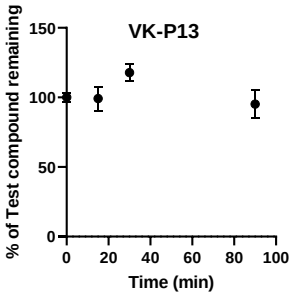

VK-P08
% of Test compound remaining
Time (min)
0
50
100
150
0
20
40
60
80
100

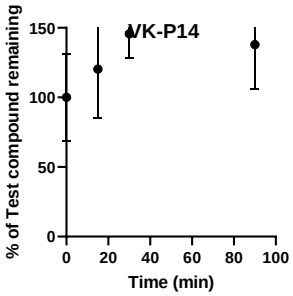

VK-P09
% of Test compound remaining
Time (min)
0
50
100
150
0
20
40
60
80
100

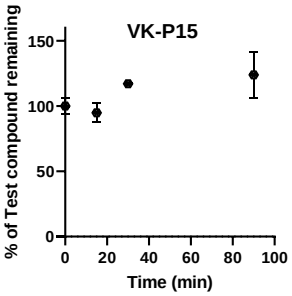

VK-P10
% of Test compound remaining
Time (min)
0
50
100
150
0
20
40
60
80
100

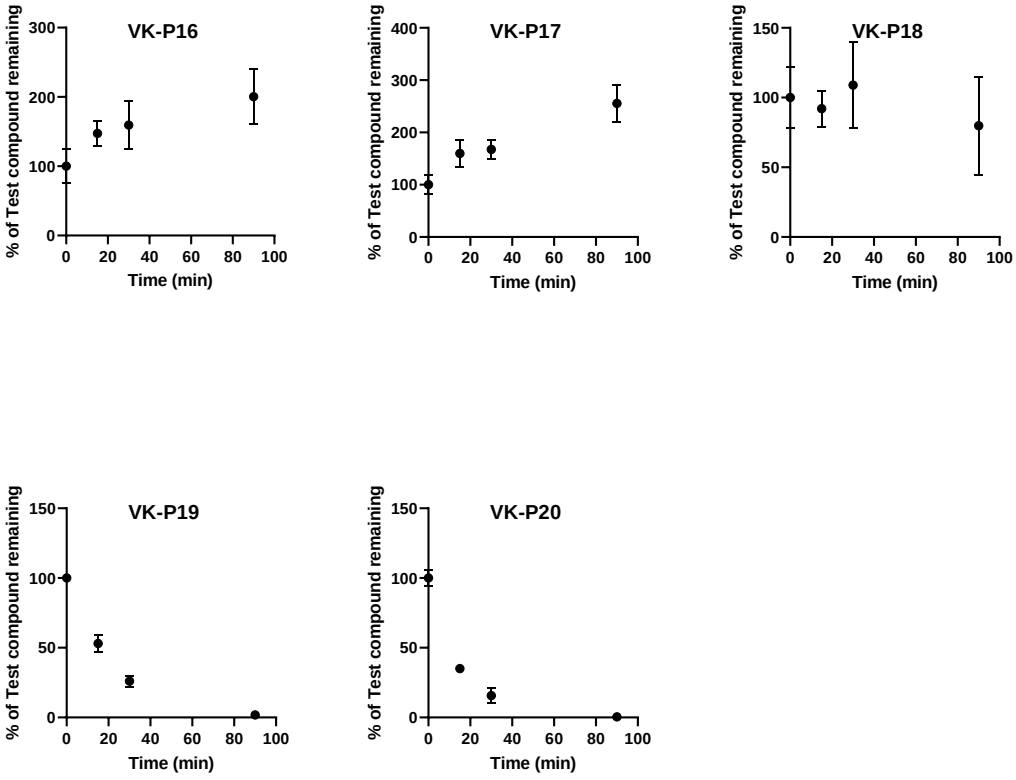

VK-P11
% of Test compound remaining
Time (min)
0
50
100
150
0
20
40
60
80
100


VK-P12
% of Test compound remaining
Time (min)
0
50
100
150
0
20
40
60
80
100


VK-P13
% of Test compound remaining
Time (min)
0
50
100
150
0
20
40
60
80
100


VK-P14
% of Test compound remaining
Time (min)
0
50
100
150
0
20
40
60
80
100


VK-P15
% of Test compound remaining
Time (min)
0
50
100
150
0
20
40
60
80
100



Graphs showing the percent of compound remaining for each compound after a 0, 15, 30, and 90 min incubation in pooled human plasma at 37 ºC relative to the amount of compound at a 0 min incubation time point. Error bars = ±SD.

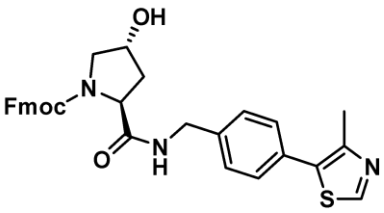
3.5 Methods

3.5.1 General Synthetic Information and Procedures

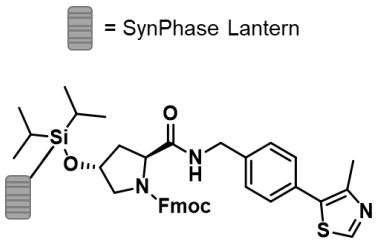
Unless otherwise stated, purchased solvents and reagents were used without further modification. Solvents were purchased from Fisher Scientific. General reagents were purchased from Fisher Scientific except for the following: HATU (Chem-Impex or Combi-Blocks), amino acids and linkers (Combi-Blocks or Oakwood), and SynPhase polystyrene lanterns (Mimotopes). A Thermo Scientific Ultimate 3000 UPLC system and Thermo Scientific Orbitrap VelosPro mass spectrometer were used to run LC/MS-based assays (PAMPA and $LogD_{(dec/w)}$) eluting with 5 – 95% ACN in H_2O with 0.1% formic acid. This system was fitted with a Thermo Hypersil GOLD C18 (30 mm x 2.1 mm, 1.9 µm particle size) column. LC/MS purity traces were collected using a Thermo Finnigan Surveyor HPLC system and Thermo Fisher Scientific Finnigan LTQ mass spectrometer. These samples were eluted with 10 – 100% ACN in H_2O with 0.1% formic acid on an Agilent Poroshell 120 EC-C18 (30 mm x 2.1 mm, 2.7 µm particle size). NMR samples were collected on Bruker 500 MHz NMR with a 5 mm BBO Smart Probe in deuterate chloroform unless otherwise stated.

3.5.2 Synthetic Methods

3.5.2.1 Loading SynPhase polystyrene L-series lantern



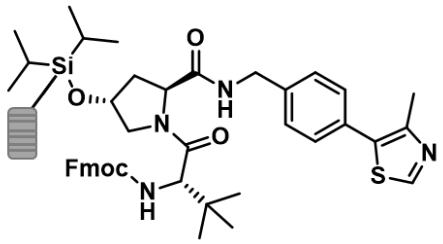
26



27

Compound **26** was synthesized and conjugated onto SynPhase polystyrene L-series lanterns (alkyl tethered diisopropylarylsilane linker, 22 µM, catalog # MIL10431000) following previously published protocols to generate compound **27**.[66,96,109]

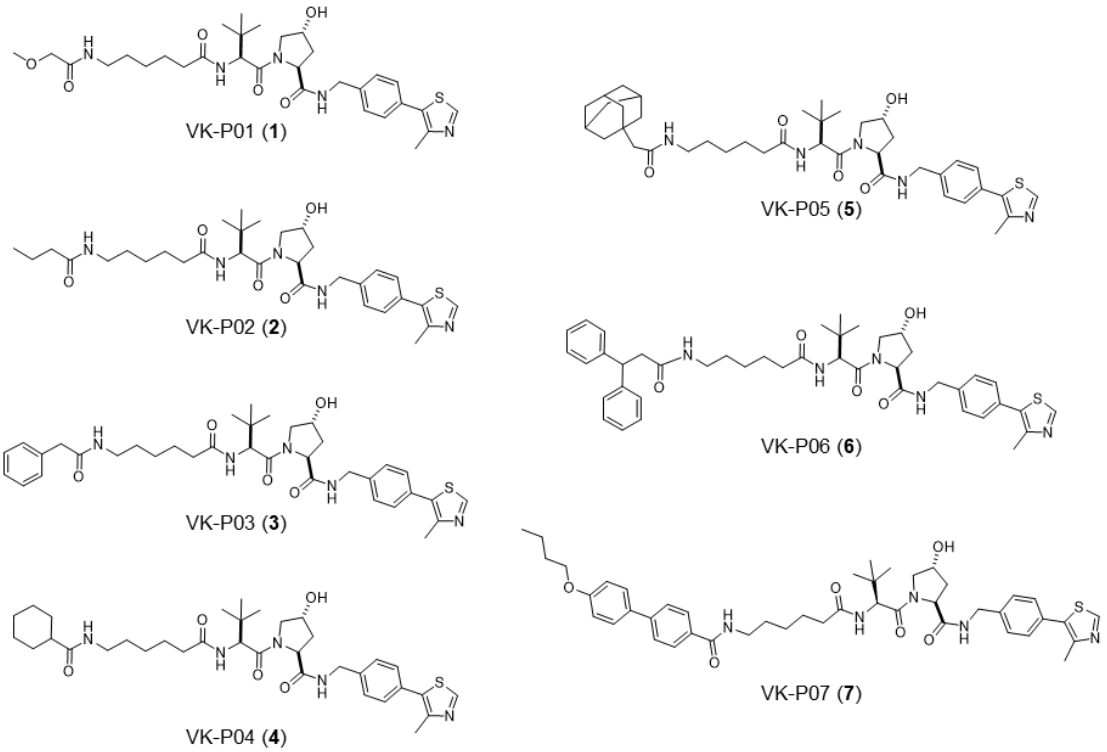
*3.5.2.2 Solid phase synthesis of compound **28***



28

Compound **28** was synthesized according to Klein *et al*. (2020)[66] on SynPhase lanterns using Fmoc deprotection and an addition of Fmoc-Tle-OH with HATU and DIPEA in DMF.

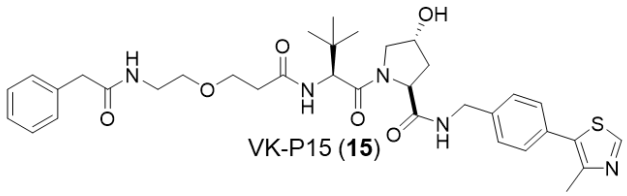
*3.5.2.3 Solid phase synthesis of compounds **1 – 7***



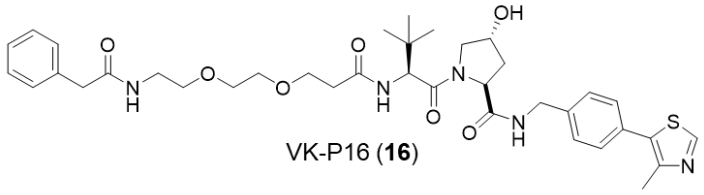
Compounds **1 – 7** were synthesized using solid phase synthesis on SynPhase polystyrene lanterns. Procedure is described for synthesis on one lantern each for **1 – 7**. A lantern with **28** Fmoc-deprotected using 2 mL of a solution of 2% piperidine and 2% DBU in DMF for 15 minutes at room temperature. The deprotection solution was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Next, a solution of Fmoc-6-aminohexanoic acid (31 mg, 0.088 mmol, 4 eq), HATU (34 mg, 0.088 mmol, 4 eq), and DIPEA (613 µL, 0.176 mmol, 8 eq) in 3 mL of DMF were added to the lantern. Reaction was mixed on a linear shaker at room temperature for 4 – 16 hours. The reaction mixture was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash).

The lantern was Fmoc-deprotected using 2 mL of a solution of 2% piperidine and 2% DBU in DMF for 15 minutes at room temperature. The deprotection solution was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Then, a solution of the capping agent (listed below) (0.088 mmol, 4 eq), HATU (34 mg, 0.088 mmol, 4 eq), and DIPEA (613 µL, 0.176 mmol, 8 eq) in 3 mL of DMF was added to the lantern. The reaction was mixed on a linear shaker for 4 – 16 hours at room temperature. Capping agents: **1** (methoxyacetic acid, 6.8 µL); **2** (butyric acid, 8.0 µL); **3** (phenylacetic acid, 12 mg); **4** (cyclohexanecarboxylic acid, 11.3 mg); **5** (2-(adamantan-1-yl)acetic acid, 17.1 mg); **6** (3,3-diphenylpropionic acid, 19.9 mg); **7** (4-butoxy-4'-biphenylcarboxylic acid, 23.8 mg). The reaction mixture was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Compounds **1 – 7** were cleaved from the lantern with 5% HF/pyridine in THF and quenched following previously published procedures with methoxytrimethylsilane.[109] The quenched cleavage solution was evaporated under reduced pressure, and the compounds were purified on a Biotage Isolera Prime flash chromatography system with a 30 g C18 column eluting with 10 – 100% ACN in H_2O, both with 0.1% TFA. Sample identify was confirmed by LC/MS.

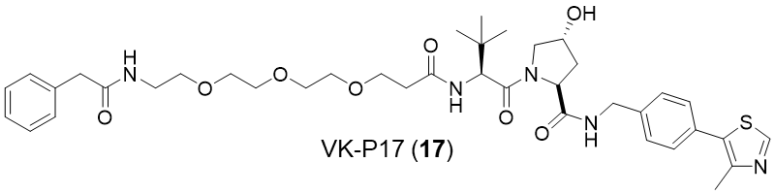
*3.5.2.4 Solid phase synthesis of compounds **15 – 17***



VK-P15 (**15**)



VK-P16 (**16**)

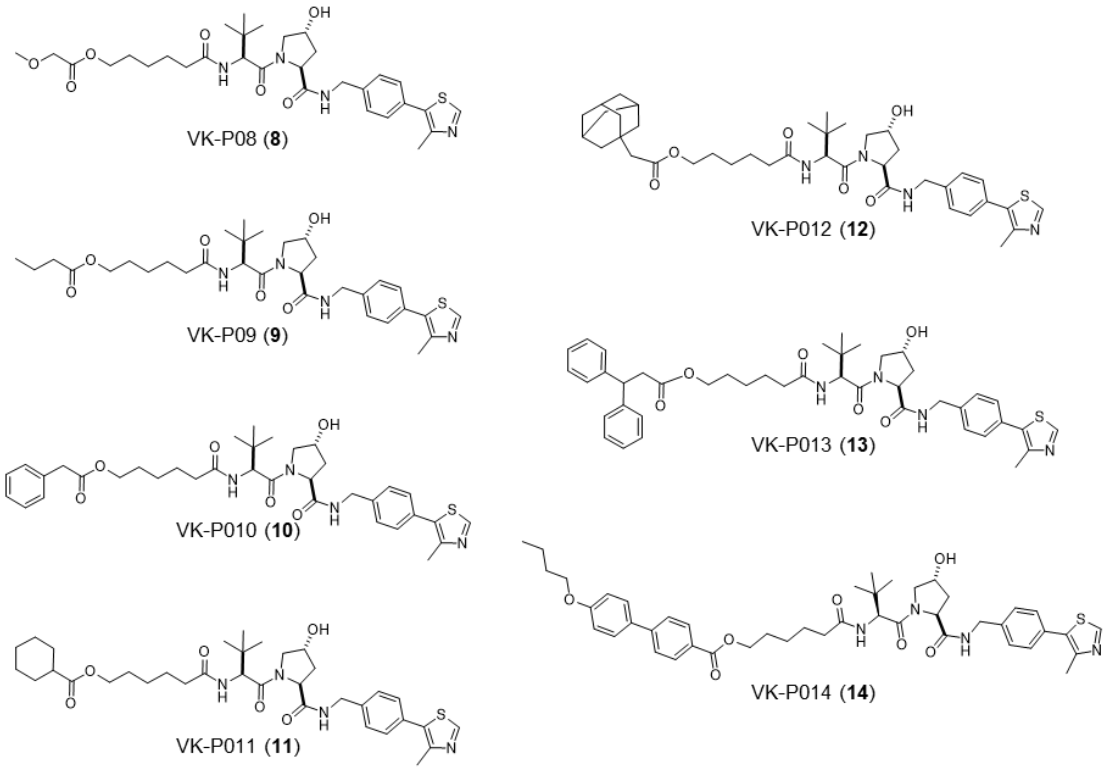


VK-P17 (**17**)

Compounds **15 – 17** were synthesized using solid phase synthesis on SynPhase polystyrene lanterns. Procedure is described for synthesis on one lantern. A lantern with **28** Fmoc-deprotected using 2 mL of a solution of 2% piperidine and 2% DBU in DMF for 15 minutes at room temperature. The deprotection solution was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Next, a solution of a linker (0.088 mmol, 4 eq), HATU (34 mg, 0.088 mmol, 4 eq), and DIPEA (613 µL, 0.176 mmol, 8 eq) in 3 mL of DMF were added to the lantern. Linker was as follows: **15** (Fmoc-6-amino-4-oxahexanoic acid, 31 mg); **16** (Fmoc-9-amino-4,7-dioxanonanoic acid, 35 mg); **17** (Fmoc-12-amino-4,7,10-trioxadodecanoic acid, 39 mg). The reaction was mixed on a linear shaker at room temperature for 4 – 16 hours. The reaction mixture was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). The lantern was Fmoc-

deprotected using 2 mL of a solution of 2% piperidine and 2% DBU in DMF for 15 minutes at room temperature. The deprotection solution was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Then, a solution of phenylacetic acid (12 mg, 0.088 mmol, 4 eq), HATU (34 mg, 0.088 mmol, 4 eq), and DIPEA (613 µL, 0.176 mmol, 8 eq) in 3 mL of DMF was added to the lantern. The reaction was mixed on a linear shaker for 4 – 16 hours at room temperature. The reaction mixture was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Compounds **15 – 17** were cleaved from the lantern with 5% HF/pyridine in THF and quenched following previously published procedures with methoxytrimethylsilane.[109] The quenched cleavage solution was evaporated under reduced pressure, and the compounds were purified on a Biotage Isolera Prime flash chromatography system with a 30 g C18 column eluting with 10 – 100% ACN in H_2O, both with 0.1% TFA. Sample identify was confirmed by LC/MS.

*3.5.2.5 Solid phase synthesis of compounds **8 – 14***

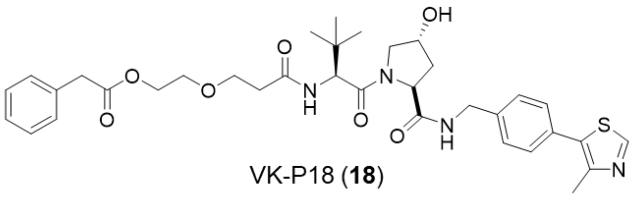


(Check linker addition steps in Yi's notebook)

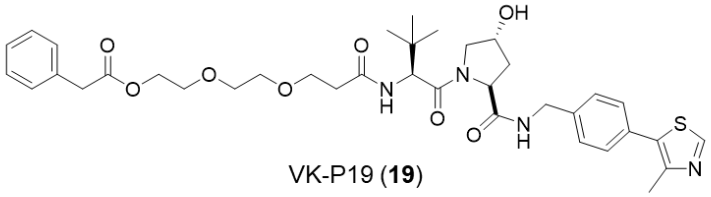
Compounds **8 – 14** were synthesized using solid phase synthesis on SynPhase polystyrene lanterns. Procedure is described for synthesis on one lantern. A lantern with **28** Fmoc-deprotected using 2 mL of a solution of 2% piperidine and 2% DBU in DMF for 15 minutes at room temperature. The deprotection solution was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Next, a solution of 6-hydroxy-hexanoic acid (11.6 mg, 0.088 mmol, 4 eq), HATU (34 mg, 0.088 mmol, 4 eq), and DIPEA (613 µL, 0.176 mmol, 8 eq) in 3 mL of DMF were added to the lantern. Reaction was mixed on a linear shaker at room temperature for 4 – 16 hours. The reaction mixture was drained, and the lantern was

rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Then, a solution of the capping agent (listed below) (0.22 mmol, 10 eq), DIC (34 µL, 0.22 mmol, 10 eq), and DMAP (0.7 mg, 0.0055 mmol, 0.25 eq) in 2 mL of dry DCM was added to the lantern. The reaction was mixed on a linear shaker for 4 – 16 hours at room temperature. Capping agents: **8** (methoxyacetic acid, 6.8 µL); **9** (butyric acid, 8.0 µL); **10** (phenylacetic acid, 12 mg); **11** (cyclohexanecarboxylic acid, 11.3 mg); **12** (2-(adamantan-1-yl)acetic acid, 17.1 mg); **13** (3,3-diphenylpropionic acid, 19.9 mg); **14** (4-butoxy-4'-biphenylcarboxylic acid, 23.8 mg). The reaction mixture was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Compounds **1 – 7** were cleaved from the lantern with 5% HF/pyridine in THF and quenched following previously published procedures with methoxytrimethylsilane.[109] The quenched cleavage solution was evaporated under reduced pressure, and the compounds were purified on a Biotage Isolera Prime flash chromatography system with a 30 g C18 column eluting with 10 – 100% ACN in H_2O, both with 0.1% TFA. Sample identify was confirmed by LC/MS.

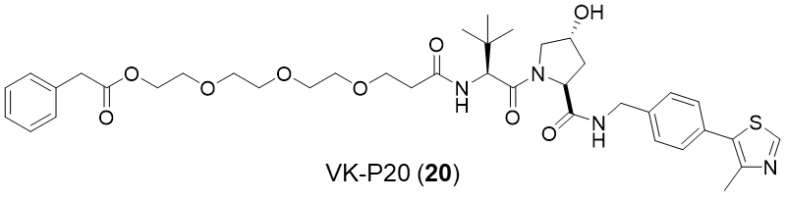
*3.5.2.6 Solid phase synthesis of compounds **18 – 20***



VK-P18 (**18**)



VK-P19 (**19**)



VK-P20 (**20**)

Compounds 1**8 – 20** were synthesized using solid phase synthesis on SynPhase polystyrene lanterns. Procedure is described for synthesis on one lantern. A lantern with **28** Fmoc-deprotected using 2 mL of a solution of 2% piperidine and 2% DBU in DMF for 15 minutes at room temperature. The deprotection solution was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Next, a solution of a linker (0.088 mmol, 4 eq), HATU (34 mg, 0.088 mmol, 4 eq), and DIPEA (613 µL, 0.176 mmol, 8 eq) in 3 mL of DMF were added to the lantern. Linker was as follows: **18** (3-(2-Hydroxyethoxy)propanoic acid, 12 mg); **19** (3-[2-(2-Hydroxyethoxy)ethoxy]propanoic acid, 16 mg); **20** (3-(2-[2-(2-Hydroxyethoxy)ethoxy]ethoxy)propanoic acid, 20 mg). The reaction was mixed on a linear shaker at room temperature for 4 – 16 hours. The reaction mixture was drained,

and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Then, a solution of phenylacetic acid (30 mg, 0.22 mmol, 10 eq), (0.22 mmol, 10 eq), DIC (34 µL, 0.22 mmol, 10 eq), and DMAP (0.7 mg, 0.0055 mmol, 0.25 eq) in 2 mL of dry DCM was added to the lantern. The reaction was mixed on a linear shaker for 4 – 16 hours at room temperature. The reaction mixture was drained, and the lantern was rinsed 3x with 2 mL of DMF and 3x with 2 mL of DCM (30 seconds per wash). Compounds **18 – 20** were cleaved from the lantern with 5% HF/pyridine in THF and quenched following previously published procedures with methoxytrimethylsilane.[109] The quenched cleavage solution was evaporated under reduced pressure, and the compounds were purified on a Biotage Isolera Prime flash chromatography system with a 30 g C18 column eluting with 10 – 100% ACN in H_2O, both with 0.1% TFA. Sample identify was confirmed by LC/MS.

*3.5.2.7 Solid phase synthesis of compounds **21***

3.5.3 Parallel artificial membrane permeability assay (PAMPA)

PAMPA[64,65] was used to determine the passive membrane permeability as described in Naylor *et al.*[7] and Klein *et al.* (2020).[66]

3.5.4 MDCK–MDR1 cell permeability assays

MDCK-MDR1 cell permeability data was collected by the CRO Quintara Discovery Inc. (QDI) in Hayward, CA.

3.5.5 $LogD_{(dec/w)}$ shake flask partition coefficient assay

The shake flask partition coefficient of each compound was determined following the procedure described in Klein *et al.* (2020).[66]

3.5.6 Lipophilic permeability efficiency (LPE) metric calculations

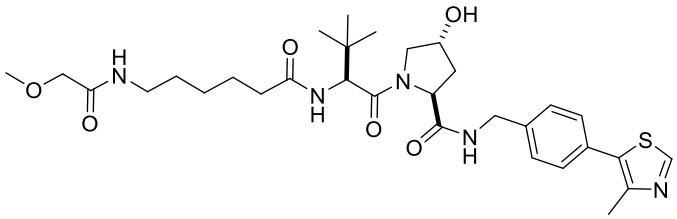
LPE was calculated using the protocol described in Naylor *et al.*[7] using the following equation: $LPE = LogD_{(dec/w)} - 1.06(ALogP) + 5.47$.

3.5.7 Plasma Stability Assay

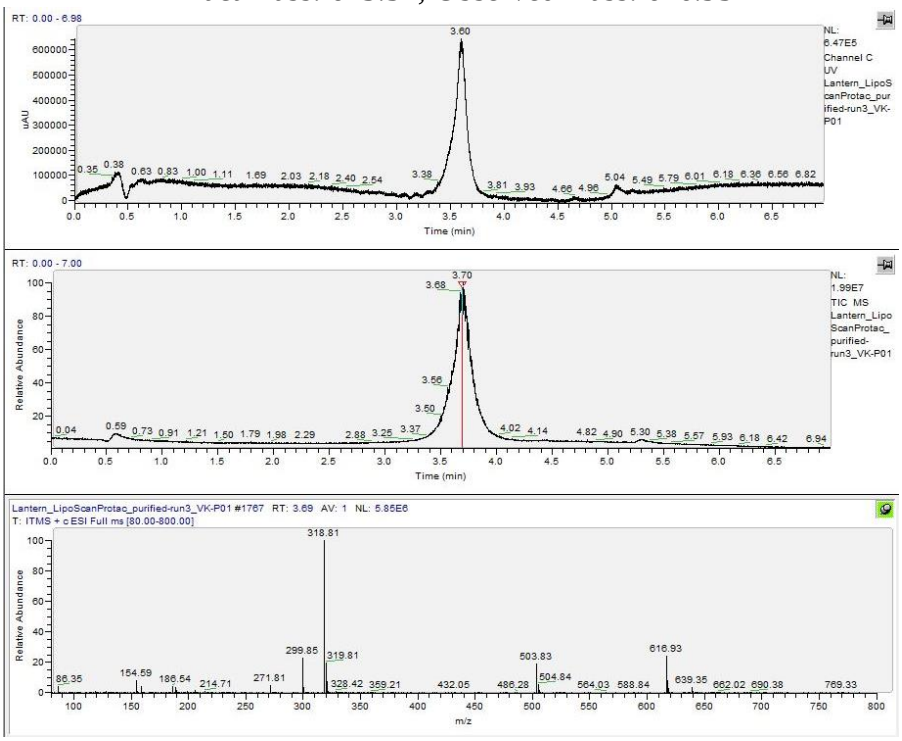
In a 1.2 mL cluster tube, test compounds (1 µL of a 1 mM DMSO stock, final DMSO concentration = 1%) were incubated in prewarmed human pooled plasma diluted to 80% with PBS pH 7.4 (100 µL) for three time points (15, 30, and 90 min) and prewarmed PBS pH 7.4 (100 µL) for a 0 min time point. At the specified time point, the reaction was terminated by protein precipitation with 400 µL of ACN. The cluster tubes were separated and vortexed for 1 min. Then, the tubes were vortexed for 15 min at 4 ºC and 12,700 rpm. The supernatant was removed and analyzed by LC/MS. Peaks were integrated using an in-house python program. The percent of compound remaining at each time point compared to the amount of compound in the 0 min time point is reported. Assay was run in triplicate. Positive control (known to degrade in plasma) and negative control (known to be stable in plasma) compounds were run alongside the test compounds as well.

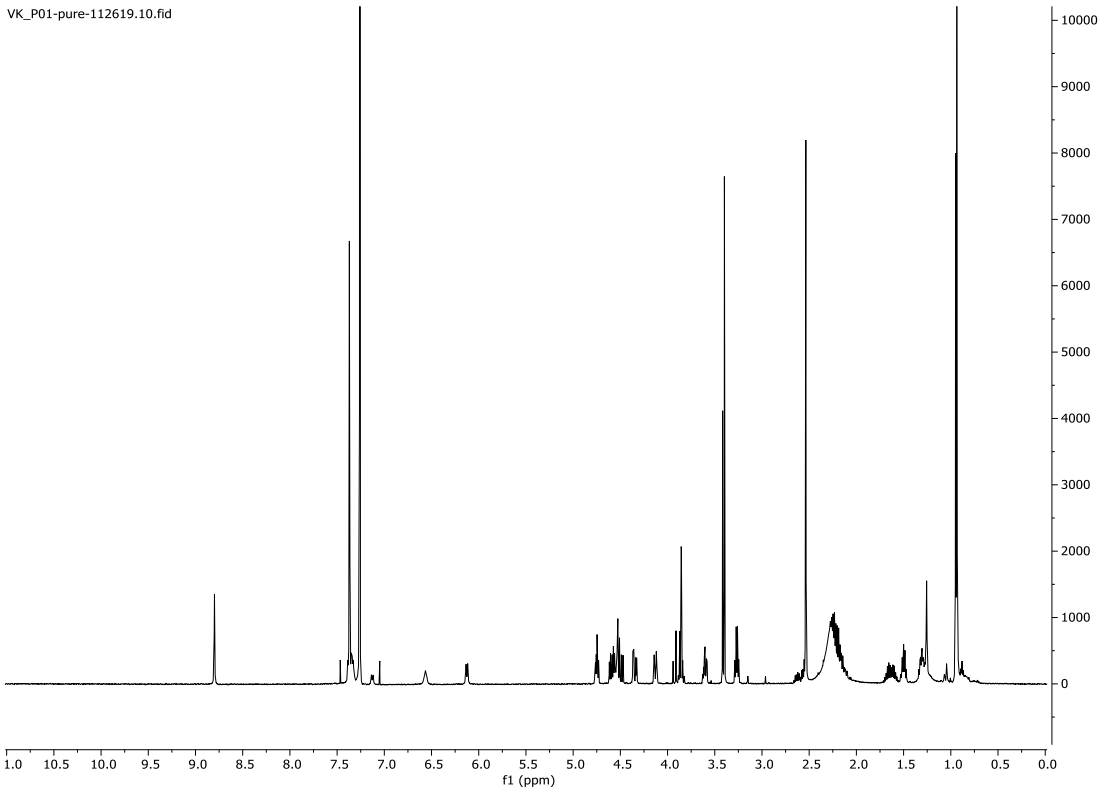
3.5.8 Spectra for synthesized compounds

*3.5.8.1 NMR and LC/MS spectra for VK-P01 (**1**)*



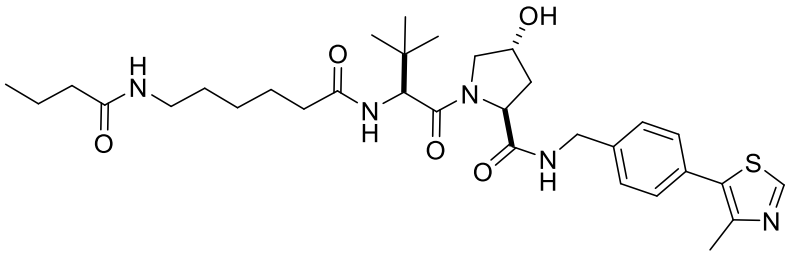
Exact mass: 615.31; Observed mass: 616.93



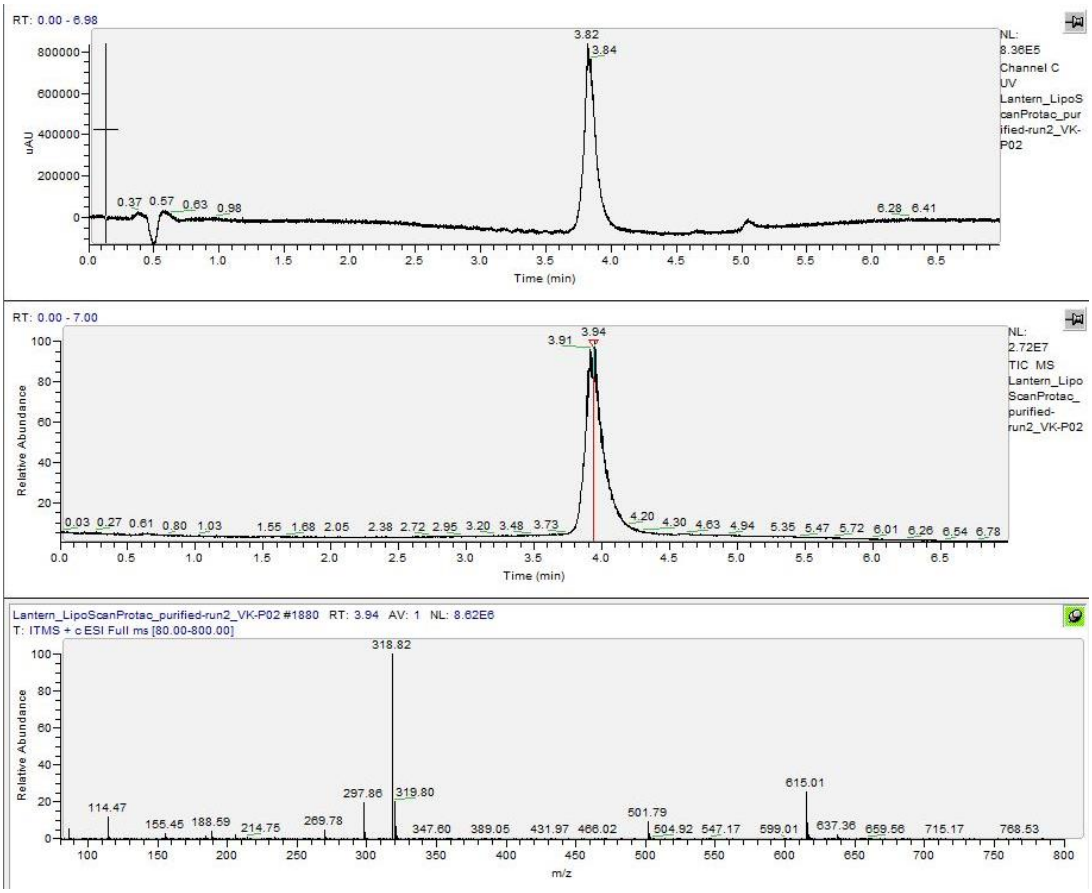


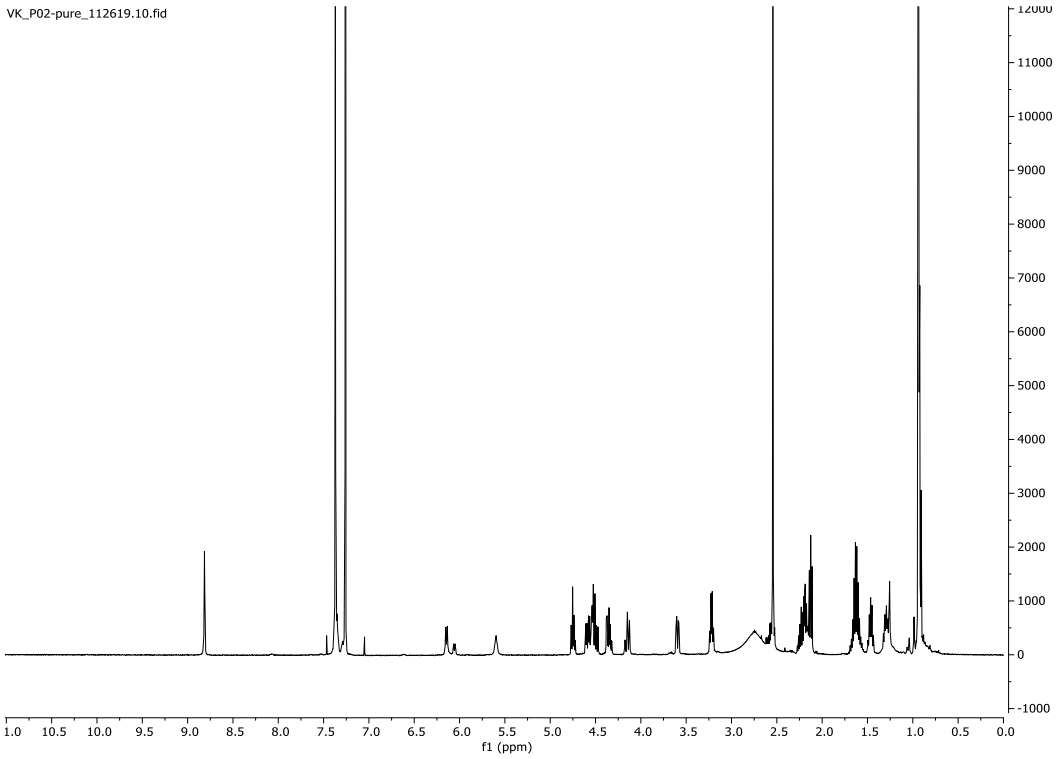
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.82 (s, 1H), 7.43 – 7.34 (m, 4H), 6.59 (s, 1H), 6.15 (d, *J* = 8.7 Hz, 1H), 4.78 (td, *J* = 8.0, 4.3 Hz, 1H), 4.61 (dt, *J* = 14.9, 6.3 Hz, 1H), 4.52 (dd, *J* = 18.6, 8.7 Hz, 1H), 4.41 – 4.33 (m, 1H), 4.15 (d, *J* = 11.2 Hz, 1H), 3.98 – 3.83 (m, 2H), 3.63 (ddd, *J* = 11.2, 7.3, 3.6 Hz, 1H), 3.43 (d, *J* = 9.1 Hz, 3H), 3.29 (q, *J* = 6.8 Hz, 1H), 2.69 – 2.56 (m, 1H), 2.56 (s, 3H), 2.23 (dtt, *J* = 22.0, 14.4, 7.5 Hz, 2H), 1.66 (ddq, *J* = 28.3, 13.9, 7.1 Hz, 1H), 1.52 (p, *J* = 7.6, 7.2 Hz, 1H), 1.34 (s, 1H), 1.38 – 1.28 (m, 1H), 1.28 (s, 1H), 0.97 (d, *J* = 6.5 Hz, 8H).

*3.5.8.2 NMR and LC/MS spectra for VK-P02 (**2**)*



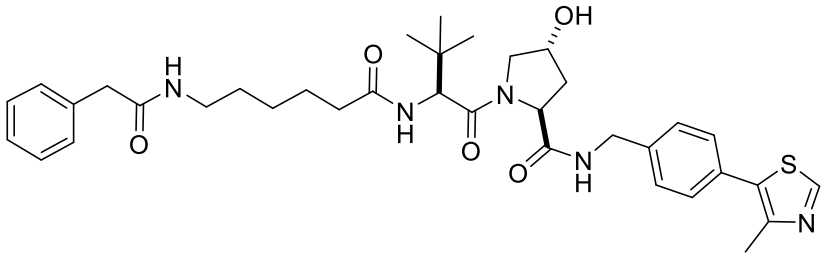
Exact mass: 613.33; Observed mass: 615.01



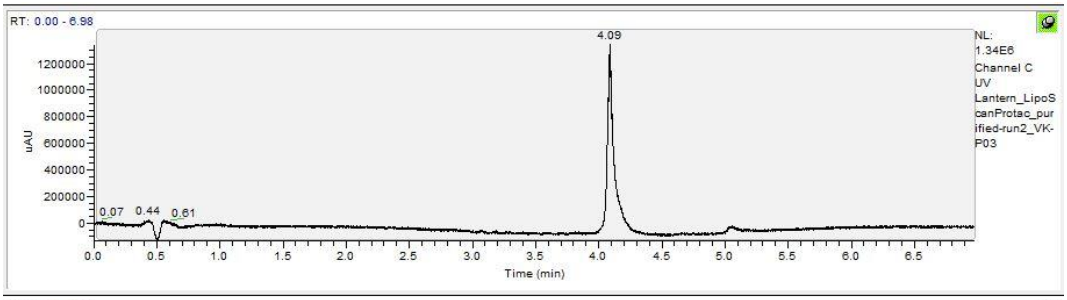


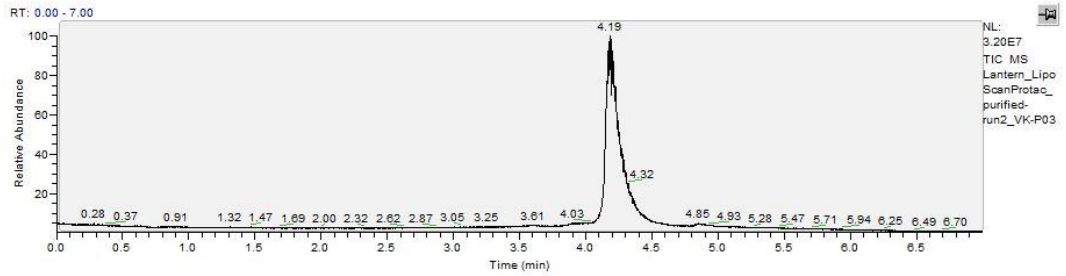
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.81 (s, 1H), 7.37 (s, 3H), 7.36 (d, *J* = 6.6 Hz, 2H), 6.14 (d, *J* = 8.8 Hz, 1H), 5.60 (s, 1H), 4.79 – 4.71 (m, 1H), 4.59 (ddd, *J* = 15.0, 6.7, 2.7 Hz, 1H), 4.54 (s, 1H), 4.50 (dd, *J* = 16.9, 8.7 Hz, 1H), 4.35 (ddd, *J* = 14.8, 9.6, 5.2 Hz, 1H), 4.15 (t, *J* = 12.1 Hz, 1H), 3.63 – 3.56 (m, 1H), 3.22 (dt, *J* = 7.9, 6.3 Hz, 1H), 2.75 (s, 7H), 2.61 – 2.51 (m, 1H), 2.54 (s, 3H), 2.29 – 2.09 (m, 4H), 2.16 (s, 1H), 1.72 – 1.54 (m, 3H), 1.47 (p, *J* = 7.1 Hz, 2H), 1.34 – 1.24 (m, 2H), 0.96 – 0.89 (m, 12H).

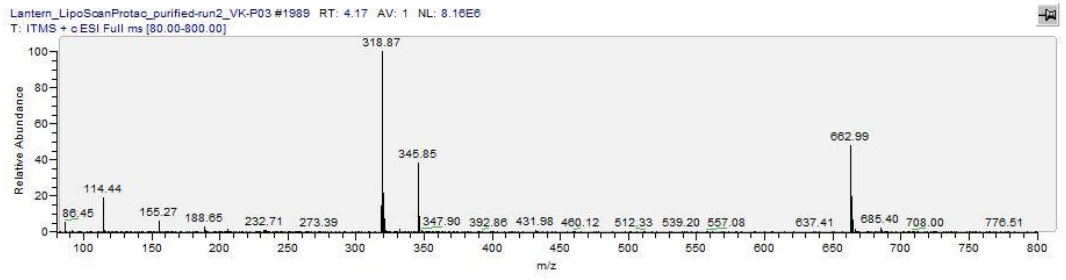
*3.5.8.3 NMR and LC/MS spectra for VK-P03 (**3**)*

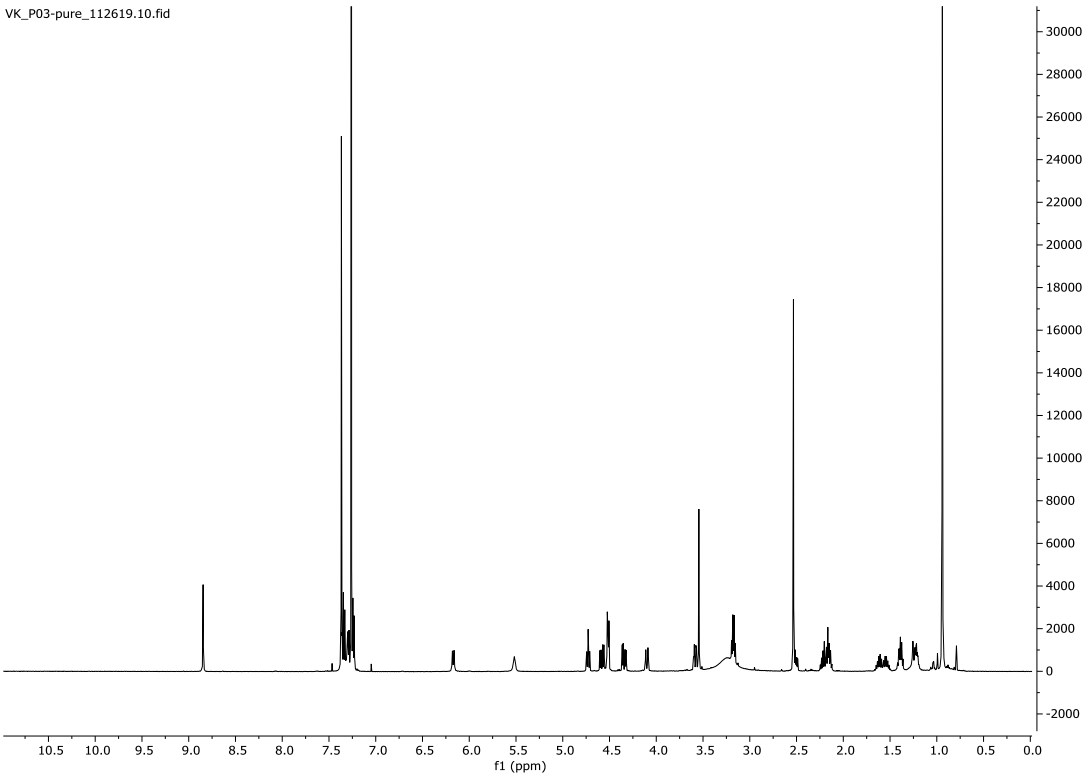


Exact mass: 661.33; Observed mass: 662.99



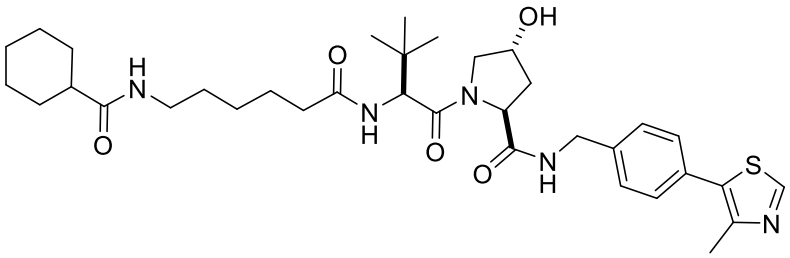




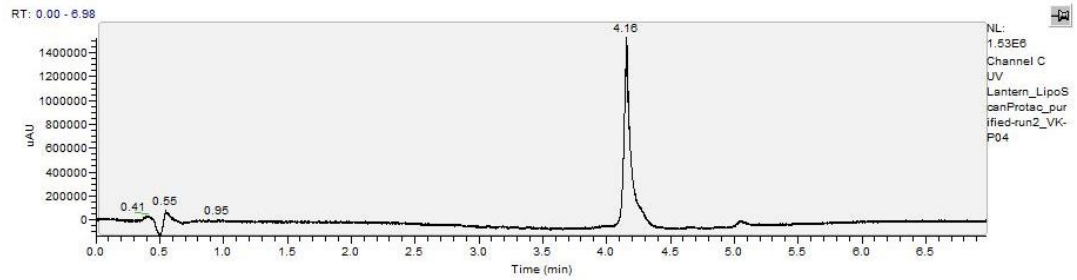


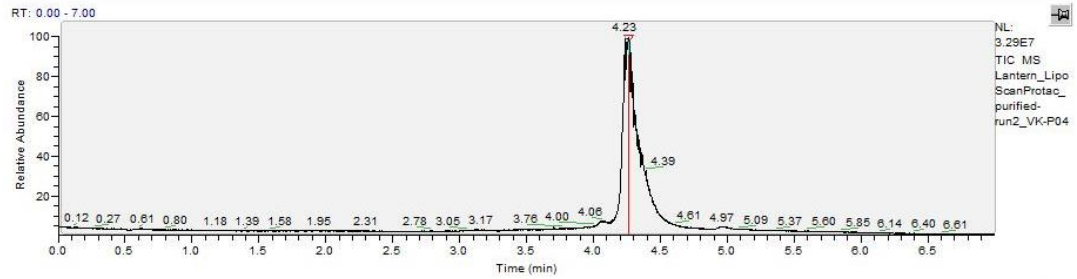
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.85 (s, 1H), 7.39 – 7.27 (m, 8H), 7.27 – 7.21 (m, 2H), 6.17 (d, *J* = 8.7 Hz, 1H), 5.52 (s, 1H), 4.73 (t, *J* = 8.1 Hz, 1H), 4.58 (dd, *J* = 15.0, 6.6 Hz, 1H), 4.53 – 4.49 (m, 2H), 4.34 (dd, *J* = 15.0, 5.3 Hz, 1H), 4.13 – 4.07 (m, 1H), 3.58 (dd, *J* = 11.4, 3.3 Hz, 1H), 3.54 (s, 2H), 3.24 (s, 8H), 3.17 (q, *J* = 6.8 Hz, 2H), 2.53 (s, 3H), 2.56 – 2.47 (m, 1H), 2.22 (dt, *J* = 14.3, 7.1 Hz, 1H), 2.17 (d, *J* = 7.5 Hz, 1H), 2.17 – 2.10 (m, 1H), 1.68 – 1.48 (m, 2H), 1.44 – 1.34 (m, 2H), 1.27 – 1.17 (m, 3H), 0.94 (s, 9H).

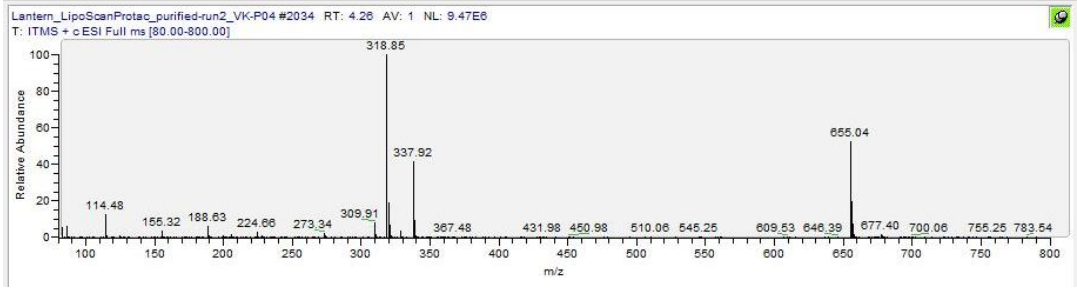
*3.5.8.4 NMR and LC/MS spectra for VK-P04 (**4**)*

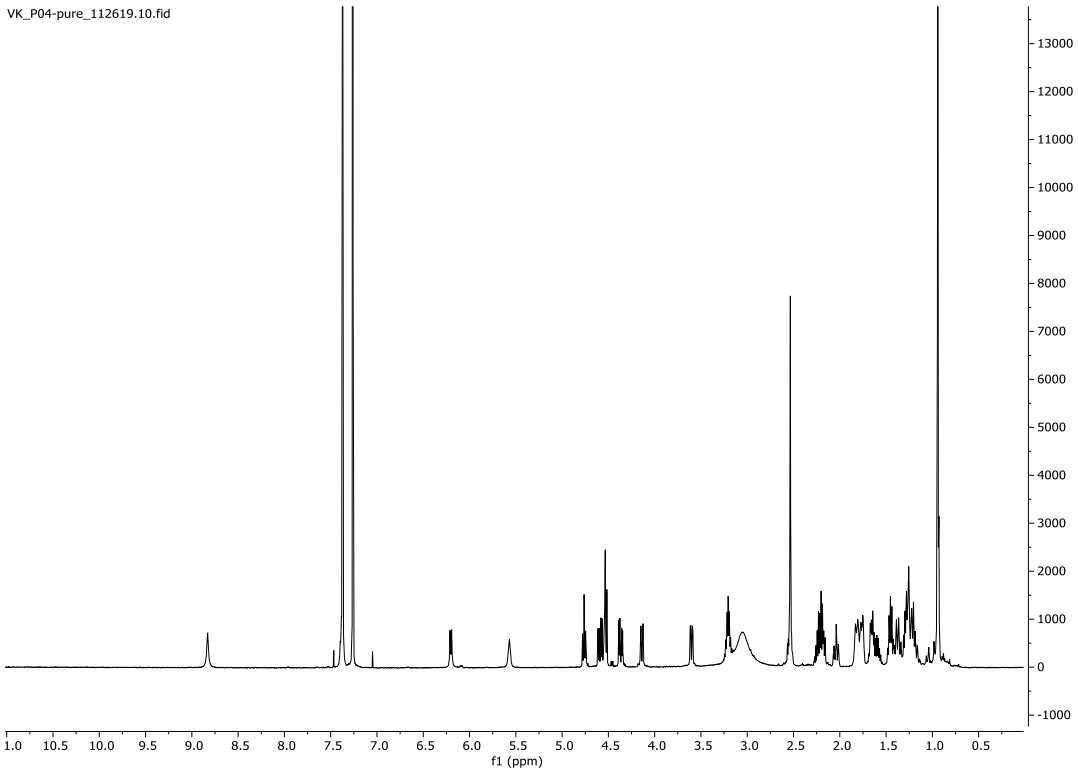


Exact mass: 653.36; Observed mass: 665.04



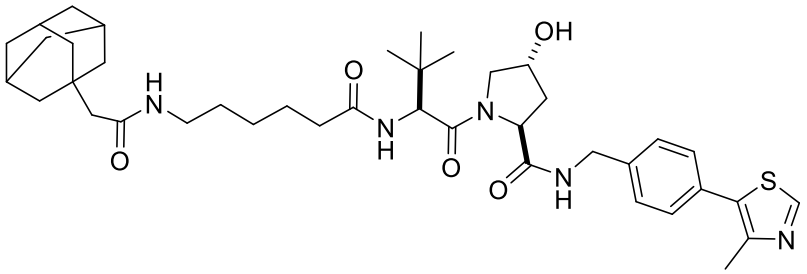




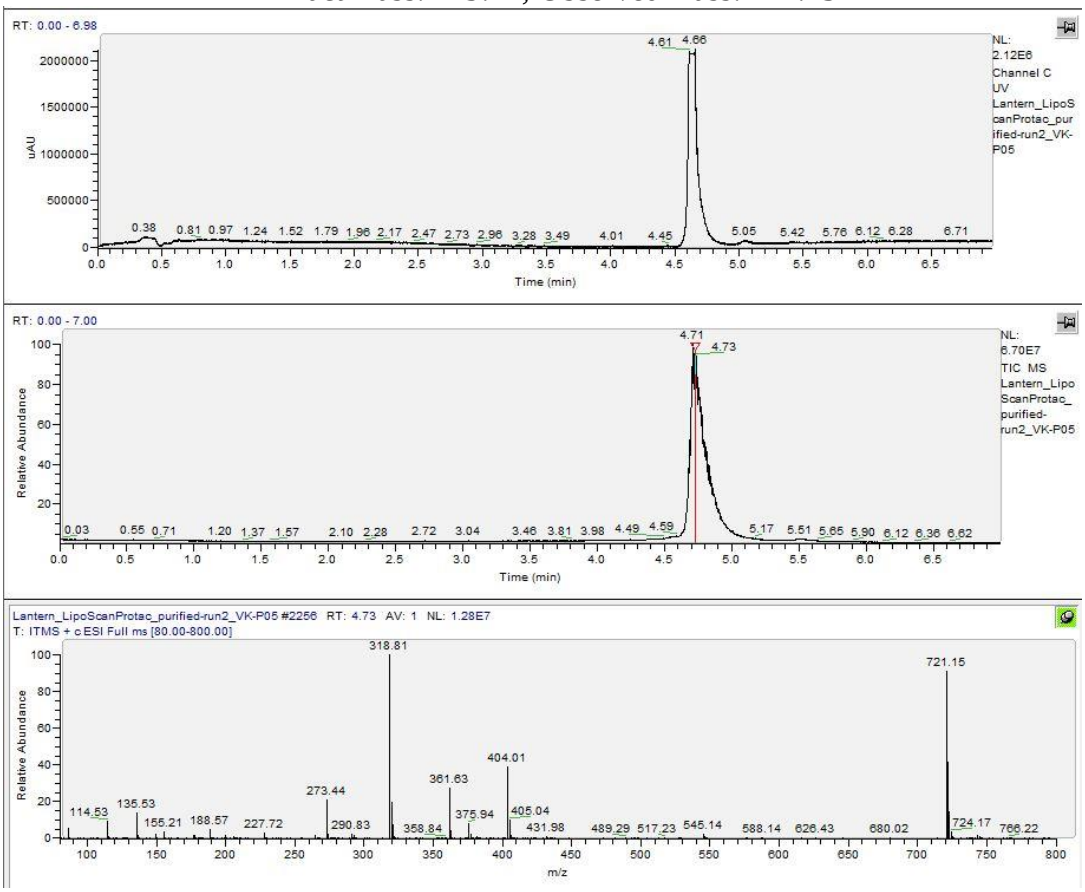


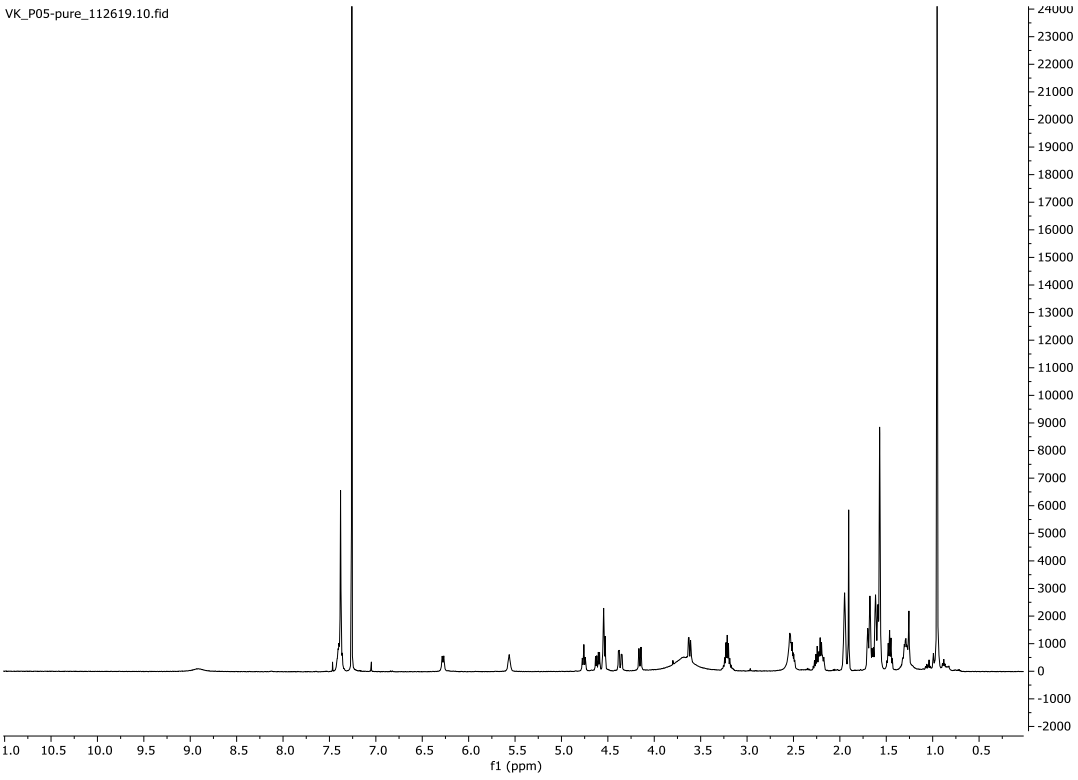
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.83 (s, 1H), 7.39 (s, 1H), 7.37 (s, 4H), 6.20 (d, *J* = 8.7 Hz, 1H), 5.57 (s, 1H), 4.76 (t, *J* = 8.1 Hz, 1H), 4.59 (dd, *J* = 15.0, 6.6 Hz, 1H), 4.53 (t, *J* = 7.0 Hz, 2H), 4.37 (dd, *J* = 15.0, 5.3 Hz, 1H), 4.14 (d, *J* = 11.5 Hz, 1H), 3.60 (dd, *J* = 11.4, 3.4 Hz, 1H), 3.28 – 3.14 (m, *J* = 6.5 Hz, 2H), 3.05 (s, 8H), 2.58 – 2.49 (m, 1H), 2.53 (s, 3H), 2.29 – 2.14 (m, 3H), 2.04 (tt, *J* = 11.8, 3.4 Hz, 1H), 1.85 – 1.73 (m, 5H), 1.66 (dd, *J* = 13.7, 7.6 Hz, 2H), 1.59 (td, *J* = 14.4, 7.2 Hz, 1H), 1.49 – 1.42 (m, 2H), 1.44 – 1.29 (m, 2H), 1.31 – 1.21 (m, 4H), 1.24 – 1.14 (m, 2H), 0.94 (s, 9H).

*3.5.8.5 NMR and LC/MS spectra for VK-P05 (**5**)*



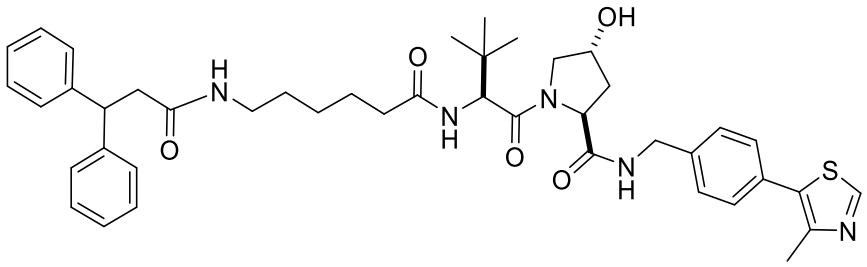
Exact mass: 719.41; Observed mass: 721.15



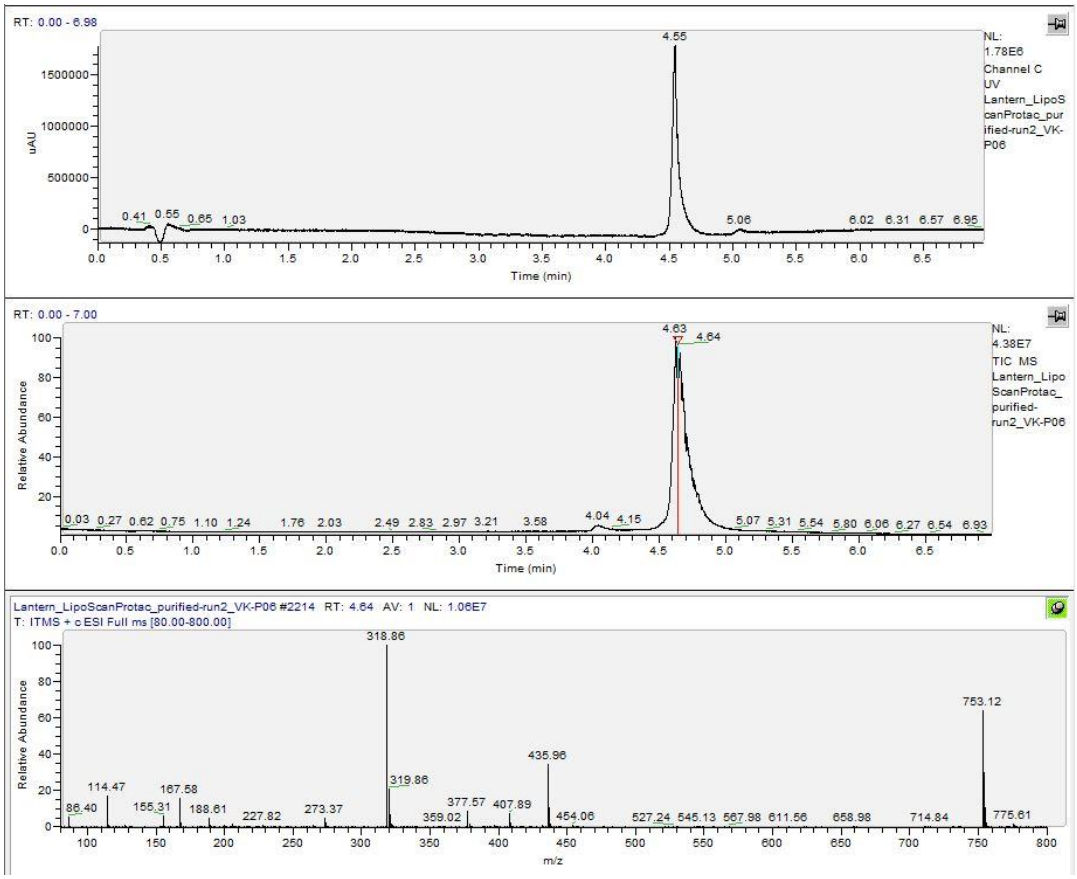


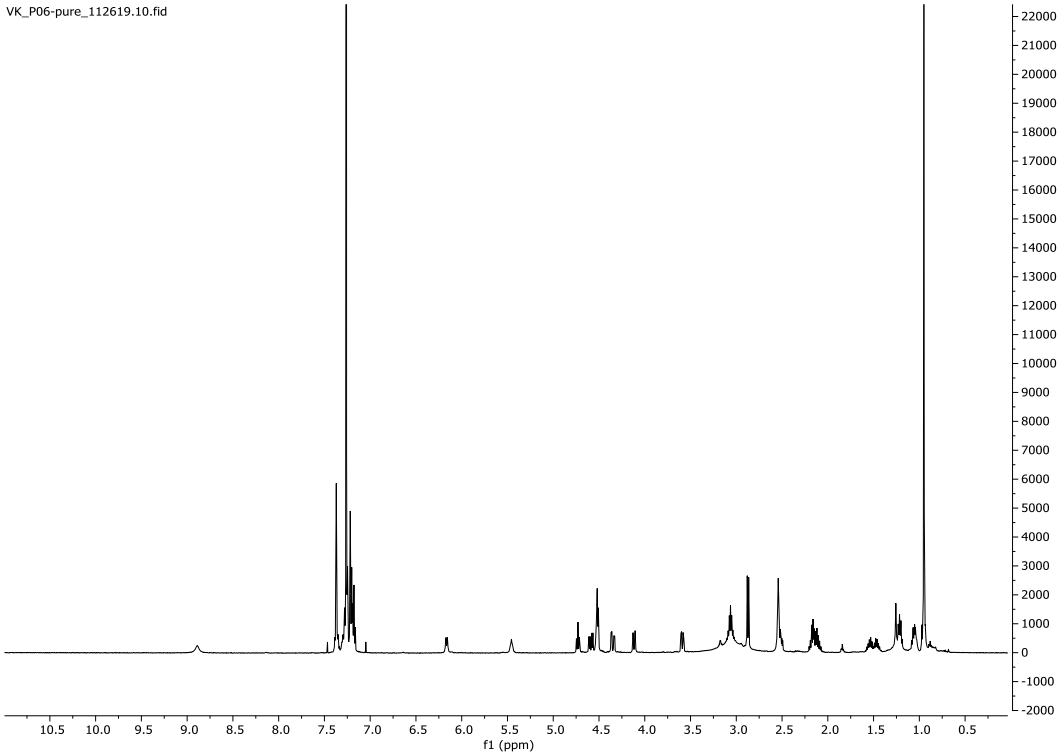
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.92 (s, 1H), 7.43 – 7.34 (m, 5H), 6.28 (d, *J* = 8.7 Hz, 1H), 5.56 (s, 1H), 4.76 (t, *J* = 8.1 Hz, 1H), 4.61 (dd, *J* = 15.1, 6.6 Hz, 1H), 4.54 (d, *J* = 8.9 Hz, 2H), 4.37 (dd, *J* = 15.1, 5.2 Hz, 1H), 4.15 (d, *J* = 11.4 Hz, 1H), 3.76 (s, 3H), 3.68 (s, 6H), 3.62 (dd, *J* = 11.3, 3.4 Hz, 1H), 3.29 – 3.14 (m, *J* = 7.0 Hz, 2H), 2.54 (s, 3H), 2.57 – 2.47 (m, 1H), 2.31 – 2.21 (m, 1H), 2.21 (s, 1H), 2.21 – 2.15 (m, 1H), 1.95 (s, 3H), 1.91 (s, 2H), 1.70 (s, 1H), 1.66 (d, *J* = 11.7 Hz, 3H), 1.62 (dd, *J* = 13.4, 10.1 Hz, 5H), 1.46 (p, *J* = 7.1 Hz, 2H), 1.30 (s, 2H), 1.33 – 1.24 (m, 2H), 0.95 (s, 8H), 0.85 (s, 1H).

*3.5.8.6 NMR and LC/MS spectra for VK-P06 (**6**)*



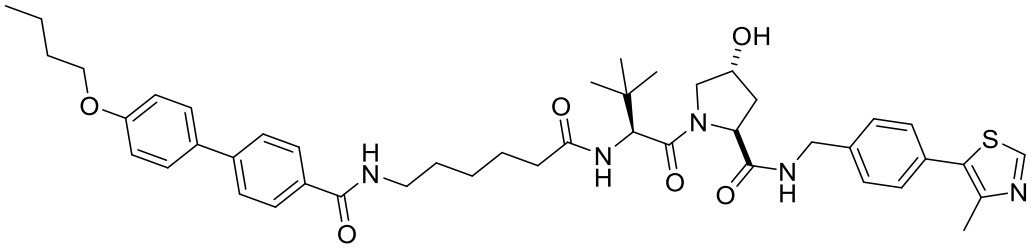
Exact mass: 751.38; Observed mass: 753.18



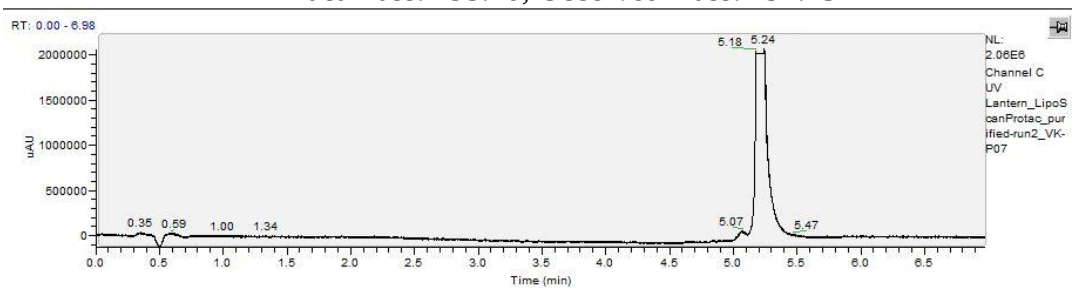


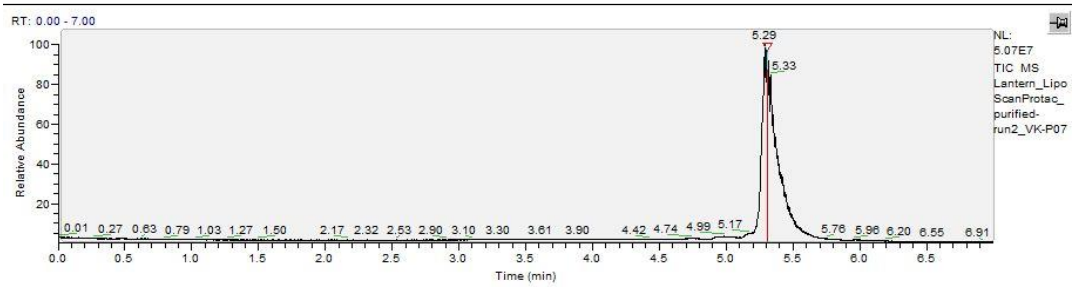
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.89 (s, 1H), 7.37 (d, *J* = 1.9 Hz, 4H), 7.29 (d, *J* = 8.5 Hz, 1H), 7.30 – 7.20 (m, 5H), 7.22 – 7.14 (m, 4H), 6.16 (d, *J* = 8.7 Hz, 1H), 5.46 (s, 1H), 4.73 (t, *J* = 8.1 Hz, 1H), 4.59 (dd, *J* = 15.1, 6.7 Hz, 1H), 4.55 – 4.49 (m, 3H), 4.35 (dd, *J* = 15.1, 5.2 Hz, 1H), 4.12 (d, *J* = 11.4 Hz, 1H), 3.59 (dd, *J* = 11.4, 3.4 Hz, 1H), 3.06 (hept, *J* = 6.7 Hz, 2H), 2.87 (d, *J* = 7.9 Hz, 2H), 2.54 (s, 4H), 2.56 – 2.48 (m, 1H), 2.13 (ddt, *J* = 28.9, 14.6, 7.4 Hz, 3H), 1.50 (ddt, *J* = 41.5, 13.9, 7.2 Hz, 2H), 1.27 – 1.17 (m, 4H), 1.10 – 1.01 (m, 3H), 0.95 (s, 9H).

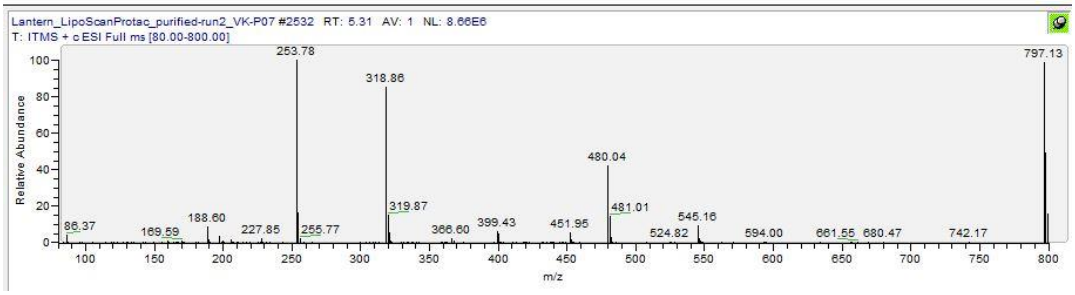
*3.5.8.7 NMR and LC/MS spectra for VK-P07 (**7**)*

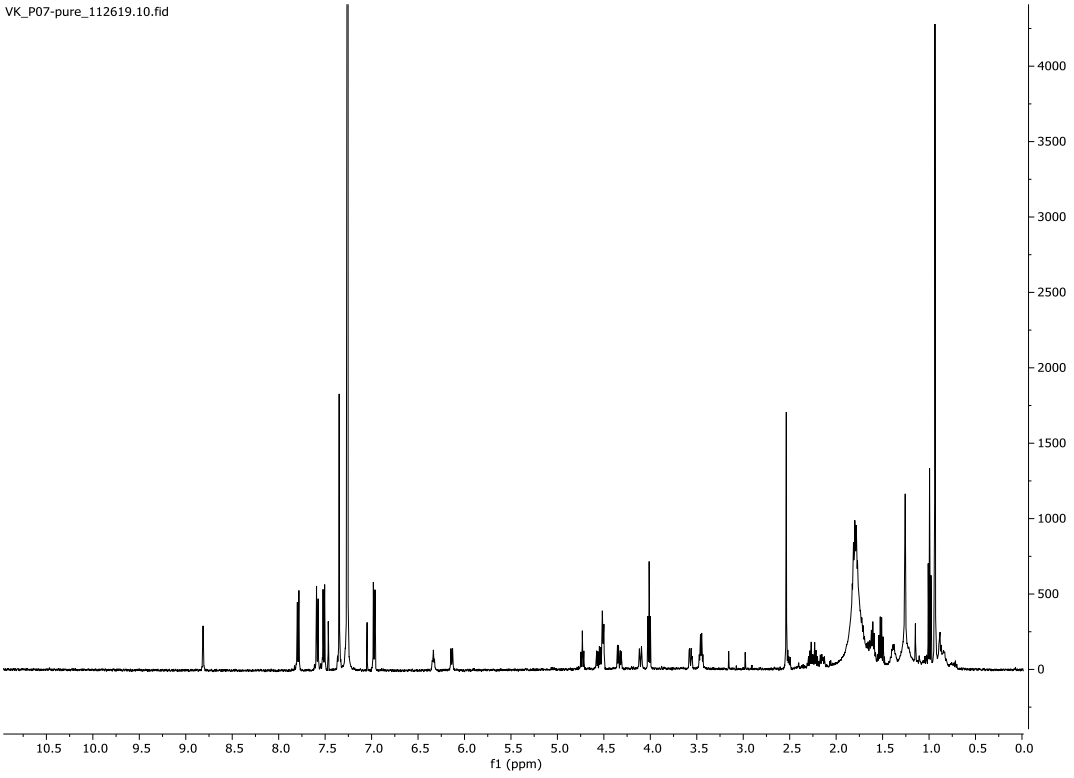


Exact mass: 795.40; Observed mass: 797.13



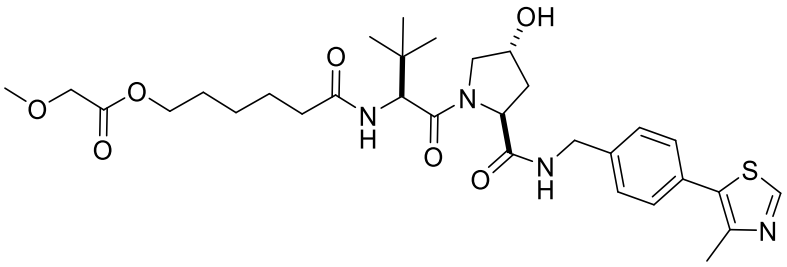




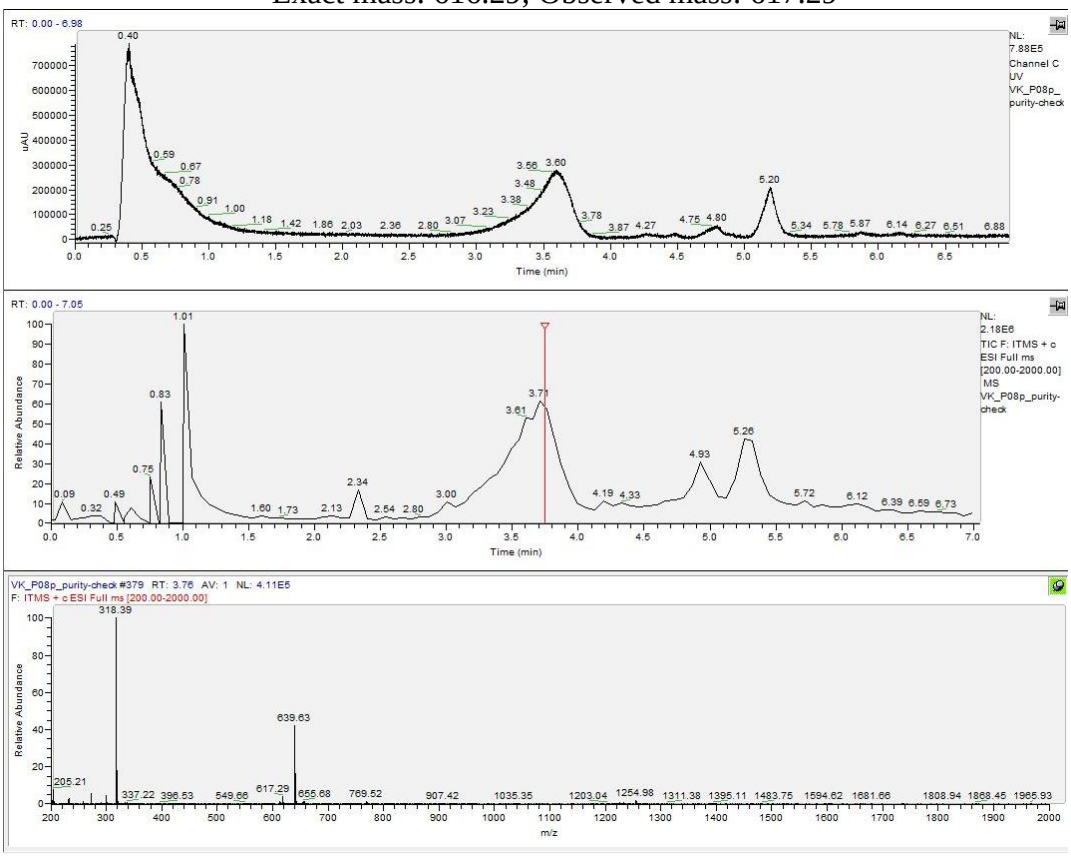


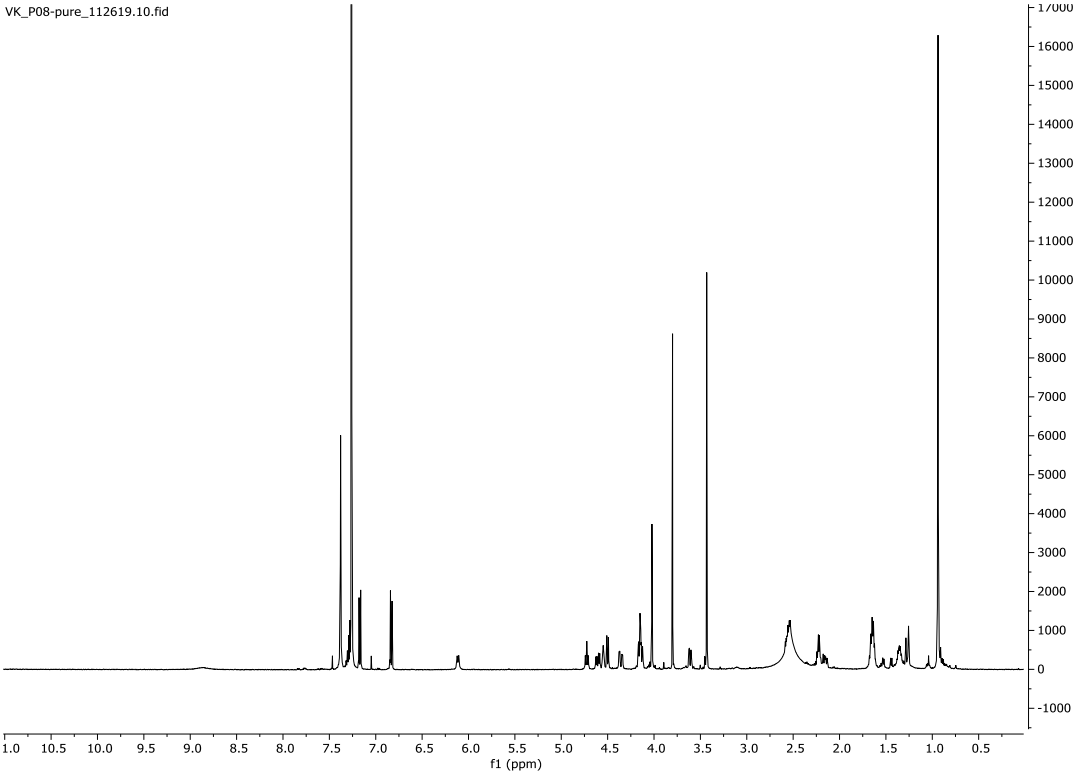
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.81 (s, 1H), 7.79 (d, *J* = 8.2 Hz, 2H), 7.63 – 7.49 (m, 5H), 7.35 (s, 4H), 6.97 (d, *J* = 8.7 Hz, 2H), 6.33 (d, *J* = 6.0 Hz, 1H), 6.14 (d, *J* = 8.7 Hz, 1H), 4.73 (t, *J* = 8.1 Hz, 1H), 4.56 (dd, *J* = 15.0, 6.6 Hz, 1H), 4.51 (d, *J* = 8.6 Hz, 2H), 4.33 (dd, *J* = 15.1, 5.2 Hz, 1H), 4.11 (d, *J* = 11.4 Hz, 1H), 4.01 (t, *J* = 6.5 Hz, 3H), 3.57 (dd, *J* = 11.4, 3.4 Hz, 1H), 3.45 (q, *J* = 6.6 Hz, 2H), 2.54 (s, 3H), 2.51 (dd, *J* = 8.4, 4.7 Hz, 1H), 2.32 – 2.11 (m, 3H), 1.84 – 1.75 (m, 1H), 1.74 – 1.46 (m, 6H), 1.39 (s, 2H), 1.26 (s, 6H), 1.15 (s, 1H), 0.99 (t, *J* = 7.4 Hz, 4H), 0.94 (s, 9H), 0.88 (t, *J* = 6.2 Hz, 2H), 0.84 (s, 3H), 0.76 (s, 1H).

*3.5.8.8 NMR and LC/MS spectra for VK-P08 (**8**)*



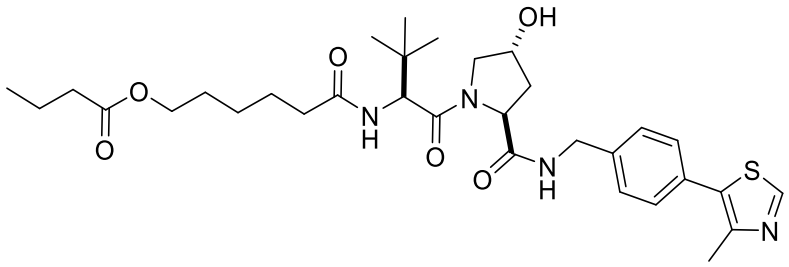
Exact mass: 616.29; Observed mass: 617.29





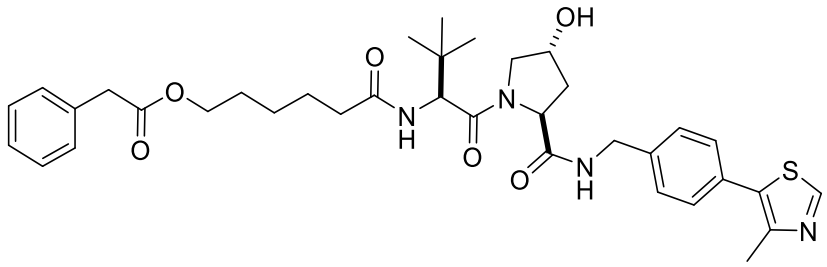
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.87 (s, 1H), 7.38 (s, 3H), 7.34 – 7.26 (m, 1H), 7.20 – 7.14 (m, 1H), 6.86 – 6.81 (m, 1H), 6.11 (d, *J* = 8.7 Hz, 1H), 4.72 (t, *J* = 8.0 Hz, 1H), 4.61 (dd, *J* = 15.0, 6.6 Hz, 1H), 4.55 (s, 1H), 4.50 (d, *J* = 8.7 Hz, 1H), 4.36 (dd, *J* = 15.0, 5.2 Hz, 1H), 4.21 – 4.11 (m, 2H), 4.02 (d, *J* = 0.9 Hz, 1H), 3.80 (s, 2H), 3.61 (dd, *J* = 11.4, 3.5 Hz, 1H), 3.43 (s, 2H), 2.56 (ddd, *J* = 12.7, 7.9, 4.5 Hz, 1H), 2.23 (dd, *J* = 7.4, 4.4 Hz, 1H), 2.23 – 2.12 (m, 1H), 1.70 – 1.59 (m, 1H), 1.47 – 1.30 (m, 1H), 1.35 (s, 1H), 1.27 (d, *J* = 14.3 Hz, 1H), 0.94 (s, 6H).

*3.5.8.9 NMR and LC/MS spectra for VK-P09 (**9**)*

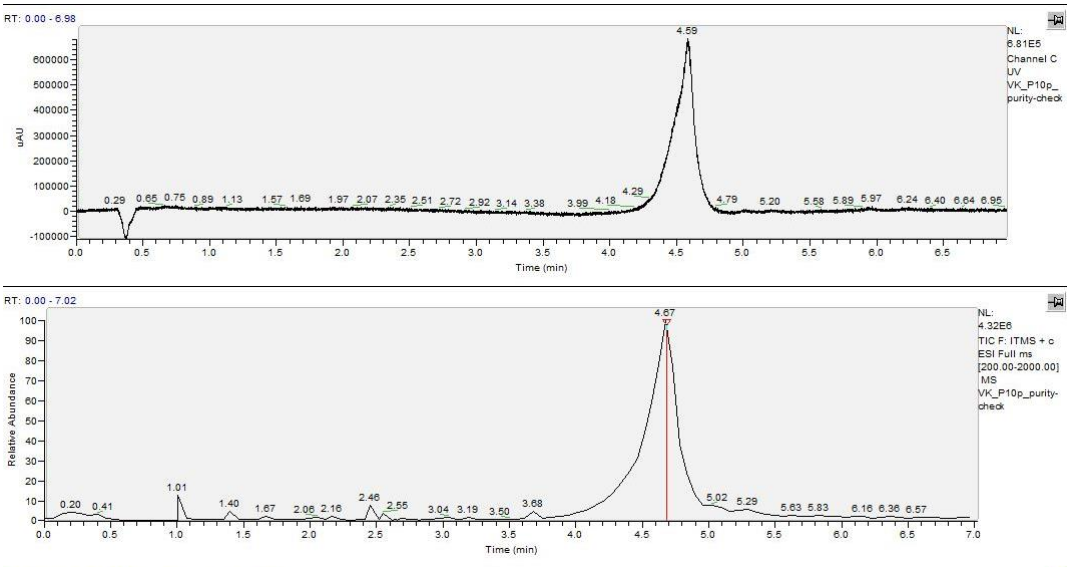


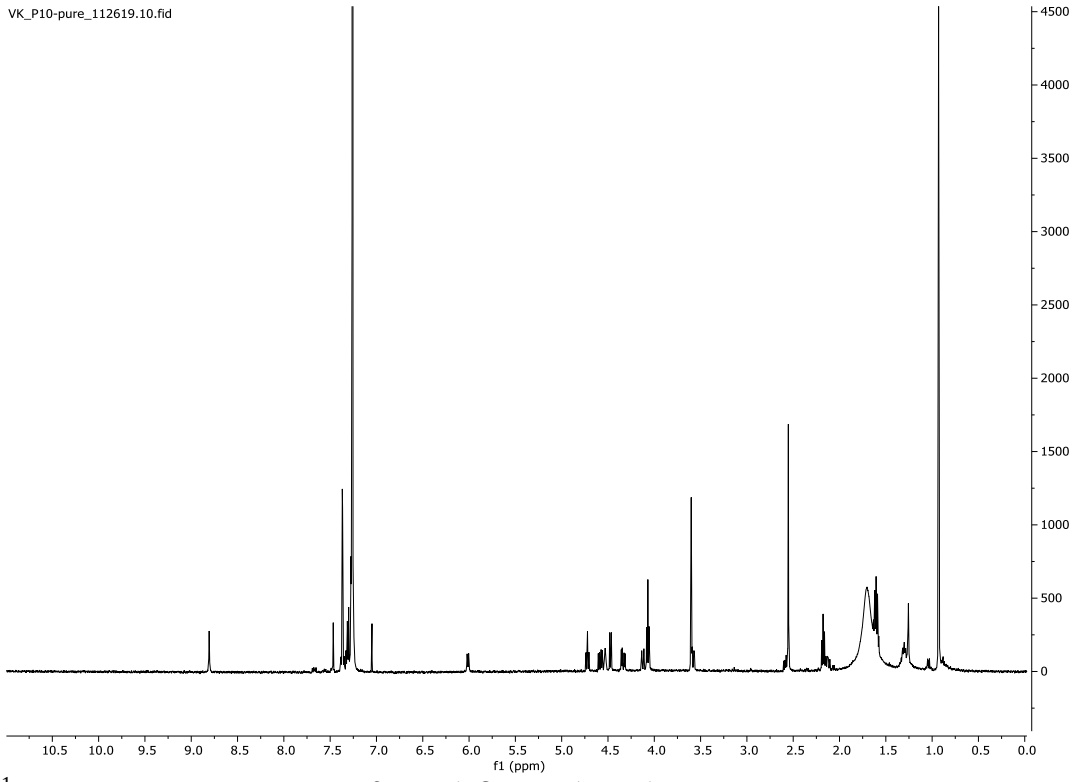
Exact mass: 614.31; Observed mass:

*3.5.8.10 NMR and LC/MS spectra for VK-P10 (**10**)*



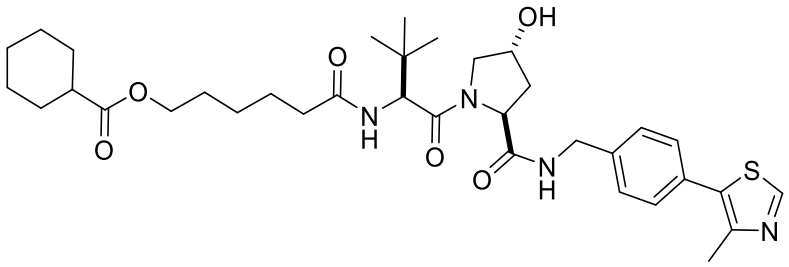
Exact mass: 662.31; Observed mass: 663.26



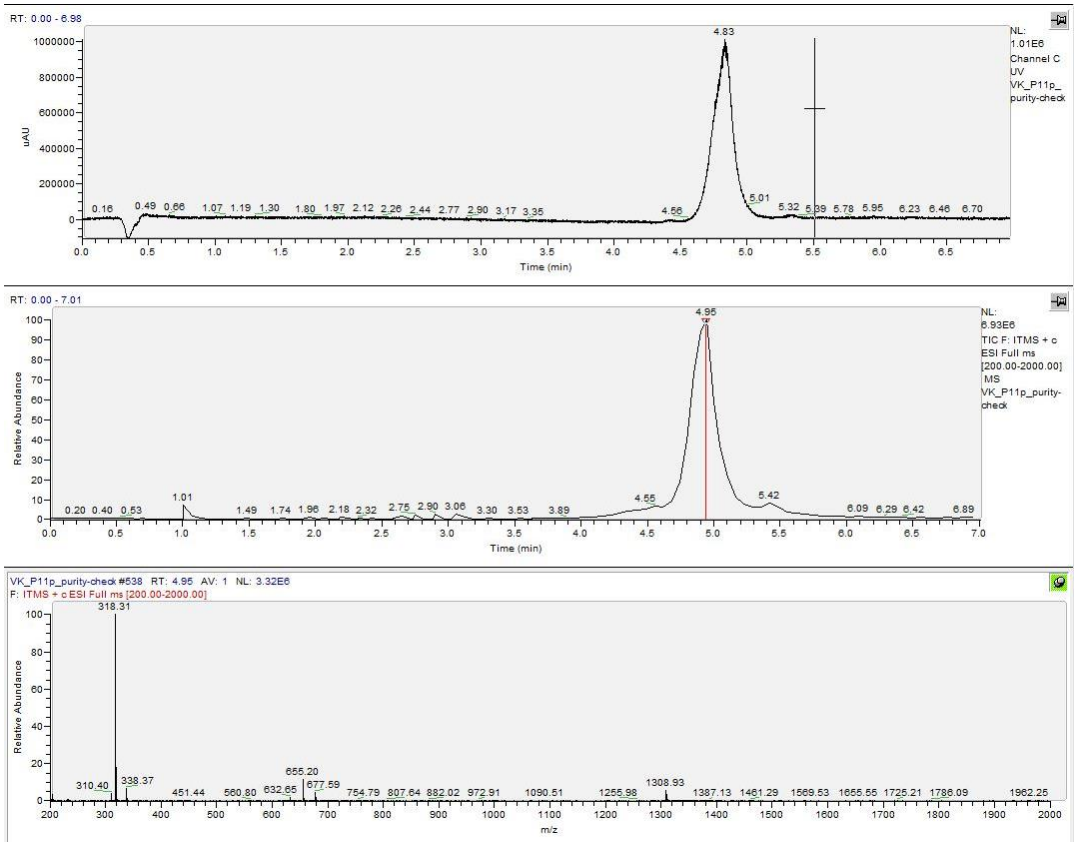


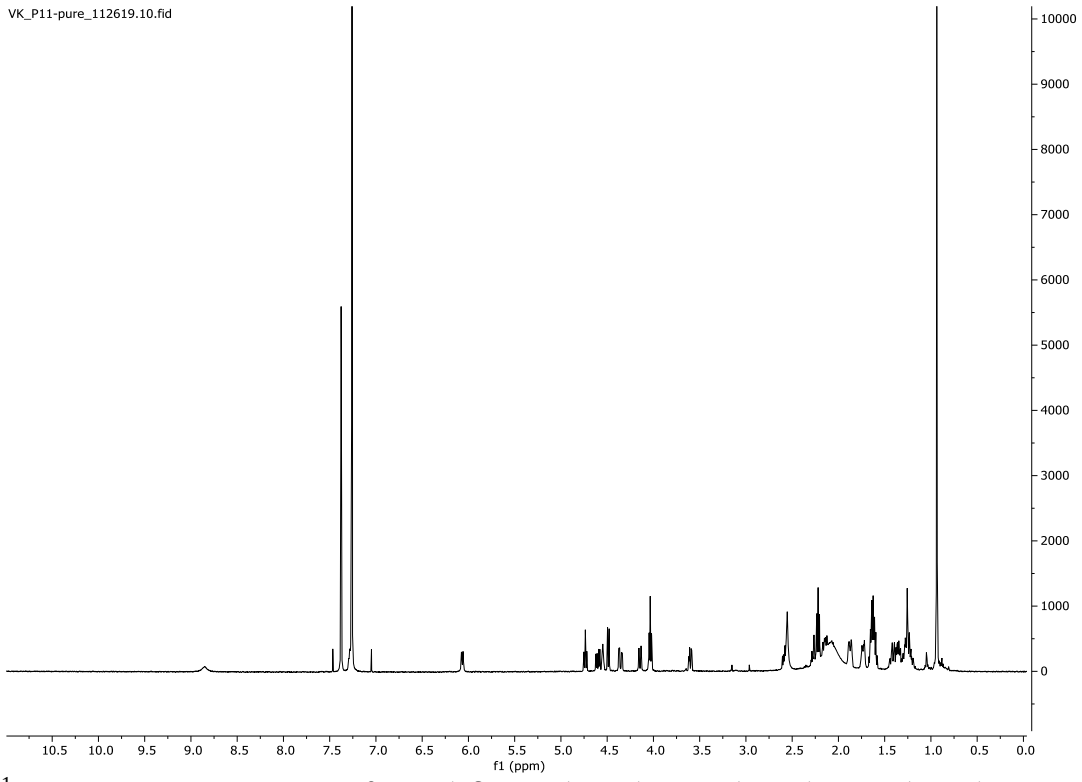
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.81 (s, 1H), 7.40 – 7.30 (m, 4H), 7.33 – 7.26 (m, 3H), 6.01 (d, *J* = 8.5 Hz, 1H), 4.72 (t, *J* = 7.9 Hz, 1H), 4.58 (dd, *J* = 15.0, 6.7 Hz, 1H), 4.53 (s, 1H), 4.47 (d, *J* = 8.6 Hz, 1H), 4.34 (dd, *J* = 14.9, 5.1 Hz, 1H), 4.12 (d, *J* = 11.6 Hz, 1H), 4.07 (t, *J* = 6.6 Hz, 2H), 3.60 (s, 2H), 3.61 – 3.55 (m, 1H), 2.62 – 2.55 (m, 1H), 2.55 (s, 3H), 2.21 – 2.09 (m, 2H), 1.60 (p, *J* = 7.1 Hz, 4H), 1.33 – 1.24 (m, 4H), 0.93 (s, 8H).

*3.5.8.11 NMR and LC/MS spectra for VK-P11 (**11**)*



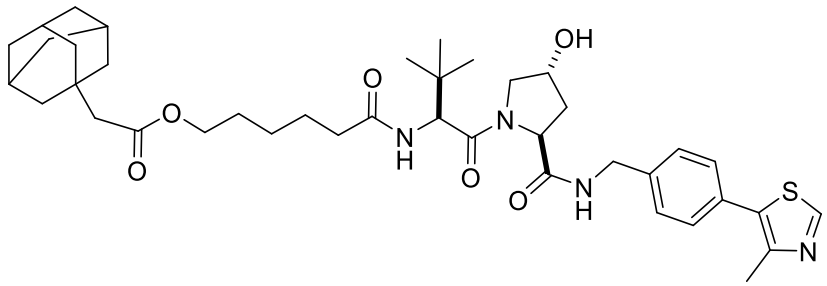
Exact mass: 654.35; Observed mass: 655.20



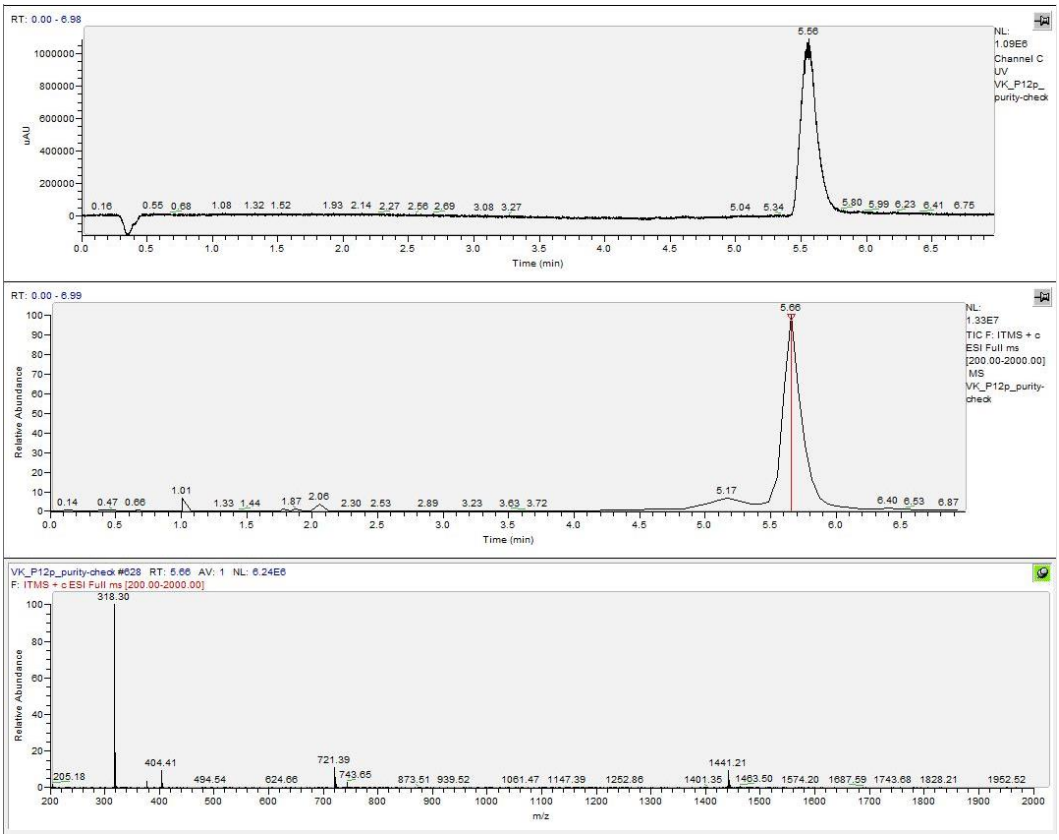


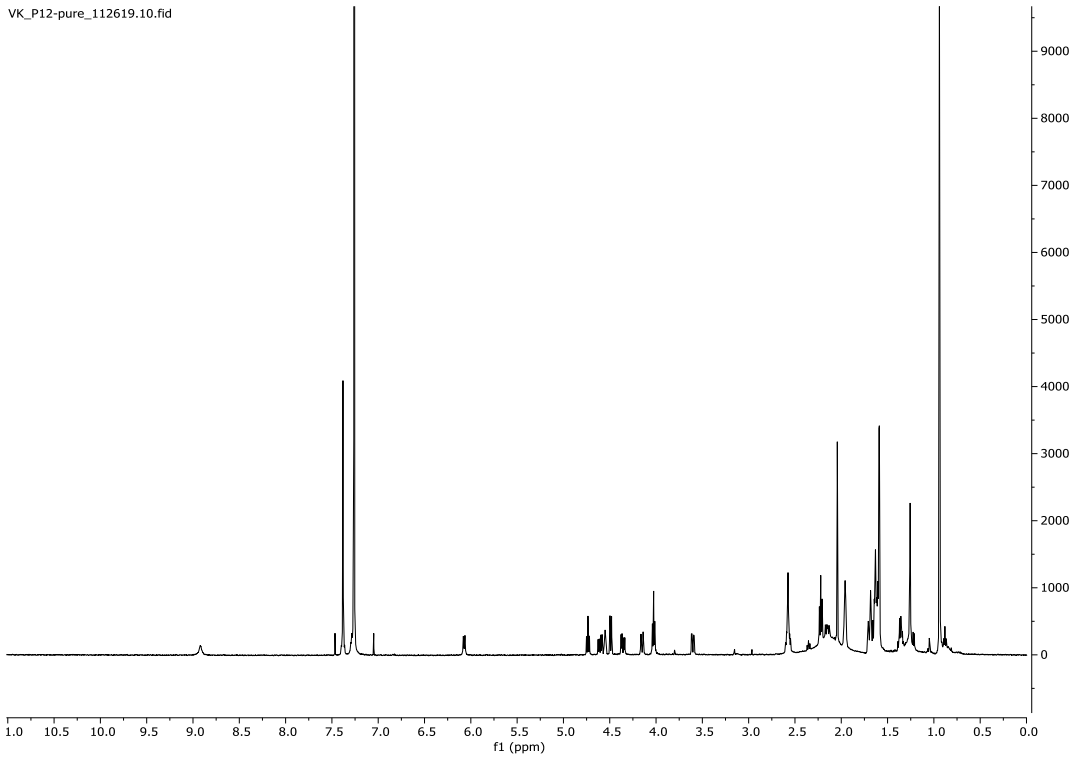
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.85 (s, 1H), 7.38 (s, 4H), 7.28 (s, 1H), 6.07 (d, *J* = 8.6 Hz, 1H), 4.74 (t, *J* = 7.9 Hz, 1H), 4.60 (dd, *J* = 15.0, 6.7 Hz, 1H), 4.55 (s, 1H), 4.49 (d, *J* = 8.6 Hz, 1H), 4.36 (dd, *J* = 14.9, 5.1 Hz, 1H), 4.15 (d, *J* = 11.5 Hz, 1H), 4.03 (td, *J* = 6.6, 1.1 Hz, 2H), 3.60 (dd, *J* = 11.4, 3.5 Hz, 1H), 2.62 – 2.54 (m, 1H), 2.55 (s, 3H), 2.27 (ddt, *J* = 11.3, 7.7, 3.6 Hz, 1H), 2.22 (t, *J* = 7.5 Hz, 2H), 1.87 (d, *J* = 13.0 Hz, 2H), 1.76 – 1.70 (m, 2H), 1.62 (dt, *J* = 14.4, 7.5 Hz, 5H), 1.42 (s, 1H), 1.41 – 1.31 (m, 2H), 1.31 – 1.27 (m, 1H), 1.26 (s, 2H), 1.24 – 1.17 (m, 1H), 0.94 (s, 8H).

*3.5.8.12 NMR and LC/MS spectra for VK-P12 (**12**)*



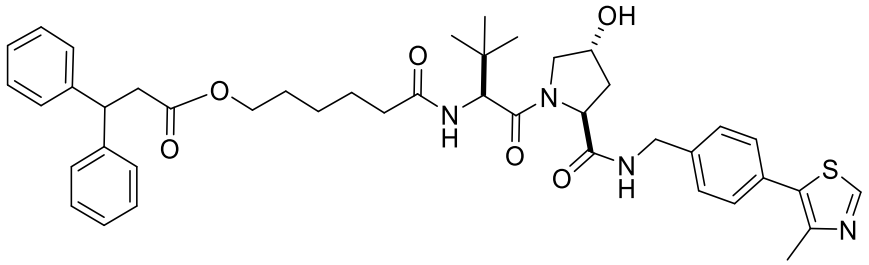
Exact mass: 720.39; Observed mass: 721.39



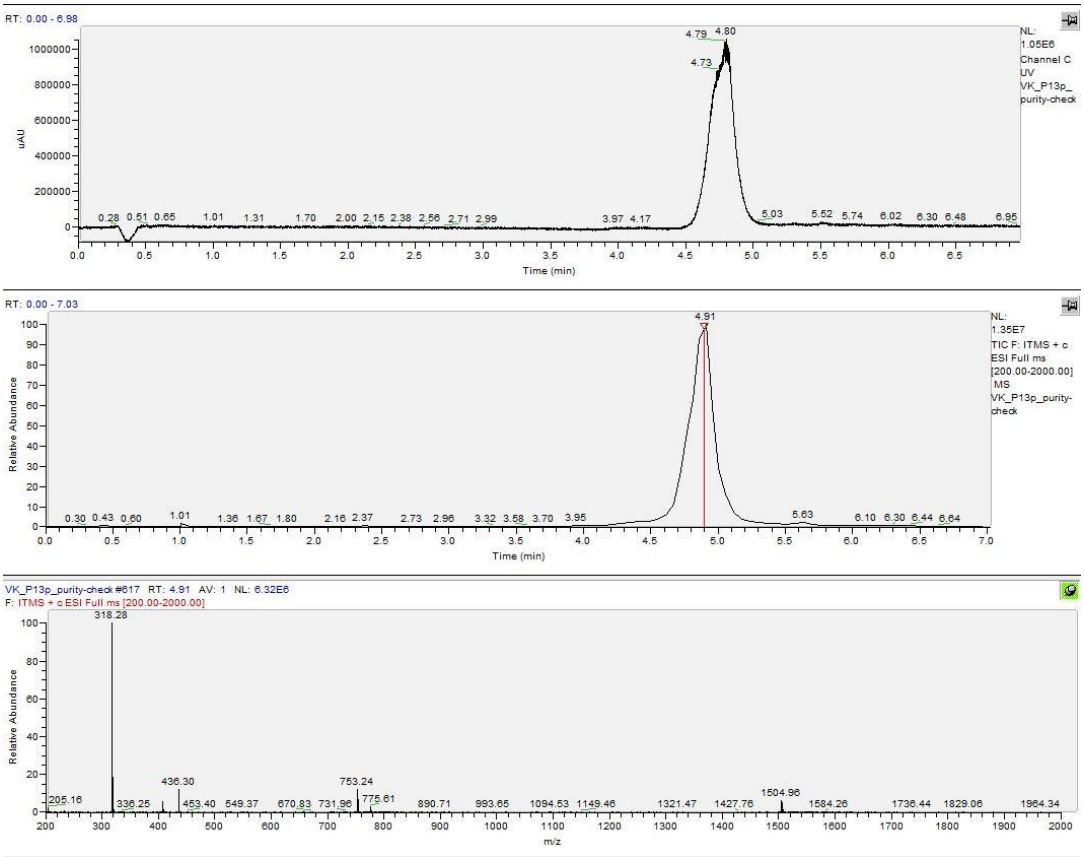


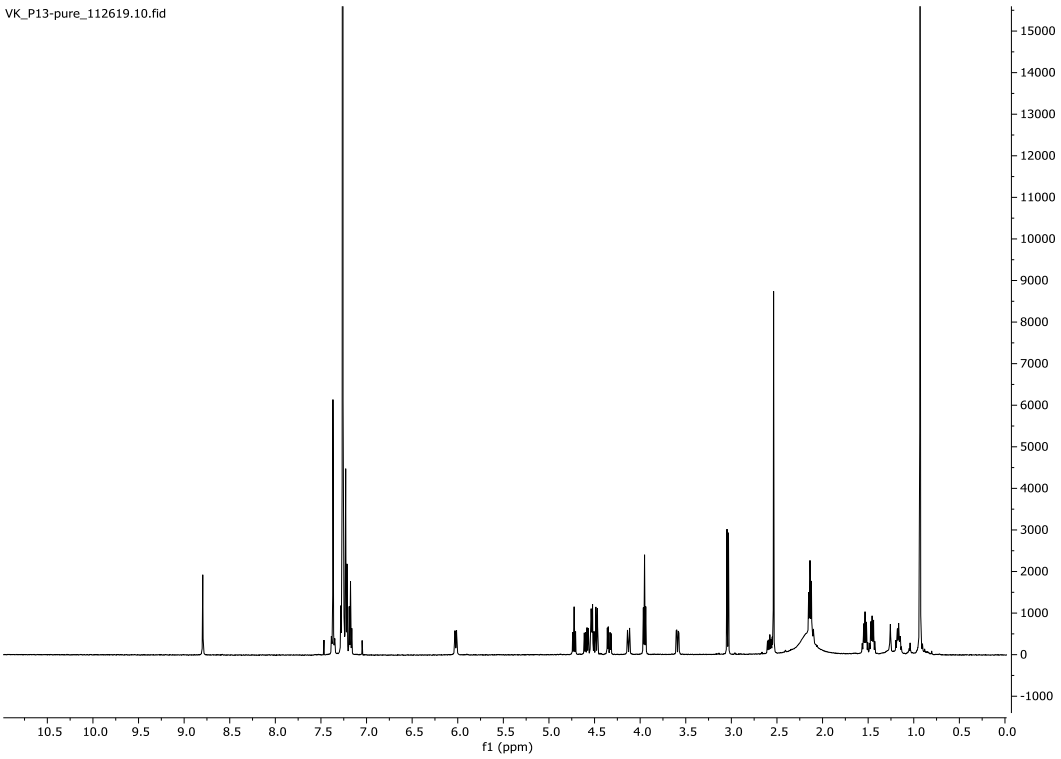
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.92 (s, 1H), 7.38 (s, 4H), 6.07 (d, *J* = 8.6 Hz, 1H), 4.74 (t, *J* = 8.0 Hz, 1H), 4.60 (dd, *J* = 15.1, 6.7 Hz, 1H), 4.55 (s, 1H), 4.49 (d, *J* = 8.7 Hz, 1H), 4.36 (dd, *J* = 15.1, 5.1 Hz, 1H), 4.15 (d, *J* = 11.5 Hz, 1H), 4.03 (td, *J* = 6.7, 1.2 Hz, 2H), 3.60 (dd, *J* = 11.4, 3.5 Hz, 1H), 2.57 (s, 3H), 2.62 – 2.53 (m, 1H), 2.25 – 2.11 (m, 4H), 2.04 (s, 2H), 1.96 (s, 4H), 1.71 (s, 1H), 1.70 – 1.61 (m, 9H), 1.59 (d, *J* = 2.7 Hz, 7H), 1.36 (qd, *J* = 8.7, 8.2, 5.3 Hz, 2H), 1.26 (s, 4H), 0.94 (s, 9H), 0.88 (t, *J* = 6.9 Hz, 0H).

*3.5.8.13 NMR and LC/MS spectra for VK-P13 (**13**)*



Exact mass: 752.36; Observed mass: 753.24

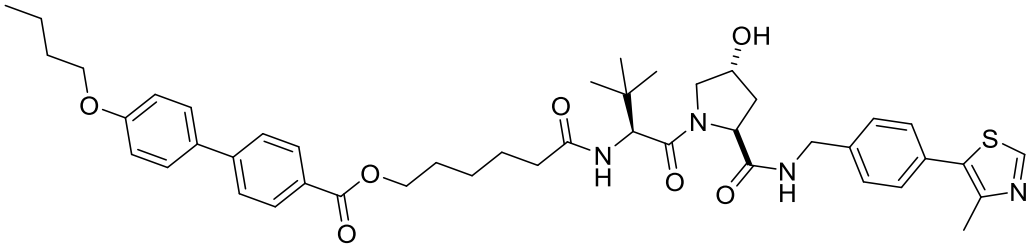




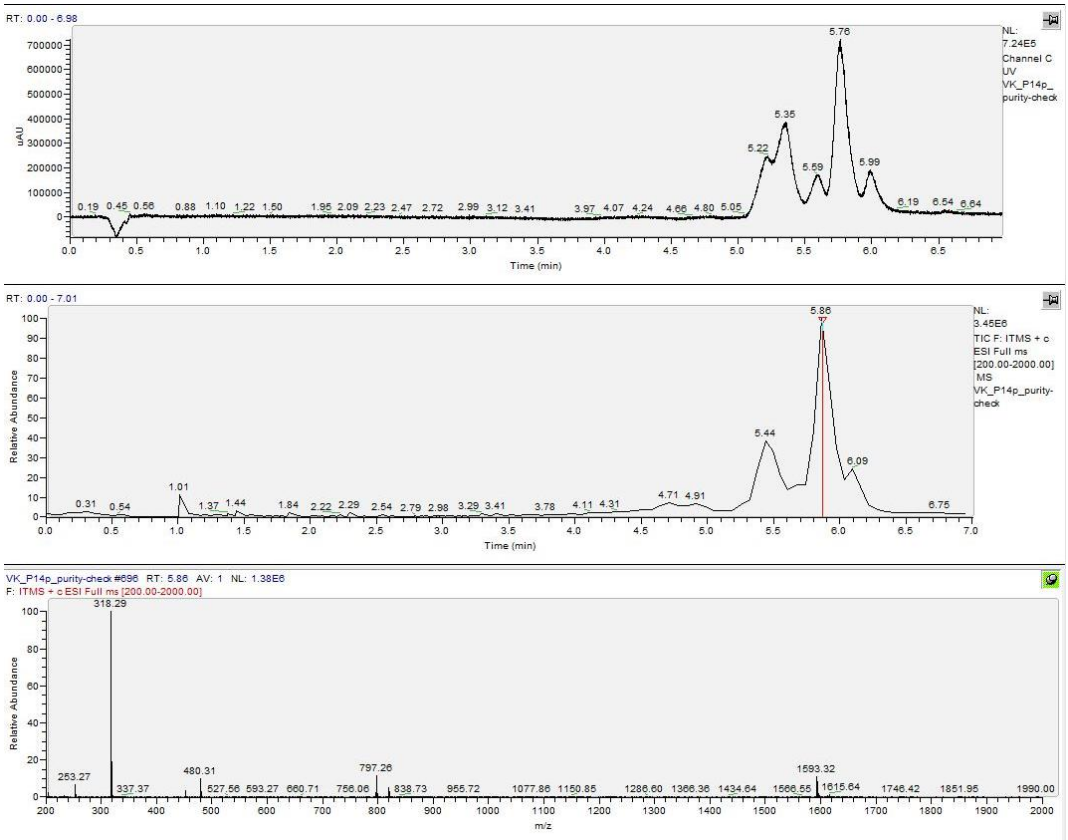
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.80 (s, 1H), 7.37 (d, *J* = 1.6 Hz, 4H), 7.30 – 7.14 (m, 11H), 6.02 (d, *J* = 8.6 Hz, 1H), 4.72 (t, *J* = 8.0 Hz, 1H), 4.63 – 4.45 (m, 5H), 4.34 (dd, *J* = 14.9, 5.2 Hz, 1H), 4.13 (d, *J* = 11.5 Hz, 1H), 3.95 (t, *J* = 6.6 Hz, 2H), 3.59 (dd, *J* = 11.4, 3.5 Hz, 1H), 3.04 (d, *J* = 8.1 Hz, 2H), 2.58 (ddd, *J* = 12.9, 7.8, 4.6 Hz, 1H), 2.54 (s, 3H), 2.17 – 2.08 (m, 3H), 1.49 (dq, *J* = 38.0, 7.2, 6.8 Hz, 4H), 1.26 (s, 1H), 1.21 – 1.12 (m, 2H), 0.93 (s, 9H).

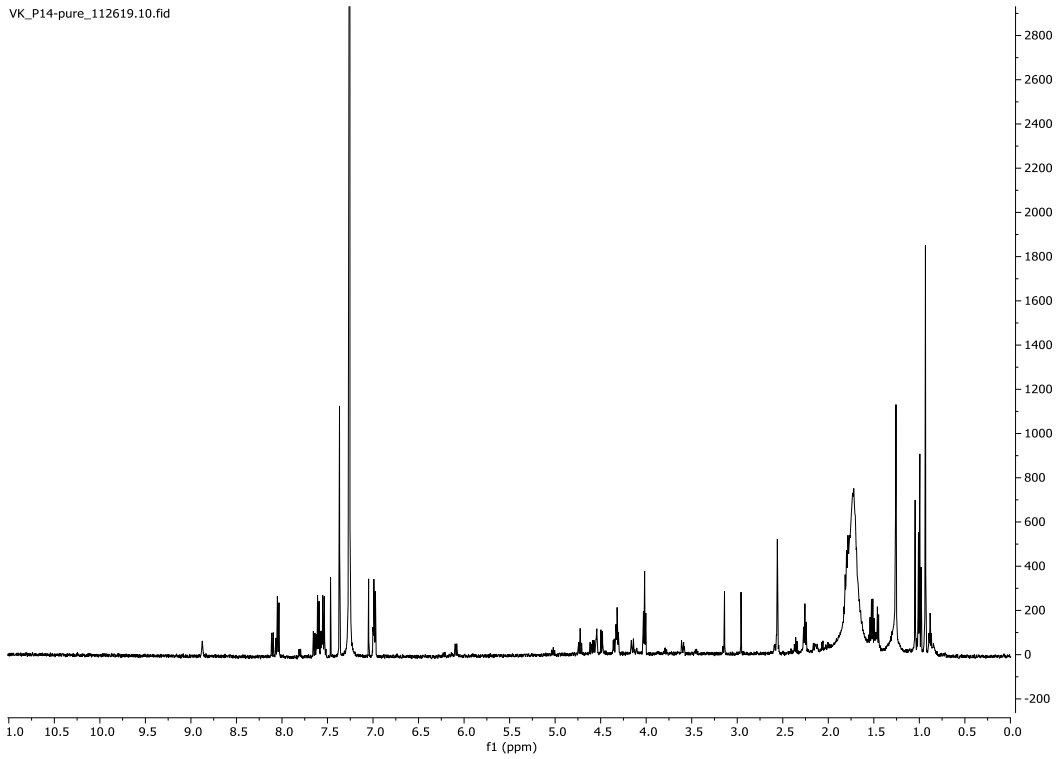
*3.5.8.14 NMR and LC/MS spectra for VK-P14 (**14**)*

VK-P14



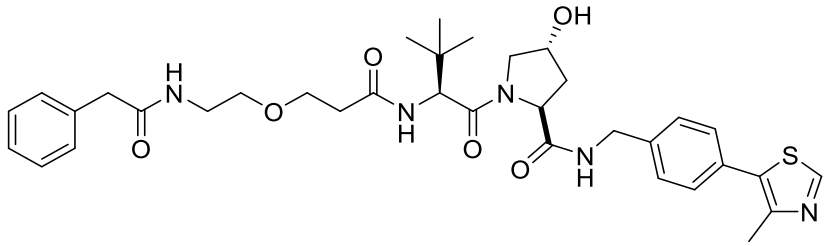
Exact mass: 796.39; Observed mass: 797.26



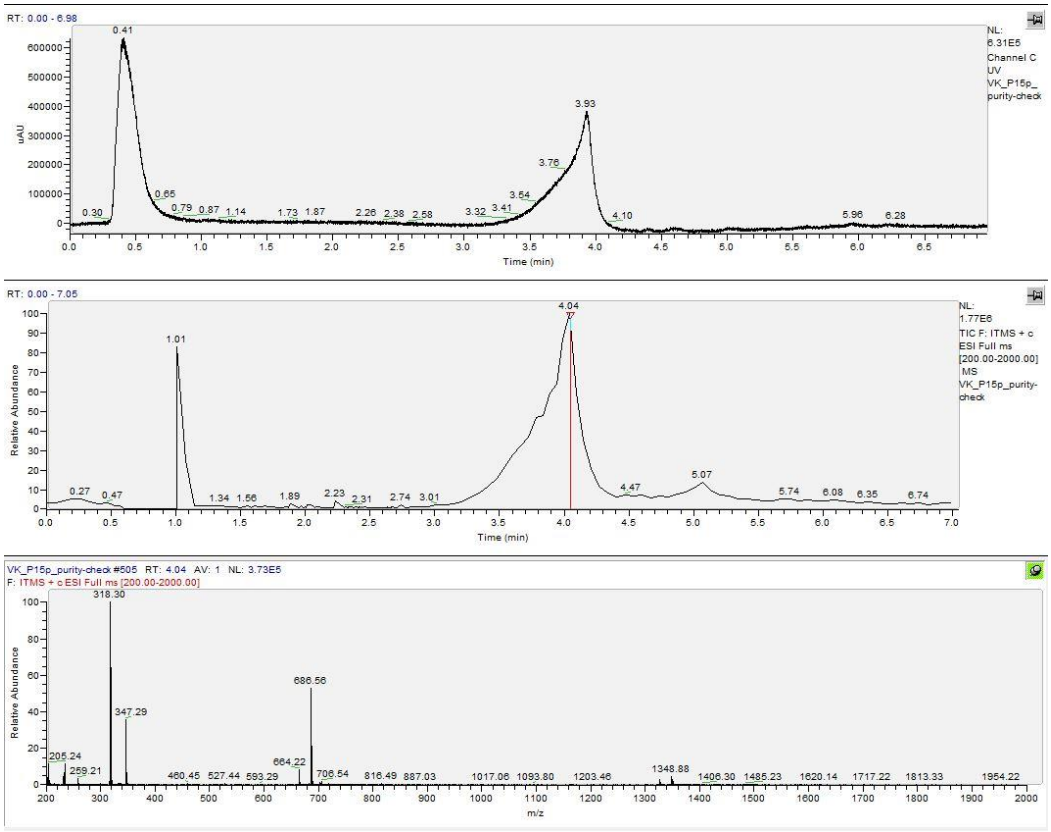


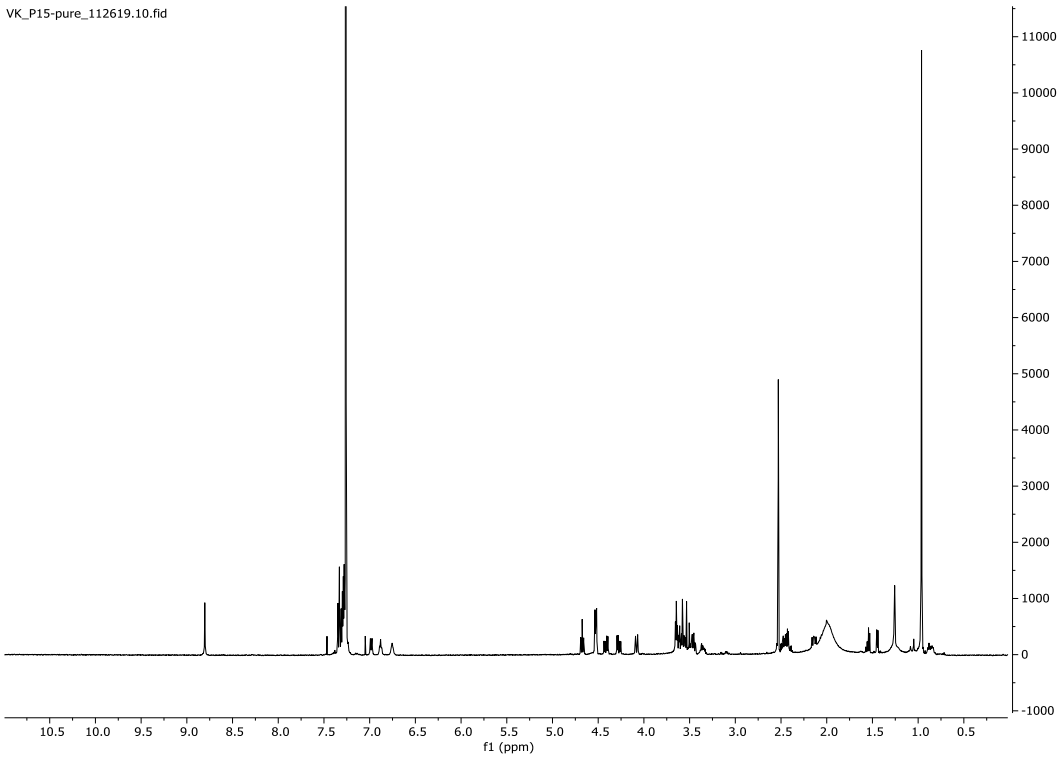
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.88 (s, 1H), 8.13 – 8.05 (m, 1H), 8.04 (d, *J* = 8.5 Hz, 2H), 7.67 – 7.50 (m, 7H), 7.37 (s, 4H), 7.02 – 6.96 (m, 4H), 6.09 (d, *J* = 8.5 Hz, 1H), 4.73 (t, *J* = 8.0 Hz, 1H), 4.59 (dd, *J* = 15.1, 6.6 Hz, 1H), 4.54 (s, 1H), 4.49 (d, *J* = 8.6 Hz, 1H), 4.38 – 4.29 (m, 3H), 4.15 (d, *J* = 11.3 Hz, 1H), 4.02 (t, *J* = 6.5 Hz, 4H), 3.60 (dd, *J* = 11.4, 3.5 Hz, 1H), 3.14 (s, 1H), 2.96 (s, 1H), 2.56 (s, 3H), 2.36 (t, *J* = 7.6 Hz, 1H), 2.26 (t, *J* = 7.4 Hz, 3H), 2.18 – 2.10 (m, 1H), 1.81 (q, *J* = 7.2 Hz, 5H), 1.53 (dt, *J* = 15.1, 7.4 Hz, 4H), 1.48 (s, 2H), 1.45 (d, *J* = 6.7 Hz, 1H), 1.26 (s, 11H), 1.06 – 0.96 (m, 10H), 0.93 (s, 9H), 0.88 (t, *J* = 6.9 Hz, 1H).

*3.5.8.15 NMR and LC/MS spectra for VK-P15 (**15**)*



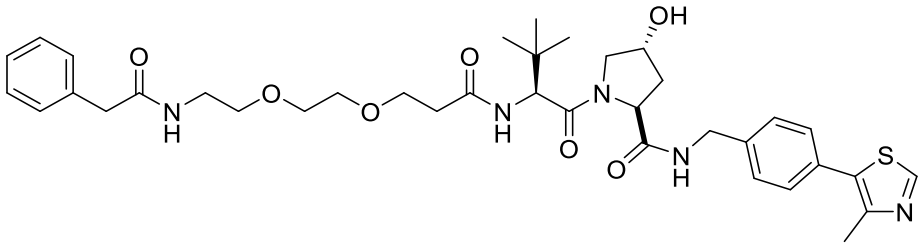
Exact mass: 663.31; Observed mass: 664.22



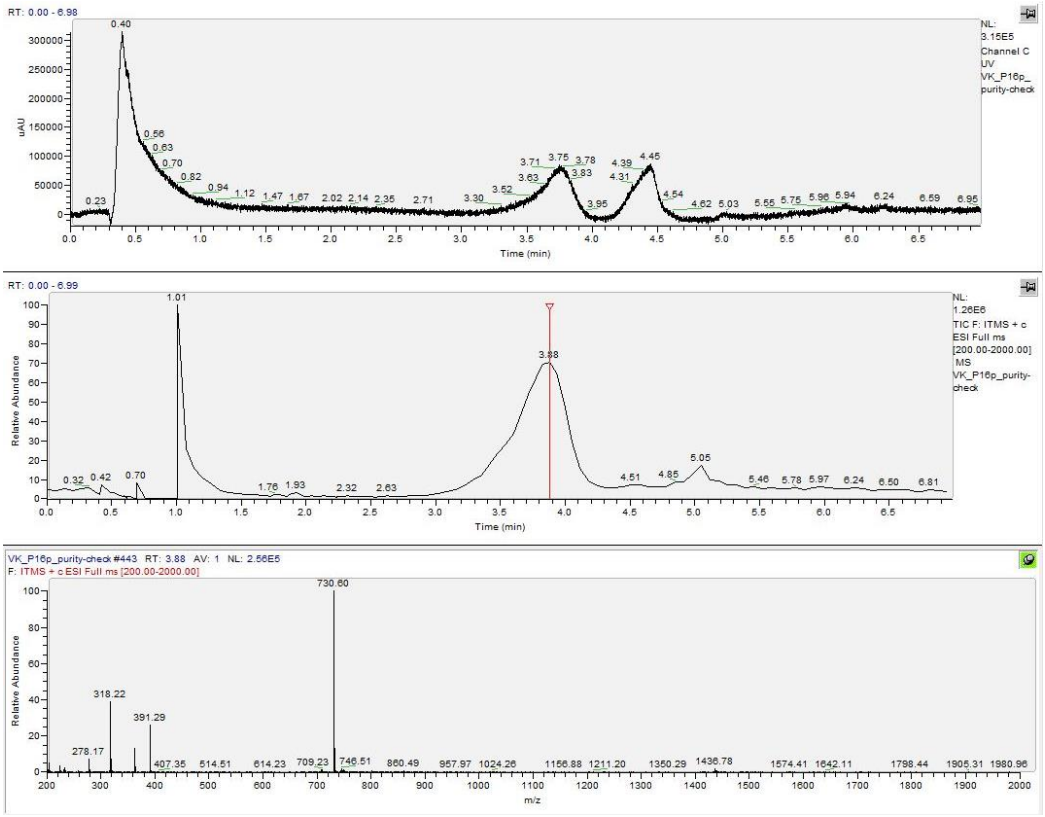


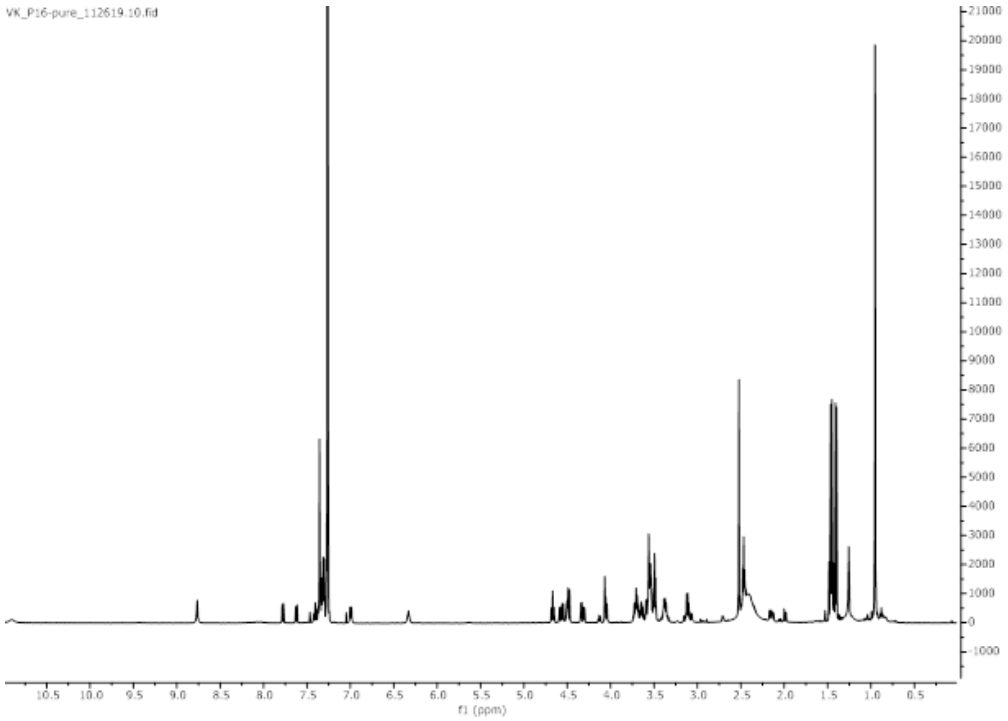
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.80 (s, 1H), 7.40 – 7.26 (m, 7H), 6.98 (d, *J* = 8.9 Hz, 1H), 6.88 (t, *J* = 5.9 Hz, 1H), 6.75 (s, 1H), 4.68 (t, *J* = 8.0 Hz, 1H), 4.53 (d, *J* = 8.8 Hz, 1H), 4.41 (dd, *J* = 15.0, 6.5 Hz, 1H), 4.27 (dd, *J* = 15.0, 5.5 Hz, 1H), 4.08 (d, *J* = 11.4 Hz, 1H), 3.67 – 3.55 (m, 4H), 3.58 – 3.42 (m, 3H), 3.39 – 3.32 (m, 1H), 2.53 (s, 3H), 2.52 – 2.36 (m, 2H), 2.18 – 2.10 (m, 0H), 1.54 (t, *J* = 7.0 Hz, 1H), 1.45 (d, *J* = 6.6 Hz, 1H), 1.26 (s, 2H), 0.96 (s, 9H).

*3.5.8.16 NMR and LC/MS spectra for VK-P16 (**16**)*



Exact mass: 707.34; Observed mass: 709.23



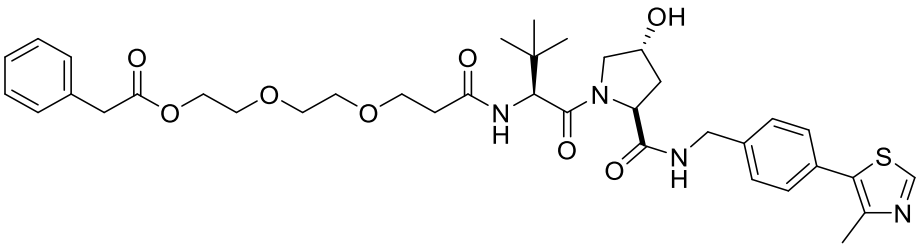


^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.90 (s, 1H), 8.77 (s, 1H), 8.07 (s, 1H), 7.78 (dd, *J* = 7.5, 1.0 Hz, 1H), 7.65 – 7.59 (m, 1H), 7.44 – 7.34 (m, 1H), 7.37 – 7.29 (m, 6H), 7.30 (d, *J* = 0.9 Hz, 1H), 7.27 (s, 2H), 7.00 (d, *J* = 8.3 Hz, 1H), 6.33 (s, 1H), 4.67 (t, *J* = 8.1 Hz, 1H), 4.57 (dd, *J* = 15.0, 6.7 Hz, 1H), 4.52 – 4.46 (m, 2H), 4.32 (dd, *J* = 15.0, 5.2 Hz, 1H), 4.09 – 4.02 (m, 2H), 3.76 – 3.61 (m, 3H), 3.64 – 3.46 (m, 9H), 3.37 (ddt, *J* = 14.0, 8.8, 4.0 Hz, 2H), 3.11 (qd, *J* = 7.4, 4.2 Hz, 2H), 2.52 (s, 3H), 2.46 (h, *J* = 4.1 Hz, 3H), 2.19 – 2.10 (m, 1H), 1.50 – 1.34 (m, 13H), 1.26 (s, 3H), 0.95 (s, 9H).

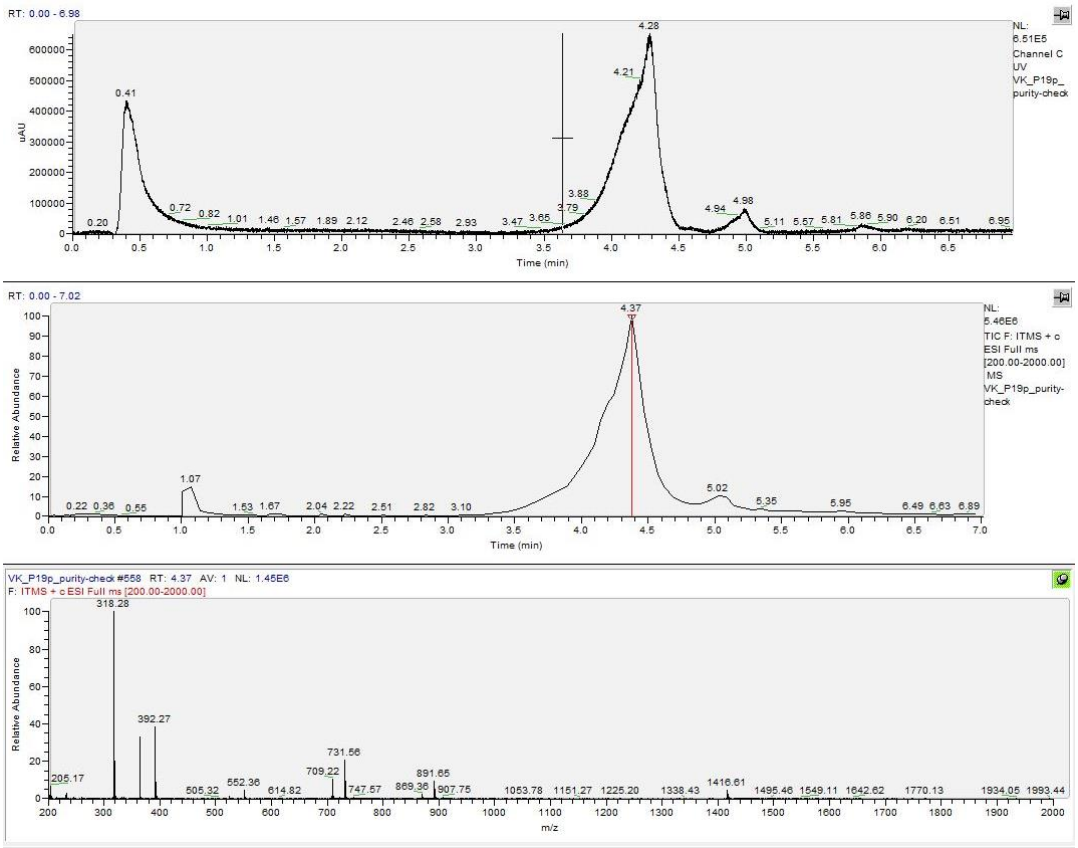
*3.5.8.17 NMR and LC/MS spectra for VK-P17 (**17**)*

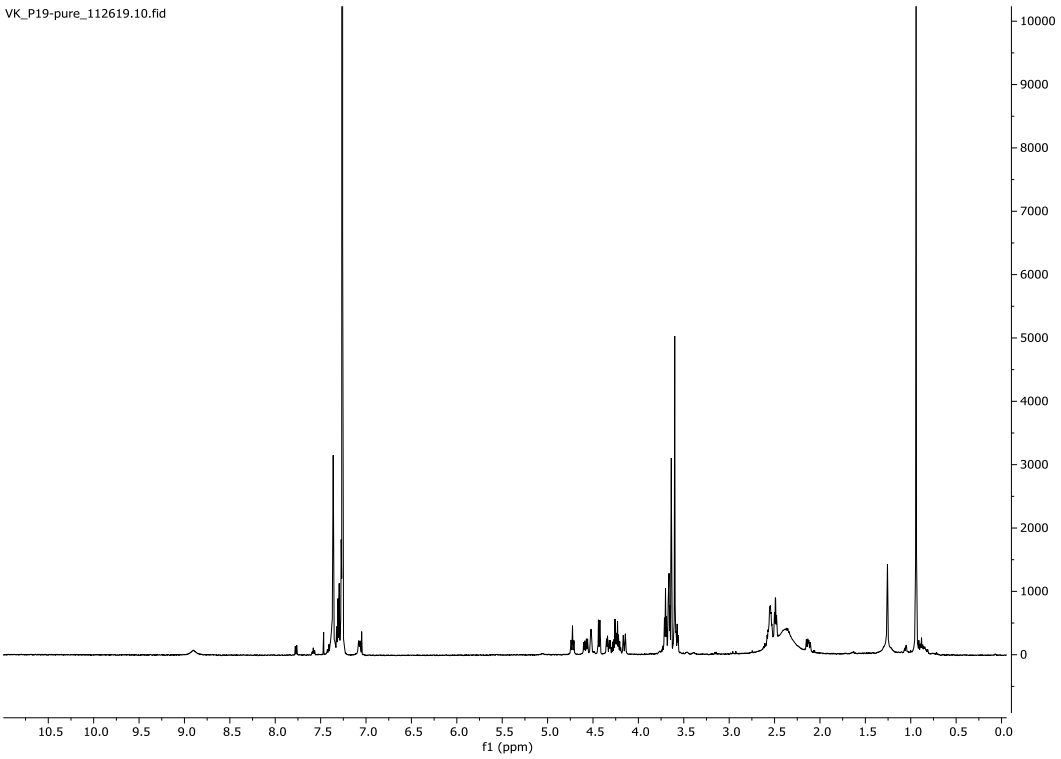
*3.5.8.18 NMR and LC/MS spectra for VK-P18 (**18**)*

*3.5.8.19 NMR and LC/MS spectra for VK-P19 (**19**)*



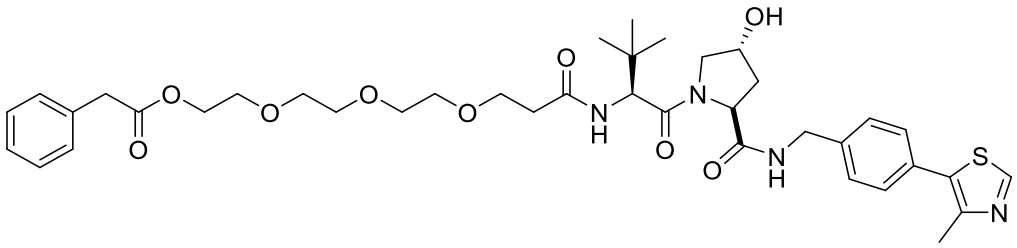
Exact mass: 708.32; Observed mass: 709.22



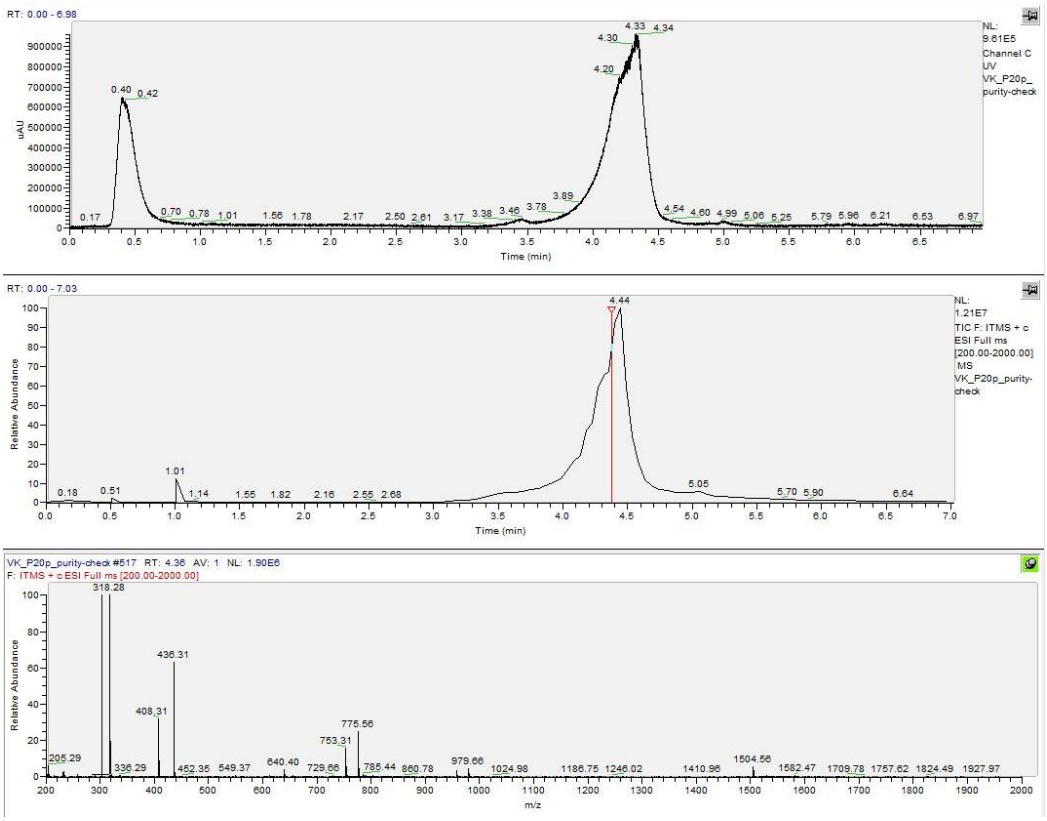


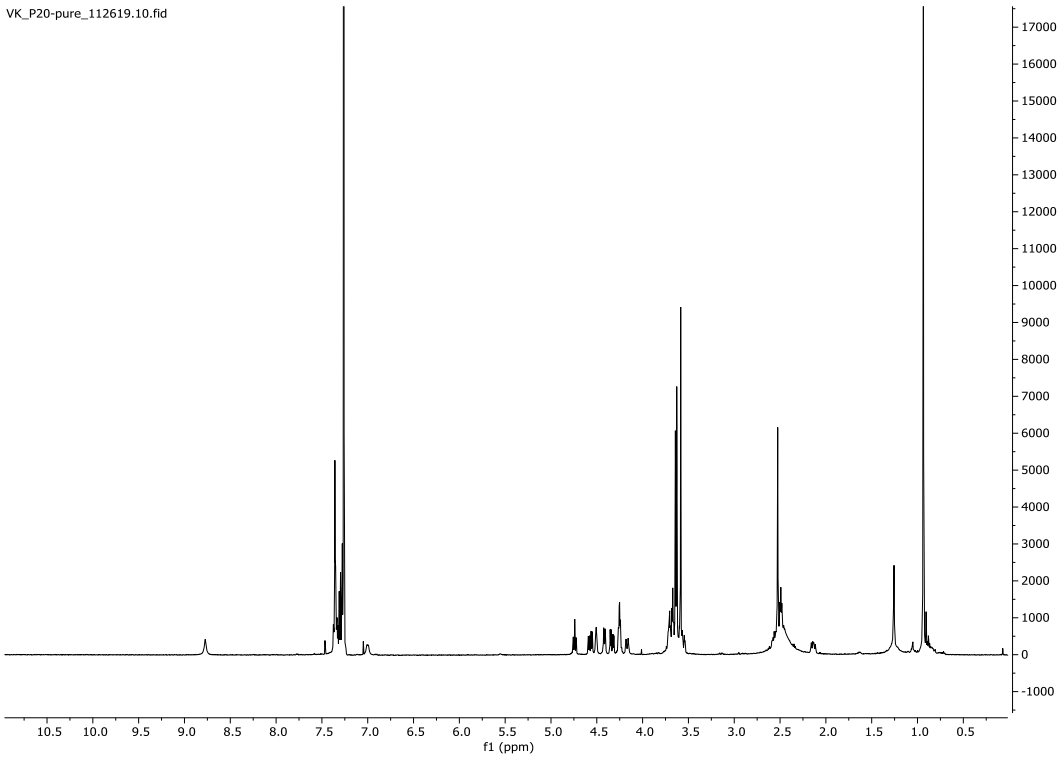
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.90 (s, 1H), 7.36 (s, 3H), 7.35 – 7.26 (m, 4H), 7.07 (d, *J* = 7.9 Hz, 1H), 4.73 (t, *J* = 8.0 Hz, 1H), 4.58 (dd, *J* = 15.1, 6.7 Hz, 1H), 4.52 (s, 1H), 4.43 (d, *J* = 8.0 Hz, 1H), 4.37 – 4.18 (m, 3H), 4.15 (d, *J* = 11.5 Hz, 1H), 3.70 (t, *J* = 5.6 Hz, 2H), 3.69 – 3.62 (m, 4H), 3.60 (s, 3H), 3.60 – 3.55 (m, 1H), 2.56 (dd, *J* = 13.6, 5.0 Hz, 3H), 2.56 – 2.46 (m, 2H), 2.13 (dd, *J* = 13.6, 8.0 Hz, 1H), 1.26 (s, 2H), 0.94 (s, 9H), 0.85 (s, 1H).

*3.5.8.20 NMR and LC/MS spectra for VK-P20 (**20**)*



Exact mass: 752.35; Observed mass: 753.31





^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.77 (s, 1H), 7.39 – 7.25 (m, 9H), 7.00 (d, *J* = 7.9 Hz, 1H), 4.74 (t, *J* = 8.0 Hz, 1H), 4.57 (dd, *J* = 14.9, 6.7 Hz, 1H), 4.51 (s, 1H), 4.42 (d, *J* = 7.9 Hz, 1H), 4.33 (dd, *J* = 15.0, 5.2 Hz, 1H), 4.25 (td, *J* = 4.4, 1.5 Hz, 2H), 4.17 (d, *J* = 11.4 Hz, 1H), 3.71 (dd, *J* = 5.5, 2.5 Hz, 1H), 3.72 – 3.63 (m, 2H), 3.63 (d, *J* = 7.5 Hz, 6H), 3.58 (s, 3H), 3.61 – 3.52 (m, 1H), 2.51 (d, *J* = 17.0 Hz, 6H), 2.14 (dd, *J* = 13.6, 8.2 Hz, 1H), 1.26 (s, 2H), 0.94 (s, 9H), 0.89 (d, *J* = 12.6 Hz, 1H).

*3.5.8.21 NMR and LC/MS spectra for VK-P21 (**21**)*

3.6 Abbreviations

ACN: acetonitrile; DBU: 1,8-Diazabicyclo[5.4.0]undec-7-ene; DCM: Dichloromethane; DIC: N,N′-Diisopropylcarbodiimide; DIPEA: N,N-Diisopropylethylamine; DMAP: 4-Dimethylaminopyridine; DMF: N,N-Dimethylformamide ; HATU: 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate; HF/pyridine: Hydrogen fluoride pyridine; TFA: Trifluoroacetic acid; THF: Tetrahydrofuran; TMS-ethanol: 2-(Trimethylsilyl)ethanol

3.7 Acknowledgements

We thank Chad E. Townsend for creation of the Python code used to process mass-directed assays.

3.8 Author Contributions

V.G.K and R.S.L designed the project and wrote the manuscript. Compounds were synthesized by V.G.K. and Y.S. PAMPA permeability, LogD, LPE, and plasma stability data were collected and processed by V.G.K. The manuscript was read and edited by all.

3.9 Funding Sources

Research reported in this publication was supported by the National Institute of General Medicine Studies of the National Institutes of Health under award number R01GM131135 (to R.S.L.). This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program (NSF DGE 1339067 to V.G.K.). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Institutes of Health or the National Science Foundation.

3.10 Notes

The authors declare the following competing financial interest(s): The A.C. laboratory receives or has received sponsored research support from Boehringer Ingelheim, Eisai Co., Nurix, Ono Pharmaceuticals, and Amphista Therapeutics. A.C. is a scientific founder, shareholder, non-executive director and consultant of Amphista Therapeutics, a company that is developing targeted protein degradation therapeutic platforms.

REFERNCES

1 Bockus, A. T., McEwen, C. M. & Lokey, R. S. Form and Function in Cyclic Peptide Natural Products: A Pharmacokinetic Perspective. *Curr. Top. Med. Chem.* **13**, 821-836, doi:http://dx.doi.org/10.2174/1568026611313070005 (2013).

2 Zorzi, A., Deyle, K. & Heinis, C. Cyclic peptide therapeutics: past, present and future. *Curr. Opin. Chem. Biol.* **38**, 24-29, doi:https://doi.org/10.1016/j.cbpa.2017.02.006 (2017).

3 Pye, C. R., Bertin, M. J., Lokey, R. S., Gerwick, W. H. & Linington, R. G. Retrospective analysis of natural products provides insights for future discovery trends. *Proc. Natl. Acad. Sci.* **114**, 5601, doi:10.1073/pnas.1614680114 (2017).

4 Marsault, E. & Peterson, M. L. Macrocycles Are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery. *J. Med. Chem.* **54**, 1961-2004, doi:10.1021/jm1012374 (2011).

5 Driggers, E. M., Hale, S. P., Lee, J. & Terrett, N. K. The exploration of macrocycles for drug discovery - An underexploited structural class. *Nat. Rev. Drug Discovery* **7**, 608-624, doi:10.1038/nrd2590 (2008).

6 Whitty, A. *et al.* Quantifying the chameleonic properties of macrocycles and other high-molecular-weight drugs. *Drug Discov. Today* **21**, 712-717, doi:https://doi.org/10.1016/j.drudis.2016.02.005 (2016).

7 Naylor, M. R. *et al.* Lipophilic Permeability Efficiency Reconciles the Opposing Roles of Lipophilicity in Membrane Permeability and Aqueous Solubility. *J. Med. Chem.* **61**, 11169-11182, doi:10.1021/acs.jmedchem.8b01259 (2018).

8 Pye, C. R. *et al.* Nonclassical Size Dependence of Permeation Defines Bounds for Passive Adsorption of Large Drug Molecules. *J. Med. Chem.* **60**, 1665-1672, doi:10.1021/acs.jmedchem.6b01483 (2017).

9 Over, B. *et al.* Structural and conformational determinants of macrocycle cell permeability. *Nat. Chem. Biol.* **12**, 1065-1074, doi:10.1038/nchembio.2203 (2016).

10 Olatunji, O. J. *et al.* The genus Cordyceps: An extensive review of its traditional uses, phytochemistry and pharmacology. *Fitoterapia* **129**, 293-316, doi:https://doi.org/10.1016/j.fitote.2018.05.010 (2018).

11 Rukachaisirikul, V. *et al.* A Cyclopeptide from the Insect Pathogenic Fungus Cordyceps sp. BCC 1788. *J. Nat. Prod.* **69**, 305-307, doi:10.1021/np050433l (2006).

12 Isaka, M., Srisanoh, U., Lartpornmatulee, N. & Boonruangprapa, T. ES-242 Derivatives and Cycloheptapeptides from Cordyceps sp. Strains BCC 16173 and BCC 16176. *J. Nat. Prod.* **70**, 1601-1604, doi:10.1021/np070357h (2007).

13 Kumar, S. *et al.* Total Synthesis and Pharmacological Investigation of Cordyheptapeptide A. *Molecules* **22**, 682, doi:10.3390/molecules22060682 (2017).

14 Turner, R. A., Hauksson, N. E., Gipe, J. H. & Lokey, R. S. Selective, On-Resin N-Methylation of Peptide N-Trifluoroacetamides. *Org. Lett.* **15**, 5012-5015, doi:10.1021/ol402341u (2013).

15 Cunningham, B. C. & Wells, J. A. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* **244**, 1081, doi:10.1126/science.2471267 (1989).

16 Furukawa, A. *et al.* Passive Membrane Permeability in Cyclic Peptomer Scaffolds Is Robust to Extensive Variation in Side Chain Functionality and Backbone Geometry. *J. Med. Chem.* **59**, 9503-9512, doi:10.1021/acs.jmedchem.6b01246 (2016).

17 Beck, J. G. *et al.* Intestinal Permeability of Cyclic Peptides : Common Key Backbone. *J. Am. Chem. Soc.* **134**, 12125−12133, doi:10.1021/ja303200d (2012).

18 Schwochert, J. *et al.* Stereochemistry Balances Cell Permeability and Solubility in the Naturally Derived Phepropeptin Cyclic Peptides. *ACS Med. Chem. Lett.* **7**, 757-761, doi:10.1021/acsmedchemlett.6b00100 (2016).

19 Wang, C. K. & Craik, D. J. Cyclic peptide oral bioavailability: Lessons from the past. *Pept. Sci.* **106**, 901-909, doi:10.1002/bip.22878 (2016).

20 Nakajima, N., Nakamura, H. & Kidera, A. Multicanonical Ensemble Generated by Molecular Dynamics Simulation for Enhanced Conformational Sampling of Peptides. *J. Phys. Chem. B* **101**, 817-824, doi:10.1021/jp962142e (1997).

21 Higo, J., Umezawa, K. & Nakamura, H. A virtual-system coupled multicanonical molecular dynamics simulation: Principles and applications to free-energy landscape of protein–protein interaction with an all-atom model in explicit solvent. *J. Chem. Phys.* **138**, 184106, doi:10.1063/1.4803468 (2013).

22 Ono, S. *et al.* Conformation and Permeability: Cyclic Hexapeptide Diastereomers. *J. Chem. Inf. Model.* **59**, 2952-2963, doi:10.1021/acs.jcim.9b00217 (2019).

23 Duan, Y. *et al.* A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999-2012, doi:10.1002/jcc.10349 (2003).

24 Khoury, G. A. *et al.* Forcefield_NCAA: Ab Initio Charge Parameters to Aid in the Discovery and Design of Therapeutic Proteins and Peptides with Unnatural Amino Acids and Their Application to Complement Inhibitors of the Compstatin Family. *ACS Synth. Biol.* **3**, 855-869, doi:10.1021/sb400168u (2014).

25 Monks, A. *et al.* Feasibility of a High-Flux Anticancer Drug Screen Using a Diverse Panel of Cultured Human Tumor Cell Lines. *JNCI, J. Natl. Cancer Inst.* **83**, 757-766, doi:10.1093/jnci/83.11.757 (1991).

26 Paull, K. D. *et al.* Display and Analysis of Patterns of Differential Activity of Drugs Against Human Tumor Cell Lines: Development of Mean Graph and COMPARE Algorithm. *JNCI, J. Natl. Cancer Inst.* **81**, 1088-1092, doi:10.1093/jnci/81.14.1088 (1989).

27 Bai, R. L. *et al.* Halichondrin B and homohalichondrin B, marine natural products binding in the vinca domain of tubulin. Discovery of tubulin-based mechanism of action by analysis of differential cytotoxicity data. *Journal of Biological Chemistry* **266**, 15882-15889 (1991).

28 Chan, J., Khan, S. N., Harvey, I., Merrick, W. & Pelletier, J. Eukaryotic protein synthesis inhibitors identified by comparison of cytotoxicity profiles. *RNA* **10**, 528-543, doi:10.1261/rna.5200204.initiation (2004).

29 Garreau de Loubresse, N. *et al.* Structural basis for the inhibition of the eukaryotic ribosome. *Nature* **513**, 517-522, doi:10.1038/nature13737 (2014).

30 Perlman, Z. E. *et al.* Multidimensional Drug Profiling By Automated Microscopy. *Science* **306**, 1194-1199, doi:10.1126/science.1100709 (2004).

31 Young, D. W. *et al.* Integrating high-content screening and ligand-target prediction to identify mechanism of action. *Nat. Chem. Biol.* **4**, 59-68, doi:10.1038/nchembio.2007.53 (2008).

32 Schulze, Christopher J. *et al.* "Function-First" Lead Discovery: Mode of Action Profiling of Natural Product Libraries Using Image-Based Screening. *Chem. Biol.* **20**, 285-295, doi:https://doi.org/10.1016/j.chembiol.2012.12.007 (2013).

33 Matson, D. R. & Stukenberg, P. T. Spindle poisons and cell fate: a tale of two pathways. *Molecular interventions* **11**, 141-150, doi:10.1124/mi.11.2.12 (2011).

34 Mukhtar, E., Adhami, V. M. & Mukhtar, H. Targeting Microtubules by Natural Agents for Cancer Therapy. *Mol. Cancer Ther.* **13**, 275, doi:10.1158/1535-7163.MCT-13-0791 (2014).

35 González-Cid, M., Larripa, I. & Slavutsky, I. Vinorelbine: cell cycle kinetics and differential sensitivity of human lymphocyte subpopulations1The authors are members of the Research Career of CONICET (National Research Council).1. *Toxicology Letters* **93**, 171-176, doi:https://doi.org/10.1016/S0378-4274(97)00089-1 (1997).

36 Arun, B., Akar, U., Gutierrez-Barrera, A. M., Hortobagyi, G. N. & Ozpolat, B. The PARP inhibitor AZD2281 (Olaparib) induces autophagy/mitophagy in BRCA1 and BRCA2 mutant breast cancer cells. *Int. J. Oncol.* **47**, 262-268, doi:10.3892/ijo.2015.3003 (2015).

37 Dieterich, D. C., Link, A. J., Graumann, J., Tirrell, D. A. & Schuman, E. M. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc. Natl. Acad. Sci.* **103**, 9482, doi:10.1073/pnas.0601637103 (2006).

38 Dieterich, D. C. *et al.* In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat. Neurosci.* **13**, 897-905, doi:10.1038/nn.2580 (2010).

39 Chehrehasa, F., Meedeniya, A. C. B., Dwyer, P., Abrahamsen, G. & Mackay-Sim, A. EdU, a new thymidine analogue for labelling proliferating cells in the nervous system. *J. Neurosci. Methods* **177**, 122-130, doi:https://doi.org/10.1016/j.jneumeth.2008.10.006 (2009).

40 MacKinnon, A. L. & Taunton, J. Target Identification by Diazirine Photo-Cross-Linking and Click Chemistry. *Curr. Protoc. Chem. Biol.* **1**, 55-73, doi:10.1002/9780470559277.ch090167 (2009).

41 Carelli, J. D. *et al.* Ternatin and improved synthetic variants kill cancer cells by targeting the elongation factor-1A ternary complex. *eLife* **4**, e10222, doi:10.7554/eLife.10222 (2015).

42 Blanch, A., Robinson, F., Watson, I. R., Cheng, L. S. & Irwin, M. S. Eukaryotic Translation Elongation Factor 1-Alpha 1 Inhibits p53 and p73 Dependent Apoptosis and Chemotherapy Sensitivity. *PLOS ONE* **8**, e66436, doi:10.1371/journal.pone.0066436 (2013).

43 Tash, J. S. *et al.* Gamendazole, an orally active indazole carboxylic acid male contraceptive agent, targets HSP90AB1 (HSP90BETA) and EEF1A1 (eEF1A), and stimulates Il1a transcription in rat Sertoli cells. *Biol. Reprod.* **78**, 1139-1152, doi:10.1095/biolreprod.107.062679 (2008).

44 Marco, E., Martín-Santamaría, S., Cuevas, C. & Gago, F. Structural Basis for the Binding of Didemnins to Human Elongation Factor eEF1A and Rationale for the Potent Antitumor Activity of These Marine Natural Products. *J. Med. Chem.* **47**, 4439-4452, doi:10.1021/jm0306428 (2004).

45 Leisch, M., Egle, A. & Greil, R. Plitidepsin: a potential new treatment for relapsed/refractory multiple myeloma. *Future Oncology* **15**, 109-120, doi:10.2217/fon-2018-0492 (2018).

46 Krastel, P. *et al.* Nannocystin A: an Elongation Factor 1 Inhibitor from Myxobacteria with Differential Anti-Cancer Properties. *Angew. Chem.* **54**, 10149-10154, doi:10.1002/anie.201505069 (2015).

47 Dalsgaard, P. W., Larsen, T. O. & Christophersen, C. Bioactive Cyclic Peptides from the Psychrotolerant Fungus Penicillium algidum. *The Journal of Antibiotics* **58**, 141-144, doi:10.1038/ja.2005.16 (2005).

48 Naylor, M. R., Bockus, A. T., Blanco, M. J. & Lokey, R. S. Cyclic peptide natural products chart the frontier of oral bioavailability in the pursuit of undruggable targets. *Curr. Opin. Chem. Biol.* **38**, 141-147, doi:10.1016/j.cbpa.2017.04.012 (2017).

49 Ovadia, O. *et al.* Improvement of drug-like properties of peptides: the somatostatin paradigm. *Expert Opin. Drug Discovery* **5**, 655-671, doi:10.1517/17460441.2010.493935 (2010).

50 Todorovic, A., Holder, J. R., Scott, J. W. & Haskell-Luevano, C. Synthesis and activity of the melanocortin Xaa-d-Phe-Arg-Trp-NH tetrapeptides with amide bond modifications. *J Pept Res* **63**, 270-278, doi:10.1111/j.1399-3011.2004.00137.x (2004).

51 Bockus, A. T. *et al.* Going Out on a Limb: Delineating The Effects of β-Branching, N-Methylation, and Side Chain Size on the Passive Permeability, Solubility, and Flexibility of Sanguinamide A Analogues. *J. Med. Chem.* **58**, 7409-7418, doi:10.1021/acs.jmedchem.5b00919 (2015).

52 Villar, E. A. *et al.* How proteins bind macrocycles. *Nat. Chem. Biol.* **10**, 723-731, doi:10.1038/nchembio.1584 (2014).

53 Gregory, R. K. & Smith, I. E. Vinorelbine - A clinical review. *Br. J. Cancer* **82**, 1907-1913 (2000).

54 Bowden, G. T., Roberts, R., Alberts, D. S., Peng, Y.-m. & Garcia, D. Comparative Molecular Pharmacology in Leukemic L1210 Cells of the Anthracene Anticancer Drugs Mitoxantrone and Bisantrene. *Cancer Res.* **45**, 4915-4920 (1985).

55 Kamiyama, M. Mechanism of Action of Chromomycin A3: III . On the Binding of Chromomycin A3 with DNA and Physicochemical Properties of the Complex 1. *The Journal of Biochemistry* **63**, 566-572 (1968).

56 Kaziro, Y. & Kamiyama, M. Mechanism of Action of Chromomycin A 3: II Inhibition of RNA Polymerase Reaction. *The Journal of Biochemistry* **62**, 424-429 (1967).

57 Sobell, H. M. Actinomycin and DNA transcription. *Proc. Natl. Acad. Sci.* **82**, 5328-5331 (1985).

58 Liaoo, L.-l., Kuochan, S. M. & Horwitz, S. B. Mode of action of the antitumor compound bruceantin, an inhibitor of protein synthesis. *Molecular Pharmocology* **12**, 167-176 (1976).

59 Jordan, M. A., Thrower, D. & Wilson, L. Mechanism of inhibition of cell proliferation by Vinca alkaloids. *Cancer Res.* **51**, 2212-2222 (1991).

60 Li, L. H. *et al.* Mechanism of action of didemnin B, a depsipeptide from the sea. *Cancer Lett.* **23**, 279-288 (1984).

61 Schiff, P. B., Fant, J. & Horwitz, S. B. Promotion of microtubule assembly in vitro by taxol. *Nature* **277**, 665-667 (1979).

62 Waring, M. J. & Wakelin, L. P. G. Echinomycin: a bifunctional intercalating antibiotic. *Nature* **252**, 653-657 (1974).

63 Bates, S. E. *et al.* Molecular targets in the National Cancer Institute drug screen. *J. Cancer Res. Clin. Oncol.* **121**, 495-500, doi:10.1007/BF01197759 (1995).

64 Kansy, M., Senner, F. & Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **41**, 1007-1010, doi:10.1021/jm970530e (1998).

65 Avdeef, A. The rise of PAMPA. *Expert Opin. Drug Metab. Toxicol.* **1**, 325-342, doi:10.1517/17425255.1.2.325 (2005).

66 Klein, V. G. *et al.* Understanding and improving the membrane permeability of VH032-based PROTACs. *ACS Med. Chem. Lett.*, doi:10.1021/acsmedchemlett.0c00265 (2020).

67 Ghose, A. K. & Crippen, G. M. Atomic Physicochemical Parameters for Three-Dimensional Structure-Directed Quantitative Structure-Activity Relationships I. Partition Coefficients as a Measure of Hydrophobicity. *J. Comput. Chem.* **7**, 565-577, doi:10.1002/jcc.540070419 (1986).

68 Cierpicki, T. & Otlewski, J. Amide proton temperature coefficients as hydrogen bond indicators in proteins. *J. Biomol. NMR* **21**, 249-261, doi:10.1023/A:1012911329730 (2001).

69 Over, B. *et al.* Impact of Stereospecific Intramolecular Hydrogen Bonding on Cell Permeability and Physicochemical Properties. *J. Med. Chem.* **57**, 2746-2754, doi:10.1021/jm500059t (2014).

70 Wang, C. K. *et al.* Rational design and synthesis of an orally bioavailable peptide guided by NMR amide temperature coefficients. *Proc Natl Acad Sci U S A* **111**, 17504-17509, doi:10.1073/pnas.1417611111 (2014).

71 Labute, P. LowModeMD—Implicit Low-Mode Velocity Filtering Applied to Conformational Search of Macrocycles and Protein Loops. *J. Chem. Inf. Model.* **50**, 792-800, doi:10.1021/ci900508k (2010).

72 Martínez, L., Andrade, R., Birgin, E. G. & Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **30**, 2157-2164, doi:10.1002/jcc.21224 (2009).

73 Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684-3690, doi:10.1063/1.448118 (1984).

74 Higo, J., Kamiya, N., Sugihara, T., Yonezawa, Y. & Nakamura, H. Verifying trivial parallelization of multicanonical molecular dynamics for conformational sampling of a polypeptide in explicit water. *Chem. Phys. Lett.* **473**, 326-329, doi:https://doi.org/10.1016/j.cplett.2009.03.077 (2009).

75 Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101, doi:10.1063/1.2408420 (2007).

76 Pettersson, M. & Crews, C. M. PROteolysis TArgeting Chimeras (PROTACs) — Past, present and future. *Drug Discovery Today: Technol.* **31**, 15-27, doi:https://doi.org/10.1016/j.ddtec.2019.01.002 (2019).

77 Maniaci, C. & Ciulli, A. Bifunctional chemical probes inducing protein–protein interactions. *Curr. Opin. Chem. Biol.* **52**, 145-156, doi:https://doi.org/10.1016/j.cbpa.2019.07.003 (2019).

78 Verma, R., Mohl, D. & Deshaies, R. J. Harnessing the Power of Proteolysis for Targeted Protein Inactivation. *Mol. Cell* **77**, 446-460, doi:https://doi.org/10.1016/j.molcel.2020.01.010 (2020).

79 Lu, J. *et al.* Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chem. Biol.* **22**, 755-763, doi:https://doi.org/10.1016/j.chembiol.2015.05.009 (2015).

80 Winter, G. E. *et al.* Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **348**, 1376, doi:10.1126/science.aab1433 (2015).

81 Zengerle, M., Chan, K.-H. & Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem. Biol.* **10**, 1770-1777, doi:10.1021/acschembio.5b00216 (2015).

82 Gadd, M. S. *et al.* Structural basis of PROTAC cooperative recognition for selective protein degradation. *Nat. Chem. Biol.* **13**, 514-521, doi:10.1038/nchembio.2329 (2017).

83 Bondeson, D. P. *et al.* Lessons in PROTAC Design from Selective Degradation with a Promiscuous Warhead. *Cell Chem. Biol.* **25**, 78-87.e75, doi:https://doi.org/10.1016/j.chembiol.2017.09.010 (2018).

84 Smith, B. E. *et al.* Differential PROTAC substrate specificity dictated by orientation of recruited E3 ligase. *Nat. Commun.* **10**, 131, doi:10.1038/s41467-018-08027-7 (2019).

85 Watt, G. F., Scott-Stevens, P. & Gaohua, L. Targeted protein degradation in vivo with Proteolysis Targeting Chimeras: Current status and future considerations. *Drug Discovery Today: Technol.* **31**, 69-80, doi:https://doi.org/10.1016/j.ddtec.2019.02.005 (2019).

86 Cantrill, C. *et al.* Fundamental aspects of DMPK optimization of targeted protein degraders. *Drug Discov. Today* **25**, 969-982, doi:10.1016/j.drudis.2020.03.012 (2020).

87 Lipinski, C. A. Lead profiling Lead- and drug-like compounds : the rule-of-five revolution. *Drug Discovery Today: Technol.* **1**, 337-341, doi:10.1016/j.ddtec.2004.11.007 (2004).

88 Edmondson, S. D., Yang, B. & Fallan, C. Proteolysis targeting chimeras (PROTACs) in 'beyond rule-of-five' chemical space: Recent progress and future challenges. *Bioorganic & Medicinal Chemistry Letters* **29**, 1555-1564, doi:https://doi.org/10.1016/j.bmcl.2019.04.030 (2019).

89 Maple, H. J., Clayden, N., Baron, A., Stacey, C. & Felix, R. Developing degraders: principles and perspectives on design and chemical space. *Medchemcomm* **10**, 1755-1764, doi:10.1039/c9md00272c (2019).

90 Peraro, L. *et al.* Cell Penetration Profiling Using the Chloroalkane Penetration Assay. *J. Am. Chem. Soc.* **140**, 11360-11369, doi:10.1021/jacs.8b06144 (2018).

91 Foley, C. A., Potjewyd, F., Lamb, K. N., James, L. I. & Frye, S. V. Assessing the Cell Permeability of Bivalent Chemical Degraders Using the Chloroalkane Penetration Assay. *ACS Chem. Biol.* **15**, 290-295, doi:10.1021/acschembio.9b00972 (2020).

92 Colletti, L. M. *et al.* Methods to measure the intracellular concentration of unlabeled compounds within cultured cells using liquid chromatography/tandem mass spectrometry. *Anal. Biochem.* **383**, 186-193, doi:https://doi.org/10.1016/j.ab.2008.08.012 (2008).

93 Gordon, L. J. *et al.* Direct Measurement of Intracellular Compound Concentration by RapidFire Mass Spectrometry Offers Insights into Cell Permeability. *J. Biomol. Screen.* **21**, 156-164, doi:10.1177/1087057115604141 (2015).

94 Mateus, A. *et al.* Prediction of intracellular exposure bridges the gap between target- and cell-based drug discovery. *Proc Natl Acad Sci U S A* **114**, E6231-E6239, doi:10.1073/pnas.1701848114 (2017).

95 Riching, K. M. *et al.* Quantitative Live-Cell Kinetic Degradation and Mechanistic Profiling of PROTAC Mode of Action. *ACS Chem. Biol.* **13**, 2758-2770, doi:10.1021/acschembio.8b00692 (2018).

96 Galdeano, C. *et al.* Structure-Guided Design and Optimization of Small Molecules Targeting the Protein–Protein Interaction between the von Hippel–Lindau (VHL) E3 Ubiquitin Ligase and the Hypoxia Inducible Factor (HIF) Alpha Subunit with in Vitro Nanomolar Affinities. *J. Med. Chem.* **57**, 8657-8663, doi:10.1021/jm5011258 (2014).

97 Frost, J. *et al.* Potent and selective chemical probe of hypoxic signalling downstream of HIF-α hydroxylation via VHL inhibition. *Nat. Commun.* **7**, 13312, doi:10.1038/ncomms13312 (2016).

98 Chan, K.-H., Zengerle, M., Testa, A. & Ciulli, A. Impact of Target Warhead and Linkage Vector on Inducing Protein Degradation: Comparison of Bromodomain and Extra-Terminal (BET) Degraders Derived from Triazolodiazepine (JQ1) and Tetrahydroquinoline (I-BET726) BET Inhibitor Scaffolds. *J. Med. Chem.* **61**, 504-513, doi:10.1021/acs.jmedchem.6b01912 (2018).

99 Maniaci, C. *et al.* Homo-PROTACs: bivalent small-molecule dimerizers of the VHL E3 ubiquitin ligase to induce self-degradation. *Nat. Commun.* **8**, 830, doi:10.1038/s41467-017-00954-1 (2017).

100 Xiang, T. X. & Anderson, B. D. The relationship between permeant size and permeability in lipid bilayer membranes. *J. Membr. Biol.* **140**, 111-122, doi:10.1007/BF00232899 (1994).

101 Doak, Bradley C., Over, B., Giordanetto, F. & Kihlberg, J. Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chem. Biol.* **21**, 1115-1142, doi:https://doi.org/10.1016/j.chembiol.2014.08.013 (2014).

102 Young, R. J., Green, D. V. S., Luscombe, C. N. & Hill, A. P. Getting physical in drug discovery II: the impact of chromatographic hydrophobicity measurements and aromaticity. *Drug Discov. Today* **16**, 822-830, doi:10.1016/j.drudis.2011.06.001 (2011).

103 Shultz, M. D. Two Decades under the Influence of the Rule of Five and the Changing Properties of Approved Oral Drugs. *J. Med. Chem.* **62**, 1701-1714, doi:10.1021/acs.jmedchem.8b00686 (2019).

104 Nielsen, D. S. *et al.* Improving on Nature: Making a Cyclic Heptapeptide Orally Bioavailable. *Angew. Chem.* **53**, 12059-12063, doi:10.1002/anie.201405364 (2014).

105 Soares, P. *et al.* Group-Based Optimization of Potent and Cell-Active Inhibitors of the von Hippel–Lindau (VHL) E3 Ubiquitin Ligase: Structure–Activity Relationships Leading to the Chemical Probe (2S,4R)-1-((S)-2-(1-Cyanocyclopropanecarboxamido)-3,3-dimethylbutanoyl)-4-hydroxy-N-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (VH298). *J. Med. Chem.* **61**, 599-618, doi:10.1021/acs.jmedchem.7b00675 (2018).

106 Ciulli, A., Maniaci, C., Hughes, S. J. & Testa, A. Small Molecules. Great Britain patent WO 2018/189554.

107 Roy, M. J. *et al.* SPR-Measured Dissociation Kinetics of PROTAC Ternary Complexes Influence Target Degradation Rate. *ACS Chem. Biol.* **14**, 361-368, doi:10.1021/acschembio.9b00092 (2019).

108 Testa, A., Hughes, S. J., Lucas, X., Wright, J. E. & Ciulli, A. Structure-Based Design of a Macrocyclic PROTAC. *Angew. Chem.* **59**, 1727-1734, doi:10.1002/anie.201914396 (2020).

109 Tallarico, J. A. *et al.* An Alkylsilyl-Tethered, High-Capacity Solid Support Amenable to Diversity-Oriented Synthesis for One-Bead, One-Stock Solution Chemical Genetics. *J. Comb. Chem.* **3**, 312-318, doi:10.1021/cc000107i (2001).

110 Van Molle, I. *et al.* Dissecting fragment-based lead discovery at the von Hippel-Lindau protein:hypoxia inducible factor 1α protein-protein interface. *Chem. Biol.* **19**, 1300-1312, doi:10.1016/j.chembiol.2012.08.015 (2012).

111 Sun, X. *et al.* PROTACs: great opportunities for academia and industry. *Signal Transduction and Targeted Therapy* **4**, 64, doi:10.1038/s41392-019-0101-6 (2019).

112 Sakamoto, K. M. *et al.* Protacs: chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *Proc Natl Acad Sci U S A* **98**, 8554-8559, doi:10.1073/pnas.141230798 (2001).

113 Schneekloth, J. S. *et al.* Chemical Genetic Control of Protein Levels: Selective in Vivo Targeted Degradation. *J. Am. Chem. Soc.* **126**, 3748-3754, doi:10.1021/ja039025z (2004).

114 Zou, Y., Ma, D. & Wang, Y. The PROTAC technology in drug development. *Cell Biochem. Funct.* **37**, 21-30, doi:10.1002/cbf.3369 (2019).

115 An, S. & Fu, L. Small-molecule PROTACs: An emerging and promising approach for the development of targeted therapy drugs. *EBioMedicine* **36**, 553-562, doi:10.1016/j.ebiom.2018.09.005 (2018).

116 Toure, M. & Crews, C. M. Small-Molecule PROTACS: New Approaches to Protein Degradation. *Angew. Chem.* **55**, 1966-1973, doi:10.1002/anie.201507978 (2016).

117 Mullard, A. Arvinas's PROTACs pass first safety and PK analysis. *Nat. Rev. Drug Discov.* **18**, 895, doi:10.1038/d41573-019-00188-4 (2019).

118 Churcher, I. Protac-Induced Protein Degradation in Drug Discovery: Breaking the Rules or Just Making New Ones? *J. Med. Chem.* **61**, 444-452, doi:10.1021/acs.jmedchem.7b01272 (2018).

119 Liu, X. *et al.* Assays and technologies for developing proteolysis targeting chimera degraders. *Future Med. Chem.* **12**, 1155-1179, doi:10.4155/fmc-2020-0073 (2020).

120 Ermondi, G., Vallaro, M. & Caron, G. Degraders early developability assessment: face-to-face with molecular properties. *Drug Discov. Today* **25**, 1585-1591, doi:https://doi.org/10.1016/j.drudis.2020.06.015 (2020).

121 Scott, D. E. *et al.* Systematic Investigation of the Permeability of Androgen Receptor PROTACs. *ACS Med. Chem. Lett.* **11**, 1539-1547, doi:10.1021/acsmedchemlett.0c00194 (2020).

122 Arnott, J. A. & Planey, S. L. The influence of lipophilicity in drug discovery and design. *Expert Opin. Drug Discovery* **7**, 863-875, doi:10.1517/17460441.2012.714363 (2012).

123 Waring, M. J. Lipophilicity in drug discovery. *Expert Opin. Drug Discovery* **5**, 235-248, doi:10.1517/17460441003605098 (2010).

124 Hann, M. M. & Simpson, G. L. Intracellular drug concentration and disposition – The missing link? *Methods* **68**, 283-285, doi:https://doi.org/10.1016/j.ymeth.2014.05.009 (2014).

125 Winiwarter, S. *et al.* Hydrogen bonding descriptors in the prediction of human in vivo intestinal permeability. *J. Mol. Graphics Model.* **21**, 273-287, doi:https://doi.org/10.1016/S1093-3263(02)00163-8 (2003).

126 Biron, E. *et al.* Improving Oral Bioavailability of Peptides by Multiple N-Methylation: Somatostatin Analogues. *Angew. Chem.* **47**, 2595-2599, doi:10.1002/anie.200705797 (2008).

127 Ovadia, O. *et al.* The Effect of Multiple N-Methylation on Intestinal Permeability of Cyclic Hexapeptides. *Mol. Pharm.* **8**, 479-487, doi:10.1021/mp1003306 (2011).

128 Rezai, T. *et al.* Conformational Flexibility, Internal Hydrogen Bonding, and Passive Membrane Permeability: Successful in Silico Prediction of the Relative Permeabilities of Cyclic Peptides. *J. Am. Chem. Soc.* **128**, 14073-14080, doi:10.1021/ja063076p (2006).

129 Thansandote, P. *et al.* Improving the passive permeability of macrocyclic peptides : Balancing permeability with other physicochemical properties. *Bioorganic & Medicinal Chemistry* **23**, 322-327, doi:10.1016/j.bmc.2014.11.034 (2015).

130 Cyrus, K. *et al.* Impact of linker length on the activity of PROTACs. *Molecular BioSystems* **7**, 359-364, doi:10.1039/c0mb00074d (2011).

131 Wang, Y., Jiang, X., Feng, F., Liu, W. & Sun, H. Degradation of proteins by PROTACs and other strategies. *Acta Pharmaceutica Sinica B* **10**, 207-238, doi:https://doi.org/10.1016/j.apsb.2019.08.001 (2020).

132 Ouellette, R. J. & Rawn, J. D. in *Organic Chemistry (Second Edition)* (eds Robert J. Ouellette & J. David Rawn) 665-710 (Academic Press, 2018).