

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Reactivity-Guided Isolation and Medicinal Chemistry of Marine Bacterial Natural Products

Permalink

<https://escholarship.org/uc/item/8vt637bt>

Author

Seiler, Grant Steven

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Reactivity-Guided Isolation and Medicinal Chemistry of Marine Bacterial Natural Products

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

in

Chemistry

by

Grant Steven Seiler

Committee in charge:

Professor Chambers Hughes, Chair
Professor Thomas Hermann, Co-Chair
Professor Edward Dennis
Professor William Gerwick
Professor Yitzhak Tor

2019

©
Grant Steven Seiler, 2019
All rights reserved

The Dissertation of Grant Steven Seiler is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California San Diego

2019

TABLE OF CONTENTS

Signature Page	iii
Table of Contents	iv
List of Figures	v
List of Schemes.....	vi
Acknowledgements.....	vii
Vita	viii
Abstract of the Dissertation.....	ix
Introduction	1
Chapter 1: “Nature’s combinatorial biosynthesis produces vatiamides A-F.”	12
Chapter 2: “An optimized tetrazine probe to detect isocyanide-containing natural products in extracts.”	97
Chapter 3: “Neolymphostin A Is a Covalent Phosphoinositide 3-Kinase (PI3K)/Mammalian Target of Rapamycin (mTOR) Dual Inhibitor That Employs an Unusual Electrophilic Vinylogous Ester.”	123
Chapter 4: “Progress toward the total synthesis of the lymphostins: Preparation of a functionalized tetrahydropyrrolo[4,3,2- <i>de</i>]quinoline and unusual oxidative dimerization.”	208
Conclusion.....	262

LIST OF FIGURES

Figure 1.1 Vatiamides (1-6) from <i>Moorea producens</i> ASI16Jul14-2 “ASI” isolated in this study	14
Figure 1.2 Vatiamide biosynthetic pathway vat in <i>Moorea producens</i> ASI16Jul14-2; biosynthetic genes for.....	16
Figure 1.3 Synthesis and testing of bromoazidocoumarin click probe. A) Synthetic route of 14 from 12	18
Figure 1.4 HPLC-UV of ASI crude extract with probe 14 added (blue) and without probe 14 (black) ...	19
Figure 1.5 Vatiamide pathway promiscuous dd interactions. A) VatM and VatP Cdd alignment	20
Figure 2.1 Representative isocyanide-containing natural products	99
Figure 2.2 The tetrazine-isocyanide [4+1] cycloaddition	100
Figure 2.3 Tetrazines 10-25 and isocyanides 26-29	102
Figure 2.4 Core tetrazine 30 , tetrazines 31-36 , and selected iminopyrazoles 32 + 26 and 35 + 26	103
Figure 3.1 Identification of neolymphostin as an electrophilic natural product. a) Cysteine-based probe 1	128
Figure 3.2 Reactivity of neolymphostin model electrophile 6 with glutathione at pH 7.4.....	129
Figure 3.3 The interaction of neolymphostin with the PI3Ks is covalent. a) Incubation studies with PI3K α	131
Figure 3.4 HDX-MS data mapped on the docked model of neolymphostin bound to PI3K α (pdbID:5SW8)	134
Figure 3.5 Competing reactivity of model electrophile 6 with <i>N</i> ^{α} -acetyl lysine and water at pH 8.1	135
Figure 3.6 Kinase selectivity of neolymphostin and blockage of AKT phosphorylation in vivo. A) At a 1 μ M.....	137
Figure 4.1 Structure of lymphostin (1a), the eponymous member of the lymphostin family of natural	211

LIST OF SCHEMES

Scheme 4.1 Retrosynthetic analysis of lymphostin (1)	212
Scheme 4.2 Synthesis of 1,3,4,5-tetrahydropyrrolo[4,3,2- <i>de</i>]quinoline 6 via intramolecular Ullman reaction	213
Scheme 4.3 Oxidation and nitration of quinoline 6 to 2-oxo-1,2-dihydropyrrolo-quinoline 12 . The ORTEP	214
Scheme 4.4 Oxidation of tetrahydropyrroloquinoline 6 to heterodimer 15 . The ORTEP drawing of 15	215

ACKNOWLEDGEMENTS

I would like to acknowledge my family for their constant support, my friends for their encouragement, and the many labmates and classmates I've worked with at UCSD, who made the whole trip truly enjoyable. Thank you all.

Chapter 1 is a reprint, in full, of published work. Nathan A. Moss, Grant S. Seiler, Tiago F. Leão, Gabriel Castro-Falcón, Lena Gerwick, Chambers C. Hughes, and William H. Gerwick. "Nature's combinatorial biosynthesis produces vatiamides A-F." *Angewandte Chemie International Edition* **2019**, *58*, 9027-9031. The dissertation author was a co-investigator and co-author of this paper.

Chapter 2 is a reprint, in full, of work that is in preparation for submission. Gabriel Castro-Falcón, Grant S. Seiler, and Chambers C. Hughes. "An optimized tetrazine probe to detect isocyanide-containing natural products in extracts." **2019**. The dissertation author was a co-primary investigator and author of this material.

Chapter 3 is a reprint, in full, of published work. Gabriel Castro-Falcón, Grant S. Seiler, Özlem Demir, Manoj K. Rathinaswamy, David Hamelin, Reece M. Hoffmann, Stefanie L. Makowski, Anne-Catrin Letzel, Seth J. Field, John E. Burke, Rommie E. Amaro, and Chambers C. Hughes. "Neolymphostin A Is a Covalent Phosphoinositide 3-Kinase (PI3K)/Mammalian Target of Rapamycin (mTOR) Dual Inhibitor That Employs an Unusual Electrophilic Vinylogous Ester." *Journal of Medicinal Chemistry* **2018**, *61*, 10463-10472. The dissertation author was a co-primary investigator and author of this paper.

Chapter 4 is a reprint, in full, of published work. Grant S. Seiler and Chambers C. Hughes. "Progress toward the total synthesis of the lymphostins: Preparation of a functionalized tetrahydropyrrolo[4,3,2-*de*]quinoline and unusual oxidative dimerization." *Journal of Organic Chemistry*. **2019**, *84*, 9339-9343. The dissertation author was the primary investigator and author of this paper.

VITA

2013	Bachelor of Arts, Carthage College
2013 – 2019	Teaching Assistant, University of California San Diego
2015	Master of Science, University of California San Diego
2019	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

Gabriel Castro-Falcón, Grant S. Seiler, and Chambers C. Hughes. "An optimized tetrazine probe to detect isocyanide-containing natural products in extracts." **2019**.

Grant S. Seiler and Chambers C. Hughes. "Progress toward the total synthesis of the lymphostins: Preparation of a functionalized tetrahydropyrrolo[4,3,2-*de*]quinoline and unusual oxidative dimerization." *Journal of Organic Chemistry*, **2019**.

Nathan A. Moss, Grant S. Seiler, Tiago F. Leão, Gabriel Castro-Falcón, Lena Gerwick, Chambers C. Hughes, and William H. Gerwick. "Nature's combinatorial biosynthesis produces vatiamides A-F." *Angewandte Chemie International Edition* **2019**, *58*, 9027-9031.

Gabriel Castro-Falcón, Grant S. Seiler, Özlem Demir, Manoj K. Rathinaswamy, David Hamelin, Reece M. Hoffmann, Stefanie L. Makowski, Anne-Catrin Letzel, Seth J. Field, John E. Burke, Rommie E. Amaro, and Chambers C. Hughes. "Neolymphostin A Is a Covalent Phosphoinositide 3-Kinase (PI3K)/Mammalian Target of Rapamycin (mTOR) Dual Inhibitor That Employs an Unusual Electrophilic Vinylogous Ester." *Journal of Medicinal Chemistry* **2018**, *61*, 10463-10472.

ABSTRACT OF THE DISSERTATION

Reactivity-Guided Isolation and Medicinal Chemistry of Marine Bacterial Natural Products

by

Grant Steven Seiler

Doctor of Philosophy in Chemistry

University of California San Diego, 2019

Professor Chambers Hughes, Chair
Professor Thomas Hermann, Co-Chair

Marine natural products have long served as leads for pharmaceutical compounds, as well as sources of compounds whose value lies in the novelty of their structures. As more natural products are discovered and described, compounds with truly novel structures or bioactivities have become increasingly difficult to find. Therefore, a reactivity-guided approach has been developed to aid in the search for natural products bearing specific functional groups. An azide-based probe for terminal alkyne-

bearing natural products was synthesized and implemented, which facilitated the discovery of the vatiamide family of natural products. Further, a suite of tetrazine-based probes for isonitrile-bearing natural products was synthesized and evaluated on the basis of reactivity, stability, and fragmentation behavior in mass spectrometry.

The marine bacterial natural product lymphostin was recognizable for its unusual structural features, but its biochemical activity, and mechanism thereof, was not fully understood. Lymphostin and its family of related analogs were evaluated for their kinase inhibition activity, and model systems were synthesized to understand lymphostin's mechanism of kinase inhibition. Lymphostin was found to be a potent and irreversible inhibitor of PI3K/mTOR, the first compound from bacteria to be shown with this activity. An effort towards the total synthesis of lymphostin was then undertaken, resulting in the discovery of an unusual heteroaryl dimerization reaction.

Introduction

Natural products, also referred to as secondary metabolites, are chemical compounds produced by living organisms. Common sources of natural products over the years have been plants^{1,2}, fungi³, bacteria⁴, and sometimes animals⁵⁻⁸. Two factors have drawn chemists to continue isolating and studying natural products: their complex chemical architectures^{9,10}, and their potential for use as drugs¹¹⁻¹⁵. A famous example of one compound that met both of those criteria is taxol. Isolated from the bark of the *taxus* tree, taxol not only showcases a topologically striking bridged and fused ring system, but also serves as a chemotherapy drug¹⁶⁻¹⁹. In fact, natural products and their direct derivatives account for 32% of FDA approved small-molecule drugs from 1981 to 2014²⁰; this portion does not include those structures that have borrowed natural product pharmacophores, or those that are considered natural product mimics. Further, it has been recognized that humans have long taken advantage of natural products in traditional medicine practices (e.g. chewing a leaf or making a tea), and that pursuing these practices might lead to the description of novel medicinal compounds^{21,22}.

Over time, natural products from marine sources have been discovered in ever higher number from ever more diverse source organisms²³⁻²⁵. Even as the catalog of known natural products continues to grow, researchers have continued to find chemical diversity and novelty^{26,27}. Indeed, the rate of discovery of natural products from marine sources remained high in recent decades²⁸. Notable marine natural products which have served as pharmaceuticals or as leads for the same include salinosporamide A^{29,30}, dolastatins 10 and 15³¹⁻³³, and halichondrin B³⁴. The study of marine natural products from microbial sources has several advantages. First, the incredible biodiversity of marine microbes bodes well for the diversity of compounds that they might produce. Second, because microbes can sometimes be cultured, metabolite production can be scaled up with the culture volume. In a field where many targets are produced in miniscule quantities, the ability to produce more sample without the need for further collection is enticing.

Longstanding methods for the detection of target compounds in extracts are UV/vis spectroscopy and bioassay-guided fractionation. The most abundant peak (usually measured by absorbance at 254 nm) in an HPLC chromatogram can often represent the most tractable compound to isolate from a crude extract. Likewise, extracts or fractions which have the strongest assayed bioactivity will be the most attractive for further purification, ultimately leading to the isolation of a single bioactive compound. While these methods have served and continue to serve researchers well, they have become somewhat limited as the catalog of characterized entities has grown. Re-discovery is an increasingly common event, and the rate of disclosure of genuinely novel frameworks (as opposed to additional members of natural product families) has stabilized²⁷. Chemists require ways to find potentially active compounds that are present in relatively low abundance, or those whose structures are truly unique, but do not lend themselves to detection by spectroscopy.

Reactivity-guided isolation has emerged recently as a method for detecting and isolating metabolites bearing specific chemical functionalities. The technique consists of adding a chemical probe to an extract or fraction, and allowing it to react with only specific functional groups, thereby covalently modifying the desired types of compounds. Features of the probe itself are designed to enhance detection of the resultant adduct by common methods of analysis. Strong UV absorbance can assist detection by spectroscopy in HPLC; a unique isotopic abundance can assist identification of adducts in mass spectrometry; fluorogenicity (“turn-on” fluorescence) can help rapidly identify fractions or extracts to further interrogate; and addition of non-polar organic groups can render highly polar natural products more lipophilic, making isolation easier.

Chemical probes have been developed to target natural products bearing a wide variety of reactive functional groups. These include alcohols³⁵, thiols³⁶⁻³⁸, amines^{36,37}, carboxylic acids^{36,37}, aldehydes and ketones^{36,37}, dehydroamino acids³⁹, and terminal alkynes^{40,41}. The alcohols, thiols, amines, and acids fall into the broad category of nucleophilic natural products, which have shared reactivity with

many primary metabolites, e.g. free amino acids. The aldehydes, ketones, and dehydroamino acids can be considered electrophilic natural products, which are chemically quite different than primary metabolites. The terminal alkyne functionality has perhaps the most subtle reactivity of those listed above, which allows for quite selective targeting of this excellent marker of secondary metabolism. Some of these probes are used as soluble substances, but many are supported on a resin, sometimes with a cleavable linker.

The Hughes lab has aimed to develop probes for compounds bearing functional groups that are most reflective secondary metabolism. These include electrophiles⁴² (to include β -lactones, β -lactams, epoxides, and α,β -unsaturated carbonyls), polyenes⁴³, and terminal alkynes⁴⁴. Screening for electrophilic natural products through the use of a thiol-based probe even helped to elucidate the provenance of the ammosamide family of natural products⁴⁵. The probes developed in our group are invariably employed as soluble small molecules, facilitating the isolation and characterization of adducts.

Microbial natural product biosynthesis proceeds by assembly line-type enzymatic machinery, which is expressed from genes that generally lie together in clusters. These are referred to as biosynthetic gene clusters (BGCs). The ordering of genes in a cluster is reflected in the ordering of biosynthetic enzymes in a pathway, and therefore in the order in which building blocks are incorporated into a growing secondary metabolite⁴⁶⁻⁴⁹. Therefore, genetic analysis of producing organisms can help predict the likely functional groups present in a natural product, or even illuminate the putative structure of a natural product⁵⁰⁻⁵³. This fact has led our group and others to selectively probe those organisms which bear genetic markers of particular functional groups. We have recently termed this approach *reactogenomics*. One elegant example of this approach was the application to organisms producing polyene-containing natural products⁴³.

Incorporation of the terminal alkyne function into a natural product requires desaturase machinery encoded by the *jamABC* cassette⁵⁴ or related machinery⁵⁵. This genetic signature can be, and

has been⁴⁴, identified by genome mining, and used to select strains for screening based upon their biosynthetic potential. Notable alkyne-bearing natural products include jamaicamides A and B⁵⁶, carmabin A⁵⁷, and ulongapeptin⁵⁸. The most obvious way to probe for terminal alkynes is to take advantage of the widely used copper-catalyzed azide-alkyne cycloaddition (CuAAC) reaction, the most well-known “click” reaction^{59–61}. The CuAAC is exquisitely selective and robust, proceeding efficiently in the presence of a wide variety of other functional groups. It is then unsurprising that several azide-based probes have been reported for the detection of alkyne-bearing natural products^{40,41}.

We decided to couple the advantageous reactivity of the azide function with the extreme analytical sensitivity of fluorescence, and implement a fluorogenic azide-based probe for terminal alkyne-bearing natural products. The use of such a probe had been previously reported⁶², but its synthesis was not. A strain of *Moorea producens* was shown through genome-mining to be capable of producing a family of related alkynylated metabolites, but traditional isolation methods failed to obtain all theoretical compounds. After probing the crude extract, not only were the expected adducts of the known family members detected by mass spectrometry, but so were those for the predicted analogs. Several adducts were isolated and characterized, helping to establish the vatiamides family of natural products⁴⁴ (Chapter 1).

Another organic functional group which is unique to secondary metabolism is the isonitrile. Again, these have the advantages of arising from known biosynthetic enzymes^{63–66}, and of undergoing a “click”-like cycloaddition with tetrazines^{67–69}. Notable isonitrile-bearing natural products include 7,20-diisocyanoadociane⁷⁰, hapalindole A⁷¹, and brasilidine A⁷². Instead of a fluorogenic approach, we decided to design tetrazine-based probes that would allow for rapid and sensitive detection of adducts by HPLC-MS. A suite of probes has been designed and synthesized, and their reactivity and MS fragmentation behavior has been evaluated (Chapter 2).

Lymphostin is a bacterial natural product isolated in 1997 from a *Streptomyces* sp⁷³. It bears two striking structural features. First, its core consists of an unusual pyrroloquinoline ring system. Second, the core is appended with a vinylogous methyl ester, an uncommon electrophile. While the biosynthesis of the tricyclic core from tryptophan is not fully understood, the provenance of the ester appendage is⁷⁴. Lymphostin was described as an immunosuppressant upon discovery (its name borrowing from the fact that it inhibits lymphocyte kinase)⁷⁵. The compound was later shown to inhibit PI3K as well⁷⁶. A close derivative, neolymphostin A, was later shown to inhibit several kinases related to cancer, including the PI3K family and mTOR (*vide infra*).

Members of the lymphostin family had been hits in prior probe screens for electrophilic natural products⁴². Structure elucidation of the resulting adduct revealed that the thiol probe had attacked the vinylogous in a 1,4-addition, prompting us to hypothesize that the lymphostins were not just potent, but also irreversible inhibitors of their kinase targets. Enzyme inhibitor dilution studies and mass spectrometry experiments helped to establish that the lymphostins covalently modify their kinase targets, and, further, that they do so at the ϵ -nitrogen of a conserved lysine in the PI3K family. Kinetics studies on a simplified model of lymphostins revealed that the family reacts quite rapidly with alkyl thiols at pH 7.1, but more slowly with primary amines at pH 8.1. Models with different electrophilic moieties failed to react with primary amines at all, suggesting that lymphostin's vinylogous ester function could be a unique handle for use in medicinal chemistry (Chapter 3).

Given lymphostin's potent kinase inhibition activity and unusual chemical structure, we found it to be a good target for total synthesis, with the intention of synthesizing analogs as well. While one total synthesis of lymphostin has already been reported⁷⁷, it suffers as a lengthy and linear sequence. We embarked on a campaign to develop a flexible route that would facilitate variation at the amide and vinylogous ester functions. While the synthesis of the natural product was not successful, a novel

method for construction of a pyrroloquinoline skeleton was reported, as well as novel oxidative heteroaryl dimerization chemistry (Chapter 4).

References

1. Georgiev MI. "From plants to pharmacy shelf: natural products revival." *Phytochemistry Reviews*. **2016**, *15*, 511-513.
2. Veeresham C. "Natural products derived from plants as a source of drugs." *Journal of Advanced Pharmaceutical Technology & Research*. **2014**, *3*, 200.
3. Schueffler A, Anke T. "Fungal natural products in research and development." *Natural Product Reports*. **2014**, *31*, 1425-1448.
4. Clardy J, Fischbach MA, Walsh CT. "New antibiotics from bacterial natural products." *Nature Biotechnology*. **2006**, *24*, 1541-1550.
5. Andersen RJ. "Sponging off nature for new drug leads." *Biochemical Pharmacology*. **2017**, *139*, 3-14.
6. König GM, Kehraus S, Seibert SF, Abdel-Lateff A, Müller D. "Natural products from marine organisms and their associated microbes." *ChemBioChem*. **2006**, *7*, 229-238.
7. Mehbub MF, Lei J, Franco C, Zhang W. "Marine sponge derived natural products between 2001 and 2010: Trends and opportunities for discovery of bioactives." *Marine Drugs*. **2014**, *12*, 4539-4577.
8. Calcabrini C, Catanzaro E, Bishayee A, Turrini E, Fimognari C. "Marine sponge natural products with anticancer potential: An updated review." *Marine Drugs*. **2017**, *15*, 1-34.
9. Hong J. "Role of natural product diversity in chemical biology." *Current Opinion in Chemical Biology*. **2011**, *15*, 350-354.
10. Skinnider MA, Magarvey NA. "Statistical reanalysis of natural products reveals increasing chemical diversity." *Proceedings of the National Academy of Sciences*. **2017**, *114*, E6271-E6272.
11. Clardy J, Walsh C. "Lessons from natural molecules." *Nature*. **2004**, *432*, 829-837.
12. Harvey AL. "Natural products in drug discovery." *Drug Discovery Today*. **2008**, *13*, 894-901.
13. Thomford NE, Senthebane DA, Rowe A, Munro D, Seele P, Maroyi A, and Dzobo K. "Natural products for drug discovery in the 21st century: Innovations for novel drug discovery." *International Journal of Molecular Sciences*. **2018**, *19*, 1-29.

14. Dias DA, Urban S, Roessner U. "A Historical overview of natural products in drug discovery." *Metabolites*. **2012**, *2*, 303-336.
15. Cheuka PM, Mayoka G, Mutai P, Chibale K. "The role of natural products in drug discovery and development against neglected tropical diseases." *Molecules*. **2017**, *22*, 1-41.
16. Wani MC, Horwitz SB. "Nature as a remarkable chemist." *Anti-Cancer Drugs*. **2014**, *25*, 482-487.
17. Kingston DGI. "Taxol, a molecule for all seasons." *Chemical Communications*. **2001**, 867-880.
18. Kingston DGI. "Recent advances in the chemistry of taxol." *Journal of Natural Products*. **2000**, *63*, 726-734.
19. Nicolaou KC, Guy RK, Potier P. "Taxoids: New Weapons." *Scientific American*. **1996**, *274*, 94-98.
20. Newman DJ, Cragg GM. "Natural Products as Sources of New Drugs from 1981 to 2014." *Journal of Natural Products*. **2016**, *79*, 629-661.
21. Efferth T, Li PCH, Konkimalla VSB, Kaina B. "From traditional Chinese medicine to rational cancer therapy." *Trends in Molecular Medicine*. **2007**, *13*, 353-361.
22. Yuan H, Ma Q, Ye L, Piao G. "The traditional medicine and modern medicine from natural products." *Molecules*. **2016**, *21*, 1-18.
23. Carroll AR, Copp BR, Davis RA, Keyzers RA, Prinsep MR. "Marine natural products." *Natural Product Reports*. **2019**, *36*, 122-173.
24. Costantino V, Fattorusso E, Menna M, Tagliatela-Scafati O. "Chemical Diversity of Bioactive Marine Natural Products: An Illustrative Case Study." *Current Medicinal Chemistry*. **2012**, *11*, 1671-1692.
25. Donia M, Hamann MT. "Marine natural products and their potential applications as anti-infective agents." *Lancet Infectious Diseases*. **2003**, *3*, 338-348.
26. Shen B. "A New Golden Age of Natural Products Drug Discovery." *Cell*. **2015**, *163*, 1297-1300.
27. Pye CR, Bertin MJ, Lokey RS, Gerwick WH, Linington RG. "Retrospective analysis of natural products provides insights for future discovery trends." *Proceedings of the National Academy of Sciences*. **2017**, *114*, 5601-5606.
28. Gerwick WH, Moore BS. "Lessons from the past and charting the future of marine natural products drug discovery and chemical biology." *Chemistry and Biology*. **2012**, *19*, 85-98.
29. Feling RH, Buchanan GO, Mincer TJ, Kauffman CA, Jensen PR, Fenical W. "Salinosporamide A: A highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus *Salinospora*." *Angewandte Chemie - International Edition*. **2003**, *42*, 355-357.

30. Fenical W, Jensen PR, Palladino MA, Lam KS, Lloyd GK, Potts BC. "Discovery and development of the anticancer agent salinosporamide A (NPI-0052)." *Bioorganic and Medicinal Chemistry*. **2009**, *17*, 2175-2180.
31. Beckwith M, Urba WJ, Longo DL. "Growth inhibition of human lymphoma cell lines by the marine products, dolastatins 10 and 15." *Journal of the National Cancer Institute*. **1993**, *85*, 483-488.
32. Pettit GR, Kamano Y, Dufresne C, Cerny RL, Herald CL, Schmidt JM. "Isolation and Structure of the Cytostatic Linear Depsipeptide Dolastatin 15." *Journal of Organic Chemistry*. **1989**, *54*, 6005-6006.
33. Pettit GR, Kamano Y, Herald CL, Tuinman A, Boettner F, Kizu H, Schmidt J, Baczynskyj L, Tomer K, and Bontems R. "The Isolation and Structure of a Remarkable Marine Animal Antineoplastic Constituent: Dolastatin 10." *Journal of the American Chemical Society*. **1987**, *109*, 6883-6885.
34. Hirata Y, Ljemura D. "Halichondrins—antitumor polyether macrolides from a marine sponge." *Pure & Appl. Chem*. **1986**, *58*, 701-710.
35. Odendaal AY, Trader DJ, Carlson EE. "Chemoselective enrichment for natural products discovery." *Chemical Science*. **2011**, *2*, 760-764.
36. Carlson EE, Cravatt BF. "Chemoselective probes for metabolite enrichment and profiling." *Nature Methods*. **2007**, *4*, 429-435.
37. Carlson EE, Cravatt BF. "Enrichment tags for enhanced-resolution profiling of the polar metabolome." *Journal of the American Chemical Society*. **2007**, *129*, 15780-15782.
38. Miles CO, Sandvik M, Nonga HE, Rundberget T, Wilkins A, Rise F, and Ballot A. "Thiol derivatization for LC-MS identification of microcystins in complex matrices." *Environmental Science and Technology*. **2012**, *46*, 8937-8944.
39. Cox CL, Tietz JI, Sokolowski K, Melby JO, Doroghazi JR, Mitchell DA. "Nucleophilic 1,4-additions for natural product discovery." *ACS Chemical Biology*. **2014**, *4*, 2014-2022.
40. Ross C, Scherlach K, Kloss F, Hertweck C. "The molecular basis of conjugated polyene biosynthesis in phytopathogenic bacteria." *Angewandte Chemie - International Edition*. **2014**, *53*, 7794-7798.
41. Jeon H, Lim C, Lee JM, Kim S. "Chemical assay-guided natural product isolation via solid-supported chemodosimetric fluorescent probe." *Chemical Science*. **2015**, *6*, 2806-2811.
42. Castro-Falcón G, Hahn D, Reimer D, Hughes CC. "Thiol Probes to Detect Electrophilic Natural Products Based on Their Mechanism of Action." *ACS Chemical Biology*. **2016**, *11*, 2328-2336.
43. Castro-Falcón G, Millán-Aguiñaga N, Roullier C, Jensen PR, Hughes CC. "Nitrosopyridine Probe to Detect Polyketide Natural Products with Conjugated Alkenes: Discovery of Novodaryamide and Nocarditrene." *ACS Chemical Biology*. **2018**, *13*, 3097-3106.

44. Moss NA, Seiler G, Leño TF, Castro-Falcón G, Gerwick L, Hughes C, and Gerwick WH. "Nature's Combinatorial Biosynthesis Produces Vatiamidines A-F." *Angewandte Chemie International Edition*. **2019**, *58*, 9027-9031.
45. Reimer D, Hughes CC. "Thiol-based probe for electrophilic natural products reveals that most of the ammosamidines are artifacts." *Journal of Natural Products*. **2017**, *80*, 126-133.
46. Fisch KM. "Biosynthesis of natural products by microbial iterative hybrid PKS-NRPS." *RSC Advances*. **2013**, *3*, 18228-18247.
47. Jensen PR. "Natural Products and the Gene Cluster Revolution." *Trends in Microbiology*. **2016**, *24*, 968-977.
48. Moore BS. "Biosynthesis of marine natural products: microorganisms (Part A)." *Natural Product Reports*. **2005**, *22*, 580-593.
49. Smanski MJ, Peterson RM, Huang SX, Shen B. "Bacterial diterpene synthases: New opportunities for mechanistic enzymology and engineered biosynthesis." *Current Opinion in Chemical Biology*. **2012**, *16*, 132-141.
50. Challis GL. "ChemInform Abstract: Genome Mining for Novel Natural Product Discovery." *ChemInform*. **2008**, *39*, 2618-2628.
51. Nett M, Ikeda H, Moore BS. "Genomic basis for natural product biosynthetic diversity in the actinomycetes." *Natural Product Reports*. **2009**, *26*, 1362-1384.
52. Boddy CN. "Bioinformatics tools for genome mining of polyketide and non-ribosomal peptides." *Journal of Industrial Microbiology and Biotechnology*. **2014**, *41*, 443-450.
53. Reen FJ, Romano S, Dobson ADW, O'Gara F. "The sound of silence: Activating silent biosynthetic gene clusters in marine microorganisms." *Marine Drugs*. **2015**, *13*, 4754-4783.
54. Haritos VS. "Biosynthesis: A terminal triple bond toolbox." *Nature Chemical Biology*. **2015**, *11*, 98-99.
55. Zhu X, Su M, Manickam K, Zhang W. "Bacterial Genome Mining of Enzymatic Tools for Alkyne Biosynthesis." *ACS Chemical Biology*. **2015**, *10*, 2785-2793.
56. Edwards DJ, Marquez BL, Nogle LM, McPhail K, Goeger DE, Roberts MA, and Gerwick WH. "Structure and biosynthesis of the jamaicamidines, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*." *Chemistry & biology*. **2004**, *11*, 817-833.
57. Hooper GJ, Orjala J, Schatzman RC, Gerwick WH. "Carmabins A and B, New Lipopeptides from the Caribbean Cyanobacterium *Lyngbya majuscula*." *Journal of Natural Products*. **1998**, *61*, 529-533.

58. Williams PG, Yoshida WY, Quon MK, Moore RE, Paul VJ. "Ulongapeptin, a cytotoxic cyclic depsipeptide from a Palauan marine cyanobacterium *Lyngbya* sp." *Journal of Natural Products*. **2003**, *66*, 651-654.
59. Kolb HC, Finn MG, Sharpless KB. "Click Chemistry: Diverse Chemical Function from a Few Good Reactions." *Angewandte Chemie - International Edition*. **2001**, *40*, 2004-2021.
60. Haldón E, Nicasio MC, Pérez PJ. "Copper-catalysed azide-alkyne cycloadditions (CuAAC): An update." *Organic and Biomolecular Chemistry*. **2015**, *13*, 9528-9550.
61. Rostovtsev V V., Green LG, Fokin V V., Sharpless KB. "A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes." *Angewandte Chemie - International Edition*. **2002**, *41*, 2596-2599.
62. Yang L, Chumsae C, Kaplan JB, Moulton KR, Wang D, Lee DH, and Zhou ZS. "Detection of Alkynes via Click Chemistry with a Brominated Coumarin Azide by Simultaneous Fluorescence and Isotopic Signatures in Mass Spectrometry." *Bioconjugate Chemistry*. **2017**, *28*, 2302-2309.
63. Wang L, Zhu M, Zhang Q, Zhang X, Yang P, Liu Z, Deng Y, Zhu Y, Huang X, Han L, Li S, and He J. "Diisonitrile Natural Product SF2768 Functions As a Chalkophore That Mediates Copper Acquisition in *Streptomyces thioluteus*." *ACS Chemical Biology*. **2017**, *12*, 3067-3075.
64. Brady SF, Clardy J. "Cloning and heterologous expression of isocyanide biosynthetic genes from environmental DNA." *Angewandte Chemie - International Edition*. **2005**, *44*, 7063-7065.
65. Brady SF, Bauer JD, Clarke-Pearson MF, Daniels R. "Natural products from *isnA*-containing biosynthetic gene clusters recovered from the genomes of cultured and uncultured bacteria." *Journal of the American Chemical Society*. **2007**, *129*, 12102-12103.
66. Harris NC, Sato M, Herman NA, Twigg F, Cai W, Liu J, Zhu X, Downey J, Khalaf R, Martin J, Koshino H, and Zhang W. "Biosynthesis of isonitrile lipopeptides by conserved nonribosomal peptide synthetase gene clusters in Actinobacteria." *Proceedings of the National Academy of Sciences*. **2017**, *114*, 7025-7030.
67. Stöckmann H, Neves AA, Stairs S, Brindle KM, Leeper FJ. "Exploring isonitrile-based click chemistry for ligation with biomolecules." *Organic and Biomolecular Chemistry*. **2011**, *9*, 7303-7305.
68. Stairs S, Neves AA, Stöckmann H, Wainman YA, Ireland-Zecchini H, Brindle KM, and Leeper FJ. "Metabolic glycan imaging by isonitrile-tetrazine click chemistry." *ChemBioChem*. **2013**, *14*, 1063-1067.
69. Imming P, Mohr R, Müller E, Overheu W, Seitz G. "[4 + 1]Cycloaddition of Isocyanides to 1,2,4,5-Tetrazines: A Novel Synthesis of Pyrazole." *Angewandte Chemie - International Edition*. **1982**, *21*, 284.
70. Baker JT, Wells RJ, Oberhänsli WE. "A New Diisocyanide of Novel Ring Structure from a Sponge." *Journal of the American Chemical Society*. **1976**, *98*, 4010-4012.

71. Moore RE, Cheuk C, Patterson GML. "Hapalindoles: new alkaloids from the blue-green alga *Hapalosiphon fontinalis*." *Journal of the American Chemical Society*. **2005**, *106*, 6456-6457.
72. Kobayashi J, Tsuda M, Nemoto A, Tanaka Y, Yazawa K, Mikami Y. "Brasilidine A, a new cytotoxic isonitrile-containing indole alkaloid from the actinomycete *Nocardia brasiliensis*." *Journal of Natural Products*. **1997**, *60*, 719-720.
73. AOTANI Y, NAGATA H, YOSHIDA M. "Lymphostin (LK6-A), a Novel Immunosuppressant from *Streptomyces* sp. KY11783: Structural Elucidation." *The Journal of Antibiotics*. **2012**, *50*, 543-545.
74. Miyanaga A, Janso JE, McDonald L, He M, Liu H, Barbieri L, Eustáquio A, Fielding E, Carter G, Jensen PR, Feng X, Leighton M, Koehn FE, and Moore BS. "Discovery and assembly-line biosynthesis of the lymphostin pyrroloquinoline alkaloid family of mtor inhibitors in *salinispora* bacteria." *Journal of the American Chemical Society*. **2011**, *133*, 13311-13313.
75. Nagata H, Ochiai K, Aotani Y, Ando K, Yoshida M, Takahashi I, and Tamaoki T. "Lymphostin (LK6-A), a Novel Immunosuppressant from *Streptomyces* sp. KY11783: Taxonomy of the Producing Organism, Fermentation, Isolation and Biological Activities." *The Journal of Antibiotics*. **1997**, *50*, 537-542.
76. Nagata H, Yano H, Sasaki K, Sato S, Nakanishi S, Takahashi I, and Tamaoki T. "Inhibition of lymphocyte kinase Lck and phosphatidylinositol 3-kinase by a novel immunosuppressant, lymphostin." *Bioscience, biotechnology, and biochemistry*. **2002**, *66*, 501-507.
77. Tatsuta K, Imamura K, Itoh S, Kasai S. "The first total synthesis of lymphostin." *Tetrahedron Letters*. **2004**, *45*, 2847-2850.

Chapter 1

“Nature’s Combinatorial Biosynthesis Produces Vatiamides A–F.”

Nathan A. Moss, Grant Seiler, Tiago F. Leão, Gabriel Castro-Falcón, Lena Gerwick, Chambers C. Hughes,
and William H. Gerwick

Abstract: Hybrid type I PKS/NRPS biosynthetic pathways typically proceed in a collinear manner wherein one molecular building block is enzymatically incorporated in a sequence that corresponds to gene arrangement. In this work, genome mining combined with the use of a fluorogenic azide-based click probe led to the discovery and characterization of vatiamides A–F, three structurally diverse alkynylated lipopeptides, and their brominated analogues, from the cyanobacterium *Moorea producens* ASI16Jul14-2. These derive from a unique combinatorial non-collinear PKS/NRPS system encoded by a 90 kb gene cluster in which an upstream PKS cassette interacts with three separate cognate NRPS partners. This is facilitated by a series of promiscuous intermodule PKS-NRPS docking motifs possessing identical amino acid sequences. This interaction confers a new type of combinatorial capacity for creating molecular diversity in microbial systems.

Numerous microbial species contain polyketide synthase/non ribosomal peptide synthetase (PKS/NRPS) biosynthetic gene clusters, which produce diverse modified lipopeptide natural products (NP).^[1] In cyanobacteria, the study of saxitoxin, cylindrospermopsin, curacin A, and apratoxin A pathways, among others, has provided much information on protein–protein interactions, novel gene function, and loading-module diversity in NP biosynthesis.^[2–8] In typical PKS/NRPS systems, activated acyl-CoAs and amino acids are sequentially coupled by individual PKS or NRPS modules. This occurs collinearly such that the order in which the genes are transcribed is coincident with the functioning of

proteins that selectively interact to elongate NPs with high fidelity. The interface between modules occurs via interacting type I or II docking domains (dd) between PKSs and communication (COM) domains between some NRPS modules.^[9] A third type of structurally distinct docking interaction between combinations of PKS and/or NRPS modules was characterized in tubulysin and rhabdopeptide/xenorpeptide (RXP) biosynthesis, and are referred to here as (C-terminus-N-terminus) β - $\alpha\beta\alpha\alpha$ type domains.^[10–12] We used a bioactivity and reactivity-guided approach to characterize six new lipopeptides and their biosynthetic pathway from the benthic cyanobacterium *Moorea producens* ASI6Jul14-2. This pathway features a unique nonselective combinatorial dd interaction that is antithetical to collinearity and represents a significant new mechanism by which prokaryotes may generate NP diversity.

An environmentally-derived culture designated ASI6Jul14-2 “ASI” from Vatia Bay, American Samoa, yielded a single non-axenic filamentous cyanobacterium after successive propagation and isolation of individual filaments in seawater BG-11 media. By 16S rRNA sequence comparison and morphology characteristics, it was identified as *Moorea producens* (Figure S1 and Table S12 in the Supporting Information).^[13] Brine shrimp toxicity and NCIH460 human lung cancer bioassay of fractionated extracts (Figure S2) from cultures of ASI led to the discovery and structure characterization of major compounds **1–4** by NMR and MS-based analysis methods (Figure 1, Table S2–S6, Figures S5–S25). Each molecule features an identical fatty acid tail that possesses a terminal alkyne or bromoalkyne, a secondary methyl group, and a vinyl chloride moiety (**7/8**) identical to that found in the jamaicamides **9/10**.^[14] This “tail” is extended with an *N,O,O*-trimethyl-L-tyrosine in **1/2**, whereas in **3/4**, it is appended by a glycine-derived amide and an extended bicyclic C10 polyketide that contains a hemiketal ring, secondary methyl group, methoxy group, and a terminating unsaturated d-lactone. The amino acid chirality of compound **1** was determined to be L (*S*) by Marfey’s analysis compared to an FDAA-derived standard of *N,O,O*-Me-L-tyrosine (Figure S35). In compound **3**, stereocenters C16, C17, C18, C20, and

C22 were assigned as S^*, S^*, R^*, S^*, R^* . $^1\text{H-NMR}$ J -coupling values were key to assigning the relative configuration of the lactol ring. NOESY correlations from the lactol ring to the lactone, as well as molecular modeling, were used to infer the relative configuration at C22 (Figures S38–S40). The configuration of the C9 branching methyl group was predicted by sequence analysis of the VatK KR domain. This motif possesses an LDD loop, although it was modified to LSD in VatK, and the catalytic region site 5 lacks a P residue. Thus, the VatK KR was determined to be a “B1”-type (Figure S37),^[15] therefore, C9 was assigned to be R in **1–6**. Compound **2** was established by HR-ESI-MS and MS²-based comparison to **1** (Figure S11), while **4** was established using a combination of MS and ^1H and ^{13}C NMR data in comparison to **3** (Table S4, Figures S23–S25).

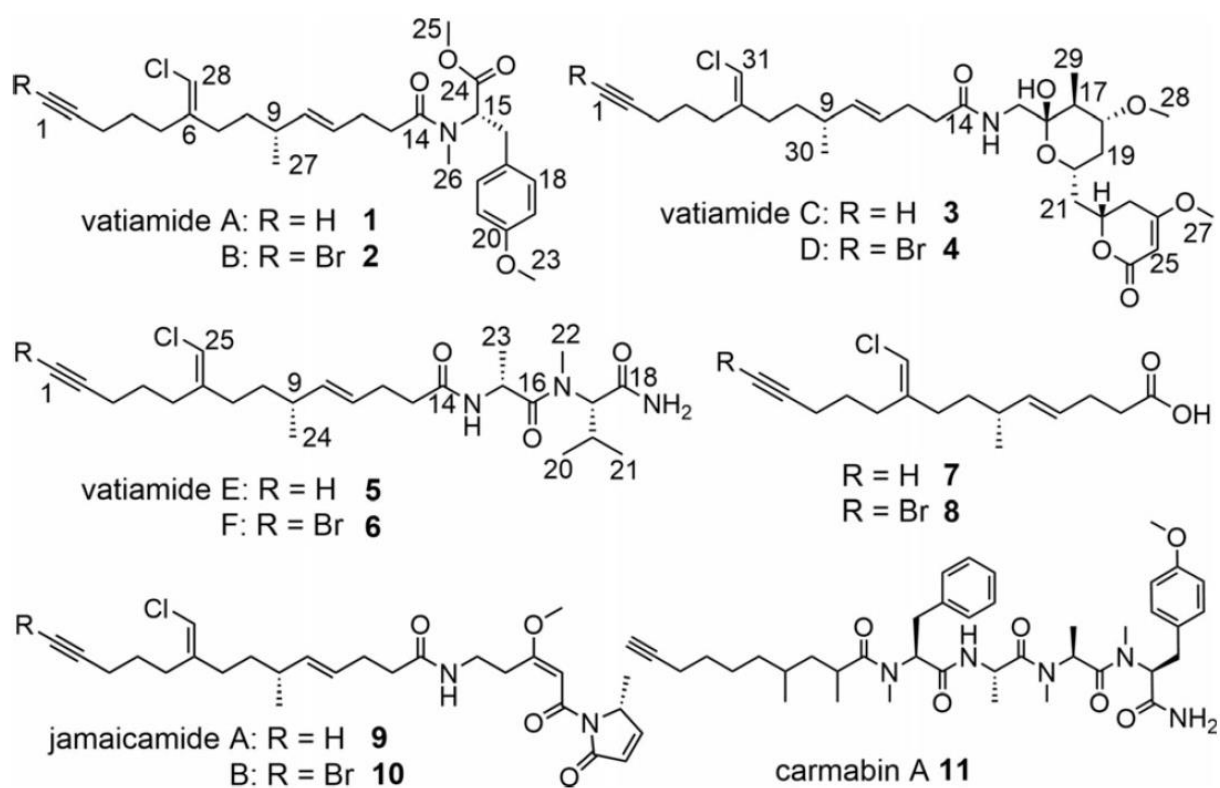


Figure 1.1. Vatiamides (**1–6**) from Moorea producers ASI16Jul14-2 “ASI” isolated in this study, jamaicamides (**9,10**) from Moorea producers JHB, and carmabin (**11**) from Moorea producers 3L. Compound **7** is produced by the organism in small quantity (Figure 4, S63). Free compound **8** was not observed.

Bioinformatic analysis of the ASI genome derived from short-read sequence data enabled assignment of the biosynthesis of **1/2**, but not **3/4** due to the fragmented nature of the assembly. The **1/2** biosynthetic pathway was evident in a 50 kb PKS/NRPS BGC *vatA-M*, which contained biosynthetic machinery homologous to that of the chloro fatty acid moiety of **9/10**, but diverging at *VatN*, an NRPS module with an adenylation (A) domain predicted to activate a tyrosine residue. This is followed by an in-module *N*-methyltransferase (N-MT) and *O*-methyltransferase (O-MT), peptidyl carrier protein (PCP), thioesterase (TE), and standalone O-MT enzyme *VatO*, in accordance with the structure of **1/2**. However, no biosynthetic genes in an orientation or domain structure matching that of **3/4** were found. Nevertheless, a second sequencing and assembly effort combining long and short reads illuminated the complete 90 kb *vat* pathway, with the original 50 kb pathway extended by 40 kb of NRPS and PKS modules immediately downstream of *VatO* (Figure 2, Table S11). *VatP*, a standalone ACP with high homology to the ACP of *VatM*, may play a role in transfer of **7/8** to subsequent NRPS modules. *VatQ* is an NRPS module predicted to incorporate Ala, and contains an epimerase domain I. Module *VatR* is an NRPS domain with A(Val), N-MT, and PCP motifs, and a terminal domain bearing homology to an NADPH-binding motif and putative amidotransferase of uncharacterized function. COM dds mediate the *VatQ/VatR* interaction. The *VatS* NRPS module contains an A(Gly), which is then followed by five PKS modules (*VatT-W*) and contains reduction and methylation domains required to biosynthesize **3/4**. Importantly, *VatT* contains an inactive KR^0 by active-site prediction, which likely enables spontaneous heterocyclization of the C20 hydroxy group with the C16 *VatS* glycine-derived carbonyl to form the lactol. A thioesterase (TE) embedded in *VatW* is believed to participate in an intramolecular cyclization to create the lactone observed in **3/4** by facilitating esterase activity on the terminal thioester C26. However, an additional uncharacterized gene, *Orf1*, possesses homology to various hydrolases, including cholyglycine hydrolase, penicillin V acylase, and *C-N* amide hydrolase, all of which catalyze hydrolytic attack of a carbonyl; thus, *Orf1* may be responsible for one or more of the heterocyclization events.

Finally, the O-MT VatX is predicted to methylate the enol functionality on the terminal δ -lactone. Owing to the exactly synchronous domain-to-structure accord between predicted and observed functionalities, we propose a non-collinear transfer of **7/8** from VatM to VatS.

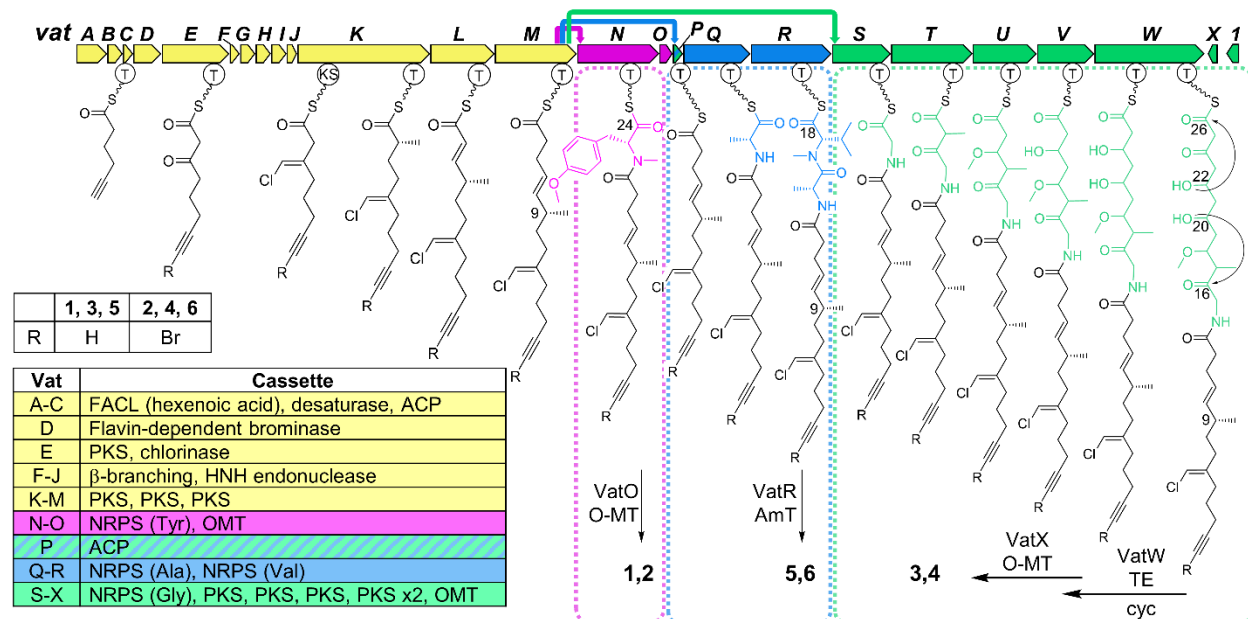


Figure 1.2. Vatiamide biosynthetic pathway *vat* in *Moorea producens* ASI16Jul14-2; biosynthetic genes for **7/8** in yellow. Three VatM/VatP cognate NRPS module partners and their subsequent PKS or NRPS cassettes highlighted as follows: pink **1/2**, green **3/4**, and blue **5/6**, with colored arrows at top indicating transmittal of **7/8** to VatN, VatP, and VatS, respectively. ACP; acyl carrier protein, AmT; amidotransferase, FACL; fatty-acid CoA ligase, OMT; O-methyltransferase, T; thiolase, TE; thioesterase (Table S11).

An alignment of the intermodule docking motifs revealed that the first 45 amino acids of the N-terminal NRPS docking motif (Ndd) on VatN, VatQ, and VatS are 100% identical, and that the final six residues of the C-terminal dd (Cdd) on VatM and VatP are identical as well. Sequence alignment of these dd matches previously structurally characterized β - $\alpha\beta\beta\alpha$ type dds.^[10, 11] Generally, since residues 24–29 on the β_2 sheet of the N-terminal NRPS dd confer interaction with the terminal 5–6 residues of the “ β_3 ” sheet of the upstream C-terminal dd, we theorized that there is no selective preference by the upstream C-terminal VatM/P dds to downstream N-terminal VatN/Q/S dds. Thus, VatM or VatP thioester-bound **7/8** may interact with either the VatN, VatQ, or VatS Ndd stochastically. This was validated by the

observed product composition wherein **7/8** appear in **1–4** but the “head” group of **1/2** (VatM→VatN transfer) differs from that of **3/4** (VatM→VatS transfer).

However, a conundrum resulted from the lack of an observed product from vatQ and vatR. To sensitively probe for these predicted products, we employed a reactivity-guided approach.^[16, 17] We synthesized fluorogenic “click”-based probe **14**, which undergoes a copper-catalyzed azide–alkyne cycloaddition (CuAAC) with terminal alkynes (Figure 3A, Schemes S1, S2, Table S9, Figures S43–S50).^[18–20] Incorporation of a bromine atom on the coumarin ring endows the probe with a characteristic isotopic pattern, thus aiding in identification of tagged compounds during LC/MS analysis of probed extracts.^[21] In addition, the probe exhibits “turn-on” fluorescence at 490 nm upon formation of a triazole adduct, thereby enabling benchtop screening for terminal alkyne-bearing NPs in crude extracts. Proof of concept was demonstrated by a model click reaction with alkyne propargyl benzoate at 90% yield, and subsequently reaction of **14** with purified **1** (Figure 3B–C, Scheme S3, Table S10, Figures S51–S60). Probing the extracts of the *Moorea producens* strains ASI, JHB (**9–10** producer),^[14] 3L (**11** producer),^[22, 23] PAL15Aug08-1 (no known alkynes), and *Moorea bouillonii* PNG5-198 (no known alkynes) clearly indicated the presence of alkynes in ASI, JHB, and 3L but not PNG5-198 or PAL15Aug08-1, which showed near baseline fluorescence in comparison to the alkyne producers (Figure 3D).

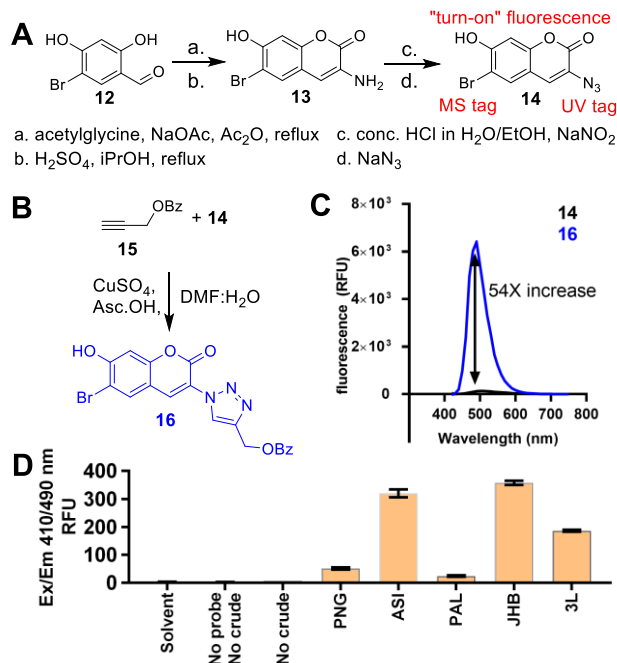


Figure 1.3. Synthesis and testing of bromoazidocoumarin click probe. A) Synthetic route of **14** from **12**. B) Synthesis of propargyl benzoate-**14** adduct. Asc.OH; ascorbic acid. C) Fluorescence increase upon formation of **16** from **15**+**14**. D) Screening Moorea extracts for alkyne with probe **14**.

Analysis of crude ASI-**14** reaction mixture by HPLC-UV-HRESI-MS/MS showed six major peaks by UV. By HRESI-MS, these peaks consisted of **13**, **14**, adducts **1**+**14**, **3**+**14**, **7**+**14**, and finally an unidentified sixth peak (**5**+**14**) with $[M+H]^+=747$, or 465 amu when a probe mass of 281 was subtracted (Figure 4, Figures S61–S63). Re-examination of unreacted ASI extract fractions revealed a minor peak with $[M+H]^+=544$ by LC/MS, thus suggesting a potential brominated analog of the 465 amu compound. This bromoalkyne was not evident in initial crude MS/MS runs due to its low abundance and poor ionization compared to **1**–**4**. Extraction of more biomass yielded sufficient material to characterize the $[M+H]^+=544$ compound (**6**), but only trace quantities of the 465 amu compound (**5**) were observed. 1D and 2D NMR revealed that partial structure **8** was linked via an amide to an alanyl residue followed by an *N*-methyl-valinamide residue, which defines compound **6**. By HRESI-MS and MS²-based comparison, the non-brominated compound **5** was also observed, thus reconciling the function of VatQ-VatR (Figures S26–S31, S34). Analysis of **6** in [D₆]DMSO indicated it was present in a 2:1 conformer ratio, likely due to

hindered rotation about the C22 *N*-Me amide. It also showed split resonances for the terminal amide protons likely due to H-bonding. The conformer ratio was reduced from 2:1 to 8:1 by $^1\text{H-NMR}$ analysis in d_3 -acetonitrile, while elevated temperature $^1\text{H-NMR}$ (80°C) resolved the split NH_2 resonance and conformers into single peaks (Figures S32, S33). Marfey's analysis of **6** indicated D-alanine and L-*N*-Me-valinamide, thus defining stereocenters 15 R and 17 S (Figure S36). We theorize that the terminal amide is formed by the undescribed domain embedded in VatR, downstream of the PCP. This domain bears homology to other amine-functionalizing enzymes, and is also found at the C-terminus of the carmabin (**11**) biosynthetic pathway in *Moorea producens* 3L; however, its enzymology and precise mechanism are unknown.^[23, 24]

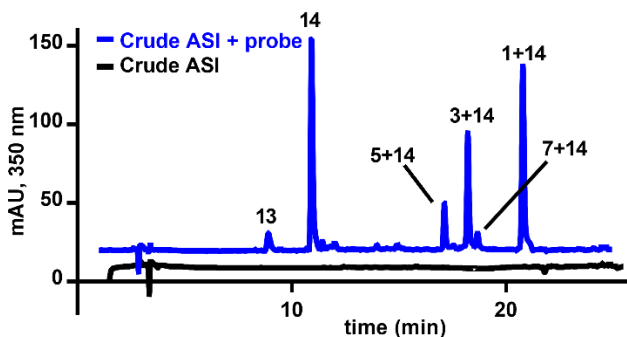


Figure 1.4. HPLC-UV of ASI crude extract with probe **14** added (blue) and without probe **14** (black).

The combinatorial biosynthesis of **1–6** reported herein is a unique and heretofore unknown method for generating chemodiversity with type I PKS/NRPS systems, and reveals that nature has selected for true combinatorial biosynthesis (Figure 5 and Figure S41). We postulate that this interaction is mediated by VatN, VatQ, and VatS N-terminal NRPS dds with identical sequence, which enable non-selective interaction with VatM and/or VatP ACP Cdd (Figure S64A). We do not anticipate module skipping to play a role, since VatO and VatR, the respective termini for the cassettes which generate **1/2** and **5/6**, do not contain C-terminal dds and have fully intact active sites. Therefore, the VatM-thioester bound **7/8** may be transferred directly to either VatN, VatQ, or VatS, thereby effectively forming three

separate assembly lines upon translation and tertiary folding (Figure S64B). Alternatively, it is possible that the small standalone VatP plays an intermediary role in the transfer of 7/8 to VatQ or VatS (Figure S64C). The GxDS phosphopantetheine (ppt) binding site of VatP ACP is found at the N-terminus of the protein; thus it is plausible that the VatP ppt arm could attack the VatM-thioester without an acyltransferase.

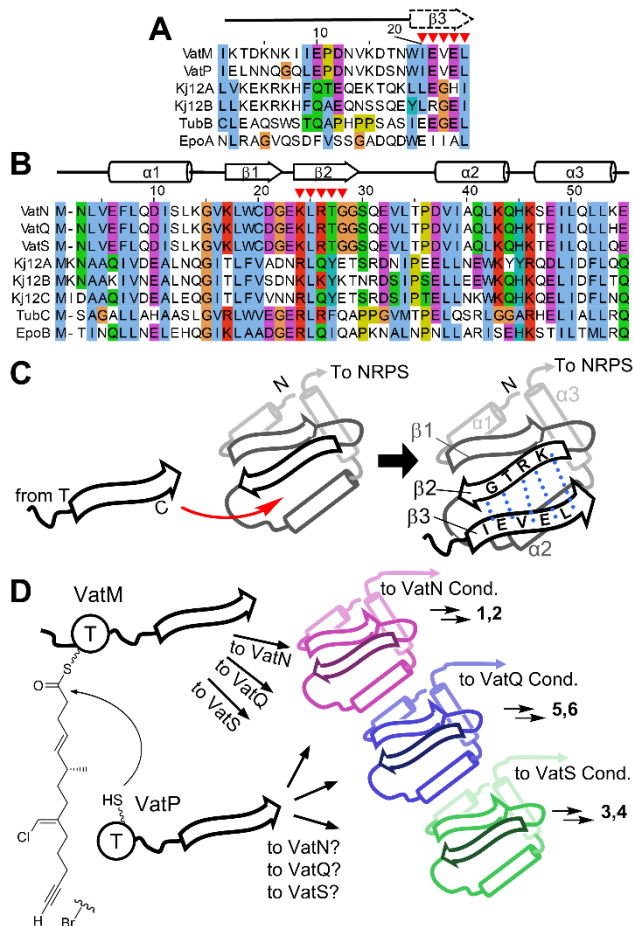


Figure 1.5. Vatiamide pathway promiscuous dd interactions. A) VatM and VatP Cdd alignment. B) VatN/Q/S Ndd aligned with characterized dds. Key interaction points in red triangles. C) Depiction of “docking”; electrostatic and steric interactions of residues in blue dots.^[10,11] D) Vat pathway promiscuity: T; thiolase, Cond; condensation domain.

There is evidence in multiple assembly-line biosynthetic pathways that significant evolutionary pressure has been placed on eliminating promiscuity, thereby ensuring that biosynthetic modules interact with their cognate dd partner through unique dd sequences.^[4,5] A BLAST search of the MIBiG

pathway database using the VatM/P Cdd and VatN/S/R Ndd revealed several pathways with at least two β - $\alpha\beta\beta\alpha$ dds between separate modules, yet no pathway features 100% homologous dds. This includes proteins other than PKSs and NRPSs, such as the standalone halogenase AerJ in aeruginosin biosynthesis, which possesses a β - $\alpha\beta\beta\alpha$ -type docking pair at both its C and N terminus (Figure S42).

Examples of stuttering,^[25] halting,^[26] out-of-order module processing,^[27] skipping,^[28–32] bidirectionality,^[33] and combinations thereof^[26] have been reported in various NP biosynthetic pathways. We report here the first characterized example to date of a native reversal of selectivity in a PKS/NRPS context. The promiscuity enabled by identical dds adds a new element to the combinatorial possibilities of both heterologous and native pathways. One may speculate that the origin in ASI is synergistic defensive toxicity imparted by the combination of more than one of the three head groups, or separate environmental targets for each. However we have only observed modest H-460 human lung cancer cytotoxicity with both pure compounds and combinations thereof, with the brominated analogs **2** and **4** being slightly more cytotoxic than the non-brominated versions **1** and **3** (Figures S3, S4).

Additional questions arise in understanding how such a protein structure is assembled over evolutionary time: either the docking motif was replicated and inserted in front of three pre-existing NRPS modules, or a single NRPS module with this dd was duplicated twice and each evolved different amino acid specificities followed by further tailoring divergence. The high homology between the VatM ACP and VatP suggests that duplication created VatP. Discovery of this full pathway only occurred after employing long-read and short-read sequencing with hybrid assembly. Assembler graphs may not be able to reconcile three nearly identical 150-bp sequences derived from short reads in a microbial genome without breaking continuity and generating “orphan” clusters. Therefore, we predict that dd-mediated combinatorial biosynthesis may be more widespread than currently recognized and may be present in numerous cyanobacterial and bacterial clades.

Acknowledgements: We thank J. G. Sanders, G. Humphrey, J. Gaffney, R. A. Salido Benitez, J. J. Minich, C. Brennan, K. Sanders, and R. Knight for their expertise in library preparation and genome sequencing. We thank E. Glukhov, P. Kanjanakantorn, M. Kissick, B. Ni, and A. M. Hoskins for their contributions to cyanobacterial culturing, bioassay, and extraction, and B. Miller for harvesting and collection. We acknowledge Y. Su, L. Gross, A. Mrse, and B. Duggan for analytical expertise. We thank M. Pierce and T. Murray for bioassay testing. This work was supported by NIH GM107550, NIH GM118815 to LG and WHG. NAM was supported by NIH Training Grant in Marine Biotechnology and Microbiome and Microbial Sciences Initiative Graduate Student Fellowship.

Chapter 1 is a reprint, in full, of published work. Nathan A. Moss, Grant S. Seiler, Tiago F. Leão, Gabriel Castro-Falcón, Lena Gerwick, Chambers C. Hughes, and William H. Gerwick. "Nature's combinatorial biosynthesis produces vatiamides A-F." *Angewandte Chemie International Edition* **2019**, *58*, 9027-9031. The dissertation author was a co-investigator and co-author of this paper.

References

- [1] M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3468 – 3496.
- [2] R. Kellmann, T. K. Mihali, J. J. Young, R. Pickford, F. Pomati, B. A. Neilan, *Appl. Environ. Microbiol.* **2008**, *74*, 4044 – 4053.
- [3] J. Muenchhoff, K. S. Siddiqui, A. Poljak, M. J. Raftery, K. D. Barrow, B. A. Neilan, *FEBS J.* **2010**, *277*, 3844 – 3860.
- [4] T. K. Mihali, R. Kellmann, J. Muenchhoff, K. D. Barrow, B. A. Neilan, *Appl. Environ. Microbiol.* **2008**, *74*, 716 – 722.
- [5] J. R. Whicher, S. S. Smaga, D. A. Hansen, W. C. Brown, W. H. Gerwick, D. H. Sherman, J. L. Smith, *Chem. Biol.* **2013**, *20*, 1340 – 1351.
- [6] Z. Chang, N. Sitachitta, J. V. Rossi, M. A. Roberts, P. M. Flatt, J. Jia, D. H. Sherman, W. H. Gerwick, *J. Nat. Prod.* **2004**, *67*, 1356 – 1367.
- [7] R. V. Grindberg, T. Isohey, D. Brinza, E. Esquenazi, R. C. Coates, W. Liu, L. Gerwick, P. C. Dorrestein, P. Pevzner, R. Lasken, et al., *PLoS One* **2011**, *6*, e18565.

- [8] M. A. Skiba, A. P. Sikkema, N. A. Moss, A. N. Lowell, M. Su, R. M. Sturgis, L. Gerwick, W. H. Gerwick, D. H. Sherman, J. L. Smith, *ACS Chem. Biol.* **2018**, *13*, 1640 – 1650.
- [9] A. Miyanaga, F. Kudo, T. Eguchi, *Nat. Prod. Rep.* **2018**, *35*, 1185 – 1209.
- [10] C. Hacker, X. Cai, C. Kegler, L. Zhao, A. K. Weickmann, J. P. Wurm, H. B. Bode, J. Wçhnert, *Nat. Commun.* **2018**, *9*, 4366.
- [11] C. D. Richter, D. Nietlispach, R. W. Broadhurst, K. J. Weissman, *Nat. Chem. Biol.* **2008**, *4*, 75 – 81.
- [12] S. E. O'Connor, C. T. Walsh, F. Liu, *Angew. Chem. Int. Ed.* **2003**, *42*, 3917 – 3921; *Angew. Chem.* **2003**, *115*, 4047 – 4051.
- [13] N. Engene, E. C. Rottacker, J. Kařtovský, T. Byrum, H. Choi, M. H. Ellisman, J. Komárek, W. H. Gerwick, *Int. J. Syst. Evol. Microbiol.* **2012**, *62*, 1171 – 1178.
- [14] D. J. Edwards, B. L. Marquez, L. M. Nogle, K. McPhail, D. E. Goeger, M. A. Roberts, W. H. Gerwick, *Chem. Biol.* **2004**, *11*, 817 – 833.
- [15] A. T. Keatinge-Clay, *Chem. Biol.* **2007**, *14*, 898 – 908.
- [16] G. Castro-Falc3n, N. Millán-Aguiñaga, C. Roullier, P. R. Jensen, C. C. Hughes, *ACS Chem. Biol.* **2018**, *13*, 3097 – 3106.
- [17] G. Castro-Falc3n, D. Hahn, D. Reimer, C. C. Hughes, *ACS Chem. Biol.* **2016**, *11*, 2328 – 2336.
- [18] K. Sivakumar, F. Xie, B. M. Cash, S. Long, H. N. Barnhill, Q. Wang, *Org. Lett.* **2004**, *6*, 4603 – 4606.
- [19] C. Ross, K. Scherlach, F. Kloss, C. Hertweck, *Angew. Chem. Int. Ed.* **2014**, *53*, 7794 – 7798; *Angew. Chem.* **2014**, *126*, 7928 – 7932.
- [20] X. Zhu, P. Shieh, M. Su, C. R. Bertozzi, W. Zhang, *Chem. Commun.* **2016**, *52*, 11239 – 11242.
- [21] L. Yang, C. Chumsae, J. B. Kaplan, K. R. Moulton, D. Wang, D. H. Lee, Z. S. Zhou, *Bioconjugate Chem.* **2017**, *28*, 2302 – 2309.
- [22] G. J. Hooper, J. Orjala, R. C. Schatzman, W. H. Gerwick, *J. Nat. Prod.* **1998**, *61*, 529 – 533.
- [23] A. C. Jones, E. A. Monroe, S. Podell, W. R. Hess, S. Klages, E. Esquenazi, S. Niessen, H. Hoover, M. Rothmann, R. S. Lasken, et al., *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8815 – 8820.
- [24] T. Leao, G. Castel3o, A. Korobeynikov, E. A. Monroe, S. Podell, E. Glukhov, E. E. Allen, W. H. Gerwick, L. Gerwick, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 3198 – 3203.
- [25] C. Olano, B. Wilkinson, C. Sánchez, S. J. Moss, R. Sheridan, V. Math, A. J. Weston, A. F. Braña, C. J. Martin, M. Oliynyk, et al., *Chem. Biol.* **2004**, *11*, 87 – 97.

- [26] A. C. Ross, Y. Xu, L. Lu, R. D. Kersten, Z. Shao, A. M. Al-Suwailem, P. C. Dorrestein, P. Qian, B. S. Moore, *J. Am. Chem. Soc.* **2013**, *135*, 1155 – 1162.
- [27] A. K. El-Sayed, J. Hothersall, S. M. Cooper, E. Stephens, T. J. Simpson, C. M. Thomas, *Chem. Biol.* **2003**, *10*, 419 – 430.
- [28] S. C. Wenzel, B. Kunze, G. Höfle, B. Silakowski, M. Scharfe, H. Blöcker, R. Müller, *ChemBioChem* **2005**, *6*, 375 – 385.
- [29] A. Ninomiya, Y. Katsuyama, T. Kuranaga, M. Miyazaki, Y. Nogi, S. Okada, T. Wakimoto, Y. Ohnishi, S. Matsunaga, K. Takada, *ChemBioChem*, **2016**, *17*, 1709 – 1712.
- [30] B. J. Beck, Y. J. Yoon, K. A. Reynolds, D. H. Sherman, *Chem. Biol.* **2002**, *9*, 575 – 583.
- [31] Z.-R. Li, J. Li, J.-P. Gu, J. Y. H. Lai, B. M. Duggan, W.-P. Zhang, Z.-L. Li, Y.-X. Li, R.-B. Tong, Y. Xu, et al., *Nat. Chem. Biol.* **2016**, *12*, 773 – 775.
- [32] H.-Y. He, H.-X. Pan, L.-F. Wu, B.-B. Zhang, H.-B. Chai, W. Liu, G.-L. Tang, *Chem. Biol.* **2012**, *19*, 1313 – 1323.
- [33] S. Li, X. Wu, L. Zhang, Y. Shen, L. Du, *Org. Lett.* **2017**, *19*, 5010 – 5013.

Supporting Information

Table of Contents

Experimental methods

1. Organism collection and culturing	30
2. DNA extraction and sequencing.....	30
3. Assembly and pathway analysis.....	31
4. Chemical extraction, fractionation, subfractionation and HPLC purification	32
5. NMR, mass spectrometry, and polarimetry.....	33
6. Computational chemical modeling	34
7. NCI H-460 MTT-stain bioassay and brine shrimp assay	34
8. Marfey's analysis of 1 and 6.....	35
9. Moorea 16S rRNA comparison.....	36
10. Synthesis of 3-amino-6-bromo-7-hydroxy-2H-chromen-2-one 13	36
11. Synthesis of probe 3-azido-6-bromo-7-hydroxy-2H-chromen-2-one 14	37
12. Synthesis of model propargyl benzoate click product 16.....	37
13. Fluorescence measurements and imaging	38

List of Schemes

Scheme S1: Synthesis of 3-amino-6-bromo-7-hydroxy-2H-chromen-2-one 13	39
Scheme S2: Synthesis of 3-azido-6-bromo-7-hydroxy-2H-chromen-2-one 14.....	39
Scheme S3: Synthesis of model propargyl benzoate click product 16	39

List of Tables

Table S1: Gradient conditions for semipreparative HPLC purification of 1-6	33
---	----

Table S2: Vatiamide A (1) NMR spectral data (500 MHz, CDCl ₃)	40
Table S3: Vatiamide A/B (1,2) mass spectrometry and optical rotation	41
Table S4: Vatiamide C (3) NMR spectral data (1H, 13C, HSQC, HMBC, COSY: 500 MHz, CDCl ₃ ; NOESY: 600 MHz, CDCl ₃). Vatiamide D (4) NMR spectral data (13C: 500 MHz, CDCl ₃).....	42
Table S5: Vatiamide C (3) NMR spectroscopy with key correlations (500 MHz, DMSO-d ₆)	44
Table S6: Vatiamide C, D (3, 4) mass spectrometry and optical rotation.....	45
Table S7: Vatiamide F (6) NMR spectroscopy with key correlations (500 MHz, DMSO-d ₆).....	46
Table S8: Vatiamide E, F (5, 6) mass spectrometry and optical rotation	47
Table S9: NMR spectral data for probe 14 (500 MHz, DMSO-d ₆)	47
Table S10: NMR spectral data for 14 + 1 adduct (500 MHz, d ₄ -methanol)	48
Table S11: Proposed functions of open reading frames in vat pathway.....	49
Table S12: Organisms and accession numbers for 16S sequence comparison	50

List of Figures

Figure S1. Analysis of 16S rRNA sequence and morphology of Moorea producens ASI16Jul14-2	51
Figure S2. Bioassay of fractionated extracts.....	52
Figure S3. NCI-H460 MTT fluorometric assay results for pure compounds 1-4, 6.	53
Figure S4. NCI-H460 MTT fluorometric assay results for mixtures of compounds 1-4, 6	54
Figure S5. High-resolution mass spectrometry of 1-6.	55
Figure S6. 1H-NMR spectrum of 1 (500 MHz, CDCl ₃).....	56
Figure S7. HSQC spectrum of 1 (500 MHz, CDCl ₃).....	56
Figure S8. HMBC spectrum of 1 (500 MHz, CDCl ₃).....	57
Figure S9. COSY spectrum of 1 (500 MHz, CDCl ₃).	57
Figure S10. NOESY spectrum of 1 (500 MHz, CDCl ₃).....	58

Figure S11. MS2 comparison of 1 and 2.	59
Figure S12. 1H-NMR spectrum of 3 (500 MHz, CDCl ₃).	60
Figure S13. 13C-NMR spectrum of 3 (500 MHz, CDCl ₃).	60
Figure S14. HSQC spectrum of 3 (500 MHz, CDCl ₃).	61
Figure S15. HMBC spectrum of 3 (500 MHz, CDCl ₃).	61
Figure S16. COSY spectrum of 3 (500 MHz, CDCl ₃).	62
Figure S17. NOESY spectrum of 3 (500 MHz, CDCl ₃).	62
Figure S18. 1H-NMR spectrum of 3 (500 MHz, DMSO-d ₆).	63
Figure S19. 13C-NMR spectrum of 3 (500 MHz, DMSO-d ₆).	63
Figure S20. HSQC spectrum of 3 (500 MHz, DMSO-d ₆).	64
Figure S21. HMBC spectrum of 3 (500 MHz, DMSO-d ₆).	64
Figure S22. NOESY spectrum of 3 (500 MHz, DMSO-d ₆).	65
Figure S23. 1H-NMR spectrum of 4 (500 MHz, CDCl ₃).	66
Figure S24. 13C-NMR spectrum of 4 (500 MHz, CDCl ₃).	66
Figure S25. MS2 comparison of 3 and 4.	67
Figure S26. 1H-NMR spectrum of 6 (500 MHz, DMSO-d ₆).	68
Figure S27. 13C-NMR spectrum of 6 (500 MHz, DMSO-d ₆).	68
Figure S28. HSQC spectrum of 6 (500 MHz, DMSO-d ₆).	69
Figure S29. HMBC spectrum of 6 (500 MHz, DMSO-d ₆).	69
Figure S30. COSY spectrum of 6 (500 MHz, DMSO-d ₆).	70
Figure S31. NOESY spectrum of 6 (500 MHz, DMSO-d ₆).	70
Figure S32. 1H-NMR spectrum of 6 in DMSO-d ₆ under increasing temperature.	71
Figure S33. 1H-NMR spectral comparison of 6 in DMSO-d ₆ and acetonitrile-d ₃	71
Figure S34. MS2 comparison of 5 and 6.	72

Figure S35. LC/MS chromatogram depicting Marfey's analysis of 1.	73
Figure S36. LC/MS chromatogram depicting Marfey's analysis of 6.	74
Figure S37. C9 stereochemistry and VatK KR analysis.	75
Figure S38. 3D structural analysis of 3.	76
Figure S39. Molecular modeling analysis with C22 stereocenter set to R* in compound 3.	77
Figure S40. Molecular modeling analysis with C22 stereocenter set to S* in compound 3.	78
Figure S41. Alignment of all PKS or NRPS docking domains in Vat pathway.	79
Figure S42. Docking domain alignments and selective PKS/NRPS pathways from MIBiG and NCBI.	80
Figure S43. 1H-NMR spectrum of 13 (500 MHz, DMSO-d6).	81
Figure S44. 13C-NMR spectrum of 13 (500 MHz, DMSO-d6).	81
Figure S45. HRMS spectrum of 13.	82
Figure S46. 1H NMR spectrum of 14 (500 MHz, DMSO-d6).	83
Figure S47. COSY NMR spectrum of 14 (500 MHz, DMSO-d6).	83
Figure S48. HSQC NMR spectrum of 14 (500 MHz, DMSO-d6).	84
Figure S49. HMBC NMR spectrum of 14 (500 MHz, DMSO-d6).	84
Figure S50. HRMS spectrum of 14.	85
Figure S51. 1H-NMR spectrum of 16 (500 MHz, DMSO-d6).	86
Figure S52. 13C-NMR spectrum of 16 (500 MHz, DMSO-d6).	86
Figure S53. HRMS spectrum of 16.	87
Figure S54. Probe and adduct spectroscopy.	88
Figure S55. 1H NMR of 14+1 azido coumarin probe adduct (500 MHz, d4-methanol).	89
Figure S56. COSY of 14+1 azido coumarin probe adduct (500 MHz, d4-methanol).	89
Figure S57. HSQC of 14+1 azido coumarin probe adduct (500 MHz, d4-methanol).	90
Figure S58. HMBC of 14+1 azido coumarin probe adduct (500 MHz, d4-methanol).	90

Figure S59. ROESY of 14+1 azido coumarin probe adduct (500 MHz, d4-methanol).....	91
Figure S60. HRMS spectrum of 14+1.	91
Figure S61. HRMS spectrum of 14+3.	92
Figure S62. HRMS spectrum of 14+5.	92
Figure S63. HRMS spectrum of 14+7.	93
Figure S64. Depiction of biosynthetic module organization.	94
References	95

1. Organism collection and culturing

A thumb-sized agglomeration of mat-forming cyanobacterial filaments was collected by hand on shallow rock substrate environs at a depth of 1-2 m in Vatia Bay, American Samoa, coordinates 14°14'45.8"S 170°40'24.3"W and given the environmental collection code ASI16Jul14-2. The filaments were transferred to breathable 30 mL culture flasks with local seawater until transport to laboratory. Repeated culturing of this environmental collection of >1 cyanobacterial species saw emergence of a single large diameter (60-100 μ M) filamentous cyanobacterium with thick rust-colored sheath and tightly stacked cells, matching the morphology of the genus *Moorea*.^[1] Single filaments of this strain were separated, drawn through agar to isolate the unicyanobacterial filament, and propagated in modified seawater BG-11 media containing a final concentration of 0.5 g/L NaNO₃.^[2] The organism was grown at 27-28°C in 16h/8h light/dark cycle under 5.5-10 μ mol photons/m²S and split for sub-culturing every 3-5 weeks. Repeated culturing and iterative scale-up of this species rendered enough biomass for chemical and genomic DNA extraction. Microscopy was performed on an Evos XL light microscope (Life Technologies) with 10X, 20X, and 40X objectives.

2. DNA extraction and sequencing

A loose clump of *Moorea* sp. ASI16Jul14-2 filaments approximately 2 cm in diameter was transferred to a Büchner funnel with filter paper and washed with DI water, then immediately transferred to mortar and pestle and pulverized under liquid nitrogen to a fine frozen powder. This powder was immediately transferred to a cold 2.0 mL Eppendorf tube on ice and immediately processed using the Qiagen bacterial gDNA isolation protocol (Qiagen), followed by cleanup using a G20 tip (Qiagen). Eluent was precipitated with isopropanol and then washed with EtOH, followed by drying in a laminar flow hood and reconstitution in 200 μ L of DNase free water.

For Illumina HiSeq sequencing, resulting gDNA was processed using a miniaturized version of the Kapa HyperPlus Illumina-compatible library prep kit (Kapa Biosystems), used for library generation. DNA

extracts were normalized to 5 ng total input per sample in an Echo 550 acoustic liquid handling robot (Labcyte Inc). A Mosquito HTS liquid-handling robot (TTP Labtech Inc was used for 1/10 scale enzymatic fragmentation, end-repair, and adapter-ligation reactions carried out using). Sequencing adapters were based on the iTru protocol,^[3] in which short universal adapter stubs are ligated first and then sample-specific barcoded sequences added in a subsequent PCR step. Amplified and barcoded libraries were then quantified by the PicoGreen assay and pooled in approximately equimolar ratios before being sequenced on an Illumina HiSeq 4000 instrument to >30X coverage.

For Pacific Biosciences long-read sequencing, prior to submission, gDNA was assessed for quality and quantity using an Agilent 4200 TapeStation with Genomic Tape and a Qubit (Thermo Fischer Scientific). Sequencing libraries were generated using SMRTbell Template Preparation Reagent Kits (Pacific Biosciences) following the 20 kb library protocol. Libraries were size selected to >6 kb using a PippinHT (Sage Sciences). Libraries were sequenced on a PacBio RS II sequencer (UCSD IGM Genomics Center, La Jolla, CA) via 4-hour movies using the DNA/Polymerase Binding Kit Version P6 V2 with C4 sequencing chemistry.

3. Assembly and pathway analysis

Initial Illumina HiSeq short-read output data was assembled via SPAdes 3.6.^[4] Darkhorse was used to bin out contigs containing non-cyanobacterial genes,^[5] followed by dissolution of contigs and reassembly of those containing only cyanobacterial genes. Submission to AntiSMASH revealed biosynthetic gene clusters and domain analysis, including the partial vat cluster, while NCBI DELTA PBLAST was used to further analyze individual modules for their domain composition, including the initial truncated **1/2** PKS/NRPS gene cluster.^[6] Hybrid assembly of the short and long-read data was performed by SPAdes 3.6 hybrid assembler tool yielding longer and more complete contigs. Pathways were analyzed in similar fashion as the short-read assembly via bioinformatics tools. PacBio-only contigs were mapped to the full vat cluster were found to span each pathway cassette with read depth of 4X –

43X, underpinning confidence in correct hybrid assembly. Average hybrid coverage of the vat pathway contig was 23.9X, and coverage of the vat pathway was not biased towards any single protein cassette. The putative VatR C-terminal aminotransferase domain as well as Orf1 were analyzed using Phyre2 protein prediction server.^[7]

4. Chemical extraction, fractionation, sub-fractionation and HPLC purification

Approximately 80 mL of packed wet biomass was collected over several growth harvests, rinsed with DI water over a Büchner funnel with glass fiber filter, and stored at -20°C until extraction. Biomass was extracted using successive rounds of 2:1 DCM:MeOH which were combined and dried by rotary evaporator. Crude extract was separated by vacuum liquid chromatography (VLC) consisting of hexane-packed TLC-grade (H) silica (Sigma-Aldrich) over a ceramic filter funnel. Fractions were eluted using gentle vacuum, from nonpolar to polar, in nine fractions (“A-I”) as follows: 100% hexane, 9:1 hexane:EtOAc, 8:2 hexane:EtOAc, 6:4 hexane:EtOAc, 4:6 hexane:EtOAc, 2:8 hexane:EtOAc, 100% EtOAc, 7:3 EtOAc:MeOH, 100% MeOH, which were dried via rotary evaporator.

Prior to individual compound purification, fractions were sub-fractionated in 10% stepwise MeCN/H₂O reverse phase gradients using 500 mg or 1 g Hypersep SPE cartridges (Agilent). Gradients and eluent % for each molecule pair are as follows: 1/2 70 to 100% acetonitrile in H₂O, elution at 90-100% acetonitrile, 3/4 50 to 100% acetonitrile in H₂O, elution at 70-90% acetonitrile, 5/6 50 to 100% acetonitrile in H₂O, elution at 50-70% acetonitrile. Sub-fractions were analyzed for 1-6 content by LC/MS as detailed in subsequent section. Sub-fractions were dried via rotary evaporator or N₂, lyophilization, reconstituted in MeCN/H₂O mixtures and submitted for semi-preparative HPLC.

Purification of individual compounds from fractions was performed by semi-preparative HPLC using a C18 Kinetex 5 µm 10x150 mm column using a reverse phase gradient of acetonitrile in H₂O, all solvents +0.1% (v/v) formic acid. HPLC gradient conditions are as below.

Table S1: Gradient conditions for semipreparative HPLC purification of **1-6**.

Compounds	1,2		3,4		5,6	
Fraction	D, E		G, H		H, I	
Flow rate (mL/min)	4.0		4.0		3.0	
UV monitor (nm)	209, 236, 276, 380		209, 228, 236, 276		209, 228, 254, 276	
Gradient	Time (min)	% MeCN	Time (min)	% MeCN	Time (min)	% MeCN
	0	80	0	65	0	80
	1	80	1	65	1	80
	5	99	16	99	22	99
	12	99	20.8	99	22.5	99
	12.5	80	21	99	29.4	80
	16.5	80	25	65	29.5	80

5. NMR, mass spectrometry, and polarimetry

¹³C-NMR and ¹H-NMR was carried out on a Varian VX500 (500 MHz – ¹H, 125 MHz – ¹³C). 2D experiments including HSQC, HMBC, COSY, NOESY, ROESY, and additional ¹H-NMR experiments were carried out on a JEOL ECZ500 500 MHz system (500 MHz – ¹H, 125 MHz – ¹³C). NOESY for **3** was also collected on a Varian NPA600 600 MHz system with 1.7 mm cryoprobe using 600 MHz for ¹H nuclei. All NMR measurements were taken at room temperature except for the analysis of **6**, which included a ¹H-NMR comparison in DMSO-d₆ at 20°C, 40°C, 60°C, and 80°C. References were made in respective experiments to solvent signals CHCl₃ (δ C = 77.16, δ H = 7.24), DMSO (δ C = 39.52, δ H = 2.50), and acetonitrile (δ C = 118.26, δ H = 1.94). HSQC, NOESY, and ROESY data from the JEOL system was processed using Delta software (JEOL), while HMBC, COSY, and ¹³C-NMR and ¹H-NMR data from both the NPA600, Varian VX500 and JEOL ECZ500 was processed in Mestrenova (Mestrelab).

Polarimetry values were obtained on a Jasco P-2000 polarimeter in a 1 dm chamber using a sodium lamp at 589 nm, at 25°C. Compounds were dissolved at 0.5 mg/mL (**2**), 1.0 mg/mL (**1**, **4**) or 2.0 mg/mL (**3**, **6**) in methanol and 130 μ L was used for each measurement. Ten readings were recorded and averaged for each compound, and subtracted from the mean of ten readings of a solvent blank for each sample.

Fractions, sub-fractions, and semi-preparative HPLC flow-through fractions were analyzed via HPLC-ESI-MS/MS on a Finnigan Surveyor HPLC coupled to an LCQ Advantage Max ion trap mass spectrometer (Thermo Fisher Scientific) using data-dependent acquisition. High-resolution mass spectrometry data of **1-6**, **14+5**, and **14+7** adduct was obtained on a 1200-series HPLC (Agilent) coupled to a 6530 qToF mass spectrometer (Agilent). Each system employed data dependent acquisition in positive ESI mode on a 4.6 mm x 100 mm Kinetex 5 μ m C18 column with a gradient of 50% acetonitrile to 99% acetonitrile in H₂O using a 3:1 split flow rate of 0.6 mL/min (0.15 mL/min to the MS source) on the LCQ system and 0.7 mL/min (unsplit) on the Agilent system. All solvents contained 0.1% (v/v) formic acid.

High-resolution data of **14+1** and **14+3** adducts were obtained via direct infusion into a LTQ Orbitrap XL mass spectrometer in negative ESI mode with Orbitrap Fourier Transform mass analyzer (Thermo Fisher Scientific).

6. Computational chemical modeling

The preferred conformation of C22 of **3** was analyzed by energy minimization using MM2 and MMFF94 force field algorithms within Chem3D (PerkinElmer).^[8,9]

7. NCI H-460 MTT-stain bioassay and brine shrimp assay

Cells from human lung cancer cell line NCI-H-460 were added at 3.33×10^4 cells/mL to a 96-well plate in Roswell Park Memorial Institute (RPMI) 1640 media, containing fetal bovine serum (FBS) and 1% penicillin/streptomycin. Cells were incubated for recovery O/N at 37°C in 5% CO₂ in 180 μ L per well. Two replicates of fractions C-I were tested at 1 μ g/mL and 10 μ g/mL. Pure compounds were initially dissolved in DMSO at 1 mg/mL, and diluted to initial working concentrations as follows: 41 μ M (**1**) – 2 dilution series of 3 replicates; 17.8 μ M (**2**) – 3 dilution series of 2 replicates; 35 μ M (**3**) – 2 dilution series of 3 replicates; 31 μ M (**4**) – 2 dilution series of 3 replicates; 29 μ M (**6**) – 2 dilution series of 3 replicates, 1 dilution series of two replicates. Synergistic initial working concentrations of combined pure compounds

were as follows: 40 μM (**1+2**) – 1 dilution series of 3 replicates; 90 μM (**1+2+6**) – 2 dilution series of 3 replicates; 72 μM (**2+4+6**) – 2 dilution series of 3 replicates; 60 μM (**4+6**) – 1 dilution series of 3 replicates; 60 μM (**1+3**) – 1 dilution series of 3 replicates. Working solutions were made through serial dilution eight times by a factor of 0.3164 in RPMI 1640 media without FBS, with 20 μL added to each well. An equal volume of RPMI 1640 media without FBS was added to wells designated as negative controls for each plate. Plates were incubated for approximately 48 h before MTT staining. Doxorubicin was used as a positive control, while an equivalent volume of DMSO was used as negative control. Plates were read at 570 and 630 nm using a SpectraMax M2 microplate reader (Molecular Devices) to determine cell viability.^[10]

Toxicity towards brine shrimp (*Artemia salina*) was assayed as per previous research.^[11–13] Brine shrimp eggs were incubated to hatching in a dark chamber for 48 h, after which 20–40 live brine shrimp in ~0.25–0.5 mL artificial seawater were transferred to 4.5 mL of artificial seawater plus 3 or 30 $\mu\text{g}/\text{mL}$ chemical extract fraction, initially dissolved in DMSO. Conditions were run in duplicate. After 24 h at 27–28 °C, the number of dead brine shrimp was counted. Acetone was used to kill the remaining brine shrimp, and % mortality was generated by dividing initial dead by final dead quantity of brine shrimp.

8. Marfey's analysis of **1** and **6**

A standard of *N,O,O*-L-tyrosine was synthesized as previously described.^[14,15] 0.5 mg of **1** was dissolved in 1 mL 6N HCl in a glass pear flask, sealed with parafilm, and incubated at 110°C O/N for 16 h. The flask was cooled and dried via rotary evaporator. The contents were resuspended in 2 mL of 1M NaHCO_3 , and 1 mL each was transferred to two separate 2.0 mL glass vials. Marfey's reagent L-FDAA (1-fluoro-2-4-dinitrophenyl-5-L-alanine amide) and D-FDAA were separately dissolved in acetone and added in 4:1 molar excess to the two separate vials, as well as to two separate vials of *N,O,O*-L-tyrosine dissolved in 1 mL 1M NaHCO_3 . The vials were incubated at 40–50°C for one hour. 1N HCl was added dropwise until the solution was neutral by pH paper. The solution was diluted 10-fold in acetonitrile, and

dispensed to LC/MS vial with syringe and needle through a nylon filter. A similar procedure was repeated for 0.5 mg of **6**, using only D-FDAA. D-FDAA was also reacted with standards of L-alanine and D-alanine, as well as standards of *N*-Me-L-valine and *N*-Me-DL-valine for retention time comparison. HPLC/MS/MS conditions for Marfey's analysis employed a 4.6 mm x 100 mm Kinetex 5 μ m C18 column using a 3:1 split flow rate of 0.6 mL/min (0.15 mL/min to the MS source) attached to a Thermo LCQ Advantage Max ion trap mass spectrometer (Thermo Fisher Scientific), using a reverse phase gradient of acetonitrile in H₂O, each solvent with 0.1% formic acid (v/v). The gradient for analysis of **1** (% acetonitrile in H₂O): 0 min – 20%, 1 min – 20%, 46 min – 99%, 52 min – 99%, 52.5 min – 20%, 60 min – 20%. The gradient for analysis of **6** (% acetonitrile in H₂O): 0 min – 20%, 1 min – 20%, 40 min – 50%, 45 min – 50%, 46 min – 99%, 50 min – 99%, 51 min – 20%, 60 min – 20%.

9. *Moorea* 16S rRNA comparison and VatK KR sequence comparison

All alignments of docking domains, 16S rRNA sequences, and the VatK KR sequence comparison were performed using MUSCLE algorithm in Geneious (Biomatters).^[16] Alignment images were prepared using Jalview.^[17] Identity matrix reflects percent homology of nucleotide sequences spanning entire 16S rRNA sequence of selected *Moorea producens*.

10. Synthesis of 3-amino-6-bromo-7-hydroxy-2*H*-chromen-2-one **13**

To a round-bottom flask equipped with a magnetic stir bar was added 5-bromo-2,4-dihydroxybenzaldehyde (2.98 g, 13.7 mmol), *N*-acetylglycine (1.7 g, 14 mmol), and NaOAc (6.8 g, 41.1 mmol). Ac₂O (27 mL) was then added, and a drying tube was affixed to the flask. The mixture was stirred for 5 min at rt and then for 2 h at 150°C. The reaction mixture was then allowed to cool to room temperature, diluted with EtOAc, and washed with a saturated Na₂CO₃ solution, water, and finally brine. The organic phase was dried over Na₂SO₄, filtered, and concentrated. To the resulting orange solid was added 2-propanol (16 mL). With stirring, H₂SO₄ (3 mL) was added dropwise, and the mixture was then stirred for 1 h at 90°C. The solution was cooled to rt, diluted with EtOAc, and washed with saturated

NaHCO₃, water, and finally brine. The organic phase was dried over Na₂SO₄ and then evaporated onto Celite. The product was purified by flash chromatography (0 to 50% acetone in hexanes). **13** was recovered (412 mg, 11.7% yield, 2 steps) as an orange solid. UV/vis: λ_{\max} = 342 nm; IR (KBr, film): $\tilde{\nu}$ = 3422, 1712, 1684 cm⁻¹; ¹H NMR: δ 10.60 (s, 1H), 7.63 (s, 1H), 6.83 (s, 1H), 6.65 (s, 1H), 5.44 (s, 2H); ¹³C NMR: δ 159.1, 152.5, 148.6, 131.6, 128.5, 115.9, 108.4, 106.4, 103.4; HR-ESI-TOFMS: m/z [M – H]⁻ calcd for C₉H₅BrNO₃ 253.9458, found 253.9459.

11. Synthesis of probe 3-azido-6-bromo-7-hydroxy-2H-chromen-2-one **14**

To a stirring mixture of **13** (100 mg, 390 μ mol) in EtOH (2 mL) and water (2 mL), conc. Aq. HCl (2 mL) was added dropwise. The mixture was cooled to 0°C in an ice bath with stirring, and a solution of NaNO₂ (54 mg, 780 μ mol) in 0.5 mL H₂O was added dropwise. The mixture was kept at 0°C with stirring for 10 minutes, and then a solution of NaN₃ (76 mg, 1.2 mmol) in 0.5 mL H₂O was added dropwise. The mixture was stirred and allowed to slowly warm for 15 minutes, at which point it was diluted with EtOAc. The organic phase was washed successively with water, a sat. NaHCO₃ solution, water, and brine. The organic layer was then dried over Na₂SO₄, filtered, evaporated onto Celite, and purified by flash chromatography (0 to 50% EtOAc in hexanes). **14** was recovered (57 mg, 52% yield) as a tan solid. UV/vis: λ_{\max} = 344 nm; IR (KBr, film): $\tilde{\nu}$ = 3182, 2131, 1712, 1609 cm⁻¹; ¹H and 2D NMR: see Table S9; HR-ESI-TOFMS: m/z (M – H)⁻ calcd for C₉H₃⁷⁹BrN₃O₃ 279.9363, found 279.9361.

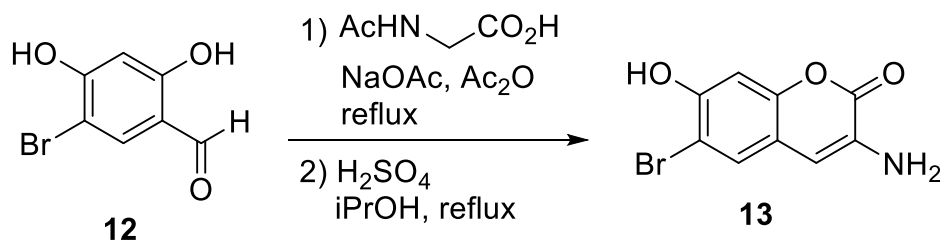
12. Synthesis of model propargyl benzoate click product **16**

To a stirring solution of probe **14** (10 mg, 35 μ mol) and propargyl benzoate (6 μ L, 40 μ mol) in 2 mL DMF and 2 mL water was added ascorbic acid (12 mg, 70 μ mol), followed by CuSO₄ (8.5 mg, 53 μ mol). The reaction was stirred for 1 hr at rt, by which point it had become cloudy and yellow and fluoresced strongly when exposed to a long-wavelength (365 nm) UV lamp. Analytical HPLC indicated consumption of the probe **14**, and the reaction was diluted with EtOAc. The organic phase was washed successively with water, a dilute NaHCO₃ solution, water, and brine. The organic phase was dried over

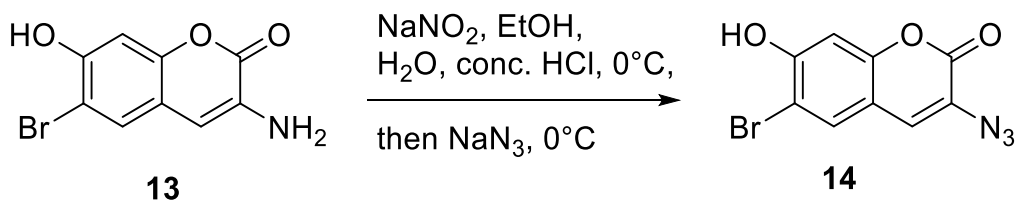
Na₂SO₄, filtered, and the solvent was removed under reduced pressure, yielding the product as a tan solid in sufficient purity for characterization (14 mg, 90% yield). UV/vis: $\lambda_{\text{max}} = 228 \text{ nm}, 348 \text{ nm}$; IR (KBr, film): $\tilde{\nu} = 3429, 3058, 2921, 1719, 1616 \text{ cm}^{-1}$; ¹H NMR: δ 8.71 (s, 1H), 8.60 (s, 1H), 8.14 (s, 1H), 7.98 (d, $J = 8 \text{ Hz}, 2\text{H}$), 7.67 (t, $J = 7 \text{ Hz}, 1\text{H}$), 7.53 (t, $J = 7 \text{ Hz}, 1\text{H}$), 6.99 (s, 1H), 5.51 (s, 2H); ¹³C NMR: δ 165.5, 158.6, 156.0, 153.5, 142.2, 135.0, 133.6, 133.0, 129.3, 128.9, 125.9, 120.1, 111.6, 107.2, 102.9, 57.7; HR-ESI-TOFMS: m/z (M + Na)⁺ calcd for C₁₉H₁₂⁷⁹BrN₃NaO₅ 463.9853, found 463.9850.

13. Fluorescence measurements and imaging

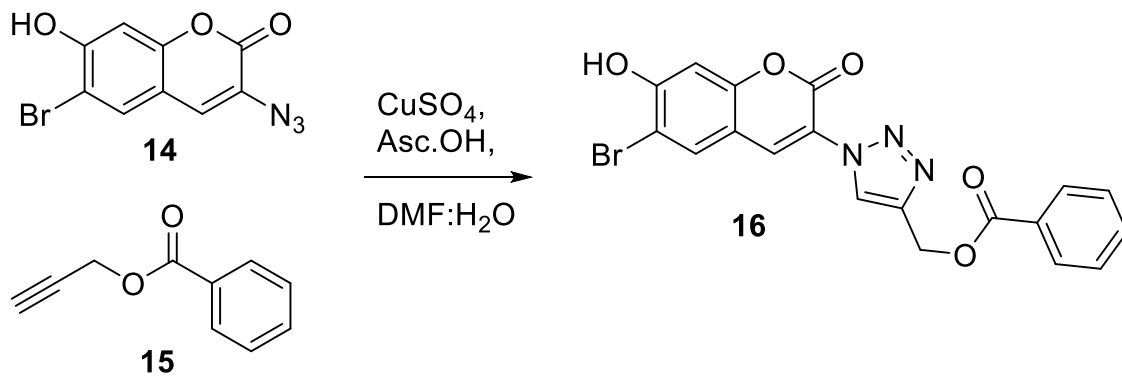
Small-scale crude extractions, without fractionation, were performed for fluorometric probe analysis. 2 x 10 mL of 2:1 DCM:MeOH was sequentially applied to a thumb-sized clump of harvested and rinsed biomass of *Moorea producens* ASI16Jul14-2 (ASI), *Moorea producens* 3L (3L), *Moorea producens* PAL15Aug08-1 (PAL), *Moorea producens* JHB (JHB), and *Moorea bouillonii* PNG5-198 (PNG). Extracts were dried and resuspended to 10 mg/mL in DMF. To carry out the click reaction, the following solutions were prepared: 1.0 mg/mL azidocoumarin probe in DMF, 2.8 mg/mL CuSO₄ in H₂O, 6.2 mg/mL ascorbic acid in H₂O. In this order, 10 μL of each of the above solutions was added to an LC vial insert: CuSO₄, then asc.OH, then crude, then probe with mixing by pipette. A negative control containing no crude or azide was created, whereby 10 μL of DMF blank was used in place of each of the omitted reagents. An additional negative control containing no crude was prepared, whereby 10 μL of DMF blank was added in place of a crude. After 1 hour, these solutions were transferred to wells in a 96-well plate, and fluorescence was measured on a SPECTRAMax M2 (Molecular devices). Excitation was at 410 nm, emission was measured at 490 nm.



Scheme S1: Synthesis of 3-amino-6-bromo-7-hydroxy-2*H*-chromen-2-one **13**

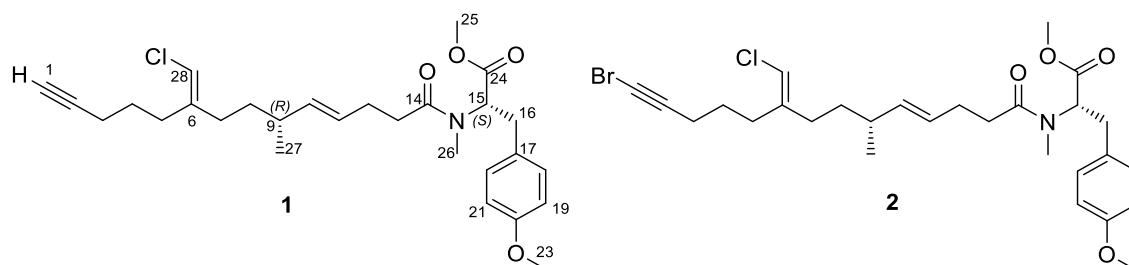


Scheme S2: Synthesis of 3-azido-6-bromo-7-hydroxy-2*H*-chromen-2-one **14**



Scheme S3: Synthesis of model propargyl benzoate click product **16**

Table S2. Vatiamide A (**1**) NMR spectral data (500 MHz, CDCl₃)



C#	δ_C	δ_H, J in Hz	Type	HMBC	COSY	key NOE
1	68.4	1.95	s		3	
2	84.2					
3	18.1	2.18	m	1, 2, 4, 5	1, 4, 5	
4	25.9	1.62	m	2, 3, 5, 6	2, 3, 5	
5	29.2	2.26	m	3, 4, 6, 7, 28	3, 4, 28	
6	141.9					
7	32.5	1.98	m	6, 8, 9, 28	8, 9, 28	28
8	34.7	1.28	m	6, 7, 9, 10, 27	7, 9, 27	28
9	36.3	1.98	m	7, 10, 11, 27	8, 10, 11, 27	
10	136.3	5.21	m	9, 11, 12	9, 11, 12, 27	
11	127.7	5.28	m	10, 12, 13	9, 10, 12	
12	27.6	2.18	m	10, 11, 13, 14	10, 11, 13	
13	33.4	2.28	m	10, 11, 12, 14	11, 12	
14	173					
15	58.2	5.25	m	14, 16, 17, 26	16, 25	
16	33.8	a – 3.27 (5.5, 14.8) b – 2.93 (11.3, 14.5)	dd dd	15, 17, 18/22, 24 15, 17, 18/22, 24	15, 17, 18/22 15, 17, 18/22	
17	128.9					
18/22	129.7	7.05 (7.9)	d	16, 17, 19/21, 20	16, 19/21	
19/21	113.9	6.78 (8.0)	d	17, 18/22, 20	18/22, 23	
20	158.4					
23	55	3.75	s	20		
24	171.7					
25	52.7	3.73	s	15, 24		
26	32.9	2.80	s	14, 15		
27	20.7	0.91 (6.0)	d	8, 9, 10	9	
28	112.9	5.76	s	4, 5, 6, 7	5, 7	7, 8

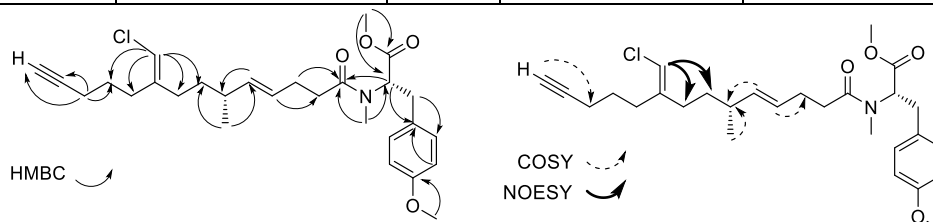
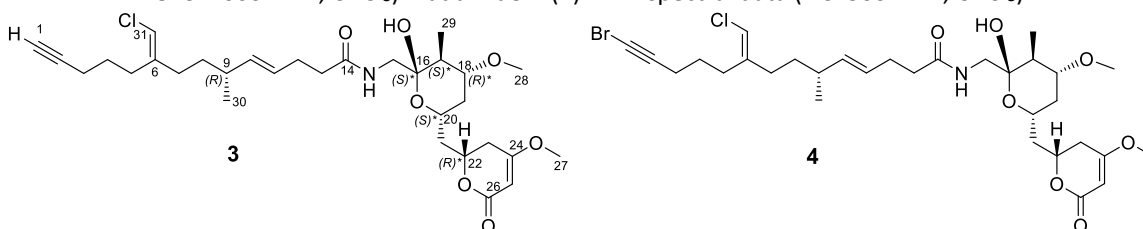


Table S3. Vatiamide A/B (**1,2**) mass spectrometry and optical rotation

Molecule, charge	1 , [M+H] ⁺	2 , [M+H] ⁺
Formula	C ₂₈ H ₃₉ ClNO ₄	C ₂₈ H ₃₈ BrClNO ₄
Calculated	488.2568	566.1673
Result	488.2557	566.1666
Δ ppm	-2.25	-1.24
[α] _D ²⁵ (c 0.1, CH ₃ OH)	7.0	-54.0

Table S4. Vatiamide C (**3**) NMR spectral data (^1H , ^{13}C , HSQC, HMBC, COSY: 500 MHz, CDCl_3 ; NOESY: 600 MHz, CDCl_3). Vatiamide D (**4**) NMR spectral data (^{13}C : 500 MHz, CDCl_3)



Molecule		3				
C#	δ_{C}	δ_{H} , J in Hz	type	HMBC	COSY	key NOE
1	68.8	1.96	t	3	3	
2	84.28					
3	18.5	2.2 (2.6, 7.2, 7.2)	dt	1, 2, 4, 5	1, 4	
4	26.2	1.63	m	2, 3, 5, 6	3, 5	
5	29.45	2.30	m	4, 6, 31	4	
6	142.02					
7	32.74	2.0	m	5, 6, 8, 9, 31	8	31
8	34.85	1.33	m	6, 7, 9, 10, 30	7, 9	31
9	36.45	2.0	m	7, 8, 10, 11, 30	7, 8, 10	
10	137.16	5.29 (7.5, 15.3)	dd	8, 9, 11, 12, 30	9, 11, 12	
11	127.45	5.34	m	9, 10, 12	10, 12	
12	28.74	2.30	m	10, 11, 13	10, 11, 13	
13	36.21	2.31	m	11, 12, 14	11, 12, 13	
14	175.26					
15	48	a – 3.88 (7.4, 13.8) b – 2.91 (4.8, 13.8)	dd dd	14, 16, 17, 18 14, 16, 17	15b, 20, NH 15a, 20, NH	29, NH 19, NH
16	98.56					
17	43.68	1.38	m	15, 16, 18, 19, 29	18, 29	19a, 15b
18	78.09	3.25 (4.5, 10.7, 10.7)	dt	16, 17, 20, 28, 29	19, 29	17, 19b
19	36.98	a – 1.12 (12, 12, 12) b – 2.05	ddd m	17, 18, 20, 21 17, 18, 20	18, 19b, 20 18, 19a, 20	17, 18, 20, 22 18, 20, 28
20	65.35	4.07 (2.5, 9.3, 2.5, 9.3)	tt	16, 18, 21	19, 21	19, 21, 22, 23
21	40.14	a – 2.08 b – 1.66 (3, 5.4)	m dd	19, 20, 22, 23 19, 20, 22, 23	20, 21b, 22 20, 21a, 22	22 20, 22
22	74.3	4.54	m	20, 21, 23, 24, 26	21, 23	19, 20, 21, 23
23	32.63	a – 2.48 (4.9, 17.2) b – 2.41 (9.3, 17.2)	dd	21, 22, 24, 25, 26 21, 22, 24, 25, 26	22, 23b, 25 22, 23a, 25	20, 27 20, 21, 27
24	172.89					
25	90.46	5.12	s	22, 23, 24, 26	23, 27	23, 27
26	167.45					
27	56.21	3.72	s	23, 24, 25		
28	56.82	3.34	s	18		
29	11.57	1.06 (6.6)	d	16, 17, 18	17	
30	20.9	0.93 (6.7)	d	8, 9, 10, 11	9	
31	112.87	5.78	s	4, 5, 6, 7, 8	5, 7	7, 8
NH		6.4 (5.6, 5.6)	t	14, 15	15	14, 15
OH		4.25	bs			

Molecule		4
C#	δ_c	Δ ppm
1	38.36	30.44
2	80.04	4.24
3	19.73	-1.23
4	26.02	0.18
5	29.4	0.05
6	141.92	0.1
7	32.75	-0.01
8	34.86	-0.01
9	36.48	-0.03
10	137.16	0
11	127.48	-0.03
12	28.75	-0.01
13	36.21	0
14	175.28	-0.02
15	48.02	-0.02
16	98.56	0
17	43.69	-0.01
18	78.1	-0.01
19	36.99	-0.01
20	65.41	-0.06
21	40.15	-0.01
22	74.33	-0.03
23	32.66	-0.03
24	172.86	0.03
25	90.49	-0.03
26	167.43	0.02
27	56.21	0
28	56.83	-0.01
29	11.58	-0.01
30	20.92	-0.02
31	112.96	-0.09

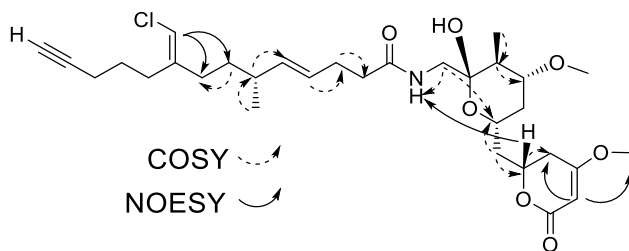
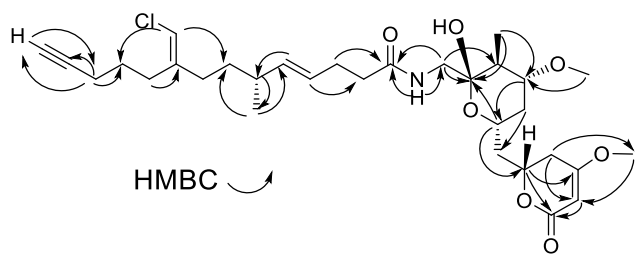


Table S5. Vatiamide C (**3**) NMR spectroscopy with key correlations (500 MHz, d_6 -DMSO)

C#	δ_C	δ_H, J in Hz	type	HMBC	key NOE
1	71.57	2.79 (3.2, 3.2)	t	2	
2	84.07				
3	17.74	2.16	d	1, 2, 4, 5	
4	25.65	1.54 (7.5, 7.5, 7.5)	dq	2, 3, 5, 6	
5	28.78	2.23	m	3, 4, 6, 7, 31	
6	142.11				
7	31.86	2.01	m	5, 6, 8, 31	
8	34.32	1.33	m	6, 7, 9, 10, 30	
9	35.68	2	m	6, 7, 8, 10, 11, 30, 31	
10	135.79	5.26 (7.3, 15.4)	dd	9, 11, 12, 30	
11	127.69	5.34	m	12, 19	
12	28.31	2.13	m	10, 11, 13, 14	
13	35.26	2.15	m	10, 11, 12, 14	
14	173.52				
15	45.47	2.89 (4.2, 13.5), 3.45 (7.8, 13.5)	dd, dd	14, 16, 17 (2.89)	
16	98.45				
17	40.44	1.33	m	15, 16, 18, 20, 21, 29	
18	77.86	3.15 (4.5, 10.7, 10.8)	td	17, 27, 29	
19	36.1	0.95, 2.1 (2.8, 5.6, 15)	m, ddd	18, 20, 21	
20	63.12	3.89	m		OH
21	40.3	1.74 (4.3, 7.7, 14.0), 1.91 (5.6, 8.7, 13.9)	ddd, ddd	19 (1.74), 20, 22, 23	
22	72.67	4.55	m		
23	31.61	2.46 (4.0, 17.2), 2.56 (1.3, 11.7, 17.1)	dd, ddd	22, 24, 25, 26 (2.56), 21 (2.56)	
24	171.96				
25	89.94	5.16 (1.3)	d	23, 24, 26	
26	166.44				
27	55.68	3.23	s	18	
28	56.24	3.72	s	24	
29	11.66	0.91 (6.6)	d	16, 18, 21	
30	20.55	0.92 (6.7)	d	8, 10	
31	112.36	6.01	s	5, 6, 7	
NH		7.52 (4.3, 7.6)	dd	14	
OH		5.73 (1.3)	d	15, 16, 21	20

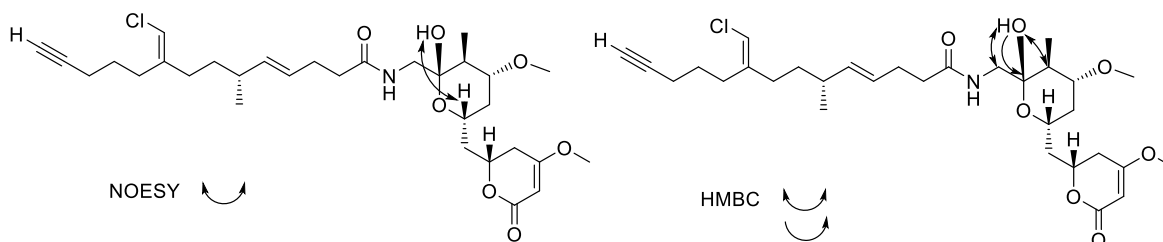
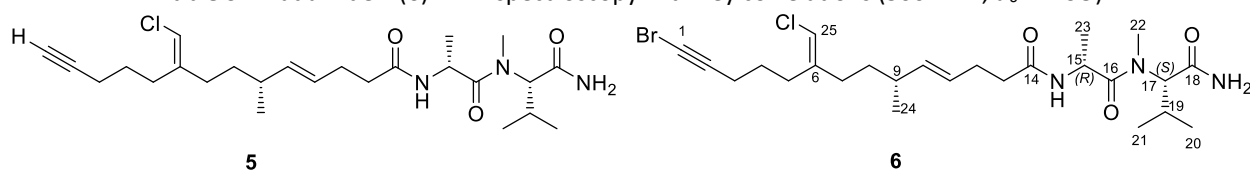


Table S6. Vatiamide C, D (**3**, **4**) mass spectrometry and optical rotation

Molecule, charge	3 , [M+Na] ⁺	4 , [M+Na] ⁺
Formula	C ₃₁ H ₄₆ ClNO ₇ Na	C ₃₁ H ₄₅ BrClNO ₇ Na
Calculated	602.2855	680.1960
Result	602.2834	680.1953
Δ ppm	3.49	1.03
[α] _D ²⁵ (c 0.1, CH ₃ OH)	-0.3	33.3

Table S7. Vatiamide F (**6**) NMR spectroscopy with key correlations (500 MHz, *d*₆-DMSO)



C#	δ_c	δ_H, J in Hz	type	HMBC	COSY	key NOE
1	40.46					
2	80.05					
3	18.83	2.24	m	1, 2, 4, 5, 6	4, 5	
4	25.43	1.54 (6.6, 6.6, 6.6, 6.6)	p	3, 5	3, 5	
5	28.66	2.20	m	3, 4, 6, 7, 25	3, 4	
6	142					
7	31.82	2.03	m	6, 8, 25	8, 25	25
8	34.3	1.33	m	7, 9, 10, 11, 24	7, 9	25
9	35.73	1.98	m	6, 7, 8, 10, 11, 24, 25	8, 10, 24	
10	135.81	5.24 (7.3, 15.3)	dd	9, 11, 12, 24	9, 11	
11	127.63	5.33	m	9, 10, 12	10, 12	
12	28.16	2.16	m	10, 11, 13, 14	11, 13	
13	35	2.12	m	10, 11, 12, 14	12	NH
14	171.13					
15	44.81	4.70 (6.9, 6.9, 6.9, 6.9)	p	14, 16, 23	NH, 23	22, 23
16	172.93					
17	60.95	4.53 (10.6)	d	15, 16, 18, 19, 20, 21, 22	19, 22	NH ₂ (b), 20, 21, 22
18	171.52					
19	26.27	2.08	m	17, 18, 20, 21	17, 20, 21	22
20	19.82	0.91 (6.3)	d	17, 19, 21	17, 19	17, 19, 22
21	18.65	0.73 (6.8)	d	17, 19, 20	17, 19, 20	17, 19, 22
22	30.38	2.98	s	16, 17	17	15, 19, 20, 21, 23
23	17.47	1.16 (7.4)	d	15, 16	15	NH, 15, 20
24	20.58	0.92 (6.3)	d	8, 9, 10	8, 9	
25	112.45	6.03	s	4, 5, 6, 7, 8, 25	5, 7	7, 8
NH		8.09 (54)	d	14, 15, 23	15, 23	13, 23
NH ₂		7.05(a), 7.16(b) ^a	s,s	17(b), 18(a, b)	17(b)	17(b), 20(a, b)

^a Note: Presence of additional shadow peaks due to conformer from rotation around N-methyl amide bond. At 25°C, peaks appear at 2:1 ratio in *d*₆-DMSO and 8:1 ratio in *d*₃-acetonitrile; single conformer resolved at 80°C in *d*₆-DMSO (Figure S29-30). Major conformer peak at 25°C reported above.

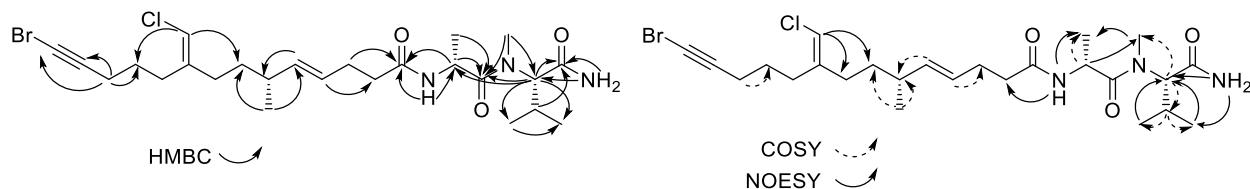
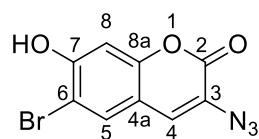


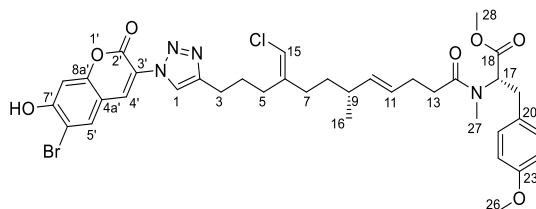
Table S8. Vatiamide E, F (**5**, **6**) mass spectrometry and optical rotation

Molecule, charge	5 , [M+Na] ⁺	6 , [M+Na] ⁺
Formula	C ₂₅ H ₄₀ ClN ₃ O ₃ Na	C ₂₅ H ₃₉ BrClN ₃ O ₃ Na
Calculated	488.2650	566.1756
Result	488.2641	566.1742
Δ ppm	1.84	2.47
[α] _D ²⁵ (c 0.1, CH ₃ OH)	N/A	-25.0

Table S9. NMR spectral data for probe **14** (500 MHz, DMSO-*d*₆)

Position #	δ _H (mult., J(Hz))	δ _C ^a	COSY	HMBC
1	---	---	---	---
2	---	156.9	---	---
3	---	122.5	---	---
4	7.56 (s)	126.3	5, 8	2, 3, 5, 8a
4a	---	112.9	---	---
5	7.87 (s)	131.2	4, 8	4, 5, 7, 8, 8a
6	---	106.6	---	---
7	---	156.2	---	---
7-OH	11.40 (s)	---	---	6, 7, 8
8	6.93 (s)	102.8	4, 5	4a, 6, 7, 8a
8a	---	151.6	---	---

Table S10. NMR spectral data for **14** + **1** adduct (500 MHz, *d*₄-MeOH)



Position #	δ_C	δ_H (mult., <i>J</i> (Hz))	COSY	HMBC	ROESY
1	124.0	8.35 (s)	----	2	3, 4
2	148.3	----	----	----	----
3	25.9	2.79 (m)	4	1, 2, 4, 5	1, 4
4	27.7	1.85 (m)	3, 5	2, 3, 5, 6	1, 3
5	30.1	2.31 (m)	4	4, 6, 7, 15	
6	143.2	----	----	----	----
7	33.3	2.06 (m)	8	8, 10, 16	15
8	35.8	1.34 (m)	7, 9	5, 6, 7, 15	10, 11, 15
9	37.5	1.99 (m)	8	8, 10, 11, 16	10, 11
10	137.5	5.27 (d, 13.0)		12	8, 9, 12, 13, 16
11	128.6	5.28 (d, 11.8)	12	9	8, 9, 12, 13, 16
12	28.7	2.14 (m)	11, 13	10, 11, 13, 14	10, 11
13	34.3	2.29 (m)	12	11, 14	10, 11
14	175.3	----	----	----	----
15	113.6	5.91 (s)	----	5, 7, 14	7, 8
16	21.1	0.95 (d, 6.7)	9	8, 10	10, 11
17	60.6	5.02 (dd, 11.1, 5.2)	19a, 19b	13, 18, 19	19a, 19b, 21, 25, 27
18	172.2	----	----	----	----
19	34.5	a 3.01 (dd, 14.6, 11.2) b 3.23 (dd, 14.6, 3.6)	17, 19b 17, 19a	17, 21, 25 17, 21, 25	17, 21, 25 17, 21, 25
20		----	----	----	----
21, 25	130.7	7.08 (d, 8.6)	22, 24	19, 21, 23, 25	17, 19, 27
22, 24	114.9	6.79 (d, 8.6)	21, 25	21, 23, 22, 24, 25	26
23	159.5	----	----	----	----
26	55.3	3.73 (s)	----	23	22, 24
27	34.4	2.82 (s)	----	17, 18	17, 21, 25
28	52.4	3.70 (s)	----	18	
1'	----	----	----	----	----
2'	161.5	----	----	----	----
3'	112.1	----	----	----	----
4'	104.2	6.86 (s)	----	2', 9', 10'	
5'	133.7	7.94 (s)	----	8', 9', 10'	
6'		----	----	----	----
7'	157.2	----	----	----	----
8'	135.6	8.41 (s)	----	5', 7', 9'	
9'	155.2	----	----	----	----
10'	109.6	----	----	----	----

Table S11. Proposed functions of open reading frames in *vat* pathway

Protein	Length	Proposed Function	Similarity	Identity	Similarity	Accession
VatA	598	Fatty-acyl AMP ligase	JamA, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	92%	95%	AAS98774.1
VatB	321	Desaturase	JamB, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	97%	98%	AAS98775.1
VatC	100	ACP	JamC, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	93%	95%	AAS98798.1
VatD	683	Flavin-dependent brominase	JamD, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	94%	97%	AAS98776.1
VatE	1737	KS AT halogenase ACP ACP ACP	JamE, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	92%	94%	AAS98777.1
VatF	79	ACP	JamF, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	94%	94%	AAS98799.1
VatG	409	KS	JamG, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	91%	94%	AAS98778.1
VatH	419	HCS	JamH, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i>)	98%	99%	AAS98779.1
VatI	434	Unknown	HNH endonuclease, <i>Moorea producens</i> PAL-8-15-08-1	96%	97%	WP_070395762.1
VatJ	254	ECH	JamI, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	96%	99%	AAS98780.1
VatK	3277	ECH ER KS AT DH CMT ER KR ACP	JamJ, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	94%	96%	AAS98781.1
VatL	1659	KS AT KR ACP	JamK, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	92%	94%	AAS98782.1
VatM	2187	KS AT DH ER KR ACP	JamL, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i> JHB)	79%	86%	AAS98783.1
VatN	2153	C A(Tyr) N-MT O-MT PCP TE	Nonribosomal peptide synthetase, <i>Moorea producens</i> 3L	72%	81%	AEF01451.1
VatO	271	O-MT	O-methyltransferase protein, <i>Planktothrix sarta</i> PCC 8927	65%	77%	CUR19792.1
VatP	79	ACP	ColF, <i>Moorea bouillonii</i> PNG5-198	94%	96%	AKQ09583.1
VatQ	1582	C A(Ala) PCP E	PuwG, <i>Cylindrospermum alatosporum</i> CCALA 988	52%	68%	AIW82284.1
VatR	1895	C A(Val) N-MT PCP AmT	BarG, <i>Lyngbya 49irsute49e</i> (<i>Moorea producens</i>)	50%	64%	AEE88297.1
VatS	1585	C A(Gly) N-MT PCP	MgcJ, <i>Okeania 49irsute</i> PAB10Feb10-1	84%	92%	AZH23792.1
VatT	2260	KS AT DH C-MT KR ACP	MgcQ, <i>Okeania 49irsute</i> PAB10Feb10-1	67%	80%	AZH23817.1
VatU	1950	KS AT O-MT KR ACP	CurL, <i>Moorea producens</i> 3L	67%	81%	AEE88278.1
VatV	1610	KS AT KR ACP	CurG <i>Moorea producens</i> 3L	62%	75%	AEE88283.1
VatW	2888	KS AT KR ACP KS AT ACP TE	ColF, <i>Moorea bouillonii</i> PNG5-198	64%	77%	AKQ09583.1
VatX	219	O-MT	StfMI, <i>Streptomyces steffisburgensis</i>	42%	60%	CAJ42328.1
Orf1	378	Putative hydrolase	Linear amide C-N hydrolase, <i>Moorea producens</i>	86%	93%	WP_071106594.1
Orf2	354	unknown	tRNA 2-selenouridine synthase MnmH <i>Moorea producens</i>	96%	97%	WP_070395342.1

Orf3	85	transposase	Transposase, <i>Leptolyngbya</i> sp. PCC 7376	59%	78%	AFY39744.1
Orf4	86	transposase	ISAs1 family transposase <i>Moorea producens</i>	88%	91%	WP_083373584.1
Orf5	185	transposase	DDE transposase family protein <i>Moorea producens</i>	85%	91%	WP_070395360.1
Orf6	1155	Regulatory	Response regulator <i>Moorea producens</i>	98%	98%	WP_071106591.1

Abbreviations: A; adenylation, ACP; Acyl carrier protein, AmT; putative amidotransferase, AT; acyltransferase, C; condensation, C-MT; C-methyltransferase, DH; dehydratase, E; epimerase, ECH; enoyl-CoA hydratase, ER; enoyl reductase, HCS; hydroxymethylglutaryl-CoA synthase, KR; ketoreductase, KS; ketosynthase, N-MT; N-methyltransferase, O-MT; O-methyltransferase, PCP; peptidyl carrier protein, TE; thioesterase

Table S12. Organisms and accession numbers for 16S sequence comparison

Organism	Accession #
<i>Moorea producens</i> 3L	NR_116539.1
<i>Moorea producens</i> JHB	CP017708.1
<i>Moorea producens</i> ASI16Jul14-2	MK605967
<i>Moorea producens</i> PAL15Aug08-1	CP017599.1

A

	ASI16Jul14-2	3L	JHB	PAL15Aug08-1
ASI16Jul14-2	-	97.7	98.6	98.9
3L	97.7	-	98.6	98.4
JHB	98.6	98.6	-	98.7
PAL15Aug08-1	98.9	98.4	98.7	-

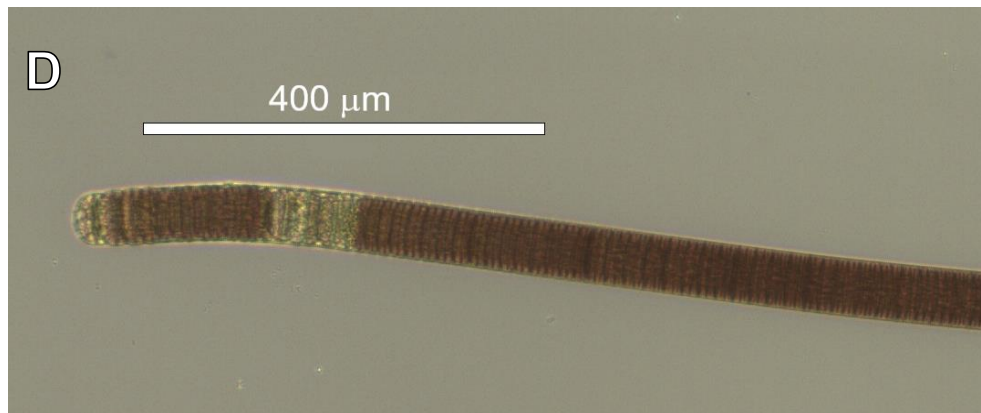
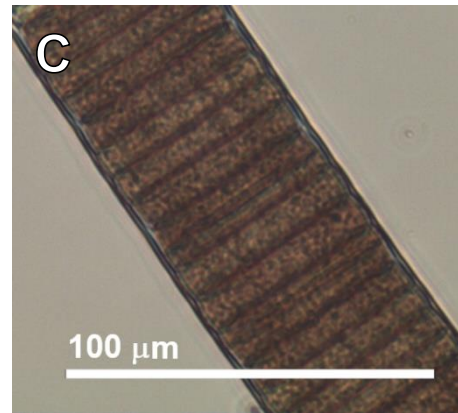
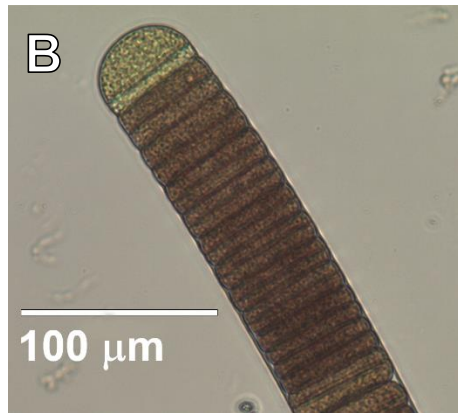


Figure S1. Analysis of 16S rRNA sequence and morphology of *Moorea producens* ASI16Jul14-2. A) Identity matrix of 16S rRNA sequences of *Moorea producens*, aligned by MUSCLE, reflecting % nucleotide identity. B) Microscopy image of ASI16Jul14-2 at 40X objective C) Microscopy image of ASI16Jul14-2 at 20X objective D) Microscopy image of ASI16Jul14-2 at 10X objective.

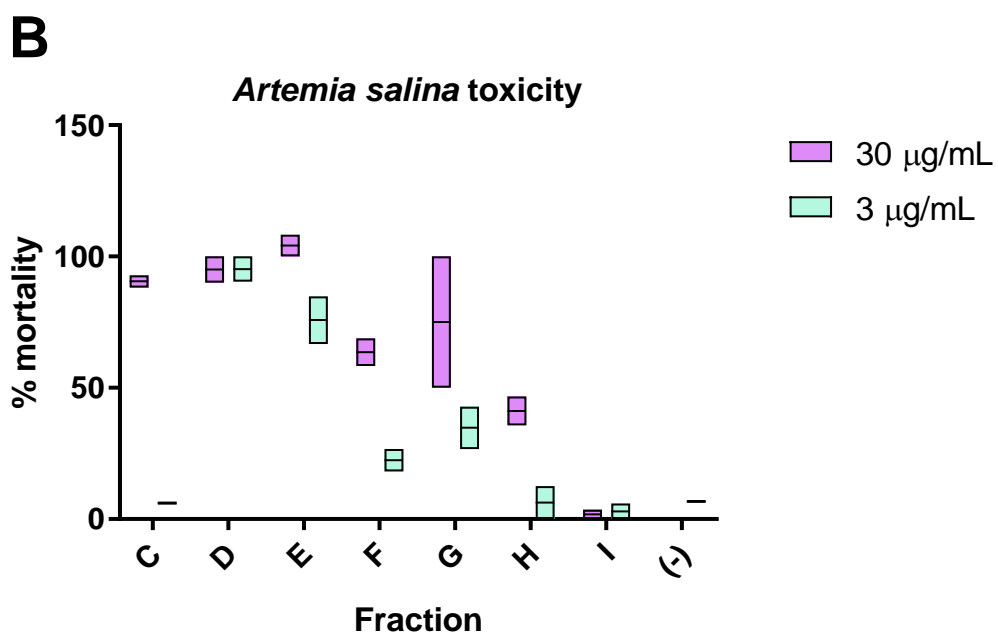
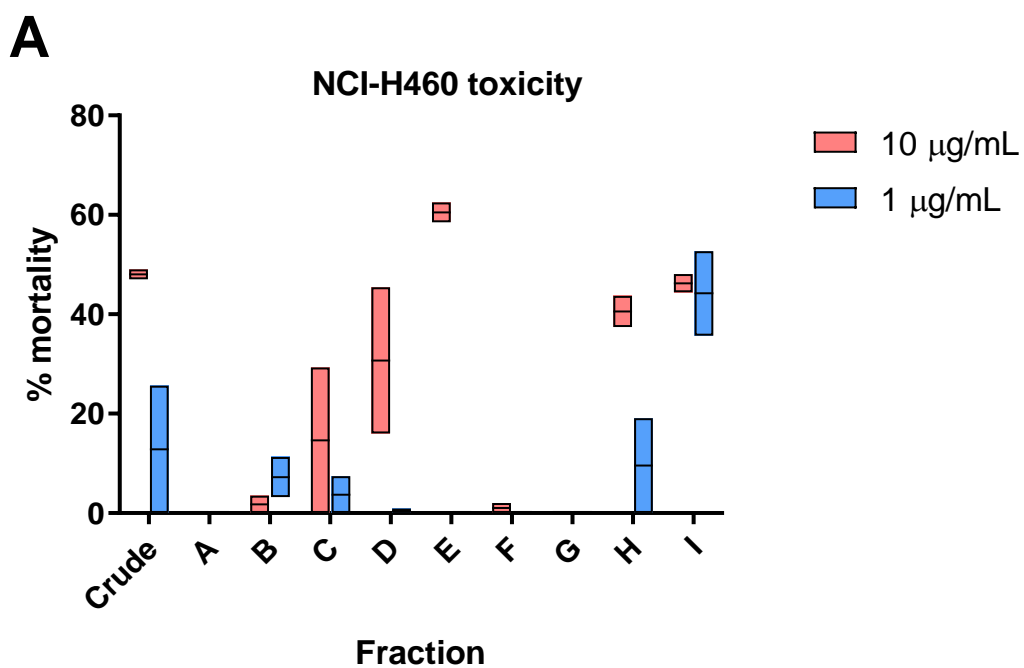


Figure S2. Bioassay of fractionated extracts. A) NCI-H460 MTT fluorometric assay results for fractions A-I and crude. B) *Artemia salina* brine shrimp toxicity for fractions C – I. DMSO used as negative control.

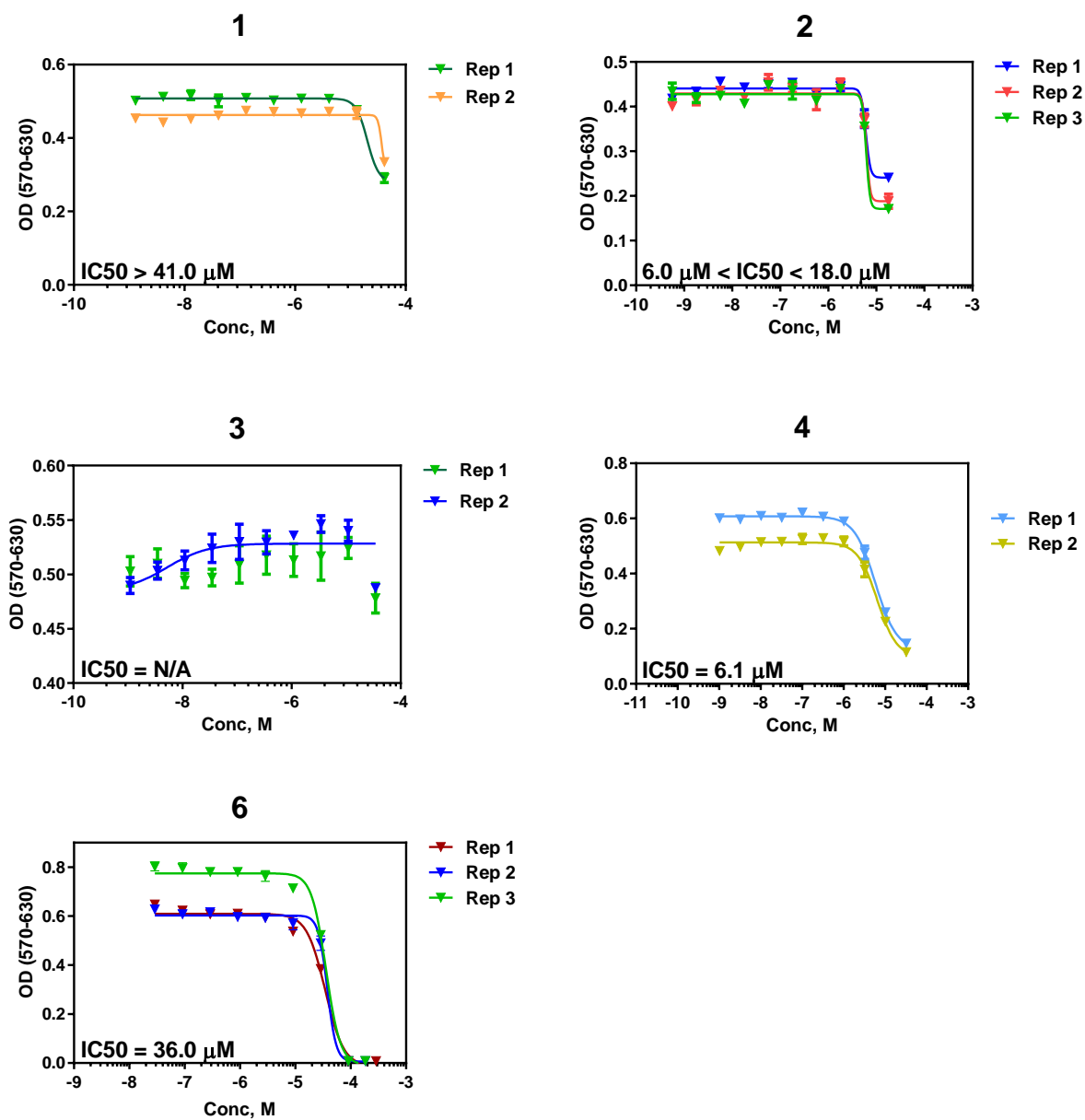


Figure S3. NCI-H460 MTT fluorometric assay results for pure compounds **1-4, 6**. Graph titles equate to compound tested. Each replicate represents a dilution curve generated from three biological replicates. **5** was not recovered in quantity sufficient for bioassay testing.

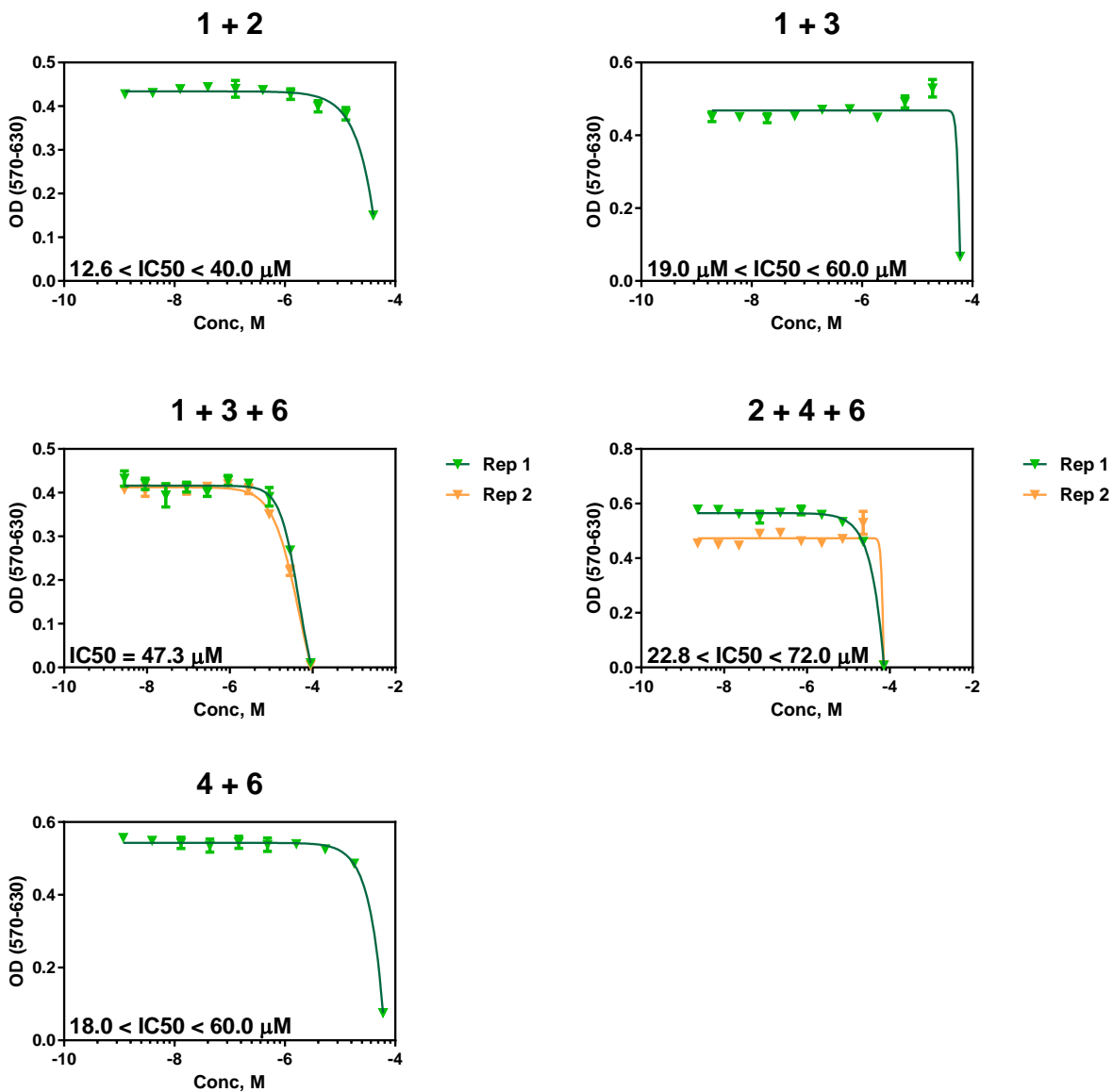


Figure S4. NCI-H460 MTT fluorometric assay results for mixtures of compounds **1-4, 6** to assess synergy. Graph titles equate to compound mixtures tested. Each replicate represents a dilution curve generated from three biological replicates. **5** was not recovered in quantity sufficient for bioassay testing.

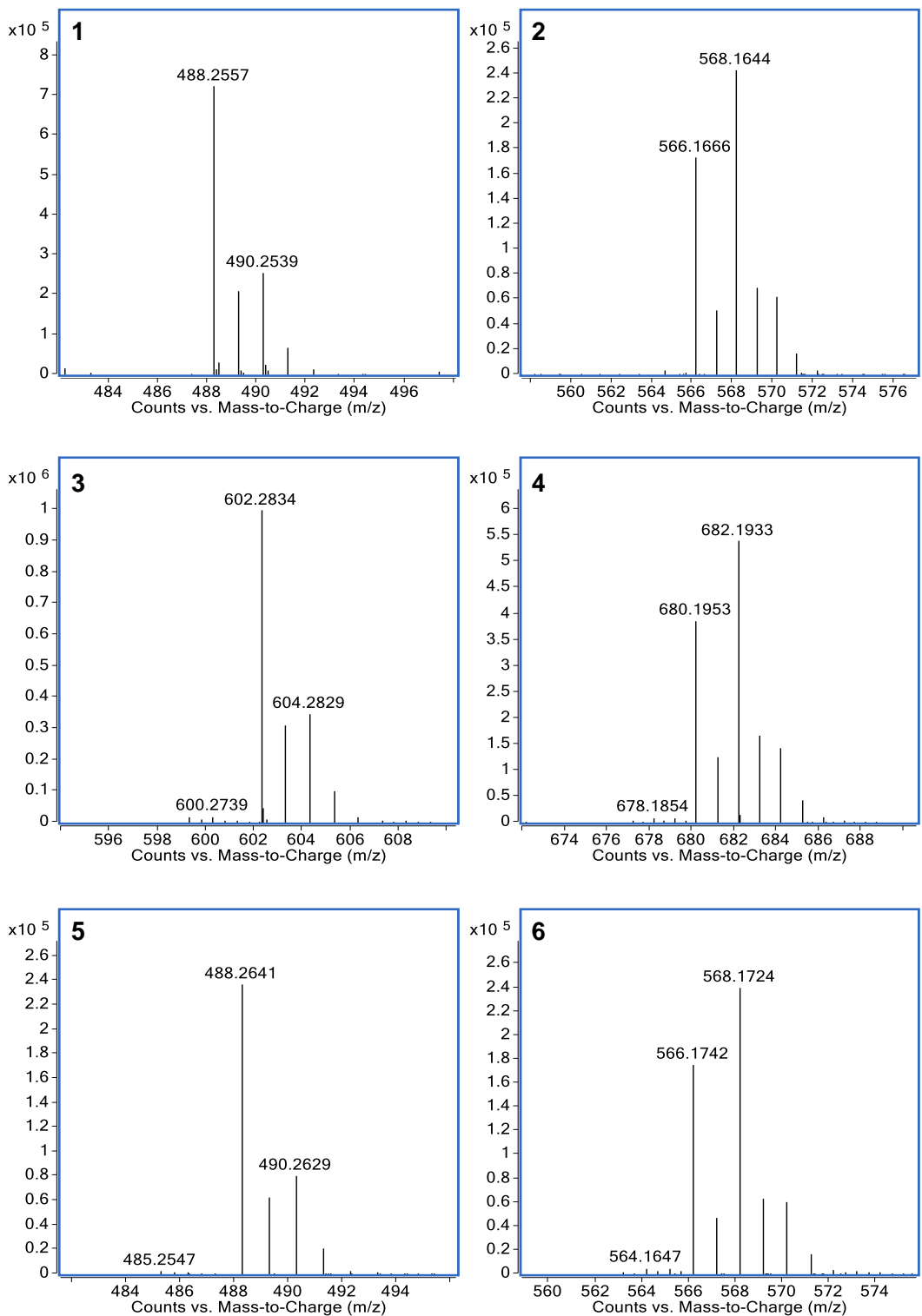


Figure S5. High-resolution mass spectrometry of **1-6**.

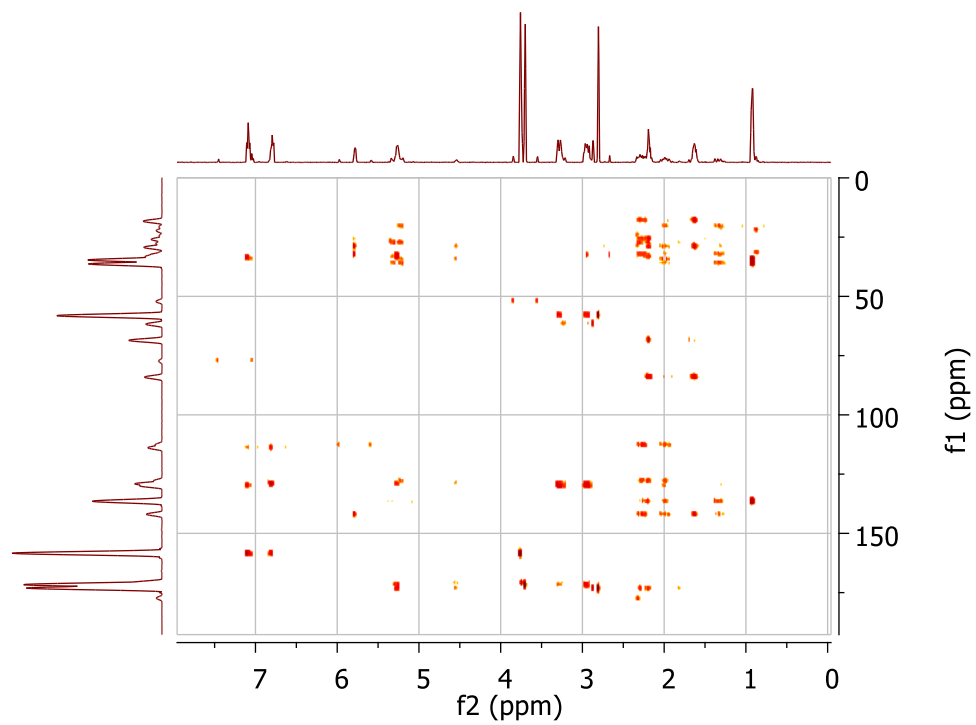


Figure S8. HMBC spectrum of **1** (500 MHz, CDCl₃).

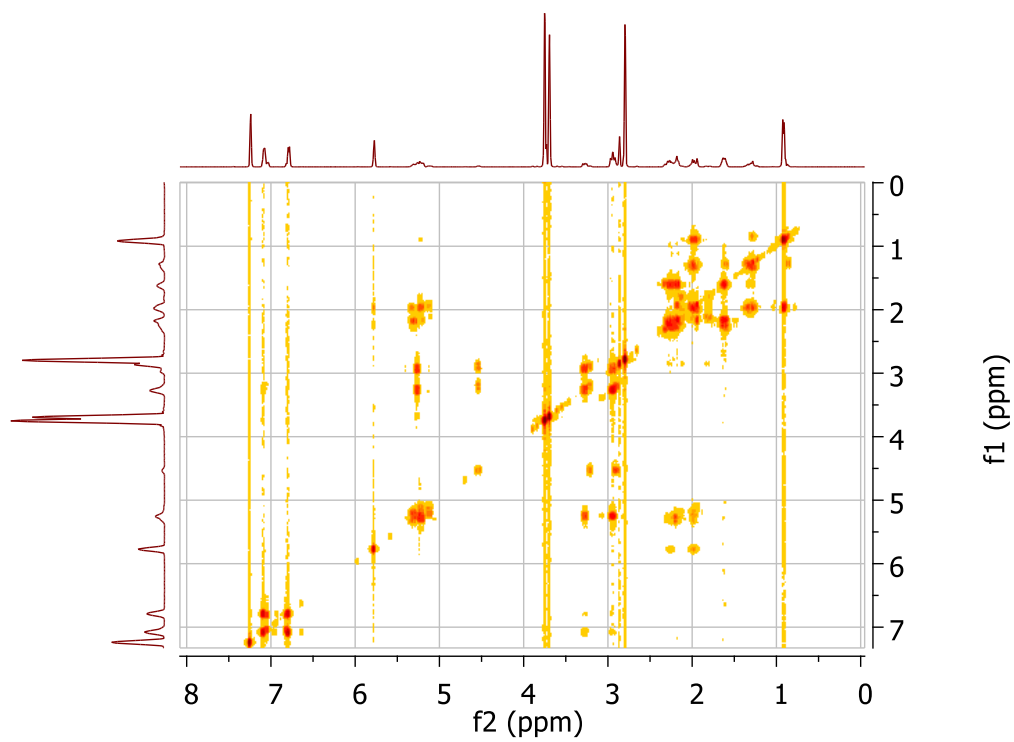


Figure S9. COSY spectrum of **1** (500 MHz, CDCl₃).

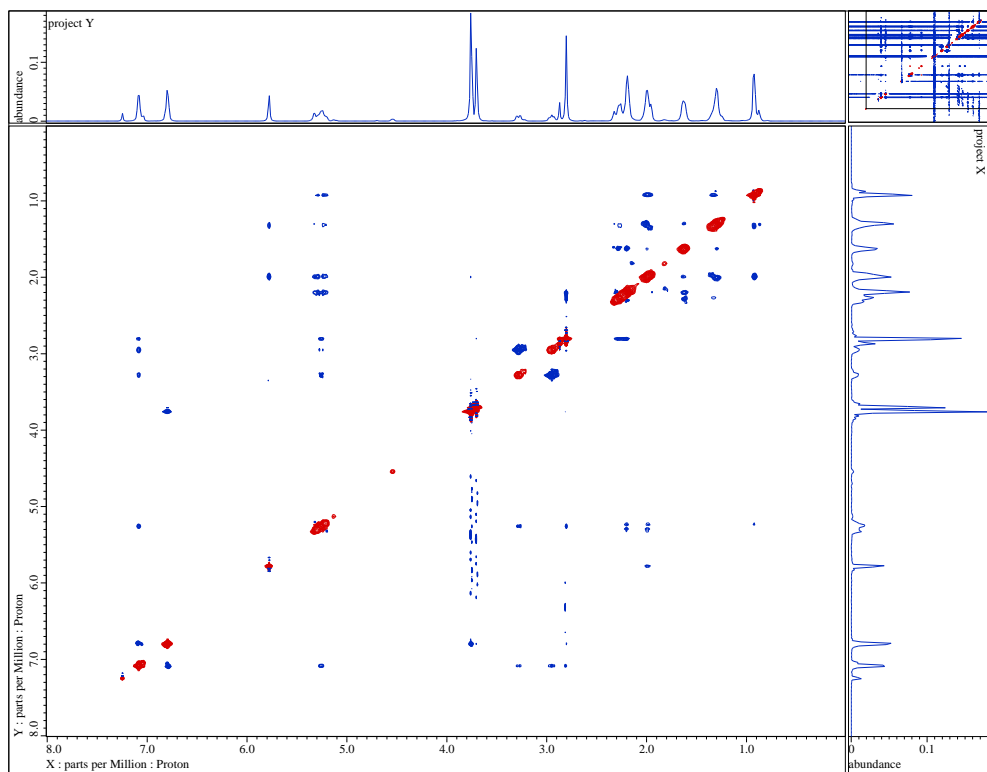


Figure S10. NOESY spectrum of **1** (500 MHz, CDCl₃).

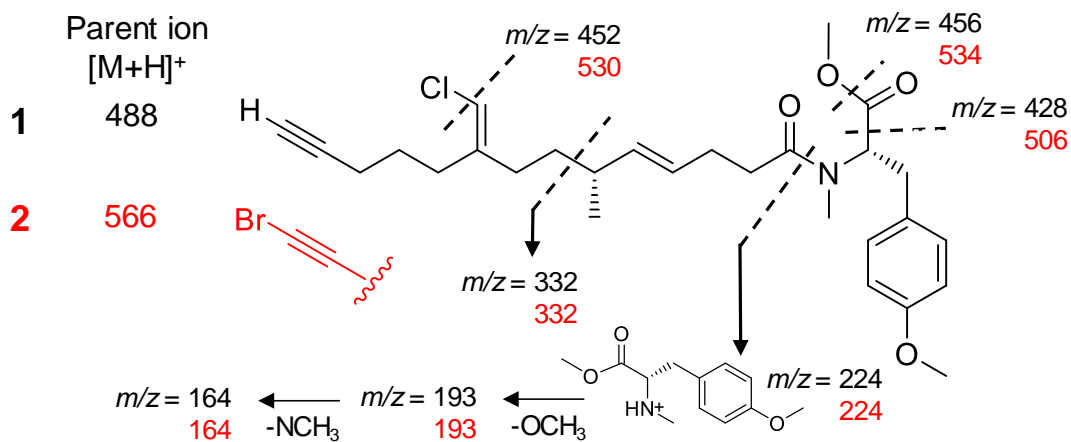
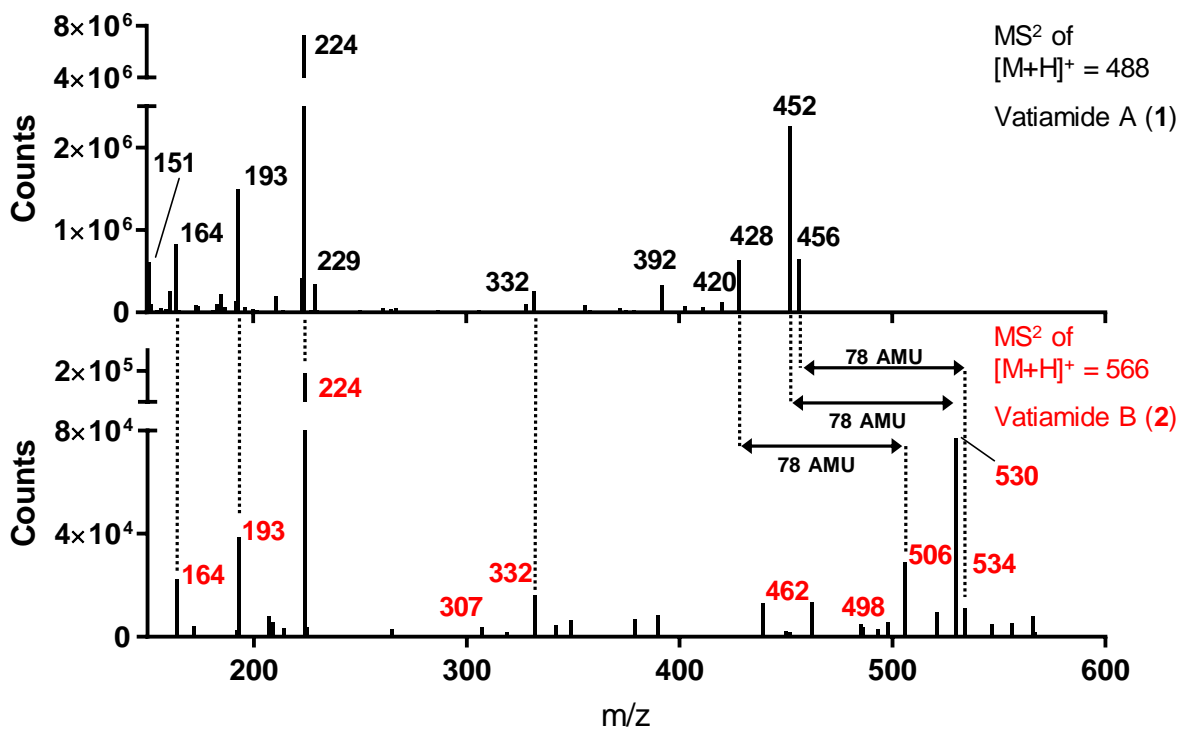


Figure S11. MS² comparison of **1** and **2**. Top spectrum, ESI-MS² of vatiamide A (**1**), bottom spectrum, ESI-MS² of vatiamide B (**2**). Major fragment ions offset by 78 amu indicated by dashed lines. Predicted fragmentation depicted in bottom panel, with 78 amu-offset masses deriving from **2** in red.

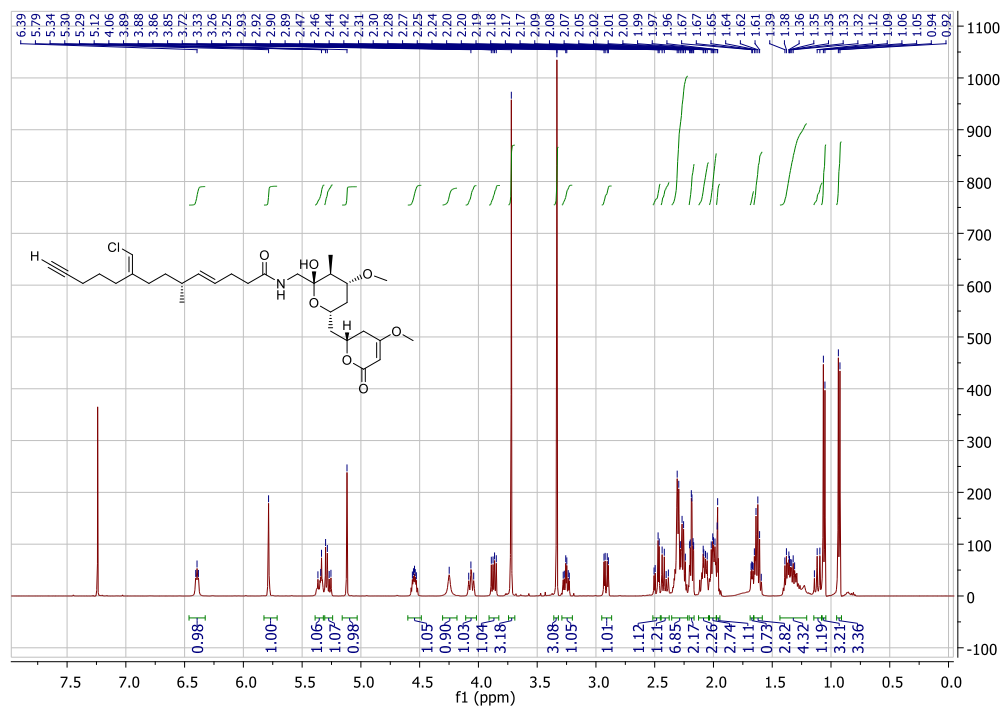


Figure S12. $^1\text{H-NMR}$ spectrum of **3** (500 MHz, CDCl_3).

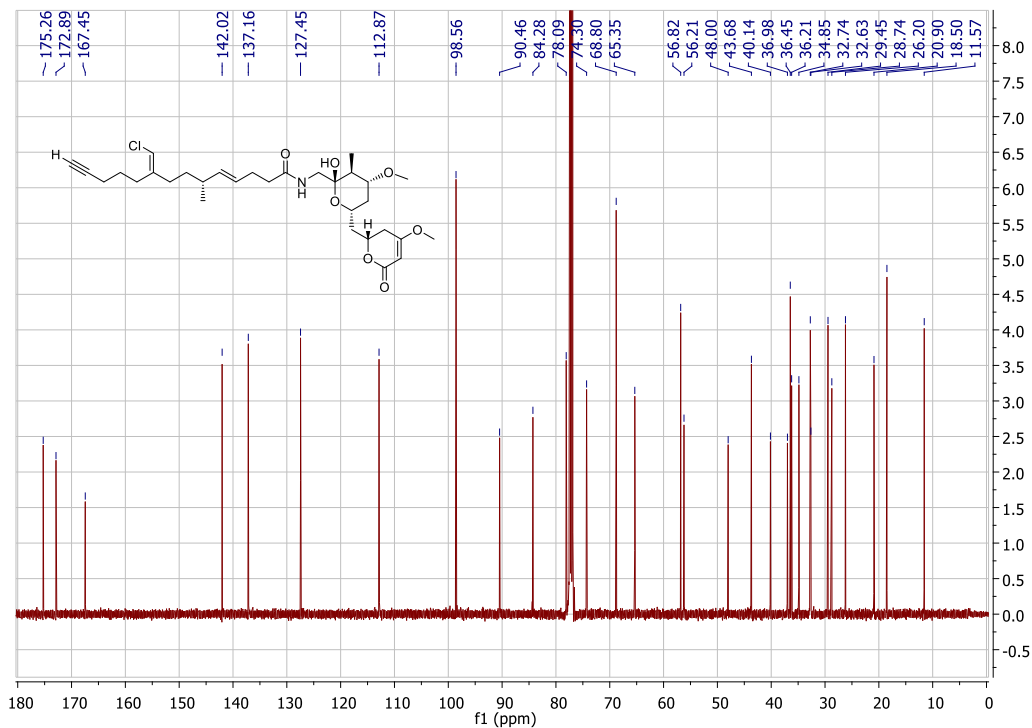


Figure S13. $^{13}\text{C-NMR}$ spectrum of **3** (500 MHz, CDCl_3).

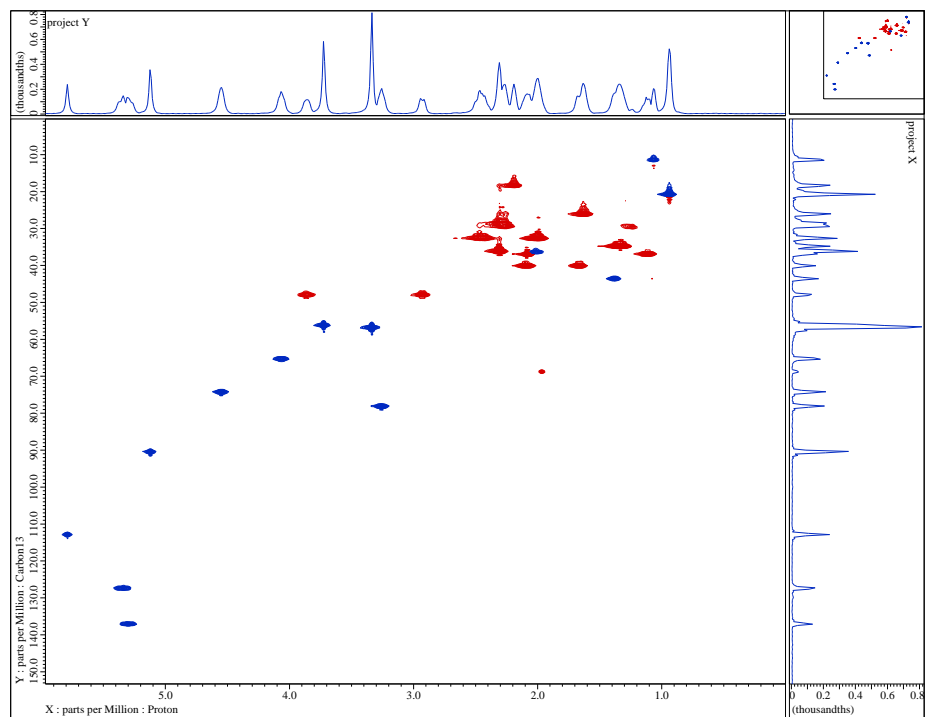


Figure S14. HSQC spectrum of **3** (500 MHz, CDCl₃).

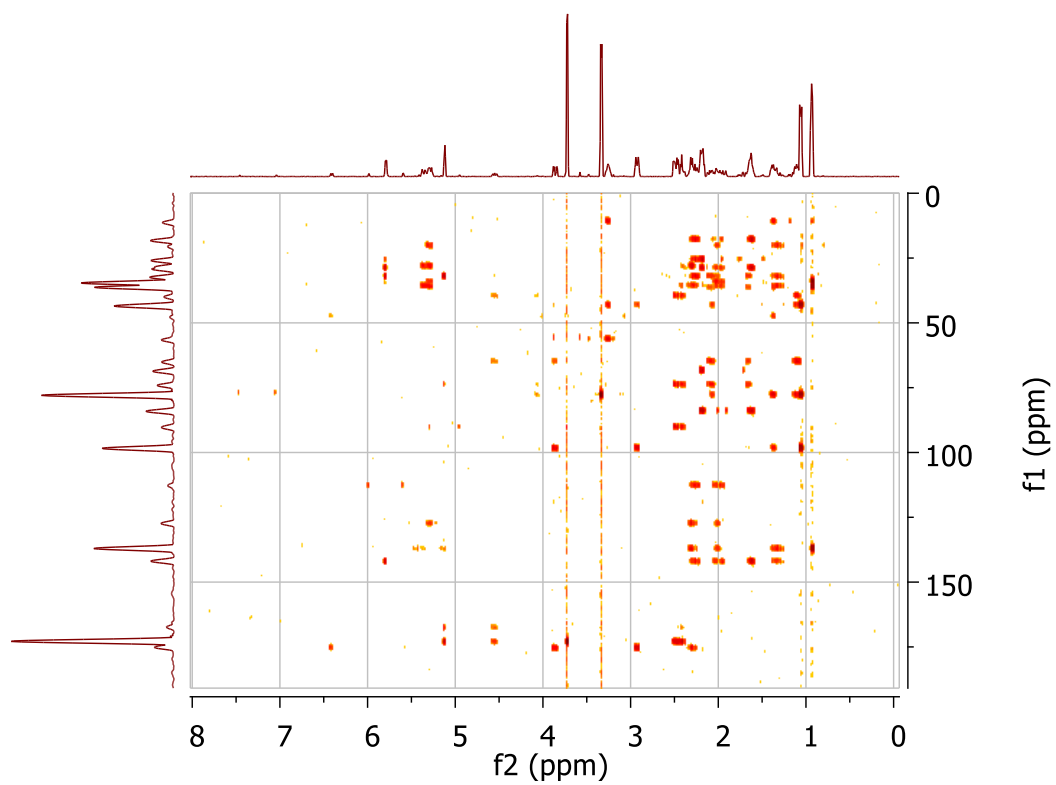


Figure S15. HMBC spectrum of **3** (500 MHz, CDCl₃).

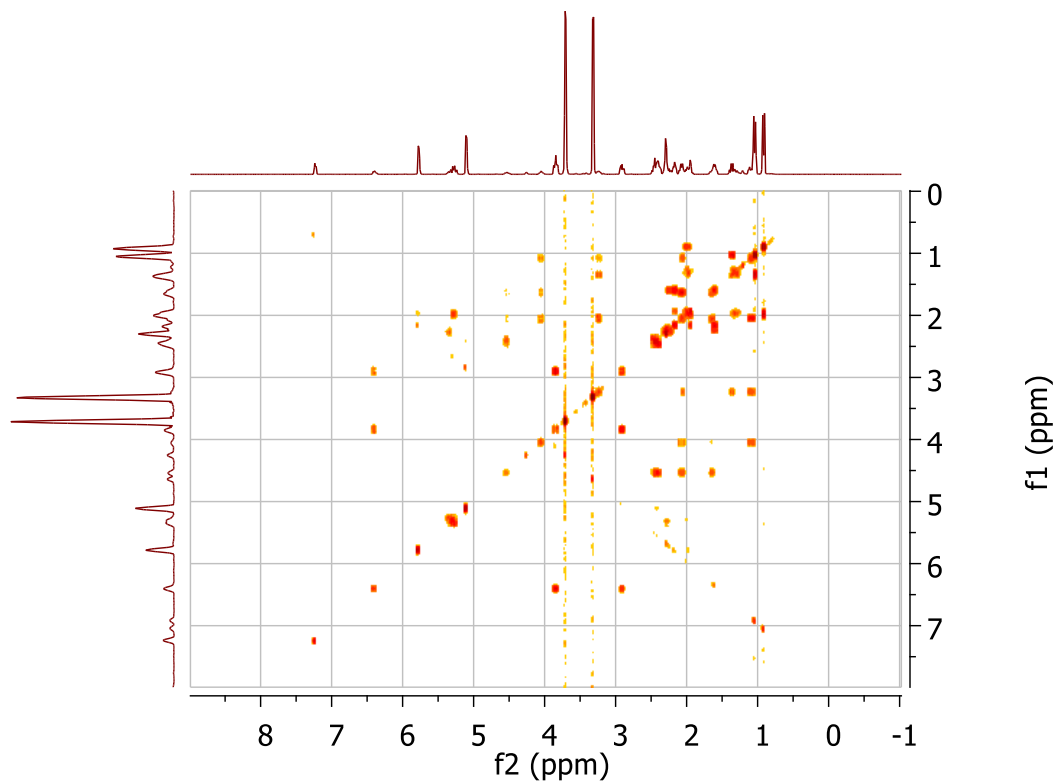


Figure S16. COSY spectrum of **3** (500 MHz, CDCl₃).

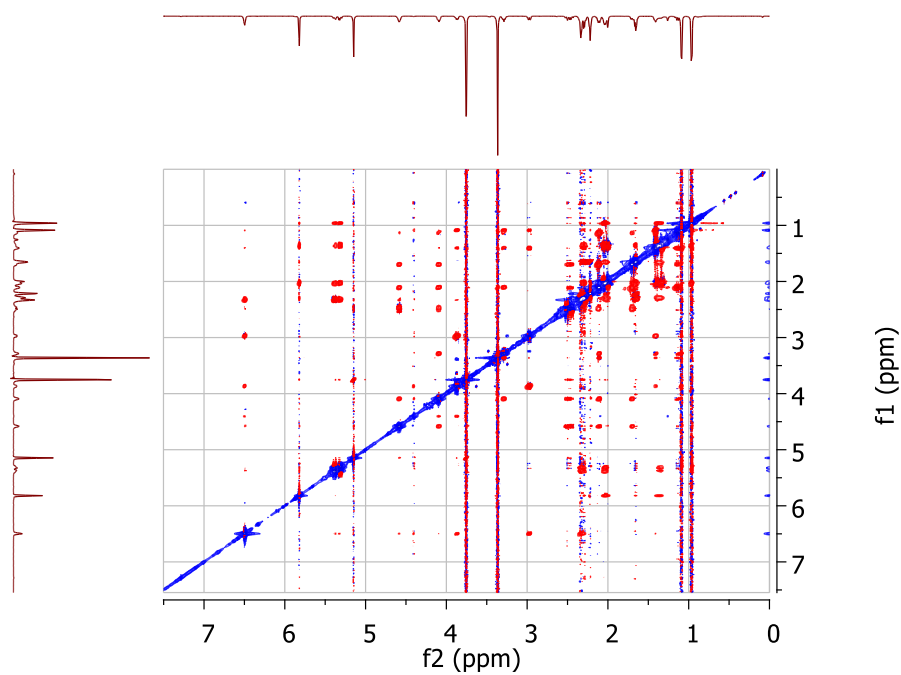


Figure S17. NOESY spectrum of **3** (500 MHz, CDCl₃).

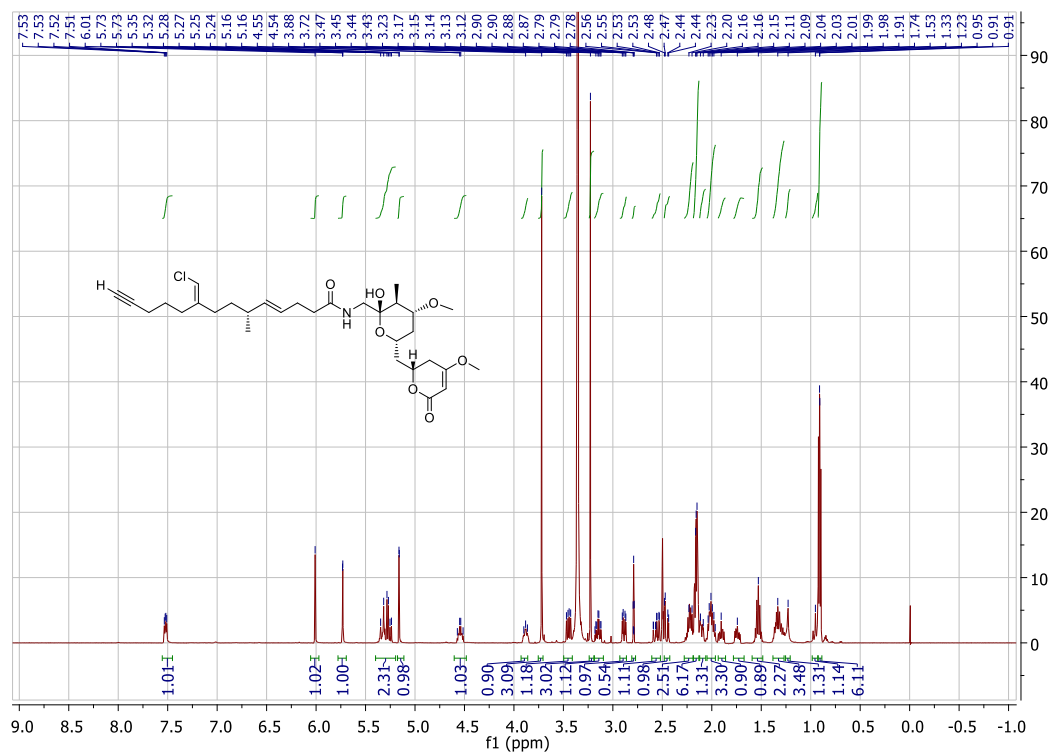


Figure S18. $^1\text{H-NMR}$ spectrum of **3** (500 MHz, $\text{DMSO-}d_6$).

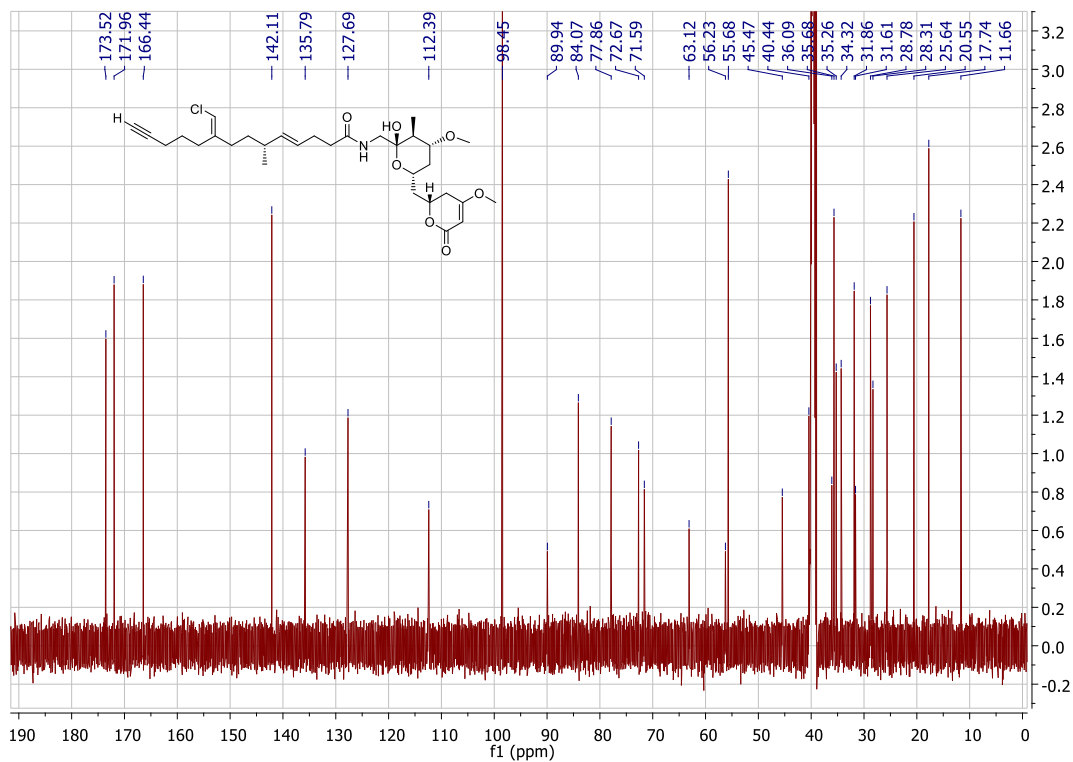


Figure S19. $^{13}\text{C-NMR}$ spectrum of **3** (500 MHz, $\text{DMSO-}d_6$).

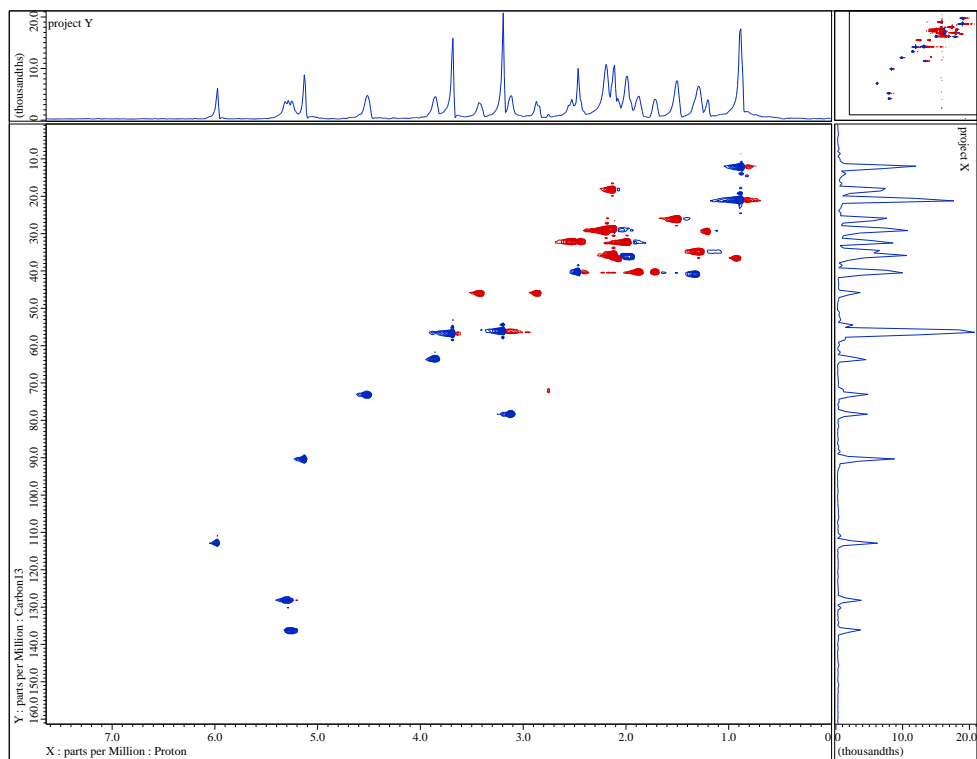


Figure S20. HSQC spectrum of **3** (500 MHz, DMSO- d_6).

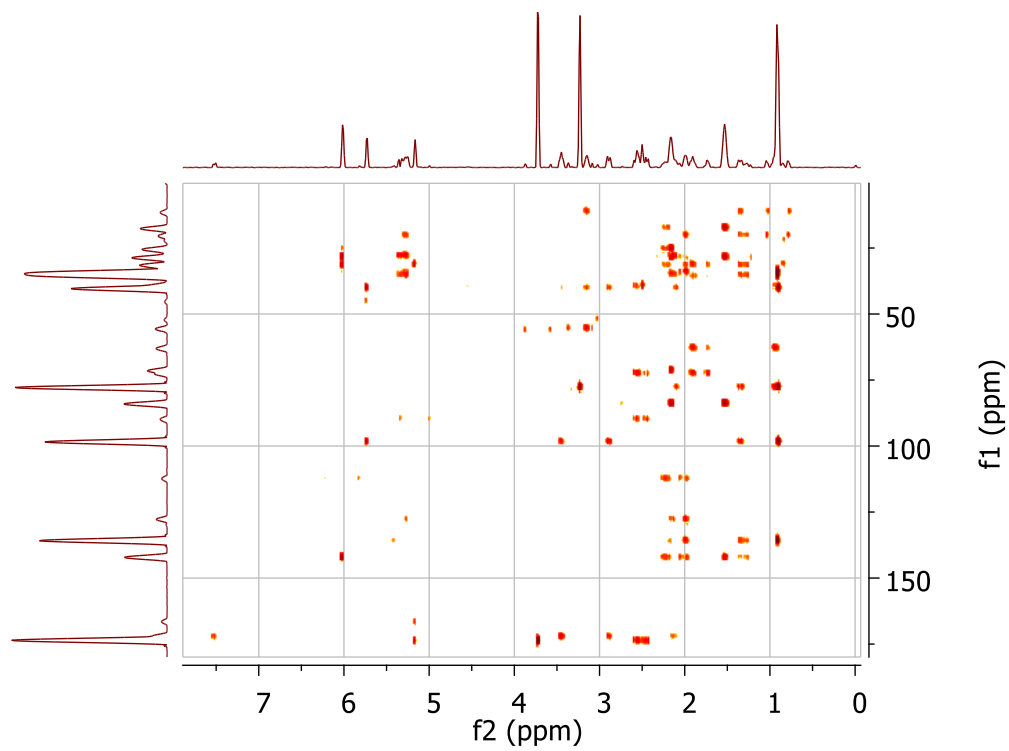


Figure S21. HMBC spectrum of **3** (500 MHz, DMSO- d_6).

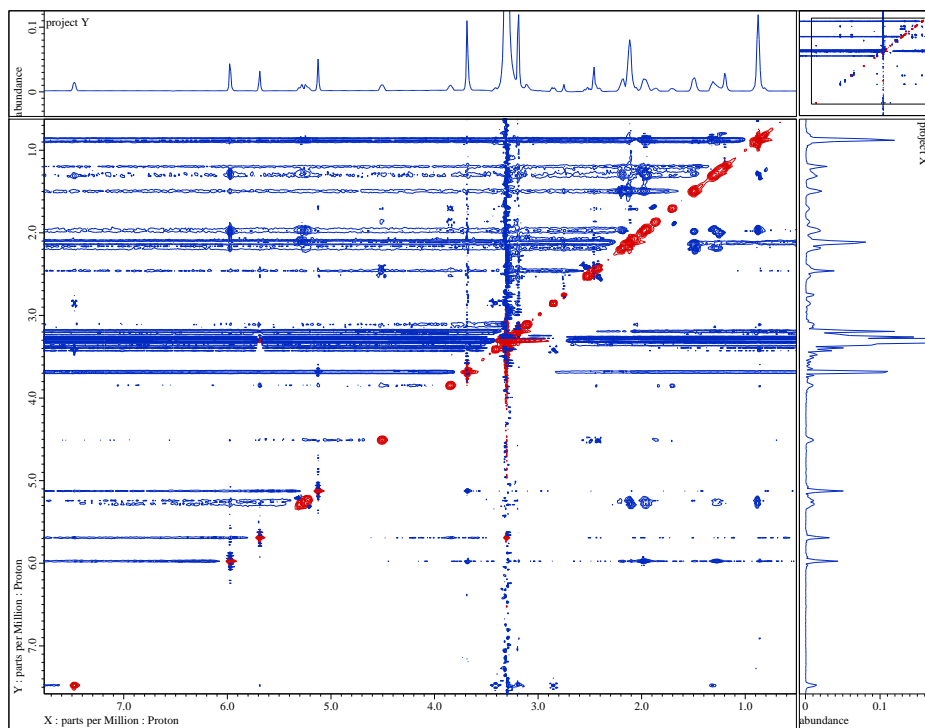


Figure S22. NOESY spectrum of **3** (500 MHz, DMSO-*d*₆).

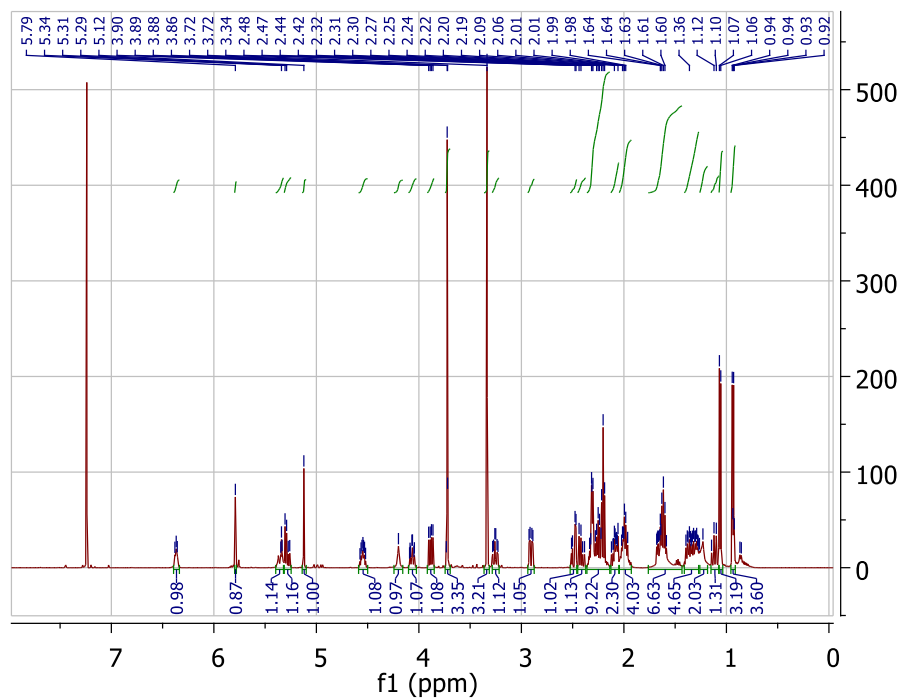


Figure S23. ^1H -NMR spectrum of **4** (500 MHz, CDCl_3).

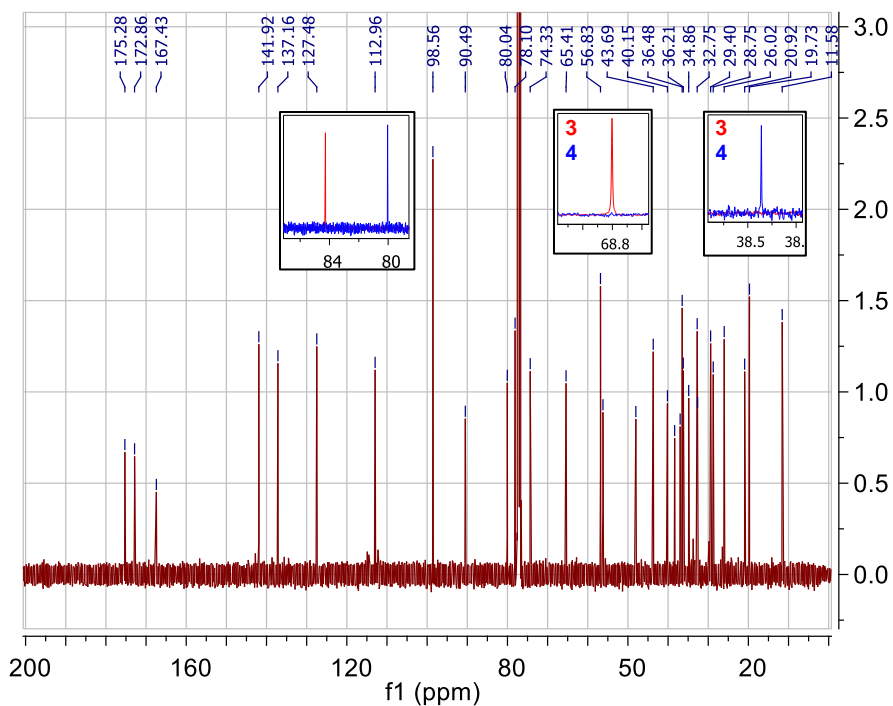


Figure S24. ^{13}C -NMR spectrum of **4**. Boxes depict the resonances for the C1-C2 alkyne for compound **3** (in red) and compound **4** (in blue; 500 MHz, CDCl_3).

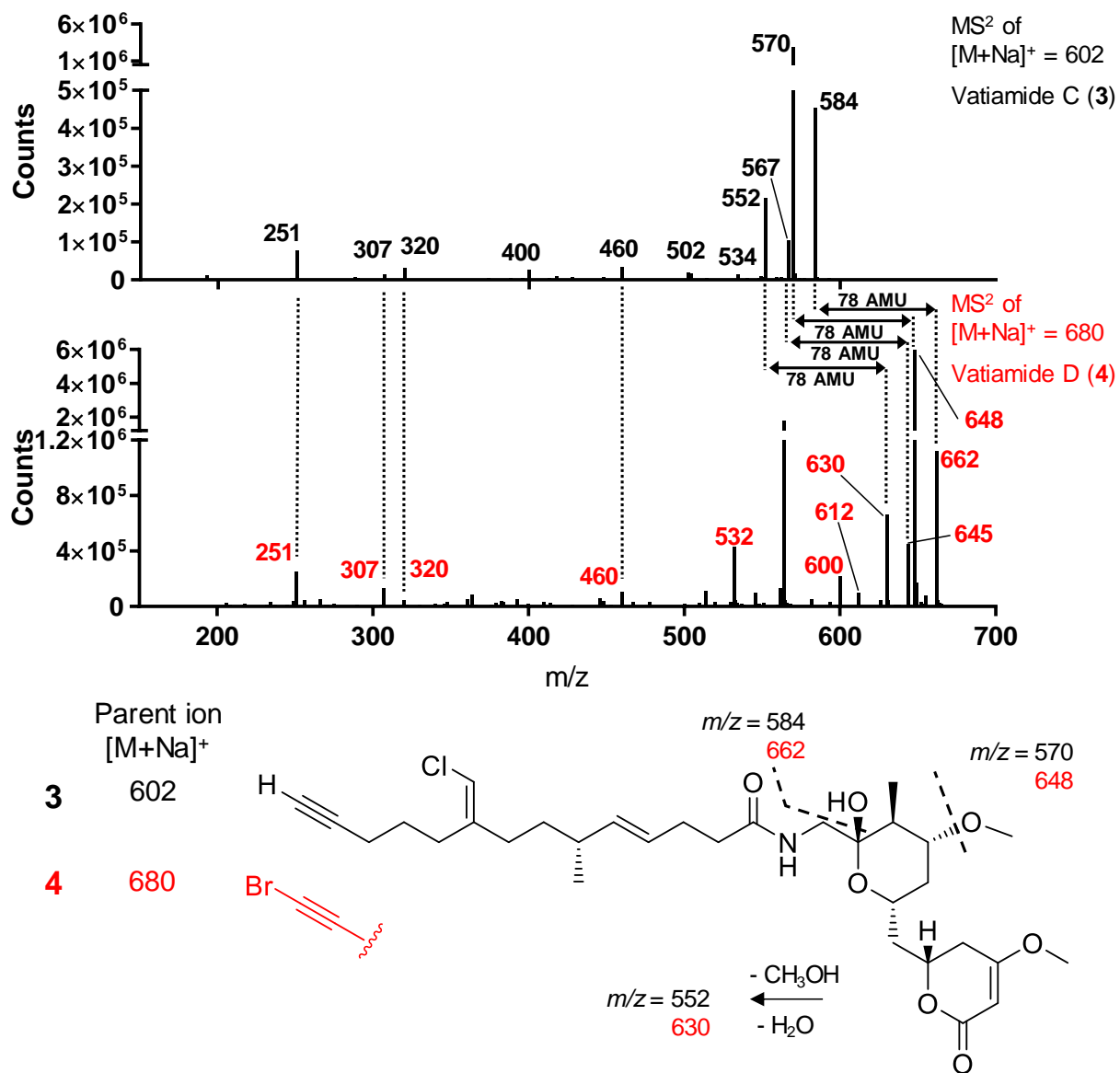


Figure S25. MS² comparison of **3** and **4**. Top spectrum, ESI-MS² of vatiamide C, bottom spectrum, ESI-MS² of vatiamide D. Major fragment ions offset by 78 amu indicated by dashed lines. Predicted fragmentation depicted in lower panel, with 78 amu-offset masses deriving from **4** in red.

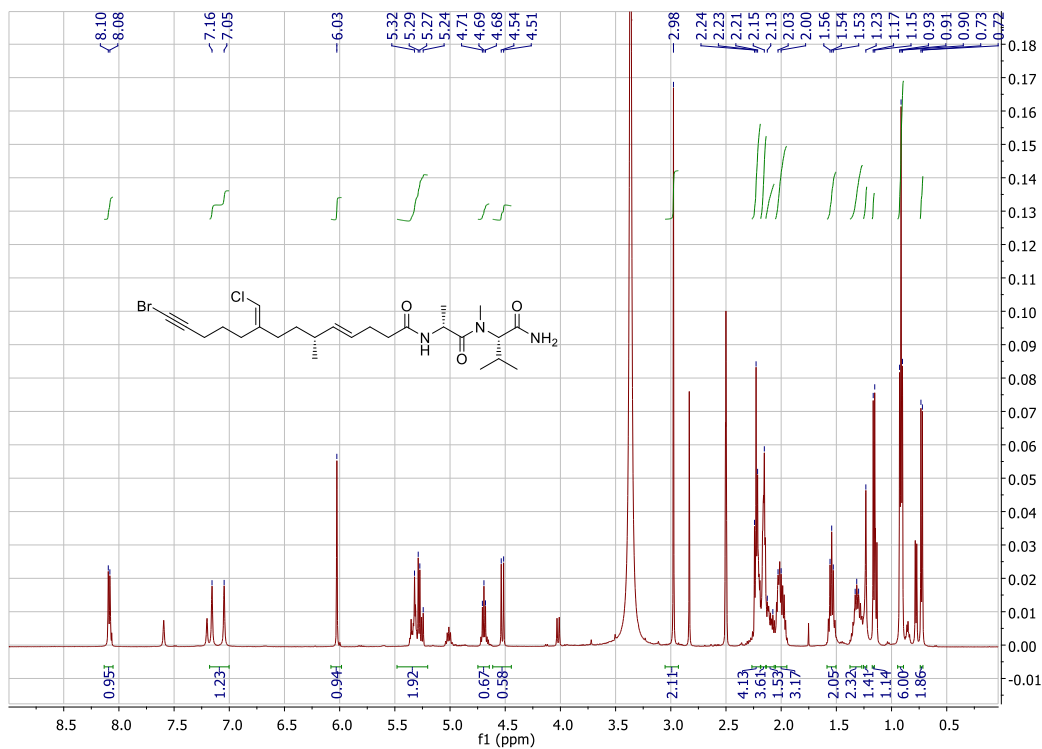


Figure S26. $^1\text{H-NMR}$ spectrum of **6** (500 MHz, $\text{DMSO-}d_6$).

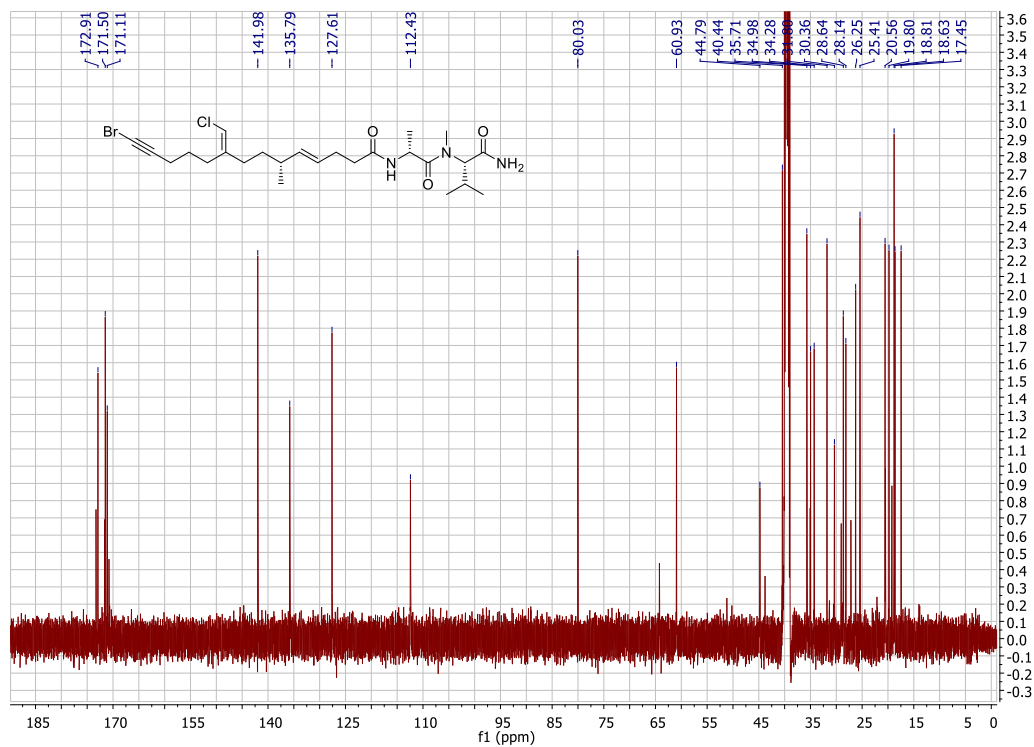


Figure S27. $^{13}\text{C-NMR}$ spectrum of **6** (500 MHz, $\text{DMSO-}d_6$).

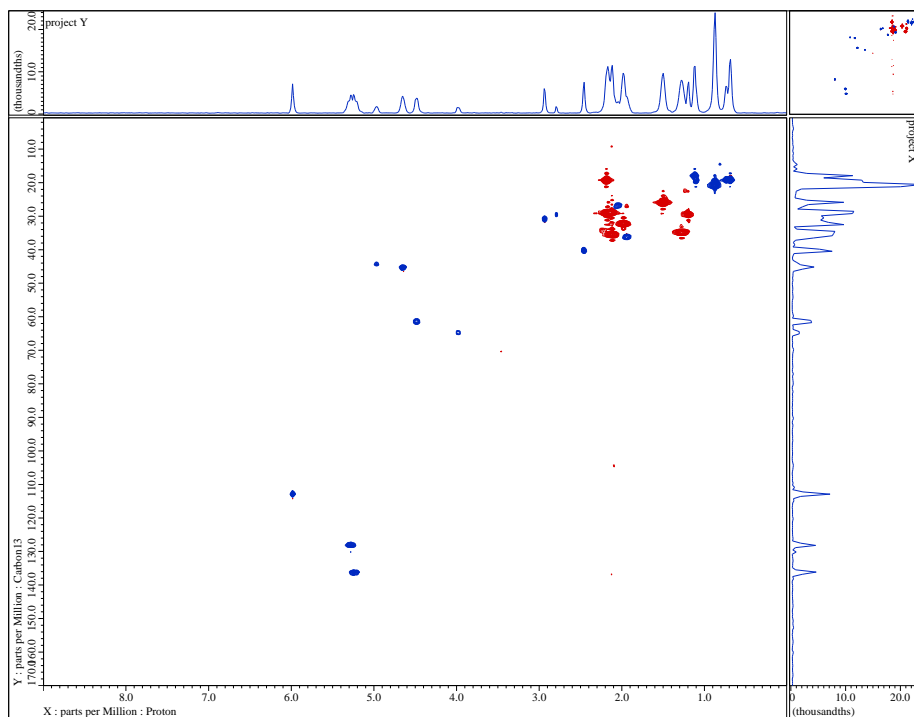


Figure S28. HSQC spectrum of **6** (500 MHz, DMSO- d_6).

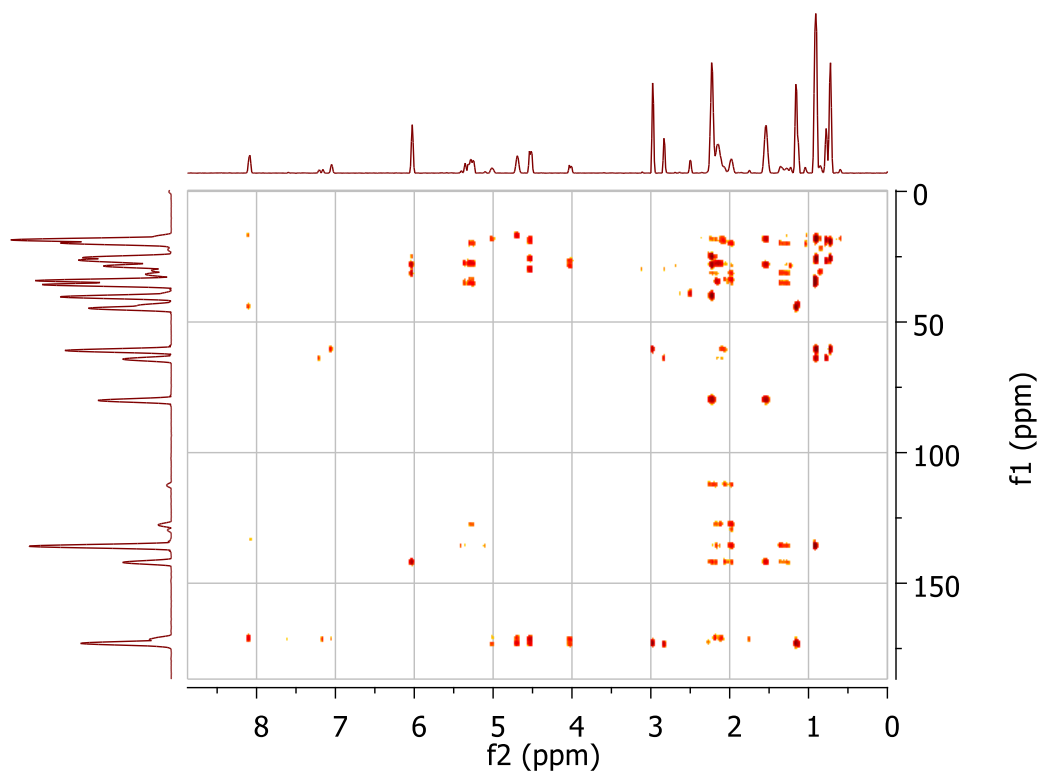


Figure S29. HMBC spectrum of **6** (500 MHz, DMSO- d_6).

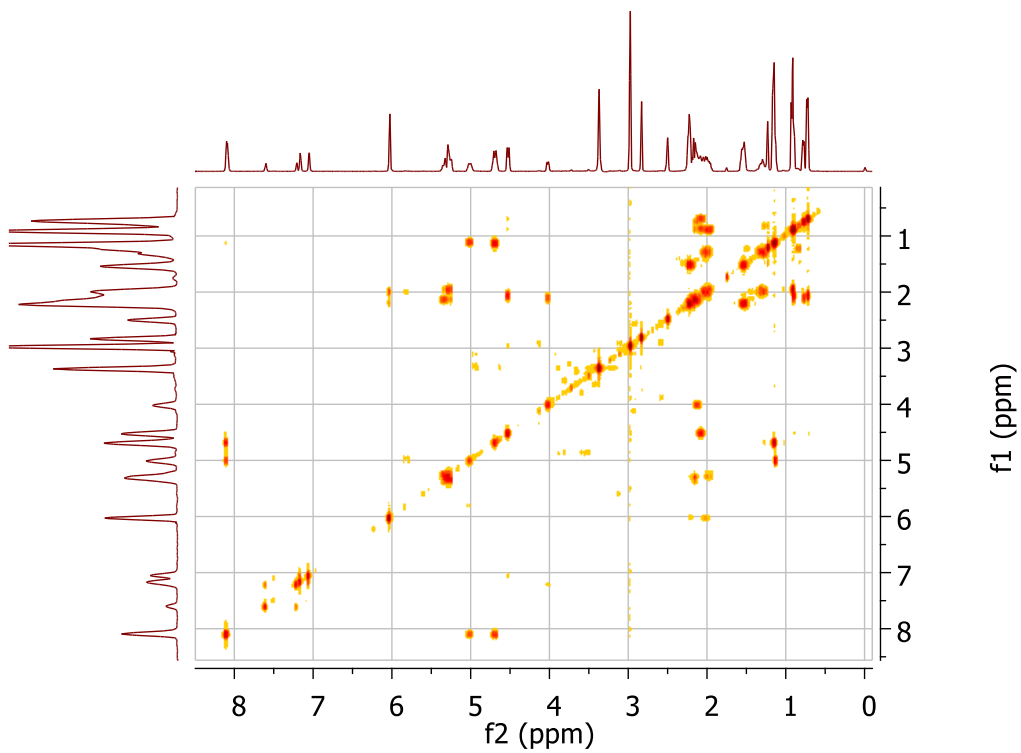


Figure S30. COSY spectrum of **6** (500 MHz, DMSO-*d*₆).

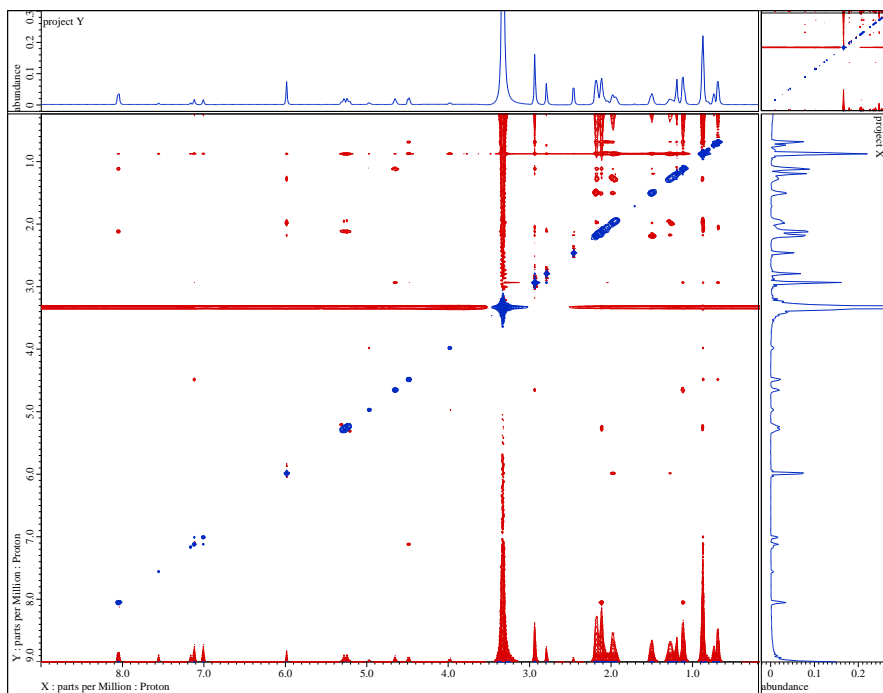


Figure S31. NOESY spectrum of **6** (500 MHz, DMSO-*d*₆).

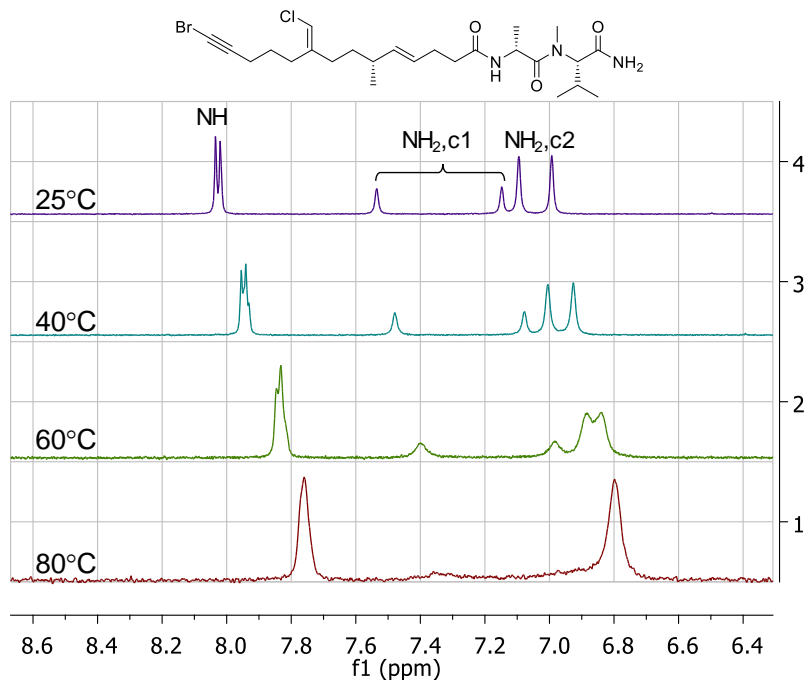


Figure S32. $^1\text{H-NMR}$ spectrum of **6** in $\text{DMSO-}d_6$ under increasing temperature. Terminal amide peaks labeled; c1 indicates conformer 1 and c2 indicates conformer 2.

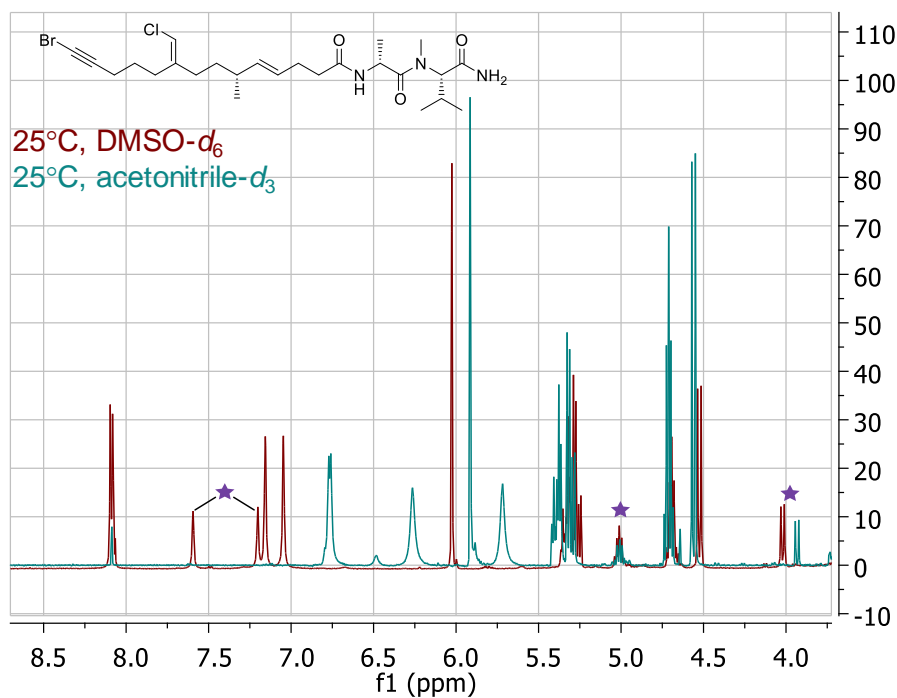


Figure S33. $^1\text{H-NMR}$ spectral comparison of **6** in $\text{DMSO-}d_6$ and $\text{acetonitrile-}d_3$. Minor conformer peaks indicated with purple stars.

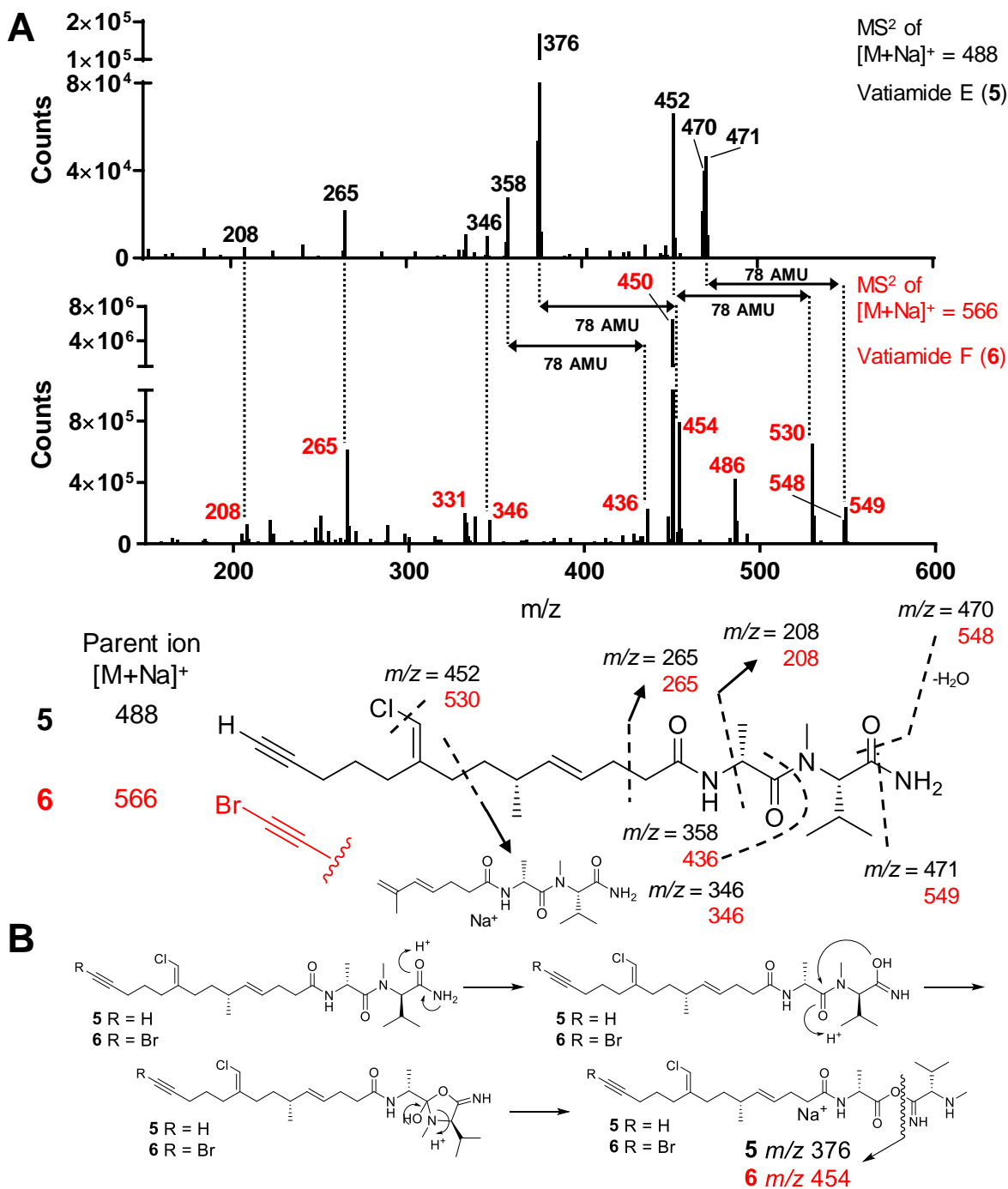


Figure S34. MS² comparison of **5** and **6**. A) Top spectrum, ESI-MS² of vatiamide E, bottom spectrum, ESI-MS² of vatiamide F. Major fragment ions shared or offset by 78 amu indicated by dashed lines. Predicted fragmentation depicted in lower panel, with shared or 78 amu-offset masses deriving from **6** in red. B) Proposed mechanism for formation of base peak $m/z = 376$ (**5**) and $m/z = 454$ (**6**).

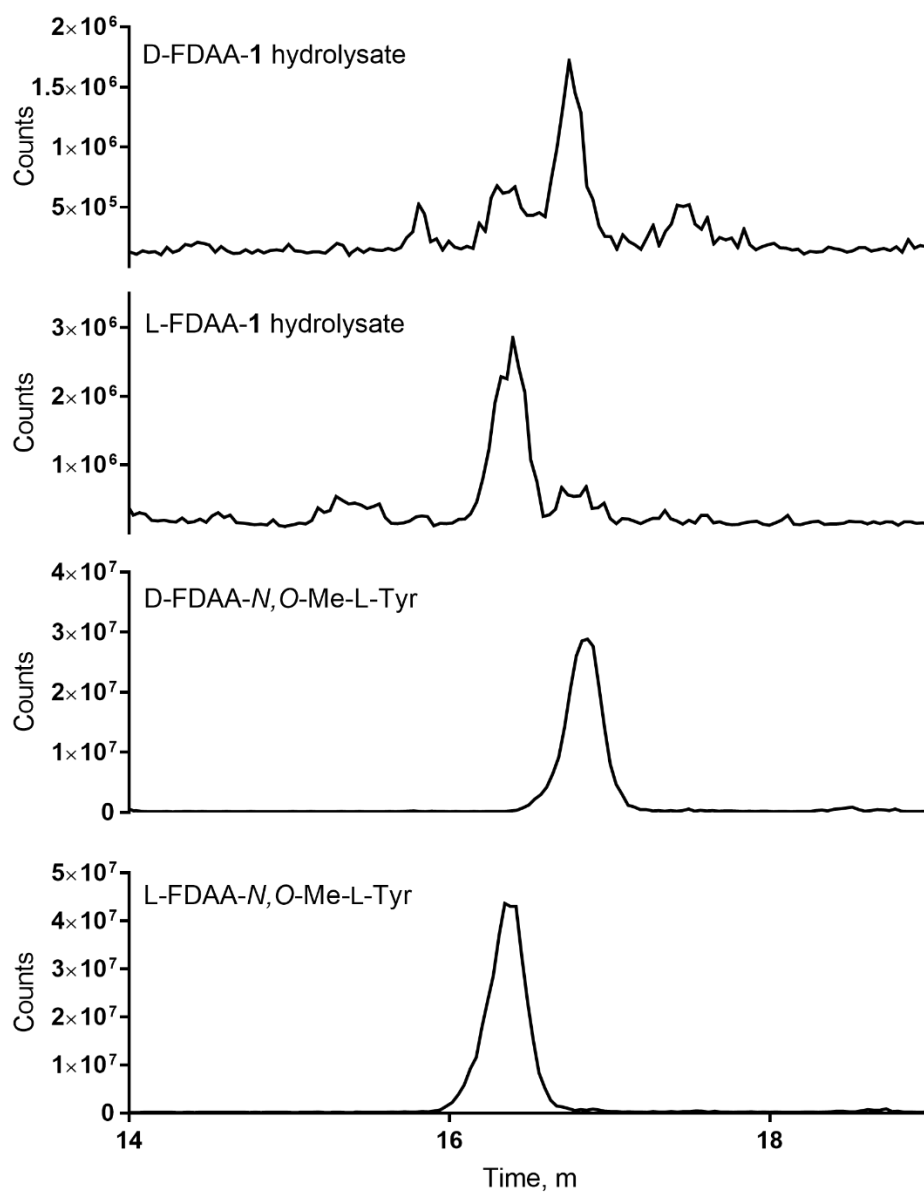


Figure S35. LC/MS chromatogram depicting Marfey's analysis of **1**. $[M+H]^+ = 462$ for each plot corresponding to L-/D-FDAA reaction with hydrolysis products of *N,O,O*-Me-L-Tyr, or **1**.

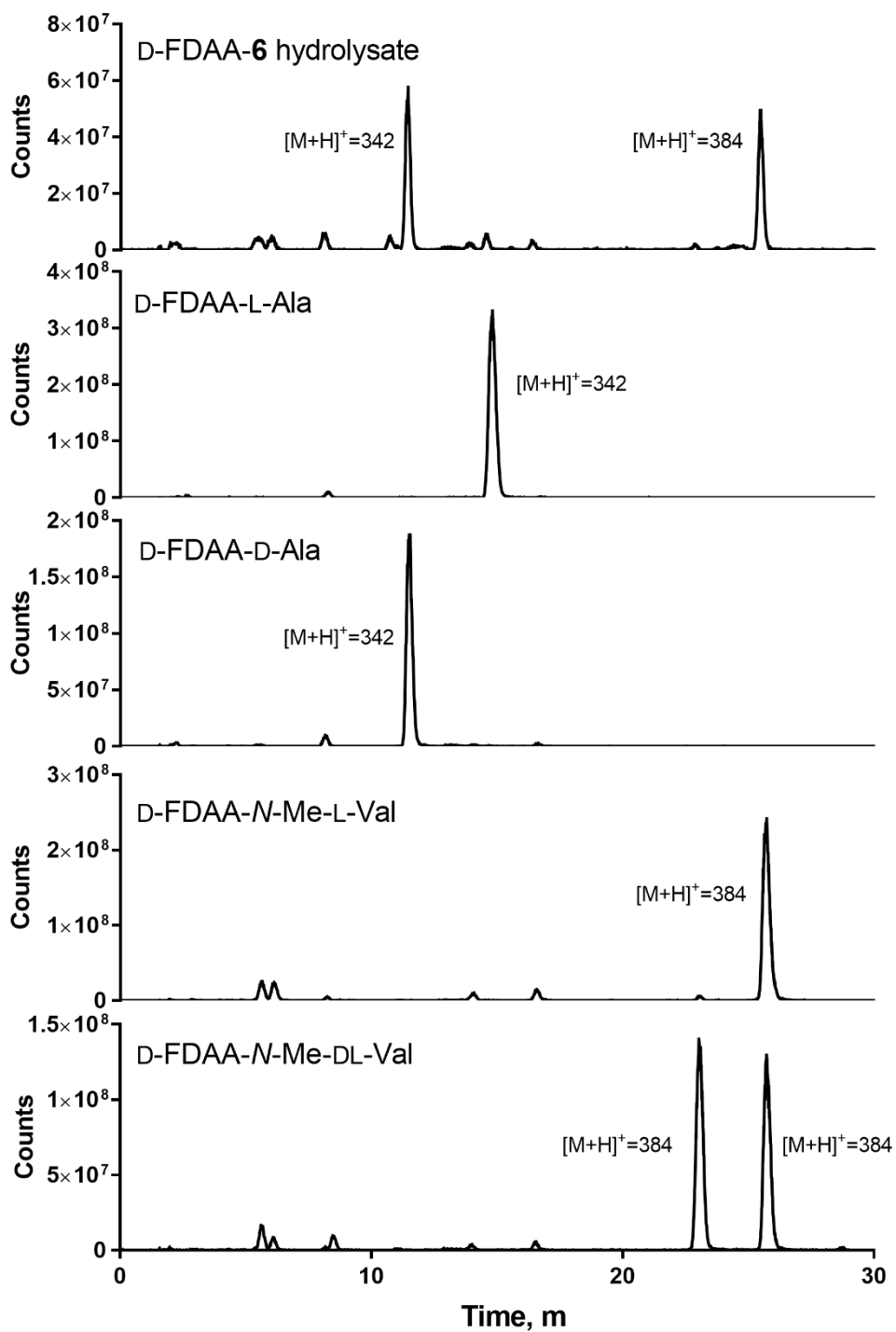


Figure S36. LC/MS chromatogram depicting Marfey's analysis of **6**. $[M+H]^+$ values indicated next to each peak, corresponding to D-FDAA reaction with Ala (342) or *N*-Me-Val (384). Hydrolysis of **6** reacted with D-FDAA shown on top plot.

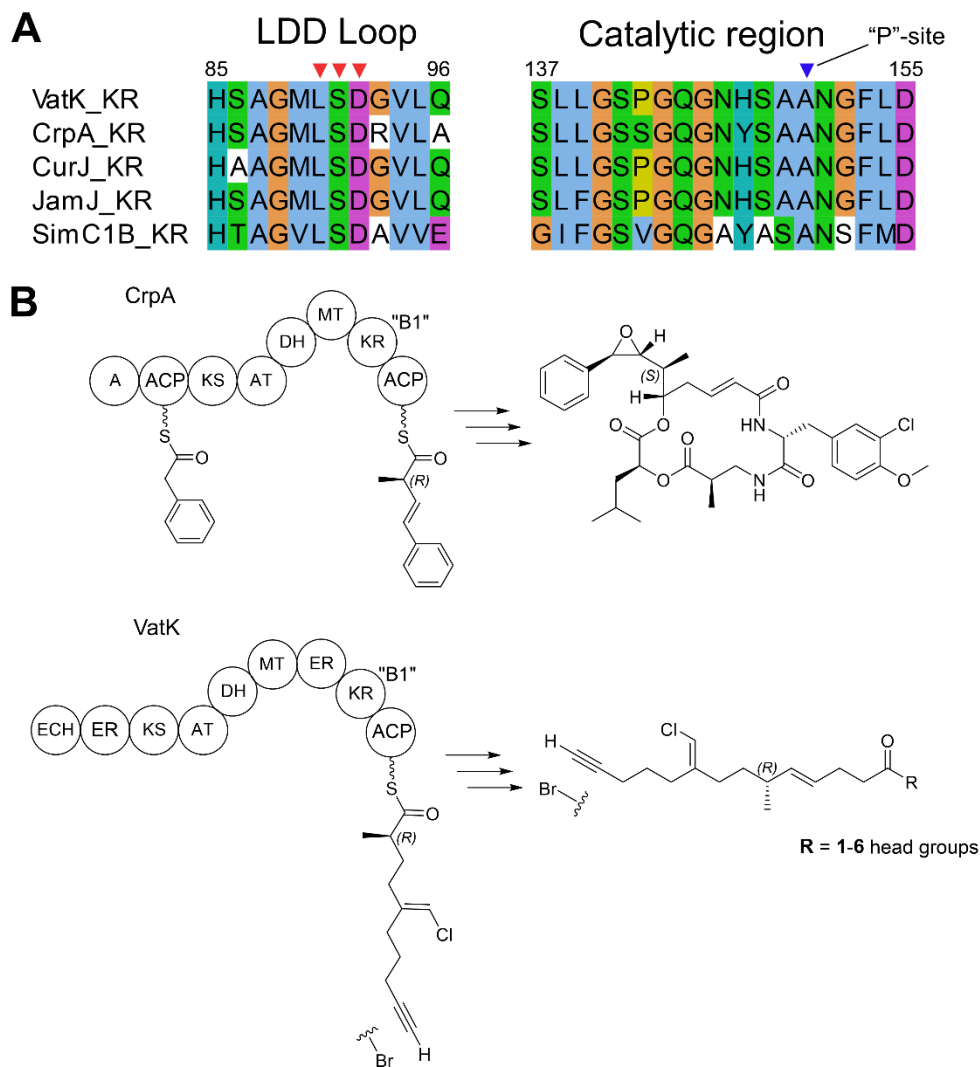


Figure S37. C9 stereochemistry and VatK KR analysis. A) Alignment of LDD loop and catalytic region of VatK and close homologs with D to S substitution. Red triangles indicate LDD motif, dictating A vs. B-type KR, and blue triangle indicates catalytic site 5 “P” position, dictating B1 (No P) vs. B2 (+P) type KR confirmation.^[18,19] Pathway and MiBIG codes as follows: CrpA: cryptophycin, BGC0000975.1, *Nostoc* sp. ATCC 53789; CurJ: curacin, BGC0000976.1, *Moorea producens* 3L; JamJ: jamaicamide, BGC0001001.1, *Moorea producens* JHB; SimC1B: simocyclinone, BGC0001072.1, *Streptomyces antibioticus*. B) Illustration of CrpA activity in cryptophycin biosynthesis and VatK activity in vatiamide biosynthesis.^[20]

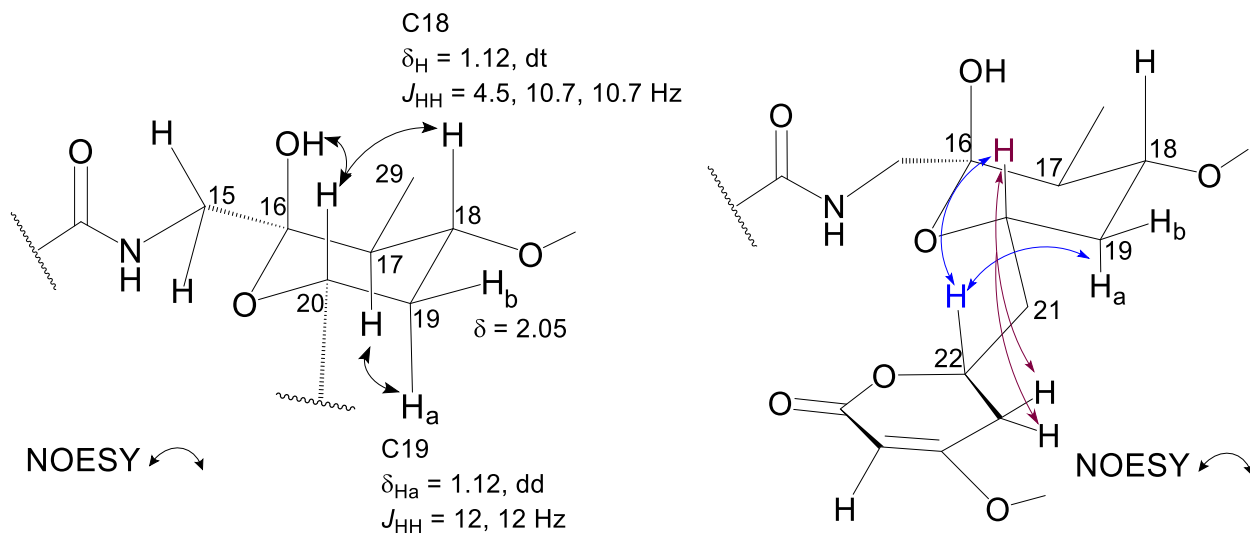


Figure S38. 3D structural analysis of **3**: determination of relative stereochemical assignment of the C16, C17, C18, C20, and C22 chiral centers. Initially, NOE correlation between the C16 hydroxy proton and C20 methine proton (in DMSO-*d*₆, 500 MHz) and C20 and C18 methine protons (CDCl₃, 500 MHz) were used to infer that they are on the same face of the chair-configured lactol. The C19 methylene protons are present at two shifts: Ha $\delta=1.12$ and Hb $\delta=2.05$, with the more deshielded proton likely equatorial owing to anisotropy. *J*-coupling constant values for the C18 methine proton, appearing as a triplet of doublets, indicated that the vicinal protons were in anti ($J = 10.7$ Hz), anti ($J = 10.7$ Hz), and gauche ($J = 4.5$ Hz) relationships. Consistent with the assignment, the C19 Ha proton showed a strong NOE correlation to the C17 proton, inferring they were on the same face of the lactol. Thus, both C19 Ha and the C17 methine proton are axial. The C18 methine proton is also axial, but on the opposite face as C17 and C19, and thus consistent with the two large 10.7 Hz couplings. C19 Hb possesses a coupling value to C18 consistent with this scenario (4.5 Hz). Continuing, C19 Ha shows two vicinal *J*-values of 12 Hz, along with a 12 Hz germinal coupling. As the C18-C19 relationship has been established, this infers that C20 methine proton is anti to the C19 axial Ha proton. This is further supported by the NOE correlation between the C16 hydroxy proton and C20 proton. Therefore, the relative chirality of the lactol ring stereocenters are C16 S*, C17 S*, C18 R*, and C20 S*.

C20 showed a strong NOE to C22, inferring they are on the same face. C22 also shows NOE to C19 Ha, however it could not be discerned whether the C22 methine also correlates by NOE to C19 Hb, as the proton signal for C19 Hb overlaps with one of the protons at C21. Both C19 protons correlate by NOE to the C21 protons, and the C23 protons correlate by NOE to the C21 protons, but not to C19 protons. These correlations are most consistent with R* configuration at C22.

To gain further support for the C22 R* assignment, molecular modeling was used (see figures S39-S40). The preferred conformations of the ring system with C22 as either R* or S* were compared. With C22 set as R* and the energy of the system minimized by MM2 or MMFF94 force field calculations, the C22 methine proton was located on nearly the same face as the C20 methine proton, and with no impeding bonds. On the other hand, setting C22 as S* and energy minimization resulted in the C22 methine proton at 180° to the C20 methine proton, and was impeded from through space interaction by C21 and its two protons (figure S40). Thus, the model supports the experimental data and C22 was assigned as R*.

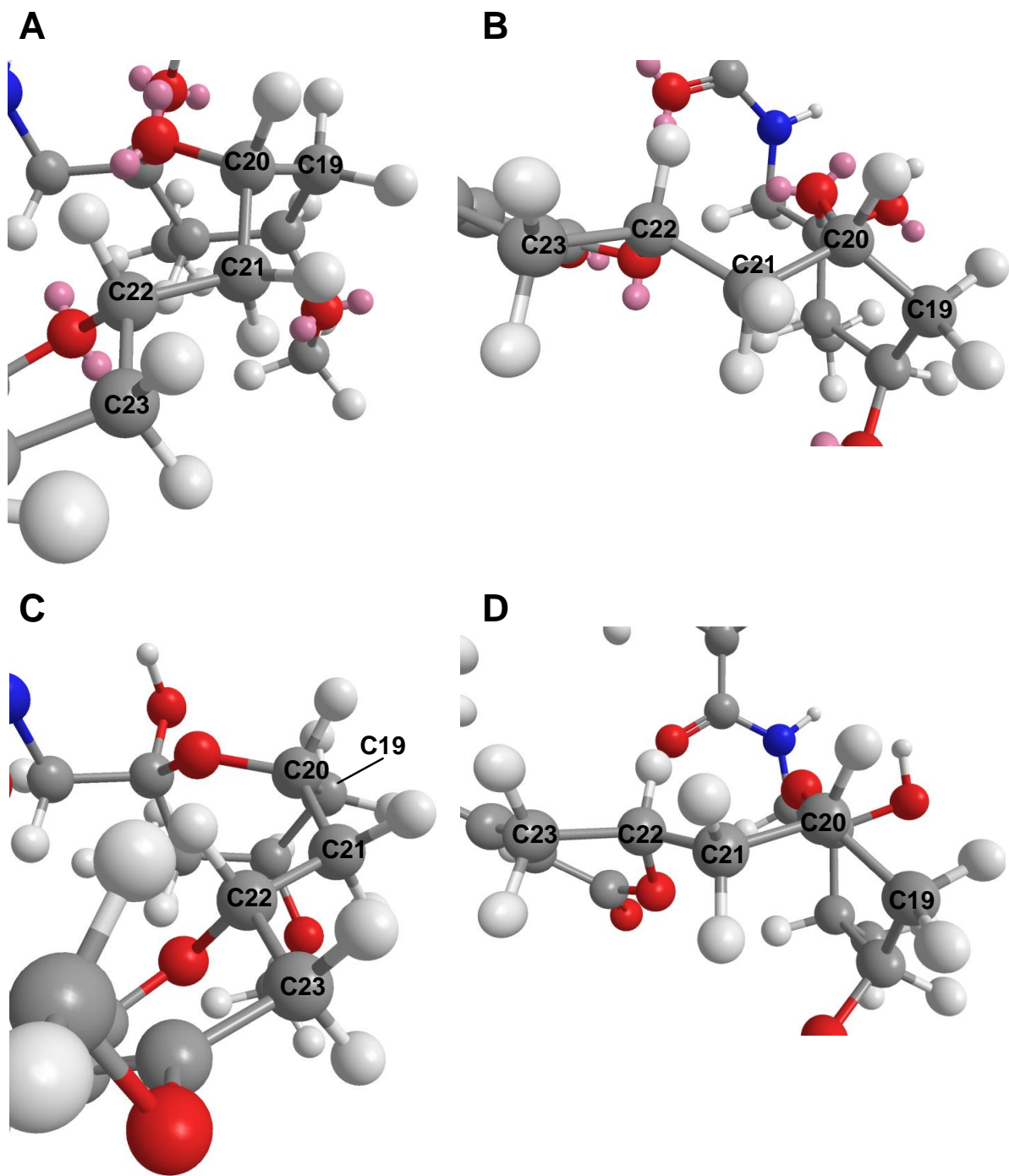


Figure S39. Molecular modeling analysis with C22 stereocenter set to R^* in compound **3**. A) MM2 force field, face view. B) MM2 force field, profile. C) MMFF94 force field, face view. D) MMFF94 force field, profile. The C20-C22 coplanar relationship in this model with C22 R^* is supported by observed NOESY correlation between these protons. (Figure S17).

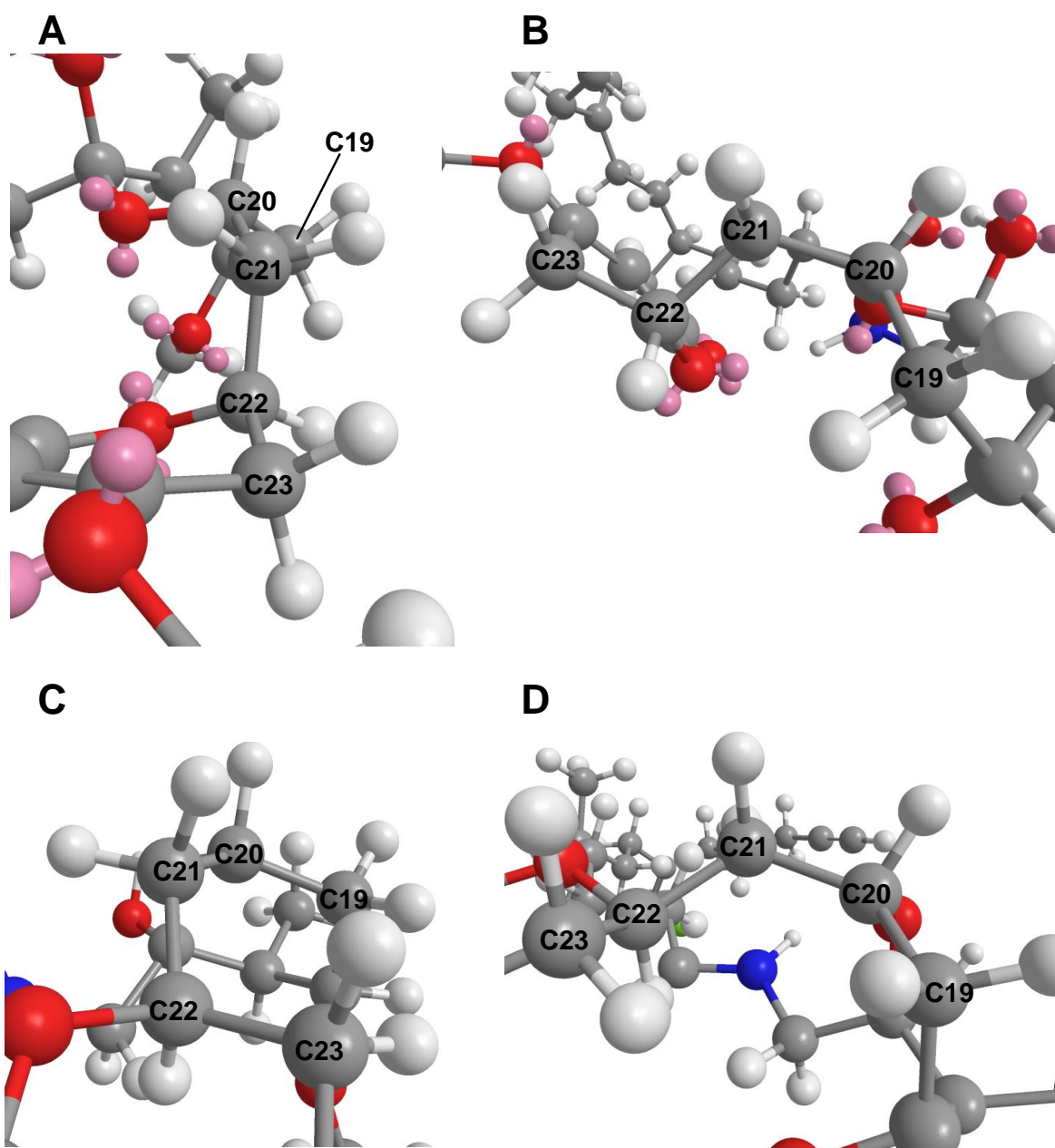


Figure S40. Molecular modeling analysis with C22 stereocenter set to S^* in compound **3**. A) MM2 force field, face view. B) MM2 force field, profile. C) MMFF94 force field, face view. D) MMFF94 force field, profile. This model with C22 S^* configuration indicates probable C20-C22 *trans* relationship, and is inconsistent with the observed NOESY correlation between these protons (Figure S17).

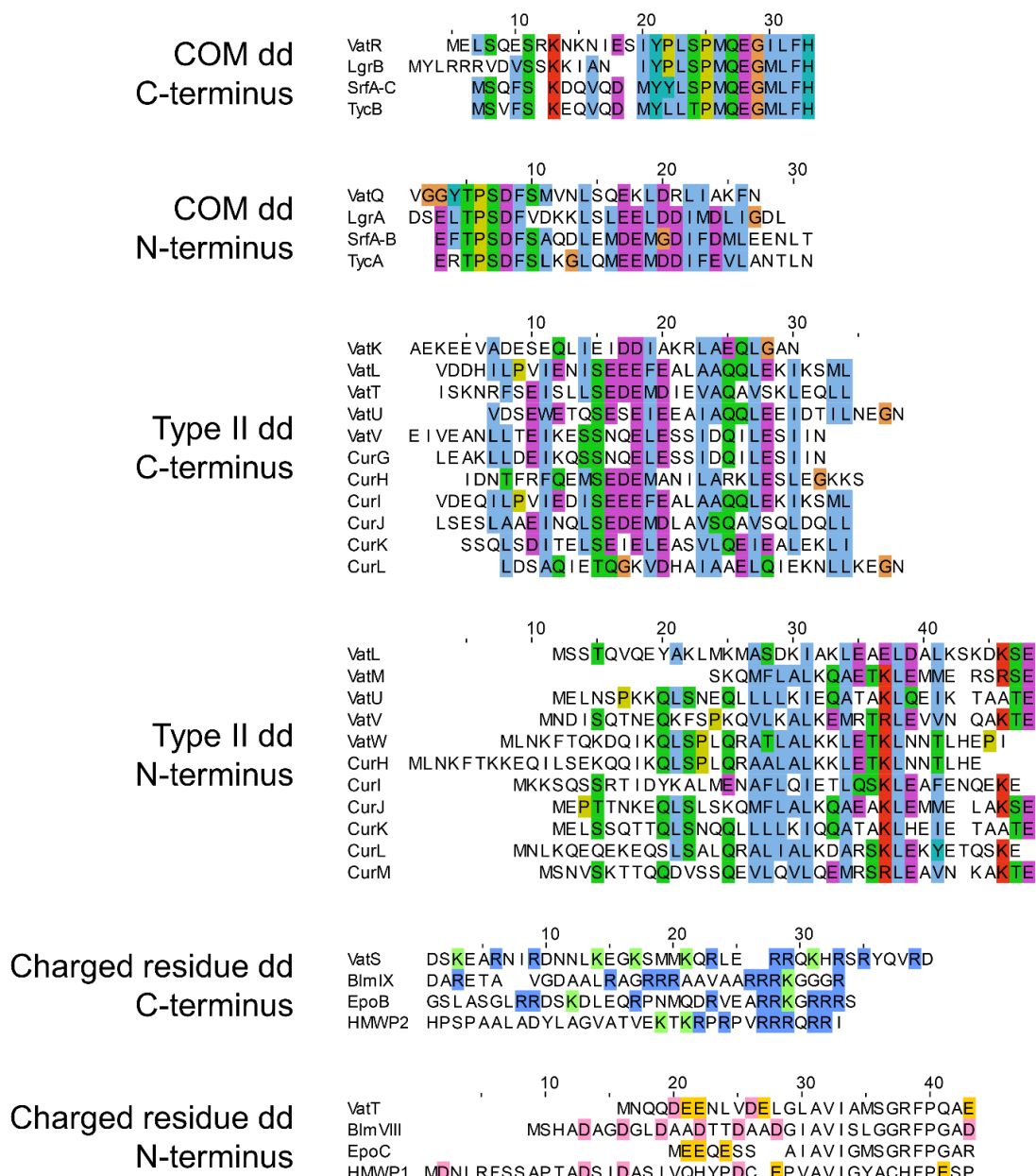


Figure S41. Alignment of all PKS or NRPS docking domains in Vat pathway with experimentally characterized examples: COM,^[21] PKS type II,^[22] and EpoB-EpoC^[23] charged-residue type.^[24] COM and Type II alignments are colored by CLUSTAL algorithm, while in the EpoB-EpoC “charged residue” type interaction, Arg and Lys residues are highlighted in blue and green respectively in the C-terminal dd alignment, while Asp and Glu are in pink and orange, respectively, in the N-terminal dd alignment. Abbreviation: dd = docking domain.

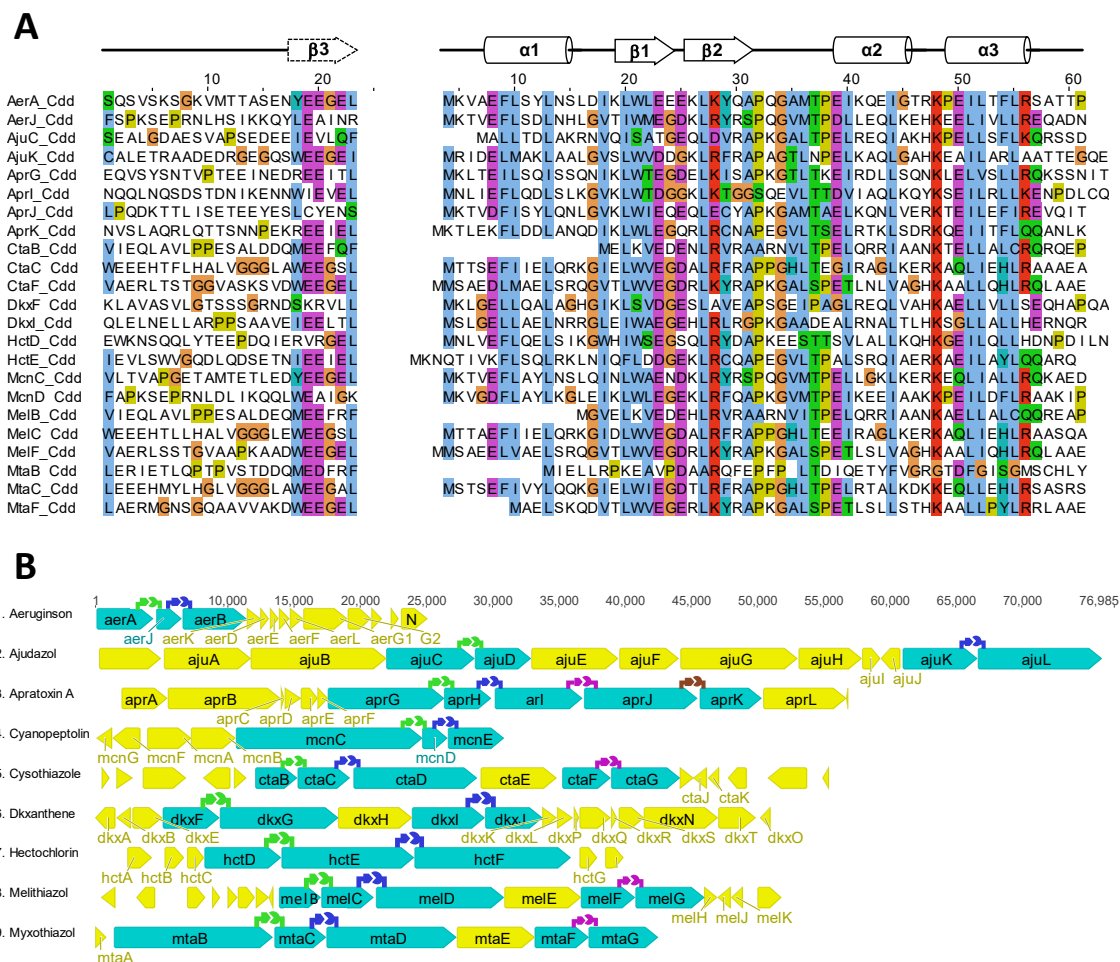


Figure S42. Docking domain alignments and selective PKS/NRPS pathways from MIBiG and NCBI. Pathways contain >1 Cdd-Ndd interactions of the β - $\alpha\beta\beta\alpha$ type within a single biosynthetic gene cluster. A) Alignment of C-terminal docking domains on left and N-terminal docking domains on right, with tertiary protein structure graphics above each alignment. B) Graphic of each pathway PKS/NRPS core, with docking domain cognate pairs depicted by colored green (first), blue (second), purple (third), or brown (fourth) cartoon above interacting genes. Separate interactions are indicated by different colors. Pathway and organism, and MIBiG/NCBI identifier is as follows: aeruginon - *Microcystis aeruginosa* NIES-98, BGC0000298; ajudazol - *Chondromyces crocatus*, BGC0000954; apratoxin A - *Moorea bouillonii* PNG5-198, NCBI: txid568701; cyanopeptolin - *Microcystis* sp. NIVA-CYA 172/5, BGC0000332; cystothiazole - *Cystobacter fuscus*, BGC0000982; dkxanthene - *Stigmatella aurantiaca* DW4/3-1, BGC0000986; hectochlorin - *Moorea producens* JHB, BGC000100; melithiazol - *Melittangium lichenicola*, BGC0001010; myxothiazol - *Stigmatella aurantiaca* DW4/3-1, BGC0001024.

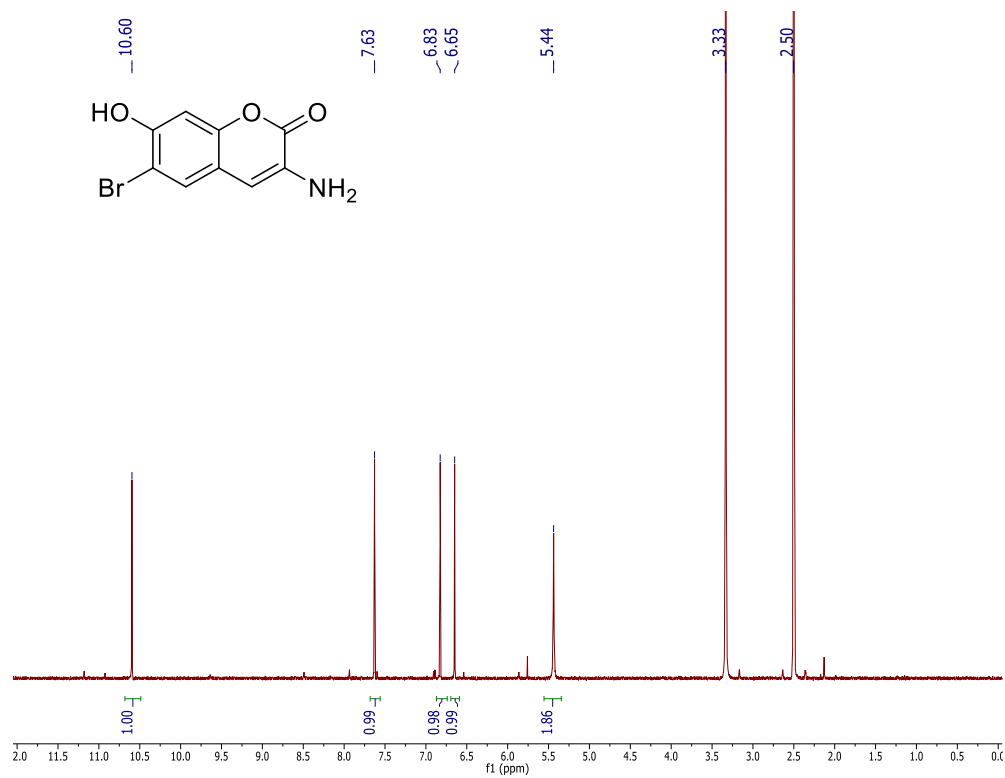


Figure S43. ¹H NMR spectrum of **13** (500 MHz, DMSO-*d*₆).

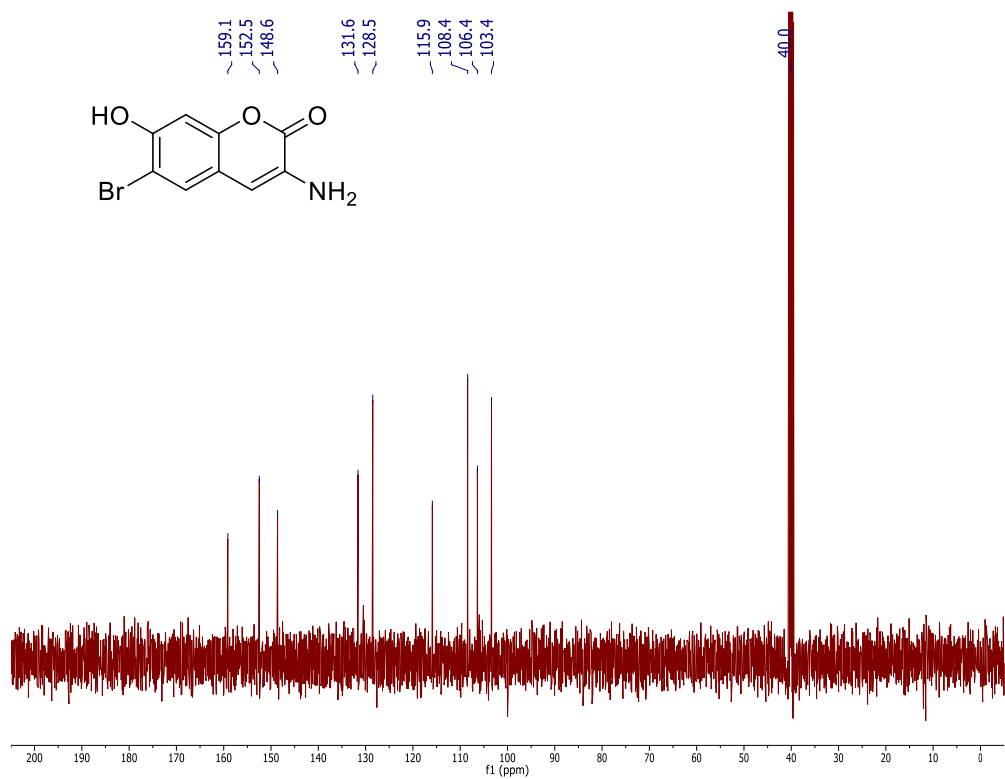


Figure S44. ¹³C NMR spectrum of **13** (500 MHz, DMSO-*d*₆).

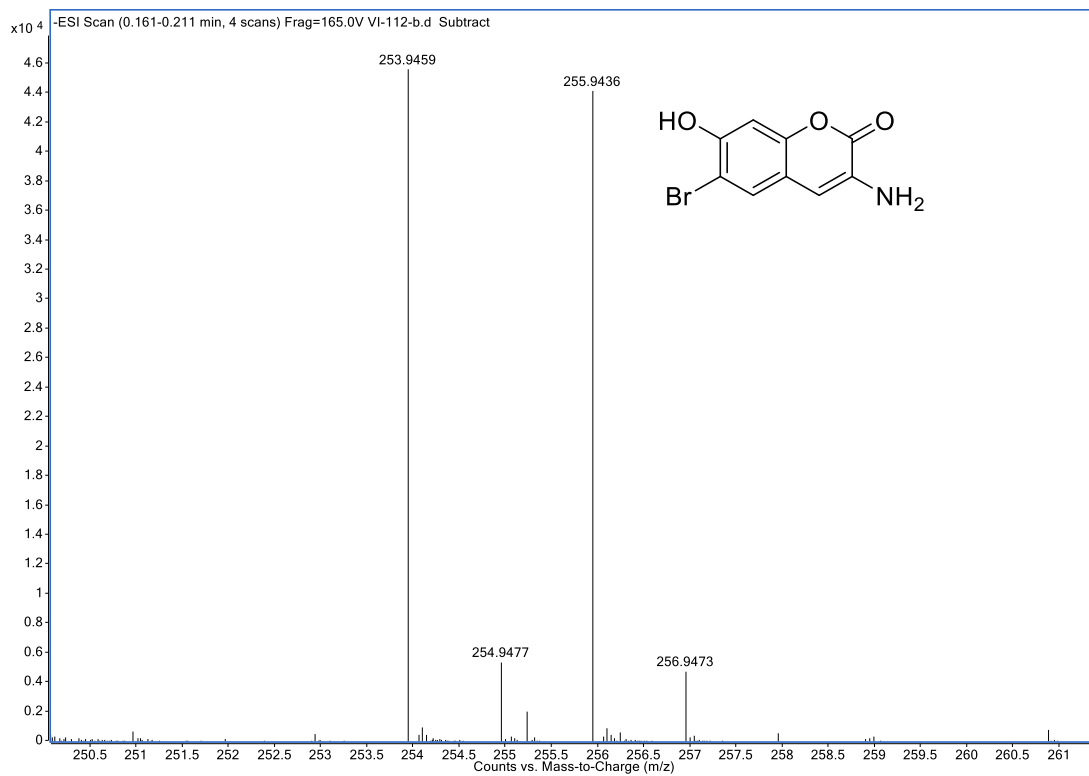


Figure S45. HRMS spectrum of **13**.

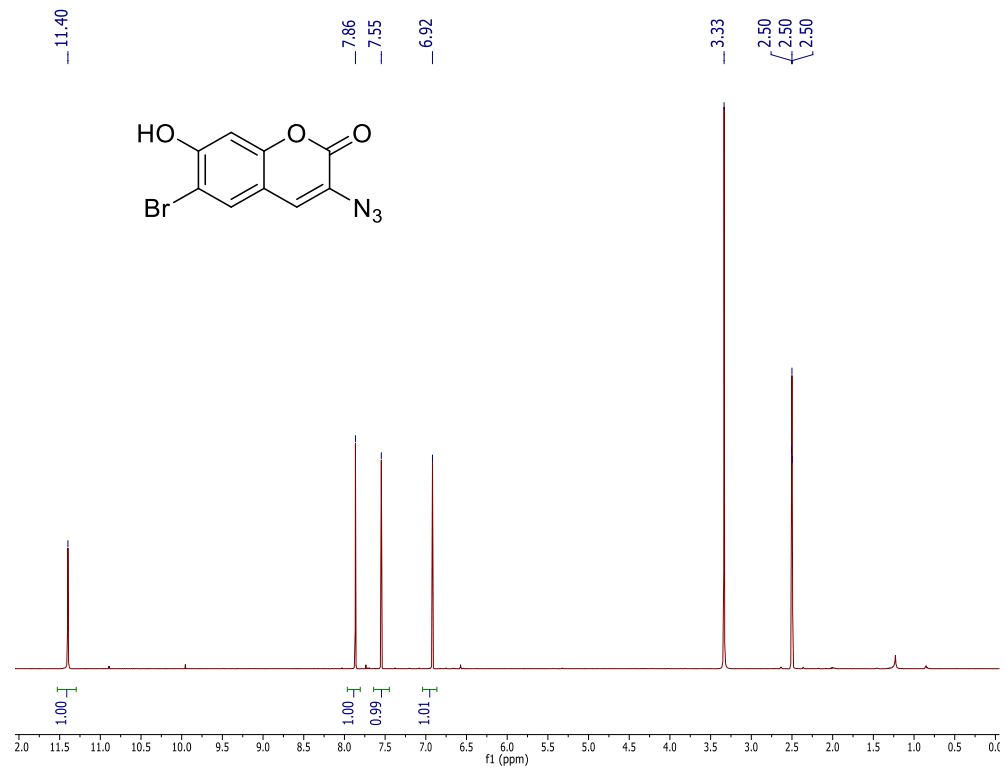


Figure S46. ¹H NMR spectrum of **14** (500 MHz, DMSO-*d*₆).

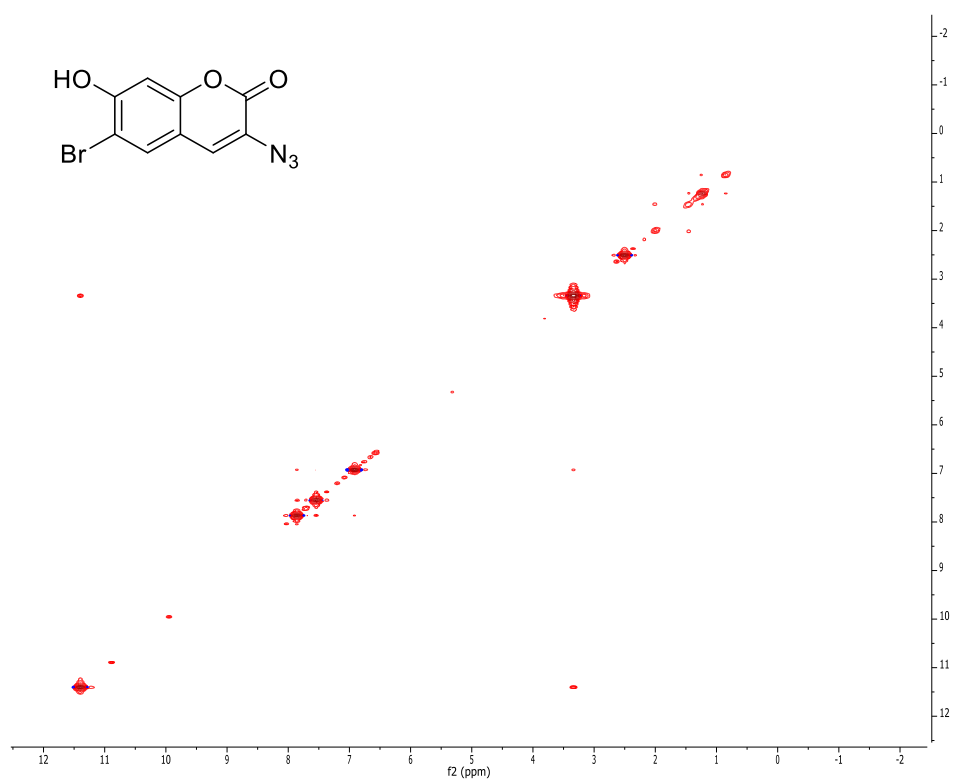


Figure S47. COSY NMR spectrum of **14** (500 MHz, DMSO-*d*₆).

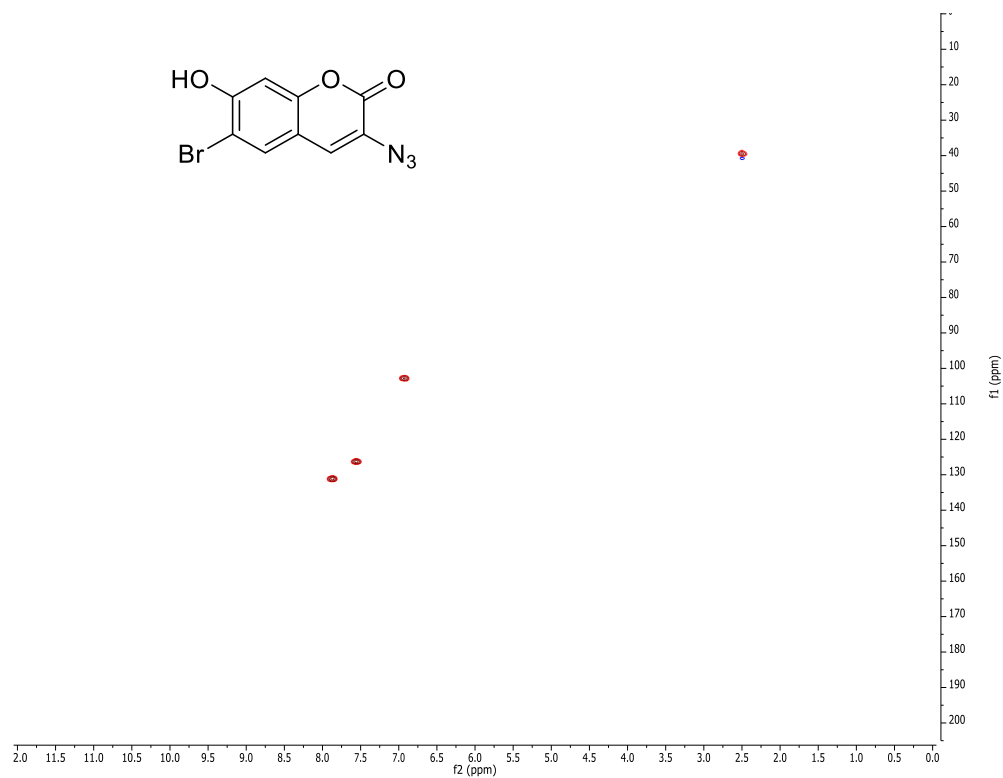


Figure S48. HSQC NMR spectrum of **14** (500 MHz, DMSO- d_6).



Figure S49. HMBC NMR spectrum of **14** (500 MHz, DMSO- d_6).

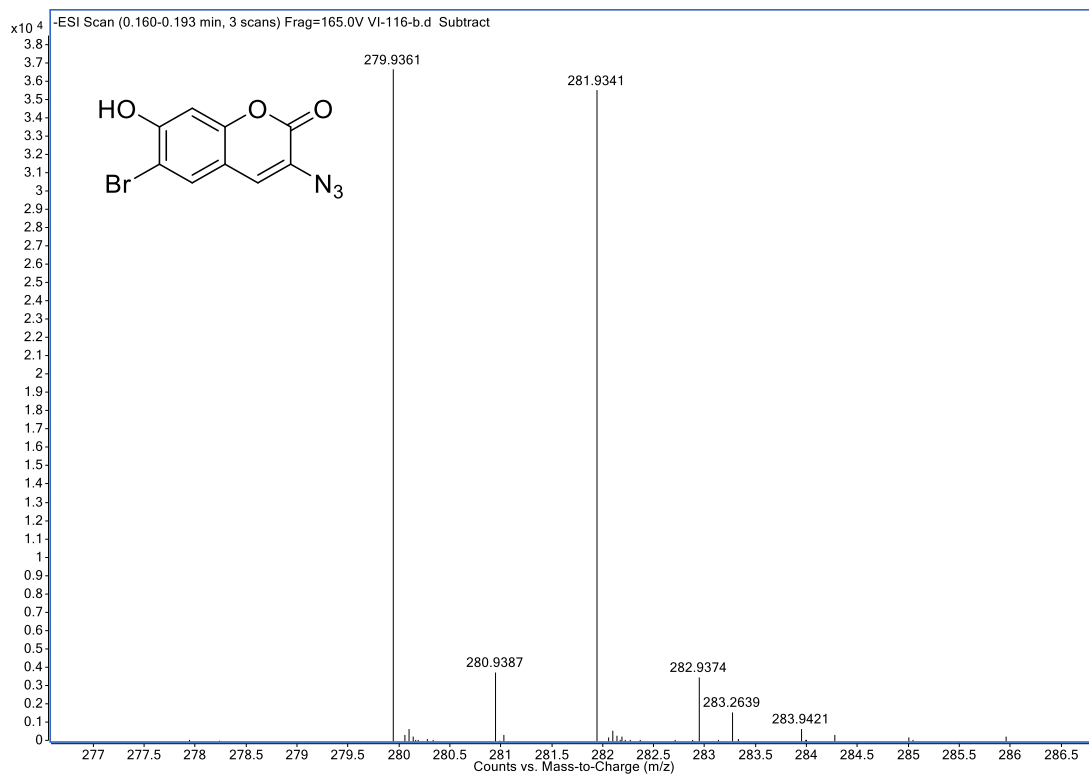


Figure S50. HRMS spectrum of 14.

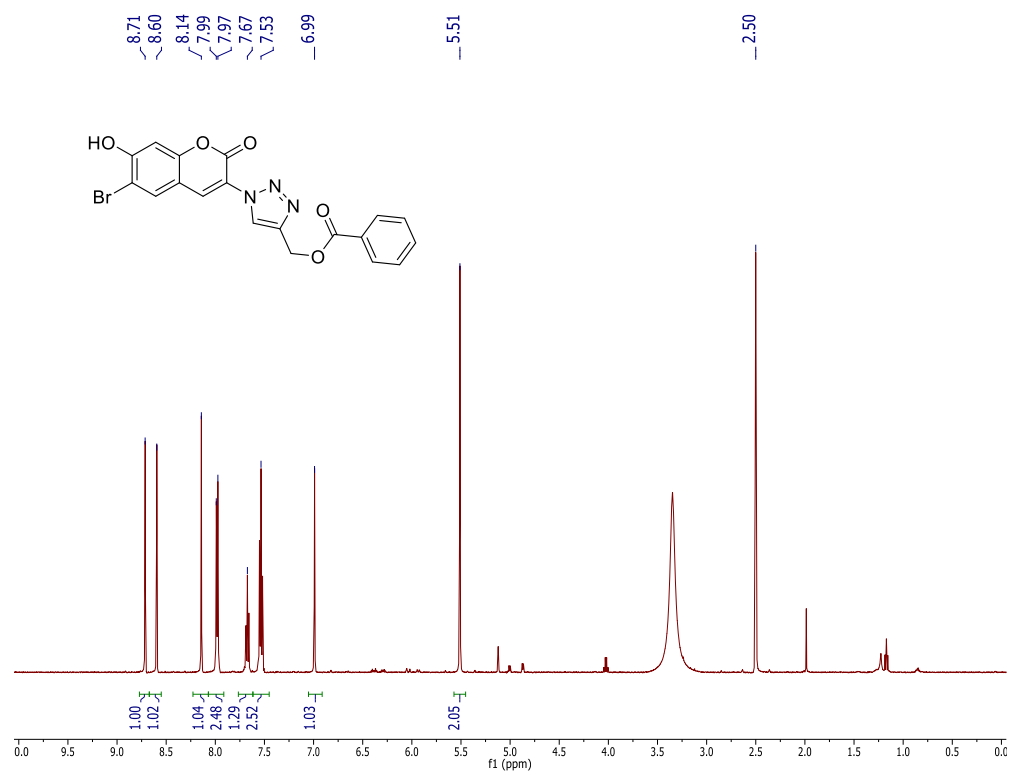


Figure S51. ¹H NMR spectrum of 16 (500 MHz, DMSO-*d*₆).

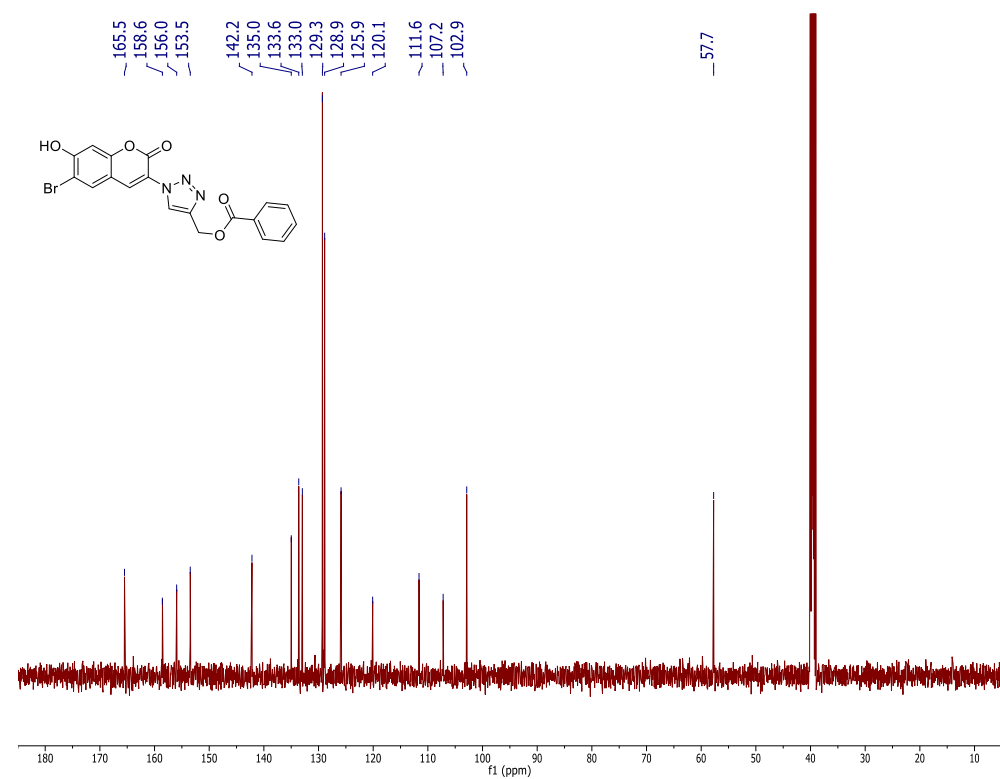


Figure S52. ¹³C NMR spectrum of 16 (500 MHz, DMSO-*d*₆).

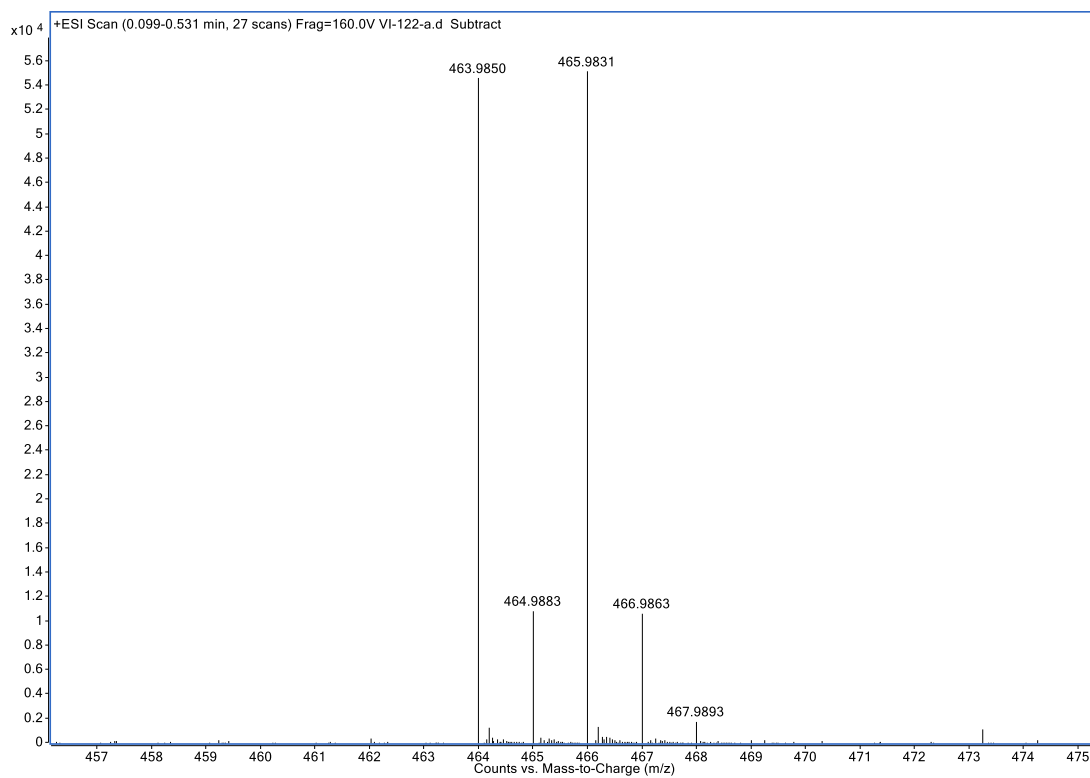


Figure S53. HRMS spectrum of **16**.

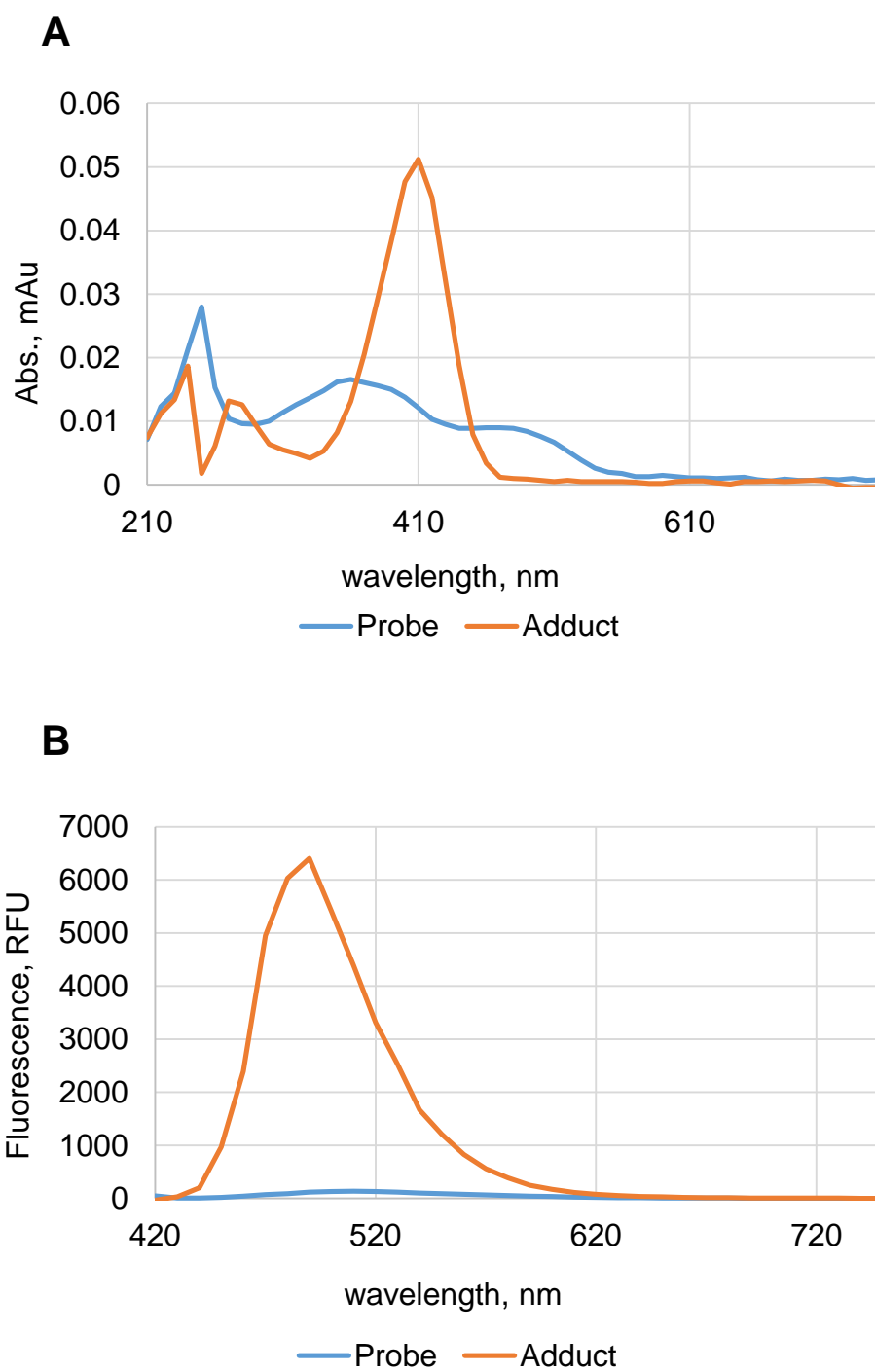


Figure S54. Probe and adduct spectroscopy. A) UV absorbance measurements for probe 14 and adduct 16. B) Fluorescence measurements for probe 14 and adduct 16, excitation at 410 nm. Concentration for each experiment was 2 μ M in 1:1 DMF:H₂O.

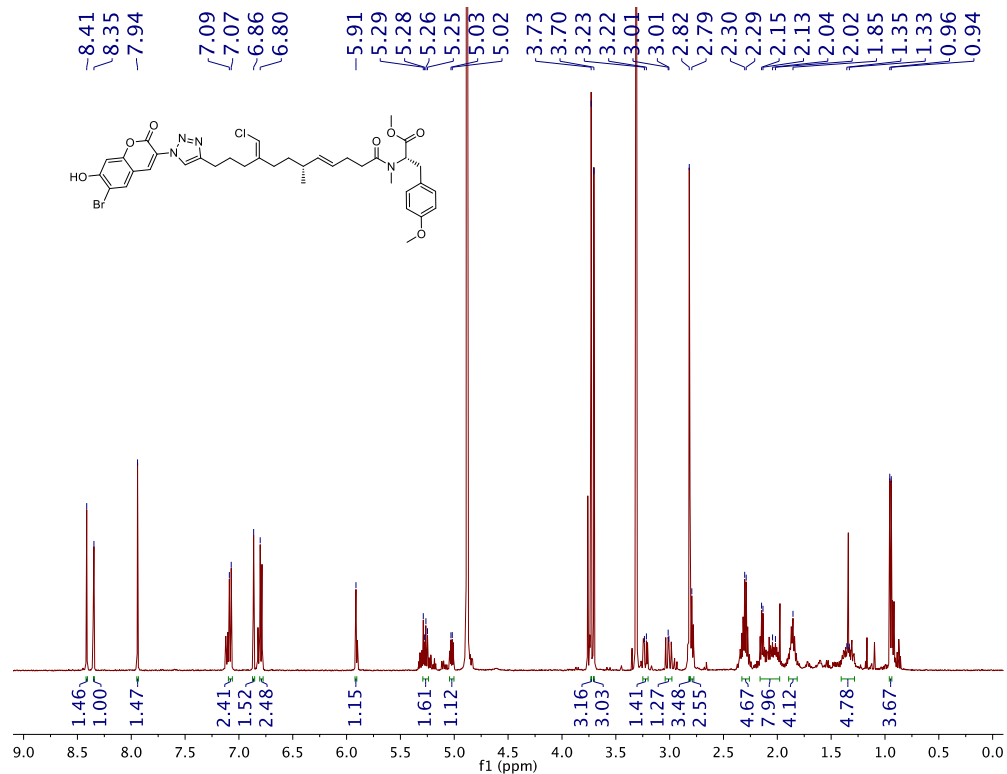


Figure S55. ^1H NMR of 14+1 azido coumarin probe adduct (500 MHz, d_4 -methanol).

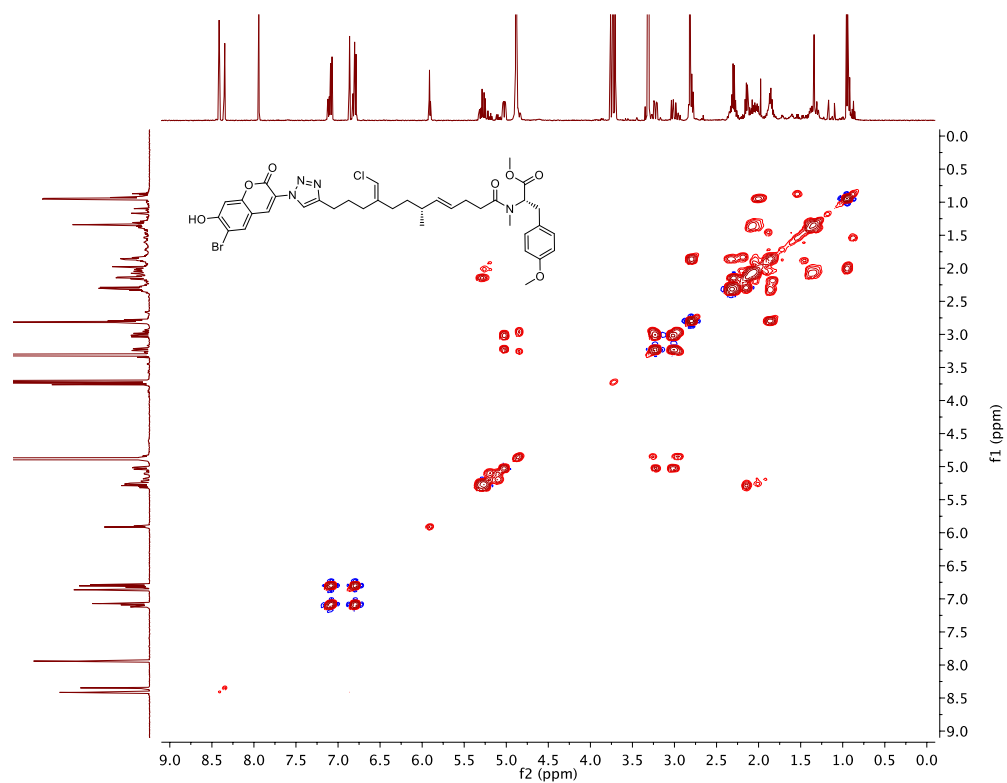


Figure S56. COSY of 14+1 azido coumarin probe adduct (500 MHz, d_4 -methanol).

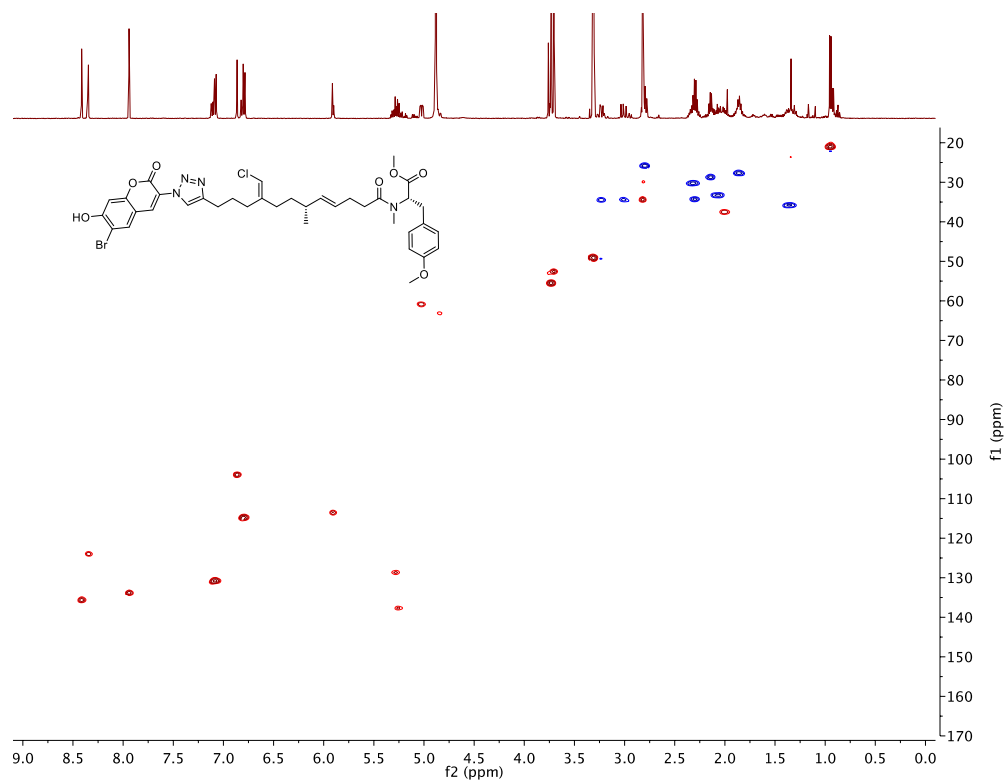


Figure S57. HSQC of **14+1** azido coumarin probe adduct (500 MHz, *d*₄-methanol).

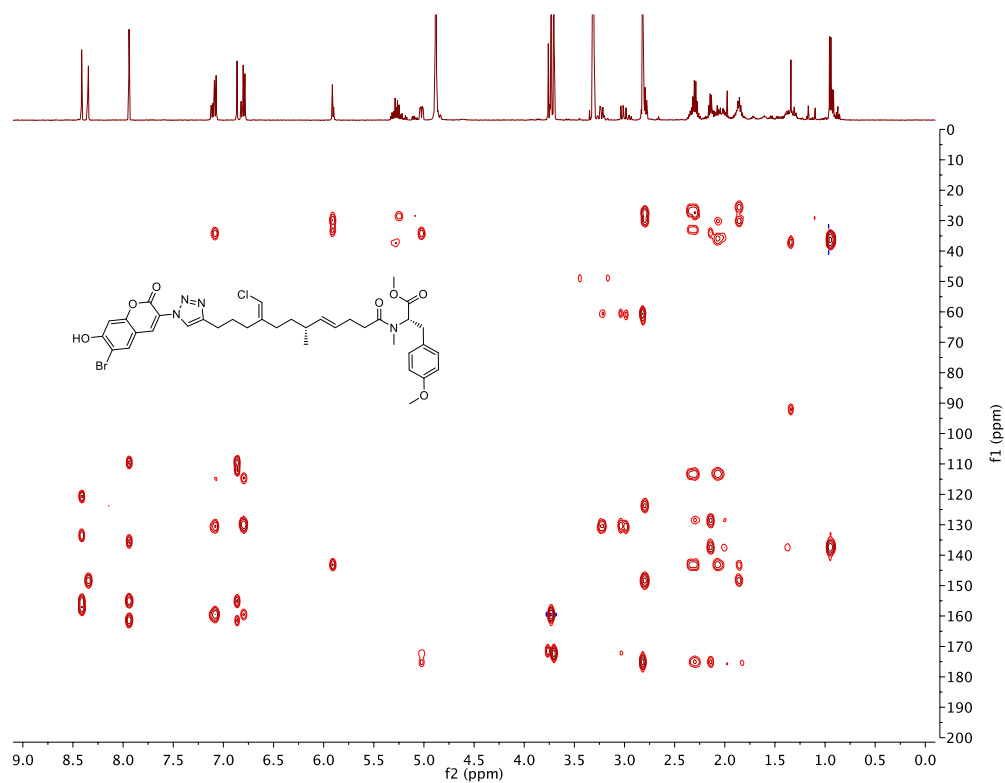


Figure S58. HMBC of **14+1** azido coumarin probe adduct (500 MHz, *d*₄-methanol).

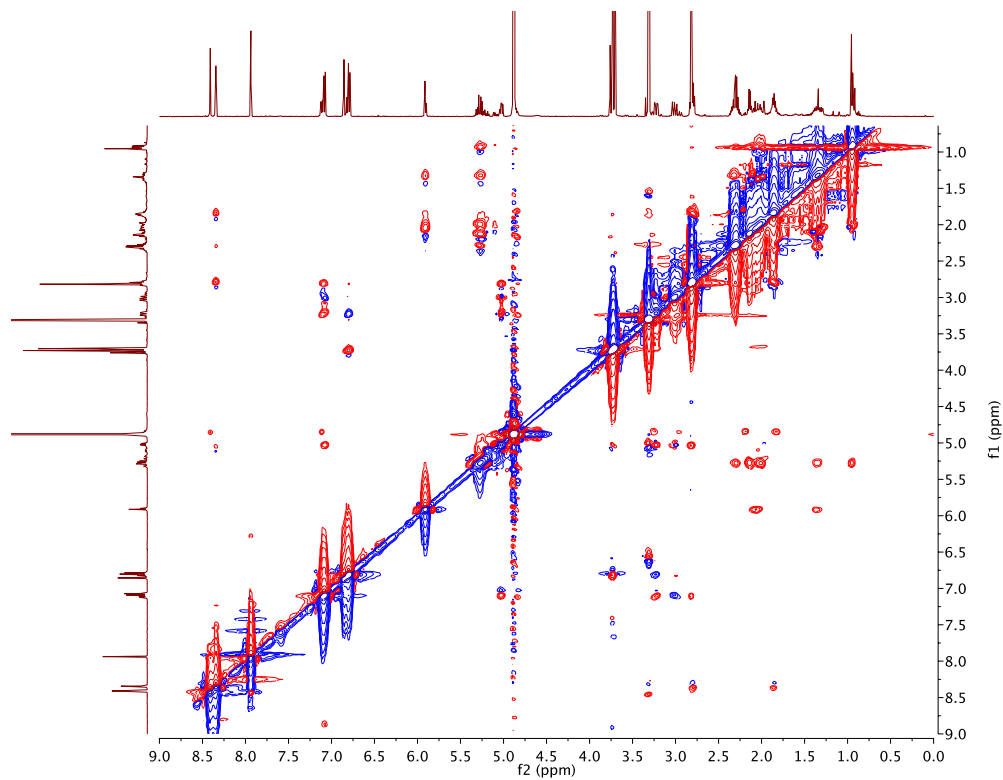


Figure S59. ROESY of **14+1** azido coumarin probe adduct (500 MHz, d_4 -methanol).

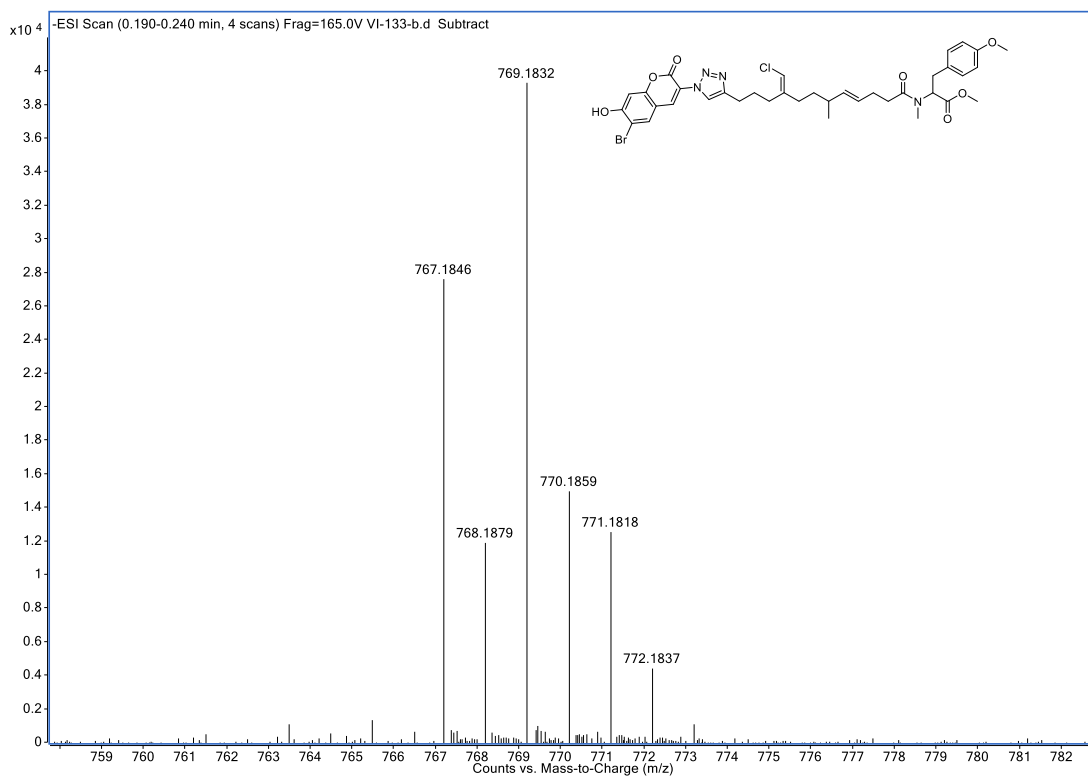


Figure S60. HRMS spectrum of **14+1**.

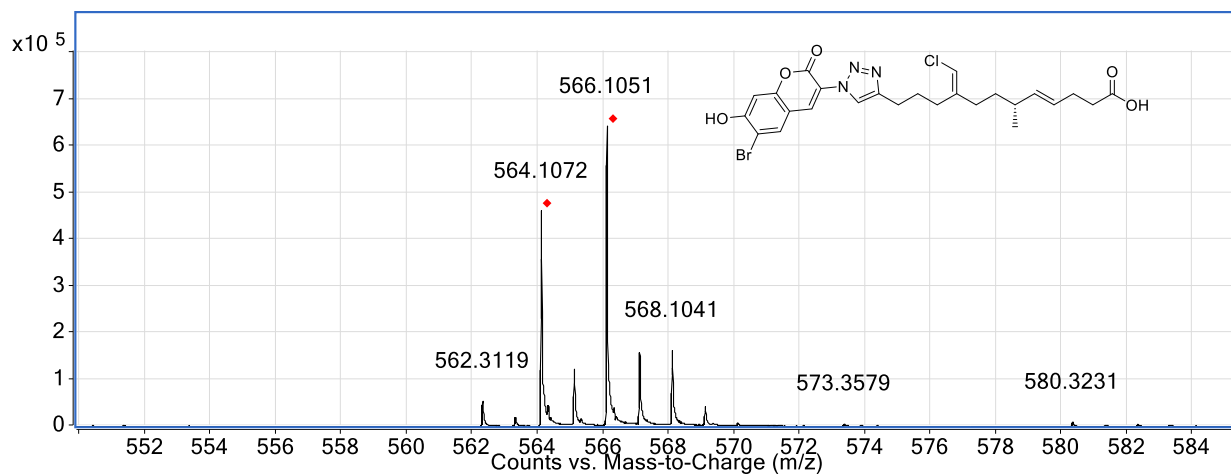


Figure S63. HRMS spectrum of **14+7**.

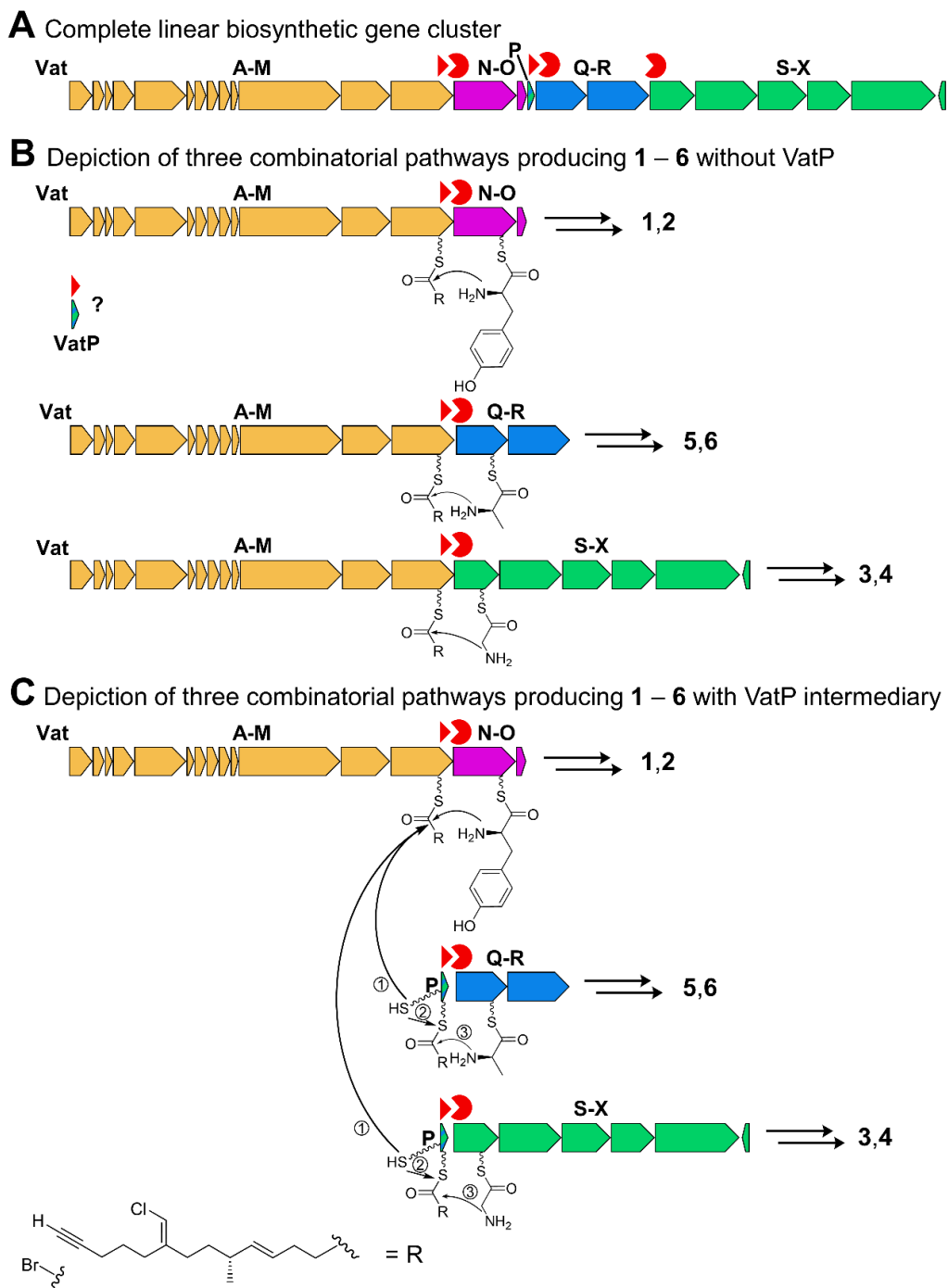


Figure S64. Depiction of biosynthetic module organization. Red triangles indicate identical Cdds, red notched circles indicate identical Ndds. A) *vat* pathway depicted in native genetic context B) Depiction of scheme whereby VatM Cdd interacts separately with Ndds at VatN-O (producing **1,2**), VatQ-R (producing **5,6**), and VatS-X (producing **3,4**). C) Depiction of scheme whereby VatM Cdd interacts with Ndd at VatN-O (producing **1,2**), while VatP ACP attacks VatM-bound substrate, then interacts with VatQ-R (producing **5,6**), and VatS-X (producing **3,4**).

References

- [1] N. Engene, E. C. Rottacker, J. Kaštovský, T. Byrum, H. Choi, M. H. Ellisman, J. Komarek, W. H. Gerwick, *Int. J. Syst. Evol. Microbiol.* **2012**, *62*, 1171-1178.
- [2] R. W. Castenholz, *Methods Enzymol.* **1988**, *167*, 68-93.
- [3] T. C. Glenn, R. Nilsen, T. J. Kieran, J. W. Finger, T. W. Pierson, K. E. Bentley, S. Hoffberg, S. Louha, F. J. Garcia-De-Leon, M. Angel del Rio Portilla, et al., *Adapterama I: Universal Stubs and Primers for Thousands of Dual-Indexed Illumina Libraries (Itru & Inext)*, **2016**.
- [4] A. Bankevich, S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, et al., *J. Comput. Biol.* **2012**, *19*, 455-477.
- [5] S. Podell, T. Gaasterland, *Genome Biol.* **2007**, *8*, R16.
- [6] G. M. Boratyn, A. A. Schaffer, R. Agarwala, S. F. Altschul, D. J. Lipman, T. L. Madden, *Biol. Direct* **2012**, *7*, 1-14.
- [7] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, *Nat. Protoc.* **2015**, *10*, 845-858.
- [8] N. L. Allinger, *J. Am. Chem. Soc.* **1977**, *99*, 8127-8134.
- [9] T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 490-519.
- [10] T. Mosmann, *J. Immunol. Methods* **1983**, *65*, 55-63.
- [11] B. Meyer, N. Ferrigni, J. Putnam, L. Jacobsen, D. Nichols, J. McLaughlin, *Planta Med.* **1982**, *45*, 31-34.
- [12] B. Austin, D. A. Allen, *Aquaculture* **1982**, *26*, 369-383.
- [13] W. H. Gerwick, P. J. Proteau, D. G. Nagle, E. Hamel, A. Blokhin, D. L. Slate, *J. Org. Chem.* **1994**, *59*, 1243-1245.
- [14] J.-C. Zhao, S.-M. Yu, Y. Liu, Z.-J. Yao, *Org. Lett.* **2013**, *15*, 4300-4303.
- [15] J. Zhao, S. Yu, H. Qiu, Z. Yao, *Tetrahedron* **2014**, *70*, 3197-3210.
- [16] R. C. Edgar, *Nucleic Acids Res.* **2004**, *32*, 1792-1797.
- [17] A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, *Bioinformatics* **2009**, *25*, 1189-1191.
- [18] A. T. Keatinge-Clay, *Chem. Biol.* **2007**, *14*, 898-908.
- [19] A. T. Keatinge-Clay, *Nat. Prod. Rep.* **2016**, *33*, 141-149.

[20] N. A. Magarvey, Z. Q. Beck, T. Golakoti, Y. Ding, U. Huber, T. K. Hemscheidt, D. Abelson, R. E. Moore, D. H. Sherman, *ACS Chem. Biol.* **2006**, *1*, 766-779.

[21] M. Hahn, T. Stachelhaus, *Proc. Natl. Acad. Sci.* **2004**, *101*, 15585-15590.

[22] J. R. Whicher, S. S. Smaga, D. A. Hansen, W. C. Brown, W. H. Gerwick, D. H. Sherman, J. L. Smith, *Chem. Biol.* **2013**, *20*, 1340-1351.

[23] F. Liu, S. Garneau, C. T. Walsh, *Chem. Biol.* **2004**, *11*, 1533-42.

[24] A. Miyanaga, F. Kudo, T. Eguchi, *Nat. Prod. Rep.* **2018**, *35*, 1185-1209.

Chapter 2

“An optimized tetrazine probe to detect isocyanide-containing natural products in extracts.”

Gabriel Castro-Falcón,^{1,#} Grant S. Seiler,^{1,2} and Chambers C. Hughes^{1,*}

¹Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093

²Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093

*e-mail: chughes@ucsd.edu

#These authors contributed equally to this work.

Isocyanide-containing natural products have been isolated from nudibranchs, sponges, cyanobacteria, bacteria, and fungi. Marine isocyanides such as axionitrile-3 (**1**), kalihinol A (**2**), and 7,20-diisocyanoadociane (**3**) and cyanobacterial isocyanides like hapalindole A (**4**), welwitindolinone A isonitrile (**5**), and fischerindole L isonitrile (**6**) are complex, polycyclic natural products that are densely substituted with chiral stereocenters (**Figure 1**).¹⁻³ As such, these compounds have captivated synthetic organic chemists since their initial discovery in the 1970s.⁴ Isocyanides from bacteria and fungi, such as brasilidine A (**7**),⁵ SF2768 (**8**),⁶ and (E)-4-(2-isocyanovinyl)phenol (**9**),⁷ are generally simpler in structure. Although their ecological role is likely related to their ability to bind metals,^{8,9} many isocyanides have been shown to have pronounced activity against the malaria parasite *Plasmodium falciparum*, in addition to other biological activities related to the treatment of disease.¹⁰

Most naturally-occurring isocyanides have been isolated using nonspecific chemical screening of extracted organisms with or without a bioassay to help guide the process.^{5-7,10-16} Heterologous expression of orphan biosynthetic gene clusters in bacteria has also led to the discovery of isocyanides, and this approach has even revealed genes specific for the production of certain isocyanides.^{8,17-19} The tell-tale signatures of the isocyanide functionality are the unique isocyanide IR stretch (2165-2110 cm^{-1}) and characteristic ^{13}C - ^{14}N coupling present in ^{13}C NMR spectra. However, these analytical methods are not readily suitable for the directed isolation of isocyanides from complex mixtures using, for example, liquid chromatography, such that at present putative isocyanide compounds must be purified and analyzed individually. Thus, an isocyanide-specific HPLC-MS method would allow for the rapid detection of isocyanide-containing natural products in complex extracts and, more generally, provide a valuable tool for biochemical and biological investigations involving isocyanide chemistry.

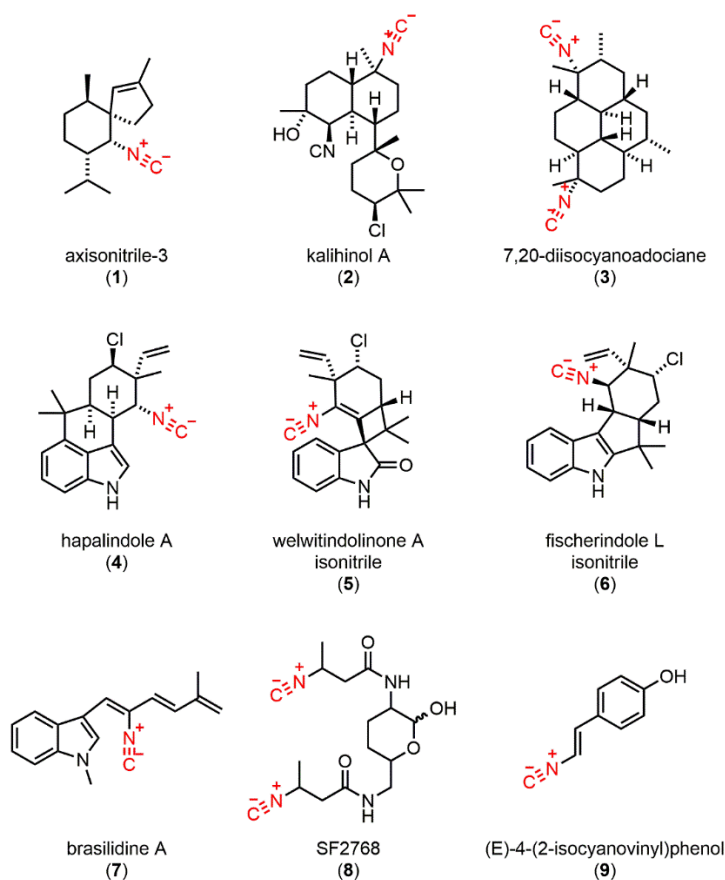


Figure 2.1. Representative isocyanide-containing natural products.

The tetrazine-isocyanide [4+1] cycloaddition was first reported by Seitz and co-workers in 1982.²⁰ Model tetrazine 3,6-di-2-pyridyl-1,2,4,5-tetrazine (**10**) reacts with isocyanide **11** in a [4+1] cycloaddition reaction to give **12**, which is then converted to 4*H*-pyrazol-4-imine **13** in a retro [4+2] cycloaddition.²¹ Tertiary isocyanides give relatively stable pyrazole adducts that hydrolyze slowly in water. When the isocyanide is primary or secondary, however, **13** can aromatize to **14**, which is then readily hydrolyzed to give 1*H*-pyrazol-4-amine **15** and carbonyl compound **16** (**Figure 2**). Certain primary isocyanides like methyl 3-isocyanopropanoate form pyrazoles that are more stable to hydrolysis than expected because **14** further tautomerizes an α,β -unsaturated ester.²¹ The utility of the cycloaddition as a biorthogonal reaction for protein labeling *in vitro* and glycan labeling *in vivo* was demonstrated in 2011 and 2013.^{21,22} More recently, the propensity of the pyrazole adducts derived from cycloaddition

with primary isocyanide groups to hydrolyze was leveraged as a prodrug strategy. Here, oxygen, nitrogen, and sulfur heteroatoms in drugs and fluorophores were masked with 3-isocyanopropyl groups and then removed *in vivo* via a cycloaddition/elimination sequence.²³

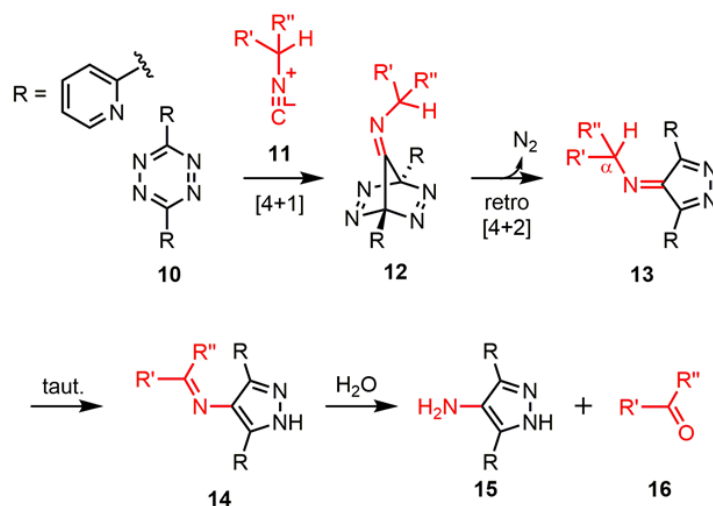


Figure 2.2. The tetrazine-isocyanide [4+1] cycloaddition.

Provided iminopyrazole hydrolysis could be minimized, we hoped that a general method for the reactivity-guided isolation of isocyanide-containing natural products could be developed using a UV-active, brominated, highly crystalline tetrazine reagent analogous to our previously reported probes.^{24–27} In this regard, we were surprised to find only one example of this reaction with a natural product in the literature, wherein commercially-available tetrazine **10** was shown to react with known isocyanide **9** in bacterial culture.²⁸ Although this reaction is noteworthy, neither tetrazine **10** nor commercially-available 3,6-diphenyl-1,2,4,5-tetrazine (**17**) are ideal probes because the resulting pyrazoles do not possess characteristic features for mass detection. Furthermore, pyrazoles derived from **10** readily hydrolyze in the case of primary and secondary isocyanides (e.g., **1**, **4**, **6**, and **8**), which is particularly a problem for analysis using reversed-phase liquid chromatography. Attempts to endow **10** and **17** with a unique isotope-based signature via chlorination (³⁵Cl:³⁷Cl 3:1) and bromination (⁷⁹Cl:⁸¹Cl 1:1) gave tetrazines **18–21** exhibiting extremely poor solubility in a wide variety of organic solvents (**Figure 2.3**).

We then prepared and evaluated 3-(5-bromopyridin-2-yl)-6-methyl tetrazine (**22**) and 3-(5-bromopyridin-2-yl) tetrazine (**23**). These tetrazines, which have greatly improved solubility in organic solvents compared to **10**, were synthesized via reaction of the 5-bromo-2-cyanopyridine, hydrazine, zinc triflate, and either acetonitrile or formamide acetate.^{29,30} To evaluate the reactivity of these reagents, we used t-butyl isocyanide (**26**), cyclohexyl isocyanide (**27**), n-butyl isocyanide (**28**), and 2-naphthyl isocyanide (**29**) as model compounds. Stock solutions of tetrazine in DMF or DMSO were added to solutions of excess isocyanide in acetonitrile at room temperature, and the reaction mixtures were analyzed 24 h later by HPLC-UV-MS using a buffered eluent (NH₄OAc, pH 7). The neutral LC conditions were critical, as the presence of formic acid led to rapid hydrolysis and noticeable side-reactivity. Tetrazines **22** and **23** reacted cleanly with **26** to give the corresponding iminopyrazoles, which possessed a distinct mass signature and long wavelength UV absorption ($\lambda_{\text{max}} = 370 \text{ nm}$). However, reaction of these tetrazines with **27** and **28** yielded aminopyrazole **15** as the only observable product, and reaction with **29** gave several unidentified side-products in addition to the desired iminopyrazole.

In order to reduce the rate of pyrazole hydrolysis, we prepared and evaluated more electron-rich tetrazines 3-(4-bromophenyl)-6-methyl tetrazine (**24**) and 3-(4-bromophenyl) tetrazine (**25**) from 4-bromobenzonitrile using the same conditions described for the synthesis of **22** and **23** (see **Figure 2.3**). Earlier reactions with **17** demonstrated that more electron-rich tetrazines would yield iminopyrazoles with improved hydrolytic stability, owing to the decreased acidity of the α -proton in **13** or simply the lack of basic pyridines in the reaction (see **Figure 2.2**). Indeed, tetrazines **24** and **25** reacted cleanly with **26-29** to give the corresponding iminopyrazoles, which again had a distinct mass signature and long wavelength UV absorption, although the iminopyrazole derived from the primary isocyanide underwent significant hydrolysis. Given that it could be obtained in higher yield than **25** and that any rate enhancement from the less sterically-hindered tetrazines was unwarranted, we decided that **24** was our most suitable tetrazine probe at this point.

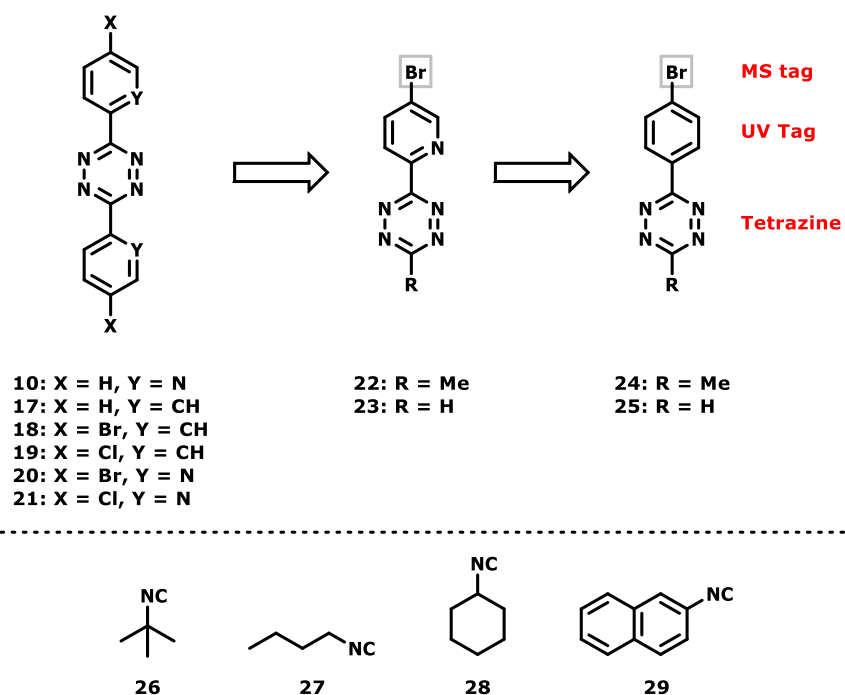


Figure 2.3. Tetrazines **10-25** and isocyanides **26-29**.

We then set out to further optimize the MS characteristics of the adducts derived from tetrazine **24** by placing substituents on the phenyl ring that would improve ionization efficiency and MS/MS fragmentation. In this way, the precursor ion of the pyrazoles would be more pronounced, lowering the detection limit, and chromatograms could be rapidly screened by searching for the characteristic product ions. To this end, we coupled various substituted amines to 4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzoic acid (**30**) with the expectation that fragmentation would readily occur across the amide bond (**Figure 2.4**). First, 2-picolylamine, 2-(2-pyridyl)ethylamine, and 3-morpholinopropylamine were coupled to **30** to give tetrazines **31-33**. The morpholino amine in **33** was then quaternized with methyl iodide to give a probe (**34**) with a permanent positive charge. Adding basic tertiary amines and charged quaternary ammonium salts is an established method in untargeted metabolomics using HPLC-ESI-MS.^{31,32} Tetrazines **35** and **36** were also synthesized via coupling of **30** to the appropriate amine. Although **35** and **36** do not possess a basic or quaternary nitrogen atom, they have a bromine atom as a

distinguishing mass feature and were expected to readily fragment to yield a tropylium ion. Finally, tetrazine **36** has a fluorine atom that could be utilized in ^{19}F NMR experiments.

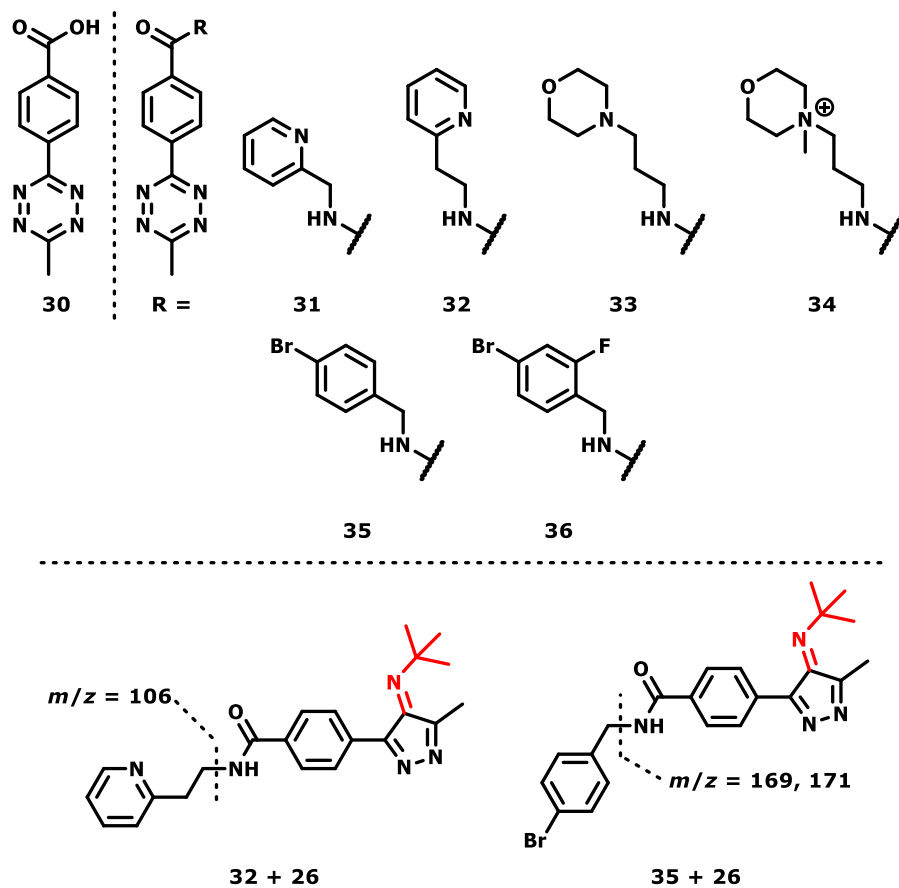


Figure 2.4: Core tetrazine **30**, tetrazines **31-36**, and selected iminopyrazoles **32 + 26** and **35 + 26**.

Experimental

Synthesis of compound **23**: To a stirring solution of 5-bromo-2-cyanopyridine (200 mg, 1.1 mmol, 1.0 equiv.) in 50% EtOH_(aq.) (2.2 mL) was added sulfur flowers (70 mg, 2.2 mmol, 2.0 equiv.), DCM (70 μ L, 1.1 mmol, 1.0 equiv.), and hydrazine (276 μ L, 8.8 mmol, 8.0 equiv.). The reaction flask was sealed and stirred at 50°C for 24 hours. The reaction was allowed to cool to 22°C, and then diluted with 100 mL H₂O and 10 mL EtOAc. NaNO₂ (760 mg, 11 mmol, 10 equiv.) was added with stirring, and conc. HCl was added slowly to the stirring mixture until gas evolution ceased. The mixture was extracted with EtOAc, and the organic phase was washed with sat. NaHCO₃, water, and brine. The organic layer was dried over Na₂SO₄, filtered, and evaporated onto Celite. The product was purified by flash chromatography (0 to 100% DCM in hexanes, 12 g column, then, 0 to 60% EtOAc in hexanes, 12 g column). 3-(5-bromopyridin-2-yl)-1,2,4,5-tetrazine (**23**) was recovered as a pink solid (20 mg, 7.7% yield). UV/Vis: λ_{max} = 308 nm, 538 nm. ¹H NMR (500 MHz, CDCl₃) δ 10.38 (s, 1H), 9.02 (s, 1H), 8.59 (d, J = 8.4 Hz, 2H), 8.15 (d, J = 8.8 Hz). ¹³C NMR (125 MHz, CDCl₃) δ 165.6, 158.5, 152.5, 148.4, 140.5, 125.4, 125.3. HR-ESI-TOFMS: m/z (M+H)⁺ calcd for C₇H₅⁷⁹BrN₅⁺ 237.9723, found 237.9723.

Synthesis of compound **25**: To a stirring solution of 4-bromobenzonitrile (200 mg, 1.1 mmol, 1.0 equiv.) in 50% EtOH_(aq.) (2.2 mL) was added sulfur flowers (70 mg, 2.2 mmol, 2.0 equiv.), DCM (70 μ L, 1.1 mmol, 1.0 equiv.), and hydrazine (276 μ L, 8.8 mmol, 8.0 equiv.). The reaction flask was sealed and stirred at 50°C for 16 hours. The reaction was allowed to cool to 22°C, and then diluted with 100 mL H₂O. NaNO₂ (760 mg, 11 mmol, 10 equiv.) was added with stirring, and conc. HCl was added slowly to the stirring mixture until gas evolution ceased. The mixture was extracted with EtOAc, and the organic phase was washed with water, sat. NaHCO₃, water, and brine. The organic layer was dried over Na₂SO₄, filtered, and evaporated onto Celite. The product was purified by flash chromatography (0 to 100% DCM in hexanes, 12 g column). 3-(4-bromophenyl)-1,2,4,5-tetrazine (**25**) was recovered as a pink solid (44 mg,

17% yield). ^1H NMR (500 MHz, CDCl_3) δ 10.24 (s, 1H), 8.51 (d, $J = 8.6$ Hz, 2H), 7.76 (d, $J = 8.6$ Hz, 2H). ^{13}C NMR (125 MHz, CDCl_3) δ 166.1, 158.0, 132.9, 130.6, 129.8, 128.6.

Synthesis of compound **31**: A mixture of pyridin-2-ylmethanamine (34 μL , 333 μmol , 1.4 equiv.) and 4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzoic acid (50 mg, 231 μmol , 1.0 equiv.) was stirred in DMF (2.3 mL). HBTU (105 mg, 278 μmol , 1.2 equiv.) was added, followed by *iPr*₂NEt (70 μL , 400 μmol , 1.6 equiv.). The reaction was stirred at 22°C for 12 hours, then diluted in EtOAc. The organic phase was washed successively with sat. aq. Na_2CO_3 , H_2O , and brine, then dried over Na_2SO_4 . The solvent was evaporated to afford the pure product as a pink solid (51 mg, 72% yield). UV/Vis: $\lambda_{\text{max}} = 270$ nm, 525 nm. ^1H NMR (CDCl_3): δ 8.69 (d, $J = 8.4$ Hz, 2H), 8.59 (d, $J = 4.7$ Hz, 1H), 8.09 (d, $J = 8.4$ Hz, 2H), 7.79 (s, br, 1H), 7.71 (t, $J = 7.7$ Hz, 1H), 7.35 (d, $J = 7.8$ Hz, 1H), 7.25 (m, overlap, 1H), 4.81 (d, $J = 4.4$ Hz, 2H), 3.13 (s, 3H). ^{13}C NMR (CDCl_3): δ 167.7, 166.6, 163.8, 155.8, 149.2, 138.1, 137.1, 134.6, 128.3, 128.1, 122.8, 122.4, 44.9, 21.4. HR-ESI-TOFMS: m/z ($\text{M}+\text{H}$)⁺ calcd for $\text{C}_{16}\text{H}_{15}\text{N}_6\text{O}$ 307.1302, found 307.1300.

Synthesis of compound **32**: A mixture of 2-(pyridin-2-yl)ethan-1-amine (40 μL , 333 μmol , 1.4 equiv.) and 4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzoic acid (50 mg, 231 μmol , 1.0 equiv.) was stirred in DMF (2.3 mL). HBTU (105 mg, 278 μmol , 1.2 equiv.) was added, followed by *iPr*₂NEt (70 μL , 400 μmol , 1.6 equiv.). The reaction was stirred at 22°C for 12 hours, then diluted in EtOAc. The organic phase was washed successively with sat. aq. Na_2CO_3 , H_2O , and brine, then dried over Na_2SO_4 . The solvent was evaporated to afford the pure product as a pink solid (69 mg, 93% yield). UV/Vis: $\lambda_{\text{max}} = 265$ nm, 525 nm. ^1H NMR (CDCl_3): δ 8.65 (d, $J = 8.2$ Hz, 2H), 8.59 (d, $J = 4.4$ Hz, 1H), 7.99 (d, $J = 8.2$ Hz, 2H), 7.89 (s, br, 1H), 7.64 (t, $J = 7.6$ Hz, 1H), 7.19-7.24 (m, overlap, 2H), 3.90 (m, $J = 5.7$ Hz, 2H), 3.13 (m, overlap, 5H). ^{13}C NMR (CDCl_3): δ 167.7, 166.4, 163.8, 160.0, 149.3, 138.5, 137.1, 134.4, 128.2, 128.0, 123.8, 122.0, 39.4, 36.4, 21.4. HR-ESI-TOFMS: m/z ($\text{M}+\text{H}$)⁺ calcd for $\text{C}_{17}\text{H}_{17}\text{N}_6\text{O}$ 321.1458, found 321.1458.

Synthesis of compound **33**: 4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzoic acid (100 mg, 463 μmol , 1.0 equiv.) and HBTU (219 mg, 579 μmol , 1.25 equiv.) were added to a flask with a stir bar. DMF (4.6 mL)

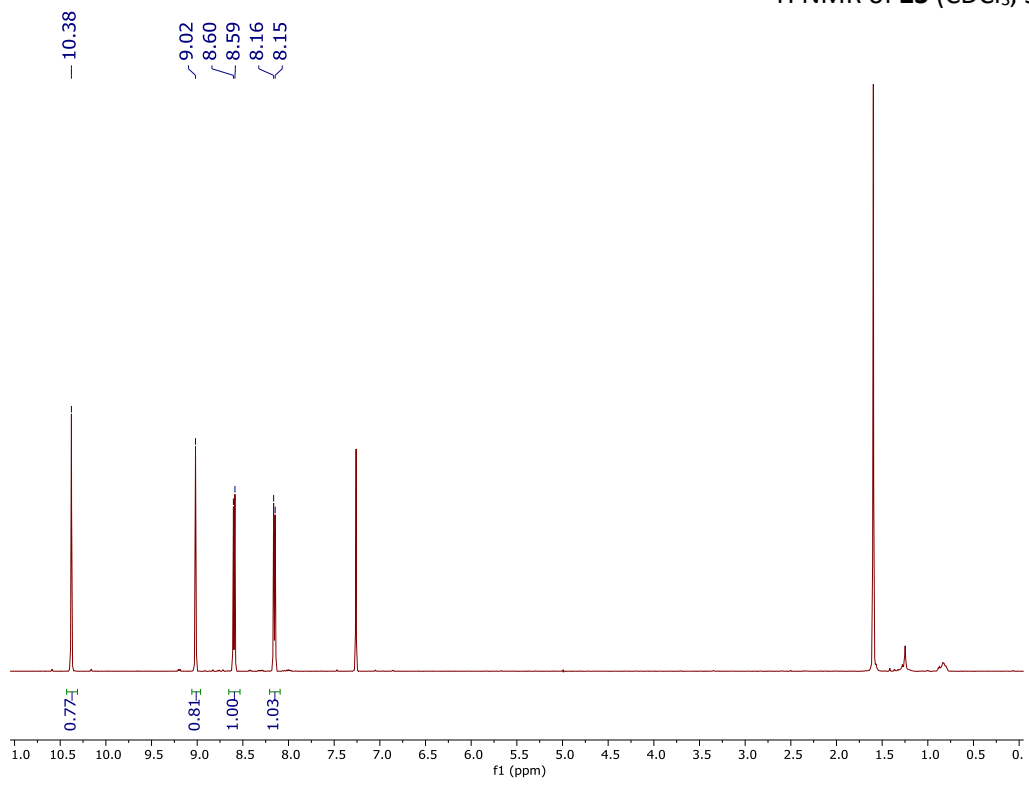
was added and the mixture was stirred to dissolve. To the flask was added 3-morpholinopropan-1-amine (68 μL , 463 μmol , 1.0 equiv.), followed by *iPr*₂NEt (121 μL , 694 μmol , 1.5 equiv.). The reaction was stirred at 22°C for 3 hr., at which point HPLC monitoring indicated reaction completion. The reaction was diluted with EtOAc, washed with sat. NaHCO₃, then with brine. The organic phase was dried over Na₂SO₄, filtered, and evaporated onto Celite. Flash chromatography (12 g column, 0 to 20% MeOH in DCM) afforded pure XYZ (89 mg, 56% yield). UV/vis. λ_{max} = 270 nm. ¹H NMR (CD₃OD) δ 8.64 (d, *J* = 8.3 Hz, 2H), 8.05 (d, *J* = 8.3 Hz, 2H), 3.71 (br. s, 4H), 3.49 (t, *J* = 7.0 Hz, 2H), 3.06 (s, 3H), 2.53 (m, 6H), 1.88 (m, *J* = 7.0 Hz, 2H). ¹H NMR (CDCl₃) δ 8.67 (d, *J* = 8.4 Hz, 2H), 8.28 (br. s, 1H), 8.02 (d, *J* = 8.4 Hz, 2H), 3.72 (t, *J* = 4.3 Hz, 4H), 3.62 (q, *J* = 5.6 Hz, 2H), 3.12 (s, 3H), 2.60 (t, *J* = 5.9 Hz, 2H), 2.50 (br. s, 3H), 1.83 (m, *J* = 6.0 Hz, 2H), 1.77 (br. s, 2H). ¹³C NMR (CDCl₃, 500 MHz) δ 167.7, 166.6, 163.7, 138.6, 134.5, 128.2, 128.0, 67.1, 59.0, 54.0, 41.0, 24.0, 21.4. HR-ESI-TOFMS: *m/z* (M+H)⁺ calcd for C₁₇H₂₃N₆O₂⁺ 343.1877, found 343.1874.

Synthesis of compound **34**: To 4-(6-methyl-1,2,4,5-tetrazin-3-yl)-*N*-(3-(piperidin-1-yl)propyl)benzamide (89 mg, 260 μmol , 1.0 equiv.) stirred in 5.2 mL 1:1 MeCN:THF was added MeI (32 μL , 520 μmol , 2.0 equiv.) at 22°C. The reaction was stirred for 12 hours, and monitoring by an adapted analytical HPLC method (0.1% formic acid replaced by 10 mM ammonium acetate) indicated completion of the reaction. The mixture was evaporated to dryness under a stream of N₂, and re-dissolved in 95% MeOH_(aq). The product was purified by preparative HPLC (20% MeCN in H₂O, 0.1% formic acid, 13 mL/min, r.t. 6 to 8 min.) to afford 1-methyl-1-(3-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzamido)propyl)piperidin-1-ium iodide as a pink solid (57 mg, 45% yield). UV/vis: λ_{max} = nm, nm. ¹H NMR (CD₃OD): δ 8.65 (d, *J* = Hz, 2H), 8.47 (br. s, 1H), 8.08 (d, *J* = Hz, 2H), 4.02 (br., overlapping, 4H), 3.4-3.6 (br., overlapping, 8H), 3.24 (s, 3H), 3.07 (s, 3H), 2.18 (br. s, 2H). ¹³C NMR (CD₃OD): δ 169.6, 169.1, 164.8, 138.7, 136.6, 129.2, 128.9, 64.6, 61.6, 61.2, 47.3, 38.0, 23.2, 21.2. HR-ESI-TOFMS: *m/z* (M+H)⁺ calcd for C₁₉H₂₇N₆O⁺ 357.2034, found 357.2030.

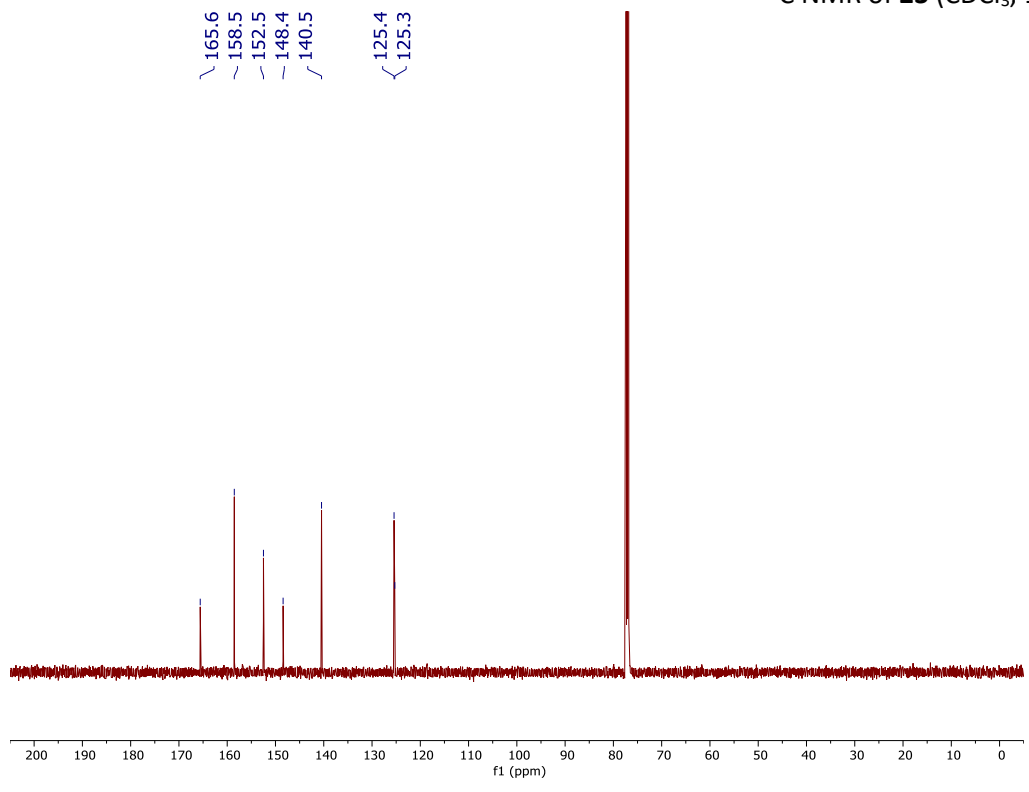
Synthesis of compound **35**: A mixture for 4-bromobenzylamine-HCl (29 mg, 129 μmol , 1.4 equiv.) and 4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzoic acid (20 mg, 93 μmol , 1.0 equiv.) was stirred in DMF (1.0 mL). HBTU (42 mg, 111 μmol , 1.2 equiv.) was added, followed by *i*Pr₂NEt (42 μL , 242 μmol , 2.6 equiv.). The reaction was stirred at 22°C for 12 hr, at which time a pink precipitate had formed. The precipitate was recovered by filtration (filter paper of sintered glass vacuum funnel) and rinsed with MeOH to afford the pure product as a pink solid (35 mg, quant.). UV/Vis: λ_{max} = 270 nm, 535 nm. ¹H NMR (DMSO-*d*₆): δ 9.31 (s, br, 1H), 8.56 (d, *J* = 8.2 Hz, 2H), 8.14 (d, *J* = 8.2 Hz, 2H), 7.53 (d, *J* = 8.2 Hz, 2H), 7.31 (d, *J* = 8.1 Hz, 2H), 4.48 (d, *J* = 5.5 Hz, 2H), 3.02 (s, 3H). ¹³C NMR (DMSO-*d*₆): δ 167.4, 165.6, 162.9, 139.0, 137.4, 134.4, 131.2, 129.6, 128.3, 127.4, 119.9, 42.2, 20.9. HR-ESI-TOFMS: *m/z* (M+H)⁺ calcd for C₁₇H₁₅⁷⁹BrN₅O 384.0454, found 384.0457.

Synthesis of compound **36**: A mixture of 4-bromo-2-fluorobenzylamine-HCl (167 mg, 694 μmol , 1.5 equiv.) and 4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzoic acid (100 mg, 463 μmol , 1.0 equiv.) was stirred in DMF (5 mL). HBTU (219 mg, 578 μmol , 1.25 equiv.) was added, followed by *i*Pr₂NEt (161 μL , 926 μmol , 2.0 equiv.). The reaction was stirred at 22°C for 4 hr., then diluted with EtOAc. The organic phase was washed with successively with 10% aq. citric acid, H₂O, sat. aq. NaHCO₃, H₂O, and brine. The organic phase was dried over Na₂SO₄, filtered, and evaporated onto Celite. Column chromatography (4 g column, 0 to 100% EtOAc in hexanes) yielded the product as a pink solid (100 mg, 53.8% yield). UV/Vis: λ_{max} = 270 nm, 540 nm. ¹H NMR (DMSO-*d*₆): δ 9.29 (s, br, 1H), 8.56 (d, *J* = 8.2 Hz, 2H), 8.14 (d, *J* = 8.2 Hz, 2H), 7.55 (d, *J* = 9.4 Hz, 1H), 7.37-7.43 (m, 2H), 4.51 (d, *J* = 3.6 Hz, 2H), 3.02 (s, 3H). ¹³C NMR (DMSO-*d*₆): δ 167.3, 165.7, 162.9, 160.0 (d, *J* = 250 Hz), 137.2, 134.5, 131.3 (d, *J* = 4.9 Hz), 128.3, 127.5 (d, *J* = 3.0 Hz), 127.4, 125.6 (d, *J* = 14.7 Hz), 120.2 (d, *J* = 9.8 Hz), 118.5 (d, *J* = 24.8 Hz), 36.4, 20.9. HR-ESI-TOFMS: *m/z* (M+H)⁺ calcd for C₁₇H₁₄BrFN₅O 402.0360, found 402.0363.

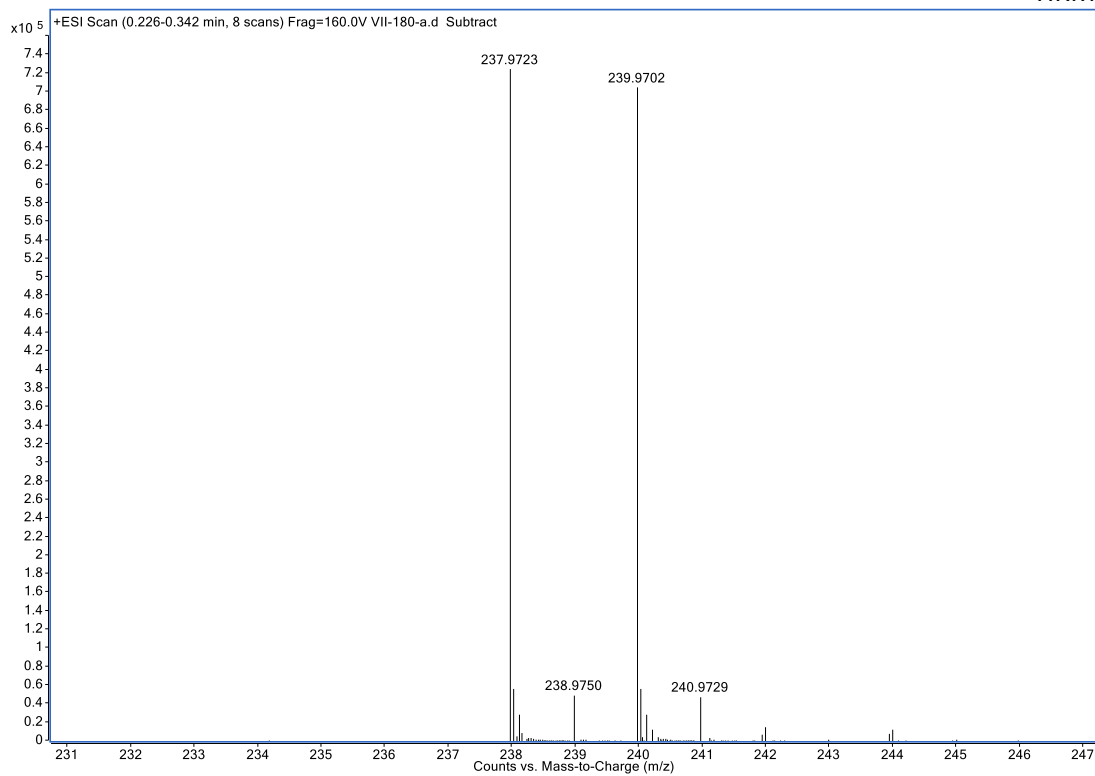
¹H NMR of **23** (CDCl₃, 500 MHz)



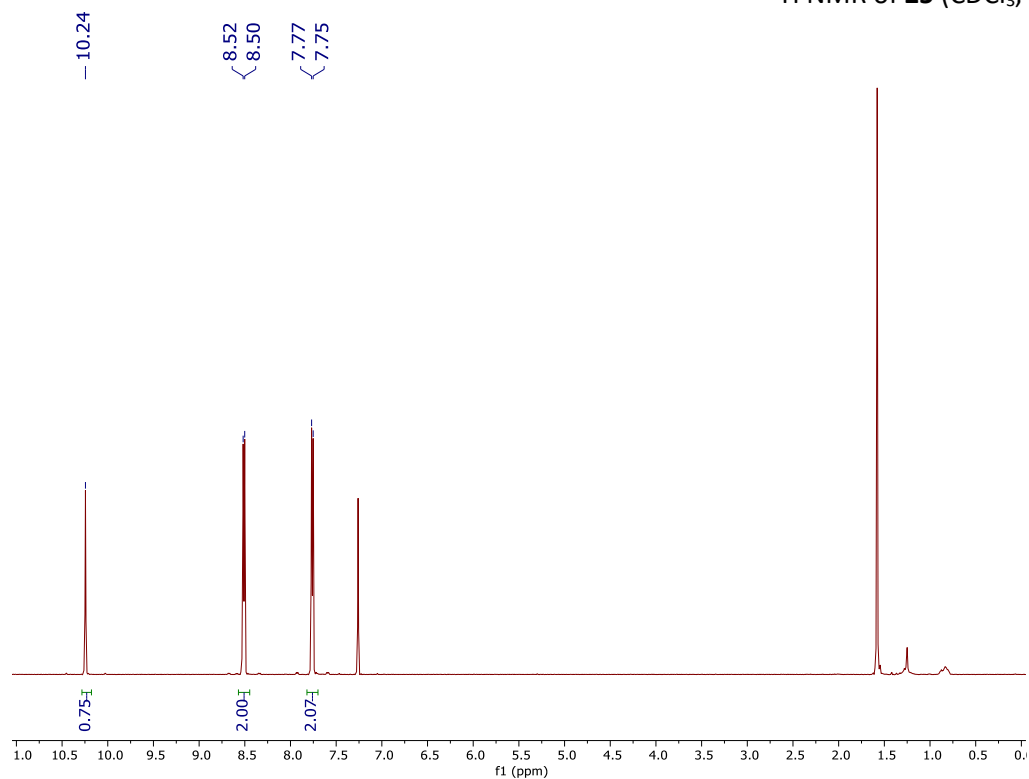
¹³C NMR of **23** (CDCl₃, 125 MHz)



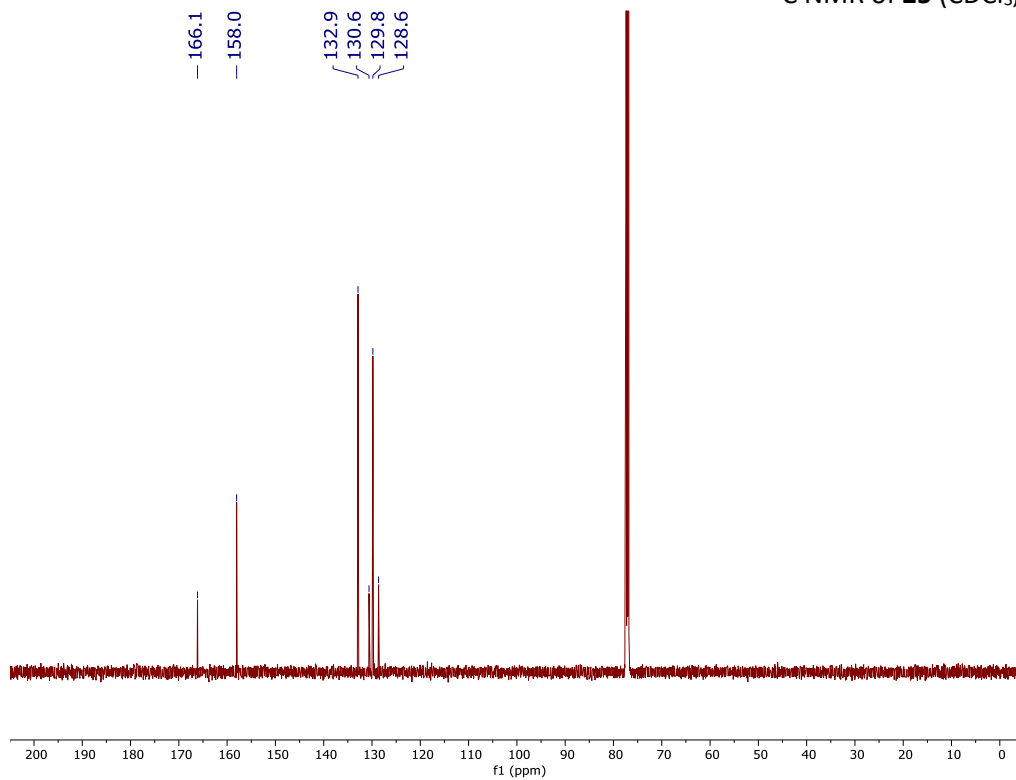
HRMS of 23



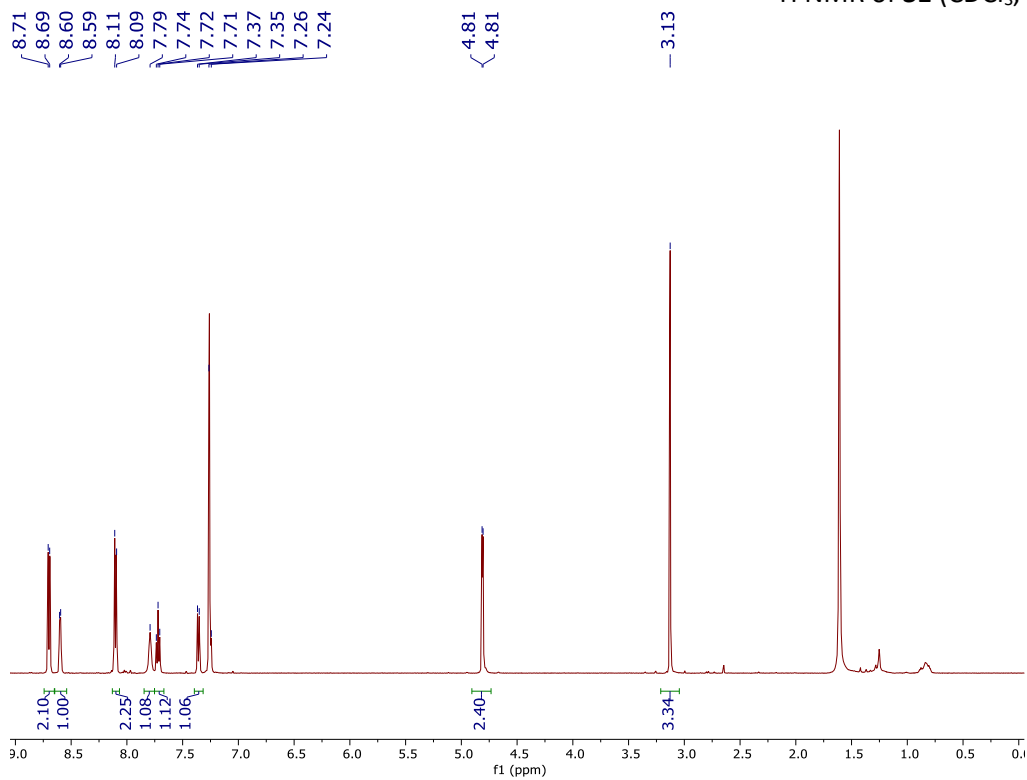
¹H NMR of 25 (CDCl₃, 500 MHz)



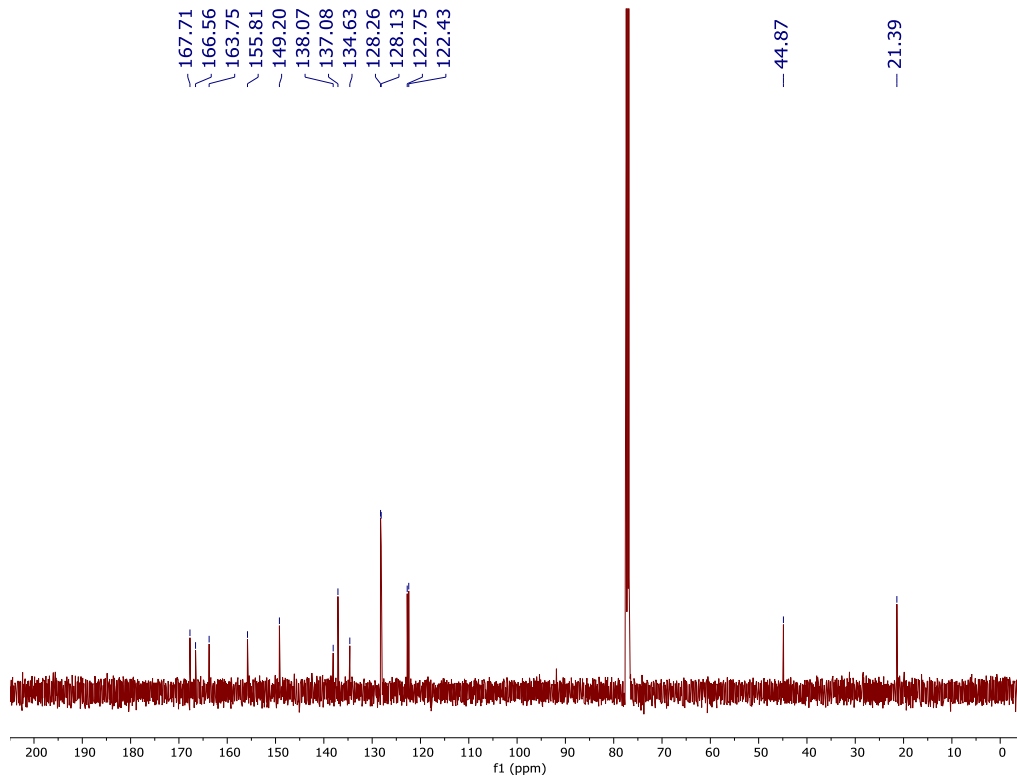
^{13}C NMR of **25** (CDCl_3 , 125 MHz)



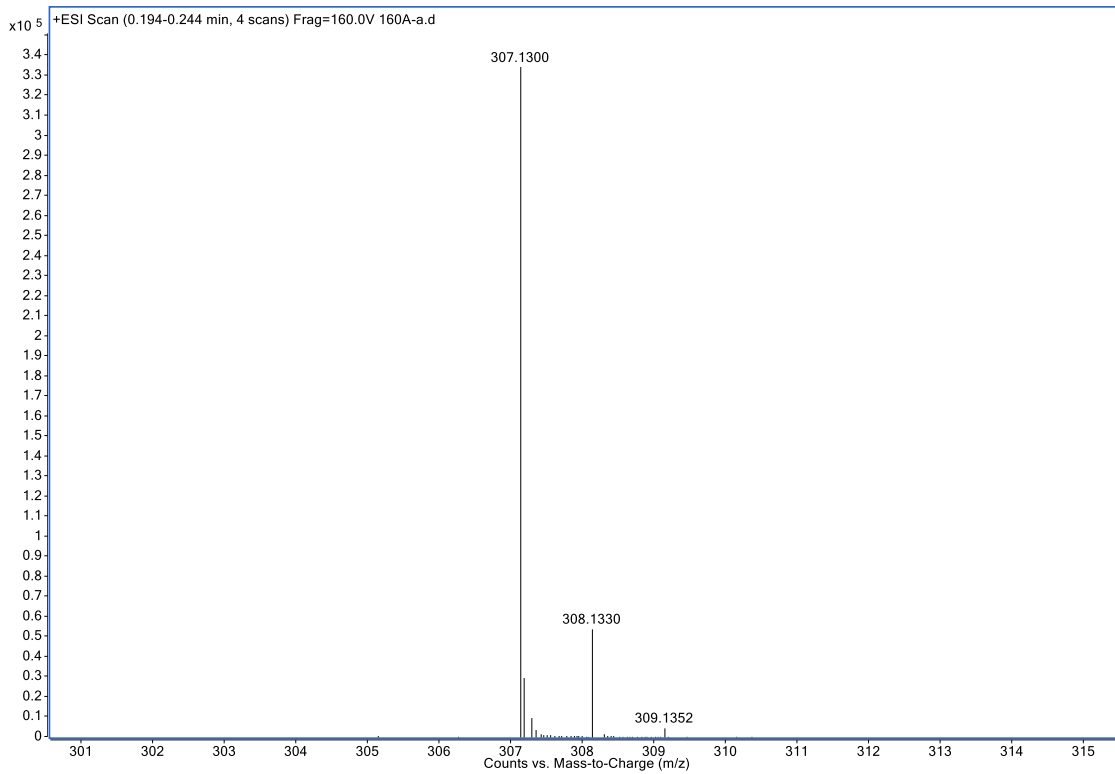
^1H NMR of **31** (CDCl_3 , 500 MHz)



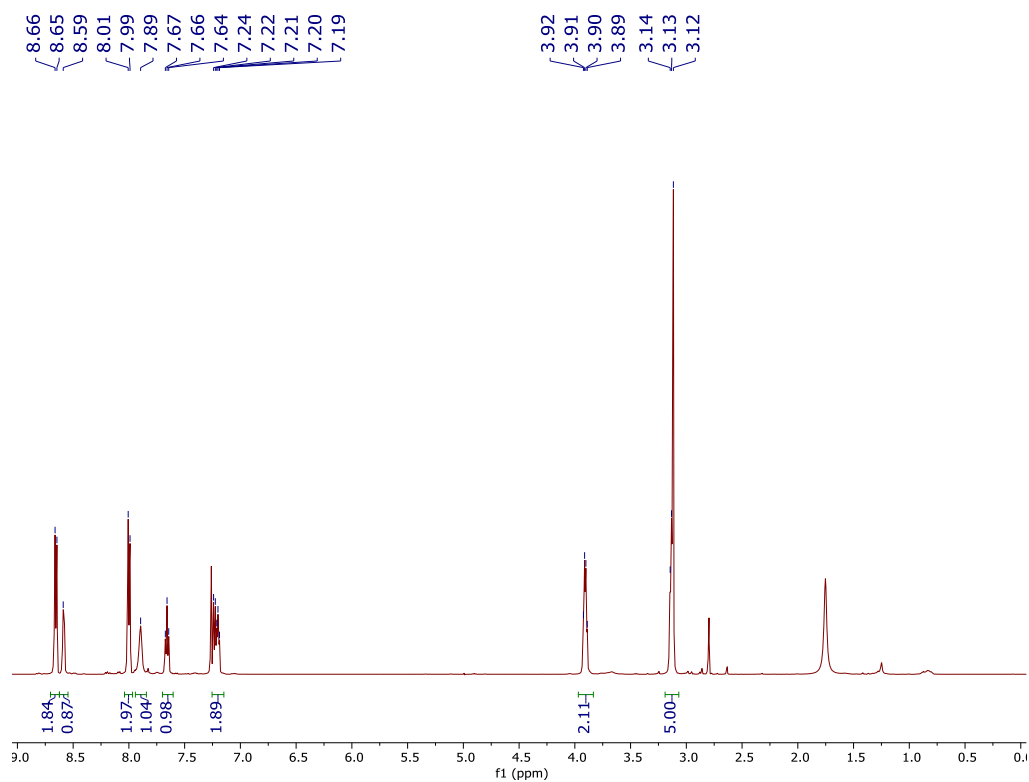
¹³C NMR of **31** (CDCl₃, 125 MHz)



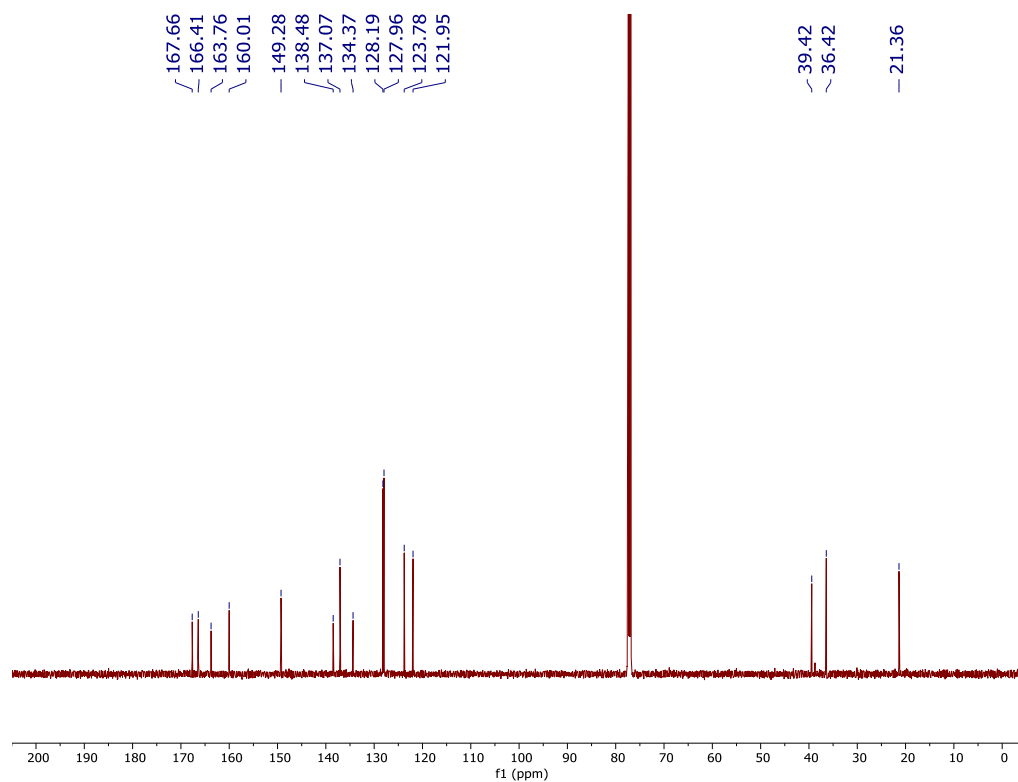
HRMS of **31**

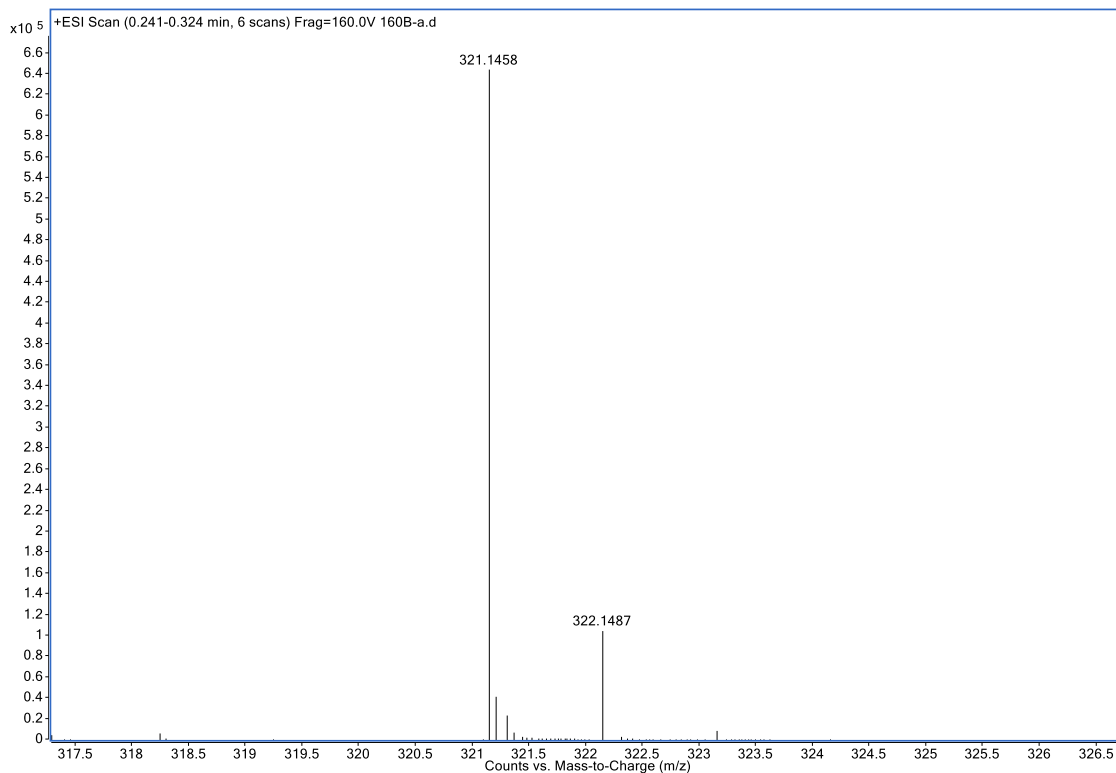


^1H NMR of **32** (CDCl_3 , 500 MHz)

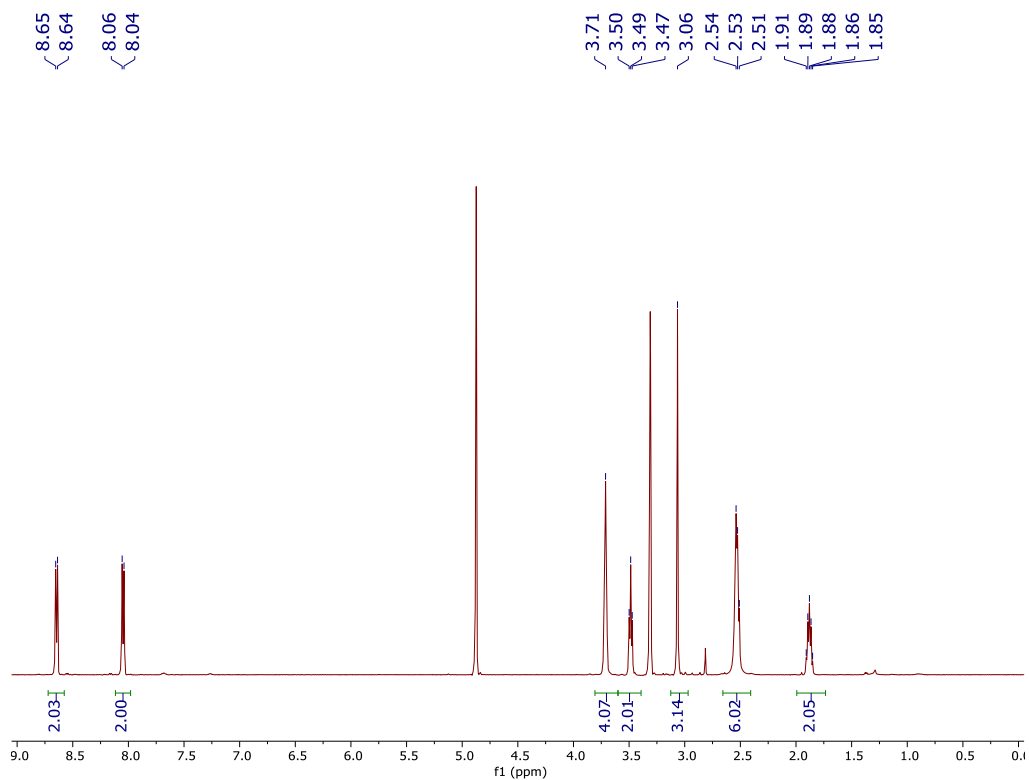


^{13}C NMR of **32** (CDCl_3 , 125 MHz)

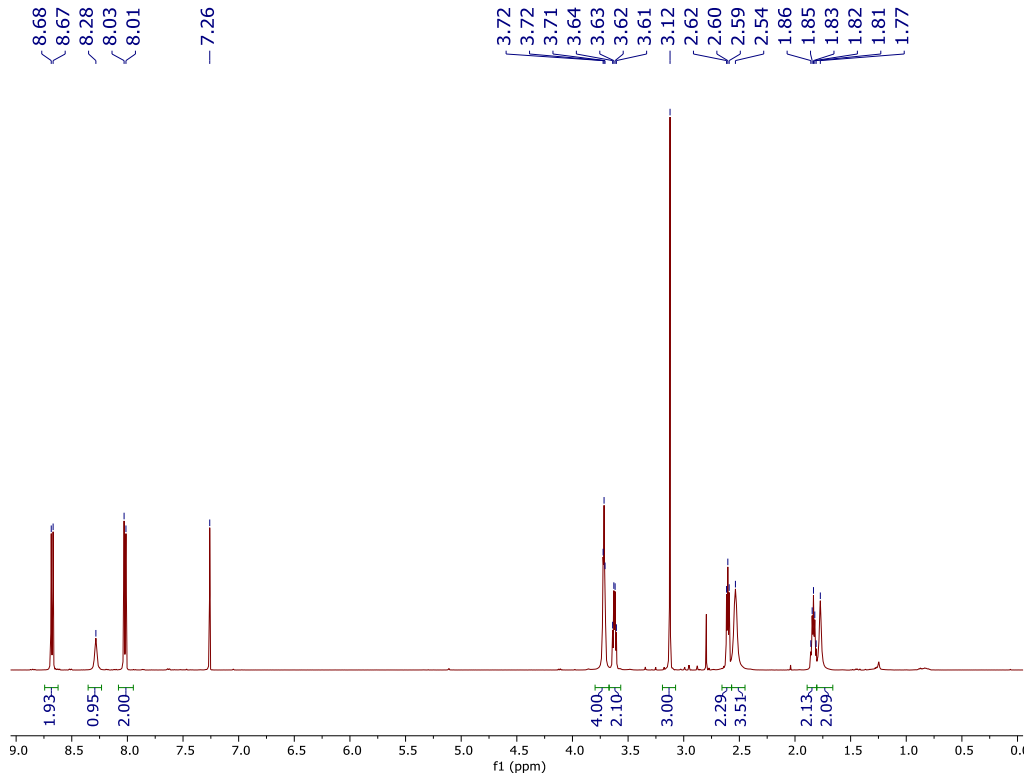




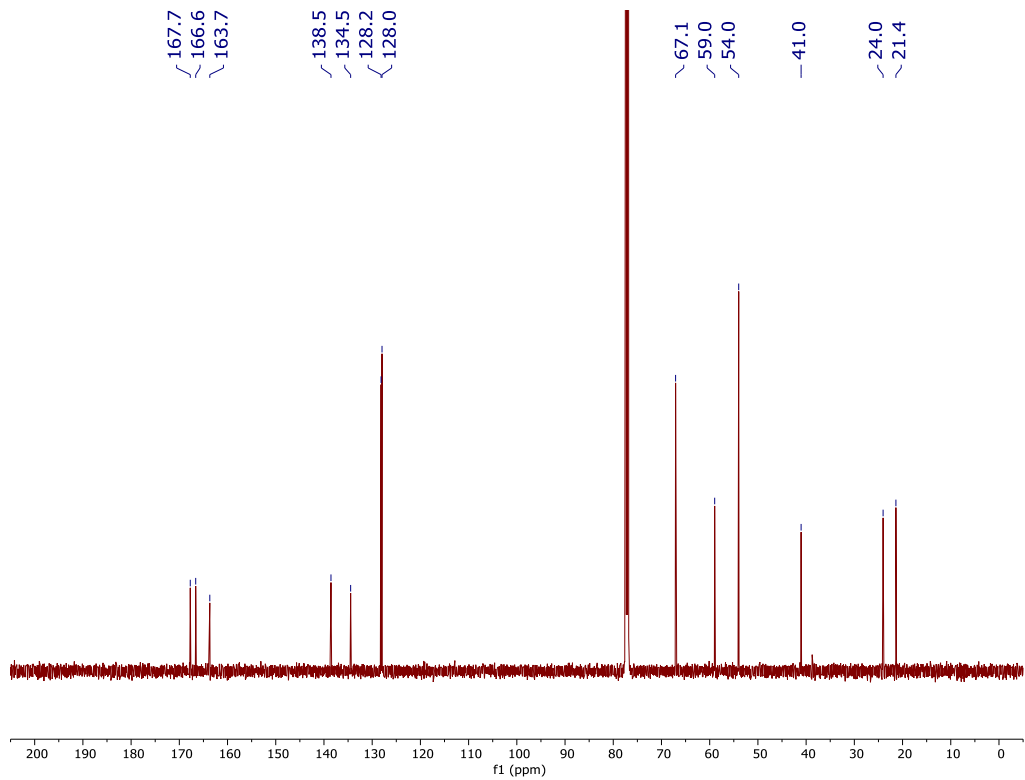
¹H NMR of **33** (CD₃OD, 500 MHz)

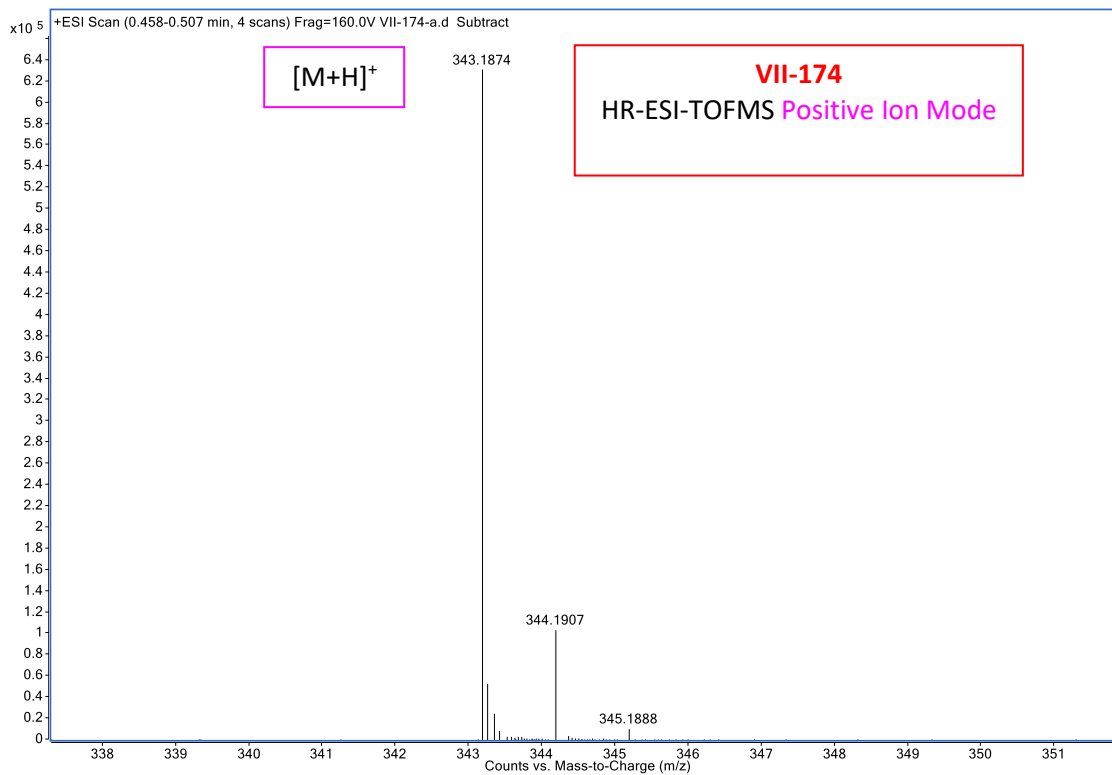
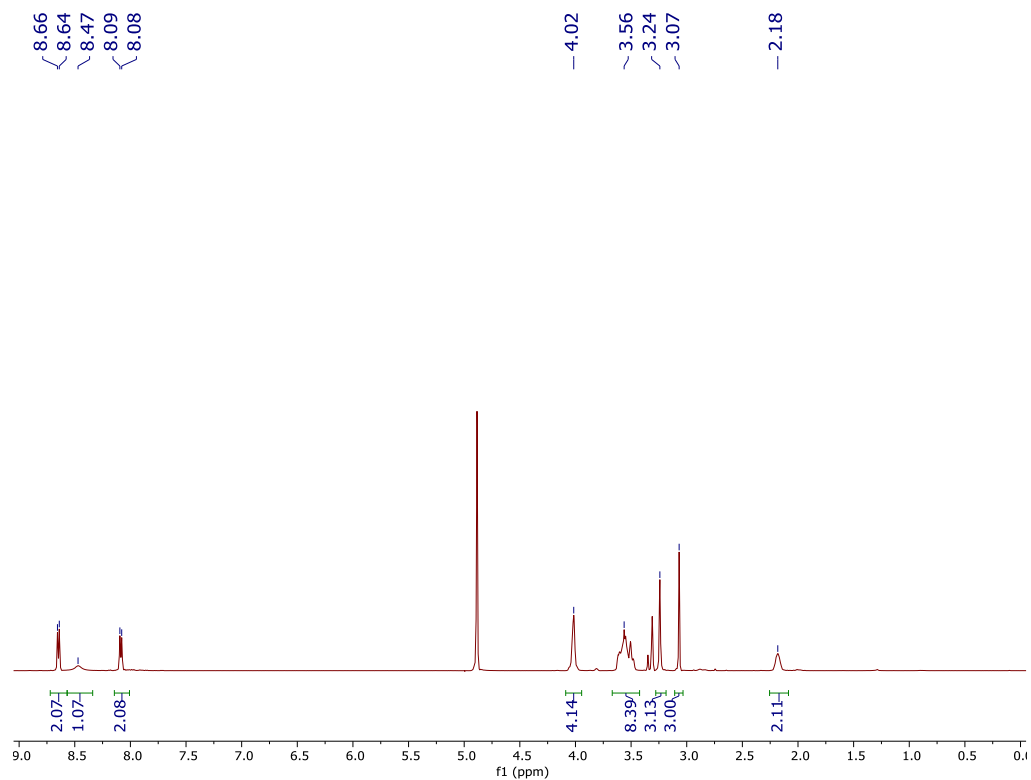


¹H NMR of **33** (CDCl₃, 500 MHz)

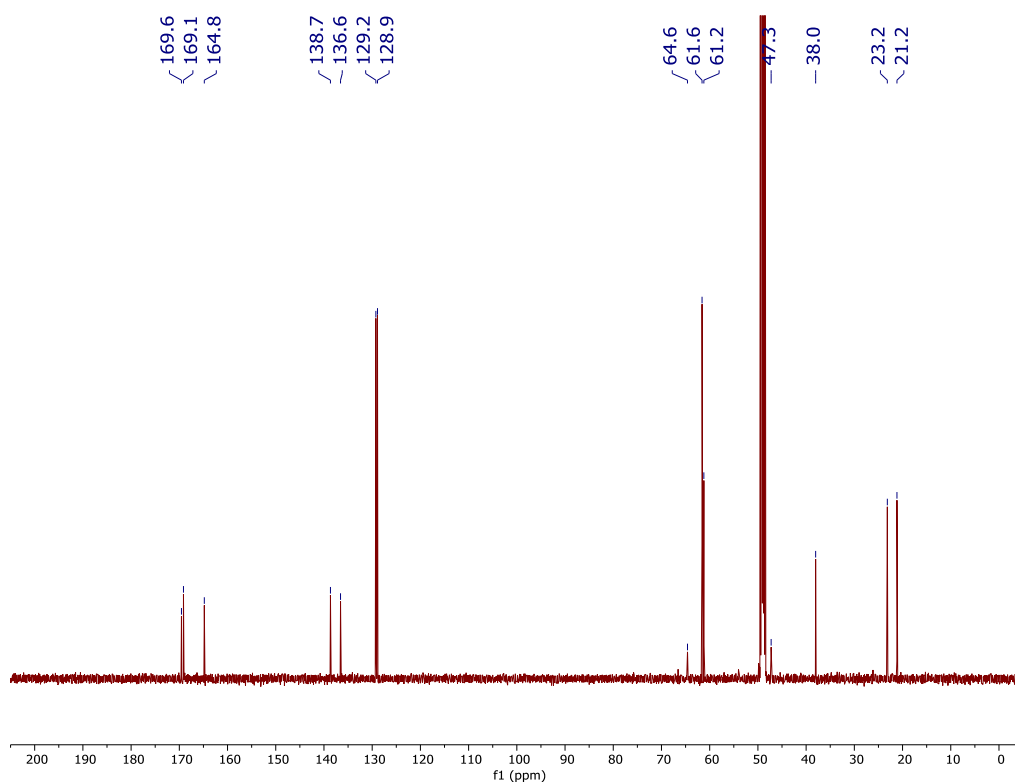


¹³C NMR of **33** (CDCl₃, 125 MHz)

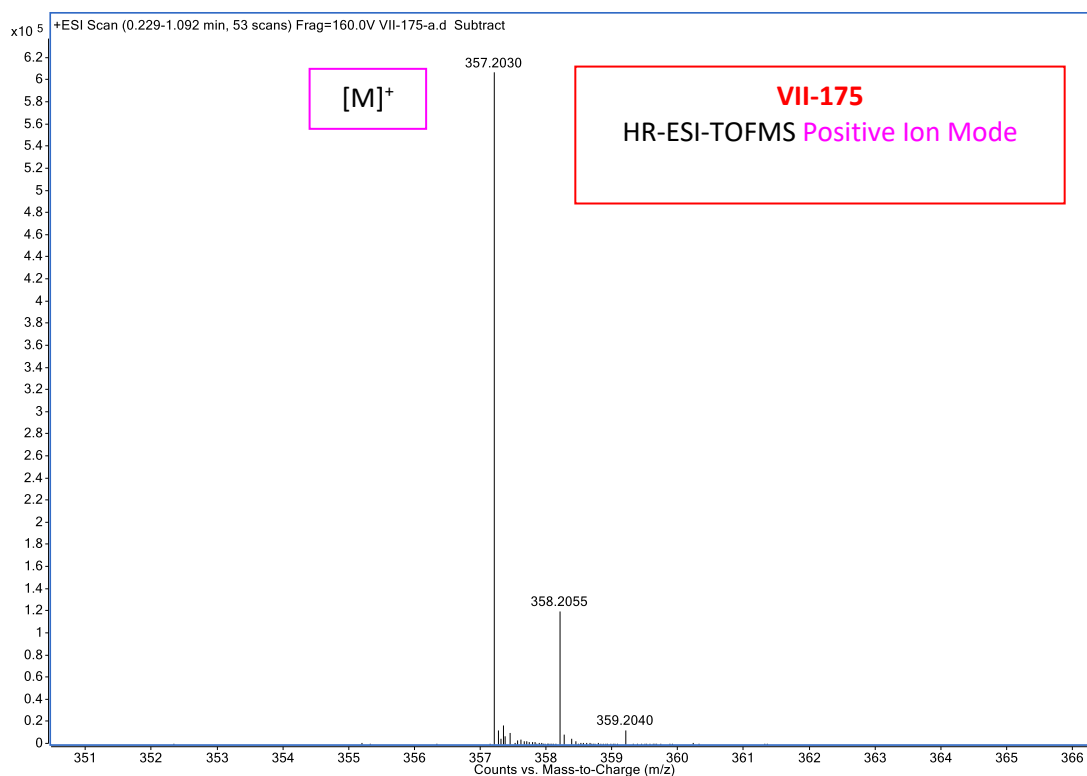


¹H NMR of **34** (CD₃OD, 500 MHz)

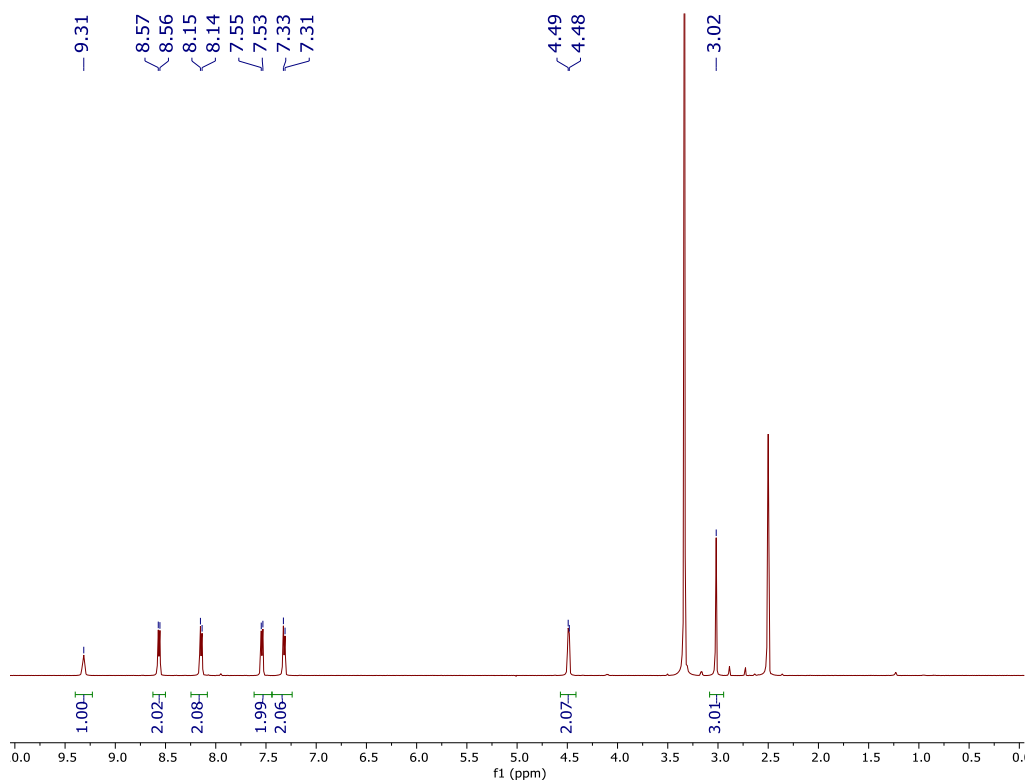
¹³C NMR of **34** (CD₃OD, 125 MHz)



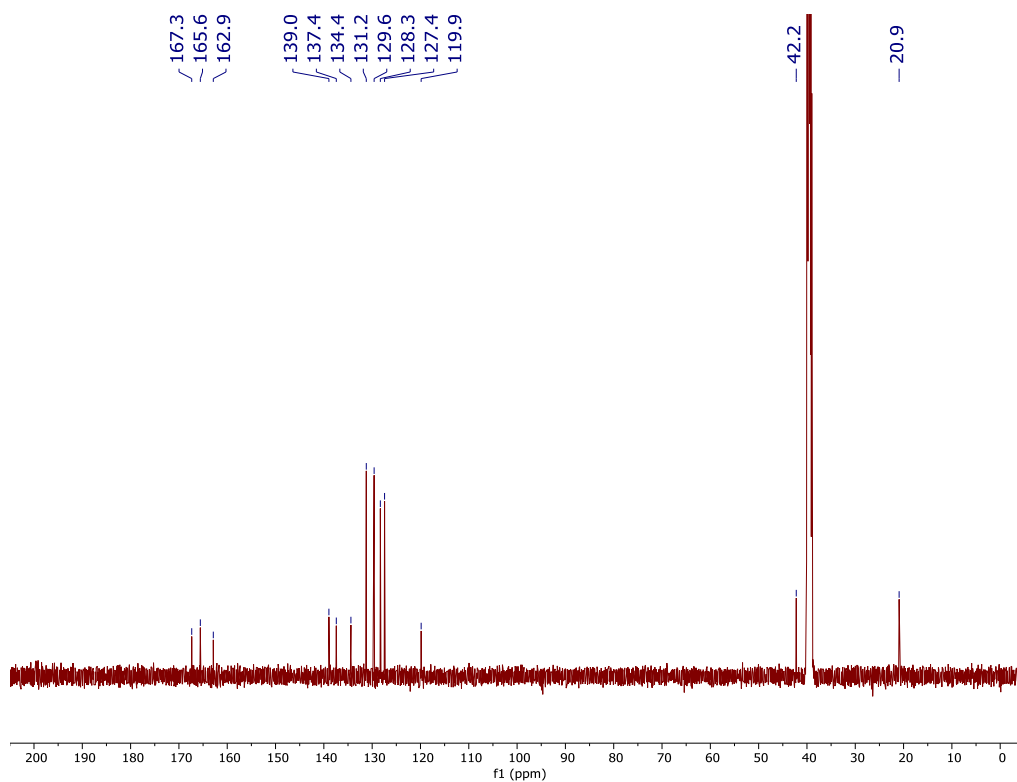
HRMS of **34**

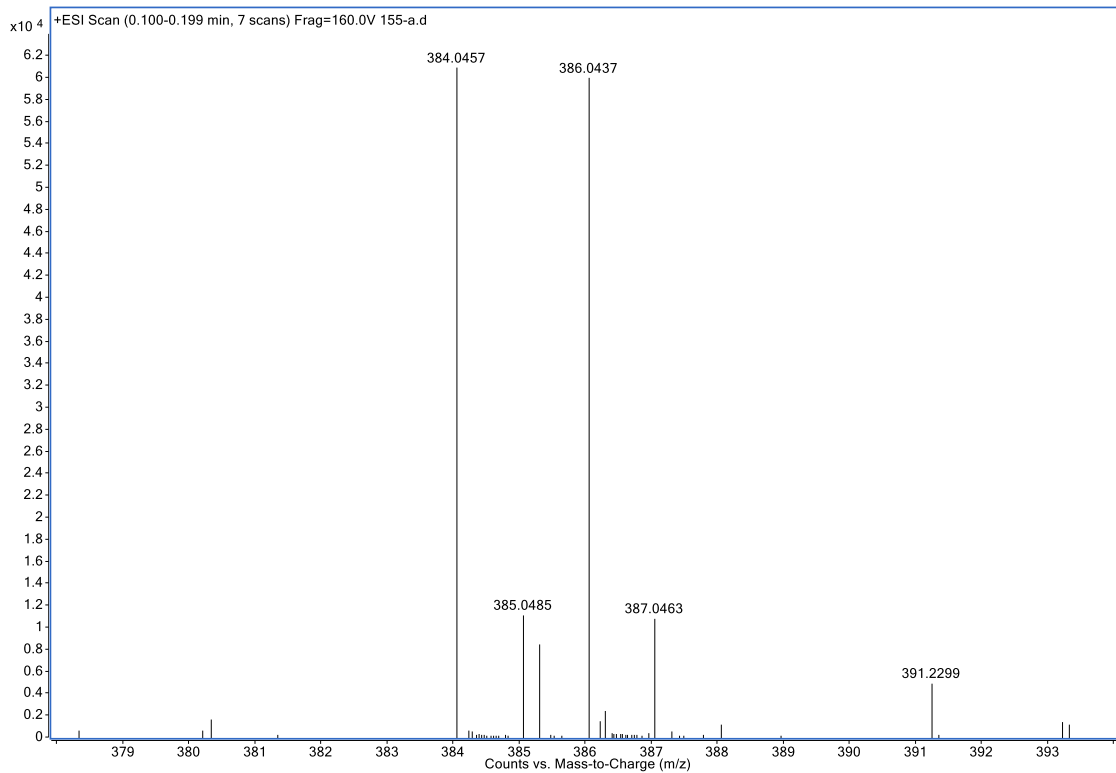


¹H NMR of **35** (DMSO-*d*₆, 500 MHz)

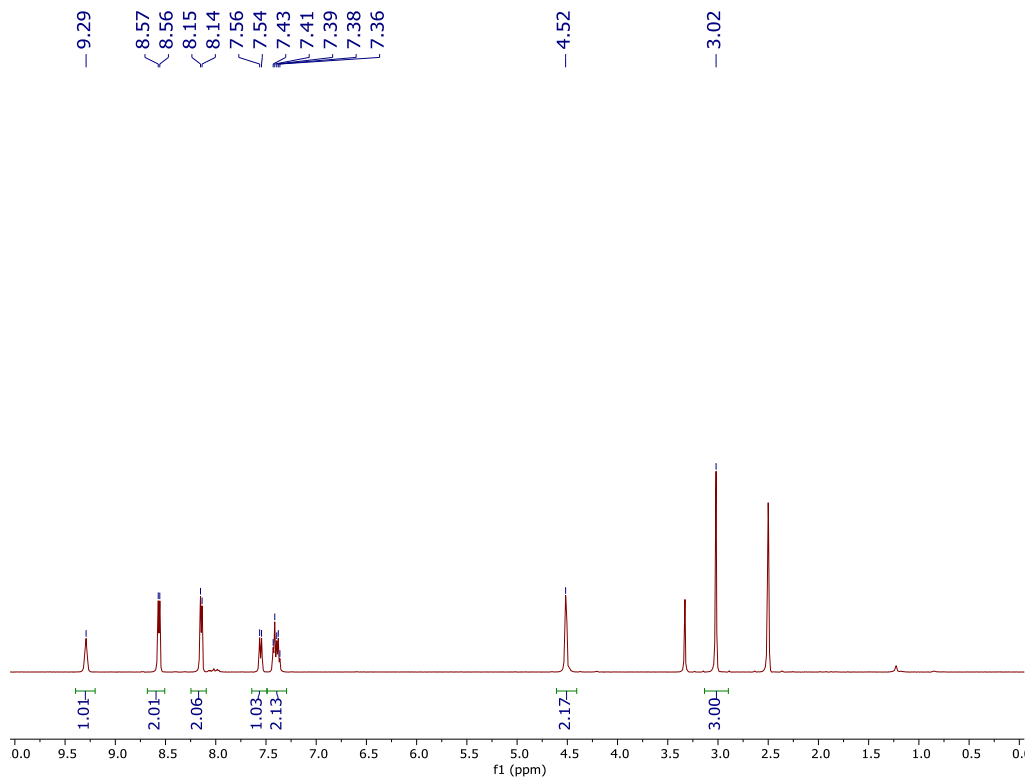


¹³C NMR of **35** (DMSO-*d*₆, 125 MHz)

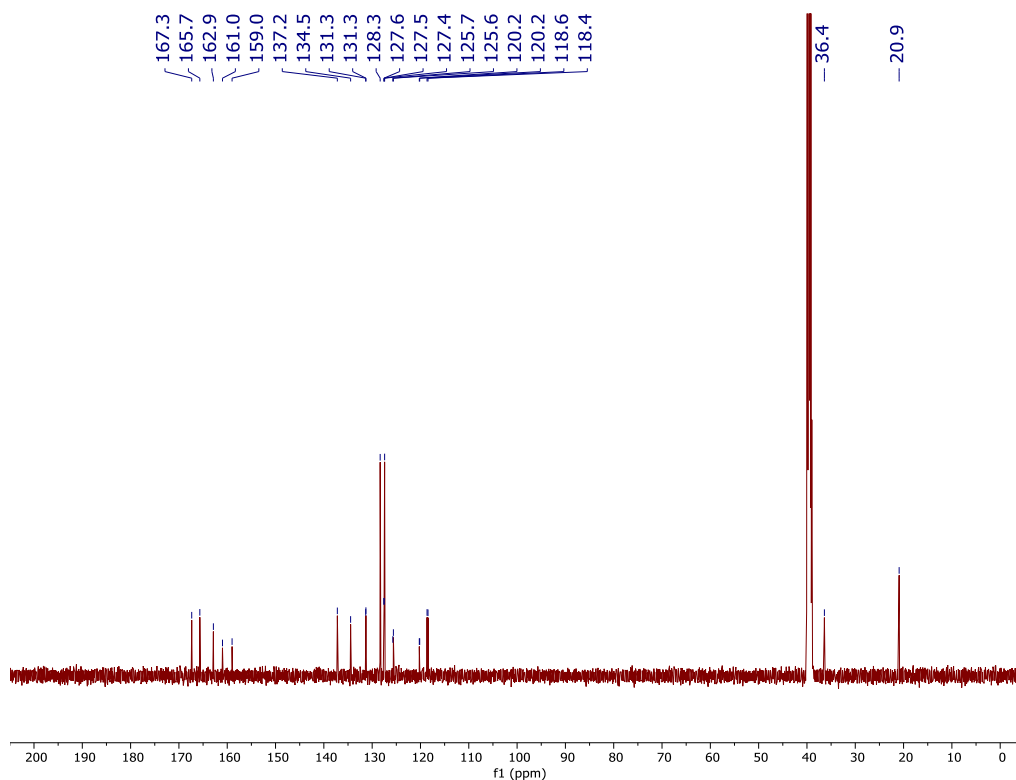




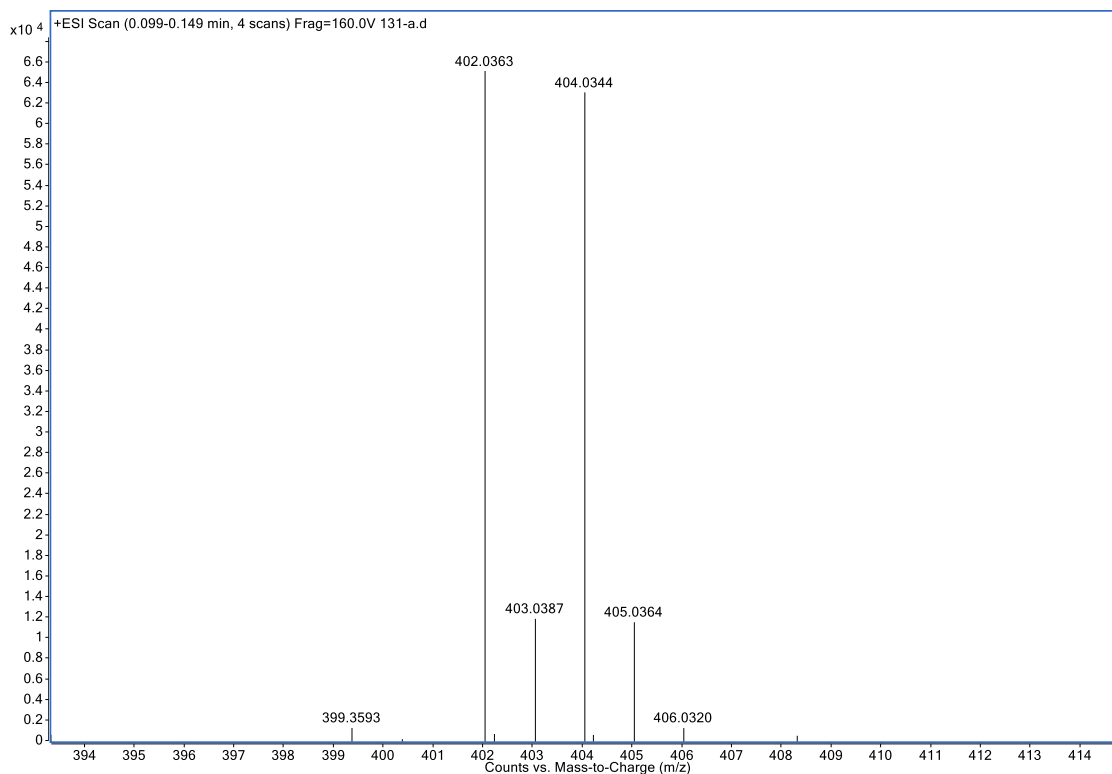
¹H NMR of **36** (DMSO-d₆, 500 MHz)



¹³C NMR of **36** (125 MHz, DMSO-*d*₆)



HRMS of **36**



References

- (1) Garson, M. J., and Simpson, J. S. (2004) Marine isocyanides and related natural products—structure, biosynthesis and ecology. *Nat. Prod. Rep.*, 21, 164–179.
- (2) Emsermann, J., Kahl, U., and Opatz, T. (2016) Marine isonitriles and their related compounds. *Mar. Drugs* 14, 1–83.
- (3) Bhat, V., Dave, A., MacKay, J. A., and Rawal, V. H. (2014) The Chemistry of Hapalindoles, Fischerindoles, Ambiguines, and Welwitindolinones. *Alkaloids Chem. Biol.* 1st ed. Elsevier Inc.
- (4) Schnermann, M. J., and Shenvi, R. A. (2015) Syntheses and biological studies of marine terpenoids derived from inorganic cyanide. *Nat. Prod. Rep.* 32, 543–577.
- (5) Kobayashi, J., Tsuda, M., Nemoto, A., Tanaka, Y., Yazawa, K., and Mikami, Y. (1997) Brasilidine A, a new cytotoxic isonitrile-containing indole alkaloid from the actinomycete *Nocardia brasiliensis*. *J. Nat. Prod.* 60, 719–720.
- (6) Tabata, Y., Hatsu, M., Amano, S., Shimizu, A., and Imai, S. (1995) SF2768, a new isonitrile antibiotic obtained from *Streptomyces*. *Sci. Rep. Meiji Seika Kaisha* 34, 1–9.
- (7) Evans, J. R., Napier, E. J., and Yates, P. (1976) Isolation of a new antibiotic from a species of *Pseudomonas*. *J. Antibiot. (Tokyo)*. 29, 850–852.
- (8) Wang, L., Zhu, M., Zhang, Q., Zhang, X., Yang, P., Liu, Z., Deng, Y., Zhu, Y., Huang, X., Han, L., Li, S., and He, J. (2017) Diisonitrile natural product SF2768 functions as a chalkophore that mediates copper acquisition in *Streptomyces thioluteus*. *ACS Chem. Biol.* 12, 3067–3075.
- (9) Sandoval, I. T., Manos, E. J., Van Wagoner, R. M., Delacruz, R. G. C., Edes, K., Winge, D. R., Ireland, C. M., and Jones, D. A. (2013) Juxtaposition of chemical and mutation-induced developmental defects in zebrafish reveal a copper-chelating activity for kalihinol F. *Chem. Biol.* 20, 753–763.
- (10) König, G. M., Wright, A. D., and Angerhofer, C. K. (1996) Novel potent antimalarial diterpene isocyanates, isothiocyanates, and isonitriles from the tropical marine sponge *Cymbastela hooperi*. *J. Org. Chem.* 61, 3259–3267.
- (11) Amano, S. I., Sakurai, T., Endo, K., Takano, H., Beppu, T., Furihata, K., Sakuda, S., and Ueda, K. (2011) A cryptic antibiotic triggered by monensin. *J. Antibiot. (Tokyo)*. 64, 703.
- (12) Moore, R. E., Cheuk, C., and Patterson, G. M. L. (1984) Hapalindoles: New alkaloids from the blue-green alga *Hapalosiphon fontinalis*. *J. Am. Chem. Soc.* 106, 6456–6457.
- (13) Chang, C. W. J., Patra, A., Roll, D. M., Matsumoto, G. K., and Clardy, J. (1984) Kalihinol-A, a highly functionalized diisocyno diterpenoid antibiotic from a sponge. *J. Am. Chem. Soc.* 106, 4644–4646.
- (14) Baker, J. T., Wells, R. J., Oberhaensli, W. E., and Hawes, G. B. (1976) A new diisocyanide of novel ring structure from a sponge. *J. Am. Chem. Soc.* 98, 4010–4012.

- (15) Di Blasio, B., Fattorusso, E., Magno, S., Mayoi, L., Pedone, C., Santacroce, C., and Sica, D. (1976) Axisonitrile-3, axisothiocyanides-3 and axamide-3. Sesquiterpenes with a novel spiro[4,5]decane skeleton from the sponge *Axinella cannabina*. *Tetrahedron* 32, 473–478.
- (16) Stratmann, K., Moore, R. E., Bonjouklian, R., Deeter, J. B., Patterson, G. M. L., Shaffer, S., Smith, C. D., and Smitka, T. A. (1994) Welwitinolides, unusual alkaloids from the blue-green alga *Hapalisophon welwitschii* and *Westiella intricata*. Relationship to fisherindoles and hapalindoles. *J. Am. Chem. Soc.* 116, 9935–9942.
- (17) Brady, S. F., and Clardy, J. (2005) Cloning and heterologous expression of isocyanide biosynthetic genes from environmental DNA. *Angew. Chem. Int. Ed. Engl.* 44, 7063–7065.
- (18) Brady, S. F., Bauer, J. D., Clarke-Pearson, M. F., and Daniels, R. (2007) Natural products from isnA-containing biosynthetic gene clusters recovered from the genomes of cultured and uncultured bacteria. *J. Am. Chem. Soc.* 129, 12102–12103.
- (19) Harris, N. C., Sato, M., Herman, N. A., Twigg, F., Cai, W., Liu, J., and Zhu, X. (2017) Biosynthesis of isonitrile lipopeptides by conserved nonribosomal peptide synthetase gene clusters in Actinobacteria. *Proc. Natl. Acad. Sci.* 114, 7025–7030.
- (20) Imming, B. P., Mohr, R., Miiller, E., Overheu, W., and Seitz, G. (1982) [4+1] cycloaddition of isocyanides to 1,2,4,5-tetrazines: A novel synthesis of pyrazole. *Angew. Chem. Int. Ed.* 21, 20133.
- (21) Henning, S., Neves, A. A., Stairs, S., Brindle, K. M., and Leeper, F. J. (2011) Exploring isonitrile-based click chemistry for ligation with biomolecules. *Org. Biomol. Chem.* 9, 7303–7305.
- (22) Stairs, S., Neves, A. A., Stöckmann, H., and Wainman, Y. A. (2013) Metabolic glycan imaging by isonitrile–tetrazine click chemistry. *ChemBioChem* 14, 1063–1067.
- (23) Tu, J., Xu, M., Parvez, S., Peterson, R. T., and Franzini, R. M. (2018) Bioorthogonal removal of 3-isocyanopropyl groups enables the controlled release of fluorophores and drugs in vivo. *J. Am. Chem. Soc.* 140, 8410–8414.
- (24) Castro-Falcon, G., Hahn, D., Reimer, D., and Hughes, C. C. (2016) Thiol probes to detect electrophilic natural products based on their mechanism of action. *ACS Chem. Biol.* 11, 2328–2336.
- (25) Reimer, D., and Hughes, C. C. (2017) Thiol-based probe for electrophilic natural products reveals that most of the ammosamides are artifacts. *J. Nat. Prod.* 80, 126–133.
- (26) Castro-Falcón, G., Seiler, G., Demir, O., Rathinaswamy, M., Hamelin, D., Hoffmann, R. M., Makowski, S., Letzel, A.-C., Field, S., Burke, J., Amaro, R. E., and Hughes, C. C. (2018) Neolymphostin A is a Covalent Phosphoinositide 3-Kinase (PI3K)/Mammalian Target of Rapamycin (mTOR) Dual Inhibitor that Employs an Unusual Electrophilic Vinylogous Ester. *J. Med. Chem.* 61, 10463–10472.
- (27) Castro-Falcón, G., Millán-Aguiñaga, N., Roullier, C., Jensen, P. R., and Hughes, C. C. (2018) Nitrosopyridine probe to detect polyketide natural products with conjugated alkenes: Discovery of novodaryamide and nocarditriene. *ACS Chem. Biol.* 13, 3097–3106.

(28) Tørring, T., Shames, S. R., Cho, W., Roy, C. R., and Crawford, J. M. (2017) Acyl histidines: New N-acyl amides from *Legionella pneumophila*. *ChemBioChem* 18, 638–646.

(29) Kozma, E., Estrada Girona, G., Paci, G., Lemke, E. A., and Kele, P. (2017) Bioorthogonal double-fluorogenic siliconrhodamine probes for intracellular super-resolution microscopy. *Chem. Commun.* 53, 6696–6699.

(30) Cserép, G. B., Demeter, O., Bätzner, E., Kállay, M., Wagenknecht, H. A., and Kele, P. (2015) Synthesis and Evaluation of Nicotinic Acid Derived Tetrazines for Bioorthogonal Labeling. *Synth.* 47, 2738–2744.

(31) Santa, T. (2013) Derivatization in LC-MS bioanalysis, in Handbook of LC-MS bioanalysis: Best practices, experimental protocols, and regulations (Li, W., Zhang, J., and Tse, F. L. S., Eds.), pp 239–248. John Wiley & Sons, Inc.

(32) Qi, B. L., Liu, P., Wang, Q. Y., Cai, W. J., Yuan, B. F., and Feng, Y. Q. (2014) Derivatization for liquid chromatography-mass spectrometry. *Trends Anal. Chem.* 59, 121–132.

Chapter 2 is a reprint, in full, of work that is in preparation for submission. Gabriel Castro-Falcón, Grant S. Seiler, and Chambers C. Hughes. “An optimized tetrazine probe to detect isocyanide-containing natural products in extracts.” 2019. The dissertation author was a co-primary investigator and author of this material.

Chapter 3

“Neolymphostin A Is a Covalent Phosphoinositide 3-Kinase (PI3K)/Mammalian Target of Rapamycin (mTOR) Dual Inhibitor That Employs an Unusual Electrophilic Vinylogous Ester.”

Gabriel Castro-Falcón,^{1,5} Grant S. Seiler,^{1,2,5} Özlem Demir,² Manoj K. Rathinaswamy,³ David Hamelin,³ Reece M. Hoffmann,³ Stefanie L. Makowski,⁴ Anne-Catrin Letzel,¹ Seth J. Field,⁴ John E. Burke,³ Rommie E. Amaro,² Chambers C. Hughes^{1*}

¹Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, USA, 92093

²Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California, USA, 92093

³Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia, Canada, V8W 2Y2

⁴School of Medicine, University of California, San Diego, La Jolla, California, USA, 92093

⁵These authors contributed equally to this work.

Abstract

Using a novel chemistry-based assay for identifying electrophilic natural products from unprocessed extracts, we identified the PI3-kinase/mTOR dual inhibitor neolymphostin A from *Salinispora arenicola* CNY-486. The method further showed that the vinylogous ester substituent on the neolymphostin core was the exact site for enzyme conjugation. Tandem MS/MS experiments on PI3K α treated with the inhibitor revealed that neolymphostin covalently modified Lys802 with a shift in mass of +306 amu, corresponding to addition of the inhibitor and elimination of methanol. The binding pose of the inhibitor bound to PI3K α was modelled, and hydrogen-deuterium exchange mass spectrometry experiments supported this model. Against a panel of kinases, neolymphostin showed good selectivity for PI3-kinase and mTOR. In addition, the natural product blocked AKT phosphorylation in live cells with an IC₅₀ of ~3 nM. Taken together, neolymphostin is the first reported example of a covalent kinase inhibitor from the bacterial domain of life.

Introduction

Natural products have played a key role in our understanding of kinase inhibition and in the development of drugs.¹ Natural product kinase inhibitors have been discovered from bacteria, fungi, plants, and marine sponges. Staurosporine, isolated in 1977 from *Streptomyces staurosporeus*, and its analogues are pan-kinase inhibitors that have been adapted to yield several clinical candidates.² Recently, a semisynthetic derivative of staurosporine, midostaurin (*N*-benzoyl staurosporine), was approved by the FDA for the treatment of hematologic cancers in combination with standard chemotherapy.³ Other ATP-competitive natural product kinase inhibitors include flavonoids such as quercetin, genistein and myricetin, terpenoids like the nakijiquinones and celastrol, and nitrogenous heterocycles like olomucine and hymenialdisine. The allosteric mTOR inhibitor rapamycin from *Streptomyces hygroscopicus* is an immunosuppressive drug used to prevent organ transplant rejection.⁴

Spurred by the discovery of natural product kinase inhibitors, the development of novel potent and selective kinase inhibitors is a key priority of pharmaceutical companies, with 41 small molecule kinase inhibitors being clinically approved for a variety of diseases (39 protein kinase inhibitors and 2 lipid kinase inhibitors) as of April 2018.⁵ Extensive efforts have gone into the development of lipid kinase inhibitors, specifically for the phosphoinositide 3-kinase (PI3K) family of enzymes, due to the key role of PI3Ks in cancer, immunodeficiencies, and metabolic disorders.⁶ There are more than 80 clinical trials of PI3K inhibitors ongoing,⁷ with specific inhibitors of the p110 δ isoform of PI3K approved by the FDA for the treatment of multiple blood cancers.⁸

Selective kinase inhibition at the ATP-binding site is difficult since this binding pocket is highly conserved.⁹ Leveraging the finding that several poorly conserved cysteine residues are present in the ATP binding pocket, a new strategy for designing selective ATP-competitive inhibitors was developed in the 1990s that involved appending an electrophilic warhead at a position on a small molecule that allows for conjugation to the cysteine residue.¹⁰⁻¹⁵ This approach has led to the FDA approval of selective

protein tyrosine kinase inhibitors ibrutinib (2013), afatinib (2013), osimertinib (2015), neratinib (2017), and acalabrutinib (2017), which all contain an electrophilic acrylamide moiety for cysteine conjugation, for the treatment of breast cancer, non-small cell lung cancer, and several blood cancers.

Covalent kinase inhibition is a strategy that also evolved in fungi and plants.¹ Wortmannin and related furanosteroids have been demonstrated to inhibit PI3K and the mammalian target of rapamycin (mTOR) in a covalent manner via conjugation to a key lysine residue.¹⁶⁻¹⁸ In addition, celastrol and parthenolide inhibit I κ B kinase via cysteine conjugation,^{19,20} and hypothemycin and related resorcylic acid lactones inhibit their kinase targets via cysteine conjugation.²¹ To date, there are no known bacterial natural products that function as covalent kinase inhibitors.

Results

In our ongoing efforts to identify naturally-occurring covalent inhibitors, we screened over 200 *Salinispora* extracts using thiol-based probe **1** (Fig. 1a).²² Probe **1** is highly conspicuous in an LC-MS chromatogram due to the UV absorbance of the benzamide ring and the isotopic pattern of the bromine atom, so adducts resulting from addition of **1** can be rapidly detected and isolated in complex extracts. To date, the probe has been demonstrated to selectively label a variety of validated electrophilic natural products in crude extracts, such as penicillin G (β -lactam), salinosporamide A (β -lactone), parthenolide (α,β -unsaturated enone), and wortmannin ($\alpha,\beta,\gamma,\delta$ -unsaturated enone), regardless of the exact identity of the nucleophilic amino acid (Cys, Thr, Ser, Lys) in the cellular target that is responsible for covalent attachment. However, the method, and others like it,²³⁻²⁵ has yet to lead to the identification of a secondary metabolite with an unestablished covalent mechanism of action.

Treatment of crude extract from *Salinispora arenicola* strain CNY-486 with **1** in triethylamine and *N,N*-dimethylformamide produced two pairs of diastereomeric adducts with a mass isotope pattern that indicated the presence of a single bromine atom (**Fig. 1b** and Supporting Information **Figs. S1-S3**).

The formation of diastereomeric products, based on previous work,²² is indicative of an electrophilic natural product containing a Michael acceptor, since the reaction involves the formation of new stereocenters with little to no diastereoselectivity. A pair of diastereomers (**2a**, **2b**) were purified by reversed phase C18 HPLC, and their structures were determined using a combination of NMR and MS techniques (Supporting Information **Tables S1** and **S2**). Retrosynthetic analysis led to the identification of neolymphostin A (**4**), which revealed that reaction occurred at the vinylogous ester substituent (C-11) of the natural product.²⁶⁻²⁹ The other diastereomeric pair of adducts (**3a**, **3b**) stemming from neolymphostin B (**5**) were also isolated and characterized (**Fig. 1b**, Supporting Information **Tables S3** and **S4**).

As the lymphostins have been reported to inhibit lymphocyte-specific tyrosine kinase (Lck), phosphoinositide 3-kinase (PI3K), and the mammalian target of rapamycin (mTOR),^{26,28,29} we first set out to determine if our isolated material showed similar activity using an active-site dependent competition binding assay described in the Supporting Information.³⁰ Indeed, neolymphostin (**4**) showed strong affinity for all four isoforms of PI3K [K_d = 0.88 (α), 5.7 (β), 3.5 (γ), and 4.9 (δ) nM] and mTOR (K_d = 10 nM) (Supporting Information **Table S5**). Binding to Lck, however, was less pronounced (K_d = 4.6 μ M). For comparison, wortmannin, a fungal natural product and well-established covalent PI3K inhibitor, was also measured in the binding assay. This metabolite showed strong affinity for PI3K [K_d = 5.4 (α), 7.6 (β), 15 (γ), and 5.5 (δ) nM] but little affinity for mTOR (K_d = 9.2 μ M) (Supporting Information **Table S6**).¹⁶⁻¹⁸

In order to quantify the electrophilicity of the vinylogous ester and to preserve precious natural product, the reactivity of model electrophile (*E*)-3-methoxy-1-(pyridin-2-yl)prop-2-en-1-one (**6**) toward glutathione was measured in 100 mM phosphate buffer at pH 7.4 according to the NMR-based method of Flanagan, et al. (**Fig. 2**).³¹ This reaction proceeded to **7** with a $t_{1/2}$ = 5.4 min at 37 °C ($t_{1/2}$ = 13 min at 25 °C) (Supporting Information **Fig. S4**). Highlighting the enhanced reactivity of the neolymphostin warhead

over an FDA-approved acrylamide-containing kinase inhibitor, ibrutinib was reported to convert to its glutathione adduct with a $t_{1/2} = 7.36$ h under similar conditions.³²

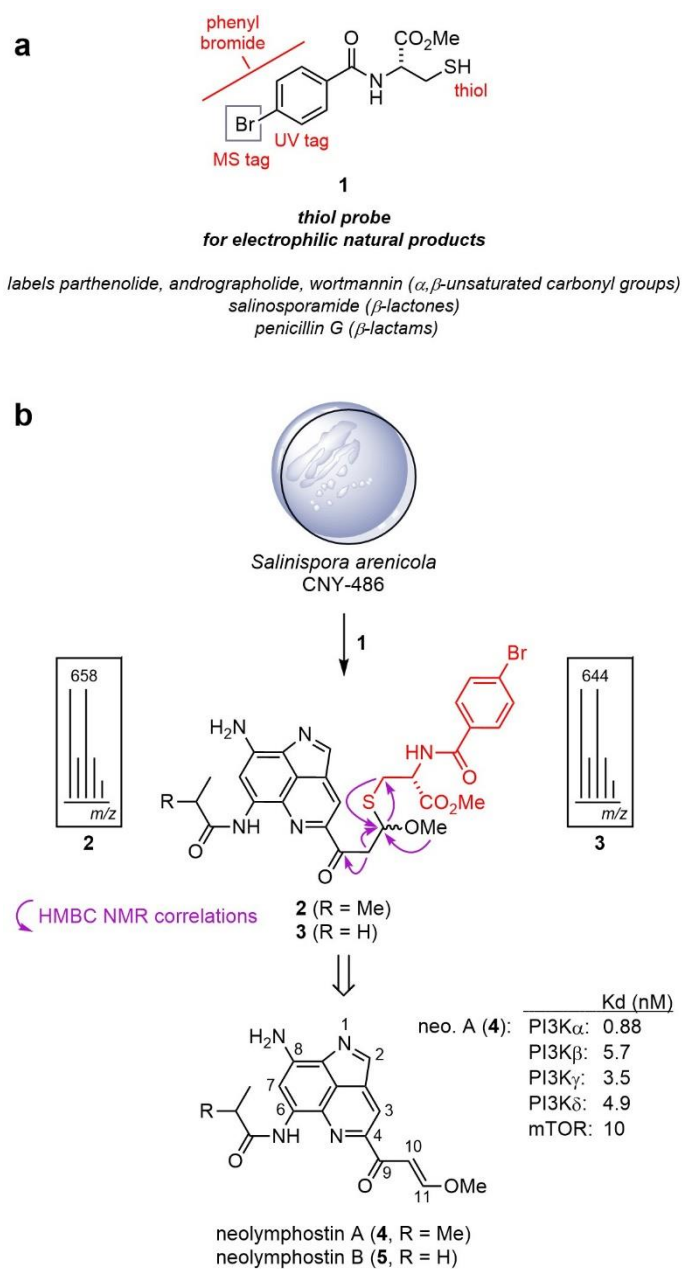


Figure 3.1. Identification of neolymphostin as an electrophilic natural product. a) Cysteine-based probe **1** for the discovery of electrophilic natural products. b) Treatment of extract from *Salinispora* strain CNY-486 with **1** produced brominated adducts **2** and **3**, both as a mixture of diastereomers, which are derived from neolymphostin A (**4**) and B (**5**).

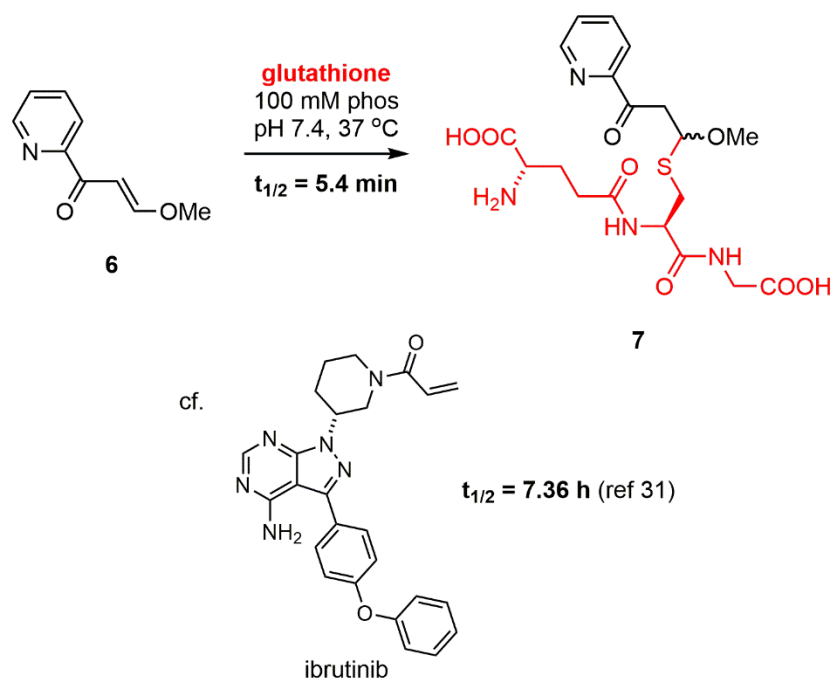


Figure 3.2. Reactivity of neolymphostin model electrophile **6** with glutathione at pH 7.4.

To rapidly assess whether the interaction of neolymphostin with the PI3Ks was reversible or irreversible, we then measured the K_d for neolymphostin-treated PI3K α and PI3K γ with and without an intermediate dilution step according to the scanKINETIC platform (**Fig. 3a** and Supporting Information **Table S7**). In study arm A and C, the inhibitor and kinase were combined and equilibrated for 6 h and 1 h, respectively, before measuring the K_d . The fact that these dissociation constants are comparable suggests rapid association kinetics. In study arm B the inhibitor and kinase were combined and equilibrated for 1 h, diluted 30-fold, and re-equilibrated before measuring the K_d , while in study arm D the pair were immediately diluted and equilibrated before measuring the K_d . Importantly, dilution after inhibitor and kinase were first combined gave dissociation constants similar to study arm C, which is expected of a covalent inhibitor. Known covalent kinase inhibitors like wortmannin show the same pattern.

To verify that the inhibition identified was due to covalent attachment and to identify the site of modification, we carried out mass spectrometry tandem MS/MS experiments on the full-length complex of PI3K α (composed of the proteins p110 α and p85 α). The tandem MS/MS data of pepsin-generated peptides was queried for modification of lysine side chains, since docking experiments suggested an appropriately positioned lysine residue was present in the binding pocket (see below). Searches were carried out for lysine modifications of either 306.11 or 338.14 amu, which correspond to the mass of neolymphostin with or without subsequent elimination of methanol. Covalent modification was detected for only a single amino acid in PI3K α , corresponding to a +306.11 amu modification on Lys802 (**Fig. 3b**), and a timecourse experiment showed that modification of the kinase was complete within 2 min (Supporting Information **Fig. S5**). This is the same catalytic residue that is covalently modified upon addition of wortmannin, revealing that bacteria and fungi produce metabolites able to target this conserved residue in the PI3K family of enzymes.¹⁶

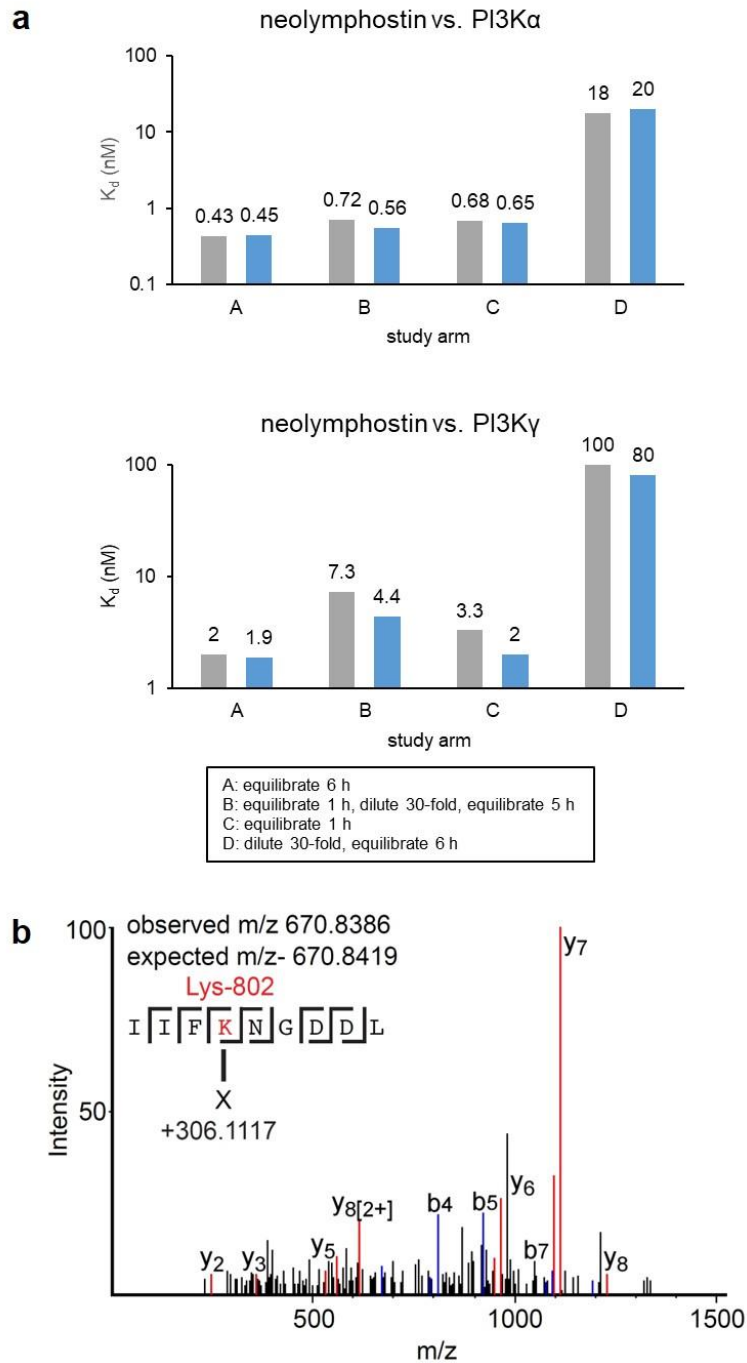


Figure 3.3. The interaction of neolymphostin with the PI3Ks is covalent. a) Incubation studies with PI3K α and PI3K γ performed in duplicate (first: grey; second: blue) b) Tandem MS/MS spectrum of a neolymphostin-modified PI3K α peptide containing Lys802. Different b and y fragments and their location are indicated, with both the observed and expected masses of the precursor annotated.

Docking neolymphostin into PI3K α and mTOR indicated favorable binding at the hinge region for both enzymes with docking scores of -7.58 kcal/mol for PI3K α and -7.45 kcal/mol for mTOR. In the best

predicted binding pose for PI3K α , hydrogen bonding occurs between the C-8 amino group of the inhibitor and the Val851 backbone carbonyl group of the kinase and between the N-1 endocyclic imine nitrogen atom of the inhibitor and the Val851 backbone amino group of the kinase (**Fig. 4** and Supporting Information **Fig. S6**). Favorable pi-pi stacking interactions between the aromatic neolymphostin core and Tyr836 are also observed. In the mTOR model, comparable hydrogen bonding interactions occur between Val2240 and the inhibitor, as well as comparable pi-pi stacking and pi-cation interactions between Trp2239 and the neolymphostin core (Supporting Information **Fig. S7**).

Homologous nucleophilic lysine residues in both enzymes, Lys802 in PI3K α and Lys2187 in mTOR, lie close to the methoxyenone arm of neolymphostin. The distance between the nitrogen atom of the lysine side-chain and the electrophilic C-11 site of neolymphostin, which corresponds to the nascent bond, is 3.29 Å in PI3K α and 6.08 Å in mTOR. Closer inspection of the neolymphostin binding pose in mTOR revealed that Lys2187 can position itself nearer to the reactive site via a slight C-C single bond rotation in its side chain. Taking active site flexibility into account is likely to generate a binding pose in mTOR similar to the case of PI3K α . When we tried induced fit docking as an alternative approach, the binding poses for both cases were similar to the binding poses obtained by initial docking experiments except for some small active site rearrangements and slight differences in the ligand position and conformation (Supporting Information **Figs. S8** and **S9**). Notably, the distance between the ϵ -amino nitrogen atom of Lys2187 in mTOR and C-11 of neolymphostin decreased to 3.98 Å. Finally, we computationally formed a chemical bond between ϵ -amino nitrogen atom of Lys802 and C-11 of neolymphostin, deleted the methoxy group that would be eliminated as a result of the chemical reaction, and energy-minimized the Lys802 adduct using Schrödinger software as depicted in **Fig. 4** (Supporting Information **Fig. S10**).

We then used hydrogen deuterium exchange mass spectrometry (HDX-MS) to experimentally verify the putative docking pose from our docking experiments, This technique measures the exchange

rate of amide hydrogen atoms with deuterated solvent, and as this exchange is dependent on secondary structure, it is a powerful tool to examine secondary structure dynamics.³³ It has previously been used to examine the binding of certain small molecule inhibitors of PI3K α .³⁴ HDX-MS experiments were performed on the full length complex of p110 α and p85 α . Deuterium exchange was carried out for both 3 and 300 s in the presence and absence of neolymphostin, with deuterium incorporation being localized through digestion of the protein into peptide fragments and subsequent MS analysis. The full set of p110 α peptides that were analyzed and their deuterium incorporation are shown in Supporting Information **Table S8**. Significant differences in H/D exchange rates (both >5% and >0.4 Da difference with a p value<0.01 from triplicate samples) was observed for only a single region, encompassing amino acids 848-856 in p110 α (blue ribbon, **Fig. 4**). This region, corresponding to the hinge region at the interface of the N-lobe and C-lobe of the p110 α kinase domain, contains the atoms that make hydrogen bonds with the inhibitor in the docked model, which is in agreement with the *in silico* model.

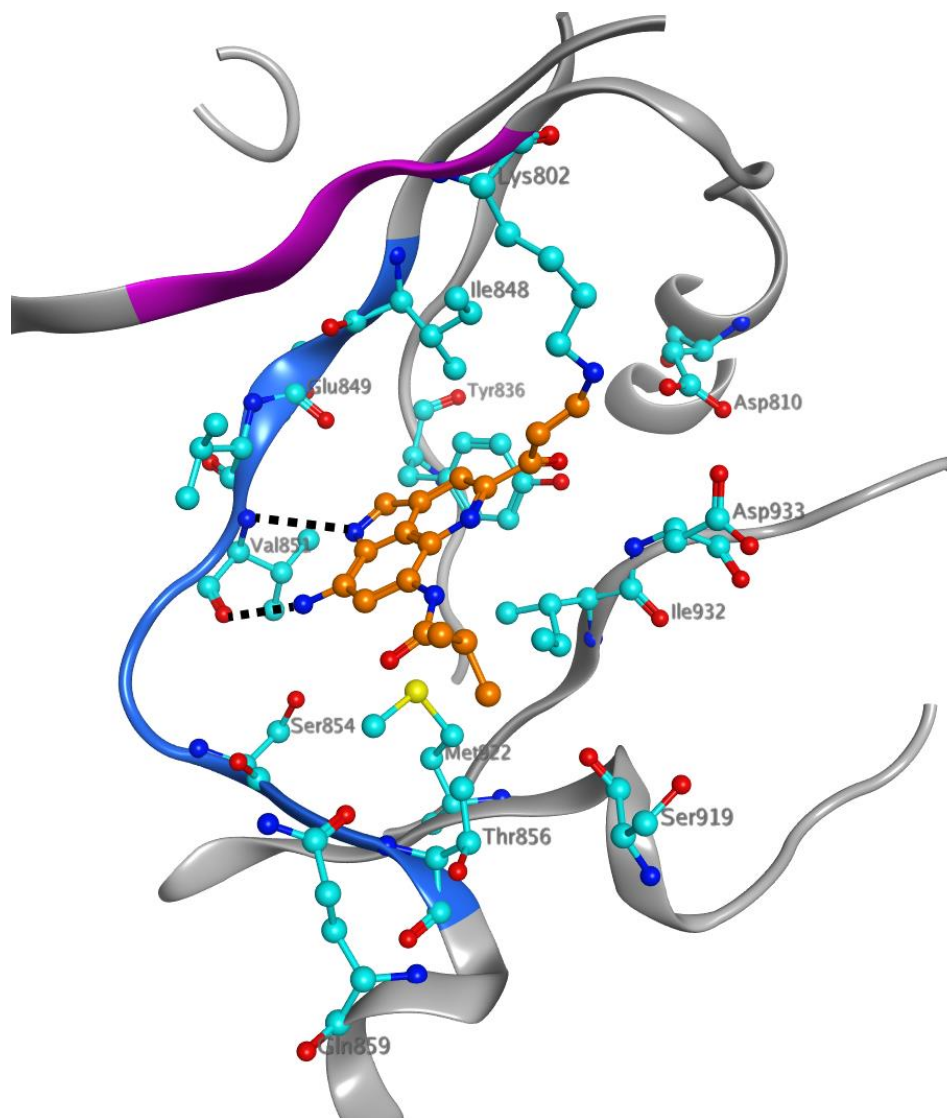


Figure 3.4. HDX-MS data mapped on the docked model of neolymphostin bound to PI3K α (pdbID:5SW8, chain A) with the covalent modification site of Lys802 indicated. PI3K α is shown in silver ribbons while amino acids that differ greater than 5% and 0.4 Da between apo- and neolymphostin-bound protein (amino acids 848-856) are highlighted in blue. Amino acids that were not covered in the HDX/MS experiment (amino acids 799-802) are highlighted in pink. The ligand and the amino acids that interact with it are shown in ball-and-stick form with N = blue, O = red, S = yellow, C(inhibitor) = orange, and C(ligand) = cyan. The black dashed lines indicate hydrogen bonds. H atoms are not shown for clarity.

To gauge the electrophilicity of the vinylogous ester toward more relevant amine nucleophiles, the reactivities of model electrophile **6** toward *N*^α-acetyl lysine was measured. In 100 mM phosphate buffer at pH 7.4, no reaction was observed in the presence of a ten-fold excess of *N*^α-acetyl lysine at 25 °C. At pH 8.1, however, *N*^α-acetyl lysine did add to **6** with a $t_{1/2} = 60$ min at 25 °C to give adduct **8**, while a competing hydrolysis reaction proceeded with a $t_{1/2} \sim 24$ h to produce **9** (Fig. 5, Supporting Information Fig. S11). Model electrophile **6** and a stoichiometric amount of *N*^α-acetyl lysine were then subjected to the same reaction conditions at higher concentration in order to unambiguously determine the structure of the covalent adduct by proton and 2D NMR analyses (Supporting Information Table S9). The pH requirement for reactivity between the two partners suggests that the pK_a of Lys802 in PI3K (and Lys2187 in mTOR) is perturbed by the local environment.

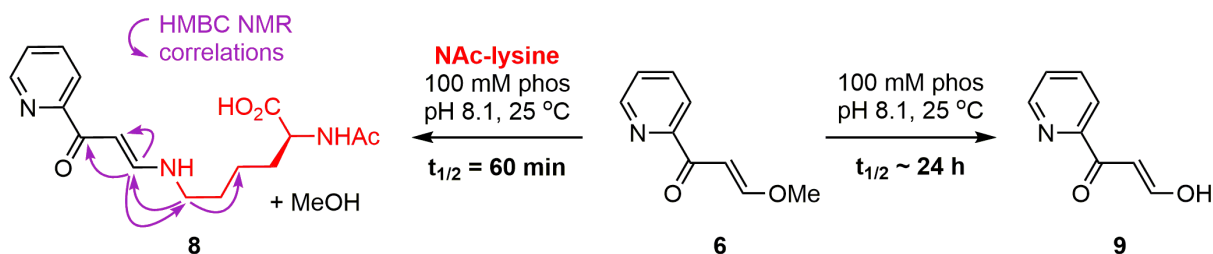


Figure 3.5. Competing reactivity of model electrophile **6** with *N*^α-acetyl lysine and water at pH 8.1.

Although we had at this point shown that neolymphostin strongly inhibits PI3K (and likely mTOR) in a covalent manner, we were interested in assessing the overall kinase selectivity of the inhibitor. So, the natural product was screened *in vitro* against a panel of 97 kinases including representatives from all seven major kinase families in addition to several atypical, lipid, pathogen, and mutant kinases (Fig. 6, Supporting Information Table S10 and Fig. S12).^{30,35} At a concentration of 1 μM neolymphostin showed strong affinity for the PI3Ks, as expected. The inhibitor also showed moderate affinity for Polo-like kinase 4 (PLK4), aurora kinases A and B (AURKA/B), and the c-Jun N-terminal kinases

1, 2, and 3 (JNK1/2/3) compared to control. The affinity of the inhibitor for AURKA/B may help counteract a recently established process in cancer cells for eluding the effects of PI3K inhibition.³⁶ For the majority of kinases the alkaloid had very little affinity compared to control. A selectivity score, S-score(35), calculated by dividing the number of kinases with affinities to neolymphostin of less than 35% compared to control by the total number of kinases tested, was equal to 0.20 (Supporting Information **Table S11**). Neolymphostin's S-score(35) is similar to that of FDA-approved drugs like erlotinib, sorafenib, and dasatinib.

Next, we examined the ability of neolymphostin to inhibit PI3K activity in cells (**Fig. 6b**). Serum-starved HeLa cells were stimulated with the growth factor EGF to activate the EGFR, which leads to class I PI3K-dependent phosphorylation of AKT and PI3K-independent phosphorylation of ERK1/2. We first observed that EGF stimulation of serum-starved cells indeed led to robust phosphorylation of EGFR, AKT, and ERK1/2. As expected, pretreatment with the known PI3K inhibitor wortmannin blocked phosphorylation of AKT on Ser473 and Thr308 without affecting phosphorylation of EGFR or ERK1/2. Similarly, pretreatment with neolymphostin inhibited AKT phosphorylation, with an estimated IC_{50} of ~ 3 nM, but had no apparent effect on EGFR or ERK1/2 phosphorylation. In contrast, AKT phosphorylation was largely unaffected by pretreatment with neolymphostinol (**10**), which lacks an electrophilic moiety (**Fig 6c**). Thus, we found that neolymphostin, but not the inactive neolymphostinol analogue, potently inhibits class I PI3K activity in cells.

Lastly, neolymphostin was tested in the National Cancer Institute 60-cell line anticancer screen (Supporting Information **Figs. S13-S16**). GI_{50} values, which indicate the concentration of compound that causes 50% inhibition of cell proliferation, ranged over two orders of magnitude from 6-840 nM. The natural product showed some selectivity toward leukemia, renal cancer, and breast cancer cell lines. At the highest concentration in the dose-response curves (50 μ M) a uniform increase in percentage growth was observed when compared to the lower 5 μ M concentration. The fact that neolymphostin appears to

Discussion and Conclusions

The structure and kinase inhibitory properties of lymphostin have been known for almost two decades but crucial information concerning the compound's mechanism of action has been absent. Here, we show that neolymphostin, a closely-related derivative of lymphostin, is a covalent kinase inhibitor that competitively binds to the ATP-binding pocket of the PI3-kinases and mTOR. The alkaloid then reacts covalently with a conserved lysine residue within the pocket that is involved in the kinases' key phospho-transfer reaction, making the development of resistance through a point mutation highly unlikely.³⁸ The vinylogous ester substituent of the lymphostins exhibits the precise electrophilicity needed to engage a pKa-perturbed amine functionality, and this conjugate addition reaction is then followed by an elimination reaction to produce a stable vinylogous amide. This study introduces a unique warhead that medicinal chemists may find useful in the development of synthetic lysine-targeting covalent inhibitors, an area which, compared to cysteine-targeting covalent inhibitors, is severely limited.^{39,40}

As both are covalent PI3K inhibitors that target the same lysine residue, the lymphostins and wortmannin are strikingly similar in terms of function. However, there are key differences between the two metabolites. First, the lymphostins are produced by bacteria while wortmannin was discovered in fungi. The fact that the same mechanism of action evolved separately in the two kingdoms of life is remarkable. Second, according to binding experiments, lymphostin is a dual PI3K/mTOR inhibitor while wortmannin selectively targets the PI3Ks. Third, although they possess a similar electrophilic moiety, the two molecules have distinct biosynthetic origins; the lymphostins are tryptophan-derived alkaloids while wortmannin is an furanosteroid-type terpenoid. As such, the lymphostins are much more "drug-like," possessing superior water solubility and bearing a combination of hydrogen bond donors and acceptors.

The pyrrolo[4,3,2-*de*]quinoline core of the lymphostins strongly resembles the purine heterocycle of ATP. In fact, the lymphostins appear to bind to the kinase ATP-binding pocket in the same

manner as ATP, forming two hydrogen bonds to the kinase hinge region via the amine substituent (H bond donor) and the endocyclic imine nitrogen (H bond acceptor).⁴¹ The natural product is a “sticky” ATP mimic that first binds like ATP and then couples covalently to a conserved lysine residue that is essential for the enzyme’s function. Given that the lymphostins possess a relatively simple ATP-like structure devoid of any chirality, the level of selectivity that neolymphostin demonstrated in the kinase panel toward the PI3Ks and mTOR is somewhat surprising. Further studies are needed to understand the structural basis for the kinase selectivity of the inhibitors.

Thiol-reactive small molecules are considered pan-assay interference compounds (PAINS) and have historically been avoided in drug discovery efforts due to concerns that they possess significant off-target liabilities.⁴²⁻⁴⁴ Nevertheless, as many naturally-occurring electrophilic compounds appear to have a defined target or function in nature, the “reactivity-guided” isolation of thiol-reactive natural products may emerge as a powerful method to uncover medicinally-relevant covalent inhibitors hidden in extracts.^{22,45}

Experimental Section

Chemistry. Reactions and compounds were analyzed with an analytical 1100 Series Agilent Technologies HPLC system coupled to an ELSD and UV/vis detector (210, 254, and 360 nm) using a Phenomenex Luna reversed-phase C18(2) column (100 × 4.6 mm, 5 μm, 100 Å) with a 10 min solvent gradient from 10% to 100% containing 0.1% formic acid and a flow rate of 1.0 mL min⁻¹ or a 20 min solvent gradient from 10% to 100% containing 0.1% formic acid and a flow rate of 0.7 mL min⁻¹. Using the same column and solvent gradients, liquid chromatography/high-resolution mass spectrometry was performed on an analytical Agilent 1260 Infinity Series LC system coupled to a 6530 Series Q-TOF mass spectrometer. Thiol **1** was synthesized as described previously.²² Neolymphostin A (**4**) was purified by normal-phase HPLC (4% CH₃OH in DCM, silica(2) Phenomenex Luna, 100 x 10 mm, 5 μ, 3 mL min⁻¹) and determined to have ≥95% purity by analytical HPLC. All other reagents and solvents were purchased commercially and were used without further purification. ¹H NMR and 2D NMR spectra were recorded at 500 MHz in DMSO-*d*₆ (residual solvent referenced to 2.50 ppm) on a Jeol 500 MHz NMR spectrometer. IR spectra were recorded on a Nicolet 100 FT-IR. Optical rotations were recorded on a Jasco P-2000 polarimeter.

Generation of extract. *Salinispora arenicola* strain CNY-486 was cultivated in three 2.8 L Fernbach flasks containing 1 L of a seawater-based A1 medium (10 g L⁻¹ starch, 4 g L⁻¹ yeast extract, 2 g L⁻¹ peptone, 75% seawater, 25% deionized (DI) water) at 230 rpm and 27 °C under artificial light. After 6 days of cultivation, Amberlite XAD-16N and XAD-7HP resins (~10 g L⁻¹ each) were added and left shaking for 1 day. The resin was filtered through cheesecloth, washed with deionized water, and extracted with ethyl acetate mixed with sodium sulfate. The ethyl acetate was removed under reduced pressure to give 246 mg of a dark red solid.

Labeling experiment. CNY-486 extract (124 mg) was treated with probe **1** (30.0 mg, 0.0943 mmol) and triethylamine (15.0 μL, 0.108 mmol) in dry and sparged DMF (2.0 mL) in a vial purged with

N₂. The reaction vessel was kept at room temperature and under an inert atmosphere using an N₂ gas source and bubbler. Analysis of the reaction mixture after 3 h by LC-MS showed formation of two pairs of isomeric products with brominated pseudomolecular ions m/z (M+H)⁺ 642/644 (1:1) and m/z (M+H)⁺ 656/658 (1:1). The solvent was removed under vacuum, and the products were purified by reversed-phased HPLC [51% CH₃CN in water, C18(2) Phenomenex Luna, 250 x 10 mm, 5 μ, 3 mL min⁻¹] to yield 0.8 mg of **3a** (t_R = 22 min), 0.8 mg of **3b** (t_R = 24 min), 0.7 mg of **2a** (t_R = 30 min) and 0.7 mg of **2b** (t_R = 32 min) as red solids. **2a**: UV/vis (CH₃CN/water/0.1% formic acid) λ_{max} 235, 330, 465 nm; ¹H NMR see Tables S1; HRESI-Q-TOF-MS m/z (M+H)⁺ 656.1178 calcd for C₂₉H₃₁BrN₅O₆S, found 656.1166 (Δ1.8 ppm). **2b**: UV/vis (CH₃CN/water/0.1% formic acid) λ_{max} 235, 330, 465 nm; ¹H NMR and 2D NMR see Table S2; HRESI-Q-TOF-MS m/z (M+H)⁺ 656.1178 calcd for C₂₉H₃₁BrN₅O₆S, found 656.1158 (Δ3.0 ppm). **3a**: UV/vis (CH₃CN/water/0.1% formic acid) λ_{max} 235, 330, 465 nm; ¹H NMR see Table S3; HRESI-Q-TOF-MS m/z (M+H)⁺ 642.1022 calcd for C₂₈H₂₉BrN₅O₆S, found 642.1004 (Δ2.8 ppm). **3b**: UV/vis (CH₃CN/water/0.1% formic acid) λ_{max} 235, 330, 465 nm; ¹H NMR and 2D NMR see Table S4; HRESI-Q-TOF-MS m/z (M+H)⁺ 642.1022 calcd for C₂₈H₂₉BrN₅O₆S, found 642.1005 (Δ2.6 ppm).

Model electrophile kinetics experiments with glutathione. The method reported by Flanagan, et al, was slightly adapted.³⁰ A 200 mM stock solution of **6** was prepared in DMSO-*d*₆. A 2.0 mM aqueous electrophile solution was then prepared by diluting an aliquot of the stock solution 1:100 in phosphate-buffered D₂O (100 mM, pH 7.4). A 20 mM glutathione solution in phosphate-buffered D₂O was freshly prepared before the kinetics experiments. In a vial was added 100 μL of 20 mM glutathione solution, followed by 100 μL of 2.0 mM electrophile solution (final concentration 10 mM for glutathione and 1.0 mM for the electrophile). The reaction mixture was briefly swirled in the vial to mix and then quickly loaded into a 3 mm O.D. NMR tube, which was then capped. The tube was loaded into a Jeol 500 MHz NMR spectrometer at 25 °C or 37 °C. The sample was locked and shimmed, and then ¹H NMR spectra were acquired every 5 min for 60 min. Each spectrum consisted of 40 transients with an acquisition time

of 2.18383 s and a relaxation delay of 5 s. The resulting array of spectra were processed in MestReNova, and the residual DMSO solvent peak was referenced to 2.50 ppm. Non-overlapping pairs of peaks corresponding to the electrophile and adduct, or those pairs that overlapped perfectly, were identified and integrated. A plot was then constructed of $\ln([SM])$ versus time for each appropriate signal, and a linear regression line was applied. For a given electrophile at a given temperature, the pseudo-first order half-life was calculated by averaging the linear regression slopes derived from each appropriate signal.

MS identification of covalent Lys802 modification. PI3K α preincubated with 4 μ M neolymphostin was brought to a final concentration of 0.8 μ M in a total of 50 μ L non-deuterated buffer (10 mM HEPES pH 7.5, 100 mM NaCl in water). After adding 20 μ L of ice-cold quench buffer (2 M guanidine-HCl, 3% formic acid), the samples were immediately frozen in liquid nitrogen and stored at -80°C . Peptides were digested and separated in a similar manner to the HDX-MS method described below. Peptide identification was carried out in PEAKS 7. The search was performed with a mass error threshold of 5 ppm for precursors and 0.1 Da for fragments. Variable modification of Lys with a mass of 306.1117 was included. The false discovery rate for peptides was set at 0.1%.

Timecourse for covalent modification of PI3K. The timecourse of covalent modification was initiated by mixing 4 μ M neolymphostin A with 2 μ M PI3K α (p110 α /p85 α) in a buffer containing 10 mM HEPES pH 7.5, 100 mM NaCl, and 5% DMSO. Different timepoints were taken by unfolding the protein in an ice-cold quench buffer (2 M guanidine HCl, 3% formic acid), with samples immediately frozen in liquid nitrogen and stored at -80°C . Samples were injected and analysed on the mass spectrometer as described in the HDX-MS section.

Docking experiments. All steps for the docking experiments in this study were performed using the Schrödinger 2017-3 suite. PI3K α (pdbID:5SW8, chain A) and mTOR (pdbID:4JT5, chain A) were first processed using Protein Preparation Wizard, which added missing hydrogen atoms, assigned protonation states of residues based on PropKA calculations, optimized H-bonds, and deleted all water

molecules farther than 5 Å from protein atoms. The receptor grid was generated for a 30 Å x 30 Å x 30 Å cubic box centered at hinge residues Val850 and Val851 for PI3K α and at hinge residues Trp2239 and Val2240 for mTOR. Neolymphostin was built in Maestro (Small-Molecule Drug Discovery Suite 2017-3: Maestro, Schrödinger, LLC, New York, NY, 2017) and docked into the active site of PI3K α and mTOR using Glide (Small-Molecule Drug Discovery Suite 2017-3: Maestro, Schrödinger, LLC, New York, NY, 2017) with standard precision docking and default parameters. For induced fit docking experiments, search space was again centered at the hinge residues Val850 and Val851 for PI3K α and at hinge residues Trp2239 and Val2240 for mTOR. The side chains of Lys802 of PI3K α and Lys2187 of mTOR were trimmed before neolymphostin docking and restored and optimized together with all residue side chains within 5 Å of the ligand by Schrödinger's Prime module after docking (Small-Molecule Drug Discovery Suite 2017-3: Prime, Schrödinger, LLC, New York, NY, 2017).

Hydrogen deuterium exchange mass spectrometry (HDX-MS) experiment. HDX experiments were performed as described in Vadas, et al.³³ The experiments were conducted in 50 μ L reaction mixtures with a final PI3K α (p110 α /p85 α) concentration of 0.4 μ M. Prior to the experiment, PI3K α was incubated at 2 μ M with neolymphostin (4 μ M final concentration) or in DMSO for 30 minutes. The final DMSO concentration was 5%. Two conditions were tested: (i) PI3K α alone and (ii) PI3K α with neolymphostin. Exchange was carried out for two time-points (3 s on ice and 300 s at 20 °C) and terminated by the addition of 20 μ L ice-cold quench buffer (2 M guanidine-HCl, 3% formic acid). Samples were immediately frozen in liquid nitrogen and stored at -80 °C. Protein samples were rapidly thawed and injected onto a UHPLC system at 2 °C. Protein was run over two immobilized pepsin columns (Applied Biosystems; porosyme, 2-3131-00) at 10 °C and 2 °C at 200 μ L min⁻¹ for 3 minutes, and peptides were collected onto a VanGuard precolumn trap (Waters). The trap was subsequently eluted in line with an Acquity 1.7 μ m particle, 100 \times 1 mm² C18 UPLC column (Waters), using a gradient of 5-36% B (buffer A 0.1% formic acid, buffer B 100% acetonitrile) over 16 minutes. Mass spectrometry experiments were

performed on an Impact II TOF-MS (Bruker) acquiring over a mass range from 150 to 2200 m/z using an electrospray ionization source operated at a temperature of 200 °C and a spray voltage of 4.5 kV.

Peptides were identified using data-dependent acquisition methods following tandem MS/MS experiments (0.5 s precursor scan from 150-2000 m/z ; twelve 0.25 s fragment scans from 150-2000 m/z). MS/MS datasets were analysed using PEAKS 7 (PEAKS STUDIO), and a false discovery rate was set at 1% using a database of purified proteins and known contaminants.

HD-Examiner Software (Sierra Analytics) was used to automatically calculate the level of deuterium incorporation into each peptide. All peptides were manually inspected for correct charge state and presence of overlapping peptides. Deuteration levels were calculated using the centroid of the experimental isotope clusters. Results for these proteins are presented as relative levels of deuterium incorporation and the only control for back exchange was the level of deuterium present in the buffer (76.92%). Changes in any peptide at any time point greater than both 5% and 0.4 Da between conditions with a paired t-test value of $p < 0.01$ was considered significant.

Model electrophile kinetics experiments with N^α -acetyl lysine. Kinetics experiments on model electrophile **6** with and without N^α -acetyl lysine were conducted in an identical manner to those with glutathione (see above), with the exception that phosphate-buffered D_2O was prepared at pH 8.1 (instead of 7.4).

Synthesis and characterization of enamine **8.** A 20 mM solution of **6** was prepared by diluting an aliquot of electrophile stock (200 mM, $DMSO-d_6$) 1:10 in phosphate-buffered D_2O (100 mM, pH 8.1). A 20 mM solution of N^α -acetyl lysine was freshly prepared in phosphate-buffered D_2O (100 mM, pH 8.1). To a vial was added 100 μL of the N^α -acetyl lysine solution, followed by 100 μL of the electrophile solution. The vial was capped, swirled to mix, and allowed to stand overnight. The reaction mixture was then loaded into a 3 mm O.D. NMR tube and characterized by 1H , COSY, HSQC, and HMBC NMR experiments on a Jeol 500 MHz spectrometer.

Kinase selectivity panel. The affinity of neolymphostin for 97 kinases was assessed using the DiscoverX KINOMEScan profiling service. A detailed description of this assay can be found in Fabian, et al., and in the Supplementary Information.³⁰

PI3K inhibition in live cells. After serum starvation overnight, HeLa cells were pretreated with the indicated concentration of wortmannin, neolymphostin A or neolymphostinol for 10 min followed by stimulation with 10ng mL⁻¹ EGF for 5 min. Cells were lysed by the addition of boiling SDS sample buffer, collected into microfuge tubes, and quantified by Pierce 660 nm Protein Assay. Equal protein lysates were boiled at 100° for 5 min, resolved by SDS-PAGE, transferred to PVDF for Western blotting, and detected by enhanced chemiluminescence. AKT, pAKT (Ser473), pAKT (Thr308), ERK1/2, pERK1/2 (Thr202/Tyr204), EGFR, pEGFR (Tyr1068) and GAPDH antibodies were purchased from Cell Signaling Technologies.

NCI 60-cell line assay. 5-Point dose-response curves were generated for leukemia, non-small cell lung cancer, colon cancer, CNS cancer, melanoma, ovarian cancer, renal cancer, prostate cancer, and breast cancer cell lines. 50% growth inhibition (GI₅₀), total growth inhibition (TGI), and 50% lethal concentration (LC₅₀) values were then calculated. A description of the NCI 60-cell line assay can be found in Shoemaker.⁴⁶

Ancillary Information

Supporting Information The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.8b00975.

Experimental procedures, compound characterization data (NMR, MS), curves for dissociation constant measurements, docking images, kinetics data, HDX-MS data, cytotoxicity data (PDF)

Molecular formula strings (CSV)

Corresponding Author

*E-mail: chughes@ucsd.edu

Author Contributions

G.C. conducted the labeling experiment on strain CNY-486. G.S. performed the glutathione and *N*-acetyl lysine kinetics experiments. O.D. conducted the docking experiments. M.R. conducted the HDX-MS experiment. D.H. conducted the MS experiments showing covalent labeling. R.M. carried out the timecourse study. S.M. showed inhibition in live cells. A.L. initially cultivated *Salinispora* CNY-486. S.F., J.B., R.A., and C.H. designed the experiments.

Notes

R.A. is a co-founder of, has equity interest in, and is on the scientific advisory board of Actavalon Inc.

Acknowledgments

This work was supported by seed funding from the Scripps Institution of Oceanography (E.W. Scripps Associates), a UCSD Academic Senate Research Award (RP42S-HUGHES), and an American Cancer Society Institutional Research Grant (ACS-IRG #15-172-45) to C.H. It was also funded in part by the Director's New Innovator Award Program NIH DP2 OD007237 to R.A. Funding and support from the National Biomedical Computation Resource (NBCR) is provided through NIH P41 GM103426. G.C.F. is a Howard Hughes Medical Institute Gilliam Fellow. We also thank the NIH for an HRLC-MS instrument (S10 OD0106400).

Abbreviations

PI3K, phosphoinositide 3-kinase; mTOR, mammalian target of rapamycin; AKT, protein kinase B; Lck, lymphocyte-specific tyrosine kinase; K_d , dissociation constant; PIKK, phosphoinositide 3-kinase-

related kinase; HDX-MS, hydrogen deuterium exchange mass spectrometry; PLK4, Polo-like kinase 4; AURK, aurora kinase; JNK, c-Jun N-terminal kinase; ERK, extracellular signal-regulated kinase; ELSD, evaporative light-scattering detector; UV/vis, ultraviolet/visible light; HRESI-Q-TOF-MS, high-resolution electrospray ionization quadrupole time-of-flight mass spectrometry; PVDF, polyvinylidene fluoride; GADPH, glyceraldehyde 3-phosphate dehydrogenase.

Chapter 3 is a reprint, in full, of published work. Gabriel Castro-Falcón, Grant S. Seiler, Özlem Demir, Manoj K. Rathinaswamy, David Hamelin, Reece M. Hoffmann, Stefanie L. Makowski, Anne-Catrin Letzel, Seth J. Field, John E. Burke, Rommie E. Amaro, and Chambers C. Hughes. “Neolymphostin A Is a Covalent Phosphoinositide 3-Kinase (PI3K)/Mammalian Target of Rapamycin (mTOR) Dual Inhibitor That Employs an Unusual Electrophilic Vinylogous Ester.” *Journal of Medicinal Chemistry* **2018**, *61*, 10463-10472. The dissertation author was a co-primary investigator and author of this paper.

References

- (1) Liu, J.; Hu, Y.; Waller, D. L.; Wang, J.; Liu, Q. Natural products as kinase inhibitors. *Nat. Prod. Rep.* **2012**, *29*, 392–403.
- (2) Omura, S.; Iwai, Y.; Hirano, A.; Nakagawa, A.; Awaya, J.; Tsuchya, H.; Takahashi, Y.; Masuma, R. A new alkaloid AM-2282 of *Streptomyces* origin. taxonomy, fermentation, isolation and preliminary characterization. *J. Antibiot. (Tokyo)*. **1977**, *30*, 275–282.
- (3) Weisberg, E.; Boulton, C.; Kelly, L. M.; Manley, P.; Fabbro, D.; Meyer, T.; Gilliland, D. G.; Griffin, J. D. Inhibition of mutant FLT3 receptors in leukemia cells by the small molecule tyrosine kinase inhibitor PKC412. *Cancer Cell* **2002**, *1*, 433–443.
- (4) Vezina, C.; Kudelski, A.; Sehgal, S. N. Rapamycin (AY-22,989), a new antifungal antibiotic. I. Taxonomy of the producing streptomycete and isolation of the active principle. *J. Antibiot. (Tokyo)*. **1975**, *28*, 721–726.
- (5) FDA-approved protein kinase inhibitors. <http://www.brimr.org/PKI/PKIs.htm> (accessed Aug 23, 2018).
- (6) Fruman, D. A.; Chiu, H.; Hopkins, B. D.; Bagrodia, S.; Cantley, L. C.; Abraham, R. T. The PI3K pathway in human disease. *Cell* **2017**, *170*, 605–635.

- (7) Bergholz, J. S.; Roberts, T. M.; Zhao, J. J. Isoform-selective phosphatidylinositol 3-kinase inhibition in cancer. *J. Clin. Oncol.* **2018**, 1–4.
- (8) Furman, R. R.; Sharman, J. P.; Coutre, S. E.; Cheson, B. D.; Pagel, J. M.; Hillmen, P.; Barrientos, J. C.; Zelenetz, A. D.; Kipps, T. J.; Flinn, I.; Ghia, P.; Eradat, H.; Ervin, T.; Lamanna, N.; Coiffier, B.; Pettitt, A. R.; Ma, S.; Stilgenbauer, S.; Cramer, P.; Aiello, M.; Johnson, D. M.; Miller, L. L.; Li, D.; Jahn, T. M.; Dansey, R. D.; Hallek, M.; O'Brien, S. M. Idelalisib and rituximab in relapsed chronic lymphocytic leukemia. *N. Engl. J. Med.* **2014**, *370*, 997–1007.
- (9) Zhang, J.; Yang, P. L.; Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer* **2009**, *9*, 28–39.
- (10) Fry, D. W.; Bridges, A. J.; Denny, W. A.; Doherty, A.; Greis, K. D.; Hicks, J. L.; Hook, K. E.; Keller, P. R.; Leopold, W. R.; Loo, J. A.; McNamara, D. J.; Nelson, J. M.; Sherwood, V.; Smaill, J. B.; Trumpp-Kallmeyer, S.; Dobrusin, E. M. Specific, irreversible inactivation of the epidermal growth factor receptor and erbB2, by a new class of tyrosine kinase inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 12022–12027.
- (11) Cohen, M. S.; Zhang, C.; Shokat, K. M.; Taunton, J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. *Science* **2005**, *308*, 1318–1321.
- (12) Pan, Z.; Scheerens, H.; Li, S. J.; Schultz, B. E.; Sprengeler, P. A.; Burrill, L. C.; Mendonca, R. V.; Sweeney, M. D.; Scott, K. C. K.; Grothaus, P. G.; Jeffery, D. A.; Spoerke, J. M.; Honigberg, L. A.; Young, P. R.; Dalrymple, S. A.; Palmer, J. T. Discovery of selective irreversible inhibitors for Bruton's tyrosine kinase. *ChemMedChem* **2007**, *2*, 58–61.
- (13) Liu, Q.; Sabnis, Y.; Zhao, Z.; Zhang, T.; Buhrlage, S. J.; Jones, L. H.; Gray, N. S. Developing irreversible inhibitors of the protein kinase cysteinome. *Chem. Biol.* **2013**, *20*, 146–159.
- (14) Kung, A.; Chen, Y.-C.; Schimpl, M.; Ni, F.; Zhu, J.; Turner, M.; Molina, H.; Overman, R.; Zhang, C. Development of specific, irreversible inhibitors for a receptor tyrosine kinase EphB3. *J. Am. Chem. Soc.* **2016**, *138*, 10554–10560.
- (15) Wu, H.; Wang, W.; Liu, F.; Weisberg, E. L.; Tian, B.; Chen, Y.; Li, B.; Wang, A.; Wang, B.; Zhao, Z.; McMillin, D. W.; Hu, C.; Li, H.; Wang, J.; Liang, Y.; Buhrlage, S. J.; Liang, J.; Liu, J.; Yang, G.; Brown, J. R.; Treon, S. P.; Mitsiades, C. S.; Griffin, J. D.; Liu, Q.; Gray, N. S. Discovery of a potent, covalent BTK inhibitor for B-cell lymphoma. *ACS Chem. Biol.* **2014**, *9*, 1086–1091.
- (16) Wymann, M. P.; Bulgarelli-Leva, G.; Zvelebil, M. J.; Pirola, L.; Vanhaesebroeck, B.; Waterfield, M. D.; Panayotou, G. Wortmannin inactivates phosphoinositide 3-kinase by covalent modification of Lys-802, a residue involved in the phosphate transfer reaction. *Mol. Cell. Biol.* **1996**, *16*, 1722–1733.
- (17) Brunn, G. J.; Williams, J.; Sabers, C.; Wiederrecht, G.; Lawrence, J. C.; Abraham, R. T. Direct inhibition of the signaling functions of the mammalian target of rapamycin by the phosphoinositide 3-kinase inhibitors, wortmannin and LY294002. *EMBO J.* **1996**, *15*, 5256–5267.

- (18) Arcaro, A.; Wymann, M. P. Wortmannin is a potent phosphatidylinositol 3-kinase inhibitor: the role of phosphatidylinositol 3,4,5-trisphosphate in neutrophil responses. *Biochem. J.* **1993**, *296*, 297–301.
- (19) Lee, J. H.; Koo, T. H.; Yoon, H.; Jung, H. S.; Jin, H. Z.; Lee, K.; Hong, Y. S.; Lee, J. J. Inhibition of NF- κ B activation through targeting I κ B kinase by celastrol, a quinone methide triterpenoid. *Biochem. Pharmacol.* **2006**, *72*, 1311–1321.
- (20) Kwok, B. H. B.; Koh, B.; Ndubuisi, M. I.; Elofsson, M.; Crews, C. M. The anti-inflammatory natural product parthenolide from the medicinal herb Feverfew directly binds to and inhibits I κ B kinase. *Chem. Biol.* **2001**, *8*, 759–766.
- (21) Schirmer, A.; Kennedy, J.; Murli, S.; Reid, R.; Santi, D. V. Targeted covalent inactivation of protein kinases by resorcylic acid lactone polyketides. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 4234–4239.
- (22) Castro-Falcon, G.; Hahn, D.; Reimer, D.; Hughes, C. C. Thiol probes to detect electrophilic natural products based on their mechanism of action. *ACS Chem. Biol.* **2016**, *11*, 2328–2336.
- (23) Miles, C. O.; Sandvik, M.; Nonga, H. E.; Rundberget, T.; Wilkins, A. L.; Rise, F.; Ballot, A. Thiol derivatization for LC-MS identification of microcystins in complex matrices. *Environmetal Sci. Technonlogy* **2012**, *46*, 8937–8944.
- (24) Cox, C. L.; Tietz, J. I.; Sokolowski, K.; Melby, J. O.; Doroghazi, J. R.; Mitchell, D. A. Nucleophilic 1,4-additions for natural product discovery. *ACS Chem. Biol.* **2014**, *4*, 2014–2022.
- (25) Rudolf, G. C.; Koch, M. F.; Mandl, F. A. M.; Sieber, S. A. Subclass-specific labeling of protein-reactive natural products with customized nucleophilic probes. *Chem. - A Eur. J.* **2015**, *21*, 3701–3707.
- (26) Aotani, Y.; Nagata, H.; Yoshida, M. Lymphostin (LK6-A), a novel immunosuppressant from *Streptomyces* sp. KY11783: Structural elucidation. *J. Antibiot. (Tokyo)*. **1997**, *50*, 543–545.
- (27) Nagata, H.; Ochiai, K.; Aotani, Y.; Ando, K.; Yoshida, M.; Takahashi, I. Taxonomy of the producing organism, fermentation, isolation. *J. Antibiot. (Tokyo)*. **1997**, *50*, 537–542.
- (28) Nagata, H.; Yano, H.; Sasaki, K.; Sato, S.; Nakanishi, S.; Takahashi, I.; Tamaoki, T. Inhibition of lymphocyte kinase Lck and phosphatidylinositol 3-kinase by a novel immunosuppressant, lymphostin. *Biosci. Biotechnol. Biochem.* **2002**, *66*, 501–507.
- (29) Miyanaga, A.; Janso, J. E.; McDonald, L.; He, M.; Liu, H.; Barbieri, L.; Eustáquio, A. S.; Fielding, E. N.; Carter, G. T.; Jensen, P. R.; Feng, X.; Leighton, M.; Koehn, F. E.; Moore, B. S. Discovery and assembly-line biosynthesis of the lymphostin pyrroloquinoline alkaloid family of mTOR inhibitors in *Salinispora* bacteria. *J. Am. Chem. Soc.* **2011**, *133*, 13311–13313.
- (30) Fabian, M. A.; Biggs, W. H.; Treiber, D. K.; Atteridge, C. E.; Azimioara, M. D.; Benedetti, M. G.; Carter, T. A.; Ciceri, P.; Edeen, P. T.; Floyd, M.; Ford, J. M.; Galvin, M.; Gerlach, J. L.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Insko, M. A.; Lai, A. G.; Lélias, J.-M.; Mehta, S. A.; Milanov, Z. V.; Velasco, A. M.; Wodicka, L. M.; Patel, H. K.; Zarrinkar, P. P.; Lockhart, D. J. A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat. Biotechnol.* **2005**, *23*, 329–336.

- (31) Flanagan, M. E.; Abramite, J. A.; Anderson, D. P.; Aulabaugh, A.; Dahal, U. P.; Gilbert, A. M.; Li, C.; Montgomery, J.; Oppenheimer, S. R.; Ryder, T.; Schuff, B. P.; Uccello, D. P.; Walker, G. S.; Wu, Y.; Brown, M. F.; Chen, J. M.; Hayward, M. M.; Noe, M. C.; Obach, R. S.; Philippe, L.; Shanmugasundaram, V.; Shapiro, M. J.; Starr, J.; Stroh, J.; Che, Y. J. Chemical and computational methods for the characterization of covalent reactive groups for the prospective design of irreversible inhibitors. *Med. Chem.* **2014**, *57*, 10072–10079.
- (32) Palkowitz, M. D.; Tan, B.; Hu, H.; Roth, K.; Bauer, R. A. Synthesis of diverse N-acryloyl azetidines and evaluation of their enhanced thiol reactivities. *Org. Lett.* **2017**, *19*, 2270–2273.
- (33) Vadas, O.; Jenkins, M. L.; Dornan, G. L.; Burke, J. E. *Using Hydrogen–Deuterium Exchange Mass Spectrometry to Examine Protein–Membrane Interactions*, 1st ed.; Elsevier Inc.: Cambridge, MA, 2017; pp. 143–172.
- (34) Masson, G. R.; Maslen, S. L.; Williams, R. L. Analysis of phosphoinositide 3-kinase inhibitors by bottom-up electron-transfer dissociation hydrogen/deuterium exchange mass spectrometry. *Biochem. J.* **2017**, *474*, 1867–1877.
- (35) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2008**, *26*, 127–132.
- (36) Donnella, H. J.; Webber, J. T.; Levin, R. S.; Camarda, R.; Momcilovic, O.; Bayani, N.; Shah, K. N.; Korkola, J. E.; Shokat, K. M.; Goga, A.; Gordan, J. D.; Bandyopadhyay, S. Kinome rewiring reveals AURKA limits PI3K-pathway inhibitor efficacy in breast cancer. *Nat. Chem. Biol.* **2018**, *14*, 768–777.
- (37) Paull, K. D.; Shoemaker, R. H.; Hodes, L.; Monks, A.; Scudiero, D. A.; Rubinstein, L.; Plowman, J.; Boyd, M. R. Display and analysis of patterns of differential activity of drugs against human tumor cell lines: development of mean graph and COMPARE algorithm. *J. Natl. Cancer Inst.* **1989**, *81*, 1088–1092.
- (38) Carrera, A. C.; Alexandrov, K.; Roberts, T. M. The conserved lysine of the catalytic domain of protein kinases is actively involved in the phosphotransfer reaction and not required for anchoring ATP. *Biochemistry* **1993**, *90*, 442–446.
- (39) Dalton, S. E.; Dittus, L.; Thomas, D. A.; Convery, A.; Nunes, J.; Bush, J. T.; Evans, J. P.; Werner, T.; Bantsche, M.; Murphy, J. A.; Campos, S. Selectively targeting the kinome-conserved lysine of PI3K δ as a general approach to covalent kinase inhibition. *J. Am. Chem. Soc.* **2018**, *140*, 932–939.
- (40) Pettinger, J.; Jones, K.; Cheeseman, M. D. Lysine-targeting covalent inhibitors. *Angew. Chemie - Int. Ed.* **2017**, *56*, 15200–15209.
- (41) Jeffrey, P. D.; Russo, A. A.; Polyak, K.; Gibbs, E.; Hurwitz, J.; Massagué, J.; Pavletich, N. P. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature*. 1995, pp 313–320.

- (42) Huth, J. R.; Mendoza, R.; Olejniczak, E. T.; Johnson, R. W.; Cothron, D. A.; Liu, Y.; Lerner, C. G.; Chen, J.; Hajduk, P. J. ALARM NMR: A rapid and robust experimental method to detect reactive false positives in biochemical screens. *J. Am. Chem. Soc.* **2005**, *127*, 217–224.
- (43) Baell, J. B.; Nissink, J. W. M. Seven year itch: Pan-assay interference compounds (PAINS) in 2017—utility and limitations. *ACS Chem. Biol.* **2018**, *13*, 36–44.
- (44) Dahlin, J. L.; Nissink, J. W. M.; Strasser, J. M.; Francis, S.; Higgins, L.; Zhou, H.; Zhang, Z.; Walters, M. A. PAINS in the assay: Chemical mechanisms of assay interference and promiscuous enzymatic inhibition observed during a sulfhydryl-scavenging HTS. *J. Med. Chem.* **2015**, *58*, 2091–2113.
- (45) Gersch, M.; Kreuzer, J.; Sieber, S. A. Electrophilic natural products and their biological targets. *Nat. Prod. Rep.* **2012**, *29*, 659–682.
- (46) Shoemaker, R. H.; Shoemaker, R. H. The NCI60 human tumour cell line anticancer drug screen. *Nat. Rev. Cancer* **2006**, *6*, 813–823.

Supporting Information

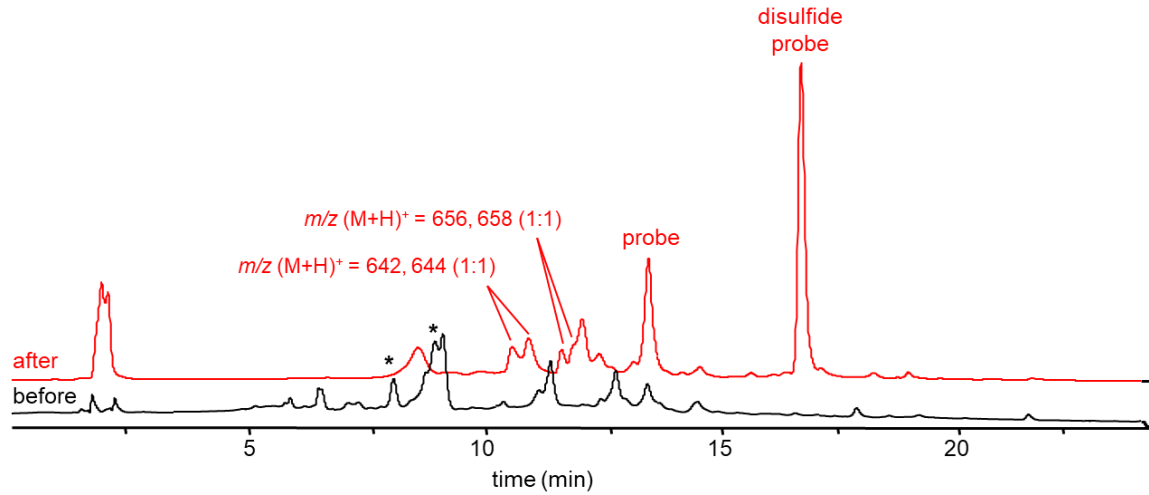
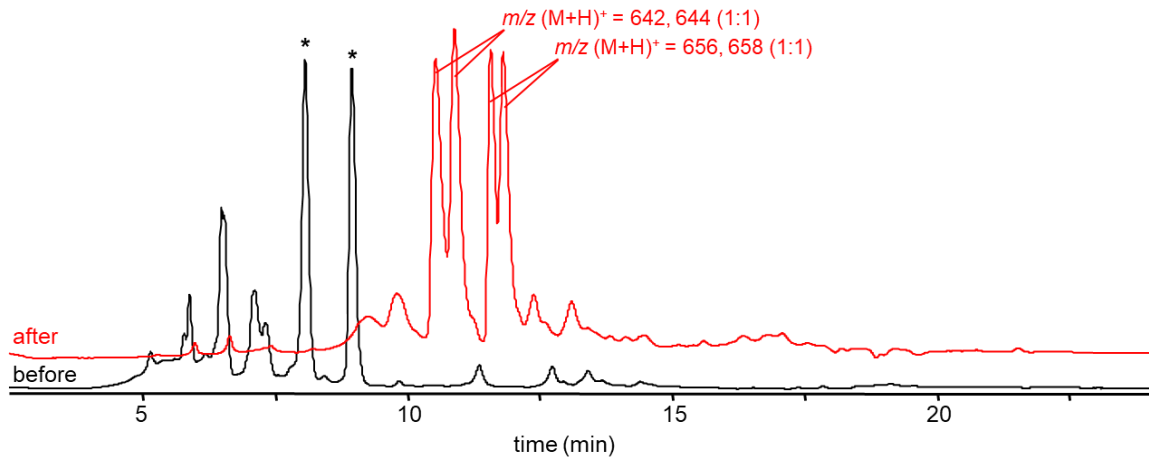


Figure S1. HPLC chromatogram (450 nm) before and after labeling reaction
C18(2) Phenomenex Luna, 100 x 4.6 mm, 5 μ , 100 Å
10-100% CH₃CN in water (0.1% FA) over 20 min, 0.7 mL min⁻¹



Asterisks (*) signal neolympostin A (8.96 min) and neolympostin B (8.08 min)

Figure S2. HPLC chromatogram (450 nm) before and after labeling reaction
C18(2) Phenomenex Luna, 100 x 4.6 mm, 5 μ , 100 Å
10-100% CH₃CN in water (0.1% FA) over 20 min, 0.7 mL min⁻¹

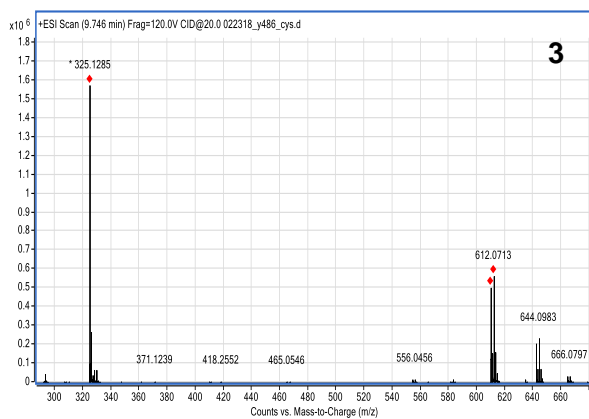
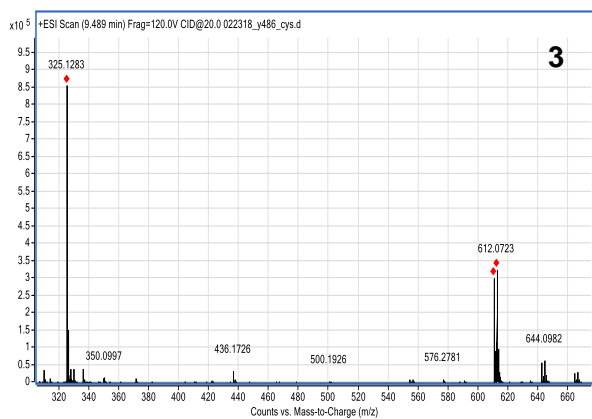
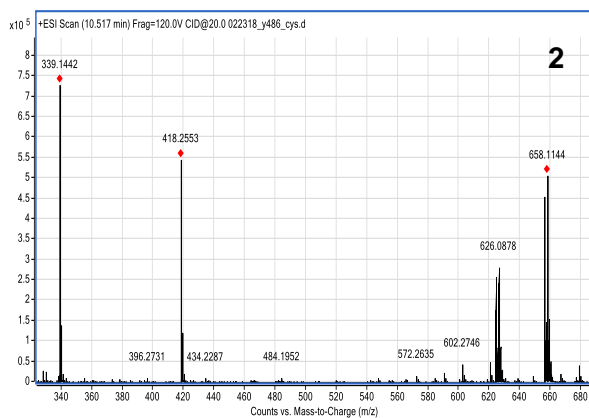
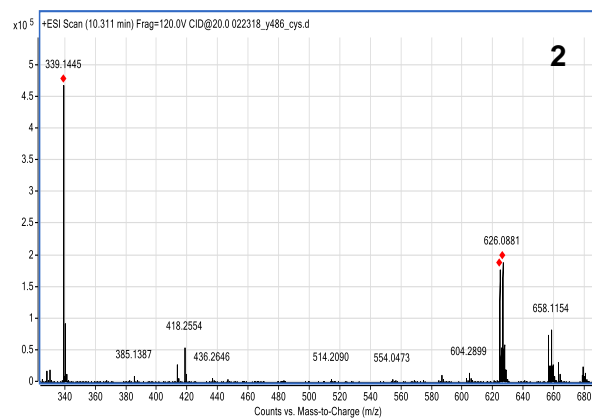
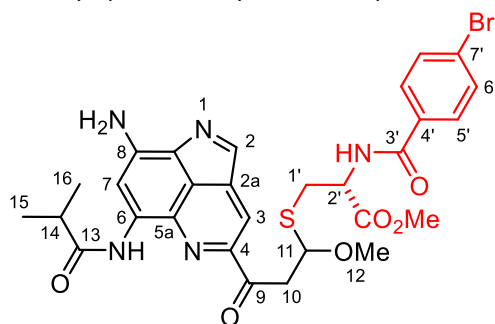


Figure S3. MS spectra of neolymphostin A/B–cysteine thiol probe adducts (**2a/2b/3a/3b**)

Table S1. NMR spectral data for neolymphostin A–cysteine thiol probe adduct (**2a**) in DMSO-*d*₆ at 500 MHz



diastereomer #1

Position # ^a	δ_{H} (mult., (J(Hz)))
1	----
2	8.33 (s)
2a	----
3	8.47 (s)
4	----
5a	----
6	----
6-NH	9.82 (s)
7	8.04 (s)
8	----
8a	----
8b	----
9	----
10	a 3.78 (dd, 16.2, 5.6) b 4.12 (dd, 16.2, 8.0)
11	5.17 (dd, 7.8, 5.7)
12	3.27 (s)
13	----
14	2.92 (sept., 6.9)
15	1.14 (d, 6.8)
16	1.13 (d, 6.8)
1'	a 3.02 (dd, 13.7, 9.4) b 3.20 (dd, 13.7, 5.1)
2'	4.65 (m)
2'-NH	8.93 (d, 7.8)
3'	----
4'	----
5'	7.70 (d, 8.5)
6'	7.59 (d, 8.5)
7'	----
8'	----
9'	3.61 (s)

^a Position numbering based on Aotani, et al. *J. Antibiot.* **1997**, *50*, 543–545.

¹H NMR (500 MHz, DMSO-d₆) of neolymphostin A–cysteine thiol probe adduct (**2a**)

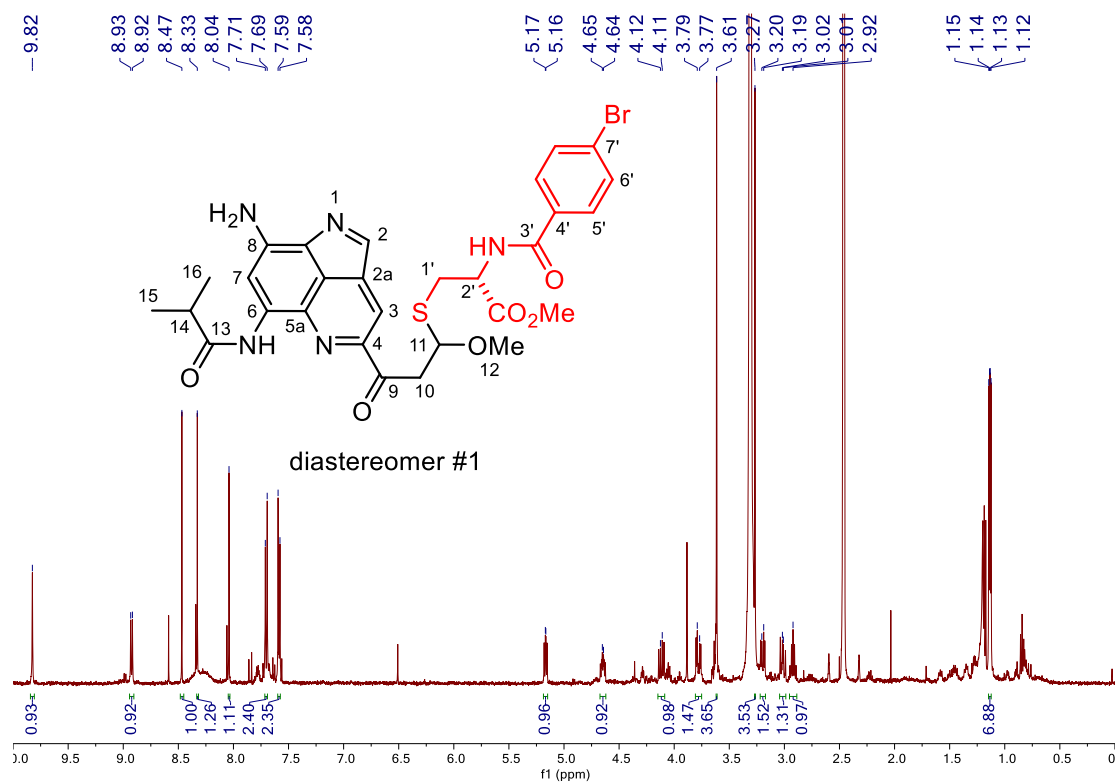
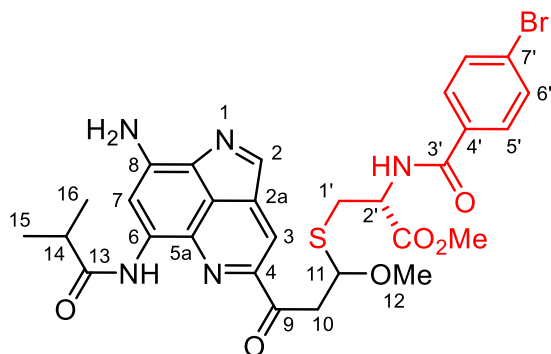


Table S2. NMR spectral data for neolymphostin A–cysteine thiol probe adduct (**2b**) in DMSO-*d*₆ at 500 MHz



diastereomer #2

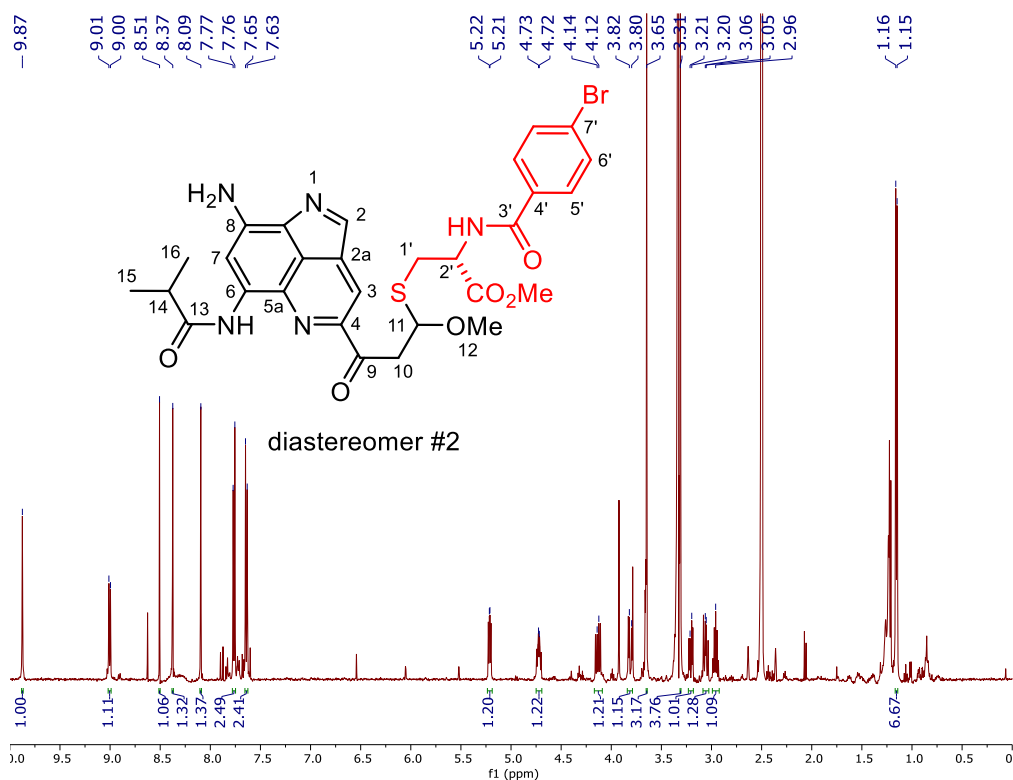
Position # ^a	δ_{H} (mult., <i>J</i> (Hz))	$\delta_{\text{C}}^{\text{b}}$	COSY	HMBC
1	----	----	----	----
2	8.37 (s)	145.8	----	2a, 8b
2a	----	136.6	----	----
3	8.51 (s)	117.6	----	4, 8b
4	----	146.2	----	----
5a	----	133.7	----	----
6	----	139.5	----	----
6-NH	9.87 (s)	----	----	5a, 7, 13
7	8.09 (s)	111.0	----	5a, 6, 8a
8	----	n.d. ^c	----	----
8a	----	124.6	----	----
8b	----	124.8	----	----
9	----	197.0	---	----
10	a 3.81 (dd, 16.5, 5.1) b 4.13 (dd, 16.5, 8.1)	44.4	10b, 11 10a, 11	9, 11 9, 11
11	5.22 (dd, 8.0, 5.1)	83.4	10a, 10b	12
12	3.31 (s)	54.8	----	11
13	----	176.1	----	----
14	2.96 (sept., 6.9)	35.1	15, 16	13, 15, 16
15	1.16 (d, 6.6)	19.3	14	13, 14, 16
16	1.16 (d, 6.6)	19.3	14	13, 14, 15
1'	a 3.06 (dd, 13.7, 9.2) b 3.21 (dd, 13.7, 5.3)	29.5	1'b, 2' 1'a, 2'	11, 8' 11, 8'
2'	4.73 (m)	53.2	1'a, 1'b, 2'-NH	8'
2'-NH	9.01 (d, 7.8)	----	2'	1', 3'
3'	----	165.7	----	----
4'	----	132.6	----	----
5'	7.77 (d, 8.6)	129.4	6'	3', 7'
6'	7.64 (d, 8.6)	131.3	5'	4', 7'
7'	----	125.2	----	----
8'	----	171.1	----	----
9'	3.65 (s)	52.0	---	8'

^a Position numbering based on Aotani, et al. *J. Antibiot.* **1997**, *50*, 543–545.

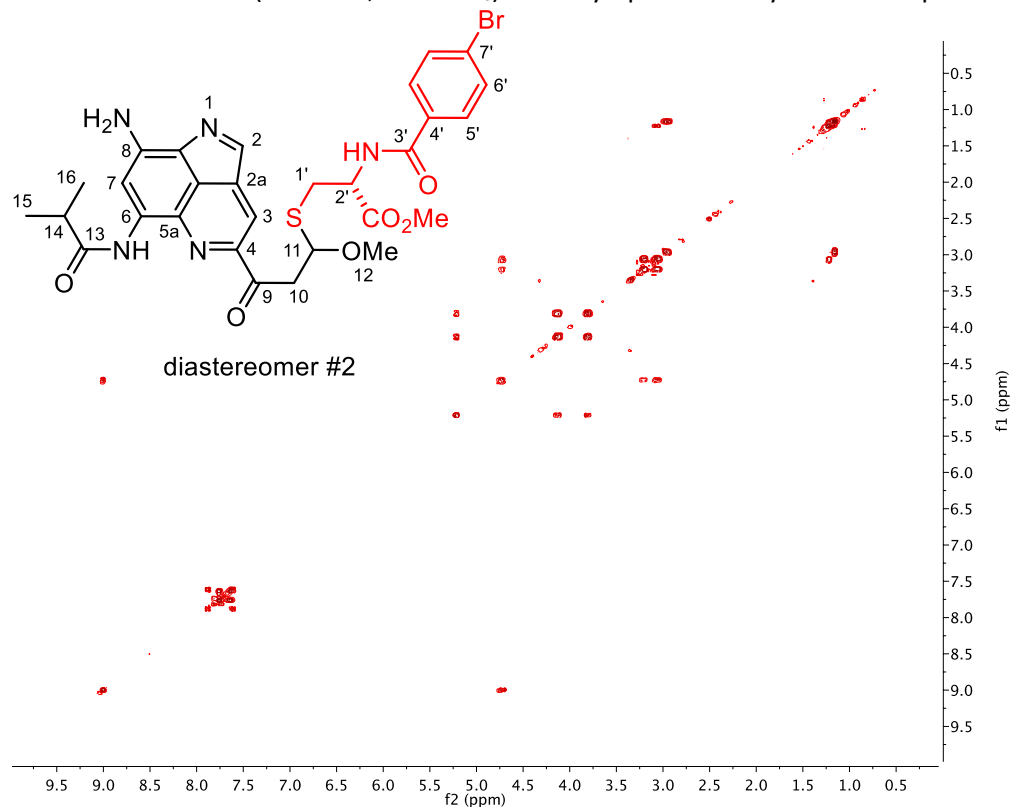
^b Carbon chemical shifts were based on HSQC and HMBC data

^c n.d. = not determined

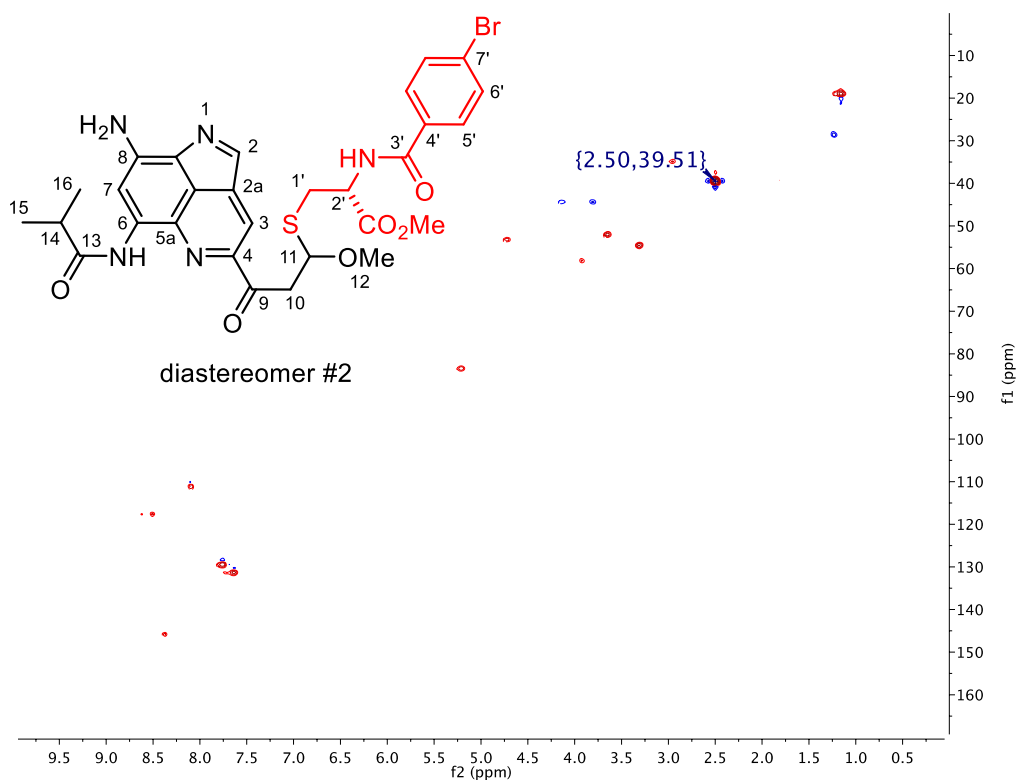
¹H NMR (500 MHz, DMSO-d₆) of neolymphostin A–cysteine thiol probe adduct (**2b**)



COSY (500 MHz, DMSO-d₆) of neolymphostin A–cysteine thiol probe adduct (**2b**)



HSQC (500 MHz, DMSO-*d*₆) of neolymphostin A–cysteine thiol probe adduct (**2b**)



HMBC (500 MHz, DMSO-*d*₆) of neolymphostin A–cysteine thiol probe adduct (**2b**)

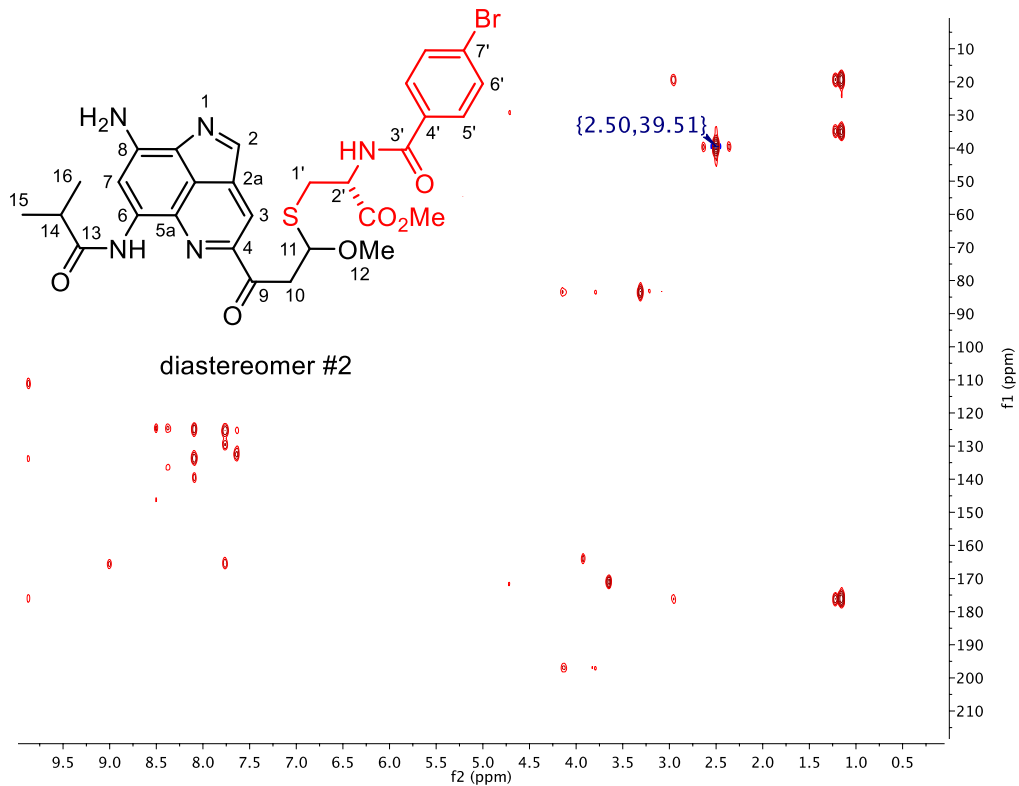
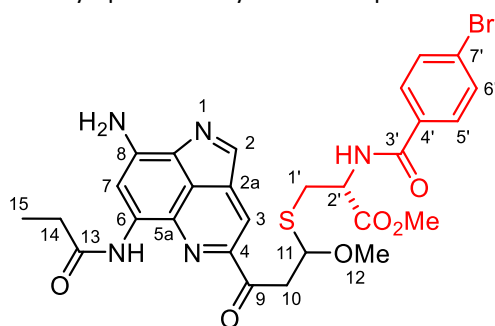


Table S3. NMR spectral data for neolymphostin B–cysteine thiol probe adduct (**3a**) in DMSO-*d*₆ at 500 MHz



diastereomer #1

Position # ^a	δ_{H} (mult., (J(Hz)))
1	----
2	8.36 (s)
2a	----
3	8.51 (s)
4	----
5a	----
6	----
6-NH	9.90 (s)
7	8.12 (s)
8	----
8a	----
8b	----
9	----
10	a 3.83 (dd, 16.2, 5.6) b 4.19 (dd, 16.2, 7.8)
11	5.18 (dd, 7.6, 5.7)
12	3.31 (s)
13	----
14	2.60 (q, 7.5)
15	1.12 (t, 7.4)
1'	a 3.07 (dd, 13.7, 9.3) b 3.27 (dd, 13.7, 5.3)
2'	4.71 (m)
2'-NH	8.96 (d, 7.8)
3'	----
4'	----
5'	7.74 (d, 8.5)
6'	7.61 (d, 8.5)
7'	----
8'	----
9'	3.65 (s)

^a Position numbering based on Aotani, et al. *J. Antibiot.* **1997**, *50*, 543–545.

¹H NMR (500 MHz, DMSO-d₆) of neolymphostin B–cysteine thiol probe adduct (**3a**)

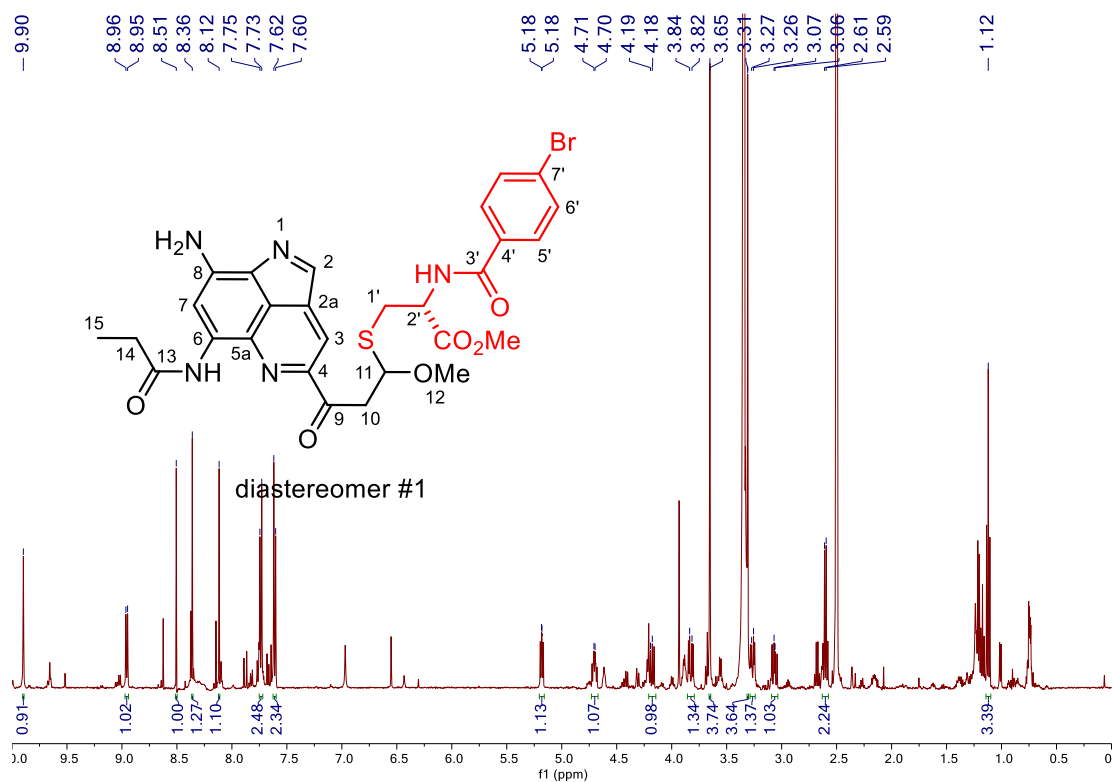
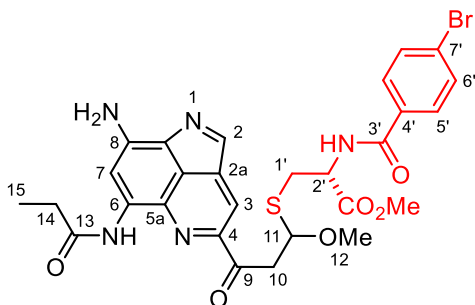


Table S4. NMR spectral data for neolymphostin B–cysteine thiol probe adduct (**3b**) in DMSO-*d*₆ at 500 MHz



diastereomer #2

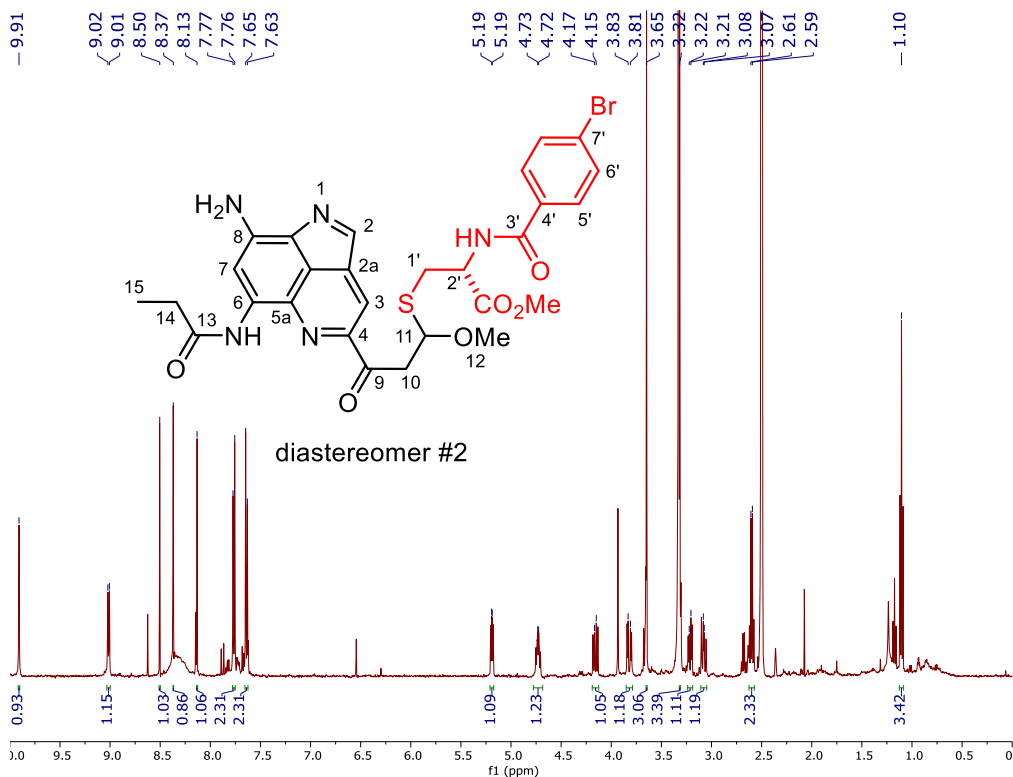
Position # ^a	δ_{H} (mult., (J(Hz)))	δ_{C} ^b	COSY	HMBC
1	----	----	----	----
2	8.37 (s)	145.5	----	2a, 8b
2a	----	136.3	----	----
3	8.50 (s)	117.4	----	4, 8b
4	----	145.4	----	----
5a	----	133.6	----	----
6	----	139.3	----	----
6-NH	9.91 (s)	----	----	5a, 7, 13
7	8.13 (s)	110.4	----	5a, 6, 8a
8	----	n.d. ^c	----	----
8a	----	124.8	----	----
8b	----	124.4	----	----
9	----	198.0	---	----
10	a 3.82 (dd, 16.8, 5.1) b 4.16 (dd, 16.8, 8.2)	44.4	10b, 11 10a, 11	9, 11 9, 11
11	5.19 (dd, 8.0, 5.1)	83.4	10a, 10b	12, 1'
12	3.32 (s)	54.6	----	11
13	----	173.2	----	----
14	2.60 (q, 7.5)	29.5	15	13, 15
15	1.10 (t, 7.5)	9.0	14	13, 14
1'	a 3.08 (dd, 13.9, 9.5) b 3.22 (dd, 13.9, 5.6)	28.9	1'b, 2' 1'a, 2'	11, 2' 11
2'	4.73 (m)	53.2	1'a, 1'b, 2'-NH	8'
2'-NH	9.02 (d, 7.8)	----	2'	3'
3'	----	165.4	----	----
4'	----	132.6	----	----
5'	7.77 (d, 8.4)	129.2	6'	3', 7'
6'	7.64 (d, 8.4)	131.1	5'	4', 7'
7'	----	125.4	----	----
8'	----	171.0	----	----
9'	3.65 (s)	52.0	---	8'

^a Position numbering based on Aotani, et al. *J. Antibiot.* **1997**, *50*, 543–545.

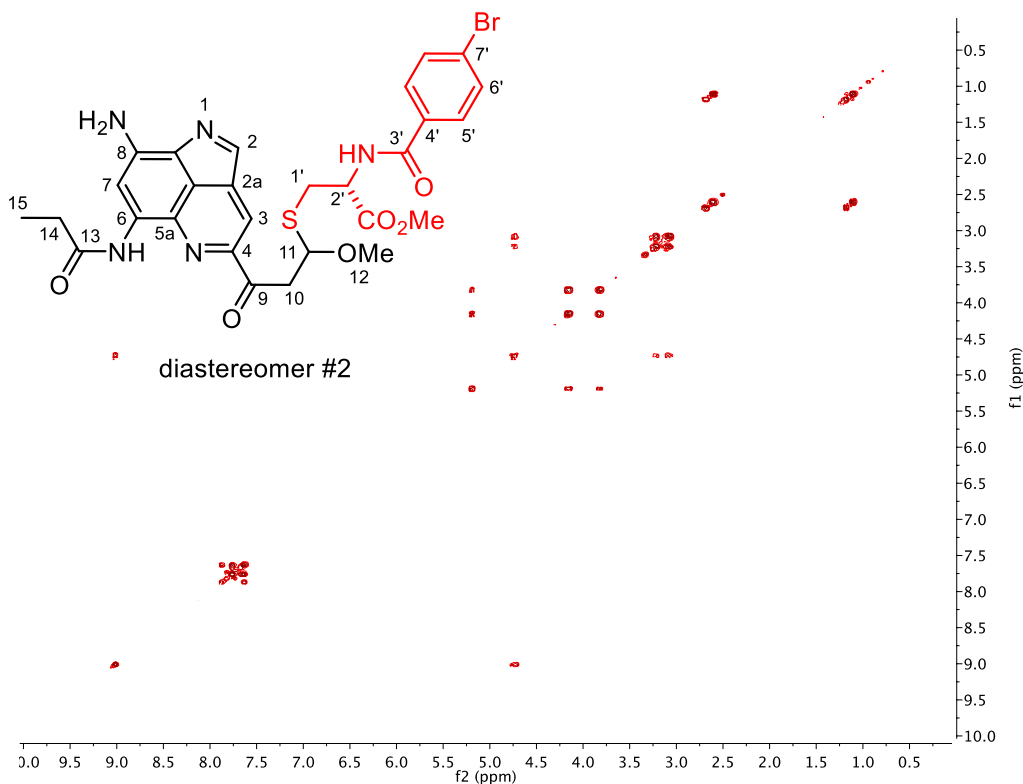
^b Carbon chemical shifts were based on HSQC and HMBC data

^c n.d. = not determined

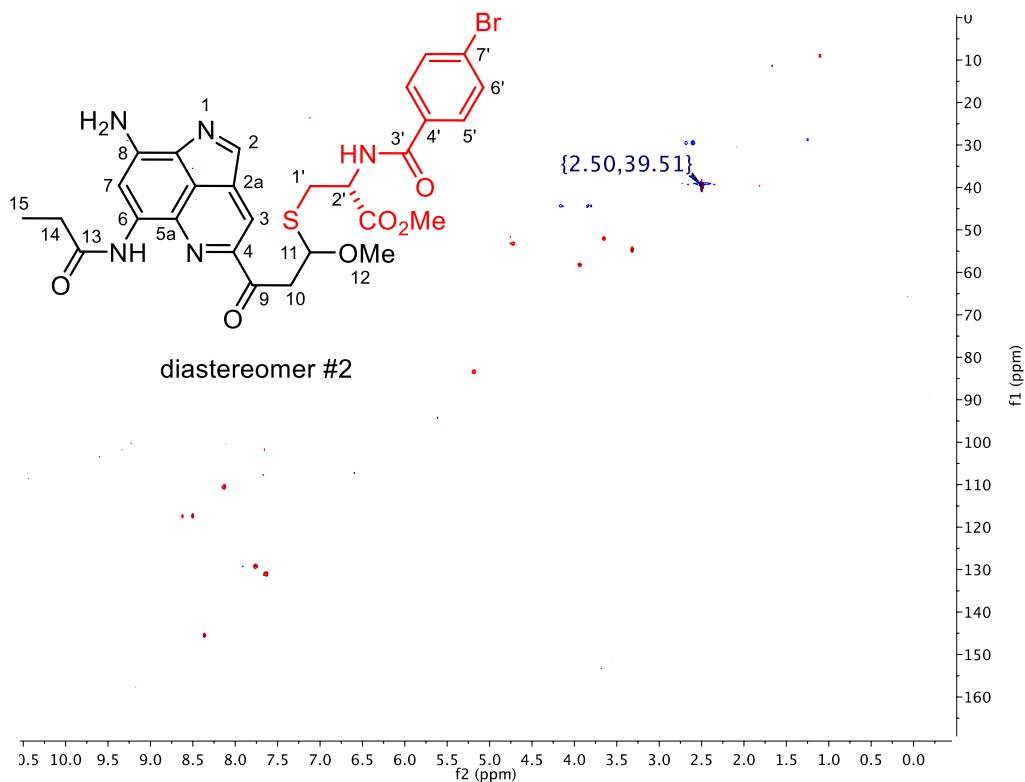
¹H NMR (500 MHz, DMSO-d₆) of neolymphostin B–cysteine thiol probe adduct (**3b**)



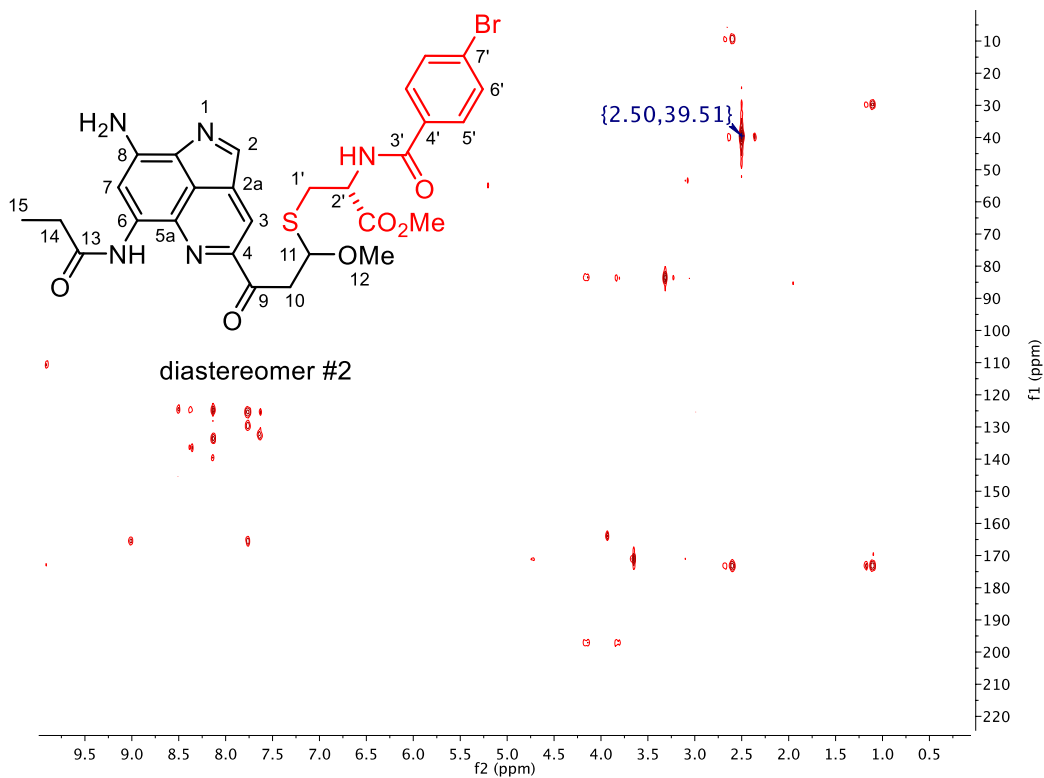
COSY (500 MHz, DMSO-d₆) of neolymphostin B–cysteine thiol probe adduct (**3b**)



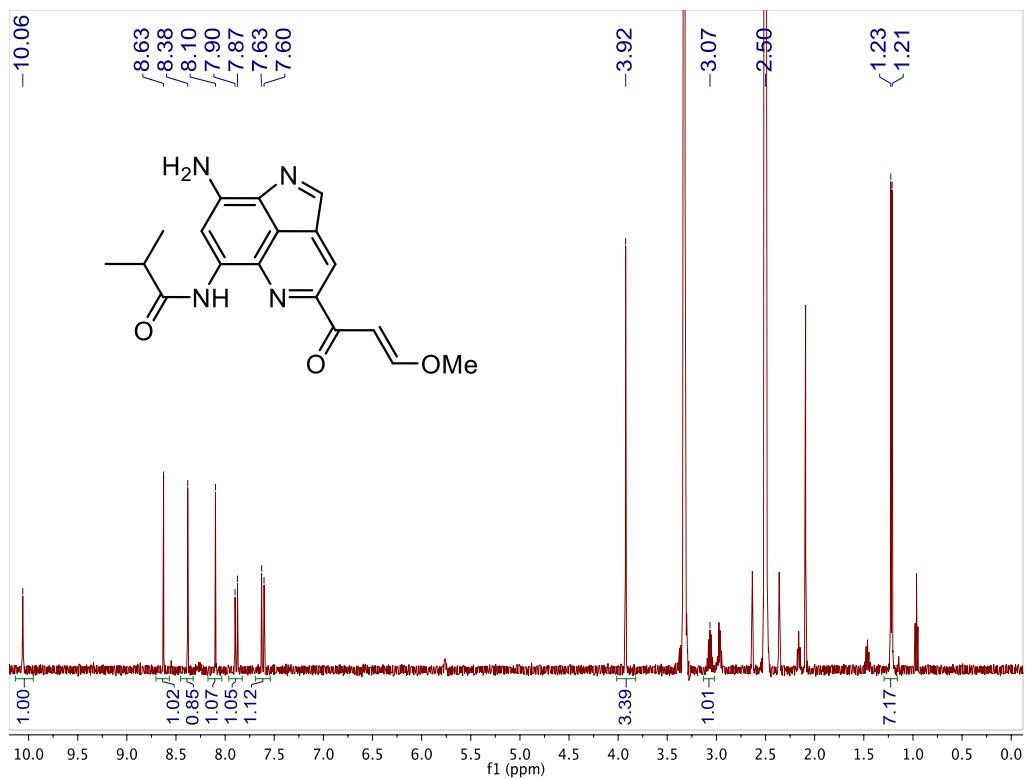
HSQC (500 MHz, DMSO- d_6) of neolymphostin B–cysteine thiol probe adduct (**3b**)



HMBC (500 MHz, DMSO- d_6) of neolymphostin B–cysteine thiol probe adduct (**3b**)

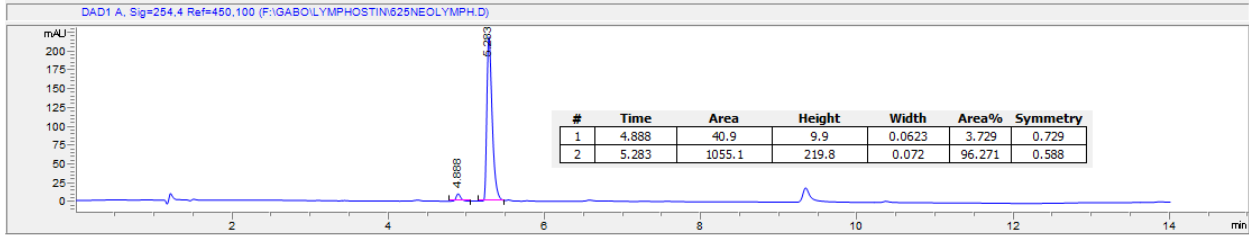
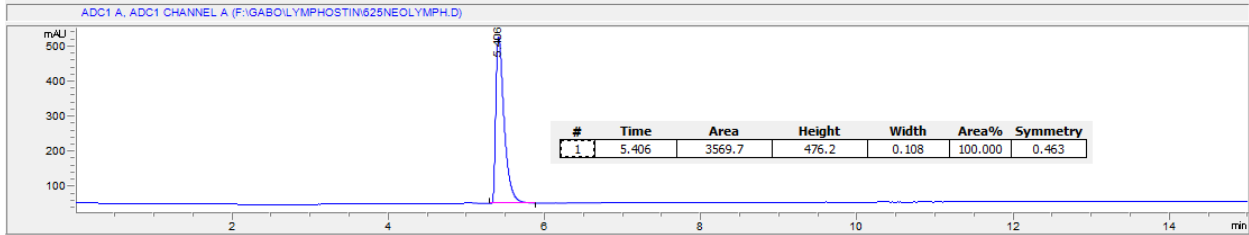


^1H NMR (500 MHz, $\text{DMSO-}d_6$) of neolymphostin A (4)

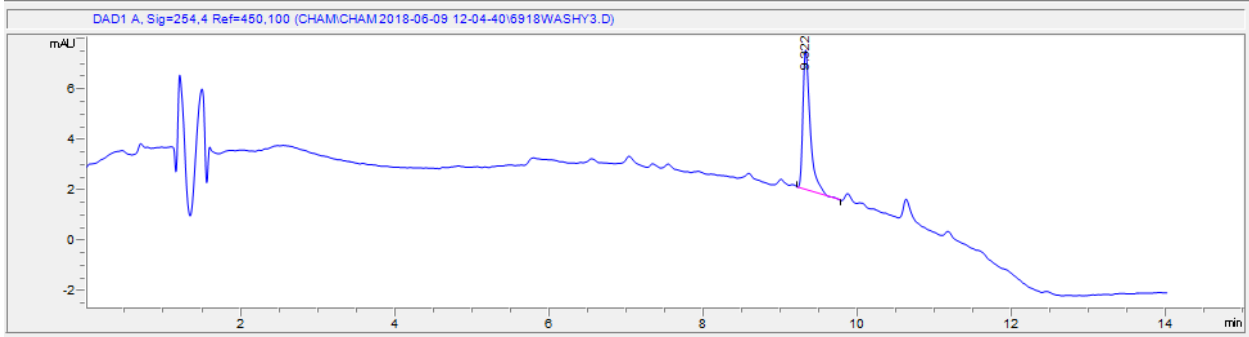


HPLC chromatogram of neolymphostin A (4)

C18(2) Luna Phenomenex, 100 x 4.6 mm, 5 μ , 100 \AA
10-100% CH₃CN in water (0.1% FA) over 10 min, 1.0 mL min⁻¹
Top: ELSD; Bottom: 254 nm



Blank run showing "ghost peak" at 9.3 min (254 nm)

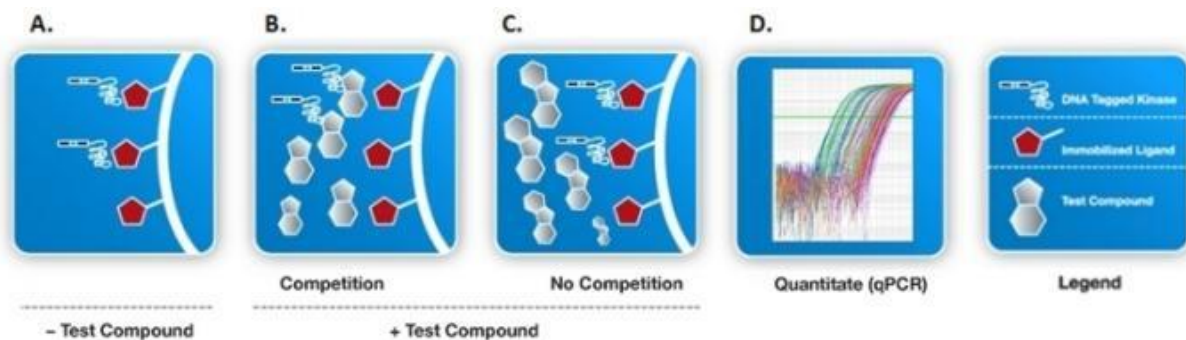


Technology Overview

The KINOMEScan™ screening platform employs a novel and proprietary active site-directed competition binding assay to quantitatively measure interactions between test compounds and more than 450 human kinases and disease relevant mutant variants. This robust and reliable assay technology affords investigators the ability to extensively annotate compounds with accurate, precise and reproducible data. KINOMEScan™ assays do not require ATP and thereby report true thermodynamic interaction affinities, as opposed to IC_{50} values, which can depend on the ATP concentration.

How KINOMEScan™ Works

Compounds that bind the kinase active site and directly (sterically) or indirectly (allosterically) prevent kinase binding to the immobilized ligand, will reduce the amount of kinase captured on the solid support (A & B). Conversely, test molecules that do not bind the kinase have no effect on the amount of kinase captured on the solid support (C). Screening "hits" are identified by measuring the amount of kinase captured in test versus control samples by using a quantitative, precise and ultra-sensitive qPCR method that detects the associated DNA label (D). In a similar manner, dissociation constants (K_d) for test compound-kinase interactions are calculated by measuring the amount of kinase captured on the solid support as a function of the test compound concentration.



Kinase assays

For most assays, kinase-tagged T7 phage strains were grown in parallel in 24-well blocks in an *E. coli* host derived from the BL21 strain. *E. coli* were grown to log-phase and infected with T7 phage from a frozen stock (multiplicity of infection = 0.4) and incubated with shaking at 32°C until lysis (90-150 minutes). The lysates were centrifuged (6,000 x g) and filtered (0.2µm) to remove cell debris. The remaining kinases were produced in HEK-293 cells and subsequently tagged with DNA for qPCR detection. Streptavidin-coated magnetic beads were treated with biotinylated small molecule ligands for 30 minutes at room temperature to generate affinity resins for kinase assays. The liganded beads were blocked with excess biotin and washed with blocking buffer (SeaBlock (Pierce), 1% BSA, 0.05 % Tween 20, 1 mM DTT) to remove unbound ligand and to reduce non-specific phage binding. Binding reactions were assembled by combining kinases, liganded affinity beads, and test compounds in 1x binding buffer (20 % SeaBlock, 0.17x PBS, 0.05 % Tween 20, 6 mM DTT). Test compounds were prepared as 40x stocks in 100% DMSO and directly diluted into the assay. All reactions were performed in polypropylene 384-well plates in a final volume of 0.02 ml. The assay plates were incubated at room temperature with shaking for 1 hour and the affinity beads were washed with wash buffer (1x PBS, 0.05 % Tween 20). The beads were then re-suspended in elution buffer (1x PBS, 0.05 % Tween 20, 0.5 µM non-biotinylated affinity ligand) and incubated at room temperature with shaking for 30 minutes. The kinase concentration in the eluates was measured by qPCR.

Binding Constants (K_ds)

Binding constants (K_ds) were calculated with a standard dose-response curve using the Hill equation:

$$\text{Response} = \text{Background} + \frac{\text{Signal} - \text{Background}}{1 + \left(K_d^{\text{Hill Slope}} / \text{Dose}^{\text{Hill Slope}} \right)}$$

The Hill Slope was set to -1.

Curves were fitted using a non-linear least square fit with the Levenberg-Marquardt algorithm.

Compound Handling

An 11-point 3-fold serial dilution of each test compound was prepared in 100% DMSO at 100x final test concentration and subsequently diluted to 1x in the assay (final DMSO concentration = 1%). Most K_ds were determined using a compound top concentration = 30,000 nM. If the initial K_d determined was < 0.5 nM (the lowest concentration tested), the measurement was repeated with a serial dilution starting at a lower top concentration. A K_d value reported as 40,000 nM indicates that the K_d was determined to be >30,000 nM.

scanKINETIC platform

This protocol offers dissociation kinetics to classify inhibitors as irreversible, reversible-slow dissociation, or reversible-rapid dissociation. Kinetic dissociation constants were measured like described above but equilibration time and dilutions were varied in a set of four different experiments, study arms A-D. In study arm A and C, the compound and kinase were combined and equilibrated for 6 h and 1 h, respectively, before measuring the K_d. In study arm B the compound and kinase were combined and equilibrated for 1 h, diluted 30-fold, and re-equilibrated for 5 h. In study arm D the compound and kinase were pre-diluted, combined, and equilibrated for 5 h

Percent Control (%Ctrl)

Neolymphostin A was screened at 1000nm against a panel of kinases and results for primary screen binding interactions are reported as '% Ctrl', where lower numbers indicate stronger hits.

$$\%Ctrl = \left(\frac{\text{test compound signal} - \text{positive control signal}}{\text{negative control signal} - \text{positive control signal}} \right) \times 100$$

test compound = neolymphostin A, negative control = DMSO (100%Ctrl), positive control = control compound (0%Ctrl)

Selectivity Score (S-scores)

Selectivity Score or S-score is a quantitative measure of compound selectivity. It is calculated by dividing the number of kinases that compounds bind to by the total number of distinct kinases tested, excluding mutant variants. This value can be calculated using %Ctrl as a potency threshold (below) and provides a quantitative method of describing compound selectivity to facilitate comparison of different compounds.

$S(35) = (\text{number of non-mutant kinases with \%Ctrl} < 35) / (\text{number of non-mutant kinases tested})$

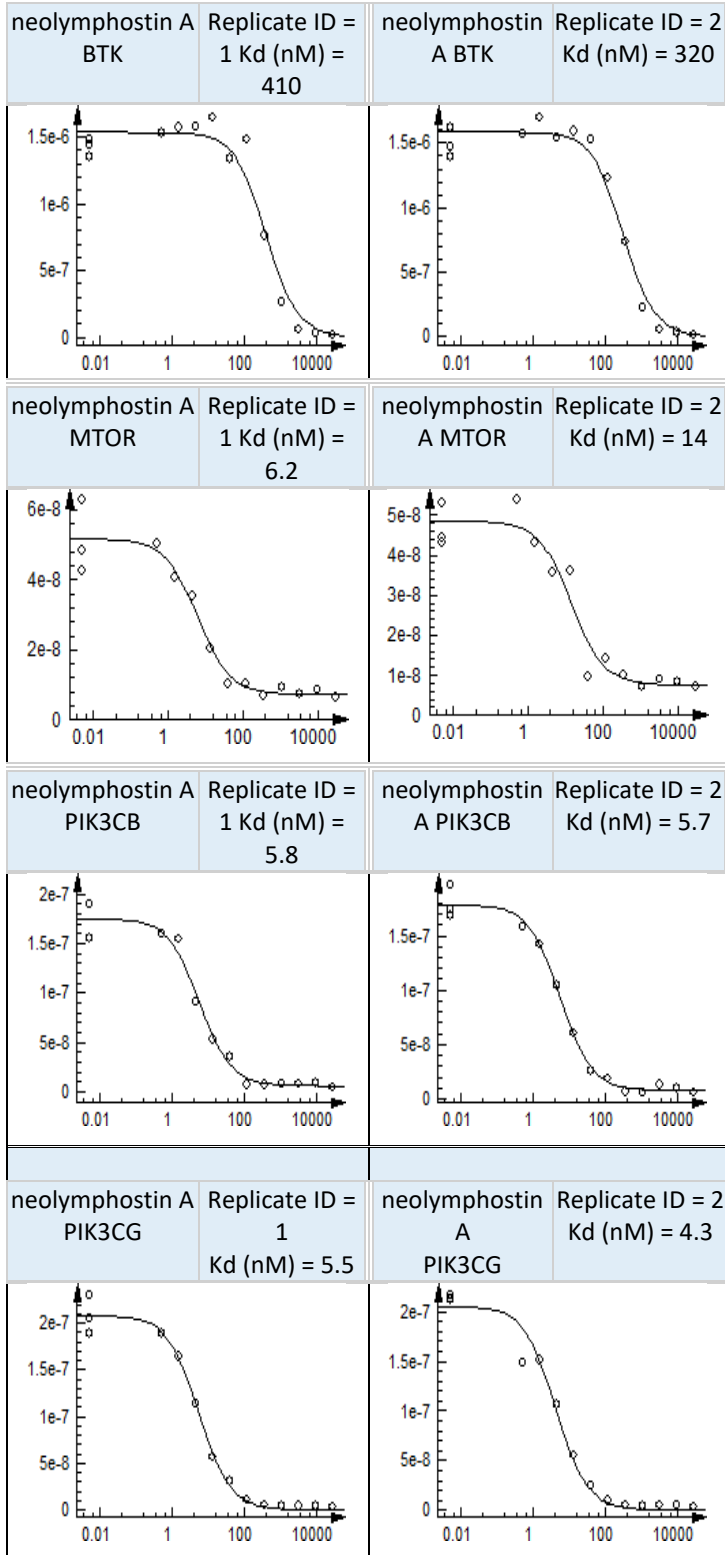
$S(10) = (\text{number of non-mutant kinases with \%Ctrl} < 10) / (\text{number of non-mutant kinases tested})$

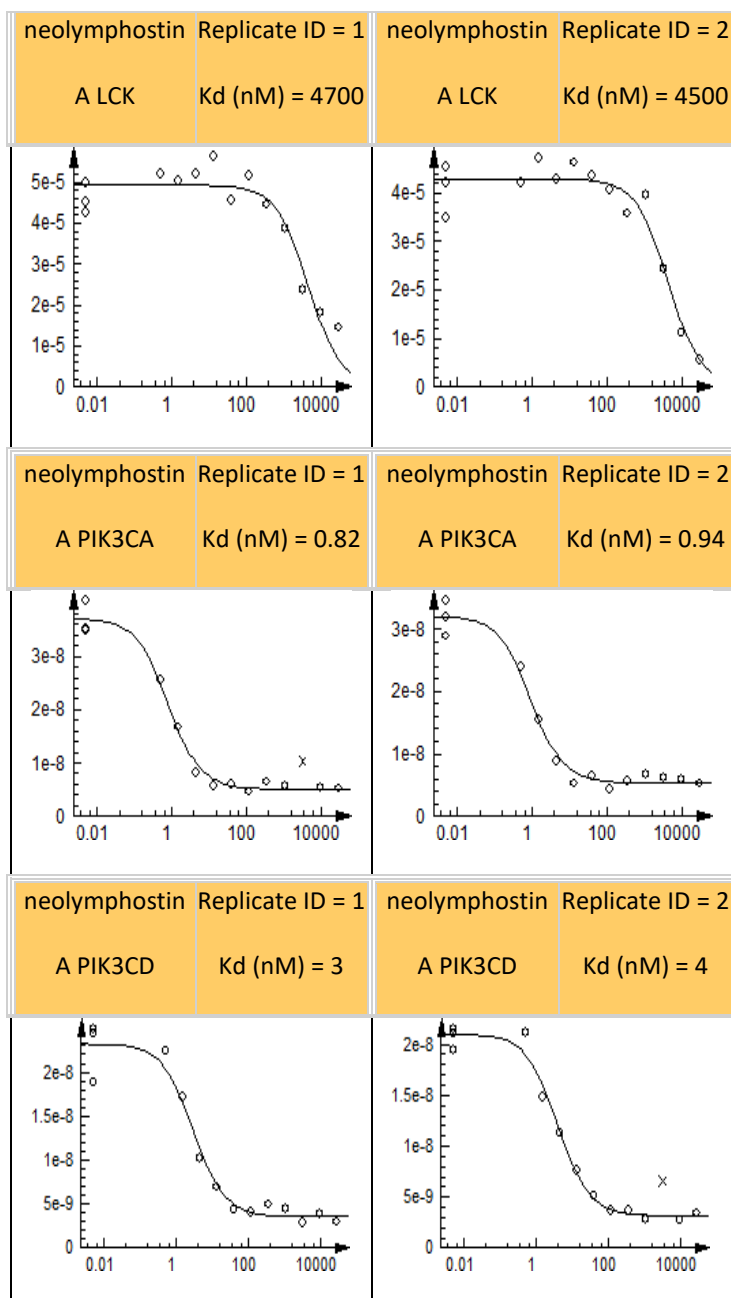
$S(1) = (\text{number of non-mutant kinases with \%Ctrl} < 1) / (\text{number of non-mutant kinases tested})$

Table S5. Dissociation constants (K_d) for neolymphostin A and selected kinases

Target	neolymphostin A
Gene Symbol	K_d (nM)
BTK	3370
LCK	4600
MTOR	10
PIK3CA	0.88
PIK3CB	5.7
PIK3CD	3.5
PIK3CG	4.9

Curve images for dissociation constants (K_d) for neolymphostin A and selected kinases



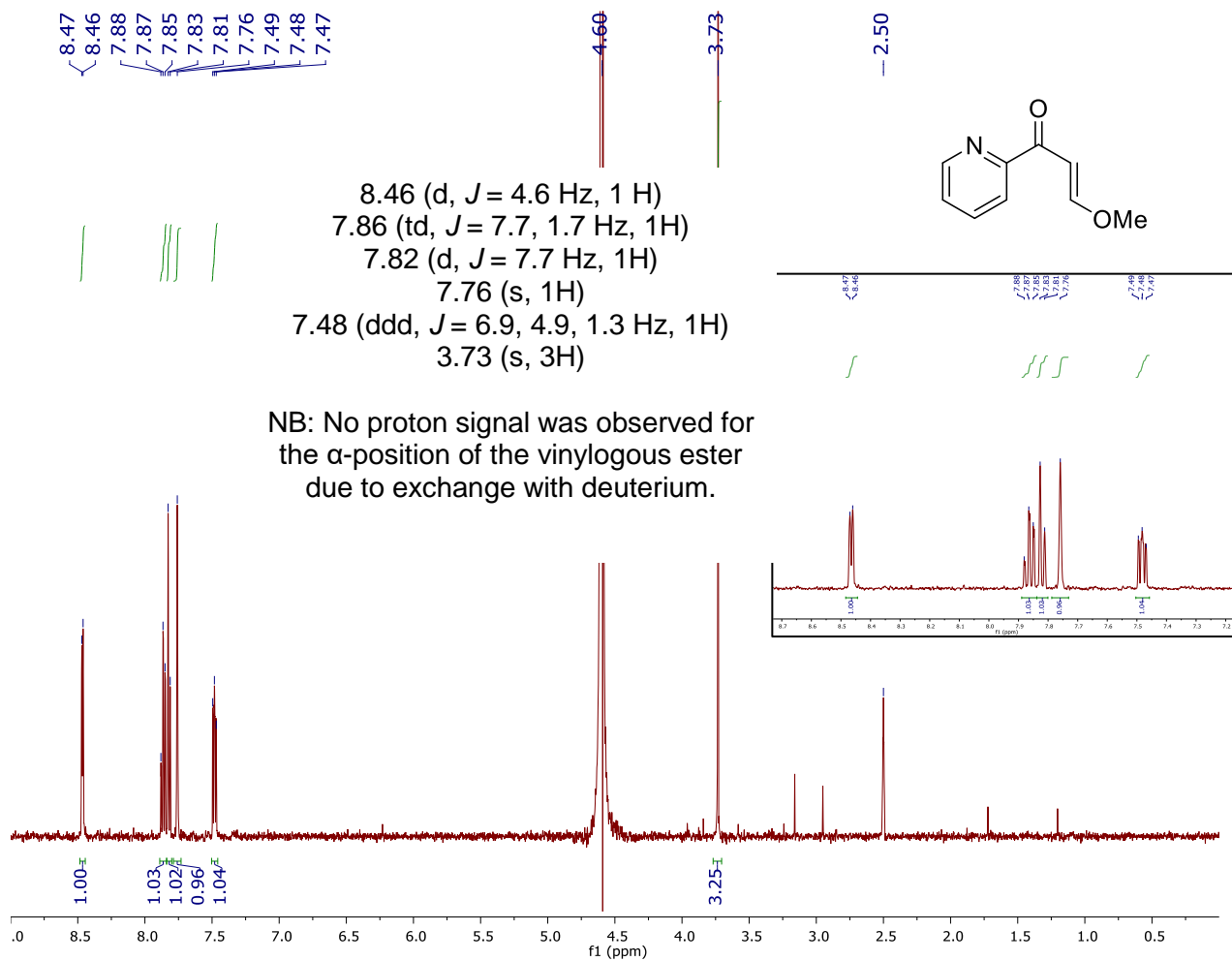


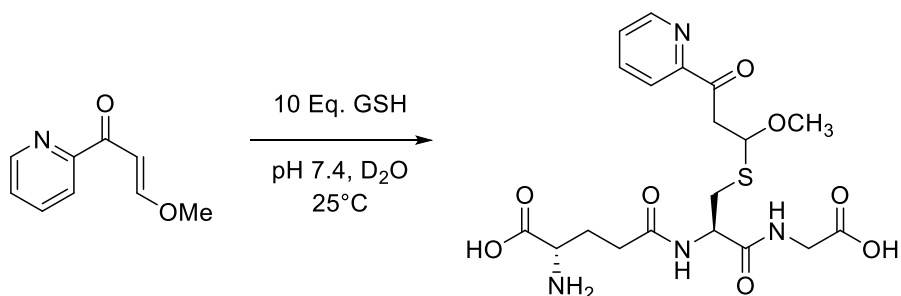
The amount of kinase measured by qPCR (Signal; y-axis) is plotted against the corresponding compound concentration in nM in log₁₀ scale (x-axis). Data points marked with an "x" were not used for K_d determination.

Table S6. Dissociation constants (K_d) for wortmannin and selected kinases

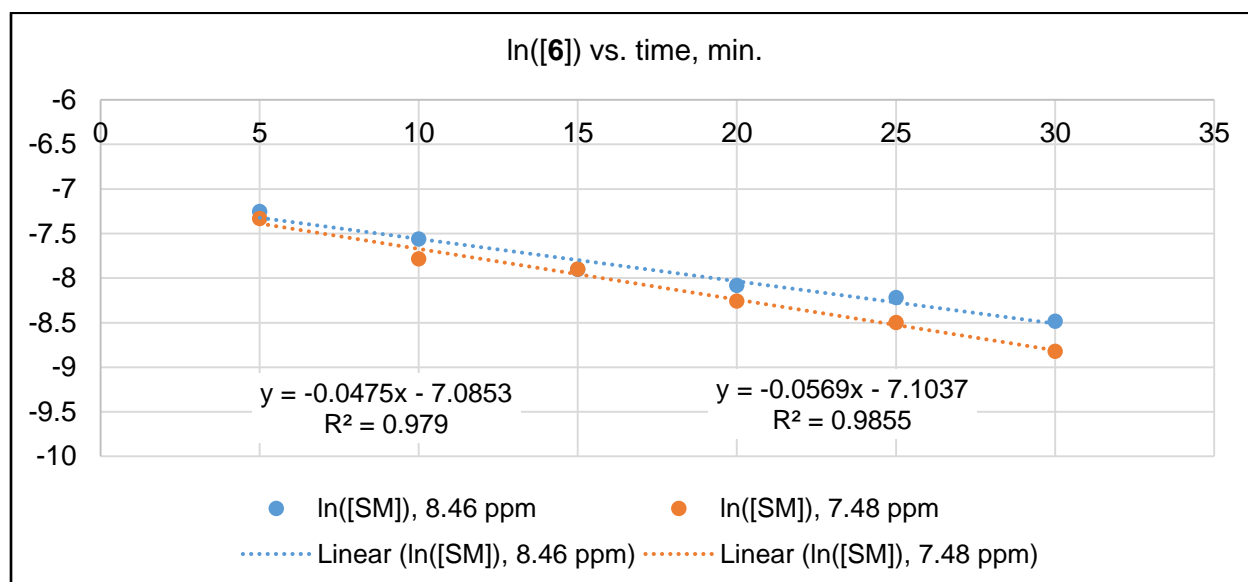
Target	neolymphostin A
Gene Symbol	K_d (nM)
MTOR	9200
PIK3CA	5.4
PIK3CB	7.6
PIK3CD	15
PIK3CG	5.5

¹H NMR (500 MHz, phosphate-buffered D₂O) of **6**





time, min.	ln([6]), 8.46 ppm	ln([6]), 7.48 ppm
5	-7.253256922	-7.331238893
10	-7.56062856	-7.785473736
15	-7.901007052	-7.898153983
20	-8.08047554	-8.259364082
25	-8.218337818	-8.496990484
30	-8.485104877	-8.822574841



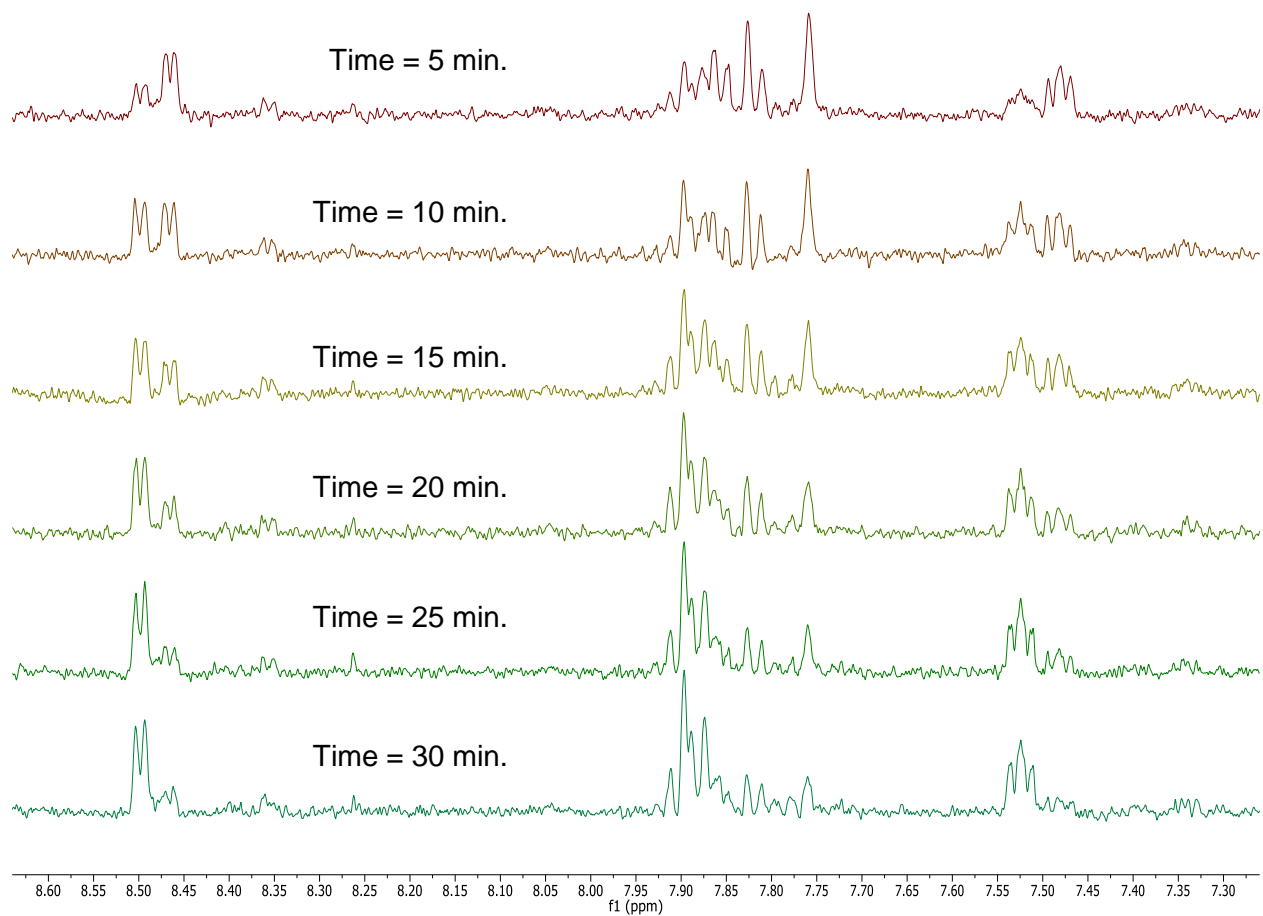
	signal at 8.46 ppm	signal at 7.48 ppm	average
slope of ln([6]) vs. time	-0.0475	-0.0569	-0.0522 ^a
calculated half-life, min.	14.6	12.2	13.3 ^b

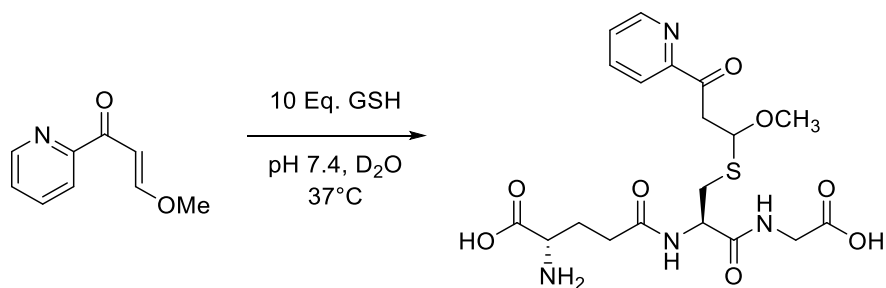
^aAverage of slopes of each linear regression line

^bCalculated by $t_{1/2} = \ln(2)/(-1 \cdot \text{average slope})$

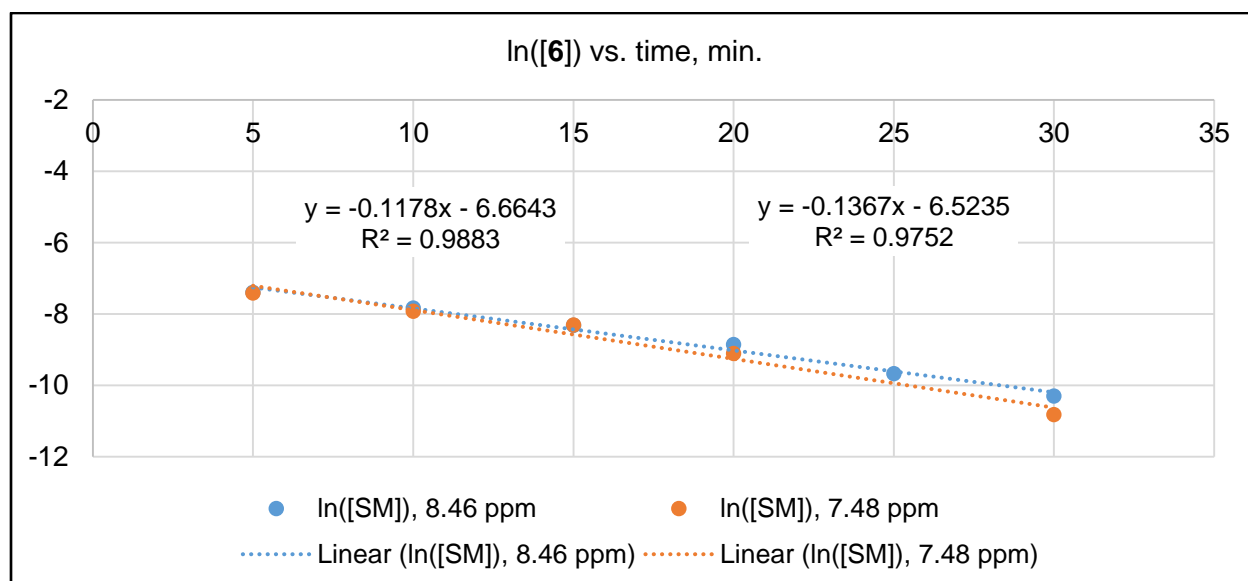
Figure S4a. ¹H NMR kinetics data for **6** reacting with GSH in phosphate-buffered D₂O, pH 7.4, 25°C

Stacked ^1H NMR spectra of **6** reacting with GSH in phosphate-buffered D_2O , pH 7.4, 25°C





time, min.	ln([6]), 8.46 ppm	ln([6]), 7.48 ppm
5	-7.385382833	-7.407998793
10	-7.830088325	-7.919356191
15	-8.317580138	-8.303619091
20	-8.853665428	-9.104979856



	signal at 8.46 ppm	signal at 7.48 ppm	average
slope of ln([6]) vs. time	-0.1178	-0.1367	-0.12725 ^a
calculated half-life, min.	5.9	5.1	5.4 ^b

^aAverage of slopes of each linear regression line

^bCalculated by $t_{1/2} = \ln(2)/(-1 \cdot \text{average slope})$

Figure S4b. ¹H NMR kinetics data for **6** reacting with GSH in phosphate-buffered D₂O, pH 7.4, 37°C

Stacked ^1H NMR spectra of **6** reacting with GSH in phosphate-buffered D_2O , pH 7.4, 37°C

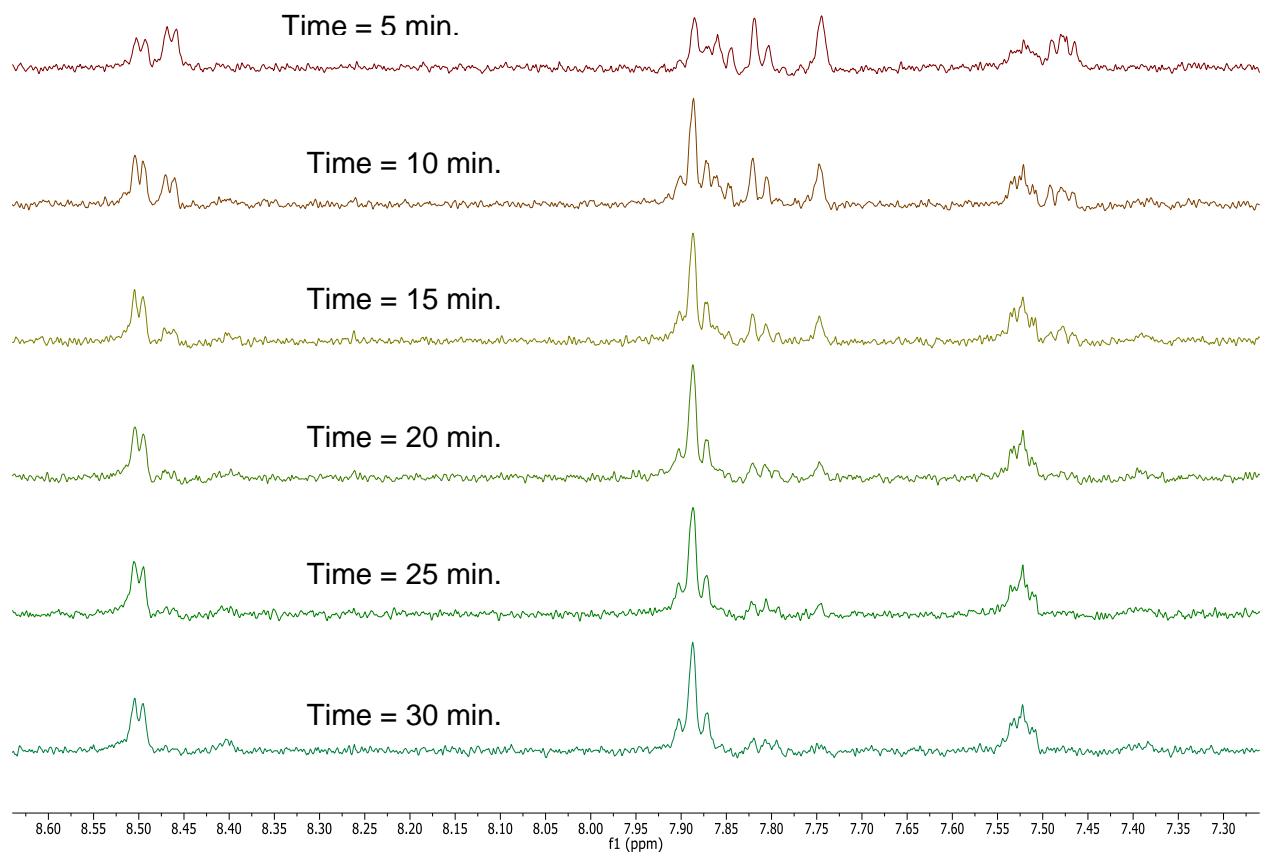
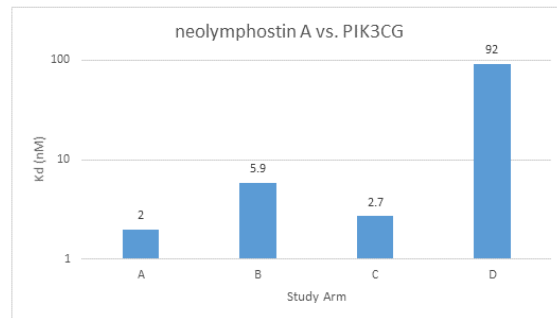
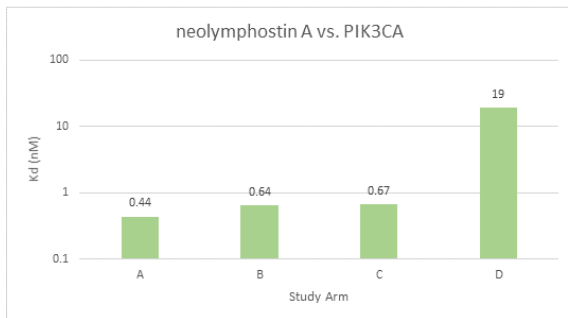
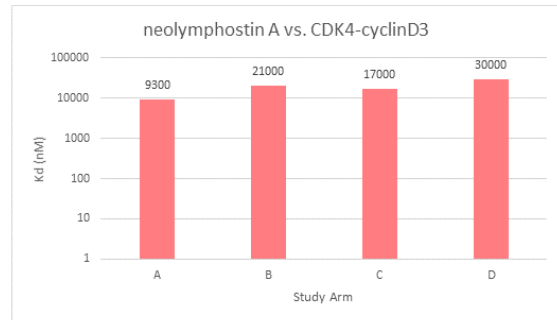
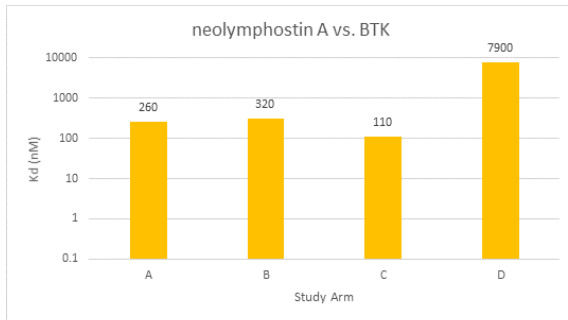
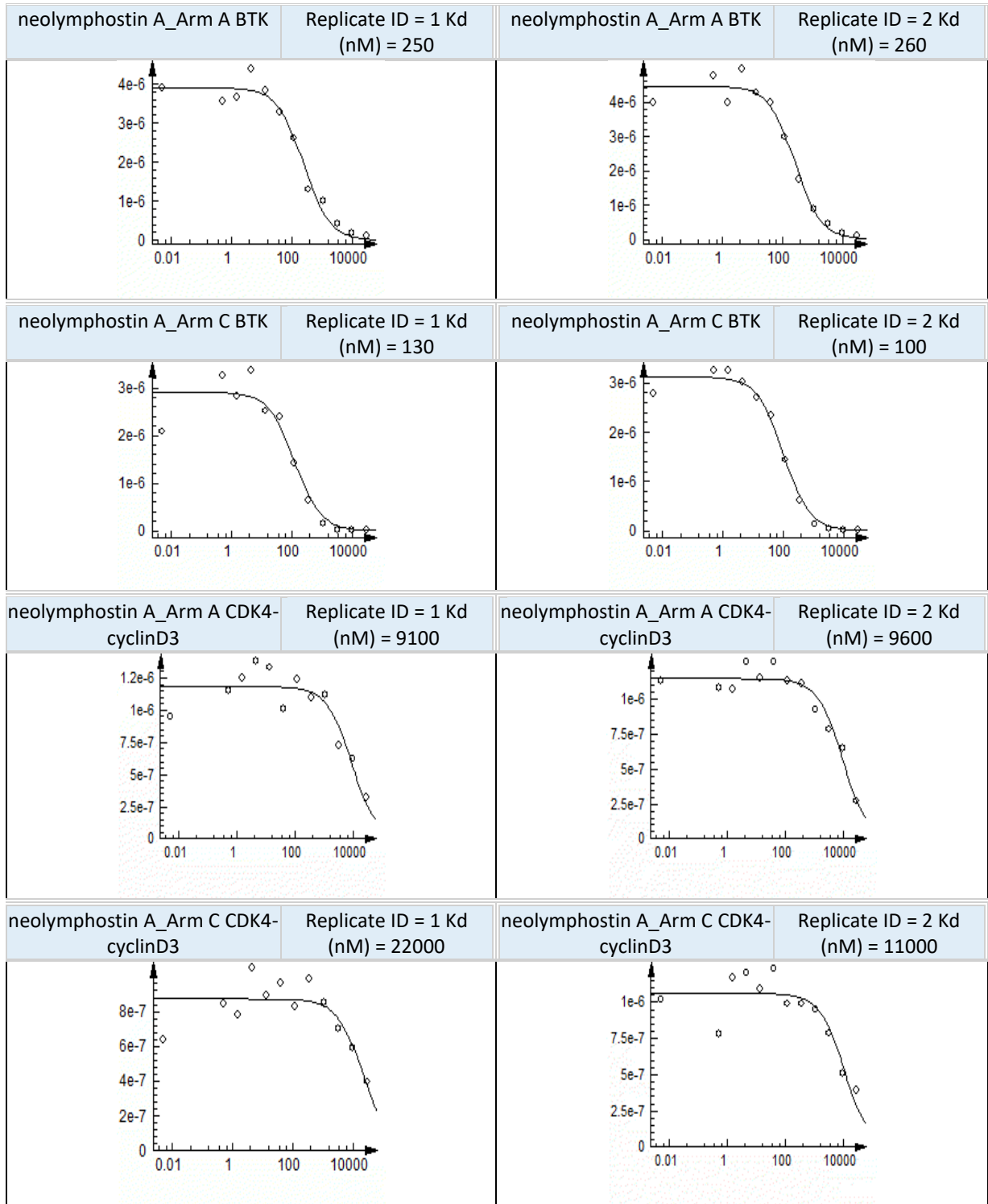


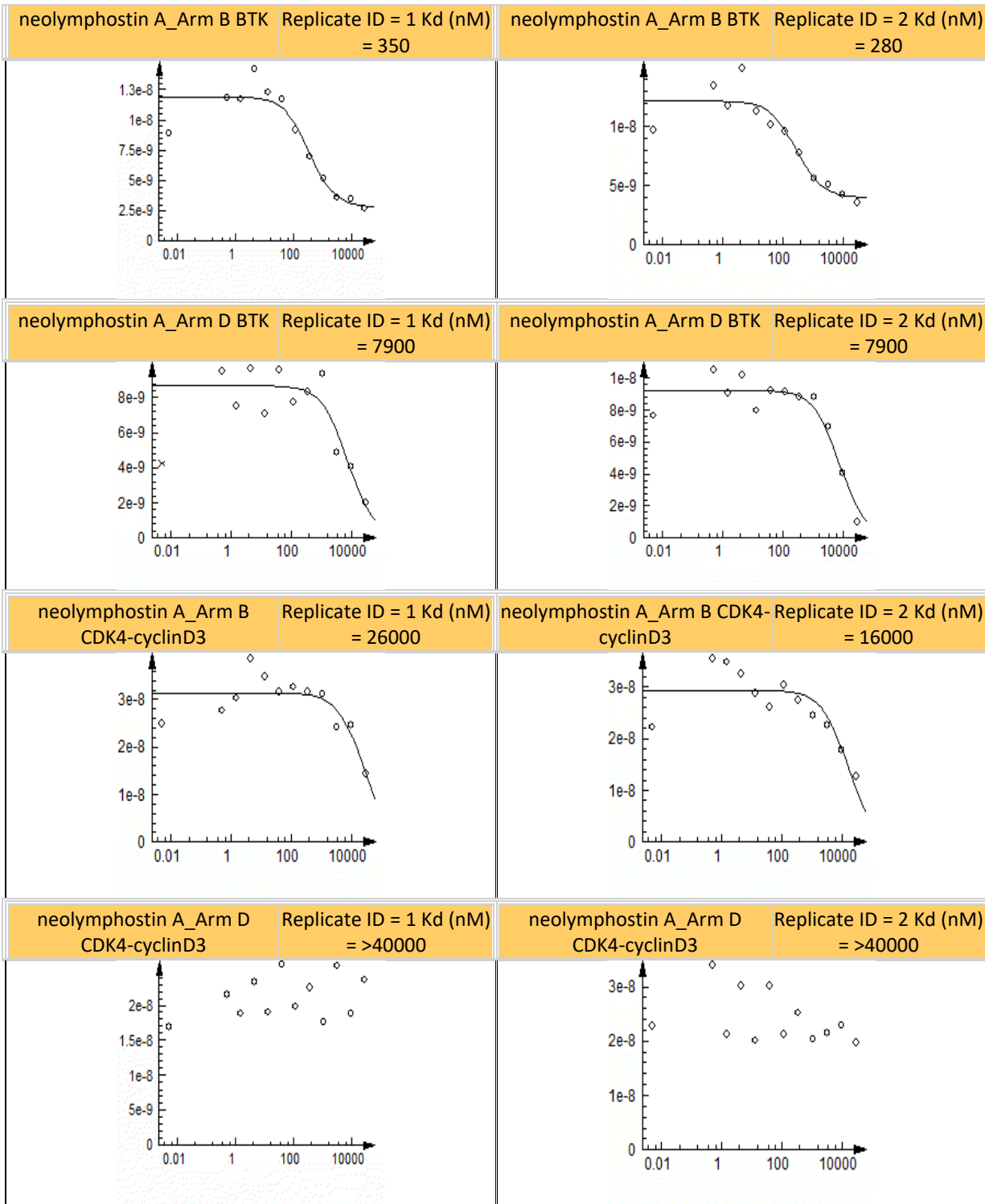
Table S7. Dissociation constants (K_d) for neolymphostin A and selected kinases under four conditioning experiments (Arms A–D)

Compound	BTK	CDK4-CYCLIND3	PIK3CA	PIK3CG
Compound name (conditioning)	K_d (nm)	K_d (nm)	K_d (nm)	K_d (nm)
Neolymphostin A (Arm A)	260	9300	0.44	2
Neolymphostin A (Arm B)	320	21000	0.64	5.9
Neolymphostin A (Arm C)	110	17000	0.67	2.7
Neolymphostin A (Arm D)	7900	>30000	19	92

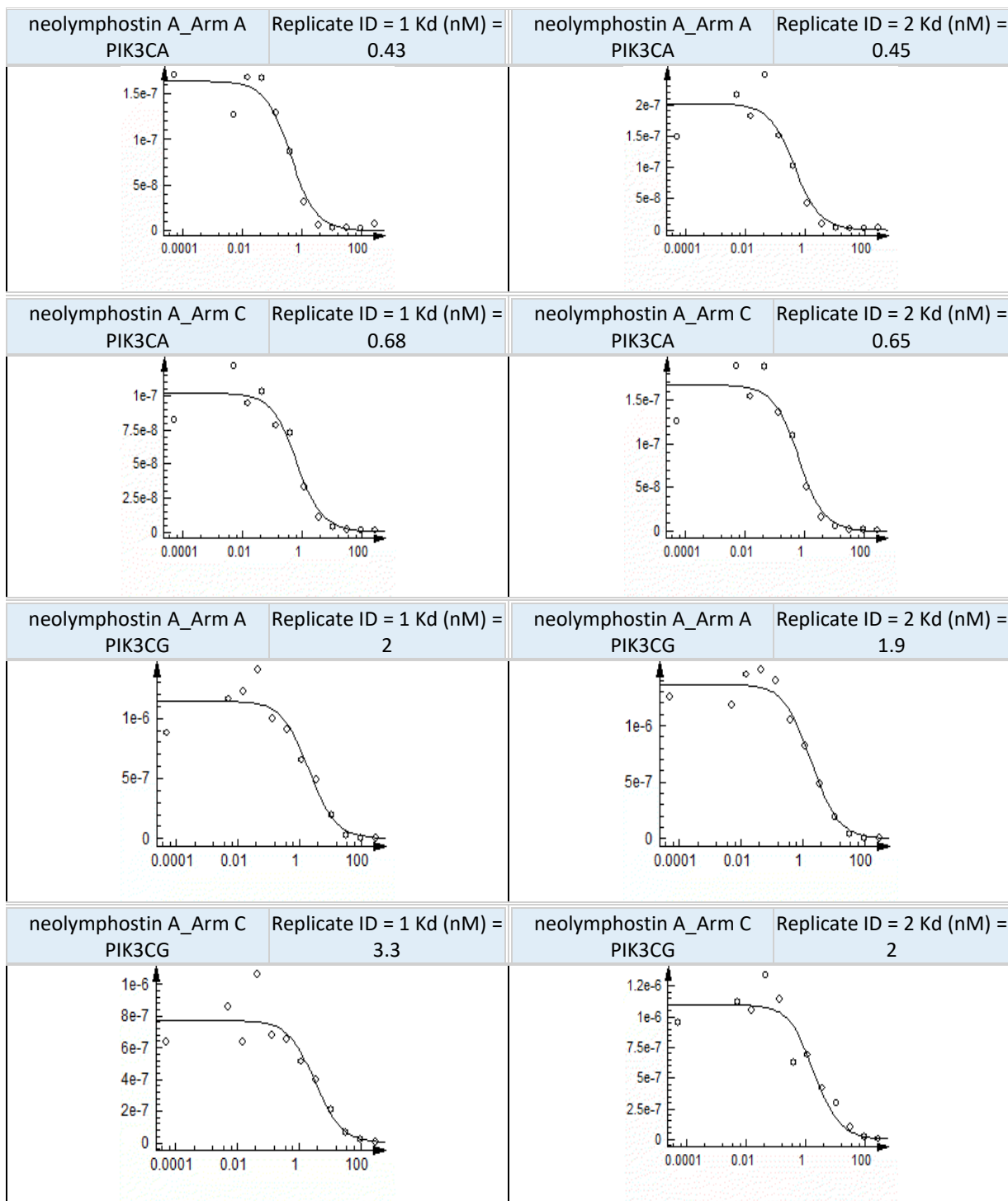


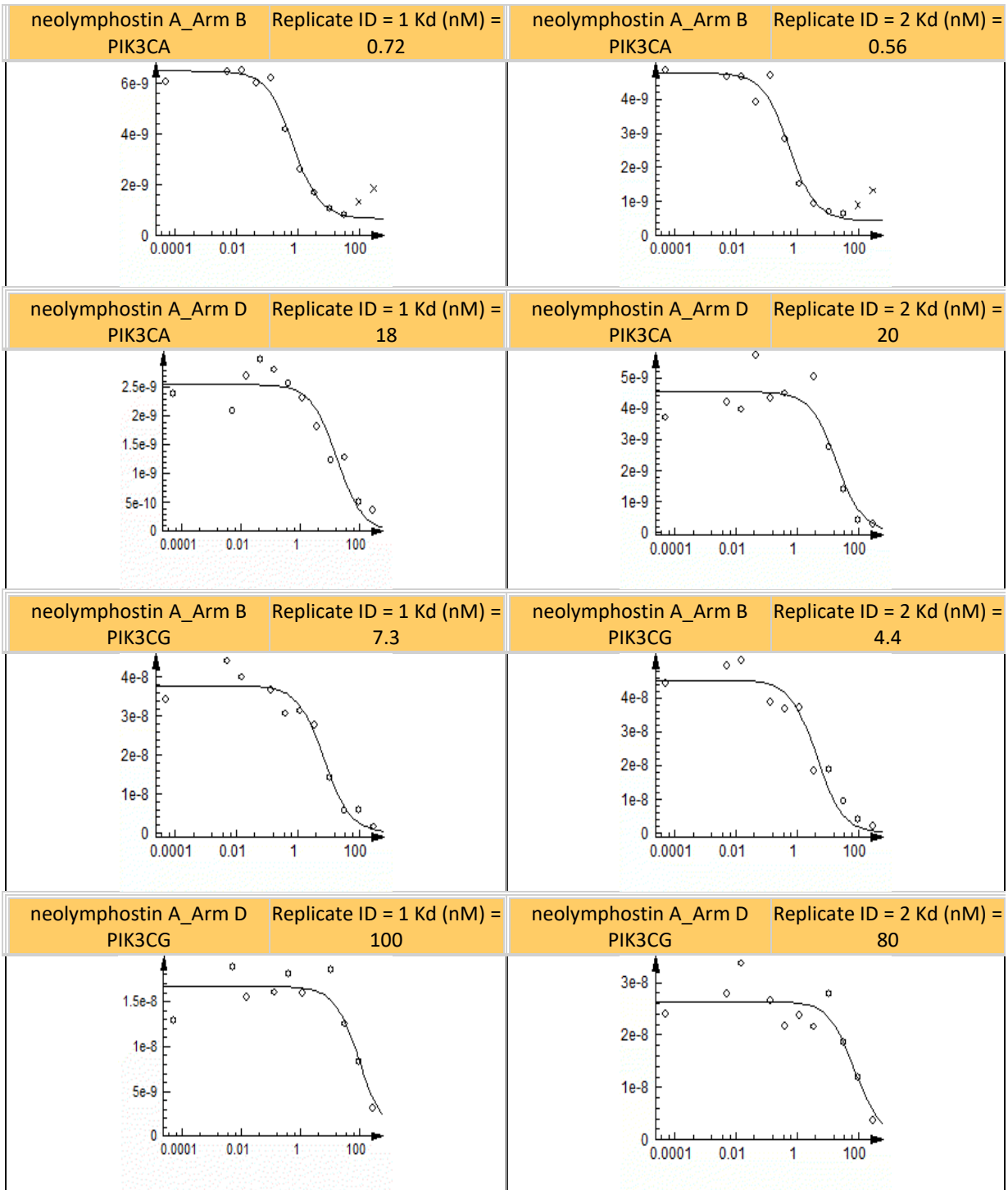
Curve images for dissociation constants (K_d) for neolymphostin A and selected kinases under four conditioning experiments (Arms A–D)





Curve images for dissociation constants (K_d) for neolymphostin A and selected kinases under four conditioning experiments (Arms A–D)...(continued)





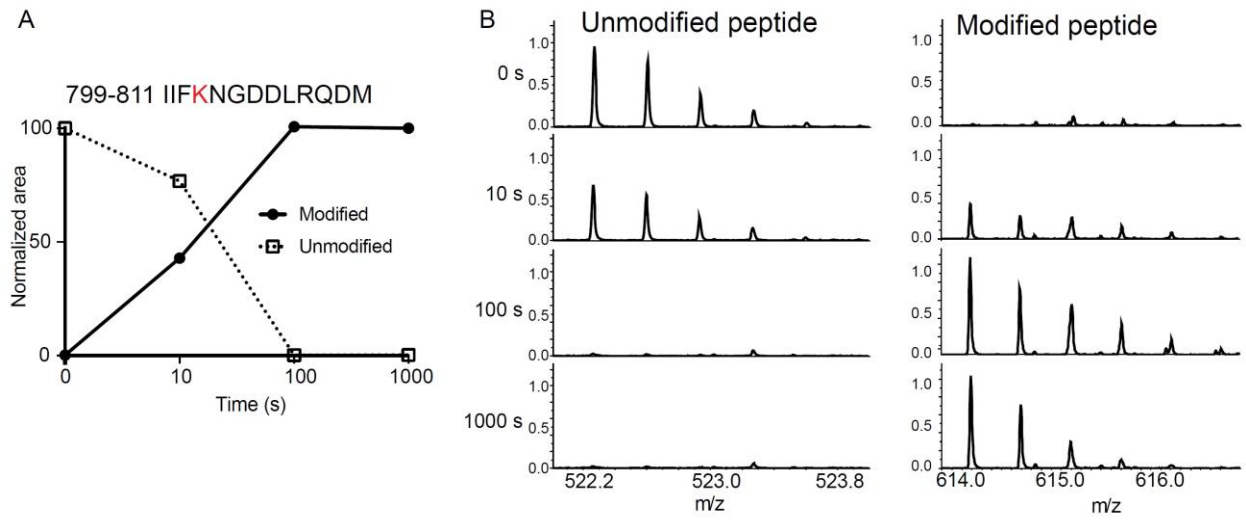


Figure S5. Timecourse for covalent modification of PI3K

A. The intensity for the modified and unmodified variant of the peptide indicated is shown over a time course of incubation with neolyphostin. B. Raw peptide traces of the unmodified and modified peptide over time.

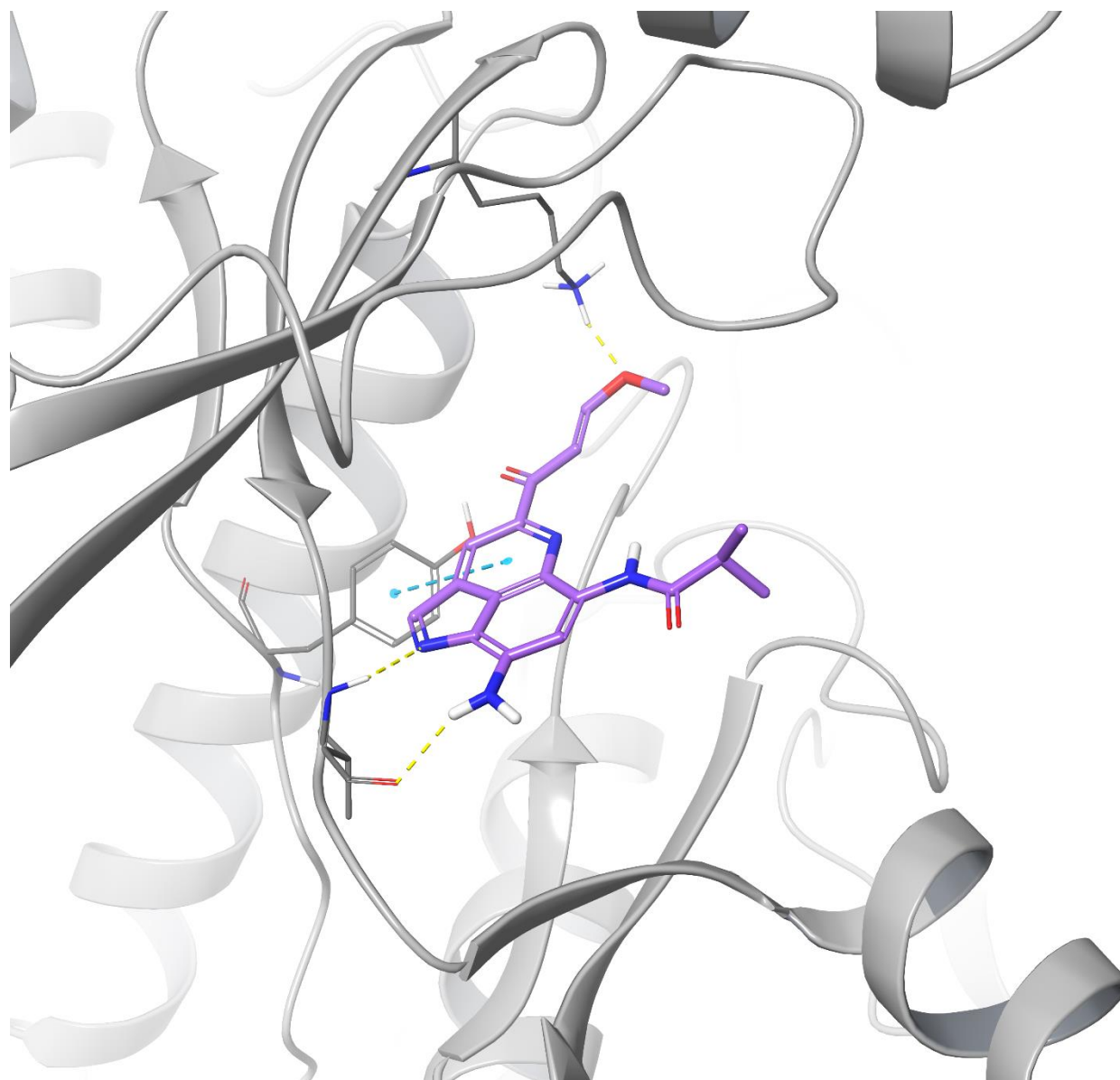


Figure S6. Neolymphostin docked into PI3K α

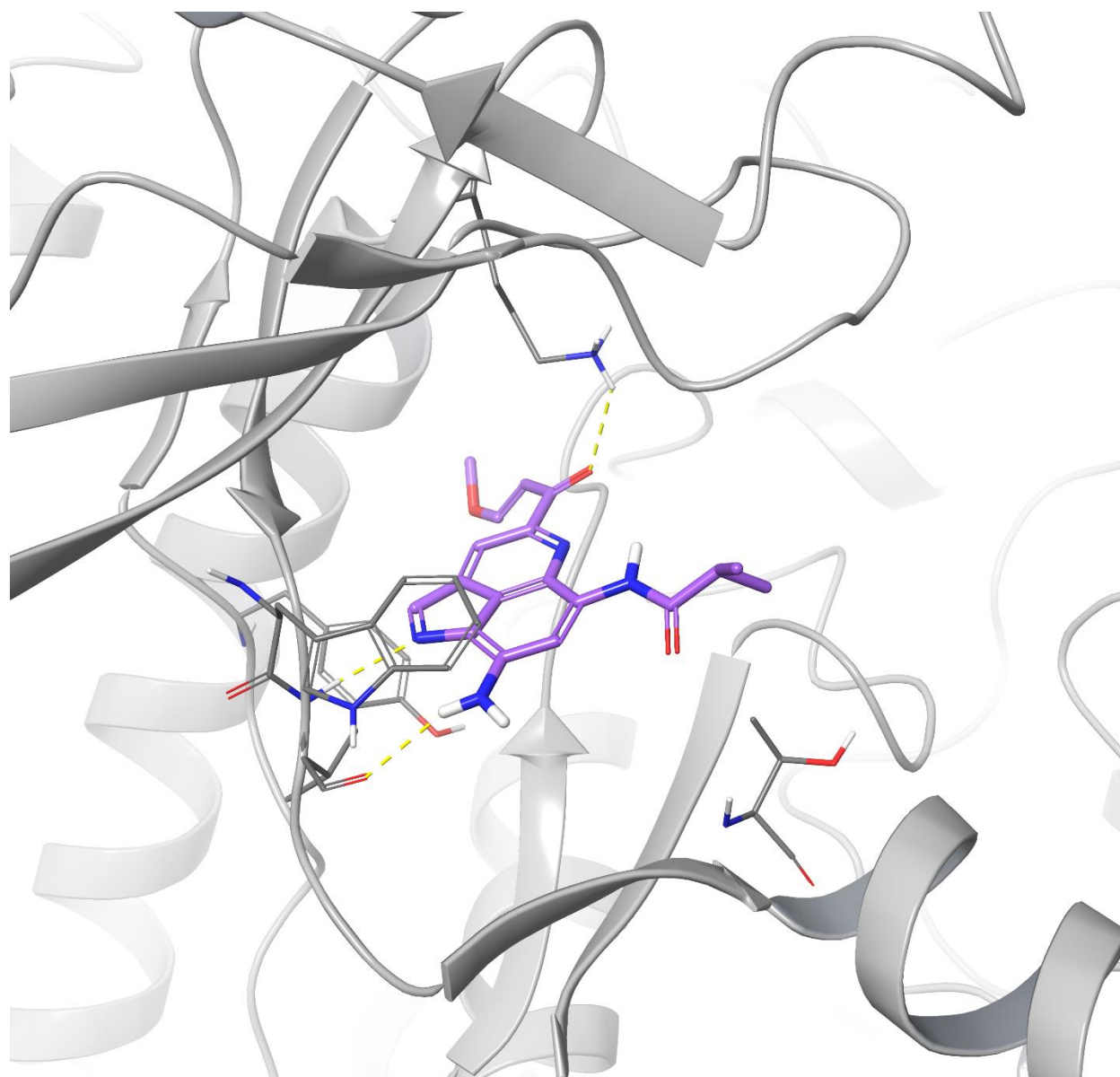


Figure S7. Neolymphostin docked into mTOR

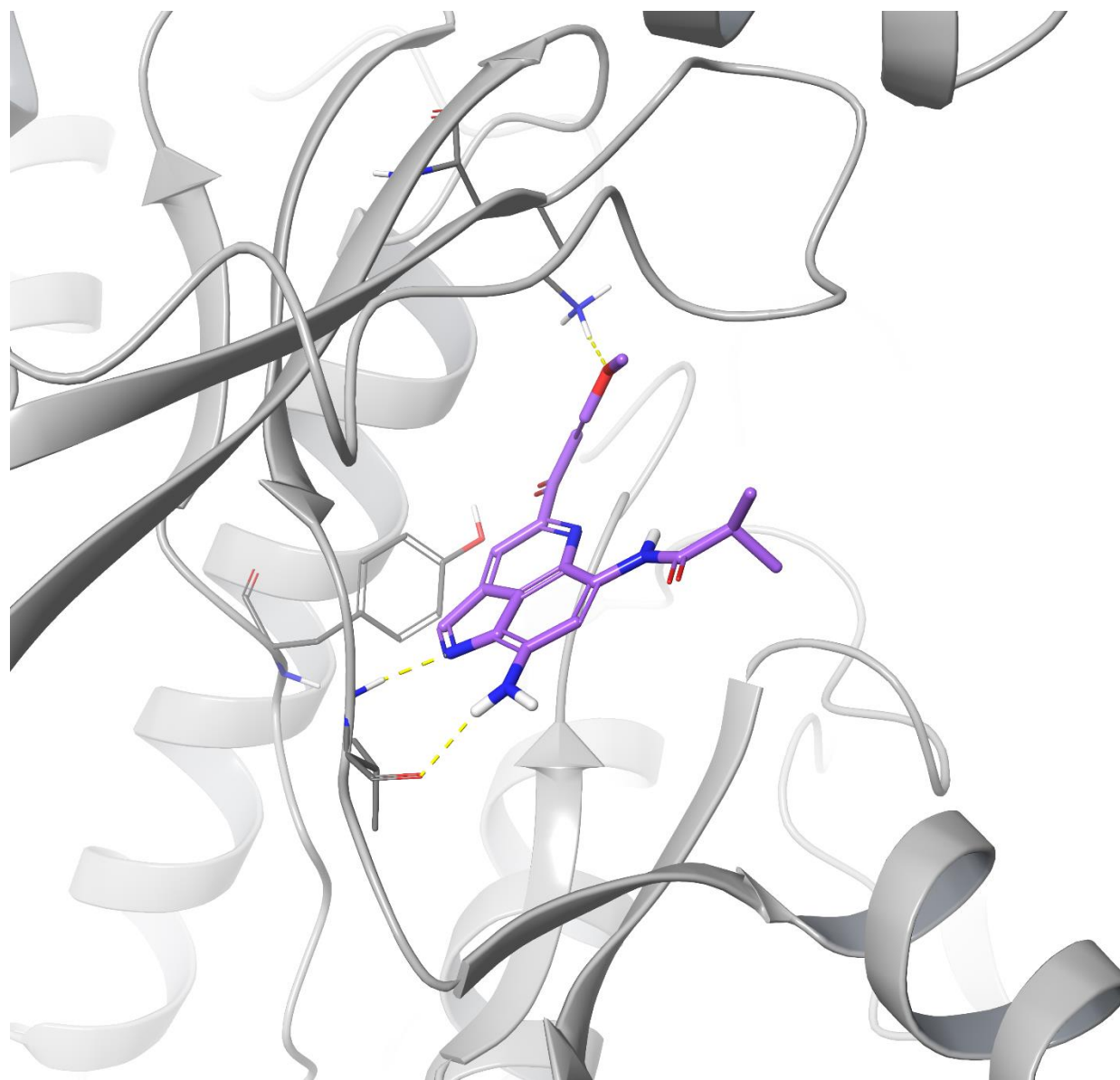


Figure S8. Neolymphostin docked into PI3K α (induced fit)

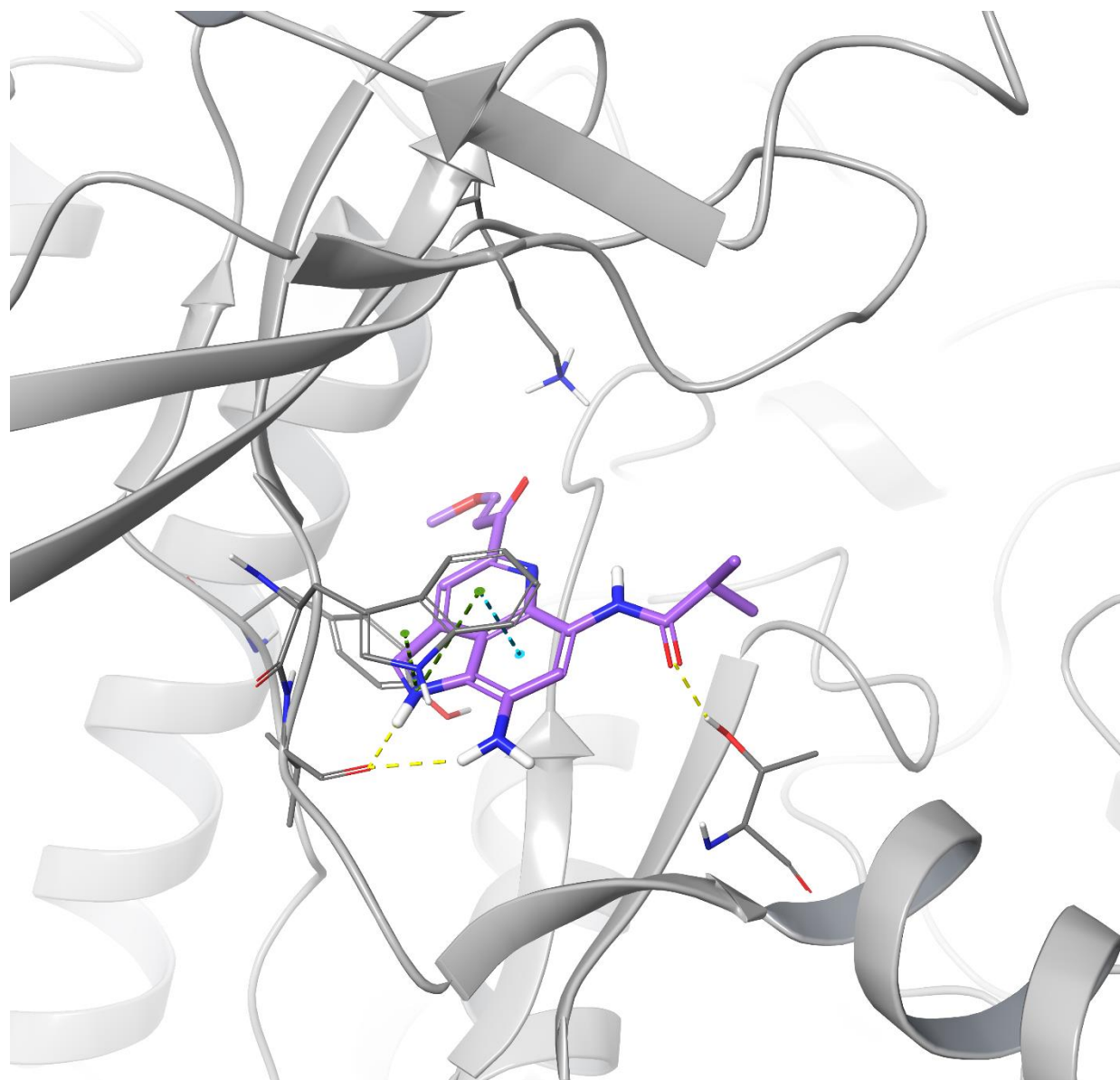


Figure S9. Neolymphostin docked into mTOR (induced fit)

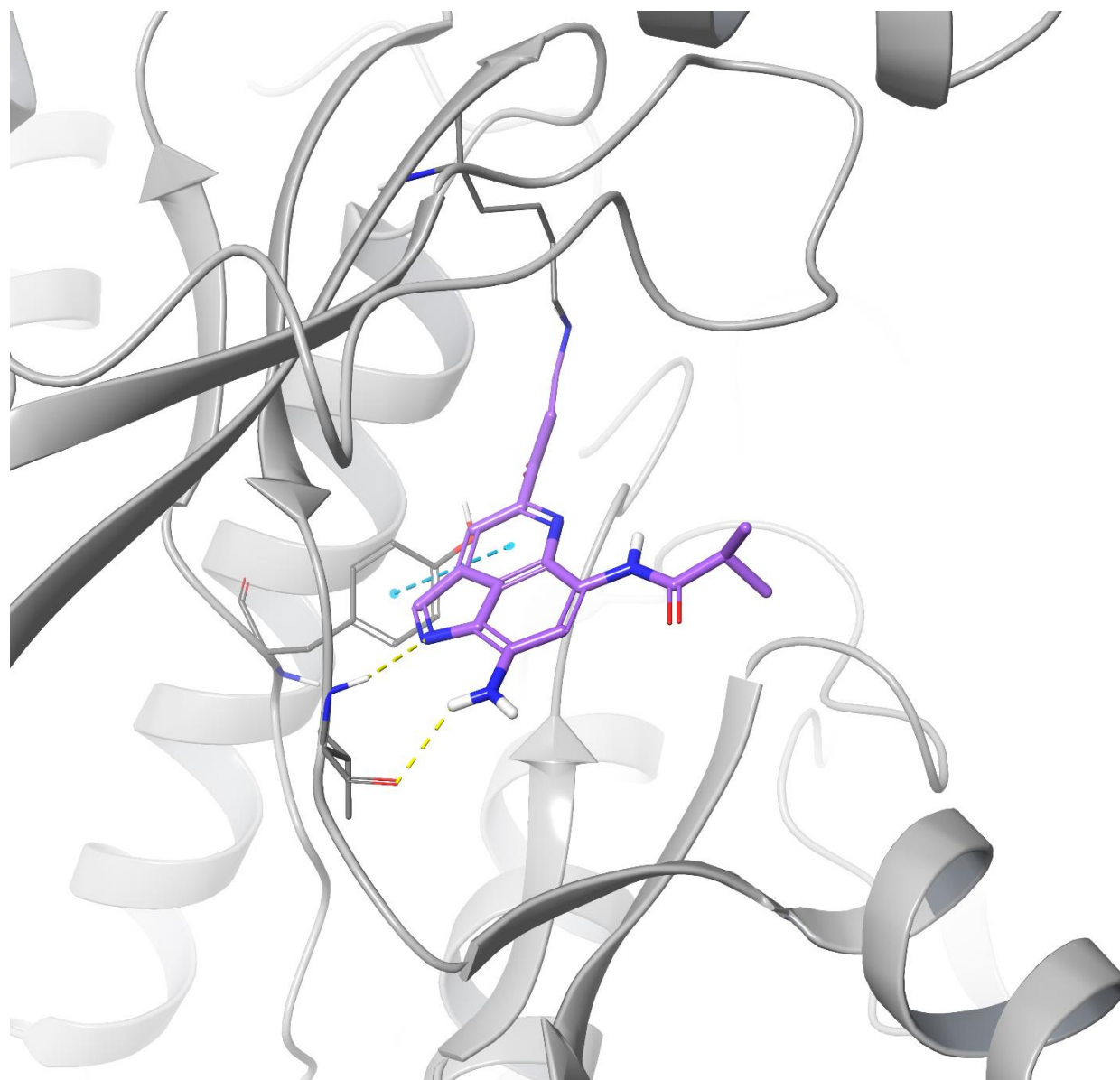
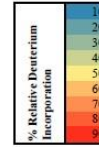
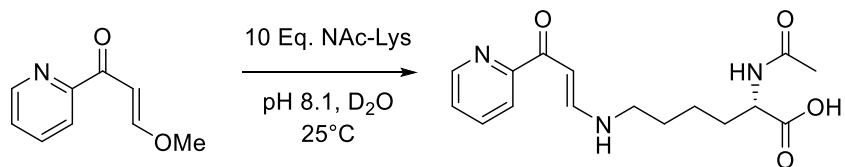


Figure S10. Neolymphostin docked into PI3K α (covalent model)

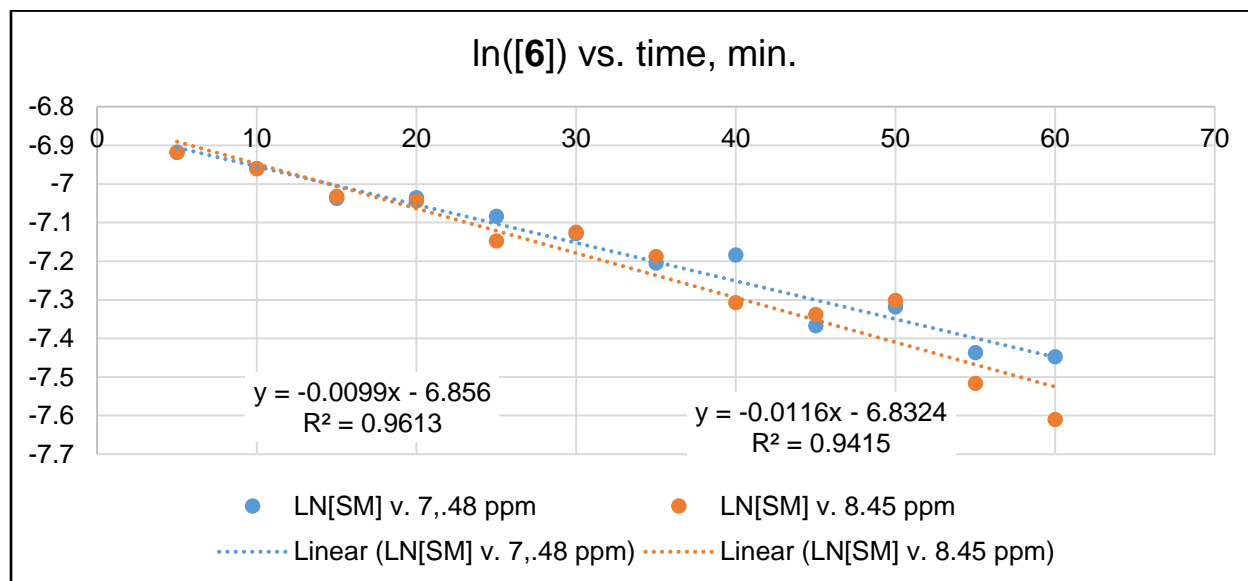
Table S8. HDX-MS experiment

Start	End	Z	RT	Sequence	PI3Ka alone				PI3Ka + neolympohtin			
					3s on ice	SD	300s	SD	3s on ice	SD	300s	SD
11	23	2	12.3	WGHLMPPRILVE	8.8	0.1	40.8	0.5	8.5	0.5	41.3	1.2
31	35	1	8.17	IVTLE	40.5	1.4	55.4	0.4	41.2	2.6	56.4	1.1
37	42	1	7.05	LREATL	1.3	0.2	4.9	0.3	1.5	0.5	5.4	0.4
43	62	4	9.34	ITIKHELFKEARKYPLHQLL	2.9	0.1	21.2	0.2	2.9	0.2	21.4	0.8
50	67	2	8.67	FKEARKYPLHQLLQDESS	6.9	0.0	34.0	0.5	6.5	0.0	35.5	2.0
69	76	1	9.08	IFVSVTQE	2.0	0.4	14.2	0.1	2.1	0.6	14.6	0.3
71	82	2	9.24	VSVTQEAREEF	1.6	0.2	21.8	0.2	1.7	0.3	22.1	1.3
77	82	1	6.99	AEREER	3.1	0.4	29.7	0.2	3.5	1.0	30.6	1.4
83	91	3	6.99	FDTRRLCD	2.2	0.4	6.8	0.3	2.5	0.7	7.9	0.2
83	92	3	9.87	FDTRRLCDL	1.0	0.1	3.8	0.1	1.0	0.0	4.4	0.3
92	99	1	13.96	LRFQFPL	0.9	0.3	1.6	0.0	1.0	0.2	2.0	0.2
93	99	2	12.8	RLFQFPL	1.5	0.2	1.7	0.2	1.7	0.3	1.9	0.2
99	103	1	6.67	LKVI	0.9	0.4	2.7	0.1	1.3	0.3	2.9	0.3
100	119	4	9.76	KVIEFVGNREEKILNREIGF	6.7	0.0	33.8	0.4	6.7	0.3	34.6	1.3
120	127	1	9.56	AIGMPVCE	5.8	0.3	35.1	0.5	6.0	0.9	36.6	0.8
128	139	2	10.93	FDMVKDPEVQDF	11.4	0.5	31.2	0.3	11.4	0.8	32.2	0.6
145	152	1	5.21	NVCKEAVD	7.2	1.3	15.0	2.4	7.6	2.7	15.8	1.0
147	152	1	3.96	KCEAVD	5.2	0.8	13.8	1.5	5.8	0.5	14.6	0.9
153	163	3	4.58	LRDLNSPHSRA	19.6	0.3	36.8	0.7	19.5	0.4	38.0	2.1
154	164	3	4.56	RDLSNSPHSRAM	13.3	0.5	34.6	0.6	13.5	0.5	35.7	1.7
165	193	5	10.35	YVYPPNVSSPELPHKHYYNKLKDGQIVIV	12.6	0.1	34.2	0.2	12.6	0.2	34.9	1.5
193	209	3	11.67	VIVYVSPNNDKQKYTL	16.9	0.4	39.9	0.3	16.4	0.1	40.4	1.6
196	223	5	9.13	VIVYVSPNNDKQKYTLKINHDCVPEQVIAE	9.7	0.1	22.7	0.0	9.3	0.4	23.4	1.0
223	233	3	5.12	EAIRKTRSM	30.1	0.1	43.2	0.5	29.6	0.4	44.2	1.7
224	233	3	4.56	AIRKTRSM	29.4	1.1	41.5	0.5	29.6	0.1	42.7	1.6
244	252	1	10.63	LEYQKYL	35.1	0.3	52.6	0.3	34.8	0.5	54.3	2.5
253	259	1	4.94	KVCGDE	3.0	0.5	12.9	0.3	3.4	1.0	14.2	0.6
260	269	2	10.82	YFLEKYLPSQ	11.8	0.3	18.7	0.1	11.8	0.7	19.2	1.0
261	275	4	9.46	FLEKYLPSQKYIRS	10.1	0.3	16.1	0.1	9.7	0.3	16.6	0.3
279	287	2	11.93	LGRMPLNML	13.0	0.1	25.0	0.5	12.9	0.3	25.4	1.0
294	301	1	10.13	YSQLPMD	36.0	0.4	65.3	0.6	36.0	0.6	66.8	1.5
302	327	4	9.64	FTMPSYRRISTATPYMNGETSKSL	38.1	0.6	45.3	0.6	37.7	0.9	46.6	2.1
328	334	1	12.03	WVINSAL	29.0	0.2	41.7	0.2	29.0	0.5	42.7	0.8
343	369	4	10.94	YVNVNIRIDIKYVIRGTIYHGGEPLCD	12.6	0.6	23.7	0.7	12.4	0.4	24.2	1.1
370	389	3	11.81	NVNTQRVPCSNPRWNEWLNLY	5.0	0.3	32.5	0.2	4.3	0.5	32.8	2.4
390	402	3	11.4	DIYVDPDPAARL	1.9	0.4	10.6	0.2	2.1	0.8	11.1	0.1
405	429	3	10.62	SICSVKGRKGAKEEHCPALWGNLNL	3.7	0.2	6.4	0.1	3.6	0.1	6.9	0.1
430	436	1	10.77	FDYDTL	1.7	0.2	9.2	0.1	2.0	0.2	10.5	0.6
437	443	1	7.53	VSGKMAL	14.7	0.3	41.5	0.4	15.0	1.3	42.4	0.8
444	474	3	13.81	NLWPPVPHGEDLLNPIGVTSNPNKTPCLE	5.5	0.2	30.1	0.9	5.3	0.2	31.2	2.1
456	474	2	9.29	LNPIGVTSNPNKTPCLE	8.6	0.3	38.7	0.5	8.7	0.6	39.5	1.6
474	478	1	10.11	ELEFD	3.4	0.5	12.7	0.6	4.8	2.6	9.4	0.5
483	491	2	10.71	VKFPDMSV	21.9	0.0	41.4	0.8	21.8	0.4	42.0	1.8
484	491	2	10.4	VKFPDMSV	27.0	0.1	51.4	0.6	27.4	0.4	52.8	1.6
492	498	1	8.23	IEEHANW	9.5	0.8	39.9	0.3	9.6	1.4	39.0	2.1
495	506	3	9.19	HANWVSVSREAGF	35.2	0.4	45.3	0.3	35.2	0.8	45.9	1.0
507	522	3	6.27	SVSHAGLSNRLARDNE	41.5	0.3	43.9	0.4	41.1	1.3	44.3	1.5
523	531	3	5.7	IRENDKEQL	14.9	0.2	44.3	0.3	15.3	0.3	45.4	1.5
532	542	2	6.99	KAISTRDLSE	7.1	0.6	29.9	0.4	7.2	0.5	31.4	1.7
546	551	1	9.19	QEKDFL	2.3	0.2	2.8	0.2	2.4	0.7	3.2	0.2
550	565	4	12.43	FLWSHRHYCVTPEIL	1.4	0.2	16.2	0.1	1.1	0.4	16.8	0.8
552	570	4	13.05	WSHRHYCVTPEILPKLL	0.3	0.1	11.6	0.2	0.3	0.2	12.0	0.6
571	583	2	8.53	SVKWSNRDEVAQM	1.4	0.2	7.6	0.1	1.5	0.3	8.1	0.3
584	600	3	11	YCLVKDWPPKPEQAME	1.2	0.1	17.5	0.8	1.3	0.3	18.7	0.8
602	632	5	13.87	LDNYPDPVVRGFAVRCKEYLTDDKLSQYL	3.0	0.1	18.2	0.7	3.0	0.1	18.6	1.5
631	635	1	12.55	YLIQL	0.4	0.5	0.7	0.3	1.0	1.8	0.9	0.4
636	649	2	13.67	VQVLKYEQLDNL	0.9	0.2	15.2	0.2	0.8	0.2	15.6	0.3
649	666	4	9.71	LVRFLKALKALNQRIGHF	0.3	0.0	0.6	0.1	0.5	0.2	0.7	0.1
667	671	1	13.61	FFVHL	1.7	0.5	2.1	0.9	0.8	0.3	2.0	0.6
672	687	2	9.21	KSEMHNKTVSQRGFL	2.5	0.1	16.9	0.3	2.5	0.3	17.1	0.5
688	692	1	7.65	LESYC	1.8	0.6	2.1	0.6	2.6	0.9	2.7	0.6
691	697	2	6.93	YCRACGM	13.9	0.2	30.3	0.5	14.1	1.1	30.7	0.7
698	709	2	6.83	YLRHNRNQVEAM	0.7	0.2	1.8	0.1	0.8	0.3	3.1	0.2
716	734	3	7.95	TDILQEKDKDETQKVMQKF	27.3	0.1	54.0	0.4	27.1	0.7	55.5	1.0
735	744	2	8.59	LVEQMRPFD	5.3	0.4	61.6	0.8	4.8	0.6	63.3	3.3
745	766	3	14.06	MDALQGLFSLPNAHQGLNLR	2.5	0.3	8.7	0.1	2.1	0.5	8.9	0.3
769	781	4	10.35	CRIMSSAKRPLWL	11.3	0.1	26.2	0.1	10.8	0.0	24.0	1.0
782	791	1	10.79	NVENPDMISE	23.0	0.2	51.4	0.2	22.3	0.5	52.3	1.5
793	797	1	6.57	LFQNN	48.0	0.8	64.9	0.2	45.7	2.5	65.0	1.8
794	798	1	4.57	FQNN	22.8	0.9	60.0	2.1	20.6	0.4	57.5	2.4
798	806	2	8.11	EIFKNGDD	5.0	0.5	18.3	0.4				
799	807	1	9.65	IFKNGDDL	6.3	0.4	22.7	0.8				
799	811	3	9.5	IFKNGDDLQDM	3.2	0.2	24.6	0.3				
799	813	3	10.3	IFKNGDDLQDMLT	2.2	0.1	16.5	0.4				
802	811	2	6.94	KNGDDLQDM	4.3	0.5	30.4	0.4				
815	821	2	9.55	QIIRIME	0.3	0.1	0.4	0.1	0.4	0.1	0.6	0.2
822	830	1	10.57	NIWQNOGLD	9.2	0.2	26.8	0.3	9.5	1.1	27.9	0.6
830	839	2	13.11	DLRMLPYGCL	1.0	0.2	10.8	1.6	1.0	1.0	9.4	0.3
840	847	1	10.93	SIGDCVGL	2.9	0.4	8.1	0.3	3.0	0.8	8.6	0.3
847	858	2	8.77	LIEVVRNSHTIM	9.6	0.4	21.7	0.2	4.5	0.5	20.5	1.2
847	859	3	8.43	LIEVVRNSHTIMQ	8.6	0.2	23.6	0.6	4.1	0.5	20.7	1.5
848	858	3	7.53	IEVVRNSHTIM	10.4	0.4	26.0	0.4	5.4	0.1	24.0	0.7
848	859	2	7.12	IEVVRNSHTIMQ	9.7	0.2	27.4	0.4	5.2	0.7	24.1	0.6
860	872	3	10.03	IQCKGGKGLALQF	63.6	0.3	66.5	0.2	63.3	0.9	66.4	1.3
873	893	2	8.12	NSHTLHQWLKDKNKNGEYDAA	8.9	0.1	23.5	0.2	8.9	0.4	24.4	0.6
880	893	2	7.8	WLDKWKNGEYDAA	8.6	0.2	32.4	0.2	9.0	0.5	33.2	0.9
892	896	1	9.46	AAIDL	2.3	1.1	4.9	0.8	2.4	1.6	3.1	0.4
909	922	2	10.83	FILGDRHNSNIM	1.8	0.2	7.7	0.1	1.8	0.3	7.8	0.4
921	930	2	10.62	IMVKDQGLF	3.2	0.5	4.5	0.8	3.4	1.0	5.4	0.5
930	956	5	10.24	FHIDFGHFLDKKKKFGYKRRVPPVL	4.1	0.1	17.0	0.2	3.9	0.3	16.9	0.6
957	961	1	10.45	TQDFL	1.5	0.1	10.4	0.1	1.8	0.5	10.7	0.5
961	980	4	9.11	LIVISGAQECTKTREFFERF	11.7	1.3	23.4	0.5	10.9	1.3	24.3	2.0
977	983	2	9.51	FERFQEM	2.0	1.6	3.4	0.9	0.4	0.6	4.3	0.6
984	989	1	8.97	CYKAYL	2.3	0.7	1.6	0.1	2.8	1.9	2.6	0.8
990	997	2	5.07	AIRQHANL	2.8	0.2	30.7	0.5	2.8	0.4	32.0	1.2
1002	1006	1	13.15	FSMML	0.9	0.4	10.5	0.2	1.1	0.8	14.2	0.4
1006	1013	1	10.77	LGSMPPEL	9.2	0.8	40.0	0.7	8.8	2.1	40.2	2.5
1014	1020	1	9.36	QSFDIA	16.2	0.6	18.9	0.3	16.4	0.4	19.5	1.2
1017	1021	1	7.84	DDIAY	3.5	0.5	24.0	0.2	3.3	1.1	24.5	0.8
1029	1037	2	6.21	DKTEQALE	10.3	0.1	32.8	0.8	10.3	0.4	33.3	1.0
1039	1055	2	7.82	FMKQMNDAHGGWTTKM	15.7	0.1	23.4	0.2	15.8	0.1	24.0	0.7
1039	1057	4	9.9	FMKQMNDAHGGWTTKMDW	15.2	0.2	22.1	0.1	15.2	0.0	22.7	0.8
1060	1068	2	3.95	HTIKQHALL	46.2	0.2	52.2	0.8	45.8	0.1	53.4	1.9





time, min.	ln([6]), 8.46 ppm	ln([6]), 7.48 ppm
5	-6.918118066	-6.918566195
10	-6.96067768	-6.960123264
15	-7.032918422	-7.037289331
20	-7.044091723	-7.03558865
25	-7.148035113	-7.083645945
30	-7.125757431	-7.128426641
35	-7.188657664	-7.205387682
40	-7.308078988	-7.184008656
45	-7.338538195	-7.3678788
50	-7.301659565	-7.318497444
55	-7.517132576	-7.437014604
60	-7.6102922	-7.448322872



	signal at 8.46 ppm	signal at 7.48 ppm	average
slope of ln([6]) vs. time	-0.0116	-0.0099	-0.01075 ^a
calculated half-life, min.	59.8	70.0	64.5 ^b

^aAverage of slopes of each linear regression line

^bCalculated by $t_{1/2} = \ln(2)/(-1 \cdot \text{average slope})$

Figure S11. ¹H NMR kinetics data for **6** reacting with N^α-acetyl lysine in phosphate-buffered D₂O, pH 8.1, 25°C

Stacked ^1H NMR spectra of **6** reacting with N^α -acetyl lysine in phosphate-buffered D_2O , pH 8.1, 25°C

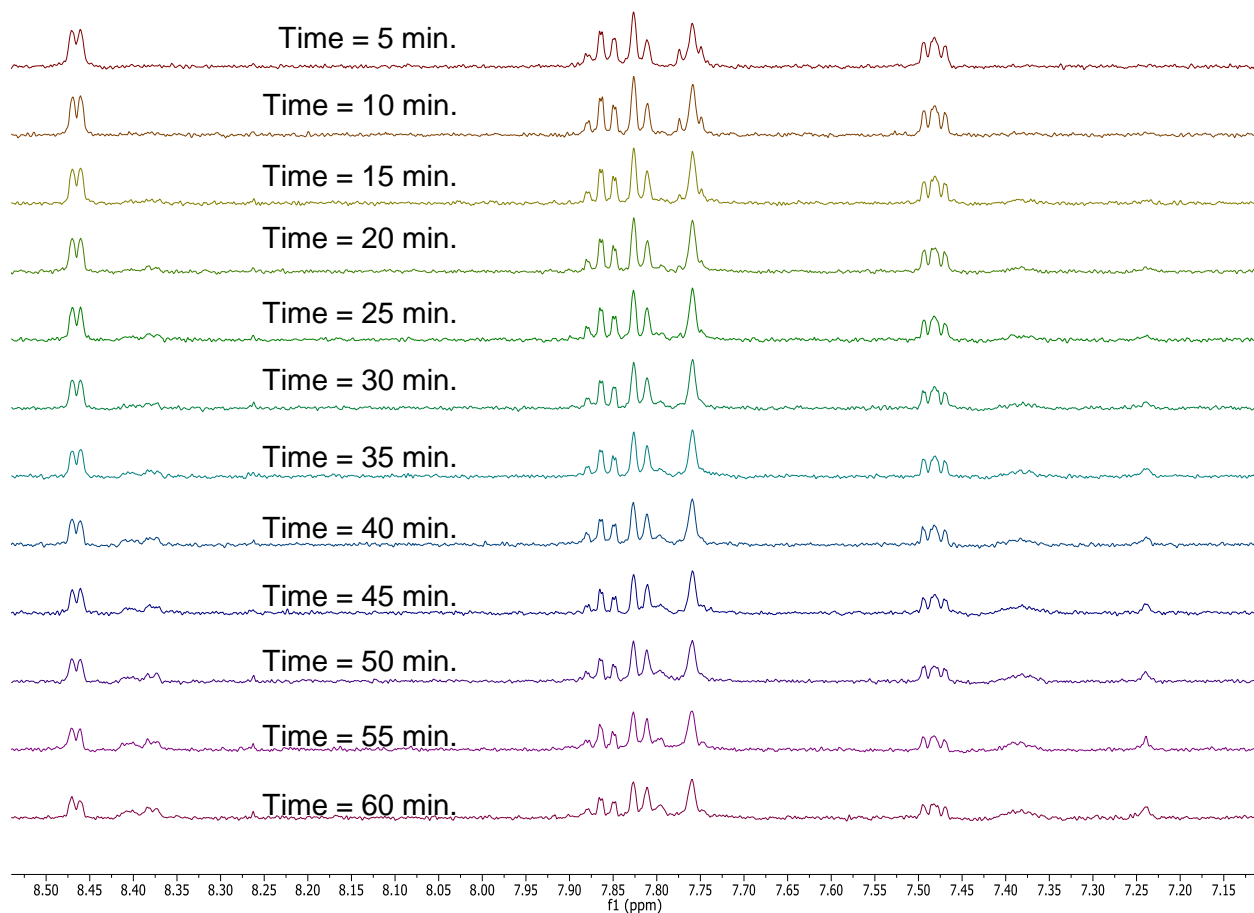
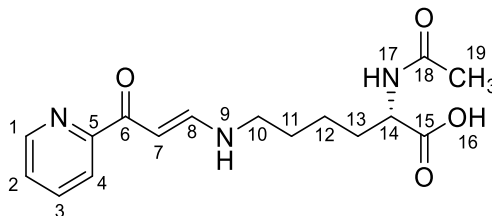


Table S9. NMR spectral data for **8** in phosphate-buffered D₂O, pH 8.1, at 500 MHz

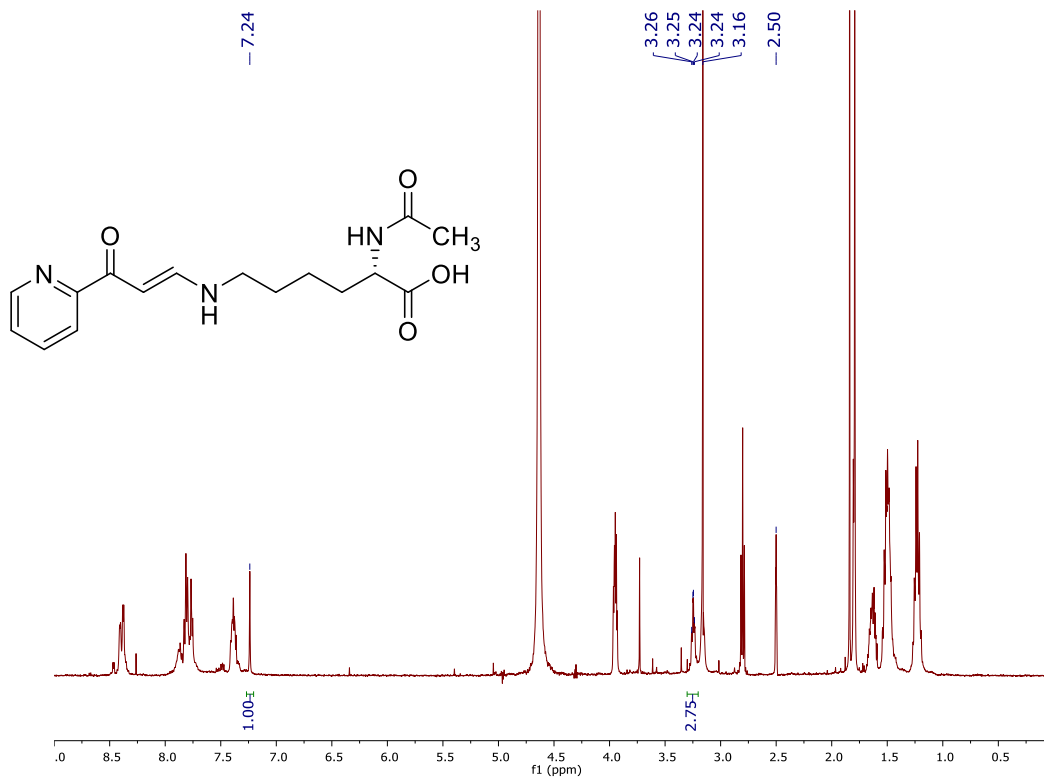


Position #	δ_H (mult., J(Hz))	δ_C^b	COSY	HMBC
1	8.39 (dd, 13.5, 4.4)	149.8	2	3
2	7.39 (m)	127.8	1, 3	4
3	7.80 (m, 7.5)	139.9	2, 4	1, 5
4	7.76 (m, 7.5)	123.5	3	2
5	---	156.0	---	---
6	---	188.0	---	---
7	--- ^a	90.6	---	---
8	7.24 (s)	160.0	---	6, 7, 10
9-NH	--- ^a	---	---	---
10	3.25 (td, 6.5, 2.3)	50.4	11	8, 11, 12
11	1.52 (m)	27.8	10, 12	10, 12
12	1.23 (m)	23.6	11, 13	10, 11, 13, 14
13	a 1.52 (m) b 1.63 (m)	32.5	14 12, 13a, 14	11, 12, 14, 15 12, 14, 15
14	3.95 (m)	56.4	13a, 13b	12, 13, 15
15	---	180.4	---	---
16-OH	--- ^a	---	---	---
17-NH	--- ^a	---	---	---
18	---	174.8	---	---
19	1.81 (s)	23.2	---	18

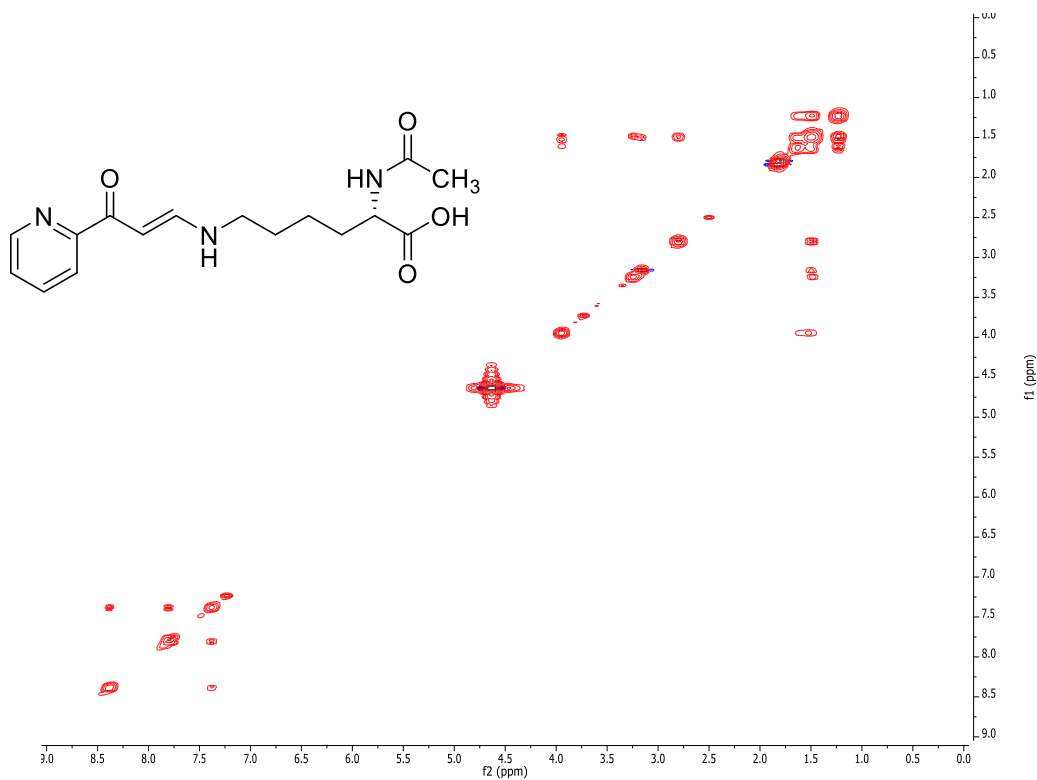
^a Proton signals were not observed due to exchange with deuterium.

^b Carbon chemical shifts were based on HSQC and HMBC data.

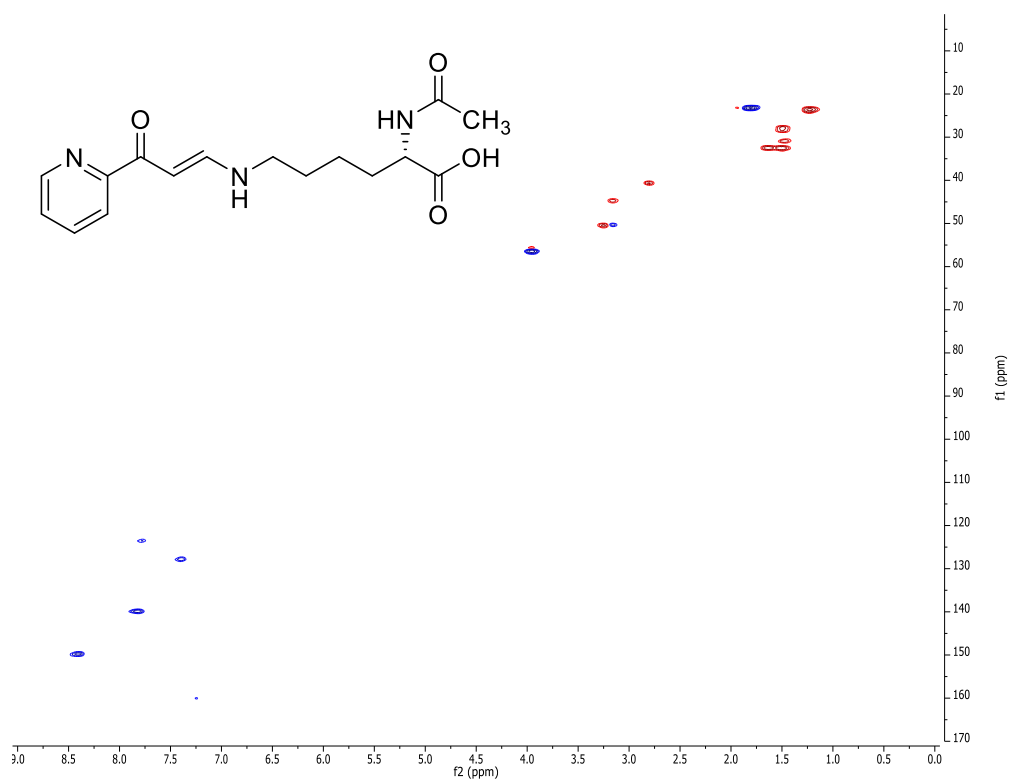
¹H NMR (500 MHz, phosphate-buffered D₂O) of **8**



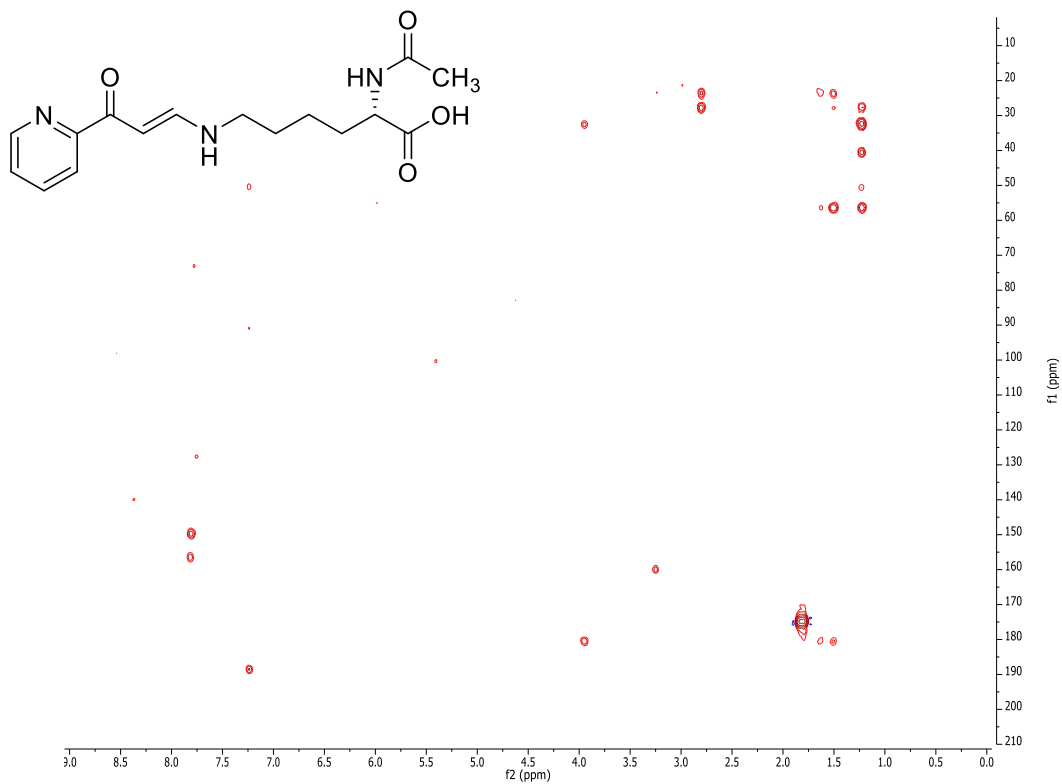
COSY NMR (500 MHz, phosphate-buffered D₂O) of **8**



HSQC NMR (500 MHz, phosphate-buffered D₂O) of **8**



HMBC NMR (500 MHz, phosphate-buffered D₂O) of **8**



¹H NMR (500 MHz, phosphate-buffered D₂O) spectrum of **6** and hydrolysis product **9**

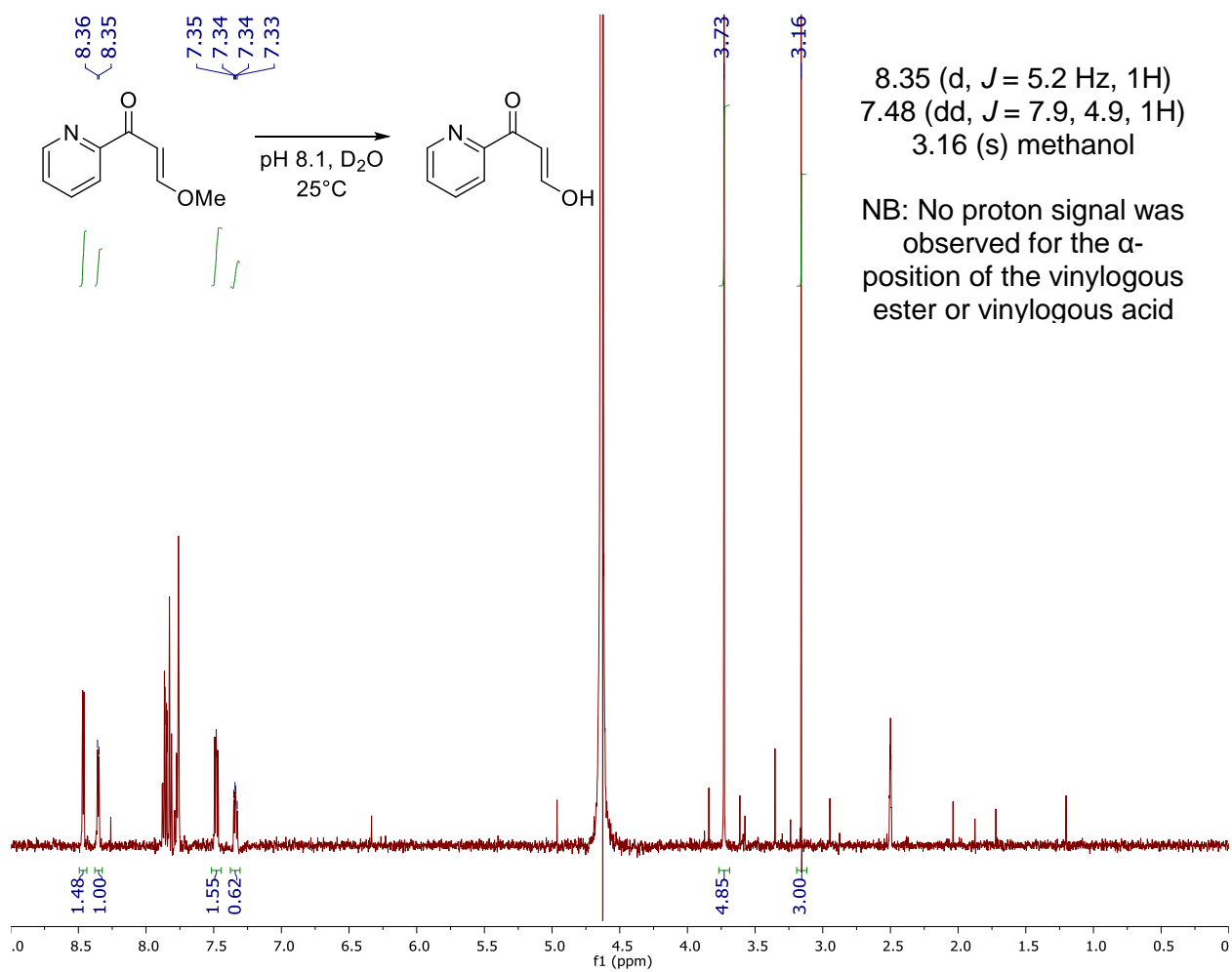


Table S10. Results of competition binding assay for neolymphostin against a panel of 97 kinases

Target	Neolymphostin A
Gene Symbol	%Ctrl @ 1000nM
ABL1(E255K)-phosphorylated	70
ABL1(T315I)-phosphorylated	75
ABL1-nonphosphorylated	39
ABL1-phosphorylated	76
ACVR1B	96
ADCK3	87
AKT1	100
AKT2	97
ALK	100
AURKA	1.5
AURKB	1.9
AXL	39
BMPR2	57
BRAF	77
BRAF(V600E)	77
BTK	15
CDK11	59
CDK2	87
CDK3	96
CDK7	50
CDK9	89
CHEK1	88
CSF1R	82
CSNK1D	75
CSNK1G2	33
DCAMKL1	83
DYRK1B	80
EGFR	58
EGFR(L858R)	57
EPHA2	96
ERBB2	28
ERBB4	90
ERK1	90
FAK	98
FGFR2	94
FGFR3	91
FLT3	74
GSK3B	89
IGF1R	96
IKK-alpha	78
IKK-beta	83
INSR	93
JAK2(JH1domain-catalytic)	62

%Ctrl Legend:

0≤x<.1	.1≤x<1	1≤x<10	10≤x<35	x≥35
--------	--------	--------	---------	------

Table S10. (continued)

Target	Neolymphostin A
Gene Symbol	%Ctrl @ 1000nM
JAK3(JH1domain-catalytic)	18
JNK1	0.25
JNK2	0.15
JNK3	0.2
KIT	84
KIT(D816V)	91
KIT(V559D,T670I)	73
LKB1	93
MAP3K4	88
MAPKAPK2	98
MARK3	100
MEK1	10
MEK2	12
MET	60
MKNK1	87
MKNK2	97
MLK1	8
p38-alpha	100
p38-beta	100
PAK1	100
PAK2	98
PAK4	100
PCTK1	70
PDGFRA	61
PDGFRB	95
PDPK1	100
PIK3C2B	0
PIK3CA	0
PIK3CG	0.55
PIM1	100
PIM2	100
PIM3	91
PKAC-alpha	89
PLK1	100
PLK3	76
PLK4	0.95
PRKCE	18
RAF1	90
RET	86
RIOK2	41
ROCK2	93
RSK2(Kin.Dom.1-N-terminal)	6.1
SNARK	27

%Ctrl Legend:

0≤x<.1	.1≤x<1	1≤x<10	10≤x<35	x≥35
--------	--------	--------	---------	------

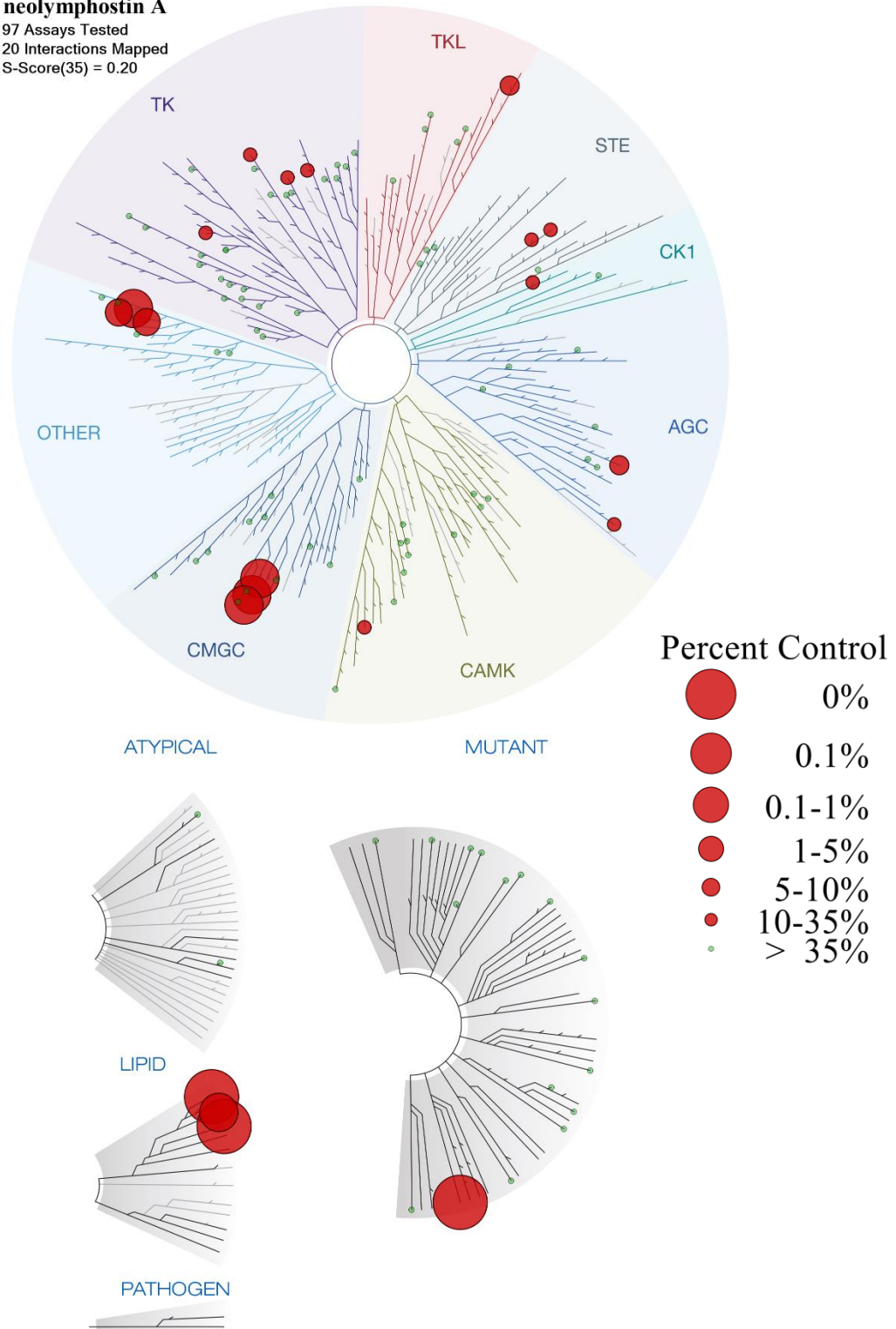
Table S10. (continued)

Target Gene Symbol	Neolymphostin A %Ctrl @ 1000nM
SRC	89
SRPK3	69
TGFBR1	97
TIE2	61
TRKA	12
TSSK1B	93
TYK2(JH1domain-catalytic)	60
ULK2	75
VEGFR2	68
YANK3	99
ZAP70	72

%Ctrl Legend:

$0 \leq x < .1$	$.1 \leq x < 1$	$1 \leq x < 10$	$10 \leq x < 35$	$x \geq 35$
-----------------	-----------------	-----------------	------------------	-------------

neolymphostin A
 97 Assays Tested
 20 Interactions Mapped
 S-Score(35) = 0.20



TREEspot™ is a proprietary data visualization software tool developed by KINOMEScan. Kinases found to bind are marked with red circles, where larger circles indicate higher-affinity binding.

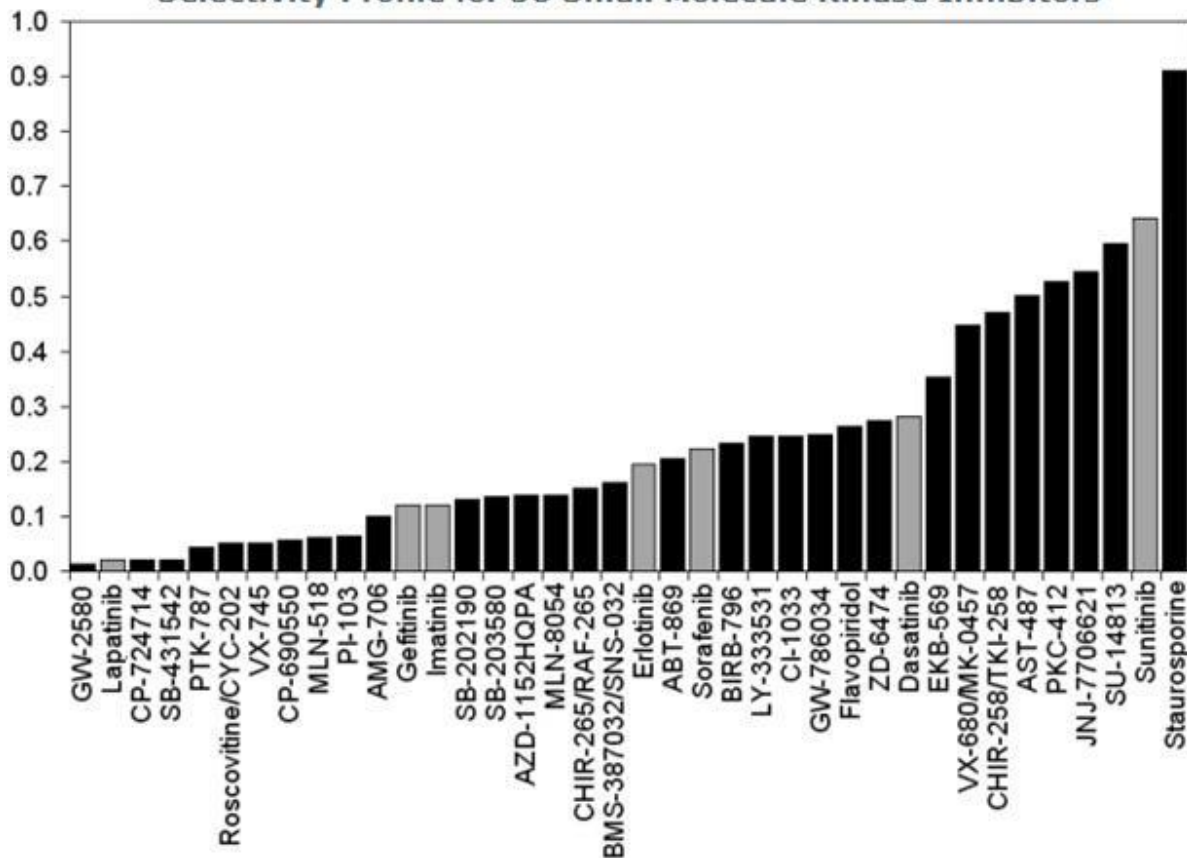
Figure S12. TREEspot™ interaction map for neolymphostin A @ 1000nM

Table S11. Selectivity S-Score for neolymphostin A against a 97 kinase panel

Compound Name	Selectivity Score Type	Number of Hits	Number of Non-Mutant Kinases	Screening Concentration (nM)	Selectivity Score
Neolymphostin A	S(35)	20	90	1000	0.222
Neolymphostin A	S(10)	11	90	1000	0.122
Neolymphostin A	S(1)	7	90	1000	0.078

Selectivity S35-Score for 38 kinase inhibitors against a panel of 287 kinases

Selectivity Profile for 38 Small Molecule Kinase Inhibitors



KINOMEScan's in vitro competition binding assay was used to evaluate 38 kinase inhibitors against a panel of 287 distinct human protein kinases (~55% of the predicted human protein kinome), and three lipid kinases. The compounds tested included 21 tyrosine kinase inhibitors, 15 serine-threonine kinase inhibitors, 1 lipid kinase inhibitor and staurosporine. $S(35) = (\text{number of non-mutant kinases with \%Ctrl} < 35) / (290 \text{ kinases tested}; 27 \text{ mutant variants were excluded from this analysis})$. Compounds approved for use in humans (as of August, 2007) are highlighted (gray bars).

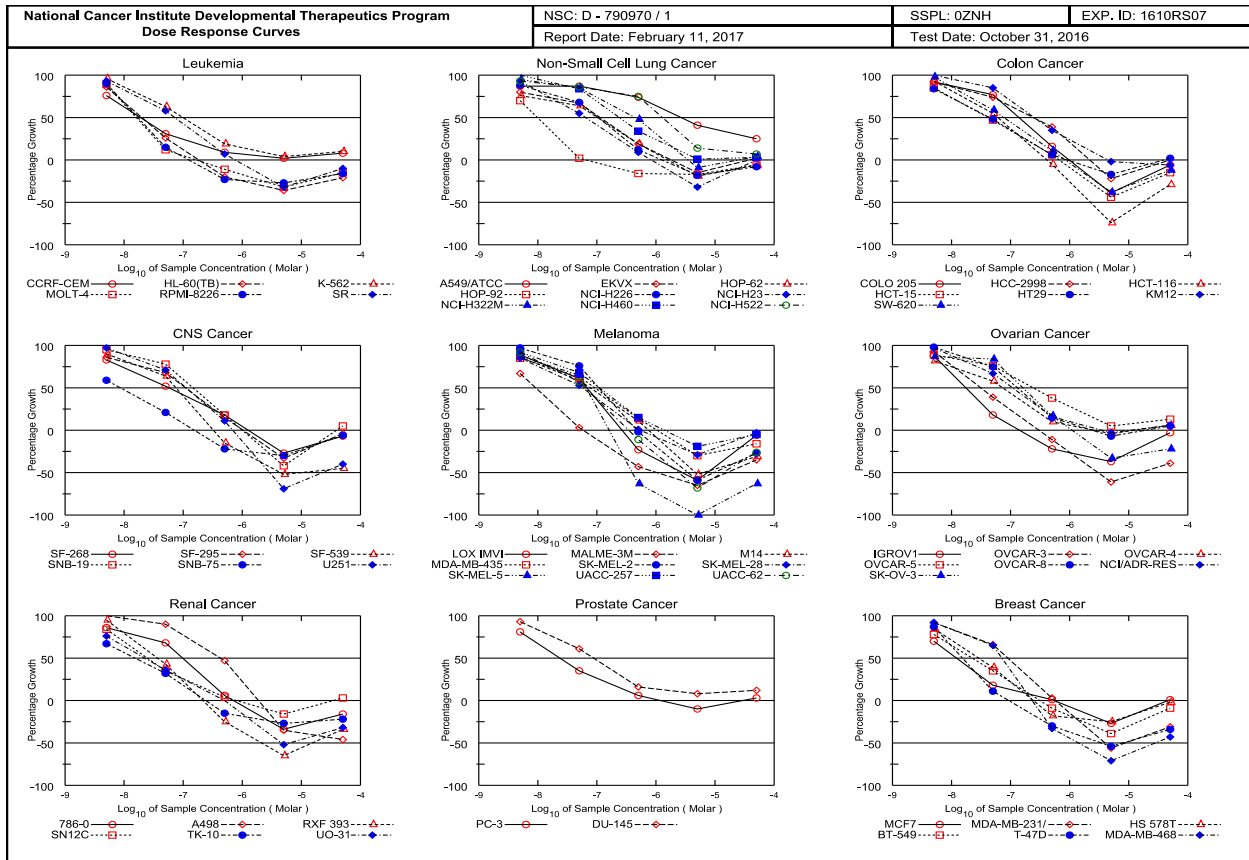


Figure S13. NCI 60-cell line dose response curves for neolymphostin A

**National Cancer Institute Developmental Therapeutics Program
In-Vitro Testing Results**

NSC : D - 790970 / 1		Experiment ID : 1610RS07					Test Type : 08					Units : Molar				
Report Date : February 11, 2017		Test Date : October 31, 2016					QNS :					MC :				
COMI : Y486.338		Stain Reagent : SRB Dual-Pass Related					SSPL : 0ZNH									
Panel/Cell Line	Time	Log10 Concentration						Percent Growth					GI50	TGI	LC50	
		Zero	Ctrl	-8.3	Mean Optical Densities	-6.3	-5.3	-4.3	-8.3	-7.3	-6.3	-5.3				-4.3
Leukemia																
CCRF-CEM	0.382	1.593	1.305	0.763	0.492	0.404	0.476	76	31	9	2	8	1.93E-8	> 5.00E-5	> 5.00E-5	
HL-60(TB)	0.809	3.021	2.720	1.384	0.649	0.521	0.640	86	26	-20	-36	-21	2.00E-8	1.84E-7	> 5.00E-5	
K-562	0.226	2.145	2.069	1.434	0.595	0.301	0.416	96	63	19	4	10	9.86E-8	> 5.00E-5	> 5.00E-5	
MOLT-4	0.563	2.041	1.872	0.737	0.503	0.383	0.477	89	12	-11	-32	-15	1.59E-8	1.66E-7	> 5.00E-5	
RPMI-8226	0.926	2.303	2.162	1.139	0.717	0.675	0.777	90	15	-23	-27	-16	1.71E-8	1.28E-7	> 5.00E-5	
SR	0.568	2.337	2.220	1.601	0.692	0.387	0.511	93	58	7	-32	-10	7.28E-8	7.57E-7	> 5.00E-5	
Non-Small Cell Lung Cancer																
A549(ATCC)	0.470	1.819	1.645	1.639	1.470	1.029	0.807	87	87	74	41	25	2.73E-6	> 5.00E-5	> 5.00E-5	
EKVX	0.853	2.298	2.016	1.824	1.121	0.725	0.882	80	67	19	-15	2	1.13E-7		> 5.00E-5	
HOP-62	0.945	2.068	1.793	1.668	1.156	0.770	0.892	76	64	19	-19	-6	1.03E-7	1.59E-6	> 5.00E-5	
HOP-92	1.002	1.574	1.405	1.014	0.837	0.834	0.939	70	2	-16	-17	-6	9.94E-9	6.42E-8	> 5.00E-5	
NCI-H226	0.768	1.285	1.222	1.122	0.832	0.627	0.705	88	68	12	-18	-8	1.06E-7	1.26E-6	> 5.00E-5	
NCI-H23	0.648	2.032	1.938	1.403	0.771	0.441	0.649	93	55	9	-32	.	6.29E-8		> 5.00E-5	
NCI-H322M	0.967	2.254	2.185	2.061	1.582	0.884	1.017	95	85	48	-9	4	4.36E-7		> 5.00E-5	
NCI-H460	0.183	1.894	1.889	1.617	0.771	0.195	0.241	100	84	34	1	3	2.41E-7	> 5.00E-5	> 5.00E-5	
NCI-H522	0.922	2.081	2.006	1.924	1.791	1.088	1.006	93	86	75	14	7	1.29E-6	> 5.00E-5	> 5.00E-5	
Colon Cancer																
COLO 205	0.580	1.888	1.785	1.587	0.787	0.357	0.546	92	77	16	-39	-6	1.38E-7	9.76E-7	> 5.00E-5	
HCC-2998	0.864	2.666	2.528	2.206	1.568	0.677	0.858	92	74	39	-22	.	2.45E-7	2.20E-6	> 5.00E-5	
HCT-116	0.252	2.145	2.016	1.296	0.240	0.065	0.179	93	55	-5	-74	-29	6.09E-8	4.16E-7	.	
HCT-15	0.330	2.224	1.941	1.213	0.442	0.185	0.281	85	47	6	-44	-15	4.08E-8	6.56E-7	> 5.00E-5	
HT29	0.228	1.362	1.181	0.778	0.293	0.189	0.256	84	48	6	-17	2	4.53E-8	.	> 5.00E-5	
KM12	0.566	2.852	2.866	2.518	1.358	0.555	0.535	101	85	35	-2	-6	2.49E-7	4.40E-6	> 5.00E-5	
SW-620	0.254	1.734	1.699	1.131	0.415	0.157	0.223	98	59	11	-38	-12	7.76E-8	8.33E-7	> 5.00E-5	
CNS Cancer																
SF-268	0.752	2.376	2.096	1.594	1.048	0.546	0.699	83	52	18	-27	-7	5.67E-8	1.25E-6	> 5.00E-5	
SF-295	0.611	2.825	2.510	2.123	0.937	0.408	0.585	86	68	15	-33	-4	1.10E-7	1.01E-6	> 5.00E-5	
SF-539	1.194	3.050	2.861	2.375	1.019	0.578	0.662	90	64	-15	-52	-45	7.46E-8	3.25E-7	.	
SNB-19	0.265	1.047	1.007	0.878	0.404	0.154	0.302	95	78	18	-42	5	1.47E-7	.	> 5.00E-5	
SNB-75	0.941	1.875	1.496	1.134	0.734	0.663	0.889	59	21	-22	-30	-6	8.74E-9	1.52E-7	> 5.00E-5	
U251	0.495	1.932	1.896	1.519	0.656	0.154	0.297	97	71	11	-69	-40	1.13E-7	6.90E-7	.	
Melanoma																
LOX IMVI	0.462	2.614	2.386	1.769	0.357	0.188	0.435	89	61	-23	-59	-6	6.72E-8	2.67E-7	.	
MALME-3M	0.588	1.044	0.892	0.602	0.333	0.204	0.385	67	3	-43	-65	-35	9.11E-9	5.79E-8	.	
M14	0.406	1.751	1.558	1.200	0.549	0.195	0.281	86	59	11	-52	-31	7.69E-8	7.39E-7	.	
MDA-MB-435	0.495	2.356	2.084	1.672	0.759	0.347	0.415	85	63	14	-30	-16	9.31E-8	1.05E-6	> 5.00E-5	
SK-MEL-2	1.000	2.074	2.037	1.816	0.982	0.412	0.729	97	76	-2	-59	-27	1.08E-7	4.74E-7	.	
SK-MEL-28	0.560	1.718	1.543	1.177	0.571	0.398	0.543	85	53	1	-29	-3	5.77E-8	5.38E-7	> 5.00E-5	
SK-MEL-5	0.727	2.593	2.345	1.932	0.270	-0.002	0.266	87	65	-63	-100	-63	6.50E-8	1.61E-7	3.96E-7	
UACC-257	1.025	1.941	1.866	1.648	1.161	0.834	0.969	92	68	15	-19	-5	1.09E-7	1.38E-6	> 5.00E-5	
UACC-62	0.899	2.891	2.729	2.025	0.799	0.292	0.663	92	57	-11	-68	-26	6.25E-8	3.42E-7	.	
Ovarian Cancer																
IGROV1	0.532	1.905	1.735	0.778	0.414	0.336	0.516	88	18	-22	-37	-3	1.73E-8	1.39E-7	> 5.00E-5	
OVCAR-3	0.441	1.409	1.341	0.821	0.393	0.170	0.271	93	39	-11	-61	-39	3.15E-8	3.03E-7	.	
OVCAR-4	0.705	1.594	1.438	1.219	0.798	0.680	0.759	82	58	10	-4	6	7.30E-8	.	> 5.00E-5	
OVCAR-5	0.733	2.242	2.077	1.876	1.304	0.807	0.932	89	76	38	5	13	2.39E-7	> 5.00E-5	> 5.00E-5	
OVCAR-8	0.522	2.037	2.009	1.658	0.747	0.487	0.592	98	75	15	-7	5	1.30E-7	.	> 5.00E-5	
NCI/ADR-RES	0.647	2.289	2.235	1.746	0.907	0.627	0.745	97	67	16	-3	6	1.07E-7	.	> 5.00E-5	
SK-OV-3	1.104	2.092	1.968	1.931	1.275	0.742	0.862	87	84	17	-33	-22	1.61E-7	1.11E-6	> 5.00E-5	
Renal Cancer																
786-0	0.636	2.392	2.154	1.836	0.736	0.419	0.533	86	68	6	-34	-16	9.81E-8	6.95E-7	> 5.00E-5	
A498	1.153	1.897	1.931	1.823	1.503	0.744	0.626	105	90	47	-35	-46	4.27E-7	1.86E-6	> 5.00E-5	
RXF 393	0.641	1.212	1.186	0.885	0.480	0.224	0.426	95	43	-25	-65	-34	3.65E-8	2.13E-7	.	
SN12C	0.522	2.076	1.835	1.068	0.588	0.438	0.564	84	35	4	-16	3	2.50E-8	.	> 5.00E-5	
TK-10	1.429	2.506	2.149	1.772	1.218	1.040	1.113	67	32	-15	-27	-22	1.51E-8	2.41E-7	> 5.00E-5	
UO-31	0.508	1.898	1.560	1.015	0.511	0.244	0.345	76	36	.	-52	-32	2.26E-8	5.04E-7	.	
Prostate Cancer																
PC-3	0.588	1.649	1.443	0.957	0.655	0.530	0.623	81	35	6	-10	3	2.32E-8	.	> 5.00E-5	
DJ-145	0.346	1.584	1.493	1.099	0.547	0.447	0.497	93	61	16	8	12	8.75E-8	> 5.00E-5	> 5.00E-5	
Breast Cancer																
MCF7	0.360	1.954	1.472	0.642	0.373	0.262	0.384	70	18	1	-27	1	1.20E-8	.	> 5.00E-5	
MDA-MB-231(ATCC 0.573)	1.251	1.199	1.023	0.595	0.255	0.394		92	66	3	-56	-31	9.08E-8	5.68E-7	.	
HS 578T	1.105	2.145	1.989	1.506	0.904	0.832	1.087	85	39	-18	-25	-2	2.83E-8	2.39E-7	> 5.00E-5	
BT-549	1.113	2.212	1.967	1.496	1.013	0.681	1.017	78	35	-9	-39	-9	2.21E-8	3.11E-7	> 5.00E-5	
T-47D	0.851	1.610	1.513	0.933	0.592	0.392	0.563	87	11	-30	-54	-34	1.53E-8	9.13E-8	.	
MDA-MB-468	0.913	1.855	1.778	1.525	0.616	0.268	0.524	92	65	-33	-71	-43	7.12E-8	2.32E-7	.	

Figure S14. NCI 60-cell line testing results for neolymphostin A

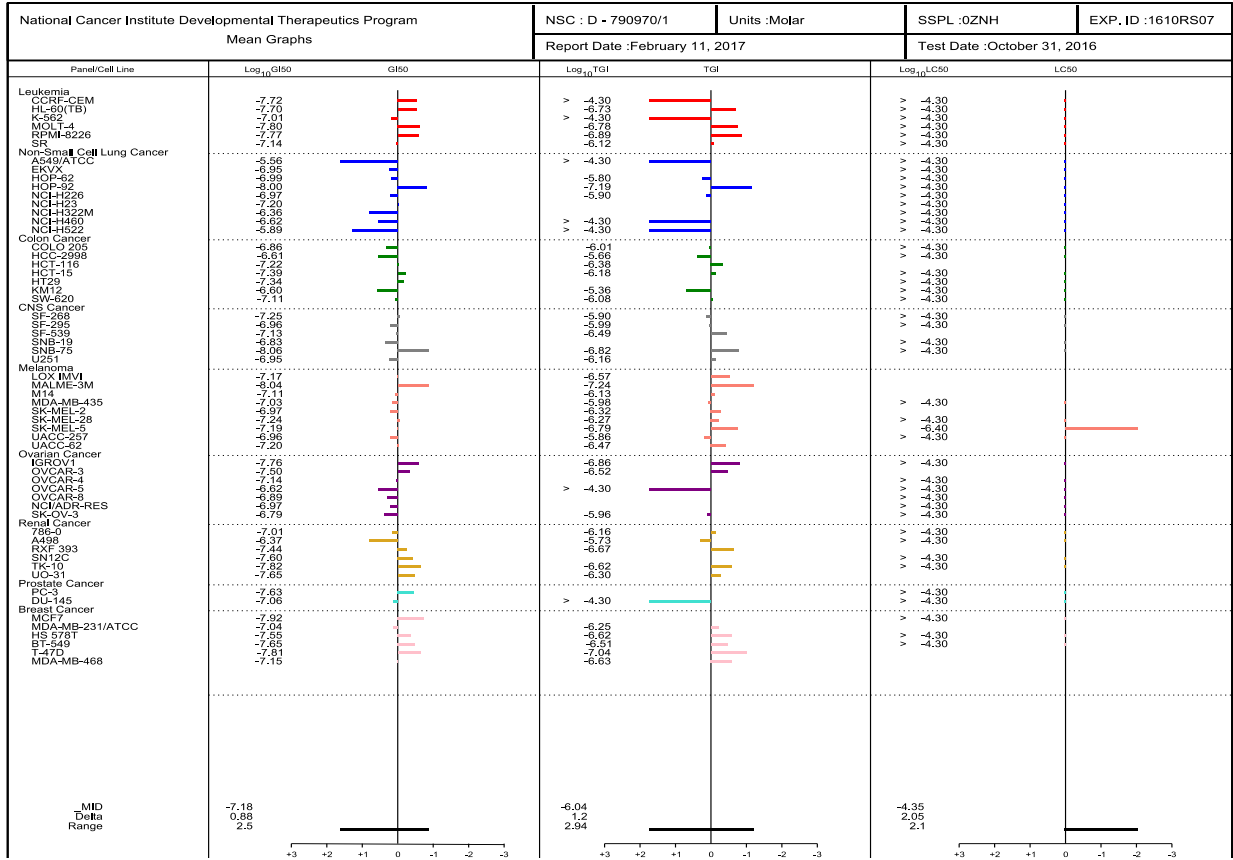


Figure S15. NCI 60-cell line mean graphs for neolymphostin A

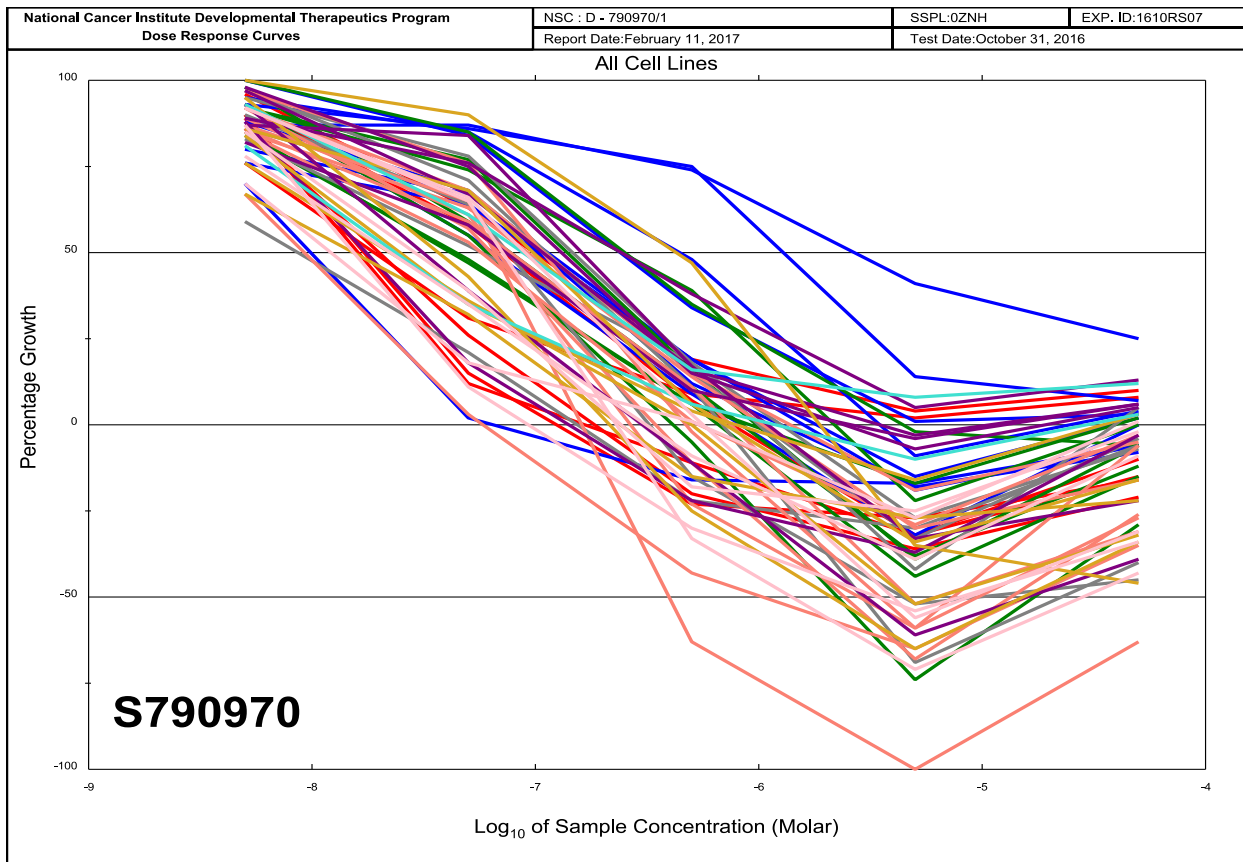


Figure S16. NCI 60-cell line dose response curves for neolymphostin A (all)

Compare results

STANDARD COMPARE Job Parameters

One item found.
1

Job Id	Status	Seed Vector descriptor For Display	Target Set Name	Minimum Standard Deviation	Minimum Correlation	Count Results To Return	Count Results To Return
3795231143875657337	COMPLETE	NEOLYMPHOSTIN A	STANDARD_AGENTS_GI50	0.05	0.2	50	50

STANDARD COMPARE Results

You can download the table below by using the buttons below the table.

Select/deselect all results on this page: [Select/Deselect All](#)

Crosstabulate the cell line data for the seed and target for the results selected below: [Crosstabulate test Results for selected vectors](#)

Use overlapping mean graphs to show the cell line data for the seed and target for the results selected below: [Graph test Results for selected vectors](#)

Graph gene distributions: [Select to graph distribution](#)

53 items found, displaying all items.
1

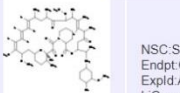




Rank	Correlation	namecode	EXTERNAL LINKS	Seed Vector ident For Display	Seed Vector descriptor For Display	Target Vector ident For Display	Target Vector descriptor For Display	Count Common Cell Lines	Seed Standard Deviation	Target Standard Deviation	
<input type="checkbox"/>	1	0.445	PUBLIC		NSC:S790970 Endpt:GI50 Expld:AVGDATA hiConc:-4.3	NEOLYMPHOSTIN A	NSC:S226080 Endpt:GI50 Expld:AVGDATA hiConc:-3.0	rapamycin	58	0.461	1.182
			DTP - chemical data DTP - all data								
<input type="checkbox"/>	2	0.429	PUBLIC		NSC:S790970 Endpt:GI50 Expld:AVGDATA hiConc:-4.3	NEOLYMPHOSTIN A	NSC:S280594 Endpt:GI50 Expld:AVGDATA hiConc:-3.3	tricirbine phosphate	58	0.461	0.864
			DTP - chemical data DTP - all data								
<input type="checkbox"/>	3	0.425	PUBLIC		NSC:S790970 Endpt:GI50 Expld:AVGDATA hiConc:-4.3	NEOLYMPHOSTIN A	NSC:S280594 Endpt:GI50 Expld:AVGDATA hiConc:-2.3	tricirbine phosphate	58	0.461	0.919
			DTP - chemical data DTP - all data								
<input type="checkbox"/>	4	0.419	PUBLIC		NSC:S790970 Endpt:GI50 Expld:AVGDATA hiConc:-4.3	NEOLYMPHOSTIN A	NSC:S280594 Endpt:GI50 Expld:AVGDATA hiConc:-4.0	tricirbine phosphate	51	0.477	0.647
			DTP - chemical data DTP - all data								
<input type="checkbox"/>	5	0.389	PUBLIC		NSC:S790970 Endpt:GI50 Expld:AVGDATA hiConc:-4.3	NEOLYMPHOSTIN A	NSC:S226080 Endpt:GI50 Expld:AVGDATA hiConc:-4.0	rapamycin	59	0.457	0.917
			DTP - chemical data DTP - all data								

Figure S17. COMPARE results

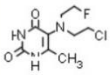
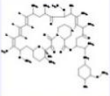
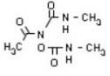
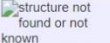
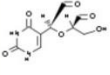
<input type="checkbox"/>	6	0.378	PUBLIC	 DTP - chemical data DTP - all data	NSC:S790970 Endpt:GI50 Expld:AVGDATA hiConc:-4.3	NEOLYMPHOSTIN A	NSC:S73754 Endpt:GI50 Expld:AVGDATA hiConc:-2.9	fluorodopan	43	0.492	0.195
<input type="checkbox"/>	7	0.367	PUBLIC	 DTP - chemical data DTP - all data	NSC:S790970 Endpt:GI50 Expld:AVGDATA hiConc:-4.3	NEOLYMPHOSTIN A	NSC:S226080 Endpt:GI50 Expld:AVGDATA hiConc:-7.0	rapamycin	59	0.457	0.778
<input type="checkbox"/>	8	0.358	PUBLIC	 DTP - chemical data DTP - all data	NSC:S790970 Endpt:GI50 Expld:AVGDATA hiConc:-4.3	NEOLYMPHOSTIN A	NSC:S253272 Endpt:GI50 Expld:AVGDATA hiConc:-2.0	caracemide	58	0.461	0.336
<input type="checkbox"/>	9	0.356	PUBLIC	 DTP - chemical data DTP - all data	NSC:S790970 Endpt:GI50 Expld:AVGDATA hiConc:-4.3	NEOLYMPHOSTIN A	NSC:S167780 Endpt:GI50 Expld:AVGDATA hiConc:-3.9	asaley	58	0.461	0.431
<input type="checkbox"/>	10	0.337	PUBLIC	 DTP - chemical data DTP - all data	NSC:S790970 Endpt:GI50 Expld:AVGDATA hiConc:-4.3	NEOLYMPHOSTIN A	NSC:S291643 Endpt:GI50 Expld:AVGDATA hiConc:-2.3	pyrimidine-5-glycodialdehyde	58	0.461	0.167

Figure S17. COMPARE results (continued)

Chapter 4

“Progress toward the total synthesis of the lymphostins: Preparation of a functionalized tetrahydropyrrolo[4,3,2-*de*]quinoline and unusual oxidative dimerization.”

Grant S. Seiler^{1,2} and Chambers C. Hughes^{1,*}

¹Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, United States

²Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, United States

*e-mail: chughes@ucsd.edu

Abstract

The lymphostins are a family of closely-related pyrrolo[4,3,2-*de*]quinoline natural products produced by *Streptomyces* and *Salinispora* actinobacteria. Neolymphostin A was recently been shown to strongly inhibit phosphoinositide 3-kinase (PI3K) and the mammalian target of rapamycin (mTOR) in a covalent manner via conjugation to a catalytic lysine residue in the ATP-binding pocket of the enzymes, making this metabolite the first reported covalent kinase inhibitor from a bacterium. A flexible and efficient synthetic route toward these alkaloids would allow for improvements in their solubility, stability, and selectivity and help to deliver a viable drug candidate. We have since established a short synthesis to methyl 8-bromo-1,3,4,5-tetrahydropyrrolo[4,3,2-*de*]quinoline-4-carboxylate via a conjugate addition/intramolecular Ullman reaction sequence. However, attempts to oxidize this intermediate to the pyrrolo[4,3,2-*de*]quinoline characteristic of the lymphostins resulted in formation of either a 2-oxo-1,2-dihydropyrrolo[4,3,2-*de*]quinoline or an unusual *N,C*-linked tetrahydropyrroloquinoline-pyrroloquinoline heterodimer. We expect that key modifications to the tetrahydropyrroloquinoline intermediate prior to oxidation should prevent these side reactions and pave the way for the completion of the synthesis.

Lymphostin (**1a**) was first isolated from the culture broths of *Streptomyces* sp. KY11783, and a description of its structure and kinase inhibitory activity against lymphocyte kinase appeared in 1997 (Figure 1).^{75,118} The natural product was later shown to inhibit phosphoinositide 3-kinase (PI3K) and the mammalian target of rapamycin (mTOR) in the low nanomolar range.^{74,119} The pyrrolo[4,3-*de*]quinoline-containing alkaloid has obvious structural similarities to adenine, and this feature contributes to its ability to compete for binding to the ATP-binding pocket of the kinases. Other, closely-related members of the lymphostin family, including the neolymphostins A-D (**1b-1e**), lymphostinol (**2a**), and the neolymphostinols A-D (**2b-2e**), have been reported from *Salinispora*.⁷⁴ Neolymphostin A was recently shown to block AKT phosphorylation in live cells with an IC₅₀ ~ 3 nM and displayed reasonably selective cytotoxicity in the low nanomolar range against leukemia and non-small cell lung cancer cell lines.¹²⁰

Although the structure and kinase inhibitory properties of the lymphostins have been known for some time, key information concerning their mechanism of action has been lacking. In 2018 we unequivocally demonstrated that the lymphostins, unlike any other bacterial natural products to date, are covalent kinase inhibitors and that the natural product-kinase interaction relies on an unusual electrophilic vinylogous ester at C-4.¹²⁰ Given this fundamental mechanistic discovery and knowing the critical role that the PI3Ks and mTOR play in cancer, immunodeficiencies, and metabolic disorders, we have since directed our efforts toward the development of a lymphostin-based drug candidate using chemical synthesis, one that particularly allows for flexibility in terms of the electrophilic C-4 warhead.¹²¹ While there is one reported synthesis of lymphostin, this route suffers from several drawbacks including its length (21 steps), overall yield (~2%), and use of toxic chemicals [e.g., Tl(OCOCF₃)₃],⁷⁷ such that an improved synthesis seems possible given the size and complexity of the molecule.

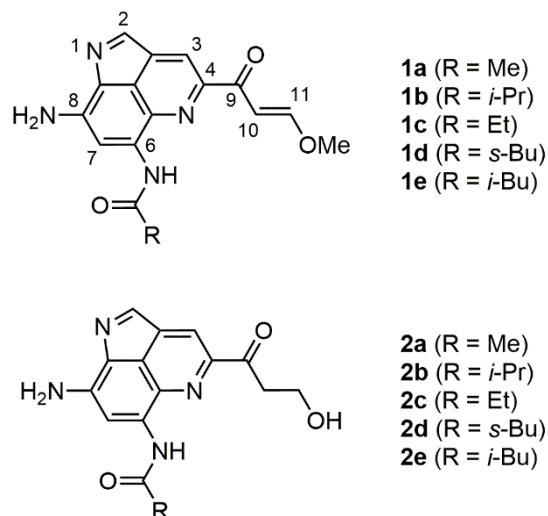
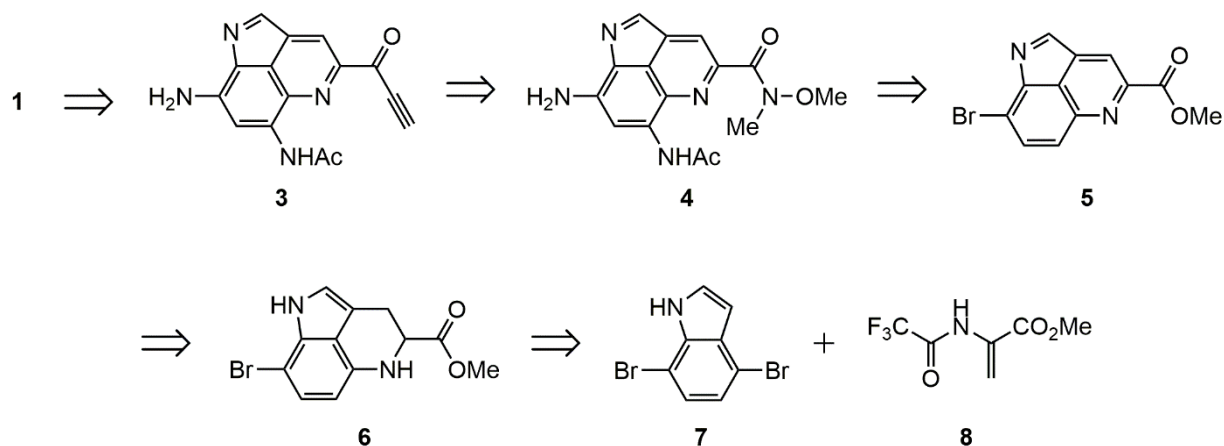


Figure 4.1. Structure of lymphostin (**1a**), the eponymous member of the lymphostin family of natural products, neolymphostins A-D (**1b-1e**), lymphostinol (**2a**), and the neolymphostinols A-D (**2b-2e**).

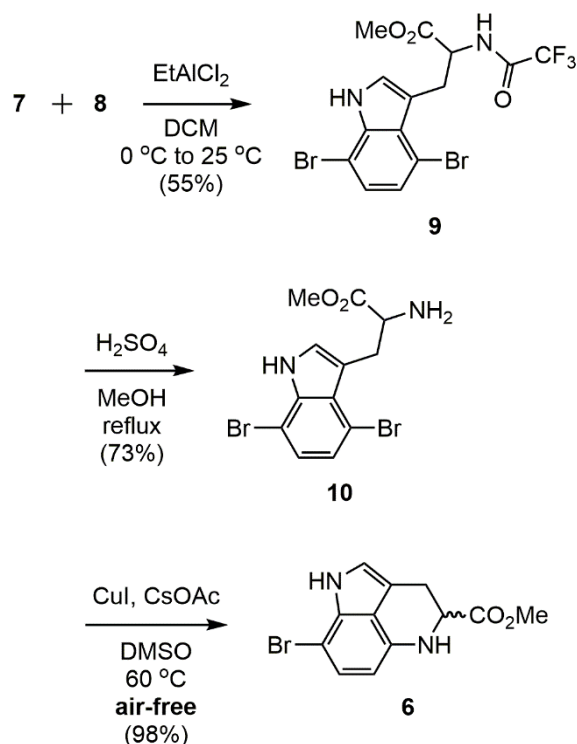
In our retrosynthetic analysis, we envisioned late-stage installment of the electrophilic vinylogous ester by first constructing ynone **3** via organometallic addition, followed by 1,4-addition of methanol (Scheme 1). Our work with a model system indicated that this approach would be both expedient and effective.¹²⁰ The requisite Weinreb amide (**4**) could be accessed by transformation of a methyl ester, and the nitrogenous groups at C-6 and C-8 could be installed by nitration and transition metal-catalyzed amination, respectively, of pyrroloquinoline **5**. This brominated tricycle would be obtained from **6** via oxidation, which in turn would be rapidly constructed via Friedel-Crafts-type alkylation of 4,7-dibromoindole (**7**) with methyl 2-(2,2,2-trifluoroacetamido)acrylate (**8**) and subsequent Ullmann cyclization. Synthetic preparations for both precursors **7** and **8** have been reported in the literature.



Scheme 4.1. Retrosynthetic analysis of lymphostin (**1**).

The synthesis commenced with 4,7-dibromoindole (**7**), prepared on large scale from 1,4-dibromobenzene following nitration and Bartoli reaction (Scheme 2).¹²² Using the method of Angelini, et al., we produced **9** in 55% yield via conjugate addition with *N*-trifluoroacetyl acrylate **8** in the presence of ethylaluminum dichloride.¹²³ Once the amine was liberated under acidic conditions, we set about trying to cleanly cyclize **10** in an intramolecular Ullman reaction. Though a sluggish reaction requiring heat and prolonged reaction times, the cyclization of similar *N*_α-acyl tryptophan derivatives has been reported using copper iodide in 1) DMSO with cesium acetate and 2) in dioxane with potassium carbonate.^{124,125} Also reported is the palladium-catalyzed cyclization of the terminal amino group of a dipeptide tryptophan derivative to form a nine-membered ring toward a synthesis of indolactam V.¹²⁵ This reaction also required heating to high temperatures (110 °C). To the best of our knowledge, direct cyclization of 4-bromo tryptophans such as **10** to the corresponding tetrahydropyrroloquinolines has not been disclosed. Notably, palladium-catalyzed cyclizations of *N*_α-methyl 4-iodo tryptamine derivatives have been reported by Buchwald, but this transformation is not directly amenable to a synthesis of lymphostin.¹²⁶ Gratifyingly, when amine **10** was heated with a CsOAc and a stoichiometric amount of CuI in DMSO, conversion to the cyclized target was detected by HPLC-MS, albeit in low yield and accompanied by extensive decomposition. After optimizing the reaction conditions, we found that if the

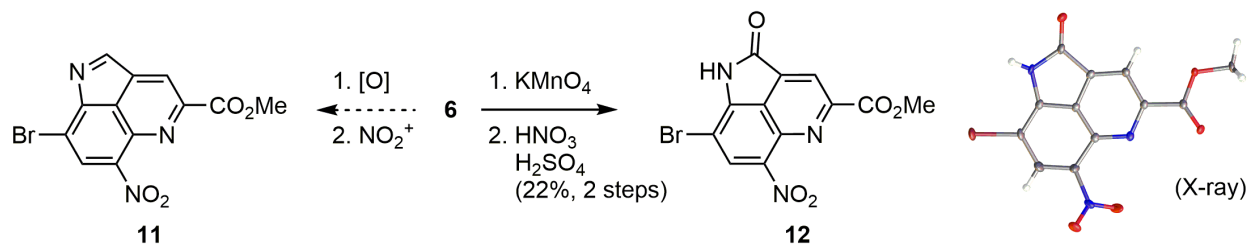
reaction was performed under air- and water-free conditions, **10** was converted to **6** in remarkable yield after several hours at 60 °C. This fundamental transformation represents a direct and efficient construction of the tricyclic skeleton shared by the pyrroloquinoline class of natural products, including the damirones, batzellines, isobatzellines, and makaluvamines.¹²⁷



Scheme 4.2. Synthesis of 1,3,4,5-tetrahydropyrrolo[4,3,2-*de*]quinoline **6** via intramolecular Ullman reaction.

As it was envisioned that tetrahydropyrroloquinoline **6** would undergo facile aromatization and nitration to pyrroloquinoline **11**, we first subjected this intermediate to a variety of oxidizing reagents (Scheme 3). Phenyliodine(III) diacetate (PIDA), I_2 , ceric ammonium nitrate (CAN), oxone, and 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) either failed to react or gave complex mixtures of uncharacterized products. However, treatment with KMnO_4 at $25\text{ }^\circ\text{C}$ and NaClO_2 at $60\text{ }^\circ\text{C}$ led to formation of the corresponding 2-oxo-1,2-dihydropyrroloquinoline in good yield. The oxidized product was further nitrated to **12** prior to its characterization in an effort to functionalize C-6 according to the

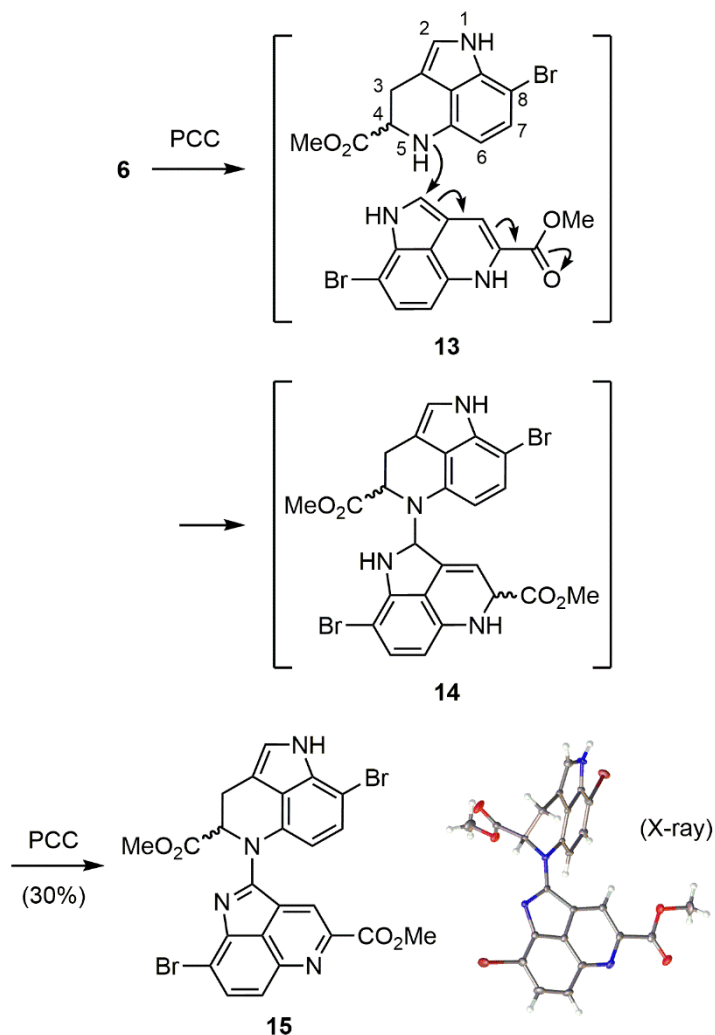
retrosynthetic scheme and to improve crystallinity. Since this intermediate contained very few protons to allow for unambiguous NMR characterization, we verified its structure using single crystal X-ray diffraction (Scheme 3) (CCDC 1888526). Unfortunately, attempts at reductive deoxygenation of lactam **12** using LiBH₄, TiCl₄ with SnCl₂, Tf₂O with Et₃SiH, and DIBAL-H all failed to afford **11**, and so we did not consider this route worth pursuing any further.



Scheme 4.3. Oxidation and nitration of quinoline **6** to 2-oxo-1,2-dihydropyrrolo-quinoline **12**. The ORTEP drawing of **12** shows thermal ellipsoids at the 50% probability level.

Treatment of intermediate **6** with other oxidants gave completely different results. Reaction with PCC resulted in the formation of an unexpected di-brominated [*m/z* (M+H)⁺ = 583, 585, 587 (1:2:1)], high-molecular weight species (**15**) with the deep purple color characteristic of a pyrroloquinoline (Scheme 4). Treatment of **6** with chloranil gave the same product, but conversion was much lower. Again, the structure was difficult to determine using NMR spectroscopy alone, owing in part to its poor solubility in most solvents, so crystals of **15** were grown from a concentrated toluene solution and then analyzed using X-ray crystallographic techniques (CCDC 1888527). This process revealed the structure of a compound with an unusual heterodimeric *N,C*-coupled hexacyclic ring system. Although **15** adopts a twisted conformation in the solid state, it does not exhibit atropisomerism in solution at 25 °C, which would presumably produce a noticeable diastereomeric mixture. We propose that conjugate addition of **6** onto the electrophilic C-2 position of oxidized $\alpha,\beta,\gamma,\delta$ -unsaturated ester **13** yielded a heterodimer (**14**) that was further oxidized with PCC to **15**. Efforts to temporarily intercept electrophile **13** with non-substrate nucleophiles such as amines, anilines, and phosphines were not

fruitful. Interestingly, the tetrahydropyrroloquinoline subunit of dimer **15** was not further oxidized by PCC, either *in situ*, or after isolation and re-subjection to the reaction conditions.



Scheme 4.4. Oxidation of tetrahydropyrroloquinoline **6** to heterodimer **15**. The ORTEP drawing of **15** shows thermal ellipsoids at the 50% probability level.

Current efforts are devoted to modification of tetrahydropyrroloquinoline **6** in order to prevent oxidation-induced dimerization and promote formation of the desired pyrrolo[4,3,2-*de*]quinoline ring system. For example, reduction of the C-4 carboxylate may prevent formation of the electrophilic species (**13**) that leads to dimerization. Alternatively, dimerization could be blocked by protection of N-5

in **6** prior to oxidation, although we realize that protection of N-5 may inadvertently preclude oxidation to the pyrroloquinoline, as seen with **15**.

Conclusions

Our proposed synthesis of the lymphostins involves the oxidation of a substituted indole intermediate, such as 1,3,4,5-tetrahydropyrrolo[4,3,2-de]quinoline **6**, into a pyrrolo[4,3,2-de]quinoline. However, this approach may be problematic given that indoles are well-known to oxidize at C-2 with a number of oxidants and give the corresponding oxindoles. Although we were able to avoid this reactivity using some oxidants like PCC, formation of an indole with electrophilic character at C-2 due to a vinylogous electron-withdrawing group (EWG) at C-3 then became the overriding issue. Simpler indoles substituted at C-3 with EWGs have been shown to react with nucleophiles at C-2, and this reactivity has been widely exploited.¹²⁸ Despite this, synthesis of 1,3,4,5-tetrahydropyrrolo[4,3,2-de]quinolines via conjugate addition/intermolecular Ullman reaction as outlined in Scheme 2 is novel and may be advantageous in other contexts, and the production of **15**, though unexpected, is a fascinating transformation in its own right giving rise to a highly complex molecular architecture.

Experimental section

General. Reactions and compounds were analyzed with an analytical 1100 series Agilent Technologies HPLC system coupled to an ELSD and UV/vis detector (210, 254, and 360 nm) using a Phenomenex Luna reversed-phase C18(2) column (100 mm × 4.6 mm, 5 μm, 100 Å) with a 10 min solvent gradient from 10% to 100% containing 0.1% formic acid and a flow rate of 1.0 mL min⁻¹. Using the same column and solvent gradients, liquid chromatography/high-resolution mass spectrometry was performed on an analytical Agilent 1260 Infinity series LC system coupled to a 6530 series Q-TOF mass spectrometer. Column chromatography was performed on a Teledyne CombiFlash Rf+ Lumen flash

chromatography system. 4,7-Dibromoindole (**7**) was prepared via nitration of 1,4-dibromobenzene and subsequent Bartoli reaction with vinylmagnesium bromide.¹²² Methyl 2-(2,2,2-trifluoroacetamido)acrylate (**8**) was prepared via reaction of L-serine methyl ester hydrochloride with trifluoroacetic anhydride and triethylamine.¹²⁹ All other reagents and solvents were purchased commercially and were used without further purification. ¹H NMR spectra were recorded at 500 MHz in CDCl₃ (residual solvent referenced to 7.26 ppm), DMSO-*d*₆ (2.50 ppm) or pyridine-*d*₅ (8.74 ppm) and ¹³C NMR spectra were recorded at 125 MHz in CDCl₃ (referenced to 77.2 ppm), DMSO-*d*₆ (39.5 ppm) or pyridine-*d*₅ (150.3 ppm), on a Jeol 500 MHz NMR spectrometer. IR spectra were recorded on a Nicolet 100 FT-IR.

Methyl 3-(4,7-dibromo-1H-indol-3-yl)-2-(2,2,2-trifluoroacetamido)propanoate (9). To 4,7-dibromoindole (**5**) (1.69 g, 6.15 mmol, 1.0 Eq.) and methyl 2-(2,2,2-trifluoroacetamido)acrylate (**6**) (1.45 g, 7.37 mmol, 1.2 Eq.) under nitrogen at 25 °C was added DCM (10 mL). The mixture was cooled to 0 °C, and a solution of EtAlCl₂ in toluene (4.7 mL, 25 wt.%, 8.6 mmol, 1.4 Eq.) was added dropwise via syringe. The reaction mixture was allowed to slowly reach 25 °C over 12 h, and then poured into a saturated aq. NaHCO₃ solution and extracted with EtOAc. The organic layer was washed with water, brine, dried over Na₂SO₄, filtered, and concentrated. The product was evaporated onto Celite and purified by flash chromatography (0-30% EtOAc in hexanes) to give **9** (1.6 g, 55% yield) as a tan solid. UV/Vis: λ_{max} = 226, 290 nm; IR (film): $\tilde{\nu}$ = 3320, 3100, 2960, 1716, 1705, 1562, 1542 cm⁻¹; ¹H NMR (CDCl₃): δ 8.41 (s, 2H), 7.21 (d, *J* = 8.4 Hz, 1H), 7.19 (d, *J* = 8.4 Hz, 1H), 7.15 (d, *J* = 2.3 Hz, 1H), 6.89 (d, *J* = 7.0 Hz, 1H), 5.01 (m, 1H), 3.78 (s, 3H), 3.76 (dd, *J* = 14.1, 5.0 Hz, 1H), 3.47 (dd, *J* = 14.1, 8.9 Hz, 1H); ¹³C{¹H} NMR (CDCl₃): δ 171.0, 156.9 (q, *J* = 37.8 Hz), 136.0, 126.1, 125.7, 125.6, 125.5, 115.7 (q, *J* = 288 Hz), 113.1, 112.3, 104.7, 54.2, 53.0, 28.1. HR-ESI-TOFMS: *m/z* (M+H)⁺ calcd for C₁₄H₁₂Br₂F₃N₂O₃ 470.9161, found 470.9165.

Methyl 2-amino-3-(4,7-dibromo-1H-indol-3-yl)propanoate (10). To a solution of amide **9** (1.6 g, 3.4 mmol) in MeOH (16 mL) at 25 °C was added conc. H₂SO₄ (8 mL) dropwise. The mixture was heated at

reflux for 2 h, and then allowed to cool to 25 °C, poured into water, and washed thrice with DCM. The solution was brought to pH 8 by addition of solid Na₂CO₃, at which point a white precipitate formed, and then extracted thrice with DCM. The combined organic extracts were washed with brine, dried over Na₂SO₄, filtered, and concentrated to give **10** (930 mg, 73% yield) as a white solid. UV/Vis: λ_{max} = 228, 290 nm. IR (film): $\tilde{\nu}$ = 3361, 3216, 2948, 1740, 1533 cm⁻¹; ¹H NMR (CDCl₃): δ 8.44 (s, 1H), 7.16 (s, 3H), 3.96 (dd, *J* = 8.5, 5.5 Hz, 1H), 3.72 (s, 3H), 3.59 (dd, *J* = 14.6, 5.5 Hz, 1H), 3.06 (dd, *J* = 14.5, 8.5 Hz, 1H); ¹³C{¹H} NMR (CDCl₃): δ 175.9, 136.2, 126.4, 125.7, 125.4, 125.3, 114.2, 113.6, 104.5, 56.0, 52.2, 31.8. HR-ESI-TOFMS: *m/z* (M+H)⁺ calcd for C₁₂H₁₃Br₂N₂O₂ 374.9338, found 374.9336.

Methyl 8-bromo-1,3,4,5-tetrahydropyrrolo[4,3,2-*de*]quinoline-4-carboxylate (6). A vial containing CuI (51 mg, 270 μmol, 1 Eq.) and CsOAc (204 mg, 1.06 mmol, 4 Eq.) was placed under high vacuum at 60°C with stirring for 1 hr. The vial was then allowed to cool to 25 °C and backfilled with nitrogen. A solution of amine **10** (100 mg, 270 μmol, 1 Eq.) in dry DMSO (2.7 mL) under nitrogen was transferred to the vial containing CuI and CsOAc via syringe. The reaction mixture was stirred at 60°C under nitrogen for 3 hr, and then allowed to cool to 25 °C and poured into EtOAc. The organic phase was washed successively with a saturated aqueous NaHCO₃ solution, water, and brine, dried over Na₂SO₄, filtered, and concentrated to give tricycle **6** (78 mg, 98% yield) as a purple foam. This material was of sufficient purity to use in subsequent reactions. When kept in a freezer and under an atmosphere of nitrogen, this material was stable for prolonged periods of time. Exposed to air, it gradually oxidized to dimer **15**. A small portion was purified by flash chromatography (0-50% EtOAc in hexanes) to obtain an analytically pure sample. UV/vis.: λ_{max} = 228, 280, 306 nm; IR (film) $\tilde{\nu}$ = 3381, 3120, 2948, 1732, 1506 cm⁻¹; ¹H NMR (CDCl₃): δ 8.03 (s, 1H), 7.11 (d, *J* = 7.8 Hz, 1H), 6.78 (s, 1H), 6.25 (d, *J* = 7.8 Hz, 1H), 4.22 (dd, *J* = 8.2, 4.6 Hz, 1H), 3.76 (s, 3H), 3.38 (dd, *J* = 15.5, 4.6 Hz, 1H), 3.22 (dd, *J* = 15.4, 8.2 Hz, 1H). ¹³C{¹H} NMR (CDCl₃): δ 172.9, 138.6, 133.1, 126.0, 118.5, 116.6, 109.7, 102.5, 92.6, 55.5, 52.6, 26.0. HR-ESI-TOFMS: *m/z* (M+H)⁺ calcd for C₁₂H₁₂BrN₂O₂ 295.0077, found 295.0073.

Methyl 8-bromo-6-nitro-2-oxo-1,2-dihydropyrrolo[4,3,2-de]quinoline-4-carboxylate (12). To a solution of tricycle **6** (114 mg, 386 μmol , 1 Eq.) in DMF (1.5 mL) at 25 °C was added KMnO_4 (61 mg, 390 μmol , 1 Eq.) portionwise. After 30 min the reaction mixture was diluted with EtOAc and washed with water. The aqueous phase was back-extracted with DCM. The combined organic extracts were washed with brine, dried over Na_2SO_4 , filtered, and concentrated. To a solution of the resulting brown solid (50 mg) in conc. H_2SO_4 (1 mL) at 0 °C was slowly added conc. HNO_3 (10 drops). After 30 min at 0 °C, ice was added to quench the reaction. This mixture was poured into EtOAc, and then washed successively with water, a saturated aq. NaHCO_3 solution, water, and brine. The organic phase was dried over Na_2SO_4 , filtered, and concentrated to give **12** (30 mg, 22% yield over 2 steps) as a red-orange residue. UV/Vis: $\lambda_{\text{max}} = 284, 360, 414 \text{ nm}$; IR (film): $\tilde{\nu} = 3220, 3072, 2921, 1739, 1636, 1533, 1348 \text{ cm}^{-1}$; $^1\text{H NMR}$ ($\text{DMSO-}d_6$): δ 12.19 (s, 1H), 8.72 (s, 1H), 8.47 (s, 1H), 4.00 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR ($\text{DMSO-}d_6$): δ 167.7, 164.4, 153.1, 145.0, 139.6, 136.3, 136.0, 135.5, 122.2, 119.2, 98.0, 53.3; HR-ESI-TOFMS: m/z (M-H) $^-$ calcd for $\text{C}_{12}\text{H}_5\text{BrN}_3\text{O}_5$ 349.9418, found 349.9416. The crystal structure data was deposited at the Cambridge Crystallographic Data Centre as CCDC 1888526.

Dimethyl 8,8'-dibromo-3',4'-dihydro-1'H-[2,5'-bipyrrolo[4,3,2-de]quinoline]-4,4'-dicarboxylate (15). To a solution of tricycle **6** (75 mg, 250 μmol , 1 Eq.) in DMF (1 mL) at 25 °C was added a solution of pyridinium chlorochromate (60 mg, 280 μmol , 1.1 Eq.) in DMF (1 mL) dropwise. After 15 min the reaction mixture was poured into water, and extracted once with EtOAc and once with DCM. The combined organic extracts were washed with brine, dried over Na_2SO_4 , filtered, and concentrated. The product was evaporated onto Celite and purified by flash chromatography (0-100% of 1:1 DCM:EtOAc in hexanes) to give **15** as a purple solid (22 mg, 30% yield). UV/Vis: $\lambda_{\text{max}} = 222, 286, 365, 530$; IR (film): $\tilde{\nu} = 3361, 3010, 2948, 1732, 1616, 1519 \text{ cm}^{-1}$; $^1\text{H NMR}$ ($\text{C}_5\text{D}_5\text{N}$): δ 12.80 (s, 1H), 8.92 (s, 1H), 7.81 (d, $J = 10.0 \text{ Hz}$, 1H), 7.79 (d, $J = 10.0 \text{ Hz}$, 1H), 7.55 (d, $J = 7.9 \text{ Hz}$, 1H), 7.37 (s, 1H), 6.66 (d, $J = 4.3 \text{ Hz}$, 1H), 3.97 (d, $J = 16.0 \text{ Hz}$, 1H), 3.95 (s, 3H), 3.82 (dd, $J = 16.0, 6.4 \text{ Hz}$, 1H), 3.42 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR ($\text{C}_5\text{D}_5\text{N}$): δ 170.7,

165.9, 163.8, 151.4, 144.6, 140.7, 137.9, 135.9, 135.0, 134.7, 131.5, 125.7, 125.0, 123.9, 122.1, 121.4, 120.7, 113.5, 107.1, 100.7, 61.2, 52.8, 52.4, 26.1. HR-ESI-TOFMS: m/z (M+H)⁺ calcd for C₂₄H₁₇Br₂N₄O₄ 582.9611, found 582.9605. The crystal structure data was deposited at the Cambridge Crystallographic Data Centre as CCDC 1888527.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: XXXXXX
¹H and ¹³C NMR spectra, HRMS data, and X-ray diffraction data.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by seed funding from the Scripps Institution of Oceanography (E.W. Scripps Associates), a UCSD Academic Senate Research Award (RP42S-HUGHES), and an American Cancer Society Institutional Research Grant (ACS-IRG 15-172-45) to C.H. We also thank the NIH for an HRLC-MS instrument (S10 OD0106400).

Chapter 4 is a reprint, in full, of published work. Grant S. Seiler and Chambers C. Hughes.

“Progress toward the total synthesis of the lymphostins: Preparation of a functionalized tetrahydropyrrolo[4,3,2-*de*]quinoline and unusual oxidative dimerization.” *Journal of Organic Chemistry*. **2019**, *84*, 9339-9343. The dissertation author was the primary investigator and author of this paper.

Notes and references

- (1) Aotani, Y.; Nagata, H.; Yoshida, M. *J. Antibiot. (Tokyo)*. **1997**, *50*, 543–545.
- (2) Nagata, H.; Ochiai, K.; Aotani, Y.; Ando, K.; Yoshida, M.; Takahashi, I. *J. Antibiot. (Tokyo)*. **1997**, *50*, 537–542.
- (3) Nagata, H.; Yano, H.; Sasaki, K.; Sato, S.; Nakanishi, S.; Takahashi, I.; Tamaoki, T. *Biosci. Biotechnol. Biochem.* **2002**, *66*, 501–507.
- (4) Miyanaga, A.; Janso, J. E.; McDonald, L.; He, M.; Liu, H.; Barbieri, L.; Eustáquio, A. S.; Fielding, E. N.; Carter, G. T.; Jensen, P. R.; Feng, X.; Leighton, M.; Koehn, F. E.; Moore, B. S. *J. Am. Chem. Soc.* **2011**, *133*, 13311–13313.
- (5) Castro-Falcón, G.; Seiler, G.; Demir, O.; Rathinaswamy, M.; Hamelin, D.; Hoffmann, R. M.; Makowski, S.; Letzel, A.-C.; Field, S.; Burke, J.; Amaro, R. E.; Hughes, C. C. *J. Med. Chem.* **2018**, *61*, 10463–10472.
- (6) Fuman, D. A.; Chiu, H.; Hopkins, B. D.; Bagrodia, S.; Cantley, L. C.; Abraham, R. T. *Cell*. **2017**, *170*, 605–635.
- (7) Tatsuta, K.; Imamura, K.; Itoh, S.; Kasai, S. *Tetrahedron Lett.* **2004**, *45*, 2847–2850.
- (8) Dobbs, A. P.; Voyle, M.; Whittall, N. *Synlett*, **1999**, *10*, 1594–1596.
- (9) Angelini, E.; Balsamini, C.; Bartoccini, F.; Lucarini, S.; Piersanti, G. *J. Org. Chem.* **2008**, *73*, 5654–5657.
- (10) Lim, H. J.; Gallucci, J. C.; Rajanbabu, T. V. *Org. Lett.* **2010**, *12*, 2162–2165.
- (11) Mari, M.; Bartoccini, F.; Piersanti, G. *J. Org. Chem.* **2013**, *78*, 7727–7734.
- (12) Peat, A. J.; Buchwald, S. L. *J. Am. Chem. Soc.* **1996**, *118*, 1028–1030.
- (13) Antunes, E. M.; Copp, B. R.; Davies-Coleman, M. T.; Samaai, T. *Nat. Prod. Rep.* **2005**, *22*, 62–72.
- (14) Bandini, M. *Org. Biomol. Chem.* **2013**, *11*, 5206–5212.
- (15) Avenozza, A.; Busto, J. H.; Canal, N.; García, J. I.; Jiménez-Osés, G.; Peregrina, J. M.; Pérez-Fernández, M. *New J. Chem.* **2007**, *31*, 224–229.

Supporting Information

Progress toward the total synthesis of the lymphostins: Preparation of a functionalized tetrahydropyrrolo[4,3,2-*de*]quinoline and unusual oxidative dimerization

Grant S. Seiler^{1,2} and Chambers C. Hughes^{1,*}

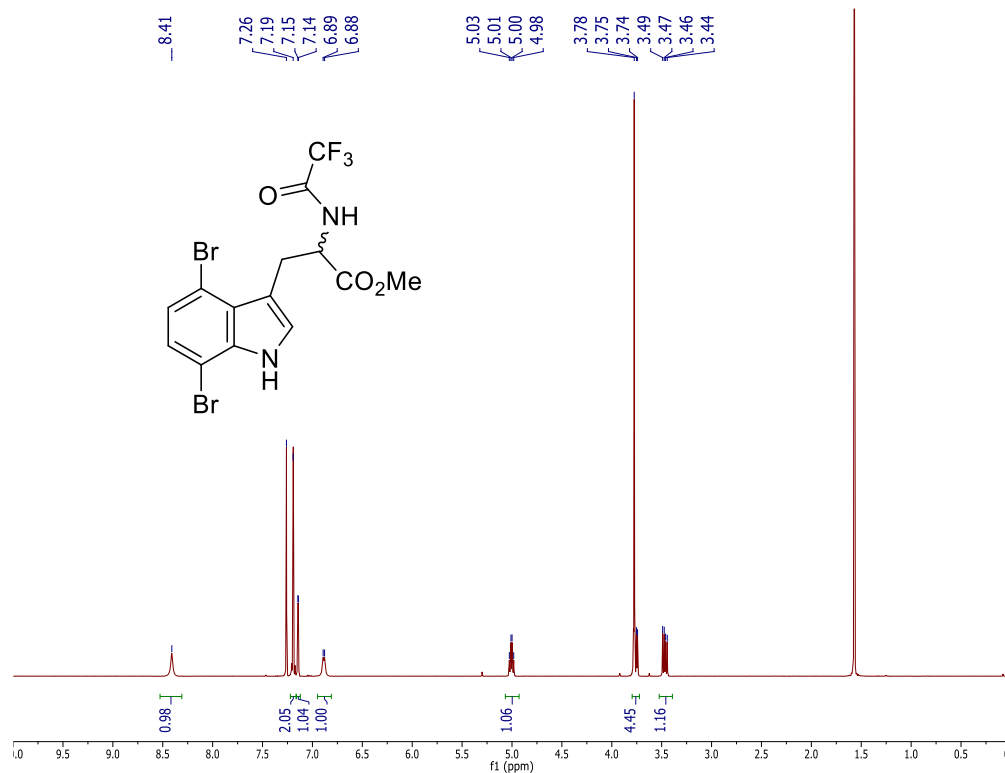
¹Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, United States

*e-mail: chughes@ucsd.edu

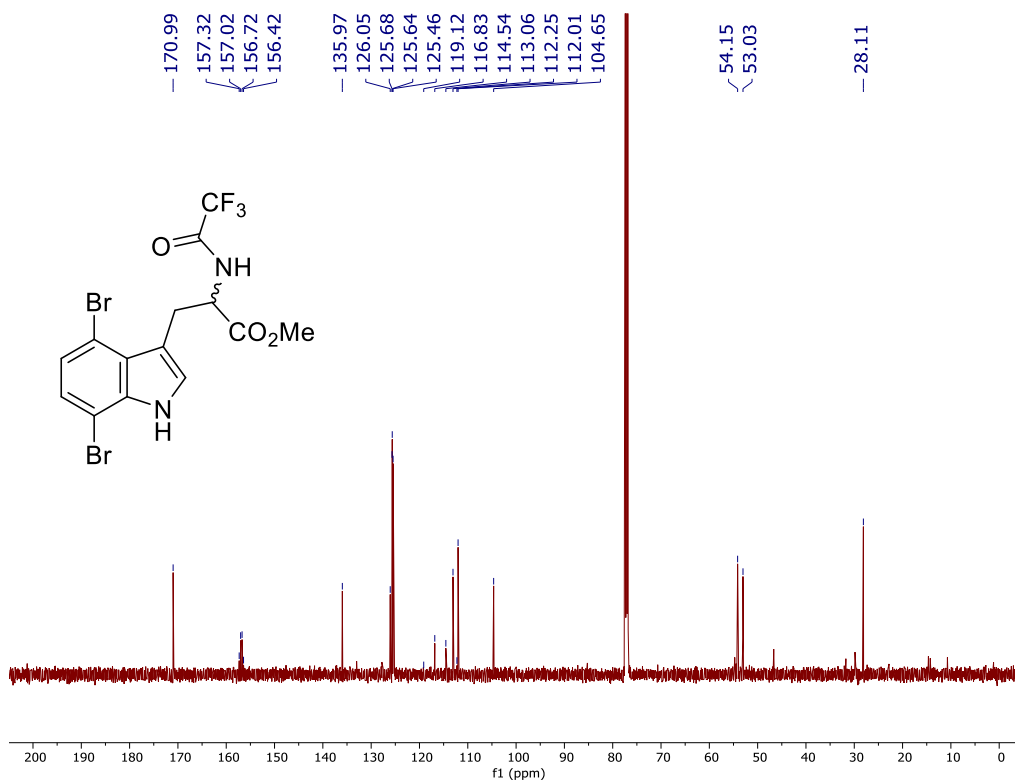
²Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, United States

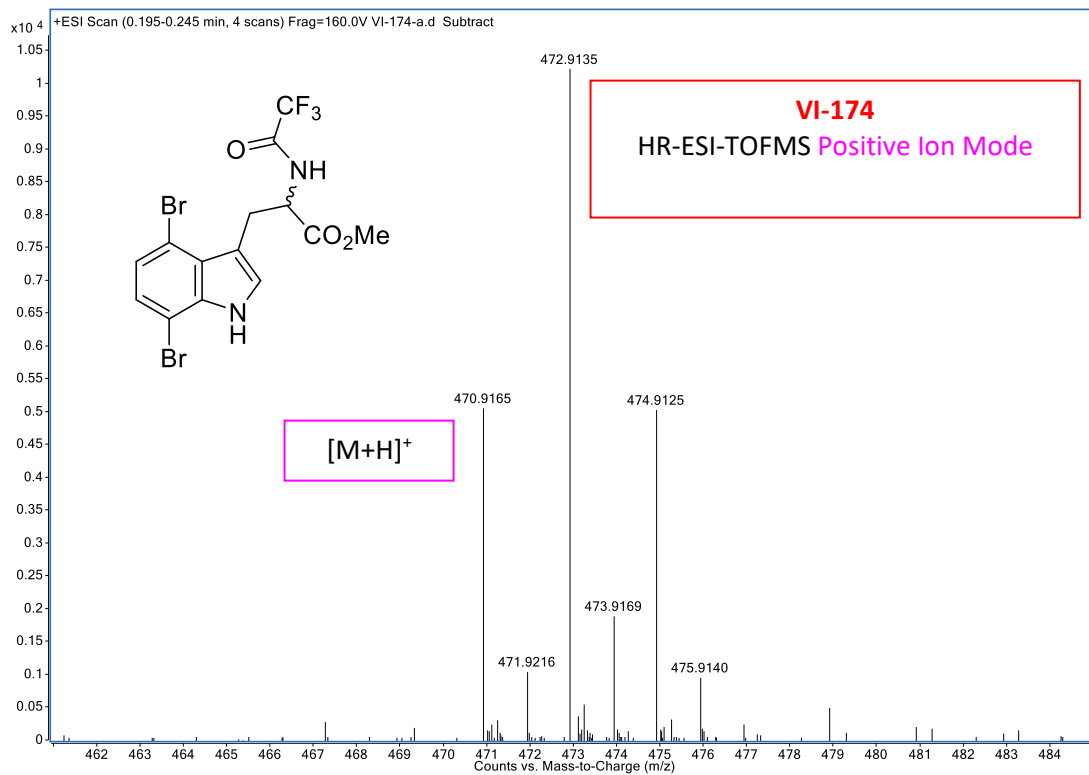
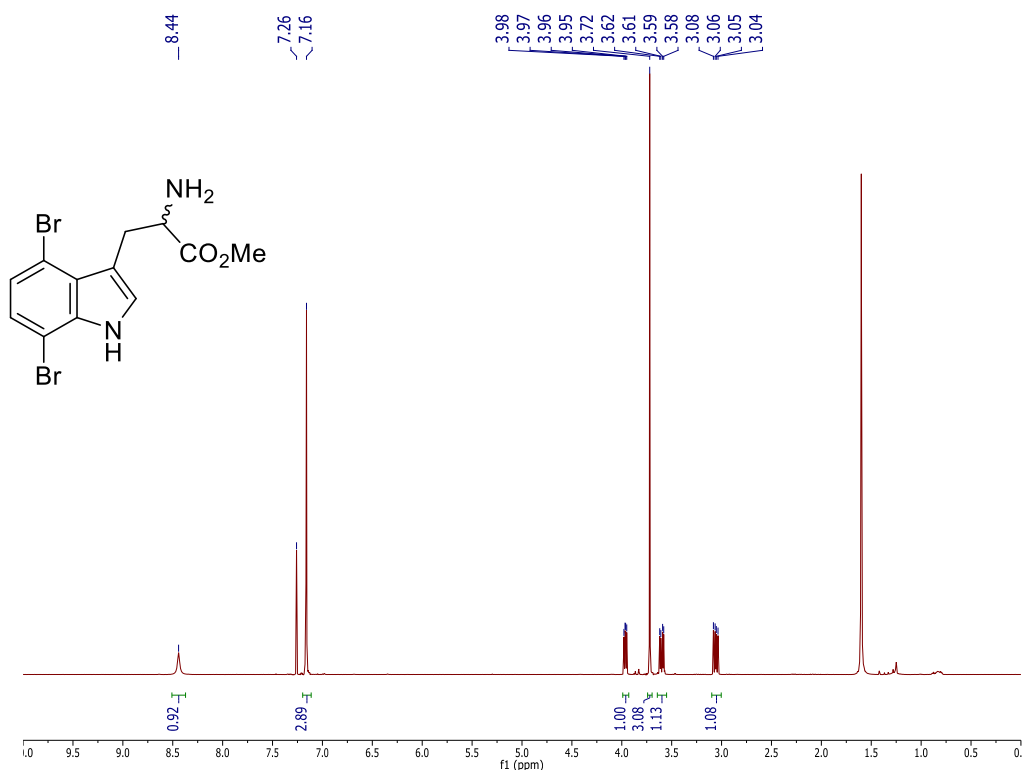
	<u>Page</u>
Table of contents	
¹ H NMR spectrum of 9 (500 MHz, CDCl ₃)	224
¹³ C NMR spectrum of 9 (125 MHz, CDCl ₃)	224
HRMS spectrum of 9	225
¹ H NMR spectrum of 10 (500 MHz, CDCl ₃)	225
¹³ C NMR spectrum of 10 (125 MHz, CDCl ₃)	226
HRMS spectrum of 10	226
¹ H NMR spectrum of 6 (500 MHz, CDCl ₃)	227
¹³ C NMR spectrum of 6 (125 MHz, CDCl ₃)	227
HRMS spectrum of 6	228
¹ H NMR spectrum of 12 (500 MHz, CD ₃ CN)	228
¹³ C NMR spectrum of 12 (125 MHz, CD ₃ CN)	229
HRMS spectrum of 12	229
X-ray crystallography report for 12 (Tables S1-S6)	230
¹ H NMR spectrum of 15 (500 MHz, C ₅ D ₅ N)	242
¹³ C NMR spectrum of 15 (125 MHz, C ₅ D ₅ N)	242
HRMS spectrum of 15	243
X-ray crystallography report for 15 (Tables S7-S12)	244

¹H NMR spectrum of **9** (500 MHz, CDCl₃)

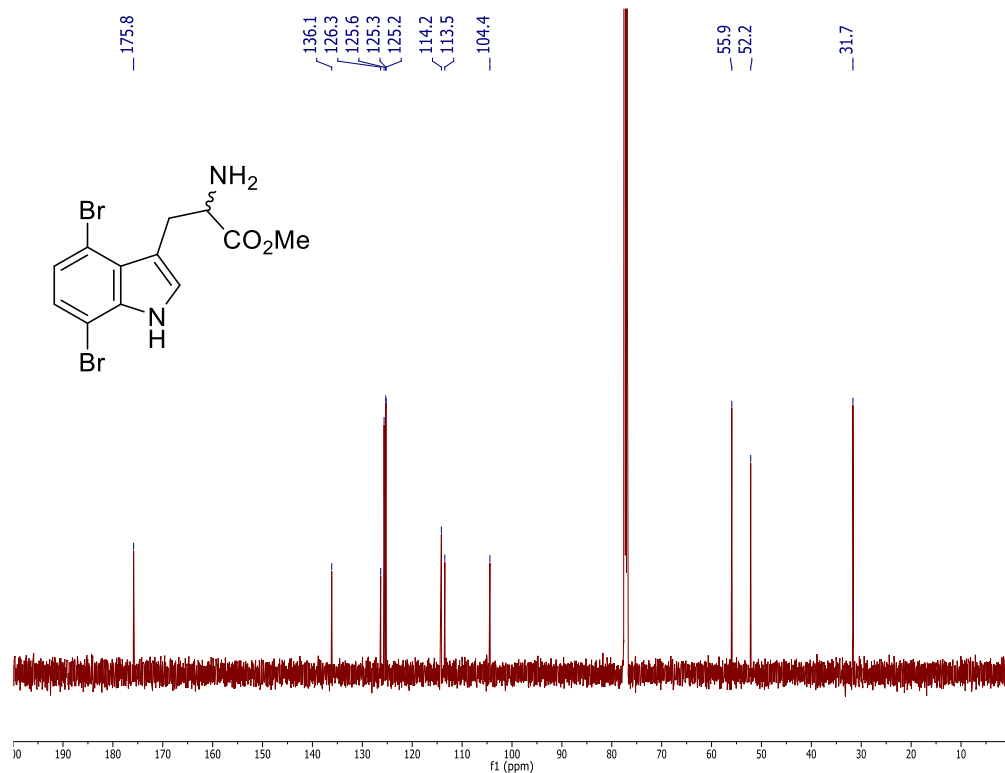


¹³C NMR spectrum of **9** (125 MHz, CDCl₃)

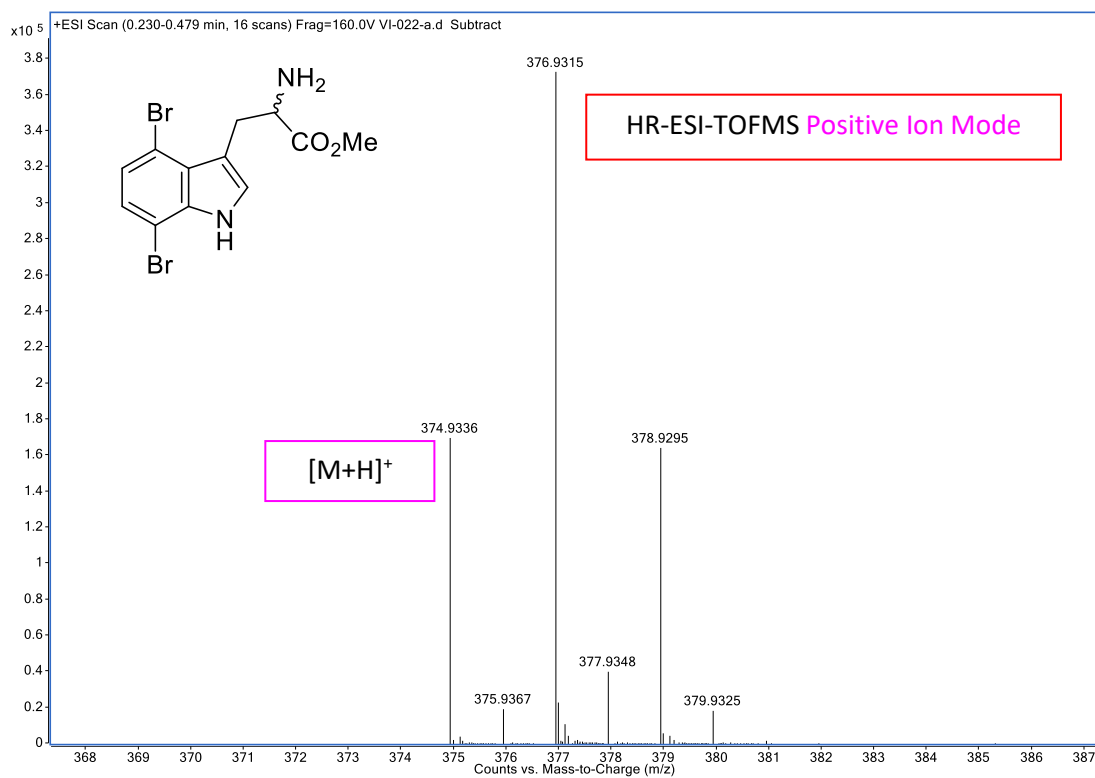


 ^1H NMR spectrum of **10** (500 MHz, CDCl_3)

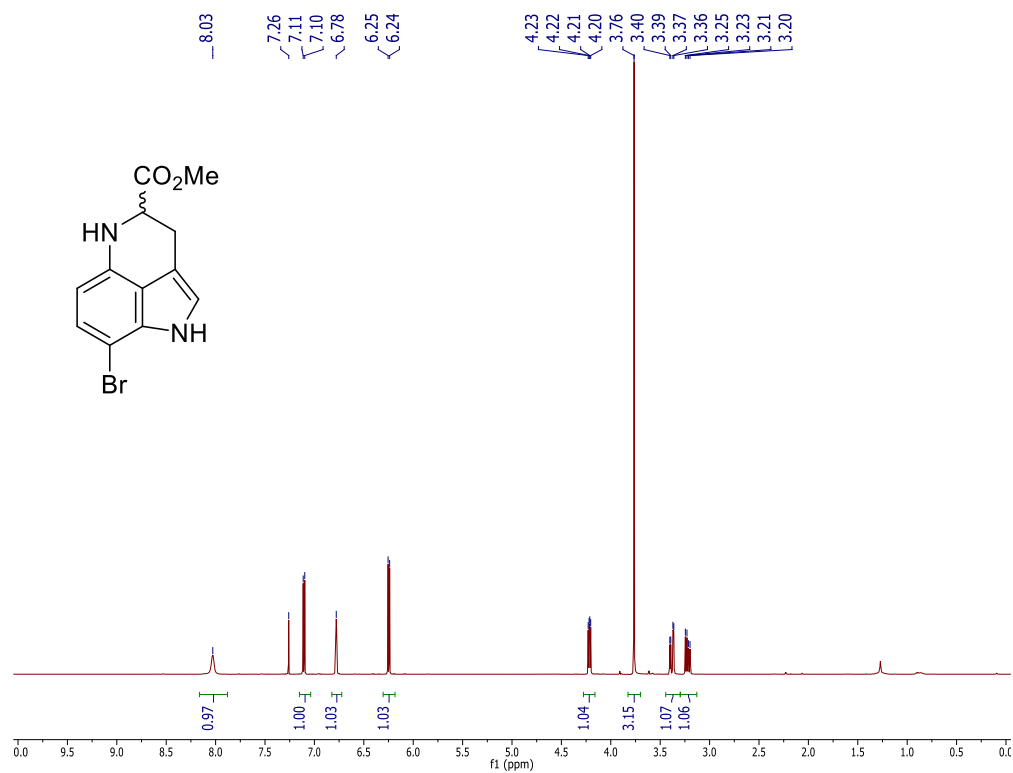
^{13}C NMR spectrum of **10** (125 MHz, CDCl_3)



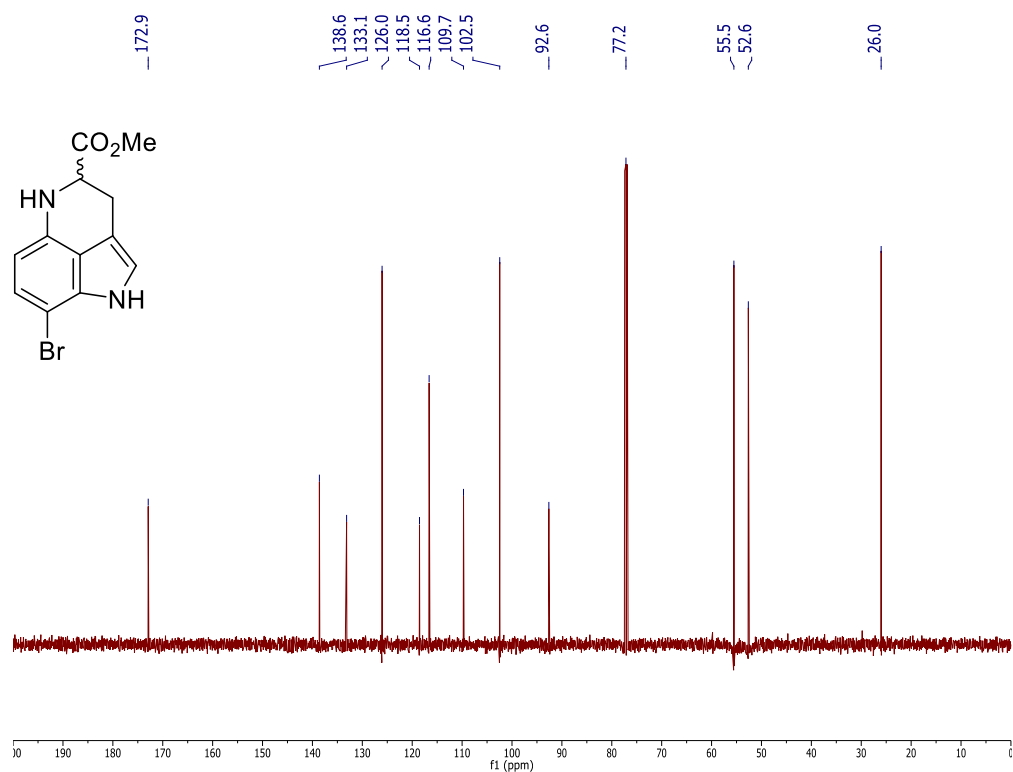
HRMS of **10**

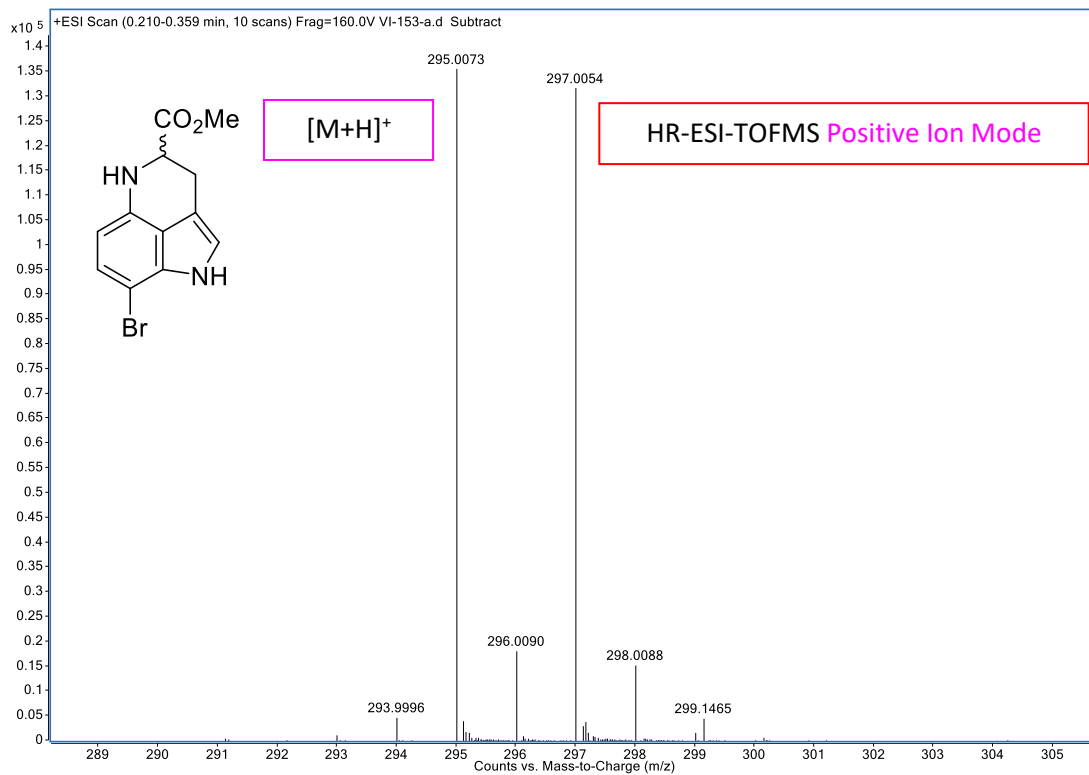
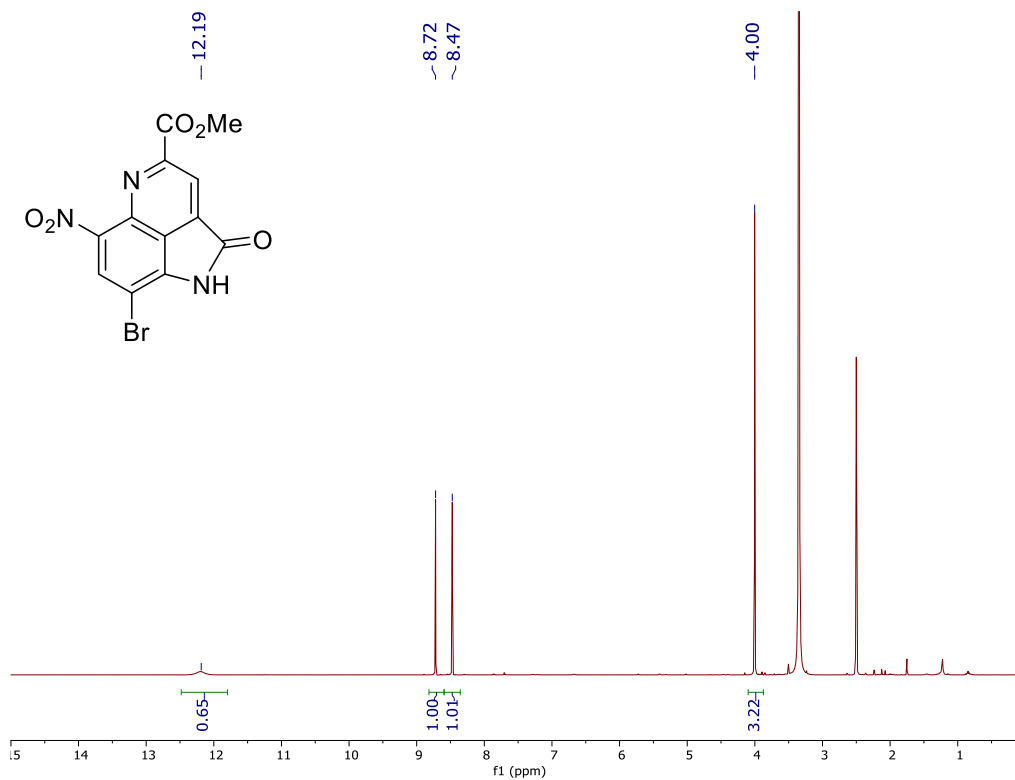


¹H NMR spectrum of **6** (500 MHz, CDCl₃)

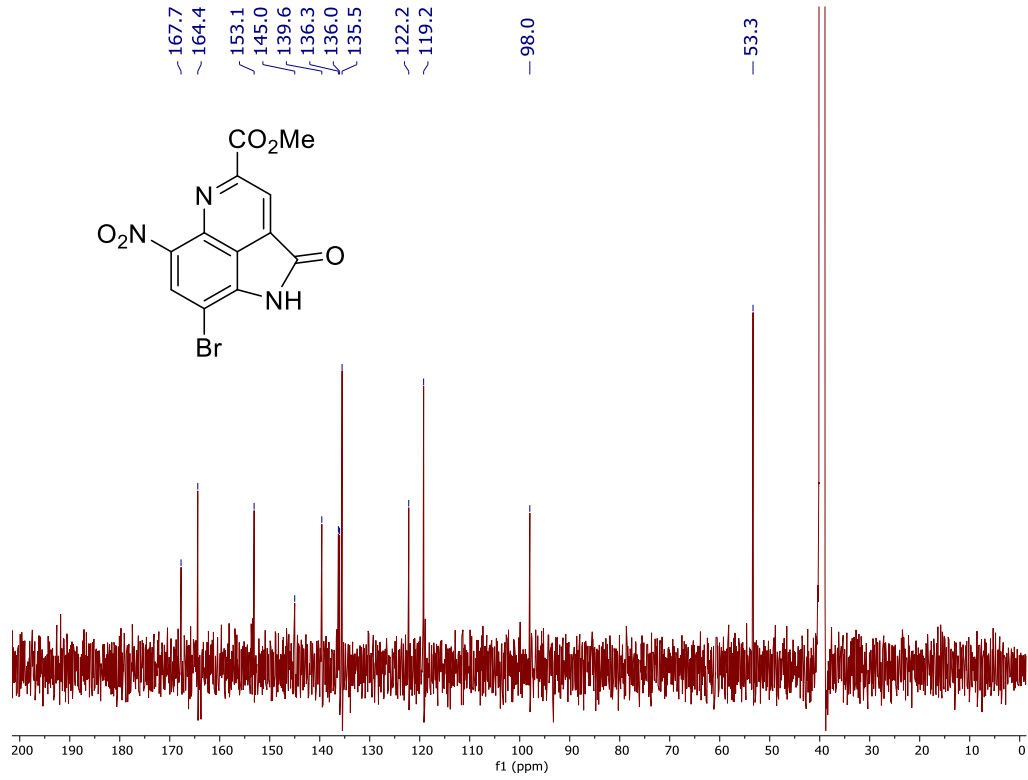


¹³C NMR spectrum of **6** (125 MHz, CDCl₃)

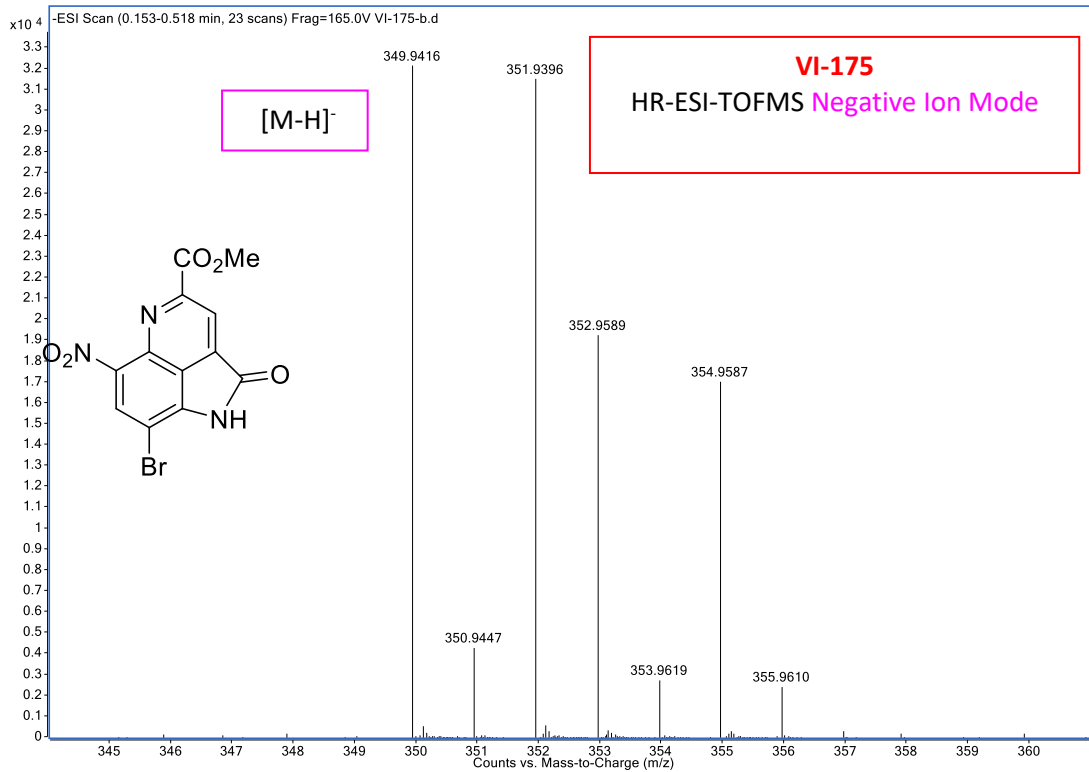


 ^1H NMR of **12** (500 MHz, DMSO- d_6)

¹³C NMR of **12** (125 MHz, DMSO-d₆)



HRMS of **12**



Experimental Summary

The single crystal X-ray diffraction studies were carried out on a Bruker Kappa APEX-II CCD diffractometer equipped with Mo K α radiation ($\lambda = 0.71073$). Crystals of the subject compound were used as received (grown from MeOH solution). A 0.250 x 0.175 x 0.150 mm light orange block was mounted on a Cryoloop with Paratone oil.

Data were collected in a nitrogen gas stream at 100(2) K using Φ and ϖ scans. Crystal-to-detector distance was 40 mm using exposure time 10s with a scan width of 0.70°. Data collection was 100.0% complete to 25.242° in Θ . A total of 15457 reflections were collected covering the indices, $-9 \leq h \leq 9$, $-11 \leq k \leq 10$, $-12 \leq l \leq 12$. 2361 reflections were found to be symmetry independent, with a R_{int} of 0.0375. Indexing and unit cell refinement indicated a primitive **Triclinic** lattice. The space group was found to be **P-1**. The data were integrated using the Bruker SAINT software program and scaled using the SADABS software program. Solution by direct methods (SHELXT) produced a complete phasing model consistent with the proposed structure.

All nonhydrogen atoms were refined anisotropically by full-matrix least-squares (SHELXL-2014). All carbon bonded hydrogen atoms were placed using a riding model. Their positions were constrained relative to their parent atom using the appropriate HFIX command in SHELXL-2014. All other hydrogen atoms (N-H) were located in the difference map. Their relative positions were NOT restrained using DFIX or AFIX commands and their thermals refined using Uiso 1.5. Crystallographic data are summarized in Table S1.

Great data and refinement.

Disorder on the Nitro group well modeled, EADP on the N3A and N3B atom.

Hydrogen bond in Table S6

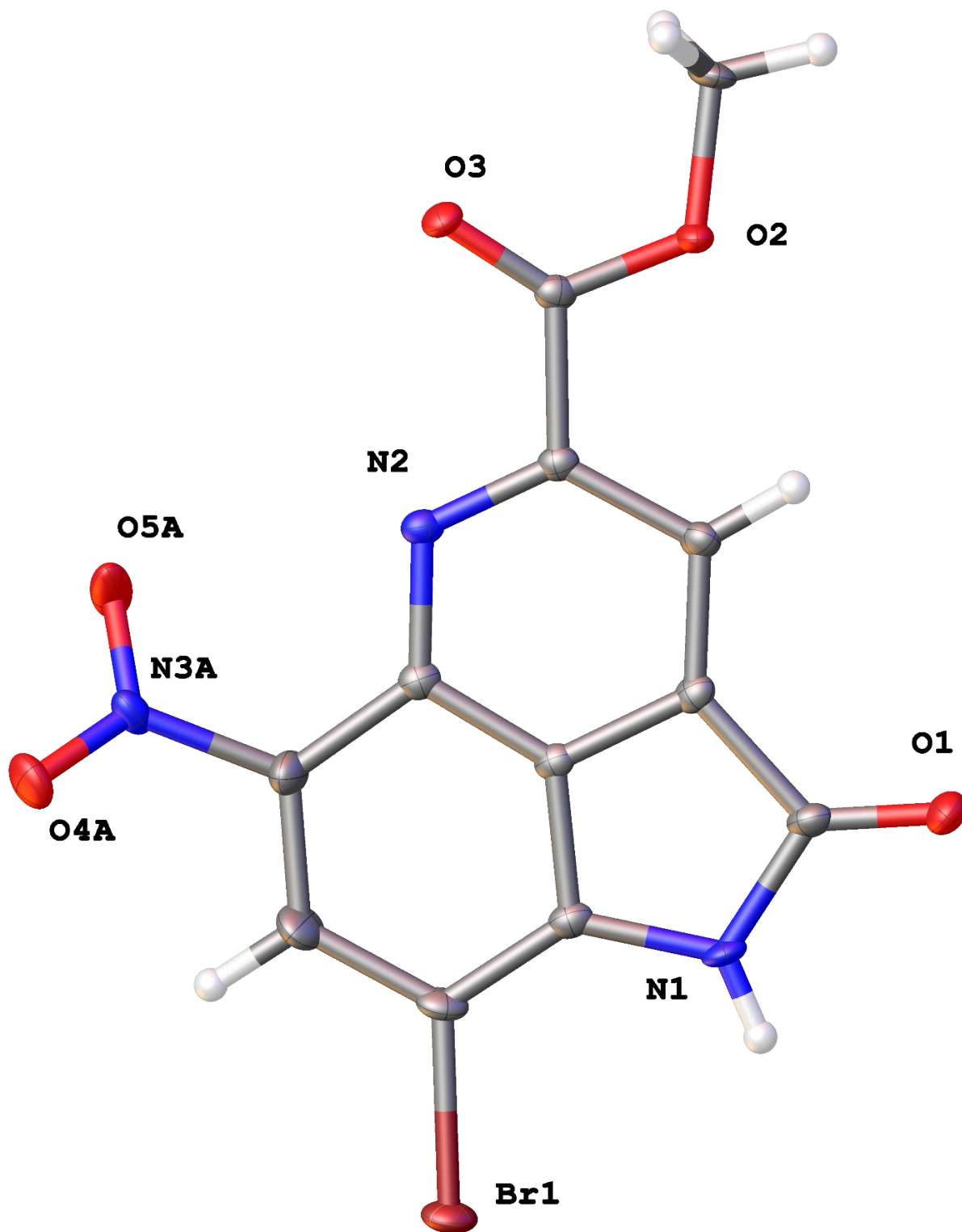


Table S1. Crystal data and structure refinement for Chamb_GS_093.

Report date 2016-04-07

Identification code chamb_gs_093

Empirical formula C₁₂ H₆ Br N₃ O₅

Molecular formula C₁₂ H₆ Br N₃ O₅

Formula weight 352.11

Temperature 100.0 K

Wavelength 0.71073 Å

Crystal system Triclinic

Space group P-1

Unit cell dimensions a = 7.3227(6) Å α = 88.721(3)°.
b = 9.1869(7) Å β = 69.390(3)°.
c = 9.8318(7) Å γ = 76.809(3)°.

Volume 601.53(8) Å³

Z 2

Density (calculated) 1.944 Mg/m³

Absorption coefficient 3.446 mm⁻¹

F(000) 348

Crystal size 0.25 x 0.175 x 0.15 mm³

Crystal color, habit orange block

Theta range for data collection 2.218 to 26.014°.

Index ranges -9 ≤ h ≤ 9, -11 ≤ k ≤ 10, -12 ≤ l ≤ 12

Reflections collected 15457

Independent reflections 2361 [R(int) = 0.0375]

Completeness to theta = 25.242° 100.0 %
Absorption correction Semi-empirical from equivalents
Max. and min. transmission 0.4908 and 0.3953
Refinement method Full-matrix least-squares on F2
Data / restraints / parameters 2361 / 0 / 216
Goodness-of-fit on F2 1.058
Final R indices [$I > 2\sigma(I)$] R1 = 0.0233, wR2 = 0.0531
R indices (all data) R1 = 0.0288, wR2 = 0.0553
Extinction coefficient n/a
Largest diff. peak and hole 0.393 and -0.321 e.Å⁻³

Table S2. Atomic coordinates ($\times 10^4$) and equivalent isotropic displacement parameters ($\text{\AA}^2 \times 10^3$)

for Chamb_GS_093. $U(\text{eq})$ is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	$U(\text{eq})$
Br(1)	6770(1)	6942(1)	9253(1)	21(1)
O(1)	7393(2)	12740(2)	7252(2)	23(1)
O(2)	8461(2)	12312(2)	1399(2)	16(1)
O(3)	7636(2)	10362(2)	593(2)	16(1)
O(4A)	5873(5)	5087(4)	4385(4)	25(1)
O(4B)	8768(5)	4626(3)	3734(4)	26(1)
O(5A)	8170(9)	5834(5)	2641(4)	38(1)
O(5B)	6878(7)	6243(4)	2847(4)	25(1)
N(1)	7191(3)	10279(2)	7785(2)	15(1)
N(2)	7596(2)	9044(2)	3160(2)	13(1)
N(3A)	7053(11)	5896(18)	3921(17)	17(2)
N(3B)	7737(11)	5879(16)	3714(15)	17(2)
C(1)	7345(3)	11502(2)	6902(2)	15(1)
C(2)	7467(3)	10903(2)	5444(2)	13(1)
C(3)	7659(3)	11474(2)	4121(2)	14(1)
C(4)	7712(3)	10468(2)	3017(2)	12(1)
C(5)	7441(3)	8487(2)	4481(2)	13(1)
C(6)	7328(3)	6996(2)	4881(2)	20(1)
C(7)	7149(4)	6578(2)	6261(3)	24(1)

C(8)	7067(3)	7589(2)	7370(2)	17(1)
C(9)	7192(3)	9019(2)	7034(2)	13(1)
C(10)	7369(3)	9418(2)	5603(2)	11(1)
C(11)	7928(3)	11019(2)	1521(2)	12(1)
C(12)	8778(3)	12991(2)	14(2)	19(1)

Table S3. Bond lengths [Å] and angles [°] for Chamb_GS_093.

Br(1)-C(8)	1.885(2)
O(1)-C(1)	1.208(3)
O(2)-C(11)	1.323(3)
O(2)-C(12)	1.450(2)
O(3)-C(11)	1.210(2)
O(4A)-N(3A)	1.228(13)
O(4B)-N(3B)	1.229(14)
O(5A)-N(3A)	1.230(14)
O(5B)-N(3B)	1.226(16)
N(1)-H(1)	0.81(3)
N(1)-C(1)	1.405(3)
N(1)-C(9)	1.387(3)
N(2)-C(4)	1.332(3)
N(2)-C(5)	1.362(3)
N(3A)-C(6)	1.491(16)

N(3B)-C(6)	1.460(15)
C(1)-C(2)	1.512(3)
C(2)-C(3)	1.365(3)
C(2)-C(10)	1.385(3)
C(3)-H(3)	0.9500
C(3)-C(4)	1.425(3)
C(4)-C(11)	1.511(3)
C(5)-C(6)	1.428(3)
C(5)-C(10)	1.390(3)
C(6)-C(7)	1.371(3)
C(7)-H(7)	0.9500
C(7)-C(8)	1.424(3)
C(8)-C(9)	1.363(3)
C(9)-C(10)	1.414(3)
C(12)-H(12A)	0.9800
C(12)-H(12B)	0.9800
C(12)-H(12C)	0.9800
C(11)-O(2)-C(12)	117.30(16)
C(1)-N(1)-H(1)	124.0(18)
C(9)-N(1)-H(1)	124.5(18)
C(9)-N(1)-C(1)	111.49(17)
C(4)-N(2)-C(5)	116.96(18)
O(4A)-N(3A)-O(5A)	123.4(13)
O(4A)-N(3A)-C(6)	122.6(11)

O(5A)-N(3A)-C(6)	113.9(9)
O(4B)-N(3B)-C(6)	118.5(11)
O(5B)-N(3B)-O(4B)	124.9(13)
O(5B)-N(3B)-C(6)	116.3(10)
O(1)-C(1)-N(1)	126.56(19)
O(1)-C(1)-C(2)	128.82(19)
N(1)-C(1)-C(2)	104.62(17)
C(3)-C(2)-C(1)	135.37(19)
C(3)-C(2)-C(10)	118.41(19)
C(10)-C(2)-C(1)	106.21(17)
C(2)-C(3)-H(3)	122.1
C(2)-C(3)-C(4)	115.88(19)
C(4)-C(3)-H(3)	122.1
N(2)-C(4)-C(3)	126.28(19)
N(2)-C(4)-C(11)	114.77(17)
C(3)-C(4)-C(11)	118.95(18)
N(2)-C(5)-C(6)	126.93(19)
N(2)-C(5)-C(10)	119.36(18)
C(10)-C(5)-C(6)	113.71(18)
C(5)-C(6)-N(3A)	122.0(6)
C(5)-C(6)-N(3B)	117.3(6)
C(7)-C(6)-N(3A)	115.9(6)
C(7)-C(6)-N(3B)	120.7(6)
C(7)-C(6)-C(5)	121.3(2)

C(6)-C(7)-H(7)	118.8
C(6)-C(7)-C(8)	122.4(2)
C(8)-C(7)-H(7)	118.8
C(7)-C(8)-Br(1)	120.03(16)
C(9)-C(8)-Br(1)	121.60(16)
C(9)-C(8)-C(7)	118.38(19)
N(1)-C(9)-C(10)	107.14(18)
C(8)-C(9)-N(1)	135.03(19)
C(8)-C(9)-C(10)	117.83(19)
C(2)-C(10)-C(5)	123.08(19)
C(2)-C(10)-C(9)	110.55(18)
C(5)-C(10)-C(9)	126.37(19)
O(2)-C(11)-C(4)	109.84(17)
O(3)-C(11)-O(2)	126.02(18)
O(3)-C(11)-C(4)	124.14(18)
O(2)-C(12)-H(12A)	109.5
O(2)-C(12)-H(12B)	109.5
O(2)-C(12)-H(12C)	109.5
H(12A)-C(12)-H(12B)	109.5
H(12A)-C(12)-H(12C)	109.5
H(12B)-C(12)-H(12C)	109.5

Table S4. Anisotropic displacement parameters ($\text{\AA}^2 \times 10^3$) for Chamb_GS_093. The anisotropic displacement factor exponent takes the form: $-2\pi^2 [h^2 a^{*2} U^{11} + \dots + 2 h k a^* b^* U^{12}]$

	U ¹¹	U ²²	U ³³	U ²³	U ¹³	U ¹²
Br(1)	28(1)	21(1)	15(1)	9(1)	-10(1)	-5(1)
O(1)	38(1)	20(1)	16(1)	1(1)	-13(1)	-12(1)
O(2)	27(1)	16(1)	10(1)	4(1)	-9(1)	-10(1)
O(3)	24(1)	17(1)	12(1)	2(1)	-10(1)	-7(1)
O(4A)	29(2)	20(2)	28(2)	5(2)	-7(2)	-16(2)
O(4B)	28(2)	13(2)	33(2)	-3(1)	-7(2)	-4(1)
O(5A)	62(4)	27(2)	15(2)	-7(2)	10(3)	-27(2)
O(5B)	40(3)	21(2)	19(2)	3(2)	-16(2)	-13(2)
N(1)	22(1)	19(1)	7(1)	2(1)	-7(1)	-6(1)
N(2)	14(1)	15(1)	11(1)	1(1)	-4(1)	-2(1)
N(3A)	19(5)	11(1)	17(4)	2(2)	0(5)	-5(5)
N(3B)	19(5)	11(1)	17(4)	2(2)	0(5)	-5(5)
C(1)	16(1)	20(1)	10(1)	2(1)	-4(1)	-6(1)
C(2)	11(1)	17(1)	12(1)	0(1)	-5(1)	-4(1)
C(3)	14(1)	16(1)	12(1)	3(1)	-4(1)	-6(1)
C(4)	10(1)	15(1)	11(1)	1(1)	-3(1)	-3(1)
C(5)	10(1)	15(1)	12(1)	1(1)	-4(1)	-2(1)
C(6)	30(1)	12(1)	17(1)	-1(1)	-10(1)	-3(1)
C(7)	39(1)	13(1)	24(1)	6(1)	-15(1)	-4(1)

C(8)	18(1)	17(1)	14(1)	7(1)	-7(1)	-2(1)
C(9)	11(1)	18(1)	11(1)	1(1)	-4(1)	-2(1)
C(10)	9(1)	15(1)	11(1)	1(1)	-4(1)	-3(1)
C(11)	11(1)	14(1)	11(1)	1(1)	-2(1)	-1(1)
C(12)	28(1)	20(1)	11(1)	9(1)	-8(1)	-11(1)

Table S5. Hydrogen coordinates ($\times 10^4$) and isotropic displacement parameters ($\text{\AA}^2 \times 10^3$) for Chamb_GS_093.

	x	y	z	U(eq)
H(1)	7140(40)	10300(30)	8620(30)	26
H(3)	7752	12481	3946	16
H(7)	7078	5577	6486	29
H(12A)	9965	12375	-733	28
H(12B)	8985	13997	101	28
H(12C)	7600	13060	-261	28

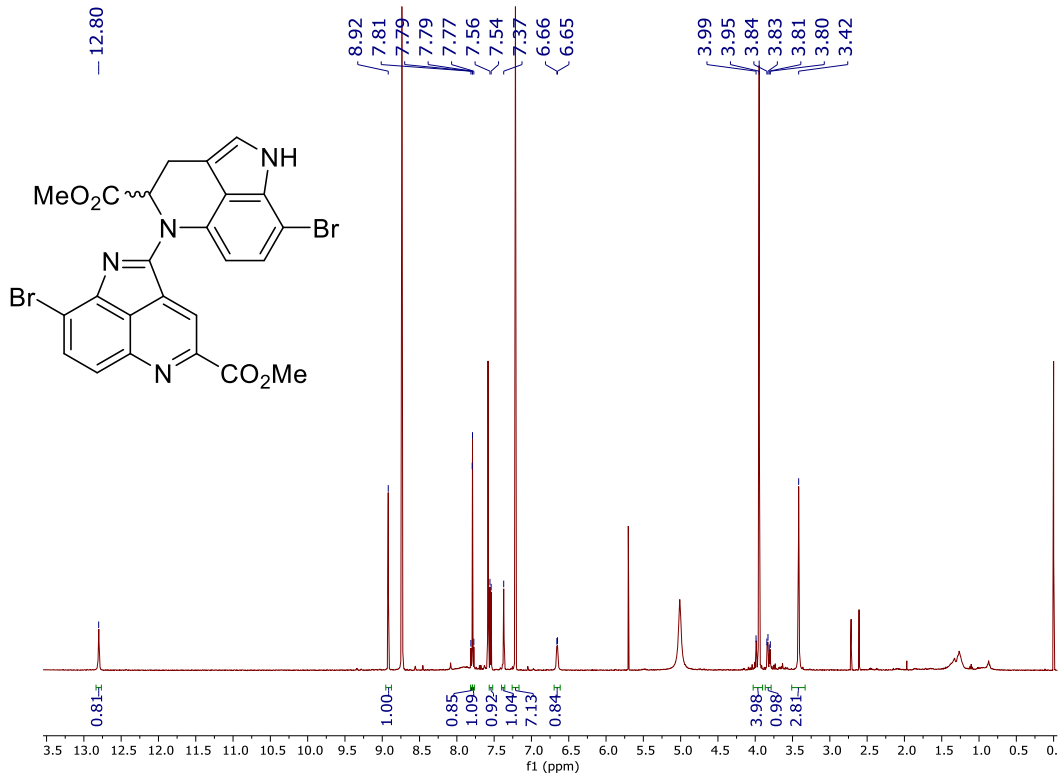
Table S6. Hydrogen bonds for Chamb_GS_093 [\AA and $^\circ$].

D-H...A	d(D-H)	d(H...A)	d(D...A)	\angle (DHA)
N(1)-H(1)...O(3)#1	0.81(3)	2.09(3)	2.896(2)	169(3)

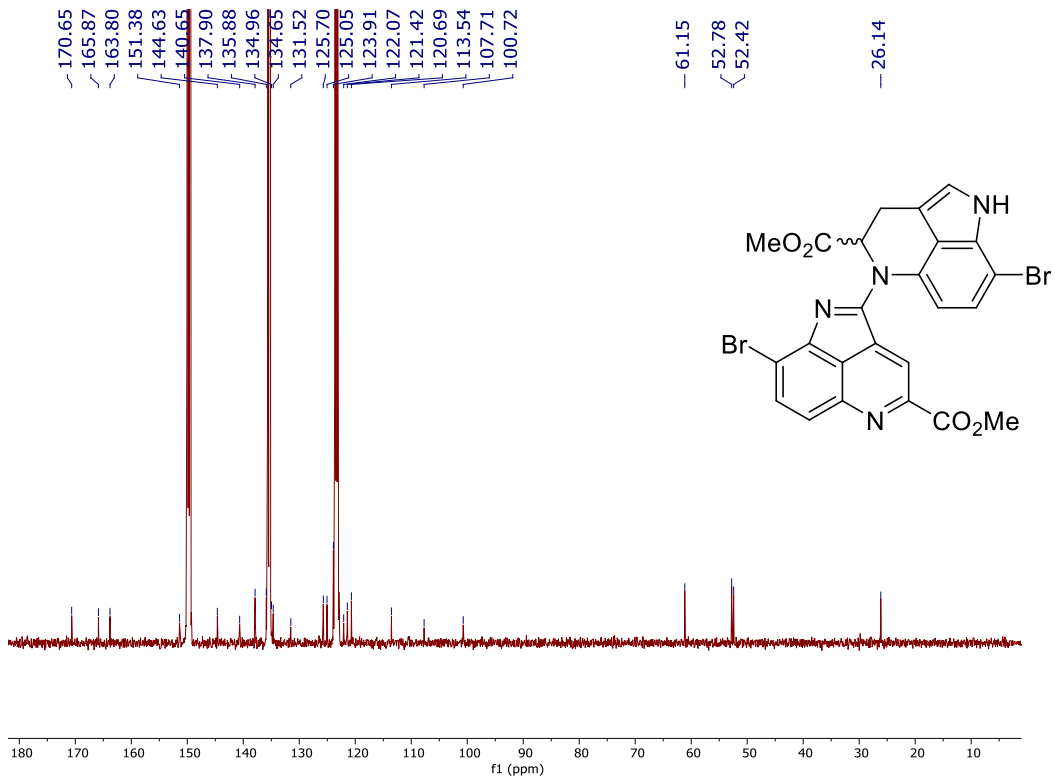
Symmetry transformations used to generate equivalent atoms:

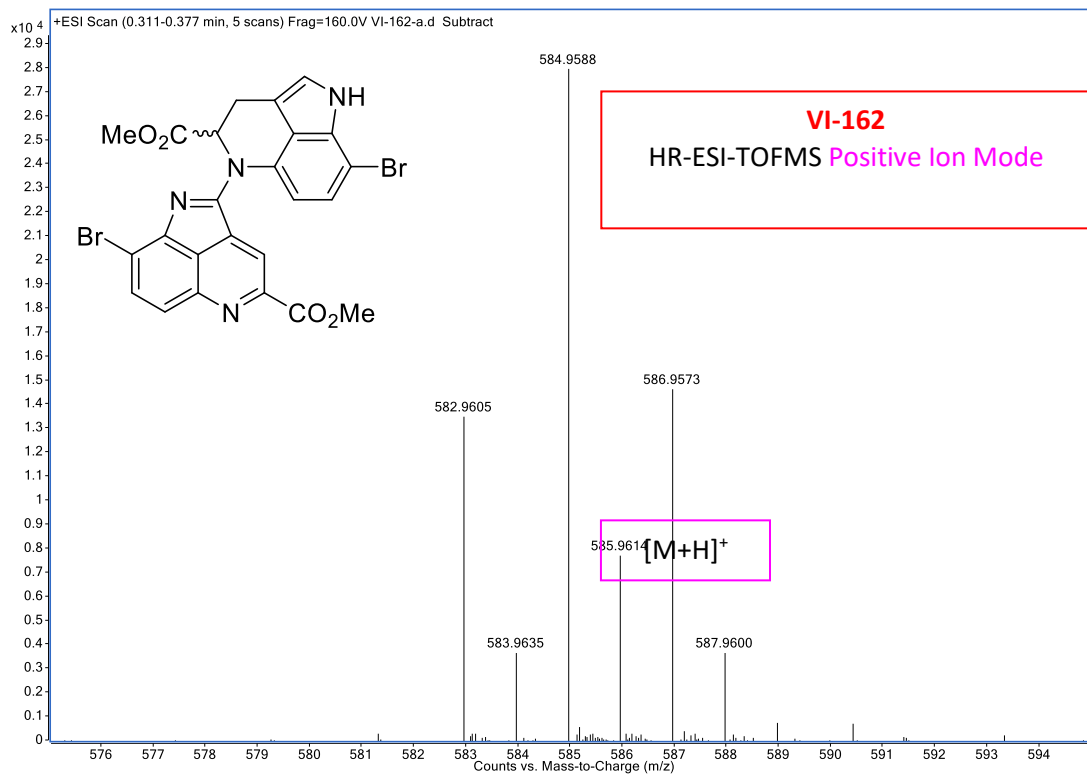
#1 x,y,z+1

¹H NMR of **15** (500 MHz, C₅D₅N)



¹³C NMR of **15** (125 MHz, C₅D₅N)





Experimental Summary

The single crystal X-ray diffraction studies were carried out on a Bruker Kappa APEX II CCD diffractometer equipped with Mo K α radiation ($\lambda = 0.71073 \text{ \AA}$). Crystals of the subject compound were used as received (grown from Toluene). A 0.115 x 0.105 x 0.050 mm piece of a dark red crystal was mounted on a Cryoloop with Paratone oil.

Data were collected in a nitrogen gas stream at 100(2) K using Φ and ω scans. Crystal-to-detector distance was 40 mm and exposure time was 15 seconds (depending on the 2θ range) per frame using a scan width of 0.60° . Data collection was 99.9% complete to 25.242° in θ . A total of 43067 reflections were collected covering the indices, $-9 \leq h \leq 8$, $-25 \leq k \leq 25$, $-18 \leq l \leq 18$. 4784 reflections were found to be symmetry independent, with a R_{int} of 0.0526. Indexing and unit cell refinement indicated a **Primitive, Monoclinic** lattice. The space group was found to be **$P2_1/c$** . The data were integrated using the Bruker SAINT Software program and scaled using the SADABS software program. Solution by direct methods (SHELXT) produced a complete phasing model consistent with the proposed structure.

All nonhydrogen atoms were refined anisotropically by full-matrix least-squares (SHELXL-2014). All carbon bonded hydrogen atoms were placed using a riding model. Their positions were constrained relative to their parent atom using the appropriate HFIX command in SHELXL-2014. . Crystallographic data are summarized in Table S7.

Excellent data and refinement, nice x-tals !

Toluene solvent in the structure, AFIX 66 and RIGU used to constrain.

Racemic space group !

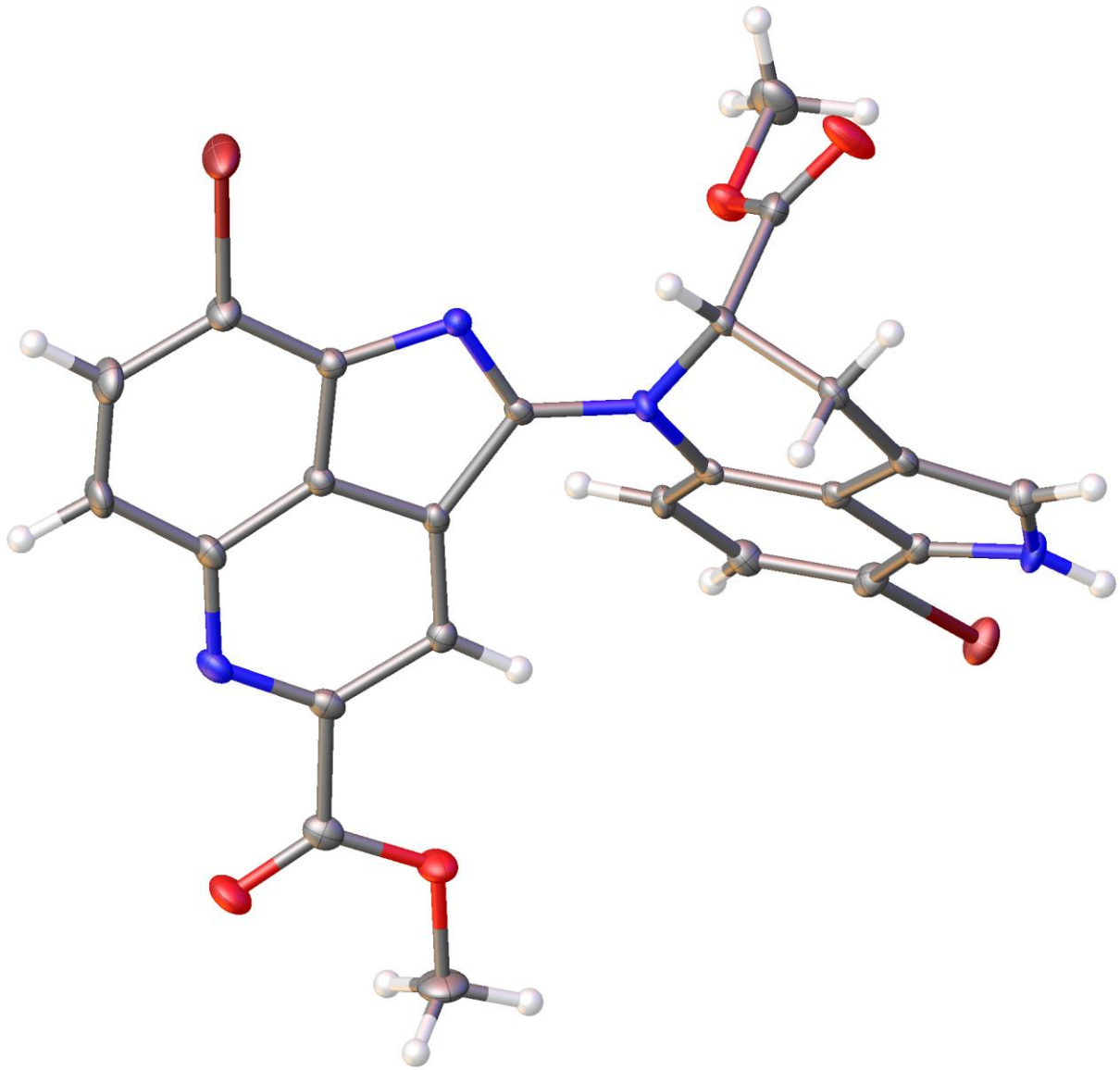


Table S7. Crystal data and structure refinement for Hughes_GS-VI-170.

Report date	2018-08-31	
Identification code	hughes_gs-vi-170	
Empirical formula	C55 H40 Br4 N8 O8	
Molecular formula	2(C24 H16 Br2 N4 O4), C7 H8	
Formula weight	1260.59	
Temperature	100.0 K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	P 1 21/c 1	
Unit cell dimensions	a = 7.8041(2) Å	α = 90°.
	b = 20.9989(7) Å	β = 90.9690(10)°.
	c = 15.3544(5) Å	γ = 90°.
Volume	2515.88(13) Å ³	
Z	2	
Density (calculated)	1.664 Mg/m ³	
Absorption coefficient	3.265 mm ⁻¹	
F(000)	1260	
Crystal size	0.115 x 0.105 x 0.05 mm ³	
Crystal color, habit	dark red block	
Theta range for data collection	2.610 to 25.680°.	
Index ranges	-9<=h<=8, -25<=k<=25, -18<=l<=18	
Reflections collected	43067	
Independent reflections	4784 [R(int) = 0.0526]	

Completeness to theta = 25.242°	99.9 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.5622 and 0.5064
Refinement method	Full-matrix least-squares on F ²
Data / restraints / parameters	4784 / 46 / 361
Goodness-of-fit on F ²	1.026
Final R indices [I>2sigma(I)]	R1 = 0.0289, wR2 = 0.0621
R indices (all data)	R1 = 0.0418, wR2 = 0.0668
Largest diff. peak and hole	0.753 and -0.748 e.Å ⁻³

Table S8. Atomic coordinates ($\times 10^4$) and equivalent isotropic displacement parameters ($\text{\AA}^2 \times 10^3$) for Hughes_GS-VI-170. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
Br(1)	8879(1)	8084(1)	7574(1)	19(1)
Br(2)	4297(1)	5295(1)	1391(1)	29(1)
O(1)	3457(2)	8141(1)	3634(1)	30(1)
O(2)	6118(2)	7745(1)	3780(1)	20(1)
O(3)	9497(2)	4001(1)	6198(1)	23(1)
O(4)	8848(2)	4952(1)	6775(1)	20(1)
N(1)	5220(3)	6694(1)	4570(1)	12(1)
N(2)	4749(3)	8030(1)	6890(1)	16(1)
N(3)	5039(3)	6031(1)	3339(1)	12(1)
N(4)	7970(3)	4354(1)	4668(1)	16(1)
C(1)	3788(3)	7103(1)	4250(2)	13(1)
C(2)	2521(3)	7251(1)	4986(2)	13(1)
C(3)	3520(3)	7520(1)	5748(2)	13(1)
C(4)	3254(3)	7944(1)	6405(2)	16(1)
C(5)	6022(3)	7668(1)	6530(2)	13(1)
C(6)	7780(3)	7597(1)	6689(2)	16(1)
C(7)	8692(3)	7195(1)	6154(2)	15(1)
C(8)	7931(3)	6871(1)	5443(2)	13(1)
C(9)	6193(3)	6939(1)	5288(2)	12(1)

C(10)	5269(3)	7337(1)	5833(2)	12(1)
C(11)	4421(3)	7724(1)	3853(2)	16(1)
C(12)	6818(4)	8331(1)	3435(2)	27(1)
C(13)	5592(3)	6148(1)	4146(2)	11(1)
C(14)	6545(3)	5591(1)	4536(2)	11(1)
C(15)	7280(3)	5396(1)	5310(2)	12(1)
C(16)	7971(3)	4768(1)	5324(2)	14(1)
C(17)	7196(3)	4541(1)	3915(2)	15(1)
C(18)	7034(3)	4161(1)	3146(2)	20(1)
C(19)	6214(3)	4410(1)	2432(2)	21(1)
C(20)	5463(3)	5032(1)	2416(2)	17(1)
C(21)	5603(3)	5410(1)	3139(2)	12(1)
C(22)	6492(3)	5151(1)	3858(2)	12(1)
C(23)	8854(3)	4518(1)	6137(2)	16(1)
C(24)	9716(4)	4763(1)	7574(2)	26(1)
C(1S)	9782(13)	10969(3)	4934(7)	46(2)
C(2S)	9795(6)	10244(2)	5015(4)	32(1)
C(7S)	8401(5)	9946(2)	5393(3)	35(2)
C(6S)	8375(6)	9288(2)	5481(3)	37(1)
C(5S)	9744(7)	8926(2)	5191(4)	45(2)
C(4S)	11139(6)	9224(2)	4813(3)	46(2)
C(3S)	11164(5)	9882(3)	4725(3)	38(2)

Table S9. Bond lengths [Å] and angles [°] for Hughes_GS-VI-170.

Br(1)-C(6)	1.894(2)
Br(2)-C(20)	1.887(3)
O(1)-C(11)	1.199(3)
O(2)-C(11)	1.331(3)
O(2)-C(12)	1.450(3)
O(3)-C(23)	1.199(3)
O(4)-C(23)	1.338(3)
O(4)-C(24)	1.446(3)
N(1)-C(1)	1.485(3)
N(1)-C(9)	1.424(3)
N(1)-C(13)	1.353(3)
N(2)-H(2)	0.8800
N(2)-C(4)	1.385(3)
N(2)-C(5)	1.375(3)
N(3)-C(13)	1.329(3)
N(3)-C(21)	1.412(3)
N(4)-C(16)	1.331(3)
N(4)-C(17)	1.354(3)
C(1)-H(1)	1.0000
C(1)-C(2)	1.546(3)
C(1)-C(11)	1.526(3)
C(2)-H(2A)	0.9900

C(2)-H(2B)	0.9900
C(2)-C(3)	1.505(3)
C(3)-C(4)	1.363(3)
C(3)-C(10)	1.422(3)
C(4)-H(4)	0.9500
C(5)-C(6)	1.398(4)
C(5)-C(10)	1.396(3)
C(6)-C(7)	1.383(4)
C(7)-H(7)	0.9500
C(7)-C(8)	1.409(3)
C(8)-H(8)	0.9500
C(8)-C(9)	1.380(3)
C(9)-C(10)	1.393(3)
C(12)-H(12A)	0.9800
C(12)-H(12B)	0.9800
C(12)-H(12C)	0.9800
C(13)-C(14)	1.504(3)
C(14)-C(15)	1.374(3)
C(14)-C(22)	1.393(3)
C(15)-H(15)	0.9500
C(15)-C(16)	1.425(3)
C(16)-C(23)	1.509(4)
C(17)-C(18)	1.428(4)
C(17)-C(22)	1.397(3)

C(18)-H(18)	0.9500
C(18)-C(19)	1.364(4)
C(19)-H(19)	0.9500
C(19)-C(20)	1.433(4)
C(20)-C(21)	1.367(3)
C(21)-C(22)	1.403(3)
C(24)-H(24A)	0.9800
C(24)-H(24B)	0.9800
C(24)-H(24C)	0.9800
C(15)-H(15A)	0.9800
C(15)-H(15B)	0.9800
C(15)-H(15C)	0.9800
C(15)-C(25)	1.528(7)
C(25)-C(7S)	1.3900
C(25)-C(3S)	1.3900
C(7S)-H(7S)	0.9500
C(7S)-C(6S)	1.3900
C(6S)-H(6S)	0.9500
C(6S)-C(5S)	1.3900
C(5S)-H(5S)	0.9500
C(5S)-C(4S)	1.3900
C(4S)-H(4S)	0.9500
C(4S)-C(3S)	1.3900
C(3S)-H(3S)	0.9500

C(11)-O(2)-C(12)	116.2(2)
C(23)-O(4)-C(24)	115.3(2)
C(9)-N(1)-C(1)	115.82(19)
C(13)-N(1)-C(1)	119.8(2)
C(13)-N(1)-C(9)	124.2(2)
C(4)-N(2)-H(2)	125.7
C(5)-N(2)-H(2)	125.7
C(5)-N(2)-C(4)	108.6(2)
C(13)-N(3)-C(21)	105.96(19)
C(16)-N(4)-C(17)	116.9(2)
N(1)-C(1)-H(1)	107.9
N(1)-C(1)-C(2)	111.21(19)
N(1)-C(1)-C(11)	112.3(2)
C(2)-C(1)-H(1)	107.9
C(11)-C(1)-H(1)	107.9
C(11)-C(1)-C(2)	109.6(2)
C(1)-C(2)-H(2A)	110.0
C(1)-C(2)-H(2B)	110.0
H(2A)-C(2)-H(2B)	108.4
C(3)-C(2)-C(1)	108.3(2)
C(3)-C(2)-H(2A)	110.0
C(3)-C(2)-H(2B)	110.0
C(4)-C(3)-C(2)	137.5(2)
C(4)-C(3)-C(10)	105.4(2)

C(10)-C(3)-C(2)	117.0(2)
N(2)-C(4)-H(4)	124.9
C(3)-C(4)-N(2)	110.2(2)
C(3)-C(4)-H(4)	124.9
N(2)-C(5)-C(6)	134.6(2)
N(2)-C(5)-C(10)	106.6(2)
C(10)-C(5)-C(6)	118.7(2)
C(5)-C(6)-Br(1)	119.97(19)
C(7)-C(6)-Br(1)	121.62(19)
C(7)-C(6)-C(5)	118.3(2)
C(6)-C(7)-H(7)	118.6
C(6)-C(7)-C(8)	122.7(2)
C(8)-C(7)-H(7)	118.6
C(7)-C(8)-H(8)	120.5
C(9)-C(8)-C(7)	118.9(2)
C(9)-C(8)-H(8)	120.5
C(8)-C(9)-N(1)	127.3(2)
C(8)-C(9)-C(10)	118.4(2)
C(10)-C(9)-N(1)	114.0(2)
C(5)-C(10)-C(3)	109.1(2)
C(9)-C(10)-C(3)	127.8(2)
C(9)-C(10)-C(5)	122.9(2)
O(1)-C(11)-O(2)	124.9(2)
O(1)-C(11)-C(1)	122.0(2)

O(2)-C(11)-C(1)	113.1(2)
O(2)-C(12)-H(12A)	109.5
O(2)-C(12)-H(12B)	109.5
O(2)-C(12)-H(12C)	109.5
H(12A)-C(12)-H(12B)	109.5
H(12A)-C(12)-H(12C)	109.5
H(12B)-C(12)-H(12C)	109.5
N(1)-C(13)-C(14)	125.2(2)
N(3)-C(13)-N(1)	122.4(2)
N(3)-C(13)-C(14)	112.3(2)
C(15)-C(14)-C(13)	140.6(2)
C(15)-C(14)-C(22)	117.2(2)
C(22)-C(14)-C(13)	102.1(2)
C(14)-C(15)-H(15)	121.9
C(14)-C(15)-C(16)	116.2(2)
C(16)-C(15)-H(15)	121.9
N(4)-C(16)-C(15)	126.7(2)
N(4)-C(16)-C(23)	113.1(2)
C(15)-C(16)-C(23)	120.1(2)
N(4)-C(17)-C(18)	125.2(2)
N(4)-C(17)-C(22)	119.3(2)
C(22)-C(17)-C(18)	115.5(2)
C(17)-C(18)-H(18)	120.5
C(19)-C(18)-C(17)	119.0(2)

C(19)-C(18)-H(18)	120.5
C(18)-C(19)-H(19)	118.4
C(18)-C(19)-C(20)	123.3(2)
C(20)-C(19)-H(19)	118.4
C(19)-C(20)-Br(2)	118.15(19)
C(21)-C(20)-Br(2)	122.57(19)
C(21)-C(20)-C(19)	119.3(2)
C(20)-C(21)-N(3)	133.7(2)
C(20)-C(21)-C(22)	116.4(2)
C(22)-C(21)-N(3)	109.8(2)
C(14)-C(22)-C(17)	123.7(2)
C(14)-C(22)-C(21)	109.8(2)
C(17)-C(22)-C(21)	126.4(2)
O(3)-C(23)-O(4)	124.5(2)
O(3)-C(23)-C(16)	124.4(2)
O(4)-C(23)-C(16)	111.2(2)
O(4)-C(24)-H(24A)	109.5
O(4)-C(24)-H(24B)	109.5
O(4)-C(24)-H(24C)	109.5
H(24A)-C(24)-H(24B)	109.5
H(24A)-C(24)-H(24C)	109.5
H(24B)-C(24)-H(24C)	109.5
H(1SA)-C(1S)-H(1SB)	109.5
H(1SA)-C(1S)-H(1SC)	109.5

H(1SB)-C(1S)-H(1SC)	109.5
C(2S)-C(1S)-H(1SA)	109.5
C(2S)-C(1S)-H(1SB)	109.5
C(2S)-C(1S)-H(1SC)	109.5
C(7S)-C(2S)-C(1S)	118.5(5)
C(7S)-C(2S)-C(3S)	120.0
C(3S)-C(2S)-C(1S)	121.5(5)
C(2S)-C(7S)-H(7S)	120.0
C(2S)-C(7S)-C(6S)	120.0
C(6S)-C(7S)-H(7S)	120.0
C(7S)-C(6S)-H(6S)	120.0
C(7S)-C(6S)-C(5S)	120.0
C(5S)-C(6S)-H(6S)	120.0
C(6S)-C(5S)-H(5S)	120.0
C(4S)-C(5S)-C(6S)	120.0
C(4S)-C(5S)-H(5S)	120.0
C(5S)-C(4S)-H(4S)	120.0
C(5S)-C(4S)-C(3S)	120.0
C(3S)-C(4S)-H(4S)	120.0
C(2S)-C(3S)-H(3S)	120.0
C(4S)-C(3S)-C(2S)	120.0
C(4S)-C(3S)-H(3S)	120.0

Table S10. Anisotropic displacement parameters ($\text{\AA}^2 \times 10^3$) for Hughes_GS-VI-170.

The anisotropic displacement factor exponent takes the form:

$$-2\pi^2 [h^2 a^{*2} U^{11} + \dots + 2 h k a^* b^* U^{12}]$$

	U ¹¹	U ²²	U ³³	U ²³	U ¹³	U ¹²
Br(1)	17(1)	23(1)	18(1)	-9(1)	0(1)	-5(1)
Br(2)	43(1)	27(1)	17(1)	-8(1)	-13(1)	13(1)
O(1)	27(1)	20(1)	45(1)	16(1)	6(1)	8(1)
O(2)	21(1)	14(1)	24(1)	5(1)	4(1)	0(1)
O(3)	27(1)	17(1)	26(1)	4(1)	-6(1)	7(1)
O(4)	27(1)	19(1)	15(1)	1(1)	-7(1)	3(1)
N(1)	15(1)	7(1)	13(1)	-2(1)	-2(1)	2(1)
N(2)	19(1)	15(1)	14(1)	-7(1)	2(1)	0(1)
N(3)	16(1)	10(1)	10(1)	0(1)	1(1)	1(1)
N(4)	19(1)	9(1)	19(1)	2(1)	-3(1)	-1(1)
C(1)	15(1)	9(1)	14(1)	1(1)	-1(1)	3(1)
C(2)	13(1)	9(1)	16(1)	1(1)	0(1)	1(1)
C(3)	15(1)	9(1)	15(1)	2(1)	1(1)	1(1)
C(4)	15(1)	14(1)	18(1)	-1(1)	2(1)	-1(1)
C(5)	17(1)	10(1)	12(1)	0(1)	2(1)	-2(1)
C(6)	21(1)	15(1)	12(1)	-2(1)	-1(1)	-4(1)
C(7)	13(1)	14(1)	19(1)	1(1)	0(1)	0(1)
C(8)	17(1)	9(1)	14(1)	-2(1)	2(1)	1(1)

C(9)	18(1)	5(1)	12(1)	1(1)	-1(1)	-3(1)
C(10)	16(1)	8(1)	12(1)	2(1)	0(1)	-3(1)
C(11)	22(1)	14(1)	13(1)	-2(1)	2(1)	3(1)
C(12)	29(2)	16(1)	36(2)	5(1)	7(1)	-5(1)
C(13)	11(1)	8(1)	12(1)	2(1)	2(1)	-1(1)
C(14)	11(1)	8(1)	13(1)	0(1)	2(1)	-1(1)
C(15)	12(1)	12(1)	13(1)	-1(1)	1(1)	-2(1)
C(16)	13(1)	12(1)	16(1)	4(1)	0(1)	-2(1)
C(17)	18(1)	8(1)	19(1)	1(1)	0(1)	-1(1)
C(18)	25(1)	10(1)	23(2)	-4(1)	0(1)	3(1)
C(19)	25(2)	16(1)	21(2)	-9(1)	-3(1)	0(1)
C(20)	21(1)	17(1)	14(1)	-2(1)	-4(1)	2(1)
C(21)	14(1)	10(1)	13(1)	0(1)	-1(1)	-1(1)
C(22)	14(1)	10(1)	13(1)	0(1)	-1(1)	-1(1)
C(23)	12(1)	16(1)	19(1)	3(1)	-1(1)	-3(1)
C(24)	30(2)	32(2)	16(2)	5(1)	-10(1)	4(1)
C(1S)	49(5)	48(3)	40(6)	-3(4)	-20(5)	2(3)
C(2S)	31(3)	50(3)	16(3)	-4(3)	-8(3)	8(3)
C(7S)	30(3)	53(3)	21(4)	-2(3)	-3(3)	12(3)
C(6S)	36(3)	49(3)	26(3)	-5(3)	-4(3)	4(3)
C(5S)	45(4)	56(4)	34(5)	-4(4)	-7(4)	17(3)
C(4S)	42(4)	70(4)	24(4)	-7(3)	-2(3)	24(3)
C(3S)	25(4)	73(4)	16(4)	1(4)	-3(3)	16(3)

Table S11. Hydrogen coordinates ($\times 10^4$) and isotropic displacement parameters ($\text{\AA}^2 \times 10^3$) for Hughes_GS-VI-170.

	x	y	z	U(eq)
H(2)	4864	8276	7351	19
H(1)	3151	6862	3786	15
H(2A)	1652	7562	4781	15
H(2B)	1922	6857	5163	15
H(4)	2196	8150	6514	19
H(7)	9880	7135	6271	18
H(8)	8603	6610	5077	16
H(12A)	6328	8410	2853	41
H(12B)	8066	8293	3398	41
H(12C)	6532	8686	3821	41
H(15)	7326	5664	5808	15
H(18)	7491	3742	3132	23
H(19)	6137	4157	1919	25
H(24A)	10918	4667	7453	39
H(24B)	9161	4384	7811	39
H(24C)	9659	5111	7998	39
H(1SA)	8867	11099	4527	69
H(1SB)	10889	11114	4717	69
H(1SC)	9581	11159	5506	69

H(7S)	7465	10193	5591	42
H(6S)	7422	9084	5739	44
H(5S)	9727	8476	5251	54
H(4S)	12075	8977	4615	55
H(3S)	12117	10085	4467	46

Table S12. Hydrogen bonds for Hughes_GS-VI-170 [\AA and $^\circ$].

D-H...A	d(D-H)	d(H...A)	d(D...A)	$\angle(\text{DHA})$
N(2)-H(2)...N(3)#1	0.88	2.10	2.978(3)	172.0

Symmetry transformations used to generate equivalent atoms:

#1 $x, -y+3/2, z+1/2$

Conclusion

The work described in this dissertation spans the disciplines of natural products discovery, chemical biology, synthetic organic chemistry, and medicinal chemistry, as applied to marine bacterial natural products.

Chapter 1 describes the development of “Reactogenomics,” an extension of the established method of reactivity-guided isolation (RGI). The process of using a fluorogenic probe to screen for terminal alkyne-bearing natural products in extracts from genome-mined organisms was shown to be useful. However, the full pipeline was not used to discover a novel natural product. Instead, the fluorogenic probe was used to detect and characterize a specific genetically predicted natural product (vatiamide E). By detecting the natural product, which was produced in extremely low titer, it could be targeted for isolation in subsequent extractions and then characterized, helping to round out the family of predicted natural products. In addition, the preparation of the relevant probe was reported in the literature, where it previously might only have been accessible by custom synthesis.

Chapter 2 describes the ongoing efforts to develop rationally designed tetrazine-based probes for the detection of isonitrile-bearing natural products. After finding the 1,4-diaryl tetrazines to be unsatisfactory for our purposes, we turned to the preparation of a panel of more complex, functional, and tractable probes. By coupling a variety of amines to a common carboxylate-bearing tetrazine core, probes with enhanced MS and MS/MS properties were synthesized. The *N*-substituents of the resultant amides include halogenated benzyl groups, which fragment to halogenated tropylium ions; alkyl pyridines, which ionize and fragment well due to the presence of a basic amine; an alkyl morpholine, which shares basic amine logic with the pyridines; a quaternary ammonium salt, which bears a permanent charge and exhibits excellent ionization and fragmentation properties; and a coumarin, which may function as a fluorogenic probe. The immediate future of the project entails synthesis of a

brominated analog of the ethylpyridine-bearing probe, which would then bear all of the functions required of a tool for use in RGI. Next, the optimized probe will be validated on authentic isonitrile-bearing natural products, and then applied to extracts to screen for novel compounds.

Chapter 3 represents the transition from RGI and screening to the fuller characterization of a compound after its identification as a hit. After identification of neolymphostin A as an electrophilic natural product, it was submitted for biochemical screening and characterization. In a striking finding, neolymphostin A exhibits an apparent k_d of less than 10 nM (and in one instance, less than 1 nM) against the PI3K family of kinases. Further, it was shown by biochemical, mass spectrometry, and computational methods to be an irreversible inhibitor of these kinases. Its electrophilicity toward an active-site lysine residue was demonstrated, and the kinetics of its reaction with biologically relevant nucleophiles were determined and compared to other electrophiles. If the lymphostins are to be advanced as useful kinase inhibitors in the future, the electrophilicity of the sensitive vinylogous ester function will need to be attenuated. Strategies to achieve this might include appending the electrophile to a less electron-withdrawing moiety than the 2-position of a pyridine-type ring, or replacement with another electrophile altogether. These efforts may be focused further by observing that the vinylogous ester as it stands appears to have been evolved for the purpose of modifying primary amines inside pK_a -perturbed binding pockets.

Chapter 4 details an effort in the total synthesis of lymphostin. The actual chemistry was done before the striking activity of the neolymphostins (described in Chapter 3) was known, when the family was only suspected to covalently modify kinase targets. However, the rationale for a total synthesis follows logically from the results of Chapter 3, and the relevant paper in the *Journal of Organic Chemistry* was published after the results of Chapter 3 had been disclosed in the *Journal of Medicinal Chemistry*. When the synthetic campaign was being planned, it was envisioned that lymphostin could be synthesized by a more efficient route than the one already reported, especially with the benefit of more

than a decade of synthetic methods development in between. Ultimately, the campaign was unsuccessful, and the pitfalls we encountered may hint at what necessitated a seemingly inefficient, but successful, total synthesis in the first place. It is entirely possible that we were not the first to encounter some of the difficulties that arose in our attempt. While the campaign failed to progress to the more advanced stage of functionalizing the pyrroloquinoline core, an unusual oxidative dimerization reaction to yield a heterodimer was described. The structure of the product of this novel bit of organic chemistry was determined by X-ray crystallography, from which was returned a model depicting a striking twisted hexacyclic biaryl structure.

As with any doctoral work, there have been personal highs and lows. While my time at UC – San Diego has lasted six years, it is sometimes hard to remember that my time in the Hughes lab has only lasted four years. Most of the days of chemistry have been characterized by frustration, punctuated with bursts of satisfaction, excitement, curiosity, and, sometimes, joy. I have on several occasions related the emotional trajectory like this: when things are going badly, it is frustrating, and when things are going well, one simply thinks, “this is how it is supposed to go.” That seems to be the nature of doing synthetic organic chemistry.

It would, however, be inaccurate to say that the graduate experience as a whole was a net emotional negative. In terms of research, at least, there is a sudden income of satisfaction that comes with the completion of a project and the publication of a paper. This payback is not necessarily felt incrementally as each successive step is completed, but it does come in large, irregular doses as one sees a C.V. slowly filled out and a graduate career taking shape. Some of us take longer than others to cash in; all three of my graduate publications occurred in my final year of school. Patience was an easy virtue to hold for one year, but the next two were quite difficult. It is easy to give the advice of “just keep grinding, and things will work out,” but that line is much harder to hold when one is entering the sixth year of graduate studies with little to show for it.

Alternatively, a PhD does not entirely consist of research results and publications. It is, after all, an academic degree, and a great deal of education takes place over the course of earning it. In part this takes place in the form of coursework, but the majority of learning is done by one's own initiative. I have learned a great deal by asking questions of my advisor, of other professors, of postdoctoral researchers, and, perhaps most importantly, of other graduate students. I am proud to have been one of the graduate students from whom my peers have sought and received advice or technical know-how. I have learned by searching and reading on my own, and by trading literature back and forth with friends. The kind of learning that gives a doctorate its unique character is that which is done at the bench, late in the evening, with no resource but a willingness to try something. There is no substitute for learning something by giving it a try (and failing many times), just as there is no substitute for real experimental results. Earning a PhD is learning how to *just figure it out*. This is a difficult process to detect in real time as it slugs along. It is easy for imposter syndrome to set in as one comes to believe that they've spent several years failing over and over again for results that should have taken a month to obtain. "I've got a few reactions that have worked, and I don't feel like much better of a chemist than when I started." In organic synthesis, especially, it can be disheartening to line up one's progress against time, since the failed reactions are rarely mentioned, and the massive tree of routes that have been conceived and attempted will likely never be seen outside one's notebook.

Finally, there are the surrounding people who make the whole thing bearable. There are labmates who will become friends for life. There are people in neighboring groups who will grow just as close. It is these people who make it worth showing up every day; I can scarcely imagine a worse experience than being in lab every day with nobody else around, and having to go through the ordeal alone. These are people who can commiserate, who can say things will be alright, and who can convince you to get out of lab, because you need it. These people know who they are, and I hope to have given to them as much as they have given me. Without them, I would be nowhere.