BIOCHEMICAL METHODS FOR PREPARATION AND STUDY OF PEPTIDE NATURAL PRODUCT LIBRARIES

Steven Robert Fleming

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Pharmaceutical Sciences in the Doctoral Program of the UNC Eshelman School of Pharmacy (Division of Chemical Biology & Medicinal Chemistry)

Chapel Hill
2020

Approved by:

Albert A. Bowers

Michael B. Jarstfer

Rihe H. Liu

Kevin M. Weeks

Leslie M. Hicks

ABSTRACT

Steven Robert Fleming: Biochemical Methods for Preparation and Study of Peptide Natural Product Libraries
(Under the direction of Albert Bowers)

mRNA Display is an increasingly popular technique in pharmaceutical sciences to make highly diverse peptide libraries to pan for protein inhibitors. The current state of the art applies Flexizyme codon reprogramming with mRNA display to introduce unnatural amino acids for peptide cyclization and to further increase library diversity. Interestingly, ribosomally synthesized and post-translationally modified peptides (RiPP) are a unique class of natural products that transform linear peptides into highly modified and structurally complex metabolites. By combining RiPP biosynthesis with mRNA display, libraries of increasingly greater diversity can be achieved, and impending selected inhibitors will have natural product-like qualities, which we expect will allow these compounds to have better drug-like properties. Herein, we have developed a platform to measure RiPP enzyme modification of mRNA display libraries to show for the first time that RiPP enzymes can modify RNA linked peptide substrates. The platform may be extrapolated to many different RiPP enzymes and provides useful measurements to determine if a RiPP enzyme is promiscuous and effective to produce highly diversified peptides. Thiopeptides are a specific class of RiPP that we would like to apply to mRNA display, because they have broad activities. Additionally, the structure of thiopeptides is primed for mRNA display protocols, consisting of a highly decorated macrocycle with a C-terminal tail that may readily accept an

mRNA tag. We present a redesigned chemoenzymatic strategy to make thiopeptides using cell free protein synthesis that is promiscuous, can synthesize different thiopeptide classes, and is mRNA display ready. Finally, bioinformatic analysis of publicly available thiopeptide gene clusters shows that many unknown thiopeptides still exist. We have organized a set of 14 new clusters for which we can apply our chemoenzymatic strategy and characterize these unknown compounds, with the hope of finding the best thiopeptide enzymes for use in mRNA display selections.

To God

ACKNOWLEDGMENTS

I would first like to thank my family. Thank you for prioritizing my education and always encouraging me to work hard and do my best. I would not be here without you.

Thank you, Dr. Bowers, for taking me into the group, for training me, and for helping to shape my future. I am very happy and thankful to be a chemical biologist and time spent in your lab has been a delight.

A special thanks to all the Bowers' lab alumni whom I've had the pleasure to interact with: Dr. Scott Allen, Dr. Walter Wever, Dr. Rachel Bleich, Dr. Paul Himes, Dr. Swapnil Ghodge, Dr. Jon Bogart, Dr. Kelly Bird, Dr. Nick Kramer, Matt Fleming, Tory Haberman, Bree Iskandar, Matt Bowler, Jarrett Pelton, Braxton Cline, and Sungwan Hwang. Lab has been a great place to be, and that is because of you all. A doubly special thanks to Dr. Paul Himes for being my lab mentor and teaching me many basic research skills. Thank you for all of the hard work you put into the development of mRNA display.

I would like to thank those on my doctoral committee whose discussions and advice have been invaluable during my Ph.D. training. Thank you, Dr. Mike Jarstfer, for serving as my chair, providing letters of recommendation, and often reaching out with opportunities to present my work. Thank you, Dr. Rihe Liu, for allowing me to rotate in your lab where my interest in mRNA display began. Thank you, Dr. Kevin Weeks, for teaching me to write an abstract. Lastly, but not least, thank you, Dr. Leslie Hicks, for collaborating with me and giving me access to your mass spectrometers. This was absolutely vital for our first publication.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

µL	microliter
µM	micromolar
µm	micrometer
Å	angstrom (1 Å = 1 x 10^{-10} meters)
aa	amino acid
ATP	adenosine 5’-triphosphate
Boc	tert-butyloxycarbonyl
ClogP	partition coefficient between n-octanol and water
CV	column volume
DCM	dichloromethane
Dha	dehydroalanine
Dhb	dehydrobutyrine
DIEA	diisopropylethylamine
DMF	dimethyl formamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EIC	extracted ion chromatogram
ESI	electrospray ionization
Glu	glutamic acid

GluC	endoproteinase GluC
H_2O_2	hydrogen peroxide
HBA	hydrogen bond acceptor
HBD	hydrogen bond donor
HCl	hydrochloric acid
HEPES	N-(2-Hydroxyethyl)piperazine-N′-(2-ethanesulfonic acid)
HPLC	high performance liquid chromatography
IC_{50}	half-maximal inhibitory concentration
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal calorimetry
KCl	potassium chloride
K_d	dissociation contant
$KHPO_4$	potassium phosphate
KOH	potassium hydroxide
L	liter
LB	luria bertani
LC	liquid chromatography
M	molar
m / z	mass / charge
MALDI	matrix-assisted laser desorption ionization
MeCN	acetonitrile
MeOH	methanol

$MgCl_2$	magnesium chloride
MHz	megahertz
min	minute (s)
mM	millimolar
MS	mass spectrometry
MW	molecular weight
MWCO	molecular weight concentrator
N	Normal
Na_2SO_4	sodium sulfate
NaCl	sodium chloride
NaOAc	sodium acetate
ng	nanogram
NGS	next generation sequencing
nm	nanometer
NMR	nuclear magnetic resonance
NRotB	number of rotatable bonds
O.D.	optical density
^oC	degrees Celsius
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
P-linker	DNA-puromycin linker
PMSF	phenylmethylsulfonyl fluoride
ppm	parts per million

PSA	polar surface area
Q	quadrupole
RF	release factor
RiPPs	ribosomally synthesized and post-translationally modified peptides
RNA	ribonucleic acid
RRE	rip recognition element
RS	recognition sequences
SAR	structure activity relationship
Se	selenium
sec	second (s)
SecPhe	phenyl selenocysteine
smSVL	saturation mutagenesis single variant library
SPPS	solid phase peptide synthesis
TAMRA	5-(and 6)-carboxytetramethylrhodamine
TCEP	tris(2-carboxyethyl)phosphine
TFA	trifluoracetic acid
TIS	triisopropylsilane
TLC	thin layer chromatography
TOF	time of flight
Tris	tris(hydroxymethyl)aminomethane
UFLC	ultra fast liquid chromatography
WT	wild type

CHAPTER 1: INTRODUCTION

1.1 Modification of Genetically Encoded Peptide Libraries and Drug Discovery

In recent years, we have seen an explosion of peptide display technologies for the discovery of protein binding peptides.[1] Genetically encoded libraries, in particular, are useful because very large and diverse libraries can be made quickly, and amplification by PCR or organism replication enables several rounds of enrichment to be conducted to reveal unambiguous, high affinity binders illuminated by next generation sequencing.[2] It is becoming increasingly clear that modified peptides perform much better as drugs than unmodified peptides.[3–5] Altering a peptide inhibitor post selection often lowers the target binding affinity, and only sometimes can careful SAR studies lead to modified peptides which retain binding and become more bioavailable.[6,7] But, to avoid such arduous SAR studies, it may be prudent to modify peptide libraries before they are triaged for a specific activity. In this manner, greater diversity can be achieved leading to better target binding affinity, and hits can be preemptively biased for better drug-like properties. Indeed, much effort and vigor are being put into developing peptide display chemistries and enzymatic reactions to achieve this goal,[8–17] especially from those in the pharmaceutical industry.

1.2 mRNA Display Has Specific Advantages Over Other Display Technologies

Some of the most popular genetically encoded peptide panning technologies are mRNA, ribosomal, yeast, phage, and bacterial display. Each of these techniques can display a peptide for selection coupled to a genetic tag for identification. Furthermore, they have an advantage over DNA-encoded libraries because hits can be amplified and subjected to multiple rounds of selection.[18] For phage, yeast, and bacterial display, libraries are prepared by a living organism

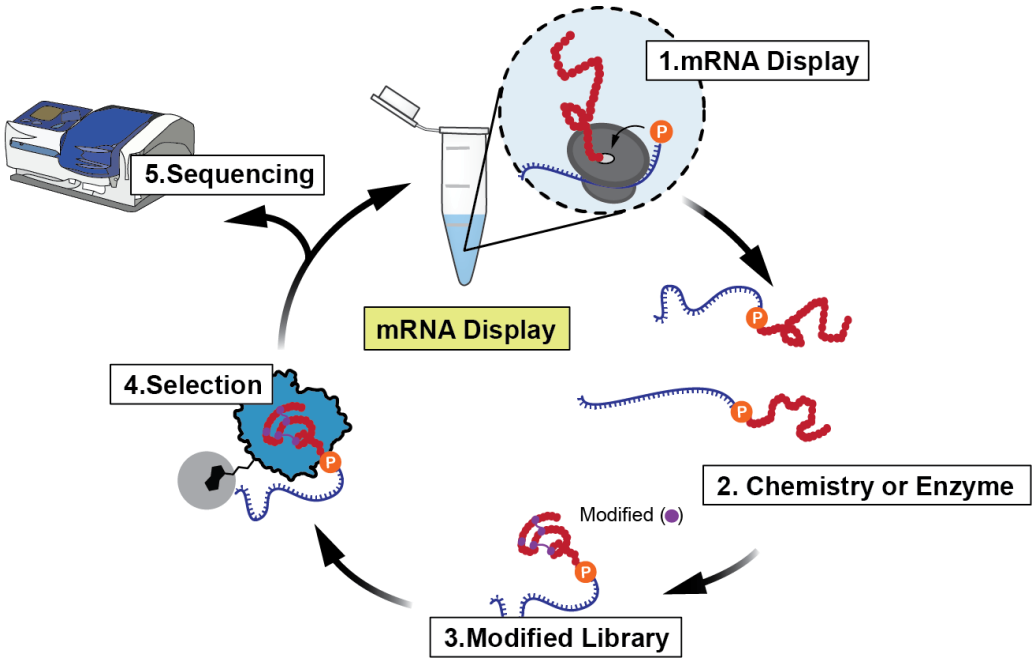


Figure 1.1. mRNA display of modified peptide libraries. (1) An mRNA library is translated to peptides by the ribosome. Each peptide is covalently linked to its cognate mRNA via puromycin (P). (2-3) Chemistry or enzymes can be used to modify peptides libraries. (4) Using an immobilized protein target, binders can be pulled out. (5) After multiple rounds, the mRNA library can be sequenced to reveal protein binders.

which immediately places a limit on library diversity due to transformation efficiency[18] (~10^9), and selections can be slow because organisms need time to replicate and grow. Additionally, very careful design is needed to make sure that the peptides are displayed in a manner that is available for panning and does not affect the livelihood of the organism, i.e. choosing the right protein within the viral coat to conjugate a peptide library to. Ribosomal display and mRNA display are valuable because they do not rely on a living organism for amplification, leading to a larger library diversity (~10^{13}). But with regards to displaying small peptides, mRNA display is the best, because a covalent bond is formed between the peptide and RNA.[19] Additionally, of all the genetically encoded display techniques, mRNA display has the smallest tag placed on the peptide, leading to less complications during the selection process, figure 1.1.

1.3 RiPP Biosynthesis and Application to mRNA Display.

Chemical modifications of peptide libraries are increasingly being used and validated, often focused on macrocyclization.[15,17,20–25] But the chemistries available remain simple and mild because they must be done in water, they must not harm living organisms, and they must not degrade RNA.[26] Conversely, enzymes are now stepping up to provide complex, regioselective, and stereoselective modifications that are inherently compatible with living organisms and RNA. Many of the enzymes used with peptide display come from promiscuous natural product biosynthesis. Indeed, a few recent reports show how lantibiotic cyclases can be used to screen libraries of lantibiotics with stereo- and regiospecific modifications;[12,13,27] a problem reported for chemically made lantibiotic libraries.[10]

Lantibiotics belong to a class of natural products called, "Ribosomally Synthesized and Post-translationally Modified Peptides" or RiPPs.[28] Biosynthesis of RiPPs begins when a small

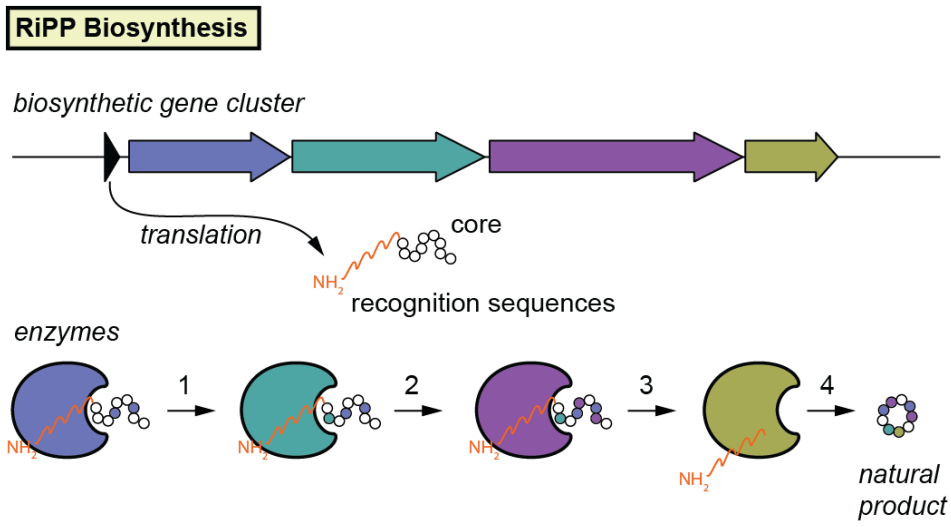


Figure 1.2. RiPP biosynthesis. A small gene called a precursor peptide containing a core and recognition sequences is translated by the ribosome. Subsequently, recognition sequences recruit biosynthetic enzymes to install chemical modifications into the core. Finally, the recognition sequences are removed, and the natural product is released.

gene is translated by the ribosome into a precursor peptide. This small peptide contains recognition sequences that recruit enzymes to install chemical modifications, and a core where all of the

chemical modifications are placed. Biosynthesis is finished when the core modifications are complete, and the recognition sequences are removed, figure 1.2. The use of recognition sequences to separate substrate affinity and substrate modification, gives RiPPs a heightened degree of promiscuity compared with other enzymes and allows RiPPs to be easily programmed for new substrates.[29] Because RiPP enzymes are promiscuous, programmable, and act on peptides, they are groomed for the transformation of peptide libraries. Lantibiotics represent the first example of this,[12,13,27] but there are many different RiPPs that can be applied in this manner. Application of new RiPPs has been slow because there are no good techniques to measure RiPP enzyme modification of a peptide library. Therefore, new experiments that quickly measure how well a RiPP can modify a diverse library of substrates would be of great use to couple more RiPPs with peptide display.

1.4 Thiopeptides are a Good Scaffold for mRNA Display Selections.

Thiopeptides are RiPPs that contain three major functional groups: azoles, dehydroalanine / dehydrobutyrine, and a central pyridine ring figure 1.3.[30] These are useful modifications to turn simple peptides into more bioavailable molecules, conforming towards "beyond rule of 5'" suggestions[31] by reducing molecular weight, hydrogen bond donors and acceptors, the number of rotatable bonds,[32] and polar surface area. Thiopeptides are privileged having a wide range of activities, such as antibacterial, antimalarial, anticancer, proteasome inhibition, transcription factor regulation, and mitophagy induction.[33–36] Privileged scaffolds are a good resource to discover new activities.[37] Finally, mature thiopeptides contain a highly decorated macrocycle, with an open C-terminal tail that we predict will readily accept an RNA tag for use in mRNA display, figure 1.3. Therefore, thiopeptides are primed for drug discovery and peptide display because they are RiPPs,

their inherent modifications can promote bioavailability, and because the scaffold is already privileged towards many bioactivities.

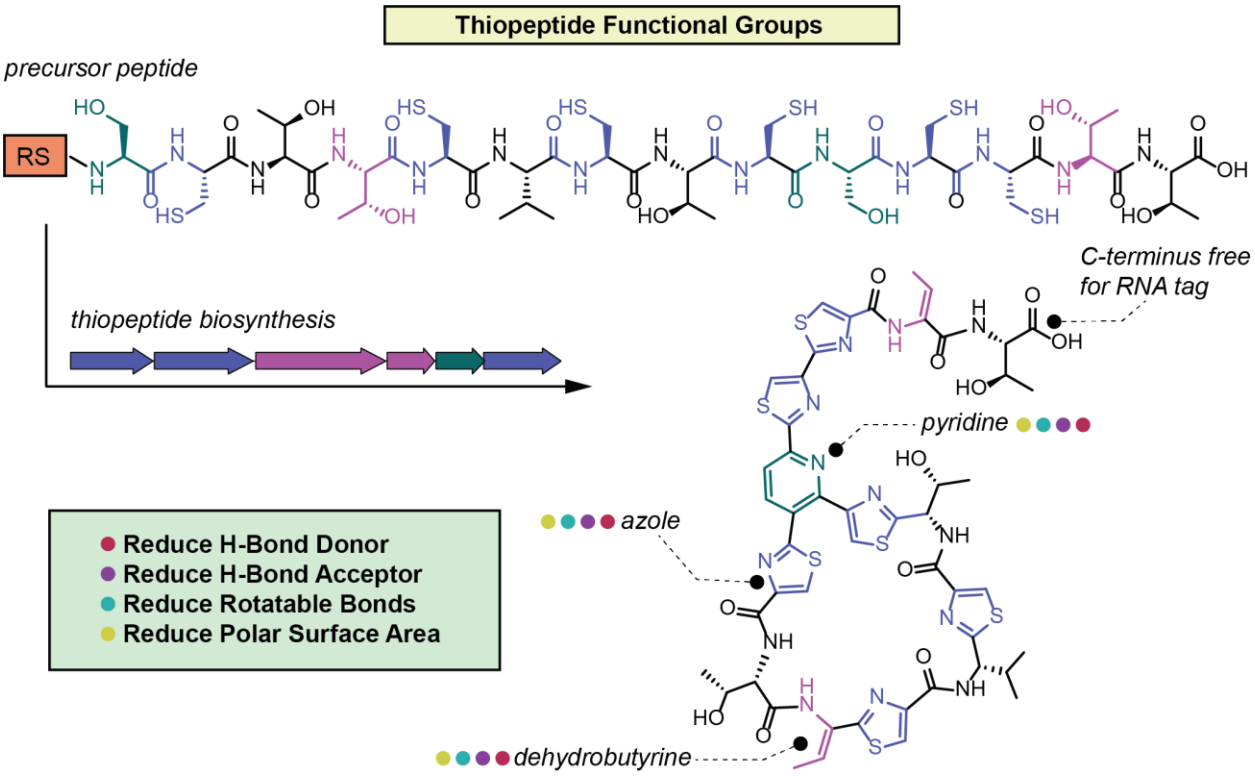


Figure 1.3. Thiopeptides are distinguished by three distinct functional groups. Thiopeptides are made from linear peptides by the installation of azoles, dehydroalanines / dehydrobutyrines, and a pyridine ring is formed during macrocyclization and RS removal. "RS" means recognition sequences.

The enzymes that make thiopeptides have been extensively studied.[38–51] In fact, the biosynthetic gene clusters for hallmark thiopeptides like Thiocillin, GE37468, and Thiostrepton have been engineered to make libraries of thiopeptide analogs, reinforcing the promiscuity of RiPPs.[32,42,52–57] But, while promiscuous to make analogs, a single thiopeptide BGC can only make its cognate class of thiopeptide, i.e. enzymes for thiocillin biosynthesis cannot make thiomuracins.[55] Additionally, the smallest thiopeptide cluster contains 5 enzymes to achieve the

characteristic backbone.[58] Often, these enzymes are found to have difficult cofactors or are unable to be expressed and purified from *E. coli*; this is an essential element for using enzymes in mRNA

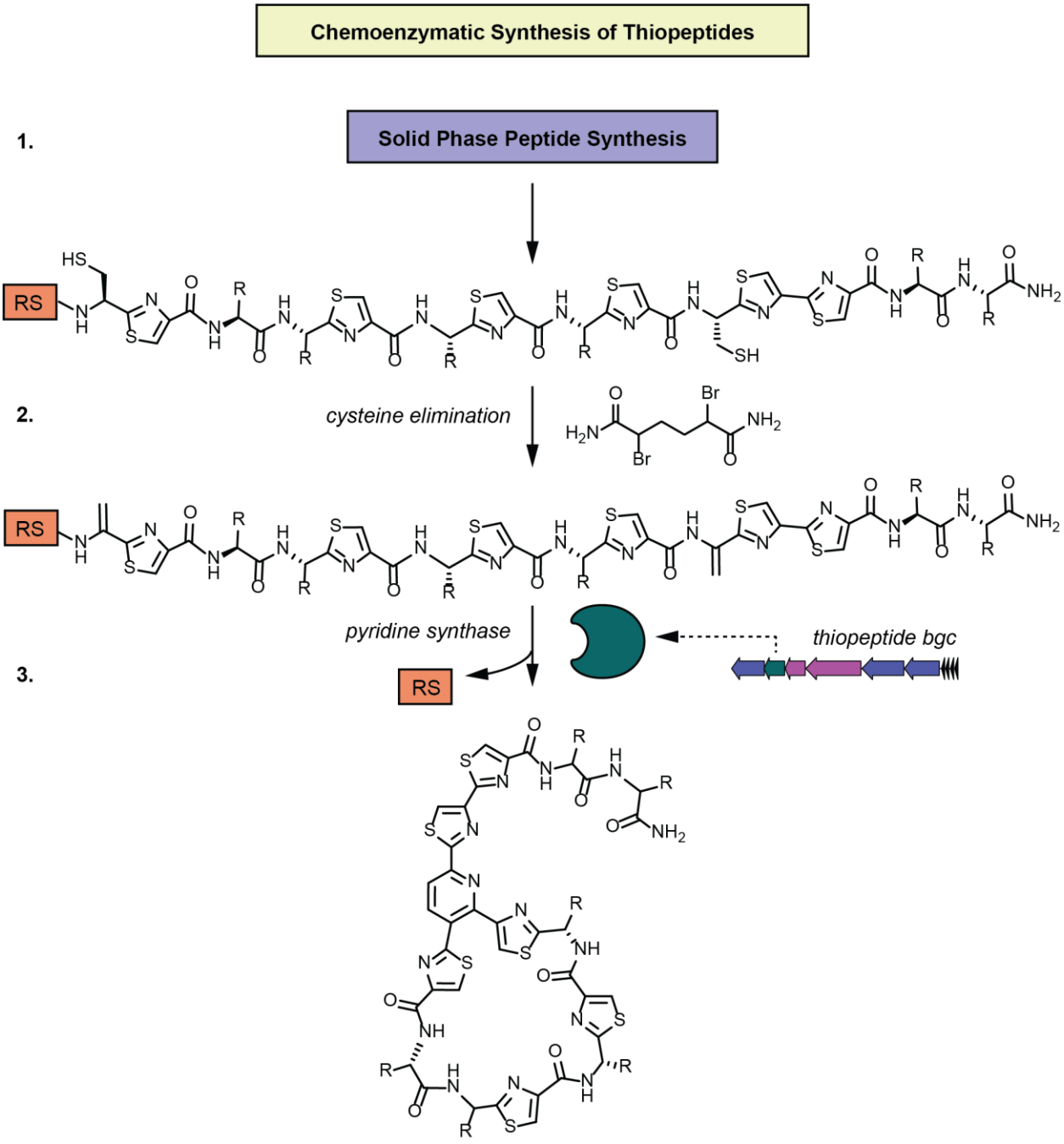


Figure 1.4. Chemoenzymatic synthesis of thiopeptides. (1) Azole containing peptides can be synthesized using solid phase peptide synthesis with a recognition sequence "RS" for a thiopeptide pyridine synthase. (2) With dibromohexaediamide, cysteines can be eliminated to dehydro-alanine. (3) A select pyridine synthase from thiopeptide biosynthesis can be expressed, purified, and used to cyclize the product from step 2 to make thiopeptides.

display. Because of this, a chemoenzymatic strategy which reduces the number of enzymes and allows the preparation of multiple structurally disparate thiopeptides could accelerate this class of natural product to best be used in peptide display technologies. Interestingly, a few recent reports use chemoenzymatic synthesis to make different thiopeptides by using only a single enzyme from the natural BGC.[43,55,59–61] This enzyme is the pyridine synthase which performs the final macrocyclization and installs the hallmark pyridine while eliminating recognition sequences, figure 1.4. Currently this chemoenzymatic synthesis is not amenable to peptide display, but, with some clever alterations it could be.

1.5 Bioinformatic Expansion of Thiopeptides and Chemoenzymatic Synthesis Could Lead to Optimal Enzymes for mRNA Display of Thiopeptides.

Most thiopeptides have been discovered using a "top down" approach; meaning the compound was discovered first based on an observed activity, extracted, and then structurally characterized without knowledge of the biosynthetic gene cluster.[62] With the rapid speed at which new genomes have been sequenced and stored in GenBank, swaths of genomic information are now available for a "bottom up" approach to discover new thiopeptides. By applying bioinformatic tools like BLAST, RODEO, and antiSMASH biosynthetic gene clusters of specific natural products can be targeted and collected. Additionally, with the help of sequence similarity networks, the diversity found therein can be visualized. This can greatly aid the discovery of new natural products, especially RiPPs, and prioritize gene clusters for characterization.

A key enzyme to chemoenzymatically make thiopeptides compatible with mRNA display is the pyridine synthase. To date there are only 5 pyridine synthases characterized in vitro: TclM (thiocillin),[59] TbtD (thiomuracin),[39] PbtD (thiomuracin),[39] and LazC (lactazole),[63] and MroD (pyritide).[61] While these enzymes may or may not work well for mRNA display, a thorough bioinformatics analysis of thiopeptide diversity may reveal the enzymes that do. For example,

there are many lantibiotic cyclases. One in particular, ProcM, has 29 diverse but natural substrates,[64] an indication of heightened promiscuity. Indeed, promiscuity was confirmed[65] and now ProcM has been used successfully in peptide display campaigns.[12,27]

In 2018 a report showed that the unknown chemical space of thiopeptides is much more diverse than what is currently known.[40] Furthermore, the same group showed the strength of chemoenzymatic synthesis by characterizing the unknown pyridine synthase, MroD, without the preceding associated enzymes.[61] This demonstrates the usefulness and power of combining bioinformatics with chemoenzymatic synthesis to illuminate new thiopeptides. But, what the 2018 report failed to do is provide any useful analysis and prioritization of new thiopeptides for chemoenzymatic discovery. Ergo, a new bioinformatic endeavor is warranted that collects what new thiopeptides that have been added to GenBank since 2018 and provides analysis of interesting new thiopeptides that may be of particular use in chemoenzymatic discovery and mRNA display.

REFERENCES

(1) Sohrabi, C.; Foster, A.; Tavassoli, A. *Nat. Rev. Chem.* **2020**, *4*, 90.

(2) Blanco, C.; Verbanic, S.; Seelig, B.; Chen, I. A. *Phys. Chem. Chem. Phys.* **2020**, *22*, 6492.

(3) Hewitt, W. M.; Leung, S. S. F.; Pye, C. R.; Ponkey, A. R.; Bednarek, M.; Jacobson, M. P.; Lokey, R. S. *J. Am. Chem. Soc.* **2015**, *137*, 715.

(4) Pye, C. R.; Hewitt, W. M.; Schwochert, J. A.; Haddad, T. D.; Townsend, C. E.; Etienne, L.; Lao, Y.; Limberakis, C.; Furukawa, A.; Mathiowetz, A. M.; Price, D. A.; Liras, S.; Lokey, S. R. *J. Med. Chem.* **2017**, *60*, 1665.

(5) Rand, A. C.; Leung, S. S.; Eng, H.; Rotter, C. J.; Sharma, R.; Kalgutkar, A. S.; Zhang, Y.; Varma, M. V.; Farley, K. A.; Khunte, B.; Limberakis, C.; Price, D. A.; Liras, S.; Mathiowetz, A. M.; Jacobson, M. P.; Lokey, S. R. *Med. Chem. Comm.* **2012**, *3*, 1282.

(6) Boehm, M.; Beaumont, K.; Jones, R. M.; Kalgutkar, A. S.; Zhang, L.; Atkinson, K.; Bai, G.; Brown, J. A.; Eng, H.; Goetz, G. H.; Holder, B. R.; Khunte, B.; Lazzaro, S.; Limberakis, C.; Ryu, S.; Shapiro, M. J.; Tylaska, L.; Yan, J.; Turner, R.; Leung, S. S.; Ramaseshan, M.; Price, D. A.; Liras, S.; Jacobson, M. P.; Earp, D. J.; Lokey, S. R.; Mathiowetz, A. M.; Menhaji-Klotz, E. *J. Med. Chem.* **2017**, *60*, 9653.

(7) Rogers, J. M.; Passioura, T.; Suga, H. *Proc. Natl. Acad. Sci.* **2018**, *115*, 10959.

(8) He, B.; Tjhung, K. F.; Bennett, N. J.; Chou, Y.; Rau, A.; Huang, J.; Derda, R. *Sci. Rep.* **2018**, *8*, 1214.

(9) Tjhung, K. F.; Kitov, P. I.; Ng, S.; Kitova, E. N.; Deng, L.; Klassen, J. S.; Derda, R. *J. Am. Chem. Soc.* **2016**, *138*, 32.

(10) Hofmann, F. T.; Szostak, J. W.; Seebeck, F. P. *J. Am. Chem. Soc.* **2012**, *134*, 8038.

(11) Schlippe, Y. V.; Hartman, M. C.; Josephson, K.; Szostak, J. W. *J. Am. Chem. Soc.* **2012**, *134*, 10469.

(12) Yang, X.; Lennard, K. R.; He, C.; Walker, M. C.; Ball, A. T.; Doigneaux, C.; Tavassoli, A.; van der Donk, W. A. *Nat. Chem. Biol.* **2018**, *14*, 375.

(13) Hetrick, K. J.; Walker, M. C.; van der Donk, W. A. *ACS Cent. Sci.* **2018**, *4*, 458.

(14) Iqbal, E. S.; Richardson, S. L.; Abrigo, N. A.; Dods, K. K.; Franco, H. E. O.; Gerrish, H. S.; Kotapati, H. K.; Morgan, I. M.; Masterson, D. S.; Hartman, M. C. T. *Chem. Commun.* **2019**, *55*, 8959.

(15) Hacker, D. E.; Abrigo, N. A.; Hoinka, J.; Richardson, S. L.; Przytycka, T. M.; Hartman, M. C. T. Direct, *ACS Comb. Sci.* **2020**, *22*, 306.

(16) Passioura, T.; Suga, H. *Chem. Comm.* **2017**, *53*, 1931.

(17) Kale, S. S.; Villequey, C.; Kong, X.-D.; Zorzi, A.; Deyle, K.; Heinis, C. *Nat. Chem.* **2018**, *10*, 715.

(18) Bozovičar, K.; Bratkovič, T. *Int. J. Mol. Sci.* **2019**, *21*, 215.

(19) Huang, Y.; Wiedmann, M. M.; Suga, H. *Chem. Rev.* **2019**, *119*, 10360.

(20) Millward, S. W.; Takahashi, T. T.; Roberts, R. W. *J. Am. Chem. Soc.* **2005**, *127*, 14142.

(21) Wang, X. S.; Chen, P. C.; Hampton, J. T.; Tharp, J. M.; Reed, C. A.; Das, S. K.; Wang, D.; Hayatshahi, H. S.; Shen, Y.; Liu, J.; Liu, W. R. *Angew. Chem. Int. Ed.* **2019**, *58*, 15904.

(22) Kawakami, T.; Ogawa, K.; Hatta, T.; Goshima, N.; Natsume, T. *ACS Chem. Biol.* **2016**, *11*, 1569.

(23) Kawakami, T.; Ishizawa, T.; Fujino, T.; Reid, P. C.; Suga, H.; Murakami, H. *ACS Chem. Biol.* **2013**, *8*, 1205.

(24) Katoh, T.; Goto, Y.; Reza, Md.; Suga, H. *Chem. Comm.* **2011**, *47*, 9946.

(25) Peraro, L.; Zou, Z.; Makwana, K. M.; Cummings, A.; Ball, H. L.; Yu, H.; Lin, Y.-S.; Levine, B.; Kritzer, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 7792.

(26) Malone, M. L.; Paegel, B. M. *ACS Comb. Sci.* **2016**, *18*, 182.

(27) Urban, J. H.; Moosmeier, M. A.; Aumüller, T.; Thein, M.; Bosma, T.; Rink, R.; Groth, K.; Zulley, M.; Siegers, K.; Tissot, K.; Moll, G. N.; Prassler, J. *Nat. Comm.* **2017**, *8*, 1500.

(28) Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; Heel, A. J. van; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Burgos, M. J. G.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; Wezel, G. P. van; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A. *Nat. Prod. Rep.* **2020**, Advance Article

(29) Burkhart, B. J.; Kakkar, N.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A. *ACS Cent. Sci.* **2017**, *3*, 629.

(30) Vinogradov, A. A.; Suga, H. *Cell Chem. Biol.* **2020**, *27*, 1032.

(31) Matsson, P.; Doak, B. C.; Over, B.; Kihlberg, J. *Adv. Drug Deliver. Rev.* **2016**, *101*, 42.

(32) Tran, H. L.; Lexa, K. H.; Julien, O.; Young, T. S.; Walsh, C. T.; Jacobson, M. P.; Wells, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 2541.

(33) Just-Baringo, X.; Albericio, F.; Álvarez, M. *Marine Drugs* **2014**, *12*, 317.

(34) Bird, K. E.; Xander, C.; Murcia, S.; Schmalstig, A. A.; Wang, X.; Emanuele, M. J.; Braunstein, M.; Bowers, A. A. *ACS Chem. Biol.* **2020**, *15*, 2164.

(35) Hegde, N. S.; Sanders, D. A.; Rodriguez, R.; Balasubramanian, S. *Nat. Chem.* **2011**, *3*, 725.

(36) Habazettl, J.; Allan, M.; Jensen, P. R.; Sass, H.-J.; Thompson, C. J.; Grzesiek, S. *Proc. Natl. Acad. Sci.* **2014**, *111*, E5498.

(37) Welsch, M. E.; Snyder, S. A.; Stockwell, B. R. *Curr. Opin. Chem. Biol.* **2010**, *14*, 347.

(38) Zhang, Z.; Hudson, G. A.; Mahanta, N.; Tietz, J. I.; van der Donk, W. A.; Mitchell, D. A. *J. Am. Chem. Soc.* **2016**, *138*, 15511.

(39) Hudson, G. A.; Zhang, Z.; Tietz, J. I.; van der Donk, W. A.; Mitchell, D. A. *J. Am. Chem. Soc.* **2015**, *137*, 16012.

(40) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. *J. Am. Chem. Soc.* **2018**, *140*, 9494.

(41) Cogan, D. P.; Hudson, G. A.; Zhang, Z.; Pogorelov, T. V.; van der Donk, W. A.; Mitchell, D. A.; Nair, S. K. *Proc. Natl. Acad. Sci.* **2017**, *114*, 12928.

(42) Duan, L.; Wang, S.; Liao, R.; Liu, W. *Chem. Biol.* **2012**, *19*, 443.

(43) Bogart, J. W.; Kramer, N. J.; Turlik, A.; Bleich, R. M.; Catlin, D. S.; Schroeder, F. C.; Nair, S. K.; Williamson, R. T.; Houk, K. N.; Bowers, A. A. *J. Am. Chem. Soc.* **2020**, *142*, 13170.

(44) Zheng, Q.; Wang, S.; Liao, R.; Liu, W. *ACS Chem. Biol.* **2016**, *11*, 2673.

(45) Liao, R.; Duan, L.; Lei, C.; Pan, H.; Ding, Y.; Zhang, Q.; Chen, D.; Shen, B.; Yu, Y.; Liu, W. *Chem. Biol.* **2009**, *16*, 141.

(46) Kelly, W. L.; Pan, L.; Li, C. *J. Am. Chem. Soc.* **2009**, *131*, 4327.

(47) Pierre, S.; Guillot, A.; Benjdia, A.; Sandström, C.; Langella, P.; Berteau, O. *Nat. Chem. Biol.* **2012**, *8*, 957.

(48) Yu, Y.; Duan, L.; Zhang, Q.; Liao, R.; Ding, Y.; Pan, H.; Wendt-Pienkowski, E.; Tang, G.; Shen, B.; Liu, W. *ACS Chem. Biol.* **2009**, *4*, 855.

(49) Badding, E. D.; Grove, T. L.; Gadsby, L. K.; LaMattina, J. W.; Boal, A. K.; Booker, S. J. *J. Am. Chem. Soc.* **2017**, *139*, 5896.

(50) Ichikawa, H.; Bashiri, G.; Kelly, W. L. *J. Am. Chem. Soc.* **2018**, *140*, 10749.

(51) Du, Y.; Qiu, Y.; Meng, X.; Feng, J.; Tao, J.; Liu, W. *J. Am. Chem. Soc.* **2020**, *142*, 8454.

(52) Young, T. S.; Dorrestein, P. C.; Walsh, C. T. *Chem. Biology.* **2012**, *19*, 1600.

(53) Bowers, A. A.; Acker, M. G.; Young, T. S.; Walsh, C. T. *J. Am. Chem. Soc.* **2012**, *134*, 10313.

(54) Acker, M. G.; Bowers, A. A.; Walsh, C. T. *J. Am. Chem. Soc.* **2009**, *131*, 17563.

(55) Wever, W. J.; Bogart, J. W.; Bowers, A. A. *J. Am. Chem. Soc.* **2016**, *138*, 13461.

(56) Bowers, A. A.; Acker, M. G.; Koglin, A.; Walsh, C. T. *J. Am. Chem. Soc.* **2010**, *132*, 7519.

(57) Zhang, F.; Kelly, W. L. *ACS Chem. Biol.* **2015**, *10*, 998.

(58) Vinogradov, A. A.; Shimomura, M.; Goto, Y.; Ozaki, T.; Asamizu, S.; Sugai, Y.; Suga, H.; Onaka, H. *Nat. Commun.* **2020**, *11*, 2272.

(59) Wever, W. J.; Bogart, J. W.; Baccile, J. A.; Chan, A. N.; Schroeder, F. C.; Bowers, A. A. *J. Am. Chem. Soc.* **2015**, *137*, 3494.

(60) Bogart, J. W.; Bowers, A. A. *J. Am. Chem. Soc.* **2019**, *141*, 1842.

(61) Hudson, G. A.; Hooper, A. R.; DiCaprio, A. J.; Sarlah, D.; Mitchell, D. A. *Org. Lett.* **2020**, XXXX, XXX, XXX-XXX

(62) Pye, C. R.; Bertin, M. J.; Lokey, S. R.; Gerwick, W. H.; Linington, R. G. *Proc. Natl. Acad. Sci.* **2017**, *114*, 5601.

(63) Fleming, S. R.; Bartges, T. E.; Vinogradov, A. A.; Kirkpatrick, C. L.; Goto, Y.; Suga, H.; Hicks, L. M.; Bowers, A. A. *J. Am. Chem. Soc.* **2019**, *141*, 758.

(64) Li, B.; Sher, D.; Kelly, L.; Shi, Y.; Huang, K.; Knerr, P. J.; Joewono, I.; Rusch, D.; Chisholm, S. W.; van der Donk, W. A. *Proc. Natl. Acad. Sci.* **2010**, *107*, 10430.

(65) Thibodeaux, C. J.; Ha, T.; van der Donk, W. A. *J. Am. Chem. Soc.* **2014**, *136*, 17513.

CHAPTER 2: EXPLORING THE POST-TRANSLATIONAL ENZYMOLOGY OF PAAA BY MRNA DISPLAY[1]

2.1 Introduction

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are an exciting family of natural products that have seen a surge in research due to exceptional versatility of their biosynthetic enzymes.[1] Recent studies have shown that, in many cases, RiPP chemistry is guided by specific interactions between a RiPP recognition element (RRE) and the peptide substrate.[2–6] The RRE binds one region of the substrate and feeds the remainder into the catalytic domain for modification. This modular strategy allows for remarkable substrate promiscuity, and as a result, RiPP enzymes are increasingly being exploited to make complex libraries of peptide derivatives.[7–16] Most recently, RiPPs have been paired with phage and yeast display to identify high affinity, RiPP-based peptide binders.[17–19] While this work hints at the extraordinary promiscuity and generality of RiPP enzymes, verifying and measuring enzymatic modification of library members remains a significant challenge. Indeed, in many display libraries where chemical modification is used, including phage,[20,21] mRNA display,[22] and even DNA-encoded libraries (DELs),[23,24] panning hits wind up being the most significant evidence that chemistry worked. In the case of RiPP enzymes, quantitative assessment, and comparison of

[1]This chapter was adapted from a communication that previously appeared in the Journal of the American Chemical Society. The original citation is as follows: Fleming, S.R., Himes, P.M., Ghodge, S.V., Goto, Y., Suga, H., Bowers, A.A. *J. Am. Chem. Soc.*, **2020**, *142* (11), pp 5024-5028.

enzymatic modifications on large display libraries could inform broadly on substrate promiscuity and mechanistic features of these enzymes, leading the way to their more effective implementation in benchtop campaigns.

Inspired by several reports of high-throughput assays for proteases and ligases,[25–28] we envisioned a coupled mRNA display assay that might allow expedient study of RiPP enzymology and promiscuity. mRNA display is a powerful peptide display technology where peptides are linked to their own encoding RNA during in vitro ribosomal translation.[29] mRNA display allows the easy incorporation of non-natural amino acids by Flexizyme reprogramming,[30] which could aid in our assay design and significantly expand library diversity one day. To substantiate mRNA display as a biochemical tool to study RiPPs, we first chose PaaA, an RRE-containing RiPP enzyme from the biosynthesis of the antibiotic Pantocin A (figure 2.1a, **1**).[31–33] PaaA is a ThiF/E1-like activating enzyme that catalyzes the double dehydration and decarboxylation of two glutamic acid residues (E16 and E17) in substrate peptide, PaaP, to form the fused-bicyclic core of the active tripeptide natural product. Interestingly, the enzymatic mechanism may proceed by either of two imide intermediates to achieve the same product (figure 2.1b).[34] Since PaaA chemically modifies two glutamic acids, we could use the indiscriminate Glu-protease GluC to cleave unmodified substrate analogues in an mRNA display library and report on PaaA activity. This approach would confirm and measure PaaA modification of substrate peptide mutants and constitute a first application of RiPP enzymes to mRNA display libraries.

2.2 Results and Discussion

To implement this approach, we first sought to confirm that PaaA could modify an mRNA displayed PaaP substrate: we established a gel shift assay, wherein treatment of 35S-Met-labeled and mRNA displayed-PaaP with PaaA rendered the peptide resistant to GluC cleavage (Appendix

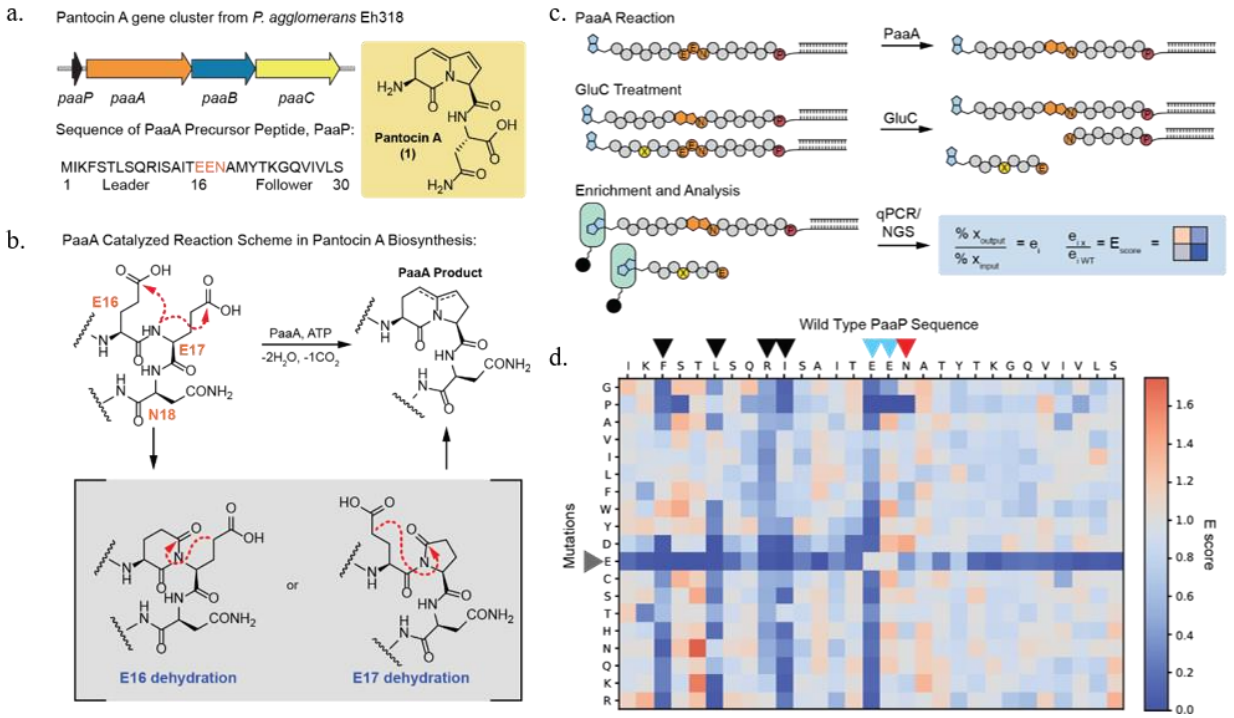


Figure 2.1. Pantocin A biosynthesis and mRNA display activity assay to study PaaA. (a) Pantocin A gene cluster, including paaA, an ATP-dependent ThiF/E1-like enzyme; paaB, α-ketogluterate-dependent iron oxidase; paaC, efflux pump; and paaP, precursor peptide. (b) Potential mechanisms for PaaA catalyzed modification of PaaP. (c) Overview of mRNA display assay with N-terminal biotinylated (blue) PaaP linked to mRNA via puromycin "P". Modified substrates are enriched by streptavidin (green) pulldown and analyzed by qPCR or next generation sequencing (NGS). "x" is any given peptide, "wt" is wild type. (d) E scores for each single mutant PaaP peptide analyzed after PaaA activity selection. Blue indicates sequence enrichment worse than WT, and orange indicates sequence enrichment better than WT. Each square represents an average (n = 3).

figure A.1a). To further optimize reaction conditions with mRNA display substrates and allow for enrichment of modified peptides from larger libraries, we adapted this assay to include affinity purification and qPCR-based quantitation (figure 2.1c). This new iteration of the assay involves four key steps: (1) display of an N-terminal biotinylated PaaP substrate via Flexizyme codon reprogramming, (2) treatment with PaaA, (3) protease treatment with GluC to remove RNA tags from unmodified substrates, and finally (4) streptavidin enrichment of only modified substrates. The assay was optimized for GluC cleavage conditions, PaaA concentration, and reaction time (Appendix figure A.1b-d). With a functioning activity assay in hand we next designed a saturation mutagenesis single variant library (smSVL) to explore how point mutations along PaaP would

affect PaaA activity. The library was treated with 1 μM PaaA for 5, 22.5, and 60 min before GluC treatment and streptavidin purification. Recovered sequences were PCR amplified and submitted for next generation sequencing (NGS). The sequencing data was processed by calculating E scores (figure 2.1c),[35] which are normalized values for the success of each mutant compared to wild type (WT). In this analysis a score >1 is better than WT and a score < 1 is worse than WT (figure 2.1d, Appendix figure A.2 and A.3). As an important assay control, we examined Glu mutants within the leader and follower. Because these mutations lie outside the core, we anticipate that they will go unmodified and thus be readily cleaved by GluC irrespective of PaaA activity at E16 and E17. Satisfyingly, these Glu mutants have consistently low E scores, demonstrating that the assay selection conditions are stringent (figure 2.1d, gray arrow). Beyond this internal control four key trends are readily apparent. (1) First, PaaA is broadly tolerant of point mutations. Of the 26 positions within the leader and follower probed by saturation mutagenesis, 22 appear numb to mutation. This is especially evident in the follower sequence as no single mutations strongly inhibited PaaA processing. (2) While the leader sequence also shows high tolerance to mutations, four positions (F4, L7, R10, and I11) consistently enriched poorly when mutated (figure 2.1d, black arrows). With increasing PaaA incubation time this result lessened suggesting that these mutations slow the reaction but do not prevent it. However, aspartic acid replacements at these positions appear to severely inhibit PaaA processing (figure 1d and Appendix figure A.3ab). Given that PaaA exhibits an RRE domain, we speculate that this FXXLXXRI motif may be involved in substrate recognition at the RRE / peptide interface where backbone hydrogen bonding typically drives β-strand interactions and hydrophobic residues fill key hotspots.[2–6] (3) The smSVL data also shows that while E16 mutations were highly susceptible to GluC cleavage, E17 mutants were strongly protected implying modification of E16 (figure 2.1d, blue arrows). These results suggest

that PaaA may exhibit a preference for activation of E16 over E17, which may further play a role in the mechanism of modification (figure 2.1b). (4) Finally, core residue N18 also displays a high tolerance to mutation (figure 2.1d, red arrow), suggesting that PaaA can be used to make new Pantocin A analogues.

To validate selection results, we chose several mutants for confirmation by in vitro enzymatic assay and mass spectrometric characterization. Each substrate was translated with NEB PURExpress, then treated with 1 μM PaaA, and reaction products were analyzed by MALDI-TOF. As seen in the smSVL data, PaaA proved promiscuous to point mutations within the leader and follower peptides (figure 2.2a, Appendix figure A.9, Appendix Table A.1). Notably, N18D was processed to the final bicycle demonstrating that PaaA can be used for preparation of novel Pantocin A analogues. Also, aligned with the smSVL data, qualitative analysis shows that PaaA struggles to process F4D, L7D, R10D, and I11D mutants, and the quadruple alanine mutant went completely unmodified. In contrast a quadruple mutant outside these four residues was accepted (K3A, T6A, S8A, T15A). These data further support that PaaA has a wide tolerance for single mutations and strengthens the importance of F4, L7, R10, and I11 for PaaA reactivity. To investigate whether these four residues are involved in PaaA recognition, we prepared synthetic variants of the N-terminal 12 residues of PaaP and a TAMRA-labeled fluorescent probe that could be used for competitive fluorescence polarization (FP) assays. In these FP assays, only WT and T6D peptides competed effectively with TAMRA-WT for PaaA binding (figure 2.2b and Appendix figure A.4). This was further substantiated by isothermal titration calorimetry (ITC, figure 2.2b and Appendix figure A.5). Together, the MALDI-TOF data and binding experiments confirm that PaaA is a promiscuous RiPP enzyme and that these four residues are essential for PaaP binding and subsequent processing. A BLAST of available PaaP homologues

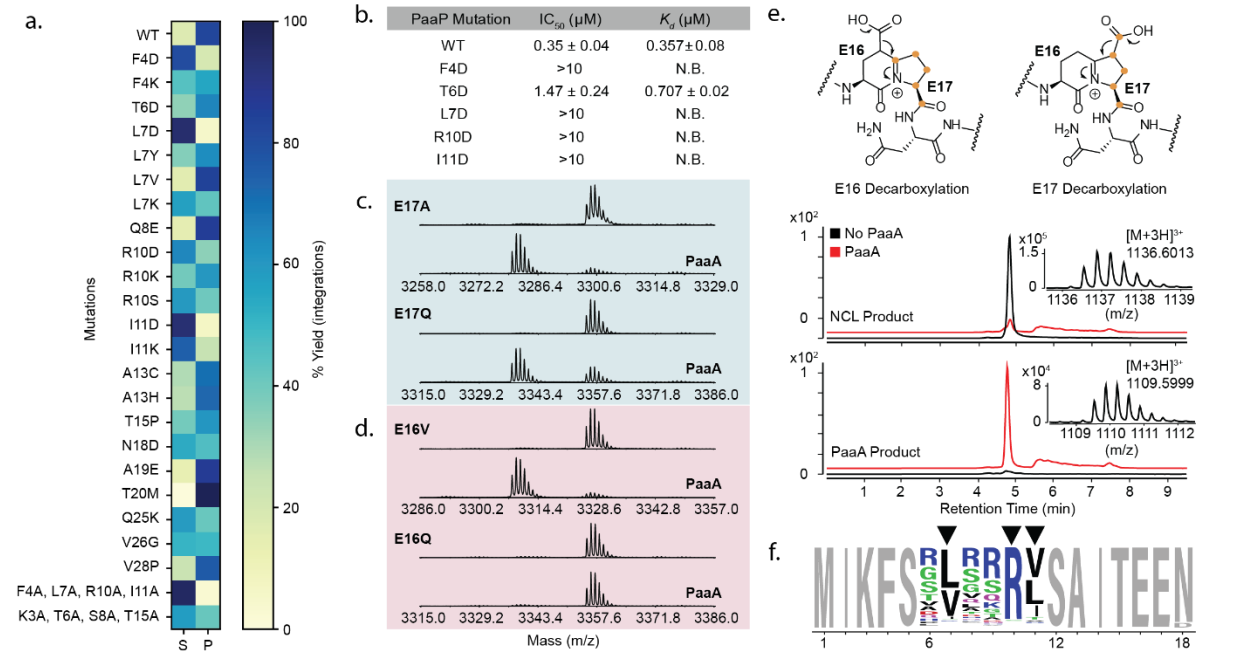


Figure 2.2. PaaA substrate promiscuity, binding, and processing. (a) MALDI-TOF analysis of PaaA activity on PaaP mutants. Total extracted ion integration areas for substrate remaining "S" and observed products "P" were summated and used to calculate percent substrate or products for display in a heatmap. (b) PaaA binding to PaaP measured by competitive fluorescence polarization (FP, mean and error for n = 3) and isothermal calorimetry (ITC, mean and error for n = 2). (c,d) MALDI-TOF analysis of PaaA catalyzed dehydration (−18 Da) of PaaP core E mutants. (e) PaaP A13C was prepared by native chemical ligation with 13C-labeled E17. Orange dots indicate labeled carbons. Reaction of E17 labeled PaaP A13C with PaaA leads to a mass shift of 81 Da indicating loss of a single 13C atom due to E17 decarboxylation. (f) Weblogo analysis of 996 most enriched sequences with PaaP NNK 6.

shows significant conservation across the central third of the precursor peptide (L7-Y21, Appendix figure A.6). F4, L7, R10, and I11 are heavily conserved, as are many others, thus highlighting the power of this approach in effectively discriminating between essential residues where bioinformatics could not.

We next turned to the difference in enrichment between E16 and E17 mutants (figure 2.1d, blue arrows). To better understand this result, we first validated several core E mutants in MALDI-TOF assays. Notably, as exemplified by E17A and E17Q (figure 2.2c), E17 mutants were readily modified to a single dehydration product, suggesting partial processing to putative imide intermediates (e.g., figure 2.1b). In good agreement with the selection data, only hydrophobic

mutations to E16 are converted to this intermediate (figure 2.2d and Appendix Table A.2). The mechanism of PaaA can be reasonably written from a first step condensation of either E16 or E17, and both routes could yield the same bicyclic Pantocin A core (figure 2.1b and Appendix figure A.7). If E16 is condensed first, then E17 must undergo decarboxylation and vice versa. The general rejection of E16 mutations suggests that the E16 side chain is the preferred initial substrate. To confirm this biosynthetic timing, we prepared a PaaA substrate selectively 13C-enriched at E17 **(**Appendix figure A.8). Treatment with PaaA transformed the peptide into the product with a mass shift consistent with loss of a single 13C label (figure 2.2e, E17 Decarboxylation, Appendix Table A.3). The loss of a 13C label agrees with a mechanism in which E16 is modified first, followed by E17 cyclization and decarboxylation.

The smSVL experiment proved informative for studying point mutations, but the power of mRNA display is its capacity to screen much larger libraries, containing multiple simultaneous mutations. To show that RiPPs and mRNA display might be compatible with larger and more diverse libraries, we prepared a ~34 million-member library where the 6 positions from T6-I11 were simultaneously randomized with NNK codons. This library was treated with 1 μM PaaA before purification, GluC treatment, and streptavidin pull-down. After enrichment, the recovered DNA was PCR amplified, submitted for NGS and analyzed by Weblogo. Significantly, positions 7, 10, and 11 showed strong enrichment to a generalized version of the natural epitope evinced by smSVL experiments: FXXBXXRB (B = V, L, or I**,** figure 2.2f). Additionally, this data shows that other residues are significantly less important for substrate recognition and processing. Such broad permissiveness has previously been hypothesized in RRE substrate interactions but not validated to this extent.

2.3 Conclusion

In summary, we successfully deployed an mRNA display based enzyme activity assay to study the RiPP enzyme PaaA. This assay provides rapid insight into the broad promiscuity, sequence dependent substrate recognition, and residue specific processing of the PaaP core all in one set of experiments. The results suggest that PaaA is broadly promiscuous outside of its core and binding epitope and might be readily adapted to synthesize new Pantocin analogues or incorporate the indolizidinone core into peptide libraries.[36,37] The present version of this assay is limited to reporting on glutamate modification because of reliance on GluC, but future iterations might probe other amino acids by exploiting alternative proteases or biorthogonal chemistries.[38] mRNA display is particularly well-suited to broadly probe peptide post-translational enzymatic chemistry in this manner because of the ready introduction of non-natural functionality through Flexizyme codon reprogramming. Perhaps most importantly, this work demonstrates the compatibility of a RiPP enzyme, here PaaA, with a C-terminal mRNA display tag, suggesting that others might also be used in this manner. These results ultimately pave the way for using RiPP enzymes to transform mRNA display libraries into more natural product-like molecules for inhibitor discovery.

2.4 Experimental.

2.4.1 Materials and General Methods

All oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), Twist Bioscience (San Francisco, CA), or Eton Bioscience (San Diego, CA). Enzymes such as Q5 DNA polymerase, T4 RNA Ligase I, T7 RNA polymerase, and DNAseI optizyme were purchased from New England Biolabs (Ipswich, MA). Taq DNA polymerase was expressed and purified. Endoproteinase GluC (from *Staphylococcus aureus* V8) was purchased from Millipore-

Sigma (Burlington, MA). Two different PURExpress translation kits were purchased and used as directed from New England Biolabs. For translation of peptides and analysis by MALDI-TOF -aa, -tRNA (E6840S) was used and for mRNA display -aa, -tRNA, -RF123 (E6850Z) was used. All amino acids used to make custom mixes for translations were purchased from Sigma-Aldrich, Alpha Aesar, or Fisher Biotech. Most protected amino acids for solid phase peptide synthesis were purchased from Chem-Impex International. ^{13}C, ^{15}N labeled protected glutamic acid was purchased from Cambridge Isotopes (Tewksbury, MA). Any other solvents, reagents, and chemicals were purchased from Fisher Scientific (Hampton, NH) unless otherwise specified. Preparatory HPLC was performed using a Shimadzu UFLC CBM-20A with a dual channel wavelength detector at 220 nm and 260 nm with a LUNA 10 µm C18 (2) 100 A, AXIA (Phenomenex) semi-preparatory column with a 15 mL / min flow rate. Peptide purification was carried out in a two-solvent system (Solvent A: 0.1 % trifluoroacetic acid and water; Solvent B: 0.1 % trifluoroacetic acid and acetonitrile) using a gradient flow from A to B. High-resolution liquid chromatography and mass spectrometry was conducted with a Kinetex 2.6 u C18 column and mass spectrometry data was collected with an Agilent 6520 Accurate-Mass Q-TOF ESI positive in high-resolution mode. All MALDI-TOF samples were first loaded onto C18 stage tips (Thermo Fisher Scientific, Waltham, MA) washed with 4 % acetonitrile and eluted directly onto a MALDI plate with 80 % acetonitrile and half saturated MALDI matrix α-cyano-4-hydroxycinnamic acid. Data was collected on an AB SCIEX TOF / TOF 5800 system in reflector, positive mode. Next generation sequencing (NGS) samples were submitted to Genewiz for Amplicon EZ analysis which provides $\geq$ 50,000 reads per sample. Sequencing data was analyzed using in-house python scripts.

2.4.2 Expression and Purification of PaaA-His_6

This procedure was adapted from the previous report.[32] Buffers were filter sterilized and made with DEPC treated water while the Ni^{2+} IMAC column was stripped and regenerated before use. PaaA (from paaA-pET28b, C-terminal 6xHis tag) was expressed in BL21 (DE3) *E. coli* cells in LB medium containing 40 μg / mL kanamycin. Cells were grown at 37 °C until an OD600 of 0.5 was attained, at which point the temperature was lowered to 18°C. At an OD600 of 0.8, protein expression was induced with the addition of 0.1 mM isopropyl-1-thio-D-galactopyranoside (IPTG) and incubated overnight at 18 °C. Cells were collected and pelleted by centrifugation. Cell pellets were suspended in Buffer A (20 mM Tris, pH 7.5, 30 mM imidazole, 300 mM NaCl, 5% Glycerol, 1 mM DTT) supplemented with 0.5 mL of 25 mg / mL T4 lysozyme, 0.5 mL of 150 mM PMSF, 80 μl of DNaseI (1 u / μl), one tablet of Pierce Protease Inhibitor Tablets (EDTA Free from Thermo Scientific), and then incubated on ice for 10 minutes. The cell pellet was then sonicated. The lysate was clarified by centrifugation at 15,000 rpm at 4 ºC for 15 minutes. The supernatant was collected and filtered through a 0.45 μm syringe filter. The flow through from the filter was then passed over a Ni^{2+} IMAC column (His-Trap HP 5-mL, GE Healthcare) coupled to an FPLC (NGC-Quest-10 Bio-Rad). The Ni^{2+} IMAC column was washed with 5 column volumes (CV) of Buffer A. Protein was eluted with Buffer B (20 mM Tris, pH=7.5, 500 mM imidazole, 300 mM NaCl, 5 % glycerol, 1 mM DTT) in a gradient of 0-14% B over 2 CV, 14-14 % B over 2 CV, and then 14-100 % B over 10 CV. Purity of eluted fractions was assessed by SDS-PAGE. Purest fractions were combined and concentrated in a 30,000 MWCO filter in an Amicon stirred cell. The concentrated protein was then exchanged into 50 mM HEPES pH 7.5, 150 mM NaCl, 5 % glycerol, and 1 mM DTT by passage over a Sephadex PD-10 column (GE Healthcare). The protein concentration was estimated by A280 using an extinction coefficient of 0.988 mg•mL-1•AU-1.

2.4.3 DNA Preparation for NEB PURExpress Translations

Genes used for NEB PURExpess transcription and translation reactions were purchased from Twist Bioscience (see Appendix A Translation Genes) and amplified using Taq polymerase (table 2.1) and the following primers:

T7 Forward (5'→3'): GAAATTAATACGACTCACTATAGGGG

T7 Terminator (5'→3'): GCTAGTTATTGCTCAGCGG

Amplification Cycles (Cycle 25x between denature, annealing, and elongation):

Table 2.1. Taq polymerase amplification cycles for translation genes.

Stage	Temperature	Time
Initial Denature	95 ºC	2 min
Denature	95 ºC	30 sec
Annealing	51 ºC	30 sec
Elongation	72 ºC	2 min
Final Elongation	72 ºC	5 min

After amplification, DNA products were column purified using Thermo Scientific GeneJET PCR Purification Kit. Purity was assessed by 3 % agarose gel and concentrations checked by nanodrop were ~100 ng / µL. Templates were stored at -20 ºC and directly used for PURExpress reactions.

2.4.4 DNA Preparation of Transcription Templates

DNA transcription templates for Flexizyme, tRNA, and PaaP NNK6 for T7 RNA polymerase reactions were prepared as previously described.[37] Briefly, each template was prepared by primer extension and / or PCR with the following primers in table 2.2.

Table 2.2. Flexizyme and mRNA display primers.

1	T7-Prom-F	GGCGTAATACGACTCACTATAG
2	eFx-extension-F	GTAATACGACTCACTATAGGATCGAAAGATTTCCGC
3	eFx-extension-R	ACCTAACGCTAATCCCCTTTCGGGGCCGCGGAAATCTTTCGATCC
4	efx-PCR-R	ACCTAACGCTAATCCCCT
5	itRNA-F.ex.49	GTAATACGACTCACTATAGGCGGGGTGGAGCAGCCTGGTAGCTCGTCGG
6	itRNA-R.ex.44	GAACCGACGATCTTCGGGTTATGAGCCCGACGAGCTACCAGGCT
7	itRNA-PCR1-R.38	TGGTTGCGGGGGCCGGATTTGAACCGACGATCTTCGGG
8	itRNA-PCR2.R.ome.20	T(mG)GTTGCGGGGGCCGGATTT*
9	T7-Display-F.46	TAATACGACTCACTATAGGGTTAACTTTAAGAAGGAGATATACATA
10	HAtag-Display-R.39	TTTCCGCCCCCCGTCCTACGAACCGGACCCTGAACCAGC
11	SVL-For	TAATACGACTCACTATAGGGTTAACTTTAAGAAGGAGATATAAATA
12	SVL-Rev	TTTCCGCCCCCCGTCCTAAGAACCAGAACCAGAACCTGC
13	PaaP-6NNK-Ex.F.94	GGGTTAACTTTAAGAAGGAGATATAAATATGATCAAGTTCTCTNNKNNKNNKNNKNNKNNKTCTGCGATCACTGAAGAAAATGCTACCTATACC
14	PaaP-HA-Gen.Ex.R102	CCAGAACCAGAACCTGCATAGTCGGGCACGTCGTATGGGTAGCTGCCGCTGCCGCTAAGGACGATAACCTGACCTTTGGTATAGGTAGCATTTTCTTCAGTG

* (mG) designates 2'Ome guanosine RNA

PCR Cycling Protocols

Table 2.3. Extension protocol (5 cycles of annealing and elongation):

Stage	Temperature	Time
Initial Denature	95 ^{o}C	2 min
Annealing	50 ^{o}C	1 min
Elongation	72 ^{o}C	1 min
Final Elongation	72 ^{o}C	5 min

Table 2.4. PCR1 protocol (5 cycles):

Stage	Temperature	Time
Initial Denature	95 ºC	2 min
Denature	95 ºC	40 sec
Annealing	50 ºC	40 sec
Elongation	72 ºC	40 sec
Final Elongation	72 ºC	5 min

Table 2.5. PCR2 protocol (12 cycles):

Stage	Temperature	Time
Initial Denature	95 ºC	2 min
Denature	95 ºC	40 sec
Annealing	50 ºC	40 sec
Elongation	72 ºC	40 sec
Final Elongation	72 ºC	5 min

Table 2.6. PCR3 protocol (25 cycles):

Stage	Temperature	Time
Initial Denature	95 ºC	2 min
Denature	95 ºC	30 sec
Annealing	57 ºC	30 sec
Elongation	72 ºC	2 min
Final Elongation	72 ºC	5 min

Table 2.7. PCR4 protocol (30 cycles):

Stage	Temperature	Time
Initial Denature	98 ºC	30 sec
Denature	98 ºC	10 sec
Annealing	57 ºC	20 sec
Elongation	72 ºC	30 sec
Final Elongation	72 ºC	2 min

eFx DNA Template Preparation

eFx Taq polymerase extension was performed with primers 2 and 3 using cycles "Extension Protocol". Subsequently, 5 µL of product was transferred to 1 mL of Taq polymerase PCR with primers 1 and 4 and amplified using cycles "PCR2 Protocol". DNA product was purified by phenol / chloroform extraction, concentrated by ethanol precipitation, air dried, and resuspended in 100 µL of water ready for T7 RNA transcription.

Initiator tRNA DNA Template Preparation

Initiator tRNA (itRNA) Taq polymerase extension was performed with primers 5 and 6 using cycles "Extension Protocol". Then 10 µL of extension product was transferred to a 200 µL Taq polymerase PCR using primers 1 and 7 using cycles "PCR1 Protocol". Subsequently 5 µL of PCR1 product was transferred to 1 mL Taq polymerase PCR with primers 1 and 8 using cycles "PCR2 Protocol". DNA product was purified by phenol / chloroform extraction, concentrated by ethanol precipitation, air dried, and resuspended in 100 µL of water ready for T7 RNA transcription.

PaaP Wild Type mRNA display DNA Template Preparation

The WT PaaP gene for mRNA display was purchased from Twist Bioscience (San Francisco, CA) and Taq PCR amplified with primers 9 and 10 using cycles "PCR 3 Protocol". PCR product was purified by Thermo Scientific GeneJET PCR Purification Kit, eluted in 20 µL of water, and directly used for T7 RNA transcription.

PaaP Saturation Mutagenesis Single Variant Library (smSVL) DNA Template Preparation

PaaP smSVL was purchased from Twist Bioscience. The library was designed to contain no "Pro" mutations at positions E16, E17, and N18 and no "met" or "Stop" mutations at all positions along the open reading frame. Primers 11 and 12 were used to amplify the library using

Q5 polymerase with cycle protocol “PCR4 Protocol”. PCR product was purified by Thermo Scientific GeneJET PCR Purification Kit, eluted in 20 µL of water, and directly used for T7 RNA transcription.

PaaP NNK6 Library DNA Template Preparation

Primers to prepare PaaP NNK6 library were purchased from Integrated DNA Technologies (Coralville, IA) and assembled via primer extension and PCR techniques. First, 100 µL Taq polymerase extension reaction was performed using primers 13 and 14 using “Extension Protocol”. All 100 µL was added to a 1 mL Taq PCR reaction with primers 11 and 12 using “PCR1 Protocol”. Product DNA was purified by phenol / chloroform extraction, concentrated by ethanol precipitation, air-dried, and resuspended in 100 µL of water ready for T7 RNA polymerase transcription.

2.4.5 Transcription with T7 RNA polymerase and Urea Gel Purification

T7 RNA transcription reactions were performed on a 1 mL scale according to NEB protocols with100 µL of prepared DNA template. The only change from NEB protocol was the addition of extra $MgCl_2$ (eFx final concentration: 30 mM; itRNA final concentration: 22.5 mM; mRNA display templates PaaP WT, PaaP smSVL: 28.5 mM; and mRNA display template PaaP NNK 6: 25 mM). Transcriptions were left at 37 ^{o}C overnight and treated with 115 µL of 10 x DNAseI buffer and 30 µL of DNAseI (Thermo Scientific, EN0521) for 1 hour at 37 ^{o}C. Upon successful transcriptions, magnesium pyrophosphate precipitates, so, 200 µL of 0.5 M ethylenediaminetetraacetic acid (EDTA) was added to clear the solution. RNA was precipitated by adding 135 µL of 3 M NaCl and 1190 µL of isopropanol and pelleted by centrifugation at 16,000 x g; 25 ^{o}C for 15 min. The supernatant was disposed of and the pellet was washed with 500 µL of 70 % ethanol. After centrifugation for 3 min at 16,000 x g the supernatant was removed, and

the pellet dried at room temperature. RNAs were resolubilized in 100 µL of water before adding RNA loading dye (2x) and loading onto a 12 % (eFx) or 8 % (itRNA and mRNA display templates) Urea-PAGE gel to run for 1 hour at 230 V in Tris-Borate-EDTA (TBE) buffer. The RNA band was visualized by UV shadowing (254 nm) on a silica-coated thin-layer chromatography plate and cut out of the gel. The removed gel slab was crushed into fine pieces and the RNA extracted twice with 0.3 M NaCl at room temperature. The combined extractions were filtered through a 0.45 µm sterile syringe filter and a 2x volume of ethanol was added to precipitate the RNA for collection by centrifugation at 4,000 rpm for 25 min (Eppendorf swing-bucket rotor A-4-81). The supernatant was removed, and the remaining pellet washed with 5 mL of 70 % ethanol followed by centrifugation at 4,000 rpm for 5 min. The final RNA pellet was recovered and dried at room temperature before resuspending in water. Both eFx and itRNA were diluted to a working stock of 250 µM and stored at -20 ^{o}C. mRNA display templates were diluted to 10 µM working stocks and stored at -80 ^{o}C. The above protocol was adjusted for only a 50 µL T7 RNA transcription reaction for PaaP WT and PaaP smSVL mRNA display templates.

2.4.6 N-Biotin-L-Phenylalanine itRNA Acylation Conditions

N-biotin-L-phenylalanine cyanomethyl ester was synthesized as in the following report.[38] Flexizyme assays were performed as described.[30] Briefly, reactions were carried out at 4 ^{o}C for 2 hours in 50 mM HEPES-KOH (pH 7.5), 600 mM $MgCl_2$, 20 % DMSO, 25 µM eFx, 25 µM itRNA, 5 mM biotin-L-phenylalanine cyanomethyl ester. Acylated tRNAs were precipitated with 0.3 M NaOAc pH 5.2 and 70 % ethanol, air-dried, and stored at -80 ^{o}C until use.

2.4.7 mRNA Display Protocols

General Strategy for Linking mRNA Templates to Puromycin

mRNA templates were covalently linked to puromycin via Y-ligation strategy adapted from the previous report.[39] Briefly, 1 µM RNA was incubated with T4 RNA ligase I (NEB, M0204S) at room temperature using manufacture protocols with 20 % DMSO and 1.5 µM of P-linker (d(pCTCCCGCCCCCCGTCC)-(SPC18)5-d(CC)-puromycin). The reaction was stopped after 30 mins by adding a 1 x volume of quenching solution (0.6 M NaCl and 10 mM EDTA). Puromycin-linked RNA was precipitated by adding 2.2 x volume of 100 % ethanol and centrifuged at 16,000 x g for 15 min. Discarding the supernatant, the resultant pellet was washed with 70 % ethanol and centrifuged again 16,000 x g for 5 min. The pellet was collected and air dried before resuspending in 1 / 10 the volume of water as the original T4 RNA ligation reaction. This was stored at -80 ^{o}C and directly used in PURExpress translation reactions (vide infra).

Peptide - RNA Fusion Formation and Reverse Transcription

mRNA display reactions were adapted from the previous report.[4] Translations were performed according to PURExpress protocol using kit (-aa, -tRNA, -RF123) with minor changes. A 19 amino acid mix devoid of methionine was prepared with 5 mM of each amino acid and added to the translation instead of the provided 20 amino acid mix. Puromycin -linked mRNA template was added to 20 % of the final translation volume in addition to 100 µM Biotin-Phe-itRNA (solubilized in 1 mM NaOAc pH 5.2). Translations were incubated at 37 ^{o}C for 30 min and then 12 min at room temperature to enhance fusion formation before adding 16.7 mM EDTA for 30 min at 37 ^{o}C to dissociate the ribosome. Before selection, RNA was reverse transcribed by adding to the displayed library 250 nM dNTPs, 2 µM reverse primer, 25 mM Tris-HCl (pH 8.3), 15 mM $Mg(OAc)_2$, 10 mM KOH, and 1 x mMLV reverse transcriptase H (-), point mutant (Promega,

M3681, supplied at 40x) and incubating at 42 ºC for 1 hour. To make 75 µL of peptide-RNA fusions, translation reaction volume was 37.5 µL and final volume after reverse transcription was 75 µL.

qPCR Analysis

To quantitate mRNA display experiments, qPCR was performed using Applied Biosystems ViiA7 qPCR Real-Time qPCR Instrument. For each sample, 1 µL was mixed with 19 µL of SsoAdvanced Universal SYBR Green Supermix, Bio-Rad (Hercules, CA). qPCR standards were made by reverse transcribing a known amount of RNA into cDNA, assuming 100 % yield and diluting to 2E9, 2E8, 2E7, 2E6, and 2E5 molecules. Experiments were initiated by heating at 50 ºC for 2 min followed by a 10 min 95 ºC incubation to activate the polymerase. Cycling then ensued between 95 ºC for 15 sec and 60 ºC for 1 min, and cycler heating acceleration was held at 1.6 ºC / s between each step. During each elongation, SYBR green fluorescence was measured with ROX as a passive reference. Standard curves were made and used to calculate mRNA display quantities.

Determination of Optimal GluC Cleavage Conditions

PaaP wild type was displayed as above and to 75 µL of prepared fusions was added Bicine buffer pH 9.0 to a final concentration of 100 mM Bicine and a final volume of 200 µL. The solution was split into two 93.75 µL halves and to each was added 93.75 µL of 2 x PaaA reaction mix (100 mM Bicine pH 9.0, 8 mM $MgCl_2$, 2 mM ATP, and 2 µM PaaA) or 2 x reaction mix (100 mM Bicine pH 9.0, 8 mM $MgCl_2$, 2 mM ATP). The reactions were incubated at room temperature for 1 hour before quenching by adding 41.5 µL 0.5 M EDTA pH 8.0. To make the solution less basic, 186 µL of 500 mM HEPES pH 7.5 and 0.05 % Tween-20 was added to each reaction (final volume: 415 µL). PaaP WT fusions for both with and without PaaA were aliquoted into 20 µL volumes

and incubated with 4 µL of Anti-HA magnetic beads (Pierce, 88837) and left rotating at 4 ºC for 1 hr. Magnetic beads were washed 3 x with wash buffer A (50 mM Ammonium Bicarbonate pH 7.8, 0.05% Tween-20) and finally resuspended in 5 µL wash buffer A. Purified PaaP WT fusions were then eluted and collected by adding 5 µL of 4 mg / mL HA Peptide (Thermo Fisher Scientific, 26184) in water for 1 hour at room temperature with rotation. The magnetic beads were pulled to the side, the supernatant collected, and to it was added 10 µL of 2 x GluC solution (50 mM Ammonium Bicarbonate pH 7.8, 0.05% Tween-20, and x µg / µL GluC, where x is 0.2, 0.02, 0.002, 0.0002, or 0 µg / µL). Each GluC treatment was done in triplicate and placed at 37 ºC for 2 hours. Samples were then placed on ice and 4 µL of magnetic Dynabeads M-280 streptavidin (Thermo Fisher Scientific, 11205D) equilibrated with wash buffer B (50 mM Tris pH 7.5, 150 mM NaCl, 0.05% Tween-20) were added and samples rotated at 4 ºC for 30 mins. After washing 3 x with 100 µL of wash buffer B, each sample was eluted by heating at 95 ºC for 5 min in 40 µL PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, and 0.1 % Triton X-100) and submitted for qPCR analysis. We noticed a significant variance in GluC activity between different purchased lots. For each GluC lot purchased, the optimal GluC concentration was determined which provided the greatest fold enrichment of fusion recovery between PaaA treated and untreated samples.

Determination of Optimal PaaA Reaction Concentration

PaaP wild type was displayed as above and to 36 µL of prepared fusions was added 54 µL of 50 mM Bicine buffer pH 9.0. and subsequently aliquoted into 18, 5 µL samples. To each aliquot was added 5 µL of 2 x PaaA reaction mix (50 mM Bicine pH 9.0, 8 mM $MgCl_2$, 2 mM ATP, and x µM PaaA where x is 0, 0.02, 0.2, 2.0, and 20). Each [PaaA] was done in triplicate and the reactions were incubated for 1 hour before adding 10 µL of 500 mM HEPES pH 7.5 and purified with 4 µL of Anti-HA magnetic beads. After rotating at 4 ºC for 1 hr magnetic beads were washed

3 x with 100 µL wash buffer B (50 mM Tris pH 7.5, 150 mM NaCl, 0.05% Tween-20), then washed once with 5 µL of wash buffer A (50 mM Ammonium Bicarbonate pH 7.8, 0.05% Tween) and finally resuspended in 5 µL wash buffer A. Purified PaaP WT fusions were then eluted and collected by adding 5 µL of 4 mg / mL HA Peptide in water for 1 hour at room temperature with rotation. The supernatant was collected by pulling the magnetic beads to the side and to it was added 10 µL of 2 x GluC solution (50 mM Ammonium Bicarbonate pH 7.8, 0.05% Tween-20, and GluC*) for 2 hours at 37 ^{o}C. Samples were moved to ice and 4 µL of magnetic Dynabeads M-280 streptavidin equilibrated with wash buffer B (50 mM Tris pH 7.5, 150 mM NaCl, 0.05% Tween-20) were added and samples rotated at 4 ^{o}C for 30 mins. After washing 3 x with 100 µL of wash buffer B, each sample was eluted by heating a 95 ^{o}C for 5 min in 40 µL PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, and 0.1 % Triton X-100) and submitted for qPCR analysis.

*see section "Determination of Optimal GluC Cleavage Conditions"

Determination of Optimal PaaA Reaction Time

PaaP wild type was displayed as above and to 36 µL of prepared fusions was added 59 µL of 50 mM Bicine buffer pH 9.0, before dispensing 18, 5 µL samples. To each was added 5 µL of 2 x PaaA reaction mix (50 mM Bicine pH 9.0, 8 mM $MgCl_2$, 2 mM ATP, and 2 µM PaaA). Reactions were incubated in triplicate for 0 (containing no PaaA), 5, 15, 30, 45, and 60 min at room temperature before adding 2.0 µL of 0.5 M EDTA pH 8.0 to quench the reaction. To make the solution less basic, 8 µL of 500 mM HEPES pH 7.5 and 0.05 % Tween-20 was added to each sample (final volume: 20 µL). PaaP-WT fusions were purified with 4 µL of Anti-HA magnetic beads rotating at 4 ^{o}C for 1 hr. Magnetic beads were then washed 3 x with 100 µL wash buffer B (50 mM Tris pH 7.5, 150 mM NaCl, 0.05% Tween-20), then washed once with 5 µL of wash buffer A (50 mM Ammonium Bicarbonate pH 7.8, 0.05% Tween-20) and finally resuspended in

5 µL wash buffer A. Purified PaaP WT fusions were then eluted and collected by adding 5 µL of 4 mg / mL HA Peptide in water for 1 hour at room temperature with rotation. The supernatant was collected by pulling the magnetic beads to the side and to it was added 10 µL of 2 x GluC solution (50 mM Ammonium Bicarbonate pH 7.8, 0.05% Tween-20, and GluC*) for 2 hours at 37 ^{o}C. Samples were moved to ice and 4 µL of magnetic Dynabeads M-280 streptavidin equilibrated with wash buffer B (50 mM Tris pH 7.5, 150 mM NaCl, 0.05% Tween-20) were added and samples rotated at 4 ^{o}C for 30 mins. After washing 3 x with 100 µL of wash buffer B, each sample was eluted by heating at 95 ^{o}C for 5 min in 40 µL PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, and 0.1 % Triton X-100) and submitted for qPCR analysis.

*see section "Determination of Optimal GluC Cleavage Conditions"

PaaP Saturated Mutagenesis Single Variant Library Selection and NGS analysis

PaaP wild type and PaaP smSVL were displayed as above and to 36 µL of prepared fusions was added Bicine buffer pH 9.0 to a final concentration of 100 mM Bicine and a final volume of 95 µL. Both PaaP WT fusions and PaaP smSVL fusions were split into 5 µL aliquots and to each was added 5 µL of 2 x PaaA reaction mix (100 mM Bicine pH 9.0, 8 mM $MgCl_2$, 2 mM ATP, and 2 µM PaaA). The reactions were incubated at room temperature for 0 (containing no PaaA), 5, 22.5, or 60 min before quenching by adding 2.0 µL 0.5 M EDTA pH 8.0. To make the solution less basic, 8 µL of 500 mM HEPES pH 7.5 and 0.05 % Tween-20 was added to each sample (final volume: 20 µL). Fusions for WT and smSVL were then incubated with 4 µL of Anti-HA magnetic beads and left rotating at 4 ^{o}C for 1 hr. Magnetic beads were washed 3 x with wash buffer A (50 mM Ammonium Bicarbonate pH 7.8, 0.05% Tween-20) and finally resuspended in 5 µL wash buffer A. Purified fusions were then eluted and collected by adding 5 µL of 4 mg / mL HA Peptide in water for 1 hour at room temperature with rotation. The supernatant was collected by pulling

the magnetic beads to the side and to it was added 10 µL of 2 x GluC solution (50 mM Ammonium Bicarbonate pH 7.8, 0.05% Tween-20, and GluC*) for 2 hours at 37 ^{o}C. Samples were then placed on ice and 4 µL of magnetic streptavidin beads equilibrated with wash buffer B (50 mM Tris pH 7.5, 150 mM NaCl, 0.05% Tween-20) were added and samples rotated at 4 ^{o}C for 30 mins. After washing 3 x with 100 µL of wash buffer B, each sample was eluted by heating at 95 ^{o}C for 5 min in 40 µL PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, and 0.1 % Triton X-100) and submitted for qPCR. Results from PaaP wild type samples were analyzed to make sure the selection functioned properly. Then 3 samples from each smSVL condition (0, 5, 22.5, and 60 min PaaA treatment) including an HA purified control with no GluC pressure (input) were PCR amplified with Amplicon EZ NGS adapters (sequence provided by Genewiz) and submitted to Genewiz for Amplicon EZ sequencing. For each NGS sample, every sequence was translated along its open reading frame. All sequences not expected from our smSVL design were removed and the percent of each peptide within input and selection samples calculated. By dividing the percentage of a single peptide within the selection sample by the input sample an enrichment score (e score) was generated. These e scores were then normalized to obtain a fitness value for each mutant (E score) by dividing each mutant e score by the WT. E scores were displayed as a heatmap for ease of visualization using in-house python scripts. Each E score is an average of three independent samples split after translation and fusion formation.

*see section “Determination of Optimal GluC Cleavage Conditions”

PaaP NNK6 mRNA Display and NGS analysis

PaaP NNK6 was displayed as above and to 12 µL of prepared fusions was added Bicine buffer pH 9.0 to a final concentration of 100 mM Bicine and a final volume of 30 µL. PaaP smSVL fusions were treated with 30 µL of 2 x PaaA reaction mix (100 mM Bicine pH 9.0, 8 mM $MgCl_2$,

2 mM ATP, and 2 µM PaaA) for 30 min before adding 6 µL of 0.5 M EDTA pH 8.0 to quench the reaction. Subsequently, 64 µL of 500 mM HEPES pH 7.5, 0.05 % Tween-20 was added and 20 µL was aliquoted for 6 samples. To each 20 µL sample was added 4 µL of Anti-HA magnetic beads and left rotating at 4 ^{o}C for 1 hr. Magnetic beads were washed 3 x with 100 µL of wash buffer A (50 mM Ammonium Bicarbonate pH 7.8, 0.05% Tween-20) and finally resuspended in 5 µL wash buffer A. Purified fusions were then eluted and collected by adding 5 µL of 4 mg / mL HA Peptide in water for 1 hour at room temperature with rotation. The supernatant was collected by pulling the magnetic beads to the side and to it was added 10 µL of 2 x GluC solution (50 mM Ammonium Bicarbonate pH 7.8, 0.05% Tween-20, and GluC*). As a control, to 3 samples was added 2x GluC solution with GluC omitted. These samples were incubated at 37 ^{o}C for 2 hours. Samples were then placed on ice and 4 µL of magnetic streptavidin beads equilibrated with wash buffer B (50 mM Tris pH 7.5, 150 mM NaCl, 0.05% Tween-20) were added and samples rotated at 4 ^{o}C for 30 mins. After washing 3 x with 100 µL of wash buffer B, each sample was eluted by heating at 95 ^{o}C for 5 min in 40 µL PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, and 0.1 % Triton X-100) and submitted for qPCR analysis before PCR amplification. Collected DNA was pooled and transcribed into RNA for ensuing selection rounds. After 4 rounds of selection, the NNK 6 library streptavidin recovery matched recovery of a WT control peptide. At this point the DNA amplification products were appended with Amplicon EZ primers containing NGS adapters via PCR and submitted for Amplicon EZ sequencing. To generate the Weblogo in figure 2f, the sequencing data was translated and peptides of length 49 – 51 amino acids were extracted. Each peptide was counted and divided by the total number of peptides to give a percent. Of the 1000 most enriched sequences 996 of them were the correct length of 50 amino acids and used in WEBLOGO assessment.

2.4.8 MALDI-TOF Experiments

MALDI-TOF Analysis of PaaA Reaction on PaaP Mutants

PaaA substrates were translated (2.5 µL) for 1 hour at 37 ^{o}C using PURExpress protocols (-aa, -tRNA). Reaction solutions were made at 1.25 x with and without PaaA (125 mM Bicine pH 9.0, 5 mM $MgCl_2$, 1.25 mM ATP, and 1.25 µM PaaA). A microliter of PURExpress solution was added to 4 µL of each reaction solution and incubated at room temperature for 22.5 min before quenching with 5 µL of 100 mM EDTA pH 8.0. Samples were desalted with C18 stage tips and analyzed by MALDI-TOF. For PaaA treated samples, extracted ion integration areas were calculated for each expected mass of unmodified and modified PaaP peptides (-1 H_2O; -2 H_2O; -1 H_2O, -1 CO_2; and -2 H_2O, -1 CO_2). All modified PaaA mass areas were summated as products. The area for the unmodified substrate and the summated area of products was divided by the total area to calculate a percent yield and visualized by a heatmap generated using python scripts.

MALDI-TOF Analysis of PaaA Reaction on E16 and E17 PaaP Mutants

Theses samples were treated as above, except reaction quenching was achieved by direct loading of the reaction directly onto C18 stage tips after 22.5 min of PaaA incubation.

2.4.9 Binding Experiments

Fluorescence Polarization Assay

Binding between PaaA and TAMRA-$PaaP_{1-12}$ was measured by fluorescence polarization using 5 nM of peptide and titrating PaaA concentration from 25 µM to 0.763 nM in binding buffer (100 mM Bicine pH 9.0, 150 mM NaCl, 2 % DMF, and 0.005% Tween-20). Binding measurements were performed in triplicate and analyzed by Tecan infinite M1000 Pro (Excitation: 530 nm, Emission: 586 nm, Gain: 152, Number of Flashes: 10, Settle Time: 0 ms, Z-Position: 20000 µm, G-Factor: 1). Background fluorescence from wells containing only PaaA was

subtracted from the mP signals to remove background fluorescence. Data was fit to a nonlinear fit, [inhibitor] vs. response, variable slope (four parameter) to calculate EC_{50}.

Competitive Fluorescence Polarization Assay

To assess how aspartic acid mutations to $PaaP_{1-12}$ affect binding to PaaA, a competitive assay was set up with 5 nM TAMRA-$PaaP_{1-12}$, 100 nM PaaA, and increasing competitor concentrations from 50 µM – 6.1 nM in binding buffer (100 mM Bicine pH 9.0, 150 mM NaCl, 1 % DMF, and 0.005% Tween-20). Each Asp mutant was performed in triplicate and FP measured by Tecan infinite M1000 Pro (Excitation: 530 nm, Emission: 586 nm, Gain: 136, Number of Flashes: 10, Settle Time: 0 ms, Z-Position: 20000 µm, G-Factor: 1). Data was analyzed using a nonlinear fit, [inhibitor] vs. response, variable slope (four parameter) to calculate IC_{50} values.

Isothermal Calorimetry Assay

PaaA was buffer exchanged and diluted into binding buffer (100 mM Bicine pH 9.0, 150 mM NaCl, 2 % DMF). Peptides were massed out and solubilized in the above binding buffer to 100 µM. Centrifugation showed no precipitation. Data was acquired using a Microcal AutoITC200. Using the above stocks. PaaA was loaded into the cell followed by 19, 2 µL titrations performed at 25 ^{o}C with stirring at 750 rpm. The final plateaued 1-6 points were subtracted from each run to account for the buffer heat of dilution. Data was analyzed using Origin software and the data was fit using one-site binding.

2.4.10 Peptide Synthesis

Solid Phase Peptide Synthesis of Aspartic Acid Mutants $PaaP_{1-12}$

All peptides were synthesized with standard solid phase peptide synthesis (SPPS) fluorenyl-methoxycarbonyl (Fmoc) procedures on Rink Amide ChemMatrix resin (0.48 mmol / g). Before synthesis, resin was preswollen for 30 min in DCM (dichloromethane) and then 30 min

1:1 DCM:DMF (dimethylformamide) using a miniblock for shaking at 670 rpm at room temperature. Deprotections were performed for 10 min at room temperature in 20 % Piperidine in DMF with shaking. For each coupling, 5 equivalents (eq) of Fmoc-Amino acid was dissolved in DMF and preactivated with 5 eq of O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) and 10 eq of Diisopropylethylamine (DIEA). After mixing, the activated amino acid was transferred to the swollen and deprotected solid support for 15 min couplings at room temperature with shaking. Between every deprotection and coupling, the resin was drained, and washed 3 x with ~5 mL of DMF. When synthesis was complete, 50 mg of resin was air dried and the peptide cleaved by addition of 2 mL of cleavage cocktail containing Trifluoroacetic acid (TFA): Triisopropylsilane (TIS): Water (95:2.5:2.5) for 1.5 hours at 37 ^{o}C. The remaining resin was filtered away and the TFA was evaporated under nitrogen before the addition of 35 mL of diethyl ether, freezing at -80 ^{o}C for 30 min, and collection by centrifugation. The supernatant was decanted, the peptide dried and resuspended in 1:1 Water:Acetonitrile and loaded onto prep HPLC for gradient purification. Fractions containing product were frozen and lyophilized. Fractions with unpurified product were solubilized in 0.1 % TFA and again loaded onto prep HPLC for re-purification. Purified fractions were frozen, lyophilized and dissolved to 5 mM in DMF. For the preparation for TAMRA labeled substrate, the final coupling was completed as described for the other amino acids but using 5(6)-TAMRA (5-(and-6)-Carboxytetramethyl-rhodamine (Chemodex, C0038; St. Gallen, Switzerland).

SPPS of N-terminal N-MeDbz PaaP$_{1-12}$

Procedure was adapted from the literature.[40] For preparation of ^{15}N and ^{13}C labeled PaaA substrate, the N-terminus was synthesized out by microwave assisted SPPS, ChemMatrix solid support (0.48 mmol / g). The resin was swollen for 20 min in DMF 70 ^{o}C, and deprotections were

performed twice with 20 % Piperidine for 3 min and then 10 min, washing 4 x with 1 mL of DMF before each coupling. Fmoc-Amino Acids-OH (5 eq, 0.1 M in DMF) were mixed with Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium (HATU, 4.9 eq, 0.2 M in DMF) and DIEA (10 eq, 0.2 M in DMF) in the reaction vessel for 5 min at 75 ^{o}C with mixing for each coupling step. Subsequently, the resin was washed 2 x with 1 mL of DMF. After the first coupling and the last deprotection any free amines were acetylated by adding Acetic Anhydride (50 eq, 5 M in DMF) and DIEA (50 eq, 0.5 M in DMF). The only deviation from this protocol was coupling of Fmoc-MeDbz-OH which was accomplished by addition of Fmoc-MeDbz-OH (5 eq, 0.1 M in DMF) with Hydroxybenzotriazole (HOBT, 5 eq, 0.5 M in DMF), 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate (TBTU, 4.9 eq, 0.6 M in DMF), and DIEA (10 eq, 0.2 M in DMF) in the reaction vessel for 5 min at 75 ^{o}C. The reaction vessel was then drained and washed 4 x with DMF and then DCM. Directly following the synthesis, the peptide was acylated with *p*-nitrophenyl chloroformate (10.0 eq, 0.25 M in DCM) for 1 hr at 37 ^{o}C. The resin was washed extensively with DCM before washing with DMF and formation of the N-acyl urea was initiated by addition of DIEA (40 eq, 1 M in DMF) for 1 hr at room temperature. After extensive washing with DMF and then DCM 100 mg of resin was dried, and the peptide was cleaved in 4 mL of cleavage cocktail TFA:TIS:Water (95:2.5:2.5) for 1 hr at 37 ^{o}C. The remaining resin was filtered away and TFA was evaporated under nitrogen before the addition of 25 mL of diethyl ether to precipitate the peptide which was collected by centrifugation. The supernatant was decanted, the peptide dried, resuspended in 1:1 Water:Acetonitrile, and loaded onto prep HPLC for gradient purification. Fractions containing product were frozen, lyophilized, and stored at 4 ^{o}C until use.

SPPS of C-terminal PaaP$_{13-30}$

Synthesis of C-terminal A13C PaaP$_{13-30}$ was accomplished by microwave assisted SPPS on Rink Amide ChemMatrix as described above for the N-terminal N-meDbz PaaP$_{1-12}$ using only natural Fmoc-AA-OH. The only deviation was the use of L-Glutamic Acid-N-Fmoc, Gamma-Tert-Butyl Ester (^{13}C5, 99%; ^{15}N, 99%; Cambridge Isotope Laboratories, Tewksbury, MA).

Native Chemical Ligation and In Situ PaaA Reaction

NCL procedure was adapted from the literature.[6] A13C PaaP$_{13-30}$ (0.47 µmol) was dissolved in 7.2 µL of DMSO and added to 230 µL of NCL reaction buffer (20 mM NaH_2PO_4 pH 7.0, 6 M Guanidine HCl, 120 mM TCEP, and 100 mM 4-Mercaptophenylacetic Acid). Then 0.72 µmol of PaaP$_{1-12}$ N-MeDbz was dissolved in 10 µL of DMSO and added to the NCL reaction. The solution was pipet mixed and sat at room temperature for 2 hours before 1 µL was removed and added to PaaA reaction solution containing 100 mM Bicine pH 9.0, 4 mM $MgCl_2$, 1 mM ATP, 1 mM TCEP, and 2.5 µM PaaA (or no PaaA for control reaction). Reactions were mixed and incubated overnight at room temperature and directly analyzed by high-resolution liquid chromatography and mass spectrometry (HR-LCMS).

REFERENCES

(1) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K-D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hem-scheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. *Nat. Prod. Rep.* **2013**, *30*, 108.

(2) Koehnke, J.; Mann, G.; Bent, A. F.; Ludewig, H.; Shirran, S.; Botting, C.; Lebl, T.; Houssen, W. E.; Jaspars, M.; Naismith, J. H. *Nat. Chem. Biol.* **2015**, *11*, 558.

(3) Grove, T. L.; Himes, P.; Hwang, S.; Yumerefendi, H.; Bonanno, J. B.; Kuhlman, B.; Almo, S. C.; Bowers, A. A. *J. Am. Chem. Soc.* **2017**, *139*, 11734.

(4) Ortega, M. A.; Hao, Y.; Zhang, Q.; Walker, M. C.; van der Donk, W. A.; Nair, S. K. *Nat.* **2015**, *517*, 509..

(5) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. *Nat. Chem. Biol.* **2015**, *11*, 564.

(6) Regni, C. A.; Roush, R. F.; Miller, D. J.; Nourse, A.; Walsh, C. T.; Schulman, B. A. *EMBO J.* **2009**, *28*, 1953.

(7) Burkhart, B. J.; Kakkar, N.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A. *ACS Cent. Sci.* **2017**, *3*, 629.

(8) Fleming, S. R.; Bartges, T. E.; Vinogradov, A. A.; Kirkpatrick, C. L.; Goto, Y.; Suga, H.; Hicks, L. M.; Bowers, A. A. *J. Am. Chem. Soc.* **2019**, *141*, 758.

(9) Himes, P. M.; Allen, S. E.; Hwang, S.; Bowers, A. A. *ACS Chem. Biol.* **2016**, *11*, 1737.

(10) Hegemann, J. D.; Bobeica, S. C.; Walker, M. C.; Bothwell, I. R.; van der Donk, W. A. *ACS Synth. Biol.* **2019**, *8*, 1204.

(11) Islam, M. R.; Shioya, K.; Nagao, J.; Nishie, M.; Jikuya, H.; Zendo, T.; Nakayama, J.; Sonomoto, K. *Mol. Microbiol.* **2009**, *72*, 1438.

(12) Young, T. S.; Dorrestein, P. C.; Walsh, C. T. *Chem. Biol.* **2012**, *19*, 1600.

(13) Pan, S.; Link, J. A. *J. Am. Chem. Soc.* **2011**, *133*, 5016.

(14) Goto, Y.; Ito, Y.; Kato, Y.; Tsunoda, S.; Suga, H. *Chem. Biol.* **2014**, *21*, 766.

(15) Ozaki, T.; Yamashita, K.; Goto, Y.; Shimomura, M.; Hayashi, S.; Asamizu, S.; Sugai, Y.; Ikeda, H.; Suga, H.; Onaka, H. *Nat. Comm.* **2017**, *8*, 14207.

(16) Wever, W.; Bogart, J.; of the American, B. A. *J. Am. Chem. Soc.* **2016**, *138*, 13461.

(17) Hetrick, K. J.; Walker, M. C.; van der Donk, W. A. *ACS Cent. Sci.* **2018**, *4*, 458.

(18) Yang, X.; Lennard, K. R.; He, C.; Walker, M. C.; Ball, A. T.; Doigneaux, C.; Tavassoli, A.; van der Donk, W. A. *Nat. Chem. Biol.* **2018**, *14*, 375.

(19) Urban, J. H.; Moosmeier, M. A.; Aumüller, T.; Thein, M.; Bos-ma, T.; Rink, R.; Groth, K.; Zulley, M.; Siegers, K.; Tissot, K.; Moll, G. N.; Prassler, J. *Nat. Comm.* **2017**, *8*, 1500.

(20) Tjhung, K. F.; Kitov, P. I.; Ng, S.; Kitova, E. N.; Deng, L.; Klas-sen, J. S.; Derda, R. *J. Am. Chem. Soc.* **2016**, *138*, 32.

(21) He, B.; Tjhung, K. F.; Bennett, N. J.; Chou, Y.; Rau, A.; Huang, J.; Derda, R. *Sci. Rep.* **2018**, *8*, 1214.

(22) Hofmann, F. T.; Szostak, J. W.; Seebeck, F. P. *J. Am. Chem. Soc.* **2012**, *134*, 8038.

(23) Malone, M. L.; Paegel, B. M. *ACS Comb. Sci.* **2016**, *18*, 182.

(24) Li, Y.; Luca, R.; Cazzamalli, S.; Pretto, F.; Bajic, D.; Scheuer-mann, J.; Neri, D. *Nat. Chem.* **2018**, *10*, 441.

(25) Sugimura, Y.; Yokoyama, K.; Nio, N.; Maki, M.; Hitomi, K. *Arch. Biochem. Biophys.* **2008**, *477*, 379.

(26) Kretz, C. A.; Dai, M.; Soylemez, O.; Yee, A.; Desch, K. C.; Siemieniak, D.; Tomberg, K.; Kondrashov, F. A.; Meng, F.; Gins-burg, D. *Proc. Natl. Acad. Sci.* **2015**, *112*, 9328.

(27) Yin, J.; Straight, P. D.; Hrvatin, S.; Dorrestein, P. C.; Bumpus, S. B.; Jao, C.; Kelleher, N. L.; Kolter, R.; Walsh, C. T. *Chem. Biol.* **2007**, *14*, 303.

(28) Matthews, D.; Wells, J. *Science* **1993**, *260*, 1113.

(29) Huang, Y.; Wiedmann, M.; Suga, H. *Chem. Rev.* **2018**, *119*, 10360.

(30) Goto, Y.; Katoh, T.; Suga, H. *Nat. Protoc.* **2011**, 6, 779.

(31) Smits, T. H.; Duffy, B.; Blom, J.; Ishimaru, C. A.; Stockwell, V. O. *Arch. Microbiol.* **2019**, *201*, 713.

(32) Ghodge, S. V.; Biernat, K. A.; Bassett, S.; Redinbo, M. R.; Bowers, A. A. *J. Am. Chem. Soc.* **2016**, *138*, 5487.

(33) Rogers, J. M.; Passioura, T.; Suga, H. *Proc. Natl. Acad. Sci.* **2018**, *115*, 10959.

(34) Atmuri, P. N.; Lubell, W. D. *J. Org. Chem.* **2020**, *85*, 1340.

(35) Cluzeau, J.; Lubell, W. D. *Biopolymers* **2005**, *80*, 98.

(36) Jetson, R. R.; Krusemark, C. J. *Angew. Chem. Int. Ed.* **2016**, 55, 9562.

(37) Goto, Y.; Katoh, T.; Suga, H. *Protoc. Ex.* **2011**, https://doi.org/10.1038/protex.2011.209 .

(38) Saito, H.; Kourouklis, D.; Suga, H. EMBO J. **2001**, *20*, 1797.

(39) Sakai, K.; Passioura, T.; Sato, H.; Ito, K.; Furuhashi, H.; Umitsu, M.; Takagi, J.; Kato, Y.; Mukai, H.; Warashina, S.; et al. *Nat. Chem. Biol.* **2019**, *15*, 598.

(40) Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. *J. Am. Chem. Soc.* **2015**, *137*, 7197.

CHAPTER 3: FLEXIZYME-ENABLED BENCHTOP BIOSYNTHESIS OF THIOPEPTIDES[1]

3.1 Introduction

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a growing family of peptide-derived natural products that exhibit natural combinatorial biosynthetic logic.[1] RiPP biosyntheses initiate from a gene-encoded precursor peptide, which contains a core region that undergoes enzymatic post-translational modification, and a leader region, which is typically responsible for recruiting and coordinating these enzymes through specific recognition sequences (RSs). These RSs have affinity for select domains of RiPP biosynthetic enzymes, increasing substrate local concentration to the otherwise promiscuous enzyme active sites and allowing the modification of diverse cores.[2,3] Natural pathways exhibit leader peptides with multiple RSs and recruit whole suites of post-translational modifying enzymes to convert precursor peptides into mature natural products. The combination of RS-programmable recruitment and promiscuous enzymes inspires recent efforts at repurposing this strategy to scaffold new-to-nature hybrid biosynthetic pathways.[4–8]

Thiopeptides are one of the most extensive natural examples of this combinatorial, RS-directed biosynthesis,[9,10] and the three class-defining transformations include the formation of azoles, dehydroamino acids, and pyridines from serine and cysteine residues. Many of these enzymes are remarkably promiscuous, and thiopeptide pathways have proven capable of

[1]This chapter was adapted from a communication that previously appeared in the Journal of the American Chemical Society. The original citation is as follows: Fleming, S.R., Bartges, T.E., Vinogradov, A.A., Kirkpatrick, C.L., Goto, Y., Suga, H., Hicks, L.M., Bowers, A.A. *J. Am. Chem. Soc.*, **2019**, *141* (2), pp 758-762.

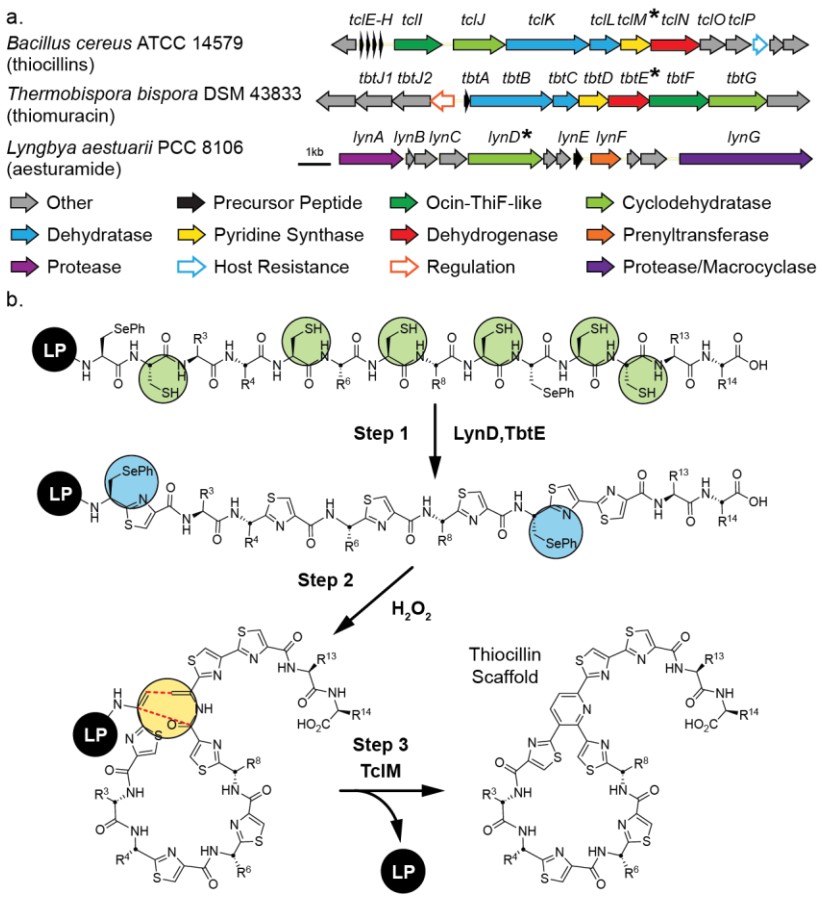


Figure 3.1. Flexizyme-enabled benchtop biosynthesis of thiopeptide scaffolds. (a) Biosynthetic gene clusters of thiocillins, thiomuracin GZ, and aesturamide. Genes for key enzymes used in this work are highlighted with asterisks. (b) Proposed hybrid route to the thiocillin core.

generating many variants.[11,12] For example, hundreds of mutants of thiocillin, GE37468, and thiostrepton have been generated by gene replacement strategies.[13–15] However, competition between the pathway enzymes for functional groups on non-native substrates can give rise to complex mixtures of products and slower processing of mutant substrates, or host toxicity can restrict production of potential compounds. The in vitro reconstitution of whole thiopeptide biosynthetic pathways, which has recently been achieved for thiomuracin, can circumvent some of these problems but relies on access to soluble, well-behaved proteins.[16,17] This can be especially challenging for the tRNA dependent lantibiotic-type dehydratases. Additionally, this strategy still does not overcome enzyme competition. Alternative strategies, such as a chemoenzymatic

approach, could allow rapid access to novel structural variants and ease characterization of new thiopeptides and thiopeptide-associated enzymes.[18]

We envisioned that carefully chosen RiPP enzymes might be combined with orthogonal chemical handles to create a flexible in vitro platform for the benchtop preparation of thiopeptides (figure 3.1).[19,20] More specifically, in vitro transcription−translation could be used to express designer hybrid leader peptide substrates displaying RSs for the cyclodehydratase, LynD, from aesturamide biosynthesis and pyridine synthase, TclM, from thiocillin biosynthesis. Both enzymes have well-defined RS motifs and broad substrate promiscuity. Additionally, LynD exhibits excellent selectivity for Cys conversion to thiazolines while ignoring Ser / Thr residues.[21,22] Alternatively, if oxazoles are of interest, then PatD, which acts on Ser / Thr / Cys, could be used in place of LynD.[20] Oxidation of thiazolines to thiazoles might be effected by the azoline oxidase, TbtE from thiomuracin biosynthesis, which acts in a leader peptide independent manner.[16] In place of lantibiotic dehydratases, we would use robust Flexizyme technology to introduce the unnatural amino acid Se-phenylselenocysteine (SecPh), which undergoes oxidative elimination with H_2O_2 to generate dehydroalanines (Dhas) for the pyridine forming cycloaddition.[23−25] Flexizymes are aptamers developed to condense a wide array of amino acid esters with tRNAs of choice (Appendix figure B.6), allowing codon reprograming in in vitro transcription−translation systems.[26,27] In total, this would cut the number of enzymes or proteins necessary to prepare a thiopeptide in vitro from six, in the case of thiocillin (Tcl IJKLMN), to three (LynD, TbtE, and TclM, figure 3.1b). Although quantities of peptide made by in vitro transcription−translation and Flexizyme reprogramming are small relative to other technologies, such as amber codon suppression / in cell expression or solid-phase peptide synthesis, the approach is rapid (~2 h), robust, and flexible for peptides of this size and complexity. Thus, we anticipate that this strategy

might ultimately enable rapid characterization of new pyridine synthases and associated enzymes and aid elucidation of new thiopeptides.

3.2 Results and Discussion

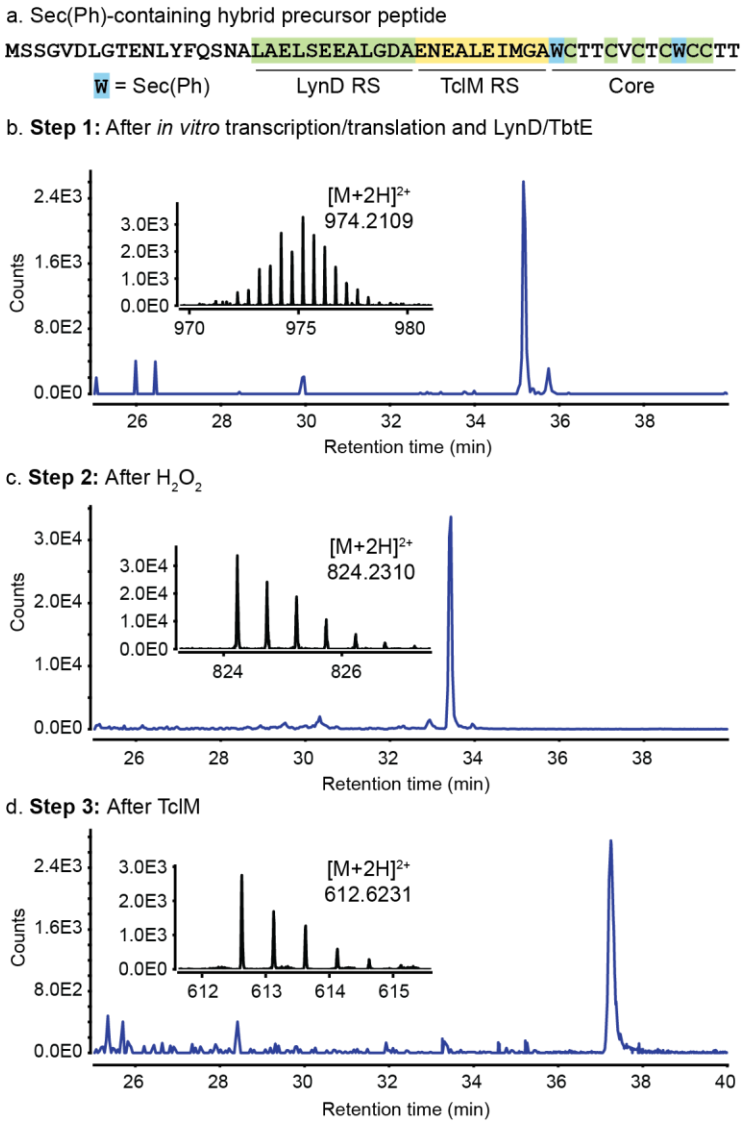


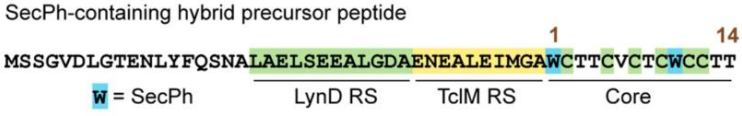
Figure 3.2. Flexizyme-enabled benchtop biosynthesis of a thiocillin scaffold. (a) Sequence of the designer precursor peptide, including a pMCSG7-derived sequence, the LynD RS, and TclM RS. The Trp codon was reprogrammed to incorporate SecPh. EICs for SecPh and hexathiazole precursor peptide after LynD/TbtE treatment (b), hexathiazole and Dha-containing product after oxidative elimination (c), and fully cyclized thiocillin core (d).

To validate this strategy, we first had to confirm that existing Flexizymes could incorporate SecPh into in vitro translated peptides and that LynD and TbtE could accommodate SecPh at

cysteine adjacent positions. We confirmed that the dinitrobenzyl (Dnb) ester-specific Flexizyme dFx could ligate the Dnb ester of SecPh by means of an in vitro microhelix assay (Appendix figure B.2).[25] Hybrid substrates were prepared by codon reprogramming the tryptophan codon – although Flexizyme allows a number of codons to be reprogrammed – because of the scarcity of Trp residues in thiopeptides, using the AsnE2 tRNA body.[28] The loaded AsnE2trp-tRNA was used to incorporate SecPh at positions Ser1 and Ser10 of the thiocillin core. For ease, initial DNA templates were prepared by cloning into plasmid pMCSG7 which necessarily incorporated an N-terminal sequence tag leading to the design of our test substrate (figure 3.2a), and transcription−translation reactions were prepared on a 2.5μL analytical scale using NEB PURExpress.[26,27] In the event, LynD was able to convert all six cysteines in several test substrates into thiazolines, which were further processed to thiazoles by TbtE as confirmed by high-resolution liquid chromatography and mass spectrometry (HR LC/MS) (figure 3.2b and Appendix figure B.4). Subsequent treatment with 1 M H_2O_2 efficiently converted both SecPh residues to Dhas (figure 3.2c). In the last proof-of-concept step, the excess H_2O_2 could be quenched by addition of tris(2-carboxyethyl)-phosphine hydrochloride (TCEP) and TclM added in situ provided complete conversion of the linear substrate into a cyclic thiopeptide (figure 3.2d), confirming that TclM is compatible with this designed leader strategy.

We next began to explore the requirements of the designer leader peptide. The N-termini of natural thiopeptide leader peptides have been implicated in an affinity enhancing interaction with pyridine synthase enzymes.[29] Although our studies have shown this interaction dispensable for pyridine synthase processing,[22] we chose to test the potential impact of changes to the N-terminus on processing of designer leader peptide substrates. Thus, we prepared two new hybrid substrates in which we replaced the original pMCSG7-derived N-terminus with two different

SecPh-containing hybrid precursor peptide



1A	Hybrid Leader Peptide Sequences	Thz	Pyr
1	MSSQLAELSEEALGDAENEALEIMGA	0.16	0.01
2	MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGA	1.00	1.00
3	MSEIKKALNTLEIEDFLAELSEEALGDAENEALEIMGA	0.76	0.69
4	MSEIKKALNTLEIEDFDAIEMVDVDAMPLAELSEEALGDAENEALEIMGA	0.73	0.22
5	MSSGVDLGTENLYFQSNALAELSEEALGDAAAAAAENEALEIMGA	0.80	0.31
6	MSSGVDLGTENLYFQSNALAELSEEALGDAGGGGGENEALEIMGA	0.13	0.11
7	MSSGVDLGTENLYFQSNALAELSEEALGDAKKKKKENEALEIMGA	0.18	0.09
8	MSSGVDLGTENLYFQSNALAELSEEALGDADDDDDENEALEIMGA	0.20	0.00
9	MSSGVDLGTENLYFQSNALAELSEEALGDAPPPPPENEALEIMGA	0.04	0.05
10	MLAELSEEALGDASEIKKALNTLEIEDFDAIEMVDVDAMPENEALEIMGA	0.04	0.03

1B	Core Amino Acid Sequences	Thz	Pyr
1	S C T T C V C T C S C C T T	✓	n/a
2	F C T T C V C T C F C C T T	✓	n/a
3	W C T T C V C T C W C C T T	✓	n/a
4	W C T T C K C T C W C C A A	✓	✓
5	W C T T C D C T C W C C A A	✓	✓
6	W C A A C V S A C W C C A A	✓	✓
7	W C T A C A C A C W C A	✓	✗
8	W C T A C A C A C W C C	✗	✗
9	W C T A C A C A C W C C A	✓	✓
10	W C T A C A C A C W C C T	✓	✗
11	W C T A C A C A C W C C A A	✓	✓
12	W C T A C A C A C W C C A G	✓	✓
13	W C T T C V C T C W C C A A N S G G V S	✓	✓

Table 3.1. Results of flexizyme-enabled benchtop biosynthesis with mutant leader peptides (1A) and cores (1B). EIC areas are relative to entry 2 as a standard. Checks indicate a detected EIC. An "X" indicates no EIC detected above the noise threshold of 1.0×10^2.

excerpts from the N-terminus of TclE, the native precursor peptide for TclM (table 3.1A, entries 3 and 4). In a third substrate, we truncated the leader peptide leaving only a short MSSQ tag before the LynD RS (table 3.1A, entry 1). The relative impact of these changes was assessed by integration of extracted ion chromatograms (EICs) for the product (table 3.1A, entry 2; figure 3.2c,d). Interestingly, the TclE fragment sequences decrease thiopeptide formation and seem to negatively impact LynD processing. Moreover, removal of the pMCSG7-derived sequence greatly reduces processing by the LynD / TbtE pair. Taken together, these results further confirm that the N-terminus of TclE is dispensable for TclM processing, but LynD may be sensitive to the location of its cognate RS within the larger peptide context. To further probe the latter aspect, we designed

a series of leader sequences in which spacers were introduced between the LynD RS and TclM RS (table 3.1A, entries 5−10).[20] In almost all cases, the substrates were converted to thiopeptides, suggesting that LynD is broadly tolerant of diverse sequence space between the RS and core, although at reduced efficiency. As an extreme example of this spacing promiscuity, a substrate with the LynD RS sequence N-terminal to the complete native TclE leader peptide was made and subsequently processed to the mature thiopeptide (table 3.1A, entry 10). This last result suggests a potentially broadly applicable strategy for circumventing reconstitution of all pathway enzymes in thiopeptide formation: express the full leader as a C-terminal fusion to LynD RS.

We next examined allowable changes to the core sequence. The TclM-containing thiocillin pathway has been shown to tolerate a wide array of changes to the core peptide in vivo; we therefore focused on changes to the core that were unproductive in those studies.[13,30] For example, Val6 of native thiocillin had been recalcitrant to charged or hydrophilic residues, such as lysine or aspartic acid. This limitation is a barrier to antibiotic development, as Val6 appears to be an ideal position for modulating the solubility.[31] In contrast, hybrid substrates bearing a V6D or V6K mutation were readily transformed to the cyclic thiopeptide in vitro (table 3.1B, entries 4 and 5). Additionally, the LynD / TbtE pair in vitro provided greater product control relative to the native thiocillin enzymes, TclI, TclJ, and TclN, as exemplified by a C7S mutant (table 3.1B, entry 6). In the in vivo system, a similar mutant gave mixtures of different modifications and apparent misprocessing. LynD, however, modifies all cysteines indiscriminately and left the newly introduced serine untouched. We focused considerably more mutagenesis on the C-terminus of the core, because studies have suggested TclM might be sensitive to C-terminal modifications,[22,32] and such modifications would be necessary for linking the current hybrid substrates with mRNA display in the future (table 3.1B, entries 7−13). Deletion of even one amino acid from the C-

terminus was unfavorable for TclM and / or LynD processing (table 3.1B, entries 7−10). In contrast, extending the C-terminus (table 3.1B, entry 13) did not significantly impact enzymatic processing. These data suggest that the hybrid strategy is amenable to C-terminal extensions and new sequences not previously accessible by in vivo engineering approaches. However, more extensive investigations will be needed to understand the limitations.

We last sought to test whether this strategy could be used to reconstitute new pyridine synthases. Recent work has suggested that only a fraction of genetically encoded thiopeptides have been isolated.[33] Of the >500 predicted thiopeptide gene clusters, the largest family comprises members that contain a close homologue of LazC, the predicted pyridine synthase from lactazole biosynthesis.[34] Additionally, while the LazC homology is high in this family, the core peptide diversity is broad, suggesting LazC and its homologues may exhibit unique substrate promiscuity (Appendix figure B.5). Despite the preponderance of predicted LazC homologues, LazC has not yet been reconstituted in vitro. Therefore, we expressed and purified LazC as its MBP-fusion and designed three new hybrid sequences as potential substrates (Appendix figure B.3). In one LazC substrate, we integrated LynD RS directly into the native lactazole leader at a site with apparent natural homology (termed LacHyb1, figure 3.3a), while in the other two, LynD RS was encoded 15 or 25 residues N-terminal to the core (termed LacHyb2 and 3, respectively). Additionally, Trp2 and oxazole-forming Ser11 were replaced with a serine and thiazole-forming cysteine, respectively (Appendix figure B.3). Both mutations were previously produced by a gene replacement strategy, suggesting that the double mutant may also be a LazC substrate and render the core compatible with Trp reprogramed SecPh incorporation. Upon treatment with LynD / TbtE, the four cysteines in each substrate readily underwent conversion to thiazoles and subsequent H_2O_2 oxidation

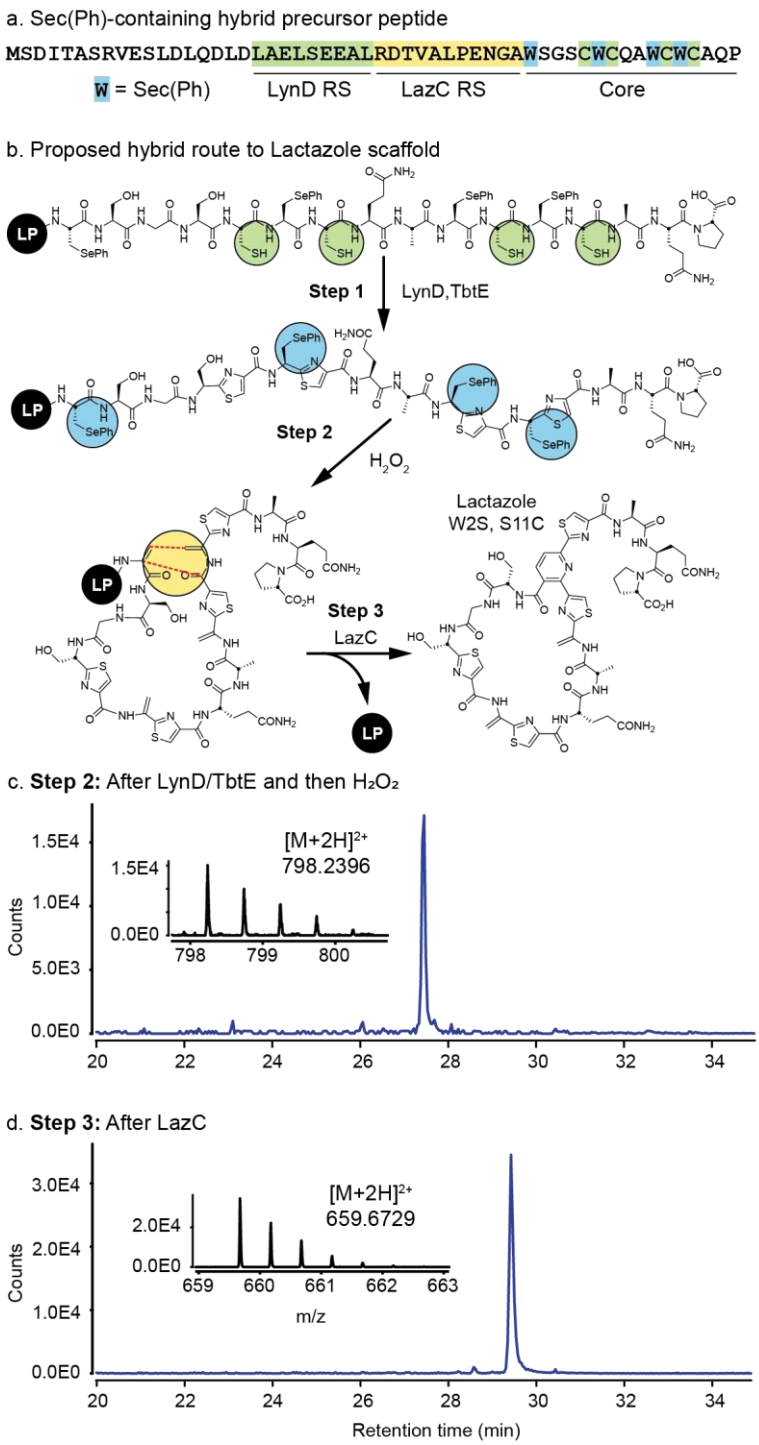


Figure 3.3. Flexizyme-enabled benchtop biosynthesis of lactazole W2S, S11C. (a) Sequence of designer precursor peptide, LacHyb1. (b) Proposed route to lactazole scaffolds. EICs for the thiazole and Dha containing product of treatment with LynD / TbtE and H_2O_2 (c) and fully cyclized lactazole core (d).

introduced the four Dhas (figure 3.3c). Lastly, treatment with LazC efficiently yielded the new, pyridine-containing masses, thus confirming activity of LazC as a pyridine synthase (figure 3.3d). This result was consistent for all three hybrid lactazole leader peptides (Appendix figure B.4.22-24 and B.4.42-44) and suggests that the hybrid in vitro strategy should work on many other, yet uncharacterized pyridine synthases and could ultimately allow elucidation of new thiopeptides.

3.3 Conclusion

In summary, we have developed a new, facile strategy to access thiopeptide backbones. This approach combines robust, Flexizyme-assisted incorporation of chemical handles into in vitro transcribed-translated peptides with three unrelated RiPP enzymes (LynD, TbtE, and TclM) by using designer leader peptides. We demonstrate the ability to make thiocillin variants previously unattainable through natural biosynthetic processes and use this strategy to reconstitute the pyridine synthase LazC to make lactazoles for the first time in vitro. We anticipate this approach will be useful in making new thiopeptide variants with therapeutic potential, in studying more pyridine synthases and associated enzymes, and may aid elucidation of new thiopeptide structures. Finally, we anticipate that this strategy will be compatible with high-throughput screening techniques, such as mRNA display, which is a current focus in our lab.

3.4 Experimental

3.4.1 General Information.

All molecular biology materials including the PURExpress kit (E6840S) were purchased from New England BioLabs (Ipswich, MA) and Thermo Fisher Scientific (Waltham, MA). DNA and primers were purchased from Integrated DNA Technologies (Coralville, IA) and Eton Bioscience, Inc (San Diego, California). Escherichia coli DH5α, BL21DE3, and BL21-CodonPlus

(DE3)-RIPL strains were used for plasmid maintenance and protein expression. Genes infused into plasmids pMCSG7 and pMCSG9 were sequenced by Genewiz (South Plainfield, NJ). Nano-LCESI-MS data was collected using nanoAcquity (Waters, Milford, MA) coupled to a TripleTOF5600 (AB Sciex, Framingham, MA). All Nano-LCESI-MS data was collected and converted to mzml files and analyzed using MZMine2.[35] Extracted ion chromatograms were generated with a 10-ppm error cutoff. Peak areas were calculated using the targeted peak detection analysis with intensity tolerance of 50%, m/z tolerance of 10 ppm, noise level of 1.0E2, and retention time tolerance of 1.0 min. For nickel affinity purification of proteins, HiTrap columns were used and purchased from GE Healthcare (Chicago, IL). All chemical reactions were performed in oven-baked round-bottom-flasks under inert argon atmosphere with stirring. Solvents, reagents, and chemicals were purchased through Fisher Scientific. N-Boc-L-Serine-beta-Lactone was purchased from Ark Pharm (Arlington Heights, IL). Spectra for 1H were recorded at room temperature with Varian Inova 400 (400 MHz). Chemical shifts are reported in δ (ppm) relative units to residual solvent peaks CDCl3 (7.26 ppm) and DMSO -d6 (2.5 ppm). Splitting patterns are assigned as s (singlet), d (doublet), t (triplet), q (quartet), quint (quintet), multiplet (m), and dd (doublet of doublets).

3.4.2 Protein Cloning.

The genes for LynD and TbtE were purchased as gBlocks from IDT®, PCR amplified with Q5 polymerase, and cloned into pMCSG7 and pMCSG9, respectively using the following primers:

LynD-Forward (5'→3'): TACTTCCAATCCAATGCGATGCAAAGCACACCACTGC

LynD-Reverse (5'→3'): TTATCCACTTCCAATGCGCTATTAGAACGGCATTGGGGTC

TbtE-Forward (5'→3'):

TACTTCCAATCCAATGCGATGACAATCCCACCCGGTCTGACGGAACGCTACGC

TbtE-Reverse (5'→3'):

TTATCCACTTCCAATGCGCTATTAAGATTCTTCCACCGTAGAACCCGAAGCTG

The PCR product was phosphorylated with T4 PNK and treated with T4 DNA polymerase to create LIC-overhangs. Plasmids pMCSG7 and pMCSG9 were linearized with SSPI, dephosphorylated with Antarctic Phosphatase, and treated with T4 DNA Polymerase to create LIC-overhangs. Prepared vectors and PCR products were incubated for 10 min at room temperature and then transformed into One-Shot® Top 10. pMCSG7-LynD was then transformed and expressed in BL21-CodonPlus (DE3)-RIPL Competent Cells (Agilent). pMCSG9-TbtE was then transformed into BL21 (DE3) electrocompetent cells.

TclM was cloned as in the following report.[18]

MBP-LazC containing plasmid was purchased from JGI.

3.4.3 Protein Expression Purification.

Expression and Purification of His6-LynD

His6-LynD protein was expressed and purified from E. coli RIPL cells as in the following report.[21] A 5 mL saturated culture was used to inoculate 1 L of Luria-Bertani (LB) medium supplemented with ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL). The culture was incubated at 37 oC and shaken 220 rpm until reaching an OD 600 of 0.6-0.8 at which point the medium was supplemented with 0.2 mM IPTG. Culture was then cooled to 18 ^{o}C and grown overnight ~20 hours. Cells were pelleted and stored at -80 ^{o}C until purification. Pellets were

resuspended in 40 mL of wash buffer (20 mM Tris pH 8.0, 200 mM NaCl, 50 mM Imidazole, and 1 mM TCEP) supplemented with 1.5 mM Phenylmethylsulfonyl fluoride (PMSF), 80 units of DNAseI, 0.25mg lysozyme and sonicated twice with intermittent pulses with a 30 % maximum amplitude for 1:30 min. The soluble protein was recovered by pelleting the cell debris by centrifugation at 4 ^{o}C 15,000 rpm for 60 mins and was filtered through a 0.45µm sterile syringe filter. The supernatant was loaded onto a 5-mL HisTrap (Ni2+) IMAC column and washed with 5 column volumes (CV) of wash buffer. Protein was eluted with elution buffer (20 mM Tris pH 8.0, 200 mM NaCl, 500 mM imidazole, 1 mM TCEP) over a gradient 0-100 % over 10 CV. Fractions collected which contained purified protein were collected and concentrated to 2.5 mL using a Centricon (10,000 Da MWCO) concentrator (EMD Millipore®). The concentrated protein was then buffer exchanged using PD-10 column (GE Healthcare Life Sciences®) into storage buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 1 mM TCEP and stored at -80 ^{o}C. A 1 L culture yielded ~22.6 mg of protein.

Expression and Purification of His6-MBP-TbtE

His6-MBP-TbtE protein was expressed in E. coli (DE3) cells as in the following report.[16] A 5 mL saturated culture was used to inoculate 1L of LB medium supplemented with ampicillin (100 µg/mL). The culture was incubated at 37 ^{o}C and shaken 220 rpm until reaching an OD 600 of 0.6-0.8 at which point the medium was supplemented with 0.2 mM IPTG. Culture was then cooled to 16 ^{o}C and grown overnight ~20 hours. Cells were pelleted and stored at -80 ^{o}C until purification. Pellets were resupended in 40 mL of wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 2.5 % glycerol, 0.1 % Triton X-100) supplemented with 1.5 mM PMSF, 80 units of DNAse I, 0.25 mg lysozyme and sonicated twice with intermittent pulses with a 30% maximum amplitude for 1:30 min. The soluble protein was recovered by pelleting the cell debris by centrifugation at 4 ^{o}C

15,000 rpm for 60 mins and was filtered through a 0.45 µm sterile syringe filter. The supernatant was loaded onto a 5-mL HisTrap (Ni2+) IMAC column and washed with 5 column volumes (CV) of wash buffer (50 mM Tris pH 7.5, 400 mM NaCl, 2.5% glycerol, 50 mM imidazole). Protein was eluted with elution buffer (50 mM Tris pH 7.5, 300 mM NaCl, 2.5% glycerol, 500 mM imidazole) over a gradient 0-100% over 10 CV. Fractions collected which contained purified protein were collected and concentrated to < 5.0 mL using a Centricon (30,000 Da MWCO) concentrator (EMD Millipore®). The concentrated protein was then buffer exchanged by diluting the concentrated protein in ~10-15 mL of storage buffer (50 mM HEPES pH 7.5, 300 mM NaCl, 2.5% glycerol), and re-concentrating. This was repeated 4 times and the protein was stored at -80 ^{o}C. A 1 L culture yielded ~33.12 mg of protein.

Expression and Purification of His6-MBP-TclM

His6-MBP-TclM was expressed and purified as in the following report.[18] Briefly, MBP-TclM protein was expressed in E. coli RIPL cells. A 5 mL saturated culture was used to inoculate 1L of LB medium supplemented with ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL). The culture was incubated at 37 ^{o}C and shaken 220 rpm until reaching an OD 600 of 0.6-0.8 at which point the medium was supplemented with 0.2 mM IPTG. Culture was then cooled to 16 ^{o}C and grown overnight ~24 hours. Cells were pelleted by centrifugation (5,000 rpm, 4 ^{o}C, 30 mins) and stored at -80 ^{o}C until purification. Pellets were resuspended in 40 mL of wash buffer (50 mM $KHPO_4$ pH 7.0, 250 mM KCl, 10 mM Imidazole, 10 % glycerol) supplemented with 1.5mM PMSF, 1 mg DNAseI, and 1 protease inhibitor tablet (ThermoFisher, 88666) and sonicated twice with intermittent pulses with a 30% maximum amplitude for 1:30 min. The soluble protein was recovered by pelleting the cell debris by centrifugation at 4 ^{o}C 15,000rpm for 60 mins and was filtered through a 0.45 µm sterile syringe filter. The supernatant was loaded onto a 5-mL HisTrap

(Ni2+) IMAC column and washed with 10 CV of wash buffer. Protein was eluted with elution buffer (50 mM $KHPO_4$ pH 7.0, 250 mM KCl, 500 mM imidazole, 10 % glycerol) over a gradient 0-100% over 10 CV. Fractions collected which contained purified protein were collected and were determined to be too impure for assays. The protein was desalted using a PD-10 column and loaded again onto a 5-mL HisTrap (Ni2+) IMAC column and washed with 10 CV of wash buffer with increased imidazole concentration (50 mM $KHPO_4$ pH 7.0, 250 mM KCl, 50 mM imidazole, 10% glycerol). Protein was eluted with elution buffer (50 mM $KHPO_4$ pH 7.0, 250 mM KCl, 500 mM Imidazole, 10% glycerol) over a gradient 0-100% over 10 CV. Fractions containing pure protein were collected and concentrated to 2.5 mL using a Centricon (50,000 Da MWCO) concentrator (EMD Millipore®). The concentrated protein was then buffer exchanged using PD-10 column (GE Healthcare Life Sciences®) into storage buffer (50 mM $KHPO_4$ pH 7.0, 250 mM KCl, 10% glycerol), concentrated to 1 mL, and stored at -80 ^oC. A 1 L culture yielded ~2.8 mg of protein.

Expression and Purification of His6-MBP-LazC

His6-MBP-LazC protein was expressed in E. coli cells. A 5 mL saturated culture was used to inoculate 1L of LB medium supplemented with ampicillin (100µg/mL). The culture was incubated at 37 ^oC and shaken 220 rpm until reaching an OD600 of 0.6-0.8 at which point the medium was supplemented with 0.2 mM IPTG. Culture was then cooled to 16 ^oC and grown overnight ~21 hours. Cells were pelleted by centrifugation (5,000 rpm, 4 oC, 30 mins) and stored at -80 ^oC until purification. Pellets were resuspended in 40 mL of wash buffer (50 mM $KHPO_4$ pH 7.0, 250 mM KCl, 50 mM imidazole, 10% glycerol) supplemented with 1.5 mM PMSF, 80 units DNAseI, and 1 protease inhibitor tablet (ThermoFisher, 88666) and lysed using a french press with a maximum pressure of 1500 psi. The soluble protein was recovered by pelleting the cell debris by centrifugation at 4 ^oC 15,000 rpm for 60 mins and was filtered through a 0.45 µm sterile syringe

filter. The supernatant was loaded onto a 5-mL HisTrap (Ni2+) IMAC column and washed with 10 CV of wash buffer. Protein was eluted with elution buffer (50 mM $KHPO_4$ pH 7.0, 250 mM KCl, 500 mM imidazole, 10% glycerol) over a gradient 0-100% over 10 CV. Fractions containing pure protein were collected and concentrated to 2.5 mL using a Centricon (50,000 Da MWCO) concentrator (EMD Millipore®). The concentrated protein was then buffer exchanged using PD-10 column (GE Healthcare Life Sciences®) into storage buffer (50 mM $KHPO_4$ pH 7.0, 250 mM KCl, 10% glycerol), and stored at -80 ^{o}C. A 1 L culture yielded ~34 mg of protein.

3.4.4 Protein Sequences.

His6-LynD

MHHHHHHSSGVDLGTENLYFQSNAMQSTPLLQIQPHFHVEVIEPKQVYLLGEQANHAL
TGQLYCQILPLLNGQYTLEQIVEKLDGEVPPEYIDYVLERLAEKGYLTEAAPELSSEVAA
FWSELGIAPPVAAEALRQPVTLTPVGNISEVTVAALTTALRDIGISVQTPTEAGSPTALNV
VLTDDYLQPELAKINKQALESQQTWLLVKPVGSVLWLGPVFVPGKTGCWDCLAHRLR
GNREVEASVLRQKQAQQQRNGQSGSVIGCLPTARATLPSTLQTGLQFAATEIAKWIVKY
HVNATAPGTVFFPTLDGKIITLNHSILDLKSHILIKRSQCPTCGDPKILQHRGFEPLKLESR
PKQFTSDGGHRGTTPEQTVQKYQHLISPVTGVVTELVRITDPANPLVHTYRAGHSFGSA
TSLRGLRNTLKHKSSGKGKTDSQSKASGLCEAVERYSGIFQGDEPRKRATLAELGDLAI
HPEQCLCFSDGQYANRETLNEQATVAHDWIPQRFDASQAIEWTPVWSLTEQTHKYLPT
ALCYYHYPLPPEHRFARGDSNGNAAGNTLEEAILQGFMELVERDGVALWWYNRLRRP
AVDLGSFNEPYFVQLQQFYRENDRDLWVLDLTADLGIPAFAGVSNRKTGSSERLILGFG
AHLDPTIAILRAVTEVNQIGLELDKVPDENLKSDATDWLITEKLADHPYLLPDTTQPLKT
AQDYPKRWSDDIYTDVMTCVNIAQQAGLETLVIDQTRPDIGLNVVKVTVPGMRHFWSR
FGEGRLYDVPVKLGWLDEPLTEAQMNPTPMPF

His6-MBP-TbtE

MHHHHHHSSGVDLGTKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKL
EEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNG
KLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIA
ADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNK
GETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKE
FLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMS
AFWYAVRTAVINAASGRQTVDEALKDAQTDYDIPGTENLYFQSNAMTIPPGLTERYALR
AGVHSAVLPDGVMRLFAWPHAESIGALSADETTLLKELAEGPREIADPALRPFVERLFR
GGWLKRTLSRGEHDLYTLDPLRRPGSRPAPPDDPVLSRFAAVRRRPSGFVIESPLAWCD
VHVHDPALLPDLLEPAGGRAGRSSLAPQIRRQALADLAWAGLVVPRGAEDGALRTRQ
WAPHELDFHQRSRLYHRGYLGDGFGGTFWARGTFDPPQARPQRYPGDPIPLHRPDLNA
LRAADPPLTAVLEDRRSVREYDDDAPMTVEQLGELLYRSARIRDVKVIDGVEYYVRKPYP
SGGSVYELEIYPVVRHVAGLAPGMYHYDAYEHVLRPVRPAGHPAVRRMLTVASHGSA
VGIRPQLLLVVSARVGRVMWKYEGMGYALILKHVGVLYQTLYCVATAMGLAPCAIGS

GDSAAFSEATGRDPLEECAVGDFLLGSRPASGSTVEES

His6-MBP-TclM

MHHHHHHSSGVDLGTKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKL
EEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNG
KLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIA
ADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNK
GETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKE

FLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMS
AFWYAVRTAVINAASGRQTVDEALKDAQTDYDIPGTENLYFQSNAMEQYHKIVLTGSN
AETMLIKNIEPVAVKFINNYKGFFYVFKYSKDFPIIDVYINNKIVTENQLNKILQNSGAKY
KIYKNSIFNETQGNFGMLGDKYLAEFFKKTNEISLNILNQNFESYNKKIEFALEIMLISAH
YNYDSIKKGYLSYASHVNGFFTRWKDPNKIRDIFHKNYLNNKEYLESKVSEIIDNNNRSS
LSELSDIITEMKKEMTTDIEKGNLHVFNIELLQKPGERDFLEKSQFHKTILNNPDFSNFMN
KDINFLGSRLITVFTYLLIRNLGIQNKDRYLLCYYIYKIIEEKYNIDTLELIRDFGKGRDNN
VEDLQRY

His6-MBP-LazC

MHHHHHHSSGVDLGTKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKL
EEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNG
KLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIA
ADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNK
GETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKE
FLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMS
AFWYAVRTAVINAASGRQTVDEALKDAQTDYDIPGTENLYFQSNAMSDPADGRGAVT
AWDVVLYHYRPDKARALREAVLPLARQAAAEGLAAHVERHWRFGPHLRLRLRGPEAR
VAGAAQRAAEALRAWAAAHPSVADRSDEQLLAEAAVAGRAELIAPPYAPLVPDNTVV
AAPADRSAEDALRALIGAESAELREELLRTGLPALDSACHFLGAHGDTPQARVQLVVTA
LAAHATAHPDGLVGAHYSVLSHLEDFLVHEDPDGSLRAAFERRWEQSGRAVTALVGRI
ADGGARDWERDWAHWSATAWSLAERRLTAGADLGGRHAEYRERAEALGDPATAER
WNAELRTRYSEFHRMLQRADPDGRMWHRPDYLINRAGTNGLYRLLAICDVRPMERYL
AAHLLVRSVPELTGHRWQTLLGAAEQPGGPEQSGAAGATGGAGRTKLEGAA

3.4.5 DNA and RNA Preparation.

DNA templates for T7 RNA polymerase reactions were prepared for tRNA and flexizyme (dFx), through extension of extension primers and then PCR reactions with forward and reverse PCR primers (see Appendix B Translation Genes). DNA extension and amplification reactions were performed with Taq Polymerase.[36]

Table 3.2. Extension protocol *(5 cycles of annealing and elongation)*

Stage	Temperature	Time
Denature	95 ºC	1 min
Annealing	50 ºC	1 min
Elongation	72 ºC	1 min
Final Elongation	72 ºC	2 min

Table 3.3. PCR1 protocol *(5 cycles of denature, annealing, and elongation)*

Stage	Temperature	Time
Initial Denature	95 ºC	1 min
Denature	95 ºC	40 sec
Annealing	50 ºC	40 sec
Elongation	72 ºC	40 sec
Final Elongation	72 ºC	2 min

Table 3.4 PCR2 protocol *(12 cycles of denature, annealing, and elongation)*

Stage	Temperature	Time
Initial Denature	95 ºC	1 min
Denature	95 ºC	40 sec
Annealing	50 ºC	40 sec
Elongation	72 ºC	40 sec
Final Elongation	72 ºC	2 min

dFx DNA Template (5'→3'):

GGCGGTAATACGACTCACTATAGGATCGAAAGATTTCCGCATCCCCGAAAGGGTAC

ATGGCGTTAGGT

AsnE2trp DNA Template (5'→3'):

GGCGTAATACGACTCACTATAGGCTCTGTAGTTCAGTCGGTAGAACGGCGGACTCCA

AATCCGTATGTCACTGGTTCGAGTCCAGTCAGAGCCGCCA

Table 3.5. dFx and AsnE2 primers

1	dFx_Ex_F	GTAATACGACTCACTATAGGATCGAAAGATTTCCGC
2	dFx_Ex_R	ACCTAACGCCATGTACCCTTTCGGGGATGCGGAAATCTTTCGATCC
3	T7_Prom_F	GGCGTAATACGACTCACTATAG
4	dFx_PCR_R	ACCTAACGCCATGTACCCT
5	AsnE2_Ex_F	GTAATACGACTCACTATAGGCTCTGTAGTTCAGTCGGTAGAACGGCGGA
6	$AsnE2_{trp}$_Ex_R	GAACCAGTGACATACGGATTTGGAGTCCGCCGTTCTACCGACT
7	T7_Prom_F	GGCGTAATACGACTCACTATAG
8	AsnE2_PCR_R	TGGCGGCTCTGACTGGACTCGAACCAGTGACATACGGA
9	AsnE2_PCR_R	T(mG)GCGGCTCTGACTGGACTC

Primers: (mG) designates 2'Ome Guanosine RNA

dFx DNA template preparation

dFx extension was performed with primers 1 and 2 using extension protocol. 5uL of the extension product was used to amplify 1 mL dFx using PCR2 protocol with primers 3 and 4. DNA was then purified and concentrated by phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. DNA was resolubilized in 100uL of water and directly used for T7 RNA transcription.

AsnE2trp tRNA DNA template preparation

AsnE2trp extension was performed with primers 5 and 6 using extension protocol. 10uL of the extension product was used to amplify 200 µL AsnE2 PCR1 reaction using primers 7 and 8. 5 µL of PCR1 product was directly added to 1 mL PCR2 reaction with primers 7 and 9. DNA was then purified and concentrated by phenol / chloroform / isoamyl alcohol extraction and ethanol precipitation. DNA was resolubilized in 100uL of water and directly used for T7 RNA transcription.

General T7 RNA polymerase transcription and RNA purification

T7 RNA transcription reactions were performed on a 1 mL scale according to NEB protocols with 100 µL of prepared DNA template. The only change from NEB protocol was the addition of extra $MgCl_2$ (dFx final concentration: 30 mM, AsnE2 tRNA final concentration: 22.5 mM) Transcriptions were left at 37 ^oC overnight and treated with 115 µL of 10x DNAseI buffer and 30 U of DNaseI (Thermo Scientific, EN0521) for 1 hour at 37 ^oC. Upon successful translations, magnesium pyrophosphate precipitates, so, 200 µL of 0.5 M ethylenediaminetetraacetic acid (EDTA) was added to clear the solution. RNA was precipitated by adding and mixing 1190 µL of isopropanol and pelleted by centrifugation at 13,000 rpm, 25 ^oC for 15 min. The supernatant was disposed of and the pellet was washed with 500 µL of 70 % ethanol, centrifuged for 3 min at 13,000 rpm. Again, the supernatant was disposed of and the pellet was dried for 5 min at room temperature. Both dFx and AsnE2 tRNA were resolubilized in 100 µL of water. 2x RNA loading dye was then added and the RNA was loaded onto a 12 % (dFx) or 8 % (AsnE2trp tRNA) Urea-PAGE gel and run for 1 hour at 230 V in TBE (Tris-Borate-EDTA) buffer. The RNA band was visualized by TLC shadowing and cut out of the gel. The removed gel slab was crushed into fine pieces and the RNA extracted twice in ~8 mL 0.3 M NaCl at room temperature. The combined

extractions were filtered through a 0.45 µm syringe filter and a 2x volume of ethanol was added to precipitate RNA. RNA was pelleted via centrifugation at 4,000 rpm for 25 min. The supernatant was removed, and the pellet washed with 5 mL of 70% ethanol and centrifuged at 4,000 rpm for 5 min and the supernatant discarded. The RNA pellets were then dried at room temperature for 5 min and resolubilized in water. Both dFx and AsnE2trp tRNA were diluted to a working stock of 250 µM and stored at -20 ºC.

DNA Preparations for Translation Templates for NEB PURExpress®

Genes with a 6xHis-Tag, were cloned into plasmid pMCSG7 (Appendix B Translation Genes). Other genes were purchased as shown from IDT® and amplified with Q5 high fidelity Polymerase or Taq polymerase. The primers used were:

T7 Forward (5'→3'): GAAATTAATACGACTCACTATAGGGG

T7 Terminator (5'→3'): GCTAGTTATTGCTCAGCGG

Q5 amplification cycles:

Perform 20 cycles starting at an annealing temp. of 58 ºC and lowering by 0.5 ºC for each subsequent cycle until 45 ºC. Then, perform 20 more cycles with an annealing temp of 45 ºC.

Table 3.6. Q5 PCR amplification

Stage	Temperature	Time
Initial Denature	98 ºC	30 sec
Denature	98 ºC	15 sec
Annealing	58 ºC- 45 ºC	20 sec
Elongation	72 ºC	1 min
Final Elongation	72 ºC	5 min

Taq amplification cycles: (25 cycles of denature, anneal, elongate)

Table 3.7. Taq PCR amplification

Stage	Temperature	Time
Initial Denature	95 ºC	2 min
Denature	95 ºC	30 sec
Annealing	51 ºC	30 sec
Elongation	72 ºC	2 min
Final Elongation	72 ºC	5 min

After amplification DNA products were column purified using Thermo Scientific GeneJET PCR Purification Kit® and analyzed by 3 % agarose gel. Concentrations were checked on nanodrop to be between 80-200 ng / µL, PCR products were stored at -20 ºC and directly used in translation reactions.

3.4.6 Flexizyme tRNA Acylation of Se-phenylselenocysteine.

Flexizyme assays were performed as described in the following report.[37] Briefly, reactions were carried out at 4 ºC for 24-72 hours in 50 mM HEPES-KOH (pH 8.0), 600 mM $MgCl_2$, 20 % DMSO, 25 µM dFx, 25 µM tRNA, 5 mM Se-phenylselenocysteine dinitrobenzyl ester.

3.4.7 Translations.

Translations were performed on a 2.5 µL scale using NEB PURExpress® Δ (aa, tRNA: NEB # E6840). For Flexizyme codon reprogramming, a 19 amino acid mix containing 5 mM of each amino acid except tryptophan was made in milliQ water and added to translation en lieu of the 20 amino acid mix provided with the kit along with 50-100 µM amino acyl-tRNA. Before adding to translation, acyl-tRNAs were thawed on ice and resuspended in 1 mM NaOAc (pH 5.2). Reactions were incubated for 1 hour at 37 ºC in an air incubator.

3.4.8 Enzyme Reactions and Oxidative Elimination.

His6-LynD / His6-MBP-TbtE Reactions

As translation reactions were incubating, a 2x enzyme mix was made containing 8 µM LynD, 6 µM TbtE, 10 mM ATP, 10 mM HEPES (pH 7.5), 150 mM NaCl, and 10 mM MgCl2. When translations finished 2x enzyme solution was added to make 1x and the reactions were incubated at 25 ºC overnight for ~16-24 hours. For MS analysis, 0.25 µL of 500 U/mL endoproteinase GluC was added and incubated at 37 ºC for > 5 hours before desalting using C18 stage tips.

Oxidative elimination and tris(2-carboxyethyl)phosphine (TCEP) Reduction

After LynD / TbtE enzyme reactions the sample was placed on ice and a 2x elimination solution was made containing 2M hydrogen peroxide and 397.89mM HEPES pH 9.0. After the solution was cooled, 2x elimination solution was added to 1x, allowed to incubate 1 hour on ice, and quenched with a final concentration of 250 mM TCEP (pH 9.0) at 37 ºC for 1 hour. For MS analysis, 0.25 µL of 500 U/mL endoproteinase GluC was added and incubated at 37 ºC for > 5 hours before desalting using C18 stage tips.

MBP-TclM/MBP-LazC Reactions

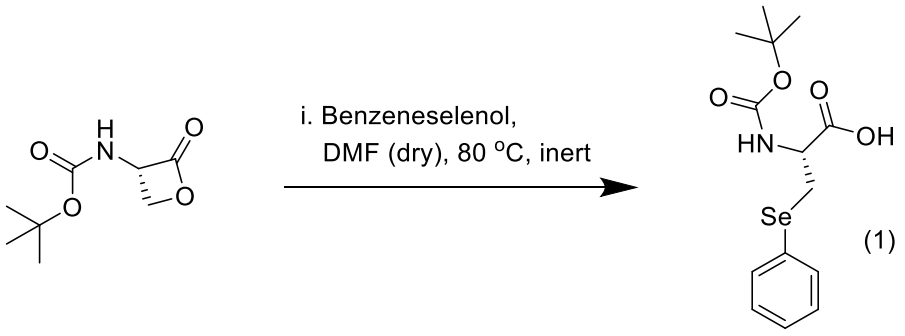
To complete processing of the thiopeptide core, pyridine synthases were added at a final concentration of 2 µM to the Dha and thiazole containing substrate and allowed to incubate at room temperature. Initial reactions were incubated overnight, but it was discovered that 1-hour was sufficient. Samples were desalted using C18 stage tips.

3.4.9 Desalting Protocol and Preparation for LC / MS.

After translation and / or enzymatic reactions, samples were desalted using Pierce® C18 stage tips according to manufacture protocols. Briefly, a large hole was made in the cap of a 1.7 mL snap cap tube to accommodate the stage tip. A stage tip was added to the tube, 15 µL of elution solution (80 % acetonitrile, 0.5 % acetic acid) was dispensed into the top of the tip and centrifuged using a tabletop centrifuge to wet the tip. 15 µL of wash solution (4% acetonitrile, 0.5% acetic acid) was added and again centrifuged through the tip using a tabletop centrifuge. Up to ~20 µL of sample was added and centrifuged to load the tip. The tip was then washed twice by adding 15 µL of wash solution and centrifuged. Finally, the tip was removed from the 1.7 mL tube, placed into a clean tube and eluted with 4 µL of elution solution via centrifugation. The eluted solution was then diluted with 6 µL of 0.1 % formic acid and submitted for nano-LC ESI-MS / MS.

3.4.10 Synthesis of Se-phenylselenocysteine dinitrobenzyl ester.

Boc-Se-phenylselenocysteine (1)

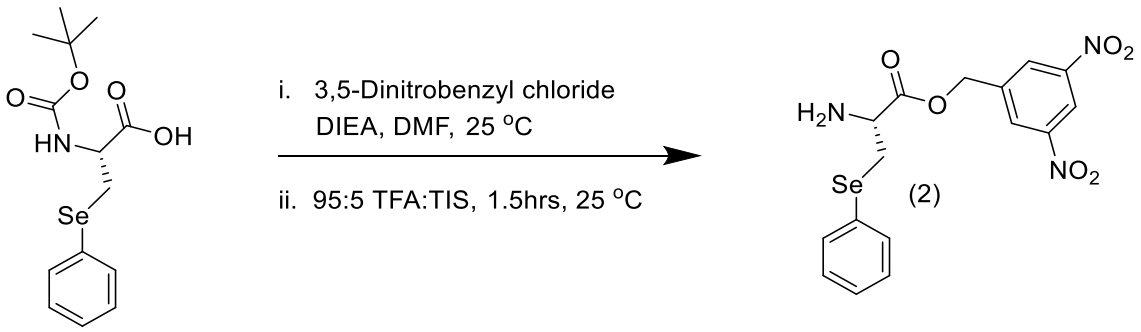


Boc-Se-phenylselenocysteine was synthesized using protocols found in the following report.[38] 508.4 mg of N-boc-serine lactone was added to a 25 mL oven-dried round bottom flask with 4.94 mL of anhydrous dimethylformamide (DMF) under inert atmosphere (argon). To it was added 375 µL (554.6 mg) of benzeneselenol before the reaction was moved to an 80 °C oil bath

where it was kept for 2 hours and the reaction monitored by TLC (9:1 DCM:MeOH). Upon completion, the reaction was quenched by the addition of 12-14 mL of 0.33M NaOH and washed with 30mL of diethyl ether. The organic layer was discarded, and the aqueous remains was acidified to pH 3.0 with 1N HCl and extracted multiple times with 50 mL ethyl acetate. All extractions were combined and dried over Na_2SO_4, filtered, and concentrated under reduced pressure to produce a soft yellow oil (840 mg).

1H NMR (400 MHz, Chloroform-d) δ 7.56 – 7.50 (m, 2H), 7.22 (p, J = 4.0, 3.5 Hz, 3H), 5.37 (d, J = 7.7 Hz, 1H), 4.64 (q, J = 6.1, 5.7 Hz, 1H), 3.46 – 3.35 (m, 1H), 3.31 (dd, J = 13.0, 5.4 Hz, 1H), 1.38 (s, 9H).

Se-phenylselenocysteine dinitrobenzyl ester (2)



Boc-Se-phenylselenocysteine dinitrobenzyl ester was synthesized as in the following report.[37] 120 mg of 1 was added to a 10-mL round bottom flask with 83 mg of 3, 5 dinitrobenzyl chloride and solubilized in 417 µL of anhydrous dimethylformamide under inert atmosphere. 67 µL (35.5 mg) of diisopropylethylamine was added and allowed to stir overnight at room temperature. The product was then extracted with 7 mL of diethyl ether and washed 2x with 4 mL 0.5 M HCl, 2x with saturated NaHCO3, and finally 2x with saturated NaCl. Finally, the organic layer was dried over sodium sulfate, filtered, and concentrated under reduced pressure. The

product was then purified by normal phase chromatography with hexanes and ethyl acetate. Purified fractions were collected and concentrated under reduced pressure. To the dried product was added 1 mL of 95:5 trifluoroacetic acid (TFA):triisopropylsilane under inert atmosphere and allowed to stir up to 1.5 hours. Reaction progress was monitored by TLC and the TFA was removed under reduced pressure. After 5 washes with DCM, the product changed from a yellow oil to a white powder. This was solubilized in minimal acetonitrile and precipitated with 10x diethyl ether. The product was collected by centrifugation, the supernatant was discarded, and the pellet dried (75 mg).

1H NMR (400 MHz, DMSO-d6) δ 8.81 (t, J = 2.1 Hz, 1H), 8.66 (dd, J = 1.8, 1.0 Hz, 2H), 7.55 – 7.49 (m, 2H), 7.31 – 7.24 (m, 3H), 5.41 – 5.36 (m, 1H), 5.24 – 5.19 (m, 1H), 4.46 (t, J = 5.9 Hz, 1H), 3.47 – 3.38 (m, 3H).

3.4.11 Liquid Chromatography and Mass Spectrometry.

Reaction products were analyzed via a nano-LC ESI MS / MS platform: nanoAcquity (Waters, Milford, MA) coupled to a TripleTOF5600 (AB Sciex, Framingham, MA). GluC treated, and undigested samples were diluted, 2.5x and 1.5x respectively, in water with 5 % acetonitrile and 0.1 % formic acid prior in LC MS / MS analysis. 5 µL of each sample was injected onto a trap column (NanoAcquity UPLC 2G-W/M Trap 5 µm Symmetry C18, 180 µm × 20 mm: Waters) coupled to an analytical C18 column (10k PSI, 100 Å, 1.8 µm, 75 µm × 250 mm: Waters). Peptide separation was carried out at a flow rate of 0.3 µL / min using a linear ramp of 5%−50% B (mobile phase A, 0.1% formic acid; mobile phase B, 0.1% formic acid in acetonitrile) over 30 min. The MS was operated in positive-ion, high-sensitivity mode with the MS survey spectrum using a mass range of 350−1600 Da in 250 ms and information-dependent acquisition of MS / MS data. The

first 20 features above 150 counts threshold and having a charge state of + 2 to + 5 were fragmented using rolling collision energy (± 5 %). Auto calibration was performed every eight samples (8 h).

REFERENCES

(1) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K-D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hem-scheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. *Nat. Prod. Rep.* **2013**, *30*, 108.

(2) Sardar, D.; Pierce, E.; McIntosh, J. A.; Schmidt, E. W. *ACS Synth. Biol.* **2015**, *4*, 167.

(3) Oman, T. J.; van der Donk, W. A. *Nat. Chem. Biol.* **2010**, *6*, 9.

(4) Sardar, D.; Lin, Z.; Schmidt, E. W. *Chem. Biol.* **2015**, *22*, 907.

(5) van Heel, A. J.; Mu, D.; Montalban-Lopez, M.; Hendriks, D.; Kuipers, O. P. *ACS Synth. Biol.* **2013**, *2*, 397.

(6) Burkhart, B. J.; Kakkar, N.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A. *ACS Cent. Sci.* **2017**, *3*, 629.

(7) Goto, Y.; Suga, H. *Curr. Opin. Chem. Biol.* **2018**, *46*, 82.

(8) Hudson, G. A.; Mitchell, D. A. *Curr. Opin. Microbiol.* **2018**, *45*, 61.

(9) Walsh, C. T.; Malcolmson, S. J.; Young, T. S. *ACS Chem. Biol.* **2012**, *7*, 429.

(10) Zhang, Q.; Liu, W. *Nat. Prod. Rep.* **2013**, *30*, 218.

(11) Lin, Z.; He, Q.; Liu, W. *Curr. Opin. Biotechnol.* **2017**, *48*, 210.

(12) Just-Baringo, X.; Albericio, F.; Alvarez, M. *Angew. Chem., Int. Ed.* **2014**, *53*, 6602.

(13) Bowers, A. A.; Acker, M. G.; Koglin, A.; Walsh, C. T. *J. Am. Chem. Soc.* **2010**, *132*, 7519.

(14) Young, T. S.; Dorrestein, P. C.; Walsh, C. T. *Chem. Biol.* **2012**, *19*, 1600.

(15) Zhang, F.; Kelly, W. L. *ACS Chem. Biol.* **2015**, *10*, 998.

(16) Hudson, G. A.; Zhang, Z.; Tietz, J. I.; Mitchell, D. A.; van der Donk, W. A. *J. Am. Chem. Soc.* **2015**, *137*, 16012.

(17) Zhang, Z.; Hudson, G. A.; Mahanta, N.; Tietz, J. I.; van der Donk, W. A.; Mitchell, D. A. *J. Am. Chem. Soc.* **2016**, *138*, 15511.

(18) Wever, W. J.; Bogart, J. W.; Baccile, J. A.; Chan, A. N.; Schroeder, F. C.; Bowers, A. A. *J. Am. Chem. Soc.* **2015**, *137*, 3494.

(19) Ozaki, T.; Yamashita, K.; Goto, Y.; Shimomura, M.; Hayashi, S.; Asamizu, S.; Sugai, Y.; Ikeda, H.; Suga, H.; Onaka, H. *Nat. Commun.* **2017**, *8*, 14207.

(20) Goto, Y.; Ito, Y.; Kato, Y.; Tsunoda, S.; Suga, H. *Chem. Biol.* **2014**, *21*, 766.

(21) Koehnke, J.; Mann, G.; Bent, A. F.; Ludewig, H.; Shirran, S.; Botting, C.; Lebl, T.; Houssen, W. E.; Jaspars, M.; Naismith, J. H. *Nat. Chem. Biol.* **2015**, *11*, 558.

(22) Wever, W. J.; Bogart, J. W.; Bowers, A. A. *J. Am. Chem. Soc.* **2016**, *138*, 13461.

(23) Okeley, N. M.; Zhu, Y.; van der Donk, W. A. *Org. Lett.* **2000**, *2*, 3603.

(24) Seebeck, F. P.; Szostak, J. W. *J. Am. Chem. Soc.* **2006**, *128*, 7150.

(25) Goto, Y.; Katoh, T.; Suga, H. *Nat. Protoc.* **2011**, *6*, 779.

(26) Asahara, H.; Chong, S. *Nucleic Acids Res.* **2010**, *38*, No. 13 e141.

(27) Zhou, Y.; Asahara, H.; Gaucher, E. A.; Chong, S. *Nucleic Acids Res.* **2012**, *40*, 7932.

(28) Iwane, Y.; Hitomi, A.; Murakami, H.; Katoh, T.; Goto, Y.; Suga, H. *Nat. Chem.* **2016**, *8*, 317.

(29) Cogan, D. P.; Hudson, G. A.; Zhang, Z.; Pogorelov, T. V.; van der Donk, W. A.; Mitchell, D. A.; Nair, S. K. *Proc. Natl. Acad. Sci.* **2017**, *114*, 12928.

(30) Tran, H. L.; Lexa, K. W.; Julien, O.; Young, T. S.; Walsh, C. T.; Jacobson, M. P.; Wells, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 2541.

(31) Harms, J. M.; Wilson, D. N.; Schluenzen, F.; Connell, S. R.; Stachelhaus, T.; Zaborowska, Z.; Spahn, C. M. T.; Fucini, P. *Mol. Cell* **2008**, *30*, 26.

(32) Bewley, K. D.; Bennallack, P. R.; Burlingame, M. A.; Robison, R. A.; Griffitts, J. S.; Miller, S. M. *Proc. Natl. Acad. Sci.* **2016**, *113*, 12450.

(33) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. *J.Am. Chem. Soc.* **2018**, *140*, 9494.

(34) Hayashi, S.; Ozaki, T.; Asamizu, S.; Ikeda, H.; Omura, S.; Oku,N.; Igarashi, Y.; Tomoda, H.; Onaka, H. *Chem. Biol.* **2014**, *21*, 679.

(35) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. *BMC Bioinformatics* **2010**, *11*, 395.

(36) Goto, Y.; Katoh, T.; Suga, H. *Protoc. Ex.* **2011**, https://doi.org/10.1038/protex.2011.209 .

(37) Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H. *Nat. Methods* **2006**, *3*, 357.

(38) Mori, T.; Higashibayashi, S.; Goto, T.; Kohno, M.; Satouchi, Y.; Shinko, K.; Suzuki, K.; Suzuki, S.; Tohmiya, H.; Hashimoto, K.; Suzuki, K.; Suzuki, S.; Tohmiya, H.; Hashimoto, K.; Nakata, M. *Chem. Asian J.* **2008**, *3*, 984.

CHAPTER 4: BIOINFORMATIC EXPANSION OF THIOPEPTIDES AND PRIORITIZATION FOR CHEMOENZYMATIC DISCOVERY

4.1 Introduction.

Thiopeptides are active and structurally complex molecules. They are part of an ever-growing class of natural products called ribosomally synthesized and post-translationally modified peptides or RiPPs.[1] As the name suggests, the biosynthesis of RiPPs begins with the ribosomal translation of a small structural gene to make a precursor peptide. For thiopeptides, core biosynthesis contains three specific peptide transformations: 1) azole formation, 2) dehydration to dehydroalanine (Dha) or dehydrobutyrine (Dhb), and 3) macrocyclization via pyridine formation, figure 4.1. Thiopeptides are quite potent and have anticancer, antiplasmodial, immunosuppressive, renin inhibitory, RNA polymerase inhibitory, antifungal, proteasome inhibitory and stimulates mitophagy.[2,3] Because of their broad therapeutic potential, much effort has been given to understand how thiopeptides are made and how their biosynthesis might be engineered to make more active or more soluble analogs that might perform well in the clinic. Additionally, the target for thiopeptides seems to be dictated by the number of atoms within the core macrocycle.[4–8] It is expected, then, that the discovery of thiopeptides with new backbone topologies will reveal thiopeptides with novel activities. However, since the report on lactazole[9] only one new class of thiopeptide-related RiPP has been discovered, the pyritides.[10]

In addition to antibiosis, thiopeptides can act as signaling molecules. For instance, a broad group of thiopeptides have been measured to bind the MerR-like stress response transcription factor, TipA, in *Streptomyces lividans* and autogenously regulate the *ptipA* promoter and possibly

other genes.[11,12] Another report by Bleich et al, demonstrates that thiopeptides also broadly stimulate biofilm production in a *kinD*-independent manner in *Bacillus subtilis*.[13] While the molecular mechanism for biofilm stimulation is still unknown, it is highly probable to be regulated

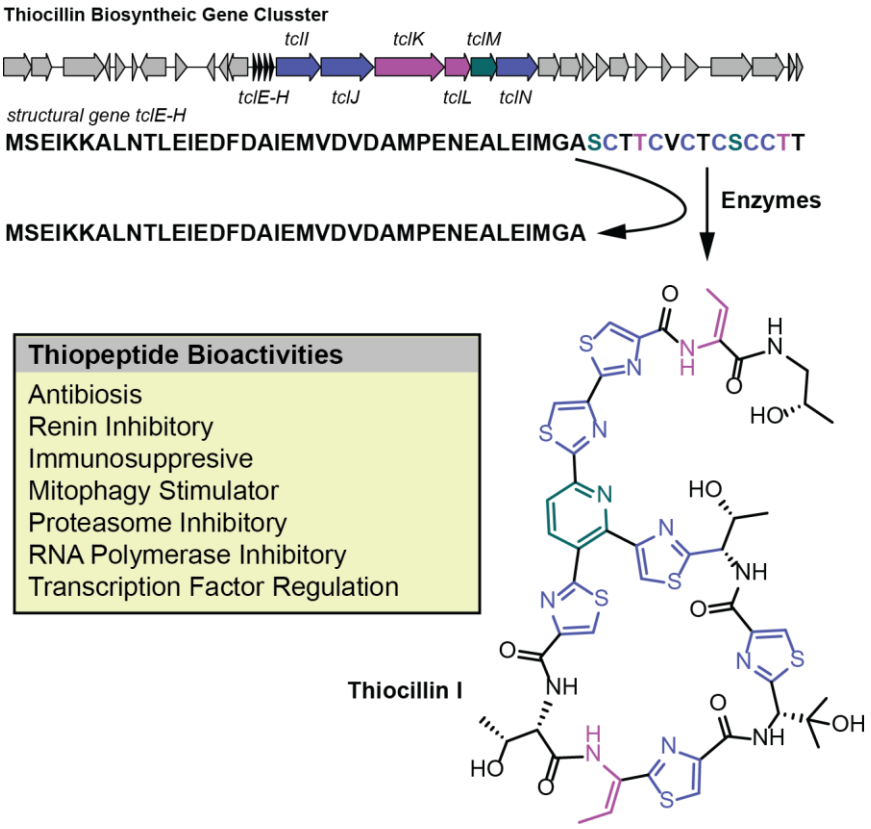


Figure 4.1. Thiopeptide biosynthetic gene cluster for thiocillin. Named enzymes are colored according to their corresponding chemical modifications of the thiocillin core scaffold.

by a TipA homolog. The pfam family for just the C-terminal thiopeptide binding domain of TipA (TipAS, PF07739) contains 2,634 sequences across 1,934 species. The most highly represented bacteria are *Actinobacteria* (705 species) and *Firmicutes* (706 species). Moreover, Habezettl et al. note that many *streptomyces* strains have a *tipA* gene but do not themselves make a thiopeptide.[14] This suggests that thiopeptide receptors may be more prolific than thiopeptide gene clusters, implying that thiopeptide producers may play a significant role in shaping bacterial communities through antibiotic activity[2] and transcription factor regulation.

New chemoenzymatic approaches are now available that might be used to structurally validate cryptic thiopeptides and avoid the headaches of heterologous expression or protein purification of many cognate biosynthetic enzymes.[10,15–19] Herein, we use the bioinformatic tool RODEO[20] to collect nucleotide sequences for all possible thiopeptide biosynthetic gene clusters currently available in GenBank and employ the RODEO thiopeptide scoring algorithm to help predict the sometimes-elusive structural genes. Using sequence similarity networks,[21] we show the current diversity of thiopeptide pyridine synthases and precusor peptides and with antiSMASH[22] visualize the concomitant biosynthetic gene clusters. We report on 14 currently uncharacterized thiopeptide BGCs that have very little similarity to already known thiopeptides and have been prioritized based on their likelihood of success in current chemoenzymatic approaches to elucidate thiopeptide structures. Furthermore, we use an SSN to visualize the diversity of TipAS homologs which further confirms that these proteins are prolific within *Actinobacteria* and *Firmicutes* and cluster based on phyla. It is likely that the different TipAs proteins cluster according to their respective thiopeptide ligand. We anticipate that our bioinformatic approach to organize cryptic thiopeptides in combination with current chemoenzymatic strategies to make thiopeptides will aid in the rapid discovery of new thiopeptides. Further studies will allow the interpretation of thiopeptide-TipAs pairings and how their interactions dictate interspecies communication within complex bacterial communities.

4.2. Diversity of the Publicly Available Pyridine Synthases from Thiopeptide Biosynthesis.

The central pyridine ring is the hallmark transformation of thiopeptides (and now pyritides), as it is not found in any other RiPP natural product. Because of this, bioinformatic studies aimed to discover new thiopeptide gene clusters use pyridine synthases to focus database searches towards thiopeptides. In 2018, one study reported that GenBank contained up to 508

unique thiopeptide pyridine synthases and organized as many thiopeptide BGCs using RODEO.[23] Therein, 9 different pyridine synthases were used in a BLAST search of the GenBank database and hits were subsequently filtered using Hidden Markov Models (HMMs) for four different thiopeptides: thiostrepton, thiomuracin, berninamycin, and lactazole. Since then, many more sequences have been added to GenBank, and I suspect that the stringent use of HMMs to filter their dataset based on highly studied thiopeptides may have greatly reduced the chances to discover new thiopeptide archetypes.

To collect any new thiopeptides that may have been added since 2018, and to encourage the collection of new thiopeptide archetypes, we collected the sequences of 24 thiopeptide pyridine synthases and applied them all to a BLAST search on GenBank without an HMM filter. All sequences with an E-value cutoff of 0.05 were collected and duplicates removed revealing 1110 predicted proteins; almost double the amount previously found. Accession numbers were submitted to RODEO with the fetch distance adjusted for 15,000 bp to gather all cooccurring biosynthetic and structural genes. RODEO is a RiPP specialized mining tool with a thiopeptide model that aids in the prediction and discovery of thiopeptide precursor peptides. One major advantage is that RODEO will also search unannotated genomic regions for thiopeptide structural genes that may be missing annotation in GenBank.

In a report by Cogan et al, pyridine synthases were found to be structurally related to Lan B C-terminal glutamyl elimination domains.[24] In fact, they both share the common ferrodoxin-like domain as found in the epimerase LsrG. Insertions and deletions within common scaffolds often lead to new activities. Indeed, pyridine synthases likely evolved divergently from Lan B C-terminal elimination domains with the addition of 3 minor insertions that were found in TbtD and PbtD to be involved in substrate recognition and likely processing. To see if these insertions are

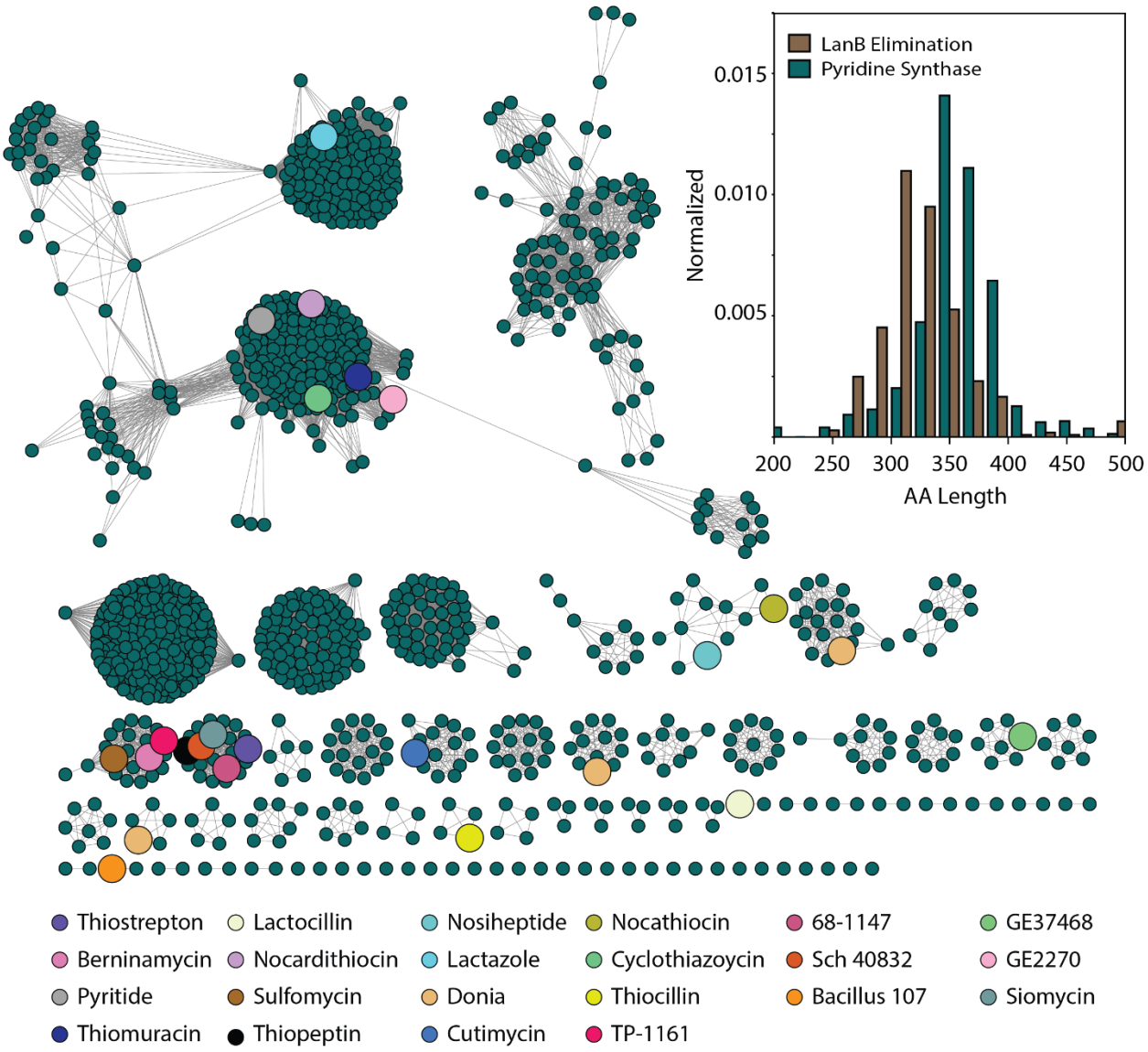


Figure 4.2. Sequence similarity network of pyridine synthases and analysis of sequence length. A group of 24 known pyridine synthases were BLASTed against the NCBI database to acquire 1110 unique putative pyridine synthases. These proteins were submitted to EFI-EST to create an SSN with an alignment score cutoff of 43. Pyridine synthase accessions were submitted to RODEO to obtain cooccurring genes. Sequence lengths of predicted pyridine synthases and neighboring split lanB elimination enzymes were measured and displayed as a histogram.

broadly found in our collection of pyridine synthases, the length of each predicted pyridine synthase was mapped out on a histogram alongside the length of RODEO collected single domain Lan B C-terminal glutamyl elimination domains, figure 4.2. Indeed, there is a clear increase in sequence length for PSs suggesting that the insertions found in TbtD and PbtD are likely common

among pyridine synthases. Additionally, this increases our confidence that the collected proteins are correctly annotated as pyridine synthases and not Lan B elimination domains, even though we did not stringently filter with HMMs. Ergo, the diversity of our 1110 pyridine synthases was visualized using a sequence similarity network, figure 4.2. Of the 35 clusters that were formed with 3 or greater sequences, only 10 contained a thiopeptide of known structure, suggesting that the yet unknown diversity of thiopeptides is much larger than the known.

4.3 Diversity of Thiopeptide Precursor Peptides.

RODEO returned 73,578 sequences which could be potential thiopeptide structural peptides. Of those, only 1674 had a thiopeptide score of 20 and were predicted by RODEO to be valid thiopeptides. But some expected sequences were missing, like the pyritide from *Micromonspora rosaria*.[10] So, for precursor peptides below a thiopeptide score of 20, the best scoring gene per biosynthetic cluster was kept. Additionally, precursors scoring between 5 and 19 were manually curated for thiopeptide-looking structural genes. Finally, any duplicates were removed to give a final list of 1,393 possible thiopeptide precursor peptides with 293 total sequences with an unvalidated thiopeptide score below 20. This was done on purpose to allow room for the discovery of new thiopeptides that may look very different than what RODEO is currently trained to find. The sequence diversity of precursor peptides collected were analyzed by sequence similarity network, figure 4.3. Due to concatenation 1,393 sequences were represented in 1085 nodes and 38 clusters with 3 or more sequences were analyzed. Only 11 clusters contain a known thiopeptide leaving 27 yet uncharacterized, further supporting that thiopeptide diversity is much larger than currently known.

Phylogenetic analysis of the species represented in this dataset shows that the recovered thiopeptides span 629 different species with the phyla *Actinobacteria* being very highly

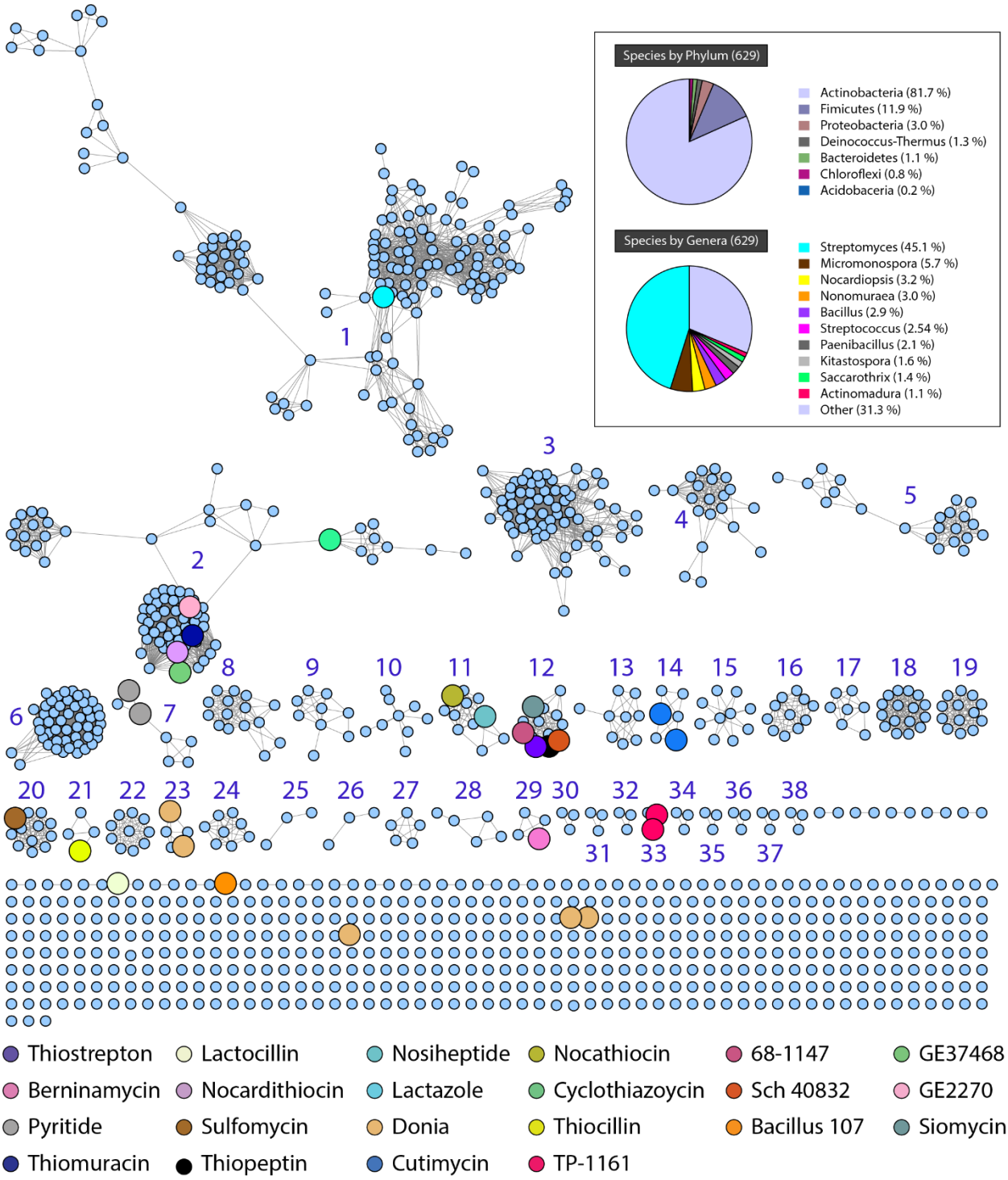


Figure 4.3. Sequence similarity network of predicted thiopeptide structural genes. Pie charts show the representative bacterial phyla and genera which contain thiopeptides. A manually curated dataset of predicted thiopeptide structural genes were collected using RODEO and a sequence similarity network was made using EFI-EST with an alignment score of 8.

Figure 4.4. WEBLOGOs of clustered thiopeptide cores. Clusters from figure 4.3 were numbered, and thiopeptide precursor sequences were extracted and aligned using MUSCLE. Alignments were uploaded to https://weblogo.berkeley.edu/logo.cgi and WEBLOGOs created. Leader peptide cleavage sites were manually determined.

represented at 81.7 %. It is therefore not surprising that almost half (45.1%) of the species are *Streptomyces*. The number of species reported is about half as many as thiopeptide precursors presented due to some species carrying multiple thiopeptide BGCs and some clusters harboring multiple predicted precursors.

Precursor peptides from each of the 38 clusters were extracted, aligned using MUSCLE, and trimmed with trimAL. For each alignment, WEBLOGOs were created and putative leader peptides, which contain only recognition sequences, were removed, figure 4.4. Average RODEO thiopeptide scores were calculated and only 4 of the 38 clusters scored below the "valid thiopeptide" threshold of 20 (Appendix table C.1**)**. These clusters were 7, 8, 31, and 35. Closer

analysis of the BGCs for 7 and 31 illuminates all the necessary traits for pyridine cyclization. In fact, a recent report characterized a member from group 7 as a pyridine cyclized peptide devoid of azoles. Because azoles are found in all other thiopeptides, group 7 RiPPs were newly termed, "Pyritides".[10] Further inspection of BGCs for cluster 8 and 35 suggest that they are not thiopeptides, but, in fact, linear azol(in)e containing peptides and lantibiotics, respectively. Each of these natural products have dehydrated amino acids as seen in thiopeptide biosynthesis and contain a Lan B C-terminal elimination domain misannotated as a pyridine synthase by our BLAST search. This highlights the difficulty of distinguishing Lan B elimination proteins from pyridine synthases. But, encouragingly, even with our more relaxed search for pyridine synthases, only 2 clusters out of 38 appear not to be thiopeptides.

It is clear from the SSN and the WEBLOGOS that structures and activities of many new thiopeptides remain to be elucidated. A goal of this study is to prioritize cryptic thiopeptide BGCs for analysis using chemoenzymatic synthesis. Current chemoenzymatic approaches to make thiopeptides favor core peptides with limited number of Ser, Thr, and Cys content. These amino acids, exclusively, are modified to dehydrated amino acids, azol(in)es, or a pyridine to make the thiopeptide backbone.[2] By inspection of putative core sequences, modifications can be predicted, but determining the type of modification must be experimentally validated via successful cyclization by the cognate pyridine synthase. Therefore, the fewer the Ser, Thr, and Cys that are present in a core the fewer the pyridine synthase substrates that need to be made and tested for cyclization. Initial inspection of the WEBLOGOs shows that many clusters would not be good candidates for chemoenzymatic discovery. For instance, clusters 5, 6, 10, 13, 15, 19, 24, 27, 34, and 37 all average > 20 CST residues per RODEO predicted core. Alternatively, some clusters, such as 1, 2, 3, 25, 30, 31, and 36, look ripe for chemoenzymatic discovery having small cores and

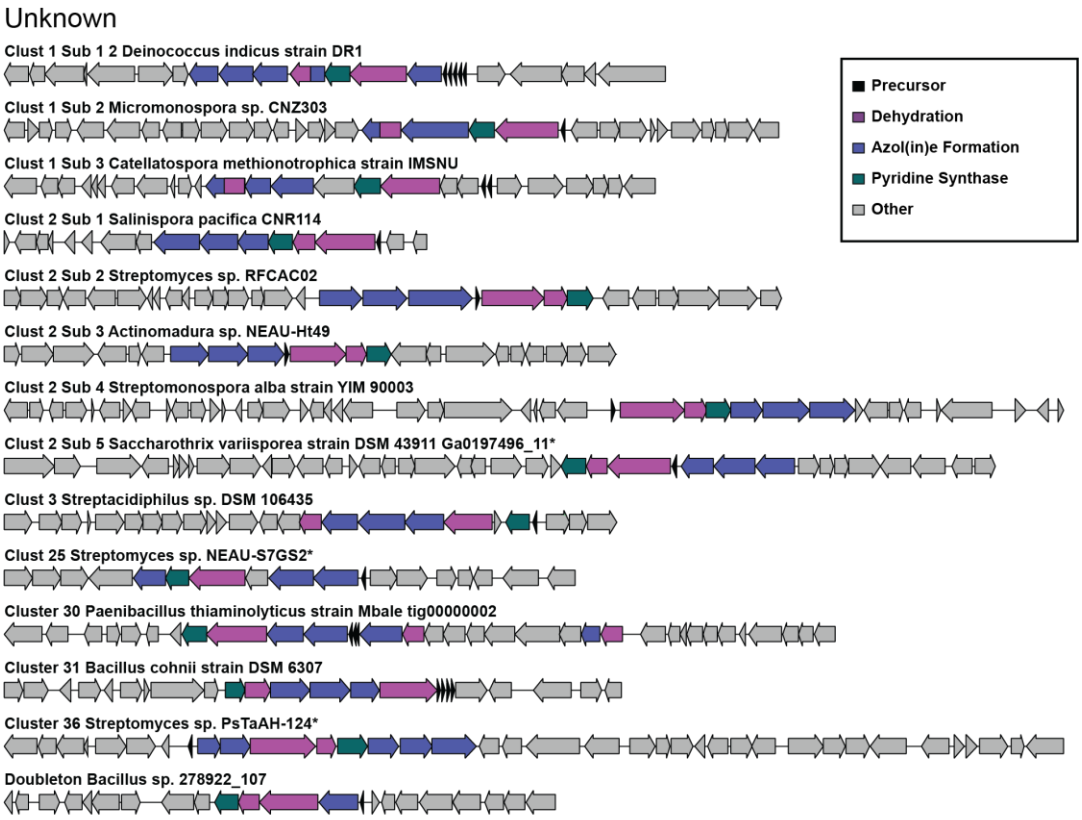
Thiopeptide	Leader	Core	RODEO Score
Thiocillin	MSEIKKALNTLEIEDFDAIEMVDVDAMPENEALEIMGA	SCTTCVCTCSCCTT	30
Berninamycin	MEQQIELDVLEISDLIAGAGENDDLAQVMAA	SCTTTSVSTSSSSSSS	33
Nosiheptide	MDAAHLSDLDIDALEISEFLDESRLEDSEVVAKVMSA	SCTTCECCCSCSS	37
Thiostrepton	MSNAALEIGVEGLTGLDVDTLEISDYMDETLLDGEDLTVTM	IASASCTTCICTCSCSS	33
Thiomuracin	MDVFELADSGVAVESLTAGHGMTEVGA	SCNCFCYICCSCSSA	27
Cyclothiazomycin	MEKELVLDLADLSVDELDVLPTSPGAGLESINVGHAMVEIGA	SNCTSTGTPASCCSCCCC	30
Lactazole	MSDITASRVESLDLQDLDLSELTVTSLRDTVALPENGA	SWGSCSCQASSSCAQPQDM	23
Pyritide	MDNVVTEAAEFADLDIVDLDLAVDEELAALSVGGLGNTEVGA	SGWLGSWVI	6
	MDNAATEATEFADLDIVNLDLPIDEELAAVSVGGLGNTEVGA	SGFFGRSWLI	6

Figure 4.5. Known thiopeptide biosynthetic gene clusters. Enzymes which make the thiopeptide core scaffold are colored by their concomitant transformation. Sidechain related enzymes are colored in light blue. Structural peptides for each BGC are bifurcated to highlight the 'leader' and 'core' sequences. For each core, pyridine cyclization is indicated by a bar connecting two serine residues.

fewer CST residues (Appendix table C.1).

4.4 Prioritization of Unknown Thiopeptide Biosynthetic Gene Clusters for Chemoenzymatic Discovery.

Known thiopeptides have been extensively studied over the past decades and several biosynthetic gene clusters have been discovered and validated, figure 4.5.[9,10,25–31] Thiocillin is a classic and well-studied thiopeptide.[13,25,32–38] While the individual enzymes have not been purified



Thiopeptide	Leader	Core	RODEO Score
Cluster 1 Sub 1	MNEPTTTAIPEALELNDLHLEDVSVAVEEDAYDLPEAGA	SLSPRFSCSIVIDLS	22
	MNEQTTTATAEVIELNDLHLEDVSVAVEEDAYALPEAGA	SVSPKYSCTVVIER	10
	MNEQNTTATSEVIELNDLHLEDVSVAVEEDAYALPEAGA	SASPKYSCTVVIKPN	10
	MNEPTTTAIPEVIELNDLHLEDVSVAVEEDAYDLPEAGA	SASPRFSCSIIIEL	9
	MNEPTTTAIPEALELNDLHLEDVSVAVEEDAYDLPEAGA	SLSPRFSCSIIIARE	8
Cluster 1 Sub 2	MTATLADLEDLNLADLEVSDLLEVHAVRDAVALPETGA	SASTSAVWRWLGYSSCAAQLPQ	21
Cluster 1 Sub 3	MADTSTTWDLIPEYDIAFDMDDLDLGPLTVTAMRDSMALPETGA	SWINTSGNASCSCCYVT	30
	MSNEDFNVDELELDDLSVTSIRDSAALPETGA	SSGSSSSSASSCCGSSSCCSSCGDNPEVAA	33
Cluster 2 Sub 1	MNAKDHTAFDSAGVEFDVDDLPLSVFELTDGGLTVESLTAGHGLVENGA	SWPSCGSSCSSTP	26
Cluster 2 Sub 2	MSDHEFTLDLDGLDVDTIDVLPADSVMTEGHGAIETGA	SCSPGYCSCYSTAMSCFVPEEI	11
Cluster 2 Sub 3	MSKDFADLSDLPVDSIDVLPADGIEALDIGHGMTEVSA	SCGHEFTRPCGSCGAPHDGELDEF	15
Cluster 2 Sub 4	MAGSKRNGKVSEQLDFDLSSVPVDVFDLADQGLTLETLTEGHGLPENGA	SSICNELMCACSSPAC	25
Cluster 2 Sub 5	MTASKPALSLDDLAVESIGVVPAVGPEDLTQGHGLTELGA	SCEDCDRMGGGSCDCCECDDLYRVIE	19
Cluster 3	MSIDFGAFDVNELEVLDVTDSVALPDMGA	SAIIIFPTVTPDGEEILGEADLGISGSLSTSCC	11
Cluster 25	MENTLLDLEQLGDEMFEVVGSESAQGEGDAN	SAWAYACCQSASSSCALS	16
Cluster 30	MKEEEIRIMLDDLELDQVEIRGLIVEDGKAIPDMGA	TIGAISCCSCCST	31
	MEELKLENQEIELSLDDLELQDVKMEVLMTNSKGLPEMGA	TLGGITCCSCCCCST	36
	LHLDDLELDDVNIIALTQDDAKGLPDMGA	TMMWVGCCSCCST	31
Cluster 31	MEIEKVMDFELEIFEVEEATVMEEMGA	SIFIIACCSSIEIKPY	20
	MAQLTNVEEVELEIFEVEEATVMEEMGA	SIGKLGCCSSIIIR	19
	MNMEQNMEEVELEIFEVEEATVMEEMGA	SIGKLGCCSSIIIR	19
	MNVEQSMEEVELEIFEVEEATVMEEMGA	SIGKLGCCSSIIIR	19
Cluster 36	MNEFSNGTPSLADAVADLALDLDLGSLAEDEHSQSLTAGHGMTEVSA	SFCCAPPPNCCNCSCS	25
Doubleton	MELKLDFNVLENLEENISFDEVKALDIEGSIGATEGA	SSAGSSLPFPPYFWFTCSA	12

Figure 4.6. Unknown thiopeptide biosynthetic gene clusters. Enzymes which make the thiopeptide core scaffold are colored by their concomitant transformation. RODEO predicted structural peptides for each BGC are bifurcated to highlight potential ‘leader’ and ‘core’ sequences. ‘*’ indicates that the full antiSMASH predicted BGC has been trimmed at the edges.

and shown to work *in vitro*, because of solubility issues, the biosynthetic gene cluster has been heterologously expressed and libraries of many different thiocillin analogs have been made. Additionally, using only the pyridine synthase, TclM, various thiocillin analogs have been produced in several reports using chemoenzymatic synthesis.[15,16,19] To make the core macrocyclic ring of thiocillin, three proteins are needed to install azoles, two enzymes to install dehydroalanines or dehydrobutyrines, and a single pyridine synthase cyclizes the peptide introducing the hallmark pyridine. This core biosynthetic structure is highly shared by other thiopeptides, including Nosiheptide, Thiomuracin, and Cyclothiazomycin. Slight deviations to this archetype are found in the number of enzymes present for these core transformations. For instance, in Thiostrepton and Berninamycin BGCs, an extra enzyme is indicated for dehydratase activity and azol(in)e formation, respectively. Interestingly, lactazole biosynthesis has only five total proteins as a single enzyme contains domains for both dehydratase and dehydrogenase activity, allowing this enzyme to participate in Dha and azole installation.[39,40] Lastly, pyritides do not contain any enzymes for azol(in)e biosynthesis and represent the shortest BGCs to make pyridine cyclized peptides.[10]

Looking though our dataset of putative thiopeptides, fourteen different precursors were prioritized and selected based on short cores, limited CST residues, increased promiscuity indicated by multiple precursor peptides, and dissimilarities from known thiopeptides, figure 4.6. Interestingly, the BGCs for most of these unknown thiopeptides contain the same protein architecture as noted for thiocillin, except for gene order. Almost, exclusively, deviations occur with the number of azol(in)e forming enzymes as seen with Berninamycin. Of note the gene cluster from Bacillus 107 contains only 1 and is the only BGC missing a predicted dehydrogenase, suggesting the formation of azolines and not azoles. As observed in lactazole biosynthesis, all cluster 1 selected BGCs contain a split enzyme predicted to participate in dehydrogenation and

dehydration. Cluster 1 Sub 1, 2 Sub 3, and 31 all contain multiple different thiopeptide cores. This is reminiscent of certain lantibiotics, like the prochlorosins, which have a heightened promiscuity at a kinetic cost.[41,42] It may also be suggestive of a two-component system as seen for lantibiotics.[43–45] Known thiopeptides are highly hydrophobic leading to poor solubility in water.[2] Notably, predicted thiopeptides cluster 2 sub 3 and 2 sub 5 contain 8 and 10 charged residues, respectively, likely allowing these thiopeptides to be very soluble. YcaO cyclodehydratases do not readily modify Cys amino acids on the C-terminus of substrates.[46] We see this in cyclothiazomycin biosynthesis where the C-terminal Cys is not cyclodehydrated, but instead is used for external cyclization giving cyclothiazomycin its secondary ring.[29] Similarly, clusters 2 sub 4 and 3 also end in a C-terminal Cys. It is likely, then, that these cysteines are not transformed into azoles and are, in fact, used for an ancillary modification.

Presenting and organizing these unknown thiopeptide gene clusters and precursor peptides enables easy design and application of chemoenzymatic synthesis for thiopeptide discovery. We expect that several new thiopeptides can be structurally validated in this manner, and new bioactivities found. One unknown that we have selected from *Bacillus sp. 107* is of particular interest as a report by Bleich et al., shows that this cryptic thiopeptide may stimulate biofilm production in *Bacillus subtilis*.[13] This activity was measured for several different thiopeptides and is hypothesized to be regulated by a TipA-like transcription factor; a protein that broadly binds thiopeptides and regulates its multi drug resistant gene expression.[14]

4.5 Diversity of Thiopeptide Binding Receptors.

Several studies now indicate that Thiostrepton and a broad range of other thiopeptides can regulate gene expression via the transcription factor TipA; a MerR-like stress response regulator.[12,14,47,48] These type of transcription factors are activated by cysteine modification. For

instance, MerR is activated by a Hg-complex which forms between 3 different cysteines.[49] In a similar fashion, thiopeptides contain thiol-reactive dehydroalanines and dehydrobutyrines that are covalently attacked by a C-terminal cysteine upon interaction with TipA. Consequent structural changes in the protein subsequently activate TipA to autogenously regulate its own transcription and possibly other genes.[14,47]

Interestingly, another report shows that thiopeptides can broadly stimulate biofilm production in *Bacillus subtilus* in an unknown, *kinD*-independent manner.[13] The structural similarities between the thiopeptides that induce *tipA* expression in *S. lividans* and biofilm production in *B. subtilis* suggests that thiopeptide sensing and gene regulation via MerR-like transcription factors might be a broad phenomenon among *Actinobacteria* and *Firmicutes*.[14]

To gain an understanding of just how widespread thiopeptide signaling via TipA may be, TipAS (the thiopeptide binding domain of TipA) was used to focus a BLAST search on NCBI and 2,311 non-redundant thiopeptide receptors were recovered. By submitting these sequences to EFI-EST an SSN was created and due to the large number of edges, proteins were concatenated at a sequence similarity of 75 %, figure 4.7. As seen with the thiopeptides, the most abundant phyla are, indeed, *Actinobacteria* and *Firmicutes* and clustering is highly phyla specific. TipAS was shown to bind a wide range of structurally distinct thiopeptides that all contain a conserved four-ring motif.[14] The discrete clustering within phyla suggests that there may exist many thiopeptide specific receptor pairings. It has also been suggested that bacterial strains that produce thiopeptides do not themselves have a tipA gene.[48] Here, it was found that 156 of the 629 unique bacteria that

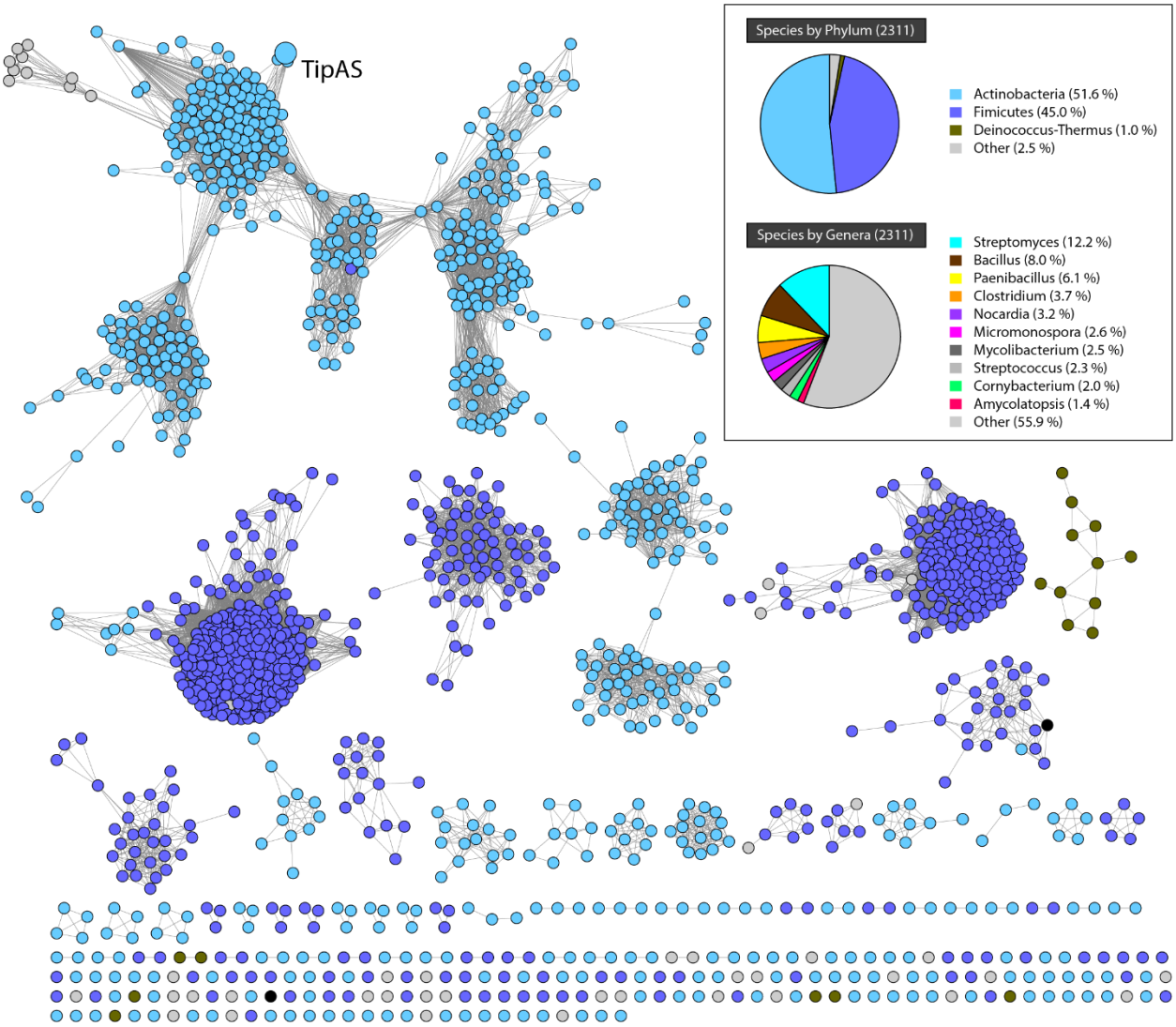


Figure 4.7. Sequence similarity network of thiopeptide receptors. A group of 2311 thiopeptide receptors were collected from NCBI using BLASTP. Pie charts show the representative bacterial phyla and genera which contain these receptors. Sequences were submitted to EFI-EST to make a sequence similarity network with an alignment score of 80 and sequences of 75% identity were grouped within the same node. Nodes are colored by phyla. Blue, purple, light brown, black, and grey are Actinobacteria, Firmicutes, Deinococcus-Thermus, Actinobacteria and Firmicutes, and other respectively. TipAS sequences were collected by Sarah Barr.

produce thiopeptides also contain a TipAS homolog but only 19 times were these TipAS genes found near a thiopeptide cluster, potentially serving as a resistance mediator. Therefore, 75 % of the bacteria which make a thiopeptide do not themselves have a thiopeptide receptor. This means that thiopeptides are likely very important interspecies metabolites that shape bacterial

communities acting both as an antibiotic and a signaling molecule. For strains that express a TipAS homolog and a thiopeptide, it will be interesting to determine if the two interact, or if the TipAS senses a disparate thiopeptide produced by bacterial neighbors. With the aid of chemoenzymatic synthesis, we expect that many new thiopeptides can be made and their associated TipAS homolog revealed, leading to a better understanding of how thiopeptides shape bacterial communities.

4.6 Conclusion.

Here, we have bioinformatically expanded the natural product class of thiopeptides. Previous attempts to do so contain ~half the number of predicted pyridine synthases, and do not organize and prioritize new thiopeptide biosynthetic gene clusters for discovery. In our assessment, we employed a somewhat more relaxed method for collecting pyridine synthases, i.e. without the use of HMMs, but we thought that the decreased stringency would better allow for the discovery of new thiopeptides. While this would also give for more negative hits, we expected that these might be filtered out as singletons during SSN creation and true thiopeptides would cluster by leader peptides as previously seen.[23] In fact, this is what we found as only 2 of the 38 clusters presented are suspected not to be thiopeptides. Furthermore, our analysis now includes at least 10 more clusters of precursor peptides with three or more nodes, including the previously missed pyritides. Of the 38 thiopeptide clusters analyzed in this study, only 11 contain a thiopeptide of known structure. It is clear from these data that the diversity of thiopeptides is much broader than currently known and with increasing genomes being added to NCBI, we expect this family of natural products to continue to grow and certain clusters to become more prolific. We also found that most thiopeptide BGCs, known and unknown, follow a similar core architecture to the well-known thiopeptide thiocillin, and that deviations are largely found in the number of azole producing enzymes. By presenting the biosynthetic gene clusters of disparate and unknown

thiopeptides, we empower the natural product community with the necessary information to unveil the structures of these compounds using chemoenzymatic synthesis or other methods.

As new, structurally diverse thiopeptides are discovered, we anticipate a commensurate revelation in TipAS guided activities. Currently, we know of thiopeptides with dual activity, like Thiostrepton; an antibiotic and TipA regulator. But thiopeptides like lactazole exist that have no antibiotic activity and, in fact, have no known activity. We hypothesize that such thiopeptides may act solely through TipA signaling to help shape interspecies bacterial communities. The clustering of TipAS homologs seen in figure 4.7 may be indicative of different thiopeptide binding partners. With the advent of new chemoenzymatic tools to illuminate cryptic thiopeptides, we anticipate that new TipAS-Thiopeptide interactions will be discovered and new thiopeptide regulated bioactivates unveiled.

4.7 Experimental

4.7.1 Collecting pyridine synthases from GenBank and RODEO.

Protein sequences for 24 different pyridine synthases were collected and BLASTed against the NCBI GenBank database with an E-value cutoff of 0.05. The search recovered 4,930 proteins, 1,110 of which were unique. Accessions for each unique predicted pyridine synthase were organized and submitted to RODEO to obtain neighboring proteins and structural genes. The RODEO module to predict thiopeptide precursor peptides was engaged, the fetch distance for nucleotide sequences was adjusted to 15,000 bp, and precursor peptides were only included if the amino acid length was between 38 and 100.

BAU84768.1 tsrE protein [Streptomyces laurentii], Thiostrepton

MSVMSTSEQKDLTVSVPWSVQEDLLLDVAAPLLDESVELGETDSWFYLRENHGGRPFL
RLRFASRSPSVERRLKSRILAHVGPTIDAGDVFTYQPYNHEHDWLGGTAGLGLAENFWT
ETTPLALDTLRATRGNRALRLAVAFDFLVCTGVMLAPHLPPSIAKFGYKAGYLSYLATF
EGYMLLIRDPEGTRAKHAQRYEKNRELLRPRLRTLVEQMSEPDGELTDVPELAREWLV

RLRDYVPALQKGFDEGRFYLYATPRKAETAKLTPSPDGLYRRPDVEWLSDLPEPPVAGI
HRAIADNTYYQGMIREDRRFLASRLAQAYTNWHLYRLGFLLADRYTLFYLIARAFEEEY
DLDAAALIRSVRPEAEVAG

WP_110629793.1 hypothetical protein [Streptomyces actuosus], Nosiheptide

MTSGPGQAPAEAAHAAGAAWLEIGLDAPADAVPALVAGVVRPLLREPAEPGAEPVPGF
FLRGVGAAQPALVVQLEVTPGTDLAEPYAARARALAAGLGLPVQVAAGRATLVPLAGS
VFAGAALGPVTRAALAAVCPALLTATEAAEQGRPALLASAAELMSAHLRAVSVSAAPG
PRQWEELREGVPLGFLSYRSHAEAFLASSRDPKAAQAMMDAKYTRAAATLERLVDGV
LTQCEERGPVVSLPARQWYEAMRAAKPAVTELFRAGTDLALDTEEQPPDTGPDGKGLS
ESAFHRIVEGSDGLRDFLDRDPSFLATRLLTSLLYLSLSSVGIALAERYFLCYAVSRACESI
FDTDALTVLSGLARTSLAS

ACS50127.1 unknown [Streptomyces hygroscopicus], Cylothiazomycin

MAPWAEAAWFGRHWLRGPHLRLNFRCRGTDWEERVRPTVTGIVTDYLRARPSAARLD
EGALAPVHARLAELEMETGPRHPWVPDNTVLERPYDHRLPVLESLRASELLAGFLSDTN
GQAFRAYERVRAGGALSLLALDLMWTTTSVAAVPFTTGGEPIERGFLSLRSHADAFLSR
TRDPVAVRAAFDDRFRRQETVLCERLRSVEAALSDGATEGDGADRSDRSDRSEAVGDV
VPFVTEWAAAVRHHQRIAHPLLASGEVSMGGAARAPRMPTRRTSEFHAVLRSDHGHED
FVRTDDWFASFRLMMNYLYLHLNRLGLKPVDRALLCHLAALTVEAVHGVDAVGSFQR
YVASVDPSSERPEWRRIGEAWAAGGAG

AGN11669.1 BerD [Streptomyces bernensis], Berninamycin

MSYGRLQDHVTARLAPAEISGVSFVHLFATIPQPVGSKYNDTFAPLIRELFAPERVGGAG
GHGPYYFVRTQDAQLGTDTLQISIEGVSDEDSTRADLHRTAERYGCAAQVDATPLDSVP
SPLWNAGFTGTGFSASSKRLFQEAAPTLVSFLNRAAETPQSPPPALGAIRLMAAHTRATL
LRSPQREIDGYEFRELLSLRLLSYRSHFEAIYLRTKDPQSFDAACARFYEQVGAGVREFIT
ACGDPDDDPADEMVRLWTKSITSESSHLAENFSDGSVVNAGHTLEDLVRKRGAPVEPT
RFHTPPSPELDRLMHRDADFLAFRLQTSLLYSCLYTLGFSLAERYVFCYVVARANEDVC
GKSMKELQDELDGLARSMASGSTKTAE

BAO57436.1 lantibiotic-type dehydratase [Streptomyces lactacystinaeus], Lactazole

MSDPADGRGAVTAWDVVLYHYRPDKARALREAVLPLARQAAAEGLAAHVERHWRFG
PHLRLRLRGPEARVAGAAQRAAEALRAWAAAHPSVADRSDEQLLAEAAVAGRAELIAP
PYAPLVPDNTVVAAPADRSAEDALRALIGAESAELREELLRTGLPALDSACHFLGAHGD
TPQARVQLVVTALAAHATAHPDGLVGAHYSVLSHLEDFLVHEDPDGSLRAAFERRWEQ
SGRAVTALVGRIADGGARDWERDWAHWSATAWSLAERRLTAGADLGGRHAEYRERA
EALGDPATAERWNAELRTRYSEFHRMLQRADPDGRMWHRPDYLINRAGTNGLYRLLAI
CDVRPMERYLAAHLLVRSVPELTGHRWQTLLGAAEQPGGPEQSGAAGATGGAGRTKL
EGAA

WP_000438834.1 hypothetical protein [Bacillus cereus], Thiocillin

MEQYHKIVLTGSNAETMLIKNIEPVAVKFINNYKGFFYVFKYSKDFPIIDVYINNKIVTEN
QLNKILQNSGAKYKIYKNSIFNETQGNFGMLGDKYLAEFFKKTNEISLNILNQNFESYNK
KIEFALEIMLISAHYNYDSIKKGYLSYASHVNGFFTRWKDPNKIRDIFHKNYLNNKEYLE
SKVSEIIDNNNRSSLSELSDIITEMKKEMTTDIEKGNLHVFNIELLQKPGERDFLEKSQFHK
TILNNPDFSNFMNKDINFLGSRLITVFTYLLIRNLGIQNKDRYLLCYYIYKIIEEKYNIDTLE
LIRDFGKGRDNNVEDLQRY

WP_028414789.1 hypothetical protein [Bacillus sp. 278922_107], Bacillus_107

MSFKAINIYLHDSSNHNLFVEEQIMPAIEKYDTLHTFKCFLLRHWRKGPHLQLHIELNKY
FSEKDLISLKEKLEKNIHSYLSDVAFSESDLEKNYEFLKIYELETNPFFPLHEDGFIEISTSV
LRNDIWGDTGVEAIRNYYCESNNLVLNSMTKIEKEKKYQNAISIMLASVKAIGDESTQQ
LSFRSHAEAFLNRFDTEDKIRKAFEHQYQSNKKNLSILFQDILTKDSVHNEWVKIYKDLI
DTNKQLIEDNIINLPKAETYLSIIEENKWVEGIEGELSPFHQYRNSLPHSETYRNTSDFYVK
RLVLNFMYLTLVQIGVKPLEKFSMCYSIARWFEDKLNKDWKKQLEEIVKI

WP_013130812.1 hypothetical protein [Thermobispora bispora], Thiomuracin

MAAGERWWRFRVDYHAGPMDDLILDGVRPAFAAFAAQAPMAYFLRHWRRGPHLRIY
VSTTREALEAVVRPAIEHVVGGYLRARPSPGMADPSAFLPLHERLAELEGEDGPLMPWS
PDNTIHAEGERPEPLTVRDVLLADFYADTTPSVYHALERVRSGASLPTIAFDLVVATAHA
LSTGGLPVARTSLRSHAEAYLARRSDGVRLRELWRDHYARNREAFTERLIAVASSAESA
ENGAHLPHVREWVRRLRPIRERARALLESGELTLEYASPAEGARDLPSLAEVSAFHRELE
SRPEWARLRDSPAFGAYRLVINCTYLHLTRLGLTPHQRFLVCHLAADAAADVYGIAAH
EEVATR

WP_067368384.1 hypothetical protein [Micromonospora rosaria], Pyratide

MTRTPPHREEPGWHSIQIRYHAENKDDLILDAVHPLLTALADAVDQPHLLRHWRRGPHL
RINLRTTPQLWAEVVRPTAQRDLTAYLHDNPSTASLREAEHLAAHRRLAVREADHGPLT
PWYPDNTVQFEPHEDRQHVLGSPALAGLLADGYARSTALTVDTLAGVRSGAVDRVGH
ALDLMFAFGHLSVPPISRGYMSFRSHVESFLGYTADPDAVRSTFDERYHRHRPALRDRL
HTTRAIVAGSRRDPLVEAWLEIVRRQKALAEPLFTSGAIDLDHLDQRGMELRHRIEFHDI
LRETTAHRREVLSQVWFRCHRLAINQLYSHLSRIGIVPTQRYLLCHLVARTVEEEYDISP
VALARQFVNDRR

WP_010777475.1 hypothetical protein [Enterococcus faecalis], Donia

MEKIHRITVALNTQNIFKQKFINEELIPQMKEFRKQNHNNKCFFHRTLEIGPVVDIYIQGL
KEEAVLMQVAIEKRLAKFRKMYSIQFSEKKIYFENQESIRRMNGFSKKAGRYQNYTVT
MDELEEVPREGEYYSEEAKKMVNSWLFEQGALIDKTTLLVEKLSELEKQKFLVSLFLTC
SKKLDGTSYRGYLSFKSHYLGFINIRRKEMENYHNMFEQYYLEHEEEFELMRERVREGQ
SFLTGHEQADELLHEWQNVIDQFMEGQKRLLKKIKLTNIFSMLGLRRYSEFHKESFKLSN
LKFYFSEQFQEYRLLVNVMYLLLPDFGFNTPKRLLASYNLISVLEGEER

WP_003649851.1 hypothetical protein [Lactobacillus paragasseri], Lactocillin

MKVYKIINGTNAKRNKFNSVSFYTYSRESNNLWCYVKSKQGTEIKLIDKKMGNVAEEII
EQFFISMGKQTIKISECSNFYNQIILLMISFLDINYNGEIFRGGQSFCSHANGFITFSSDPKM
AKQRLEQYYLKNKDILINIVNLYCKGKIGKFEKNNMKNIFVSLDEEVKSSIRDNKIYFINY
SQNNLLKKSNFHVRFYEKHKSLFSNEQFKKERFMTICFYQYLYFCLKINYKTRSELDYLI
YRALEDFYNKKYLNIVNR

WP_002518053.1 hypothetical protein [Cutibacterium acnes], Cutamycin

MNPKNSKVTELWFPDNSDSPCTLLLDLSRASAEGTEPSPQYYCRTWDPTTKDPGVRVGI
LDPNRLTHQTAFLGGSGITHHFLVMEKPPSWPLLWNAGFAGLKLSTISKQLFTASYPWM
VKWALQASAWTDDERLLALTKLLVAHADASLPPVDKTEMLGRLPLKLSLRLITYRSHY
ESIRARTIDPKAMDTAVNRAFMKYQDELLRWSQSLTCSSSYREKEWSNPNLSEWSAIIY
CHLPQIRQLLSGGGLTVAGPTREDIEASLGHSIPDTRFHERLDPAMQYYLNHDQDFTAYR
FCTGFLYSNLYSLGIPLRLRYLVCGLAARFNEELSDTTWRDLRDLFHEAATVAYYTRDT
DGSGIR

WP_002963036.1 hypothetical protein [Streptococcus sobrinus], donia

MKELYIIFESYEDLFRVQQRYFLSNFINQGMILFSKSSTKKSLTFVSEDCREFDTLLGINRQ
CTRVDISDFNSKIYFPYFLDTDFFVKNYKLFFQGVVSLIQESDYWDLDTEHKRYLIEELLC
TVADQHADGVSHGYLSFYSNYLYYLSQLRAIADKKSYQKIKKRIEFVSDLDRGHFKEEL
VTFPKLSKNLGMVNKELVKNVEKLDLRQLPSPYDFFKNSKVHLEYSEFHKNVFSNPLLL
KRYSDIHFVSYRIIMGFFFKVLPLLGISLNERNHILYLFVRHVEEYFNVDWKKQINESIKW
EECNVNPKR

WP_002998788.1 hypothetical protein [Streptococcus downei], donia

MEYKLSVFINNLAFSNMLISKIYSLLNEKSMSNVWIKKSIQRGHHIELYSQNLADLKELES
FLAGALENYTPVNIEESKLKKQIMTVSSAENTEAKLDIQPDGTLLLEEIEAAAEESSLSSL
ETELAVEKKKMELILYLEKIGFFNKSESQQNILLTKLFLNLGLLFNNNLKMGYLSMKSNI
VFFNLQLKNLKDTMSQENYERYYQLVNSLSSDEQKFIHEKIETFLAENEESYFKIFVEKL
YQFFTDELKKGHLYYQNMSDAESFLQQARERGQLTEFHKNFFTNQEFLSQHSSLPFVLY
RYTVDVLYQVMPLLNVNPLRKQKITKMVCDKVETGYGITWEDSYAMMQKLFKKKV

WP_110565432.1 hypothetical protein [Micromonospora arborensis], 68-1147

MTADRFGDVTVNVPWSSQEDLLLDVAAPTLDESVERGETESWFYLRETKEAKSYLRLK
FRTQSPSVEARLRARISEAVGEAGRENLFEYRPYNKEHDWLGGAGALGLAESFWTETTP
LALESLRATRGNRALRMALALDMLACSGVFLASRLPESVSRFGYRAGYLSYLATFEGYL
LLIRDPEAARARHARRYDLNRDALRPRIRRLVEIMQDPDGDLDGLPEATLRYMNSLRSH
LPPLQQGFDEGRFYLYATPRKDDFNKLTPAPDGLYRRPDLPWLADAPEAPVAGIHRAIA
DNPYYQGMIREDRRFQASRVAQAYGHWHLFRLGFLLSDRYTLCYLVARAFEEEFGLDA
AELIRSVKPEAEVS

ACS83769.1 hypothetical protein [Nonomuraea sp. WU8817], GE2270

MSWRRVDVAYHDPDLDGLILATRPLLAGTPGRGWFQRHWVRGPHLELWFDAAQPSWE
RIRDVLEPWLRVNPSRARIDRDRLLAQHRHLAAAERIDEPLLPFYADNTLHRAAPRSRA
HVLGGPAAEELFHDFHTTASAVAYDELDAVRAGESRLVMALDLMVAAAHAHAEGGV
RGGFVSFRSHAEAFLASAPGLRERWDAEYAARAGALRARITAVVAGIPRGRAWAGLLD
RFADRGDELIASGALLVEPAGPDAVARPDTAFHRALRGNRTWHEEVLRSAPFRRYRLLL
NLTYLQLSRLGVNAVQRALLCHFAASAVEQEYGVSAIEIAMGGA

AEM00619.1 putative aza-diels-alderase protein [Streptomyces sp. ATCC 55365], GE37468

MADRAGVELRQDWLADNSLTWTTGALAATGTAETTLQSLLEDFHYAATVPALRLLSA
APGERLGLACADLMAVTAQEFGRGGLVSAALSFRSHAEAYLNLEAAPDERAAWDAAA
RASAPALRRRLLAVATGPDRPAYARDWLGLITPLVRAAEQAQRRGELALPTLAEGFSSD
LTERSAFHRGLAGSTSWEDVRTSDWFVLYRFAINLLYLQLSRLGVKPVGRYRLCHLVAT
ALDQREAPTTTAEEGDS

WP_147287926.1 hypothetical protein [Nocardia pseudobrasiliensis], nocardithiocin

MHWLSTDWHSVRVFYYRPDKSDLLLHGVRPLFDRLRPQVEAVSFTRHWRQGPHVRLN
FLTDETTFTETIWPTVQRDMRAYLAEAPSTTVLDIEQELRRHVRMAELEHERGPLTPWH
SDNTVVLSEFDPRTDIMGDAGRADALARFYSDTTDLAFDCYDACRTRSALLSRALDLMI
AAGHIGAAGGIRQGFVSFRSHAEGFLAATDDGAALRDRWDANFRHSSDRLRDRVTALV
DAIDRAQFSEPFLADWARIMTCCRAFAEETAAAGVSPVPPPPPPAPGATRRLGEASAFHH
RLFDNPTWATMQQSAEFAVYRLILNFTYLHLTQLGVTPVERVLLCHLAAEAVEAAYGIS
AYEIVGQPYTEKAVG

ADR01090.1 NocO [Nocardia sp. ATCC 202099], nocathiocin

MTWTELVFPATQDEQPGLVAGVLAPLLADLDRPGLFLRELGPEGATRLLLQVRDAPPDL
PTRTAALPVQPTAVRAATVAPLGGPVFDGPGLDETTRGFLADTAPVAVDLSTRPDRTGA
ALTLMTAHLAAVADPARSGDGPPLSFLSFRSHAEAFLATTRDPNAARHAFDTRYTDHRT
TVEAAVRAILLDGDPGDAAPWSAAARAAKPRFTAGFASADLVAHTGYTRDHLRERTDF
ADNPFHSRAGASEQLQAYLGGDPSFLATRLLTSLLYVTLHSSGVSLMQRYFLCHAIAKA
CESIYHVDSMSLLADLAVG

WP_110565432.1 hypothetical protein [Micromonospora arborensis], Sch 18640

MTADRFGDVTVNVPWSSQEDLLLDVAAPTLDESVERGETESWFYLRETKEAKSYLRLK
FRTQSPSVEARLRARISEAVGEAGRENLFEYRPYNKEHDWLGGAGALGLAESFWTETTP
LALESLRATRGNRALRMALALDMLACSGVFLASRLPESVSRFGYRAGYLSYLATFEGYL
LLIRDPEAARARHARRYDLNRDALRPRIRRLVEIMQDPDGDLDGLPEATLRYMNSLRSH
LPPLQQGFDEGRFYLYATPRKDDFNKLTPAPDGLYRRPDLPWLADAPEAPVAGIHRAIA
DNPYYQGMIREDRRFQASRVAQAYGHWHLFRLGFLLSDRYTLCYLVARAFEEEFGLDA
AELIRSVKPEAEVS

WP_007074886.1 hypothetical protein [Micromonospora sp. ATCC 39149], Sch 40832

MSNNSFQLTVNVPWSVQEDILLDVAAPLLDHSMELGETERWFYLRENHRGTPYLRLRI
WCSTRSVVERLKAHIATQAVGAAAPNEMFVDRPYNYEHDWLGGPGGLDMLESFWAET
TPLAVRSLRLTRGDRALRMAHAFDLLVCTGALLGPRLPGRLGELGFRAGYLSYLATYE
GYLLLIRDPEAVRAKQNRRYEMNRDRLRARARELVEALQEPEGDLSGLPEFAGEWTHV
LRSYLPRIQQGFDEGRLFLYANPRNGELVKLAPSPDGLYRRPEVPWLADAPTPAVAGIH
RAIADNPYFQGMIREDRRFLASRLAQGYTHWHLYRLGFTLADRYTLFHLVARSFEEEFG
LDAVEIIRSVQPESEAAR

ACN80649.1 SioL [Streptomyces sioyaensis], siomycin

MSAMSSSEQKDLTVSVPWSVQEDLLLDVAAPTLEESVALGETESWFYLRENYGGRPFL
RLRFATRSPSVERRLKSRIMEHIGSAASDDPFEHQPYNHEHDWLGGEAGLGLAEAFWTE
TTPLALRTLRATRGDRALRLAAAFDFLVCSGVLLAPHLPPPVAKFGFKAGYLSYLATFE
GYMLLIRDPEGTRAKHAQRYEQNRNLLRPRLRALVEQMQDPEGDLADVPELAREWAV
RLRGYLPAIRKGFDEGRFYLYATPRKAETAKLTPSPDGLYRRPKVEWLADLPEAPVAGI
HRAIADNTYYQGMIREDRRFLASRLAQAYTNWHLYRLGFLLADRYTLFYLIARAFEEEY
DLDAAELIRSVQPEAEVSG

QJC58228.1 SulJ [Streptomyces viridochromogenes], sulfomycin

MHVFVTIPLDGGDRTLPQVVRGFITDLFGPAQPAWPGDFYFVRTWDQTTGRDCLQISLG
SDTATADEARDRVQRAADRHGLRPATAEVPFEEIPSPLWNSGIGGDGFDAAAKRLYRA
VAPVLARSVSEIGGDTAKAYLLALRLMVANSGATLLESEQRQMSSRDFGELLSLRMLSY
RSHYEGAKTAQAKDPDAFERRCAAYYEKFGSAAREFIGDCASDPLPAADDQPARAWV
DVVRQHFGTLRDECRAGLIAHDGPTLDDLHQTPDAALSRSSFHSKELSHEMSDLLHRNP
DFLAYRIQTSLLYSCLYTMGFNLAERFLFCYLLARANEELTGKSTQELAEGLDAVAKEL
ANH

WP_110669044.1 hypothetical protein [Streptomyces tateyamensis], thiopeptin

MSTAGPQRDLTINVPWSVQEEVLLDVAAPTLDQSVANGETRSWFYLRESFDGRPYLRL
RFDTDSPSVSRRLAARIGEEVAPAMREGLFEHRPYNREHDWLGGPAGLPLAERFWTETT
PLALEAITATRGDRALRMAMTFDLLVASGALLASRLPPSVGKHGFKAGYLSYLATFEGY
MLLIRDPEGTRAKHAQRWEANRELLRPRLRALVEAMRDPDGDLAELPPMAARWTGTL
RDYLPAIQQGFDEDRFYLYATPRTSEAAKITASPDGLYRRPEVPWLADLPEPPVAGIHRA
IADNQYYQGMIREDRRFLASRLAQAYTNWHLFRLGFQLADRYTLFYLVARAFEEEYEL
DAAELIRSVRPQAEVAG

ADO67771.1 hypothetical protein [Nocardiopsis sp. TFS65-07], TP-1161

MSKYDLSVADRYRSMVSELFSRGGRLPIGDHYFLRTWDQRTAGECVQAVIEVEDAEAP
GVLDVVNGVARENGFEADVRETPPAEVPSPLWNVGFAGISFDRVSKALYRDASPLASEF
AGAARESVGGAYRLALRLMAEHDKATLVGSEQRGVSPLGFGDLLSTRLLSFRSHYEGM
AARVRDRESFEGRCEKYYEDFGHYARGLVSSCMESENPTDDPHMRQWALNISETTKPL

REEVREGRIVNVGLTLDDLNRDRETLLSPAGFHAFGADSEAFQYLMNSDPDFMAFRIQA
GLLYSCLHTIGFSLPERFVLCYILGRANEDVSGKSANELQERLMSVAEGMYARTDPARE
GQWS

4.7.2 Curating Thiopeptide Precursor Peptides.

RODEO provides a score for each thiopeptide. If the score is 20 or greater, then the precursor peptide is predicted to be a valid thiopeptide. All of these precursors were collected. For some thiopeptide gene clusters, no precursor scored 20 or above. In this case, the best scoring precursor or precursors were kept. Additionally, all predicted precursors with a thiopeptide score between 5 and 19 were manually analyzed and sequences with hallmark thiopeptide characteristics were added. The final dataset was filtered for duplicates and 1,393 predicted thiopeptide precursors remained.

4.7.3 Creating Sequence Similarity Networks.

All sequences for pyridine synthases, thiopeptide precursors and TipAS homolgos were organized in fasta format and submitted to EFI-EST to make a sequence similarity network. For each dataset, the default settings were left unchanged. For pyridine synthases, precursors and TipAS homologs alignment scores of 43, 8, and 80 were employed, respectively. All were concatenated at 100 % sequence identity except for TipAS homologs which were concatenated at 75% sequence identity. All networks were visualized with CytoScape with an organic layout.

4.7.4 WEBLOGOs for Predicted Thiopeptide Cores.

Sequences from each of the 38 selected clusters from the precursor SSN (figure 4.3), were extracted, organized into fasta format, and aligned using MUSCLE. Alignments were trimmed using trimAL[50] before submitting to https://weblogo.berkeley.edu/logo.cgi to make WEBLOGOs. Core regions were manually selected based on patterns from characterized thiopeptides.

4.7.5 Thiopeptide Predicted Gene Clusters with antiSMASH

Nucleotide accessions obtained from RODEO were used to download 708 available RefSeq and GenBank assemblies. These were submitted to antiSMASH and all predicted gene clusters containing an accession number for a predicted pyridine synthase were collected. For ease of viewing different gene clusters, the antiSMASH files were input into BiG-SCAPE[51] and selected gene clusters were downloaded as .SVG files. Enzyme activities were predicted based on GenBank, RODEO, and antiSMASH annotations or BLAST analysis of nearest homologs.

REFERENCES

(1) Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; Heel, A. J. van; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Burgos, M. J. G.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; Wezel, G. P. van; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A. *Nat. Prod. Rep.* **2020**, Advance Article

(2) Just-Baringo, X.; Albericio, F.; Álvarez, M. Thiopeptide Antibiotics: Retrospective and Recent Advances. *Marine Drugs* **2014**, *12*, 317.

(3) Bird, K. E.; Xander, C.; Murcia, S.; Schmalstig, A. A.; Wang, X.; Emanuele, M. J.; Braunstein, M.; Bowers, A. A. *ACS Chem Biol.* **2020**, *15*, 2164.

(4) Heffron, S. E.; Jurnak, F. *Biochemistry* **2000**, *39*, 37.

(5) Parmeggiani, A.; Nissen, P. *FEBS Lett.* **2006**, *580*, 4576.

(6) Harms, J. M.; Wilson, D. N.; Schluenzen, F.; Connell, S. R.; Stachelhaus, T.; Zaborowska, Z.; Spahn, C.; Fucini, P. *Molecular Cell* **2008**, *30*, 26.

(7) Walter, J. D.; Hunter, M.; Cobb, M.; Traeger, G.; Spiegel, P. C. *Nucleic Acids Res.* **2012**, *40*, 360.

(8) Selva, E.; Montanini, N.; Stella, S.; Soffientini, A.; Gastaldo, L.; Denaro, and M. Targeted Screening for Elongation Factor Tu Binding Antibiotics. *J. Antibiotics* **1997**, *50*, 22.

(9) Hayashi, S.; Ozaki, T.; Asamizu, S.; Ikeda, H.; Ōmura, S.; Oku, N.; Igarashi, Y.; Tomoda, H.; Onaka, H. *Chem Biol.* **2014**, *21*, 679.

(10) Hudson, G. A.; Hooper, A. R.; DiCaprio, A. J.; Sarlah, D.; Mitchell, D. A. *Org. Lett.* **2020**, XXXX, XXX, XXX-XXX

(11) Dong, L.; Nakashima, N.; Tamura, N.; Tamura, T. *FEMS Microbiol. Lett.* **2004**, *237*, 35.

(12) Murakami, T.; Holt, T. G.; Thompson, C. J. *J. Bacteriol.* **1989**, *171*, 1459.

(13) Bleich, R.; Watrous, J. D.; Dorrestein, P. C.; Bowers, A. A.; Shank, E. A. *Proc. Natl. Acad. Sci.* **2015**, *112*, 3086.

(14) Habazettl, J.; Allan, M.; Jensen, P. R.; Sass, H.-J.; Thompson, C. J.; Grzesiek, S. *Proc. Natl. Acad. Sci.* **2014**, *111*, E5498

(15) Wever, W. J.; Bogart, J. W.; Baccile, J. A.; Chan, A. N.; Schroeder, F. C.; Bowers, A. A. *J. Am. Chem. Soc.* **2015**, *137*, 3494.

(16) Wever, W. J.; Bogart, J. W.; Bowers, A. A. *J. Am. Chem. Soc.* **2016**, *138*, 13461.

(17) Bogart, J. W.; Kramer, N. J.; Turlik, A.; Bleich, R. M.; Catlin, D. S.; Schroeder, F. C.; Nair, S. K.; Williamson, R. T.; Houk, K. N.; Bowers, A. A. *J. Am. Chem. Soc.* **2020**, *142*, 13170.

(18) Bogart, J. W.; Bowers, A. A. *J. Am. Chem. Soc.* **2019**, *141*, 1842.

(19) Fleming, S. R.; Bartges, T. E.; Vinogradov, A. A.; Kirkpatrick, C. L.; Goto, Y.; Suga, H.; Hicks, L. M.; Bowers, A. A. *J. Am. Chem. Soc.* **2019**, *141*, 758.

(20) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. *Nat. Chem. Biol.* **2017**, *13*, 470.

(21) Zallot, R.; Oberg, N.; Gerlt, J. A. *Biochemistry* **2019**, *58*, 4169.

(22) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. *Nucleic Acids Res.* **2019**, *47*, W81.

(23) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. *J. Am. Chem. Soc.* **2018**, *140*, 9494.

(24) Cogan, D. P.; Hudson, G. A.; Zhang, Z.; Pogorelov, T. V.; van der Donk, W. A.; Mitchell, D. A.; Nair, S. K. *Proc. Natl. Acad. Sci.* **2017**, *114*, 12928.

(25) Brown, L. C. W.; Acker, M. G.; Clardy, J.; Walsh, C. T.; Fischbach, M. A. *Proc. Natl. Acad. Sci.* **2009**, *106*, 2549.

(26) Liao, R.; Duan, L.; Lei, C.; Pan, H.; Ding, Y.; Zhang, Q.; Chen, D.; Shen, B.; Yu, Y.; Liu, W. *Chem. Biol.* **2009**, *16*, 141.

(27) Yu, Y.; Duan, L.; Zhang, Q.; Liao, R.; Ding, Y.; Pan, H.; Wendt-Pienkowski, E.; Tang, G.; Shen, B.; Liu, W. *ACS Chem. Biol.* **2009**, *4*, 855.

(28) Kelly, W. L.; Pan, L.; Li, C. *J. Am. Chem. Soc.* **2009**, *131*, 4327.

(29) Wang, J.; Yu, Y.; Tang, K.; Liu, W.; He, X.; Huang, X.; Deng, Z. *Appl. Environ. Microbiol.* **2010**, *76*, 2335.

(30) Morris, R. P.; Leeds, J. A.; Naegeli, H.; Oberer, L.; Memmert, K.; Weber, E.; LaMarche, M. J.; Parker, C. N.; Burrer, N.; Esterow, S.; Hein, A. E.; Schmitt, E. K.; Krastel, P. *J. Am. Chem. Soc.* **2009**, *131*, 5946.

(31) Malcolmson, S. J.; Young, T. S.; Ruby, J. G.; Skewes-Cox, P.; Walsh, C. T. *Proc. Natl. Acad. Sci.* **2013**, *110*, 8483.

(32) Acker, M. G.; Bowers, A. A.; Walsh, C. T. *J. Am. Chem. Soc.* **2009**, *131*, 17563.

(33) Bowers, A. A.; Walsh, C. T.; Acker, M. G. *J. Am. Chem. Soc.* **2010**, *132*, 12182.

(34) Bowers, A. A.; Acker, M. G.; Koglin, A.; Walsh, C. T. *J. Am. Chem. Soc.* **2010**, *132*, 7519.

(35) Bowers, A. A.; Acker, M. G.; Young, T. S.; Walsh, C. T. *J. Am. Chem. Soc.* **2012**, *134*, 10313.

(36) Luo, X.; Zambaldo, C.; Liu, T.; Zhang, Y.; Xuan, W.; Wang, C.; Reed, S. A.; Yang, P.-Y.; Wang, R. E.; Javahishvili, T.; Schultz, P. G.; Young, T. S. *Proc. Natl. Acad. Sci.* **2016**, *113*, 3615.

(37) Tran, H. L.; Lexa, K. H.; Julien, O.; Young, T. S.; Walsh, C. T.; Jacobson, M. P.; Wells, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 2541.

(38) Aulakh, V. S.; Ciufolini, M. A. *J. Am. Chem. Soc.* **2011**, *133*, 5900.

(39) Vinogradov, A. A.; Shimomura, M.; Goto, Y.; Ozaki, T.; Asamizu, S.; Sugai, Y.; Suga, H.; Onaka, H. *Nat. Commun.* **2020**, *11*, 2272.

(40) Vinogradov, A. A.; Shimomura, M.; Kano, N.; Goto, Y.; Onaka, H.; Suga, H. *J. Am. Chem. Soc.* **2020**, *142*, 13886.

(41) Thibodeaux, C. J.; Ha, T.; van der Donk, W. A. *J. Am. Chem. Soc.* **2014**, *136*, 17513.

(42) Li, B.; Sher, D.; Kelly, L.; Shi, Y.; Huang, K.; Knerr, P. J.; Joewono, I.; Rusch, D.; Chisholm, S. W.; van der Donk, W. A. *Proc. Natl. Acad. Sci.* **2010**, *107*, 10430.

(43) McClerren, A. L.; Cooper, L. E.; Quan, C.; Thomas, P. M.; Kelleher, N. L.; van der Donk, W. A. *Proc. Natl. Acad. Sci.* **2006**, *103*, 17243.

(44) Singh, M.; Chaudhary, S.; Sareen, D. *Mol. Microbiol.* **2020**, *113*, 326.

(45) Xin, B.; Zheng, J.; Liu, H.; Li, J.; Ruan, L.; Peng, D.; Sajid, M.; Sun, M. *Front. Microbiol.* **2016**, *7*, 1115.

(46) Burkhart, B. J.; Schwalen, C. J.; Mann, G.; Naismith, J. H.; Mitchell, D. A. *Chem. Rev.* **2017**, *117*, 5389.

(47) Kahmann, J. D.; Sass, H.-J.; Allan, M. G.; Seto, H.; Thompson, C. J.; Grzesiek, S. *EMBO J.* **2003**, *22*, 1824.

(48) Yun, B.-S.; Hidaka, T.; Kuzuyama, T.; Seto, H. *J. Antibiotics* **2001**, *54*, 375.

(49) Chiu, M. L.; Viollier, P. H.; Katoh, T.; Ramsden, J. J.; Thompson, C. J. *Biochemistry* **2001**, *40*, 12950.

(50) Capella-Gutiérrez, S.; Silla-Martínez, J. M.; Gabaldón, T. *Bioinformatics* **2009**, *25*, 1972.

(51) Navarro-Muñoz, J. C.; Selem-Mojica, N.; Mullowney, M. W.; Kautsar, S. A.; Tryon, J. H.; Parkinson, E. I.; Santos, E. L. C. D. L.; Yeong, M.; Cruz-Morales, P.; Abubucker, S.; Roeters, A.; Lokhorst, W.; Fernandez-Guerra, A.; Cappelini, L. T. D.; Goering, A. W.; Thomson, R. J.; Metcalf, W. W.; Kelleher, N. L.; Barona-Gomez, F.; Medema, M. H. *Nat. Chem. Biol.* **2020**, *16*, 60.

CHAPTER 5: CONCLUSION

5.1 Concluding Remarks.

Genetically encoded peptide libraries are becoming increasingly important, and there has been a sustained effort to diversify these libraries with modifications. Of particular note is the success of chemical cyclization strategies leading to many tight binding inhibitors.[1–4] While chemical techniques have been valuable, there is a limited number of reactions that are display compatible.[5] To increase the types of peptide transformations available to peptide display, the power of enzymes as library biocatalysts is being explored. Specifically, RiPPs are being targeted for the preparation of peptide libraries because they are promiscuous, programmable, and act on peptides.[6] Lantibiotics are the first class of RiPP to be successfully applied for the discovery of novel inhibitors,[7–9] but there exist many other types of RiPPs that may perform well in peptide display selections.

mRNA display has a few advantages over other display technologies; namely, (1) diversity is not dependent on transformation efficiency, (2) libraries can be quickly prepared, (3) the tag is small, (4) unnatural amino acids can be readily incorporated, and (5) selections are fast.[10,11] Because of these advantages, my work has focused on uniting RiPP biochemistry with mRNA display. The initial experiments presented with PaaA showed for the first time that RiPP enzymes could modify RNA linked substrates, provided assays to measure peptide library modification, and paves the way for developing RiPP modified mRNA display libraries for inhibitor discovery.

Thiopeptides are another class of RiPP expected to perform well in peptide display selections. This is because thiopeptides are privileged, having many known bioactivities;[12] their structure lends to the technique, having an open C-terminus; and their functional groups may one day be helpful for tuning the bioavailability of a peptide library.[13] To date thiopeptides have not been used with display, because no versatile display-ready synthesis exists. Ergo, my work has focused on overcoming this limitation by designing and implementing a new chemoenzymatic strategy to make thiopeptides that is display compatible. My strategy uses a couple of highly promiscuous RiPP enzymes from disparate biosyntheses[14,15] paired with nucleotide compatible chemistry[16,17] to make the complex multi-azole / Dha substrate for thiopeptide pyridine synthases,[18] that complete the transformation to the thiopeptide core. In this manner, different thiopeptide classes can be made in an mRNA display-ready fashion by simply exchanging the pyridine synthase.

Currently, there are 5 pyridine synthase that have been reconstituted in vitro[15,19–21] that we can use for the display of thiopeptides. But it is not certain that these will be the best enzymes for mRNA display. Previous chemoenzymatic synthesis and my reported synthesis readily lends themselves to the speedy reconstitution of new thiopeptides via a kindred pyridine synthase.[21] To advance the discovery of uncharacterized pyridine synthases, my work has also focused on using current bioinformatic tools[22,23] to collect a database of thiopeptide BGCs from which we have prioritized 14 gene clusters for chemoenzymatic discovery. This work primes experiments for the realization of new thiopeptides and will provide a toolbox of the best pyridine synthases to be used for peptide display.

Overall, my work has greatly advanced the diversification of peptide libraries by uniting RiPP enzymes with mRNA display. While we have prioritized thiopeptides, many other RiPPs

can be used in this manner, and we anticipate that the work herein will be of significant use and will encourage others to develop new RiPP biochemistry for peptide library modification.

REFERENCES

(1) Kale, S. S.; Villequey, C.; Kong, X.-D.; Zorzi, A.; Deyle, K.; Heinis, C. *Nat. Chem.* **2018**, *10*, 715.

(2) Millward, S. W.; Takahashi, T. T.; Roberts, R. W. *J. Am. Chem. Soc.* **2005**, *127*, 14142.

(3) Passioura, T.; Suga, H. *Chem. Comm.* **2017**, *53*, 1931.

(4) Hacker, D. E.; Abrigo, N. A.; Hoinka, J.; Richardson, S. L.; Przytycka, T. M.; Hartman, M. C. T. Direct, *ACS Comb. Sci.* **2020**, *22*, 306.

(5) Malone, M. L.; Paegel, B. M. *ACS Comb. Sci.* **2016**, *18*, 182.

(6) Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; Heel, A. J. van; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Burgos, M. J. G.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; Wezel, G. P. van; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A. *Nat. Prod. Rep.* **2020**, Advance Article

(7) Urban, J. H.; Moosmeier, M. A.; Aumüller, T.; Thein, M.; Bosma, T.; Rink, R.; Groth, K.; Zulley, M.; Siegers, K.; Tissot, K.; Moll, G. N.; Prassler, J. *Nat. Comm.* **2017**, *8*, 1500.

(8) Yang, X.; Lennard, K. R.; He, C.; Walker, M. C.; Ball, A. T.; Doigneaux, C.; Tavassoli, A.; van der Donk, W. A. *Nat. Chem. Biol.* **2018**, *14*, 375.

(9) Hetrick, K. J.; Walker, M. C.; van der Donk, W. A. *ACS Cent. Sci.* **2018**, *4*, 458.

(10) Huang, Y.; Wiedmann, M. M.; Suga, H. *Chem. Rev.* **2019**, *119*, 10360.

(11) Goto, Y.; Katoh, T.; Suga, H. *Nat. Protoc.* **2011**, *6*, 779.

(12) Just-Baringo, X.; Albericio, F.; Álvarez, M. *Marine Drugs* **2014**, *12*, 317.

(13) Matsson, P.; Doak, B. C.; Over, B.; Kihlberg, J. *Adv. Drug Deliver. Rev.* **2016**, *101*, 42.

(14) Koehnke, J.; Mann, G.; Bent, A. F.; Ludewig, H.; Shirran, S.; Botting, C.; Lebl, T.; Houssen, W. E.; Jaspars, M.; Naismith, J. H. *Nat. Chem. Biol.* **2015**, *11*, 558.

(15) Hudson, G. A.; Zhang, Z.; Tietz, J. I.; van der Donk, W. A.; Mitchell, D. A. *J. Am. Chem. Soc.* **2015**, *137*, 16012.

(16) Okeley, N.; Zhu, Y.; van der Donk, W. A. *Org. Lett.* **2000**, *2*, 3603.

(17) Hofmann, F. T.; Szostak, J. W.; Seebeck, F. P. *J. Am. Chem. Soc.* **2012**, *134*, 8038.

(18) Wever, W. J.; Bogart, J. W.; Baccile, J. A.; Chan, A. N.; Schroeder, F. C.; Bowers, A. A. *J. Am. Chem. Soc.* **2015**, *137*, 3494.

(19) Wever, W. J.; Bogart, J. W.; Bowers, A. A. *J. Am. Chem. Soc.* **2016**, *138*, 13461.

(20) Fleming, S. R.; Bartges, T. E.; Vinogradov, A. A.; Kirkpatrick, C. L.; Goto, Y.; Suga, H.; Hicks, L. M.; Bowers, A. A. *J. Am. Chem. Soc.* **2019**, *141*, 758.

(21) Hudson, G. A.; Hooper, A. R.; DiCaprio, A. J.; Sarlah, D.; Mitchell, D. A. *Org. Lett.* **2020**, XXXX, XXX, XXX-XXX

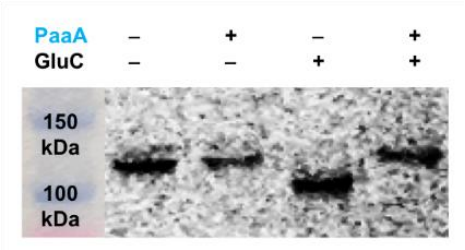
(22) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. *Nucleic Acids Res.* **2019**, *47*, W81.

(23) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. *Nat. Chem. Biol.* **2017**, *13*, 470.

APPENDIX A. SUPPLEMENTARY FIGURES AND TABLES FOR CHAPTER 2

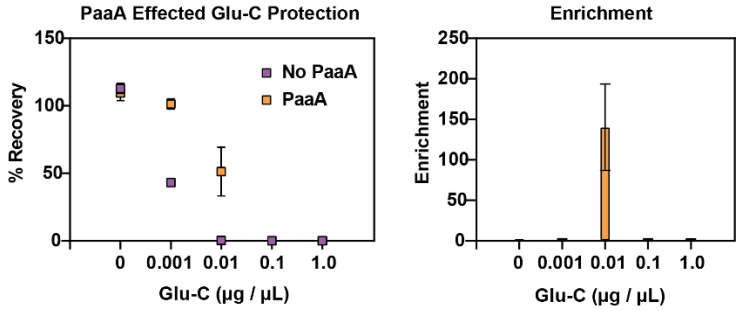
Figure A.1 PaaA Effected GluC protection of PaaP. (a) PaaP (MIKFSTLSQRISAITEEN-AMYTKGQVIVLS) was displayed with 35S-methionine. Treatment with GluC cleaves off the N-terminus noted by band shift that is protected in the presence of PaaA. This data was collected by Dr. Paul Himes. (b) Diagram of working PaaP-display construct for streptavidin pullout-based selections. (c) A dot plot showing the difference in streptavidin bead recovery between PaaA treated and non-treated PaaP-mRNA fusions with increasing µg of GluC measured by qPCR. Error bars represent the standard error of the mean (SEM, n=3). PaaA treated streptavidin recoveries were divided by non-PaaA treated streptavidin recoveries to calculate a fold enrichment effected by PaaA per GluC treatment. Error bars represent the SEM, n=3. (d) Dot plots showing the dependence of PaaA concentration and reaction time on streptavidin recovery of PaaP-mRNA fusions. Error bars represent the SEM, n=3.

a.



b.

c.



d.

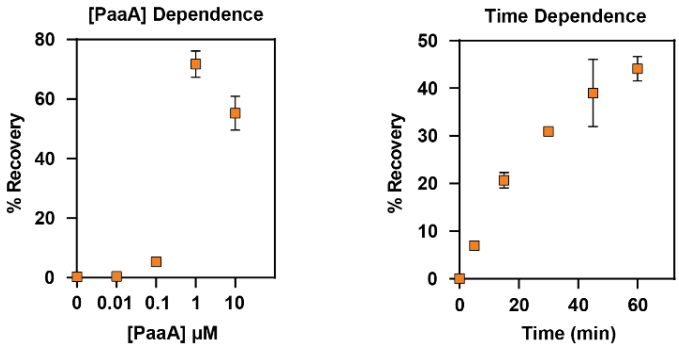
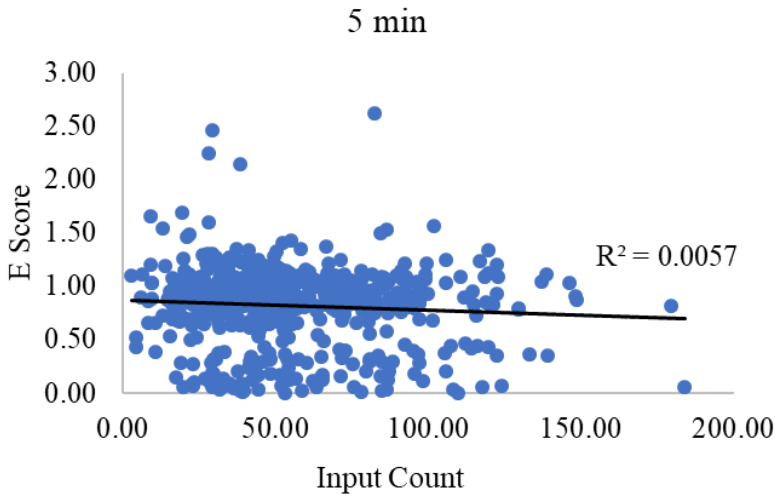
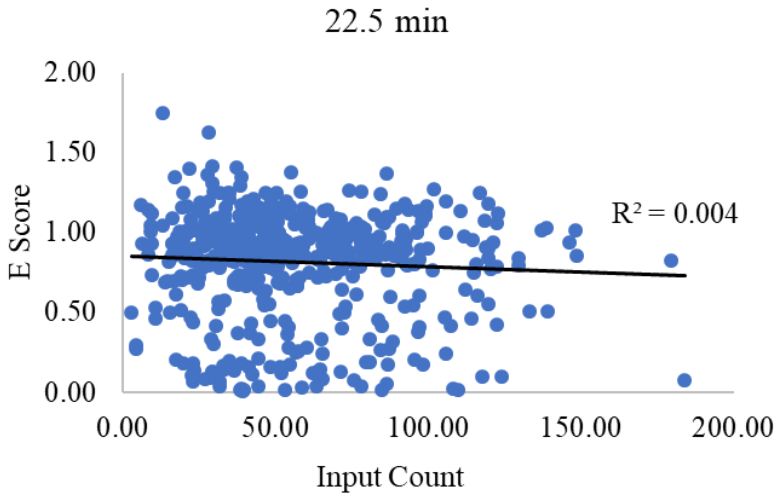


Figure A.2. NGS analysis of E scores. No correlation between input counts and E scores is found for the smSVL dataset (a) 5 min, (b) 22.5 min, and (c) 60 min.

a.



b.



c.

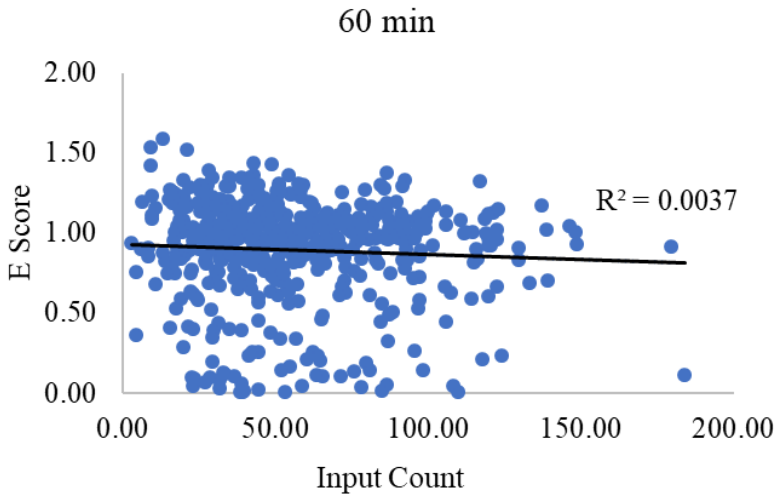
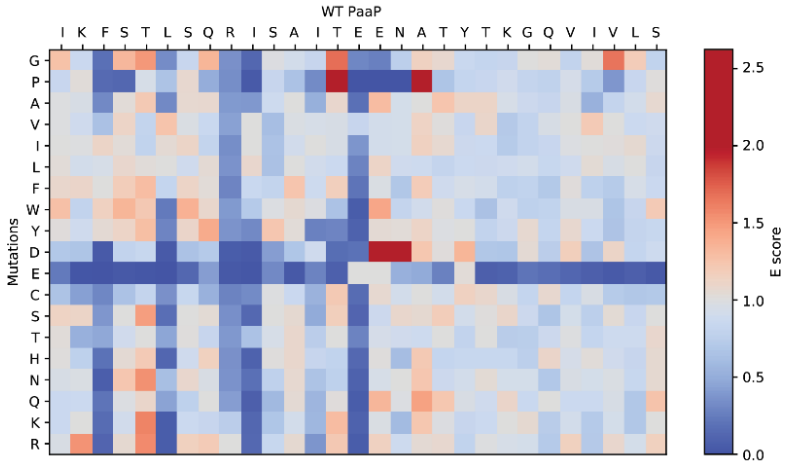
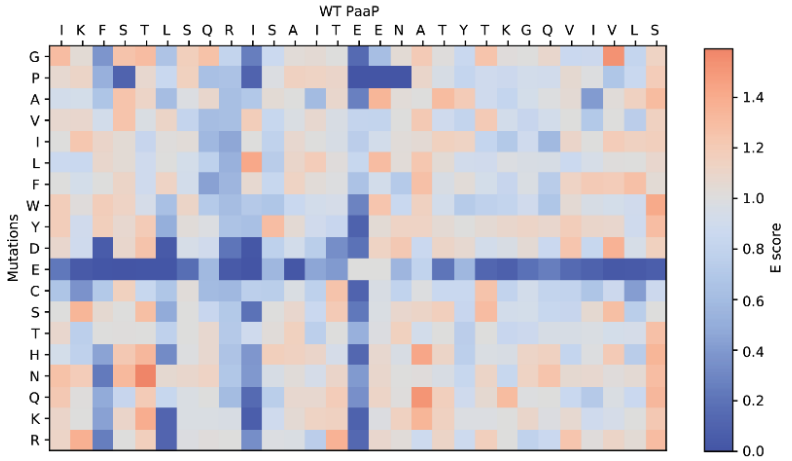


Figure A.3. PaaP smSVL with 5 min and 60 min PaaA Treatments. PaaP-mRNA fusions of smSVL were treated with PaaA for (a) 5 min and (b) 60 min before purification, GluC treatment, and streptavidin enrichment. Each square represents an average (n=3). (c) smSVL streptavidin recoveries were monitored by qPCR for PaaA reaction times of 0, 5, 22.5, and 60 mins and compared to a control PaaP-WT-mRNA display. Error bars represent the SEM (n=3).

a.



b.



c.

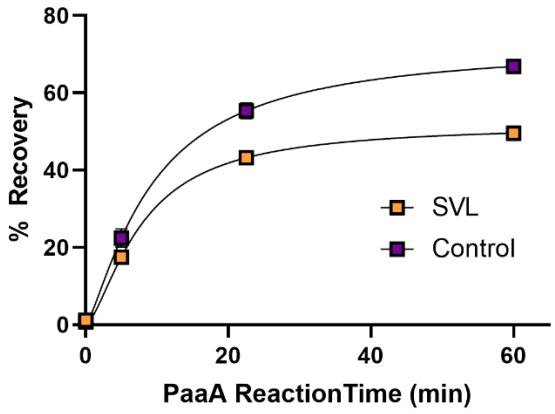
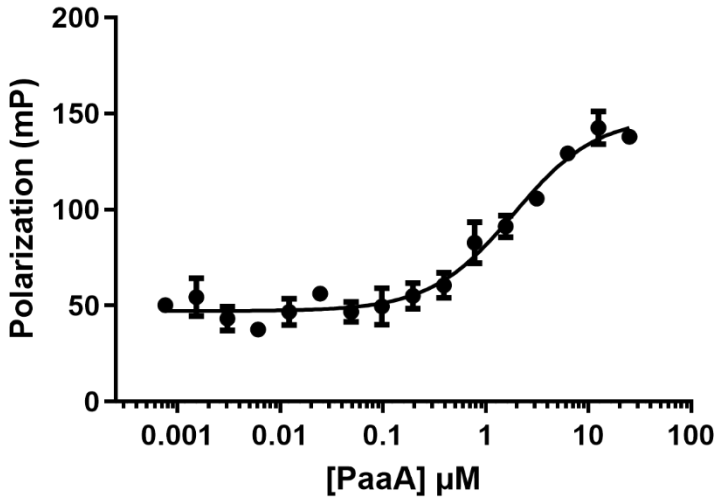


Figure A.4. Fluorescence Polarization of Tamra-WT-PaaP$_{1-12}$ and PaaA. (a) Binding between TAMRA-PaaP$_{1-12}$ was measured by fluorescence polarization by keeping peptide constant at 5 nM and titrating PaaA from 25 µM - 0.763 nM. EC_{50} = 1.869 ± 0.485 µM. (b) Six different aspartic acid mutations were synthesized and tested for competitive binding with 5 nM TAMRA-PaaP$_{1-12}$ at 100 nM PaaA. Competing peptides were titrated from 50 - 0.0061 µM. WT IC_{50} = 0.3482 ± 0.044 µM; T6D IC_{50} = 1.471 ± 0.238 µM. IC_{50} values for other mutants were not determined as they are > 10 µM. EC_{50} and IC_{50} errors calculations and error bars represent the SEM (n=3).

a.



b.

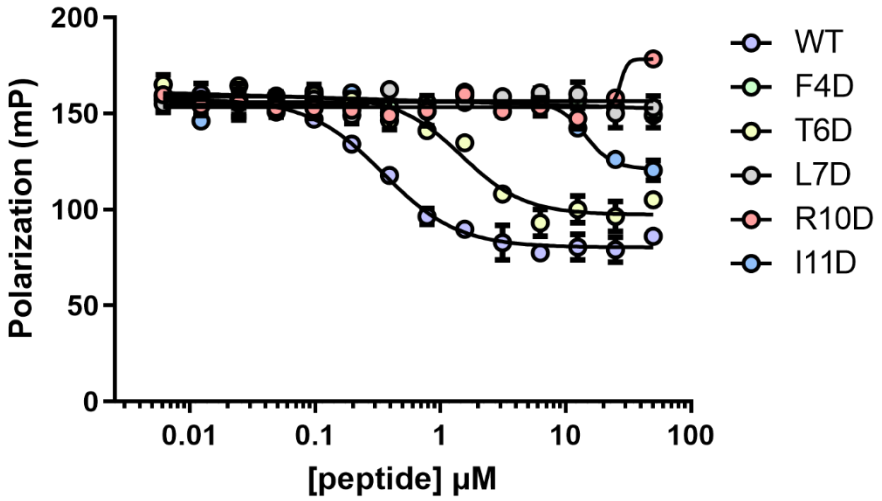
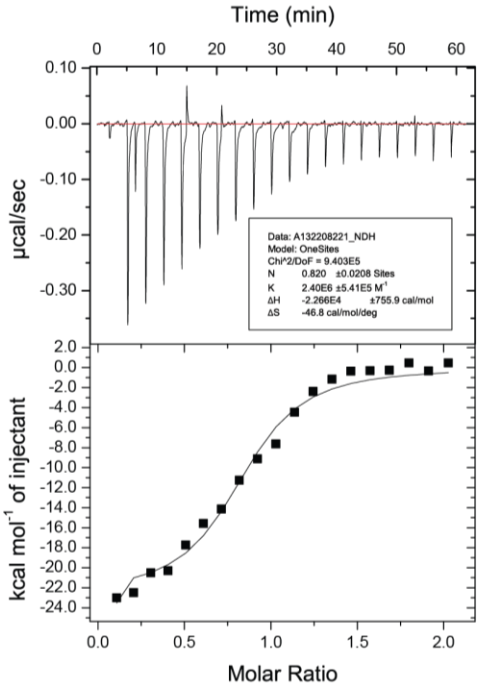
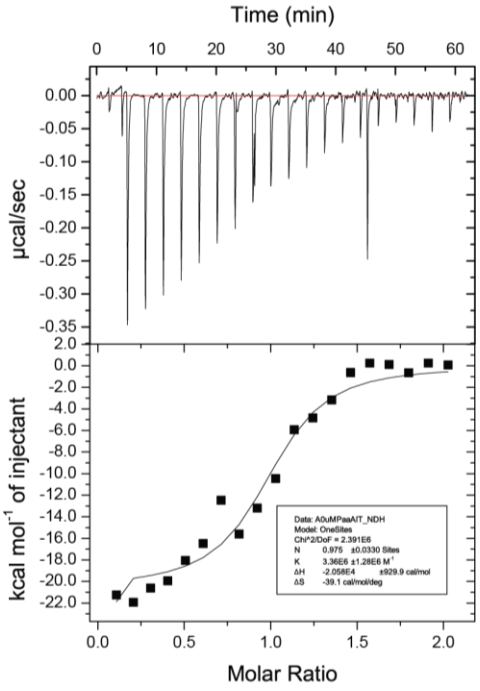


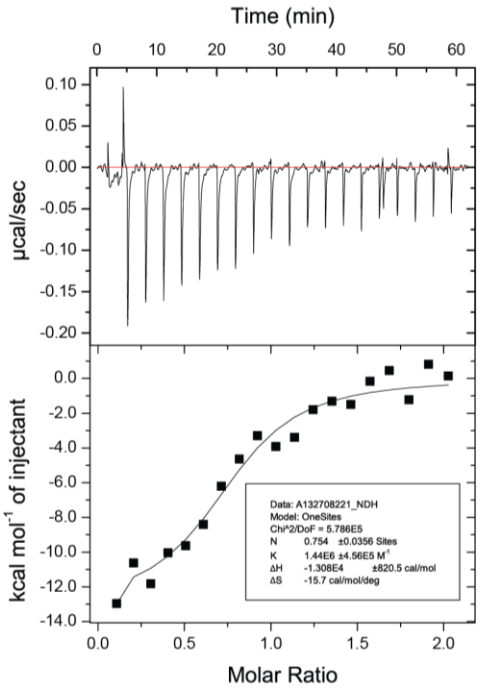
Figure A.5. Isothermal Calorimetry (ITC) of PaaP Peptides. To further substantiate the FP data, ITC experiments were run for each synthesized PaaP peptide. The (a) WT and (b) T6D peptides were run in duplicate and mutants (c) F4D, (d) L7D, (e) R10D, and (f) I11D were performed once since no binding isotherms were detected. WT K_d = 0.357 ± 0.08; T6D K_d = 0.707 ± 0.02. Error is the standard deviation between K_d measurements (n=2).

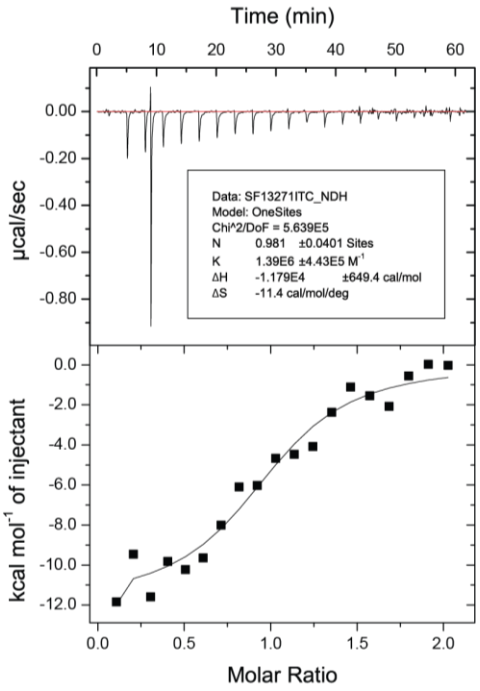
a.



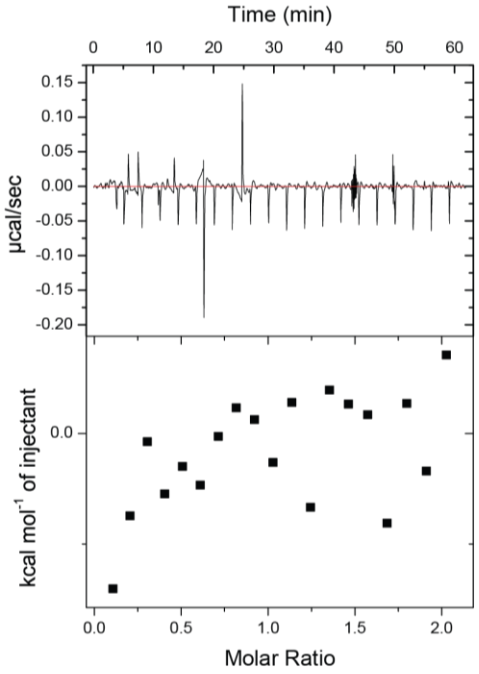


b.

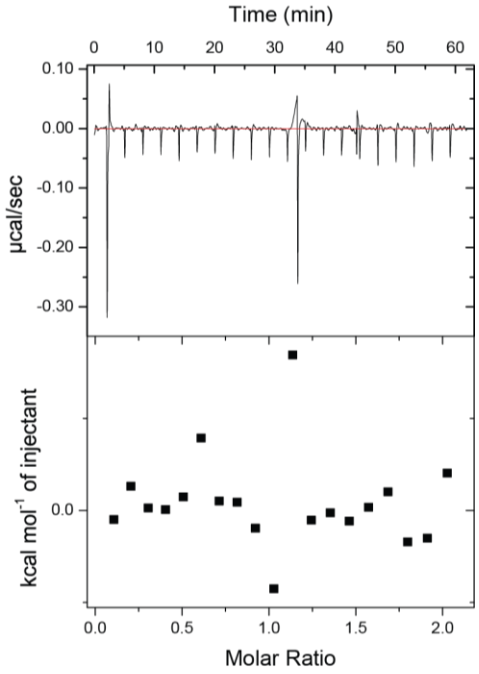




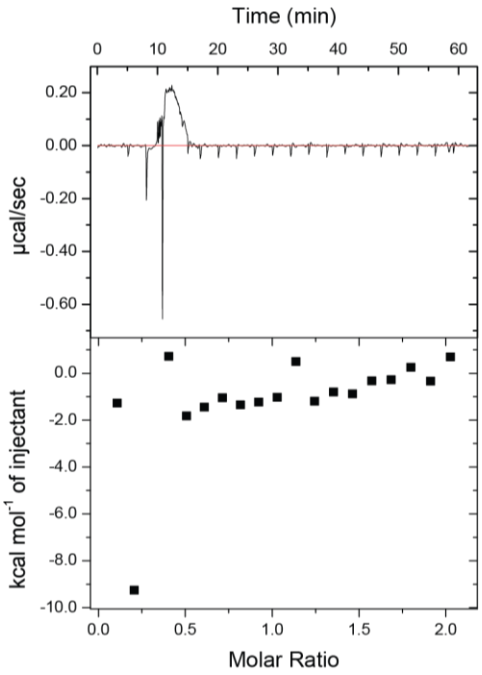
c.


Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio

d.


Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio

e.


Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio

f.

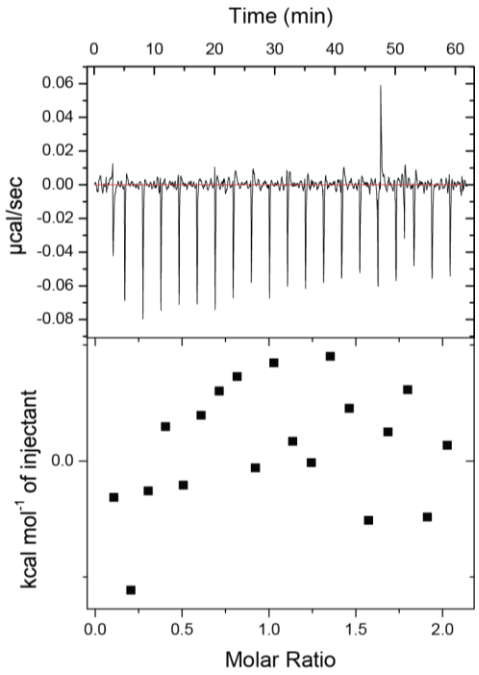

Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio

Figure A.6. Alignment of PaaP Homologs. PaaP was used in a BLAST search on NCBI and all sequences of an e-value cut off of 5.0 were obtained. Sequences obviously not PaaP homologs were removed and extended N-terminal sequence was trimmed (42 total sequences). All homologs were aligned according to the core glutamic acids and visualized using JalView. Blue highlights are shaded by percent identity. Black arrows represent the positions discovered by the smSVL data to be important for PaaA reactivity.

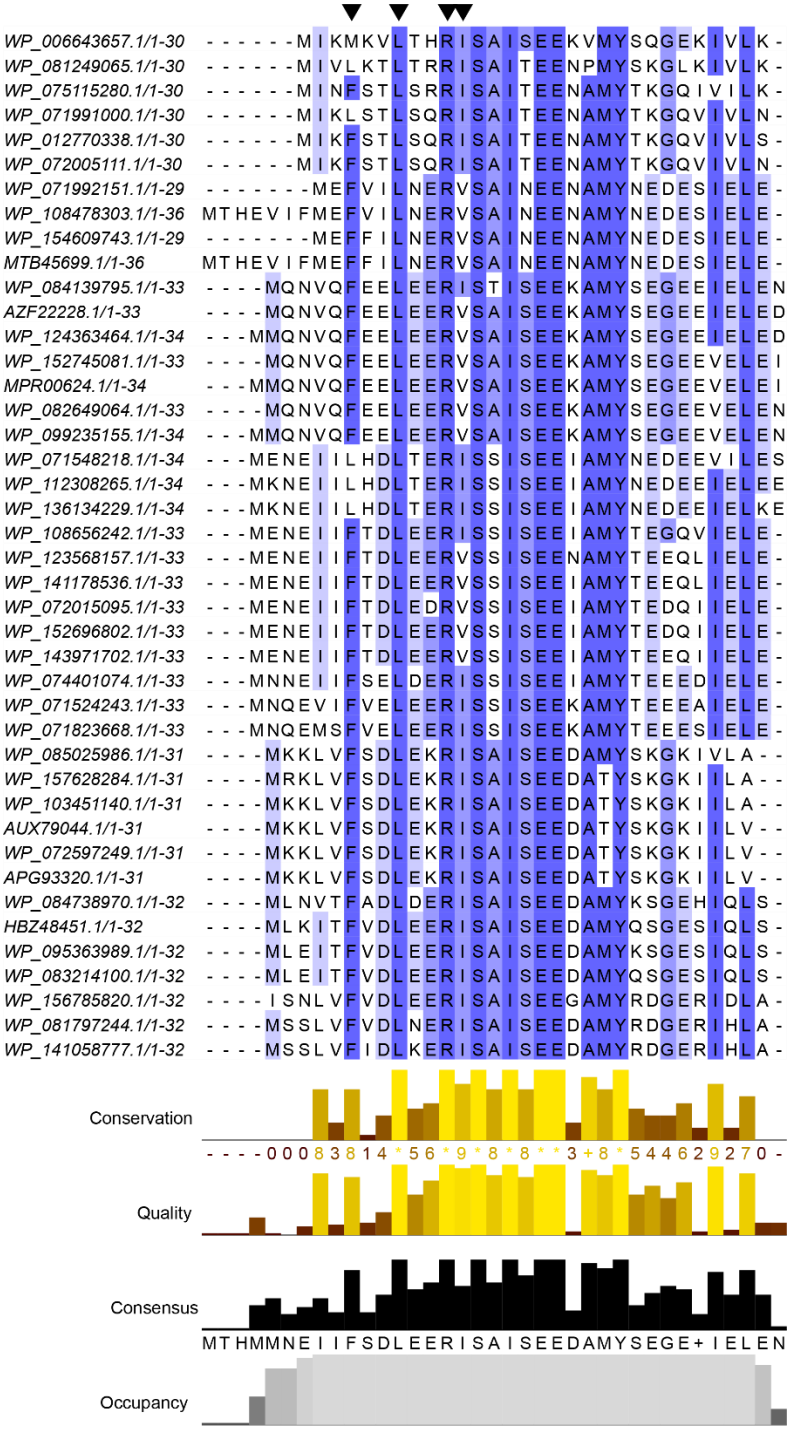


Figure A.7. PaaA Reaction Mechanisms.

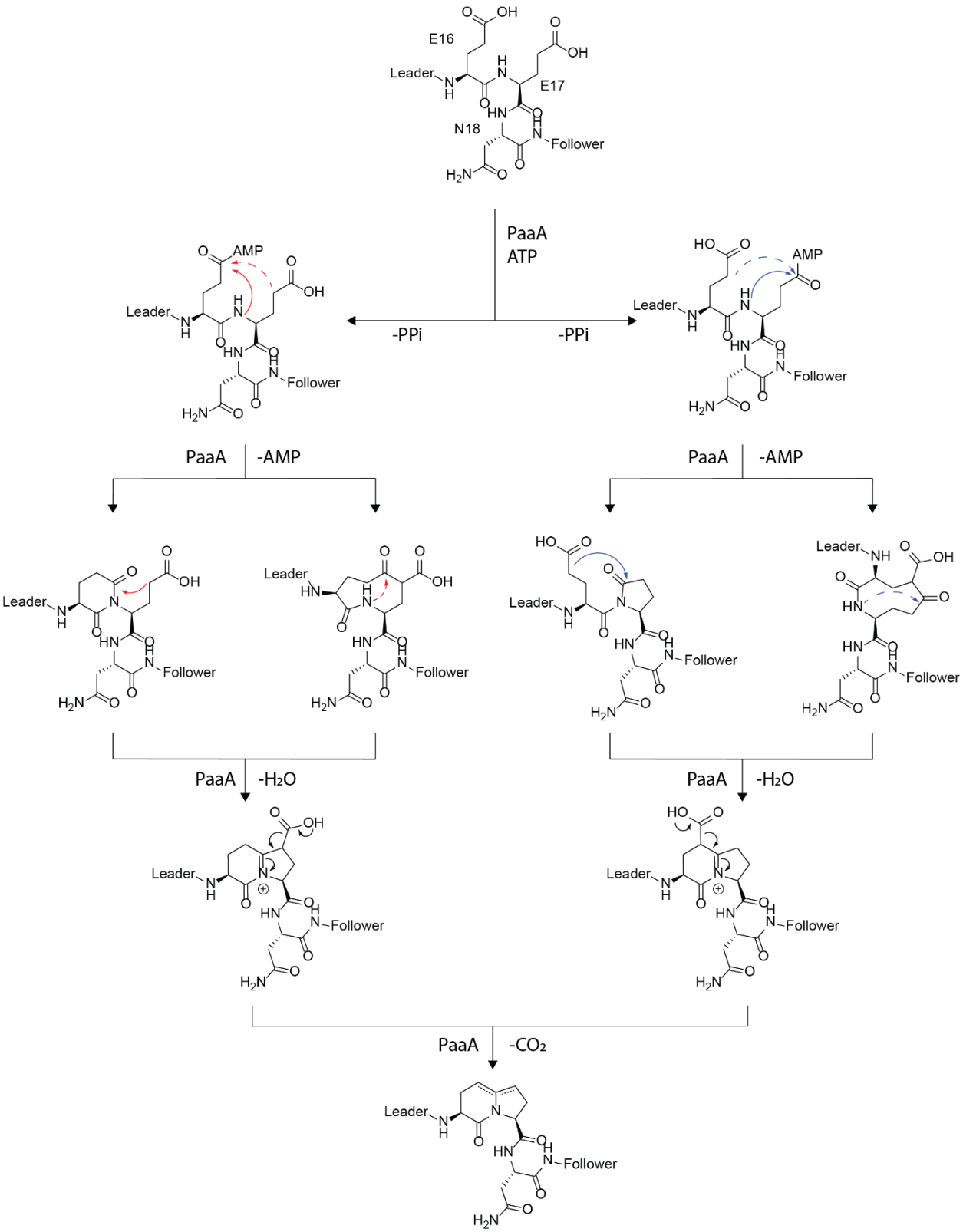


Figure A.8. PaaP A13C NCL Reaction Scheme. The following procedure was adapted from a previous report[6] to synthesize E17 labeled PaaA substrate Ac-MIKFSTLSQRISCITEE*NATYT-KGQVIVLS. Trifluoroacetic acid = TFA, Triisopropylsilane = TIS, Tris(2-carboxyethyl)phosphine = TCEP. Numbers next to atoms designate isotope. "*" following an amino acid designates isotope labeling

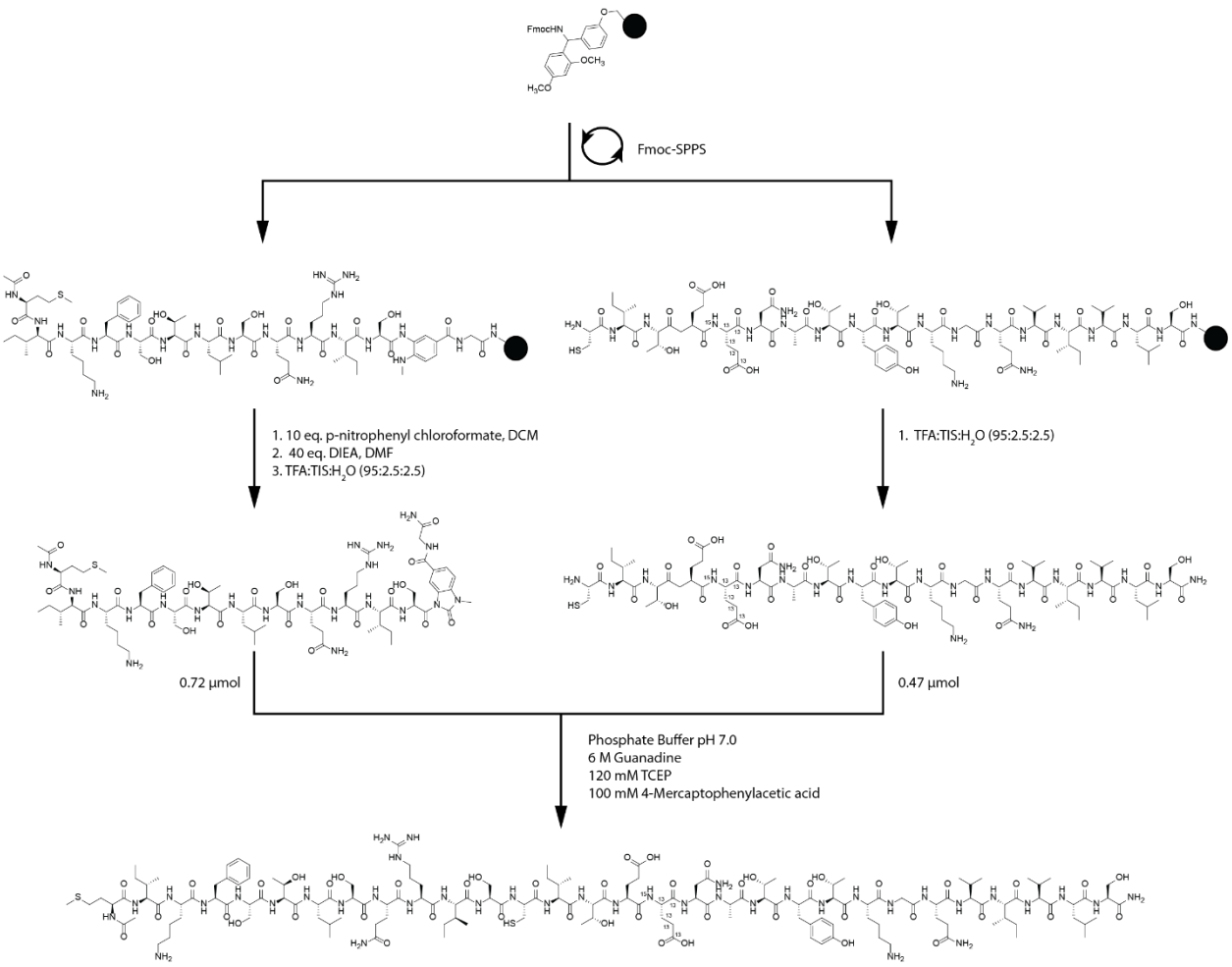
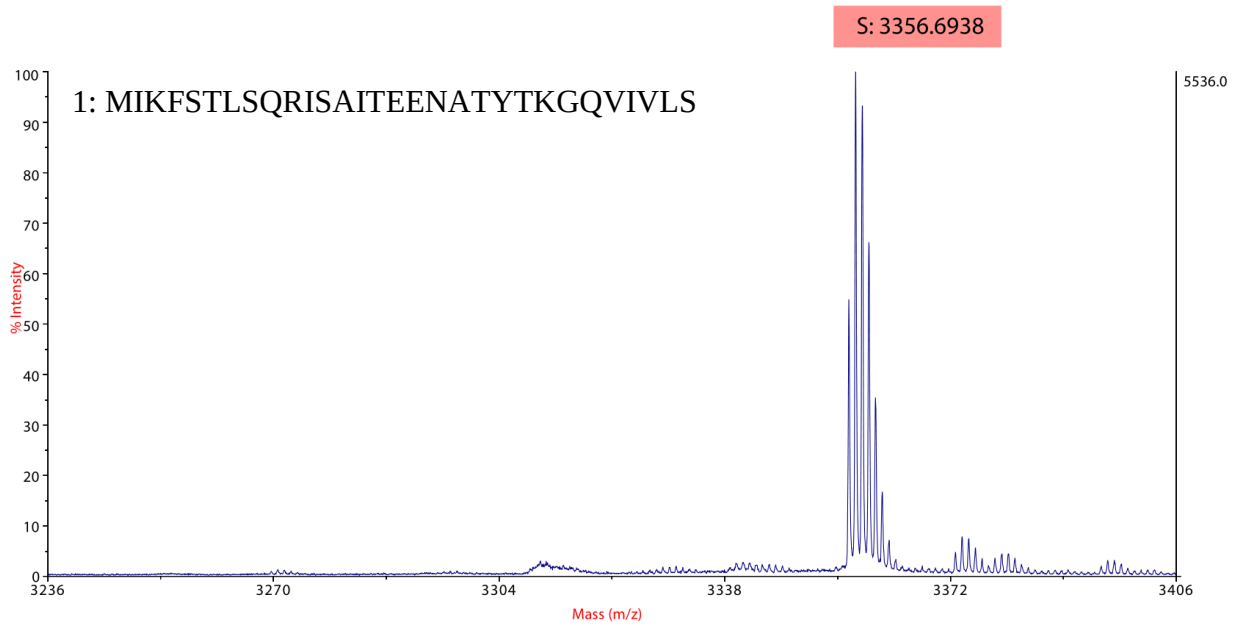
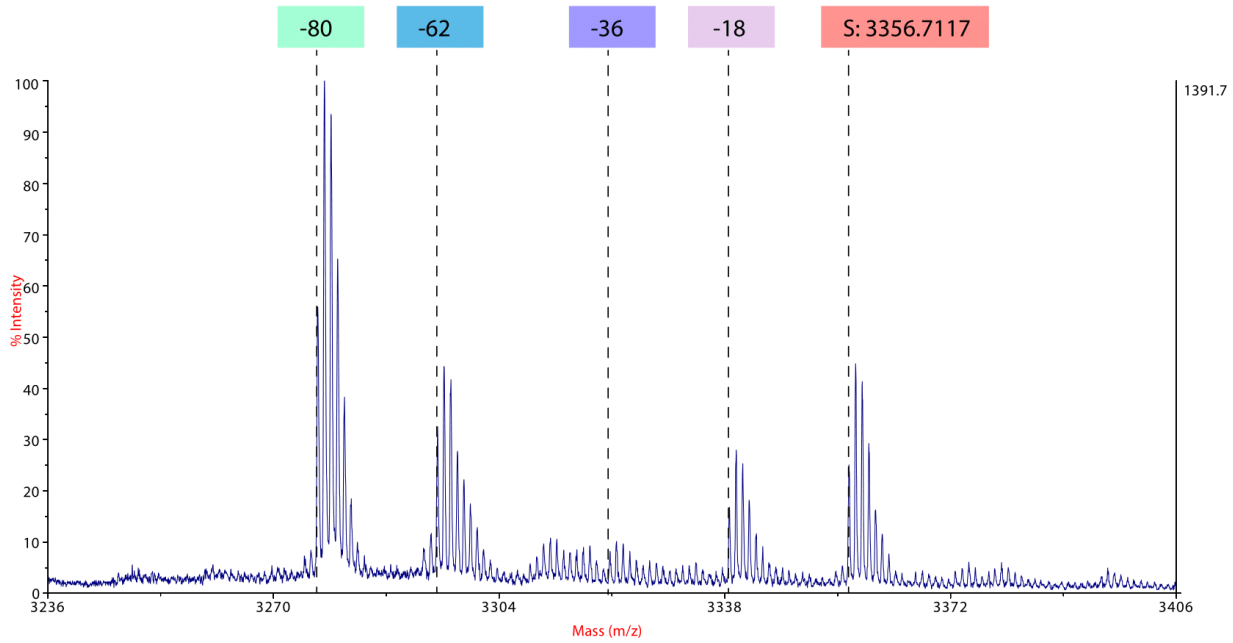
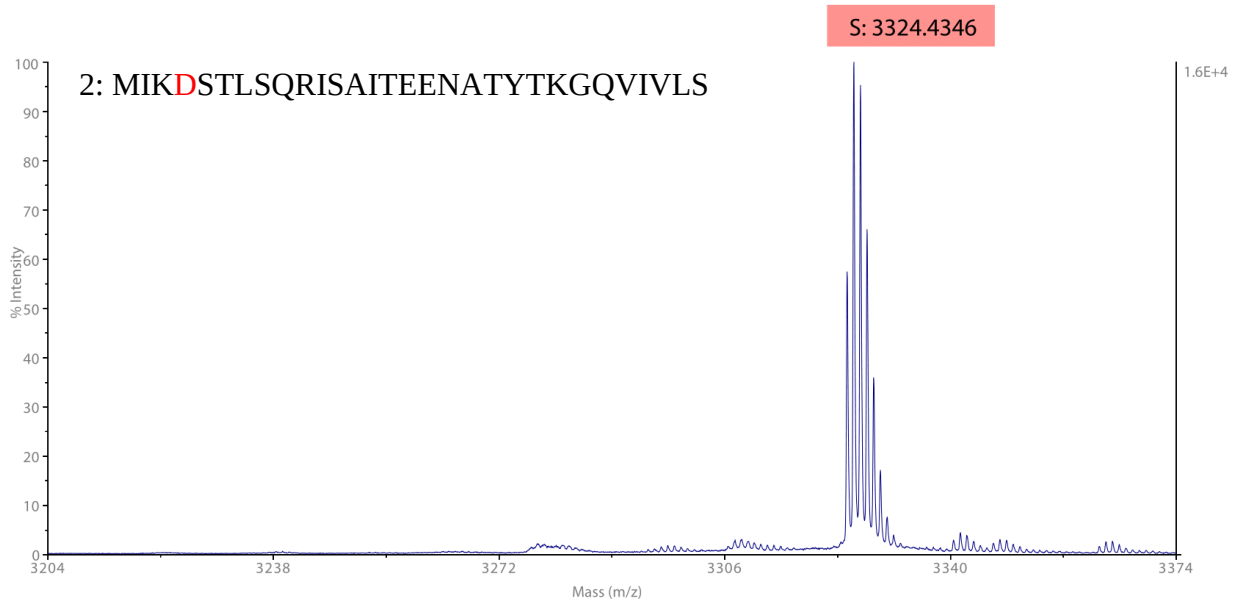
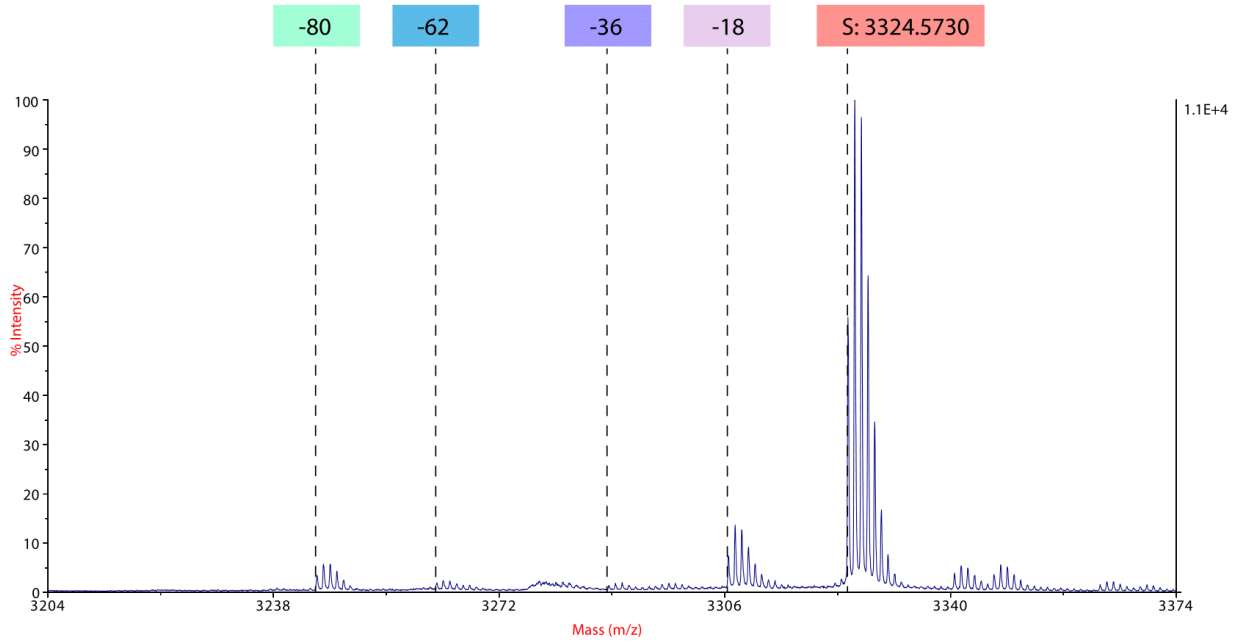


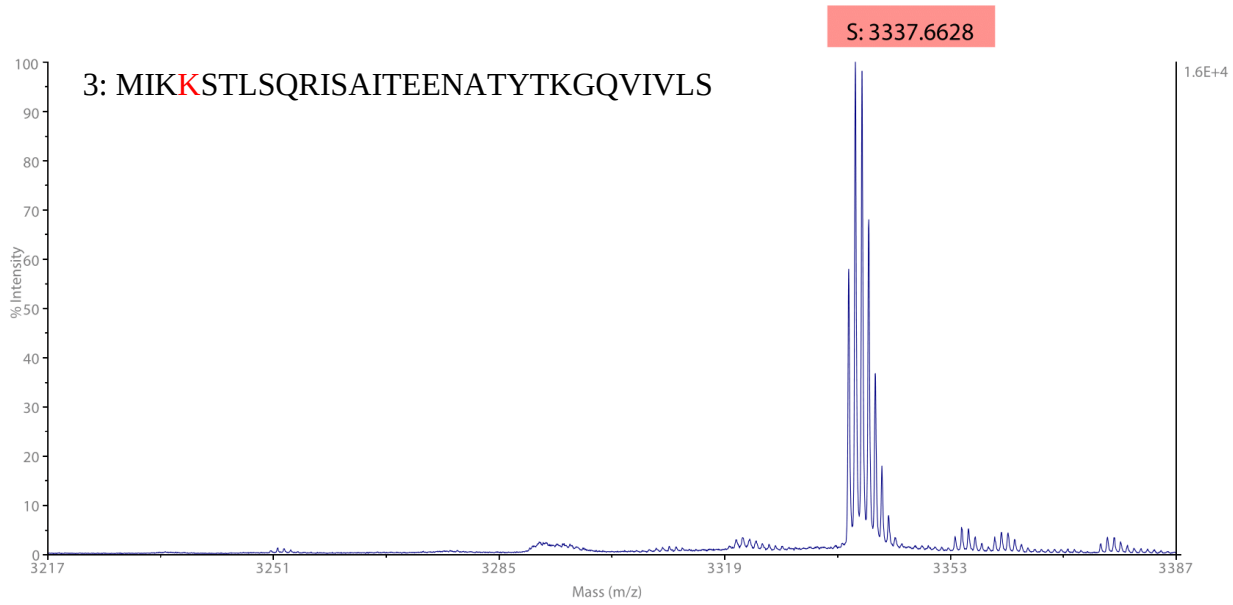
Figure A.9. MALDI-TOF Traces. PaaA substrates were transcribed and translated using NEB PURExpress and treated with 1 μM PaaA, 4 mM $MgCl_2$, 1 mM ATP and 100 mM Bicine pH 9.0 for 22.5 min before quenching with 50 mM EDTA. Reactions were desalted with C18 stage tips and analyzed by MALDI-TOF. For expected and observed masses see Table A.1. For each peptide, top traces are translation controls and bottom traces are analysis of the PaaA reaction. All masses are measured with a formylated N-terminus $[M+H]^{+1}$. S = substrate, -18 = - H_2O; -36 = -2 H_2O; -62 = -1 H_2O, -1 CO_2; - 80 = -2 H_2O, -1 CO_2 (final product).

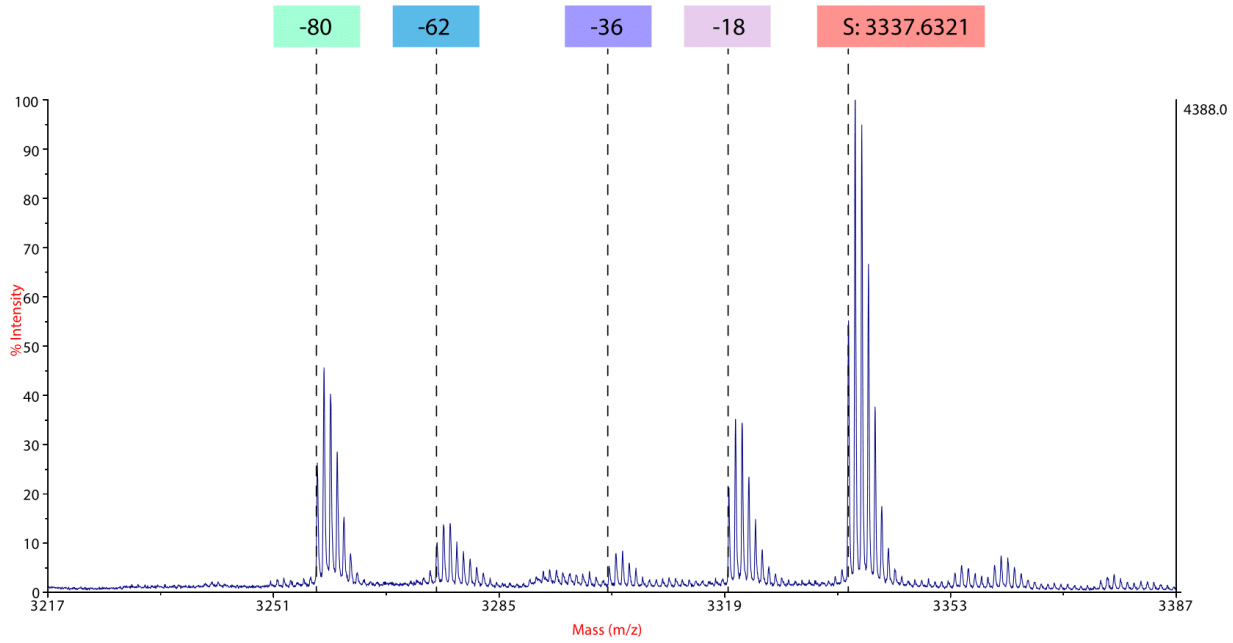

S: 3356.6938
1: MIKFSTLSQRISAITEENATYTKGQVIVLS
5536.0
% Intensity
Mass (m/z)
3236
3270
3304
3338
3372
3406

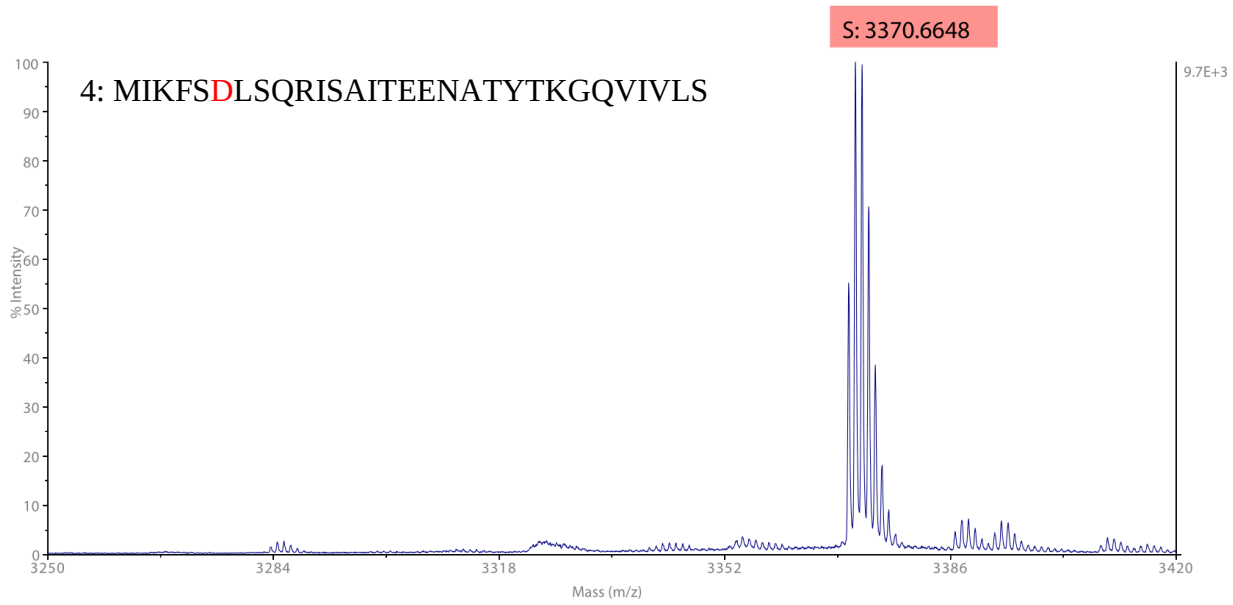

-80
-62
-36
-18
S: 3356.7117
1391.7
% Intensity
Mass (m/z)
3236
3270
3304
3338
3372
3406

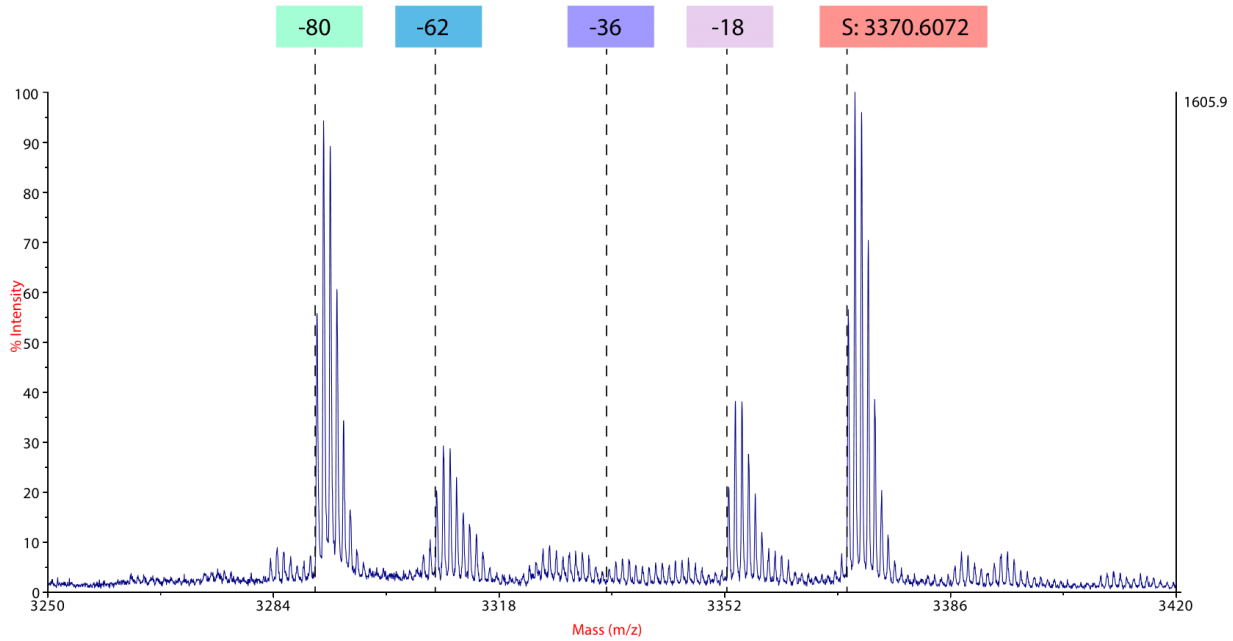

S: 3324.4346
2: MIKDSTLSQRISAITEENATYTKGQVIVLS
1.6E+4
% Intensity
Mass (m/z)
3204
3238
3272
3306
3340
3374

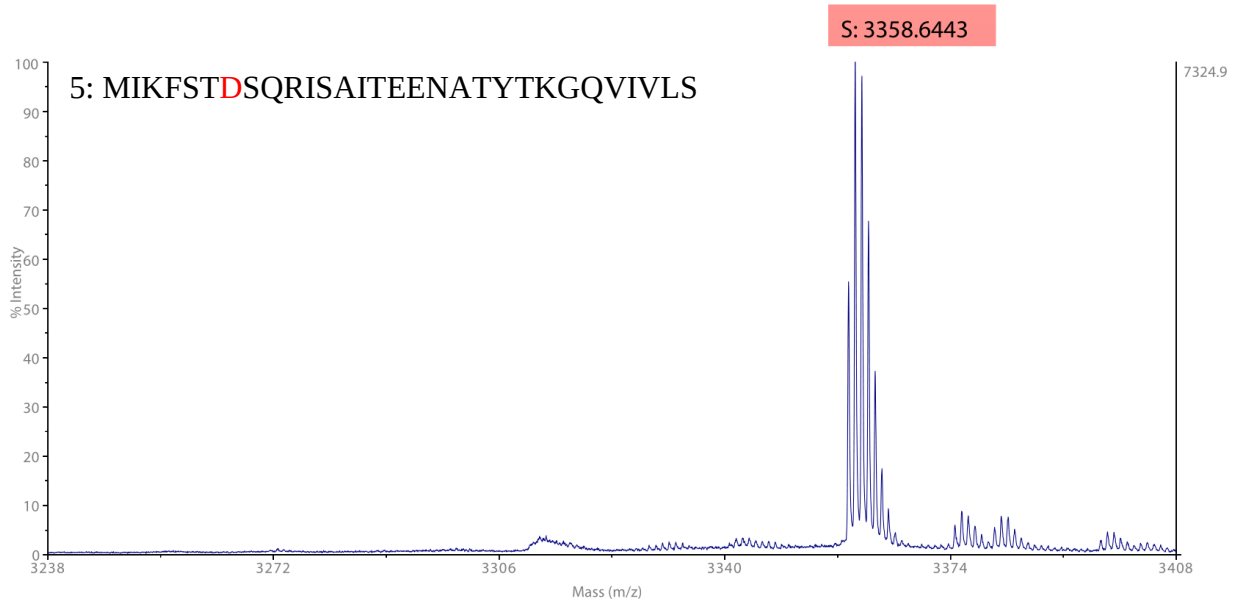

-80
-62
-36
-18
S: 3324.5730
1.1E+4
% Intensity
Mass (m/z)
3204
3238
3272
3306
3340
3374

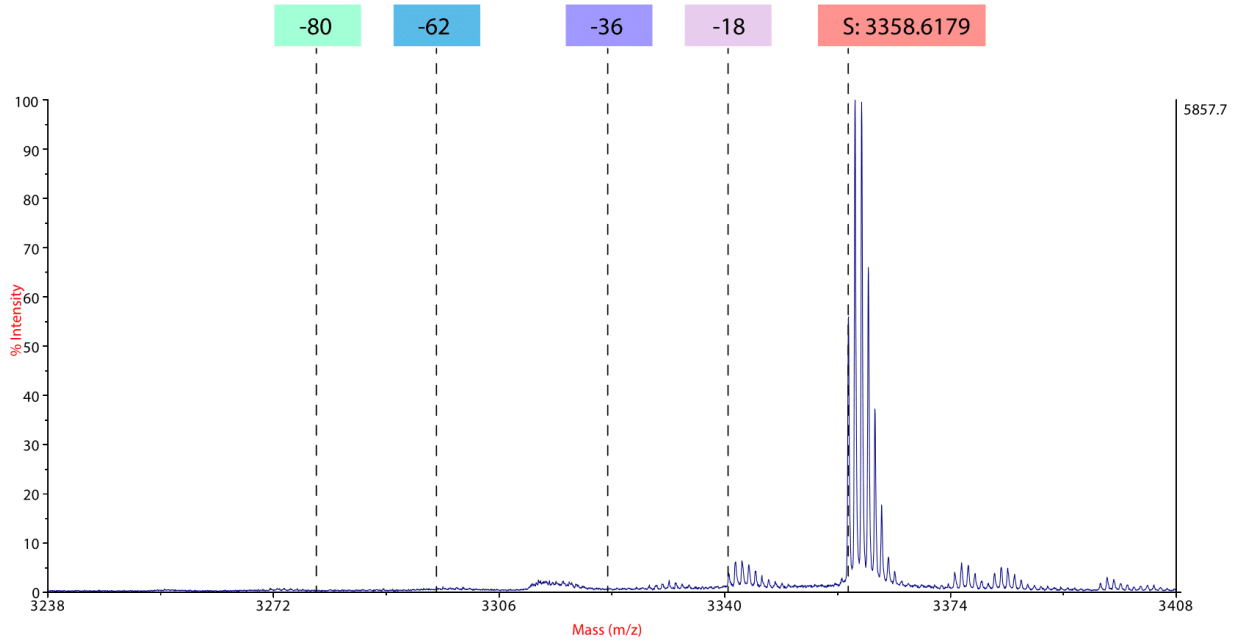

S: 3337.6628
3: MIKKSTLSQRISAITEENATYTKGQVIVLS
1.6E+4
% Intensity
Mass (m/z)
3217
3251
3285
3319
3353
3387

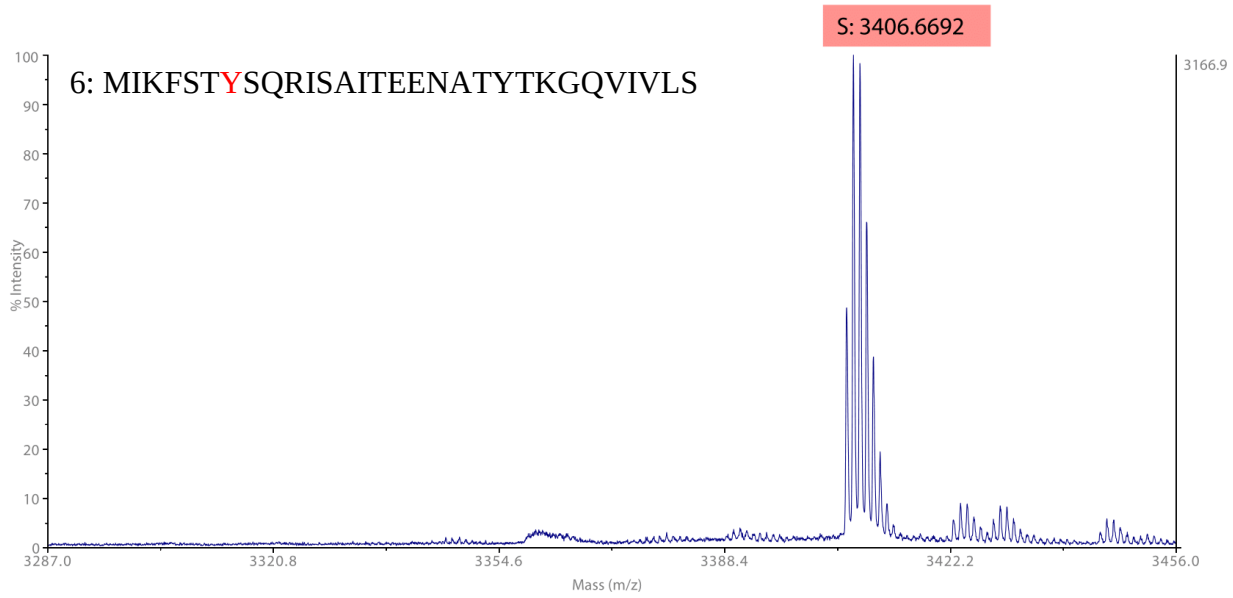

-80
-62
-36
-18
S: 3337.6321
4388.0
% Intensity
Mass (m/z)
3217
3251
3285
3319
3353
3387

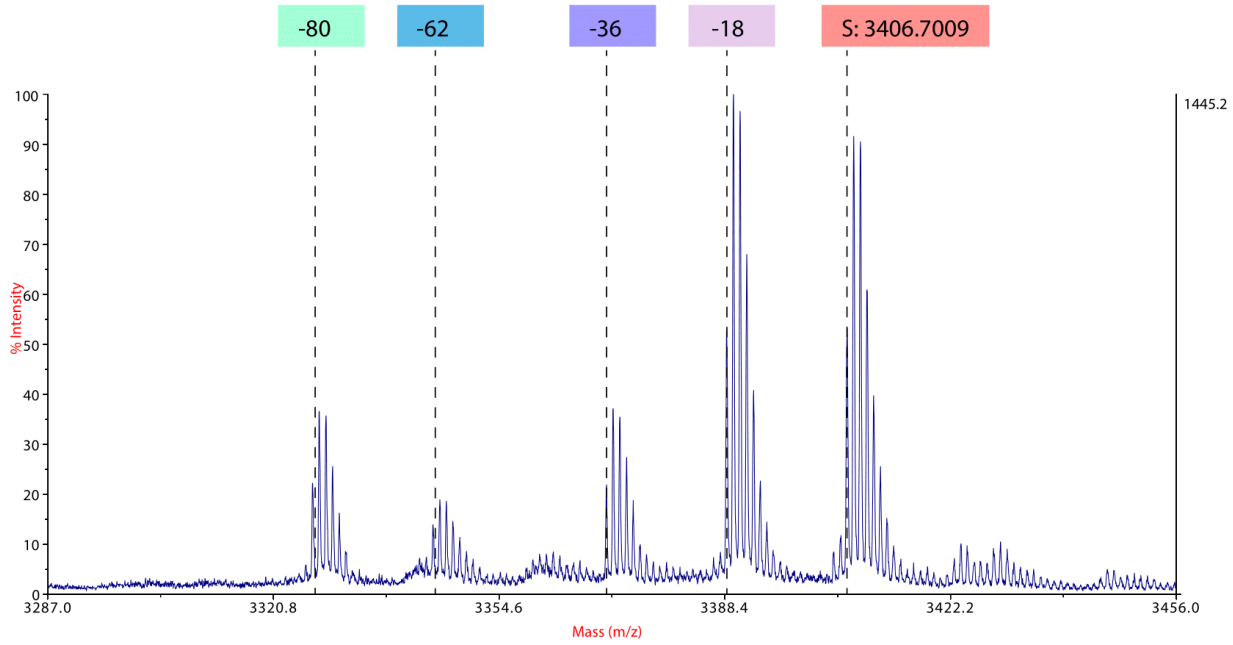

S: 3370.6648
4: MIKFSDLSQRISAITEENATYTKGQVIVLS
9.7E+3
% Intensity
Mass (m/z)
3250
3284
3318
3352
3386
3420


-80
-62
-36
-18
S: 3370.6072
1605.9
% Intensity
Mass (m/z)
3250
3284
3318
3352
3386
3420

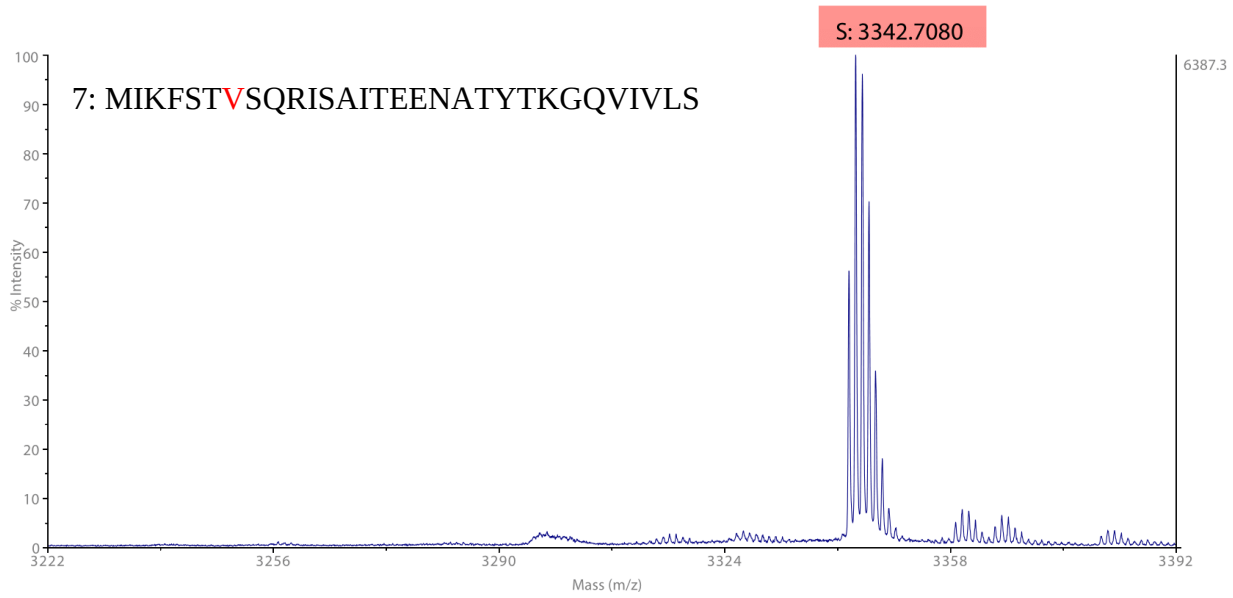

S: 3358.6443
5: MIKFSTDSQRISAITEENATYTKGQVIVLS
7324.9
% Intensity
Mass (m/z)
3238
3272
3306
3340
3374
3408

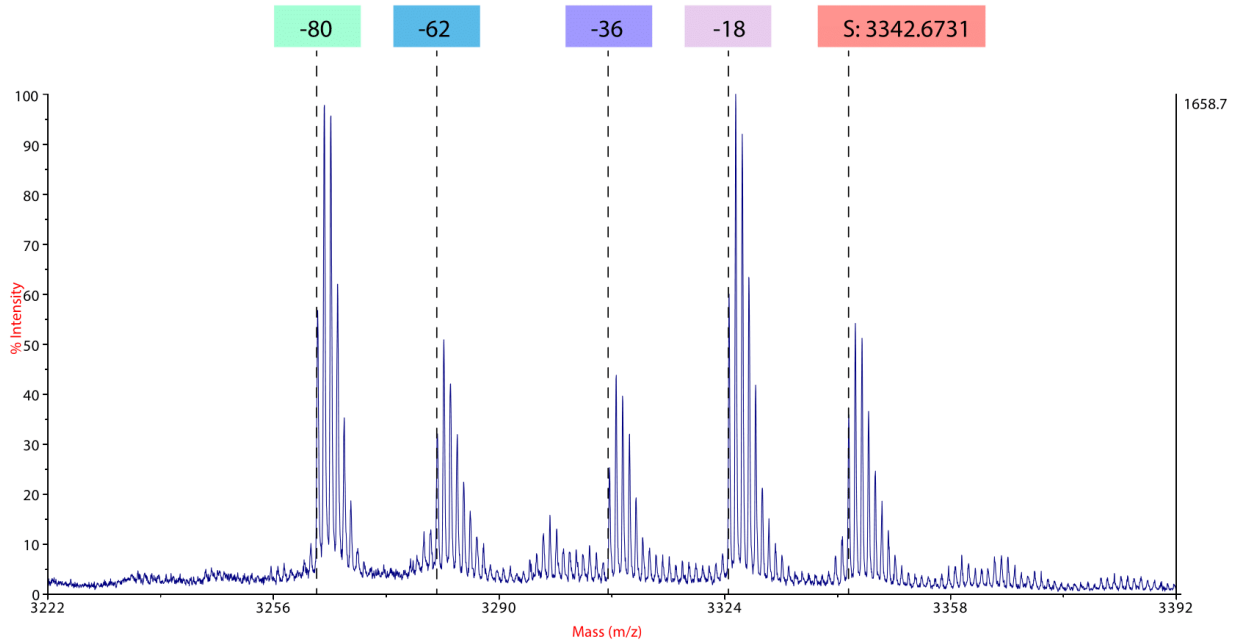

-80
-62
-36
-18
S: 3358.6179
5857.7
% Intensity
Mass (m/z)
3238
3272
3306
3340
3374
3408

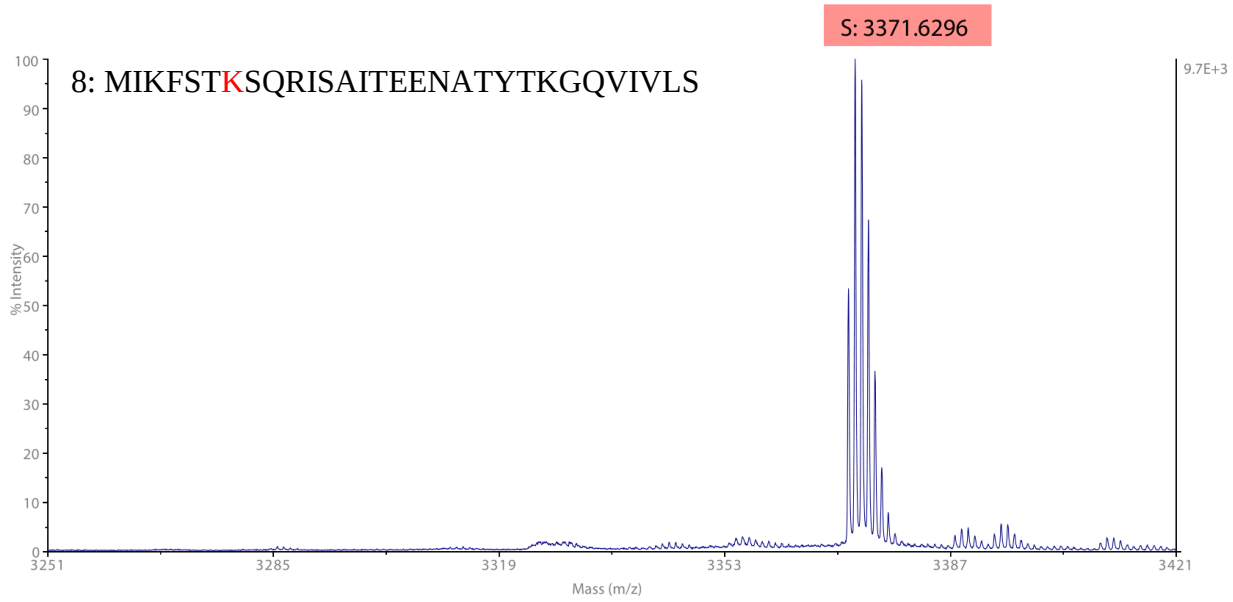

S: 3406.6692
6: MIKFSTYSQRISAITEENATYTKGQVIVLS
3166.9
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3287.0
3320.8
3354.6
3388.4
3422.2
3456.0
Mass (m/z)

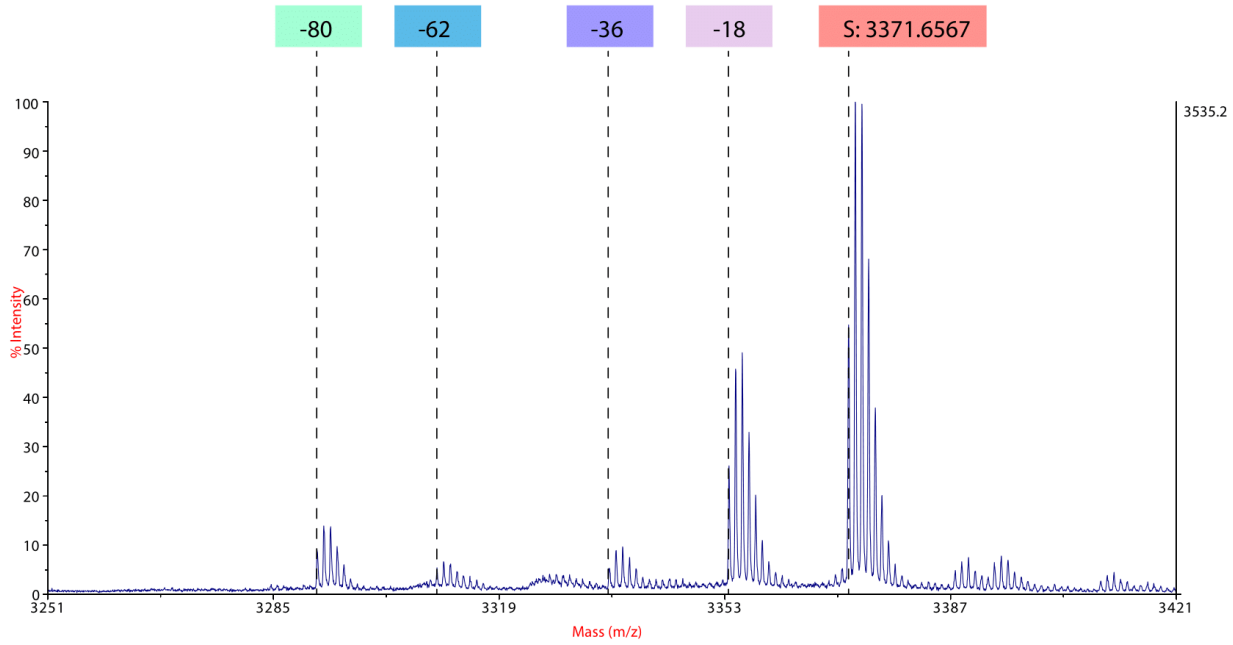

-80
-62
-36
-18
S: 3406.7009
1445.2
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3287.0
3320.8
3354.6
3388.4
3422.2
3456.0
Mass (m/z)

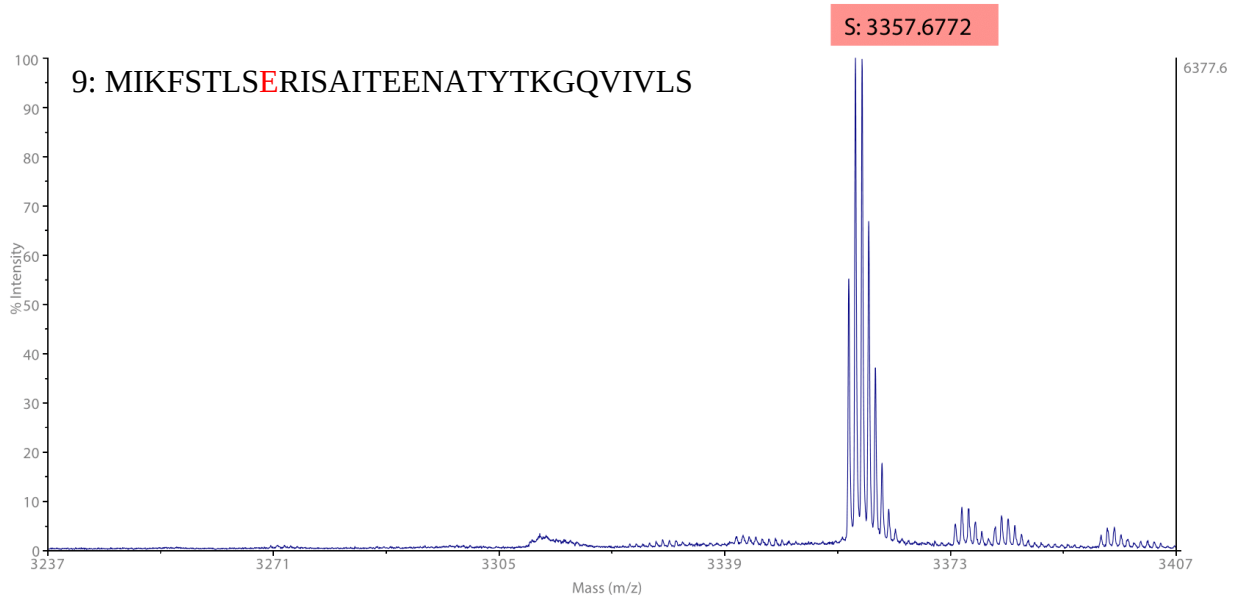
-

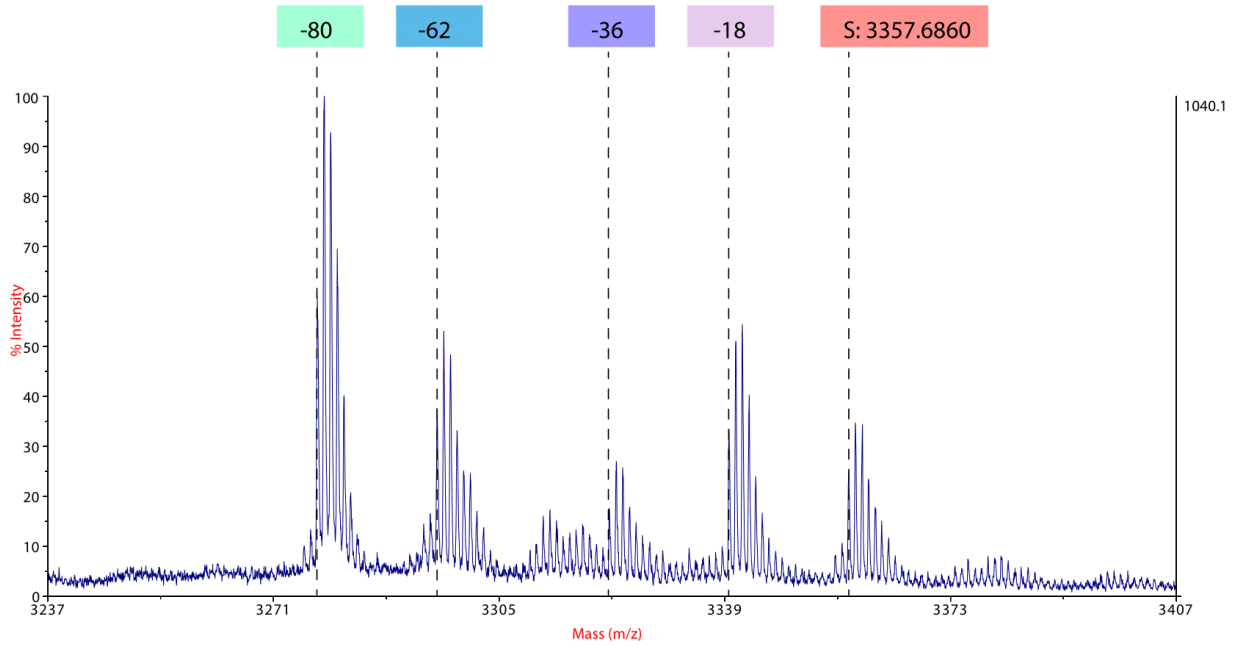

S: 3342.7080
7: MIKFSTVSQRISAITEENATYTKGQVIVLS
6387.3
% Intensity
Mass (m/z)

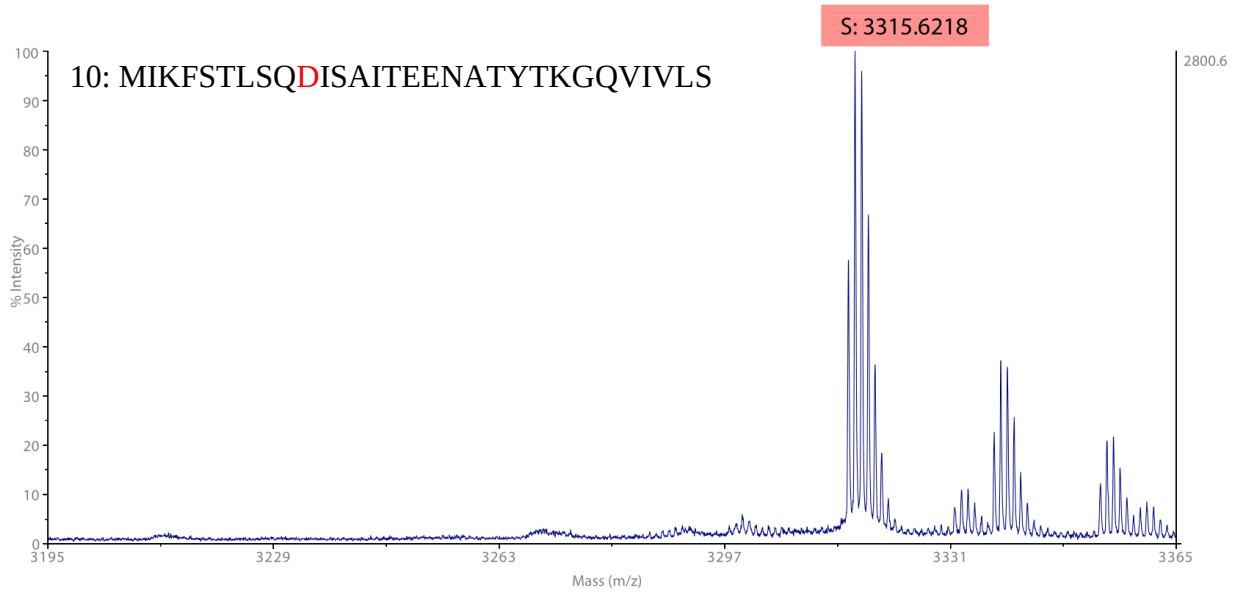

-80
-62
-36
-18
S: 3342.6731
1658.7
% Intensity
Mass (m/z)

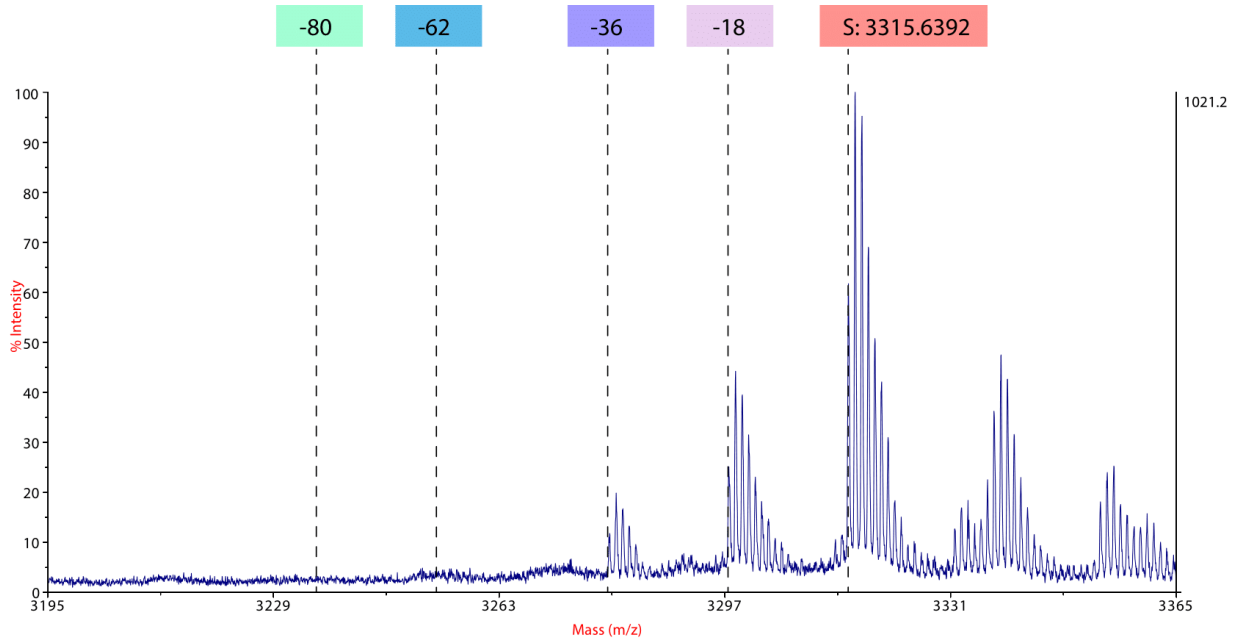

S: 3371.6296
8: MIKFSTKSQRISAITEENATYTKGQVIVLS
9.7E+3
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3251
3285
3319
3353
3387
3421
Mass (m/z)

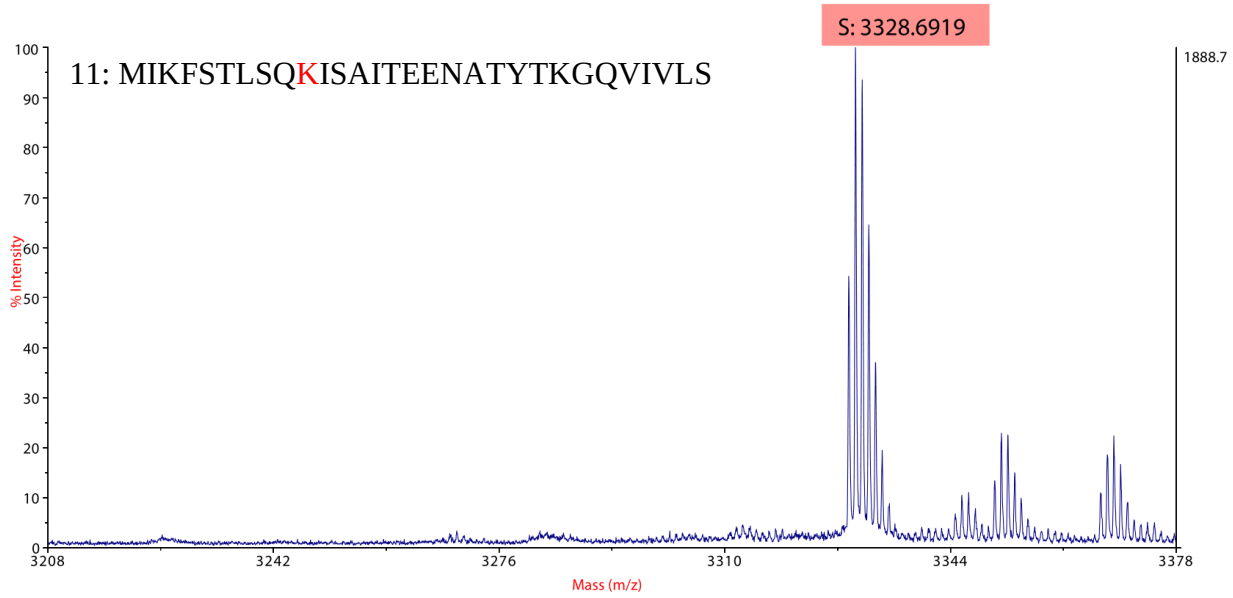

-80
-62
-36
-18
S: 3371.6567
3535.2
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3251
3285
3319
3353
3387
3421
Mass (m/z)

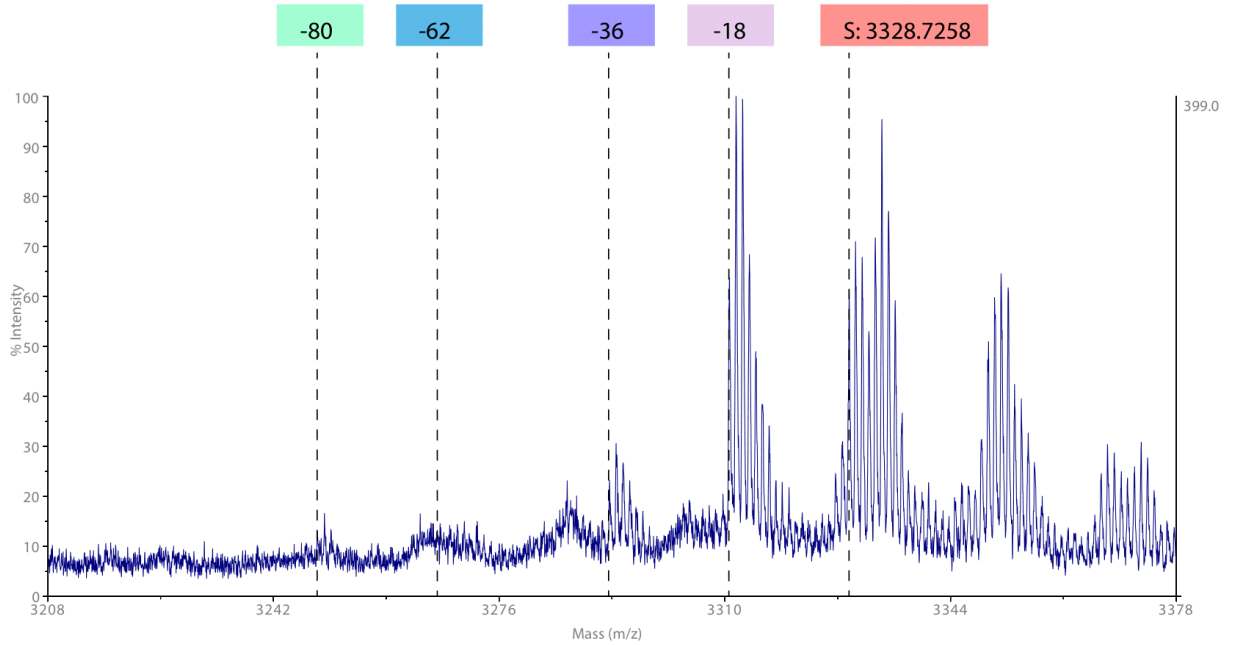

S: 3357.6772
9: MIKFSTLSERISAITEENATYTKGQVIVLS
6377.6
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3237
3271
3305
3339
3373
3407
Mass (m/z)

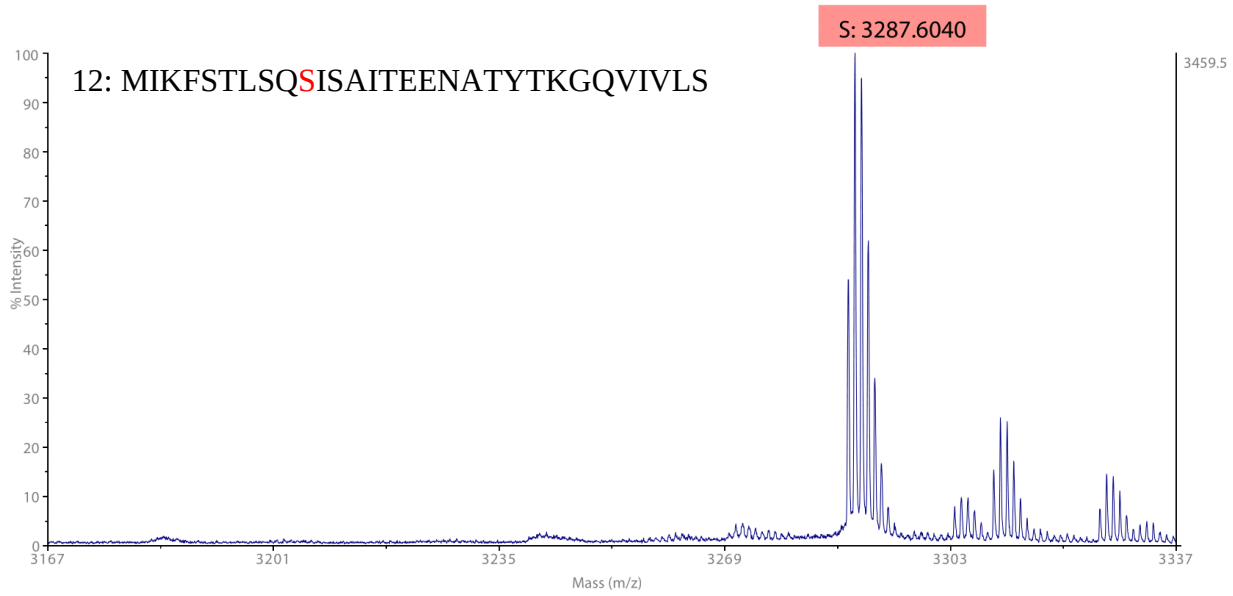

-80
-62
-36
-18
S: 3357.6860
1040.1
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3237
3271
3305
3339
3373
3407
Mass (m/z)

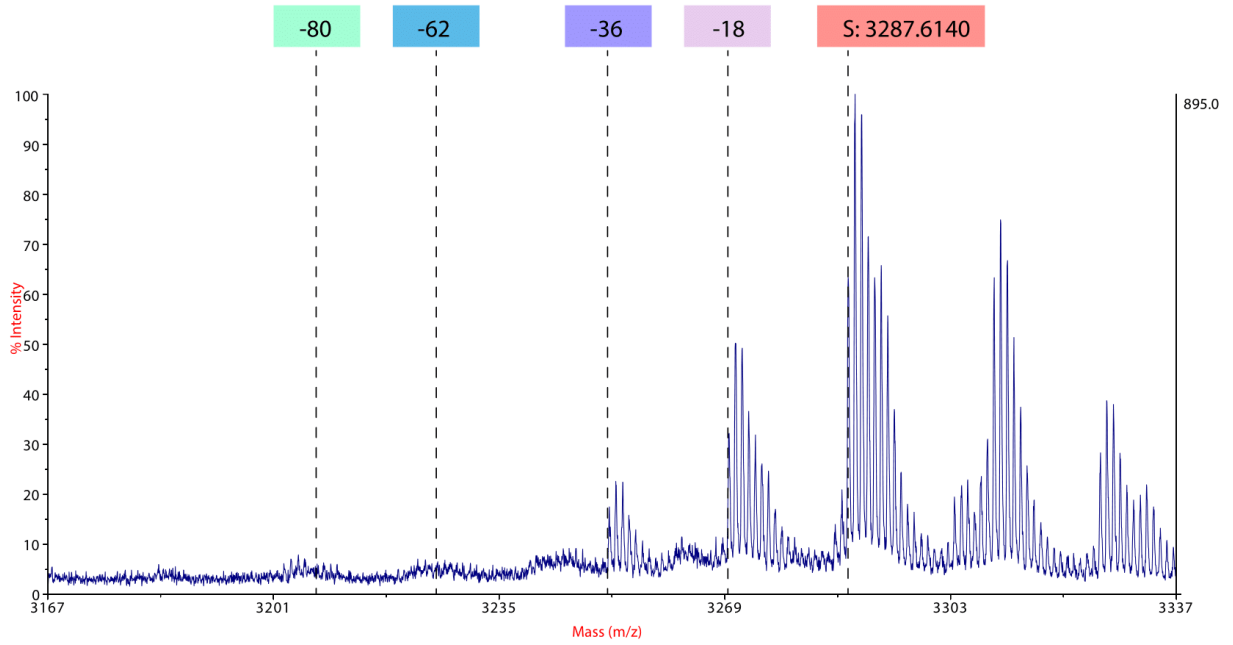

S: 3315.6218
10: MIKFSTLSQDISAITEENATYTKGQVIVLS
2800.6
% Intensity
Mass (m/z)

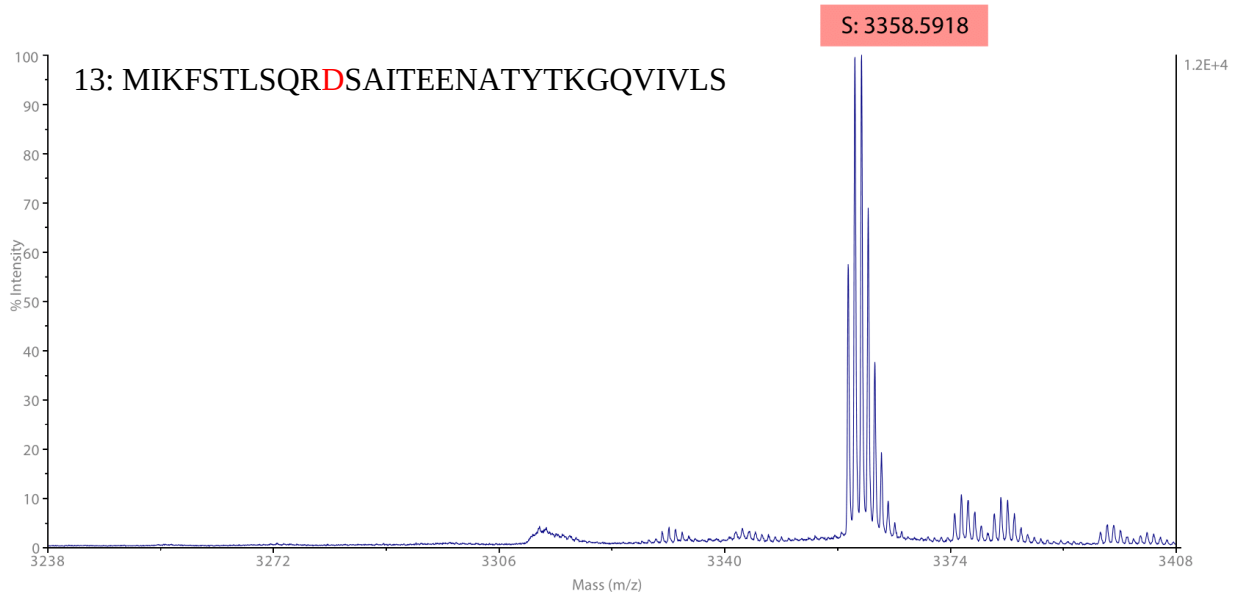

-80
-62
-36
-18
S: 3315.6392
1021.2
% Intensity
Mass (m/z)

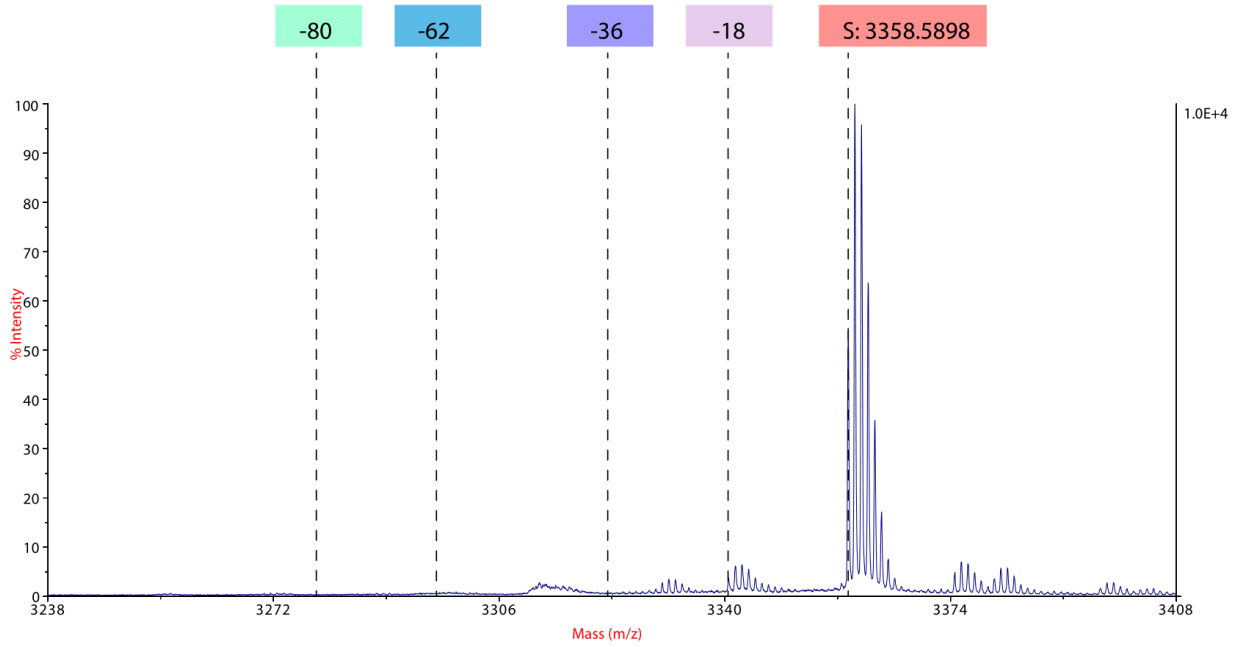

S: 3328.6919
11: MIKFSTLSQKISAITEENATYTKGQVIVLS
1888.7
% Intensity
Mass (m/z)

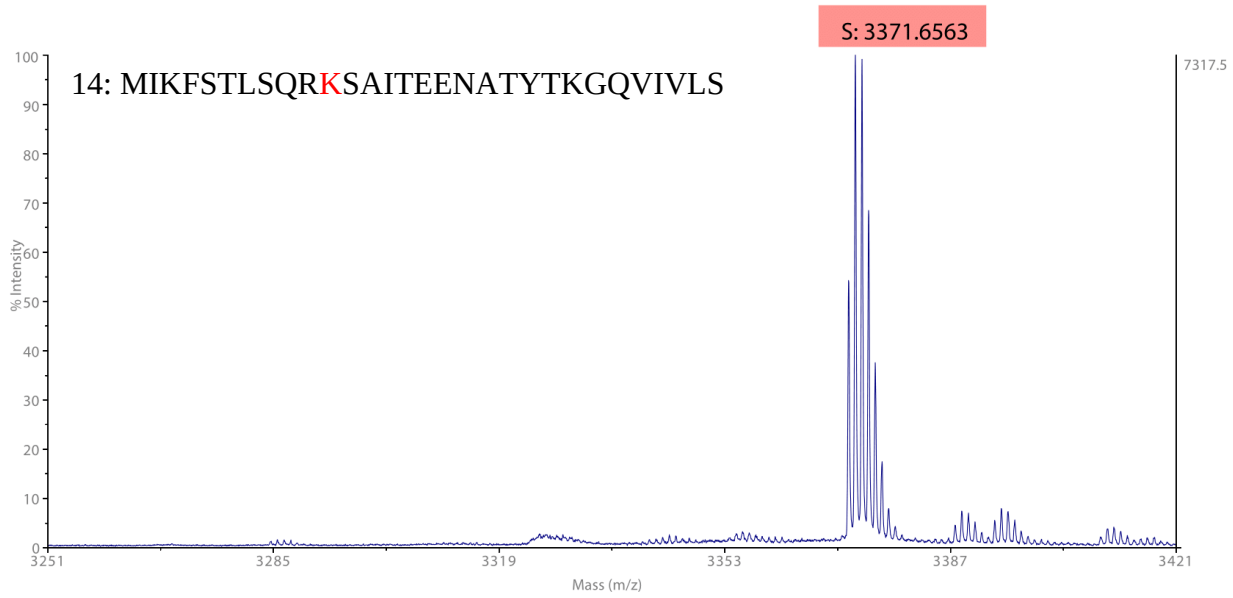

-80
-62
-36
-18
S: 3328.7258
399.0
% Intensity
Mass (m/z)

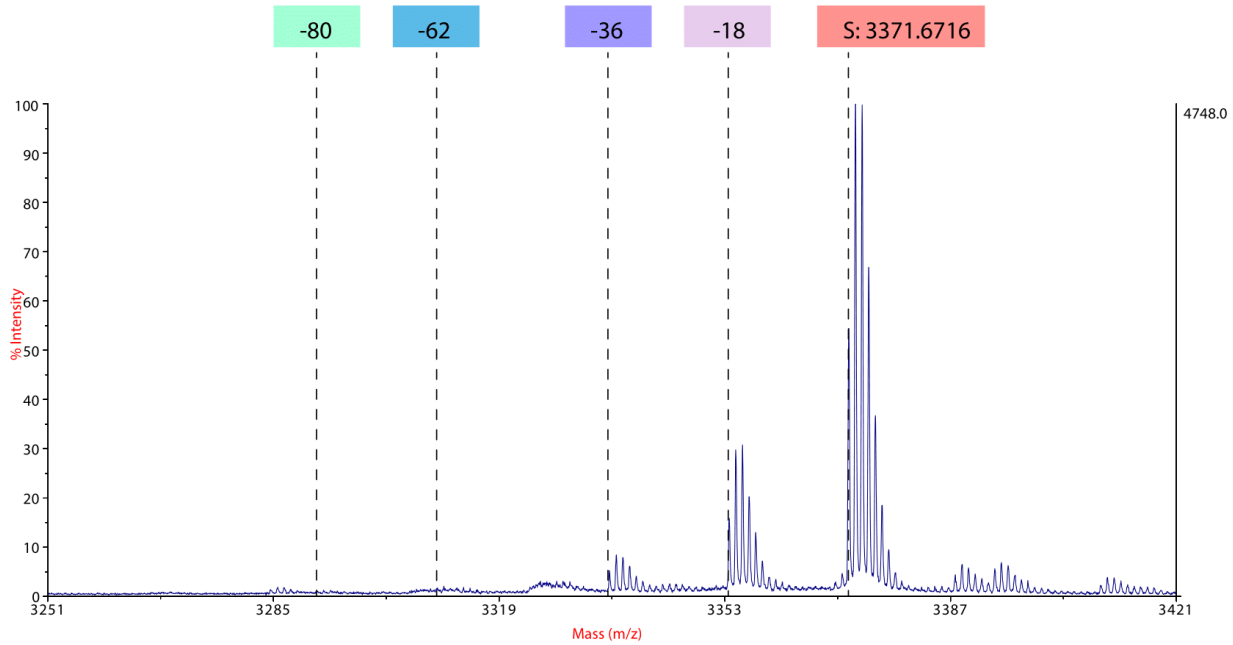

S: 3287.6040
12: MIKFSTLSQSISAITEENATYTKGQVIVLS
3459.5
% Intensity
Mass (m/z)
3167
3201
3235
3269
3303
3337

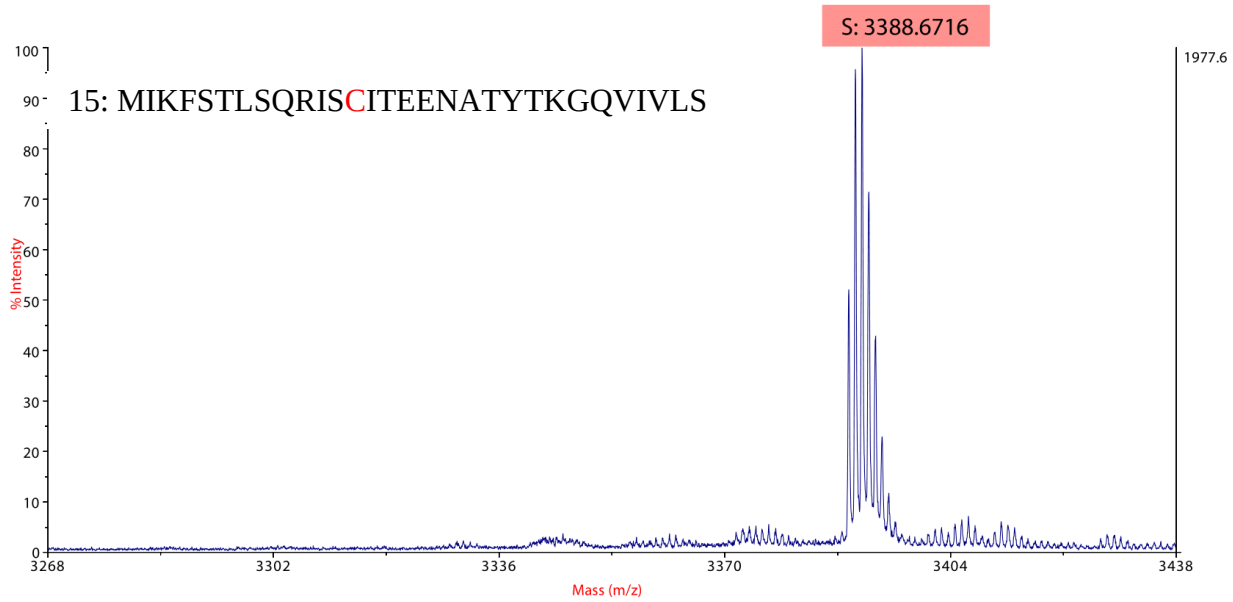

-80
-62
-36
-18
S: 3287.6140
895.0
% Intensity
Mass (m/z)
3167
3201
3235
3269
3303
3337

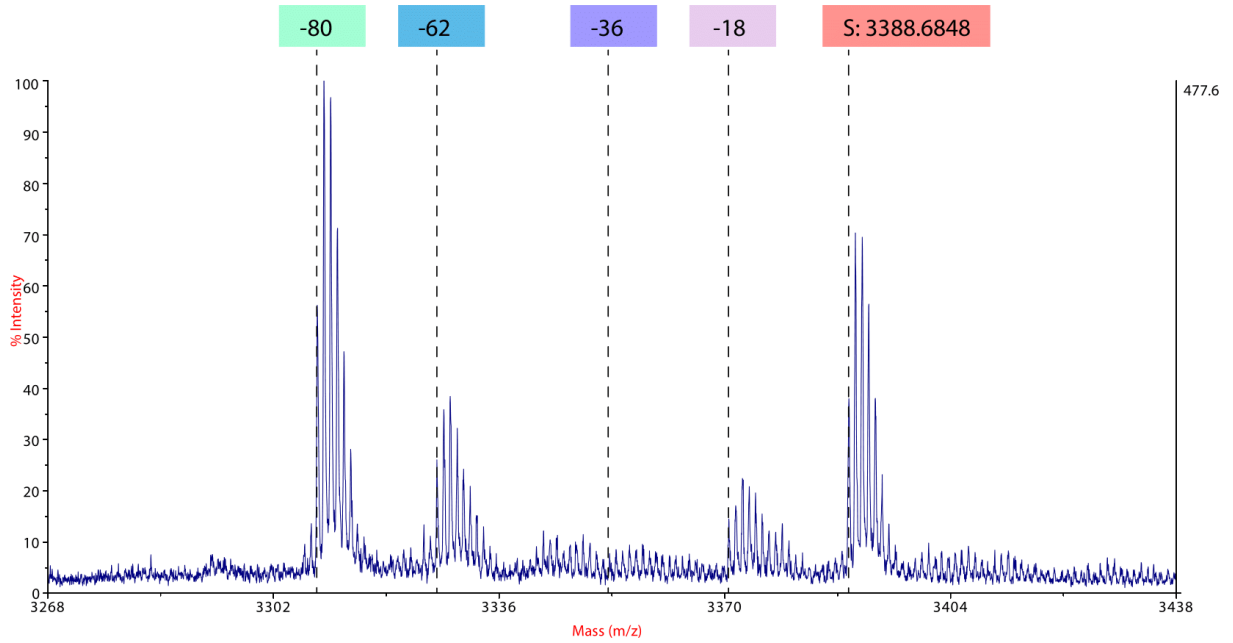

S: 3358.5918
13: MIKFSTLSQRDSAITEENATYTKGQVIVLS
1.2E+4
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3238
3272
3306
3340
3374
3408
Mass (m/z)

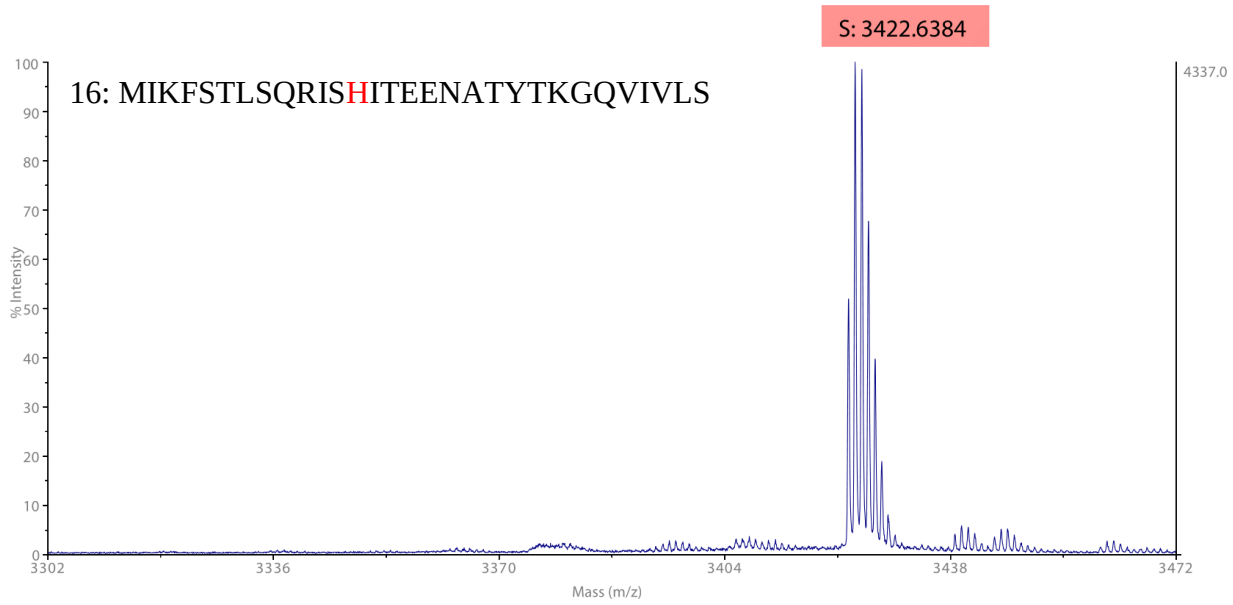

-80
-62
-36
-18
S: 3358.5898
1.0E+4
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3238
3272
3306
3340
3374
3408
Mass (m/z)

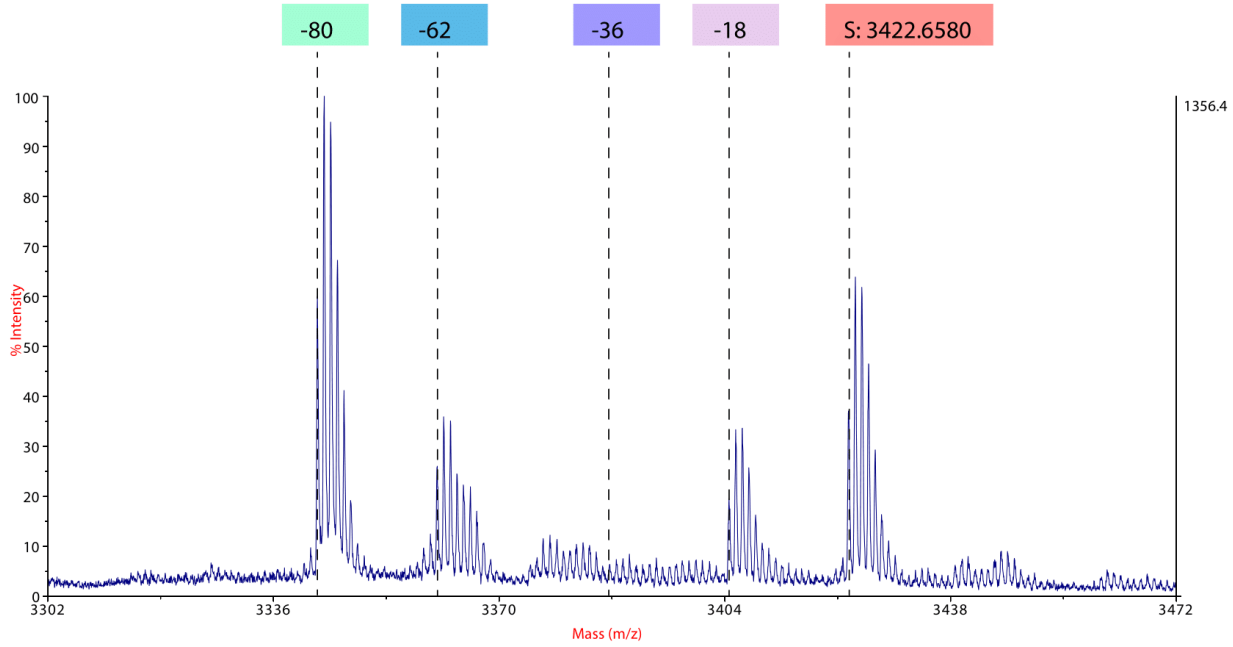

S: 3371.6563
14: MIKFSTLSQRKSAITEENATYTKGQVIVLS
7317.5
% Intensity
Mass (m/z)
3251
3285
3319
3353
3387
3421

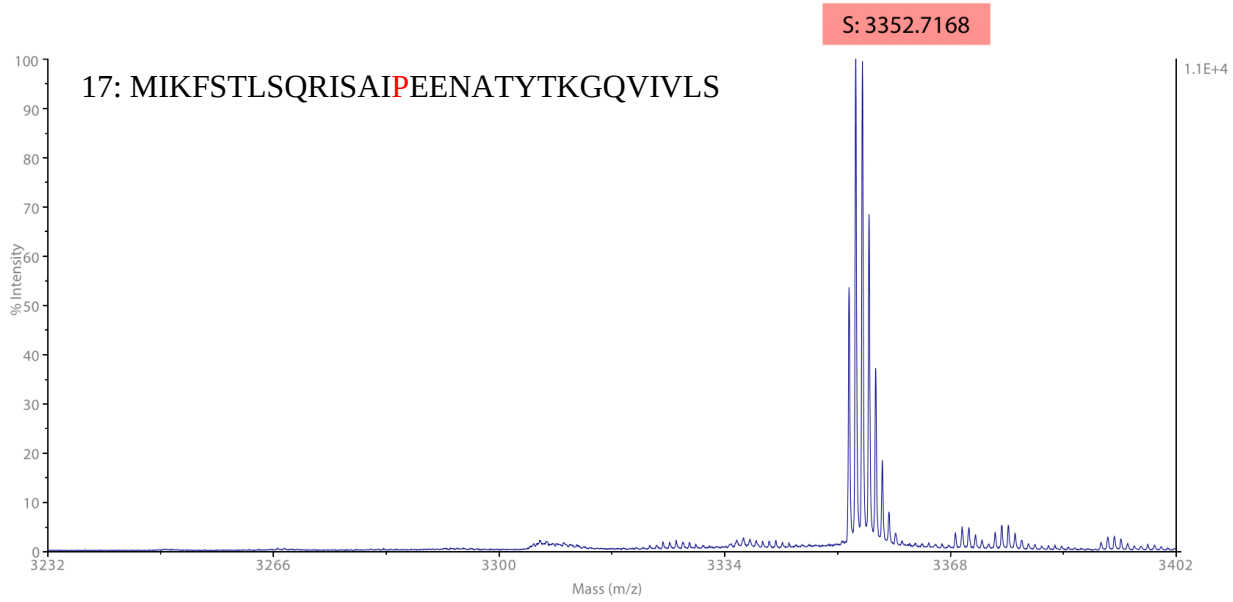

-80
-62
-36
-18
S: 3371.6716
4748.0
% Intensity
Mass (m/z)
3251
3285
3319
3353
3387
3421

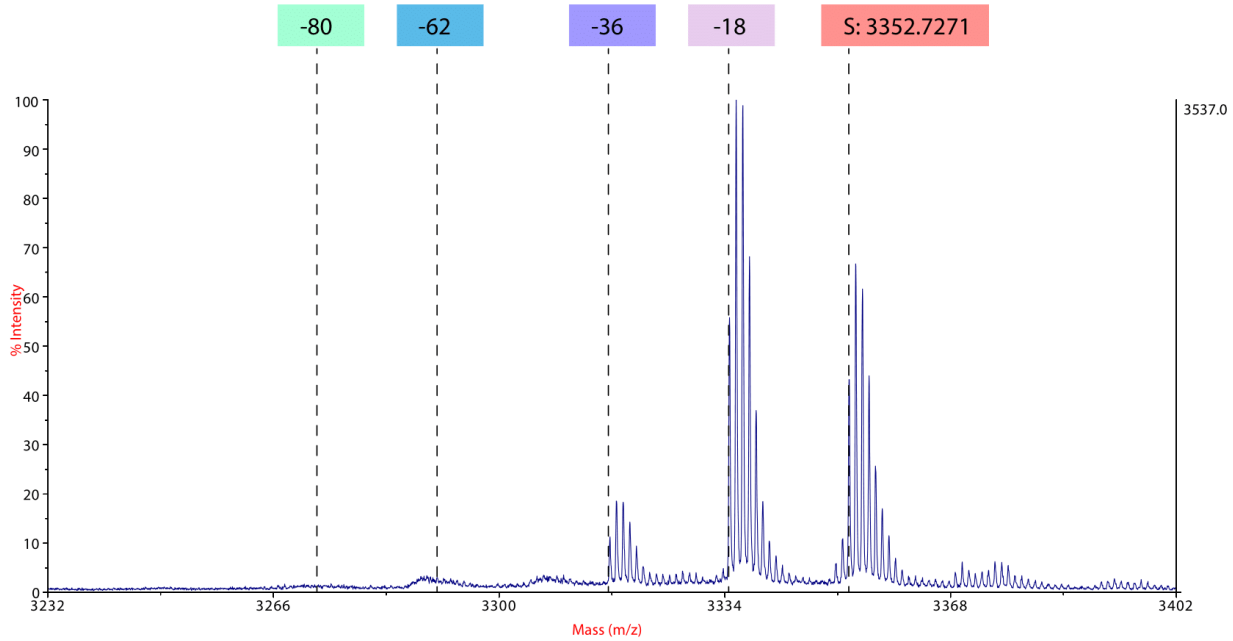

S: 3388.6716
15: MIKFSTLSQRISCITEENATYTKGQVIVLS
1977.6
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3268
3302
3336
3370
3404
3438
Mass (m/z)

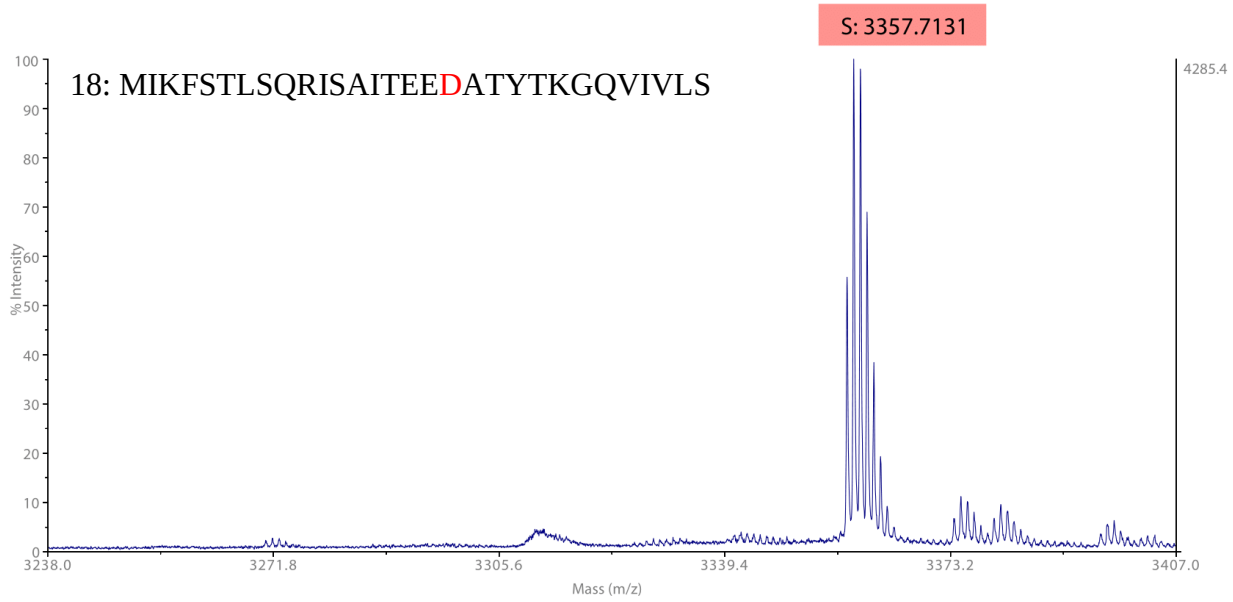

-80
-62
-36
-18
S: 3388.6848
477.6
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3268
3302
3336
3370
3404
3438
Mass (m/z)

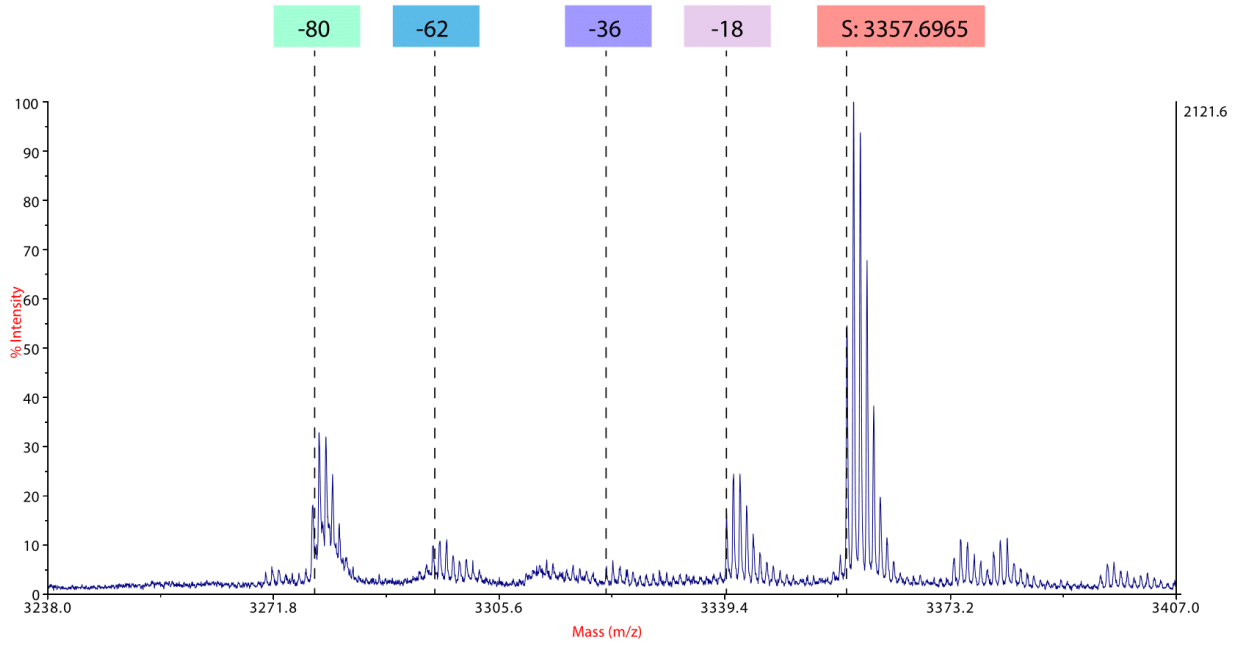

S: 3422.6384
16: MIKFSTLSQRISHITEENATYTKGQVIVLS
4337.0
% Intensity
Mass (m/z)
3302
3336
3370
3404
3438
3472

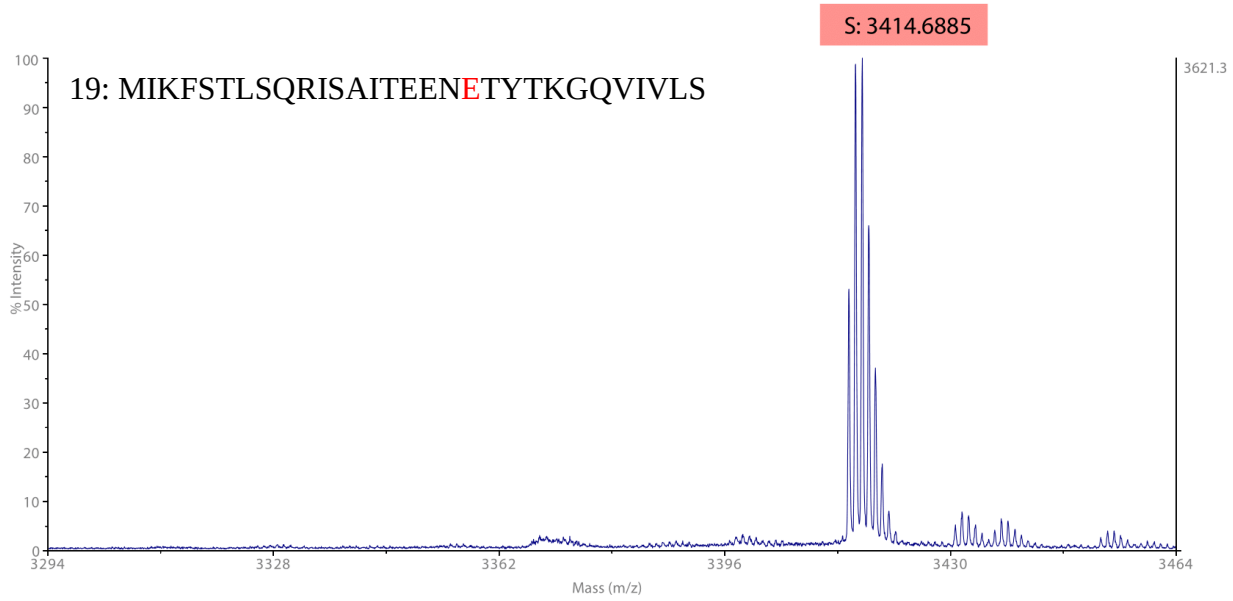

-80
-62
-36
-18
S: 3422.6580
1356.4
% Intensity
Mass (m/z)
3302
3336
3370
3404
3438
3472

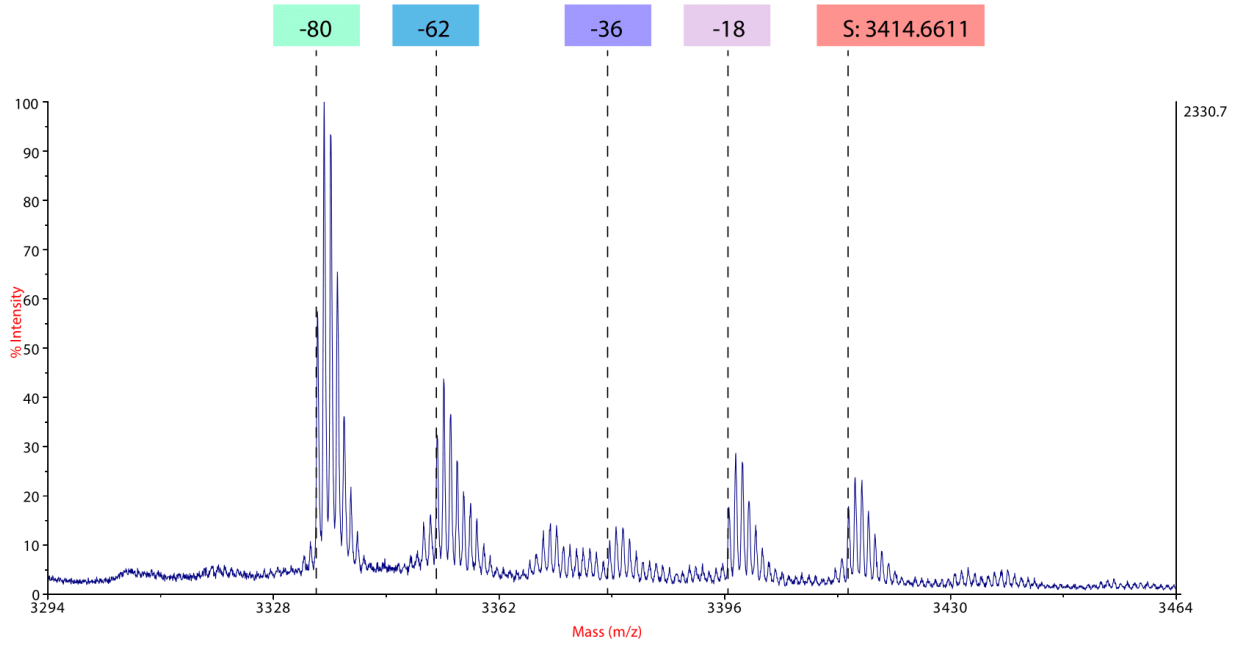

S: 3352.7168
17: MIKFSTLSQRISAIPEENATYTKGQVIVLS
1.1E+4
% Intensity
Mass (m/z)

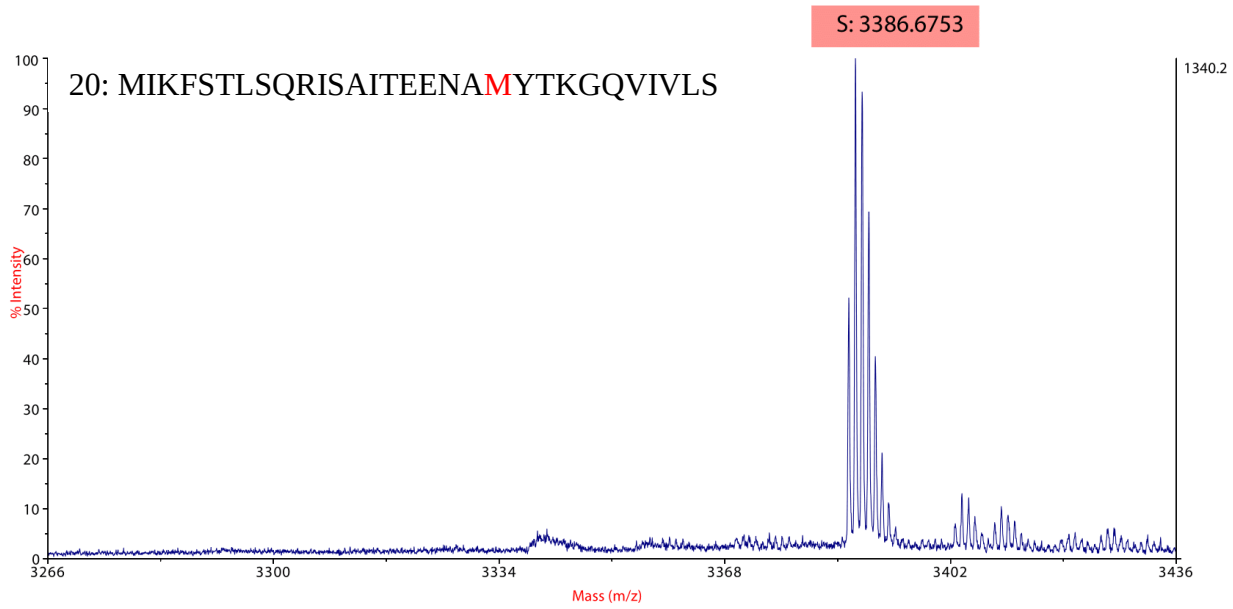

-80
-62
-36
-18
S: 3352.7271
3537.0
% Intensity
Mass (m/z)

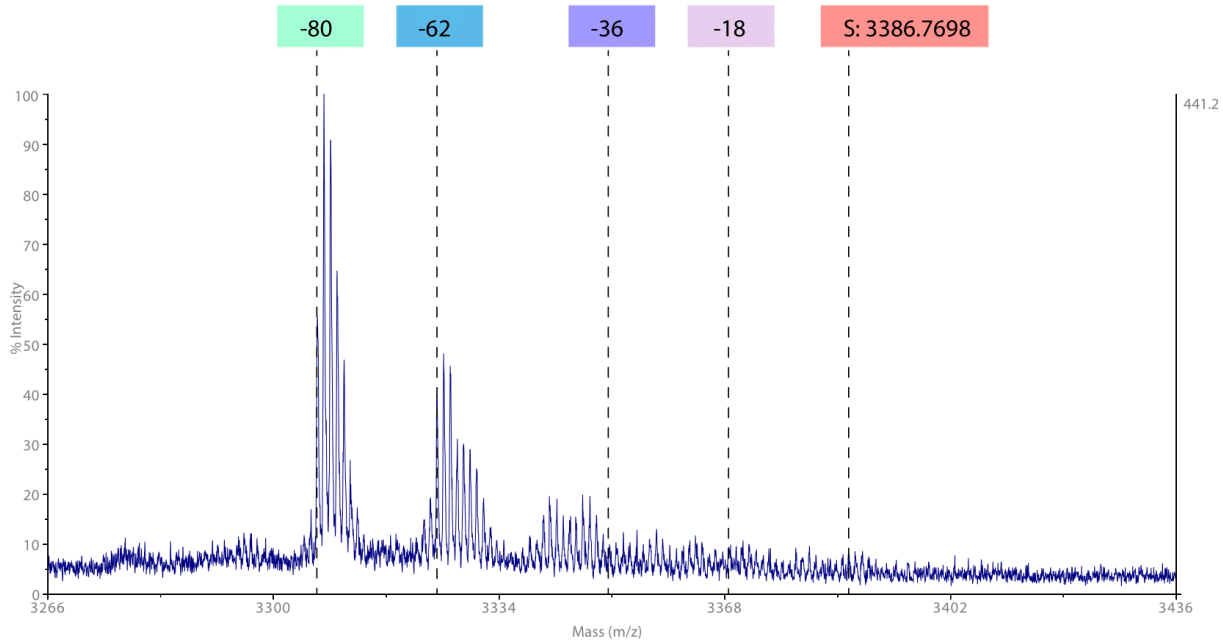

S: 3357.7131
18: MIKFSTLSQRISAITEEDATYTKGQVIVLS
4285.4
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3238.0
3271.8
3305.6
3339.4
3373.2
3407.0
Mass (m/z)

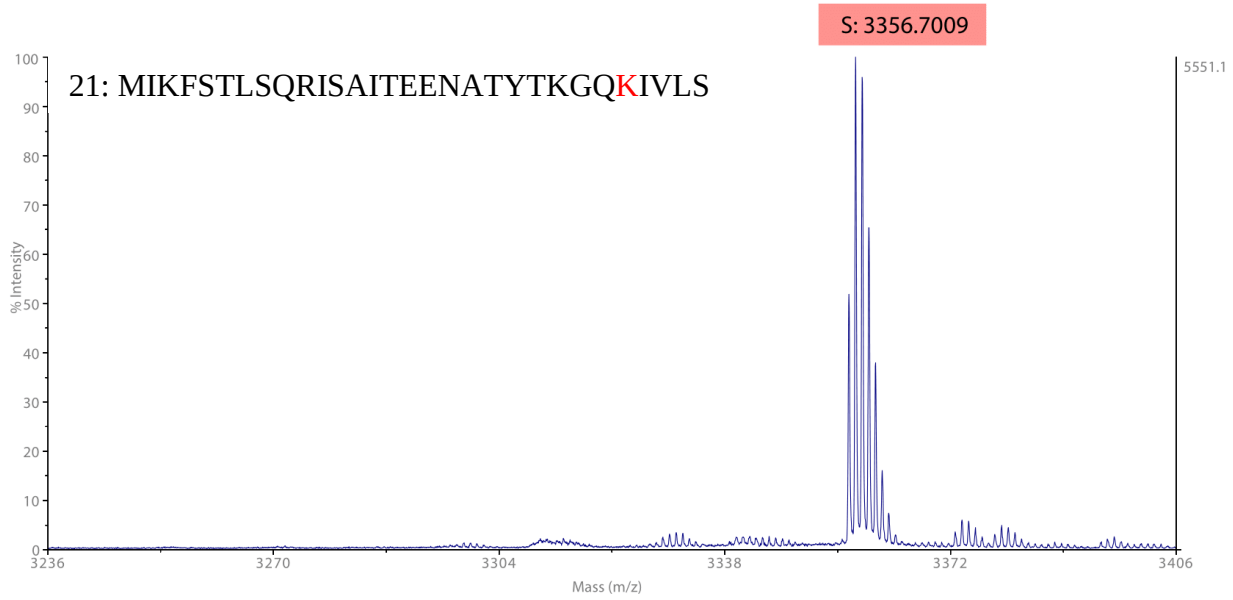

-80
-62
-36
-18
S: 3357.6965
2121.6
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3238.0
3271.8
3305.6
3339.4
3373.2
3407.0
Mass (m/z)

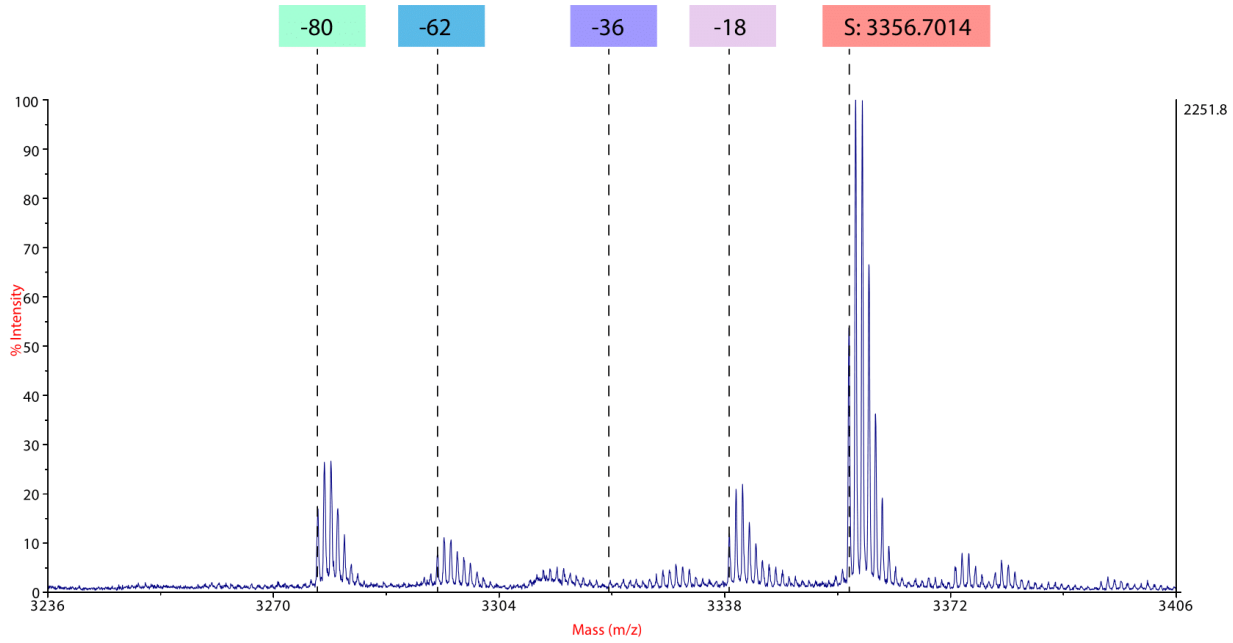

S: 3414.6885
19: MIKFSTLSQRISAITEENETYTKGQVIVLS
3621.3
% Intensity
Mass (m/z)
3294
3328
3362
3396
3430
3464

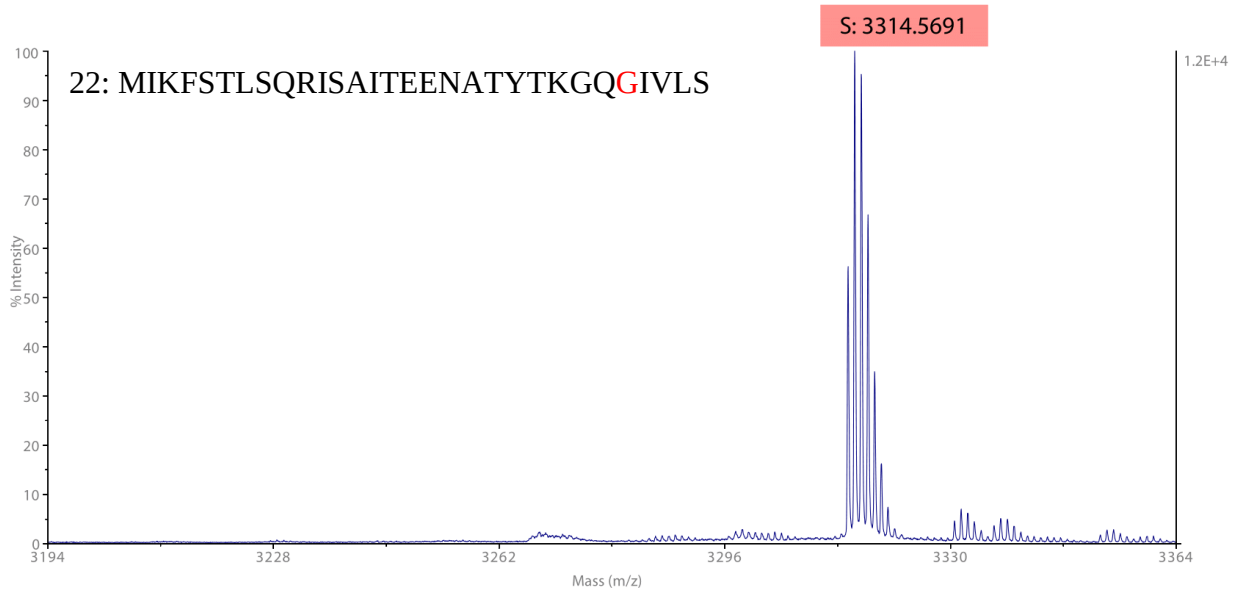

-80
-62
-36
-18
S: 3414.6611
2330.7
% Intensity
Mass (m/z)
3294
3328
3362
3396
3430
3464

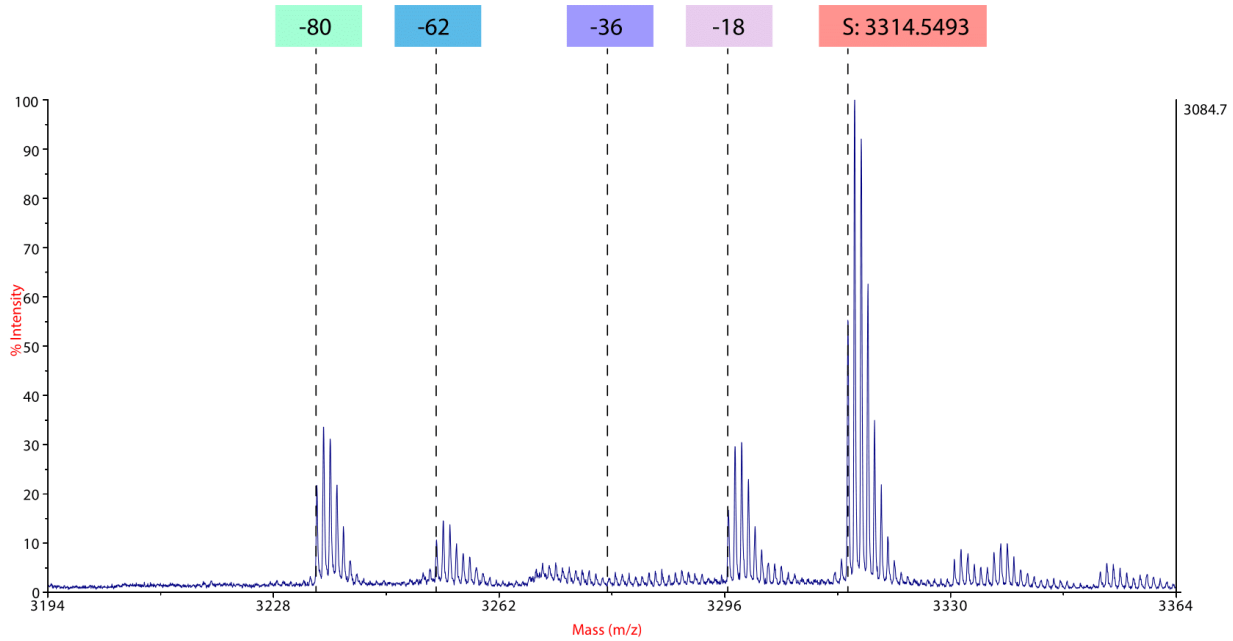

S: 3386.6753
20: MIKFSTLSQRISAITEENAMYTKGQVIVLS
1340.2
% Intensity
Mass (m/z)

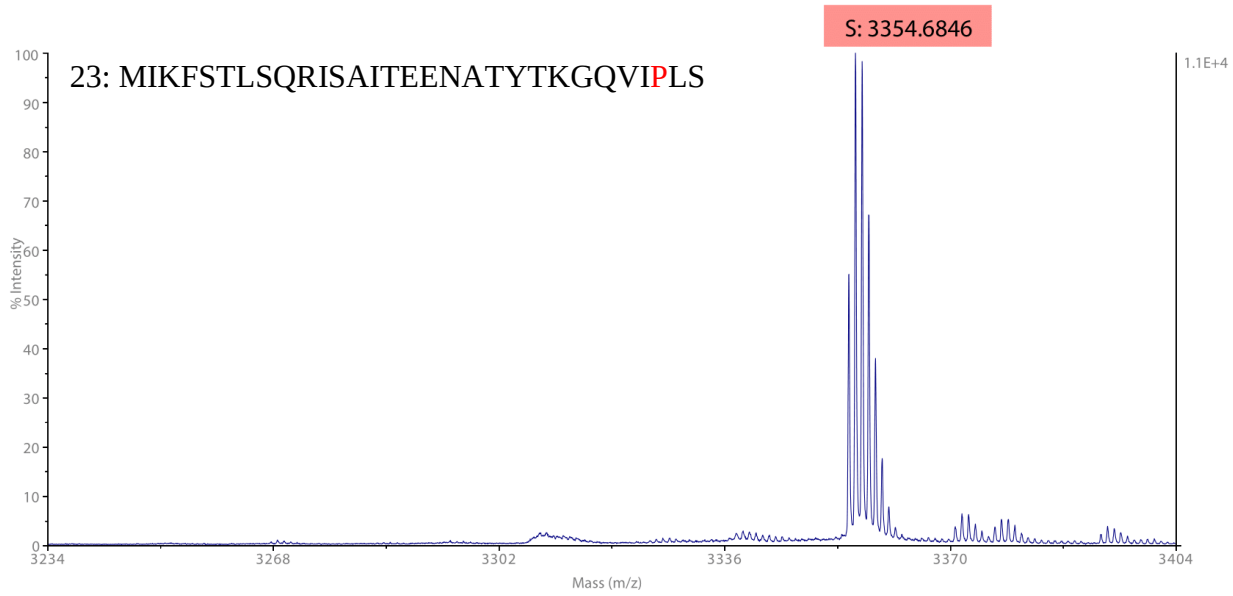

-80
-62
-36
-18
S: 3386.7698
441.2
% Intensity
Mass (m/z)

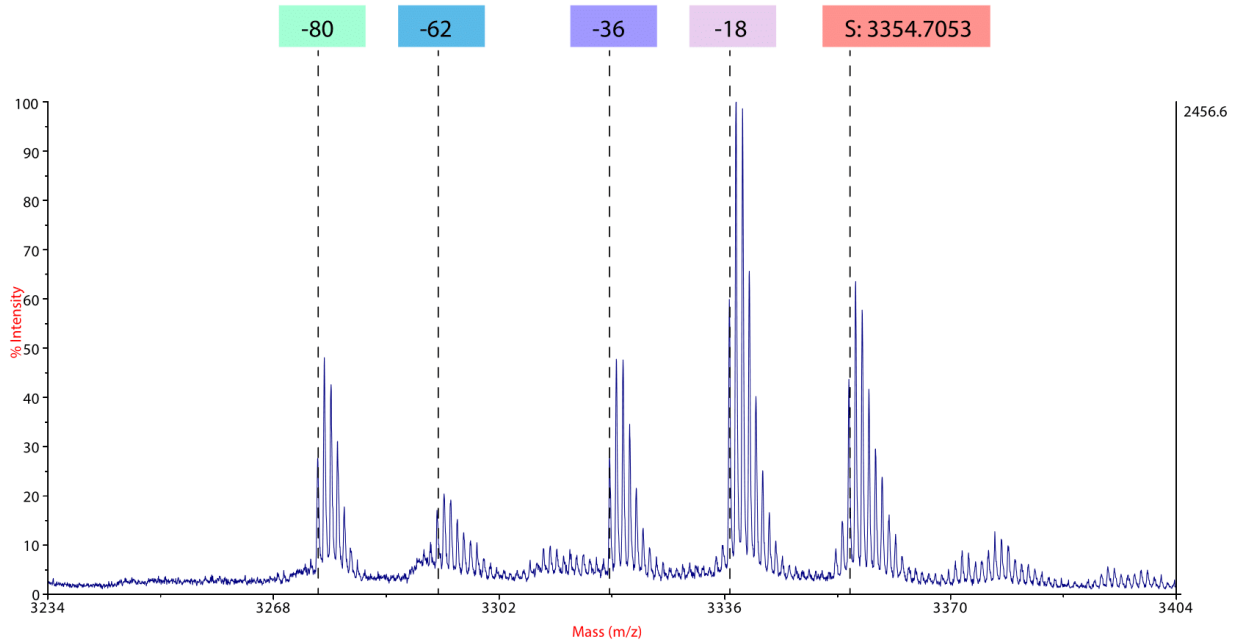

S: 3356.7009
21: MIKFSTLSQRISAITEENATYTKGQKIVLS
5551.1
% Intensity
Mass (m/z)

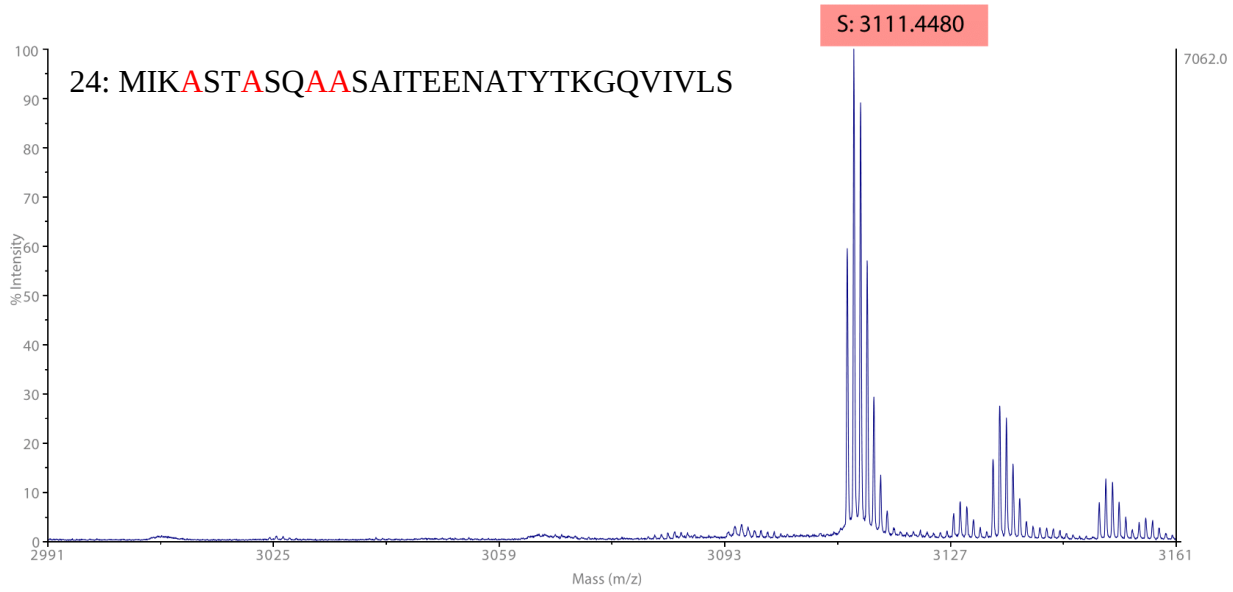

-80
-62
-36
-18
S: 3356.7014
2251.8
% Intensity
Mass (m/z)

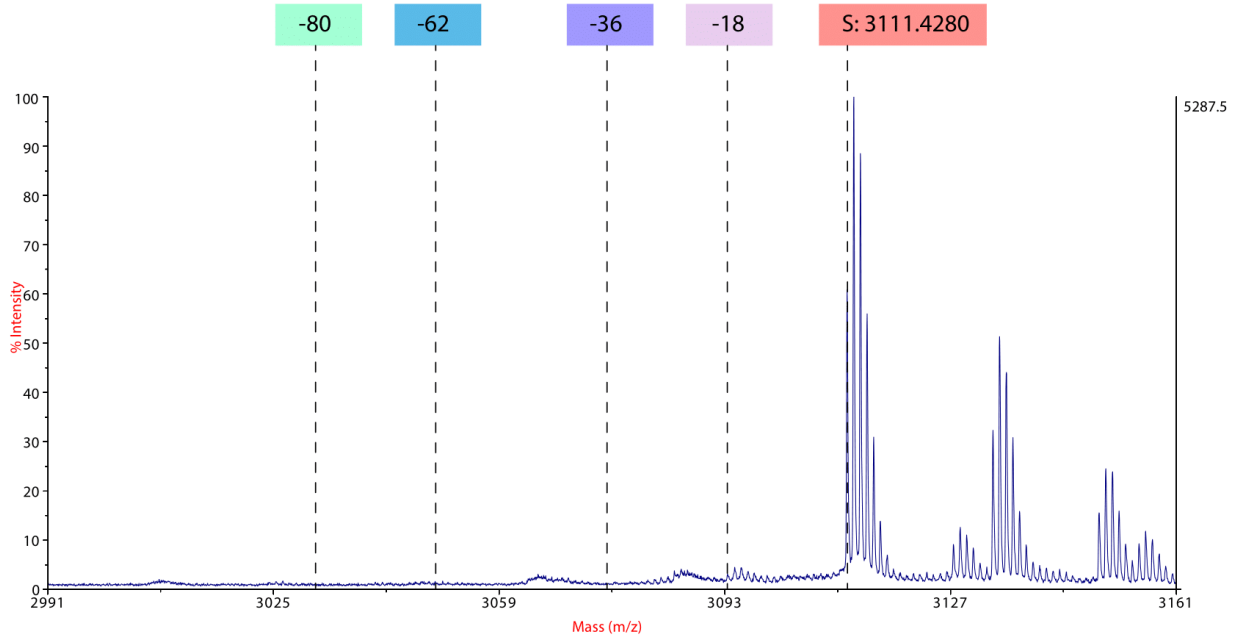

S: 3314.5691
22: MIKFSTLSQRISAITEENATYTKGQGIVLS
1.2E+4
% Intensity
Mass (m/z)

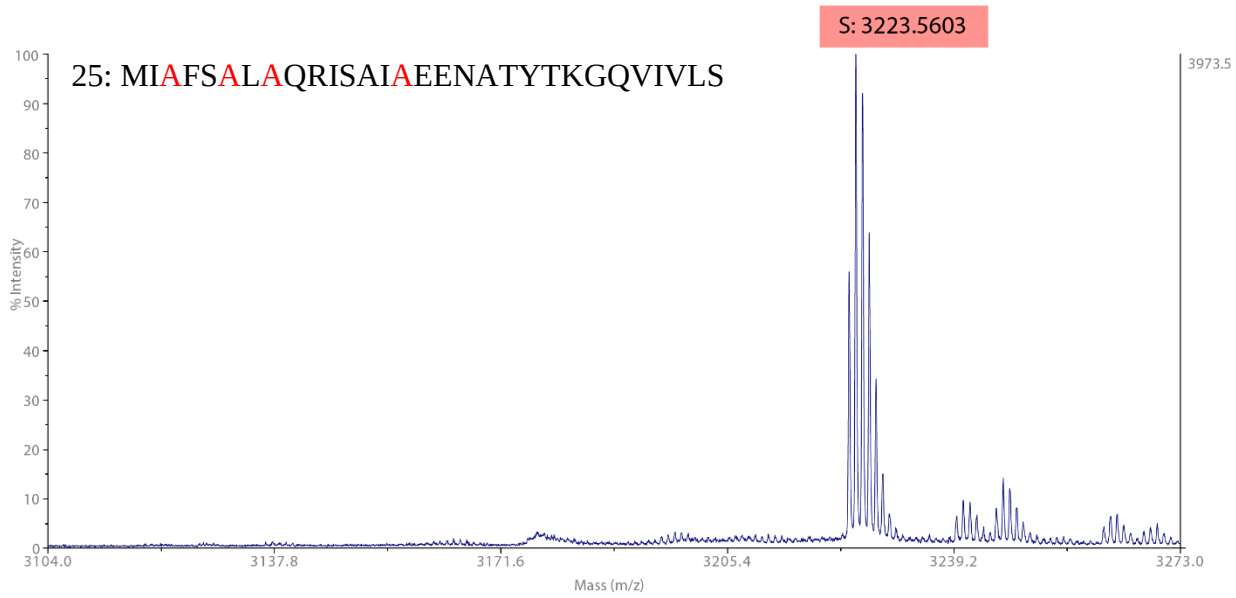

-80
-62
-36
-18
S: 3314.5493
3084.7
% Intensity
Mass (m/z)


S: 3354.6846
23: MIKFSTLSQRISAITEENATYTKGQVIPLS
1.1E+4
% Intensity
Mass (m/z)
3234
3268
3302
3336
3370
3404


-80
-62
-36
-18
S: 3354.7053
2456.6
% Intensity
Mass (m/z)
3234
3268
3302
3336
3370
3404


S: 3111.4480
24: MIKASTASQAASAITEENATYTKGQVIVLS
7062.0
% Intensity
Mass (m/z)
2991
3025
3059
3093
3127
3161


-80
-62
-36
-18
S: 3111.4280
5287.5
% Intensity
Mass (m/z)
2991
3025
3059
3093
3127
3161


S: 3223.5603
25: MIAFSALAQRISAIAEENATYTKGQVIVLS
3973.5
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3104.0
3137.8
3171.6
3205.4
3239.2
3273.0
Mass (m/z)

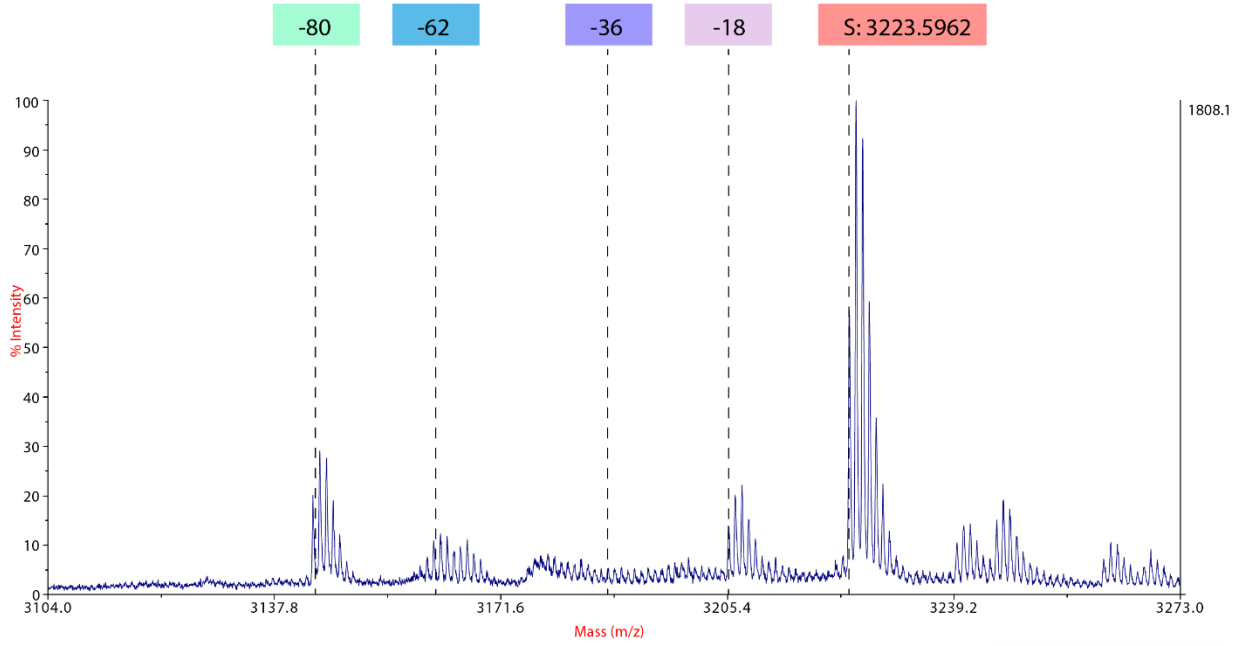

-80
-62
-36
-18
S: 3223.5962
1808.1
100
90
80
70
60
50
40
30
20
10
0
% Intensity
3104.0
3137.8
3171.6
3205.4
3239.2
3273.0
Mass (m/z)

Table A.1. MALDI-TOF Masses and Integration Areas of PaaA Reaction Products. All peptide masses are measured in Daltons, include an N-terminal formyl group and are protonated $[M+H]^{1+}$. "-" indicates not observed.

Gene Name	Unmodified	Observed	Error	EIC Integration
WT	3356.7770	3356.7117	-0.0653	3323.7595
F4D	3324.7355	3324.5730	-0.1625	51800.9883
F4K	3337.8036	3337.6321	-0.1715	20183.3262
T6D	3370.7563	3370.6072	-0.1491	8115.5410
L7D	3358.7199	3358.6179	-0.1020	30735.2559
L7Y	3406.7563	3406.7009	-0.0554	7898.8286
L7V	3342.7614	3342.6731	-0.0883	5106.0410
L7K	3371.7879	3371.6567	-0.1312	18042.0156
Q8E	3357.7610	3357.6860	-0.0750	2548.8987
R10D	3315.7029	3315.6392	-0.0637	6325.0918
R10K	3328.7709	3328.7258	-0.0451	2147.5251
R10S	3287.7079	3287.6140	-0.0939	6006.4946
I11D	3358.7199	3358.5898	-0.1301	50172.5039
I11K	3371.7879	3371.6716	-0.1163	25183.5840
A13C	3388.7491	3388.6848	-0.0643	1784.1371
A13H	3422.7988	3422.6580	-0.1408	5394.7520
T15P	3352.7821	3352.7271	-0.0550	13464.7822
N18D	3357.7610	3357.6965	-0.0645	10541.2402
A19E	3414.7825	3414.6611	-0.1214	4823.1382
M20T	3386.7698	-	-	-
Q25K	3356.8134	3356.7014	-0.1120	11373.5938
V26G	3314.7301	3314.5493	-0.1808	15089.8584
V28P	3354.7614	3354.7053	-0.0561	9591.4844
F4A, L7A, R10A, I11A	3111.5878	3111.4280	-0.1598	28027.8398
K3A, T6A, S8A, T15A	3223.7031	3223.5962	-0.1069	11647.6045

Gene Name	-1 H_2O	Observed	Error	EIC Integration
WT	3338.7665	3338.6812	-0.0853	1876.3193
F4D	3306.7250	3306.5535	-0.1714	6723.4258
F4K	3319.7931	3319.6157	-0.1774	7544.8604
T6D	3352.7458	3352.5825	-0.1633	2886.2915
L7D	3340.7094	3340.6047	-0.1047	1749.9618
L7Y	3388.7458	3388.6836	-0.0622	7232.4409
L7V	3324.7509	3324.6462	-0.1046	8057.5601
L7K	3353.7774	3353.6367	-0.1406	7785.0503
Q8E	3339.7505	3339.6516	-0.0988	3025.9424
R10D	3297.6924	3297.6440	-0.0484	2535.9700
R10K	3310.7604	3310.6997	-0.0606	2605.1704
R10S	3269.6974	3269.6370	-0.0603	2818.8687
I11D	3340.7094	3340.5691	-0.1402	3375.2825
I11K	3353.7774	3353.6619	-0.1154	6773.6841
A13C	3370.7386	3370.6155	-0.1231	544.7452
A13H	3404.7883	3404.6523	-0.1359	2846.8960
T15P	3334.7716	3334.7024	-0.0691	17477.9102
N18D	3339.7505	3339.6780	-0.0725	3023.4785
A19E	3396.7720	3396.6418	-0.1301	4482.0005
M20T	3368.7593	-	-	-
Q25K	3338.8029	3338.6790	-0.1238	2500.1147
V26G	3296.7196	3296.5432	-0.1764	4750.1860
V28P	3336.7509	3336.6934	-0.0574	13877.6465
F4A, L7A, R10A, I11A	3093.5773	3093.4541	-0.1231	879.2079
K3A, T6A, S8A, T15A	3205.6926	3205.5906	-0.1020	2809.2393

Gene Name	-2 H_2O	Observed	Error	EIC Integration
WT	3320.7559	3320.7031	-0.0528	859.2036
F4D	3288.7144	3288.5408	-0.1736	845.0863
F4K	3301.7825	3301.6057	-0.1768	1680.1243
T6D	3334.7352	-	-	-
L7D	3322.6988	-	-	-
L7Y	3370.7352	3370.6699	-0.0653	2302.5500
L7V	3306.7403	3306.6340	-0.1063	3884.3833
L7K	3335.7668	3335.6230	-0.1438	1760.0526
Q8E	3321.7399	3321.6387	-0.1012	1384.8456
R10D	3279.6818	3279.6138	-0.0680	863.6547
R10K	3292.7498	3292.6814	-0.0684	666.4230
R10S	3251.6868	3251.6350	-0.0518	1253.6515
I11D	3322.6988	-	-	-
I11K	3335.7668	3335.6338	-0.1330	1741.8193
A13C	3352.7280	-	-	-
A13H	3386.7777	-	-	-
T15P	3316.7610	3316.6892	-0.0718	3220.2356
N18D	3321.7399	3321.6707	-0.0692	587.8158
A19E	3378.7614	3378.6584	-0.1030	2075.9150
M20T	3350.7487	-	-	-
Q25K	3320.7923	3320.7163	-0.0760	347.0039
V26G	3278.7090	3278.5784	-0.1306	638.4860
V28P	3318.7403	3318.6704	-0.0699	6275.6924
F4A, L7A, R10A, I11A	3075.5667	-	-	-
K3A, T6A, S8A, T15A	3187.6820	-	-	-

Gene Name	-1 H_2O, -1 CO_2	Observed	Error	EIC Integration
WT	3294.7767	3294.7229	-0.0538	4031.1494
F4D	3262.7352	3262.6104	-0.1247	1474.7930
F4K	3275.8033	3275.6460	-0.1573	3868.3882
T6D	3308.7560	3308.6177	-0.1383	3165.8730
L7D	3296.7196	-	-	-
L7Y	3344.7560	3344.7151	-0.0409	1426.3845
L7V	3280.7611	3280.6948	-0.0662	4892.7886
L7K	3309.7876	3309.6484	-0.1391	1027.7512
Q8E	3295.7607	3295.6841	-0.0766	3267.9446
R10D	3253.7026	-	-	-
R10K	3266.7706	-	-	-
R10S	3225.7076	-	-	-
I11D	3296.7196	-	-	-
I11K	3309.7876	-	-	-
A13C	3326.7488	3326.6663	-0.0825	960.8293
A13H	3360.7985	3360.6846	-0.1139	3255.9944
T15P	3290.7818	-	-	-
N18D	3295.7607	3295.7017	-0.0590	1154.3992
A19E	3352.7822	3352.6794	-0.1028	7580.5332
M20T	3324.7695	3324.6689	-0.1006	1478.0989
Q25K	3294.8131	3294.7190	-0.0941	1439.7134
V26G	3252.7298	3252.5757	-0.1541	2622.8750
V28P	3292.7611	3292.7024	-0.0586	3359.0815
F4A, L7A, R10A, I11A	3049.5875	-	-	-
K3A, T6A, S8A, T15A	3161.7028	3161.5918	-0.1110	1606.0957

Gene Name	-2 H_2O, -1 CO_2	Observed	Error	EIC Integration
WT	3276.7661	3276.6963	-0.0698	9133.7402
F4D	3244.7246	3244.5747	-0.1499	3804.2710
F4K	3257.7927	3257.6287	-0.1640	11306.4912
T6D	3290.7454	3290.5815	-0.1639	9257.2480
L7D	3278.7090	-	-	-
L7Y	3326.7454	3326.6863	-0.0591	2727.9836
L7V	3262.7505	3262.6643	-0.0862	10029.3389
L7K	3291.7770	3291.6172	-0.1598	2826.6680
Q8E	3277.7501	3277.6621	-0.0880	7170.7930
R10D	3235.6920	-	-	-
R10K	3248.7600	-	-	-
R10S	3207.6970	-	-	-
I11D	3278.7090	-	-	-
I11K	3291.7770	-	-	-
A13C	3308.7382	3308.6440	-0.0942	2814.1140
A13H	3342.7879	3342.6484	-0.1395	8514.3945
T15P	3272.7712	-	-	-
N18D	3277.7501	3277.6929	-0.0572	4361.3672
A19E	3334.7716	3334.6594	-0.1122	16647.4512
M20T	3306.7589	3306.6543	-0.1046	2770.1619
Q25K	3276.8025	3276.6887	-0.1138	3713.7407
V26G	3234.7192	3234.5571	-0.1621	6547.8315
V28P	3274.7505	3274.6931	-0.0574	7198.2451
F4A, L7A, R10A, I11A	3031.5769	-	-	-
K3A, T6A, S8A, T15A	3143.6922	3143.5735	-0.1187	3850.2832

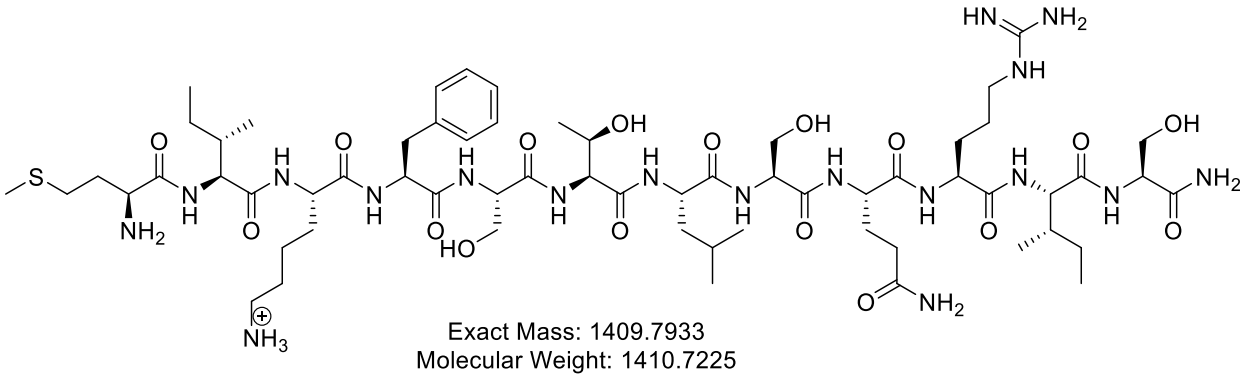
Table A.2. MALDI-TOF Mass Analysis of PaaA Reactions on Core E Mutants. All peptide masses are measured in Daltons, include an N-terminal formyl group and are protonated $[M+H]^{1+}$. "-" indicates not observed. "*" Only trace amounts observed.

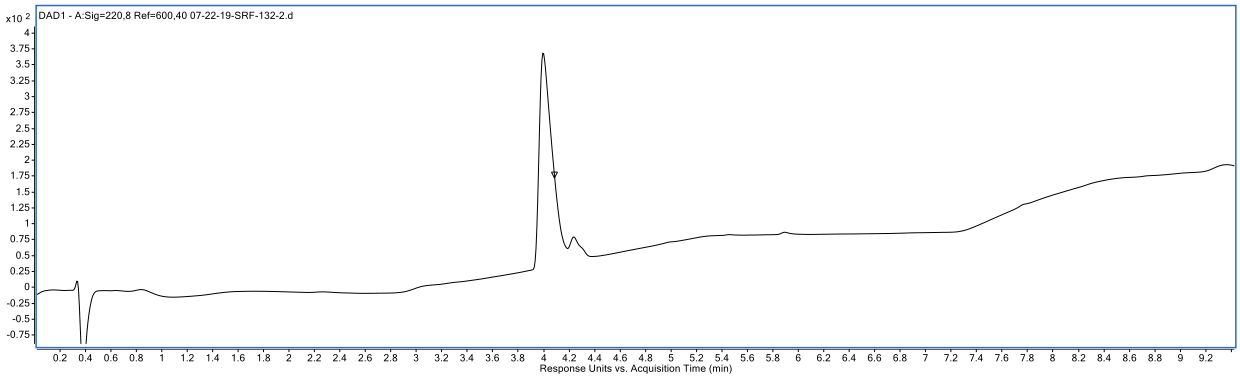
No PaaA						
Gene Name	**Unmodified**	**Observed**	**Error**	**-1 H_2O**	**Observed**	**Error**
E16V	3326.8028	3326.0938	-0.7090	3308.7923	-	-
E17A	3298.7715	3298.2246	-0.5469	3280.7610	-	-
E16Q	3355.7930	3355.2129	-0.5801	3337.7825	-	-
E17Q	3355.7930	3355.1023	-0.6907	3337.7825	-	-
PaaA						
Gene Name	**Unmodified**	**Observed**	**Error**	**-1 H_2O**	**Observed**	**Error**
E16V	3326.8028	3326.0979	-0.7049	3308.7923	3308.0933	-0.6989
E17A	3298.7715	3298.1704	-0.6011	3280.7610	3280.1570	-0.6039
E16Q	3355.7930	3355.1912	-0.6018	3337.7825	3337.1838*	-0.5987
E17Q	3355.7930	3355.1023	-0.6907	3337.7825	3337.0637	-0.7188

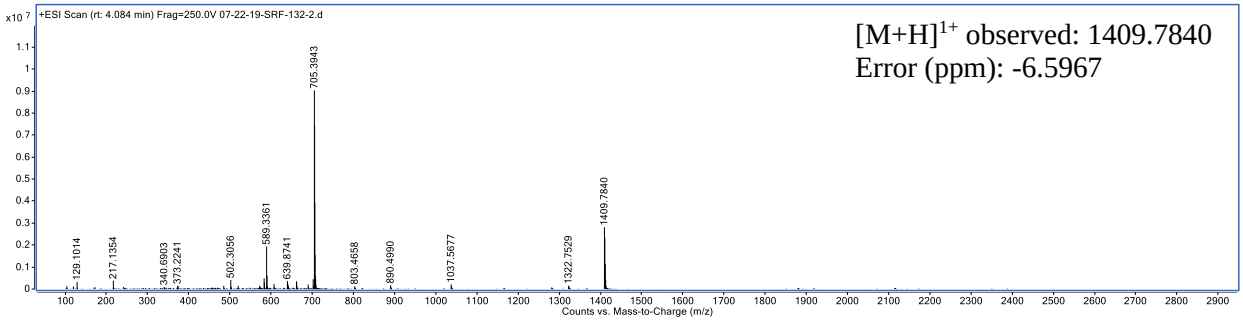
Table A.3. Calculated mass error for 13C-enriched PaaP A13C modification with PaaA. Masses are reported in Daltons as $[M+3H]^{3+}$.

PaaP A13C	Expected (m/z)	Observed (m/z)	Error (PPM)
Substrate	1136.6030	1136.6013	-1.4957
Product	1109.5983	1109.5999	1.4420

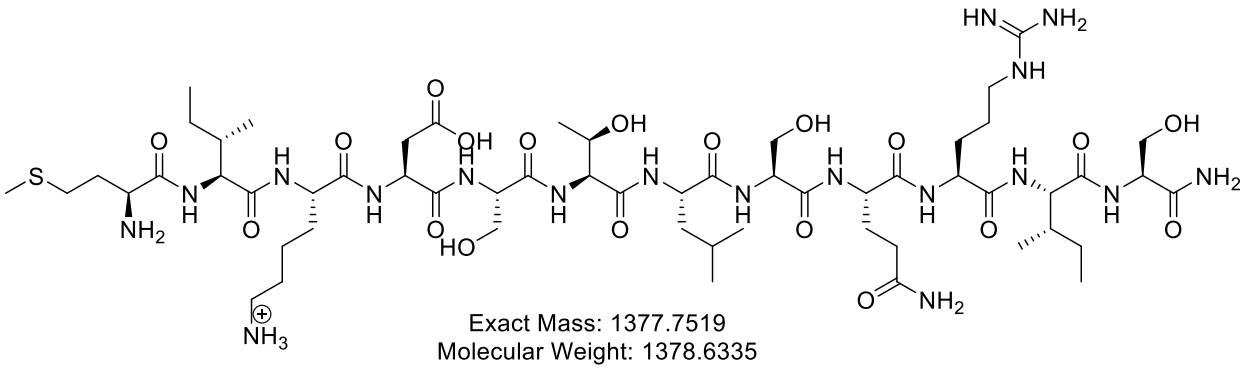
Figure A.10. Peptide Characterization
Wild Type



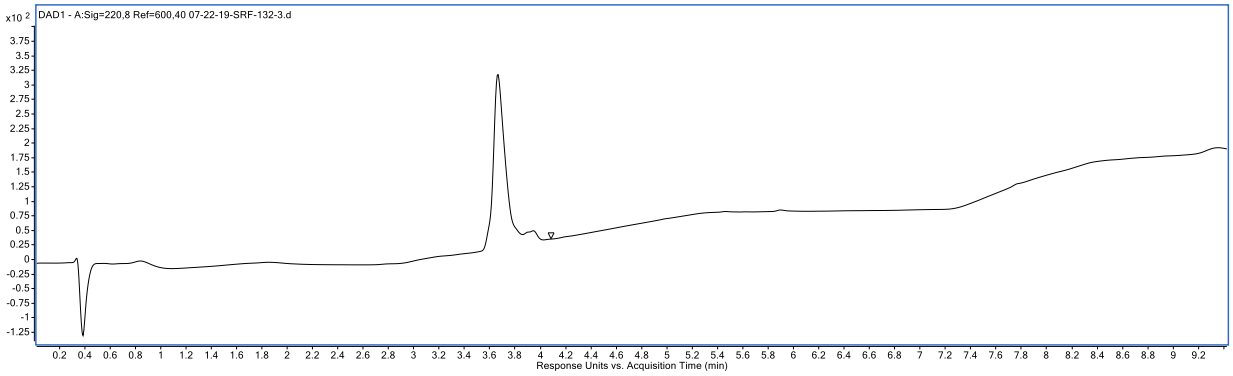


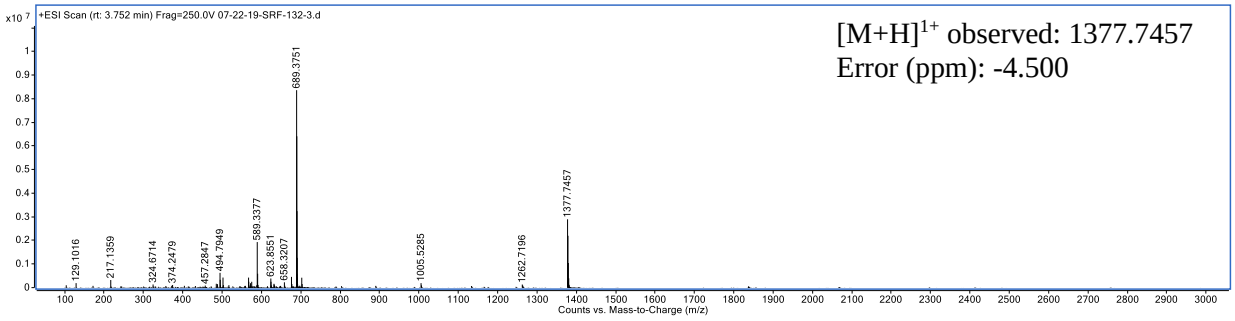


F4D



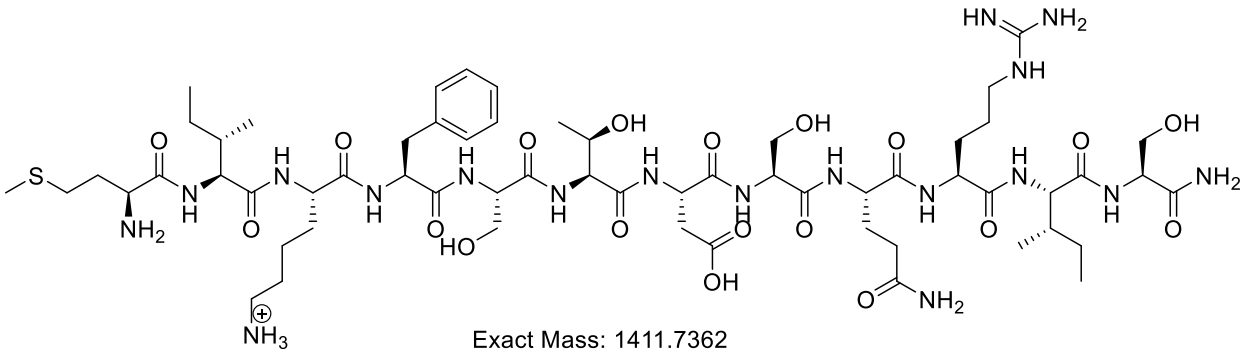
Exact Mass: 1377.7519
Molecular Weight: 1378.6335



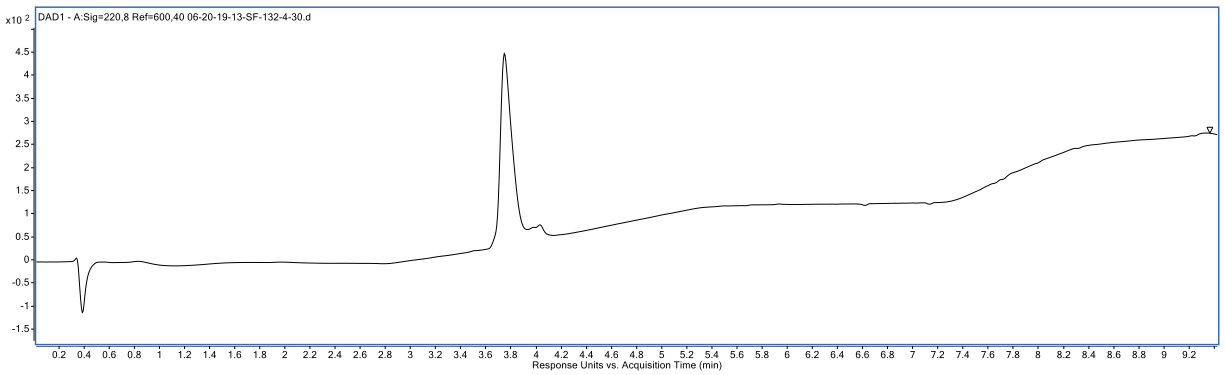


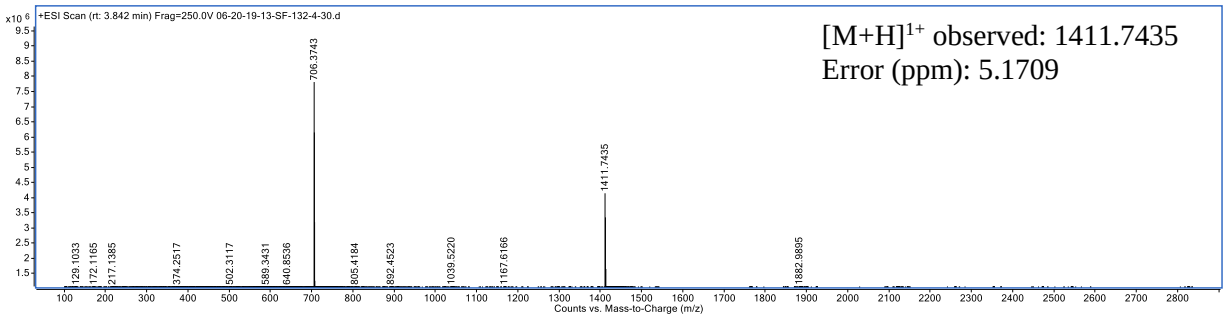
$[M+H]^{1+}$ observed: 1377.7457
Error (ppm): -4.500

L7D



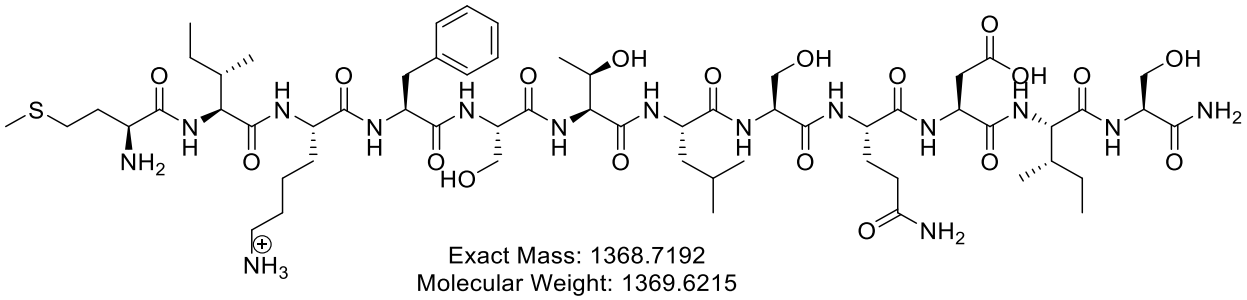
Exact Mass: 1411.7362
Molecular Weight: 1412.6505



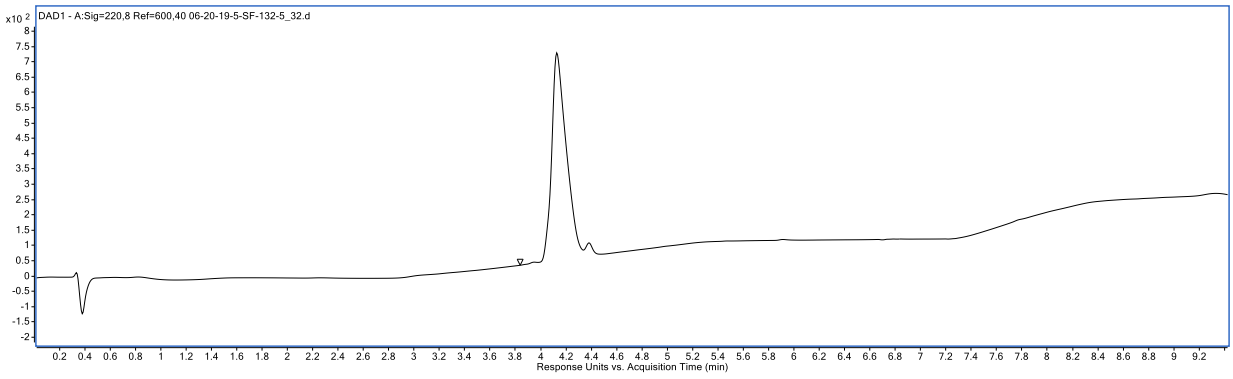


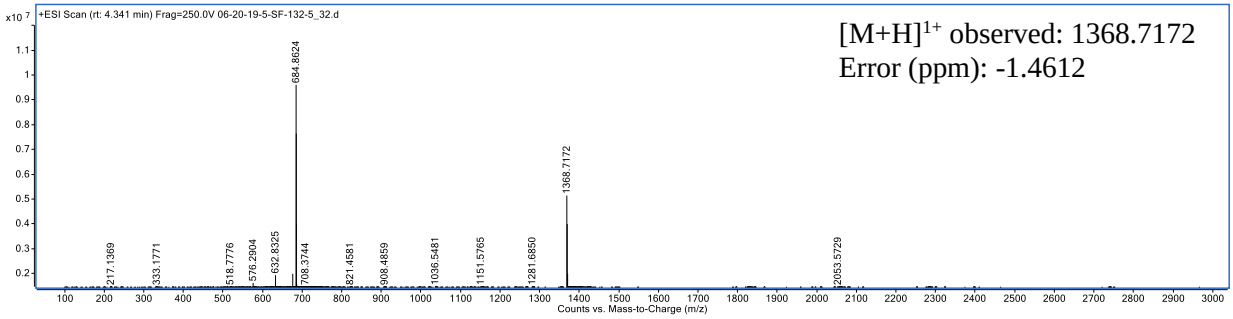
$[M+H]^{1+}$ observed: 1411.7435
Error (ppm): 5.1709

R10D



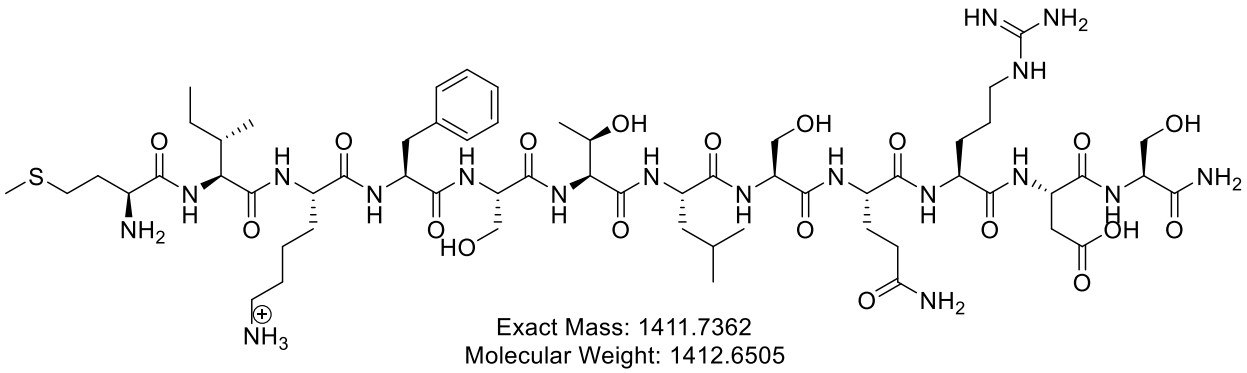
Exact Mass: 1368.7192
Molecular Weight: 1369.6215



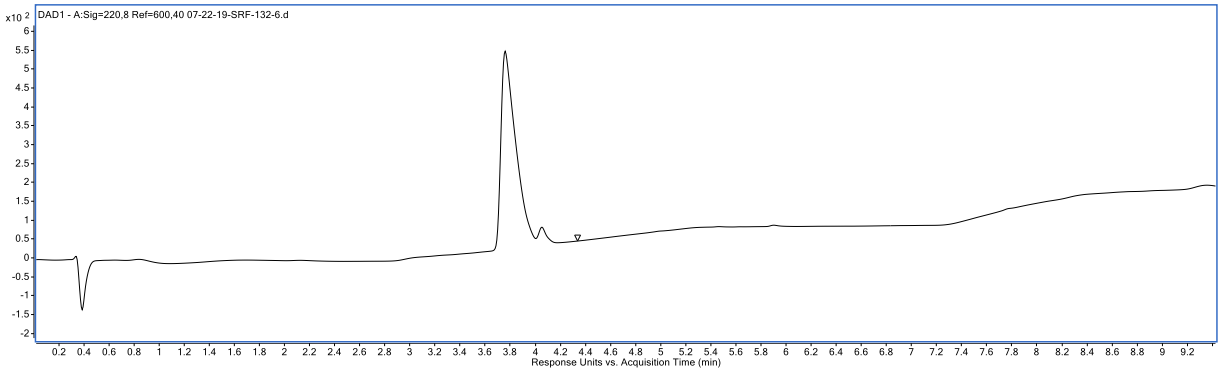


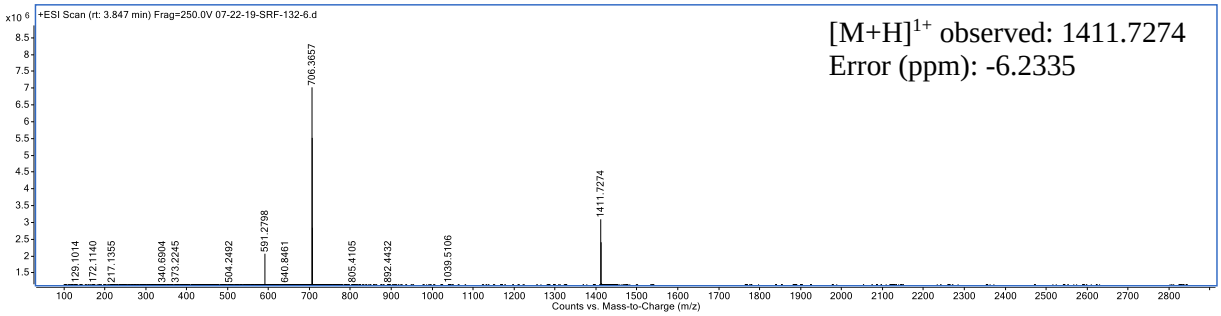
$[M+H]^{1+}$ observed: 1368.7172
Error (ppm): -1.4612

I11D



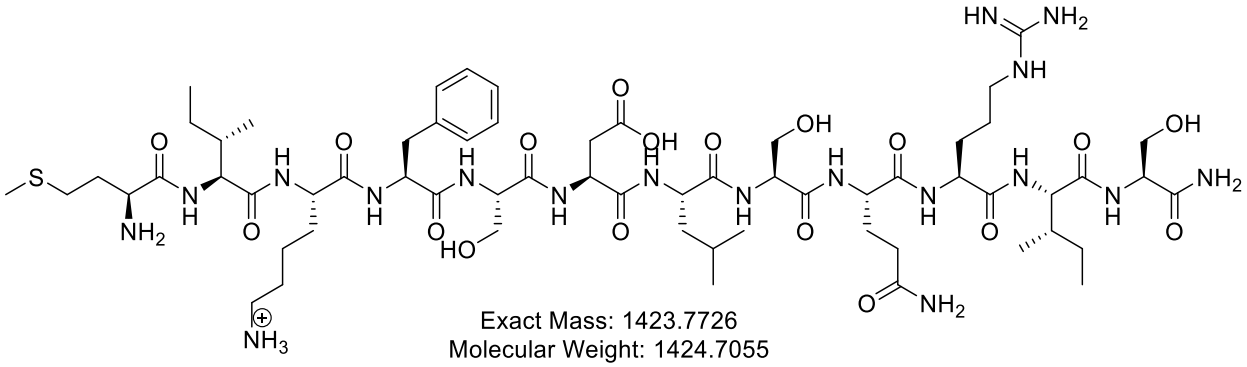
Exact Mass: 1411.7362
Molecular Weight: 1412.6505

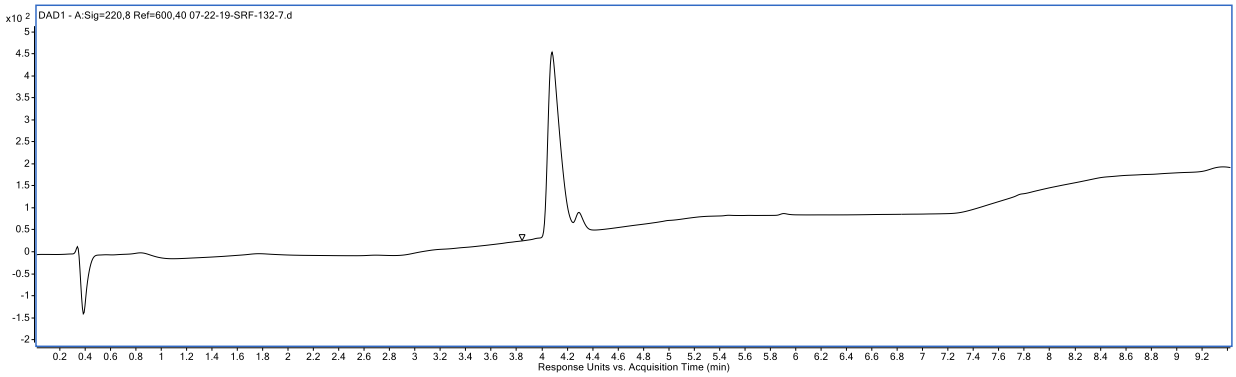


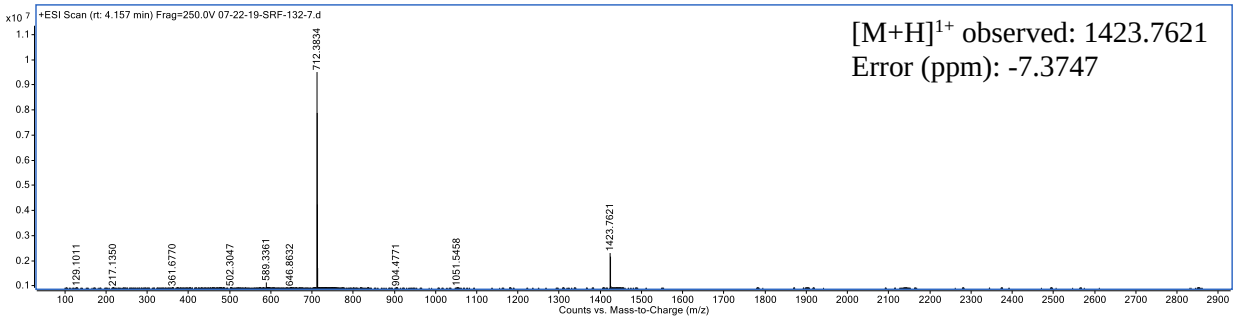


$[M+H]^{1+}$ observed: 1411.7274
Error (ppm): -6.2335

T6D

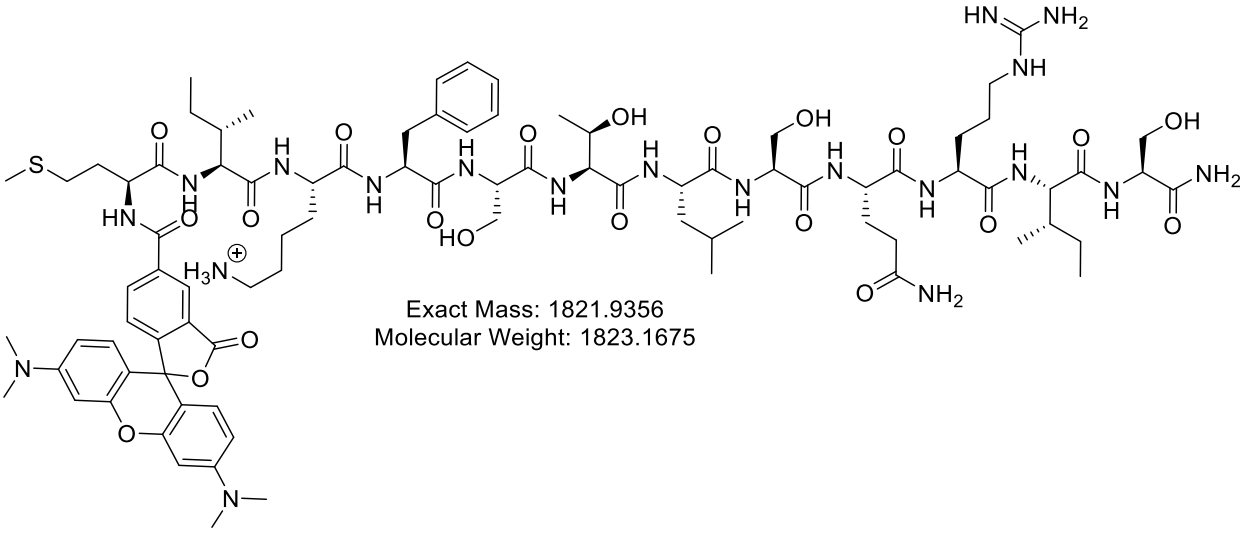


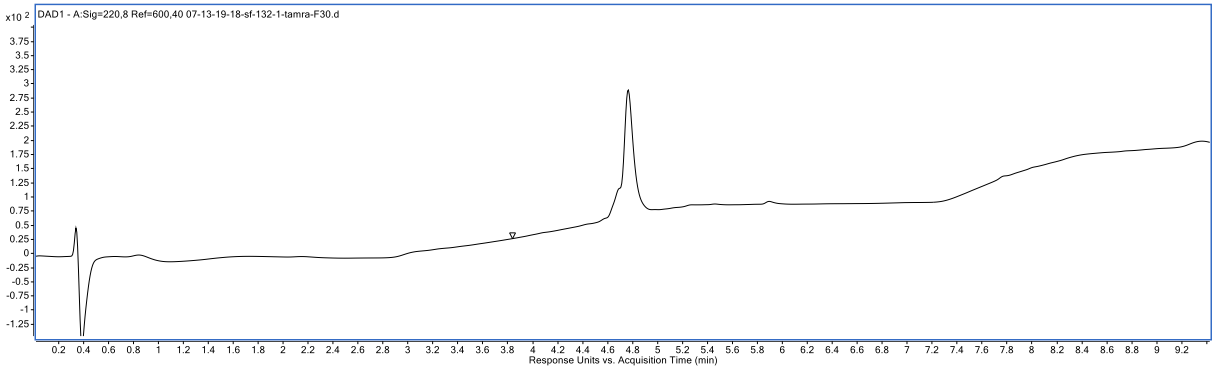




$[M+H]^{1+}$ observed: 1423.7621
Error (ppm): -7.3747

TAMRA-WT


Exact Mass: 1821.9356
Molecular Weight: 1823.1675


DAD1 - A:Sig=220,8 Ref=600,40 07-13-19-18-sf-132-1-tamra-F30.d
Response Units vs. Acquisition Time (min)

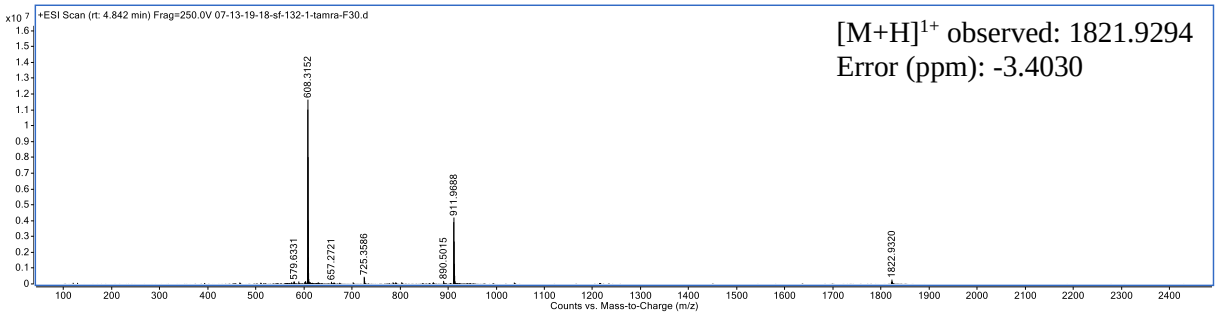

+ESI Scan (rt: 4.842 min) Frag=250.0V 07-13-19-18-sf-132-1-tamra-F30.d
608.3152
579.6331
657.2721
725.3586
890.5015
911.9688
1822.9320
Counts vs. Mass-to-Charge (m/z)
[M+H]1+ observed: 1821.9294
Error (ppm): -3.4030

Table A.4. Translation Genes

PaaP WT

MIKFSTLSQRISAITEENATYTKGQVIVLS
GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAATTCTCTACGT
TATCTCAGCGCATTTCAGCCATCACTGAAGAAAACGCCACGTACACGAAAGGGCAG
GTGATTGTTCTTTCATAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP F4D

MIKDSTLSQRISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAAGACAGTACGC
TTTCGCAGCGCATCAGCGCAATTACAGAAGAAAACGCTACGTATACCAAAGGACAG
GTCATCGTGTTGTCCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP F4K

MIKKSTLSQRISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAAAAATCCACGC
TTTCCCAACGTATTTCTGCGATTACGGAAGAGAACGCCACCTATACTAAAGGACAAG
TGATCGTCCTGTCATAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCTC
CGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGAA
AGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP T6D

MIKFSDLSQRISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAGTTCTCAGACT
TGAGTCAACGTATCAGTGCGATTACCGAAGAGAACGCCACTTATACCAAGGGGCAA
GTAATTGTCCTGAGTTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP L7D

MIKFSTDSQRISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAGTTCAGCACAG
ATTCACAACGCATTTCAGCAATTACAGAAGAGAACGCTACTTATACAAAGGGCCAG
GTCATTGTGCTGAGTTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP L7Y

MIKFSTYSQRISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAATTCTCCACCT
ACAGTCAACGCATCAGCGCTATTACAGAAGAAAATGCCACGTATACTAAAGGACAA
GTCATTGTGCTGAGCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP L7V

MIKFSTVSQRISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAGTTCAGTACCG
TTTCTCAGCGTATTAGCGCGATTACAGAGGAGAATGCAACGTACACAAAGGGTCAG
GTCATTGTTTTAAGCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP L7K

MIKFSTKSQRISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAGTTTTCCACGA
AGTCACAACGCATTTCGGCAATTACAGAAGAGAACGCCACGTATACCAAGGGTCAA
GTGATCGTACTGTCCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP Q8E

MIKFSTLSERISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAATTCAGCACGT
TATCGGAGCGCATCAGTGCTATCACAGAGGAAAACGCAACTTACACGAAAGGTCAG
GTCATTGTACTGTCATAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP R10D

MIKFSTLSQDISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAGTTCTCTACCT
TGTCACAGGACATCTCCGCTATCACAGAAGAAAATGCAACCTACACCAAGGGCCAG
GTCATTGTACTTTCGTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP R10K

MIKFSTLSQKISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAATTCTCCACCT
TGAGCCAAAAAATTTCTGCAATCACGGAAGAAAATGCGACGTACACTAAAGGCCAG
GTAATTGTCTTATCGTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP R10S

MIKFSTLSQSISAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAATTTTCAACGT
TATCCCAATCTATTTCAGCCATCACGGAAGAAAATGCTACATACACGAAAGGTCAA
GTGATTGTCCTGAGCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP I11D

MIKFSTLSQRDSAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAATTTTCTACCTT
GAGCCAACGTGATTCTGCGATCACAGAAGAAAACGCAACATACACTAAAGGTCAAG
TAATCGTGCTGTCATAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCTC
CGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGAA
AGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP I11K

MIKFSTLSQRKSAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAGTTTAGTACCC
TTAGCCAGCGCAAATCAGCAATTACTGAAGAGAATGCAACTTATACGAAAGGGCAA
GTAATTGTCTTGTCATAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP A13C

MIKFSTLSQRISCITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAATTCTCTACCC
TGTCACAACGCATCTCATGTATCACAGAAGAAAACGCTACTTACACAAAAGGCCAG
GTAATTGTACTTTCCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP A13H

MIKFSTLSQRISHITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAGTTCTCTACCT
TGTCGCAACGTATTTCCCATATTACCGAGGAGAACGCTACCTATACCAAGGGACAGG
TCATCGTACTTTCCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCTC
CGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGAA
AGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP T15P

MIKFSTLSQRISAIPEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAGTTTAGCACCT
TGTCGCAGCGTATTTCTGCCATCCCCGAGGAGAACGCCACATATACTAAGGGGCAA
GTTATCGTATTGTCATAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP N18D
MIKFSTLSQRISAITEEDATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAGTTTAGCACTC
TGTCCCAGCGCATCAGTGCGATCACTGAGGAAGATGCAACCTATACGAAGGGGCAA
GTGATCGTCCTTTCGTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP A19E

MIKFSTLSQRISAITEENETYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAATTCTCTACTCT
TTCTCAACGTATTTCGGCCATCACCGAAGAGAATGAGACTTACACTAAAGGGCAAGT
GATTGTCTTGTCTTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCTCC
GTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGAAA
GGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP M20T

MIKFSTLSQRISAITEENAMYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAATTCTCTACTTT
ATCCCAACGTATCTCCGCAATCACCGAGGAAAACGCAATGTATACTAAGGGGCAGG
TAATTGTACTTAGCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCTC
CGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGAA
AGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP Q25K

MIKFSTLSQRISAITEENATYTKGKVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAGTTTTCCACGT
TGTCACAGCGCATTTCTGCTATCACCGAAGAGAACGCCACTTACACCAAAGGCAAA
GTAATTGTGTTATCCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP V26G

MIKFSTLSQRISAITEENATYTKGQGIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAATTTTCTACAT
TGAGTCAACGCATTTCGGCCATTACGGAGGAGAACGCCACATATACTAAAGGTCAG
GGTATCGTCTTAAGCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP V28P

MIKFSTLSQRISAITEENATYTKGQVIPLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAATTCAGTACTC
TGTCGCAACGTATTTCGGCGATTACTGAGGAGAATGCGACATACACAAAAGGACAA
GTGATCCCACTTTCCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP F4A, L7A, R10A, I11A

MIKASTASQAASAITEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATTAAGGCGTCCACGG
CTTCCCAGGCGGCTTCCGCCATTACGGAAGAGAATGCAACGTACACCAAGGGGCAA
GTTATTGTTTTGTCATAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP K3A, T6A, S8A, T15A

MIAFSALAQRISAIAEENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCGCATTCTCCGCGT
TGGCCCAACGTATTTCAGCTATCGCCGAAGAGAACGCCACCTACACCAAAGGCCAA
GTAATCGTCCTGAGTTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP E16V

MIKFSTLSQRISAITVENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAATTCTCCACTT
TGAGCCAGCGCATCTCGGCCATTACTGTAGAGAATGCGACGTACACAAAAGGTCAG
GTTATTGTTCTTTCGTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP E17A

MIKFSTLSQRISAITEANATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAATTTAGTACGT
TAAGCCAACGTATTTCTGCGATCACGGAGGCTAACGCAACATATACGAAAGGTCAA
GTGATCGTCCTTAGCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP E16Q

MIKFSTLSQRISAITQENATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAGTTTTCAACGC
TGTCGCAACGCATTAGCGCGATCACTCAAGAAAACGCGACGTACACTAAGGGCCAG
GTGATCGTCTTGAGTTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGA
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP E17Q

MIKFSTLSQRISAITEQNATYTKGQVIVLS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCAAATTCAGCACGC

TTTCGCAACGCATTAGTGCTATTACGGAACAAAATGCAACCTATACTAAAGGCCAGG
TGATTGTTCTTAGTTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCTC
CGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGAA
AGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

PaaP WT-Display

MIKFSTLSQRISAITEENATYTKGQVIVLSGSGSYPYDVPDYAGSGSGS

TAATACGACTCACTATAGGGTTAACTTTAAGAAGGAGATATACATATGATCAAGTTT
TCGACTCTTAGCCAACGTATTTCTGCTATTACGGAAGAAAATGCTACCTACACTAAA
GGGCAGGTAATCGTCCTGTCTGGATCTGGCAGCTACCCGTATGACGTACCTGATTAT
GCTGGTTCAGGGTCCGGTTCGTAGGACGGGGGGCGGAAA

PaaP smSVL (Each peptide contains 1 randomized NNK codon between I2 to S30)

MIKFSTLSQRISAITEENATYTKGQVIVLSGSGSYPYDVPDYAGSGSGS

TAATACGACTCACTATAGGGTTAACTTTAAGAAGGAGATATAAATATGATCAAGTTC
TCTACTCTGTCTCAGCGCATTTCTGCTATCACTGAAGAAAATGCTACCTATACCAAA
GGTCAGGTTATCGTCCTTAGCGGCAGCGGCAGCTACCCATACGACGTGCCCGACTAT
GCAGGTTCTGGTTCTGGTTCTTAGGACGGGGGGCGGAAA

APPENDIX B. SUPPLEMENTARY TABLES AND FIGURES FOR CHAPTER 3

Figure B.1. Purification of His_6-LynD, His_6-MBP-TbtE, His_6-MBP-TclM, and His_6-MBP-LazC. *lynD* was cloned into pMCSG7 as an N-terminal His_6 fusion. The remaining genes were cloned into pMCSG9 as N-terminal His_6 maltose binding protein fusions. Proteins were expressed in *E. coli* DE3 cells and affinity column purified. Below is an SDS-Page analyzing the purified proteins.

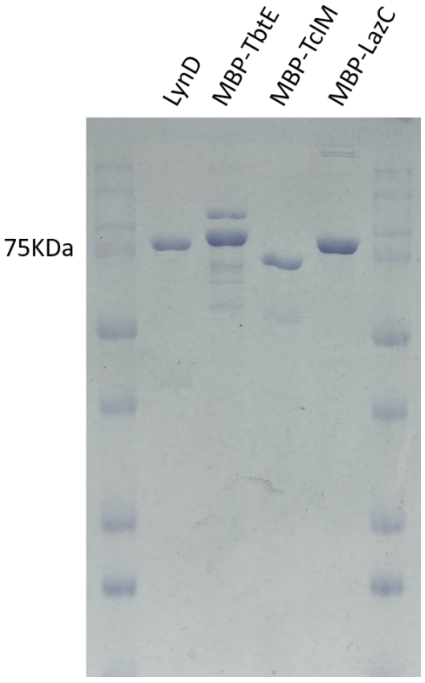


Figure B.2. Optimization of Sec(Ph) acylation of microhelix.[1] 6 µL of 500 mM HEPES-KOH pH 8.0, 6 µL of 250 µM of dFx, 6 µL of 250 µM microhelix, and 18 uL of water were pipetted together and heated for 2 min at 95 °C. The solution was cooled for 5 min at room temperature before 12 µL of 3 M $MgCl_2$ was added and mixed. After another 5 min incubation at room temperature, 10 µL of 25 mM Sec(Ph)-DBE in DMSO was added (or DMSO only for negative sample), mixed and the solution placed on ice. 10 µL aliquots were removed at 6, 24, 48, and 72 hours and quenched with 40 µL 0.3 M AcONa and ethanol precipitated. Pellets were dried and resuspended in 1.5 µL of 10 mM NaOAc pH 5.2 and further diluted with 14 µL of acid page RNA loading buffer. 1 µL of each sample was loaded onto a 20% Urea-Acid-PAGE gel and run for 150 min at 120 V in 50 mM AcONa pH 5.2 as running buffer. The gel was soaked in ethidium bromide and visualized by typhoon FLA 9000.

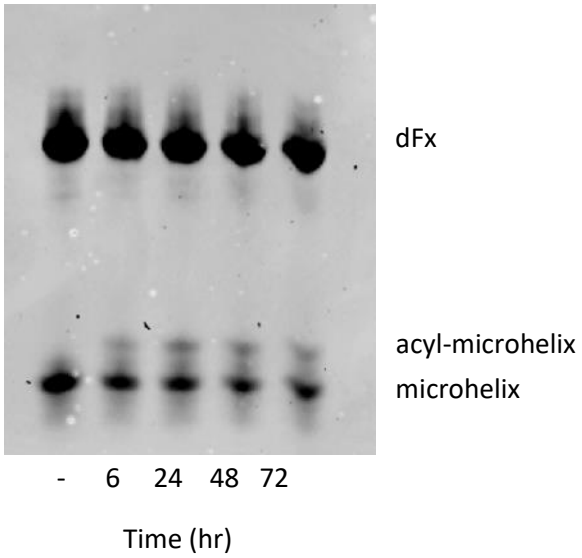


Figure B.3. Design of LazC substrates. Shown is a wild type sequence for the lactazole and three newly designed precursor peptides to make lactazole analogs using our developed strategy. The LynD RS is colored in blue, amino acids colored in red are transformed into Dhas, amino acids colored in green are transformed into thiazoles / oxazoles, and the tryptophan codon is reprogrammed for Sec(Ph).

WT: MSDITASRVESLDLQDLDLSELTVTSLRDTVALPENGASWGSCSCQASSSCAQP

LacHyb1: MSDITASRVESLDLQDLDLAELSEEALRDTVALPENGAWSGSCWCQAWCWCAQP

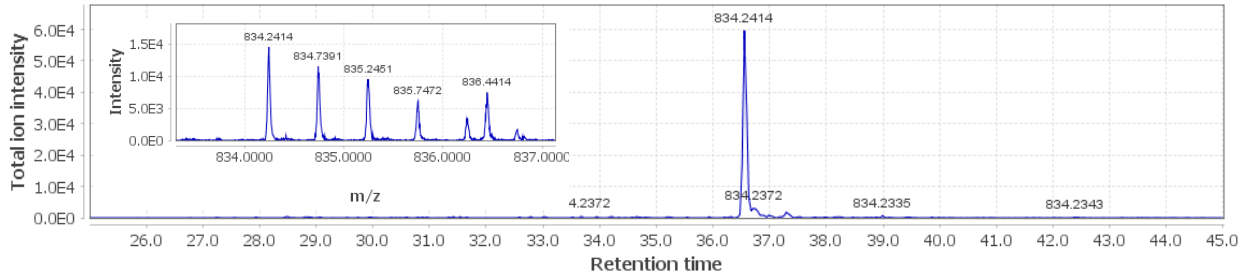
LacHyb2: MSDITASRVESLAELSEEALGDAVTSLRDTVALPENGAWSGSCWCQAWCWCAQP

LacHyb3: MLAELSEEALGDALQDLDLSELTVTSLRDTVALPENGAWSGSCWCQAWCWCAQP

Figure B.4.1-B.4.44 Nano-LCESI-MS Traces. Designed precursor peptides were made as described earlier. Condition 1a: Samples were translated, treated with 4 µM LynD / 3 µM MBP-TbtE (~20hrs, 25 °C) and then 0.125 U endoproteinase GluC (37 °C, overnight). Condition 1b: Samples were translated with tryptophan reprogrammed for Sec(Ph), treated with 4 µM LynD / 3 µM MBP-TbtE (~20hrs, 25 °C), 1 M H_2O_2 (1hr, 4 °C), 500 mM TCEP (1 hr, 37 °C), and 0.125 U endoproteinase GluC (37 °C, >5 hrs). Condition 2: Samples were translated with tryptophan reprogrammed for Sec(Ph), treated with 4 µM LynD / 3 µM MBP-TbtE (~20hrs, 25 °C), 1 M H_2O_2 (1hr, 4 °C), 500 mM TCEP (1 hr, 37 °C), and 2 µM MBP-TclM or MBP-LazC (1-22 hrs, 25°C) All reactions were desalted using c18 stage tips and analyzed by LC-ESI-Q/TOF MS. All traces were generated as extracted ion chromatograms (XIC) within 10 ppm error between 25 and 45 minutes. In spectra with multiple major peaks an arrow indicates the peak used to generate MS spectra. Additional peaks were found to not contain the correct exact mass or charge state.

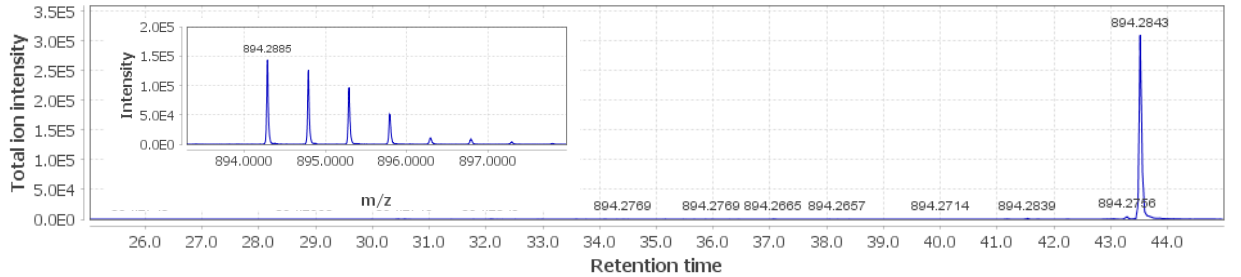
B.4.1. (Table 3.1B, entry 1) Condition 1a
MHHHHHHSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGASCTTCVCTCSCCTT

Calc. $[M+2H]^{2+}$: 834.2385 Obs. $[M+2H]^{2+}$: 834.2414 Chemical Formula: $C_{65}H_{92}N_{18}O_{20}S_7^{2+}$



B.4.2. (Table 3.1B, entry 2) Condition 1a
MHHHHHHSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAFCTTCVCTCFCCTT

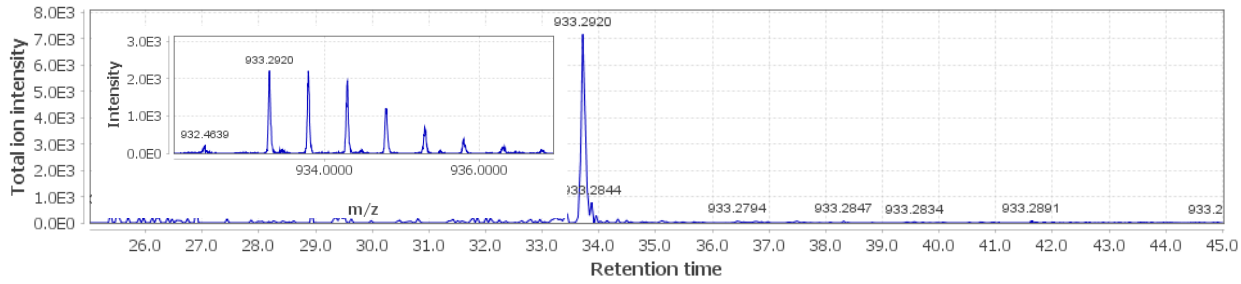
Calc. $[M+2H]^{2+}$: 894.2749 Obs. $[M+2H]^{2+}$: 894.2885 Chemical Formula: $C_{77}H_{100}N_{18}O_{18}S_7^{2+}$



B.4.3. (Table 3.1B, entry 3) Condition 1a

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCVCTCWCCTT

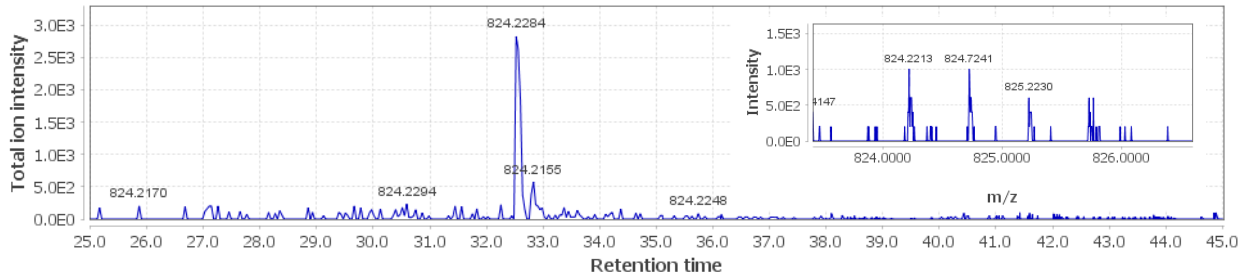
Calc. $[M+2H]^{2+}$: 933.2858 Obs. $[M+2H]^{2+}$: 933.2920 Chemical Formula: $C_{81}H_{102}N_{20}O_{18}S_7^{2+}$



B.4.4. (Table 3.1A, entry 1) Condition 1b

MSSQLAELSEEALGDAENEALEIMGAWCTTCVCTCWCCTT

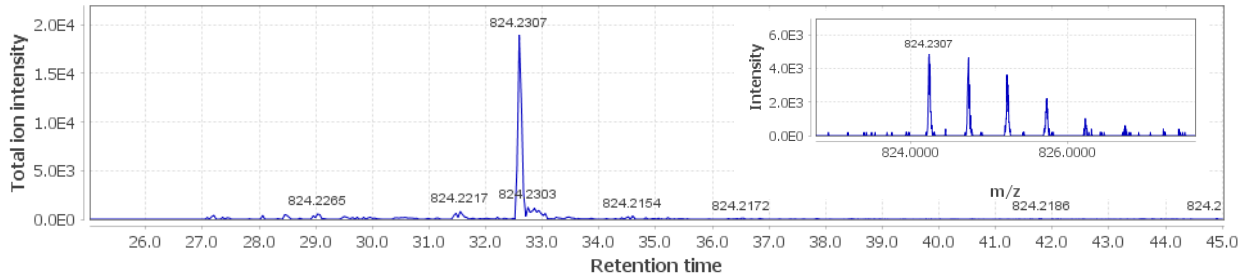
Calc. $[M+2H]^{2+}$: 824.2254 Obs. $[M+2H]^{2+}$: 824.2213 Chemical Formula: $C_{65}H_{88}N_{18}O_{19}S_7^{2+}$



B.4.5. (Table 3.1A, entry 3) Condition 1b

MSEIKKALNTLEIEDFLAELSEEALGDAENEALEIMGAWCTTCVCTCWCCTT

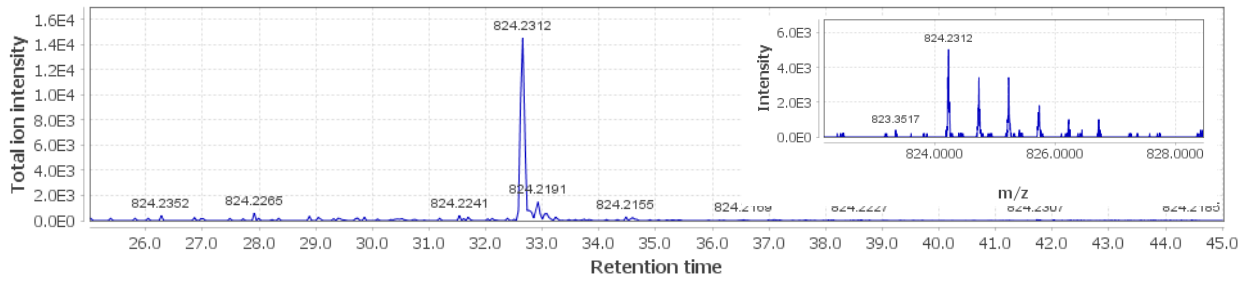
Calc. $[M+2H]^{2+}$: 824.2254 Obs. $[M+2H]^{2+}$: 824.2307 Chemical Formula: $C_{65}H_{88}N_{18}O_{19}S_7^{2+}$



B.4.6. (Table 3.1A, entry 4) Condition 1b

MSEIKKALNTLEIEDFDAIEMVDVDAMPLAELSEEALGDAENEALEIMGAWCTTCVCTCWCCTT

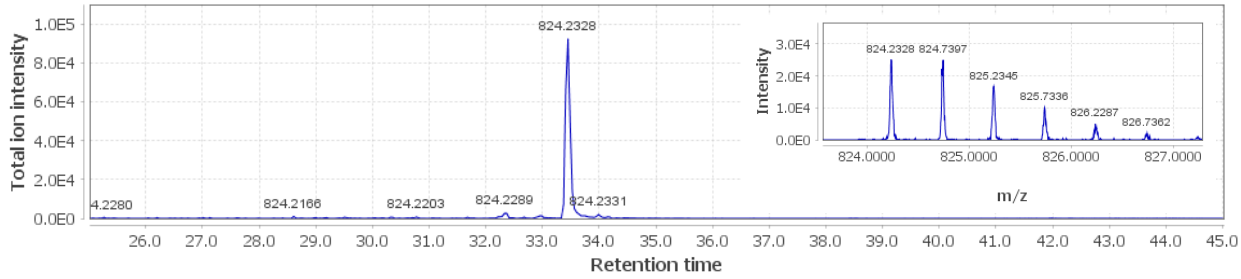
Calc. $[M+2H]^{2+}$: 824.2254 Obs. $[M+2H]^{2+}$: 824.2312 Chemical Formula: $C_{65}H_{88}N_{18}O_{19}S_7^{2+}$



B.4.7. (Table 3.1A, entry 5) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAAAAAAENEALEIMGAWCTTCVCTCWCCTT

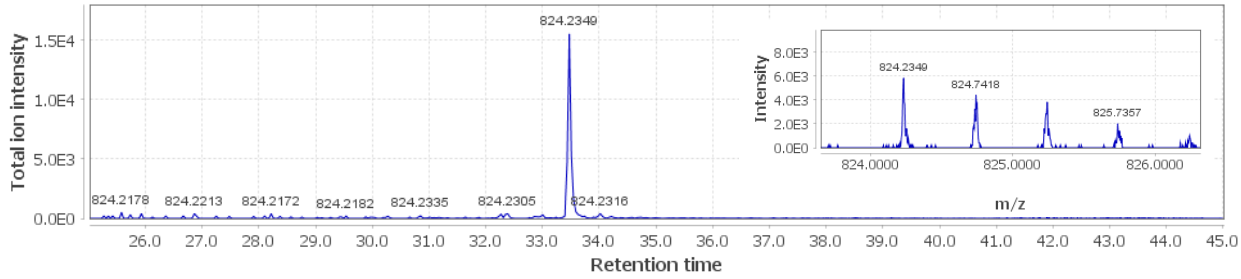
Calc. $[M+2H]^{2+}$: 824.2254 Obs. $[M+2H]^{2+}$: 824.2328 Chemical Formula: $C_{65}H_{88}N_{18}O_{19}S_7^{2+}$



B.4.8. (Table 3.1A, entry 6) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAGGGGGENEALEIMGAWCTTCVCTCWCCTT

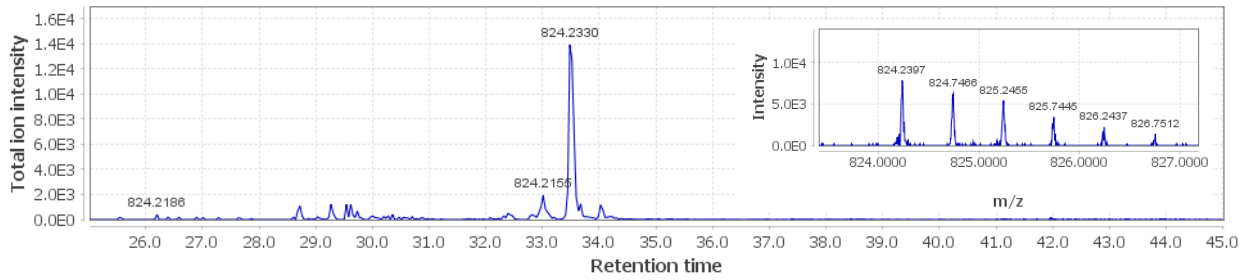
Calc. $[M+2H]^{2+}$: 824.2254 Obs. $[M+2H]^{2+}$: 824.2349 Chemical Formula: $C_{65}H_{88}N_{18}O_{19}S_7^{2+}$



B.4.9. (Table 3.1A, entry 7) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAKKKKKENEALEIMGAWCTTCVCTCWCCTT

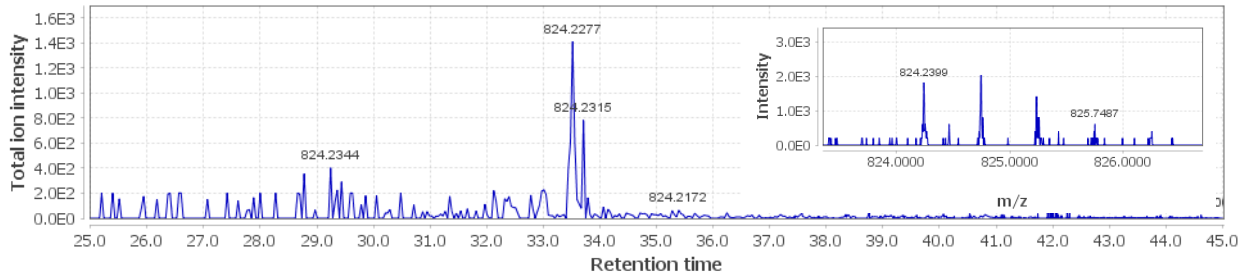
Calc. $[M+2H]^{2+}$: 824.2254 Obs. $[M+2H]^{2+}$: 824.2397 Chemical Formula: $C_{65}H_{88}N_{18}O_{19}S_7^{2+}$



B.4.10. (Table 3.1A, entry 8) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDADDDDDENEALEIMGAWCTTCVCTCWCCTT

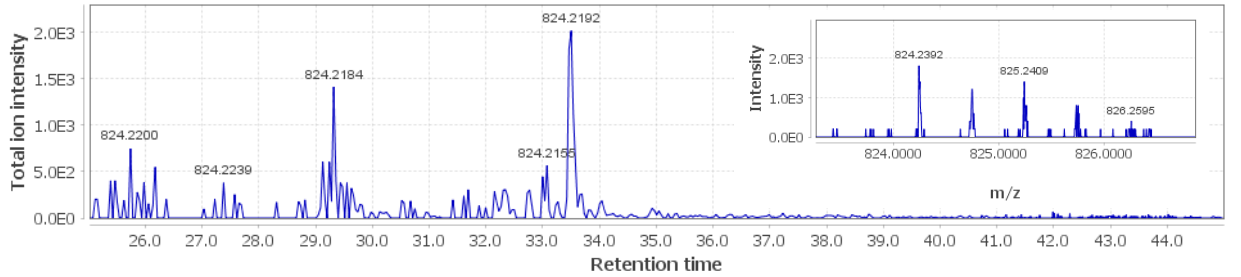
Calc. $[M+2H]^{2+}$: 824.2254 Obs. $[M+2H]^{2+}$: 824.2399 Chemical Formula: $C_{65}H_{88}N_{18}O_{19}S_7^{2+}$



B.4.11. (Table 3.1A, entry 9) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAPPPPPENEALEIMGAWCTTCVCTCWCCTT

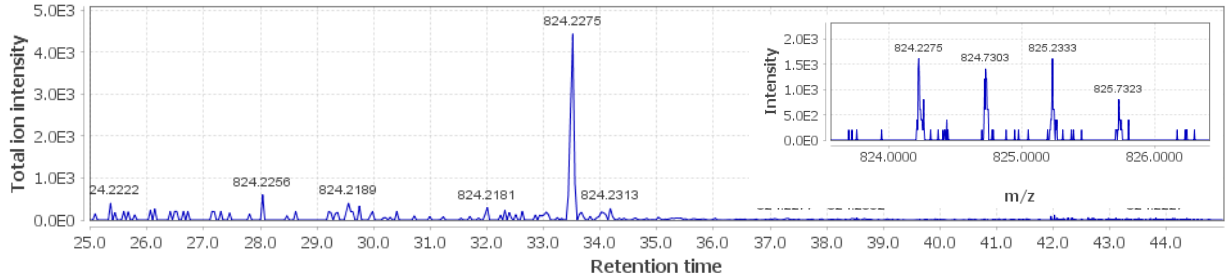
Calc. $[M+2H]^{2+}$: 824.2254 Obs. $[M+2H]^{2+}$: 824.2392 Chemical Formula: $C_{65}H_{88}N_{18}O_{19}S_7^{2+}$



B.4.12. (Table 3.1A, entry 10) Condition 1b

MLAELSEEALGDASEIKKALNTLEIEDFDAIEMVDVDAMPENEALEIMGAWCTTCVCTCWCCTT

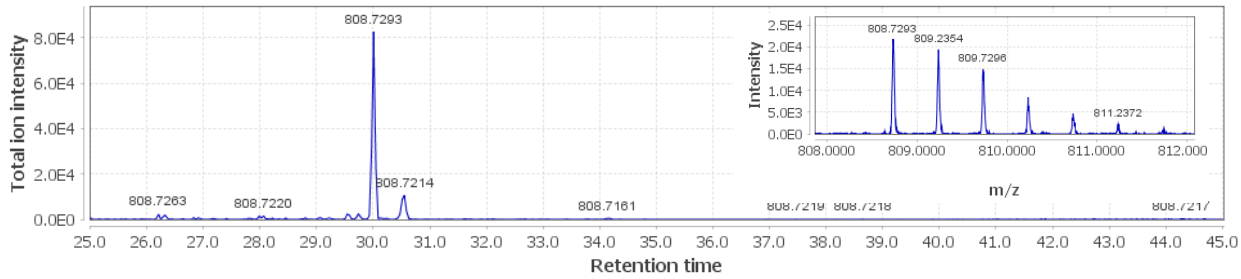
Calc. $[M+2H]^{2+}$: 824.2254 Obs. $[M+2H]^{2+}$: 824.2275 Chemical Formula: $C_{65}H_{88}N_{18}O_{19}S_7^{2+}$



B.4.13. (Table 3.1B, entry 4) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCKCTCWCCAA

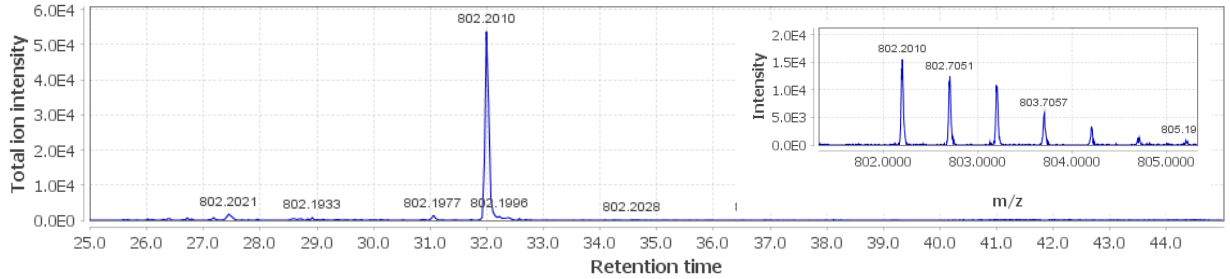
Calc. $[M+2H]^{2+}$: 808.7281 Obs. $[M+2H]^{2+}$: 808.7293 Chemical Formula: $C_{64}H_{87}N_{19}O_{17}S_7^{2+}$



B.4.14. (Table 3.1B, entry 5) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCDCTCWCCAA

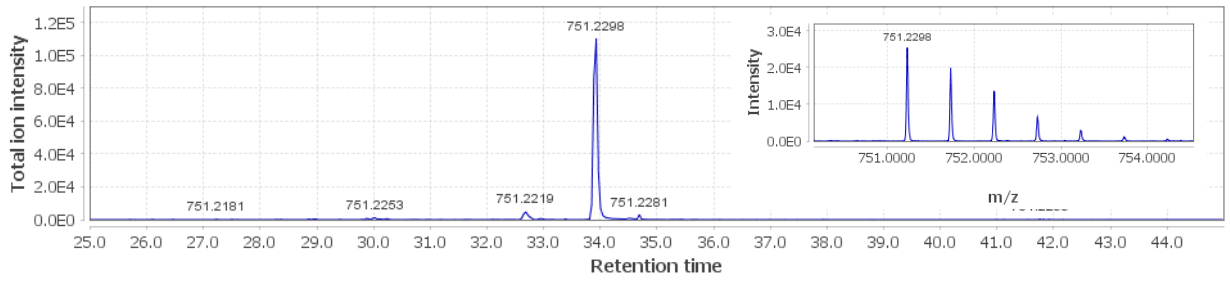
Calc. $[M+2H]^{2+}$: 802.1941 Obs. $[M+2H]^{2+}$: 802.2010 Chemical Formula: $C_{62}H_{80}N_{18}O_{19}S_7^{2+}$



B.4.15. (Table 3.1B, entry 6) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACVSACWCCAA

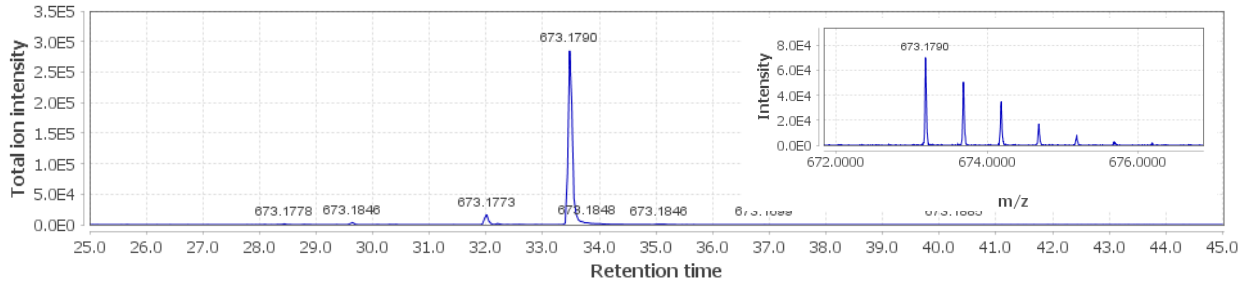
Calc. $[M+2H]^{2+}$: 751.2234 Obs. $[M+2H]^{2+}$: 751.2298 Chemical Formula: $C_{60}H_{82}N_{18}O_{16}S_6^{2+}$



B.4.16. (Table 3.1B, entry 7) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCA

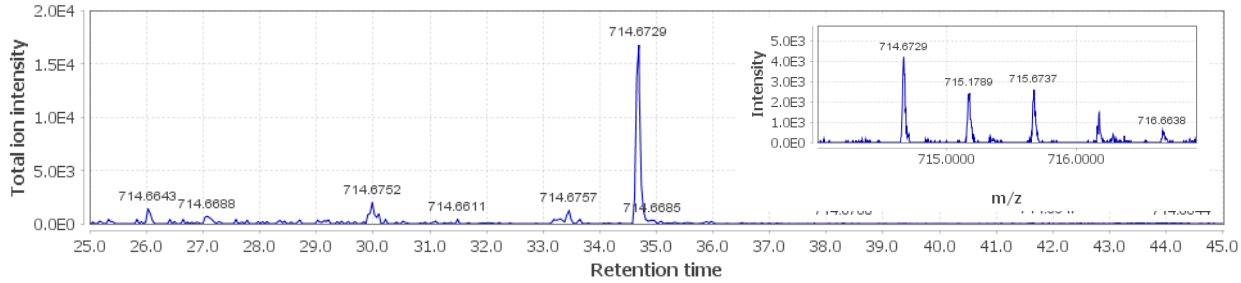
Calc. $[M+2H]^{2+}$: 673.1783 Obs. $[M+2H]^{2+}$: 673.1790 Chemical Formula: $C_{53}H_{70}N_{16}O_{14}S_6^{2+}$



B.4.17. (Table 3.1B, entry 9) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCA

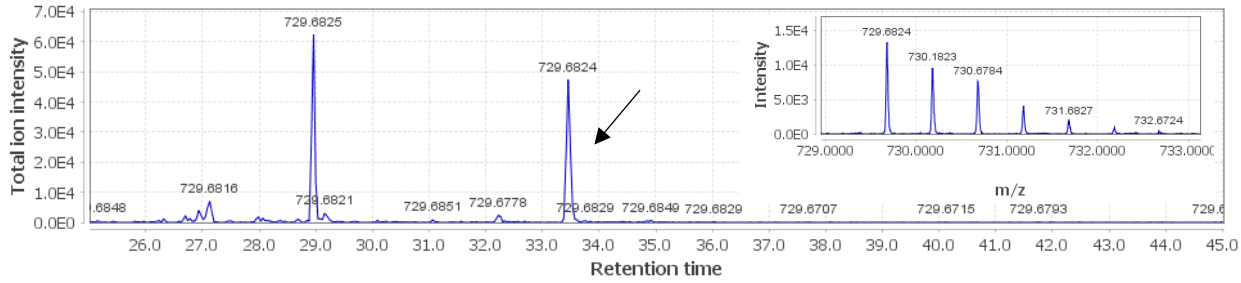
Calc. $[M+2H]^{2+}$: 714.6700 Obs. $[M+2H]^{2+}$: 714.6729 Chemical Formula: $C_{56}H_{71}N_{17}O_{14}S_7^{2+}$



B.4.18. (Table 3.1B, entry 10) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCT

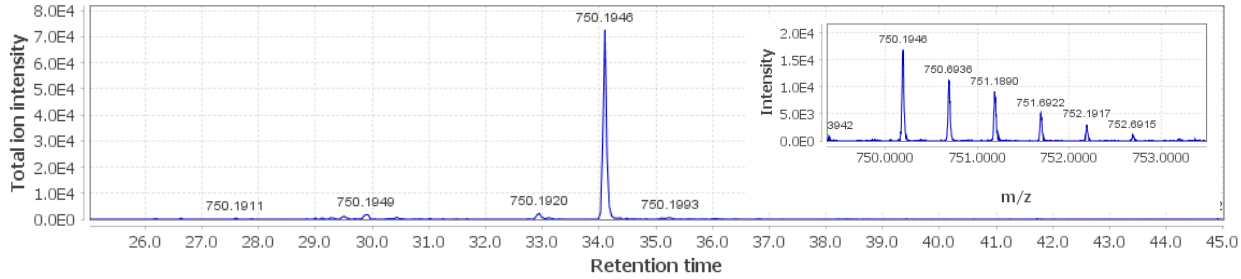
Calc. $[M+2H]^{2+}$: 729.6753 Obs. $[M+2H]^{2+}$: 729.6824 Chemical Formula: $C_{57}H_{73}N_{17}O_{15}S_7^{2+}$



B.4.19. (Table 3.1B, entry 11) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCAA

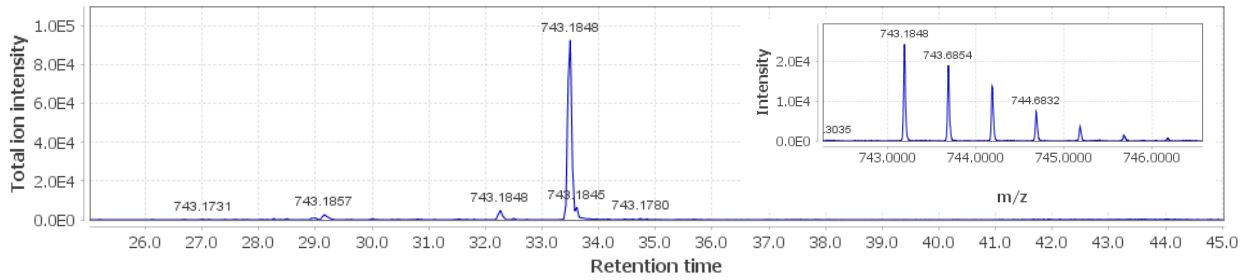
Calc. $[M+2H]^{2+}$: 750.1886 Obs. $[M+2H]^{2+}$: 750.1946 Chemical Formula: $C_{59}H_{76}N_{18}O_{15}S_7^{2+}$



B.4.20. (Table 3.1B, entry 12) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCAG

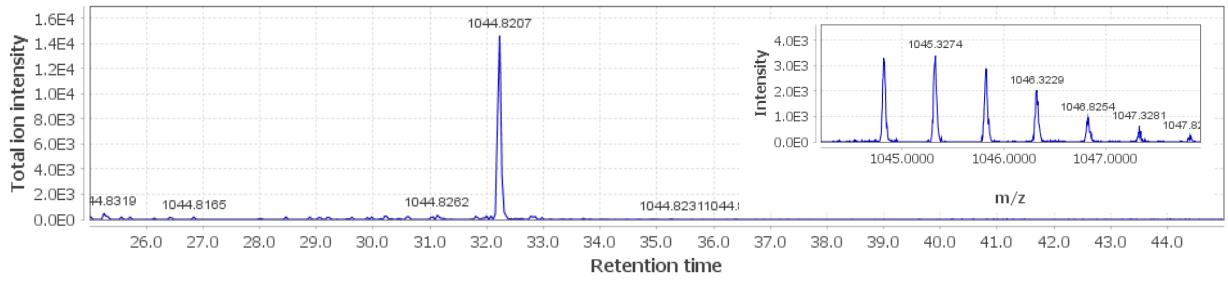
Calc. $[M+2H]^{2+}$: 743.1808 Obs. $[M+2H]^{2+}$: 743.1848 Chemical Formula: $C_{58}H_{74}N_{18}O_{15}S_7^{2+}$



B.4.21. (Table 3.1B, entry 13) Condition 1b

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCVCTCWCCAANSGGVS

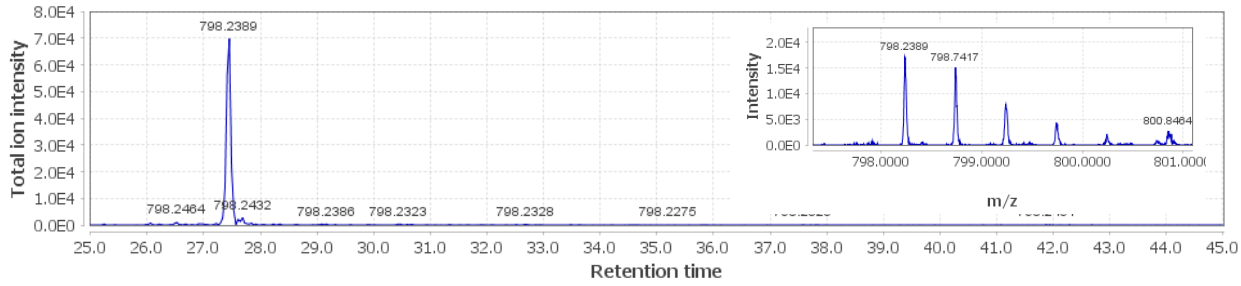
Calc. $[M+2H]^{2+}$: 1044.8240 Obs. $[M+2H]^{2+}$: 1044.8207 Chemical Formula: $C_{82}H_{115}N_{25}O_{26}S_7^{2+}$



B.4.22. (figure S3, LacHyb1) Condition 1b

MSDITASRVESLDLQDLDLAELSEEALRDTVALPENGAWSGSCWCQAWCWCAQP

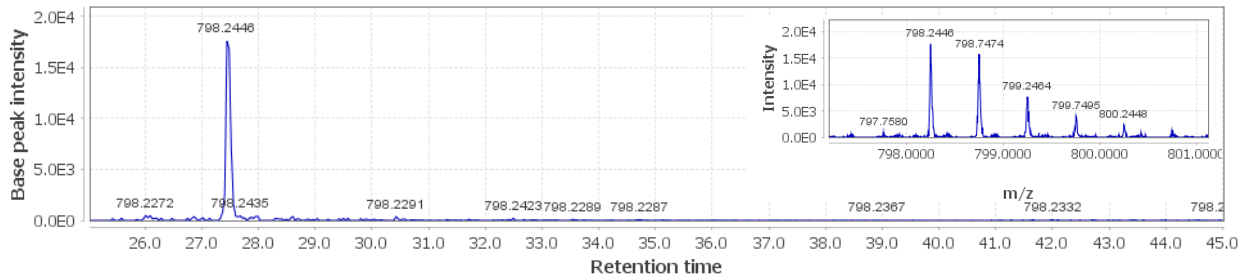
Calc. $[M+2H]^{2+}$: 798.2370 Obs. $[M+2H]^{2+}$: 798.2389 Chemical Formula: $C_{62}H_{80}N_{22}O_{21}S_4^{2+}$



B.4.23. (figure S3, LacHyb2) Condition 1b

MSDITASRVESLAELSEEALGDAVTSLRDTVALPENGAWSGSCWCQAWCWCAQP

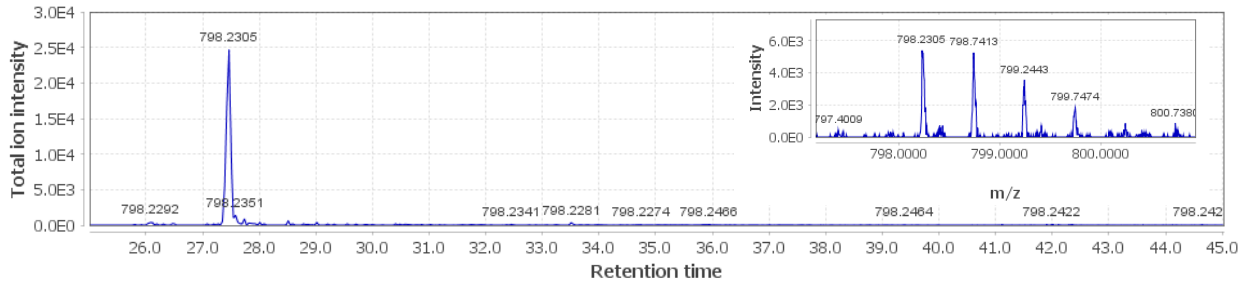
Calc. $[M+2H]^{2+}$: 798.2370 Obs. $[M+2H]^{2+}$: 798.2446 Chemical Formula: $C_{62}H_{80}N_{22}O_{21}S_4^{2+}$



B.4.24. (figure S3, LacHyb3) Condition 1b

MLAELSEEALGDALQDLDLSELTVTSLRDTVALPENGAWSGSCWCQAWCWCAQP

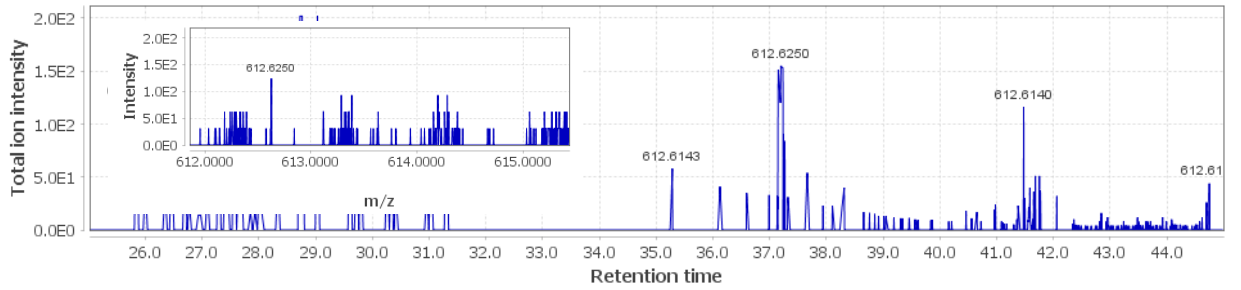
Calc. $[M+2H]^{2+}$: 798.2370 Obs. $[M+2H]^{2+}$: 798.2305 Chemical Formula: $C_{62}H_{80}N_{22}O_{21}S_4^{2+}$



B.4.25. (Table 3.1A, entry 1) Condition 2

MSSQLAELSEEALGDAENEALEIMGAWCTTCVCTCWCCTT

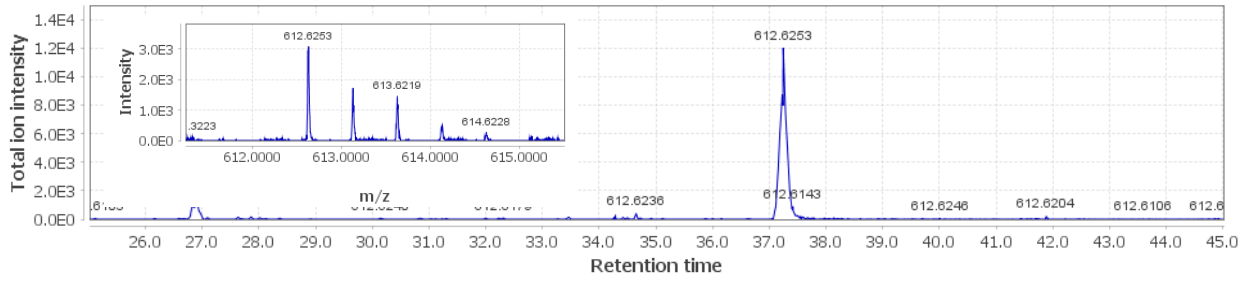
Calc. $[M+2H]^{2+}$: 612.6178 Obs. $[M+2H]^{2+}$: 612.6250 Chemical Formula: $C_{49}H_{55}N_{13}O_{13}S_6^{2+}$



B.4.26. (Table 3.1A, entry 3) Condition 2

MSEIKKALNTLEIEDFLAELSEEALGDAENEALEIMGAWCTTCVCTCWCCTT

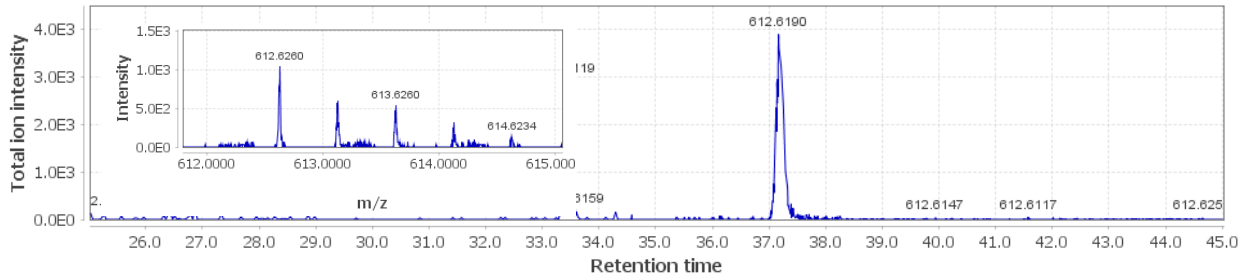
Calc. $[M+2H]^{2+}$: 612.6178 Obs. $[M+2H]^{2+}$: 612.6253 Chemical Formula: $C_{49}H_{55}N_{13}O_{13}S_6^{2+}$



B.4.27. (Table 3.1A, entry 4) Condition 2

MSEIKKALNTLEIEDFDAIEMVDVDAMPLAELSEEALGDAENEALEIMGAWCTTCVCTCWCCTT

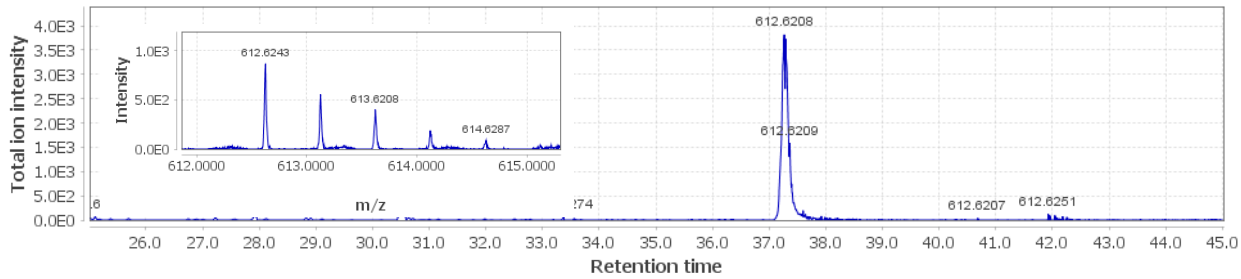
Calc. $[M+2H]^{2+}$: 612.6178 Obs. $[M+2H]^{2+}$: 612.6260 Chemical Formula: $C_{49}H_{55}N_{13}O_{13}S_6^{2+}$



B.4.28. (Table 3.1A, entry 5) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAAAAAAENEALEIMGAWCTTCVCTCWCCTT

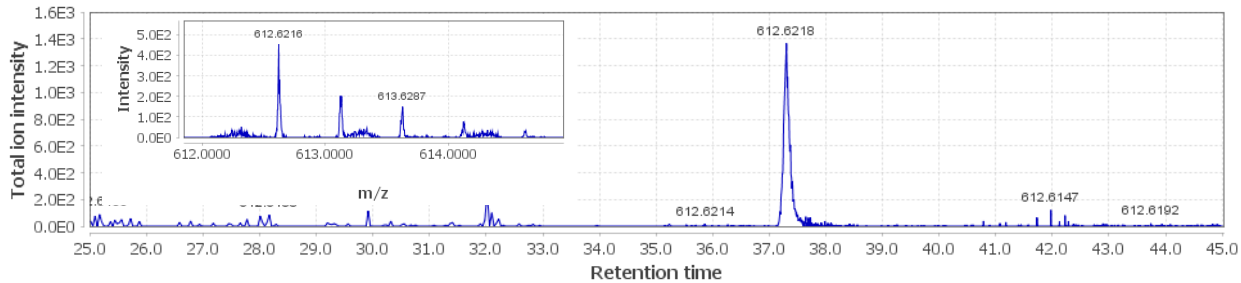
Calc. $[M+2H]^{2+}$: 612.6178 Obs. $[M+2H]^{2+}$: 612.6243 Chemical Formula: $C_{49}H_{55}N_{13}O_{13}S_6^{2+}$



B.4.29. (Table 3.1A, entry 6) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAGGGGGENEALEIMGAWCTTCVCTCWCCTT

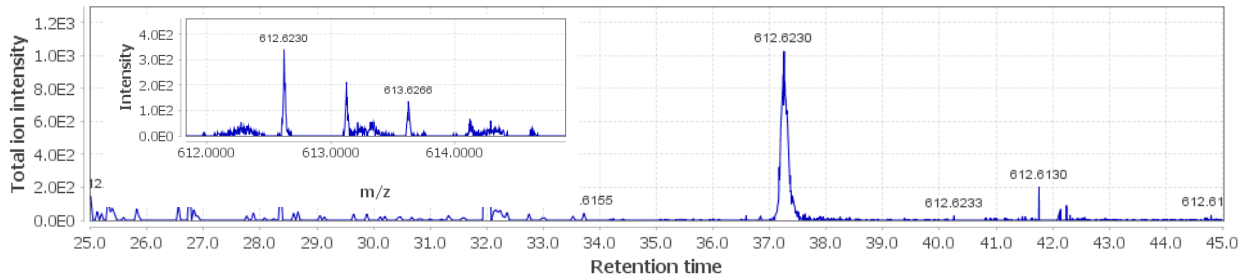
Calc. $[M+2H]^{2+}$: 612.6178 Obs. $[M+2H]^{2+}$: 612.6216 Chemical Formula: $C_{49}H_{55}N_{13}O_{13}S_6^{2+}$



B.4.30. (Table 3.1A, entry 7) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAKKKKKENEALEIMGAWCTTCVCTCWCCTT

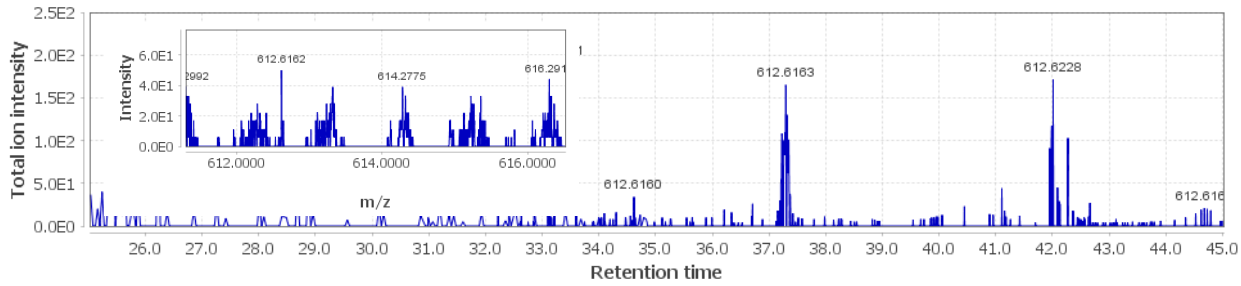
Calc. $[M+2H]^{2+}$: 612.6178 Obs. $[M+2H]^{2+}$: 612.6230 Chemical Formula: $C_{49}H_{55}N_{13}O_{13}S_6^{2+}$



B.4.31. (Table 3.1A, entry 8) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDADDDDDENEALEIMGAWCTTCVCTCWCCTT

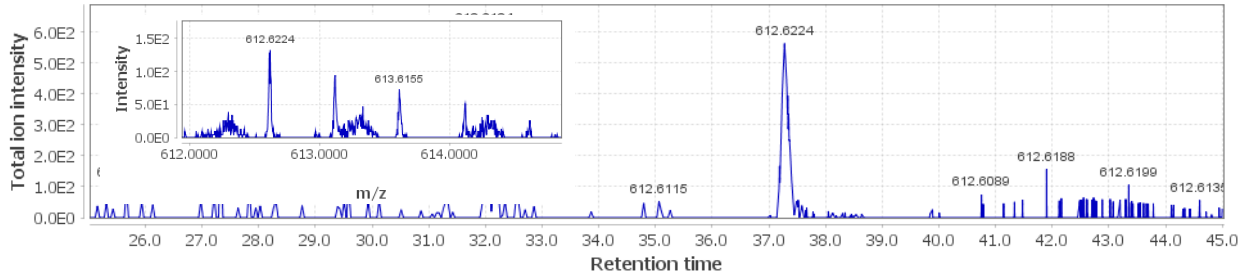
Calc. $[M+2H]^{2+}$: 612.6178 Obs. $[M+2H]^{2+}$: 612.6162 Chemical Formula: $C_{49}H_{55}N_{13}O_{13}S_6^{2+}$



B.4.32. (Table 3.1A, entry 9) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAPPPPPENEALEIMGAWCTTCVCTCWCCTT

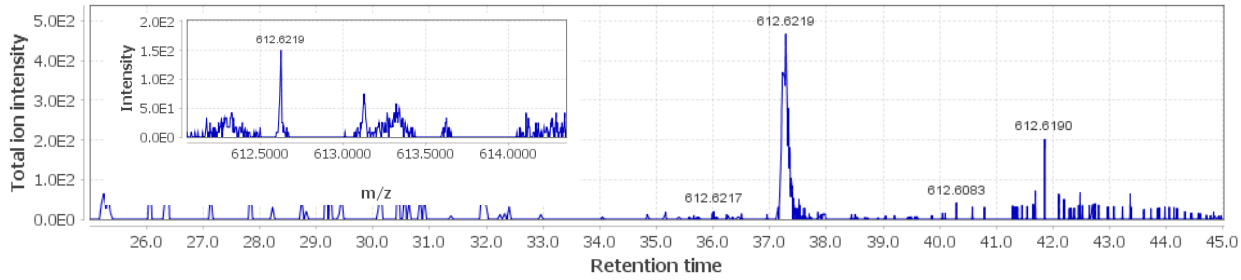
Calc. $[M+2H]^{2+}$: 612.6178 Obs. $[M+2H]^{2+}$: 612.6224 Chemical Formula: $C_{49}H_{55}N_{13}O_{13}S_6^{2+}$



B.4.33. (Table 3.1A, entry 10) Condition 2

MLAELSEEALGDASEIKKALNTLEIEDFDAIEMVDVDAMPENEALEIMGAWCTTCVCTCWCCTT

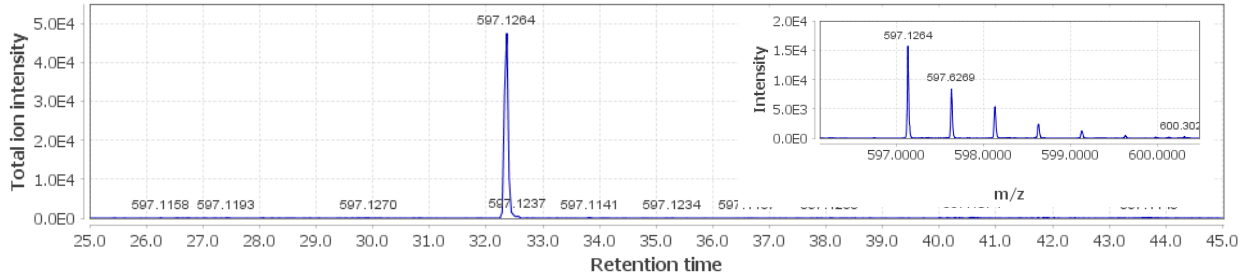
Calc. $[M+2H]^{2+}$: 612.6178 Obs. $[M+2H]^{2+}$: 612.6219 Chemical Formula: $C_{49}H_{55}N_{13}O_{13}S_6^{2+}$



B.4.34. (Table 3.1B, entry 4) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCKCTCWCCAA

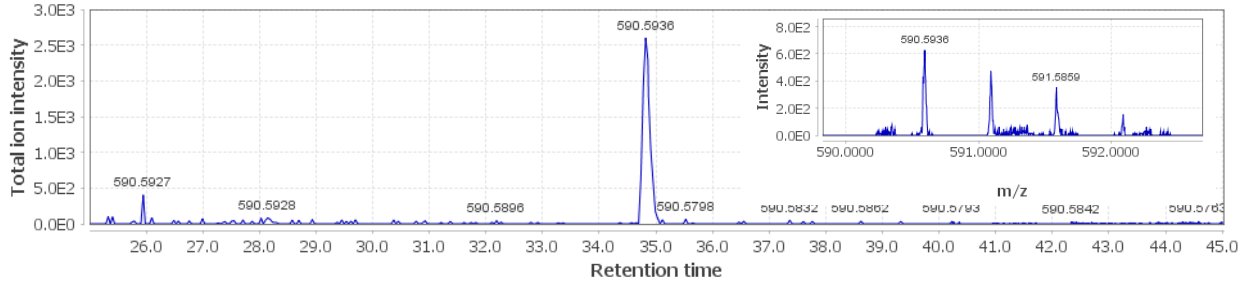
Calc. $[M+2H]^{2+}$: 597.1172 Obs. $[M+2H]^{2+}$: 597.1264 Chemical Formula: $C_{48}H_{54}N_{14}O_{11}S_6^{2+}$



B.4.35. (Table 3.1B, entry 5) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCDCTCWCCAA

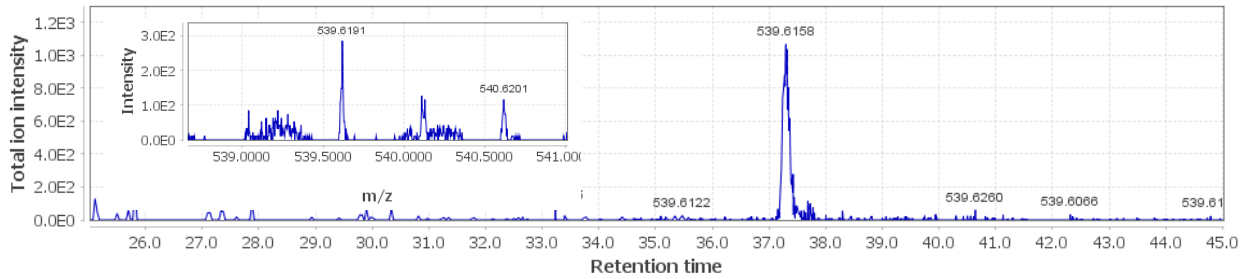
Calc. $[M+2H]^{2+}$: 590.5862 Obs. $[M+2H]^{2+}$: 590.5936 Chemical Formula: $C_{46}H_{47}N_{13}O_{13}S_6^{2+}$



B.4.36. (Table 3.1B, entry 6) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACVSACWCCAA

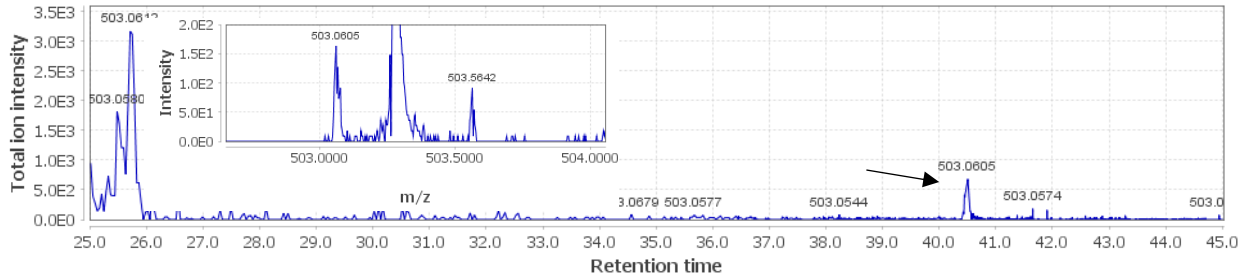
Calc. $[M+2H]^{2+}$: 539.6162 Obs. $[M+2H]^{2+}$: 539.6191 Chemical Formula: $C_{44}H_{49}N_{13}O_{10}S_5^{2+}$



B.4.37. (Table 3.1B, entry 9) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCA

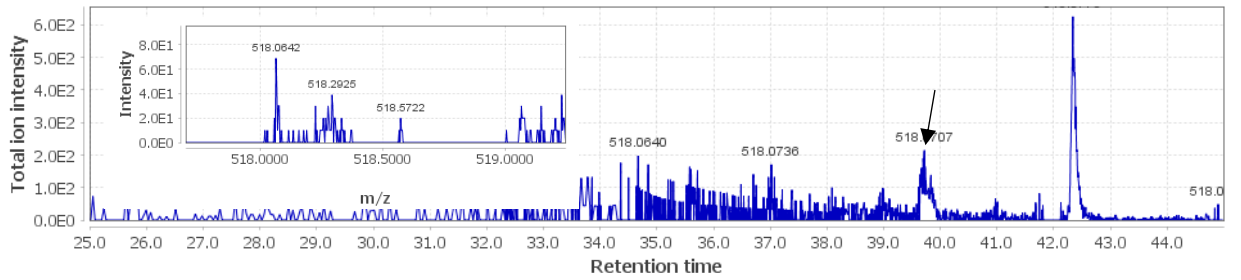
Calc. $[M+2H]^{2+}$: 503.0625 Obs. $[M+2H]^{2+}$: 503.0605 Chemical Formula: $C_{40}H_{38}N_{12}O_8S_6^{2+}$



B.4.38. (Table 3.1B, entry 10) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCT

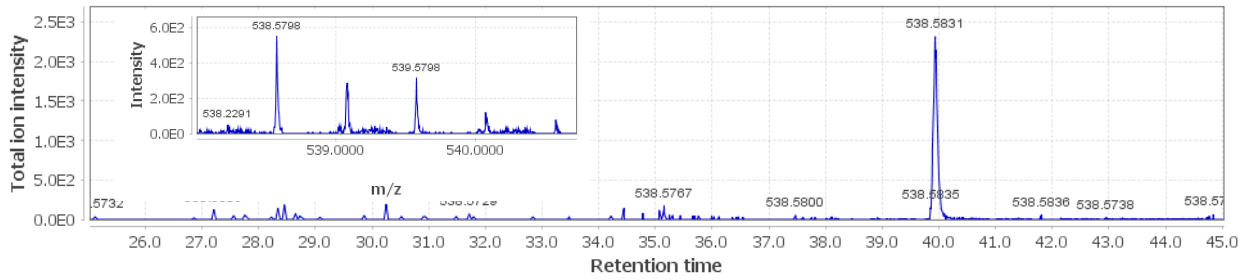
Calc. $[M+2H]^{2+}$: 518.0677 Obs. $[M+2H]^{2+}$: 518.0642 Chemical Formula: $C_{41}H_{40}N_{12}O_9S_6^{2+}$



B.4.39. (Table 3.1B, entry 11) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCAA

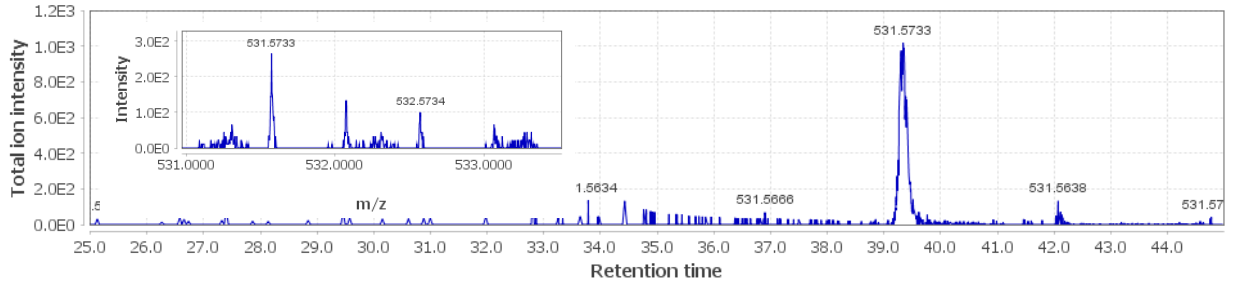
Calc. $[M+2H]^{2+}$: 538.5810 Obs. $[M+2H]^{2+}$: 538.5798 Chemical Formula: $C_{43}H_{43}N_{13}O_9S_6^{2+}$



B.4.40. (Table 3.1B, entry 12) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCAG

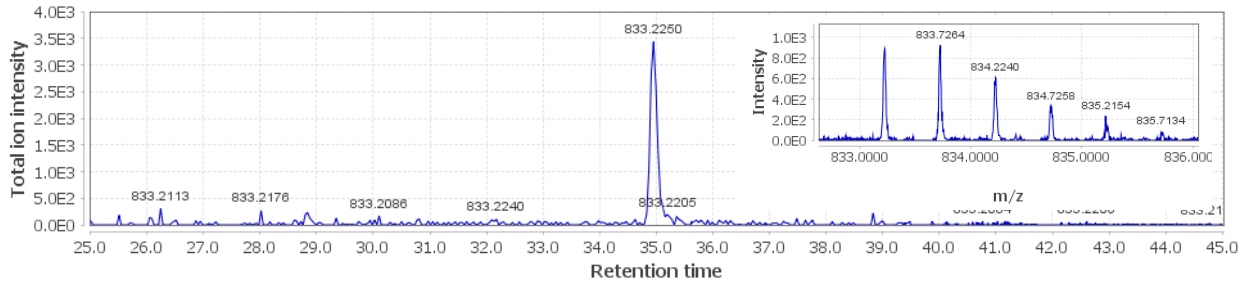
Calc. $[M+2H]^{2+}$: 531.5732 Obs. $[M+2H]^{2+}$: 531.5733 Chemical Formula: $C_{42}H_{41}N_{13}O_9S_6^{2+}$



B.4.41. (Table 3.1B, entry 13) Condition 2

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCVCTCWCCAANSGGVS

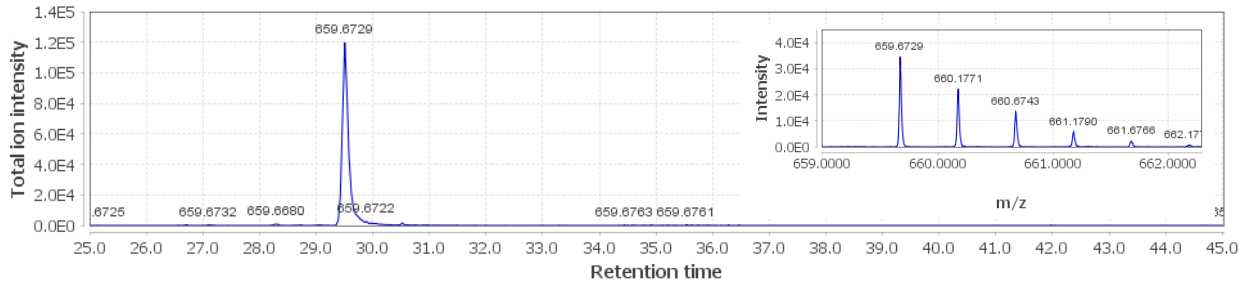
Calc. $[M+2H]^{2+}$: 833.2168 Obs. $[M+2H]^{2+}$: 833.2250 Chemical Formula: $C_{66}H_{82}N_{20}O_{18}S_6^{2+}$



B.4.42. (figure S3, LacHyb1) Condition 2

MSDITASRVESLDLQDLDLAELSEEALRDTVALPENGAWSGSCWCQAWCWCAQP

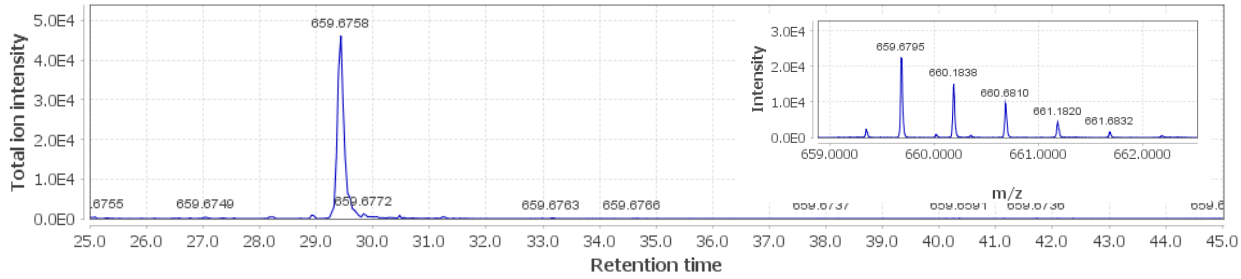
Calc. $[M+2H]^{2+}$: 659.6677 Obs. $[M+2H]^{2+}$: 659.6729 Chemical Formula: $C_{53}H_{61}N_{17}O_{16}S_4^{2+}$



B.4.43. (figure S3, LacHyb2) Condition 2

MSDITASRVESLAELSEEALGDAVTSLRDTVALPENGAWSGSCWCQAWCWCAQP

Calc. $[M+2H]^{2+}$: 659.6677 Obs. $[M+2H]^{2+}$: 659.6795 Chemical Formula: $C_{62}H_{80}N_{22}O_{21}S_4^{2+}$



B.4.44. (figure S3, LacHyb3) Condition 2

MLAELSEEALGDALQDLDLSELTVTSLRDTVALPENGAWSGSCWCQAWCWCAQP

Calc. $[M+2H]^{2+}$: 659.6677 Obs. $[M+2H]^{2+}$: 659.6698 Chemical Formula: $C_{62}H_{80}N_{22}O_{21}S_4^{2+}$

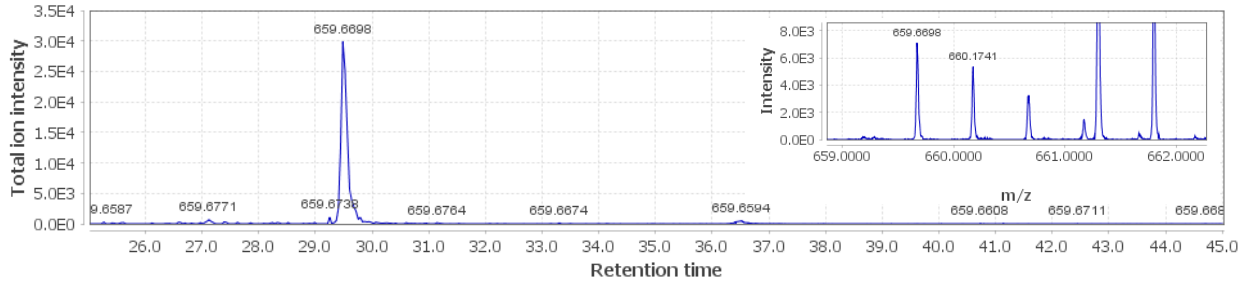
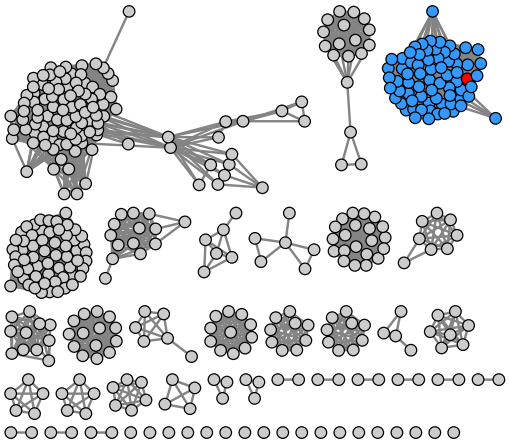
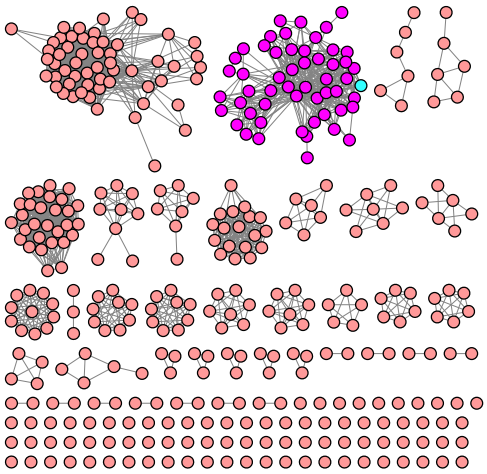


Figure B.5. Lactazole Bioinformatics. (a) Accession numbers from the following report[2] of 508 thiopeptide pyridine synthase homologs were used to bulk download their respective sequences in fasta file format from NCBI. These sequences were then uploaded to Enzyme Function Initiative Enzyme Similarity Tool (EFI-EST) (https://efi.igb.illinois.edu/) to generate a sequence similarity network (SSN) with 100% sequence conflation and an alignment score of 43. LazC is highlighted in red and its cluster of close homologs is highlighted in blue. (b) Full thiopeptide precursor peptides[2] were used to generate a sequence similarity network with alignment score 8 and sequences with 100 % identity were conflated. Lactazole homologs are colored in magenta and lactazole is colored in turquoise. (c) The cores of the lactazole homologs were aligned using ClustalW, Muscle. Sequences were then aligned to the first serine residue and cores larger than 30 residues or beginning with threonine were removed. These sequences were analyzed by Weblogo.

a.



b.



c.

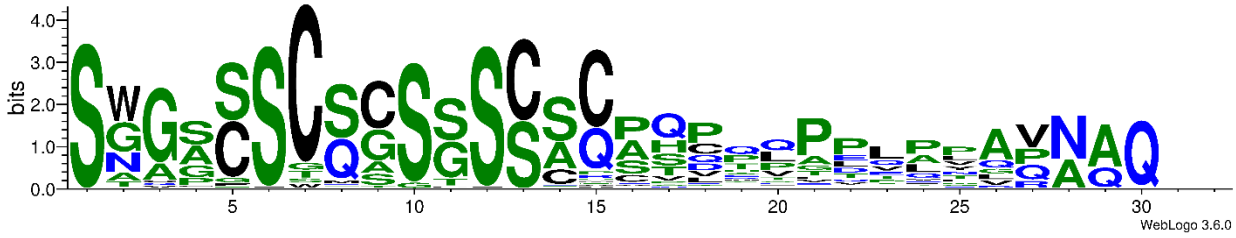


Figure B.6. Flexizyme Reaction. (a) Aminoacyl tRNA synthetases (aaRS) are enzymes which activate amino acids and acylate them onto tRNAs.[3] They are often specific for their evolved amino acid substrate and tRNA. (b) Flexizymes are laboratory evolved RNAs which catalyze the acylation of synthetically activated amino acids onto a tRNA of choice. dFX is a flexizyme which recognizes a wide range of amino acids activated with dinitrobenzyl ester (Dnb).[1]

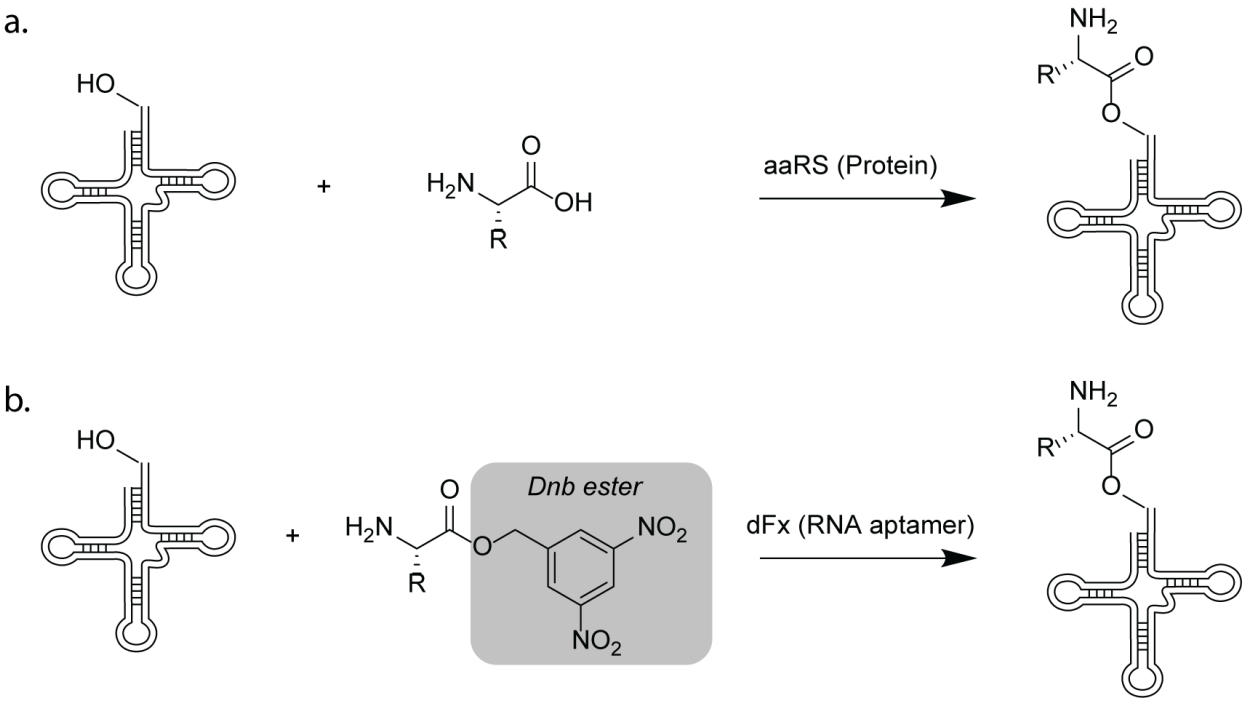


Figure B.7. NMR Spectra of Synthetic Compounds.

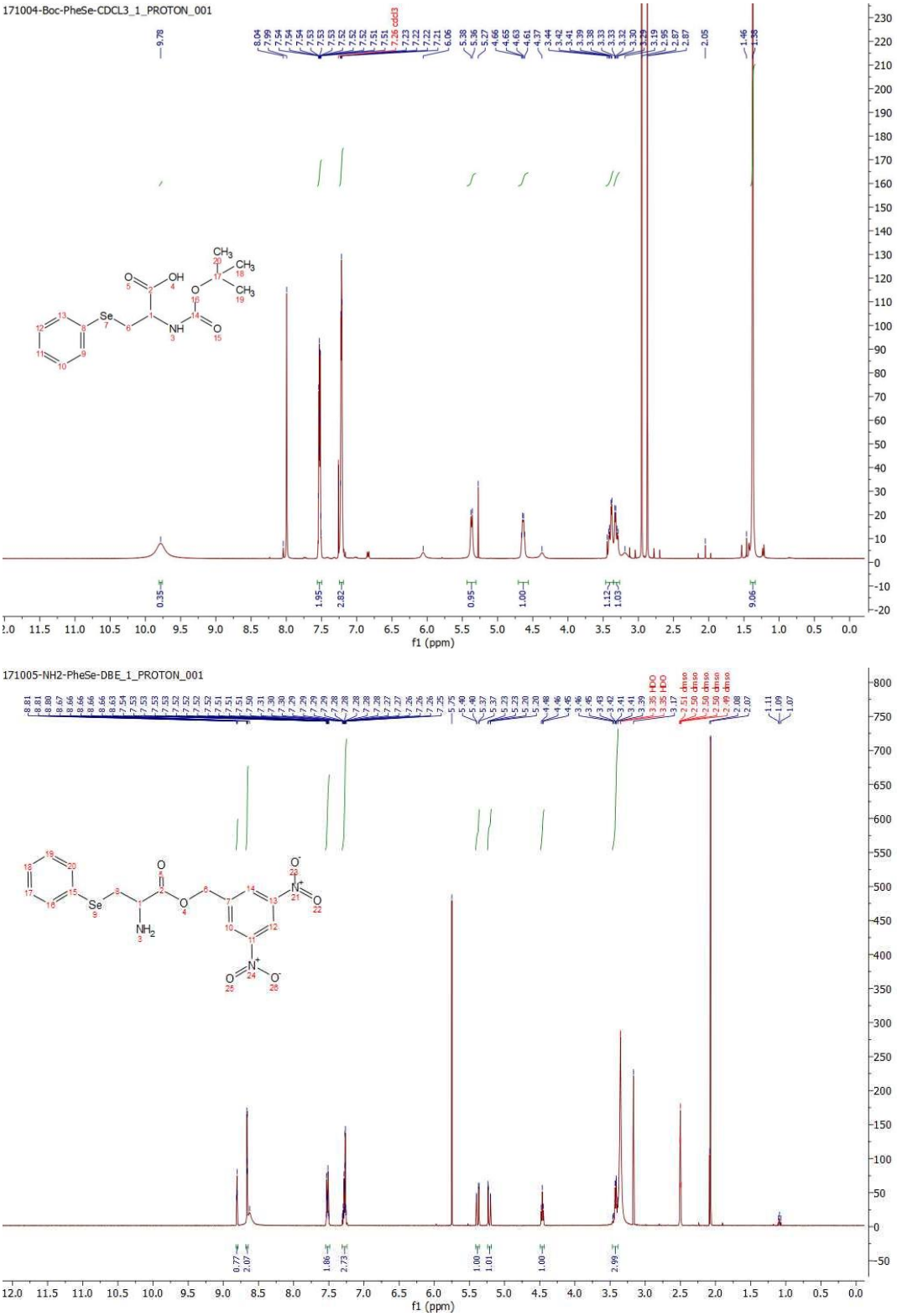


Table B.1. Translation Genes

Table 3.1A, Entry 1:

MSSQLAELSEEALGDAENEALEIMGAWCTTCVCTCWCCTT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGCTCACAGCTGGCAG
AACTTTCAGAAGAGGCTCTTGGCGATGCTGAAAATGAAGCCCTTGAGATCATGGGG
GCGTGGTGCACTACATGCGTGTGCACGTGTTGGTGCTGCACTACATAATAGCGCATT
GGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTC
GAGTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCAC
CGCTGAGCAATAACTAGCATAACC

Table 3.1A, Entry 2 and Table 3.1B, Entry 3:

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCVCTCWCCTT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTTCTGGTGTAGATC
TGGGTACCGAGAACCTGTACTTCCAATCCAATGCGCTGGCCGAGTTATCCGAAGAGG
CACTGGGCGACGCGGAAAACGAAGCGCTTGAAATTATGGGAGCGTGGTGTACGACA
TGCGTATGTACATGCTGGTGTTGTACAACTTAATAGCGCATTGGAAGTGGATAACGG
ATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTG
CTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTA
GCATAACC

Table 3.1A, Entry 3

MSEIKKALNTLEIEDFLAELSEEALGDAENEALEIMGAWCTTCVCTCWCCTT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTGAGATTAAAAAAG
CCTTAAATACCTTAGAGATCGAGGATTTCTTAGCAGAACTGAGTGAAGAAGCGTTAG
GGGACGCTGAAAATGAAGCGTTAGAAATCATGGGTGCATGGTGTACCACATGTGTG
TGTACCTGTTGGTGTTGCACTACGTAATAGCGCATTGGAAGTGGATAACGGATCCGA
ATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACA
AAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAA
CC

Table 3.1A, Entry 4

MSEIKKALNTLEIEDFDAIEMVDVDAMPLAELSEEALGDAENEALEIMGAWCTTCVCTC
WCCTT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCCGAGATTAAAAAAG
CGCTTAATACCTTAGAGATTGAGGATTTCGACGCTATCGAAATGGTTGACGTTGATG

CGATGCCCCTTGCGGAACTGTCAGAAGAAGCGCTGGGCGATGCGGAAAACGAAGCA
TTGGAAATCATGGGGGCCTGGTGTACGACCTGTGTATGTACGTGCTGGTGTTGCACA
ACATAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAG
CTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGA
GTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

Table 3.1A, Entry 5

MSSGVDLGTENLYFQSNALAELSEEALGDAAAAAAENEALEIMGAWCTTCVCTCWCC
TT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGTTCGGGGGTGGACT
TGGGCACCGAGAATTTATATTTCCAATCGAACGCACTTGCTGAGTTGTCTGAAGAGG
CCCTTGGGGACGCGGCTGCCGCAGCGGCGGAGAACGAGGCCCTTGAAATTATGGGT
GCTTGGTGTACCACTTGCGTATGTACCTGTTGGTGCTGTACAACCTAATAGCGCATT
GGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTC
GAGTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCAC
CGCTGAGCAATAACTAGCATAACC

Table 3.1A, Entry 6

MSSGVDLGTENLYFQSNALAELSEEALGDAGGGGGENEALEIMGAWCTTCVCTCWCC
TT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCATCAGGGGTCGACC
TTGGTACGGAAAACCTGTATTTCCAGTCAAACGCCCTTGCCGAGCTGTCGGAGGAAG
CCTTAGGTGATGCCGGAGGGGGAGGAGGTGAAAATGAGGCATTGGAGATTATGGGA
GCGTGGTGTACTACCTGCGTCTGCACATGCTGGTGCTGCACTACTTAATAGCGCATT
GGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTC
GAGTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCAC
CGCTGAGCAATAACTAGCATAACC

Table 3.1A, Entry 7

MSSGVDLGTENLYFQSNALAELSEEALGDAKKKKKENEALEIMGAWCTTCVCTCWCC
TT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGTTCCGGTGTAGACC
TTGGGACGGAAAATCTGTACTTCCAAAGCAATGCGCTTGCTGAGCTGTCTGAGGAGG
CGTTAGGGGATGCCAAGAAAAAGAAAAAGGAAAATGAGGCTCTTGAGATCATGGG
AGCTTGGTGTACAACGTGCGTTTGCACTTGTTGGTGCTGTACCACGTAATAGCGCAT
TGGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACT
CGAGTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCA
CCGCTGAGCAATAACTAGCATAACC

Table 3.1A, Entry 8

MSSGVDLGTENLYFQSNALAELSEEALGDADDDDDENEALEIMGAWCTTCVCTCWCC
TT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGCAGCGGCGTTGATC
TGGGGACCGAAAACTTGTACTTCCAAAGCAACGCTTTAGCTGAGCTTTCTGAGGAAG
CCCTGGGTGACGCCGATGACGATGATGATGAAAATGAGGCTCTGGAGATCATGGGC
GCATGGTGTACTACGTGTGTCTGCACCTGCTGGTGCTGCACTACATAATAGCGCATT
GGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTC
GAGTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCAC
CGCTGAGCAATAACTAGCATAACC

Table 3.1A, Entry 9

MSSGVDLGTENLYFQSNALAELSEEALGDAPPPPPENEALEIMGAWCTTCVCTCWCCTT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTTCCGGTGTTGACC
TGGGCACAGAAAATTTGTATTTCCAATCAAACGCCCTTGCAGAACTGTCAGAAGAA
GCTCTTGGTGACGCGCCTCCACCACCGCCTGAAAATGAAGCATTAGAAATTATGGGT
GCATGGTGTACTACCTGCGTATGTACGTGCTGGTGCTGCACTACCTAATAGCGCATT
GGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTC
GAGTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCAC
CGCTGAGCAATAACTAGCATAACC

Table 3.1A, Entry 10

MLAELSEEALGDASEIKKALNTLEIEDFDAIEMVDVDAMPENEALEIMGAWCTTCVCTC
WCCTT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCTGGCAGAACTTTCGG
AGGAGGCCCTTGGGGACGCTAGTGAGATTAAAAAAGCACTTAATACATTAGAGATT
GAGGACTTCGATGCCATTGAAATGGTCGATGTGGACGCCATGCCGGAAAATGAGGC
TTTAGAAATCATGGGGGCATGGTGTACTACATGTGTTTGTACCTGTTGGTGTTGCAC
CACCTAATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAA
GCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTG
AGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

Table 3.1B, Entry 1

MHHHHHHSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGASCTTCVCTCSCC
TT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCACCATCATCATCATC

ATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCAATGCGCTGG
CGGAACTGTCGGAAGAGGCACTGGGCGACGCGGAAAACGAAGCGCTTGAAATTATG
GGAGCGTCATGTACGACATGCGTATGTACATGCAGTTGTTGTACAACTTAATAGCGC
ATTGGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCA
CTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGA
AGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

Table 3.1B, Entry 2

MHHHHHHSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAFCTTCVCTCFCC
TT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCACCATCATCATCATC
ATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCAATGCGCTGG
CCGAGTTATCCGAAGAGGCACTGGGCGACGCGGAAAACGAAGCGCTTGAAATTATG
GGAGCGTTTTGTACGACATGCGTATGTACATGCTTTTGTTGTACAACTTAATAGCGC
ATTGGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCA
CTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGA
AGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACC

Table 3.1B, Entry 4

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCKCTCWCCAA

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTTCTGGTGTAGATC
TGGGTACCGAGAACCTGTACTTCCAATCCAATGCGCTGGCCGAGTTATCCGAAGAGG
CACTGGGCGACGCGGAAAACGAAGCGCTTGAAATTATGGGAGCGTGG TGT ACA AC
C TGT AAG TGT ACG TGC TGG TGT TGC GCT GCGTAATAGCGCATTGGAAGTGGAT
AACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATC
CGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAA
TAACTAGCATAACC

Table 3.1B, Entry 5

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCDCTCWCCAA

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTTCTGGTGTAGATC
TGGGTACCGAGAACCTGTACTTCCAATCCAATGCGCTGGCCGAGTTATCCGAAGAGG
CACTGGGCGACGCGGAAAACGAAGCGCTTGAAATTATGGGAGCGTGG TGT ACT AC
T TGT GAC TGT ACT TGC TGG TGC TGT GCA GCATAATAGCGCATTGGAAGTGGAT
AACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATC
CGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAA
TAACTAGCATAACC

Table 3.1B, Entry 6

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCAACVSACWCCAA

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTTCTGGTGTAGATTT
GGGAACCGAGAACCTTTATTTTCAAAGCAATGCCCTTGCGGAGTTATCCGAGGAAGC
ATTGGGTGATGCAGAAAACGAAGCACTGGAAATCATGGGGGCATGGTGCGCGGCTT
GCGTGTCGGCGTGTTGGTGTTGTGCAGCTTAATAGCGCATTGGAAGTGGATAACGGA
TCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGC
TAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAG
CATAACC

Table 3.1B, Entry 7

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCA

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATG TCT TCC GGT GTA GA
T TTA GGT ACA GAA AAC CTT TAC TTC CAA TCT AAT GCG CTT GCT GAG CTT TC
T GAG GAG GCT CTG GGCGAT GCC GAG AAT GAG GCC CTT GAG ATC ATG GGA G
CT TGG TGC ACG GCC TGT GCG TGC GCT TGT TGG TGT GCGTAATAGCGCATTGGA
AGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAG
TGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGC
TGAGCAATAACTAGCATAACC

Table 3.1B, Entry 8

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCC

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGTTCTGGTGTAGACT
TGGGAACGGAGAACTTGTATTTTCAATCGAATGCTTTGGCGGAGCTTAGCGAAGAG
GCCCTGGGGGATGCGGAGAACGAAGCATTAGAGATCATGGGGGCATGGTGTACGGC
CTGTGCCTGCGCGTGCTGGTGCTGCTAATAGCGCATTGGAAGTGGATAACGGATCCG
AATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTAAC
AAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATA
ACC

Table 3.1B, Entry 9

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCA

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGTAGTGGTGTCGATC
TGGGGACCGAGAATTTGTATTTTCAATCTAATGCACTGGCGGAATTAAGCGAAGAA
GCACTTGGGGACGCCGAAAACGAAGCGTTGGAGATCATGGGTGCGTGGTGTACGGC
CTGCGCGTGTGCCTGTTGGTGTTGCGCGTAATAGCGCATTGGAAGTGGATAACGGAT

CCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCT
AACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGC
ATAACC

Table 3.1B, Entry 10

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCT

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTTCTGGGGTTGATC
TTGGTACGGAGAACTTATACTTCCAATCAAATGCCTTAGCGGAATTATCCGAAGAGG
CACTTGGGGATGCAGAGAACGAAGCCCTGGAGATTATGGGCGCATGGTGTACCGCC
TGCGCGTGCGCGTGTTGGTGTTGTACCTAATAGCGCATTGGAAGTGGATAACGGATC
CGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGCTA
ACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCA
TAACC

Table 3.1B, Entry 11

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCAA

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCGAGCGGTGTCGATT
TGGGTACTGAGAACTTGTACTTTCAAAGTAACGCCTTAGCTGAATTATCAGAGGAAG
CTCTTGGTGACGCTGAGAACGAAGCCCTTGAAATTATGGGAGCTTGGTGCACTGCGT
GTGCTTGTGCCTGTTGGTGCTGCGCTGCATAATAGCGCATTGGAAGTGGATAACGGA
TCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTGC
TAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAG
CATAACC

Table 3.1B, Entry 12

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTACACACWCCAG

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTTCCGGTGTTGATTT
GGGAACGGAGAATTTATATTTTCAATCGAACGCTCTGGCGGAACTTAGCGAGGAGG
CTTTGGGGGACGCTGAAAATGAGGCACTGGAGATCATGGGAGCCTGGTGTACTGCA
TGTGCGTGTGCTTGCTGGTGTTGCGCCGGCTAATAGCGCATTGGAAGTGGATAACGG
ATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTG
CTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTA
GCATAACC

Table 3.1B, Entry 13

MSSGVDLGTENLYFQSNALAELSEEALGDAENEALEIMGAWCTTCVCTCWCCAANSG
GVS

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTTCTGGTGTAGATC
TGGGTACCGAGAACCTGTACTTCCAATCCAATGCGCTGGCCGAGTTATCCGAAGAGG
CACTGGGCGACGCGGAAAACGAAGCGCTTGAAATTATGGGAGCGTGG TGT ACC AC
G TGT GTA TGT ACG TGT TGG TGT TGT GCC GCT AAC AGT GGC GGG GTA TCGTA
ATAGCGCATTGGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGC
GGCCGCACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGG
CTGCTGCCACCGCTGAGCAATAACTAGCATAACC

LacHyb1

MSDITASRVESLDLQDLDLAELSEEALRDTVALPENGAWSGSCWCQAWCWCAQP

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTGACATCACCGCCA
GTCGCGTAGAGTCGCTTGACCTGCAAGACCTTGACTTGGCTGAGTTATCAGAGGAAG
CGTTACGTGATACGGTTGCTCTTCCTGAAAACGGAGCTTGGTCTGGGAGTTGTTGGT
GCCAAGCGTGGTGCTGGTGTGCCCAACCGTAATAGCGCATTGGAAGTGGATAACGG
ATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTG
CTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTA
GCATAACC

LacHyb2

MSDITASRVESLAELSEEALGDAVTSLRDTVALPENGAWSGSCWCQAWCWCAQP

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATG AGT GAT ATT ACA GC
C AGT CGC GTA GAG AGT TTA GCG GAA TTG TCC GAA GAA GCA CTT GGA GAC G
CA GTA ACA TCC TTG CGCGAT ACT GTT GCG TTG CCT GAG AAT GGT GCG TGG T
CA GGA TCA TGT TGG TGC CAG GCT TGG TGC TGG TGC GCT CAA CCATAATAGCG
CATTGGAAGTGGATAACGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGC
ACTCGAGTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTG
CCACCGCTGAGCAATAACTAGCATAACC

LacHyb3

MLAELSEEALGDALQDLDLSELTVTSLRDTVALPENGAWSGSCWCQAWCWCAQP

GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCTTGCAGAACTTAGCG
AGGAGGCACTGGGAGATGCGCTTCAAGACCTTGACTTATCCGAACTTACAGTCACCA
GCCTGCGTGACACCGTCGCTCTTCCCGAGAACGGTGCCTGGTCAGGCTCATGTTGGT
GCCAAGCCTGGTGTTGGTGTGCTCAACCCTAATAGCGCATTGGAAGTGGATAACGG
ATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGTGAGATCCGGCTG
CTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTA
GCATAACC

REFERENCES

(1) Goto, Y.; Katoh, T.; Suga, H. *Nat. Protoc.* **2011**, *6*, 779.

(2) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. *J. Am. Chem. Soc.* **2018**, *140*, 9494.

(3) Rajendran, V.; Kalita, P.; Shukla, H.; Kumar, A.; Tripathi, T. *Int. J. Biol. Macromol.* **2018**, *111*, 400.

APPENDIX C. SUPPLEMENTARY TABLES AND FIGURES FOR CHAPTER 4

Table C.1. Precursor SSN table of statistics. For each cluster in the precursor ssn (figure 4.3), the average RODEO predicted thiopeptide score was calculated, the number of peptides in each cluster was counted, and the average number of cys, ser, and thr residues are reported.

Cluster	Thiopeptide Score Ave.	Thiopeptide Score Stdev.	'C' Ave.	'C' Stdev.	'S' Ave.
clust_1	26.67	6.04	3.79	1.72	8.13
clust_2	26.12	5.80	4.35	1.94	5.87
clust_3	32.68	4.35	3.39	1.47	8.84
clust_4	32.33	1.69	4.74	0.86	10.41
clust_5	30.33	1.63	8.04	0.62	16.13
clust_6	30.81	1.68	2.90	0.34	13.40
clust_7	13.43	7.55	0.14	0.38	2.29
clust_8	11.81	2.54	2.00	0.63	9.81
clust_9	30.56	2.88	7.78	2.64	4.11
clust_10	30.18	1.59	8.18	0.53	15.29
clust_11	32.38	4.17	5.38	0.51	4.46
clust_12	35.26	1.85	5.00	0.00	6.00
clust_13	31.67	1.58	4.67	0.50	12.11
clust_14	29.41	1.99	1.00	0.00	9.00
clust_15	31.59	1.06	8.00	0.00	14.88
clust_16	32.75	1.60	8.50	0.52	2.67
clust_17	23.93	1.82	3.00	0.00	6.07
clust_18	36.73	0.92	7.88	0.33	4.12
clust_19	30.50	0.94	7.00	0.00	18.93
clust_20	35.47	2.33	2.40	0.63	8.40
clust_21	30.00	0.58	6.00	0.00	2.43
clust_22	28.00	0.00	10.00	0.00	4.40
clust_23	29.14	2.27	6.00	0.00	11.86
clust_24	32.63	2.07	7.63	0.74	16.75
clust_25	21.00	5.18	3.00	0.00	7.50
clust_26	31.00	0.00	5.00	0.00	10.00
clust_27	26.20	1.92	5.40	2.07	12.40
clust_28	27.83	0.75	6.00	0.00	9.50
clust_29	33.00	0.63	1.00	0.00	10.50
clust_30	28.75	3.30	4.00	0.00	2.50
clust_31	19.00	0.00	2.00	0.00	3.00
clust_32	22.67	0.58	0.67	0.58	12.33
clust_33	29.50	2.74	3.00	0.00	6.50
clust_34	30.00	1.73	8.00	2.65	14.00
clust_35	2.33	1.15	0.67	1.15	6.00
clust_36	25.20	0.84	6.00	0.00	3.00
clust_37	30.00	1.00	6.00	1.73	20.67
clust_38	26.75	0.50	3.00	0.00	6.00

Cluster	'S' Stdev.	'T' Ave.	'T' Stdev.	'CST' Ave.	'CST' Stdev.	Total Peptide
clust_1	3.04	1.77	1.97	13.69	5.33	197
clust_2	2.22	1.42	1.35	11.64	2.94	105
clust_3	1.76	1.79	1.11	14.02	2.05	97
clust_4	1.12	2.48	0.70	17.63	1.15	27
clust_5	1.60	5.21	1.02	29.38	1.66	24
clust_6	1.50	8.15	1.01	24.45	1.95	105
clust_7	0.49	0.71	0.76	3.14	1.35	7
clust_8	1.81	3.33	1.24	15.14	1.49	21
clust_9	2.09	1.56	0.53	13.44	2.13	9
clust_10	0.47	5.71	0.77	29.18	0.53	17
clust_11	0.52	2.00	0.00	11.85	0.38	13
clust_12	0.88	3.26	0.99	14.26	1.66	19
clust_13	1.36	7.33	1.00	24.11	1.36	9
clust_14	0.00	3.00	1.02	13.00	1.02	22
clust_15	0.49	3.41	0.71	26.29	0.69	17
clust_16	0.49	0.58	0.51	11.75	0.97	12
clust_17	0.27	1.43	0.51	10.50	0.65	14
clust_18	0.33	1.15	0.46	13.15	0.37	26
clust_19	0.27	6.50	0.52	32.43	0.51	14
clust_20	0.91	4.33	1.91	15.13	1.51	15
clust_21	1.13	5.29	0.76	13.71	1.89	7
clust_22	0.51	0.00	0.00	14.40	0.51	15
clust_23	1.77	2.00	0.00	19.86	1.77	7
clust_24	0.71	4.63	1.19	29.00	1.31	8
clust_25	0.84	0.33	0.82	10.83	0.98	6
clust_26	0.00	3.00	0.00	18.00	0.00	6
clust_27	2.79	5.20	1.30	23.00	5.10	5
clust_28	0.55	2.00	1.10	17.50	1.05	6
clust_29	0.55	3.33	0.52	14.83	0.41	6
clust_30	0.58	2.00	0.00	8.50	0.58	4
clust_31	0.00	0.00	0.00	5.00	0.00	3
clust_32	1.15	3.67	0.58	16.67	1.53	3
clust_33	0.55	4.00	0.00	13.50	0.55	6
clust_34	2.65	4.33	2.31	26.33	2.89	3
clust_35	1.73	2.67	0.58	9.33	1.15	3
clust_36	0.00	0.00	0.00	9.00	0.00	5
clust_37	3.21	9.67	0.58	36.33	1.15	3
clust_38	0.00	1.25	0.50	10.25	0.50	4

Table C.2. Sequences of RODEO predicted thiopeptide precursor peptides. Each entry contains a unique number, "X_" followed by the NCBI accession number for the associated pyridine synthase. Additionally, the bacterial species is presented, followed by the genome nucleotide sequence for the predicted peptide, and the RODEO thiopeptide score (TS).

1_WP_093941962.1 Actinoalloteichus hoggarensis 3540491 3540491 TS=39
MMNANSVASVDVGIEGLTGLDVDTLEIGDYLDETVLDTTDLTETMTSSSSCTTCICTCSC
SS

2_WP_083975908.1 Kitasatospora azatica KCTC 9699 517547 517547 TS=39
MEGFDNLVLDDSAFELVDLGEVVPASVTAGYGLTELSASKGSSSCCCCCPCCSCT

3_EJJ05679.1 Streptomyces auratus AGR0001 3879982 3879982 TS=39
MDAVEELDLNALEISDLIEDSNYSDETLSQAMAASTTTTGCTTSSCSSSSS

4_WP_091614370.1 Amycolatopsis saalfeldensis 714664 714664 TS=38
MSAPTSIQANVGVAGLTGLDTDTLEISDYLDETLLDSHDLTETMASSSSCTTCICTCSCSS

5_WP_133205963.1 Arthrobacter sp. JH1-1 25821 25821 TS=38
MTITLPDLDFDGIEVLSVRESVAVPETGATSGSSSSNSSSCCGSSSCCSS

6_WP_083975908.1 Kitasatospora azatica KCTC 9699 517775 517775 TS=38
MEGFDNLVLDDSTFELVDLGEVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

7_WP_007074886.1 Micromonospora sp. ATCC 39149 4703451 4703451 TS=38
MDANATTVGVDGLTGLDIDTFEIGDYLDETVLDTADLTETMTSSSSCTTCICTCSCSS

8_WP_153464103.1 Streptomyces kaniharaensis 4883725 4883725 TS=38
MEGFEDLVLDDGAFELVDLGEVVPTSVTAGYGLTELSASGGQSSCCCCCPCCSCT

9_WP_153464103.1 Streptomyces kaniharaensis 4883953 4883953 TS=38
MEGFEDLVLDDGAFELVDLGEVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

10_WP_073444392.1 Streptomyces noursei 494011 494011 TS=38
MDSTQELDLNALEISDLIEDSNYSDETLSQAMAASTTTTGCTTSSCSSSSS

11_WP_073444392.1 Streptomyces noursei 494258 494258 TS=38
MDSTQELDLNALEISDLIEDSNYSDETLSQAMAASTTTTGCTTSSCSSSSS

12_WP_067348540.1 Streptomyces noursei ATCC 11455 4423573 4423573 TS=38
MDSTQELDLNALEISDLIEDSNYSDETLSQAMAASTTTTGCTTSSCSSSSS

13_WP_067348540.1 Streptomyces noursei ATCC 11455 4423821 4423821 TS=38
MDSTQELDLNALEISDLIEDSNYSDETLSQAMAASTTTTGCTTSSCSSSSS

14_WP_168712275.1 Streptomyces sp. A1277 164519 164519 TS=38
MSTLANETPAGDLFASFDIAELEVLEVSDSVGLPEMGASSSSLSVASEAAADVSLCCSSS
SSTCCCC

15_WP_073741099.1 Streptomyces sp. CB02115 809522 809522 TS=38
MFDSFDIEELDTMEIADGVALPEMAASSLPVTCCSSSSSSCCCC

16_WP_058927704.1 Streptomyces sp. CdTB01 9686259 9686259 TS=38
MEEFDELVLDDSAFELVDLGKVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

17_WP_058927704.1 Streptomyces sp. CdTB01 9686492 9686492 TS=38
MEEFDELVLDDSAFELVDLGKVVPASVTAGYGLTELSASGGQSSCCCCCPCCSCT

18_WP_047180975.1 Streptomyces sp. MNU77 2814790 2814790 TS=38
MRSTLETADVDAMFDSFDIEELDTMEIADGVALPEMAASSLPVTCCSSSSSSCCCC

19_WP_097963940.1 Streptomyces sp. or20 78898 78898 TS=38
MDVDAMFDSFDIEELDTMEIADGVALPEMAASSLPVTCCSSSSSSCCCC

20_WP_050360717.1 Streptomyces sp. or3 87798 87798 TS=38
MFDSFDIEELDTMEIADGVALPEMAASSLPVTCCSSSSSSCCCC

21_WP_109601719.1 Actinoplanes xinjiangensis 67568 67568 TS=37
MSAELRATETNAADLSALDIESLEISEFLDDNRLEDSEVVAKVMSASCTTCECCCSCSS

22_WP_034091981.1 Streptacidiphilus albus 87319 87319 TS=37
MEGFDDLVLDDSSFELVDLGAVVPESVTAGYGLTELSASGGQSSCCCCCPCCSCT

23_WP_034091981.1 Streptacidiphilus albus 87541 87541 TS=37
MEGFDDLVLDDSSFELVDLGAVVPESVTAGYGLTELSASGGASSCCCCCPCCSCT

24_WP_110629793.1 Streptomyces actuosus 4914932 4914932 TS=37
MDAAHLSDLDIDALEISEFLDESRLEDSEVVAKVMSASCTTCECCCSCSS

25_WP_030578091.1 Streptomyces globisporus subsp. globisporus 151064 151064 TS=37
MDAVESLDLNALEISDLIEDSNYSDETLSQAMAASTTTTGCTTSSCSSSSS

26_WP_120757588.1 Streptomyces klenkii 92932 92932 TS=37
MEKNGIEGLDLSDLDVESLEFSADSEGGLESLRMGHGLTEVGASCCCSTSSCSCST

27_WP_037746661.1 Streptomyces mirabilis 82938 82938 TS=37
MEEFDELVLDDSAFELVDLGKVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

28_WP_037746661.1 Streptomyces mirabilis 83159 83159 TS=37
MEEFDELVLDDSAFELVDLGKVVPASVTAGYGLTELSASGGQSSCCCCCPCCSCT

29_WP_097239509.1 Streptomyces sp. 1114.5 4871787 4871787 TS=37
MEDFDNLVLEDSAFELVDLGEVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

30_WP_097239509.1 Streptomyces sp. 1114.5 4872016 4872016 TS=37
MEDFDNLVLEDSAFELVDLGEVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

31_WP_148643628.1 Streptomyces sp. CB01881 1909152 1909152 TS=37
MTEHHRAGRTESPADPTELEIEELSVLELGDAIALPEMGASNWGGWFCCSSSSSATCCC

32_WP_148643628.1 Streptomyces sp. CB01881 1909525 1909525 TS=37
MSESINTPQVESAEVPAVDVEELSVLEVGDAAALPEMGASSGGTMCCSSSSSTCCC

33_WP_148643628.1 Streptomyces sp. CB01881 1909820 1909820 TS=37
MSESVNTPQVESAEVPAVDVEELSVLEVGDAAALPEMGASSGGTMCCSSSSSTCCC

34_WP_087767808.1 Streptomyces sp. CS057 2994009 2994009 TS=37
MRSTLETADVDAMFDSFDIEELDTMEIADGVALPEMAASSIPVTCCSSSSSSCCCC

35_WP_051094648.1 Streptomyces sp. SID4913 68505 68505 TS=37
MSTLASETPAGGDLFASFDIDELEVLEVSDSVGLPEMGASSSSLSVASEAAADVSLCCSS
SSSTCCCC

36_WP_079277171.1 Streptomyces sp. TSRI0261 218301 218301 TS=37
MRSTLETADVDAMFDSFDIEELDTMEIADGVALPEMAASSLPATCCSSSSSSCCCC

37_WP_104482524.1 Actinokineospora auranticolor 89137 89137 TS=36
MNDHKGLASLADEILRLETETFEISDYADTNDVVLAACTSSSSTSTCSSTTSTTSCSA

38_WP_151570259.1 Actinomadura rudentiformis 75901 75901 TS=36
MNETDDSGTRSLLLEDLGLLEAETFEVRDHSDVAEELAGTCSCTSTTSSCSSSSGTSCCD

39_WP_151570259.1 Actinomadura rudentiformis 76134 76134 TS=36
MSDTDEDRGTHTVLQAELGLLEAETFEISDYVETIEEVAGTTSCTSTCSSCSSSSSTSCS

40_WP_027346917.1 Hamadaea tsunoensis DSM 44101 222558 222558 TS=36
MPTQDLELMDLDLDLEIVEISSVRDAVALPESGASSGTSSCGSSSCCGSCSCCCCL

41_WP_033822905.1 Kitasatospora sp. MBT63 24103 24103 TS=36
MEGFDDLVLDDGAFELVDLGEVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

42_WP_033822905.1 Kitasatospora sp. MBT63 24328 24328 TS=36
MEGFDDLVLDDGAFELVDLGEVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

43_WP_145908795.1 Kitasatospora viridis 310559 310559 TS=36
MEGFDDLVLDDSTFELVDLGVVVPASVTAGYGLTELAASSGGSSCCCCTPCCSCT

44_WP_145908795.1 Kitasatospora viridis 310781 310781 TS=36
MEGFDDLVLDDSTFELVDLGVVVPASVTAGYGLTELAASSGGSSCCCCTPCCSCT

45_WP_145908795.1 Kitasatospora viridis 311005 311005 TS=36
MEGFNDDLVLDDSAFELVDLGAVVPASVTAGYGLTELSASKGTSSCCCCTPCCSCT

46_WP_140008476.1 Nocardioides plantarum 1542643 1542643 TS=36
MKNANAFAAFDIAELEVLETAQGMALPEMGASSGSIGTSSSSSSTCSAC

47_WP_143802042.1 Paenibacillus thiaminolyticus 1669884 1669884 TS=36
MEELKLENQEIELSLDDLELQDVKMEVLMTNSKGLPEMGATLGGITCCSCCCCST

48_WP_152645972.1 Streptacidiphilus albus JL83 9685908 9685908 TS=36
MEGFDDLVLDDSSFELVDLGAVVPESVTAGYGLTELSASGGASSCCCCCPCCSCT

49_WP_152645972.1 Streptacidiphilus albus JL83 9686130 9686130 TS=36
MEGFDDLVLDDSSFELVDLGAVVPESVTAGYGLTELSASGGQSSCCCCCPCCSCT

50_WP_037604491.1 Streptacidiphilus rugosus AM-16 1514789 1514789 TS=36
MEGFEDLVLDDGAFELVDLGEVVPASVTAGYGLTELSASGGQSSCCCCCPCCSCT

51_WP_037604491.1 Streptacidiphilus rugosus AM-16 1515016 1515016 TS=36
MEGFEDLVLDDGAFELVDLGEVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

52_WP_020274901.1 Streptomyces afghaniensis 772 19644 19644 TS=36
MHLSELDVDALEISEFLDDSRLEDSEVVAKVMSASCTTCECCCSCSS

53_WP_167523738.1 Streptomyces brevispora 5176660 5176660 TS=36
MTILASETPAGGDLFASFDIDELEVLEVSDSVGLPEMGASSSSLSVASEAVGDISLCCSSS
SSTCCCC

54_RXS65595.1 Streptomyces sioyaensis 24176 24176 TS=36
MSTAAIVGQEIGVDGLTGLDVDALEISDYMDETLLDGEDLSVTMVSSASCTTCICTCSCS
S

55_ACN80649.1 Streptomyces sioyaensis 18074 18074 TS=36
MSTAAIVGQEIGVDGLTGLDVDALEISDYMDETLLDGEDLSVTMVSSASCTTCICTCSCS
S

56_AQM75226.1 Streptomyces sp. 6618 6618 TS=36
MDAELNGLDIDALEISEFLDENRLEDSEVVAKVMSASCTTCECCCSCSS

57_WP_097239513.1 Streptomyces sp. 1331.2 7264409 7264409 TS=36
MEDFDNLVLEDSAFELVDLGEVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

58_WP_097239513.1 Streptomyces sp. 1331.2 7264638 7264638 TS=36
MEDFDNLVLEDSAFELVDLGEVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

59_WP_093484014.1 Streptomyces sp. 2314.4 4587597 4587597 TS=36
MSTAAIVGQEIGVEGLTGLDVDALEISEYMDETLLDGEDLSVTMVSSASCTTCICTCSCS
S

60_WP_096213443.1 Streptomyces sp. 2323.1 4190398 4190398 TS=36
MSTAAIVGQEIGVEGLTGLDVDALEISEYMDETLLDGEDLSVTMVSSASCTTCICTCSCS
S

61_WP_159032229.1 Streptomyces sp. 2333.5 3859662 3859662 TS=36
MSTAAIVGQEIGVEGLTGLDVDALEISEYMDETLLDGEDLSVTMVSSASCTTCICTCSCS
S

62_WP_123493429.1 Streptomyces sp. 844.5 1163845 1163845 TS=36
MEGFDNLVLDDSAFELVDLGEVVPASVTAGYGLTELSASGGASSCCCCCPCCSCT

63_WP_123493429.1 Streptomyces sp. 844.5 1164062 1164062 TS=36
VNDMEGFDNLVLDDSAFELVDLGEVVPASVTAGYGLTELSASGGQSSCCCCCPCCSCT

64_OKI60309.1 Streptomyces sp. CB00072 834561 834561 TS=36
VPGTGGKTHAPPVFRPAGQELGGTHKMRSTLETVDVDAMFDSFDIEELDTMEIADGVA
LPEMAASSIPVTCCSSSSSSCCCC

65_WP_147246070.1 Streptomyces sp. ISID311 378458 378458 TS=36
MSTAAIVGQEIGVEGLTGLDVDALEISEYMDETLLDGEDLSVTMVSSASCTTCICTCSCS
S

66_RXS86842.1 Streptomyces sp. TM32 24169 24169 TS=36
MSTAAIVGQEIGVDGLTGLDVDALEISDYMDETLLDGEDLSVTMVSSASCTTCICTCSCS
S

67_WP_037764021.1 Streptomyces sp. VN1 3712729 3712729 TS=36
MSTVAIDSADIEQVFDSFDVTELEVLEVSQGVALPEMGASGGTVGTSSSSSTCSSSTC

68_WP_125824502.1 Streptomyces sp. W1SF4 192883 192883 TS=36
MANTTAVADFDQLFTSFDIEELEVLDVADGVALPEMGASSGNALCCSSSSSSSSSCC

69_WP_030891366.1 Streptomyces varsoviensis 52660 52660 TS=36
MEGFDDLVLDDSAFELVDLGEVVPESVTAGYGLTELSASGGASSCCCCCPCCSCT

70_WP_094236693.1 Tumebacillus algifaecis 2652710 2652710 TS=36
MEIKALEVTGLEVMDLNDAMALPETAASSGSSSSSGCSTCGSSSCCGSC

71_WP_094236693.1 Tumebacillus algifaecis 2652905 2652905 TS=36
MEIKALEVTGLEVMDLNDAMALPETAASSGSSSSSGCSTCGSSSCCGSC

72_WP_094236693.1 Tumebacillus algifaecis 2653102 2653102 TS=36
MEIKALEVTGLEVMDLNDAMALPETAASSGSSSSSGCSTCGSSSCCGSC

73_WP_087458049.1 Tumebacillus avium 3770985 3770985 TS=36
MEIKALEVTGLEVMDLNDAMALPETAASSGSSSSSGCSTCGSSSCCGSC

74_WP_087458049.1 Tumebacillus avium 3771181 3771181 TS=36
MEIKALEVTGLEVMDLNDAMALPETAASSGSSSSSGCSTCGSSSCCGSC

75_WP_087458049.1 Tumebacillus avium 3771376 3771376 TS=36
MEIKALEVTGLEVMDLNDAMALPETAASSGSSSSSGCSTCGSSSCCGSC

76_WP_016697827.1 Actinoalloteichus spitiensis RMV-1378 45 45 TS=35
MSADLPALNIESLEISEFLDDNRLEDGEVVAKVMSASCTTCECCCSCSS

77_WP_173027111.1 Arthrobacter sp. NEB 688 749885 749885 TS=35
MSKDTQDITFDSFDVQDVEVLEAADGTALPEMGASSGSIGTSSSSSSTCSAC

78_WP_094181683.1 Cellulomonas sp. PSBB021 3338019 3338019 TS=35
MTTNDIANLSGEITALESENFEISDYADAAEAILASCSSSSSTSTCSSTTSTTSCTA

79_WP_141322329.1 Cellulomonas uda 30387 30387 TS=35
MTTNDIANLSGEITALESENFEISDYADAAEAILASCSSSSSTSTCSSTTSTTSCTA

80_WP_043588221.1 Clavibacter cf. michiganensis LMG 26808 41659 41659 TS=35
MEKTMSMKISLPDLDFDSVEIMSVREAVAVPETGATSGSSSSNSSSCCGSSSCCSSS

81_WP_147362287.1 Clavibacter michiganensis 24430 24430 TS=35
MEKTMSMKISLPDLDFDSVEIMSVREAVAVPETGATSGSSSSNSSSCCGSSSCCSSS

82_WP_051750259.1 Phycicoccus jejuensis 399039 399039 TS=35
MSKDTQDITFDSFDVQDVEVLEAADGTALPEMGASSGSIGTSSSSSSTCSAC

83_GAT69286.1 Planomonospora sphaerica 29424 29424 TS=35
MSTATSQSIGVESLTGLDVDMLEISDYIDETLLDTADLTVTMVSSASCTTCICTCSCSS

84_WP_131102100.1 Streptomonospora sp. M2 1433823 1433823 TS=35
MDATNELDLNALEISDLIEDSSYNDETLSQVMAASCTTTGCTTSSSSSSSS

85_WP_086561064.1 Streptomyces africanus 28640 28640 TS=35
MAIETEREVFQVDAAEQLDLNALEISDLIEDSSYSDETLSQVMAASCTTTGCTTSSSSSSS
S

86_WP_167745813.1 Streptomyces azureus 263530 263530 TS=35
MSNASIGQDIGVEGLTGLDVDALEISDYMDEALLDSEDLTVTLVSSVSCTTCICSCSCSS

87_WP_101254739.1 Streptomyces barkulensis 228575 228575 TS=35
MSSVAETEIFDSFDIEELEVLDVSDSVALPEMGASSGVTMTETDAIEGTFCCSSSSSSSCC
CC

88_WP_099198214.1 Streptomyces cinnamoneus 6082075 6082075 TS=35
MNNDILEMDELDFQLEELSVLDIADARALPELGASNGSIGCSSSTCSSTCCC

89_WP_099198214.1 Streptomyces cinnamoneus 6082271 6082271 TS=35
MDELVFQLEELSVLDVSDAVALPEMGASSGWGCSSSTCSSCCC

90_WP_121802851.1 Streptomyces griseocarneus 11806 11806 TS=35
MTQKAELATLAQELLELESETFEISDYSDASEVVLAGSTSSSSTSTCSSTTSTTSCSA

91_WP_120721641.1 Streptomyces hundungensis 3238693 3238693 TS=35
MNNIEELDLNALEISDLIEDSGNNDESLSQIMAASCTTSGCACSSSSSSSS

92_WP_093850009.1 Streptomyces pini 52058 52058 TS=35
MSSVAETEIFDSFDIEELEVLDVSDSVALPEMGASSGVTMTETDAIEGTFCCSSSSSSSCC
CC

93_WP_063837860.1 Streptomyces pluripotens 7488978 7488978 TS=35
MSTLANETPAGDLFASFDIAELEVLEVSDSVGLPEMGASSSSMSVAAESTSASLCCSSSSS
TCCCC

94_WP_073741099.1 Streptomyces sp. CB02115 809912 809912 TS=35
MDVDAMFDSFDIEELETMELADGVALPEMAASTIPVACCSSSSSSCC

95_WP_073741099.1 Streptomyces sp. CB02115 810474 810474 TS=35
MSKILESVDADAMFDSFDIEELGTMELADGVALPEMGASKGDFTCCSSSSSSCC

96_WP_073741099.1 Streptomyces sp. CB02115 810851 810851 TS=35
MNADALFDSFDIEELDTMEVADGVALPEMGASTGNTMCCSSSSSSCC

97_WP_073931449.1 Streptomyces sp. CB02400 57383 57383 TS=35
MTAKDIKNDADNAFEGFDADELETLEVADGVALPEMGASSGSIGTSSSSSSTCSAC

98_WP_073931449.1 Streptomyces sp. CB02400 57589 57589 TS=35
MTAKDIKNDADNAFEGFDADELETLEVADGVALPEMGASSGSIGTSSSSSSTCSAC

99_WP_047180975.1 Streptomyces sp. MNU77 2815181 2815181 TS=35
MDVDAMFDSFDIEELDTLEIADGVALPEMGASTIPVACCSSSSSSCC

100_WP_047180975.1 Streptomyces sp. MNU77 2815736 2815736 TS=35
MDVDAMFDSFDIEELDTMELADGVALPEMGASKGDFTCCSSSSSSCC

101_WP_047180975.1 Streptomyces sp. MNU77 2816107 2816107 TS=35
MNADALFDSFDIEELDTMEVADGVALPEMGASTGNTMCCSSSSSSCC

102_WP_030264332.1 Streptomyces sp. NRRL B-24484 225882 225882 TS=35
LLLLPVLLLHLSGGRSSRRVPDVPSSVTAGYGLTELSASGGASSCCCCCPCCSCT

103_WP_097963940.1 Streptomyces sp. or20 79468 79468 TS=35
MDADAMFDSFDIEELGTMELADGVALPEMGASKGDFTCCSSSSSSCC

104_WP_097963940.1 Streptomyces sp. or20 79842 79842 TS=35
MNADALFDSFDIEELDTMEVADGVALPEMGASTGNTMCCSSSSSSCC

105_WP_050360717.1 Streptomyces sp. or3 86854 86854 TS=35
MNADALFDSFDIEELDTMEVADGVALPEMGASTGNTMCCSSSSSSCC

106_WP_050360717.1 Streptomyces sp. or3 87231 87231 TS=35
MSKILETVDADAMFDSFDIEELGTMELADGVALPEMGASKGDFTCCSSSSSSCC

107_EHM28179.1 Streptomyces sp. W007 206 206 TS=35
MSMAIEAMNADALFDSFDIEELDTMEVADGVALPEMGASTGNTMCCSSSSSSCC

108_WP_052889351.1 Thermogemmatispora carboxidivorans 2855515 2855515 TS=35
MHGELAQLELDGLTLDDLEIEDLKLSSEGGLELLTTGHGMVEIGASSGTAACCSCCLVC
CCCW

109_WP_052889351.1 Thermogemmatispora carboxidivorans 2855805 2855805 TS=35
MEDLKQELSHLNLEELQLDDLSIQPLEADHSGLEALTLGHGMIEIGASILPTACCSCCIPC
CCCC

110_WP_165898146.1 Tumebacillus sp. BK434 341556 341556 TS=35
MEIKALEVTGLEVMDLNDAMALPETAASSGSSSSSGCSTCGSSSCCGSC

111_WP_165898146.1 Tumebacillus sp. BK434 341752 341752 TS=35
MEIKALEVTGLEVMDLNDAMALPETAASSGSSSSSGCSTCGSSSCCGSC

112_WP_165898146.1 Tumebacillus sp. BK434 341948 341948 TS=35
MEIKALEVTGLEVMDLNDAMALPETAASSGSSSSSGCSTCGSSSCCGSC

113_WP_015802628.1 Actinosynnema pretiosum subsp. pretiosum 4541473 4541473 TS=34
MDLTFDESELDLGDLAVTAMRDAVALPETGASTAACSCSSTSCCCCQQPPTPELPQV

114_WP_066450536.1 Bacillus gottheilii 4313252 4313252 TS=34
MKNELNLDLDFEVIELDDVTALPETAASSGSSGDNVYSTCGSSSCCSSCCT

115_WP_066450536.1 Bacillus gottheilii 4313642 4313642 TS=34
MKNELNLDLDFEVIELDDVTALPETAASSGSSGDNVYSTCGSSSCCSSCCT

116_WP_151575809.1 Bacillus mesophilum 51812 51812 TS=34
MNNDMNLDLDFEVIELDDVTALPETAASSGSSGDNVYSTCGSSSCCSSCCT

117_WP_151575809.1 Bacillus mesophilum 52002 52002 TS=34
MKVDVNLDLDFEVIELDDVTALPETAASSGSSGDNVYSTCGSSSCCSSCCT

118_WP_151575809.1 Bacillus mesophilum 52201 52201 TS=34
MKNEMNLDLDFEVIELDDVTALPETAASSGSSGDNVYSTCGSSSCCSSCCT

119_WP_098212038.1 Bacillus sp. AFS075034 7326 7326 TS=34
MNNNNDAKADLEKLNSGMGMTEVGASLLCSCSCPCSCCSSSSI

120_WP_098212038.1 Bacillus sp. AFS075034 7547 7547 TS=34
MNNNNDAKADLEKLNSGMGMTEVGASLLCSCSCPCSCCSSSSI

121_WP_143406617.1 Exiguobacterium sp. AT1b 7220 7220 TS=34
MENNLDVELELDFEVIELDDVTALPETAASSGSSGDNVYSTCGSSSCCSSCCT

122_WP_143406617.1 Exiguobacterium sp. AT1b 7413 7413 TS=34
MENNLDIELELDFEVIELDDVTALPETAASSGSSGDNVYSTCGSSSCCSSCCT

123_WP_143406617.1 Exiguobacterium sp. AT1b 7608 7608 TS=34
MENNLDIELELDFEVIELDDVTALPETAASSGSSGDNVYSTCGSSSCCSSCCT

124_WP_106538471.1 Haloactinopolyspora alba 107149 107149 TS=34
MAHTTLGMELLELEAVTFEVDEITDPSADMAASSSSCSCSSCSSCSTSSCCSCSTSTSSCG

125_WP_027344871.1 Hamadaea tsunoensis DSM 44101 46280 46280 TS=34
MSNEDFNVEELELDEIAVTSIRDSAALPETGASSGSSSSSSSCCGSCSCCGSCADTNEPA
VQ

126_WP_089022234.1 Micromonospora coriariae 6768840 6768840 TS=34
MPDLTEELRALETETFEIEDVDAIDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGG
G

127_WP_120331704.1 Micromonospora globbae 32725 32725 TS=34
MPITLNDELRELETETFEIEEVEDGGEALAASSSSCSCSSCCSCSTSSCCSTSTSTSSCG

128_WP_120331704.1 Micromonospora globbae 33387 33387 TS=34
MPELTEELRALESETFEIEDVDAIDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

129_WP_120688751.1 Micromonospora musae 112953 112953 TS=34
MSIDETTAAQDDKTFDSFEIADLEVLEVAQGVALPELGASSGSIGTSSSSSTCSSSTC

130_WP_120675210.1 Micromonospora musae 208089 208089 TS=34
MSIDETTAAQDDKTFDSFEIADLEVLEVAQGVALPELGASSGSIGTSSSSSTCSSSTC

131_WP_067140220.1 Microtetraspora malaysiensis 49059 49059 TS=34
MNALSFDVEELALDELSVTTVRDGVALPDTGASACCFSWMTSSSCCCCTEIE

132_WP_067140220.1 Microtetraspora malaysiensis 51434 51434 TS=34
MSMTTTRFDLADLDLSGLQVTVMRDAVAVPEGGATSGSSSSDSSSTCGSSSCCSSVAQQ

133_WP_167477798.1 Nocardia arthritidis 8632159 8632159 TS=34
MENNLPALDINALEISEFLDDSRMDDRDVVAKVMSASCTTCECSCSCSS

134_WP_067608788.1 Nocardiopsis listeri NBRC 13360 5300 5300 TS=34
MRKSTGLSDLQGELELLESETFEVLDLVEAREEMLGSTTSTTSCCSSCSSSSCSSTSCAGG

135_WP_087097343.1 Nocardiopsis sp. JB363 9265 9265 TS=34
MSRTTALSNLQGELELLESETFEVLDLVEAREEMLGSTTSTTSCCSSCSSSSCSSTSCAGG

136_WP_111153483.1 Paenibacillus dendritiformis 191 191 TS=34
MEELKLENQEIELSLDDLELQDVKMEVLMTNSKGLPEMGATLGGITCCSCCCCST

137_WP_111149723.1 Paenibacillus sambharensis 123808 123808 TS=34
MEKNMTALVDEELSFEIIELEDITALPETAASSGSSSSDTCSTCGSCSCSSCCT

138_WP_107280457.1 Streptomyces albus subsp. albus 227304 227304 TS=34
MHMSTNVMVEDFSAFDVEELEVLDVADATALPEMGASSATWACCSSSSSSSCTCSCC

139_KUJ37316.1 Streptomyces albus subsp. albus 45481 45481 TS=34
MSVELKDELNALESATFEIEEMTDEAVELAWSSSSCSCSSCCSCSTSSCCSCSSSTS

140_WP_157841653.1 Streptomyces atroolivaceus 434477 434477 TS=34
MEQQIELDVLEISDLIAGAGENDDLAQVMAASCTTTSVSTSSSSSSS

141_AKJ11416.1 Streptomyces incarnatus 3575800 3575800 TS=34
METETEREVFQVDAAEQLDLNALEISDLIEDSSYSDETLSQVMAASCTTTGCTTSSSSSSS

142_WP_031037067.1 Streptomyces olivaceus 349083 349083 TS=34
MSTVAIDSADIEQVFDSFDVTELEVLEVSQGVALPEMGASGGTVGTSSSSSTCSSSTC

143_WP_135337719.1 Streptomyces palmae 35357 35357 TS=34
MSVELKNELNALESATFEIEEMTDEAVELAWSSSSCSCSSCCSCSTSSCCSCSSSTS

144_WP_096213443.1 Streptomyces sp. 2323.1 4190685 4190685 TS=34
MSTVAAVGQDVGVEGLTGLDAGALEISETRTRRCSTARAAVTLISSVSCTTCICSGRCSS

145_WP_097223039.1 Streptomyces sp. Ag82_G6-1 3689019 3689019 TS=34
MTAKDIKNQDADKAFEGFDADELETLEVTDGVALPEMGASSGSIGTSSSSSSTCSAC

146_WP_097223039.1 Streptomyces sp. Ag82_G6-1 3689258 3689258 TS=34
MTAKDIKNQDADKAFEGFDADELETLEVTDGVALPEMGASSGSIGTSSSSSSTCSAC

147_WP_087767808.1 Streptomyces sp. CS057 2994943 2994943 TS=34
MDVDAMFDSFDIEELGTMELADGVALPEMGASKGDFTCCSSSSSSCC

148_WP_087767808.1 Streptomyces sp. CS057 2995314 2995314 TS=34
MNADALFDSFDIEELDTMEVADGVALPEMGASTGNTMCCSSSSSSCC

149_WP_127355247.1 Streptomyces sp. LAM7114 74848 74848 TS=34
MNTVSVNSADIEQVFDSFDVTELEVLEVSQGVALPEMGASGGNGGTSSSSSTCSSSTC

150_WP_030812058.1 Streptomyces sp. NRRL F-2799 40412 40412 TS=34
MSVELRDELNALESATFEIEEMTDESVELAWSSSSCSCSSCCSCSTSSCCSCSSSTS

151_WP_018560187.1 Streptomyces sp. SID8377 12175 12175 TS=34
MSVELKDELSALESATFEIEEMTDESVELAWSSSSCSCSSCCSCSTSSCCSCSSSTS

152_WP_019547680.1 Streptomyces sulphureus DSM 40104 63021 63021 TS=34
MSNVTVDSIDIEQVFDSFDVTELEVLEVSQGVAQPEMGASGGTVGTSSSSSTCSSSTC

153_WP_110669044.1 Streptomyces tateyamensis 30401 30401 TS=34
MSAPTEIQNLGVVGLTGLDTDTLEISDYLDESLLDEHDLTVTMVASASCTTCICTCSCSS

154_QJC58228.1 Streptomyces viridochromogenes 31046 31046 TS=34
MFQVESAEQLDLNALEISDLIEDSSYSDETLSQVMAASCTTTGCTTSSSSSSSS

155_WP_069802514.1 Thermogemmatispora onikobensis 151400 151400 TS=34
MEDLRQELSHLNLDELQLDDLSIQPLEADHSGLEALTLGHGMIEIGASILPTACCSCCIPC
CCCC

156_WP_151570259.1 Actinomadura rudentiformis 75633 75633 TS=33
MRQEKQDHTLSVLKDELGLLEAETFEIRDHNDVVEELAGTTSCTSTTSSCTSSSGTSCCD

157_WP_151570259.1 Actinomadura rudentiformis 76381 76381 TS=33
MSADLVDLRNDLELLEAETFEILDYEEAPEVLADCCSTSSCSTSSSTTSCTSTASCA

158_WP_014690724.1 Actinoplanes sp. SE50 4175479 4175479 TS=33
MENIDEELDLNALEISDLIEDSGTEDETLSQIMAASCTTSGCACSSSSSSSS

159_WP_098207202.1 Bacillus cereus 13832 13832 TS=33
MEKMNKENVNFFDDLEITSLDITEVTDAISIPETGASSGSNGTYLCGSSSCCSSCCS

160_WP_098207202.1 Bacillus cereus 14083 14083 TS=33
MEKMNKENVNLFDDLEITGLDITEVTDAISIPETGASSGSNGTYLCGSSSCCSSCCS

161_WP_098481582.1 Bacillus cereus 122446 122446 TS=33
MEKMNKENVNLFDDLEITGLDITEVTDAISIPETGASSGSNGTYLCGSSSCCSSCCS

162_WP_098481582.1 Bacillus cereus 122697 122697 TS=33
MEKMNKENVNFFDDLEITGLDITEVTDAISIPETGASSGSNGTYLCGSSSCCSSCCS

163_WP_066450536.1 Bacillus gottheilii 4313452 4313452 TS=33
MCTQHAAPALAALLAVLDKRTRTIKEVKKVKVDVNLDLDFEVIELDDVTALPETAASS
GSSGDNVYSTCGSSSCCSSCCT

164_WP_002169253.1 Bacillus mycoides 4862089 4862089 TS=33
MEKMNKENVNFFDDLEITGLDITEVTDAISIPETGASSGSNGTYLCGSSSCCSSCCS

165_WP_002169253.1 Bacillus mycoides 4862335 4862335 TS=33
MEKMNKENVNFFDDLEITGLDITEVTDAISIPETGASSGSNGTYLCGSSSCCSSCCS

166_WP_044053021.1 Bacillus velezensis 13527 13527 TS=33
MKKEKNDLLNNLEITGLDVTEITDSISIPETGATSGSGGHSTCGSSSCCSSCCS

167_WP_166378646.1 Catellatospora methionotrophica 633632 633632
TS=33
MSNEDFNVDELELDDLSVTSIRDSAALPETGASSGSSSSSASSCCGSSSCCSSCGDNPEVA
A

168_WP_128211777.1 Clostridium sp. CT4 1171657 1171657 TS=33
MENKNNIFSEFDIEVSELEFIDVMPSTAASRAGIFPLSTSTCGSSSSSTCSSSCS

169_WP_128211777.1 Clostridium sp. CT4 1171900 1171900 TS=33
MEEKKAQEINNIIDDFDIEVIELDNVKAVPSTAASSGNGWSSTSTCGSSSSSTCSACCS

170_WP_128211777.1 Clostridium sp. CT4 1172140 1172140 TS=33
MEEKKVQEINNILDDFDLEVIEMDNVQAVPSTAASGGNGSSTSTCGSSSSSTCSSSCA

171_WP_128211777.1 Clostridium sp. CT4 1172370 1172370 TS=33
MNKENVKELNNILDDFDIEVIEMDNVQAVPSTAASGGNGSSTSTCGSSSSSTCSSSCA

172_WP_128211777.1 Clostridium sp. CT4 1172600 1172600 TS=33
MNKENVKELNNILDDFDIEVIEMDNVQAVPSTAASGGNGSSTSTCGSSSSSTCSSSCA

173_WP_119949388.1 Frankiales bacterium YIM 75000 984240 984240 TS=33
MAKNTQQPTTDDVFGAFDVDGFETLDVAQGTALPEMGASSGSIGTSSSSSSTCSAC

174_WP_173132718.1 Kibdelosporangium sp. 4NS15 157026 157026 TS=33
MSTSRRLDFDLENLPMDVFDIVDSGLTVESLTAGHGMAENGASCSAPGGGCSGCSGTT
WDCCCSCGTD

175_WP_109243206.1 Kocuria rosea 3432150 3432150 TS=33
MDRKPTDLVDLPMDVFELEDQGMDITSLTAGHGMTEVGASTNCFCYPCCSCSAPSSSA

176_WP_172604317.1 Kocuria rosea 727572 727572 TS=33
MDRKPTDLVDLPMDVFELEDQGMDITSLTAGHGMTEVGASTNCFCYPCCSCSAPSSSA

177_WP_017833576.1 Kocuria sp. UCD-OTCP 60951 60951 TS=33
MDRKPTDLVDLPMDVFELEDQGMDITSLTAGHGMTEVGASTNCFCYPCCSCSAPSSSA

178_WP_053593010.1 Lysinibacillus sp. FJAT-14222 145802 145802 TS=33
MKKHAEYLFEALDEEALEEDISIEVMDISDSLGLPDMAASKGTTVCCSCCVVCCCCGTE

179_WP_110565432.1 Micromonospora arborensis 105583 105583 TS=33
MTGLDIDTLEISDYLDESMLDTHDLTETMIASASCTTCICTCSCSS

180_SCF23618.1 Micromonospora carbonacea 466536 466536 TS=33
MSISVDPQTELGVEELTGLDIDTLEISDYLDESMLDTHDLTETMIASASCTTCICTCSCSS

181_SCL44848.1 Micromonospora citrea 462425 462425 TS=33
MKTSDELDFELDDLPMDVFDLAESGLTIESLTAGHGMPEHGASLPFCSCSATCSCCPSSS
S

182_WP_120331704.1 Micromonospora globbae 33125 33125 TS=33
MPDSVRSELRDLQAETFEVEDIADLTADLMDICSSSTSTSSCSSCSTSSCCSCTSSSCSSTS

183_WP_091639110.1 Micromonospora pallida 640313 640313 TS=33
MPELSEELRALETETFEIEDVDAVDAMVMAWSSSSCSCSSCCSCSTSSCCSTSTSTGCGG
G

184_WP_123600439.1 Micromonospora sp. Llam0 379505 379505 TS=33
MSDEIIDVEDLELDELSVSSIRDSAALPETGASSGSSSCSSSSCCGASSCCASCAEVEQSAL

185_WP_089009171.1 Micromonospora viridifaciens 6444094 6444094 TS=33
MENELSTLDIDDLEISEFLDESRLEDSEVVAKVMSASCTTCECSCSCSS

186_ADR01090.1 Nocardia sp. ATCC 202099 38581 38581 TS=33
MSADLSALNIDSLEISEFLDDSRLEDSEVVAKVMSASCTTCECSCSCSS

187_WP_111149723.1 Paenibacillus sambharensis 124018 124018 TS=33
MENTKTVSVLDDLTFEIIELDDVTALPETAASSGSSSSDTCSTCGSSSCSSCSSCT

188_WP_101628591.1 Schaalia turicensis 9369 9369 TS=33
MSAMNEIANLEVDTFEILEMVELDEAAMKAWCSCSTSSCCGCSSSSCGSTTSTSCSCGST
SSCA

189_WP_131102100.1 Streptomonospora sp. M2 1433577 1433577 TS=33
MDATNELDLNALEISDLIEDLNHSEETLSQVMGASCVGCACTCSSTSSSS

190_WP_078843712.1 Streptomyces albus subsp. albus 3810 3810 TS=33
MDAKNELATLANDILELESETFEISDYSDAVEVVLAGSTSSSSTSTCSSTTSTTSCSA

191_WP_167745813.1 Streptomyces azureus 263197 263197 TS=33
MSNASIGQEIGVEGLTGLDVDALEISDYVDETLLDGEDLTVTMIASASCTTCICTCSCSS

192_AGN11669.1 Streptomyces bernensis 10654 10654 TS=33
MEQQIELDVLEISDLIAGAGENDDLAQVMAASCTTTSVSTSSSSSSS

193_WP_037804196.1 Streptomyces caelestis 60329 60329 TS=33
GFDADELETLEVADGVALPEMGASSGSIGTSSSSSSTCSAC

194_WP_125043512.1 Streptomyces chrestomyceticus JCM 4735 766532 766532 TS=33
MDANLVLDLDDLSVDQLDILPTAPGSTLESINVGHAMVEIGASNCTSKGSPASCCSCCCC

195_WP_051820544.1 Streptomyces flavochromogenes 106147 106147 TS=33
MSIDDIKNGDDAAFDNFDVAELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

196_WP_051820544.1 Streptomyces flavochromogenes 106354 106354 TS=33
MSIDDIKQGQDADKAFENFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

197_WP_120752968.1 Streptomyces klenkii 318540 318540 TS=33
MAQQAELATLAQEILELESETFEISDYSDASEVVLAGSTSSSSTSTCSSTTSTTSCSA

198_BAU84768.1 Streptomyces laurentii 4051564 4051564 TS=33
MSNAALEIGVEGLTGLDVDTLEISDYMDETLLDGEDLTVTMIASASCTTCICTCSCSS

199_ACN52295.1 Streptomyces laurentii 8076 8076 TS=33
MSNAALEIGVEGLTGLDVDTLEISDYMDETLLDGEDLTVTMIASASCTTCICTCSCSS

200_WP_069569928.1 Streptomyces lydicus 4027407 4027407 TS=33
MQTELVLDLADLSADELEILPTSPGASLETVNVGHAMVEIGASNCTSTGTPASCCSCCCC

201_WP_129298148.1 Streptomyces lydicus 75510 75510 TS=33
MQDKLVLDLADLSADELEILPTSPGESLEAVNVGHAMVEIGASNCTSRGTPASCCSCCC
C

202_WP_067348540.1 Streptomyces noursei ATCC 11455 4424069 4424069
TS=33
MQELDLNALEISDLIDDLGHSEEELSQVMAASCVGCAGSCSSTSSSS

203_WP_070390316.1 Streptomyces olivaceus 3660694 3660694 TS=33
MSTVAIDSADIEQVFDSFDVTELEVLEVSQGVALPEMGASGGTVGTSSSSSTCSSSTC

204_WP_053557716.1 Streptomyces pristinaespiralis 2256680 2256680
TS=33
MSIDDIKNNGDDAAKAFESFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

205_WP_053557716.1 Streptomyces pristinaespiralis 2256895 2256895
TS=33
MSINDIKNDDADKAFESFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

206_WP_125051547.1 Streptomyces rimosus subsp. paromomycinus 732273
732273 TS=33
MDANLVLDLDDLSVDQLDILPTAPGSTLESINVGHAMVEIGASNCTSKGSPASCCSCCCC

207_QBP39279.1 Streptomyces sp. 13023 13023 TS=33
MEQQIELDVLEISDLIAGAGENDDLAQVMAASCTTSSVSTSSSSSSS

208_WP_116513357.1 Streptomyces sp. AcE210 2918363 2918363 TS=33
MSINDIKNGDDAEKAFDSFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

209_WP_116513357.1 Streptomyces sp. AcE210 2918608 2918608 TS=33
MSINDIKNGDDAEKAFDSFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

210_OKI60309.1 Streptomyces sp. CB00072 835509 835509 TS=33
LRAANRLEIVHRRSRGRRSEPGGRPKMSKIMETMDVDAMFDSFDIEELGTMELADGVA
LPEMGASKGDFTCCSSSSSSCC

211_OKI60309.1 Streptomyces sp. CB00072 835883 835883 TS=33
MSMAIEAMNADALFDSFDIEELDTMEVADGVALPEMGASTGNTMCCSSSSSSCC

212_WP_100591758.1 Streptomyces sp. CB01635 193600 193600 TS=33
MSINDIKNGDDAEKAFDSFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

213_WP_100591758.1 Streptomyces sp. CB01635 193845 193845 TS=33
MSINDIKNGDDAEKAFDSFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

214_WP_073913239.1 Streptomyces sp. CB02009 398669 398669 TS=33
MSIDDIKQGQDADKAFENFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

215_WP_073913239.1 Streptomyces sp. CB02009 398876 398876 TS=33
MSIDDIKNGDDAAFDNFDVAELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

216_WP_071279171.1 Streptomyces sp. CC53 37175 37175 TS=33
MAKSELATLADEILELESETFEISDYSDAVEVVLAGSCSSSSTSTCSSTTSTTSCSA

217_WP_058043695.1 Streptomyces sp. MBT76 2287677 2287677 TS=33
MAQQAELATLAQEILELESETFEISDYSDASEVVLAGSTSSSSTSTCSSTTSTTSCSA

218_WP_084769183.1 Streptomyces sp. MOE7 558467 558467 TS=33
MNDDLVLDLADLSVDDLDILPASPGASLETVNVGHAMVEIGASNCTSTGTPASCCSCCC
C

219_WP_065486866.1 Streptomyces sp. PTY087I2 139070 139070 TS=33
MAPKTELATLADEILELESETFEISDYSDVAEVVLAGSTSSSSTSTCSATTSTTSCSA

220_WP_153452235.1 Streptomyces sp. RB5 35851 35851 TS=33
MMDQIKDVDVFEAFDIDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

221_WP_018515651.1 Streptomyces sp. SID5594 132742 132742 TS=33
MAPKTELATLADEILELESETFEISDYSDAVEVVLAGSTSSSSTSTCSSTTSTTSCSA

222_WP_149184555.1 Streptomyces sp. TRM49041 37595 37595 TS=33
MAKTELATLADEILELESETFEISDYSDAVEVVLAGSCSSSSTSTCSSTTSTTSCSA

223_WP_159694946.1 Streptomyces sp. Tu 2975 2049893 2049893 TS=33
MSIDDIKNNGDDAAKAFESFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

224_WP_159694946.1 Streptomyces sp. Tu 2975 2050108 2050108 TS=33
MSINDIKNDDADKAFESFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

225_WP_051674763.1 Streptomyces sp. URHA0041 18010 18010 TS=33
MAVELKDELSALESATFEIEEMTDESIELAFSSSSCSCSSCCSCSTSSCCSCSSSTS

226_WP_162929078.1 Streptomyces sp. YIM 130001 199867 199867 TS=33
MEQQIELDVIEISDLIEGAGENDDVAQVMASSCTTSSVSTSSSSSS

227_APE26458.1 Streptomyces venezuelae 5537821 5537821 TS=33
MSIDDIKNGDDAAFDNFDVAELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

228_WP_150508245.1 Streptomyces venezuelae 2510362 2510362 TS=33
MSIDDIKNGDDAAFDNFDVAELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

229_WP_150508245.1 Streptomyces venezuelae 2510569 2510569 TS=33
MSDDIKQGQDADKAFENFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

230_PWW57157.1 Streptomyces venezuelae 109058 109058 TS=33
MSIDDIKNGDDAAFDNFDVAELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

231_PWW57157.1 Streptomyces venezuelae 109265 109265 TS=33
MSDDIKQGQDADKAFENFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

232_CCA58419.1 Streptomyces venezuelae ATCC 10712 5538722 5538722 TS=33
MSDDIKQGQDADKAFENFDVDELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

233_CCA58419.1 Streptomyces venezuelae ATCC 10712 5538929 5538929 TS=33
MSIDDIKNGDDAAFDNFDVAELETLEVAQGVALPEMGASSGSIGTSSSSSSTCSAC

234_WP_093656411.1 Streptomyces wuyuanensis 176289 176289 TS=33
MEQQIELDVLEISDLIAGAGENDDLAQVMAASCTTSSVSTSSSSSSS

235_WP_043471952.1 Streptomyces xinghaiensis 286787 286787 TS=33
MDAKTELATLANDILELESETFEISDYSDAVEVVLAGSTSSSSTSTCSSTTSTTSCSA

236_WP_084012656.1 Thermobifida halotolerans 79874 79874 TS=33
MSTQSLGMELMELEAVTFEVDEIADLSDQLAACSSSCSCSSCCSCSTSSCCSTSTSTSSCG

237_WP_084012656.1 Thermobifida halotolerans 80199 80199 TS=33
MSTQSLGMELMELEAVTFEVDEIADLSDQLAACSSSCSCSSCCSCSTSSCCSTSTSTSSCG

238_WP_084012656.1 Thermobifida halotolerans 80524 80524 TS=33
MSSPSLGLELMELEAVTFEVEDIADLGQEAAAWCSSSCSCSSCCSCSTSSCCSTSTSTSSCG

239_WP_151728560.1 Thermogemmatispora aurantia 521626 521626 TS=33
MHGELAQLELEGLSLDSLEIEDLKLSNEGGLELLTTGHGMVEIGASSGQTACCSCCIVCCCCW

240_WP_151728560.1 Thermogemmatispora aurantia 521971 521971 TS=33
MHGELAQLELEGLSLDNLEIEDLKLSSEGGLELLTTGHGMVEIGASSGQTACCSCCIVCCCCPCLG

241_WP_052889351.1 Thermogemmatispora carboxidivorans 2855170 2855170 TS=33
MHGELAQLELDGLTLDDLEIEDLKLSPEGGLELLTTGHGMVEIGASSGQGACCSCCLLCCCCCPCLG

242_RAQ94249.1 Thermogemmatispora tikiterensis 406067 406067 TS=33
MNELHGELAQLELEGLSLDNLEIEDLKLSSEGGLELLTTGHGMVEIGASSGQTACCSCCIVCCCCPCLG

243_RAQ94249.1 Thermogemmatispora tikiterensis 406412 406412 TS=33
MQDLHGELAQLELEGLSLDSLEIEDLKLSNEGGLELLTTGHGMVEIGASSGQTACCSCCIVCCCCW

244_WP_083378345.1 Varibaculum massiliense 891306 891306 TS=33
MSALNEIANLEVDTFEILEMVELDEAAMKAWCSCSTSSCCGCSSSSCGSTTSTSCSCGSTSSCA

245_WP_129626314.1 Yimella sp. RIT 621 45714 45714 TS=33
MNKLNLTAELAELEAETFEIADYADDAEMVLAGCTSSSSTSTCSSTTSTTSCSA

246_WP_129626314.1 Yimella sp. RIT 621 45926 45926 TS=33
MNKLNLTAELAELEAETFEIADYADDAEMVLAGCTSSSSTSTCSSTTSTTSCTA

247_WP_129626314.1 Yimella sp. RIT 621 46138 46138 TS=33
MNKLNLTAELAELEAETFEIADYADDAEMVLAGCTSSSSTSTCSSTTSTTSCTA

248_WP_151570259.1 Actinomadura rudentiformis 76613 76613 TS=32
MAMLEAETFEIHDHVEATEDLLGWTSCDCSTSTGSSTSSTSTSSSTTSCTSTASCA

249_OLO56992.1 Actinomyces oris 389 389 TS=32
VKYVNNVIDFAAIEISDLVEDAVEGGELPSQVMAASTTTCGCACSSCSSTCS

250_WP_085945035.1 Actinosynnema mirum DSM 43827 4523946 4523946 TS=32
MDLTFDESELDLGDLAVTAMRDAVALPETGASTAACSCSSTSCCCCQQPPTPELPQV

251_ATE55167.1 Actinosynnema pretiosum 4512414 4512414 TS=32
MDLTFDESELDLGDLAVTAMRDAVALPETGASTAACSCSSTSCCCCQQPPMPELPQV

252_PET35443.1 Bacillus anthracis 17458 17458 TS=32
MKEQKELKNEEFELDVEFLDLDEVSAIPETTASSGTSSCSASSTCGSSSCCGSC

253_PET35443.1 Bacillus anthracis 17658 17658 TS=32
MKEQKELKELKNEEFELDVEFLDLDEVSAIPETTASSGSSGCSASSTCGSSSCCGSC

254_PET35443.1 Bacillus anthracis 17858 17858 TS=32
MKEQKELKELKNEEFELDVEFLDLDEVSAIPETTASSGTSSCSASSTCGSSSCCGSC

255_PET35443.1 Bacillus anthracis 18058 18058 TS=32
MKEQKELKELKNEEFELDVEFLDLDEVSAIPETTASSGTSSCSASSTCGSSSCCGSC

256_WP_066417934.1 Bacillus cohnii 4130843 4130843 TS=32
MDEQKKEEIKMEKNLNLELDFEVIELDDVTALPETAASSGSSGDNVYSTCGSSSCCSSCC
T

257_WP_066417934.1 Bacillus cohnii 4131041 4131041 TS=32
MEKNLNLELDFEVIELDDVTALPETAASSGSSGDNVYSTCGSSSCCSSCCT

258_WP_062352644.1 Bacillus kwashiorkori 33888 33888 TS=32
MENALNLDLDFEVIELDDVTALPETAASSGSSGDKVYSTCGSSSCCSSCCT

259_WP_062352644.1 Bacillus kwashiorkori 34098 34098 TS=32
MENTLNLDLDFEVIELDDVTALPETAASSGSSGDKVYSTCGSSSCCSSCCT

260_WP_113938616.1 Bacillus mycoides 19804 19804 TS=32
MKNNFEPKKEKLDLEEMLAAANLEIIELDQAEALPSTAASSGSSSGNTCSTCGSSSSSTI

261_WP_113938616.1 Bacillus mycoides 20277 20277 TS=32
MKNNFEPKKEKLDLEEMLAAANLEIIELDQAEALPSTAASSGSSSGNTCSTCGSSSSSTI

262_PEU19231.1 Bacillus sp. AFS014408 32936 32936 TS=32
MLKNDVLDLSGIEVENLEFSFMNNNNDAKADLEKLNSGMGMTEVGASLLCSCSCPCSC
CSSSSI

263_PEU19231.1 Bacillus sp. AFS014408 33157 33157 TS=32
MLKNDVLDLSGIEVENLEFSFMNNNNDAKADLEKLNSGMGMTEVGASLLCSCSCPCSC
CSSSSI

264_WP_098200741.1 Bacillus wiedmannii 159 159 TS=32
MLAAANLEIIELDQAEALPSTAASSGSSSGNTCSTCGSSSSSTI

265_WP_071803305.1 Couchioplanes caeruleus 3433567 3433567 TS=32
MEDLKLDLEDLEVESLAFGGETDADLQSLGMGHGMTEVGASAGCCCCTCSCCCCPCG

266_WP_002518053.1 Cutibacterium acnes 1500 1500 TS=32
MENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

267_WP_115884642.1 Cutibacterium acnes 574732 574732 TS=32
MENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

268_AEH29193.1 Cutibacterium acnes 6609 945694 945694 TS=32
MENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

269_EFT77539.1 Cutibacterium acnes HL030PA1 583847 583847 TS=32
MENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

270_WP_010777475.1 Enterococcus faecalis 22340 22340 TS=32
MEKELKTVDVEVEELDELFGEVELIEIDETITLSETAASSGISSCNSCSCCASTSCCSSSSVS
FT

271_WP_069647476.1 Enterococcus ureasiticus 63661 63661 TS=32
MDTEIEELDELFGEVELIEIDETITLSETAASSGISSCNSCSCCASTSCCSSSSVSFT

272_WP_154775326.1 Erwinia sp. CPCC 100877 37208 37208 TS=32
MEQEKELQNLLEELEIETIEMDNIQAIPTAAASSGSGKGTLCGSSSSSTCSSCCS

273_WP_116950971.1 Jiangella sp. KE2-3 80982 80982 TS=32
VTALTTELLALETATFEVEEISDVAQDAAGWCSSSTSTSSCSSCSTSSCCGCSSCSCSSTS

274_WP_073810225.1 Kitasatospora sp. CB01950 385477 385477 TS=32
MADTSLASLAQEILNLESETFEISDYSDTSEVMLGSSTSCSSTSTCSSTTSTTSCTA

275_WP_090047425.1 Lechevalieria fradiae 364393 364393 TS=32
MKDLSFDLDDLELGELAVTSMRDSVALPESGASGAPSSCSCGSSSCSSCTQPQLPPA

276_WP_056374170.1 Microbacterium sp. Leaf161 667587 667587 TS=32
MSFHVDSSRASSLPTRDRDAVAQASTSCTTTSSSTGTMM

277_WP_089022234.1 Micromonospora coriariae 6769318 6769318 TS=32
MPITLNDELRELETETFEIEEVEDGGEALAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

278_WP_088993330.1 Micromonospora echinaurantiaca 1890504 1890504
TS=32
MPITLNDELRELETETFEIEEVEDGGEALAASSSSCSCSSCCSCSTSSCCSTSTSTSSCG

279_WP_088993330.1 Micromonospora echinaurantiaca 1891167 1891167 TS=32
MPELTEELRALETETFEIEDVDAVDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

280_WP_091639110.1 Micromonospora pallida 640564 640564 TS=32
MAPQHPASVALPESVRSELRNLQTETFEVEDIADLSADVMDICSSSTSTSSCSSCSTSSCCSCTSSSCSSTS

281_WP_167184099.1 Micromonospora sp. CNZ280 593216 593216 TS=32
MPELTEELRALETETFEIEDVDAVDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

282_WP_167184099.1 Micromonospora sp. CNZ280 593881 593881 TS=32
MPITLNDELRELETETFEIEEVEDGGEALAASSSSCSCSSCCSCSTSSCCSTSTSTSSCG

283_WP_123604194.1 Micromonospora sp. Llam0 5994064 5994064 TS=32
MENELSTLDIDDLEISEFLDESRLEDSEVVAKVMSASCTTCECSCSCSS

284_WP_041561842.1 Nocardiopsis alba ATCC BAA-2165 2806778 2806778 TS=32
MNKDIDLSAIEISDLISETEQSDDALSQVMAASCTTTGCACSSSSSST

285_WP_026125251.1 Nocardiopsis alba DSM 43377 57923 57923 TS=32
MNKDIDLSAIEISDLISETEQSDDALSQVMAASCTTTGCACSSSSSST

286_ADO67771.1 Nocardiopsis sp. TFS65-07 14551 14551 TS=32
MNKDIDLSAIEISDLISETEQSDDALSQVMAASCTTTGCACSSSSSST

287_EGL44723.1 Propionibacterium sp. 434-HC2 10496 10496 TS=32
MENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

288_KOX20424.1 Saccharothrix sp. NRRL B-16348 23605 23605 TS=32
MDLTFDDSELDLGDLAVTAMRDAVALPETGASAAACSCSSTSCCCCQQPQLPTLPV

289_WP_158061821.1 Serinicoccus sp. W204 2752617 2752617 TS=32
MQNTDTLDLGLDLADLDLDVLEVSQVRDAMALPETGASSGSSSCHSTSCCGTCSCSSCCT

290_WP_114033190.1 Sphaerisporangium album 139279 139279 TS=32
MPVRRAEEGSVSDLDRELIALETVTFEVEDIADLGRDAADWCSSSTSTSSCSSCSTSSCCSCTSSSCSSSS

291_WP_124728414.1 Staphylospora marina 2497352 2497352 TS=32
MEIRDLEITGLEVLDLSDAAALPETGASSGSTSCSASSTCGSSSCCASC

292_WP_124728414.1 Staphylospora marina 2497541 2497541 TS=32
MEIRDLEITGLEVLDLSDAAALPETGASSGSSSCSASSTCGSSSCCGSC

293_WP_034090612.1 Streptacidiphilus albus 50110 50110 TS=32
MSDSTANVGFDLQELDLGDLTVTSMRDTVALPENGASTQSCSCESSSSCTMPHPQVVTT
LQ

294_WP_129806648.1 Streptomyces albidoflavus 35368 35368 TS=32
MTPKTELATLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

295_WP_128826424.1 Streptomyces albidoflavus 137209 137209 TS=32
MTPKTELATLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

296_WP_071337239.1 Streptomyces albidoflavus 155064 155064 TS=32
MTPKTELATLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

297_WP_010643005.1 Streptomyces albidoflavus 138167 138167 TS=32
MTPKTELATLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

298_WP_129809367.1 Streptomyces albidoflavus 143068 143068 TS=32
MTPKTELATLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

299_WP_129814060.1 Streptomyces albidoflavus 131410 131410 TS=32
MTPKTELATLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

300_WP_040246700.1 Streptomyces albus 1272677 1272677 TS=32
MEKTPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

301_WP_030578091.1 Streptomyces globisporus subsp. globisporus 151316
151316 TS=32
MNSSEELDLNALEISDLIDELGRDNETLSQVMAASCVGTACACSSTSSSS

302_WP_137310449.1 Streptomyces lasalocidi 7721055 7721055 TS=32
MSVDLKDELSALESATFEIEEMTDESVELAWSSSSCSCSSCCSCSTSSCCSCSSSTS

303_WP_123498760.1 Streptomyces sp. 844.5 4774965 4774965 TS=32
MENRKSDALATLTQEILELESETFEITDYADASEVMNGTCSSTTSSCCSKVT

304_WP_037796980.1 Streptomyces sp. ADI91-18 1298906 1298906 TS=32
MEKSPLASLAEEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

305_WP_123075909.1 Streptomyces sp. ADI95-16 1164220 1164220 TS=32
MEKSPLASLAEEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

306_WP_100582969.1 Streptomyces sp. CB02120-2 57948 57948 TS=32
MEKSPLASLAEEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

307_WP_073794041.1 Streptomyces sp. CB03578 172641 172641 TS=32
MEKSPLASLAEEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

308_WP_087767808.1 Streptomyces sp. CS057 2994390 2994390 TS=32
MDVDAMFASFDIEELDTLEIADGVALPEMGASYLNAPLCCSSSSSCC

309_WP_075986497.1 Streptomyces sp. FR-008 757852 757852 TS=32
MTPKTELATLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

310_WP_159312970.1 Streptomyces sp. GF20 4571275 4571275 TS=32
MTPKTELATLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

311_WP_138053671.1 Streptomyces sp. ICN19 3402272 3402272 TS=32
MCCDLRHPVAGCRRSIAGVRFEVHRRKEVKAVEQQIELDVLEISDLIAGAGENDDLAQV
MAASCTTSSVSTSSSSSSS

312_WP_168542416.1 Streptomyces sp. LD120 108164 108164 TS=32
MEKTPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

313_WP_073775987.1 Streptomyces sp. MJM1172 56069 56069 TS=32
MEKSPLASLAEEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

314_WP_069813260.1 Streptomyces sp. TP-A0874 138483 138483 TS=32
MPTTLKEELHALESATFEIEEMNDSAAVELAWSSSSCSTSSCCSCSTSSCCSTSTSTS

315_OKJ11516.1 Streptomyces sp. TSRI0261 216643 216643 TS=32
MSMAIEAMNADSLFDSFDIEELDTMEVADGVALPEMGASTGNTMCCSSSSSSCC

316_OKJ11516.1 Streptomyces sp. TSRI0261 217018 217018 TS=32
LHAANRLEIVHRRSRGRRPELGGTTKMSKIMETADVDAMFDSFDIEELGTMELADGVAL
PEMGASKGDWTCCSSSSSSCC

317_WP_159749194.1 Streptomyces tubercidicus 1781776 1781776 TS=32
MQDKLVLDLADLSADELEILPTSPGESLEAVNVGHAMVEIGASNCTSRGTPASCCSCCC
C

318_APE26458.1 Streptomyces venezuelae 5537614 5537614 TS=32
VVLGKAGAAELGSALLVLQPNGCGKEFGKEFTMSDDIKQGQDADKAFENFDVDELETL
EVAQGVALPEMGASSGSIGTSSSSSSTCSAC

319_WP_069802514.1 Thermogemmatispora onikobensis 151695 151695 TS=32
MHGELAQLELDGLALDELEIEDLKLSPEGGLELLTTGHGMVEIGASSGQAACCSCCLLC
CCCW

320_WP_069802514.1 Thermogemmatispora onikobensis 152040 152040 TS=32
MHGELAQLELDGLTLDDLEIEDLKLSPEGGLELLTTGHGMVEIGASSGQAACCSCCLVC
CCCPCLG

321_WP_132119980.1 Actinocrispum wychmicini 46992 46992 TS=31
MKDLTFDIDDLRLDDLAVTSMRDAVAIPESGASNVGSSCSCGSSSCCCCQHPTKDPVIIP
PNWA

322_WP_151570259.1 Actinomadura rudentiformis 73230 73230 TS=31
MVSKMCVPDSGMMSSSPIPVTASFSVTATTDSSCMGAPHTVSPRCWPHI

323_WP_173391936.1 Actinomadura sp. DSM 109109 20674 20674 TS=31
MSDMALPDLGDLEIEDLEFSTDQEGGLDALRMGHDVSEVGASCCCSTSSCSCST

324_WP_098207032.1 Bacillus cereus 43346 43346 TS=31
MNKEIQNQANAIDFLELGDVEVLDVEGSLGMPETGASSAPVSWLCCSSSSSCCS

325_WP_098207032.1 Bacillus cereus 43571 43571 TS=31
MQNQANAIDFLELGDVEVLDVEGSLGMPETGASSAPVSWLCCSSSSSCCS

326_WP_113938616.1 Bacillus mycoides 19118 19118 TS=31
MKNKFEPKKEKLDLEEMLAAANLEIIELDQAEALPSTAASSGSSSGNTCSTCGSS

327_WP_119908502.1 Carnobacterium divergens 2438529 2438529 TS=31
MEKELSTKDFDLEVELLDLDEVSAIPETTASSGSTSCSASSTCGSTSCCGSC

328_WP_119908502.1 Carnobacterium divergens 2438746 2438746 TS=31
MERELSVNETTTEDFDLEVELLDLDEVSAIPETTASSGSTSCSASSTCGSTSCCGSC

329_WP_135025206.1 Carnobacterium divergens 79190 79190 TS=31
MEKELSTKDFDLEVELLDLDEVSAIPETTASSGSTSCSASSTCGSTSCCGSC

330_WP_135025206.1 Carnobacterium divergens 79407 79407 TS=31
MERELSVNETTTEDFDLEVELLDSDEVSAIPETTASSGSTSCSASSTCGSTSCCGSC

331_WP_135081547.1 Carnobacterium divergens 79186 79186 TS=31
MEKELSTKDFDLEVELLDLDEVSAIPETTASSGSTSCSASSTCGSTSCCGSC

332_WP_135081547.1 Carnobacterium divergens 79403 79403 TS=31
MERELSVNETTTEDFDLEVELLDSDEVSAIPETTASSGSTSCSASSTCGSTSCCGSC

333_WP_122641943.1 Clostridiales bacterium Marseille-P5551 579434 579434 TS=31
MSEFEKAIEGLEIEELKVSDMMVSESMTEEDAKQIMGASCTTCTCTCSCCTT

334_WP_123701223.1 Curtobacterium sp. PhB130 874603 874603 TS=31
MERNSIIDLPDLDFSGLEVLSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

335_WP_132046886.1 Curtobacterium sp. PhB136 310908 310908 TS=31
MERNSIIDLPDLDFSGLEVLSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

336_KFC17662.1 Cutibacterium acnes HL201PA1 1258 1258 TS=31
MENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

337_GAE70556.1 Cutibacterium acnes JCM 18909 10240 10240 TS=31
VNRMENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

338_ESS85761.1 Cutibacterium acnes P6 10263 10263 TS=31
MENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

339_WP_034345793.1 Deinococcus misasensis DSM 22328 149116 149116 TS=31
MSQNLNNLNDVDFNLDTLEISEVRDSTGLAETGASSGSSSCGSSSCCGSSSCCCNLEASE
QVAP

340_WP_034345793.1 Deinococcus misasensis DSM 22328 149387 149387 TS=31
MNQKKDIMQDLEFDLDTLEISEVRDSTGLAETGASSGSSSCGSSSCCGSSSCCCSMEAQ

341_WP_157458697.1 Deinococcus sp. HMF7620 17894 17894 TS=31
MTNINMTDIDFSDLNIEALEVTEVRDSTALAETGASSGSSSCSATSTCGSSSCCCGSVEAQ
LSE

342_WP_107138905.1 Deinococcus sp. OD32 24443 24443 TS=31
MTINTPMTDIDFSDLNIEALEVTEVRDSTALAETGASSGSSSCSATSTCGSSSCCCGSVEV
VAE

343_WP_154775326.1 Erwinia sp. CPCC 100877 37395 37395 TS=31
MEQEKELQNLLEELEIETIEMDNIQAIPTAAASSGSGKGTLCGSSSSSTCSSCCG

344_ABD13543.1 Frankia casuarinae 5019526 5019526 TS=31
VTNVSTPYGCGCVPYRSDRRITLSAYRLTQSMMKERTMSALASEKPMDDKLFAGFDIEE
LEVLEVSDSVALPEMGASSGSGILCCSSSSSSSSSCC

345_WP_008487053.1 Idiomarina xiamenensis 10-D-4 128223 128223 TS=31
MLTENINFDDLNVDDVNLDGLEVANVSNAMALPETGASSGTSSCGSCSTCGSSSCCGSCGGGDVSVQQEIKAN

346_WP_116950971.1 Jiangella sp. KE2-3 80198 80198 TS=31
MSNETLGLELLELEAVTFEIGEISDPSLEMAATSSSCSCSSCSSCSTSSCCSTSTSTSSCG

347_WP_116950971.1 Jiangella sp. KE2-3 80474 80474 TS=31
MSNATLGLELLELEAVTFEIDELTDPSVELAATSSSCSCSSCSSCSTSSCCSCSTSTSSCG

348_WP_116950971.1 Jiangella sp. KE2-3 80855 80855 TS=31
VDADLACGQRISRRRSPCWCSSSSSTCCRRSSSSRSSSSCTSRRRPARRR

349_WP_116950971.1 Jiangella sp. KE2-3 80730 80730 TS=31
VTQMALRDELLVLETATFEVEDITDVGQDAAGWCSSSCSCSSCSSCSTSSCCSTSTSTGCGGG

350_WP_056374170.1 Microbacterium sp. Leaf161 678286 678286 TS=31
MSETMELADIDVMFDEIEIIELDETVTLSETAASSVSSSCNSCSCCASTSCCSVNVFVGL

351_WP_091261732.1 Micromonospora chaiyaphumensis 176127 176127 TS=31
MPELTEELRALETETFEIEDVDAVDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

352_WP_111214312.1 Micromonospora craterilacus 8830 8830 TS=31
MPELTDELRTLETETFEIEDVDAIDSMVMAWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

353_WP_088993330.1 Micromonospora echinaurantiaca 1890910 1890910 TS=31
MPDSVRSELRDLQTETFEVEDIADLSADLMDICSSSTSTSSCSSCSTSSCCSCTSSSCSSTS

354_WP_091247327.1 Micromonospora matsumotoense 238935 238935 TS=31
MSGQADLDLIAQEILELETETFAISDYTDVTEAVLASTSSSATTSTTSSTTSSTSTASSTSCCG

355_WP_120675702.1 Micromonospora musae 201578 201578 TS=31
MRIPSKGEAVKKSDELDFELEDLPMDVFDLAESGLTIESLTAGHGMPEHGASLPFCSCSATCSCCPSSSS

356_WP_120689006.1 Micromonospora musae 111493 111493 TS=31
MRIPSKGEAVKKSDELDFELEDLPMDVFDLAESGLTIESLTAGHGMPEHGASLPFCSCSATCSCCPSSSS

357_WP_091191776.1 Micromonospora narathiwatensis 939594 939594 TS=31
MPELTEELRALETETFEIEDVDAVDEMVMAWSSSSCSCSSCCSCSTSSCCSTSTSTGCGGG

358_WP_091639110.1 Micromonospora pallida 640971 640971 TS=31
MPITLNHELRELETETFEIEEVEDGAEALAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

359_WP_151488655.1 Micromonospora sp. ALFpr18c 19502 19502 TS=31
MPDLTEELRALETETFEIEDVDAIDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

360_WP_073835304.1 Micromonospora sp. CB01531 77059 77059 TS=31
MPELTEELRALETETFEIEDVDAIDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

361_WP_167184099.1 Micromonospora sp. CNZ280 593477 593477 TS=31
MPESVRSELRDLQTETFEVEDIADLSADLMDICSSSTSTSSCSSCSTSSCCSCTSSSCSSTS

362_WP_165438069.1 Micromonospora sp. CNZ295 4959453 4959453 TS=31
MPELTEELRALESETFEIEDVDAIDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

363_WP_165438069.1 Micromonospora sp. CNZ295 4960122 4960122 TS=31
MPITLNDELRELETETFEIEEVEDGGEALAASSSSCSCSSCCSCSTSSCCSTSTSTSSCG

364_WP_046564724.1 Micromonospora sp. HK10 74413 74413 TS=31
MPELTEELRALETETFEIEDVDAIDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

365_WP_123600439.1 Micromonospora sp. Llam0 379236 379236 TS=31
MLAELSTTWDATQDYDLDIDVDGLDLGPLTVTAMRDSVALPETGASSAGDGGRASCSCCYVT

366_WP_148799196.1 Micromonospora sp. MP36 43526 43526 TS=31
MPELTEELRALETETFEIEDVDAIDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

367_WP_165521730.1 Micromonospora zingiberis 473342 473342 TS=31
MPELSEELRTLETETFEIEDVDAIDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

368_WP_165521730.1 Micromonospora zingiberis 474003 474003 TS=31
MPITLNDELRELETETFEIEEVTDGGETLAASSSSCSCSSCCSCSTSSCCSTSTSTSSCG

369_WP_171023219.1 Micromonosporaceae bacterium KJ-029 73986 73986 TS=31
MPELTDELRTLETETFEIEDVDAIDAMVMAWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

370_WP_017606430.1 Nocardiopsis alkaliphila YIM 80379 349 349 TS=31
MTTKEIPTIELPALDFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

371_WP_073380611.1 Nocardiopsis flavescens 140959 140959 TS=31
MQDSPSIELPDLDFSDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

372_WP_067608788.1 Nocardiopsis listeri NBRC 13360 4896 4896 TS=31
MSNTFTSLRHELEQLESETFEVLDFADSPEEMLSGTTSTSSTSCCSSCTSCTSTASSCA

373_WP_017544320.1 Nocardiopsis prasina DSM 43845 325 325 TS=31
MALQETPTIDLPDLNFDDMEVMSVREAVAVPETGATSGSSSCTSPSCCGSSSCCSSGSCG

374_WP_150240357.1 Nocardiopsis quinghaiensis 131025 131025 TS=31
MAMQDSPVIDLPDLDFDSMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

375_WP_077692331.1 Nocardiopsis sinuspersici 176148 176148 TS=31
MAMQDSPVIDLPDLDFDSMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

376_WP_159944718.1 Nocardiopsis sp. FR26 17161 17161 TS=31
MAMQESPAIDLPDLDFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

377_WP_110052479.1 Nocardiopsis sp. L17-MgMaSL7 30664 30664 TS=31
MALQETPTIDLPDLNFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

378_WP_073704895.1 Nocardiopsis sp. TSRI0078 219471 219471 TS=31
MAMQDSPVIDLPDLDFNDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

379_WP_033352405.1 Nocardiopsis xinjiangensis YIM 90004 436999 436999 TS=31
MSHSLSGLRHELEQLESETFEVMDFAEAPEEMLGNTTSTSSTSCCSSCTSCTSTASSCA

380_WP_033352405.1 Nocardiopsis xinjiangensis YIM 90004 437216 437216 TS=31
MSHSLSGLRHELEQLESETFEVVDFAEASEEMLAGTTSTSSTSCTSSCCSCTSCAS

381_WP_033352405.1 Nocardiopsis xinjiangensis YIM 90004 437557 437557 TS=31
MTKKTGSDHLSGLRGELELLESETFEVLDLVEAREEMLGSTTSTTSCTSSCSSSSCSSTSCSGG

382_WP_106246561.1 Nonomuraea fuscirosea 63112 63112 TS=31
MVEKPLSDLFQELEQLGSETFEVSDYTEVDIHLSPVAAATSCCGSTSTTSTSCSSGTQSGE
C

383_SEM92144.1 Nonomuraea pusilla 31086 31086 TS=31
MSVSKGLDFDLANLPMDVFELADSGLGLESLTAGHGLTETGASTSALCSCSVACSSSS

384_WP_165977801.1 Nonomuraea sp. KC712 16169 16169 TS=31
MDVFDLTDSGLTIESLTAGHGMTENGASTTCVCSSGCSS

385_WP_082535677.1 Nonomuraea sp. NBRC 110462 239301 239301
TS=31
MDVFELADSGLGLESLTAGHGLTETGASTSALCSCSVACSSSS

386_WP_143782003.1 Ornithinimicrobium sp. H23M54 564036 564036
TS=31
MTTRFPHEELVQLEVETFEVLELAEVAEDMAAWCSCSTSSCCSTSSSSCSSTCSCSCASC
A

387_WP_131104644.1 Ornithinimicrobium sp. HY008 14872 14872 TS=31
MSAPTLTTTIELPDLDFDGLEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

388_WP_036732400.1 Paenibacillus sp. FSL R7-277176317 176317 TS=31
MKDLNKELEALDLELIELDQVEAIPSTAASSGSGSSSTCGSSSCSSCSSSCA

389_WP_036732400.1 Paenibacillus sp. FSL R7-277176531 176531 TS=31
MKDLNKELEDLDLELIELEQVEAIPSTAATSGSGSSSTCGSSSCSSCSSSCA

390_WP_036732400.1 Paenibacillus sp. FSL R7-277176748 176748 TS=31
MKSFDKELEDLDLELIELDQVEAIPSTAATSGSGSSSTCGSSSCSSCSSSCA

391_WP_160499259.1 Paenibacillus sp. HJL G12 149025 149025 TS=31
MSQTLNKELNKELENLDLELIELDQAEAIPSTAASSGSGSSSTCGSSSCSSCSSSCA

392_WP_160499259.1 Paenibacillus sp. HJL G12 149253 149253 TS=31
MSQSLNKELNKDLENLDLELIELDQVEAIPSTAASSGSGSSSTCGSSSCSSCSSSCA

393_WP_160499259.1 Paenibacillus sp. HJL G12 149483 149483 TS=31
MSNLNKELNKELENLDLELIELDQVEAIPSTAATSGSGSSSTCGSSSCSSCSSSCA

394_WP_036707317.1 Paenibacillus sp. P1XP2 14921 14921 TS=31
MSDLNMELNKELENLDLELIELDQAEAIPSTAASSGSGSSSTCGSSSCSSCSSSCA

395_WP_036707317.1 Paenibacillus sp. P1XP2 15149 15149 TS=31
MNQLNKELNKELENLDLELIELDQAEAIPSTAASSGSGSHSTCGSSSCSSCSSSCA

396_WP_160038800.1 Paenibacillus sp. USDA918EY 150023 150023 TS=31
MNQLNKELNKELENLDLELIELDQAEAIPSTAASSGSGSSSTCGSSSCSSCSSSCA

397_WP_160038800.1 Paenibacillus sp. USDA918EY 150247 150247 TS=31
MSDLNMELNKELENLDLELIELDQAEAIPSTAASSGSGSSSTCGSSSCSSCSSSCA

398_WP_143802042.1 Paenibacillus thiaminolyticus 1669713 1669713 TS=31
MKEEEIRIMLDDLELDQVEIRGLIVEDGKAIPDMGATIGAISCCSCCST

399_WP_143802042.1 Paenibacillus thiaminolyticus 1670089 1670089 TS=31
LHLDDLELDDVNIIALTQDDAKGLPDMGATMMWVGCCSCCST

400_WP_163743638.1 Phytoactinopolyspora halotolerans 30727 30727 TS=31
MSNKALNLELVDLEAVTFEVEELADPSMDMAASSSSCSTSSCSSCSTSSCCSCSTSTSSC
G

401_WP_165968252.1 Saccharopolyspora sp. 7K50277049 77049 TS=31
MEDLKLDLESLQVESLAFGADADHDLQSLGMGHGMTEVGASAGCCCCTCSCCCCPCG

402_WP_147133977.1 Stackebrandtia albiflava 993006 993006 TS=31
MNPLIKDIAELEAQTFEIDDVAQLSQDEAVTWSSCSTSSCCGCSSSSCGSCCGCSSSCGST
SG

403_WP_147133977.1 Stackebrandtia albiflava 993241 993241 TS=31
MNPLIKDIAELEAQTFEIDDVAQLSQDEAVTWSSCSTSSCCGCSSSSCGSCCGCSCSCGST
SK

404_WP_142043611.1 Stackebrandtia endophytica 4635036 4635036 TS=31
MSELNDQIALLEAETFEVIDSTPSEDFAGSSSTSSSTSSTTCSCSTSSTCSTSSSSCSSTSSCG

405_WP_124728414.1 Staphylospora marina 2496897 2496897 TS=31
MDLSDAAALPETGASSGESSCSVASSTCGSSSCCSSCW

406_WP_016397512.1 Streptococcus pneumoniae 14110 14110 TS=31
MDKQRDLDISFLDDFEFEVIELDDVSVLPETAASSGSSGNNVHSTCGSCSCCSSCCT

407_WP_149561037.1 Streptococcus salivarius 45881 45881 TS=31
MKNEKNLLEELEIDILELDDVEGIPATAASSGSGGSSTCGSTSCSSCSSSCA

408_WP_149561037.1 Streptococcus salivarius 46089 46089 TS=31
MKDEKNLLEELEIDILELDDVEGIPATAASSGSGGSSTCGSTSCSSCSSSCA

409_WP_149561037.1 Streptococcus salivarius 46298 46298 TS=31
MKDEKNLLEELEIDILKLDDVEGIPATAASSGSGGSSTCGSTSCSSCSSSCA

410_WP_149561037.1 Streptococcus salivarius 46507 46507 TS=31
MKDEKNLLEELEIDILELDDVEGIPATAASSGSGGSSTCGSTSCSSCSSSCA

411_WP_149561037.1 Streptococcus salivarius 46716 46716 TS=31
MKDEKNLLEELEIDILELDDVEGIPATAASSGSGGSSTCGSTSCSSCSSSCA

412_WP_149561037.1 Streptococcus salivarius 46925 46925 TS=31
MKDEKNLLEELEIDILELDDVEGIPATAASSGSGGSSTCGSTSCSSCSSSCA

413_WP_149561037.1 Streptococcus salivarius 47134 47134 TS=31
MKDEKNLLEELEIDILELDDVEGIPATAASSGSGGSSTCGSTSCSSCSSSCA

414_WP_040270458.1 Streptomonospora alba 91856 91856 TS=31
MDKAPAVDFGDLDLGDLSFDSVEVATVRDAVALPETGASSGSSSCESCSCCAVCSCCSV
CST

415_WP_040270458.1 Streptomonospora alba 92143 92143 TS=31
MNKAPTVDFGDLDLGDLSFDSVEVATVRDAVALPETGASSGSSSCDSCSCCVGCSCCSV
CST

416_WP_040270458.1 Streptomonospora alba 92427 92427 TS=31
MDKTQTVDFGDLDLGDLSFDSVEVATVRDAVALPETGASSASSSCDSCSCCVGCSCCSV
CST

417_WP_086715354.1 Streptomyces angustmyceticus 63764 63764 TS=31
MQDKLVLDLADLSADELEILPTSPGESLEAVNVGHAMVEIGASNCTSRGTPASCCSCCC
C

418_WP_044375373.1 Streptomyces badius 4532734 4532734 TS=31
MTPKTELATLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

419_WP_114931476.1 Streptomyces cavourensis 2845046 2845046 TS=31
MAPKTELTTLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

420_WP_142229169.1 Streptomyces cavourensis 2850763 2850763 TS=31
MAPKTELTTLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

421_WP_059265103.1 Streptomyces corchorusii 19103 19103 TS=31
MEKELVLDLADLSVDELDVLPTSPGAGLESINVGHAMVEIGASNCTSTGTPASCCSCCC
C

422_WP_010059797.1 Streptomyces globisporus C-1027 3132781 3132781 TS=31
MTPKTELATLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

423_WP_030589129.1 Streptomyces globisporus subsp. globisporus 250464 250464 TS=31
MTPKTELATLADEIMELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

424_WP_014677376.1 Streptomyces hygroscopicus subsp. jinggangensis TL01 9779518 9779518 TS=31
MEKELVLDLADLSVDELDVLPTSPGAGLESINVGHAMVEIGASNCTSTGTPASCCSCCC
C

425_WP_051840688.1 Streptomyces lavendulae subsp. lavendulae 841344 841344 TS=31
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

426_WP_033265171.1 Streptomyces lydicus 3558536 3558536 TS=31
MNDDLVLDLADLSVDDLDILPASPGASLETVNVGHAMVEIGASNCTSTGTPASCCSCCC
C

427_WP_077191410.1 Streptomyces lydicus 443547 443547 TS=31
MNDDLVLDLADLSVDDLDILPASPGASLETVNVGHAMVEIGASNCTSTGTPASCCSCCC
C

428_WP_030803226.1 Streptomyces mediolani 118583 118583 TS=31
MTPKTELATLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

429_WP_152264682.1 Streptomyces mobaraensis 168365 168365 TS=31
MEAKLELDLGDLSVEELDVLPTSPGAGLESINVGHAMVEIGASNCTSRGTPASCCSCCC
C

430_WP_046505522.1 Streptomyces odonnellii 27349 27349 TS=31
MADKSTLASLADEIMELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

431_WP_099872715.1 Streptomyces sp. 76 419113 419113 TS=31
MNDDLVLDLADLSVDDLDILPASPGASLETVNVGHAMVEIGASNCTSTGTPASCCSCCC
C

432_WP_136228001.1 Streptomyces sp. A0592 75269 75269 TS=31
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

433_WP_037796980.1 Streptomyces sp. ADI91-18 1315523 1315523 TS=31
LDRERTATTRSSSSSRSACRSRSPCWCSSSSSRPRTSCTPARCWWRRPRWRP

434_WP_137234306.1 Streptomyces sp. BPSDS2 76482 76482 TS=31
MAPKTELATLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

435_OKI60309.1 Streptomyces sp. CB00072 834948 834948 TS=31
VPVAVGHHTRRRSSGRRARAWGYTEMSKIMETMDVDAMFDSFDIEELETMEISDGIALP
EMGASYLGAPACCSSSSSCC

436_WP_145501173.1 Streptomyces sp. CFMR 7 92154 92154 TS=31
MSLPEMSTTFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TA

437_WP_053557900.1 Streptomyces sp. CFMR 7 91753 91753 TS=31
MSLPEMSTTFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TA

438_WP_053630189.1 Streptomyces sp. MMG1064 92673 92673 TS=31
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

439_WP_045322258.1 Streptomyces sp. NRRL F-4428 39374 39374 TS=31
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

440_WP_030959257.1 Streptomyces sp. NRRL S-378 133471 133471
TS=31
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

441_EYU71052.1 Streptomyces sp. PCS3-D2 1213911 1213911 TS=31
MQKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

442_WP_065486866.1 Streptomyces sp. PTY087I2 138728 138728 TS=31
MAPKTELATLADEILELESETFEISDYSDVAEVVLAGSTSCSSTSTCSSTTSTTSCSA

443_WP_018492813.1 Streptomyces sp. SID8356 164062 164062 TS=31
MAPKTELTTLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

444_WP_028417601.1 Streptomyces sp. SID8359 376777 376777 TS=31
MAPKTELTTLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

445_WP_079273369.1 Streptomyces sp. TN58 6865695 6865695 TS=31
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

446_OKJ11516.1 Streptomyces sp. TSRI0261 217955 217955 TS=31
MSTAGLTAGCPTAWGHIKMSKILESVDADAMFDSFDIEELETMELADGVALPEMGAST
APAPLCCSSSSSACC

447_WP_030899367.1 Streptomyces virginiae 306816 306816 TS=31
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

448_WP_031151278.1 Streptomyces xanthophaeus 218828 218828 TS=31
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

449_WP_031152038.1 Streptomyces xanthophaeus 322083 322083 TS=31
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

450_WP_016906219.1 Streptomyces xiaopingdaonensis 6646 6646 TS=31
MSNVTVDSMDIEQVFDSFDVTELEVLEVSQGVAQPEMGASGGTVGTSSSSSTCSSSTC

451_WP_046775770.1 Streptomyces yangpuensis 205357 205357 TS=31
MDKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

452_WP_112425990.1 Thermogemmatispora tikiterensis 389607 389607 TS=31
VAETFCSGAFSSASNVRRCSSSGSRSISVRSSASRSISPNTSSLSVSPSSSSLPTSSSSSTASP
TC

453_WP_126711131.1 Verrucosispora sp. FIM060022 1450631 1450631 TS=31
MPELTEELRALETETFEIEDVDAIDAMVMDWSSSSCSCSSCCSCSTSSCCSSSTSTGCGGG

454_WP_102658257.1 Verrucosispora sp. ts21 8803 8803 TS=31
MPELTEELRALETETFEIEDVDAIDAMVMDWSSSSCSCSSCCSCSTSSCCSTSTSTGCGGG

455_GES11969.1 Acrocarpospora macrocephala 13739 13739 TS=30
MQELAREFAVLDLGTFEIEEIADLGSAFASVAEVPADGGCSTSSCCSCSTSSCCSCSCSDT
SCSSSSCSASSCY

456_WP_170224452.1 Actinokineospora cianjurensis 1159877 1159877 TS=30
MSVPNLDFDLENLPMDVFDLGQNGLTIESLTSGHGMVENGASVPCPSVCSTSTISCSSS

457_WP_124934044.1 Actinomyces bowdenii 1620 1620 TS=30
MNNVIDFAAIEISDLIEDAVDGGELPSQVMAASTTTSGCACSSCSSTCS

458_WP_126382269.1 Actinomyces howellii 2257 2257 TS=30
MNTFDIPMDVFELEDSGMDVEPLTAGHGMTEVGASTNCFCYPCCSCSAPSSSSA

459_WP_130475466.1 Amycolatopsis suaedae 339145 339145 TS=30
MRKWGMSMEGLKLELESLEVESLSFGGDAERDLQSLGMGHGMTEVGASAGCCCCTCS
CCCCPCG

460_WP_098481569.1 Bacillus cereus 18661 18661 TS=30
MQNQANAIDFLELGDVEVLDVEGSLGMPETGASSAPVSWLCCSSSSSCCS

461_WP_098481569.1 Bacillus cereus 18886 18886 TS=30
MNKEIQNQANAIDFLELGDVEVLDVEGSLGMPETGASSAPVSWLCCSSSSSCCS

462_WP_000438834.1 Bacillus cereus ATCC 14579 5000697 5000697 TS=30
MSEIKKALNTLEIEDFDAIEMVDVDAMPENEALEIMGASCTTCVCTCSCCTT

463_WP_000438834.1 Bacillus cereus ATCC 14579 5000901 5000901 TS=30
MSEIKKALNTLEIEDFDAIEMVDVDAMPENEALEIMGASCTTCVCTCSCCTT

464_WP_000438834.1 Bacillus cereus ATCC 14579 5001105 5001105 TS=30
MSEIKKALNTLEIEDFDAIEMVDVDAMPENEALEIMGASCTTCVCTCSCCTT

465_WP_000438834.1 Bacillus cereus ATCC 14579 5001309 5001309 TS=30
MSEIKKALNTLEIEDFDAIEMVDVDAMPENEALEIMGASCTTCVCTCSCCTT

466_WP_159145314.1 Bacillus sp. 9J 1612729 1612729 TS=30
MTNLKKALNSIEIEELDVTEMVDAEAMSEEDATQIMGASCTTCVCTCSCCTT

467_WP_166378646.1 Catellatospora methionotrophica 633900 633900 TS=30
MADTSTTWDLIPEYDIAFDMDDLDLGPLTVTAMRDSMALPETGASWINTSGNASCSCC
YVT

468_WP_123701223.1 Curtobacterium sp. PhB130 887874 887874 TS=30
MSGSMLELPDLDFDGLEVLSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSAGSCG

469_WP_132046886.1 Curtobacterium sp. PhB136 297637 297637 TS=30
MSGSMLELPDLDFDGLEVLSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSAGSCG

470_WP_016622713.1 Enterococcus faecalis 104938 104938 TS=30
MEKELKTVDVEVEELDELFGEVELIEIDETITLSETAASSGISSCNSCSCCASTSCCSSSSVS
FT

471_WP_049761140.1 Frankia sp. B2 6065 6065 TS=30
MDDKLFAGFDIEELEVLEVSDSVALPEMGASSGSGILCCSSSSSSSSCC

472_SFL62282.1 Gracilibacillus orientalis 561620 561620 TS=30
MATENNVLEEELMDFNDVDLDEAFEGVEVIELDESVTLSETAASSGISSCTSCSCCASTS
CCSVKIT

473_WP_027344871.1 Hamadaea tsunoensis DSM 44101 46000 46000 TS=30
MAEGSITWDLVPEYDFAFDLDDLDLGPLTVTAMRDSMALPETGASWVNTSGNASCSCC
YVT

474_WP_100586424.1 Kitasatospora sp. CB02891 32685 32685 TS=30
MADTSLASLAQEILNLESETFEISDYSDTSEVMLGSSTSCSSTSTCSSTTSTTSCTA

475_WP_077018665.1 Kribbella sp. ALI-6-A1658587 1658587 TS=30
MTDLTTELLSLETATFEVEEIADVGADSASWCSSSTSTSSCSSCSTSSCCGCSSCSCSSTS

476_WP_132283874.1 Kribbella sp. VKM Ac-2568 58548 58548 TS=30
MTDLTTELLSLETATFEVDEIADVGQDAADWCSSSTSTSSCSSCSTSSCCGCSSCSCSSTS

477_WP_129894060.1 Ktedonosporobacter rubrisoli 10890218 10890218 TS=30
MLSEEIQEKIGQLDAEIGGELHQLEAETFEVLDFMDAEQNTPPSMCTSTSCTSTSSTSSCT
STTSCS

478_WP_093590332.1 Lentzea waywayandensis 1164236 1164236 TS=30
MKDLSFDPDDLELGDLAVTSLRDSVALPESGASGGASSCSCGSSSCSSCTQPPVQNA

479_WP_115858067.1 Lysobacter sp. zong2l5 1314682 1314682 TS=30
MNNDFDFDSLDLASLNIEGMEVVTLKEAMALPETGASSAVSICSSCSSCGSSSCFNDVAT
VE

480_WP_145876768.1 Massilia flava 4141542 4141542 TS=30
MKNDIKAAIPADTTDNKADSLADIAQENFEVEDINDLNMFAPLSDISLCSSSSTSSCSSCC
S

481_WP_091261732.1 Micromonospora chaiyaphumensis 175865 175865 TS=30
MPDSVRSELRDLQTETFEIEDIADLSAELMDICSSSTSTSSCSSCSTSSCCSCTSSSCSSTS

482_WP_089022234.1 Micromonospora coriariae 6788875 6788875 TS=30
VGYGTGITAMSGPVTPASASVPSSAWASPTRPTVSSRGCCTPAGCSGHACRRYSPCTDA
TYQRNGT

483_WP_121160063.1 Micromonospora pisi 7445194 7445194 TS=30
MIEMLDLADLDLDDITVTSMRDSVALPETGASSGSSSCSASSTCGSSSCCASCGGGGSVN
PDDGSTGG

484_WP_151488655.1 Micromonospora sp. ALFpr18c 19240 19240 TS=30
MSDSVRGELRDLQTETFEVEDIADLSADHMDICSSSTSTSSCSSCSTSSCCSCTSSSCSSTS

485_WP_073835304.1 Micromonospora sp. CB01531 76798 76798 TS=30
MPDSVRSELRDLQTETFEVEDIAELSADLMDICSSSTSTSSCSSCSTSSCCSCTSSSCSSTS

486_WP_165438069.1 Micromonospora sp. CNZ295 4959714 4959714 TS=30
MPTQPRTPAALPDSVRTELRDLQTETFEVEDIADLSADLMDICSSSTSTSSCSSCSTSSCCS
CTSSSCSSTS

487_WP_046564724.1 Micromonospora sp. HK10 74151 74151 TS=30
MSDSVRSELRDLQTETFEVEDIADLSADLMDICSSSTSTSSCSSCSTSSCCSCTSSSCSSTS

488_WP_148799196.1 Micromonospora sp. MP36 43264 43264 TS=30
MPDSVRSELRDLQTETFEVEDIAELSADLMDICSSSTSTSSCSSCSTSSCCSCTSSSCSSTS

489_WP_165521730.1 Micromonospora zingiberis 473604 473604 TS=30
MPTQPPPSGSLPDSVRSELRDLQTETFEVEDIAELSADLMDICSSSTSTSSCSSCSTSSCCS
CTSSSCSSTS

490_WP_171023219.1 Micromonosporaceae bacterium KJ-029 73726 73726 TS=30
MPTQPPTSSSLSDSVRSELRDLQTETFEVEDIAELSDDLMGICSSSTSTSSCSSCSTSSCCSC
TSSSCSSTS

491_WP_067138951.1 Microtetraspora malaysiensis 273569 273569 TS=30
MSELARELVALETVTFEVEDIPDLGRDAVDWCSSSTSTSSCSSCSTSSCCSCTSSSCGSTS

492_WP_073380611.1 Nocardiopsis flavescens 140484 140484 TS=30
MPETTALDIDLPELNFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSGGSCG

493_WP_067608788.1 Nocardiopsis listeri NBRC 13360 5107 5107 TS=30
MTIGNIQGLRHELEQLESETFEVLDFVDSPEEMLAGTTSTSSTSTCSSCTSCTSCASF

494_WP_150240357.1 Nocardiopsis quinghaiensis 131359 131359 TS=30
MPHSTAPDIDLPELDFDNMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSGGSCG

495_WP_077692331.1 Nocardiopsis sinuspersici 175814 175814 TS=30
MPHSTAPDIDLPELDFDSMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSGGSCG

496_WP_159944718.1 Nocardiopsis sp. FR26 16834 16834 TS=30
MPHSMAPDIDLPELDFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSGGSC
G

497_WP_087097343.1 Nocardiopsis sp. JB363 9019 9019 TS=30
MNNTFTSLRHELEQLESETFEVLDFADSPEEMLSGTTSTSSTSTCSSCTSCTSTASSCA

498_WP_110052479.1 Nocardiopsis sp. L17-MgMaSL7 30354 30354 TS=30
MPHSTAPEIDLPELNFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSGGSCG

499_WP_073704895.1 Nocardiopsis sp. TSRI0078 219136 219136 TS=30
MPHSTAPDIDLPELDFDNMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSGGSCG

500_WP_017609023.1 Nocardiopsis xinjiangensis YIM 90004 151893 151893
TS=30
MSESMNLRLDDLLVESVDFADTDEAASLETLTTGQGMTEIGASCGICGSSSCC

501_WP_090933862.1 Nonomuraea jiangxiensis 249694 249694 TS=30
VPSNTLLALDDLDIEAIHVVEIPGIADSLAAGHSMPETGASACCSIVCSSCCCSC

502_WP_143782003.1 Ornithinimicrobium sp. H23M54 563820 563820 TS=30
VKKTMDPFLTTEIASLEAETFEVLELTEVAEDMAAWSSCSTSSCCSTSSSSCGSTSSCSCS
CASTSSCA

503_WP_131104644.1 Ornithinimicrobium sp. HY008 34403 34403 TS=30
MTLRRTLRVRVISSSSGSSCLCSSTNRWTRIASGRSSLECTISSAMSS

504_WP_163743638.1 Phytoactinopolyspora halotolerans 30209 30209 TS=30
MTATETSLTSELLTLESVTFEVEEIADLGQDGAGWCSSSTSTSSCSSCSTSSCCGCSSCSCS
STS

505_WP_162452872.1 Phytoactinopolyspora sp. XMNu-373 96501 96501 TS=30
MTQTATSLSSELLALETVTFEVEEIADVGQDAAGWCSSSTSTSSCSSCSTSSCCGCSSCSC
SSTS

506_WP_093271394.1 Saccharopolyspora shandongensis 37017 37017 TS=30
VGYGMEDLKLELESLQVESLAFGADADRDLQSLGMGHGMTEVGASAGCCCCTCSCCC
CPCG

507_WP_158847633.1 Saccharothrix deserti 82460 82460 TS=30
MDLTFDVEELALDDLAVVAMRDAVALPETGASGGSCSGGSSCCCCPSPQLPD

508_WP_156077241.1 Saccharothrix sp. NRRL B-16314 1639 1639 TS=30
MDLTFDDSELDLGDLAVTAMRDAVALPETGASAQACSCSSTSCCCCQVPQLPA

509_WP_170184966.1 Saccharothrix texasensis 1633738 1633738 TS=30
MDLTFDDSELDLGDLAVTAMRDAVALPETGASAAACSCSSTSCCCCQQPQLPTLPTQG

510_WP_147133977.1 Stackebrandtia albiflava 992554 992554 TS=30
MSPLNDQIAVLEAETFEVVDFTPAEDLAASTSTSCACSTTSTCSCSTSSCCSTTSSSCSSTS
SCA

511_WP_124728414.1 Staphylospora marina 2497100 2497100 TS=30
MDMEIRHLEVTGLEVMELSDAAALPETGASSGITPCTYCSTCGSSSCCASCQ

512_WP_152645972.1 Streptacidiphilus albus JL83 9721254 9721254 TS=30
VPLCAARPGHAEGPVQSSCPARSGSAHCSRFRPVTTNSAFGIVHTTCCSSRSRSRWVRSP
TSSQVSSSTTGTTGCPPPSCARSQASRCGR

513_WP_152645972.1 Streptacidiphilus albus JL83 9695602 9695602 TS=30
VCAEGAREGGERGPPNAAGTGGTSSTTACTPSSLGSSPSSSVTSASGKQSRAVAFWSPAT
TSACATWIPASRQASR

514_WP_002998788.1 Streptococcus downei MFe28173335 173335 TS=30
MKDENKLLEELEIDILELDDVEGVPATAATSGSGGSSTCGSTSCSSCSSSCA

515_WP_002998788.1 Streptococcus downei MFe28 173531 173531 TS=30
MKDENKLLEELEIDILELDDVEGAPATAASSGSGGSSTCGSTSCSSCSSSCA

516_WP_039697800.1 Streptococcus equinus ATCC 33317 124433 124433 TS=30
MKDEMKLFEELEIDILELDDVEGVPATAASSGSGGHSTCGSTSCSSCSSSCA

517_WP_115267823.1 Streptococcus hyointestinalis 106414 106414 TS=30
MKDEKNLLEELEIDILELDDVEGVPATAATSGSGGHSTCGSTSCSSCSSSCA

518_WP_115267823.1 Streptococcus hyointestinalis 106620 106620 TS=30
MKDEKNLLEELEIDILELDDVEGVPATAATSGSGGHSTCGSTSCSSCSSSCA

519_WP_115267823.1 Streptococcus hyointestinalis 106826 106826 TS=30
MKDEKNLLEELEIDILELDDVEGVPATAATSGSGGHSTCGSTSCSSCSSSCA

520_WP_138068656.1 Streptococcus porcinus 1302514 1302514 TS=30
MFDNDINLLDAAIENLDVSEFEVLSNDVLSEANEKTVIGASCTTCVCTCSSCCSA

521_WP_102210137.1 Streptococcus sp. UMB0029 58021 58021 TS=30
MKDEKKLFEDLEIDILELDDVEGVPATAASSGSGGFSTCGSTSCSSCSSSCA

522_WP_102210137.1 Streptococcus sp. UMB0029 58212 58212 TS=30
MKDEKKLLEELEIDILELDDVEGVPATAATSGSGGYSTCGSTSCSSCSSSCA

523_WP_070012657.1 Streptomyces abyssalis 130664 130664 TS=30
MADKSSLDSLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

524_WP_008409260.1 Streptomyces albidoflavus 139180 139180 TS=30
MTPKTELATLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

525_WP_067299370.1 Streptomyces griseochromogenes 265028 265028 TS=30
MGNNEEYFVDVNDLSIDVFDVVEQGGAVTALTADHGMPEVGASTNCFCYICCSCSSN

526_WP_093786373.1 Streptomyces guanduensis 135754 135754 TS=30
MADISKLDAPLASLAQEILELESETFEITDYSDASEVMLGSSTSCSSTSTCSSTTSTTSCTA

527_WP_055593610.1 Streptomyces hirsutus 11408 11408 TS=30
MNKDLSTLADEILELEAETFEISDYSDASEVVLAGCTSTSSTSTSSSTCSTTSCSA

528_ACS50127.1 Streptomyces hygroscopicus 15724 15724 TS=30
MEKELVLDLADLSVDELDVLPTSPGAGLESINVGHAMVEIGASNCTSTGTPASCCSCCC
C

529_WP_159491596.1 Streptomyces libani subsp. libani 1166040 1166040 TS=30
MEDNEFDFNLEDIPSDVFELADRGLTVESLTSGHGLVENGASSPSCGSSCSSLP

530_WP_127152319.1 Streptomyces lydicus 3419629 3419629 TS=30
LATIWWPLTGSMPSSKSPLGLASSSSAYSTVSSSGARSSGTSARWCPSASFSPRSTYGP

531_WP_004943480.1 Streptomyces mobaraensis NBRC 13819 = DSM 40847 39195 39195 TS=30
MEAKLELDLGDLSVEELDVLPTSPGAGLESINVGHAMVEIGASNCTSRGTPASCCSCCC
C

532_WP_023541200.1 Streptomyces niveus 3558 3558 TS=30
MAPKTGLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

533_WP_070195834.1 Streptomyces oceani 56745 56745 TS=30
MNDKTQLASLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

534_WP_055471433.1 Streptomyces pathocidini 10908 10908 TS=30
MAADLVLDLADLSVDELDILPTSPGASLESINVGHAMVEIGASNCTSTGTPASCCSCCCC

535_WP_028801239.1 Streptomyces sp. 142MFCol3.1 25970 25970 TS=30
MADSPLASLAQEILDLESETFEITDYSDASEVMLGSSTSCSSTSTCSSTTSTTSCTA

536_WP_084994964.1 Streptomyces sp. 196(2019) 25659 25659 TS=30
MAPKTELTTLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

537_WP_069628351.1 Streptomyces sp. 4R-3d 204148 204148 TS=30
MAPKTGLATLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

538_WP_123498760.1 Streptomyces sp. 844.5 4790335 4790335 TS=30
LESSSAGGAYWASRPPSRSTATRSPMVMASSMSWVTKTMVLPSSDWIRSSSSCSPLRTT
GSTAEKGSSISITGGSAARARATPTRCCWPPESW

539_WP_097215169.1 Streptomyces sp. Ag82_G6-1 922656 922656 TS=30
MAPKSELASLAQEILELESETFEISDYSDESEVLLAGSTSCSSTSTTSSTTSTTSCSA

540_KOV59016.1 Streptomyces sp. AS58 36004 36004 TS=30
MADKSTLAALADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

541_AEM00619.1 Streptomyces sp. ATCC 55365 7407 7407 TS=30
MGNNEEYFIDVNDLSIDVFDVVEQGGAVTALTADHGMPEVGASTNCFCYICCSCSSN

542_WP_073863645.1 Streptomyces sp. CB01249 720373 720373 TS=30
MEKNLSTLADEILELEAETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCTA

543_WP_073771703.1 Streptomyces sp. CB02366 111651 111651 TS=30
MAPKTELTTLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

544_WP_073931449.1 Streptomyces sp. CB02400 78983 78983 TS=30
MSTESGTATSSTAPSARPCAVSRASPPSSRTSPRSSTGSSAWSGSCSSVCGPREPCRTRRT
PWPSSPPSRRRPAPSSSTV

545_OKJ76123.1 Streptomyces sp. CB02460 145965 145965 TS=30
MEKNLSTLADEILELEAETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCTA

546_WP_067398575.1 Streptomyces sp. F-3 8965 8965 TS=30
MNKDLSTLADEILELEAETFEISDYSDASEVVLAGCTSTSSTSTSSSTCSTTSCSA

547_WP_053673516.1 Streptomyces sp. H036 59225 59225 TS=30
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

548_WP_053686323.1 Streptomyces sp. IGB124 14779 14779 TS=30
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

549_WP_129766430.1 Streptomyces sp. L-9-10 233439 233439 TS=30
MADKSPLASLADEIMELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

550_WP_018548801.1 Streptomyces sp. LaPpAH-108 345529 345529
TS=30
MADNNLASLAQEILDLESETFEITDYSDASEVMLGSSTSCSSTSTCSSTTSTTSCTA

551_WP_063345887.1 Streptomyces sp. MJM8645 100940 100940 TS=30
MSDHFEDVLALDDLDLGELTVTALRDTVALPETGASGAPSSSSCGSSSCAAPHPPTQTF

552_WP_109886853.1 Streptomyces sp. NEAU-S7GS2 824486 824486
TS=30
MEDTEFDFNLEDIPSDVFELADRGLTVESLTSGHGLVENGASSPSCGSSCSSLP

553_WP_051841832.1 Streptomyces sp. NRRL F-5193 103012 103012
TS=30
MNKDLSTLADEILELEAETFEISDYSDASEVVLAGSTSTSSTSTCSSTTSTTSCSA

554_WP_030822034.1 Streptomyces sp. NRRL S-104 23504 23504 TS=30
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

555_WP_030197398.1 Streptomyces sp. NRRL S-87 68957 68957 TS=30
MSRTPNPQDPLELQDLDLDADLDLTDLTVTSLRDTAALPENGASWGSCSCQASSSCAQP
HVTTPEVL

556_AID54696.1 Streptomyces sp. NRRL WC-3908 3807 3807 TS=30
MDANLVLDLDDLSVDQLDILPTAPGSTLESINVGHAMVEIGASNCTSKGSPASCCSCCCC

557_WP_109378529.1 Streptomyces sp. NWU339 69135 69135 TS=30
MNKDLSTLADEILELEAETFEISDYSDASEVVLAGCTSTSSTSTSSSTCSTTSCSA

558_SCK09537.1 Streptomyces sp. SceaMP-e96 736171 736171 TS=30
MEDTEFDFNLEDIPSDVFELADRGLTVESLTSGHGLVENGASSPSCGSSCSSLP

559_WP_161271514.1 Streptomyces sp. SID10115 34892 34892 TS=30
MSIDTMNSGDVEQIFEGFDIDELEVLEVSQGVALPEMGASGGNSGTSSSSSTCSSCTC

560_WP_164601129.1 Streptomyces sp. SID13031 131471 131471 TS=30
MTDLTAELLSLETATFEVEEIADVGMDAASWCSSSTSTSSCSSCSTSSCCGCSSCSCSSTS

561_WP_164616055.1 Streptomyces sp. SID14515 41819 41819 TS=30
MAPKTELATLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

562_WP_093589027.1 Streptomyces sp. SID4917 9688 9688 TS=30
MADKSTLASLADEIMELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

563_WP_093636282.1 Streptomyces sp. SID4951 736171 736171 TS=30
MEDTEFDFNLEDIPSDVFELADRGLTVESLTSGHGLVENGASSPSCGSSCSSLP

564_WP_103507787.1 Streptomyces sp. SM13 9436 9436 TS=30
MAPKTELATLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

565_WP_051265163.1 Streptomyces sp. TAA204 244623 244623 TS=30
MAEKSSLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

566_WP_028435937.1 Streptomyces sp. TAA486 167189 167189 TS=30
MAEKSSLDSLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

567_OKJ11516.1 Streptomyces sp. TSRI0261 217579 217579 TS=30
MFSAGPEGDRVPAAVGNHTHRRPPGRWARAWEYTEMSKIMETVDVDAMFDSFNIEEL
ETMEIADGVALPEMGASYLNAPLCCSSSSSCC

568_WP_167933484.1 Streptomyces sp. ventii 2135 2135 TS=30
MTQKPELATLADEILELESETFEISDYSDTAEVVLAGSTSCSSTSTCSSTTSTTSCSA

569_WP_148587905.1 Streptomyces sp. WAC01526 37221 37221 TS=30
MEDNEFDFNLEDIPSDVFELADRGLTVESLTSGHGLVENGASSPSCGSSCSSLP

570_WP_125815636.1 Streptomyces sp. WAC07149 207544 207544 TS=30
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

571_WP_045862234.1 Streptomyces sp. WMMB 714 304917 304917 TS=30
MTEKSSLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

572_WP_053706163.1 Streptomyces sp. XY413 14680 14680 TS=30
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

573_WP_053627034.1 Streptomyces sp. XY593 88182 88182 TS=30
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

574_WP_053679289.1 Streptomyces sp. XY66 14707 14707 TS=30
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

575_WP_053760581.1 Streptomyces sp. ZFG47 1590886 1590886 TS=30
MADKSTLAALADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

576_WP_069922943.1 Streptomyces subrutilus 6493249 6493249 TS=30
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

577_APE26458.1 Streptomyces venezuelae 5537340 5537340 TS=30
LGEVQQAEQVEDEEDEVPIEPLEAPISGRATPWATSRVSSSSTSKFSNALSASWPCLMSSL
MVNSFPNSFPQPFGCNTNNADPSSAAPAFPSTT

578_APE26458.1 Streptomyces venezuelae 5566907 5566907 TS=30
MICSAPMSSTPWTSSASSSSPSPAPCWPSARTSTSSASPSSPRSPRWAAVSSVT

579_PWW57157.1 Streptomyces venezuelae 79972 79972 TS=30
MICSAPMSSTPWTSSASSSSPSPAPCWPSARTSTSSASPSSPRSPRWAAVSSVT

580_WP_150205637.1 Streptomyces venezuelae 336768 336768 TS=30
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

581_CCA58419.1 Streptomyces venezuelae ATCC 10712 5567985 5567985 TS=30
MICSAPMSSTPWTSSASSSSPSPAPCWPSARTSTSSASPSSPRSPRWAAVSSVT

582_WP_151728560.1 Thermogemmatispora aurantia 521336 521336 TS=30
MEDLKQELAHLTLDELQLDELNIQPLTADRDGLEALTLGHGMIEIGASVLPEGCCSCCIP
CCCCW

583_WP_112425992.1 Thermogemmatispora tikiterensis 406702 406702 TS=30
MEDLKQELAHLTLDELQLDELNIQPLTADRDGLEALTLGHGMIEIGASVLPEGCCSCCIP
CCCCW

584_WP_132943275.1 Tumebacillus sp. BK434 314807 314807 TS=30
VTAAVSGVRSACADTISSMRYCGSSCSVPFQSRSSSCCSSGGSSGS

585_WP_126711131.1 Verrucosispora sp. FIM060022 1450371 1450371 TS=30
MPTQPPTSGSLADSVRSELRDLQTETFEVEDIADLSADLMDICSSSTSTSSCSSCSTSSCCS
CTSSSCSSTS

586_WP_102658257.1 Verrucosispora sp. ts21 9066 9066 TS=30
MPTQPPTSASLADSVRSELRDLQTETFEVEDIADLSADLMDICSSSTSTSSCSSCSTSSCCS
CTSSSCSSTS

587_WP_108074041.1 Vitiosangium sp. GDMCC 1.1324 100184 100184 TS=30
MNNEIKMENLDLDSLELPKIQELEISDSLAMPEAGASVGDSNCCSCCCCCA

588_WP_148872464.1 [Eubacterium] rectale 14082 14082 TS=29
MEDKKVLDDVFEGIEIVELDETTTLSETAASSGISSCNSCSCCLSTSCCSIHFGA

589_GES11969.1 Acrocarpospora macrocephala 13418 13418 TS=29
VNGGETTLSIALMRQEIAALETATFEVEEVADLAAAYASAEVPIGDGGEFSGATSSSCSC
SSCCSSSTSSCCSCSCSTTSCAE

590_WP_132119980.1 Actinocrispum wychmicini 46786 46786 TS=29
MKELTFDVEDLRLDDIAVTSMRDAVALPESAASSSCNSCSCGSSSCCCCQQELPPQLA

591_WP_169735453.1 Actinokineospora inagensis DSM 44258 10441 10441 TS=29
MAPENTPQGLESLDFDFSALEISDFIDESRLEDGETISTVMAASCTSCECCCSTS

592_ABC02780.1 Actinomadura melliaura 11777 11777 TS=29
VSSNTLLALDDLDIEAIHVVEIPGIADSLAAGHSMPETGASACCSIVCSSCCCSC

593_WP_026413554.1 Actinomadura oligospora ATCC 43269 173983 173983 TS=29
MNGPEELDFDLRDVPMDVFDLAESGLTIESLTAGHGMPEHGASLPFCSCSATCSCCPSSS

594_OLO48685.1 Actinomyces oris 10210 10210 TS=29
MSDSTIIDGLEFQDLSDAFAGIEVLELDETMTLSETAASSGISSCNSCSCCLSTSCCSVHIG
A

595_WP_147680663.1 Actinomyces ruminicola 2150 2150 TS=29
MDVFELEDGSVDVEPLTAGHGMSEVGASTNCFCYPCCSCSAPSSSSSSSA

596_WP_044856426.1 Amycolatopsis orientalis 4923074 4923074 TS=29
MNLNLDDLNLDELDVTSLRDTVALPETAASSSNNGSGTNTGSCTDPA

597_WP_002151253.1 Bacillus cereus HuA4-10 37128 37128 TS=29
MLDNKESKLNLVDELDIEILELDDVESIPSSAASSGSGGSSTCGSSSCSSCSSSCA

598_WP_002151253.1 Bacillus cereus HuA4-10 37518 37518 TS=29
MRNNKDVNDNLLDELDIEILELDDVESIPSSAATSGSGGSSTCGSSSCSSCSSSCA

599_WP_142325983.1 Bacillus thuringiensis 47413 47413 TS=29
MRNNKDVNDNLLDELDIEILELDDVESIPSSAATSGSGGSSTCGSSSCSSCSSSCA

600_WP_142325983.1 Bacillus thuringiensis 47608 47608 TS=29
MLDNKESKLNLVDELDIEILELDDVESIPSSAASSGSGGSSTCGSSSCSSCSSSCA

601_WP_104129685.1 Cryobacterium sp. N21 28794 28794 TS=29
MADSIKGLPIDVFELMDEGLNVNSLTAGHGMDEVGASTNCFCYPCCSCSAPSSSA

602_PZA02665.1 Cutibacterium acnes 1906450 1906450 TS=29
MENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

603_KPG63735.1 Cutibacterium acnes 10289 10289 TS=29
VNRMENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

604_KPG65283.1 Cutibacterium acnes 1271 1271 TS=29
VNRMENETLDLDVMELADIIGSASDQDDMAQVMAASCTTTSVSTSSSSSS

605_WP_002518053.1 Cutibacterium acnes 1889 1889 TS=29
MDRNNLYHLSREKVNTNENEALNLNVMEFSDMISDATERDDVTQIMAASCVETSVSSS
STSSS

606_WP_115884642.1 Cutibacterium acnes 575121 575121 TS=29
MDRNNLYHLSREKVNTNENEALNLNVMEFSDMISDATERDDVTQIMAASCVETSVSSS
STSSS

607_AEH29193.1 Cutibacterium acnes 6609 945305 945305 TS=29
MDRNNLYHLSREKVNTNENEALNLNVMEFSDMISDATERDDVTQIMAASCVETSVSSS
STSSS

608_EFT77539.1 Cutibacterium acnes HL030PA1 583458 583458 TS=29
MDRNNLYHLSREKVNTNENEALNLNVMEFSDMISDATERDDVTQIMAASCVETSVSSS
STSSS

609_WP_148872464.1 Eubacterium rectale 14082 14082 TS=29
MEDKKVLDDVFEGIEIVELDETTTLSETAASSGISSCNSCSCCLSTSCCSIHFGA

610_ABD13543.1 Frankia casuarinae 5019309 5019309 TS=29
MPEPDDAPISGSATLSETSSTSSSSMSKPANSLSSMGFSLARALIVLSFIMDCVSL

611_WP_161449277.1 Glutamicibacter soli 133831 133831 TS=29
MNTTENPGLALPDLDFEGLEVMSVREAVAIPETGATSGSSSCTSTSCCGSSSCCSSGSCG

612_WP_054262645.1 Janthinobacterium sp. CG23_2 586196 586196 TS=29
LSSMLSKKSWRSSRRTSTRVAPACLRTFDSASCTMNSSCSCCSRGKGAPSTSATSSASMP
VWRRNLSTVASSAPIRLPSTRRLRKLCSSSRTSS

613_WP_058858331.1 Kocuria flava 1670162 1670162 TS=29
MDRTPHDIVDLPMDVFELEDQGMDVTSLTAGHGMTEVGASTNCFCYPCCSCSAPSSSA

614_WP_077018665.1 Kribbella sp. ALI-6-A1658851 1658851 TS=29
MALREELLVLETATFEIEDIIDLGQDAAGWCSSSCSCSSCSSCSTSSCCSTSTSTGCGGG

615_WP_077018665.1 Kribbella sp. ALI-6-A1659152 1659152 TS=29
MSTQSLGLELLDLEAVTFEVDEITDPSLDLAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

616_WP_132283874.1 Kribbella sp. VKM Ac-2568 58806 58806 TS=29
VTTMALREELLVLETATFEIEDIVDLGQDAAGWCSSSCSCSSCSSCSTSSCCSTSTSTGCG
GG

617_WP_132283874.1 Kribbella sp. VKM Ac-2568 59106 59106 TS=29
MSTQTLGLELLDLEAATFEVDEITDPSLDLAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

618_WP_025357078.1 Kutzneria albida DSM 43870 3981847 3981847 TS=29
MANVGSLRFDIADLPMDVFELADRGLTVESLTVGHGMAENGASDSSGFCSGFCGGGSS
SSSCSCGGGGFCGGGGSCSCGGSCSH

619_WP_091261732.1 Micromonospora chaiyaphumensis 175464 175464 TS=29
MPITLNDELRELETETFEIEEVEDGGEALAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

620_WP_111214312.1 Micromonospora craterilacus 9492 9492 TS=29
MLITLNDELRELETETFEIEDVEDGAEALAATSSSCSCSSCCSCSTSSCCSCSTSCSSCG

621_WP_091191776.1 Micromonospora narathiwatensis 939039 939039 TS=29
MPITLNDELRELETETFEIEEVEDGAEALAATSSSCSCSSCCSCSTSSCCSCSTSTSSCG

622_WP_151488655.1 Micromonospora sp. ALFpr18c 18831 18831 TS=29
MPITLNDELRELETETFEIEEVEDGGEALAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

623_WP_073835304.1 Micromonospora sp. CB01531 76394 76394 TS=29
MPITLNDELRELETETFEIEEVEDGGEALAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

624_WP_046564724.1 Micromonospora sp. HK10 73752 73752 TS=29
MPITLNDELRELETETFEIEEVEDGGEALAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

625_WP_148799196.1 Micromonospora sp. MP36 42859 42859 TS=29
MPITLNDELRELETETFEIEEVEDGGEALAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

626_WP_171023219.1 Micromonosporaceae bacterium KJ-029 73326 73326 TS=29
MPITLNDELRELETETFEIEEIEDGAEALAATSSSCSCSSCCSCSTSSCCSCSTSCSSCG

627_WP_067138951.1 Microtetraspora malaysiensis 273796 273796 TS=29
MTMALRDELLALETATFEIADIGDVGQDAAGWCSSSCSCSSCCSCSTSSCCSTSTSTGCG
GG

628_WP_067138951.1 Microtetraspora malaysiensis 274022 274022 TS=29
MSKTLGLELLELEAATFEVVDNIDPSMDMAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

629_WP_067138951.1 Microtetraspora malaysiensis 273819 273819 TS=29
LARRVQLWISRSWTCWCWSSSSWTCCTSSSWNSCTTSWWQPCPSTGRCCRPPRRWPPR
ARAARGRASCSSCAFLI

630_WP_042284370.1 Nocardiopsis alba 2300445 2300445 TS=29
MATKATAVIDLPELDFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSC
G

631_WP_041561753.1 Nocardiopsis alba ATCC BAA-2165 1721536 1721536 TS=29
MATKETAVIDLPELDFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

632_WP_036567597.1 Nocardiopsis alba DSM 43377 162972 162972 TS=29
MATKETAVIDLPELDFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

633_WP_017590583.1 Nocardiopsis ganjiahuensis DSM 45031 5569 5569 TS=29
MAIQESPTIELPDLNFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

634_WP_165977801.1 Nonomuraea sp. KC712 16774 16774 TS=29
MDVFDLADSGLTVESLTAGHGMAENGASWLCSCVCSSSASSS

635_WP_077720345.1 Novibacillus thermophilus 2607694 2607694 TS=29
MSDLKHSLNELEIEELDVSEMMDSDSLSEQEATQVMGASCTTCVCTCSCCTT

636_WP_111153483.1 Paenibacillus dendritiformis 362 362 TS=29
MDLKEEEIRIMLDDLELDQVEIRGLIVEDGKAIPDMGATIGAISCCSCCST

637_WP_163743638.1 Phytoactinopolyspora halotolerans 30486 30486 TS=29
MALHEELLDLETTTFEVEEITDLGQDAAGWCSSSCSCSSCSSCSTSSCCSTSTSTGCGGG

638_EGL44723.1 Propionibacterium sp. 434-HC2 10107 10107 TS=29
MMEFSDMISDATERDDVTQIMAASCVETSVSSSSTSSS

639_WP_125229216.1 Pseudopropionibacterium propionicum 4216 4216 TS=29
MSLVKEMQELEAETFEIQEIDVSNEIELAAVSCTSCSSCSCASTTSCSCSTSSCCGCTSCST
ASTS

640_WP_106620269.1 Saccharothrix carnea 30430 30430 TS=29
MDLTFDDSELDLGDLAVTAMRDAVALPETGASAAACSCSGTSCCCCQQPQLPTTLG

641_WP_106620217.1 Saccharothrix carnea 54473 54473 TS=29
MELSFEVDDLVVDDLSVTVMRDAVALPEAGASGGSCCGGSSCCCCPNPM

642_WP_106620217.1 Saccharothrix carnea 54659 54659 TS=29
MELNFEVDDLVLDDLSVTAMRDAVALPEAGASGGSCSGGSSCCCSVQPPDPTF

643_WP_015102784.1 Saccharothrix espanaensis DSM 44229 6068085
6068085 TS=29
MDNDHVLLDLSDLDVETLDFASNTAGLESLRMGHGLTEVGASCCCSTSACSCST

644_WP_170231938.1 Saccharothrix saharensis 1279763 1279763 TS=29
MDLTFDDSELDLGDLAVTAMRDAVALPETGASAAACSCSGTSCCCCQQPELPAQG

645_WP_123741114.1 Saccharothrix texasensis 99953 99953 TS=29
MNEFDLTFDAEDLVLDDLAVTTMRDTVALPESGASIGASCGSSCCCQPQLPQPTF

646_WP_080677888.1 Salinispora pacifica CNT003 71747 71747 TS=29
MNAKDRTAFDSAGVEFDVDDLPLSVFELTDGGLTVESLTAGHGLVENGASWPSCGSSC
SSTP

647_WP_101628591.1 Schaalia turicensis 9622 9622 TS=29
METLTLDEIADLEVETFEVIDLKELSEDLANCCSTTSCACSTTSCCSCSTSSTCSTSCSCGS
TSSCA

648_WP_114033190.1 Sphaerisporangium album 138817 138817 TS=29
MSKTLGLELMELEAATFEVVDNVDPSMDMAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

649_WP_114033190.1 Sphaerisporangium album 139043 139043 TS=29
MALSQELLALETATFEIADISDLGQDAAGWCSSSCSCSSCSSCSTSSCCSCSTSTGCGGG

650_WP_147133977.1 Stackebrandtia albiflava 992259 992259 TS=29
MESSLKNDLSALETETFEIQDTTELALDQPAWSSCSTSSCCGCSSCSCSSTSCSSTSSTSTT
SCSG

651_WP_142043611.1 Stackebrandtia endophytica 4635284 4635284 TS=29
MSELNDQIALLEAETFEVIDSTPSEDFAGSSSTSSSTSSTTCSCSTSSTCSTTSSSCSSTSSCG

652_WP_143935226.1 Streptococcus dysgalactiae subsp. equisimilis 126151 126151 TS=29
MDKQRDLDISFLDDFEFEVIELDDVSVLPETAASSGSSGDNVYSTCGSCSCCSSCCT

653_WP_111692612.1 Streptococcus equi subsp. zooepidemicus 1476993 1476993 TS=29
MDKQRDLDISFLDDFDFEVIELDDVSVLPETAASSGSSGDNVFSTCGSCSCCSSCCT

654_WP_061411505.1 Streptococcus gordonii 246488 246488 TS=29
MKHLDEEQVAALEEISQLDLSILDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

655_WP_040270454.1 Streptomonospora alba 128532 128532 TS=29
VRPRSTLPAASGARGSSPRSWPPSPPARCWTPRASATPECSAARTWPPSPPPTRSRSPTTG
TAGRCARRARGARALPSCSSSRCCPTSSTTAPPSTCTP

656_WP_103420827.1 Streptomyces bacillaris 46082 46082 TS=29
MSLPEMSTAFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TA

657_WP_142230511.1 Streptomyces cavourensis 7641547 7641547 TS=29
MSLPEMSTTFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TA

658_WP_121798699.1 Streptomyces griseocarneus 133441 133441 TS=29
MNNEDYFVDVSDLSIDVFEVVEQGGAVTALTADHGMPEVGSSTNCFCYICCSCSSN

659_WP_055552665.1 Streptomyces kanamyceticus 2636 2636 TS=29
MSIDTMNSGDVEQIFEGFDIDELEVLEVSQGVALPEMGASGGNSGTSSSSSTCSSCTC

660_WP_120752968.1 Streptomyces klenkii 353376 353376 TS=29
MSAASSSGHARLKAVSSSLFRRSDLPSWSGETSPSCSASCLSTRISSSSSQTCSRASLCGVP
GSAASNAAAQSSAVSRMSDQWVVTVIPSAGGSS

661_WP_097239702.1 Streptomyces sp. 1331.2 7559497 7559497 TS=29
MLALDGLDLGELTVTALRDTVALPETGASGAPSSSSCGSSSCATPYPPVYPY

662_WP_030086229.1 Streptomyces sp. B226SN101 35539 35539 TS=29
MAPKTELATLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

663_WP_127468904.1 Streptomyces sp. B27 82265 82265 TS=29
MSLPEMSTAFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TA

664_WP_114933936.1 Streptomyces sp. CAI-24 72925 72925 TS=29
MSLPEMSTTFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TA

665_WP_073863645.1 Streptomyces sp. CB01249 727012 727012 TS=29
MSGSPRSSLLGGSSSVRTGSGPGAGAVPCSSTGTLSCPSVAVSLPDAFSVT

666_WP_100591758.1 Streptomyces sp. CB01635 171674 171674 TS=29
VRSTRPRAASTGRRRPCSSSSETASSGVKASSAPRSSVSSPAAR

667_OKJ76123.1 Streptomyces sp. CB02460 139559 139559 TS=29
MSGSPGGSLRGGSSSRTGSGPGAGAVPCSSTGTLSCPSVAVSLPDALSVT

668_WP_093751374.1 Streptomyces sp. DvalAA-19 12772 12772 TS=29
MSLPEMSTTFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TA

669_WP_131544036.1 Streptomyces sp. IBTA2 289671 289671 TS=29
MSLPEMSTTFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TA

670_WP_138048765.1 Streptomyces sp. NEAU-C151 53167 53167 TS=29
VETPHSSASAGSSAQAGCWAAAAASTSNPPSAAPAKSPWPSGTWRAVGCSWASSARSS
SSAPSDSTTASTGAYSS

671_WP_172629511.1 Streptomyces sp. NL15-2K 57954 57954 TS=29
MGNNEEYFVEIDDLSIDVFDVVEQGGAVTALTADHGMPEVGASTNCFCYICCSCSSN

672_WP_164601129.1 Streptomyces sp. SID13031 131687 131687 TS=29
VITMALCEELLVLETATFEVEDISDLGQDALGWCSSSCACSSCTSCSTSGCCSSSTSTSSS

673_WP_164601129.1 Streptomyces sp. SID13031 130908 130908 TS=29
MSTQSLGLELLDLEAVTFEVDEITDPSLDLAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

674_OKJ88320.1 Streptomyces sp. TSRI0107 18161 18161 TS=29
MSDTSSTPGLDLADLDLGDLTVTSMRDTVALPEGGASNGASSCSCSSSSCAQPQLPVPL

675_WP_125823540.1 Streptomyces sp. W1SF4 6944096 6944096 TS=29
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

676_WP_125595288.1 Streptomyces sp. WAC07061 6957 6957 TS=29
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

677_WP_083378345.1 Varibaculum massiliense 891053 891053 TS=29
METLTLDEIADLEVETFEVIDLKELSEDLANCCSTTSCACSTTSCCSCSTSSTCSTSCSCGS
TSSCA

678_AEB45324.1 Verrucosispora maris AB-18-032 3678651 3678651 TS=29
MPITLNDELRELETETFEIEEVADGEEMLAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

679_WP_126711131.1 Verrucosispora sp. FIM060022 1449973 1449973 TS=29
VKGGGNMPITLNDELRELETETFEIEEVADGEEMLAATSSSCSCSSCCSCSTSSCCSTSTS
TSSCG

680_WP_102658257.1 Verrucosispora sp. ts21 9466 9466 TS=29
MPITLHDELRELETETFEIEEVADGEEMLAATSSSCSCSSCCSCSTSSCCSTSTSTSSCG

681_WP_132119980.1 Actinocrispum wychmicini 26586 26586 TS=28
VVPPPTSADTTQLPPSSVALSRIDHSPTPVACCSGNPIPSSTTSMCRSSSTHRRTAQDRAQ
A

682_WP_085945035.1 Actinosynnema mirum DSM 43827 4548757 4548757 TS=28
MGTAAQRSSPRRTTTRSGWTTGTPGSARDSSPQGTSSATTTAAGTPSGRWCREPRSPPS
ATSAWPTCTPTTCSPGTPRSWCTTRGARPGPTS

683_WP_037311645.1 Amycolatopsis orientalis DSM 40040 = KCTC 9412 45509 45509 TS=28
MNLNLDDLNLDELDVTSLRDTVALPETAASSGTNSGSCGNTGSCTDPV

684_WP_173027111.1 Arthrobacter sp. NEB 688 728099 728099 TS=28
VRCSAGSPGPGRSTASRPPSTSSAAASSACCARSSCRWSSSRSSSASPGCAR

685_WP_098223293.1 Bacillus thuringiensis 36167 36167 TS=28
MLDNKESKLNLVDELDIEILELDDVESIPSSAASSGSGGSSTCGSSSCSSCSSSCA

686_WP_098223293.1 Bacillus thuringiensis 36362 36362 TS=28
MRNNKDVNDNLLDELDIEILELDDVESIPSSAATSGSGGSSTCGSSSCSSCSSSCA

687_WP_043588221.1 Clavibacter cf. michiganensis LMG 26808 39204 39204 TS=28
MARTSVAVGTLGEWQSSPSNVVATDCQRGRTGRYPTSSTSSCCLSVASASCMQALCKL
RI

688_WP_147362287.1 Clavibacter michiganensis 21975 21975 TS=28
MARTSVAVGTLGEWQSSPSNVVATDCQRGRTGRYPTSSTSSCCLSVASASCMQALCKL
RI

689_WP_010073638.1 Clostridium cellulovorans 743B 4413289 4413289 TS=28
MNSNLLIEELELNVDSAFEGIDIIELDETMTLSETAASSGVSSCNSCSCCGSTSCCSIGTINI
T

690_WP_010073638.1 Clostridium cellulovorans 743B 4413536 4413536 TS=28
MNSNLLIEELELNVDSAFEGIDIIELDETMTLSETAASSGVSSCNSCSCCGSTSCCSIGTIKI
T

691_WP_010073638.1 Clostridium cellulovorans 743B 4413787 4413787 TS=28
MNADLLIEELEMNVDSAFEGIDIIELDETMTLSETAASSGVSSCNSCSCCGSTSCCSIKTISI
T

692_KFC17662.1 Cutibacterium acnes HL201PA1 1647 1647 TS=28
MDRNNLYHLSREKINTNENEALNLNVMEFSDMIGDATERDDVTQMMAASCVETSVSSS
STSSS

693_GAE70556.1 Cutibacterium acnes JCM 18909 9851 9851 TS=28
MDRNNLYHLSREKINTNENEALNLNVMEFSDMIGDATERDDVTQMMAASCVETSVSSS
STSSS

694_ESS85761.1 Cutibacterium acnes P6 9874 9874 TS=28
MDRNNLYHLSREKVNTNENEALNLNVMEFSDMISDATERDDVTQIMAASCVETSVSSS
STSSS

695_WP_010777475.1 Enterococcus faecalis 22558 22558 TS=28
MKAEELEVENQLEELDELFGEVELIEIDETITLSETAASSGISSCNSCSCCASTSCCSNSSISI
T

696_ALS37682.1 Enterococcus rotai 2405418 2405418 TS=28
MEVLNSQVELNDLDEMFGEVELIEIDETITLSETAASSGLSSCDSCSCCLSTSCCALT

697_WP_069647476.1 Enterococcus ureasiticus 63443 63443 TS=28
MEELDELFGEVELIEIDETITLSETAASSGISSCNSCSCCASTSCCSNSSISIT

698_KFB02771.1 Frankia sp. Allo2 3422 3422 TS=28
MSALASEKPMDDKLFAGFDIEELEVLEVSDSVALPEMGASSGSGILCCSSSSSSSSCC

699_WP_035858852.1 Kitasatospora cheerisanensis KCTC 2395 791461 791461 TS=28
MADTSLASLAEEILNLESETFEISDYSDTSEVMLGSSTSCSSTSTCSSTTSTTSCTA

700_WP_104819511.1 Kitasatospora sp. MMS16-BH015 7556958 7556958 TS=28
MIEPAPAFELEDLDLGDLTVTSMRDTIALPEGGASNGGSSTSCGSSSCAVPQLPIHQY

701_PLC13238.1 Kocuria flava 3281681 3281681 TS=28
MDRTPHDIVDLPMDVFELEDQGMDVTSLTAGHGMTEVGASTNCFCYPCCSCSAPSSSA

702_WP_089952438.1 Lechevalieria xinjiangensis 164750 164750 TS=28
MKDLSFDLDDLELGDLAVTTMRDSVALPESGASGAPSSCSCGSSSCSSCHQPQLPTLPA

703_WP_030478671.1 Lentzea albidocapillata 513449 513449 TS=28
MKDLSFDPDDLELGDLAVTSLRDSVALPESGASGGASSCSCGSSSCSSCTQPPQLPAQNA

704_WP_090064780.1 Lentzea flaviverrucosa 148747 148747 TS=28
MKDLSFDPDDLELGDLAVTSLRDSVALPESGASGGASSCSCGSSSCSSCTQPPVQNA

705_WP_090003615.1 Lentzea violacea 56214 56214 TS=28
MKDLSFDPDDLELGDLAVTSLRDSVALPESGASGGASSCSCGSSSCSSCTQPPVQTA

706_WP_091636923.1 Lysobacter sp. cf310 323849 323849 TS=28
MNSDFDFDSLDLASLNIEGMEVVTLKEAMALPETGASSVVSSSNSCSSCGSSSCANDNIA
E

707_WP_115858067.1 Lysobacter sp. zong2l5 1283977 1283977 TS=28
LPMRWLRRTMSAAARKPSPRGTTMSTTPMRCSSSTGYTRASSIIRSACTGVTRRVRKV

708_WP_121160063.1 Micromonospora pisi 7448986 7448986 TS=28
MPGEEWTSVELDVFDLDDLDLDDITVTSMRDAVALPETGMSSSGDAAAACSCCGSSSC
CPNIDTYQPY

709_WP_165438069.1 Micromonospora sp. CNZ295 4975208 4975208
TS=28
LPIGTDSPVSADSSASSREDSSSRRSAVTTSPPSSRTTSPTTRSSAGTRRRSPARRTVAVTC
PRVRSDSTDRAARNSVTKPIAPLSRSTTTIAAPSAT

710_WP_132401450.1 Micromonospora sp. KC207 12669 12669 TS=28
MATGNAPQTTAPDSLESMDFDFSSLEISDFIDESRLEDGETISNVMAASCTSCECCCSTS

711_WP_017535514.1 Nocardiopsis alba 179889 179889 TS=28
MSIETHLSNSLGDLPVEEFSTTDEGVLGFGANQPGGIESLGSTLAMMEIGASGGGGCCSC
CSCCPCCCCS

712_WP_017535514.1 Nocardiopsis alba 180360 180360 TS=28
MSIETGPLSGSLGDLPVEEFSTSSEGVVGFGVNQPGGIESLGSTLAMMEIGASGGGECCS
CCSCCPCCCCS

713_WP_017535514.1 Nocardiopsis alba 180801 180801 TS=28
MSIETHLSNSLGDLPVEEFSTTDEGVLGFGANQPGGIESLGSTLAMMEIGASGSGGGCCS
CCSCCPCCCCS

714_WP_161111330.1 Nocardiopsis alba 3775210 3775210 TS=28
MSIETHLSNSLGDLPVEEFSTTDEGVLGFGADQPGGIESLGSTLAMMEIGASGSGGGCCS
CCSCCPCCCCS

715_WP_161111330.1 Nocardiopsis alba 3775651 3775651 TS=28
MSIETGPLSGSLGDLPVEEFSTSSEGVVGFGVDQPGGIESLGSTLAMMEIGASGGGECCS
CCSCCPCCCCS

716_WP_161111330.1 Nocardiopsis alba 3776122 3776122 TS=28
MSIETHLSNSLGDLPVEEFSTTDEGVLGFGADQPGGIESLGSTLAMMEIGASGGGGCCSC
CSCCPCCCCS

717_WP_042284370.1 Nocardiopsis alba 2300103 2300103 TS=28
MPHSTAPAIDLPELDFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSGGSCG

718_AFR05803.1 Nocardiopsis alba ATCC BAA-2165 3209157 3209157 TS=28
MSIETHLSNSLGDLPVEEFSTTDEGVLGFGANQPGGIESLGSTLAMMEIGASGGGGCCSC
CSCCPCCCCS

719_WP_036567597.1 Nocardiopsis alba DSM 43377 163315 163315
TS=28
MPHSTAPAIDLPELDFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSGGSCG

720_WP_017590583.1 Nocardiopsis ganjiahuensis DSM 45031 5191 5191 TS=28
MPHSTAPEIDLPELNFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSGGSCG

721_WP_017545984.1 Nocardiopsis prasina DSM 43845 37115 37115 TS=28
MSIESGHLSASLGDLPVEEFSSSGEGVVGFGSNQPGGIESLGSTLAMMEIGASGGGECCS
CCSCCPCCCCS

722_WP_017545984.1 Nocardiopsis prasina DSM 43845 37518 37518 TS=28
MSMKADQISAGLGDLPVEEFSTSSEGVLGFGSDQPGGIESLGSTLAMMEIGASGSGGGC
CSCCSCCPCCCCS

723_WP_159944718.1 Nocardiopsis sp. FR26 46370 46370 TS=28
LALPASMSLRTFSGVGPCSSMVRTSSAPSWRRCQLVCSRVLSSSCTAESSS

724_WP_110049192.1 Nocardiopsis sp. L17-MgMaSL7 182587 182587
TS=28
MSMKADQISAGLGDLPVEEFSTSSEGVLGFGSDQPGGIESLGSTLAMMEIGASGSGGGC
CSCCSCCPCCCCS

725_WP_110049192.1 Nocardiopsis sp. L17-MgMaSL7 183124 183124
TS=28
MSIESGHLSASLGDLPVEEFSSSGEGVVGFGSNQPGGIESLGSTLAMMEIGASGGGECCS
CCSCCPCCCCS

726_WP_049570703.1 Nocardiopsis sp. SBT366 14147 14147 TS=28
MSIESGHLSASLGDLPVEEFSSSGEGVVGFGSNQPGGIESLGSTLAMMEIGASGGGECCS
CCSCCPCCCCS

727_WP_049570703.1 Nocardiopsis sp. SBT366 14550 14550 TS=28
MSMKADQISAGLGDLPVEEFSTSSEGVLGFGSDQPGGIESLGSTLAMMEIGASGSGGGC
CSCCSCCPCCCCS

728_WP_017580750.1 Nocardiopsis valliformis DSM 45023 83227 83227 TS=28
MSIDAKGLSSSLGDLPVEEFSASSEGVVGFGANQPGGIESLGSTLAMMEVGASGSGECCS
CCSCCPCCCCS

729_WP_017580750.1 Nocardiopsis valliformis DSM 45023 83756 83756 TS=28
MSIKADQLSAGLGDLPVEEFSTSSEGVLGFGTDSPGGIESLGSTLAMMEVGASGGGGCC
SCCSCCPCCCCS

730_WP_143782003.1 Ornithinimicrobium sp. H23M54 564335 564335 TS=28
MSINSLGRELKELEVETFEVIDIADASQDMDGATSTSSCACSTTSCCSCSTSSCCGCTSCS
CGSTSSCA

731_WP_125229216.1 Pseudopropionibacterium propionicum 3951 3951 TS=28
MEVETSEMTKELETLEAETFEIEEIPEQDFELAGLCSCSTSTSSTCSSSTSTSSTCGCSSCSC
SSTSS

732_WP_165968252.1 Saccharopolyspora sp. 7K50272327 72327 TS=28
LIGGSGRAPGRSSLTPGTSVGSPLGGASVRCSSSAVCCCAGQGRVA

733_WP_081911977.1 Sphingomonas wittichii 121655 121655 TS=28
MDANMGDSAEFGIESIDLTDLAKGLDLEGLQVISASNSLAMPEGGASCCGGSGYSCCCT
CCCP

734_WP_147133977.1 Stackebrandtia albiflava 992768 992768 TS=28
MNSLIKDIAELEAQTFEIVDVAELSQDEALTWSSCSTSSCCGTTSTSCGSTSSCA

735_WP_142043611.1 Stackebrandtia endophytica 4634342 4634342 TS=28
MESSLAKDLSALEMETFEIQDTAQLDTDQPAWSSCSTSSCCGTSSCSCGATSCSGSATSS
TSTTSCSG

736_WP_019777477.1 Streptococcus sobrinus 491401 491401 TS=28
MVSLTENEQNTLDELTQLDLTVLDLDESELVSAAAASSGSSSSCSISTCGSSSCCA

737_WP_002963036.1 Streptococcus sobrinus 510977 510977 TS=28
MVSLTENEQNTLDELTQLDLTVLDLDESELVSAAAASSGSSSSCSISTCGSSSCCA

738_WP_021673885.1 Streptococcus sobrinus W1703 11831 11831 TS=28
MVSLTENEQNTLDELTQLDLTVLDLDESELVSAAAASSGSSSSCSISTCGSSSCCA

739_WP_077353877.1 Streptococcus sp. NCTC 10233 1435817 1435817 TS=28
MKSLLNIDDSMIESLNLSDFEVIDSIELSEADEKAILGASCTTCTCCCSCCG

740_WP_061921828.1 Streptomyces bungoensis 42185 42185 TS=28
MGNNEEYFVDVNDLSIDVFDVVEQGGAVTALTADHGMPEVGASTNCFCYICCSCSSN

741_WP_043497397.1 Streptomyces glaucescens 163923 163923 TS=28
MSDTGVTPDLDLADLDLGDLTVTSMRDTVALPEGGASNGASSCSCSSSSCAQPQLPVPL

742_WP_093786373.1 Streptomyces guanduensis 170854 170854 TS=28
VTSAPSTRSSPSASRWSTTSSGARSTGTTTWTTSSRRPCCGSSGGWATCASPPRSAPGW
WRSPSARSATASRPAKPPGTSAPAWTRPNSSPTRPPTSPR

743_WP_014677376.1 Streptomyces hygroscopicus subsp. jinggangensis TL01 9769932 9769932 TS=28
MSPPEPASGSSPRPGVCCTRPSSTTPWPPQALTHTSGTRSTTGCS

744_WP_052713429.1 Streptomyces katrae 9723 9723 TS=28
MDNLTAEDFASFELEDLEILDVVDGVALPEMGASNLGSWLCCSSSSSGSSCC

745_WP_052713429.1 Streptomyces katrae 10006 10006 TS=28
MDKNLGTEDFASFELEEIEVLDVAEGAALPEVGASSGGWLCCSSSSSSSCC

746_WP_127149161.1 Streptomyces lydicus 8083624 8083624 TS=28
MEDTEFDFNLEDIPSDVFELADRGLTVESLTSGHGLVENGASSPSCGSSCSSLP

747_WP_125051547.1 Streptomyces rimosus subsp. paromomycinus 759211 759211 TS=28
VSTGTATAPALAMPRSTVTSSAVRGSSTATRSPGPTPCACRWCAYWLARASSSR

748_WP_121176943.1 Streptomyces sp. 1114.5 5179019 5179019 TS=28
MLALDGLDLGELTVTALRDTVALPETGASGAPSSSSCGSSSCATPYPPVYPY

749_WP_094213841.1 Streptomyces sp. 2R 21615 21615 TS=28
MSRPEMATSFDLSALDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TV

750_WP_123498760.1 Streptomyces sp. 844.5 4774758 4774758 TS=28
MSKNTTDSIATLNQEILELESETFEITDYADASEVLNGSTCSSTTSSCCAGAKK

751_WP_123502812.1 Streptomyces sp. 844.5 400664 400664 TS=28
MNDNNALASLTQEILELESDTFEITDYADAAEVLNGTCSSTCSSCCSKLN

752_WP_007268049.1 Streptomyces sp. C 7097257 7097257 TS=28
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

753_WP_099879114.1 Streptomyces sp. CNZ279 1461704 1461704 TS=28
VLAMTCGGSSISSISKVSASSASSSSTSMCQLPCQIRELRGHPPLW

754_WP_111002325.1 Streptomyces sp. NTH33 17188 17188 TS=28
LTGRASLVQLVPASAAAAPAGHGRSVQEPPYTKGGLEMPIELKDELLSLESATFEIEEMT
DPADVELAWSSSSCSTSSCCSCSTSSCCSCSTSTS

755_QCW77982.1 Streptomyces sp. S6 2932986 2932986 TS=28
MAPKTELTTLADEILELESETFEISDYSDAVEVVLAGSTSSSSTSTCSSTTSTTSCSA

756_WP_086160989.1 Streptomyces sp. SCSIO 03032 5057142 5057142 TS=28
MTSRDTPASFDLDGLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCAGCSGCSVQPPTSLQ
M

757_WP_093847691.1 Streptomyces sp. SID4934 8603 8603 TS=28
MTPKTELATLADEILELESETFEISDYSDAAEVVLAGSTSCSSTSTCSSTTSTTSCSA

758_WP_028417601.1 Streptomyces sp. SID8359 340752 340752 TS=28
VRSARTGRCPWTGPASSTPPVRPRSPRPSSAPTSRCSSPPAPSTAASATARTTSSTARCRSP
STTASTRCS

759_WP_093858544.1 Streptomyces sp. TLI_053 814009 814009 TS=28
MSDHLQDALAFDGLDLDLDLGELTVTALRDSIALPEGGASSNPASSSCQSSSNSSVIIHPP
ATRW

760_WP_093858544.1 Streptomyces sp. TLI_053 837836 837836 TS=28
VRTTSATTDRAPARPGTPSSCRASATTTTTSSAPSAGPPGCSRCWTGRTTARAATAPAST
PTAPPGRPA

761_WP_099897193.1 Streptomyces sp. TLI_171 413815 413815 TS=28
MSDHLGGAPAPTDLRFDDLDLDLDLGELTVTALRDTVALPEGGASGAPSSTSCGSSSCA
VPHPPTQPL

762_WP_151484583.1 Streptomyces sp. TRM68295 15504 15504 TS=28
MEKQIAGAVVSAVDSPSLDLGDLDVESLEFSTDSEGGLESLRMGHGLTEVGASCCCSTS
VCSCST

763_WP_079277171.1 Streptomyces sp. TSRI0261 189676 189676 TS=28
VCGRAWAGLSGSAPRPGPVVSCRCCWGGPSSPPSSVPGSPGSPRCWSSGTPTPGRRRGS
SSSPCSSGPI

764_WP_030652003.1 Streptomyces sp. WAC05950 14934 14934 TS=28
MEKSPLASLADEILELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

765_WP_030891366.1 Streptomyces varsoviensis 24450 24450 TS=28
VRRPGATRAICTSPGVCCAPTPSCAATACCRSSPPSSAR

766_WP_063836007.1 Streptomyces yeochonensis CN732 453587 453587 TS=28
MADISKLDAPLASLAQEILELESETFEITDYSDASEVMLGSSTSCSSTSTCSSTTSTTSCTA

767_WP_132943275.1 Tumebacillus sp. BK434 317138 317138 TS=28
MNRVVPVSTLSCCRESVSALESINWVVAPQKVWLLGRGRSVGNWSGSSSPPSCCFQYSS
CCCKAACCSGSFCQYTKSLYWIGSSWSGEGRPFFKVA

768_WP_132943275.1 Tumebacillus sp. BK434 319910 319910 TS=28
MCSSCPTLSNFARKSGPCAKLNGVRASSSATARASATLCTSHTGSAMVTSSVTNCFGCP
STILNVVRRDSCRSTTSAKHFPSTSASRFPLIQTTTGML

769_GES11969.1 Acrocarpospora macrocephala 12277 12277 TS=27
MINPRIELAALEIETFEIAEIADLDSAFASAEMPVDSSFSCSVPTWCSTSSSCSIVTSCGPGS

770_WP_054221078.1 Actinobacteria bacterium OV450 73473 73473 TS=27
VRCCCTNSPTWPTGTSPSPTPPSRCGGSSWPSCCCPTCSGRGTSSTGSSRTAGCRYWTGR
CCSPRSRSLWSTSPARTSCAAVRSTPT

771_WP_146884718.1 Deinococcus cellulosilyticus NBRC 106333 = KACC 11606 109305 109305 TS=27
VAMICPVFLAGQWAFTPRSRTFLTTTGSSSSKSSCSMTVR

772_WP_057236393.1 Kitasatospora sp. Root187 171612 171612 TS=27
MSTAVDFAAFDVNELAVLDANDAVALPDMGASVIIIFGYTEDGEEILTEDTLNNADAGS
GSLSTSSSCC

773_WP_057236393.1 Kitasatospora sp. Root187 196437 196437 TS=27
VSTLVRGPKASPAAARPAQRSPASVTRVASPRLCSWTTARNGRCSSSFSQLSCHSSESSA
WLGSSRRSETRTPASSPAAISAAHPRSAPAC

774_QBD82834.1 Ktedonosporobacter rubrisoli 10645921 10645921 TS=27
MSKEQIEHLLTLEAELSELEVETFEIEDIFDPGQVAAQGPVKTSGGSGGGGSGGCCSTVST
STSSCCSCTSCSH

775_WP_004896705.1 Lactobacillus johnsonii 10285 10285 TS=27
MKNQKQKKLENLNDLSDAFDDIEVIELDETMTLSETAASSGSSSCNSCSCCASTSCCSVH
IHL

776_WP_004896705.1 Lactobacillus johnsonii 36218 36218 TS=27
MKNQKQKKLENLNDLSDAFDDIEVIELDETMTLSETAASSGSSSCNSCSCCASTSCCSVH
IHL

777_WP_082267664.1 Lactobacillus sakei 18926 18926 TS=27
VRVVEELNKAQQESLETLNDLDLSLLDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

778_WP_109636348.1 Lechevalieria deserti 560338 560338 TS=27
MKDLSFDLDDLELGDLAVTSMRDSVALPESGASGGPSSCSCSSGSCSSCHQPQLPALTA

779_WP_167748232.1 Microbispora bryophytorum 311868 311868 TS=27
MSNESLKLDLRGLDVDSIDVLSADSIVAEGHGAVETGASSAAGYCSSYLSATSCWAPEA

780_WP_169947016.1 Microbispora sp. H11081 907547 907547 TS=27
MSNDSLKLDLRGLDVDSIDVLSADSIVAEGHGAVETGASSAAGYCSSYLSATSCWTPEA

781_WP_120688749.1 Micromonospora musae 88134 88134 TS=27
VLVPVPTFSRSSGGGSGSIPSVRAGCSSARSAPTGPTSSTPRWRCAGRSKGSSSTCRPGRG
NGSRSSAAPRLAHSPSCSTGCPLRRSAPNTPGSATSR

782_WP_091639110.1 Micromonospora pallida 611203 611203 TS=27
LRSAAGRALMVLVETCRCAPTSCANCACSVLVIRPAARAPCSSASSWDSSSRARLSSIRA
AVSWATCRSSTPPSSS

783_WP_117665680.1 Micromonospora sp. MW-13 4312 4312 TS=27
MENELLTLDIDDLEISEFLDESRLEDSDVVAKVMSASCTTCECSCSCSS

784_WP_157320809.1 Nesterenkonia alkaliphila 53342 53342 TS=27
MERTDNFDLPMDVFELEDQGMDVKSLTAGHGMTEVGASTNCFCYPCCSCSAPSSSA

785_WP_147287926.1 Nocardia pseudobrasiliensis 196295 196295 TS=27
MSAQGKDPNEIRRRFEELPMEVFQLDGSGLPIESLTDGHGMTEVGASCTSCVCICSCCT

786_WP_123201682.1 Nocardiopsaceae bacterium YIM 96095 51969 51969 TS=27
MPEKFNLELDDLQVDGLELPTGDGASLETMTTGQGLTEMGASDCGCWCCSCC

787_WP_041561753.1 Nocardiopsis alba ATCC BAA-2165 1736845 1736845 TS=27
VWAATTRGRSTWATTPATCTTRCTTPSTRERSAVFRSRRWSSGCWASSSCPSPSPSFSGS
SPSSRC

788_AFR06332.1 Nocardiopsis alba ATCC BAA-2165 2818836 2818836 TS=27
MNHEIDLSAIEISDLVHEVEQGEDAFSQVMAASCVGTACTCSSTSSST

789_WP_026125251.1 Nocardiopsis alba DSM 43377 45865 45865 TS=27
MNHEIDLSAIEISDLVHEVEQGEDTFSQVMAASCVGTACTCSSTSSST

790_WP_049566838.1 Nocardiopsis sp. SBT366 102996 102996 TS=27
MALQETPTIDLPDLNFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSSGSCG

791_ADO67771.1 Nocardiopsis sp. TFS65-07 26609 26609 TS=27
MNHEIDLSAIEISDLVHEVEQGEDTFSQVMAASCVGTACTCSSTSSST

792_WP_017607927.1 Nocardiopsis xinjiangensis YIM 90004 62791 62791 TS=27
VWTPGRSTTGCSTPTGSWWATRSTTGTPAPTGWPSGCSSTCRRRSCTRSTACRSSPSTPS
SNWWRPPGAPTWTRRATCCSCPT

793_ACS83784.1 Nonomuraea sp. Bp3714-39 4826 4826 TS=27
MDLSDLPMDVFELADDGVAVESLTAGHGMTEVGASCNCFCYICCSCSSA

794_WP_131104644.1 Ornithinimicrobium sp. HY008 34745 34745 TS=27
VRTSLARSITTRCPSSSRYPASPVCSHPCSSTAFAVSSGWRW

795_WP_125229216.1 Pseudopropionibacterium propionicum 4472 4472 TS=27
MSDDTVASNLQEIEAETFEIQDVSDADLLLAPSSCTSTASTSCCSTTSCALCSCMNPWK

796_WP_132489383.1 Saccharopolyspora sp. 7K50259770 59770 TS=27
VIELHNGPRSMAGMSPASSARSRSNSSARLPMSSQCSSTQSRHLASRSCAKPGWG

797_WP_168587984.1 Saccharopolyspora sp. ASAGF58 5489123 5489123 TS=27
LTIQLSDESDLGTDMEDLKLDLESLEVESLSFGGDADRDLQSLGMGHGMTEVGSSAGCC
CCTCSCCCCPCG

798_KOX21547.1 Saccharothrix sp. NRRL B-16348 45171 45171 TS=27
VTRWRCPRRVRPAAGPAVAAPAAAARTPASAETNRVREDIMELSFEVDDLVLDDLSVT
AMRDAVALPEAGASGGSCSGGSSCCCSVQPPEPTF

799_KOX21547.1 Saccharothrix sp. NRRL B-16348 44815 44815 TS=27
MPDIAFDARELHLDDLSVTAMRDAVALPEAGASGASCSCNSSCCCQQEPIPTFG

800_WP_142043611.1 Stackebrandtia endophytica 4635141 4635141 TS=27
LSSDMVFHSSQVRDYRARTPTAISRRKTSNCSSNSWSSRSTTSSRSRWSTRCSSWCSKNL
RSPHSAWSR

801_WP_034090612.1 Streptacidiphilus albus 49862 49862 TS=27
MSDSTANTGFDLQELDLGDLTVTSMRDTVALPENGASWGGCSCQGSSSCASPTPTPGPV
DL

802_WP_125367351.1 Streptococcus gordonii 251954 251954 TS=27
MKHLDKEQAAALEEISQLDLSILDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

803_WP_061601555.1 Streptococcus gordonii 252258 252258 TS=27
MKHLDKEQAAALEEISQLDLSILDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

804_WP_125375864.1 Streptococcus gordonii 252100 252100 TS=27
MKHLDKEQAAALEEISQLDLSILDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

805_WP_071519546.1 Streptococcus parauberis 415936 415936 TS=27
MKHLDKEQAAALEEISQLDLSILDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

806_KWT63928.1 Streptomyces albus subsp. albus 203921 203921 TS=27
VSSCPGPGGRRTRWCPGSGTPSSRAWPGGGTSCSSSPTPCCARKDRSSPRST

807_WP_044375373.1 Streptomyces badius 4496887 4496887 TS=27
VIRTCGQVSRIIRAQYPPGSSQDSSSVWSSMTCCGGTFIARSSCSICHSGCGVPGAERRISC
CSLPSGARNRSTAW

808_WP_067007690.1 Streptomyces cellostaticus 49392 49392 TS=27
MSRTPNQQDALELQDLALDTDLDLTDLTVTSLRDTAALPENGASWGSCSCQGSSSCAQ
PQVTTPVVL

809_WP_114056112.1 Streptomyces globosus 3876221 3876221 TS=27
MTDPTEGVLALDDLDLDLDLGELTVTALRDSVALPETGASGGVSSSSCGSSSCAQPRLP
ELPY

810_WP_070195834.1 Streptomyces oceani 38868 38868 TS=27
VMSPPPWTPCSRTWSTRVTGASWTACCASTNARWPPPSTPTPGTSSTTSRSIPSTQRASW
TDSPTTTPSSPSTPA

811_WP_121025185.1 Streptomyces sp. 3212.4 6074033 6074033 TS=27
MPDHREDLLAFEDLDLGELTVTALRDTVALPETGASGGASSCSCGSSSCAVPRLPDLPY

812_WP_121025185.1 Streptomyces sp. 3212.4 6074259 6074259 TS=27
MSRTPNQQAALELQDVALDMDLDLTDLTVTSLRDTAALPENGASWGSCSCQGSSSCAQ
PQVETPVVL

813_WP_116150976.1 Streptomyces sp. 3212.5 6077626 6077626 TS=27
MPDHREDLLAFEDLDLGELTVTALRDTVALPETGASGGASSCSCGSSSCAVPRLPDLPY

814_WP_116150976.1 Streptomyces sp. 3212.5 6077852 6077852 TS=27
MSRTPNQQAALELQDVALDMDLDLTDLTVTSLRDTAALPENGASWGSCSCQGSSSCAQ
PQVETPVVL

815_WP_168093022.1 Streptomyces sp. AA8 7248 7248 TS=27
MAQQAELATLAQEILELESETFEISDYSDASEVVLAGSTSSSSTSTCSSTTSTTSCSA

816_WP_073917951.1 Streptomyces sp. CB00455 89220 89220 TS=27
MAPSTGAPGFELEDLDLGDLTVTSMRDTVALPEGGASNGGSSCSCGSSSCAHPQLPDLP
AA

817_WP_073917951.1 Streptomyces sp. CB00455 89439 89439 TS=27
MSGTESTSAFSLQDLELDLSDLTVTSMRDTAALPEGGASWGSCSCQGSSSCAQPTDSMP
PAA

818_WP_089510236.1 Streptomyces sp. NBS 14/10 294273 294273 TS=27
MENTSLLDLEQLGDEMFEVVESDTTEDNGGAGSAWAYACCHSASSSCALS

819_ADZ45316.1 Streptomyces sp. NRRL 30471 3283 3283 TS=27
MEETEFDFDLEDIPSDVFELADRGLTVESLTSGHGLVENGASSPSCGSSCSSLP

820_WP_111002325.1 Streptomyces sp. NTH33 17387 17387 TS=27
VEQEQQLLVEQLLLDQASSTSAGSVISSISNVALSRESSSSFSSMGISNPPFV

821_WP_018566130.1 Streptomyces sp. PsTaAH-124 168411 168411
TS=27
VKSSPTPSSANGCWPGACSWSSTSWRRPAAPGDTCSSPVTTRSPGASPDWSNAATPTRS
CSATTCARRSSSRSTAARATTTSPATFSPPCAASG

822_KIF07648.1 Streptomyces sp. RSD-27 20670 20670 TS=27
MDQNLGTEDFASFELEEVEVLDVADGVALPEVGASSGGWLCCSSSSSSCC

823_KIF07648.1 Streptomyces sp. RSD-27 20954 20954 TS=27
VSRESSESERGIRMDMNLTAQDFASFELENLEVLDVVDGVALPDMGASNLGSWLCCSSS
SGSSCC

824_WP_084990321.1 Streptomyces sp. S8 51634 51634 TS=27
MSLPEMSTAFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLE
TA

825_WP_159396242.1 Streptomyces sp. Sge12 74713 74713 TS=27
MNDITRSSGFDLDDLDLGELTVTSMRDTVALPEGGASNGGSSCSCGSSSCAHPQLPDLP
V

826_WP_099219969.1 Streptomyces sp. SID8350 23794 23794 TS=27
MSLPEMSTAFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLE
TA

827_WP_018487806.1 Streptomyces sp. SID8356 761224 761224 TS=27
MSLPEMSTAFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLE
TA

828_WP_069922943.1 Streptomyces subrutilus 6467301 6467301 TS=27
VVNGSASGSGRSSSISSPDSRGDTSLTVRMTSGRASATIAAIRSSGWSSSSGT

829_WP_013130812.1 Thermobispora bispora DSM 43833 610003 610003 TS=27
MDVFELADSGVAVESLTAGHGMTEVGASCNCFCYICCSCSSA

830_WP_108074040.1 Vitiosangium sp. GDMCC 1.1324 85769 85769 TS=27
LSARSFVRVKMSAVSSSCANNVSSSADFSSGFTTYSRCSARSAGVPLRVTSTRTGR

831_WP_155357449.1 Acrocarpospora macrocephala 16260 16260 TS=26
MSKIRNELAALESEVFEIDEIPDLMEALASAEMPIASSTDGLSCTSTNSCTSSASTNSCCSG

832_WP_124934044.1 Actinomyces bowdenii 27770 27770 TS=26
VSAPRRAASASSSRPTSTSWATAPCPCTTTSRSPSSCTRR

833_WP_130475459.1 Amycolatopsis suaedae 304303 304303 TS=26
MTSPLPAIGSDTQQLEPELWLISRVGSWDRTTSTMRELSYSPQPSLNGTHCTIDVMPLSW
STIRRSSASKFTRLSAEVSSTASSDGMSCHTSSPSRSAQ

834_WP_091243560.1 Aquimonas voraii 104928 104928 TS=26
MDNNNAIGFDDIDVESMNIDGLEVVSLKDALALPETGASSGISSCNSCSTCGSTSTCGSC
NGGDGGSGGNGGNGGSPPVTIQP

835_WP_002151253.1 Bacillus cereus HuA4-10 37323 37323 TS=26
MGGRKXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXSGSGGSSTCGSSSCSSC
SSSCA

836_PZA02665.1 Cutibacterium acnes 1906061 1906061 TS=26
MMEFSDMIGDATERDDVTQMMAASCVETSVSSSSTSSS

837_KPG63735.1 Cutibacterium acnes 9900 9900 TS=26
MDRNNLYHLSREKINTNENEALNLNVMEFSDMIGDATERDDVTQMMAASCVETSVSSS
STSSS

838_KPG65283.1 Cutibacterium acnes 1660 1660 TS=26
MDRNNLYHLSREKINTNENEALNLNVMEFSDMIGDATERDDVTQMMAASCVETSVSSS
STSSS

839_WP_016622713.1 Enterococcus faecalis 104720 104720 TS=26
MKAEELEVENQLEELDELFGEVELIEIDETITLSETAASSGISSCNSCSCCASTSCCSNSSISI
T

840_WP_033255127.1 Kitasatospora phosalacinea 48089 48089 TS=26
MSTAVDFAAFDVNELAVLDANDAVALPDMGASVIIIFGYTEDGEEILTEDSLNSEAAGS
GSLSTSSSCC

841_WP_104819511.1 Kitasatospora sp. MMS16-BH015 7557172 7557172 TS=26
MSENTNGAEFELEELDLSDISVSSMRDSSALPEGGASWGSCSCQGSSSCQPVQQPPTLPV

842_WP_043718787.1 Kutzneria sp. 744 4494430 4494430 TS=26
MEGWSPTGANPCSAPLTGCPCPRSCYPSSMCAGQRRRSTTRTTWSRRCTSPTPAASTWP
GRPTAATWSC

843_WP_082267664.1 Lactobacillus sakei 18733 18733 TS=26
LHELLPLLAAAEETKSDSSRSKSDKSKSFNVSKDSCCALFNSSTTLTSNY

844_WP_090064780.1 Lentzea flaviverrucosa 130060 130060 TS=26
VPGVSASCCHRVGVTASPPSSTSAGLCSAPSSPSATSCCA

845_WP_135208110.1 Massilia sp. MC02 14057 14057 TS=26
MTDKTKAVDFDNIDVASLDIEGLEVVTLKDAMALPETGASSGISSCNSCSSCGSSSCVA
MDSAQVAE

846_WP_056374162.1 Microbacterium sp. Leaf161 692820 692820 TS=26
VTGCTMAWAVRSSARSKSRRASSASVDCSASPSKSTSSAGAAIRLSVNRSSTPTCTRSCC
AAVTPVSRSSSISRTSSRTRSATPGC

847_SCL44848.1 Micromonospora citrea 448174 448174 TS=26
MAASAGRGRAARAGCRRTTTSCTACGSSMAATRWATTSCGASSAPASPRRPPWPY

848_WP_145817302.1 Micromonospora sagamiensis2643153 2643153 TS=26
VPRRSSSSSKRSSSRTTTGLVCPRTDQGLRSGASSQWSTSSSRRGNGSTSSIMGRRPSSGP
PRASSGSRSRRCRRSSSSDRSGSSVAIGVMSSSSS

849_WP_145817300.1 Micromonospora sagamiensis2623456 2623456 TS=26
MDRTAFDNAGAEFDVDDLPLSVFELADGGMTVESLTAGHGLVENGASWPSCGSSCSSL
P

850_WP_049566838.1 Nocardiopsis sp. SBT366 102666 102666 TS=26
MPHSTAPEIDLPELNFDDMEVMSVREAVAVPETGATSGSSSCTSTSCCGSSSCCSGGSCG

851_AQZ61457.1 Nonomuraea sp. ATCC 55076 1833806 1833806 TS=26
MSAENLKLDLRGLEIDSIDVVSADSIVAEGHGTIETGASSSPGYCSSYLTAMSCWAPEL

852_WP_168006649.1 Nonomuraea sp. FMUSA5-5 110856 110856 TS=26
MSAENLKLDLRGLEIDSIDVVSADSIVAEGHGTIETGASSSPGYCSSYLTAMSCWAPEL

853_WP_165974660.1 Nonomuraea sp. KC310 8543 8543 TS=26
VPSLPTGTRSPSSPTWPPAYACRRTATGASSTACTAAPGWTTATSPCPSSSTRSSTASAPP
TTPSSRSPWTWGSRA

854_WP_084779544.1 Planobispora rosea 800510 800510 TS=26
MNLNDLPMDVFEMADSGMEVESLTAGHGMPEVGASCNCVCGFCCSCSPSA

855_PZS18042.1 Pseudonocardiales bacterium 12032 12032 TS=26
MIDQAIDLGFDLGDLDLTAVSVTVMTDAAALPETGASSGSSSCDSHSTCGSSSCCILL

856_WP_158847633.1 Saccharothrix deserti 82727 82727 TS=26
MYELSFEVEDLVLDDLAVTAMRDGVALPETGASNSCSCGSCCCCQQDPGLPPAMG

857_KOX21547.1 Saccharothrix sp. NRRL B-16348 44984 44984 TS=26
VTRSRYRRRARPVRRARATARAAANKNPSPPSAEWEFVVELSFEVDDLVVDDLSVTAM
RDAVALPEAGASGGGSCCGSSCCCCPDPSFR

858_KOX21547.1 Saccharothrix sp. NRRL B-16348 24216 24216 TS=26
VTGTQSPPSIASRTPSGLRPPACSRCATNSSSKAGAEFHTVTPCSSARSHQCCGSLSRPAS
GTTTAPPAIVMPRNS

859_WP_080678157.1 Salinispora pacifica CNR114 16618 16618 TS=26
MNAKDHTAFDSAGVEFDVDDLPLSVFELTDGGLTVESLTAGHGLVENGASWPSCGSSC
SSTP

860_WP_050956578.1 Staphylococcus aureus 26219 26219 TS=26
MSEFNTDVLDNIDLDSLELASEEISEKEEKEIMGASCTTCVCTCSCCTT

861_WP_034090612.1 Streptacidiphilus albus 49561 49561 TS=26
MSDSTANTGFDLQDLDLSDLAVTSMRDTVALPENAASAGASSCQASSSCSVYAVPERE
VTFDRNQQ

862_WP_034090612.1 Streptacidiphilus albus 53454 53454 TS=26
VTRASTCGRSAVTTPSRTTTSTTPASTTPATARGSTSARPTATGASTAAACRTPPTTT

863_WP_086573078.1 Streptomyces alboverticillatus 20236 20236 TS=26
MSDTTGALGFELEDLDLGELTVTSMRDTIALPEGGASNGGSSCSCGSSSCATPQLPHLPQ

864_WP_067007690.1 Streptomyces cellostaticus 58128 58128 TS=26
MVEPNGHGGVTDATALKVSAGTRSGSPTTRTSKAAPASSGRTTAPRGSGSTRLSPKSAP
EPTRTSCFRSCTPSGSASERSPPASTSRWRPPSPWPST

865_GFE19186.1 Streptomyces glebosus 8044601 8044601 TS=26
MEDTEFDFDLADIPSDVFELADRGLTVESLTSGHGLVENGTSSPSCGSSCSSLP

866_WP_067306578.1 Streptomyces griseochromogenes 4886556 4886556 TS=26
MSRTPNQQDGLDLQDLALDADLDLSDLTVTSLRDTAALPENGASWGSCSCQGSSSCAQ
PQVDTPVVL

867_ACS50127.1 Streptomyces hygroscopicus 25950 25950 TS=26
VLDEPARPDIEQPPAPSGMTARKSGVCSARPAACSNPSRNDCGNTSATRRKSSASNSRCS
SSDSNP

868_WP_159402072.1 Streptomyces katrae 1264596 1264596 TS=26
MTAGNHTPDFALDDLDLDLGDLTVTAMRDTVALPEGGASWGSCSCQGSSSCAQPAQPT
PV

869_GFE29173.1 Streptomyces libani subsp. rufus 1575119 1575119 TS=26
MEDTEFDFDLADIPSDVFELADRGLTVESLTSGHGLVENGTSSPSCGSSCSSLP

870_WP_066952401.1 Streptomyces lushanensis 8147 8147 TS=26
MADKSILASLADEIMELESETFEISDYSDASEVVLAGSTSCSSTSTCSSTTSTTSCSA

871_WP_127149161.1 Streptomyces lydicus 8053065 8053065 TS=26
VNATSSSPQSLSWATASVRPSAAAVGSAGSSQVATLCAARQRVRAVSSAMSSPGRSGSS
ASSTALQVSSSERAGSLSAPRSSSRPSSTSRRRTARSPSV

872_WP_070200355.1 Streptomyces nanshensis 69317 69317 TS=26
MSLPEMSTAFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TA

873_WP_070200355.1 Streptomyces nanshensis 70436 70436 TS=26
MSLPEMSTAFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLD
TA

874_WP_085923094.1 Streptomyces platensis 492510 492510 TS=26
MEDTDFDFDLADIPSDVFELADRGLTVESLTSGHGLVENGTSSPSCGSSCSSLP

875_WP_165273764.1 Streptomyces sp. 196(2019) 47660 47660 TS=26
MSLPEMSTAFDLSTLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPLPLE
TA

876_WP_119097970.1 Streptomyces sp. 3211.1 178123 178123 TS=26
MSENSATTGYNDLDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQVE
TPVV

877_WP_121014721.1 Streptomyces sp. 3211.6 1437650 1437650 TS=26
MTAGNHTPDFALDDLDLDLGDLTVTAMRDTVALPEGGASWGSCSCQGSSSCAQPAQPT
PV

878_WP_123502812.1 Streptomyces sp. 844.5 435417 435417 TS=26
MPSGRNDRSPRLRACSSAIRSSPSCGLSAPVLASSPRSSLSAASATAAPSPP

879_WP_079424625.1 Streptomyces sp. Ag109_G2-6 143267 143267 TS=26
MTAGNHTPDFALDDLDLDLGDLTVTAMRDTVALPEGGASWGSCSCQGSSSCAQPAQPT
PV

880_WP_159030320.1 Streptomyces sp. CB01201 18426 18426 TS=26
MNDNSNGSPSLAAAVADLALDLDLDSLAGDEHSQSLTAGHGMTEVSASFCCAPPPNCC
NCSCS

881_WP_074003108.1 Streptomyces sp. CB02056 185101 185101 TS=26
MSDTTTSRTAATEGLDLQDLDLSELTVTSLRDTVALPENGASWGSCSCQGSSSCAQPQL
PDVPTA

882_WP_109783537.1 Streptomyces sp. CG 926 106665 106665 TS=26
MNDATQSAGFDLDDLDLGELTVTSMRDTVALPEGGASNGGSSCSCGSSSCAHPQLPELP
L

883_WP_109783537.1 Streptomyces sp. CG 926 106870 106870 TS=26
MSDTAHTPATELHDLELDLGDLTVTSMRDTAALPEGGASWGSCSCQGSSSCAQPQDLT
ALEV

884_WP_147972293.1 Streptomyces sp. col6 132914 132914 TS=26
MEDLGDLEFNLDDIPADVFELADSGLTVESLTSGHGLAENGASAPSCGSSCSSLP

885_WP_129766430.1 Streptomyces sp. L-9-10 198474 198474 TS=26
VSTRTVRPALRTTHDCNAITVLSLSRLQWSGASHCSLSFHTSALTSGRNSEVGRNGSCHS
TTRMTSTSPSTMLCTRSSSFTMCSTLRTVNDPLNAHPAQ

886_WP_167151757.1 Streptomyces sp. MBT27 594430 594430 TS=26
MNEYSNGSPSLADAVADLALDLDLDSLAGDEHSQSLTAGHGMTEVSASFCCAPPPNCC
NCSCS

887_WP_030231853.1 Streptomyces sp. NRRL S-350 537771 537771 TS=26
MSRSAAAETGLDLQDLDLSELTVTSLRDTVALPENGASWGSCSCQGSSSCAQPQTMVTP
IA

888_WP_030231853.1 Streptomyces sp. NRRL S-350 538032 538032 TS=26
MPDPARGELALGEILALDGLDGLDLGELTVTALRDTVALPETGASGVPGSTSCGSSSCA
VPHPPTQSF

889_WP_168477216.1 Streptomyces sp. RPA4-5 762538 762538 TS=26
MEDTEFDFDLADIPSDVFELADRGLTVESLTSGHGLVENGTSSPSCGSSCSSLP

890_WP_144384198.1 Streptomyces sp. SAJ15 118013 118013 TS=26
MSDTSGETLDFELQDLDLGELSVTSMRDTVALPEGGASNGGSSCSCGSSSCATPQLPHLP
Y

891_WP_093858544.1 Streptomyces sp. TLI_053 813766 813766 TS=26
MSDMSSDFGLDLQDLDLSELTVTALRDTVALPENGASHGSCSCQASSSCVQPNLSATELI

892_OKJ88320.1 Streptomyces sp. TSRI0107 14832 14832 TS=26
MVALNLGASDSDRLSHVRSPCAVYCPWQKKYMMSSSRSCRARSSSSSLRYHRSVRSVR
SSSSACTSAAPATDIRSSSALRLLTGSVRWETYQ

893_WP_125650645.1 Streptomyces sp. WAC06614 8636 8636 TS=26
MANTTAVADFDQLFTSFDIEELEVLDVADGVALPEMGASSGNALCCSSSSSSSSCC

894_WP_125600994.1 Streptomyces sp. WAC07061 10731 10731 TS=26
MANTTAVADFDQLFTSFDIEELEVLDVADGVALPEMGASSGNALCCSSSSSSSSCC

895_GES11969.1 Acrocarpospora macrocephala 12485 12485 TS=25
MIKMSAELAALESETFEIAEIADVDSAFASAEVPVFDAAPCSCMCSTGCTHWTSTSTSCS
STS

896_GES11969.1 Acrocarpospora macrocephala 27048 27048 TS=25
MRVDVPTGEVIRVDVPVDGVIARDVGTGGVARVGPRGGPSSPRSSRASASVRDTSLSRY
AWYDSYSSRARINSSSSSSKSCQTGDNGGSDGSSWLDA

897_WP_054221078.1 Actinobacteria bacterium OV450 41202 41202 TS=25
MSEFSSTTGSDLDLDLDLDLSDLTVTALSDTAALPENGGSWGSCSCQGSSSCAQPQVET
PVV

898_WP_156214266.1 Actinomadura sp. NEAU-AAG5 359391 359391 TS=25
MNAPEELDFDLQDVPMDVFDLAESGLTIESLTAGHGMPEHGASLPFCSCSATCSCCS

899_WP_091295323.1 Amycolatopsis xylanica 91349 91349 TS=25
MSDFELDSLDLGELTVTSLRDTVALPETGASWGSCSCQGSSSCSQQPVEQ

900_WP_119908502.1 Carnobacterium divergens 2418348 2418348 TS=25
VRISLRSSVENVCSCSSNSTSNSSCTLSCPPTVTKALIL

901_WP_135025206.1 Carnobacterium divergens 59010 59010 TS=25
VRISLRSSVENVCSCSSNSTSNSSCTLSCPPTVTKALIL

902_WP_135081547.1 Carnobacterium divergens 59007 59007 TS=25
VRISLRSSVENVCSCSSNSTSNSSCTLSCPPTVTKALIL

903_WP_066361445.1 Herbidospora mongoliensis 155826 155826 TS=25
MHPRLPAPAVNARHPVVAAMPAAVATIVLALSETQSVTAPNHPPSSSSAGFGSSFGCGL
GAGVFCCSFFLCSWMFCSRVSSLFLTDCF

904_WP_051969186.1 Kitasatospora azatica KCTC 9699 209218 209218
TS=25
MSDATGTTDNPNFELQDLDLSDLTVTSMRDTAALPEGGASWGSSSCQGSSSCAQPQVP
VVA

905_AUG81639.1 Kitasatospora sp. MMS16-BH015 8236899 8236899 TS=25
MSNAVDFAAFDVNELAVLDAKDAVALPEMGASVLFIFNYTEDGEEILTEDTLNGSGSAS
TSSSCC

906_AUG81639.1 Kitasatospora sp. MMS16-BH015 8237132 8237132 TS=25
MSNAVDFAAFDVNELAVLDAKDAVALPEMGASVLFIFNYTEDGEEILTEDTLNGSGSAS
TSSSCC

907_AUG81639.1 Kitasatospora sp. MMS16-BH015 8237377 8237377 TS=25
MSNAVDFAAFDVNELAVLDAKDAVALPEMGASVLFIFAYTEDGEEVLTEDSLNASGSA
STSSSCC

908_WP_117485553.1 Kitasatospora xanthocidica 608259 608259 TS=25
MSDTTTSRTAATEGLDLQDLDLSELTVTSLRDTVALPENGASWGSCSCQGSSSCAQPQL
PDVPTA

909_WP_156858452.1 Klebsiella pneumoniae subsp. pneumoniae 20582 20582 TS=25
MNLNDLPMDVFEMADSGMEVESLTAGHGMPEVGASCNCVCGFCCSCSPSA

910_PLC13238.1 Kocuria flava 3271519 3271519 TS=25
VLPGRAGDPGVEVLRAPGGRLPLATTKPPAGTVSVTASRHACHSSRPSSGPCSTKRNCR
PVEVSSTVTFSRVSPRIGTTTVGTRSASSSRR

911_WP_132283874.1 Kribbella sp. VKM Ac-2568 39885 39885 TS=25
MVRSIAGAASYRPVGSAPKCLTDKDFAGCFSGWATSSKSATSSTYHATGMFSGTWNRT
LSPALYVLPSSTRCSS

912_WP_043718787.1 Kutzneria sp. 744 4494309 4494309 TS=25
MPLSWMSDSRLSTLANTCSARRVNECSSCDDASRWLHGYSSASHCARLRPGMCCITRK
SCSHIAR

913_WP_043718787.1 Kutzneria sp. 744 4493899 4493899 TS=25
MTGASSSLICIRATSWSTSTTPSRWWTSSSPSMSASGGAGRRSGRPGSAPRPTARVSRST
TTRSRRCGCGCSCR

914_WP_090064780.1 Lentzea flaviverrucosa 129109 129109 TS=25
LRSAAWSNRSTTVRSPKSLAGLSSGVTTEPPTAPSRPSARVSSSSTRRYPSNNRCANAGS
VTSGRYSSLTVSSSPGVTTHVNG

915_WP_172686362.1 Macrococcus caseolyticus 13329 13329 TS=25
MSEFQTNNIEGLDVTDLEFISEEVTEKDEKEIMGASCTTCVCTCSCCTT

916_AXA92875.1 Massilia sp. YMA4 4223094 4223094 TS=25
MRITLSLHLPVLNLKSILKETHMQNVDTTVDLDIDFDFEILETTEATAMEEMGASSTICA
GSTCSKQAALLARGC

917_WP_169982732.1 Microbispora sp. H10836 280451 280451 TS=25
MSNDSLKLDLRGLDVDSIDVLSADSIVAEGHGAVETGASSAAGYCSSYLSATSCWTPEA

918_WP_165521730.1 Micromonospora zingiberis 441311 441311 TS=25
VPAASSARASRPRSCCAPSGSSPTVPRCSHRRRPPRWSAGSSPSCGRRRTAGSGASTCSPT
GSARCSCSSRTGCPTTTSPTGCSSHR

919_WP_160167435.1 Nocardiopsis chromatogenes YIM 90109 44510 44510 TS=25
MNNRLDLDVQDLPMDVFDLGSTGLELETLTTGHGMQEHAASTGSMPCSFNTVCCCAC
CSQTCSTPAPYAFENGSDELE

920_WP_017607927.1 Nocardiopsis xinjiangensis YIM 90004 92550 92550 TS=25
MSEESTDFNLDDIPADVFELSDKGLTVESLTSGHGLVENGASSPSCGCSCSSLP

921_SEL93819.1 Nonomuraea pusilla 101434 101434 TS=25
MESKLDLSDLPMDVFEMADSGMEVESLTAGHGMPEVGASCNCVCGFCCSCSPSA

922_WP_166427735.1 Nonomuraea sp. 6K102 963 963 TS=25
MSENLKLDLRGLDVDSIDVVSADSIVAEGHGTVETGASSSPGYCSSYLTAMSCWAPEV

923_WP_168006649.1 Nonomuraea sp. FMUSA5-5 135463 135463 TS=25
LKGVTCPLPPCCPHPRRAPSSRDSAGPSRSSPWPSSSTRSTSTSCSSRSRRSAPTWASPARP
SNSS

924_WP_111178107.1 Nonomuraea sp. KC333 40836 40836 TS=25
MDLNDLPMDVFEMADSGMEVESLTAGHGMPEVGASCNCVCGFCCSCSPSA

925_WP_165977389.1 Nonomuraea sp. KC712 49587 49587 TS=25
MSENLKLDLRGLDVDSIDVVSADSIVAEGHGTIETGASSSPGYCSSYLTAMSCWAPEL

926_ACS83769.1 Nonomuraea sp. WU8817 9635 9635 TS=25
MSELESKLNLSDLPMDVFEMADSGMEVESLTAGHGMPEVGASCNCVCGFCCSCSPSA

927_WP_141786322.1 Ornithinicoccus hortensis 5843 5843 TS=25
MENNSIELPDLDFDGIEVMSVREAVAVPETGATSGSSSCSSTSCCGSSSCCSSGSCG

928_WP_111156021.1 Paenibacillus dendritiformis 11847 11847 TS=25
VKDLELLLEDLDVHLFQLKIEGATSKDEKALPEMGASGCWSCCCPLNT

929_KOX21547.1 Saccharothrix sp. NRRL B-16348 57953 57953 TS=25
MTSAGSSRGGRSSAPPTTTSPPTSPWCPAGTASRSSTAASPT

930_WP_147133977.1 Stackebrandtia albiflava 979460 979460 TS=25
MFTLATRLGDAMLAKLAPRTEASAGCPTESWYDYRCTGARLYRRRCTTSSNCTTSCGI
WSYISQCP

931_WP_096560225.1 Staphylococcus felis 7985 7985 TS=25
MSEFNTDILDNVDLDSLELASEEISEKEEKEIMGASCTTCVCTCSCCTT

932_WP_125444147.1 Streptococcus gordonii 7537 7537 TS=25
MKHLDKEQAAALEEISQLDLSILDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

933_WP_125366556.1 Streptococcus gordonii 6847 6847 TS=25
MKHLDKEQAAALEEISQLDLSILDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

934_WP_040274418.1 Streptomonospora alba 144175 144175 TS=25
MAGSKRNGKVSEQLDFDLSSVPVDVFDLADQGLTLETLTEGHGLPENGASSICNELMCA
CSSPAC

935_WP_067007690.1 Streptomyces cellostaticus 29073 29073 TS=25
VIVVLLAVDAMTPTACCTGSSRSTWPTPVSICAVPPADASPWMS

936_WP_114056112.1 Streptomyces globosus 3875968 3875968 TS=25
MPESIAGPAPEGLDLDLDLDLSDLAITALRDTPALPESGASYGSCSCQGSSSCAQPHTDTP
IG

937_WP_014677376.1 Streptomyces hygroscopicus subsp. jinggangensis TL01
9800790 9800790 TS=25
VRNARASPDTRAMVASSAIPTSRRGPPAIPTPDCSPAARSSSCLTTATSSLTGPTP

938_WP_037634310.1 Streptomyces katrae 22546 22546 TS=25
MSGASATTGYNHLDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQV
ETPVV

939_WP_100661070.1 Streptomyces lavendulae subsp. lavendulae 7025552
7025552 TS=25
MSDTAPDAAFDLQDLELDLGDLSVTSMRDTAALPEGGASWGSCSCQGSSSCAQPQTDT
PTA

940_WP_159487034.1 Streptomyces libani subsp. libani 3801736 3801736 TS=25
VLPGVAPAVVLPGPAIRGARTAQMSAAGSGSSILACSQSTSWMSSSPSRNSFHACGLVH
SC

941_WP_129298148.1 Streptomyces lydicus 53128 53128 TS=25
MRSSTMSATTSSRTSATAKPCSSWTTPASSRKGPVRQVSSVSTPGLPDAQRTARSECSSP
MQPAAAGH

942_QBP39279.1 Streptomyces sp. 24439 24439 TS=25
VPPTRTTCCGSSASAPSRSTWSARSRRSTTRRVCRSTTSTSRSSSGRCCAV

943_WP_097239702.1 Streptomyces sp. 1331.2 7559217 7559217 TS=25
MSRTAATTGLDLQDLDLSELTVTSLRDTVALPENGASWGSCSCQASSSCAQPQVMVDTI
G

944_WP_159395493.1 Streptomyces sp. 3211 61247 61247 TS=25
MNDITRSSGFDLDDLDLGELTVTSMRDTVALPEGGASNGGSSCSCGSSSCAHPQLPDLP
V

945_WP_133924344.1 Streptomyces sp. BK141 56239 56239 TS=25
MSRTSNQQDALELQDLALDTDLDLTDLTVTSLRDTAALPENGASWGSCSCQGSSSCAQ
PQVPTPVGL

946_WP_133924344.1 Streptomyces sp. BK141 56482 56482 TS=25
MSDQREDLLALEDLDLGDLDLGELTVTALRDTVALPETGASGGASSCSCGSSSCAQPRL
PDLPY

947_PJN01608.1 Streptomyces sp. CB01201 1718 1718 TS=25
VVLIVVPTTRRLPSNPNRTRRARVRSQCGRPWSSAIRKSTSTCRASPLRSAQNASSTCPTS
SAYTSPSSSASGIRSPGPTPLYRRSPSSQLNSSASSS

948_WP_173260350.1 Streptomyces sp. CWH03 27479 27479 TS=25
VTCSATACSSPSTSGWSRPSTTAISSSTPTRTRPSRTPSAAGCSSCPAPRGPTTTPSCCHRA
AASTPGPPSRSRSTRRSATPSASSQG

949_WP_133895671.1 Streptomyces sp. KS 21 553032 553032 TS=25
MSEASATTGYNHLDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQVE
TPVV

950_WP_089510236.1 Streptomyces sp. NBS 14/10 302416 302416 TS=25
LPSAVELSSVRAHGGTGSGTTETTRTTRTTRTTRTTRRKHARWNC

951_WP_018566130.1 Streptomyces sp. PsTaAH-124 186733 186733 TS=25
MNEFSNGTPSLADAVADLALDLDLGSLAEDEHSQSLTAGHGMTEVSASFCCAPPPNCC
NCSCS

952_WP_093642064.1 Streptomyces sp. SID4951 333868 333868 TS=25
MENTLLDLEQLGDEMFEVVGSESAQGEGDANSAWAYACCQSASSSCALS

953_WP_093858544.1 Streptomyces sp. TLI_053 825826 825826 TS=25
MAILDLQTMELPETEAAPIDDTLASTSSLSVLNCGTSTVSTLICL

954_WP_099897193.1 Streptomyces sp. TLI_171 414087 414087 TS=25
MSDLSRSVETETVLDLQDLDLSELTVTSLRDTVALPENGASWGSCSCQGSSSCAQPQQL

955_WP_037895024.1 Streptomyces sp. Tu 6176 186197 186197 TS=25
MNEFSNGTPSLADAVADLALDLDLGSLAEDEHSQSLTAGHGMTEVSASFCCAPPPNCC
NCSCS

956_WP_148588519.1 Streptomyces sp. WAC0152641416 41416 TS=25
MENTLLDLEQLGDEMFEVVGSESAQGEGDANSAWAYACCQSASSSCALS

957_WP_116427025.1 Streptomyces spongiicola 74347 74347 TS=25
MAVVATTTSGASSTSSLVAVSNSPSSARGTMRRAGAWSTVAPRRCSSAASSSARRAAV
TPTVKPVSGATAGAPCCSWYMPTEPGKGA

958_WP_167536884.1 Streptomyces subrutilus 7061483 7061483 TS=25
MIRATSATRFELDELDLDLGDLTVTSMSDTVALPEGGASWGTCSCQGSSSCAQPDQTGV
PVP

959_WP_110669416.1 Streptomyces tateyamensis 32947 32947 TS=25
MAGTTRTSGTTGSTGTESFELQDLDLSGLTVSSLRESAALPEGGASWGSSSCQVSSSCSM
PPVPESA

960_WP_110669416.1 Streptomyces tateyamensis 33171 33171 TS=25
MSDTADRPHFELEDLDLSDLTVSSLRDSAALPEGGASWGSSSCQGSSSCAVQQHLLPA

961_WP_170198917.1 Thermopolyspora flexuosa 5011691 5011691 TS=25
MAENLFDLDGLKVDSIEVLSADSIVAEGHGLIETGASSYSPYCSSYMQACSCWAPEQ

962_GES11969.1 Acrocarpospora macrocephala 12926 12926 TS=24
MNRTKAELAALETETFEIEVIADMDSTLASVEMPIDPGYSSCTYCSTSCSCWTSTSTSCS

963_ADH93701.1 Actinokineospora fastidiosa 2038 2038 TS=24
MSADLSALNIDSLEISEFLDDSRLEDSEVVAKVMSASCTTCECSCSCSS

964_WP_026413431.1 Actinomadura oligospora ATCC 43269 433432 433432 TS=24
VHPLPQPGLATGASSSTSTTTAASASSRHPCPHCDGHGRREAPST

965_WP_156214266.1 Actinomadura sp. NEAU-AAG5 330731 330731 TS=24
VCPSRMFCLASPIATVGPAAKIRACSSARSSSLSGATTSLTTPMRSASSASMIRPV

966_WP_126890039.1 Curtobacterium sp. HSID17257 98924 98924 TS=24
MATDSNFDFDLNDLAAEVFDLSDEGLTVESLTAGHGLVETGASFPSCGSSCSSLP

967_WP_146884718.1 Deinococcus cellulosilyticus NBRC 106333 = KACC 11606 109573 109573 TS=24
LPCMAPPPETASWTPCGKAPVRKKIMTTSGWPCDACGQTCKRFSRRWTTRCPSMAVCT
ACPASCNFNSTCTPSSRPWSSLLQST

968_WP_083529318.1 Kocuria flava 5085 5085 TS=24
MDRTPHDIVDLPMDVFELEDQGMDVTSLTAGHGMTEVGASTNCFCYPCCSCSAPSSSA

969_WP_134385526.1 Massilia plicata 3211847 3211847 TS=24
MQNVDTAVDFDFEILETNEATAMEEMGASSTICAGSTCSKQDSSLVAAC

970_WP_091191776.1 Micromonospora narathiwatensis 970150 970150 TS=24
LNRAMSAPCTRSVWARNASPVPTSTTRYVCSSCRVASSLRRYQTLITASTTFSPTMGHTC
TAAL

971_WP_123200200.1 Nocardiopsaceae bacterium YIM 96095 9314 9314 TS=24
MDKNEPAFEVDDLPVDVFSLSGESLTVESLTSGHGTPENGASCFEDDPTGSLSCTS

972_WP_143782003.1 Ornithinimicrobium sp. H23M54 563686 563686 TS=24
VGIASSCFLRSLWCGQVRHNCWSRHNCRSSWTCCHSWSSTSSSNLTWSTSSRPPCPQQP
P

973_WP_111156021.1 Paenibacillus dendritiformis 38 38 TS=24
LHLDDLELDDVNIIALTQDDAKGLPDMGATMMWVGCCSCCST

974_PZS18042.1 Pseudonocardiales bacterium 9726 9726 TS=24
MDELRLDDISVIRMHDAIALPETGASPASSMASSWACHSCGSCCCQE

975_WP_160487251.1 Rathayibacter iranicus NCPPB 2253 = VKM Ac-1602 12656 12656 TS=24
MNEKESFQKIDAQVFDLSDDGLTVESLTAGHGLVETGASWPSCGCSCSSLP

976_WP_073888311.1 Saccharothrix sp. CB00851 85649 85649 TS=24
LPVRRASASAQCSRCLAVTTSPRAASRTPSSYWARPSSGHACTSSSSCTASLSRSTSWPS
ASTTARAHNARPTVRDSPHSTAASYS

977_KOX21547.1 Saccharothrix sp. NRRL B-16348 22864 22864 TS=24
VSCWPPTRWARWWLGPAAPSPSGRPRTSHRRAGATTSSHRCRSARNSCGTSRSSCPTPP
STTSATRCRGTRTSTSACSSGC

978_KOX21547.1 Saccharothrix sp. NRRL B-16348 58916 58916 TS=24
VAPTGTSTSCALTERSTDTVGTAPAGTTTACTSASARVGSGGTTRATGTASRWTPTTTST
ACPRTAN

979_KOX21547.1 Saccharothrix sp. NRRL B-16348 61313 61313 TS=24
VLPAMYMLWVPRPSKTLGSTSTGYTRFGSSSHAWSSTTRNVPPLSASSYRTPCEVAGPE
VSVTGAESTM

980_WP_082403981.1 Saccharothrix sp. NRRL B-16348 7090 7090 TS=24
VACTPRRTPPTPPTRRPCGGSSPGWTGRCAGSCTPRCTWTTRRSPTWTTSASTRCSARSS
AARSCSTS

981_WP_170184969.1 Saccharothrix texasensis 1659168 1659168 TS=24
VEGSASRPVTSTVVTPGKSAVDTTWRNPARSATAVSRPAGNLGSSTVYAAPALSTASTA
VTRVASRSSSTPTTSPGTTPASTSRCARRFAAASSSA

982_WP_170184969.1 Saccharothrix texasensis 1659585 1659585 TS=24
VVTACSVSRPAMTAGSVTWPGCSTHAAPCRSGPNSSSANASHETGEHCNHTSSAVNST

983_WP_153468430.1 Streptomyces kaniharaensis 350213 350213 TS=24
MNVDLDTTLDLADLDLGDLTVTAMRDTVALPESGASHSGSSCSCGSSSCGVPRMPMD
M

984_WP_159402072.1 Streptomyces katrae 1264835 1264835 TS=24
MRLSPPSRETHVSDTARTPEFGFDDLDLDLGDLTVTAMRDTVALPEGGASWGSCSCQG
SSSCNKPPVMDSLLD

985_WP_121176943.1 Streptomyces sp. 1114.5 5178739 5178739 TS=24
MSRTAATTGLDLQDLDLSELTVTSLRDTVALPENGASWGSCSCQASSSCAQPQVMVDTI
G

986_WP_123510766.1 Streptomyces sp. 2132.2 554524 554524 TS=24
MSENAATPGYDDHDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQV
ETPTV

987_WP_166682199.1 Streptomyces sp. 25 86288 86288 TS=24
MSENSATTGYNDLDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQVE
TPVV

988_WP_121014721.1 Streptomyces sp. 3211.6 1437889 1437889 TS=24
MRLSPPSRETHVSDTARTPEFGFDDLDLDLGDLTVTAMRDTVALPEGGASWGSCSCQG
SSSCNKPPVMDSLLD

989_WP_099888076.1 Streptomyces sp. 61 592821 592821 TS=24
MSENAATPGYDDHDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQV
ETPTV

990_WP_136227307.1 Streptomyces sp. A1547 11388 11388 TS=24
MSEASATTGYNHLDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQVE
TPVV

991_WP_079424625.1 Streptomyces sp. Ag109_G2-6 143028 143028
TS=24
MRLSPPSRETHVSDTARTPEFGFDDLDLDLGDLTVTAMRDTVALPEGGASWGSCSCQG
SSSCNKPPVMDSLLD

992_WP_132761556.1 Streptomyces sp. BK038 28494 28494 TS=24
MSENAATPGYDDHDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQV
ETPTV

993_WP_138053671.1 Streptomyces sp. ICN19 3375966 3375966 TS=24
MRTSRSVWTRPRRTTAASSASSTCSAAARCCRSRSSRSASCRTSRRASSCNC

994_WP_063345887.1 Streptomyces sp. MJM8645 101157 101157 TS=24
MSASTSTPGLDLQDLDLSELTVTALRDTVALPENGGSWGSCSCQASSSCAQPNLPGADS
F

995_WP_138048765.1 Streptomyces sp. NEAU-C151 26795 26795 TS=24
MELEFDLDDIPADVFDLADSGLTVESLTAGHGLVENGASFPSCGSSCSSLP

996_WP_030721924.1 Streptomyces sp. NRRL F-2580 22002 22002 TS=24
MTDRTQSSGFDLDELDLGELTVTSMRDTVALPEGGASNGGSSCSCGSSSCAHPQMPELP
V

997_WP_030873335.1 Streptomyces sp. NRRL F-2747 1646 1646 TS=24
MSETSATPRSNHLDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQVE
TPTA

998_WP_064530944.1 Streptomyces sp. SAT1 223091 223091 TS=24
MNEFSNGTPSLADAVADLALDLDLGSLAEDEHSQSLTAGHGMTEVSASFCCAPPPNCC
NCSCS

999_WP_093739484.1 Streptomyces sp. SID4948 20001 20001 TS=24
MDEFKQSSPSLAAAVADLALELDLASLAGDEQSQSLTAGHGMTEASASFCCAPPPNCCN
CSCS

1000_OKJ88320.1 Streptomyces sp. TSRI0107 17921 17921 TS=24
MSDLARTALDLQDLDLDLSDLTVTAMRDTAALPEGGASWGSCSCQGSSSCAQPQPQLE
TGVVDAG

1001_WP_053684588.1 Streptomyces sp. WM4235 17505 17505 TS=24
MDRATPGAGFALDDLDLDLGDLTVTSMRDTVALPEGGASWGSCSCQGSSSCAQPERPE
VPTL

1002_WP_053684588.1 Streptomyces sp. WM4235 17933 17933 TS=24
MSDIARTPEFALQDLDLDLGDLTVTSMRDTVALPEGGASWGSCSCQGSSSCAQPTLDAG
ELAG

1003_WP_053694256.1 Streptomyces sp. WM6372 24875 24875 TS=24
MSENSATTGYNDLDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQVE
TPVV

1004_WP_053790560.1 Streptomyces sp. XY332 11335 11335 TS=24
MSEASATTGYNHLDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQVE
TPVV

1005_WP_150529706.1 Streptomyces vinaceus 7059287 7059287 TS=24
MSENAATPGYDDHDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQV
ETPTV

1006_PYR75691.1 Acidobacteria bacterium 17892 17892 TS=23
MGSAVWHVVCGLSARCVLGVSHSPASIRRLSVTRQGEGNAMKYENLRIDLSDIELEDIQ
VLAQEGAQGMPEFAASTSGCCGNCTGTSTSAGDELDDAE

1007_OLB64233.1 Actinobacteria bacterium 13_2_20CM_2_72_6 4250 4250 TS=23
MIREGIAMAVDKALANALEDLTLEPLGVEGDALEDSLTAGYSMPEVGASDILPPLLCCS
CYCC

1008_WP_133205963.1 Arthrobacter sp. JH1-1 42751 42751 TS=23
LTESSEEYRFVSWWAGGASSPVALSRTLRNCATRSSSSTRGTCPVGSSRGFGSRSHCLKR

1009_WP_066415890.1 Bacillus cohnii 1461934 1461934 TS=23
MIDEQQPSFPMDAPISSITVASSTSNISNSTSSMFCSIFIFHSSFNSLLLH

1010_WP_066415890.1 Bacillus cohnii 1462124 1462124 TS=23
MDEQQPSFPIDAPISSITVASSTSKISNSTSSMLCSTFIFYSSFDSL

1011_WP_112225817.1 Lechevalieria atacamensis 1261940 1261940 TS=23
MKDLSFDLDDLELGDLAVTSMRDSVALPESGASGGPSSCSCSSGSCSSCHQPQLPALTA

1012_WP_017625649.1 Nocardiopsis chromatogenes YIM 90109 6749 6749 TS=23
MVGLVPRRRAPRPERSPWTAMSCRSSSSPSPSAASPRSTASPSTCPTAASSA

1013_EGL44723.1 Propionibacterium sp. 434-HC2 19792 19792 TS=23
MTRLSSDGSISCSLSSWRRTSYCSQEMSSGFTRERSATSSH

1014_WP_104263765.1 Rathayibacter iranicus12625 12625 TS=23
MNEKESFQKIDAQVFDLSDDGLTVESLTAGHGLVETGASWPSCGCSCSSLP

1015_WP_106620269.1 Saccharothrix carnea 19875 19875 TS=23
VSSVVSGTTASSMNASLRAHSSRQCACACGSAPTRWIESTHHCSAARPSSGSVNISLAPR
TYSRSHVARSASSSASSPR

1016_WP_170184969.1 Saccharothrix texasensis 1662741 1662741 TS=23
MARRTVANTSGCSSSTWSNRLAEASKTVTPCSTTSAATTAGSTSRPGSSTDVPPCSSGPS
SWVAKASHETDAICSHTSPGPNPA

1017_WP_043497397.1 Streptomyces glaucescens 163690 163690 TS=23
MSDLARTALDLQDLDLDLSDLTVTAMRDTAALPEGGASWGSCSCQGSSSCAQPQPQLD
TGVVDAG

1018_BAO57436.1 Streptomyces lactacystinaeus 3149 3149 TS=23
MSDITASRVESLDLQDLDLSELTVTSLRDTVALPENGASWGSCSCQASSSCAQPQDM

1019_WP_127152319.1 Streptomyces lydicus 3418973 3418973 TS=23
LPAGTGRCRLSPRPTPPSSSPPPSSRWSSWRPSSSPRSSSSPRSSSSPRSSSSPRSSWRPSSSP
RSSSWPAPSWPPSSPAPWRASPRAARSPARR

1020_WP_158721406.1 Streptomyces sp. NRRL S-241 28456 28456 TS=23
MNDITRSSGFDLDDLDLGELTVTSMRDTVALPEGGASNGGSSCSCGSSSCAHPQLPDLP
V

1021_WP_159396242.1 Streptomyces sp. Sge12 74922 74922 TS=23
MSDAARTPGVELHDLELDLGDLTVTSMRDTAALPEGGASWGSCSCQGSSSCAQPHDAV
ALQA

1022_WP_053790560.1 Streptomyces sp. XY332 28943 28943 TS=23
MGLPVSSTDSSRKVRSGAVIAGLTRGTCEWSCVKGGVSSVFGIVTLRRASFAASSSIRCS
SDSDSFAPSARATSMVSGLASGSTPKCSSATFSAAAES

1023_WP_167536884.1 Streptomyces subrutilus 7061751 7061751 TS=23
MSDRDSARTPEFALQDLSLDLGDLSVTSMSDTVALPEGGASWGTCSCQGSSSCAHPQQ
NVGELAG

1024_WP_093941962.1 Actinoalloteichus hoggarensis 3524021 3524021 TS=22
MVPTNTPVSVPWRDVGASPARSTASQEVSSSSRCCGSIADASRGDTSKKSASKPLAPSRK
PPCLV

1025_WP_026413431.1 Actinomadura oligospora ATCC 43269 414887 414887 TS=22
VSESASAAPASWAALSSGGRSSGSTAPTCHSSRTGALLTRSTTCSKGLSRTTAPARTGSA
GTAAGSAGTGGGSARCSRTLWVTSSTTGPSSAS

1026_WP_151570259.1 Actinomadura rudentiformis 41637 41637 TS=22
LTAPASSVATTEWSSSFQASPSVCTLSGLPFISSWWRPL

1027_ATE55167.1 Actinosynnema pretiosum 4524172 4524172 TS=22
MDVLSHSSEAGITRVLTGLAAEFGDRLPESVVTGTVLQARRDLQGQIAPESLEELLHRLA
HYRLRELPDAGATTPA

1028_WP_066415890.1 Bacillus cohnii 1461795 1461795 TS=22
MHNYLIMIDEQQPSFPMDAPISSITVASSTSKISNSTSSTFVS

1029_PZA02665.1 Cutibacterium acnes 1918159 1918159 TS=22
VLEVPAAGQDVNSSWPASAISTNSSVWRQRSQESSPAGCCS

1030_WP_088248048.1 Deinococcus indicus 115677 115677 TS=22
MNEPTTTAIPEALELNDLHLEDVSVAVEEDAYDLPEAGASLSPRFSCSIVIDLS

1031_WP_099749137.1 Deinococcus sp. UR1 75067 75067 TS=22
MNEQTKAAIPEALELNDLHLEDVSVAVEEDAYDLPEAGASVSPRFSCSIIIDLS

1032_WP_061289420.1 Herbidospora cretacea 457091 457091 TS=22
MEKLPLDLSGLEVEAIDVLSADSIVMEGHGAVETGASSAMGYCSSYMQANSCWAPEA

1033_WP_083975908.1 Kitasatospora azatica KCTC 9699 503155 503155 TS=22
MPSARRTPTSPPTWANGSPSTARCSPAAACWSSSTTSVWSPTCCRCCPPARAARPW

1034_WP_158515427.1 Kitasatospora sp. MBT63 4116 4116 TS=22
MTDHQNGLSFELEDLDLGELTVTSMRDTVALPETGASNGGSSCSCGSSSCAQPQLPTLP
Y

1035_WP_117669869.1 Micromonospora sp. MW-13 112425 112425 TS=22
MSEKAATLADLEDLNLADLEMSDLEVHAVRDAVALPETGASASTSVVWQWLGYSSCA
AEVVLQ

1036_WP_148443734.1 Nonomuraea sp. PA05 966 966 TS=22
MSSENLKLDLRGLEIDSIDVVSADSIVAEGHGTIETGASSSPGYCSSYLTAMSCWAPEL

1037_WP_138671795.1 Nonomuraea turkmeniaca 3130 3130 TS=22
MSNEILPLDLSGLEVASIDVLSADSIVVEGHGAIETGASSYSPYCSSYMQACSCWVPEQ

1038_WP_142106441.1 Pseudonocardia cypriaca 4986 4986 TS=22
MSEHKPSLDLTGLVVDTVEVVPAGSLDAVAYGHGAPELGASCWCNCPGGLSTSCHVTA
GSSIDMPDEGVEI

1039_KOX20090.1 Saccharothrix sp. NRRL B-16348 165246 165246 TS=22
MGERCADCVRGPKGAPQRPFPKPCPPLTPVLASFFPERRWMVNDFNLTFDVEDLALDD
LAVTVMRDAVGLPESGASQGSGSCGSSCCCVSPPVPDPSF

1040_WP_086573078.1 Streptomyces alboverticillatus 19971 19971 TS=22
MSDLARNTDFALQDLELDLSELTVTSMRDTAALPEGGASWGSCSCQASSSCAHPQLET
GMPDLG

1041_WP_010472146.1 Streptomyces somaliensis DSM 40738 26564 26564 TS=22
MSDLTPAPGFDLQDLELDLGDLTVTSMRDTAALPEGGASWGSCSCQASSSCAQPQVET
GPLAAG

1042_WP_173260350.1 Streptomyces sp. CWH03 54324 54324 TS=22
MDRMLDTDVLELVLEGERPELEVLPAGYAPGSSVGSAGSISCASCPAASISSASTASSH

1043_WP_030721924.1 Streptomyces sp. NRRL F-2580 21796 21796 TS=22
MSDTARTPEFELHDLELDLGDLTVTSMRDTSALPEGGASWGSCSCQGSSSCAQPQYSVA
LEA

1044_WP_146884718.1 Deinococcus cellulosilyticus NBRC 106333 = KACC 11606
79082 79082 TS=21
MHQDRPLDDLDLRDLDLSEVEVESEDDTYALPEAGASVSWNSCSSIRPK

1045_WP_146884718.1 Deinococcus cellulosilyticus NBRC 106333 = KACC 11606
79265 79265 TS=21
MSHDKKQEELNIQDLDLNEVEVTTEDDSYALPEAGASIGYNSCSSVKVSK

1046_WP_066368573.1 Herbidospora mongoliensis 394279 394279 TS=21
MATNEPFHLDLDSLDVMTVELPGEDLVKALGMGLGNTEVGASGYMRTSWVV

1047_WP_173132718.1 Kibdelosporangium sp. 4NS15 177713 177713
TS=21
VARAANCTSSAAPTPWSRSAATASNWPKWSGAWRSSPASRRPARCCCPRTSIRHSRCSS
CWRRTWSPSTSWSWRRSAWRHCRTTWCRGRSACWTSCRSP

1048_WP_089003128.1 Micromonospora echinofusca 6940870 6940870 TS=21
MTATLADLEDLNLADLEVSDLLEVHAVRESVALPETGASASTSAVWQWLGYSSCAAQL
PQ

1049_WP_120331704.1 Micromonospora globbae 13964 13964 TS=21
LPTTLAKAVTNGAIRWTVRSSSSRLRCRTATTSRSLRSAGSS

1050_SCL60890.1 Micromonospora peucetia 2524243 2524243 TS=21
MPDMTATLADLEDLNLADLEVSDLLEVHAVRDAVALPETGASASTSAVWQWLGYSSC
AAQVLQ

1051_WP_151454116.1 Micromonospora sp. AMSO12t 25312 25312 TS=21
MTATLADLEDLNLADLEVSDLLEVHAVRDAVALPETGASASTSAVWQWLGYSSCAAQ
LPQ

1052_WP_132239898.1 Micromonospora sp. CNZ303 24934 24934 TS=21
MTATLADLEDLNLADLEVSDLLEVHAVRDAVALPETGASASTSAVWRWLGYSSCAAQ
LPQ

1053_WP_101413023.1 Micromonospora sp. CNZ309 4636843 4636843
TS=21
MRDMTATLADLEDLNLADLEVSDLLEVHAVRESVALPETGASASTSVVWQWLGYSSC
AAQLPQ

1054_WP_041561842.1 Nocardiopsis alba ATCC BAA-2165 2778821 2778821
TS=21
VAQEGRSGPLSALLSLRSASSASSSESCPSSLTATSLRARSSSRSGSTAVWASRETVAGTS
ASRTRRESEMGSSTLPSSSWMESPRAWA

1055_WP_026125251.1 Nocardiopsis alba DSM 43377 85684 85684 TS=21
VAQEGRSGPLSALLSLRSASSASSSESCPSSLTATSLRARSSSRSGSTAVWASRETVAGTS
ASRTRRESEMGSSTLPSSSWMESPRAWA

1056_WP_125645884.1 Nonomuraea sp. WAC 01424 61354 61354 TS=21
VSRRERRPFGNGGSMDLHEEALDLDALDVATVELPGSEVLVEAVAMGLGNTEIGASGC
TSGKSWLI

1057_WP_170231932.1 Saccharothrix saharensis 1241693 1241693 TS=21
MSSSTARHRRSPSSLPVSSHSPHTAPSCGTCSTCAPTTWR

1058_KOX34204.1 Saccharothrix sp. NRRL B-16348 140577 140577 TS=21
VRESGSGPYPRHVNASGSTSVSSAPSSPSAADTTTRASSGGQRRASSAMNSSVVGSTACT
SSNTRHSGGSPASARALACATRNVTTPWLMQRTICAP

1059_KOX34204.1 Saccharothrix sp. NRRL B-16348 154168 154168 TS=21
MIQSPGSDLMSVSKPVLNLDDLLVETIGVVPAPGPEELTQGHGVTELGASCDCGGECDC
DCDCTGSCEAIA

1060_WP_129806648.1 Streptomyces albidoflavus 69044 69044 TS=21
LATVSSSGGLPASRSRNASKSSPQGSSSRSTRSCTDQPSWSRASS

1061_WP_128826424.1 Streptomyces albidoflavus 171078 171078 TS=21
LATVSSSGGLPASRSRNASKSSPQGSSSRSTRSCTDQPSWSRASS

1062_WP_067306578.1 Streptomyces griseochromogenes 4886329 4886329 TS=21
MPDHSGDLLALDDPELGELTVTALRDTVALPETGASGGAGSGSCGSSSRAQPRLPDLPY

1063_WP_153464091.1 Streptomyces kaniharaensis 4865715 4865715 TS=21
VTSCPASGWSGAQINTMSSLASCSDSYRPGGSGPSATPTPMRPCSSAPARHSGLLTTTVN
SPSGSRRRSSRITAGSR

1064_WP_120757588.1 Streptomyces klenkii 93202 93202 TS=21
MDNYVHFSPCWPACELCSSTVPCCIPELSSDHATHTRVVLPQARHGSNSLLLWSTSGYL
RPA

1065_WP_159395493.1 Streptomyces sp. 3211 61457 61457 TS=21
MSDAARTPGVELHDLELDLGDLTVTSMRDTAALPEGGASWGSCSCQGSSSCAQPHDAV
ALQA

1066_WP_075986497.1 Streptomyces sp. FR-008 723962 723962 TS=21
LATVSSSGGLPASRSRNASKSSPQGSSSRSTRSCTDQPSWSRASS

1067_WP_167151757.1 Streptomyces sp. MBT27 624219 624219 TS=21
VPRSRDMAIQPPSVTSHPYSGRCPMTVTTTDRARSSSSQTPRPAITRSKKT

1068_WP_073775987.1 Streptomyces sp. MJM1172 88822 88822 TS=21
MRTIASSTVSLSYPAVRSTDAASSAESGPRRIVSSEGTESISGCRSSSSFTSRCSAGSAASPI
SWARAPLPSTYAT

1069_WP_030343478.1 Streptomyces sp. NRRL S-1022 280429 280429 TS=21
MSLGQNKQDALELQDLALDTDLDLTDLTVTSLRDTAALPENGASWGSCSCQGSSSCAQ
PQDNGPVVL

1070_WP_110669416.1 Streptomyces tateyamensis 33407 33407 TS=21
MSDTADLRRFELHDLDLGDLTVSSMRDTVALPEGGASYGGPCSCGSCGHPQLPTPTPW
GSAG

1071_AUS77075.1 Actinoalloteichus sp. AHMU CJ021 115457 115457 TS=20
VPGREERGWSPHPDMRTGAAMRPCQDCRSASSNHAHATRSSPIGASSSR

1072_AUS77075.1 Actinoalloteichus sp. AHMU CJ021 117906 117906 TS=20
MWVEYSELGVLHRLRDRTTTRVAVWCGRVFDAGEVRGVPHGTGGKPCLRCWERELLT
WMNRRRSGGSSSSGDALPGPLA

1073_AUS77075.1 Actinoalloteichus sp. AHMU CJ021 118778 118778 TS=20
VIPFARGRLPLSWDDGRGTSSGFTYISCCCSVGAGPAACLRARVRSSAC

1074_WP_104482524.1 Actinokineospora auranticolor 53863 53863 TS=20
VRGISFGNPVLPPESWRNATLAGSGSGSSRAAVADRSSNAADPTTWTWRTEGTSAATA
CARAS

1075_WP_091614370.1 Amycolatopsis saalfeldensis 695241 695241 TS=20
VESASRLCRCLMTHPPPGRAACANGPRSSPPAARSSSARATSGPASTPSPPGPGSPSARCT
TTSATRTGCSTR

1076_WP_173027111.1 Arthrobacter sp. NEB 688 749713 749713 TS=20
VLRSPGWSGERSWCPPWPSPRDGRGRWSVVPPACPDSSPTGARAAVCTTTTAR

1077_WP_066415890.1 Bacillus cohnii 1461750 1461750 TS=20
VIHDFGEGGEKMEIEKVMDFELEIFEVEEATVMEEMGASIFIIACCSSIEIKPY

1078_WP_043588228.1 Clavibacter michiganensis 10249 10249 TS=20
VGADAWMVPSASTANQPSLPETRSMLCTRTGRPCSSAWRVSHSMASAYELAPEKAMSP
RRRTSPVAIRSSSSQGIVVMRAS

1079_WP_106538471.1 Haloactinopolyspora alba 124215 124215 TS=20
VPVSLLAAITVTSTVSSVRAAASCEGSTTPSVSTGRYVTRTPCRRSRDRQVSSTAGCSVT

1080_WP_027344871.1 Hamadaea tsunoensis DSM 44101 38999 38999 TS=20
MRRVRLGGRAASFTSDNCGDALGGAVRSRWSSRPSSSSSSPSPSPSPR

1081_WP_054262645.1 Janthinobacterium sp. CG23_2 568118 568118 TS=20
MNESQVVVSNDDLQNLEIETFQIDEMSDLDSKAPLNIYTSSTCCCG

1082_WP_173132718.1 Kibdelosporangium sp. 4NS15 177408 177408 TS=20
MQNAASSSSSTDSTSCASTTNTASAGSTSAGSSSAQAAATPGNCATRRSTSASSTRLPRIL
TRESARPTKYSSPPSPR

1083_WP_158515427.1 Kitasatospora sp. MBT63 4348 4348 TS=20
MSDTTAQHAFDLEDLDLSDLTVSSLRDTAALPETGASWGSCSCQGSSSCAQPEVDTLAF

1084_WP_145908791.1 Kitasatospora viridis 294787 294787 TS=20
LSRPEAKESTVSGSSPKPAAATACASEASPAASLATSPSSSTRTVSPRSRTRSCAVRSPAR
SSSSRTAVSARSTRLSRWLVTGMP

1085_WP_150929942.1 Microbispora sp. Gxj-6 113102 113102 TS=20
LAEYRGTRRTPPETGGSMAMHEEALNLDALDLDSLDVATVELPGSDLLVEAVTMGLGN
TEVGASGAWTSRTSWLV

1086_WP_110565432.1 Micromonospora arborensis 125716 125716 TS=20
VTATRPAGSPRPGTARTGSCRRTCWSTTSSRPARPTCTTPRSRSTRSATGCRTPGSGMGR
PAPPPSRWRPTAPSRSVST

1087_WP_165947742.1 Micromonospora sp. CNZ303 22051 22051 TS=20
MNEHSPEFAISDLPVDVFELSTAGLEVESLTAGHGMVEHGGSNACNSMVLACSCGQCA
CSSGSCHEVW

1088_WP_123600439.1 Micromonospora sp. Llam0 394152 394152 TS=20
MPTNPFFAVVAAIAIAAPATQAAVEAQQPDETTSQSMVVFELPTESTGMHDDSSSVNFA
CAFN

1089_WP_167477798.1 Nocardia arthritidis 8606460 8606460 TS=20
MAWQRLFPQLTTSSCSRTPTTGIPFEGIPVTSFAATLIG

1090_ADR01080.1 Nocardia sp. ATCC 202099 3415 3415 TS=20
MRSTYPGDWVNSLGEPPWFPANTTRSPRTAFCSTLRAVGGALSSPHSPSSSALVGTALLT
CNKSIPSVARWRLPPS

1091_WP_087097343.1 Nocardiopsis sp. JB363 41601 41601 TS=20
VHSEESARDQRRAPEEAPYTSNGTSSSFRTTTLSPSRGSAASASWLSGDASLFKSPCLSFV
R

1092_WP_143802042.1 Paenibacillus thiaminolyticus 1637569 1637569 TS=20
LPSSTSIPADMAVNAFVQEPIANSVSGVTFSPSSTFRRPKPSAIAICCPWMTAIASPGTFHC
SICC

1093_WP_051750259.1 Phycicoccus jejuensis 398867 398867 TS=20
VGRDGHERMPPMPRAALAVMVWRAVVVSSMAVSSGWSGPAVGRPSGMPDSSPAGAR
AAVCTTTTAR

1094_WP_051750259.1 Phycicoccus jejuensis 376795 376795 TS=20
MGDEGLSSHQRQVLPRRQRWPPSPVTRSPPTCAARRRCTSTTWPAAGSGTAAPVPRSGS
GRCRPRSARCRSTTASSRSSCRCAAPTSSS

1095_WP_142106441.1 Pseudonocardia cypriaca 0 0 TS=20
MTTVLLAEDDAAIAEPLSRALQREGYAVEVATDGAAALERVRRGQVDLLVLDLGLPG
MDGLEVCRRVRLDDPDLPVLMLTART

1096_WP_073888311.1 Saccharothrix sp. CB00851 83142 83142 TS=20
MSHVDAPDVEMIGVVPPVGLEELTEGYGLTELGACGGTCLCDDCDGTCLCDDCDGGTD
SGSAIA

1097_WP_037604491.1 Streptacidiphilus rugosus AM-16 1500923 1500923 TS=20
LPFMSPIGCSGGFPTDSSRPTSTAPKPCRRLPTTSSPASSGRSASHRLASPSIPTSARPCTAA
CSPSGGC

1098_WP_037604491.1 Streptacidiphilus rugosus AM-16 1502363 1502363 TS=20
VRPRRTPTSAWRFSSWATTPPLSGTIGGRSRSTCGPPTASPSARRSPISRSPCSSSAASTRR
PTRHDAPSR

1099_WP_040246700.1 Streptomyces albus 1271860 1271860 TS=20
MPPPARRRTAAATTCTSARPPGPRCSSPGPGSPGGAPSSSSASPPRRPWRAPGPC

1100_WP_120757588.1 Streptomyces klenkii 93377 93377 TS=20
VDQRSSEFEPCRACGRTTLVWVAWSDESSGMQHGTVEEHSSQAGQQGEK

1101_WP_037796980.1 Streptomyces sp. ADI91-18 1315959 1315959 TS=20
LMTSMYGPSGIWATSAVTKPSRTASSPPTMPVTTASRRNAAPQWRMTSTRSSS

1102_WP_148643628.1 Streptomyces sp. CB01881 1892339 1892339 TS=20
MTSPPPSATATGSGTPVPDARRRSGFVALGMVGTLALALSGCSSSSSSPPSAASTPPR

1103_WP_058043695.1 Streptomyces sp. MBT76 2269461 2269461 TS=20
VSPVSWLAIVLLGPTPTTVLPFITSVFINRTVSKARRTTSSAASRDISSMPTAGPNMACSSS
SATERTGKLSGA

1104_WP_053707100.1 Streptomyces sp. NRRL B-3648 119288 119288 TS=20
MSLGQNKQDALELQDLALDADLDLTDLTVTSLRDTAALPENGASWGSCSCQGSSSCAQ
PQDNGPVVL

1105_WP_030903859.1 Streptomyces sp. NRRL S-515 2011 2011 TS=20
MSDTAHTPAAELHDLELDLGDLTVTSMRDTAALPEGGASWGSCSCQGSSSCAQPQDVT
ALEV

1106_WP_030903859.1 Streptomyces sp. NRRL S-515 2217 2217 TS=20
MNDATQSAGFDLDDLDLGELTVTSMRDTVALPEGGASNGGSSCSCGSSSCAHPQLPELP
L

1107_WP_125824502.1 Streptomyces sp. W1SF4 213521 213521 TS=20
VSGDRVSAMTRHEPSGAANTRVRPRCSSSRATSSVVVSSAASAGSGPATAAAGGRAAR
AAANRSAKAKTSFSSSVSR

1108_WP_125815390.1 Streptomyces sp. WAC07149 84665 84665 TS=20
LASPTASSPGPPASQTTGSGFGWAALAGTIATPSLMVRPPGSARFSGTVSVPQRAPSSSAT
GCCVFGQGPGS

1109_WP_033226943.1 Streptomyces virginiae 13762 13762 TS=20
MSDTAHTPAAELHDLELDLGDLTVTSMRDTAALPEGGASWGSCSCQGSSSCAQPQDVT
ALEV

1110_WP_033226943.1 Streptomyces virginiae 13967 13967 TS=20
MNDATQSTGFDLDDLDLGELTVTSMRDTVALPEGGASNGGSSCSCGSSSCAHPQLPELP
L

1111_WP_116178934.1 Kutzneria buriramensis 158169 158169 TS=19
MFMYSKGPVSTRRATTVRAARSSSRSDTQLHTRPCSSGCMVSRNLT

1112_WP_170199045.1 Saccharothrix variisporea 46041 46041 TS=19
MTASKPALSLDDLAVESIGVVPAVGPEDLTQGHGLTELGASCEDCDRMGGGSCDCCEC
DDLYRVIE

1113_WP_048778129.1 Streptococcus gordonii 1236 1236 TS=19
MKHLDKEQAAALEEISQLDLSILDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

1114_WP_125400247.1 Streptococcus gordonii 265 265 TS=19
MKHLDKEQAAALEEISQLDLSILDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

1115_WP_103344440.1 Streptococcus parauberis 577 577 TS=19
MKHLDKEQAAALEEISQLDLSILDLDESDLVSSAAASSGSSSCKVTSTCGSSSCCA

1116_WP_116178935.1 Kutzneria buriramensis 125970 125970 TS=18
MVTEKVQYALGDLSMDVFTLTDRGLTVESLTAGHGMAENEASSSSHCSGSCSCGGGSC
SCGDGGWDGGWGWGGGHGGGHGHW

1117_WP_158977512.1 Streptomyces roseus 3142 3142 TS=18
MSENSATPGYDDLDLDLDLDLSELTVTALSDTAALPENGASWGSCSCQGSSSCAQPQVE
TPNV

1118_WP_037873311.1 Streptomyces sp. NRRL S-37 33184 33184 TS=18
VPPRSLPTATGGRTGPRRPTAGSRAVRSPAATERRTTTRCCT

1119_WP_052395034.1 Kutzneria sp. 744 4443490 4443490 TS=17
MPPTTSCRSYDSRPAEITSFHRGPGAVGPVRQGGGRVAEGDPASRRATRRTTTTTKAPE
ARTRMCC

1120_WP_164903976.1 Nonomuraea polychroma 8293219 8293219 TS=17
MSDGYPTAAQKEALRLICDHGRLHTEELGHHLVSARRSSTNPGFAPAIARMAGTLAWR
LEVQGFIAETGDEWTTTADGRRLISCSSEPE

1121_WP_152265679.1 Streptomyces mobaraensis 77311 77311 TS=17
MENAANAALLDLEQLGDEMFEVVQDDAVESDDNAAWAYACCRSASSSCALS

1122_ARP73782.1 Streptomyces pluripotens 7247870 7247870 TS=17
VADLPGQPLVERVRQEARQGQEGRARWCTTTWSAVSALPQWQSQLCSSGWAVIPRKV
TARPSIAASRRSRARDRSRFIMDVRLRGLRELAPPEDR

1123_WP_109891469.1 Streptomyces sp. NEAU-S7GS2 5447447 5447447
TS=17
LPLGVSGQLSTTTNAEGVMYDGSSVAVSSRTAAGLNGGSSGACWT

1124_WP_018090006.1 Streptomyces sp. SID8375 259491 259491 TS=17
LPLGVSGQLSTTTNAEGVMYDGSSAAVSSRTAAGLNGGPSGACWT

1125_WP_089317077.1 Actinomadura mexicana 1715 1715 TS=16
VAPRSAPLEPSAQGSSMTNSSSIPTARCGASVGSRRGVDC

1126_WP_122195758.1 Actinomadura sp. NEAU-Ht49 4910 4910 TS=16
MTAGSLLVREGHSGTCGTWDAGGRTRDTERQKRLPSSRGLGFS

1127_WP_156046005.1 Herbidospora cretacea197191 197191 TS=16
VRRSPAAFPVSGPTPCSPERHSPSPDSTSPSWARNTISTTPTCLRSTTSASSSTARSSIPATR

1128_WP_156046005.1 Herbidospora cretacea226789 226789 TS=16
LSPLVARPPADPLRRRCAHRSGREVTTTTARPRLGGAMPAAITSHSRFVNC

1129_WP_062436515.1 Herbidospora daliensis 157084 157084 TS=16
VCGLPCTVTSTSCSTTTTSGSPPWCRETRRTRTTSAGEDPQLRPYPARISSRPV

1130_WP_151456132.1 Micromonospora sp. AMSO12t 5164 5164 TS=16
MKKEMSGPNFAISDLPVDVFELSIEGLEVESLTAGHGMPEHGASGHLSPVVSGCSCGSA
CYFHY

1131_WP_073685807.1 Streptococcus salivarius 31244 31244 TS=16
LTFEIKLLLFWFPSLETTPSISYTKTVFRSSPKHSHCFSLRGILWRGCFSPSSPYLLFHKTSL
FFENIVVVTIGTSCCSNGFFLF

1132_WP_042800563.1 Streptomyces sp. C 1161653 1161653 TS=16
MTYPTRWVAPAAFLRATTQVCRTSGCSARQASTSSVSSRMPRILTWLSARPRKSSEPSG
A

1133_WP_109200274.1 Streptomyces sp. CS014 113841 113841 TS=16
MSLPEMATAFDLDSLDLGEITVTAMRDTAGIGESTYSLSSSSSTSCQGCSGCSVQPPASID
AA

1134_WP_109178846.1 Streptomyces sp. CS131 2972410 2972410 TS=16
VQAHEVSVRDQISGGSDVAGSRLLAGQREGPDTWDAPGGSRLRPGGDSRARNRPGGPG
RPSCRARRSSTVAHRSSPYRTQLCRGSSQE

1135_WP_125815392.1 Streptomyces sp. WAC07149 132625 132625 TS=16
MPGGDRPPGRVRCCTGERMPCFATTSRGGEECESPIRVSASHTTPA

1136_WP_157556458.1 Herbidospora yilanensis 46816 46816 TS=15
MTIVPEIARETGEQMANDEPLHLDLNSLDVTTVELPGEDLVEALGMGLGNTEVGASGIG
RTSWLI

1137_WP_003649851.1 Lactobacillus paragasseri JV-V03 6157 6157 TS=15
MKDLFDIDEKSLENLDLGDFKILDSEDLSEKDEKSILGASCTTCTCCCSCCA

1138_SEG89460.1 Nonomuraea solani 499933 499933 TS=15
VTVGFGLTIAIQVGVDTPCPTRPSMRRVDHHGRARPITASAENSVRRSHLQVLGEERSSL
KRSLKQCCAVAQSS

1139_WP_148033048.1 Nonomuraea sp. C10 5232706 5232706 TS=15
MSAENLKLDLRGLDVDSIEVVSADSLVVEGHGAIETGASSALGYCSSYLTATSCWAPEA

1140_AXI76532.1 Streptacidiphilus bronchialis 563785 563785 TS=15
VPPTGSRAWDRTPAARRSCCGCPTATTTAWCCPRTHGSGRPTM

1141_WP_071967152.1 Streptomyces cinnamoneus 239197 239197 TS=15
LLAGKWGRVDRFGVEHESGPRGVSSTTSSRVSETRKPASSRTMRPACRRGPVGLNSTPP
TDTSPELRPRQSSTGSCTTVT

1142_WP_051879301.1 Streptomyces sp. NRRL B-24720 27804 27804 TS=15
MVYQLTLENQQLREQSSSPTAIVRALPLTPSPCRPGPAHLALAHQPVAAPGVLPEAGSSG
GTTTGTGRAPARAPWPRSRVDPARAASGERPVTEG

1143_WP_114024659.1 Candidatus Streptomyces philanthi 72894 72894 TS=14
VGAPGVEWSARASQVGTRPSWAFLPSRGVSSSSTDLYCVTKPGGSQT

1144_WP_049161721.1 Lactobacillus gasseri 3701 3701 TS=14
MKDLFDIDEKSLENLDLGDFKILDSEDLSEKDEKSILGASCTTCTCCCSCCA

1145_AFR07166.1 Nocardiopsis alba ATCC BAA-2165 1704912 1704912 TS=14
MTRLLSWPVSCSEERPRVIVVRDSDVIAKPSIRGCSSSLDSCSSKERTRKAAR

1146_AFR07166.1 Nocardiopsis alba ATCC BAA-2165 1753581 1753581 TS=14
VAPPHPTPGPPTSSGSSTGKRNRRPGPWPPPHLEGNAGRACSSRPVRPRSPSSWSPPVSSC
GPSPKRNR

1147_WP_127937127.1 Nonomuraea polychroma 8861525 8861525 TS=14
MPCAAPPDAYLLADGHPLREHAWRTDRAGTSSSATPRWACCSPSGPAWSSRAAEHLTG
IRRFTCGFFGK

1148_KOT43230.1 Streptomyces caelestis 37519 37519 TS=14
LPAAPGGTEERCAPDSCEERSSGVTGPGPLHDRHWPSVT

1149_KJY39143.1 Streptomyces katrae 4093 4093 TS=14
VVTTGRAASPPRAATRESSSAPSSSSGTHTSATRNSAAGPRTTNEVPSAPK

1150_BAU81447.1 Streptomyces laurentii 554629 554629 TS=14
MLTHANLIANICQNEQVLPQTTAETVPAVLPFFHDSGSQGHH

1151_WP_031510723.1 Streptomyces megasporus 38447 38447 TS=14
VRPGGPPGRVPVEGESRGIATVDRLLAWQWHRPDPSPRAQPSSLRATTCREASMAGRC
RTT

1152_WP_031510723.1 Streptomyces megasporus 67462 67462 TS=14
VVSSLADARIQGGNAPGLTLHQSLIRQRGRRNRACCWRLFFRGTWLTRSLPRSSARFPCF
NSRCRNRSSPC

1153_WP_026411568.1 Actinomadura oligospora ATCC 43269 237815 237815 TS=13
MGGSTQPSTAETRETRLVRAPYATAAASAVINLIPLSPSIHITYRRCRYQPDGTTRHSPGT
SHLMAVGGSEPGGPQGRWVVTKRSGASTGARGSSSGS

1154_WP_026411568.1 Actinomadura oligospora ATCC 43269 280008 280008 TS=13
MPSENPLAGFLSTEWRPTSSSTASTRRVGMPLLRARQSRWLWALRPPCTALASSSAPTW
CSGAPRPA

1155_WP_157429535.1 Actinomadura oligospora ATCC 43269 236750 236750 TS=13
VPSAIAASTAQPGSLSWAQSENRQSAASSSISANTCWMPLSPSWTPTARRPGVSMRTPPS
GSRISSRWTVV

1156_WP_157429535.1 Actinomadura oligospora ATCC 43269 237398 237398 TS=13
VSTFAARTWAWDVRDADDRTMAVRRGSNATSSGRSPSRTAAQSPVQGGSAAVDQAAR
TQPSAVTTVAWPRSTRATRPGVRPSERCAVKAFSYASSHP

1157_WP_094862908.1 Amycolatopsis antarctica 325770 325770 TS=13
MARAGEGHTNAGALLGPCLPVRPGSATTPADRGLVPAPAGTRRPSPADRPQGT

1158_WP_034336238.1 Deinococcus misasensis DSM 22328 90312 90312 TS=13
MLTEAIIPVQSNPTPQMLRALSSLNETEVLSEVKTVKKPTKLGDSLISCWANPC

1159_WP_034336238.1 Deinococcus misasensis DSM 22328 95543 95543 TS=13
MWVLLVEVQEGVASGSLMLGSKPVFDKGKGGRSGCCPEGVLLTDIFIEHSDIFCEVLT

1160_WP_026928618.1 Glycomyces tenuis 308468 308468 TS=13
LRRAGSTPPGSTPAAPWSPTPNRRPGSTLGPDPLRAAPRPSGRGAGPSAPALRPGTSGRS
QWEPSRYSRSPSSVCC

1161_WP_026928618.1 Glycomyces tenuis 311897 311897 TS=13
LDLSSARPRGAERSSAPRTRRPPRLRPGAAPRCGRLLHRLHHPRESHR

1162_WP_026928618.1 Glycomyces tenuis 312204 312204 TS=13
VNTVVPLTLAPVNQASLLNCTSANRASSLNCAAPNLAAPSNRALSNQTSPSNRALENSA
SPSKYVLLN

1163_WP_063494187.1 Lactobacillus plantarum 7305 7305 TS=13
VQSGLPFSMIYSTTKIKKLLVMRIANGQRKLCEKLIRLYTAMPKQEKLWHS

1164_WP_067368384.1 Micromonospora rosaria 26161 26161 TS=13
VILVDHGPEQVGEAVRPGTSSVQVRRAFIPPTAAWMPSYRSGRHCREHARDTTTVDLID
DTSVGQVEARQGRRDPRP

1165_WP_040691311.1 Nocardiopsis lucentensis DSM 44048 7330 7330 TS=13
VSVHADLPTSNKSMLLPSSTVRQWGKAHLAYGRVMCGAVAGQEREIDRRGG

1166_KAA1017585.1 Pseudonocardia sp. EV170527-09 4711 4711 TS=13
VAALKFSVDPISPPISTMSSSTVVTRSATGMPACAAIRPRPRPCPPWTAFTTRRILGSAVSC
R

1167_WP_070437603.1 Streptococcus sp. HMSC10E12 15670 15670 TS=13
LSNFWGAVQKLFAHGEDNIILVTNFYDYSLFISMKSSMYPFINSSSVV

1168_WP_125051747.1 Streptomyces rimosus subsp. paromomycinus 951775
951775 TS=13
VARSRWTRPSVFGASTSSQSSAVLPVSEASRRTPAACTTPDSSPACSRTRVRRWRTSSQR
ETSALS

1169_WP_138054944.1 Streptomyces sp. ICN19 5251424 5251424 TS=13
VDAQSTPNPEGPHPFKTAQPRSASPSTTSPDGTPRAGPGRPGTEQLVSCASNPSVDLPRL
GDRGSGRRSSALAVRACAGTT

1170_WP_138054944.1 Streptomyces sp. ICN19 5272952 5272952 TS=13
LRHMVYEIIAASSTRFPCFDSRCRNRSSPCCFFKTRPLGEGRSHRT

1171_WP_109293928.1 Streptomyces spongiicola 1931785 1931785 TS=13
MNDITPATGLALDLRLDQEAPELEVLPTSHSPGSTVGCAATGSSISTPSGCFSSAGTASSA

1172_WP_109293928.1 Streptomyces spongiicola 1932129 1932129 TS=13
MDRLPETDVLELVLEGERPELEVLPAEYAPGSSVGCAGSISCASCPAATISSGSTASSH

1173_WP_109293928.1 Streptomyces spongiicola 1945308 1945308 TS=13
MPHSRPPAPLDPIRLPDRTHYGKYPTGSKACQTAAVQARFTPNERALLCIPCESINSSDTT
TPT

1174_WP_109293928.1 Streptomyces spongiicola 1958314 1958314 TS=13
MQVGGAAPVGRRARSGLTGARGRWIRVGTDPFRRYGESRTDRRRNATVSSSRVCCWT

1175_WP_037914261.1 Streptomyces yeochonensis CN732 3598508 3598508
TS=13
VVPERSGVERAGSAGGAGGAGLELSAAVSAGEASSSPLSRPAVTAASAGPGPSRRASAA
VSRSRPAVPSALSATRVRTVPVAVTARSRVRSGATTR

1176_WP_037914261.1 Streptomyces yeochonensis CN732 3604590 3604590
TS=13
MSIDFGAFDVTELEVLEAGDGVALPDMGASIVVIVAATPDGEEILGEADLGVSGSLSTSC
C

1177_WP_098777088.1 Bacillus cereus 667 667 TS=12
LENDLPTKFFFKMVPLEYKDLVKSQDDDNYENTDAKPLYMDLSNPIFVKVFRKLTTTIK
YGLLIEEVLPDLGEYIDNSEQENYVEEYILELTQKVCKGI

1178_WP_098777088.1 Bacillus cereus 24831 24831 TS=12
LPFPPGGVVLASSVFIRIEVGCDLLPAASFANTENVYSVLGFKFLIVMLFVVVLHFSSPAV
ISYSVTPTLSVDAVQFN

1179_WP_028414789.1 Bacillus sp. 278922_107 416169 416169 TS=12
VKETEKSILAQVKNEPLINIFDSWFVFLLFNYVFTCPLFPSIPSMCETVS

1180_WP_028414789.1 Bacillus sp. 278922_107 443549 443549 TS=12
MELKLDFNVLENLEENISFDEVKALDIEGSIGATEGASSAGSSLPFPPYFWFTCSA

1181_WP_009582377.1 Fulvivirga imtechensis AK7 38911 38911 TS=12
VLARTPSGFKKRSSSRLVKYRKAESPGDEHLWGKHSADIGACFRDTT

1182_WP_090945800.1 Nonomuraea jiangxiensis 12225 12225 TS=12
MSAENFTLDLRGLDVDSIDVVSADSIVVEGHGAIETGASSALGYCSSYLTATSCWAPEE

1183_WP_027734140.1 Streptomyces sp. CNR698 45136 45136 TS=12
MNDITPVSGPALDLRLDQEAPELEVLPASHSPGSTVGTAATGSSISTPSGCFSSAGTASSA

1184_WP_027734140.1 Streptomyces sp. CNR698 45463 45463 TS=12
MDRLPDTDVLELVLEGEPPELEVLPAGHAPSSTLGSAGSISCASCPAASASSASTASSH

1185_WP_018891562.1 Streptomyces sp. CNT302 63945 63945 TS=12
MDRLPDTDVLELVLEGEPPELEVLPAGHAPSSTLGSAGSISCASCPAASASSASTASSH

1186_WP_018891562.1 Streptomyces sp. CNT302 64272 64272 TS=12
MNDITPVPGPALDLRLDQEAPELEVLPASHSPGSTVGTAATGSSISTPSGCFSSAGTASSA

1187_WP_170812553.1 Streptomyces sp. PKU-MA01144 10589 10589 TS=12
MDRLPDTDVLELVLEGEPPELEVLPAGHAPSSTLGSAGSISCASCPAASASSASTASSH

1188_WP_170812553.1 Streptomyces sp. PKU-MA01144 10916 10916 TS=12
MNDITPVPGPALDLRLDQEAPELEVLPASHSPGSTVGTAATGSSISTPSGCFSSAGTASSA

1189_WP_129842995.1 Streptomyces sp. RFCAC02 5132401 5132401 TS=12
MTEGVVIVAPRERQESSSAGSPATCRPAGFTRHCTGLDRQSCAFAGNPPFRPD

1190_WP_161178111.1 Streptomyces sp. SID4985 1924 1924 TS=12
LLSGVLISPFLAVSGAAGTTAGAAPPGGLWGWGTACGAVCGADAVSGAATAARRSRSE
RPSPPSAHSSSTAASRAGGSDQASTA

1191_WP_161178111.1 Streptomyces sp. SID4985 1991 1991 TS=12
VPTAVPRSARCCCSARCWASPTSSRCRRSTATRRSPTPSAWSNSRATAVRSWRSRYGAC
SGAVRASPRTRRGCGCCTAASRC

1192_KOG56891.1 Streptomyces virginiae 7157 7157 TS=12
LVAWTALAVSVTVTFTWYSPDLSYVWEWCSSAGPWAVAVSVFPSPQSMTTPRAKSAL
VIRVVSSMPWAV

1193_WP_119313826.1 Meiothermus terrae 14190 14190 TS=11
VEVSITRGAAGPVVAVGREPVARSPPEALRCCRAARRRSPRDPANVGDFT

1194_WP_119313826.1 Meiothermus terrae 16603 16603 TS=11
VRRSANTPPPAHRNSTPASSRRSRGASSTPDKGGDEQGLRVKVVREVLAIGSHGTPSFLT
PCRPRSSSKRKE

1195_WP_119313826.1 Meiothermus terrae 16347 16347 TS=11
LLLDLGRQGVRKEGVPCDPIASTSRTTLTRRPCSSPPLSGVDDAPLERLLEAGVEFLWAG
GGVFALRRTGGFGCAPLKRL

1196_WP_151462301.1 Micromonospora sp. AMSO31t 32525 32525 TS=11
MACTCAGPSVGCSSCPPPHRSSSRVRVVRAPRGSPTSSVNIR

1197_WP_117665659.1 Micromonospora sp. MW-13 14251 14251 TS=11
VRGQYGEHCAGLGCCRRDGSVVMSGGERAKDCDTQRSSPY

1198_SCL48150.1 Micromonospora yangpuensis 836745 836745 TS=11
VAGGSRSLSPIACRTSCSRCASRASSSAVTVSADRRAVMRSSCSHHTAAA

1199_WP_149666213.1 Pseudonocardia sp. EV170527-09 327 327 TS=11
VAELDAGVGGGELPVHLTVVGVGGVLPGCEFGVEGVDVGDASVGALAGQGREFDLGD
VEPGTVSGLSSAGCST

1200_WP_149666213.1 Pseudonocardia sp. EV170527-09 2491 2491 TS=11
VKPAISMAASRQPFPASATQQRFNHALSLPILRYSSILATQDPRSQSYTPSSIQCIRGSVPG
VTT

1201_WP_162824113.1 Streptacidiphilus bronchialis 592495 592495 TS=11
MSIDFGAFDVNELEVLDVTDSVALPDMGASAIIIFPTVTPDGEEILGEADLGISGSLSTSCC

1202_WP_162824113.1 Streptacidiphilus bronchialis 606799 606799 TS=11
MRLRAADAVPERATTCCTRTTRPLYASGGRSSRGRCSPTEGRPTGLRPEAEADLSRGNR
SHA

1203_WP_167515856.1 Streptomyces huasconensis 25346 25346 TS=11
MRAARSSRSRTSSSPSTTARSPSWAPPPCAAPRRRSSSPAAPTSCAAPSSRSA

1204_WP_100865210.1 Streptomyces sp. Ag109_G2-1 32801 32801 TS=11
LGPDVARKAAPLAYSRHVACTPREVFTAPHRPGTTRTRGKRRRPSCAGSSAPSRPRRPT
RTRPNSPPSRPGSGTAAPTSTACCAGPAGRSAATGSP

1205_PIG63282.1 Streptomyces sp. CNZ279 1456744 1456744 TS=11
LFSDHFWPLRLKTQGCEPNSRHLPGEFTPRSPSPFAFVTPCG

1206_PIG63282.1 Streptomyces sp. CNZ279 1475724 1475724 TS=11
LYNVGNGDSWHDSTHAGPCPGTRRRQPNRARSGQRAGGVGHG

1207_WP_161371729.1 Streptomyces sp. SID5770 7375 7375 TS=11
LRALSESVGRYGSGIDLAGPYQKRDGGGSASRRLPRLSALDGRPPGSAPRADERAAGRP
AARSSECADA

1208_WP_146894940.1 Adhaeribacter aerolatus 241014 241014 TS=10
MKNQKLKLDELKVQSFVTDFDKEQGQTQEINGGTGRAYCIDPSQLIICLGPRLTITRNISC
FAATCNYALCNPEQIPTKIPTDFPVTTTTWDGGTIRF

1209_WP_025028233.1 Bacillus mannanilyticus JCM 10596 35897 35897 TS=10
LHSIAEETRRTLSRAELKDECKNHLREGELMKKVLNVVFKKSGTTVVPGDKMSW

1210_WP_107724852.1 Desmospora activa DSM 45169 565287 565287 TS=10
MMNQTSHLAPLLIDLLLSWNRLAPTIPLRIAENLPQENPCPYPLSDFSCFGRHRRRSSSYP
QHCPFTRIR

1211_WP_107724852.1 Desmospora activa DSM 45169 569583 569583 TS=10
MKFQLDVTELEKMEKDLSIEEIQAIAIPGTTGNTEGAASVGSTLPVPPYFWFTCSA

1212_WP_028594895.1 Paenibacillus assamensis DSM 18201 26155 26155 TS=10
VTYPPREWPSSTNLSSSSLSANAIISSAICSIVYAPFGAVDFP

1213_WP_028544504.1 Paenibacillus taiwanensis DSM 18679 161371 161371 TS=9
MNAKTILNADYILKSPPSCTMLPWMRIGICCFYLCLSSYSSEAAGYGVHE

1214_WP_121832244.1 Streptomyces sp. S1 6004 6004 TS=10
LAAARAAVRVRRAPPPRDGPGTVAGPGYRAPGERRRPDRGGRTERVGRSARERRCAAA
GNGRAPAPVRRCAG

1215_WP_033226294.1 Streptomyces virginiae 5190 5190 TS=10
VGGGGPGGGGGGGGGGSGCELRPDKGPGGGPVGCSEEGGREFTTLAPRHRAVGKPW

1216_WP_033226294.1 Streptomyces virginiae 28459 28459 TS=10
VSGESGPWSSRSGPVTIGVSVSRTDSGLRLRTSKEITVRSSVAPRVSTCDIRSGLRGDGRH
LFICVVRQAISGRRVAAHARRGRQGARL

1217_KIC48916.1 Tateyamaria sp. ANG-S1 679855 679855 TS=10
MWLTQIGKSREGTTPAIALSSSWLSRYDCATKGRIAAPIGRDVNGT

1218_WP_084467419.1 Actinokineospora inagensis DSM 44258 8378 8378 TS=9
MPGSRVGTAGIAPTGEGLEPEGSPAKPSSRFDGGRTRHHWSSSYRSDRRTGSSTAWTRL
CGGVGYRRKRWRSCIRPHWQGSARRRPVSAQTLRAALRR

1219_WP_122194823.1 Actinomadura sp. NEAU-Ht49 50770 50770 TS=9
VPEHQRPATRLDSGVVQRGSSSTHLQRQRCRERTGLGGATTESFPHAAAQNRSGHEKR
QPRSVPPRSS

1220_WP_120609584.1 Corallococcus sp. CA053C 10973 10973 TS=9
MNGGGSAWGHLAPLGISGMDATPWRRTSGLSNPIQGSDIRCSRMKSGSSRRTGIVPRHG
GWRAPSGLPLEAEQDRRPWAGQWPQACSSP

1221_WP_147012481.1 Luteibaculum oceani 270507 270507 TS=9
MGFNPEVRIALFPFLERMGNKKAPRCGVGLSEVHMWSHLGSNQGPLDYESSALTS

1222_WP_155472526.1 Massilia buxea 23762 23762 TS=9
MHGCLRWSRDVDLRVLCGTFDQSCTKTKFLVNGGVGRTNEKPPERRIFLSEGEVGCIST
TYSLRICSSTTFNARKLWTCQKFCVRGQIS

1223_WP_091433994.1 Micromonospora yangpuensis 824705 824705 TS=9
LTAPWSIGSHRASSNAWLVGERDAWPCGSSDRLQIILDPTRGHRPGHRARCVLRLARGG
CWRGISTRTTSPEWPARLGAAPAEAGDGTAADAR

1224_RYZ37930.1 Myxococcaceae bacterium 16808 16808 TS=9
MSTAPVGASAHSSDVAPKPSNPSRRMRLRPNASDSEPATSINAPSVMRYASTTHCWSVR
PPPRSRTMAGSATLMMAPSRNVTSEARTAMARTSRC

1225_WP_169027295.1 Paenibacillus aquistagni 31448 31448 TS=9
MEASASFSGQAICKYARAMAGSGRANTGNPQPSTPVPSSSGLVIDISAMHAAISCCAGCS
ASA

1226_KUJ64075.1 Streptomyces albus subsp. albus 63 63 TS=9
MSDLLHRSPDFLAYRIQASLLYSCLHTMGFGLAERYLFCYILARANEEVAGRTTEELAQ
GLDAVAQEMAAQQGATTHRLMQAP

1227_WP_167972679.1 Streptomyces lonarensis 11278 11278 TS=9
MTAVTSDTAPAPAAVPTTAGPFAAARDRLERLLAERHAREGRPPLLGA

1228_WP_030019170.1 Streptomyces monomycini 164156 164156 TS=9
VAIAPAAAAASAMPEFRPRAPKGAKRCPASPASSTRPCRSVKSSATTCWKR

1229_WP_159036914.1 Streptomyces sp. Ag109_G2-1 39 39 TS=9
VDEFQVGFAEGGPAAVGADLVGDEEGWRLALNPPTGRIWAARAPGPRARRICPETARR
PRGPGCRGSGW

1230_WP_159036914.1 Streptomyces sp. Ag109_G2-1 667 667 TS=9
MVDLGPLVPRGTRHAPGGGRVTHLPRLRHRDLVLRRRRWLLPPGTVERLRADLTADGT
VPAQAAARWRAALDSPSRSSCTPPEPPPPAGPPRTS

1231_WP_159036914.1 Streptomyces sp. Ag109_G2-1 26323 26323 TS=9
VGSLVPGPMEPRTQRGLPSSAVKASADSRAMRAPASESS

1232_WP_051779832.1 Streptomyces sp. NRRL S-241 16742 16742 TS=9
VPPSVRGHNLRGPGPPTASAAASSAQRRTTRKKTAGESSPPAFFALFQVSFCMEKSACRR
LRSVSVLSGIFAPSPGGSGRRRRV

1233_WP_051779832.1 Streptomyces sp. NRRL S-241 24271 24271 TS=9
LQSWTSRSHVDGKMTDNLGRPSSSRVQAGSRVNFSNIPGVNSDKGKCLW

1234_WP_078937765.1 Streptomyces virginiae 11766 11766 TS=9
VEYEMTAAVGAGAEQDSAPPSSGAGHHGGRVCPAPQQAPPAGHRAEVPRRDRGKALT
CSFLQGLSC

1235_WP_119930010.1 Streptosporangiaceae bacterium YIM 75507 88159 88159 TS=9
LSREVSNISEGPLTPRASQFRRTPPARPSGCSRTNASAPSTPSSSPSDISSTMSLRGAGPVRS
ARTVSIIAATAAPSSDAPGPMAAES

1236_WP_157163121.1 Actinoalloteichus spitiensis RMV-1378 3967 3967 TS=8
VSVSRFGSSCFPGEVDPCHQKWEGVRYRTTLVRARLLGDGSPVIAGAGWTAEGTRPGF
RRSTTGSRTIRSRRPPQPTTKCAAVAVHA

1237_WP_150091428.1 Adhaeribacter sp. DK36 16913 16913 TS=8
MKNQKLNLDELKVQSFVTDFDKEQEQTQDVNGGGSYQIYCVSQRTLQKDCFFVPVTTT
KFTDIWSLGCSYGCPNTIDISDVIDNPVVRF

1238_WP_157554633.1 Herbidospora sakaeratensis 1005 1005 TS=8
VVPRRHGSPDSLNNGHRLRPAQGTSSLGDAIETLAGAAACASSSPTYSTTPAARNTPSRA
AGERGAAPRCSPAGPGVADGSWSAATSG

1239_WP_062351038.1 Herbidospora yilanensis 176204 176204 TS=8
VLLSRIEADRSTTITTSSAVTLPVEPCDSMDAHALAGDDGTGADVH

1240_WP_082527048.1 Kitasatospora sp. MBT63 462 462 TS=8
MPFGLAGDQDQQAAEVHPPRHGKLRGFAAMCQGRARGWAS

1241_WP_027890929.1 Meiothermus chliarophilus DSM 9957 6064 6064 TS=8
LDGGVLAIASSWSGLEPAHVGPEERICPSPEPESRTGTD

1242_WP_027890929.1 Meiothermus chliarophilus DSM 9957 6082 6082 TS=8
MDNTKNTPKPEDIQADFNLEDLDLNDVKVLAESDSYALPEAGASIGKFSCTIVITK

1243_WP_132402251.1 Micromonospora sp. KC207 12591 12591 TS=8
VAAPAAAPAGSARFRSTTSRSPTWKSAGGPSCGSWRRRCRRWDSSA

1244_WP_028559081.1 Paenibacillus pinihumi DSM 23905 = JCM 16419 548961
548961 TS=8
MSSLASSRARTKSRNTSSSSVGICTAVNSPPGITGLIFTRLICLFLSGLLHVLAGQLPSAR

1245_WP_144994345.1 Paenibacillus sp. N4 68041 68041 TS=8
MVDLGEITSTALVRESAGLQQILYTGNLITYRSTKEDFSSSKYYAPTIITYLYNSKVSIANC
VIRLLLSISNLLMLQTTLSKKRPFSM

1246_AXK31734.1 Streptomyces armeniacus 651322 651322 TS=8
VPAGIPSGQPAGAGGADRAGVRDAAGAPSSAPYAAVAPGTRPASRQAAASAPVSFLRG
LCLRSVRRLVSTACSSTIGGSCGAAA

1247_AXK31734.1 Streptomyces armeniacus 657279 657279 TS=8
LTLFSAREIPSSVTSLIRECFAFPARRGECGGSRPAHSARQAWSPSRSSCS

1248_WP_125043618.1 Streptomyces chrestomyceticus JCM 4735 995917 995917 TS=8
MWASAPPTAPTARRSPGSRTSATTPRASTAPARSSTSSASPASPSRTSGPSTWSASTGAPTPTG

1249_WP_125043618.1 Streptomyces chrestomyceticus JCM 4735 995848 995848 TS=8
VPHCWPPGCRWSRWSCSDRRRRCCGCGCSGSSSGRCRRPSRGRPCARPRCSRW

1250_WP_125043617.1 Streptomyces chrestomyceticus JCM 4735 927648 927648 TS=8
VRSVVSVVSSESGGGGGGGPGPSAVRTSVAPLSVAAWRSSSGAEAAFSGTPTAPSSASAR

1251_WP_069755267.1 Streptomyces sp. EN16 22553 22553 TS=8
VPDSFSGCGSRWDRHSKCSSSRFVSPGGRRTTASVRADGSRSHHSRRG

1252_WP_097871220.1 Streptomyces sp. rh34 1161 1161 TS=8
MSELMWGGVALMGGGLLAANVRGVADRFQAMSYAYRSWPSSVMTCRVIGGVFALVGAGILIAAGL

1253_WP_125542282.1 Streptomyces sp. WAC05292 4845 4845 TS=8
LLSASMAHPFLPAGCPAAASGLASHAPPVVRAAHEWLGGPNGPTLEG

1254_OBZ08797.1 Bacillus sp. FJAT-27264 44139 44139 TS=7
LHILDLNQIPYSTCPLADISSQSVSLSSMKSRDLRVLPTS

1255_WP_171413357.1 Corallococcus exercitus 21824 21824 TS=7
MGADGTPSASLSSAARTRRPRPRSSTQGSVSSASSCLETDEETPVPRSSVQGRASSTSSFAMTDAGMPRLHS

1256_WP_161479596.1 Herbidospora sp. NEAU-GS84 280730 280730 TS=7
VSTTIFWTTTVGSTIQNLSVWAPGFVMVCVVTVRLTLSEVLRSHCLFIRVSSEPVAAMTSTTWGAASSMETAYLRPSVGSTLLRGDR

1257_WP_161479596.1 Herbidospora sp. NEAU-GS84 281144 281144 TS=7
VRPMPGIVFHEPTVLDGRAASRETYPPASTKPFTGPDGFFTTARTTEVEPFGPVTVMGTSVVPRGARDQSTFSVLRTCLSSRPLEGSARNHGTDGLIW

1258_WP_161479596.1 Herbidospora sp. NEAU-GS84 281775 281775 TS=7
LATSRTRSCRRSSSSTTPGSPSRWRGSRTGSAATPPRTSAPARPMRSPRCCTTTRPSSRAASRNCSTG

1259_TDC02021.1 Micromonospora fluostatini 7533 7533 TS=7
MSFMRRPRVRSLLATAAVAGGLVLGLATPAQADYTSPIYPTLKACNDARPKYSSSWTR
PQACYAMYNWDGTKVIGYAFLVKTRS

1260_WP_109900069.1 Micromonospora sp. S4605 25267 25267 TS=7
VARAGEPPPLRCTDARCPARVGACCGQPAGRSRSERIRRKPPCPPPHPTPQLRRFPVFAM

1261_WP_017625650.1 Nocardiopsis chromatogenes YIM 90109 24198 24198 TS=7
VPSMSTLLPTDGPHGTMSLPLTGRSRRVGPEISPSGVRCGGSGRN

1262_WP_017625650.1 Nocardiopsis chromatogenes YIM 90109 26592 26592 TS=7
MPTTGSSSVERAVPGGCGRSPKTRTAGGVPTGAPTGAASAWPSP

1263_WP_054708907.1 Paenibacillus pinihumi DSM 23905 = JCM 16419 487 487 TS=7
LVEHLTFNQGVEGSSPSWLTMFYQTRACGGIGRRTRLRI

1264_WP_054708907.1 Paenibacillus pinihumi DSM 23905 = JCM 16419 5615 5615 TS=7
LILESPISFIGMRAGIFNVHRMCSVRSEDTAYRKSYYGTAYTEVLKHANPQWYCALKLR
FT

1265_WP_054708907.1 Paenibacillus pinihumi DSM 23905 = JCM 16419 27275 27275 TS=7
LLSSNESSVTHTPFFDIIKPLTTVFHVMLMIIVLMPSCQLILILFRNYLFAIHTHYEKMPA

1266_WP_161027165.1 Pseudoduganella sp. DS3 88741 88741 TS=7
MTDKIKSEEVVKAELEISGAIKDLESQKEGFLARLSDDLLEDVNGGSVHLNGHGNVVVT

1267_WP_161027165.1 Pseudoduganella sp. DS3 119409 119409 TS=7
VPSKWRAPSEADLSRRTAQLLRTCFPAQATTLDLMSAIWIIGA

1268_WP_161027165.1 Pseudoduganella sp. DS3 132816 132816 TS=7
VRSGMRDKLTDSKAKCATQKQHLHFCRSSPYNLAPKRTSCSVCEISK

1269_WP_061377893.1 Streptococcus pneumoniae 6632 6632 TS=7
MKLLFFFTILCYYTKFPSSKQCSSDGNNGSSGLNNRSEFCFSIVLDI

1270_WP_061377893.1 Streptococcus pneumoniae 24581 24581 TS=7
MSIQGFMWIDIRRIVIDDFSVLNHDNSHLDNFSLTTCASCL

1271_WP_000817629.1 Streptococcus pneumoniae 106654 106654 TS=7
VPLVALVYLVAISSIQIRERFFSTSSIPCQVSPSGLCTHYFGTIWGSATQN

1272_WP_000817629.1 Streptococcus pneumoniae 119897 119897 TS=7
MENNDSFTKLKESTQKLFDAQKKRLNNEDRIETTKNNVIAKHCQTVLSFLVLTSFFVKN
CVK

1273_WP_000817629.1 Streptococcus pneumoniae 139909 139909 TS=7
LLEIDLTVLIDLSCSYFILIYSFEMVDRTDEVSSKHGFEVVDKEKLMWFEEVFEECKKILV
S

1274_WP_136290219.1 Streptococcus pyogenes 13128 13128 TS=7
MLIYNKLTSFCPGLVTRQTNEENFVFRNLFFIELNIILS

1275_WP_030019169.1 Streptomyces monomycini 93240 93240 TS=7
VPRCWSPGCRWSRWSCSDRRRRCCGCGCSGSSSGRCWRPSRGRPCARPRCSRW

1276_WP_136016861.1 Streptomyces sp. 1AS2c 10303 10303 TS=7
VEDLSLDGDRAAGGSPCDRPQVIVAQPGPVAGGFRHGVRAAAPALPESNPRPVGMSNTI
SASE

1277_WP_097932819.1 Streptomyces sp. rh206 22728 22728 TS=7
MATQRARVGDLPQKGSGLEVTVNEGGNRQHRRIPILQAHVVPDSFSGCGSRRDRHSKC
SSSRFVSPDGRRTTASVRADGSRSHHSRRGRP

1278_WP_076971822.1 Streptomyces sparsogenes DSM 40356 9881 9881 TS=7
VRCGSMTAGTPGPENSCASAVSATALPAGSRYAEGWCKAHRRRWCWPQWQPVAAAK
ERSGPGKGTPRIRKAGVLET

1279_PYR60860.1 Acidobacteria bacterium 23173 23173 TS=6
VPARAPARSRLPWALSKGLSWAFYLSPSAFDPVPSSPCLLPSRQPRPACHGPFHPFGQSP
EHRPASSSSPWPSSRSPGS

1280_WP_171433399.1 Corallococcus exercitus 22474 22474 TS=6
MAVFLEIGWASVRRRPVQRCGGQEHGACHAPHALLEETSSRNCSRAG

1281_WP_120524662.1 Corallococcus exercitus 13257 13257 TS=6
MKTLKRSLLLAACLLGTVGCGVEAPDESAPAQEDTQEVSAQAGCIPGQTRTIKVGCCTP
TLERRQNQICTNSSGSWSNSGSSYCGGSTLNCYGG

1282_WP_154116038.1 Paenibacillus sp. LC-T2 42100 42100 TS=6
LDIPDSCRTGFSVVPQHRCHRSSRFAAYNTTCLDQSVHLRYSSFRTEYAGTSKSAALTVH
TNPVHNDVPSHPDVQE

1283_WP_154116038.1 Paenibacillus sp. LC-T2 63633 63633 TS=6
VNSYGKGRENTGFVVTARNVWTSGRCCPQISFILPPFADEIRGQRRALTLLQFQNFSSFL
QPPSLLCLESSSNHKAPHF

1284_WP_154116038.1 Paenibacillus sp. LC-T2 88986 88986 TS=6
MIESFFFIITISYLLLVIFTTIQLYSNSSSLLRSFIGYSKISPANVTPNRASNIRLR

1285_WP_061745861.1 Streptococcus pneumoniae 8668 8668 TS=6
LLEIDLTVLIDLSCSYFILIYSFEMVDRTDEVSSKHGFEVVDKEKLMWFEEVFEECKKILV
S

1286_WP_087671776.1 Streptococcus pneumoniae 8796 8796 TS=6
LLEIDLTVLIDLSCSYFILIYSFEMVDRTDEVSSKHGFEVVDKEKLMWFEEVFEECKKILV
S

1287_WP_061827610.1 Streptococcus pneumoniae 8691 8691 TS=6
LLEIDLTVLIDLSCSYFILIYSFEMVDRTDEVSSKHGFEVVDKEKLMWFEEVFEECKKILV
S

1288_WP_061760958.1 Streptococcus pneumoniae 8698 8698 TS=6
LLEIDLTVLIDLSCSYFILIYSFEMVDRTDEVSSKHGFEVVDKEKLMWFEEVFEECKKILV
S

1289_VPA84996.1 Streptococcus pneumoniae 38287 38287 TS=6
LLEIDLTVLIDLSCSYFILIYSFEMVDRTDEVSSKHGFEVVDKEKLMWFEEVFEECKKILV
S

1290_VFH36036.1 Streptococcus pneumoniae 307128 307128 TS=6
LLEIDLTVLIDLSCSYFILIYSFEMVDRTDEVSSKHGFEVVDKEKLMWFEEVFEECKKILV
S

1291_VFH36036.1 Streptococcus pneumoniae 320479 320479 TS=6
MNIKKRVLSAGLTFASVLLLAACGQSGSDTKTYSSTFSGNPTTFNYLLDYYTDNIVN

1292_VFH36036.1 Streptococcus pneumoniae 333777 333777 TS=6
MGQLHFITKLLDIKDPNIKILDIINMDTHKEIIAKLDYEAPSCPECGSQLKKYDFQKPSKIP
YLETTGMPL

1293_WP_097867410.1 Streptomyces sp. rh34 10403 10403 TS=6
LRPSTNSEASLHPSMSGLQRSMCSSRNRSGSSSLLANADRH

1294_WP_147469035.1 Lactobacillus reuteri 2273 2273 TS=5
LLEVYRLLITLLDNSSFFVREFIVDFAKFNRRMTKPMFVDLRQPVSIKAFLQYLKNVEWV
VLEKSSPDVFKNLDEKLCEFIYTTNISSKEGEQH

1295_WP_164258878.1 Streptomyces anulatus 24363 24363 TS=5
VDVPDVSAAAGPTSIPIWRDAAPKTHYSHQIFLPRGNRSAGNRATRGPGPARPSRTGSSI
HRTRTASVTCCIGKTPSASGPSRLGTPEPPGTIKPSA

1296_WP_014174828.1 Streptomyces milbemycinicus 3537 3537 TS=5
VDRGGAQQASGEGPAFWPGTGRPPWCVEPGSALDTALRRLEGATDG

1297_WP_053680623.1 Streptomyces sp. WM4235 22005 22005 TS=5
LPYPVSTFTGRAPARANSRKTASSRSSSTGRPNSASLARALPNSCAAYVVFSTARKSRAA
CSWKTGRKCGTDTASRPASSIQASRAAFGLVFA

1298_KOG87603.1 Streptomyces varsoviensis 2490 2490 TS=5
MDAVAAGSAANDRARKPLYVDVGSPPLMLAFERLARDPGSAAVFYEVAPGPETAITDH
KGLPRVTEYVIELNCRGRQGPHGEDGRNGQGGPGRRKGEGK

1299_WP_057090276.1 Enterococcus faecalis 8062 8062 TS=4
MDEVESNCTHLLFLKKGKVVECGKLSDILSKYDTNMNEFYLEKNKGAY

1300_SBO97612.1 Nonomuraea gerenzanensis 7664584 7664584 TS=4
VHVPAVPVCPVGKCARIHRVVLLVSNGGGPAEFGCGVSRDGCGWWTVG

1301_SBO97612.1 Nonomuraea gerenzanensis 7708541 7708541 TS=4
LLDGQVPGAVGRVEAGVAGQQNGASDGGAQGVHSVHGVHGPNATVGKDPI

1302_WP_036835402.1 Pontibacillus litoralis JSM 072002 25452 25452 TS=4
MLPTMFITSATFGRGRRLSIIAIGAFKRSAILRALVTPPWSGDTITTSSKFFSSK

1303_PNY21853.1 Streptococcus parauberis 34 34 TS=4
MQKRYSKEFKETLIAFYHSGQSVTQLSKEYDVAPATIYKWIDLYSKSNESSVSKADFLEL
KRQLAKVKEERDILKKVLTIFAEKKK

1304_RME10422.1 Bacteroidetes bacterium 1064 1064 TS=3
LLPSFYKYPANADARAILHDIFFAIVVPCLSMYQNLDVRKKKKPSRRYTSERL

1305_WP_155203238.1 Cytophagales bacterium RKSG123 1009 1009 TS=3
MIGKAPASIKIEEMTPSSKELLSFNGEDLVSEFVVQWSC

1306_WP_155169938.1 Fulvivirga kasyanovii 7650 7650 TS=3
MGQKIKQYEACRRKSVNLTCYYCIPVCNCFCRLLLTTVVCISPEIPPDH

1307_WP_142380484.1 Streptococcus parauberis 295 295 TS=3
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

1308_WP_164268899.1 Streptomyces anulatus 4461 4461 TS=3
VLYTVEGHCCVTILMLRPLRGDLGGLSAGAKVLSSRRPGPEGPEPPGTAAFAGFRPPAV
RGRRAAGGPAAHDRCQVPSTARCPGPGAAAL

1309_WP_097923571.1 Streptomyces sp. wa1063 7996 7996 TS=3
MLRPLRGDLGELSAGAKVLSSRRPGPEASKPPGTAASAGLRSSTVRGPRHGGRSGGT

1310_WP_107428220.1 Streptomyces pharetrae CZA14 3638 3638 TS=1
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXX

1311_WP_003950320.1 Streptomyces sp. SID4915 3698 3698 TS=1
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

1312_WP_097910655.1 Streptomyces sp. wa1071 3319 3319 TS=1
MSGSAHCSLSGARRGGIVSPAGPRHTTSGGRHLPQRLPIMNGYAA

1313_WP_097910655.1 Streptomyces sp. wa1071 3460 3460 TS=1
MLRPLRGGLGELSAGAKVLSSRRPGPEASKPPGTAASAGLRSSTVRGPRHGGRSGGT

1314_WP_168090536.1 Streptomyces bohaiensis 1654 1654 TS=0
VAAGSPAAPAQAAARTAGGTPRDGGAPAAVRHDAPRSASPARWRDRSTRS

1315_WP_158102173.1 Streptomyces glaucescens 3885 3885 TS=0
LAELFDLGHVMSRVARDRFVARYGEGGTCRVPWDFGADNAAAWEEAGRHLDRPSGD
PGLPTGLTELAALREEFTTLARQAAEP

1316_KJY27547.1 Streptomyces katrae 1465 1465 TS=0
VLVGTHARRAPLFGPLRRPSSVGRRPSSVRRPVAPSSHGRRPVR

1317_KIF07466.1 Streptomyces sp. RSD-27 9 9 TS=0
MAGEPFGQAAQVQGVAQVQEVRFGAAEAGQEVLGGPPGGGGGGRVQEDLFGEVQGG
APAGGGLRRDLAAGGEPSPS

1318_WP_052395034.1 Kutzneria sp. 744 4472101 4472101 TS=15
MVTEKVQYALGDLSMDVFKLTDRGLTVESLTAGHGMAENEASSSSHCSGSCSCGGGSC
SCGDGGWDGGWGGGHGGHGHW

1319_WP_152265679.1 Streptomyces mobaraensis 91538 91538 TS=11
MSTQWDDRFEQVLRRNLPDLPAEDALLPEADLMDLGLDSMGMISLLMDLETEYGVRIP
ESELTFEAFASVAGLWAVVGSVTSAA

1320_WP_152265679.1 Streptomyces mobaraensis 103112 103112 TS=11
LALCSALIARLPFERCDVPLEPFTGPHRRVSRAGPPRARPGRPNGSGAAPRVSPGPLRAL
GGAWCTSATQ

1321_WP_109891469.1 Streptomyces sp. NEAU-S7GS2 5414894 5414894 TS=16
MENTLLDLEQLGDEMFEVVGSESAQGEGDANSAWAYACCQSASSSCALS

1322_WP_018090006.1 Streptomyces sp. SID8375 226983 226983 TS=16
MENTLLDLEQLGDEMFEVVGSESAQGEGDANSAWAYACCQSASSSCALS

1323_WP_089317077.1 Actinomadura mexicana 10573 10573 TS=10
MDKLDLDGLKVESIEVVSLEGLESGHGMTEIDASCGAFCICATISCQAPVAE

1324_WP_089317077.1 Actinomadura mexicana 1389 1389 TS=8
MEFARCLALLGRCLRRLTSGPGRSSGLLSVSPAPMGTSWGSLIKRCSCTS

1325_WP_122195758.1 Actinomadura sp. NEAU-Ht49 17560 17560 TS=15
MSKDFADLSDLPVDSIDVLPADGIEALDIGHGMTEVSASCGHEFTRPCGSCGAPHDGEL
DEF

1326_WP_156046005.1 Herbidospora cretacea221841 221841 TS=14
MDKLPLDLSGLEVEAIDVLSADSIVMEGHGAIETGASSAMGYCSSYMQANSCWAPEA

1327_SEG89460.1 Nonomuraea solani 490488 490488 TS=10
MPPVSTLDFDVHDLPMDVFDLTVSGLEVQTLTLGRGYDIVNGPFTNCHGVSGCVCGCG
SIDKPHADPDDDIDEE

1328_SEG89460.1 Nonomuraea solani 493090 493090 TS=10
MSSAVIWAGGCWEWEAVGAGLPGGEPDAAGVRCRRFRTRAARAEGGAGSRSTSGPGV
RAGRSGSSAP

1329_WP_071967152.1 Streptomyces cinnamoneus 232120 232120 TS=8
MHTGGNAMLPQLDFDALDITNIELPEDEQVFDSLADGSAHSEIGASCFRFCYHCSYVI

1330_WP_051879301.1 Streptomyces sp. NRRL B-24720 9616 9616 TS=11
MAASTPYRLQLVVSTWSHLANGGDPGPIYDHGLGALSGSVGGTFAGSSIRTPVWIC

1331_AFR07166.1 Nocardiopsis alba ATCC BAA-2165 1721193 1721193 TS=11
MADTDKEVLMPHSTAPAIDLPELDFDDMEVMSVRDGGSCG

1332_WP_127937127.1 Nonomuraea polychroma 8873330 8873330 TS=13
MSNDILPLDLSGLEVASIDVLSADSIVVEGHGAIETGASSYSPYCSSYMQACSCWVPEQ

1333_KOT43230.1 Streptomyces caelestis 37726 37726 TS=13
MEQGARSQNVCNLLPVLASSDRRPVAIMKWAGARHTTRPFLARVRCAPLLRTAGGGR
QRDMTVARRDTAGHSHVVRQARVTGPRRS

1334_KJY39143.1 Streptomyces katrae 8228 8228 TS=9
MCQQSRRACGSRPVVGSSRNSSSGRPTTARARSRRRRWPPEREPRTAWGRSARSTSSSA
SRTGSGSE

1335_BAU81447.1 Streptomyces laurentii 524649 524649 TS=11
VVFTGGCPPCSWTCSRTGCRPAWRRRAVRVAPGGERGAGRGASPPGFSYSNP

1336_WP_031510723.1 Streptomyces megasporus 34758 34758 TS=7
VPRGSPRGSRELPKGPRRTRGRLSGHYHSQCPCSFDPGTRGP

1337_WP_157429535.1 Actinomadura oligospora ATCC 43269 262724 262724 TS=8
MSKDFADLNDLPVDSIDVLPADGVERLDIGHGMTEMAASCGHREVGFPHCGSCRAPYD
GEHDES

1338_WP_094862908.1 Amycolatopsis antarctica 309976 309976 TS=12
MLSDELGALEAEAFEVIDIEEPARMLAGTTVTSPSTCAVPTGGSNTENR

1339_WP_034336238.1 Deinococcus misasensis DSM 22328 103796 103796 TS=10
MKNEQAQDIQINLEDLNLQDVETLGESDTYGLPEGGASLSVGRGCGSCSVKPALNEVQ

1340_WP_034336238.1 Deinococcus misasensis DSM 22328 104038 104038 TS=10
MKNEKVQDIQINLEDLNLQDVETLGESDTYGLPEGGASLSVGRGCGSCSVKPALNEVQ

1341_WP_034336238.1 Deinococcus misasensis DSM 22328 104280 104280 TS=10
MKNEKVQDIQINLEDLNLQDVETLGESDTYGLPEGGASLSVGRGCGSCSVKPALNEVQ

1342_WP_026928618.1 Glycomyces tenuis 308895 308895 TS=12
MKKSSSPVSMHIPFGSGVNRWLPPLLSCASSMSQRTSYSSPS

1343_WP_067368384.1 Micromonospora rosaria 11581 11581 TS=6
MDNVVTEAAEFADLDIVDLDLAVDEELAALSVGGLGNTEVGASGWLGSWVI

1344_WP_067368384.1 Micromonospora rosaria 11870 11870 TS=6
MDNAATEATEFADLDIVNLDLPIDEELAAVSVGGLGNTEVGASGFFGRSWLI

1345_WP_040691311.1 Nocardiopsis lucentensis DSM 44048 7509 7509 TS=11
MEPLPSGVRRGVGVTLAGGVAAARHRLPGSPGAALGRVVQVRSGSRGRACSYARTCT
NT

1346_WP_125051747.1 Streptomyces rimosus subsp. paromomycinus 953410 953410 TS=7
MFVLTDPSPQAARPVAPSLRNAVCRARTSTGSPSAVPVPCASM

1347_WP_125051747.1 Streptomyces rimosus subsp. paromomycinus 953932 953932 TS=6
MASTARWTATSELEQAVSTGSLGPRRSSTYEMRLASIEWALPVEVNASGSTPYVVSSVA

1348_WP_138054944.1 Streptomyces sp. ICN19 5266269 5266269 TS=12
MNDITPASGPVLDLRLDQEAPELEVLPAAHSPGSTVGCAGTGSSISTPSGCFSSAGTASSV

1349_WP_138054944.1 Streptomyces sp. ICN19 5265428 5265428 TS=10
MNELDLRIEQEHPEIEVLPDTRSAGSCAGSASTAGSFSCPAGSIGSVGSASSAG

1350_WP_138054944.1 Streptomyces sp. ICN19 5265674 5265674 TS=10
MNDLDLRLDQEMPELEVLPEEYSAGNVSCAATVGSGSCPAMTWGTASTLSSK

1351_WP_138054944.1 Streptomyces sp. ICN19 5265953 5265953 TS=10
MDRLPDTDVLELVLERELPELEVLPAGYAPGSTLGSAGSISCASCPAASISSASTASSH

1352_WP_109293928.1 Streptomyces spongiicola 1932652 1932652 TS=12
MNELDLRIEQESLEIEVLPDTQSAGNCAGTASTAGSFSCPAGSIGSAGSASST

1353_WP_109293928.1 Streptomyces spongiicola 1932405 1932405 TS=11
MNELDLRLDQEMPELEVLPEEYSAGNVSSAATVGSGSCPAMTWGTASTLSSK

1354_WP_098777088.1 Bacillus cereus 13343 13343 TS=9
MGDTQKGTLRVRKVPEHVLLHSPYAGIIQQVQMVNDGLVVILSVCIHKLPCVYYSSTSK
STLILVK

1355_WP_009582377.1 Fulvivirga imtechensis AK7 12043 12043 TS=8
MKQKKLKLEALKLKSFITTEERLKAQTIKGGAPISAYGNCEPLPYSELCIVTCLCSNTCM
TDLCGDTGGGASGICDGTTTQQQK

1356_WP_009582377.1 Fulvivirga imtechensis AK7 11539 11539 TS=7
MKKNKLNLSELRVNSFVTDADAVNQETIKGGVIDSCVPECGGTGFPTIPASPLCASHNSC
QLLCDGDVMTY

1357_WP_009582377.1 Fulvivirga imtechensis AK7 8355 8355 TS=6
MKKKKLSLNKLKIESFITTDVLRQLHAGNSIADRRSDSDCPKNHSDQSFCSCDSCVSCTC
HTEFICTNNISDCYSIQPCDGQPY

1358_WP_009582377.1 Fulvivirga imtechensis AK7 37092 37092 TS=7
VSIRKTYGDPSLRTNDKFSLHADSQMTASYCSVCNFCVYCDFS

1359_WP_027734140.1 Streptomyces sp. CNR698 46020 46020 TS=11
MNELDLRIEQENPEVEVLPDTQSAGSCAGSAGTAGSLSCPGGSIGTVGTASSSG

1360_WP_027734140.1 Streptomyces sp. CNR698 45761 45761 TS=10
MNDLDLRLDQEMPELEVLPEEYSAGNVSCAATVGSGSCPAMTWGCASTLSSK

1361_WP_027734140.1 Streptomyces sp. CNR698 28121 28121 TS=6
MRSLVAASNSPSSARGTIRRAGAWSTVAPRRCSSAPSSSARRAAVTPTVKPESGATAGA
PCCSWYMPTEPGKEA

1362_WP_018891562.1 Streptomyces sp. CNT302 63388 63388 TS=11
MNELDLRIEQESPEVEVLPDTQSAGSCAGSAGTAGSLSCPGGSIGTVGTASSSG

1363_WP_018891562.1 Streptomyces sp. CNT302 63647 63647 TS=10
MNDLDLRLDQEMPELEVLPEEYSAGNVSCAATVGSGSCPAMTWGCASTLSSK

1364_WP_018891562.1 Streptomyces sp. CNT302 81145 81145 TS=6
MRSLVAASNSPSSARGTIRRAGAWSTVAPRRCSSAPSSSARRAAVTPTVKPESGATAGA
PCCSWYMPTEPGKEA

1365_WP_170812553.1 Streptomyces sp. PKU-MA01144 10032 10032 TS=11
MNELDLRIEQESPEVEVLPDTQSAGSCAGSAGTAGSLSCPGGSIGTVGTASSSG

1366_WP_170812553.1 Streptomyces sp. PKU-MA01144 10291 10291 TS=10
MNDLDLRLDQEMPELEVLPEEYSAGNVSCAATVGSGSCPAMTWGCASTLSSK

1367_WP_170812553.1 Streptomyces sp. PKU-MA01144 28336 28336 TS=6
MRSLVAASNSPSSARGTIRRAGAWSTVAPRRCSSAPSSSARRAAVTPTVKPESGATAGA
PCCSWYMPTEPGKEA

1368_WP_129842995.1 Streptomyces sp. RFCAC02 5116202 5116202 TS=11
MSDHEFTLDLDGLDVDTIDVLPADSVMTEGHGAIETGASCSPGYCSCYSTAMSCFVPEEI

1369_WP_119313826.1 Meiothermus terrae 6888 6888 TS=10
MDNTKNTPKPEDIQADFNLEDLDLNDVKVLAESDSYALPEAGASIGKFSCTIVITK

1370_SCL48150.1 Micromonospora yangpuensis 811256 811256 TS=5
MEDMAVEATEFADLDIVDLDLPIDEELAAVSIGGLGNTEVGASGFWGRSWLI

1371_WP_149666213.1 Pseudonocardia sp. EV170527-09 304 304 TS=6
VLVACCQAASSVLRVSMSVMRRSGHWRARVESSISAMLSQEPCRA

1372_PIG63282.1 Streptomyces sp. CNZ279 1482235 1482235 TS=10
MTAPSDGPHAPRGPVAVDSEQPLARRQVALVDLLDRLLGTGAVLTGSLTLGIADVDLV
RIDLRALISSVNARVPSPWEGGQPL

1373_WP_161371729.1 Streptomyces sp. SID5770 9893 9893 TS=8
MGITRILSVKVGKIRPEVQKSRSTGTNAPARRLREETRGRARGAGSPPVPAPDLAISLA

1374_WP_028544504.1 Paenibacillus taiwanensis DSM 18679 185530 185530 TS=6
MDLEMVLEVSLTRASKGFVIIVDKVIMLDFGSSTCSDSPQNYCGNYGA

1375_WP_121832244.1 Streptomyces sp. S1 9925 9925 TS=8
MGITRILSVKVGKIRPEVQKSRSTGTNAPARRLREETRGRARGAGSPPVPAPDLAISLA

1376_KIC48916.1 Tateyamaria sp. ANG-S1 703427 703427 TS=7
VFVFLNAAAFVENLPQPTRGENSKTDAKSHPGCWSVARWRVTSCPKT

1377_WP_084467419.1 Actinokineospora inagensis DSM 44258 8148 8148 TS=7
LAGFRATATGLRANPAGRTPQVMAWLCRSVRPTAASSRCSTRSTPAYCLLWTFRRCVE
QVRQASSSCGPLWSVARYPASTSGWSGSAAGTRSRSPATR

1378_WP_122194823.1 Actinomadura sp. NEAU-Ht49 40673 40673 TS=5
MGMLYSATATSANAPCPPDTGPAPPAAARGEPQEPAHRCTGRFRATGHSRTTAGHSSA

1379_RYZ37930.1 Myxococcaceae bacterium 17111 17111 TS=6
MAKIHRQPTESTSSPPTSGPAMNAPPVQPVHEPIALACSGPEKKDRMSASEPGTSSAPVA
P

1380_WP_054708907.1 Paenibacillus pinihumi DSM 23905 = JCM 16419 15487 15487 TS=5
MDVGCLHLTARFQHIWEVLEAGLRGQLDGDGNSGCKSGCSEAKTSIINN

1381_WP_066415890.1 Bacillus cohnii 1461919 1461919 TS=19
MLINYLKGGDHMAQLTNVEEVELEIFEVEEATVMEEMGASIGKLGCCSSIIIR

1382_WP_066415890.1 Bacillus cohnii 1462100 1462100 TS=19
MNMEQNMEEVELEIFEVEEATVMEEMGASIGKLGCCSSIIIR

1383_WP_066415890.1 Bacillus cohnii 1462281 1462281 TS=19
MNVEQSMEEVELEIFEVEEATVMEEMGASIGKLGCCSSIIIR

1384_WP_088248048.1 Deinococcus indicus 115228 115228 TS=10
MNEQTTTATAEVIELNDLHLEDVSVAVEEDAYALPEAGASVSPKYSCTVVIER

1385_WP_088248048.1 Deinococcus indicus 115445 115445 TS=10
MNEQNTTATSEVIELNDLHLEDVSVAVEEDAYALPEAGASASPKYSCTVVIKPN

1386_WP_088248048.1 Deinococcus indicus 115888 115888 TS=9
MNEPTTTAIPEVIELNDLHLEDVSVAVEEDAYDLPEAGASASPRFSCSIIIEL

1387_WP_088248048.1 Deinococcus indicus 116105 116105 TS=8
MNEPTTTAIPEALELNDLHLEDVSVAVEEDAYDLPEAGASLSPRFSCSIIIARE

1388_WP_099749137.1 Deinococcus sp. UR1 74633 74633 TS=19
MNEQTTTAIPEALELNDLHLEDVSVAVEEDAYDLPEAGASASPRFSCSIIIQL

1389_WP_099749137.1 Deinococcus sp. UR1 75301 75301 TS=10
MNEQTTTATAEALELNDLHLEDVSVAVEEDAYALPEAGASVSPKYSCTVVIEK

1390_WP_099749137.1 Deinococcus sp. UR1 74847 74847 TS=8
MNEQTKAAIPEALELNDLHLEDVSVAVEEDAYDLPEAGASVSPRFSCSIVIAKE

1391_WP_146884718.1 Deinococcus cellulosilyticus NBRC 106333 = KACC 11606 79451 79451 TS=19
MSHDKKQEELNIQDLDLNEVEVTTEDDSYALPEAGASIGYNSCSSVKPAAK

1392_WP_146884718.1 Deinococcus cellulosilyticus NBRC 106333 = KACC 11606
79630 79630 TS=19
MSHEKKQEELNIQDLDLNEVEVTTEDDSYALPEAGASIGYNSCSSVKLK

1393_WP_146884718.1 Deinococcus cellulosilyticus NBRC 106333 = KACC 11606
79807 79807 TS=19
MSHEKKQEELNIQDLDLNEVEVTTEDDSYALPEAGASIGYNSCSSVKPK