UNIVERSITY OF CALIFORNIA,
IRVINE

X-ray Crystallographic Studies of Oligomers Derived from Amyloidogenic Peptides and Proteins

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by

Ryan K. Spencer

Dissertation Committee:
Professor James S. Nowick, Chair
Professor Kenneth Shea
Professor Andrej Lupták

2015

DEDICATION

To the memory of my grandfather Robert Spencer. Thank you for believing in me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

ACKNOWLEDGMENTS

I thank Prof. James S. Nowick for being such a wonderful and patient mentor to me during my Ph.D. studies. He provided unmatched support, guidance, and encouragement for the projects that I present in this dissertation. He fostered an environment of thoughtful discussion, exploration, and creativity that helped these projects flourish.

During the developement of these projects I was lucky to work with two outstanding undergraduate students, Arya Baghkhanian and Hao Li. Arya suffered with me through the difficult development of the *N*-methylated macrocycles. Hao helped develop the Aβ derived macrocycles, which ultimately turned into a publication in *JACS*, and with the β_2m macrocycles.

I thank Profs. Thomas Poulos and Celia Goulding and Drs. Nicholas Chim and Huiying Li for their advice and guidance in X-ray crystallography. I thank Dr. Phil Dennison for providing support for the NMR experiments and John Greaves for providing support for the mass spectrum experiments. I would also like to thank all of the professors and staff in the UCI Department of Chemistry for always being helpful and providing an unparalleled learning environment. I would also like to thank the National Institute of Health (NIH) and the National Science Foundation (NSF) for supporting these projects.

I would not have been able to accomplish what I have accomplished without the love and support of my family and the Maida family. They continuously provided a stable foundation that I am extremely grateful for.

I would not be who I am, as a man and as a scientist, without the love, patience, and support of my fiancée, Diana Maida. I have grown and learned so much over this period of time because of her and I will forever be indebted to her.

CURRICULUM VITAE

Ryan K. Spencer

EDUCATION

Doctor of Philosophy in Chemistry **2015**
University of California, Irvine *Irvine, California*

Master of Science in Biochemistry **2010**
California State University, Northridge *Northridge, California*

Bachelor of Science in Biochemistry **2005**
University of California, San Diego *San Diego, Californiai*

Bachelor of Arts in Political Science **2005**
University of California *San Diego, California*

RESEARCH EXPERIENCE

Graduate Research Assistant **2010–2015**
University of California, Irvine *Irvine, California*

Graduate Research Assistant **2007–2010**
California State University, Northridge *Northridge, California*

TEACHING EXPERIENCE

Teaching Assistant - Organic Spectroscopy **2011–2014**
University of California, Irvine *Irvine, California*

Teaching Assistant - General/Organic Chemistry **2007–2010**
California State University, Northridge *Northridge, California*

HONORS AND AWARDS

Reagent's Dissertation Fellowship **2014**
University of California, Irvine

Harold W. Moore Award **2014**
University of California, Irvine

Outstanding Graduate Fellowship Award **2009**
California State University, Northridge

Outstanding Graduate Student Teaching Assistance Award **2009**
California State University, Northridge

Sandra L. Jewett Scholarship in Biochemistry **2008**
California State University, Northridge

PUBLICATIONS

"X-ray Crystallographic Structures of Oligomers of Peptides Derived from β_2-Microglobulin." Spencer, R. K.; Kreutzer, A. G.; Salveson, P. J.; Li, H.; Nowick, J. S. *submitted to J. Am. Chem. Soc.* **2015**

"A Newcomer's Guide to Peptide Crystallography." Spencer, R. K.; Nowick, J. S. *accepted Israel J. Org. Chem.* **2015**

"A Fibril-Like Assembly of Oligomers of a Peptide Derived from β-Amyloid." Pham, J. D.; Spencer, R. K.; Chen, K. H.; Nowick, J. S. *J. Am. Chem. Soc. 136*: 12682–12690. **2014**

"X-ray Crystallographic Structures of Trimers and Higher-Order Oligomeric Assemblies of a Peptide Derived from $A\beta_{17–36}$." Spencer, R. K.; Li, H.; Nowick, J. S. *J. Am. Chem. Soc. 136*: 5595–5598. **2014**

"Recipe for β-Sheet: Foldamers Containing Amyloidogenic Peptide Sequences." Spencer, R.; Chen, K.; Manuel, G.; Nowick, J. S. *Eur. J. Org. Chem. 2013*: 3523–3528. **2013**

"Heterodivalent Linked Macrocyclic β-Sheets with Enhanced Activity against Aβ Aggregation: Two Sites Are Better Than One." Cheng, P.-N.; Spencer, R.; Woods, J. R.; Glabe, C. G.; Nowick. J. S. *J. Am. Chem. Soc. 134*: 14179–14184. **2012**

"Carbon Tether Rigidity as a Stereochemical Tool Directing Intramolecular Radical Cyclizations." Melikyan, G. G.; Voorhees, E.; Wild, C.; Spencer, R.; Molnar, J. *Tetrahedron Lett. 51*: 2287–2290. **2010**

"Inter- vs intramolecular couplings of cobalt-complexed propargyl radicals: Questioning the consensus." Melikyan, G. G.; Spencer, R. *Tetrahedron 66*: 5321–5328. **2010**

"1,3-Steric Induction in Intermolecular Radical Coupling Reactions." Melikyan, G. G.; Spencer, R. *Organometallics 29*: 3556–3562. **2010**

"Cobalt-Complexed Propargyl Cations: Generation under Neutral Conditions and Spontaneous, High Temperature Conversion to Propargyl Radicals." Melikyan, G. G.; Sepanian, R.; Spencer, R.; Rowe, A.; Toure, P. *Organometallics 28*: 5541–5549. **2009**

"Selective Synthesis of *meso*-1,5-Cyclodecadiynes" Melikyan, G. G.; Spencer, R.; Abedi, E. *J. Org. Chem. 74*: 8541–8546. **2009**

PRESENTATIONS

X-ray Crystallographic Studies of Oligomers Derived from Amyloidogenic Peptides and Proteins. **2014**
Bristol-Meyers Squibb (BMS) Symposium *Irvine, California*

Amyloid Origami: Supramolecular Assemblies of Amyloidogenic Peptides. **2014**
ACS Division of Organic Chemistry Graduate Research Symposium *Irvine, California*

Crystallographic Structures of Trimers and Higher-Order Oligomeric Assemblies of a Peptide Derived from $A\beta_{17-36}$. **2014**
Peptide Gordon Research Conference (GRC) *Ventura, California*

New Class of β-Sheet Macrocycles. **2013**
American Chemical Society (ACS) National Meeting *New Orleans, Louisiana*

Inter- and intramolecular radical C-C bond formation mediated by a transition metal core. **2010**
Thesis Defense, Department of Chemistry & Biochemistry, CSUN *Northridge, California*

Stereoselective synthesis of *meso*-1,5 cyclodecadiynes. **2010**
CSUN Student Research Symposium *Northridge, California*

Radical carbon-carbon bond formation stereodirected by an alpha-metal core. **2009**
Organometallics Gordon Research Conference (GRC) *Newport, Rhode Island*

Controlling stereoselectivity of cobalt-mediated inter- and intramolecular coupling reactions. **2009**
9^{th} Annual Sigma Xi Symposium *Northridge, California*

Radical carbon-carbon bond formation stereodirected by an alpha-metal core. **2009**
CSUN Student Research Symposium *Northridge, California*

Intramolecular radical reactions enabled by transition metals. **2009**
National American Chemical Society (ACS) Meeting *Washington, D.C.*

Impact of aromatic substitution and reducing agent on diastereoselectivity of intra- and intermolecular cobalt mediated radical reaction. **2008**

Western Region American Chemical Society (ACS) Meeting *Las Vega, Nevada*

PROFESSIONAL MEMBERSHIPS

American Chemical Society (ACS)

ABSTRACT OF THE DISSERTATION

X-ray Crystallographic Studies of Oligomers Derived from
Amyloidogenic Peptides and Proteins

By

Ryan K. Spencer

Doctor of Philosophy in Chemistry

University of California, Irvine, 2015

Professor James S. Nowick, Chair

This dissertation describes a new class of macrocylic peptides which I invented as a tool for understanding the folding of β-sheets and the structures of amyloid oligomers. I designed these macrocylic peptides to incorporate various sequences from amyloidogenic peptides and proteins, to fold to adopt β-sheets, and to further assemble into oligomers. This dissertation details of their formation of β-sheets by NMR spectroscopy and oligomer structures determined by X-ray crystallography.

In Chapter 2, I describe the development of a new class of macrocyclic β-sheets that contains an *N*-methyl amino acid and an amyloidogenic pentapeptide sequence from Aβ, tau, the B chain of insulin, and prion protein. These *N*-methylated macrocycles fold to adopt β-sheets in solution. In Chapter 3, I expand these *N*-methylated β-sheet peptides to incorporate heptapeptide sequences from β-amyloid ($A\beta_{17\text{-}36}$). I used these peptides to study the structures of the Aβ oligomers using X-ray crystallography. In Chapter 4, I explain the techniques that I adopted from protein crystallography to solve the X-ray crystallographic structures of these β-sheet peptides. In Chapter 5, I describe a range of oligomer structures that are accessible to a macrocycle containing a sequence from β_2-microglobulin ($\beta_2m_{63\text{-}69}$) and various *N*-methyl amino acids. The structures formed by these β-sheet peptides of-

fer insights into the structures formed by amyloidogenic peptides and proteins in amyloid diseases.

Amyloid diseases such as Alzheimer's disease, Parkinson's disease, and type II diabetes share common features of toxic soluble protein oligomers. There are no structures at atomic resolution of oligomers formed by full-length amyloidogenic peptides and proteins, and only a few structures of oligomers formed by peptide fragments. The paucity of structural information provides a fundamental roadblock to understanding the pathology of amyloid diseases and developing preventions or therapies. Here, I describe the developement of a new class of β-sheet macrocycles to study the structures of oligomers formed by amyloidogenic peptides and proteins. Macrocycles containing heptapepeptide sequences from $A\beta_{17\text{-}36}$ fold to adopt β-sheets and associate further to form triangular trimers and higher-ordered oligomers. These triangular trimers are unprecedented and represent a new motif that full-length amyloidogenic peptides and proteins may adopt.

Macrocyclic peptides derived from $\beta_2m_{63\text{-}69}$ form β-sheets that further associate into hexamers, octamers, and dodecamers: the hexamers are trimers of dimers; the octamers are tetramers of dimers; and the dodecamers contain two trimer subunits surrounded by three pairs of β-sheets. These structures illustrate a common theme in which dimer and trimer subunits further associate to form a hydrophobic core. The seven X-ray crystallographic structures not only illustrate a range of oligomers that a single amyloidogenic peptide sequence can form, but also how mutation can alter the size and topology of the oligomers. A cocrystallization experiment in which a dodecamer-forming peptide recruits a hexamer-forming peptide to form mixed dodecamers demonstrates that one species can dictate the oligomerization of another. These findings should also be relevant to the formation of oligomers of full-length peptides and proteins in amyloid diseases.

Chapter 1

Introduction

Chemical model systems that mimic β-sheets are important because β-sheets are key structural components in many proteins, are involved in protein-protein interactions, and are also a major structural conformation in many amyloid diseases.[1–3] In amyloid disease such as Alzheimer's, Parkinson's, Creutzfeldt-Jacob, and type II diabetes, proteins and peptides adopt β-sheet conformations that further aggregate to form soluble oligomers and insoluble fibrils.[3–16] Research had focused on the formation and structure of the insoluble fibrils because they were believed to be the causative components of many of these diseases, however it appears that the soluble oligomers, formed by the aggregation of a few peptides and proteins, are responsible for these diseases.[5,7,17–32] Despite the intense focus over the last decade on these soluble oligomers, very little information has been obtained at atomic resolution. My research has focused on using the chemical model systems I invented to gain insight into the structure of these oligomers at atomic resolution.

The macrocycles I created evolved gradually from a project that focused on the development of macrocyclic peptides to inhibit the formation of amyloid-β (Aβ) oligomers and fibrils. Over the last decade, the Nowick group developed macrocyclic peptides that incor-

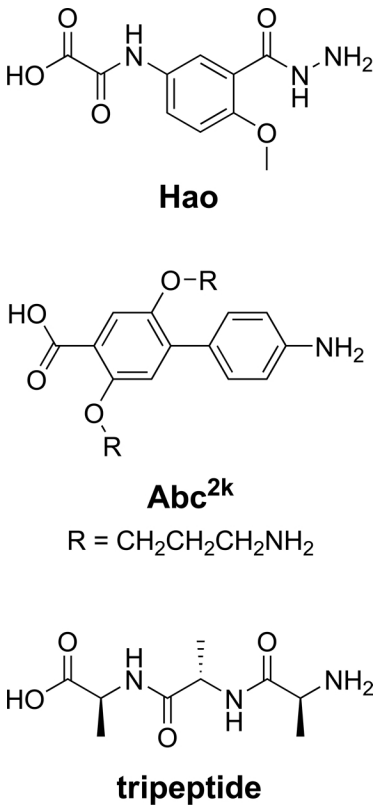
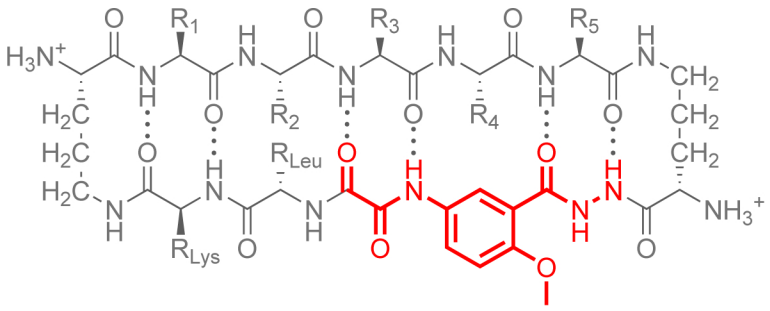


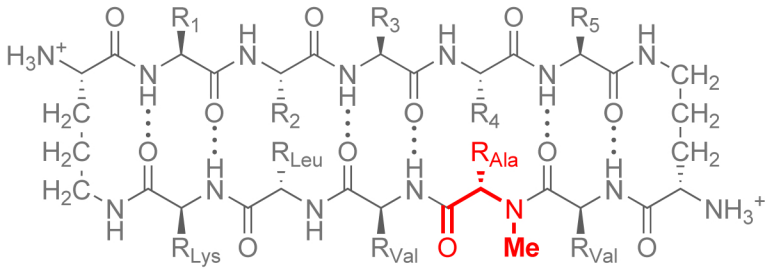
Figure 1.1: Chemical drawing of *Hao*, Abc^{2K}, and a tripeptide.

porate amyloidogenic sequences and fold to adopt β-sheets.[3,33–39] These macrocycles contain a peptide sequence from various amyloidogenic peptides and proteins and a second sequence that acts as a template strand to reinforce β-sheet formation. The template strand utilized a tripeptide mimic, *Hao*, which helps reinforce β-sheet formation and block aggregation.[40] I began my research by challenging the importance of using the *Hao* tripeptide to help stabilize β-sheet formation. I replaced the *Hao* tripeptide with Abc^{2K} (amino biphenyl carboxylic acid)[41] and three natural amino acids (Figure 1.1) and determined their impact on β-sheet formation.

The incorporation of the Abc^{2K} linker into a macrocyclic scaffold was straightforward but resulted in poorly folded β-sheet macrocycles. The incorporation of a simple triple alanine amino acid sequence instead of *Hao* proved to be troublesome. I was able to synthesize the linear peptide and the cyclic peptide but upon deprotection and isolation, the macrocycle



Hao containing macrocycle



***N*-methyl amino acid** containing macrocycle

Figure 1.2: Chemical drawing of a macrocycle containing *Hao* and a macrocycle containing an *N*-methyl alanine.

aggregated and could not be isolated. This was not surprising because β-sheets containing amyloidogenic sequences are prone to aggregate. β-Sheets can aggregate either along the hydrogen bonding edges of the amide backbone or through hydrophobic contacts of the side chains of the amino acids. Macrocycles containing the triple alanine could aggregate along either exposed hydrogen bonding edge and through hydrophobic contacts among the side chains of the amino acids. I incorporated a single *N*-methyl amino acid along the template strand to prevent association along a single hydrogen bonding edge in the hope that it would prevent aggregation and permit isolation. The methyl group on the *N*-methyl amino acid removes a hydrogen bond donor, provides steric bulk, and interrupts hydrogen bonding with the C=O of the *N*-methyl amino acid. This innocuous modification was enough to prevent aggregation and allowed me to isolate considerable quantities, often in the tens of miligrams. Figure 1.2 shows a macrocycle containing *Hao* and an analogous macrocycle containing an *N*-methyl alanine.

I was surprised that macrocycles containing a single *N*-methyl amino acid were well-behaved and well-folded, even better than similar macrocycles containing *Hao*. I began exploring the limitations of these *N*-methylated macrocycles by incorporating sequences from various amyloidogenic peptides and proteins and studying their ability to fold into β-sheets. In Chapter 2, I describe the developement of macrocycles that contain sequences from Aβ , tau, the B chain of insulin, and human prion protein (hPrP) and an *N*-methyl alanine in the template strand.[42] These macrocycles do not aggregate in solution and fold to adopt β-sheets.

Using *N*-methyl amino acids in macrocyclic peptides not only offered all of the benefits of the tripeptide mimic *Hao* (templating β-sheet formation and blocking aggregation) but also allowed for the reincorporation of the side chains of three amino acids that was not previously possible in macrocycles that contained *Hao*. This is important when studying oligomers formed by amyloidogenic peptides and proteins because the β-sheets formed by amyloidogenic peptides and proteins contain amino acids with side chains which have a large impact on aggregation and on the structure of the oligomers that are formed.

My focus shifted from challenging the necessity of *Hao* to create well folded β-sheets, to designing and utilizing *N*-methylated macrocycles to study the structures of oligomers formed by amyloidogenic peptides and proteins. Amyloidogenic peptides and proteins aggregate to form soluble oligomers comprising a few peptides or protein monomers. These peptides and proteins often fold to adopt β-sheets that then associate further through hydrogen-bonding and hydrophobic interactions to form soluble oligomers. I used the *N*-methylated macrocycles to mimic the β-sheets formed by amyloidogenic peptides and proteins and then asked the question: What structures form when these β-sheets associate further?

I wanted to study these oligomer structures at atomic resolution using X-ray crystallography. The Nowick group, in collaboration with the Eisenberg and Goulding groups, had success crystallizing a handful of *Hao* containing macrocyclic peptides. I adopted the crystal-

lization techniques employed by these laboratories to crystallize *N*-methylated macrocyclic peptides. I wanted to create a fast method for designing, synthesizing, crystallizing, and solving the X-ray crystallographic structures of peptides containing amyloidogenic sequences. I developed methods for quickly incorporating any *N*-methyl amino acid into a macrocycle but I had yet to develop a method for crystallizing and solving the X-ray crystallographic structures of oligomers formed by macrocyclic peptides. Solving the structures of peptide crystals offered some unique challenges, in particular, determining the phases of the electron density map. I detail how I grew peptide crystals and determined the X-ray crystallographic phases in Chapter 4. I relied heavily on incorporating an *p*-iodophenylalanine as an innocuous surrogate for phenylalanine and tyrosine residues within amyloidogenic sequences. I utilized the anomalous edge of the iodine on an in-house X-ray diffractometer with a Cu rotating anode to collect diffraction data and solve the crystal structures rapidly. This process became so efficient that from an initial peptide design to a solved crystal structure took only a few weeks. I used these techniques to study the X-ray crystallographic structures of an oligomer derived from $A\beta_{17\text{-}36}$ (Chapter 3)[43] and a range of oligomers derived from β_2m (Chapter 5).

This new class of *N*-methylated macrocyclic peptides are easy to synthesize, incorporate multiple amyloidogenic sequences, and crystallize easily. These macrocycles are valuable tools for studying the structure of oligomers formed by amyloidogenic peptides and proteins and provide unprecedented insight into the structures of these oligomers at atomic resolution. These structure may serve as a basis for designing drugs and therapies for preventing and curing Alzheimer's disease, Parkinson's disease, type II diabetes, prion diseases, and many other amyloidogenic diseases. I believe these *N*-methylated macrocycles will continue to provide insight into the structures of oligomers formed by amyloidogenic peptides and proteins.

References

1. Cerf, E.; Sarroukh, R.; Tamamizu-Kato, S.; Breydo, L.; Derclaye, S.; Dufrne, Y. F.; Narayanaswami, V.; Goormaghtigh, E.; Ruysschaert, J. M.; Raussens, V. *Biochem. J.* **2009**, *421*, 415–423.

2. Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T. *Nat. Chem.* **2011**, *3*, 172–177.

3. Buchanan, L. E.; Dunkelberger, E. B.; Tran, H. Q.; Cheng, P. N.; Chiu, C. C.; Cao, P.; Raleigh, D. P.; de Pablo, J. J.; Nowick, J. S.; Zanni, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 19285–19290.

4. Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16742–16747.

5. Walsh, D. M.; Klyubin, I.; Fadeeva, J. V.; Cullen, W. K.; Anwyl, R.; Wolfe, M. S.; Rowan, M. J.; Selkoe, D. J. *Nature* **2002**, *416*, 535–539.

6. Lührs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Dbeli, H.; Schubert, D.; Riek, R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17342–17347.

7. Haass, C.; Selkoe, D. J. *Nat. Rev. Mol. Cell. Biol.* **2007**, *8*, 101–112.

8. Paravastu, A. K.; Leapman, R. D.; Yau, W. M.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18349–18354.

9. Gurlo, T.; Ryazantsev, S.; Huang, C. J.; Yeh, M. W.; Reber, H. A.; Hines, O. J.; O'Brien, T. D.; Glabe, C. G.; Butler, P. C. *Am. J. Pathol.* **2010**, *176*, 861–869.

10. Smith, D. P.; Radford, S. E.; Ashcroft, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 6794–6798.

11. Bertini, I.; Gonnelli, L.; Luchinat, C.; Mao, J.; Nesi, A. *J. Am. Chem. Soc.* **2011**, *133*, 16013–16022.

12. Lee, J.; Culyba, E. K.; Powers, E. T.; Kelly, J. W. *Nat. Chem. Biol.* **2011**, *7*, 602–609.

13. Celej, M. S.; Sarroukh, R.; Goormaghtigh, E.; Fidelio, G. D.; Ruysschaert, J. M.; Raussens, V. *Biochem. J.* **2012**, *443*, 719–726.

14. Fändrich, M. *J. Mol. Biol.* **2012**, *421*, 427–440.

15. Hamley, I. W. *Chem. Rev.* **2012**, *112*, 5147–5192.

16. Tay, W. M.; Huang, D.; Rosenberry, T. L.; Paravastu, A. K. *J. Mol. Biol.* **2013**, *425*, 2494–2508.

17. Chromy, B. A.; Nowak, R. J.; Lambert, M. P.; Viola, K. L.; Chang, L.; Velasco, P. T.; Jones, B. W.; Fernandez, S. J.; Lacor, P. N.; Horowitz, P.; Finch, C. E.; Krafft, G. A.; Klein, W. L. *Biochemistry* **2003**, *42*, 12749–12760.

18. Kayed, R.; Head, E.; Thompson, J. L.; McIntire, T. M.; Milton, S. C.; Cotman, C. W.; Glabe, C. G. *Science* **2003**, *300*, 486–489.

19. Eakin, C. M.; Attenello, F. J.; Morgan, C. J.; Miranker, A. D. *Biochemistry* **2004**, *43*, 7808–7815.

20. Hoyer, W.; Cherny, D.; Subramaniam, V.; Jovin, T. M. *J. Mol. Biol.* **2004**, *340*, 127–139.

21. Ivanova, M. I.; Sawaya, M. R.; Gingery, M.; Attinger, A.; Eisenberg, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10584–10589.

22. Bitan, G.; Fradinger, E. A.; Spring, S. M.; Teplow, D. B. *Amyloid* **2005**, *12*, 88–95.

23. Zhao, W.-Q.; De Felice, F. G.; Fernandez, S.; Chen, H.; Lambert, M. P.; Quon, M. J.; Krafft, G. A.; Klein, W. L. *FASEB J.* **2008**, *22*, 246–260.

24. Bernstein, S. L.; Dupuis, N. F.; Lazo, N. D.; Wyttenbach, T.; Condron, M. M.; Bitan, G.; Teplow, D. B.; Shea, J. E.; Ruotolo, B. T.; Robinson, C. V.; Bowers, M. T. *Nat. Chem.* **2009**, *1*, 326–331.

25. Schmidt, M.; Sachse, C.; Richter, W.; Xu, C.; Fandrich, M.; Grigorieff, N. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 19813–19818.

26. Yankner, B. A.; Lu, T. *J. Biol. Chem.* **2009**, *284*, 4755–4759.

27. Yu, L. *et al. Biochemistry* **2009**, *48*, 1870–1877.

28. Ahmed, M.; Davis, J.; Aucoin, D.; Sato, T.; Ahuja, S.; Aimoto, S.; Elliott, J. I.; Van Nostrand, W. E.; Smith, S. O. *Nat. Struct. Mol. Biol.* **2010**, *17*, 561–567.

29. Sakono, M.; Zako, T. *FEBS J.* **2010**, *277*, 1348–1358.

30. Benilova, I.; Karran, E.; De Strooper, B. *Nat. Neurosci.* **2012**, *15*, 349–357.

31. Larson, M. E.; Lesne, S. E. *J. Neurochem.* **2012**, *120 Suppl 1*, 125–139.

32. Stroud, J. C.; Liu, C.; Teng, P. K.; Eisenberg, D. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 7717–7722.

33. Liu, C.; Sawaya, M. R.; Cheng, P.-N.; Zheng, J.; Nowick, J. S.; Eisenberg, D. *J. Am. Chem. Soc.* **2011**, *133*, 6736–6744.

34. Zheng, J.; Liu, C.; Sawaya, M. R.; Vadla, B.; Khan, S.; Woods, R. J.; Eisenberg, D.; Goux, W. J.; Nowick, J. S. *J. Am. Chem. Soc.* **2011**, *133*, 3144–3157.

35. Cheng, P. N.; Liu, C.; Zhao, M.; Eisenberg, D.; Nowick, J. S. *Nat. Chem.* **2012**, *4*, 927–933.

36. Cheng, P. N.; Spencer, R.; Woods, R. J.; Glabe, C. G.; Nowick, J. S. *J. Am. Chem. Soc.* **2012**, *134*, 14179–14184.

37. Liu, C.; Zhao, M.; Jiang, L.; Cheng, P.-N.; Park, J.; Sawaya, M. R.; Pensalfini, A.; Gou, D.; Berk, A. J.; Glabe, C. G.; Nowick, J. S.; Eisenberg, D. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 20913–20918.

38. Pham, J. D.; Chim, N.; Goulding, C. W.; Nowick, J. S. *J. Am. Chem. Soc.* **2013**, *135*, 12460–12467.

39. Pham, J. D.; Spencer, R. K.; Chen, K. H.; Nowick, J. S. *J. Am. Chem. Soc.* **2014**, *136*, 12682–12690.

40. Nowick, J. S.; Chung, D. M.; Maitra, K.; Maitra, S.; Stigers, K. D.; Sun, Y. *J. Am. Chem. Soc.* **2000**, *122*, 7654–7661.

41. Gothard, C. M.; Nowick, J. S. *J. Org. Chem.* **2010**, *75*, 1822–1830.

42. Spencer, R.; Chen, K. H.; Manuel, G.; Nowick, J. S. *Eur. J. Org. Chem.* **2013**, 3523–3528.

43. Spencer, R. K.; Li, H.; Nowick, J. S. *J. Am. Chem. Soc.* **2014**, *136*, 5595–5598.

Chapter 2

Recipe for β-Sheets: Foldamers Containing Amyloidogenic Peptide Sequences.[a]

Introduction

Controlling and understanding β-sheet folding and aggregation is important, because β-sheet formation and aggregation is central to amyloid formation in amyloid diseases.[1–5] In Alzheimers disease, type II diabetes, prion disease, and many other diseases, amyloidogenic peptides and proteins aggregate to form oligomers and fibrils consisting of networks of β-sheets. Chemical systems that mimic and inhibit β-sheet aggregation can provide biophysical insights into amyloid diseases and lead to chemical probes and drugs with which to study and treat these diseases.

[a]Reproduced with permission from Spencer, R.; Chen, K. H.; Manuel, G.; Nowick, J. S. *Eur. J. Org. Chem.* **2013**, 3523–3528. Copyright 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Macrocyclization is an effective strategy for inducing β-sheet folding in peptides.[6–13] Our laboratory has developed macrocyclic β-sheets in which the unnatural amino acid template *Hao*[14] and δ-linked ornithine turn units[15] are combined with α-amino acids to create well-behaved β-sheets that fold without participating in uncontrolled aggregation.[16–19] We have used these macrocyclic β-sheets to study and inhibit amyloid aggregation and to study β-sheet assembly.[20–22] While the δ-linked ornithine turn can easily be implemented by other laboratories using the commercially available Boc-Orn(Fmoc)-OH building block, the amino acid Hao is not commercially available and poses a challenge to others who wish to use this template. To address this challenge, we set out to determine whether a commercially available *N*-methyl amino acid could achieve the same function.[23–30]

Here I introduce a new class of peptide-based macrocycles that fold to adopt β-sheet structures (Figure 2.1). The macrocycles **1** are readily synthesized from commercially available amino acid building blocks and tolerate a variety of peptide sequences. Each macrocycle contains a pentapeptide β-strand (R_1–R_5) with a sequence from an amyloidogenic peptide linked by two turn units to a pentapeptide template strand (R_6–R_{10}) containing a single *N*-methyl amino acid. I designed this new class of macrocyclic β-sheets to incorporate natural sequences from β-sheet forming peptides and proteins and to fold into β-sheet structures. β-Sheets are notoriously prone to aggregation. I incorporated an *N*-methyl amino acid into the template strand to block aggregation by disrupting intermolecular hydrogen bonding. I incorporated the *N*-methyl amino acid at the R_7 position to not interfere with the intramolecular hydrogen bonding while blocking intermolecular hydrogen bonding. Residues R_6, R_8, and R_{10} participate in intramolecular hydrogen bonding with residues R_5, R_3 and R_1, precluding *N*-methylation of these residues. Incorporation of the *N*-methyl amino acid at the R_9 position should also be possible.

I selected *N*-methyl-L-alanine (Ala_{NMe}) for the R_7 residue because it is inexpensive and easy to integrate into standard Fmoc-based solid-phase peptide synthesis (SPPS). I flanked

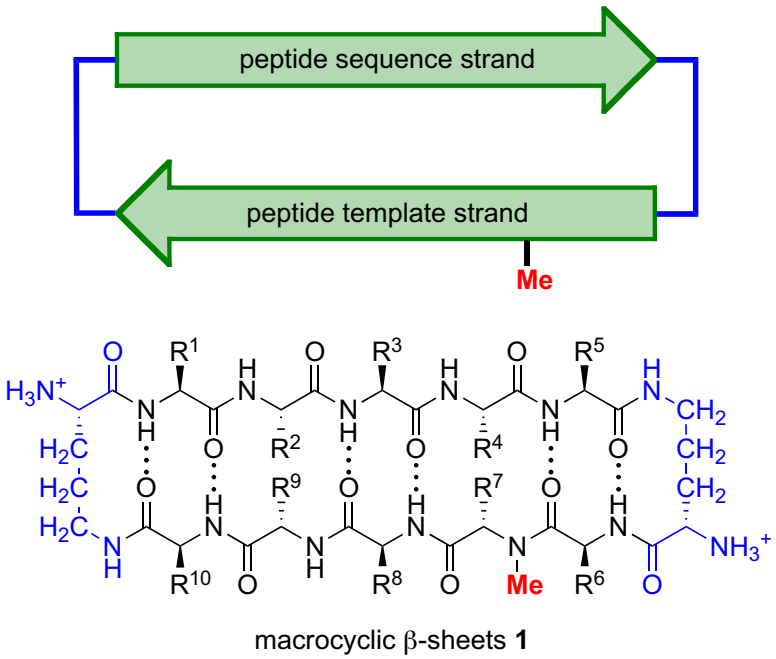


Figure 2.1: Cartoon representation (top) and chemical structure (bottom) of macrocyclic β-sheet **1**. The macrocycle consists of a peptide sequence strand (R_1–R_5) connected by two β-linked ornithine turn units to a peptide template strand (R_6–R_{10}) containing an *N*-methyl amino acid.

the Ala_{NMe} with Val residues in the R_6 and R_8 positions to enhance β-sheet formation. I incorporated Lys at the R_{10} position to ensure water solubility of the macrocycles. I varied the residue at the R_9 position to minimize overlap of α-resonances in the ^{1}H NMR spectrum and to enhance β-sheet formation. I used two δ-linked ornithine amino acids ($^{\delta}$Orn) to connect the amyloidogenic pentapeptide strand (R_1–R_5) to the pentapeptide template strand containing the Ala_{NMe} (R_6–R_{10}). The $^{\delta}$Orn residue mimics a β-turn in a β-hairpin and orients the two pentapeptide strands into an antiparallel β-sheet structure.

Results and Discussion

Sequences derived from amyloidogenic peptides tend to aggregate and pose unique challenges in synthesis and handling. To evaluate whether macrocycles of the general structure **1** could

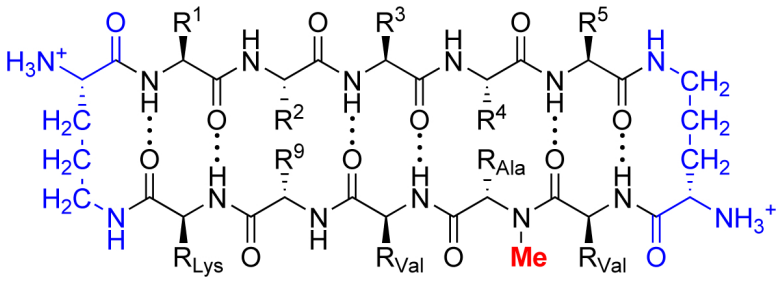


Table 2.1: Amyloidogenic Sequences (R_1–R_5) and Residue R_9 in Macrocyclic β-Sheets **1a**–**e**.

	Sequence	R_1	R_2	R_3	R_4	R_5	R_9
1a	$A\beta_{16\text{-}20}$	Lys	Leu	Val	Phe	Phe	Val
1b	$tau_{306\text{-}310}$	Val	Gln	Ile	Val	Tyr	Leu
1c	B $insulin_{12\text{-}16}$	Val	Glu	Ala	Leu	Tyr	Leu
1d	$hPrP_{118\text{-}122}$	Ala	Gly	Ala	Val	Val	Leu
1e	$hPrP_{118\text{-}122}$	Ala	Gly	Ala	Val	Val	Tyr

incorporate a variety of amyloidogenic peptide sequences, adopt folded structures, and not aggregate, I synthesized macrocycles **1a**–**e**. I grafted sequences derived from amyloid β-peptide ($A\beta_{16–20}$), tau protein ($tau_{306–310}$), the B-chain of insulin (B $insulin_{12–16}$), and human prion protein ($hPrP_{118–122}$) into positions R_1–R_5 (Table 2.1).

Macrocyclic β-sheets **1a**–**e** were synthesized by preparing the corresponding linear peptides on 2-chlorotrityl resin, cleaving the protected linear peptides from the resin, macrocyclizing in solution, removing the protecting groups, and purifying by reverse-phase HPLC (Scheme 2.1). The protected linear peptides **2a**–**e** were prepared by standard Fmoc-based SPPS using an excess of amino acids (4 equiv) and coupling reagent (HCTU, 4 equiv). Coupling Val_6 to the free secondary amino group of the Ala_{NMe} residue required longer reaction times (1 h), double coupling, and the use of HATU and HOAt (4 equiv each) as coupling reagents. The full-length linear peptides were cleaved from the resin with a solution of hexafluoroisopropanol (HFIP) in methylene chloride (1:4). Macrocyclization was performed with a combination of HBTU (5 equiv), HOBt (5 equiv), and diisopropylethylamine (DIPEA, 14 equiv) in DMF at ca. 1 mM to give protected cyclic peptides **3a**–**e**. The synthesis and cyclization typically proceed smoothly, with one major peptide product of ca. 60% purity by HPLC analysis of the crude cyclic peptide. Protected cyclic peptides **3a**–**e** were globally

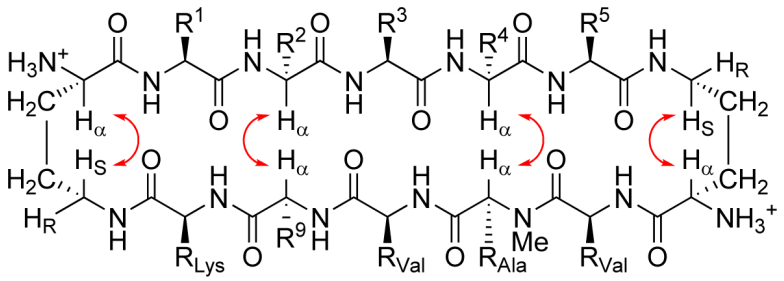
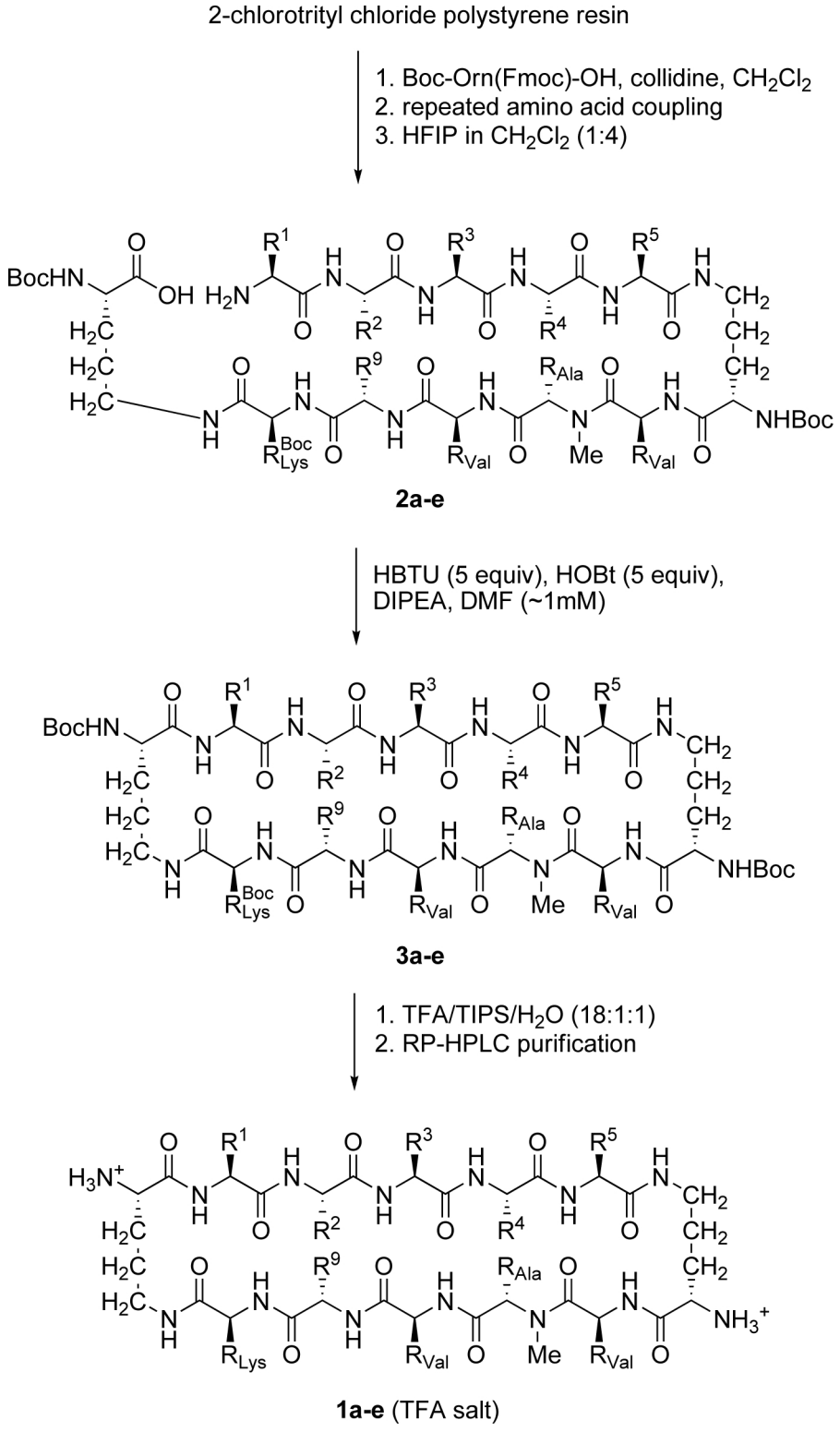


Figure 2.2: NOE correlations observed in macrocyclic β-sheets **1a**–**e**. In macrocycle **1d**, an NOE correlation between the α-proton of R_2 and the α-proton of R_9 was not observed. NOEs were identified in ^{1}H NMR ROESY spectra in 2 mM D_2O solution at 278–290 K. The temperatures at which the experiments were performed were chosen to minimize overlap of the α-proton resonances with the residual HOD peak.

deprotected using a solution of TFA/TIPS/H_2O (18:1:1). The crude peptides were purified by RP-HPLC and lyophilized to give macrocyclic β-sheets **1a**–**e**. Synthesis on a 0.1 millimole scale typically affords 10-25 mg of purified peptide as the TFA salt.

^{1}H NMR ROESY experiments show that macrocycles **1a**–**e** fold to adopt β-sheet structures in aqueous (D_2O) solution. Macrocycles **1a**–**c** and **1e** exhibit characteristic NOEs between the α-protons of residues R_2 and R_9 and between the α-protons of residues R_4 and Ala_{NMe}. These NOEs reflect the folding of the macrocycles to bring residues that are remote in sequence into spatial proximity (Figure 2.2). Macrocycle **1d** exhibits NOE correlations between the α-proton of Val_4 and the α-proton of Ala_{NMe} but not between the α-proton of Gly_2 and the α-proton of Leu_9. These data suggest that macrocycle **1d** is partially folded into a β-sheet structure. All of the macrocycles exhibit characteristic NOEs between each Orn δ-proton and the corresponding diastereotopic *pro*-S δ-proton ($H_{\delta S}$). These NOEs reflect the formation of the well-defined turn structures we have observed previously for δ-linked ornithine.[15,16] Figure 2.3 displays these characteristic NOE correlations in the ROESY spectrum of macrocycle **1a**.

The magnetic anisotropies of the diastereotopic Orn δ-protons in the ^{1}H NMR spectrum reflect the folding of the $_{\delta}$Orn turn units.[15] In a largely folded macrocycle, with well-folded turn structures, the *pro*-S δ-protons appear about 0.6 ppm downfield of the *pro*-R δ-protons.



Scheme 2.1: Synthesis of Macrocyclic β-Sheets **1a**–**e**.

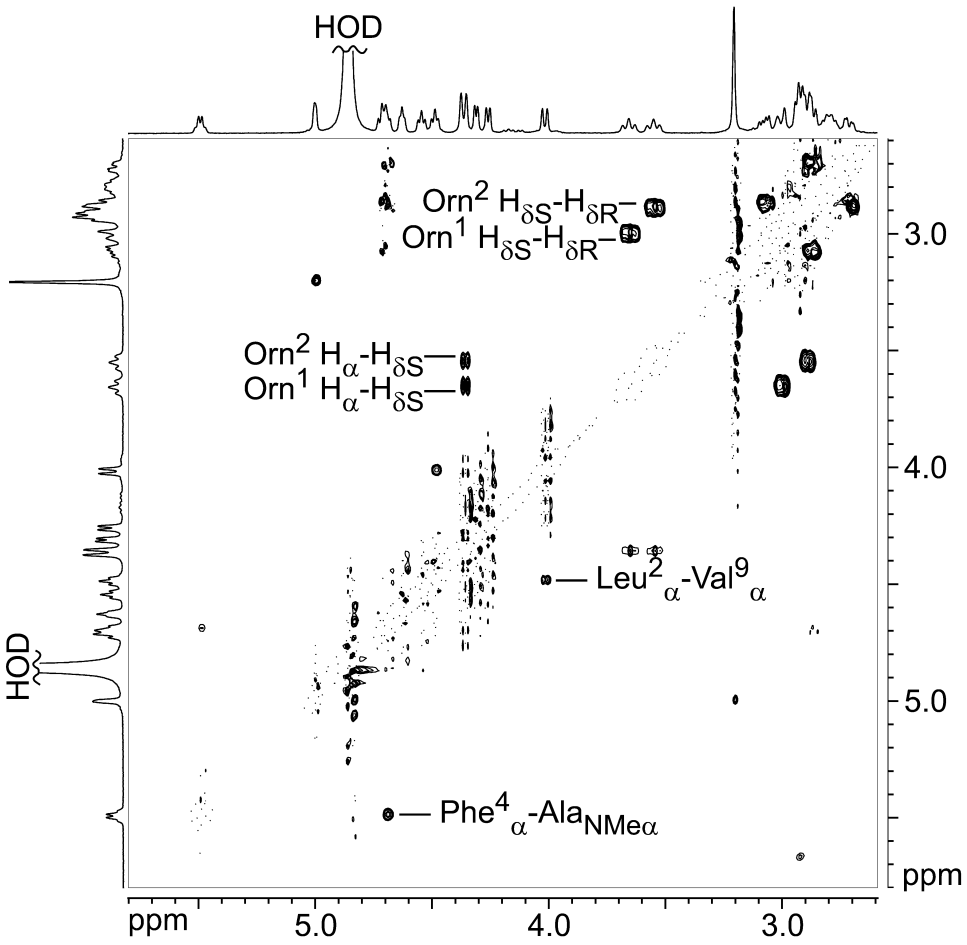


Figure 2.3: Key NOE crosspeaks in the ^{1}H NMR ROESY spectrum of macrocyclic β-sheet **1a** in D_2O (2 mM solution at 500 MHz and 290 K with 150-ms spin-lock mixing time). Important NOEs occur between the α-proton of Leu_2 and Val_9 ($Leu_{2\alpha}$-$Val_{9\alpha}$), the α-proton of Phe_4 and Ala_{NMe} ($Phe_{4\alpha}$-$Ala_{NMe\alpha}$), the α-proton of each Orn and the corresponding *pro*-S δ-proton (Orn^1 H_α-$H_{\delta S}$ and Orn^2 H_α-$H_{\delta S}$), and the diastereotopic Orn δ-protons (Orn^1 $H_{\delta S}$-$H_{\delta R}$ and Orn^2 $H_{\delta S}$-$H_{\delta R}$).

In a partially folded macrocycle, with an ensemble of folded and unfolded states, the Orn δ-protons exhibit less magnetic anisotropy. A macrocycle that is about 50% folded exhibits a magnetic anisotropy of about 0.3 ppm, while a macrocycle that is largely unfolded exhibits little magnetic anisotropy. Table 2.2 summarizes the magnetic anisotropies of the Orn δ-protons of macrocycles **1a**–**e**. Macrocycles **1a**–**c** exhibit large magnetic anistropies and are well folded, while macrocycles **1d** and **1e** exhibit smaller magnetic anisotropies and are less well folded. Macrocycle **1e** exhibits a slightly larger magnetic anisotropy than homologue

Table 2.2: Magnetic Anisotropies of the Orn α-Protons of Macrocyclic β-Sheets **1a**–**e**.[ab]

	Sequence	Orn[1]	Orn[2]	Avg.
1a	$A\beta_{16–20}$	0.65	0.65	0.65
1b	$tau_{306–310}$	0.65	0.60	0.63
1c	B $insulin_{12–16}$	0.62	0.56	0.59
1d	$hPrP_{118–122}$	0.25	0.13	0.19
1e	$hPrP_{118–122}$	0.36	0.28	0.33

[a] ^{1}H NMR spectra in 2 mM D_2O solution at 278-290 K.
[b] Orn[1] and Orn[2] were assigned arbitrarily.

1d, suggesting that a suitable choice of residues at the R_9 position can enhance folding. In macrocycle **1e**, the Tyr residue at R_9 may interact constructively with the Gly residue at R_2 to mitigate the poor β-sheet forming properties of the latter. this is consistent with a model in which the Tyr residue at R_9 interacts with the Gly residue at R_2, the *pro*-S α-proton of the Gly residue of macrocycle **1e** shifts upfield and the diastereotopic Gly α-protons exhibit greater magnetic anisotropy (0.44 ppm in **1e** vs. 0.13 ppm in **1d**).

The coupling patterns of the diastereotopic Orn δ-protons in the ^{1}H NMR spectrum also reflect the folding of the $^{\delta}$Orn turn units.[15,16] In a well-folded $^{\delta}$Orn turn unit, the Orn side chain adopts a well-defined conformation in which the *pro*-S δ-proton appears as a broad triplet (J ≈ 13-14 Hz) at ca. 3.6 ppm. In D_2O solution, each *pro*-S Orn δ-proton is split by three coupling partners—the *pro*-R Orn δ-proton and the *pro*-S and *pro*-R Orn δ-protons. The triplet-like coupling pattern arises from a large geminal coupling between the *pro*-S δ-proton and the *pro*-R δ-proton and a large vicinal coupling between the *pro*-S δ-proton and the *pro*-R δ-proton. In a well-folded $^{\delta}$Orn turn unit, the *pro*-S δ-proton and the *pro*-R δ-proton are nearly antiperiplanar, and exhibit a large coupling constant. The *pro*-S δ-proton and the *pro*-S δ-proton are nearly gauche and only have small coupling, which is seen as broadening of the resonances rather than actual splitting. In a less well-folded $^{\delta}$Orn turn unit, the *pro*-S δ-proton shifts upfield and appears as a broadened doublet of doublet of doublets. The well-folded macrocycles **1a**–**c** show the former coupling behavior, while the less well-folded macrocycles **1d** and **1e** show the latter (Figure 2.4).

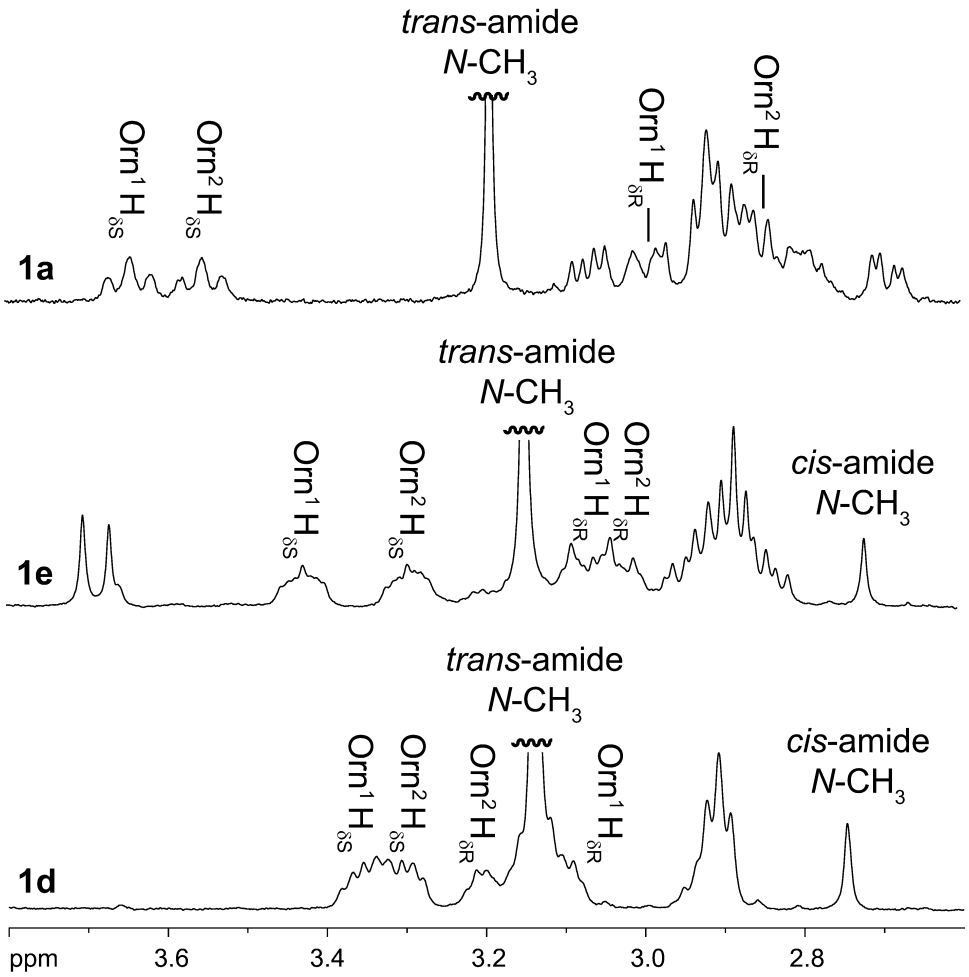


Figure 2.4: Expansions of the ^{1}H NMR spectra of macrocyclic β-sheets **1a** (top), **1e** (middle), and **1d** (bottom) in D_2O (2 mM solution at 500 MHz and 298 K). The Orn δ-proton resonances and Ala_{NMe} N-methyl resonances are marked.

Tertiary amides, such as peptides containing proline, generally show mixtures of *trans*- and *cis*-amide rotamers in the ^{1}H NMR spectrum. While the ^{1}H NMR spectrum of well-folded macrocyclic β-sheets **1a**–**c** show little or no ($\leq$ 5%) *cis*-amide rotamer, the less well-folded β-sheets **1d** and **1e** show 10–15% (Figure 2.4).[b] The methyl resonance of the *trans*-amide rotamer appears at 3.1–3.2 ppm, while that of the *cis*-amide rotamer appears at 2.7–2.8 ppm. The greater degree of β-sheet folding of macrocycles **1a**–**c** shifts the conformational equilibria of these peptides toward the *trans*-amide rotamer.

[b]The two N-methyl resonances of the Ala_{NMe} residue exhibit an EXSY crosspeak in the ROESY spectrum of macrocycle **1d** at 343 K in D_2O. The presence of an EXSY crosspeak confirms that these peaks arise from two amide rotamers.

Conclusions

Macrocyclic peptides that fold to form β-sheets should now be accessible to other laboratories capable of doing peptide synthesis with commercially available amino acid building blocks. The combination of amyloidogenic peptide strands, δ-linked ornithine turn units, and *N*-methyl amino acid-based templates constitutes a recipe for β-sheets.[31] Other *N*-alkyl amino acids and a wide variety of peptide sequences should also be tolerated. Larger macrocyclic β-sheets containing longer peptide sequences can also be envisioned. I am currently exploring the scope and limitations of using *N*-methyl amino acids, δ-linked ornithine turns units, and α-amino acids to create well-behaved macrocycles that fold to adopt β-sheet structures. I anticipate using these macrocyclic β-sheets to study and inhibit amyloid aggregation and to study β-sheet assembly.[32]

Materials and Methods

General methods

Reagent-grade solvents, chemicals, amino acids, and resin were used as received, with the exception that methylene chloride was dried by passage through a column of alumina under argon. Automated solid-phase peptide synthesis was carried out on a PS3^{TM} Peptide Synthesizer (Protein Technologies, Inc.). Analytical reverse-phase HPLC was performed on an Agilent Zorbax SB-C18 column (50 mm x 4.6 mm) with a gradient of 5-100% CH_3CN in H2O with 0.1% TFA and a flow of 1.0 mL/min over 20 minutes. Preparative reverse-phase HPLC purification was carried out on a Zorbax SB-C18 PrepHT (21b x 250 mm, 7-m particle size) column from Agilent on a Beckmann system with a flow of 15.0 mL/min. UV detection (214 nm) was used for analytical and preparative HPLC. Water and acetonitrile were used as the

solvents. Both solvents contained 0.1% biochemical grade TFA.

^{1}H NMR experiments of peptides **1a-e** were performed in D_2O solution. Solutions of peptides **1a**–**e** were prepared gravimetrically by dissolving an appropriate weight of each peptide in an appropriate volume of solvent. In calculating molecular weights, all amino groups were assumed to be protonated as the TFA salts. Important ^{1}H NMR resonances of peptides **1a**–**e** were assigned by TOCSY and ROESY experiments. Presaturation of the residual HOD resonance was performed in ROESY experiments. TOCSY and ROESY spectra were collected with 2048 data points in the f2 domain and 512 data points in the f1 domain. Data were processed to a 1024 x 1024 real matrix with a Qsine weighting function and with forward linear prediction in the f1 domain. The data were processed with the Bruker XwinNMR software.

^{1}H NMR EXSY experiments of peptide **1d** were performed in D_2O at 343 K using a ROESY pulse sequence. (The temperature was selected to give slight broadening of the peaks associated with chemical exchange.) Data were collected with 2048 data points in the f2 domain and 256 data points in the f1 domain. Data were processed to a 1024 x 1024 real matrix with a Qsine weighting function and with forward linear prediction in the f1 domain. The data were processed with Bruker Topspin NMR software.

Synthesis of macrocyclic β-sheets 1a–e.

Macrocyclic β-sheets **1a**–**e** were synthesized in a fashion similar to that which was reported previously for other macrocylic β-sheet peptides.[19] The syntheses involved: loading of the resin, peptide coupling, cleavage of the peptide from the resin, cyclization of the linear peptide, and global deprotection and purification. The synthesis of macrocycle **1a** is representative of the procedures used for the synthesis of macrocycles **1a**–**e**.

Synthesis of macrocycle 1a.

*Loading of the resin.*2-Chlorotrityl chloride resin (300 mg, 1a2 mmol/g) was added to a Bio-Rad Poly-Prep chromatography column (10 mL, 0.8 x 4.0 cm). The resin was suspended in dry CH_2Cl_2 (10 mL) and allowed to swell for 30 min. The solution was drained from the resin and a solution of Boc-Orn(Fmoc)-OH (0.50 equiv, 70 mg, 0.17 mmol) in 20% 2,4,6-collidine in dry CH_2Cl_2 (5 mL) was added immediately and the mixture was gently agitated for 12 h. The solution was then drained and a mixture of CH_2Cl_2/MeOH/DIPEA (17:2:1, 10 mL) was added immediately, and the mixture was gently agitated for 1 h to cap the unreacted 2-chlorotrityl chloride sites. The resin was washed with DMF (2 times) and dried by passing nitrogen through the vessel. The resin loading was determined to be 0.12 mmol (0.39 mmol/g, 70% based on Boc-Orn(Fmoc)-OH) by UV analysis of the Fmoc cleavage product. Loadings of 0.11–0.13 mmol (65% to 75%, based on Boc-Orn(Fmoc)-OH) were typically observed in various repetitions of this procedure.

Peptide coupling. The PS-2-chlorotrityl-Orn(Fmoc)-Boc generated from the previous step was transferred to a solid-phase peptide synthesizer reaction vessel and submitted to cycles of automated peptide coupling using Fmoc-protected amino acid building blocks. The linear peptide was synthesized from the C-terminus to the N-terminus. Each coupling consisted of: *i* Fmoc-deprotection with 20% piperidine in DMF for 3 min; *ii* washing with DMF (3 times); *iii* coupling of the amino acid (0.5 mmol, 4 equiv) in the presence of HCTU; *iv* washing with DMF (6 times). Each amino acid coupling step took 20 min for residues R_{10}–R_8 (Lys-Val-Val), *N*-methyl-L-alanine, and R_5-R_1 (Phe-Phe-Val-Leu-Lys). The R_6 Val residue (after the *N*-methyl-L-alanine) was double coupled (0.5 mmol, 4 equiv) and mixed for 1 h per coupling with HATU (4 equiv) and HOAt (4 equiv). After the last amino acid was coupled, the resin was transferred from the reaction vessel of the peptide synthesizer to a Bio-Rad Poly-Prep chromatography column. The terminal Fmoc group was removed with 20% piperidine in DMF. The suspension was gently agitated for 20 min. The DMF was drained thorougly by

passing nitrogen gas through the vessel.

Cleavage of the peptide from the resin. Linear peptide **2a** was cleaved from the resin by agitating the resin for 30 min with a solution of hexafluoroisopropanol (HFIP) in CH_2Cl_2 (1:4, 5 mL). The suspension was filtered, and the filtrate was collected in a 250-mL round-bottom flask. The resin was washed with additional hexafluoroisopropanol (HFIP) in CH_2Cl_2 (1:4, 5 mL) and CH_2Cl_2 (2 x 10 mL). The combined filtrates were concentrated by rotary evaporation to give a white solid. The white solid was further dried using a vacuum pump to afford the crude protected linear peptide **2a**, which was cyclized without further purification.

Cyclization of linear peptide. Crude protected linear peptide **2a** was dissolved in dry DMF (125 mL). HOBt (81 mg, 0.60 mmol, 5 equiv) and HBTU (227 mg, 0.6 mmol, 5 equiv) were added to the solution. The reaction mixture was then stirred under nitrogen for 20 min. DIPEA (0.3 mL, 1.7 mmol, 14 equiv) was added to the solution and the mixture was stirred under nitrogen for 24 h. The mixture was concentrated under reduced pressure to afford crude protected cyclic peptide **3a**.

Global deprotection and purification of the cyclic peptide. Protected cyclic peptide **3a** was dissolved in a mixture of TFA/triisopropylsilane (TIPS)/H_2O (18:1:1, 10 mL) in a 250-mL round-bottom flask equipped with a nitrogen-inlet adaptor. The solution was stirred for 1 h. The reaction mixture was then concentrated by rotary evaporation under reduced pressure to afford the deprotected cyclic peptide as a yellow oil. The oil was dissolved in H_2O (5 mL), and the solution was filtered through a 0.45 µm syringe filter and purified by reverse-phase HPLC (gradient elution with 20-50% CH_3CN over 40 min). The pure fractions were lyophilized to afford peptide **1a** (15 mg, 9% yield over three steps, based on initial resin loading). The following yields were obtained for peptides **1b**–**1e**: **1b** (25 mg, 16% yield), **1c** (12 mg, 7% yield), **1d** (25 mg, 19% yield), and **1e** (12 mg, 8% yield).

References

1. Tzotzos, S.; Doig, A. J. *Protein Science* **2010**, *19*, 327–348.

2. Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006**, *75*, 333–366.

3. Bartolini, M.; Andrisano, V. *Chembiochem* **2010**, *11*, 1018–1035.

4. Greenwald, J.; Riek, R. *Structure* **2010**, *18*, 1244–1260.

5. Fändrich, M. *J. Mol. Biol.* **2012**, *421*, 427–440.

6. Bisang, C.; Jiang, L.; Freund, E.; Emery, F.; Bauch, C.; Matile, H.; Pluschke, G.; Robinson, J. A. *J. Am. Chem. Soc.* **1998**, *120*, 7439–7449.

7. Athanassiou, Z.; Dias, R. L.; Moehle, K.; Dobson, N.; Varani, G.; Robinson, J. A. *J. Am. Chem. Soc.* **2004**, *126*, 6906–6913.

8. Robinson, J. A. *Acc. Chem. Res.* **2008**, *41*, 1278–1288.

9. Syud, F. A.; Espinosa, J. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1999**, *121*, 11577–11578.

10. Freire, F.; Gellman, S. H. *J. Am. Chem. Soc.* **2009**, *131*, 7970–7972.

11. Almeida, A. M.; Li, R.; Gellman, S. H. *J. Am. Chem. Soc.* **2012**, *134*, 75–78.

12. Woys, A. M.; Almeida, A. M.; Wang, L.; Chiu, C. C.; McGovern, M.; de Pablo, J. J.; Skinner, J. L.; Gellman, S. H.; Zanni, M. T. *J. Am. Chem. Soc.* **2012**, *134*, 19118–19128.

13. Park, J. H.; Waters, M. L. *Org. Biomol. Chem.* **2013**, *11*, 69–77.

14. Nowick, J. S.; Chung, D. M.; Maitra, K.; Maitra, S.; Stigers, K. D.; Sun, Y. *J. Am. Chem. Soc. 122*, 7654–7661.

15. Nowick, J. S.; Brower, J. O. *J. Am. Chem. Soc.* **2003**, *125*, 876–877.

16. Woods, R. J.; Brower, J. O.; Castellanos, E.; Hashemzadeh, M.; Khakshoor, O.; Russu, W. A.; Nowick, J. S. *J. Am. Chem. Soc.* **2007**, *129*, 2548–2558.

17. Khakshoor, O.; Demeler, B.; Nowick, J. S. *J. Am. Chem. Soc.* **2007**, *129*, 5558–5569.

18. Khakshoor, O.; Lin, A. J.; Korman, T. P.; Sawaya, M. R.; Tsai, S. C.; Eisenberg, D.; Nowick, J. S. *J. Am. Chem. Soc.* **2010**, *132*, 11622–11628.

19. Cheng, P. N.; Nowick, J. S. *J. Org. Chem.* **2011**, *76*, 3166–3173.

20. Cheng, P. N.; Liu, C.; Zhao, M.; Eisenberg, D.; Nowick, J. S. *Nat. Chem.* **2012**, *4*, 927–933.

21. Liu, C.; Zhao, M.; Jiang, L.; Cheng, P.-N.; Park, J.; Sawaya, M. R.; Pensalfini, A.; Gou, D.; Berk, A. J.; Glabe, C. G.; Nowick, J. S.; Eisenberg, D. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 20913–20918.

22. Zheng, J.; Liu, C.; Sawaya, M. R.; Vadla, B.; Khan, S.; Woods, R. J.; Eisenberg, D.; Goux, W. J.; Nowick, J. S. *J. Am. Chem. Soc.* **2011**, *133*, 3144–3157.

23. Kobayashi, K.; Granja, J. R.; Ghadiri, M. R. *Angew. Chem. Int. Ed.* **1995**, *34*, 95–98.

24. Rajarathnam, K.; Clark-Lewis, I.; Sykes, B. D. *Biochemistry* **1995**, *34*, 12983–12990.

25. Soto, C.; Kindy, M. S.; Baumann, M.; Frangione, B. *Biochem. Biophys. Res. Commun.* **1996**, *226*, 672–680.

26. J. Doig, A. *Chemical Communications* **1997**, 2153–2154.

27. Adessi, C.; Frossard, M. J.; Boissard, C.; Fraga, S.; Bieler, S.; Ruckle, T.; Vilbois, F.; Robinson, S. M.; Mutter, M.; Banks, W. A.; Soto, C. *J. Biol. Chem.* **2003**, *278*, 13905–13911.

28. Kokkoni, N.; Stott, K.; Amijee, H.; Mason, J. M.; Doig, A. J. *Biochemistry* **2006**, *45*, 9906–9918.

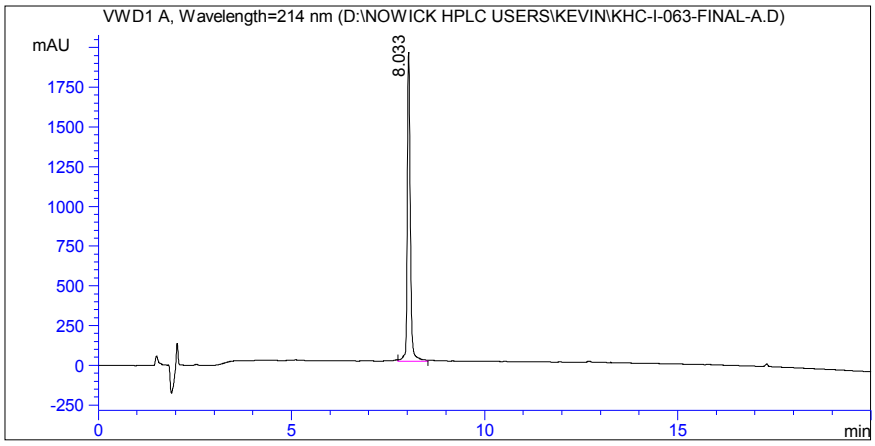
29. Amijee, H.; Bate, C.; Williams, A.; Virdee, J.; Jeggo, R.; Spanswick, D.; Scopes, D. I.; Treherne, J. M.; Mazzitelli, S.; Chawner, R.; Eyers, C. E.; Doig, A. J. *Biochemistry* **2012**, *51*, 8338–8352.

30. Chatterjee, J.; Rechenmacher, F.; Kessler, H. *Angew. Chem., Int. Ed.* **2013**, *52*, 254–269.

31. Mayo, K. H.; Ilyina, E.; Park, H. *Protein Sci.* **1996**, *5*, 1301–1315.

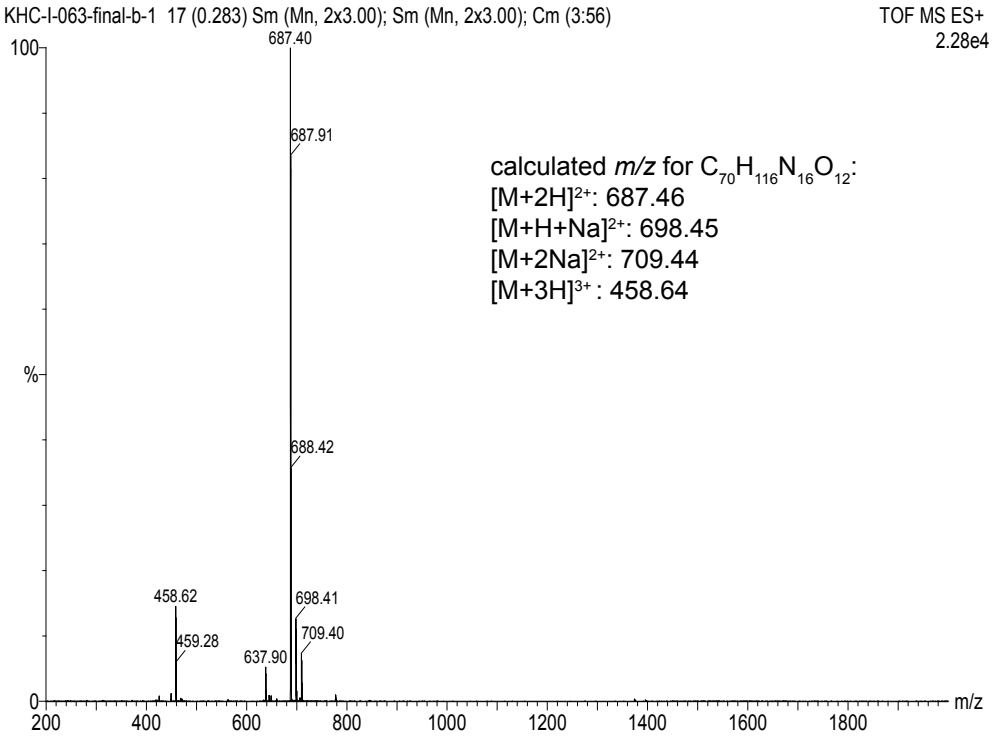
32. Cheng, P.-N.; Pham, J. D.; Nowick, J. S. *J. Am. Chem. Soc.* **2013**, *135*, 5477–5492.

HPLC and MS ESI$^+$ TOF of macrocycle **1a**

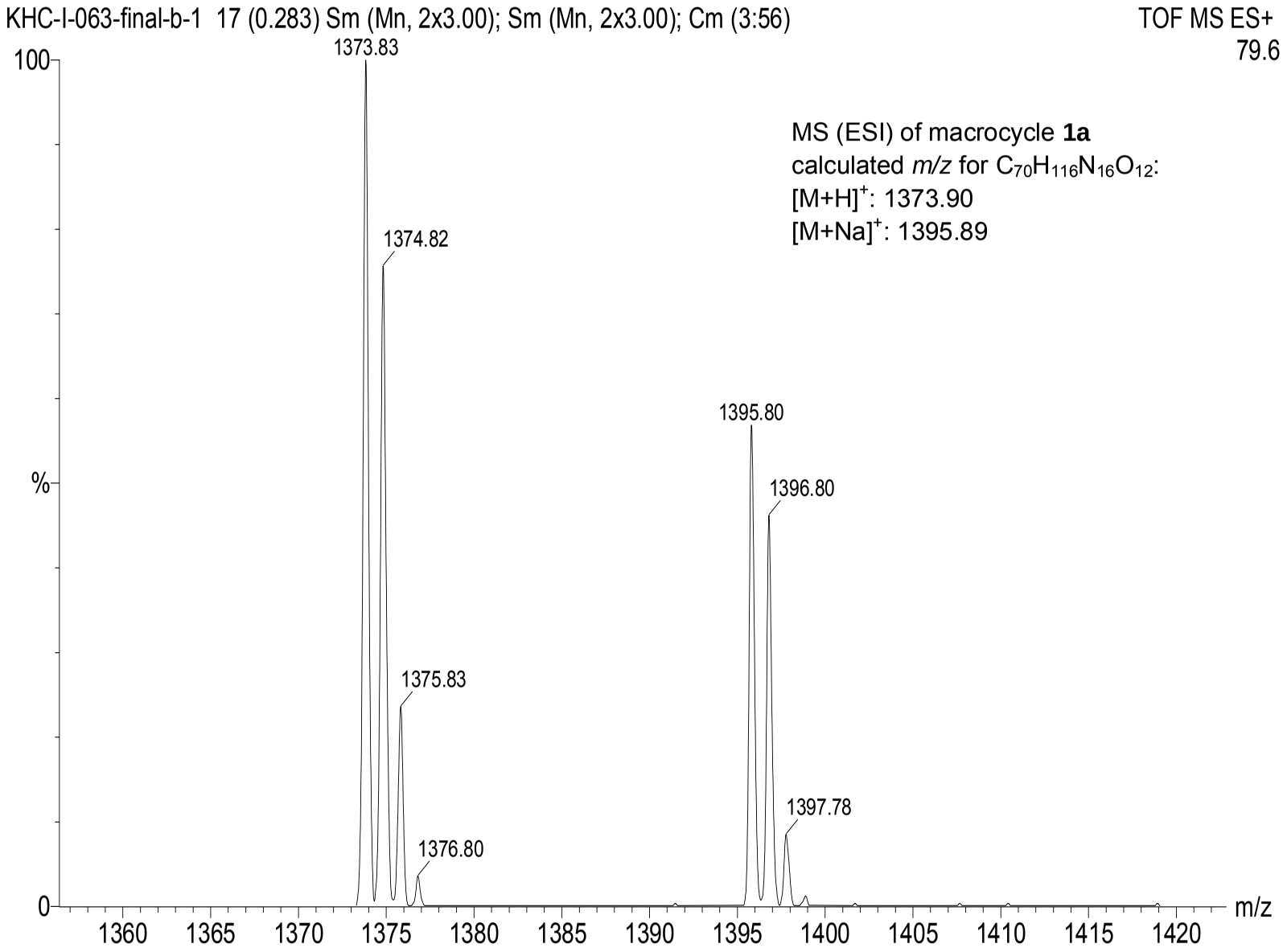


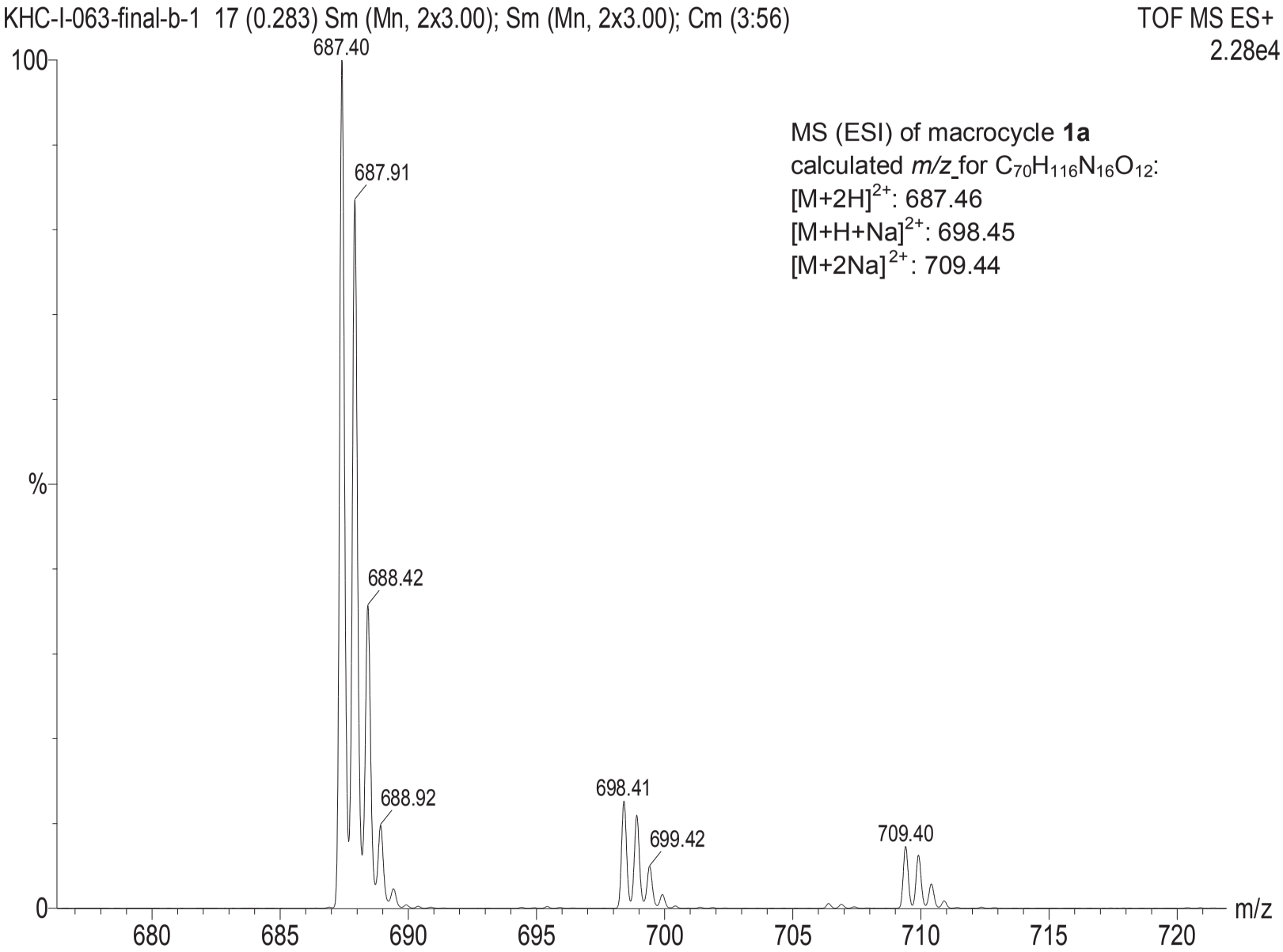
Signal 1:VWD1 A, Wavelength=214 nm

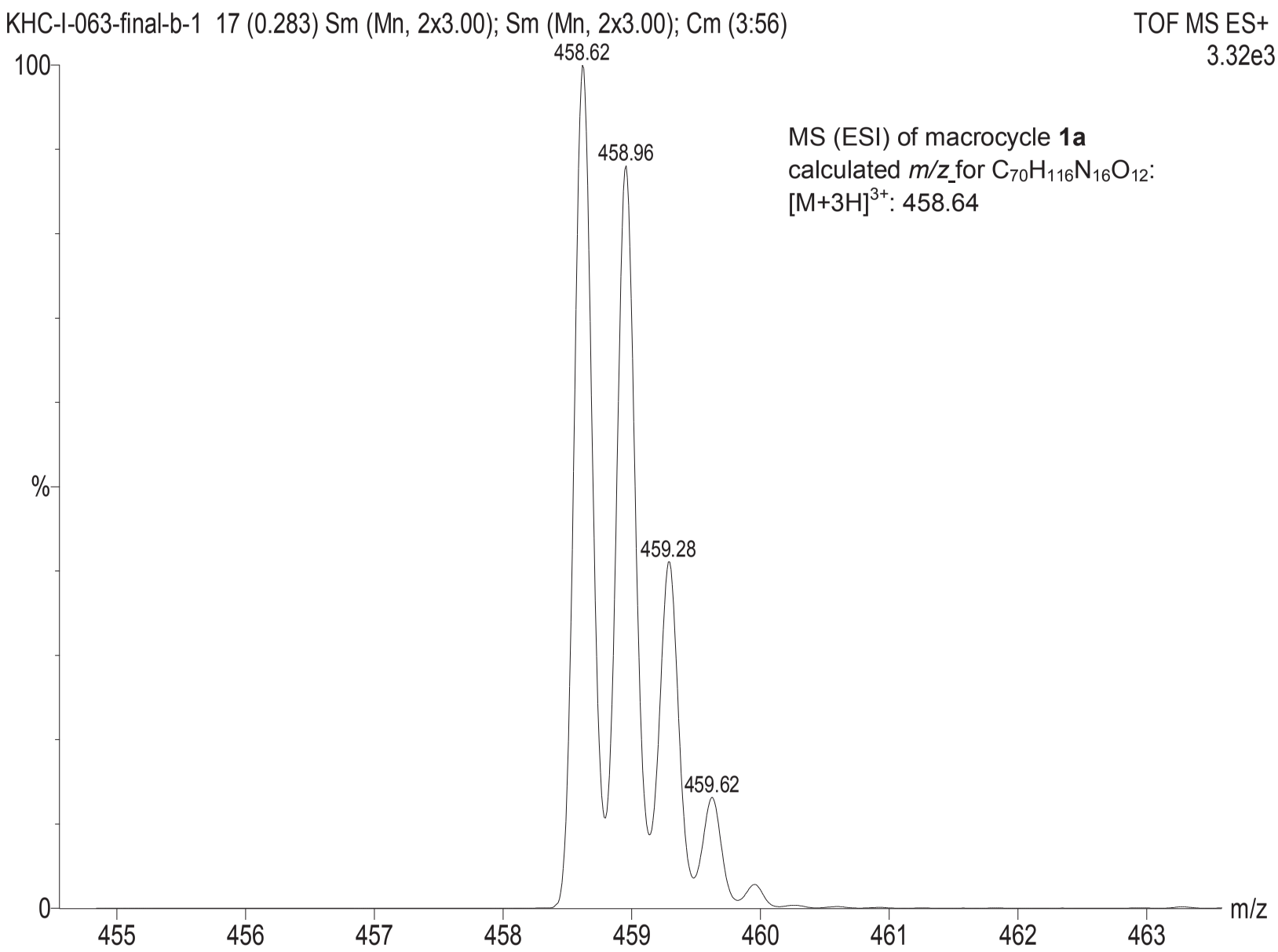
Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	8.033	VV	0.082	10583.563	100.000	100.000



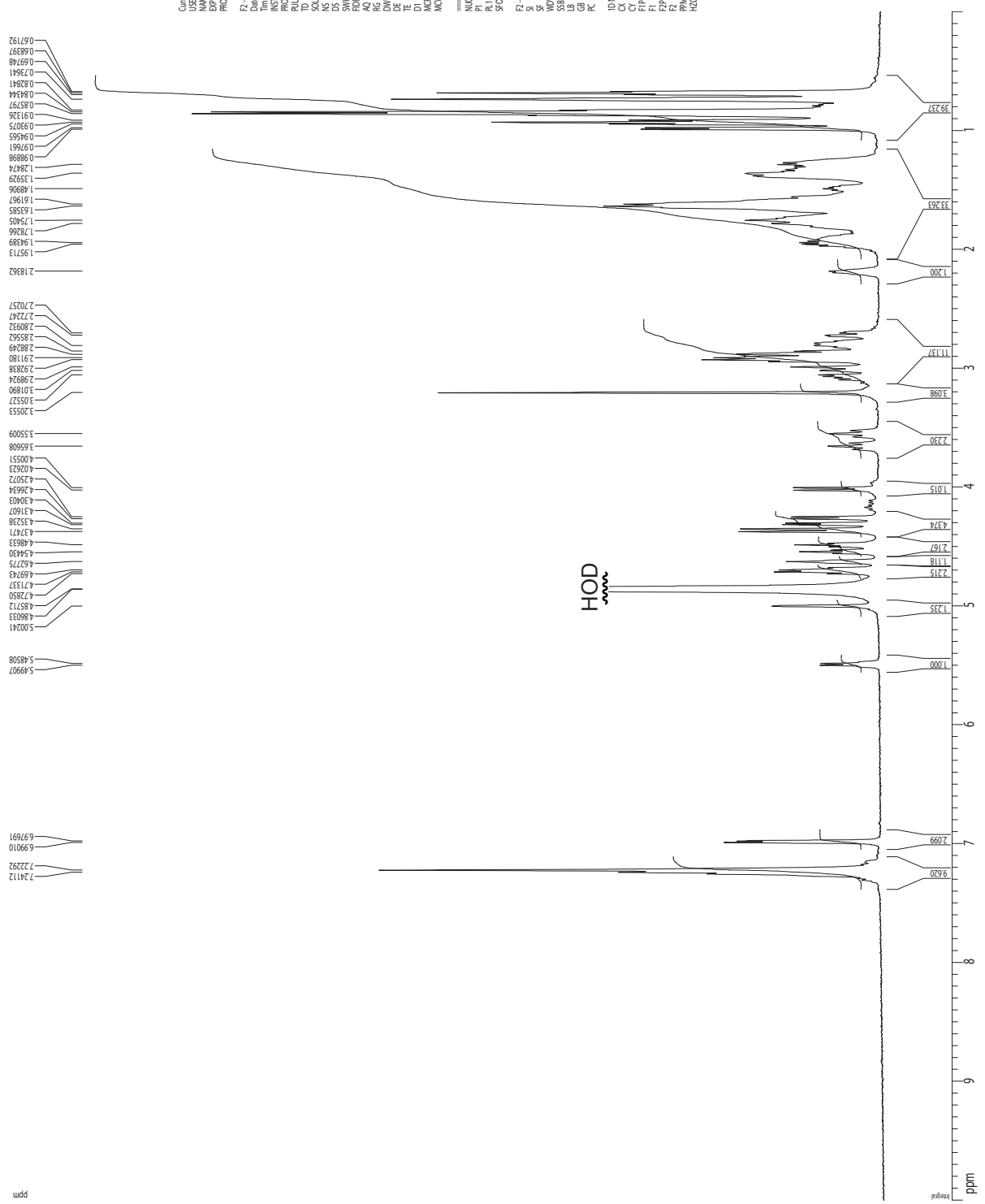
calculated *m/z* for $C_{70}H_{116}N_{16}O_{12}$:
$[M+2H]^{2+}$: 687.46
$[M+H+Na]^{2+}$: 698.45
$[M+2Na]^{2+}$: 709.44
$[M+3H]^{3+}$: 458.64


KHC-I-063-final-b-1 17 (0.283) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:56)
TOF MS ES+
79.6
MS (ESI) of macrocycle 1a
calculated m/z for C70H116N16O12:
[M+H]+: 1373.90
[M+Na]+: 1395.89
1373.83
1374.82
1375.83
1376.80
1395.80
1396.80
1397.78
100
%
0
m/z
1360
1365
1370
1375
1380
1385
1390
1395
1400
1405
1410
1415
1420


KHC-I-063-final-b-1 17 (0.283) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:56)
TOF MS ES+
2.28e4
100
%
0
687.40
687.91
688.42
688.92
698.41
699.42
709.40
680
685
690
695
700
705
710
715
720
m/z
MS (ESI) of macrocycle 1a
calculated m/z for C70H116N16O12:
[M+2H]2+: 687.46
[M+H+Na]2+: 698.45
[M+2Na]2+: 709.44


KHC-I-063-final-b-1 17 (0.283) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:56)
TOF MS ES+
3.32e3
100
%
0
458.62
458.96
459.28
459.62
MS (ESI) of macrocycle 1a
calculated m/z for C70H116N16O12:
[M+3H]3+: 458.64
455
456
457
458
459
460
461
462
463
m/z

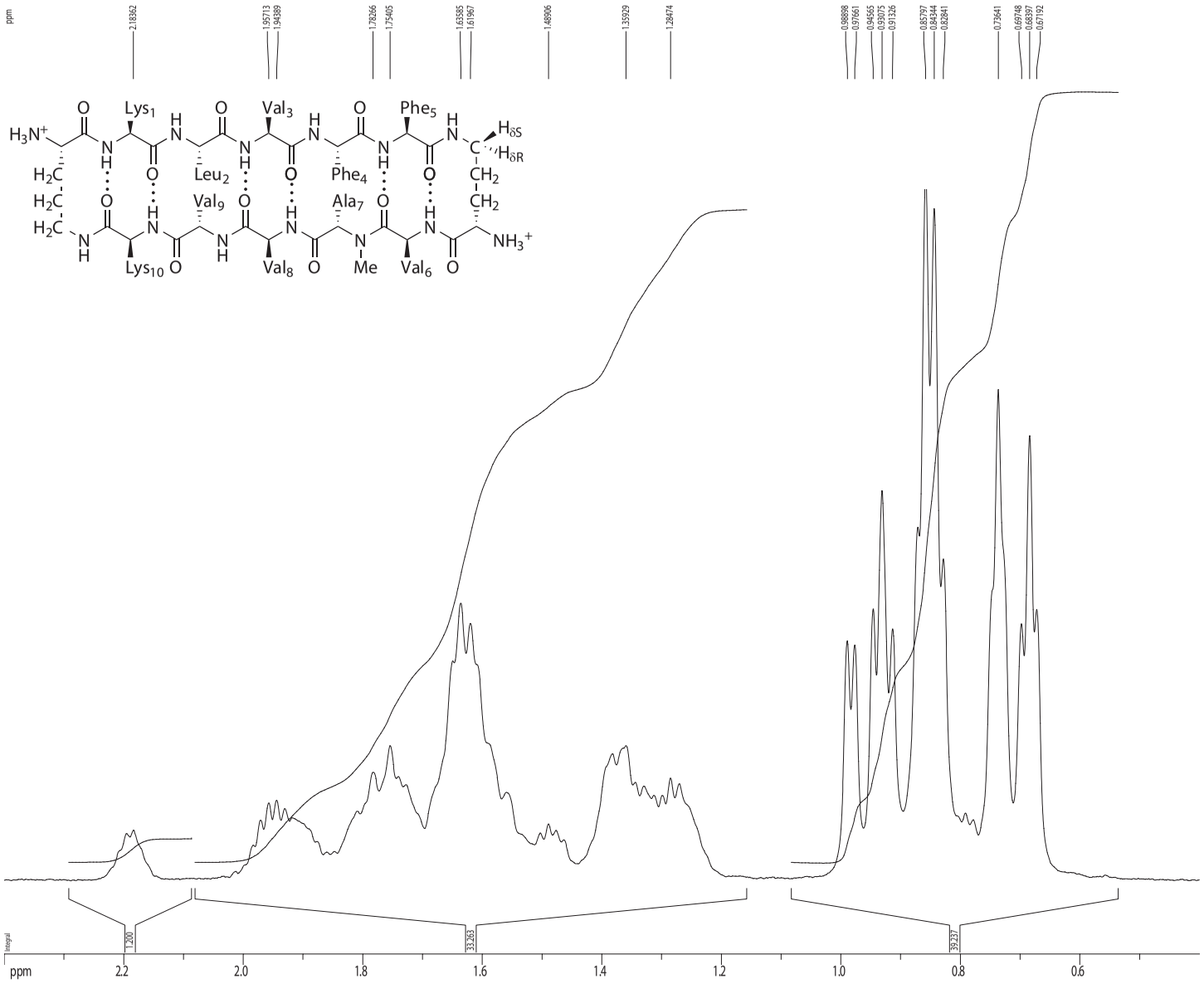
^{1}H NMR 1D of macrocycle **1a**, 2 mM in D_2O, 500 MHz at 290 K



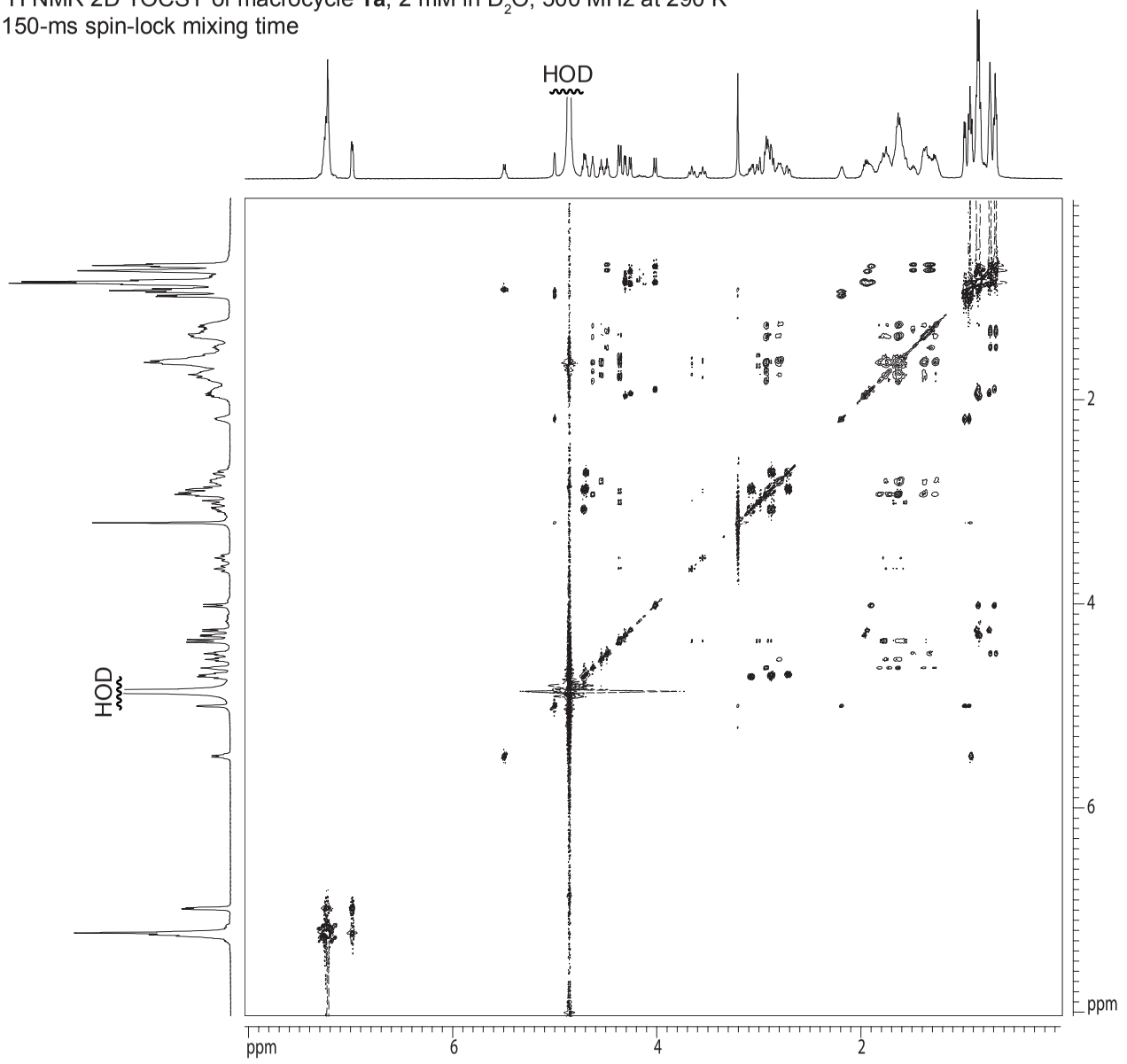
^{1}H NMR 1D of macrocycle **1a**, 2 mM in D_2O, 500 MHz at 290 K

^{1}H NMR 1D of macrocycle **1a**, 2 mM in D_2O, 500 MHz at 290 K

^{1}H NMR 1D of macrocycle **1a**, 2 mM in D_2O, 500 MHz at 290 K



^{1}H NMR 2D TOCSY of macrocycle **1a**, 2 mM in D_2O, 500 MHz at 290 K
150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME KHC-I-63
EXPNO 10
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130122
Time 20.37
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG tocsytp.wu
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 4006.410 Hz
FIDRES 1.956255 Hz
AQ 0.2556404 sec
RG 4.5
DW 124.800 usec
DE 6.00 usec
TE 290.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15211743 sec
d12 0.00002000 sec
IN0 0.00012480 sec
L1 65

======== CHANNEL f1 =======
NUC1 1H
P1 7.50 usec
p5 23.34 usec
P6 35.00 usec
p7 70.00 usec
P17 2500.00 usec
PL1 1.60 dB
PL10 15.20 dB
SFO1 500.2220009 MHz

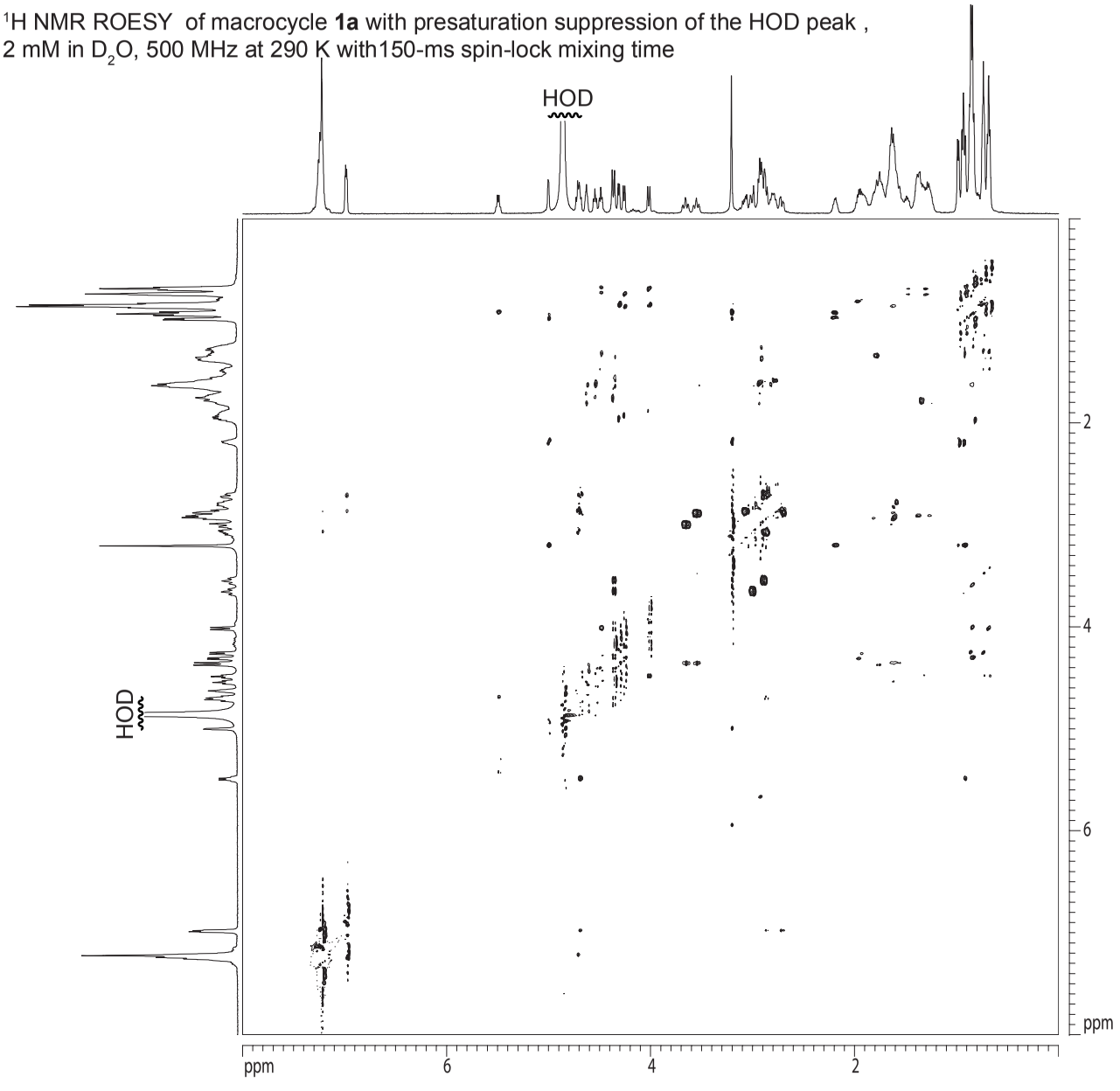
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.222 MHz
FIDRES 7.825020 Hz
SW 8.009 ppm
FnMODE undefined

F2 - Processing parameters
SI 1024
SF 500.2199738 MHz
WDW QSINE
SSB 3
LB 0.00 Hz
GB 0
PC 4.00

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2199738 MHz
WDW QSINE
SSB 3
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 8.036 ppm
F2LO 4019.55 Hz
F2PHI 0.026 ppm
F2HI 13.14 Hz
F1PLO 8.036 ppm
F1LO 4019.55 Hz
F1PHI 0.026 ppm
F1HI 13.14 Hz
F2PPMCM 0.53395 ppm/cm
F2HZCM 267.09399 Hz/cm
F1PPMCM 0.53395 ppm/cm
F1HZCM 267.09399 Hz/cm

^{1}H NMR ROESY of macrocycle **1a** with presaturation suppression of the HOD peak , 2 mM in D_2O, 500 MHz at 290 K with150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME KHC-I-63
EXPNO 6
PROCNO 2

F2 - Acquisition Parameters
Date_ 20130112
Time 21.03
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesyprtp.2.prd
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 5482.456 Hz
FIDRES 2.676980 Hz
AQ 0.1868276 sec
RG 322.5
DW 91.200 usec
DE 6.00 usec
TE 290.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15000001 sec
d11 0.03000000 sec
d12 0.00002000 sec
d13 0.00000300 sec
IN0 0.0000912 0 sec
L4 300

======= CHANNEL f1 =======
NUC1 1H
P1 7.50 usec
P25 250.00 usec
PL1 1.60 dB
PL9 54.00 dB
PL11 26.60 dB
SFO1 500.2224049 MHz

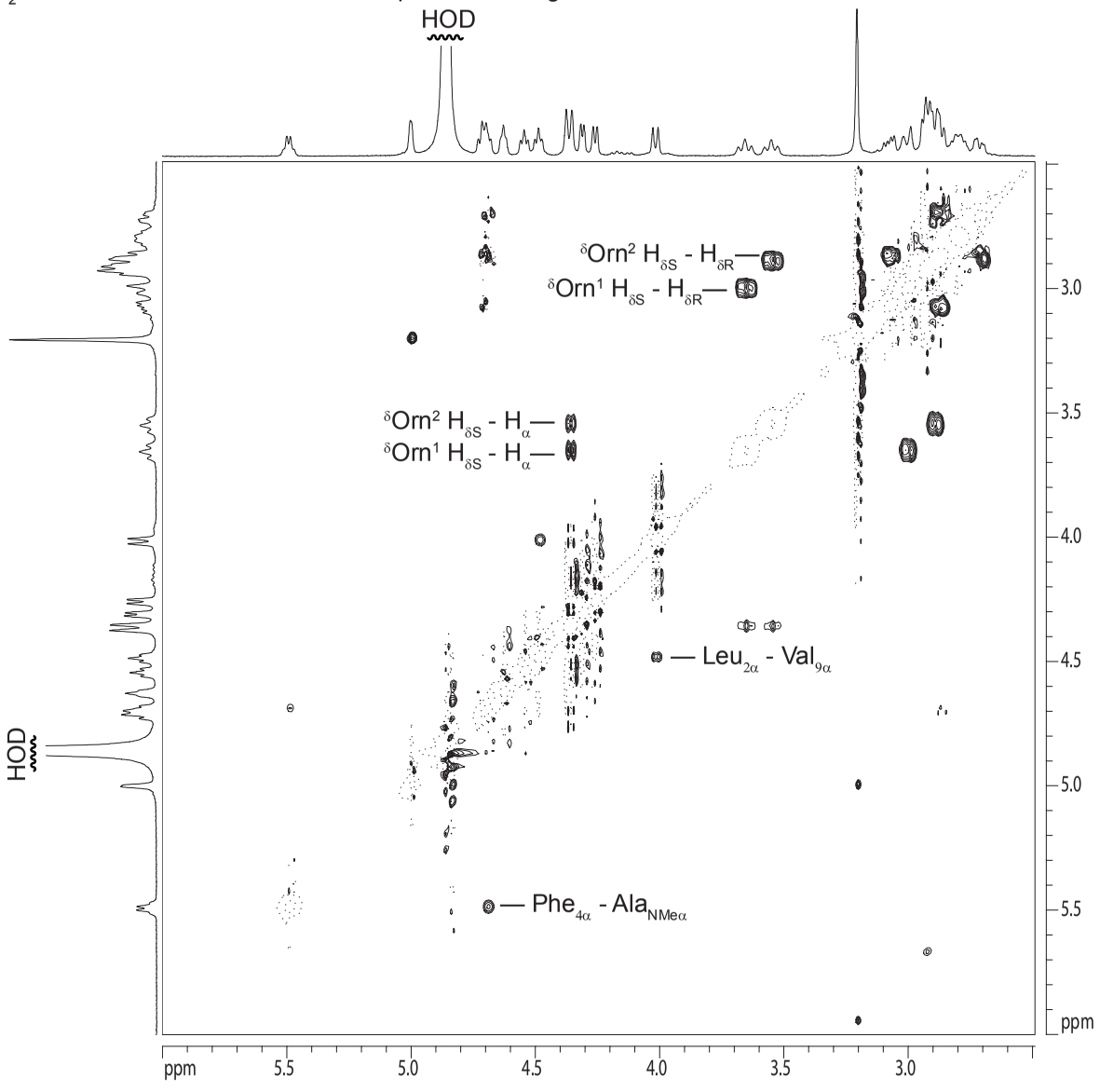
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 10.707922 Hz
SW 10.960 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2199738 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 1.40

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2199738 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 8.000 ppm
F2LO 4001.76 Hz
F2PHI 0.000 ppm
F2HI 0.00 Hz
F1PLO 8.000 ppm
F1LO 4001.76 Hz
F1PHI 0.000 ppm
F1HI 0.00 Hz
F2PPMCM 0.53333 ppm/cm
F2HZCM 266.78400 Hz/cm
F1PPMCM 0.53333 ppm/cm
F1HZCM 266.78400 Hz/cm

^{1}H NMR ROESY of macrocycle **1a** with presaturation suppression of the HOD peak , 2 mM in D_2O, 500 MHz at 290 K with150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME KHC-I-63
EXPNO 6
PROCNO 2

F2 - Acquisition Parameters
Date_ 20130112
Time 21.03
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesyprtp.2.prd
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 5482.456 Hz
FIDRES 2.676980 Hz
AQ 0.1868276 sec
RG 322.5
DW 91.200 usec
DE 6.00 usec
TE 290.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15000001 sec
d11 0.03000000 sec
d12 0.00002000 sec
d13 0.00000300 sec
IN0 0.00009120 sec
L4 300

======= CHANNEL f1 =======
NUC1 1H
P1 7.50 usec
P25 250.00 usec
PL1 1.60 dB
PL9 54.00 dB
PL11 26.60 dB
SFO1 500.2224049 MHz

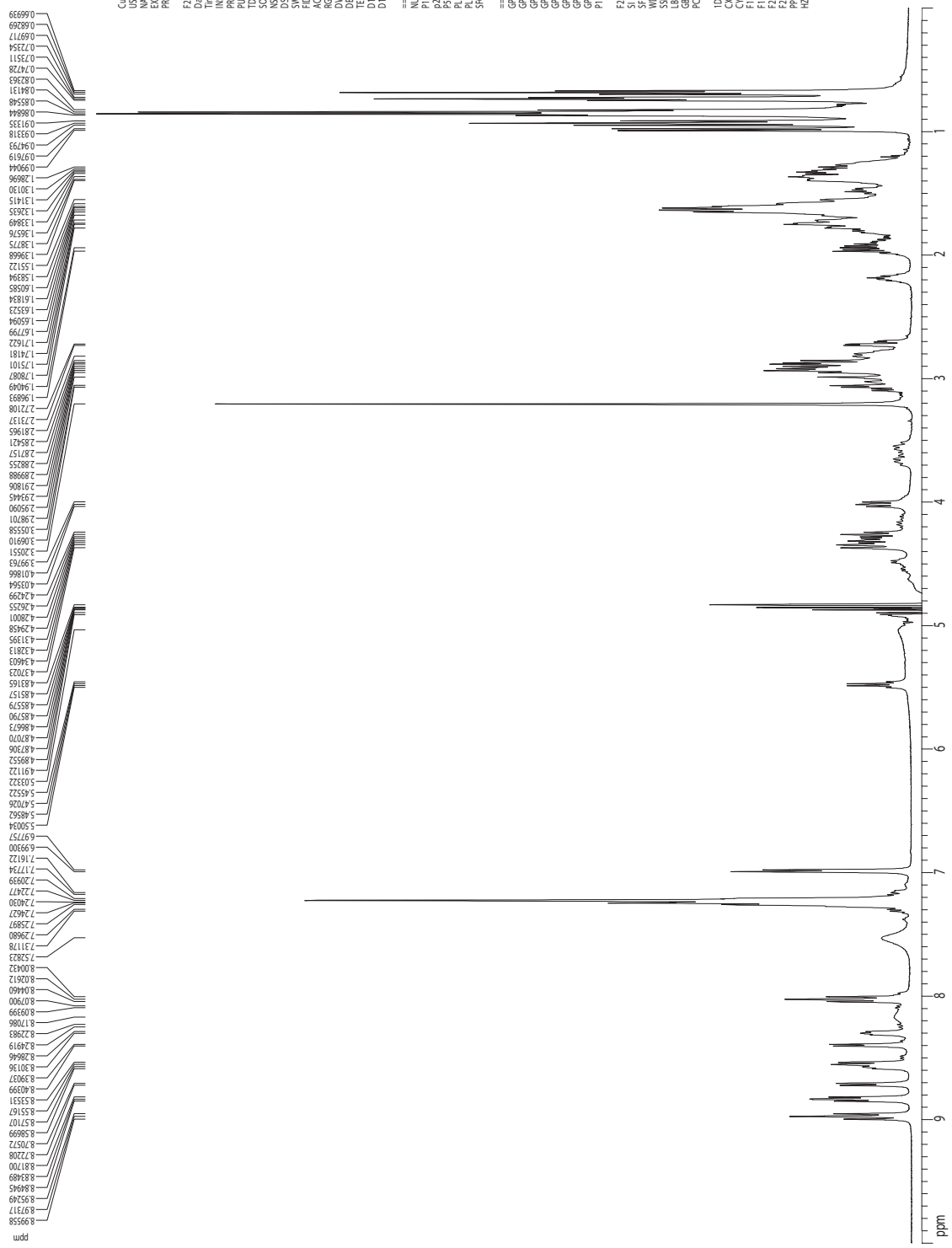
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 10.707922 Hz
SW 10.960 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2199738 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 1.40

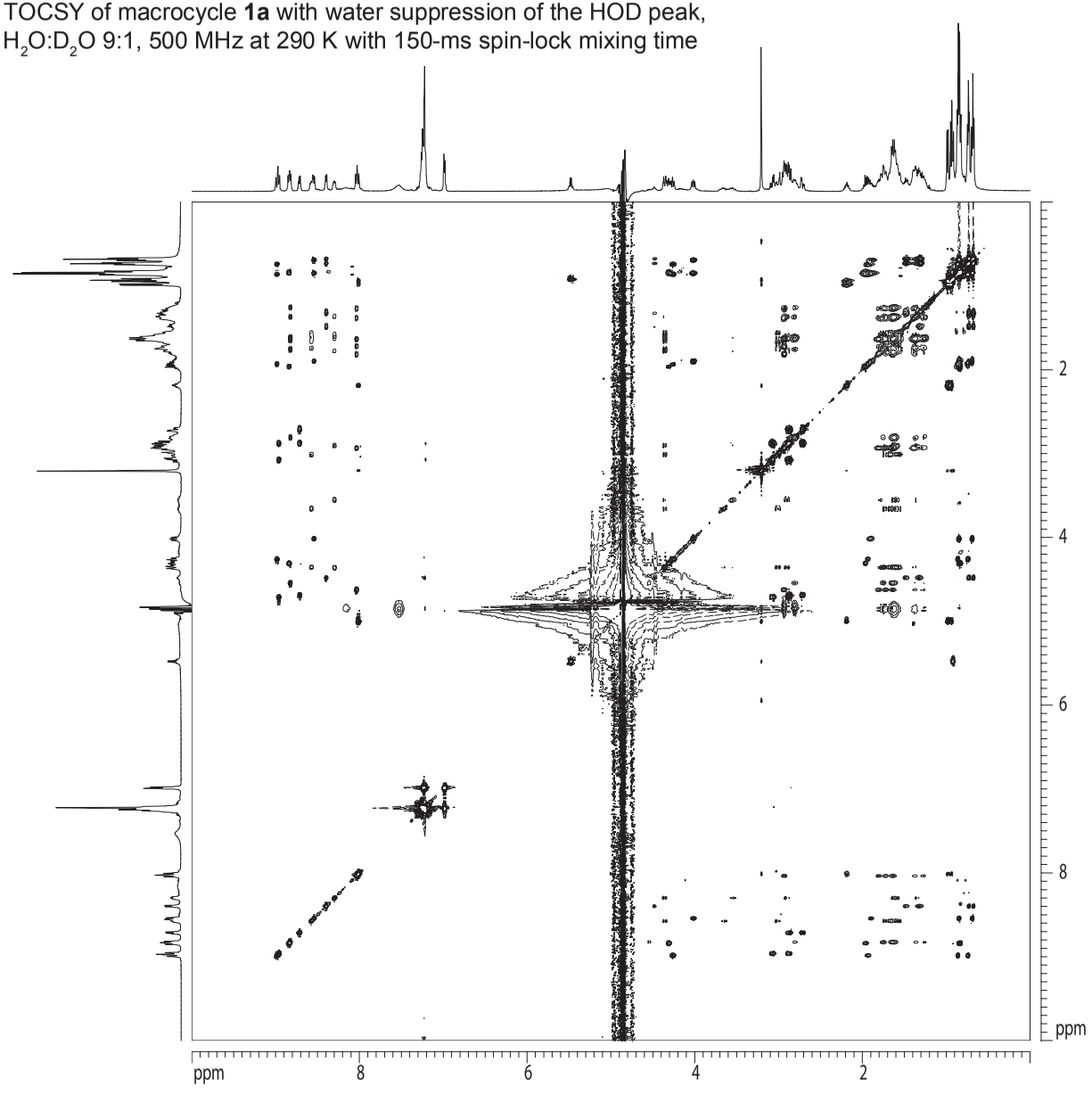
F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2199738 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 6.001 ppm
F2LO 3001.78 Hz
F2PHI 2.490 ppm
F2HI 1245.68 Hz
F1PLO 6.001 ppm
F1LO 3001.78 Hz
F1PHI 2.490 ppm
F1HI 1245.68 Hz
F2PPMCM 0.23404 ppm/cm
F2HZCM 117.07327 Hz/cm
F1PPMCM 0.23404 ppm/cm
F1HZCM 117.07327 Hz/cm

^{1}H NMR of macrocycle **1a** with water suppression, 2 mM in H_2O:D_2O 9:1, 500 MHz at 290K.



^{1}H NMR TOCSY of macrocycle **1a** with water suppression of the HOD peak, 2 mM in H_2O:D_2O 9:1, 500 MHz at 290 K with 150-ms spin-lock mixing time



Current Data Parameters

USER	spence
NAME	KHC-I-63_H2O
EXPNO	4
PROCNO	1

F2 - Acquisition Parameters

Date_	20130327
Time	3.28
INSTRUM	cryo500
PROBHD	5 mm CPTC1 1H-
PULPROG	tocsytpnw.wu
TD	2048
SOLVENT	D2O
NS	16
DS	16
SWH	5482.456 Hz
FIDRES	2.676980 Hz
AQ	0.1868276 sec
RG	228
DW	91.200 usec
DE	6.00 usec
TE	290.3 K
d0	0.00000300 sec
D1	2.00000000 sec
d8	0.14985408 sec
D16	0.00075000 sec
IN0	0.00009120 sec
L1	64

====== CHANNEL f1 ======

NUC1	1H
P1	7.50 usec
p2	15.00 usec
p5	23.34 usec
P6	35.00 usec
p7	70.00 usec
P15	2200.00 usec
P17	2500.00 usec
PL1	1.60 dB
PL2	45.60 dB
PL10	15.20 dB
SFO1	500.2224014 MHz

===== GRADIENT CHANNEL ====

GPNAM1	sine.100
GPNAM2	sine.100
GPX1	0.00 %
GPX2	0.00 %
GPY1	0.00 %
GPY2	0.00 %
GPZ1	20.00 %
GPZ2	15.00 %
P16	1000.00 usec

F1 - Acquisition parameters

ND0	2
TD	512
SFO1	500.2224 MHz
FIDRES	10.707922 Hz
SW	10.960 ppm
FnMODE	TPPI

F2 - Processing parameters

SI	1024
SF	500.2199739 MHz
WDW	QSINE
SSB	3
LB	0.00 Hz
GB	0
PC	1.40

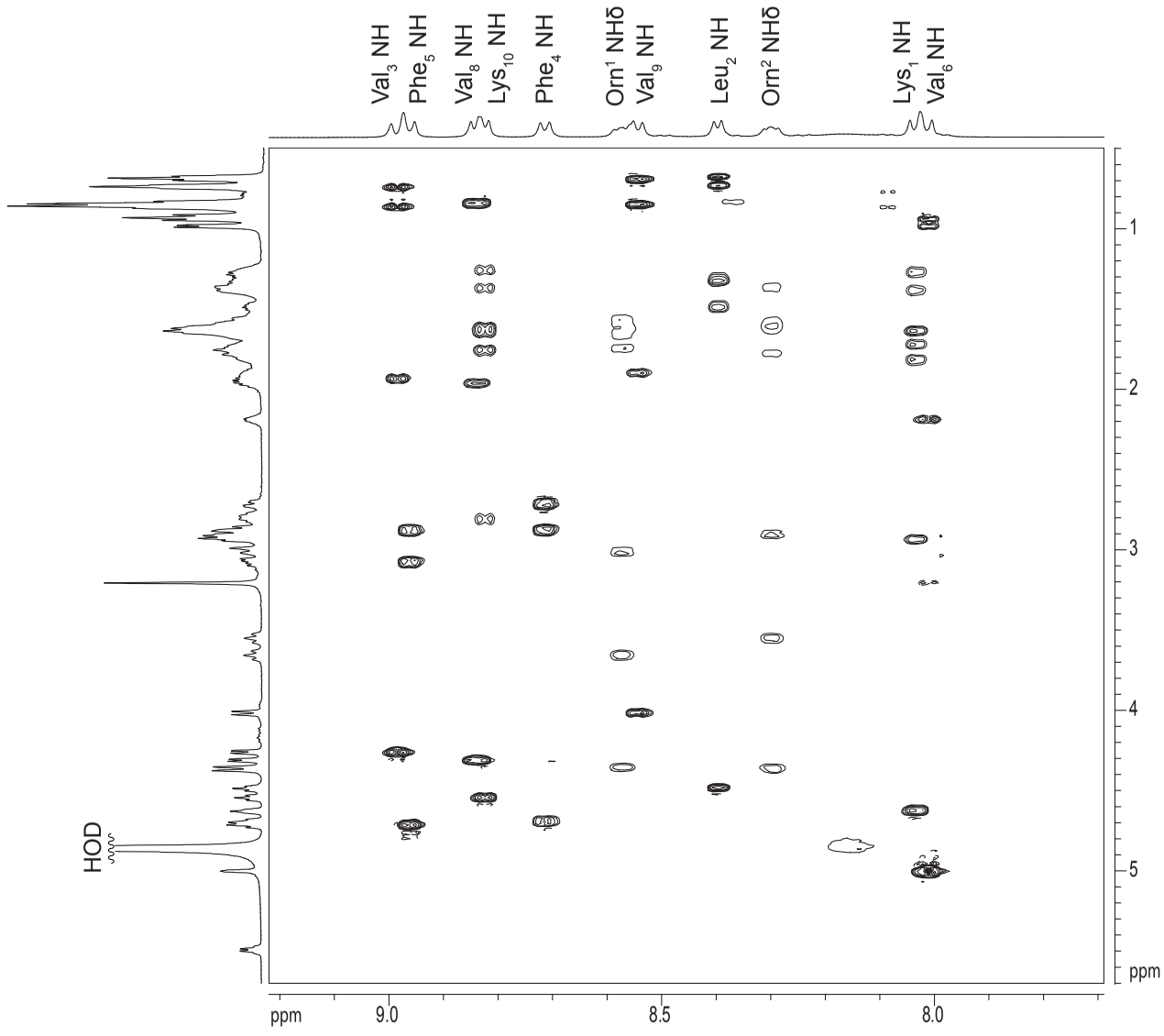
F1 - Processing parameters

SI	1024
MC2	TPPI
SF	500.2199739 MHz
WDW	QSINE
SSB	3
LB	0.00 Hz
GB	0

2D NMR plot parameters

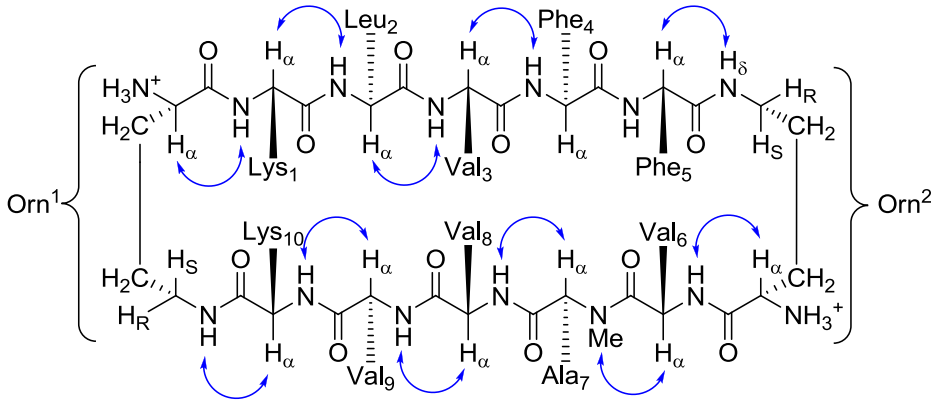
CX2	15.00 cm
CX1	15.00 cm
F2PLO	10.000 ppm
F2LO	5002.20 Hz
F2PHI	0.000 ppm
F2HI	0.00 Hz
F1PLO	10.000 ppm
F1LO	5002.20 Hz
F1PHI	0.000 ppm
F1HI	0.00 Hz
F2PPMCM	0.66667 ppm/cm
F2HZCM	333.47998 Hz/cm
F1PPMCM	0.66667 ppm/cm
F1HZCM	333.47998 Hz/cm

^{1}H NMR TOCSY of macrocycle **1a** with water suppression of the HOD peak, 2 mM in H_2O:D_2O 9:1, 500 MHz at 290 K with 150-ms spin-lock mixing time




```
Current Data Parameters
USER      spence
NAME      KHC-I-63_H2O
EXPNO     4
PROCNO    1

F2 - Acquisition Parameters
Date_     20130327
Time      3.28
INSTRUM   cryo500
PROBHD  5 mm CPTCI 1H-
PULPROG   tocsytpnw.wu
TD        2048
SOLVENT   D2O
NS        16
DS        16
SWH       5482.456 Hz
FIDRES    2.676980 Hz
AQ        0.1868276 sec
RG        228
DW        91.200 usec
DE        6.00 usec
TE        290.3 K
d0        0.00000300 sec
D1        2.00000000 sec
d8        0.14985408 sec
D16       0.00075000 sec
IN0       0.00009120 sec
L1        64

======== CHANNEL f1 =======
NUC1      1H
P1        7.50 usec
p2        15.00 usec
p5        23.34 usec
P6        35.00 usec
p7        70.00 usec
P15       2200.00 usec
P17       2500.00 usec
PL1       1.60 dB
PL2       45.60 dB
PL10      15.20 dB
SFO1      500.2224014 MHz

====== GRADIENT CHANNEL =====
GPNAM1    sine.100
GPNAM2    sine.100
GPX1      0.00 %
GPX2      0.00 %
GPY1      0.00 %
GPY2      0.00 %
GPZ1      20.00 %
GPZ2      15.00 %
P16       1000.00 usec

F1 - Acquisition parameters
ND0       2
TD        512
SFO1      500.2224 MHz
FIDRES    10.707922 Hz
SW        10.960 ppm
FnMODE    TPPI

F2 - Processing parameters
SI        1024
SF        500.2199739 MHz
WDW       QSINE
SSB       3
LB        0.00 Hz
GB        0
PC        1.40

F1 - Processing parameters
SI        1024
MC2       TPPI
SF        500.2199738 MHz
WDW       QSINE
SSB       3
LB        0.00 Hz
GB        0

2D NMR plot parameters
CX2       15.00 cm
CX1       15.00 cm
F2PLO     9.220 ppm
F2LO      4611.90 Hz
F2PHI     7.689 ppm
F2HI      3846.28 Hz
F1PLO     5.700 ppm
F1LO      2851.25 Hz
F1PHI     0.497 ppm
F1HI      248.55 Hz
F2PPMCM   0.10204 ppm/cm
F2HZCM    51.04110 Hz/cm
F1PPMCM   0.34687 ppm/cm
F1HZCM    173.51373 Hz/cm
```

Sequential inter-residue NH-α NOE correlations observed in macrocyclic β-sheet **1a**. NOEs were identified in ^{1}H NMR ROESYspectrum with water suppression of the HOD peak in 2 mM H_2O:D_2O (9:1) solution at 290 K. An inter-residue NOE correlation between Phe_4 and Phe_5 could not be observed due to overlap.

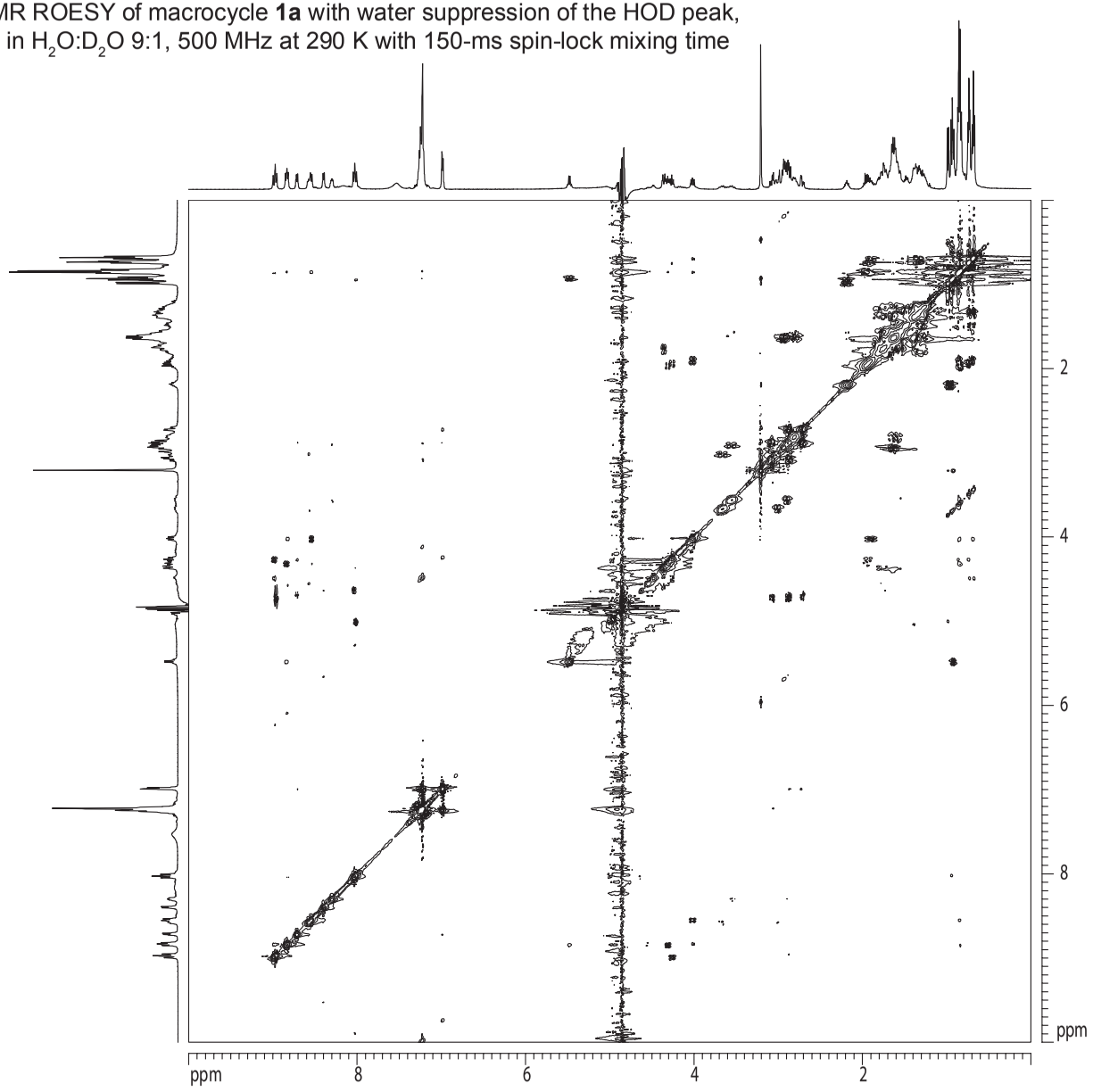


Summary of $^{3}J_{NH\text{-}\alpha}$ are shown in the table below:

Residue	$^{3}J_{NH\text{-}\alpha}$
Lys_1	10.9 Hz
Leu_2	6.8 Hz
Val_3	11.2 Hz
Phe_4	8.2 Hz
Phe_5	10.3 Hz
Val_6	10.9 Hz
Val_8	9.2 Hz
Val_9	8.4 Hz
Lys_{10}	9.0 Hz

The large $^{3}J_{NH\text{-}\alpha}$ coupling constants and sequential NH-α NOEs is consistent with a β-strand conformation.

^{1}H NMR ROESY of macrocycle **1a** with water suppression of the HOD peak, 2 mM in H_2O:D_2O 9:1, 500 MHz at 290 K with 150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME KHC-I-63_H2O
EXPNO 3
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130326
Time 22.02
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesytp.2nw.wu15
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 5482.456 Hz
FIDRES 2.676980 Hz
AQ 0.1868276 sec
RG 1024
DW 91.200 usec
DE 6.00 usec
TE 290.4 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15155201 sec
D16 0.00020000 sec
IN0 0.00009120 sec
L4 512
P15 74.00 usec

======== CHANNEL f1 ========
NUC1 1H
P1 7.50 usec
p2 15.00 usec
P5 2200.00 usec
p25 148.00 usec
PL1 1.60 dB
PL2 45.60 dB
PL11 26.60 dB
SFO1 500.2224014 MHz

====== GRADIENT CHANNEL ====
GPNAM1 SINE.100
GPNAM2 SINE.100
GPX1 0.00 %
GPX2 0.00 %
GPY1 0.00 %
GPY2 0.00 %
GPZ1 20.00 %
GPZ2 15.00 %
P16 1000.00 usec

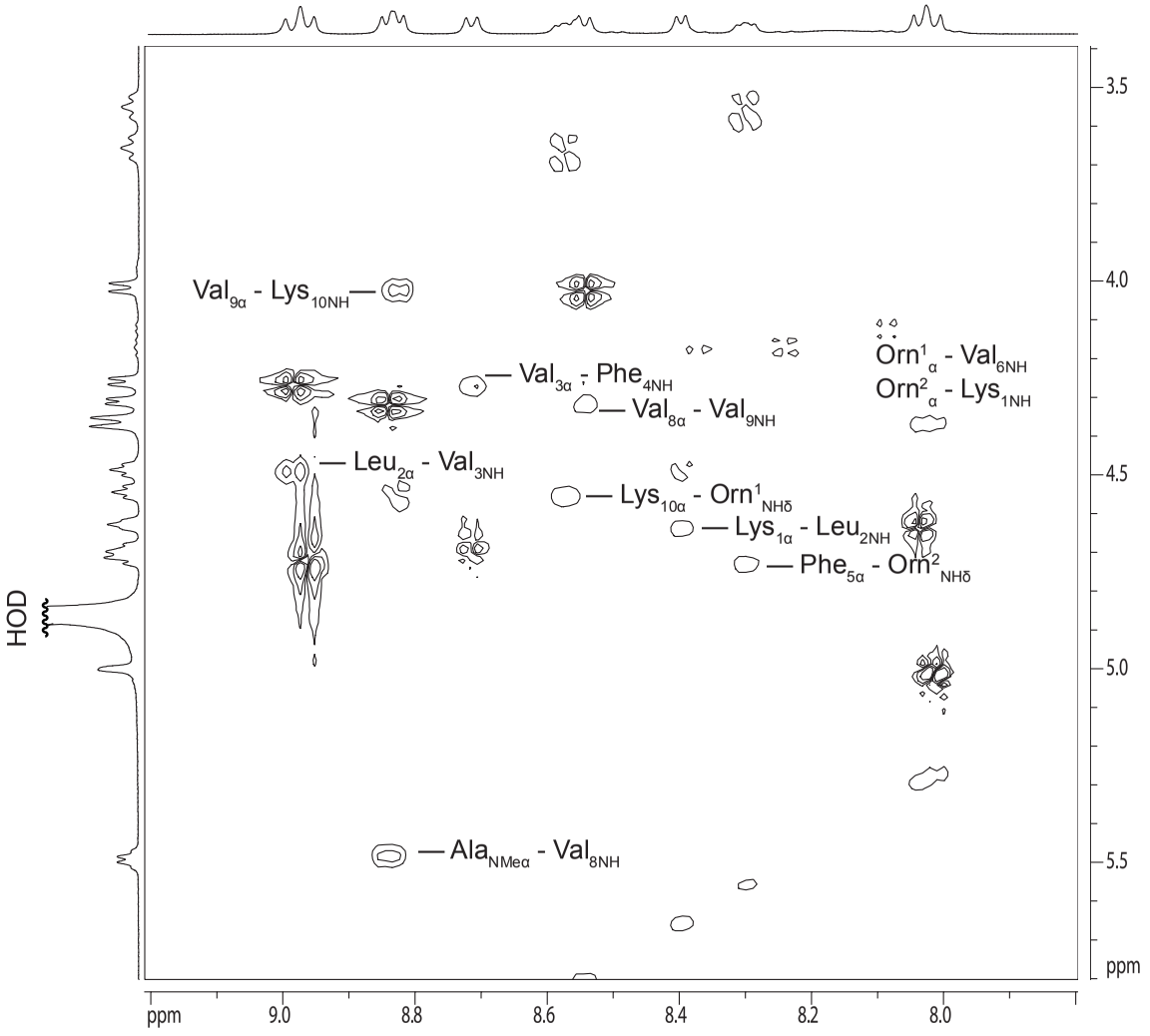
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 10.707922 Hz
SW 10.960 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2199739 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 1.00

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2199739 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 10.000 ppm
F2LO 5002.20 Hz
F2PHI 0.000 ppm
F2HI 0.00 Hz
F1PLO 10.000 ppm
F1LO 5002.20 Hz
F1PHI 0.000 ppm
F1HI 0.00 Hz
F2PPMCM 0.66667 ppm/cm
F2HZCM 333.47998 Hz/cm
F1PPMCM 0.66667 ppm/cm
F1HZCM 333.47998 Hz/cm

^{1}H NMR ROESY of macrocycle **1a** with water suppression of the HOD peak,
2 mM in H_2O:D_2O 9:1, 500 MHz at 290 K with 150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME KHC-I-63_H2O
EXPNO 3
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130326
Time 22.02
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesytp2nw.wu15
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 5482.456 Hz
FIDRES 2.676980 Hz
AQ 0.18682760 sec
RG 1024
DW 91.200 usec
DE 6.00 usec
TE 290.4 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15155201 sec
D16 0.00020000 sec
IN0 0.00009120 sec
L4 512
P15 74.00 usec

====== CHANNEL f1 ======
NUC1 1H
P1 7.50 usec
p2 15.00 usec
P5 2200.00 usec
p25 148.00 usec
PL1 1.60 dB
PL2 45.60 dB
PL11 26.60 dB
SFO1 500.2224014 MHz

===== GRADIENT CHANNEL ====
GPNAM1 SINE.100
GPNAM2 SINE.100
GPX1 0.00 %
GPX2 0.00 %
GPY1 0.00 %
GPY2 0.00 %
GPZ1 20.00 %
GPZ2 15.00 %
P16 1000.00 usec

F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 10.707922 Hz
SW 10.960 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2199739 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 1.00

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2199738 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 9.209 ppm
F2LO 4606.54 Hz
F2PHI 7.796 ppm
F2HI 3899.82 Hz
F1PLO 5.802 ppm
F1LO 2902.26 Hz
F1PHI 3.394 ppm
F1HI 1697.62 Hz
F2PPMCM 0.09419 ppm/cm
F2HZCM 47.11486 Hz/cm
F1PPMCM 0.160055 ppm/cm
F1HZCM 80.30941 Hz/cm

Diffusion-Ordered Spectroscopy (DOSY) Experiments and Diffusion Constant Measurements of Macrocycle 1a

DOSY experiments comprised of a series of 16 or 32 pulsed field gradient spin-echo experiments with a diffusion delay (Δ) of 75-ms and a diffusion gradient length (δ) of 2.5-ms in which the gradient strength was incremented to allow ca. 2%–95% signal attenuation with a linear ramp.[1] Data were processed to give a pseudo-2D spectrum using Topspin, and the diffusion coefficient was read from the spectrum after phasing and converted from the logarithmic value to the linear value. Calibration of the gradients was performed on the basis of the diffusion coefficient of residual HOD in D_2O (1.9×10^{-9} m^2/s at 298 K).[2]

Diffusion Coefficients (*D*) of macrocycle **1a** in D_2O at 298K.

macrocycle	*D*	concentration
1a	20.8×10^{-11} m^2/s	1 mM
1a	21.8×10^{-11} m^2/s	10 mM

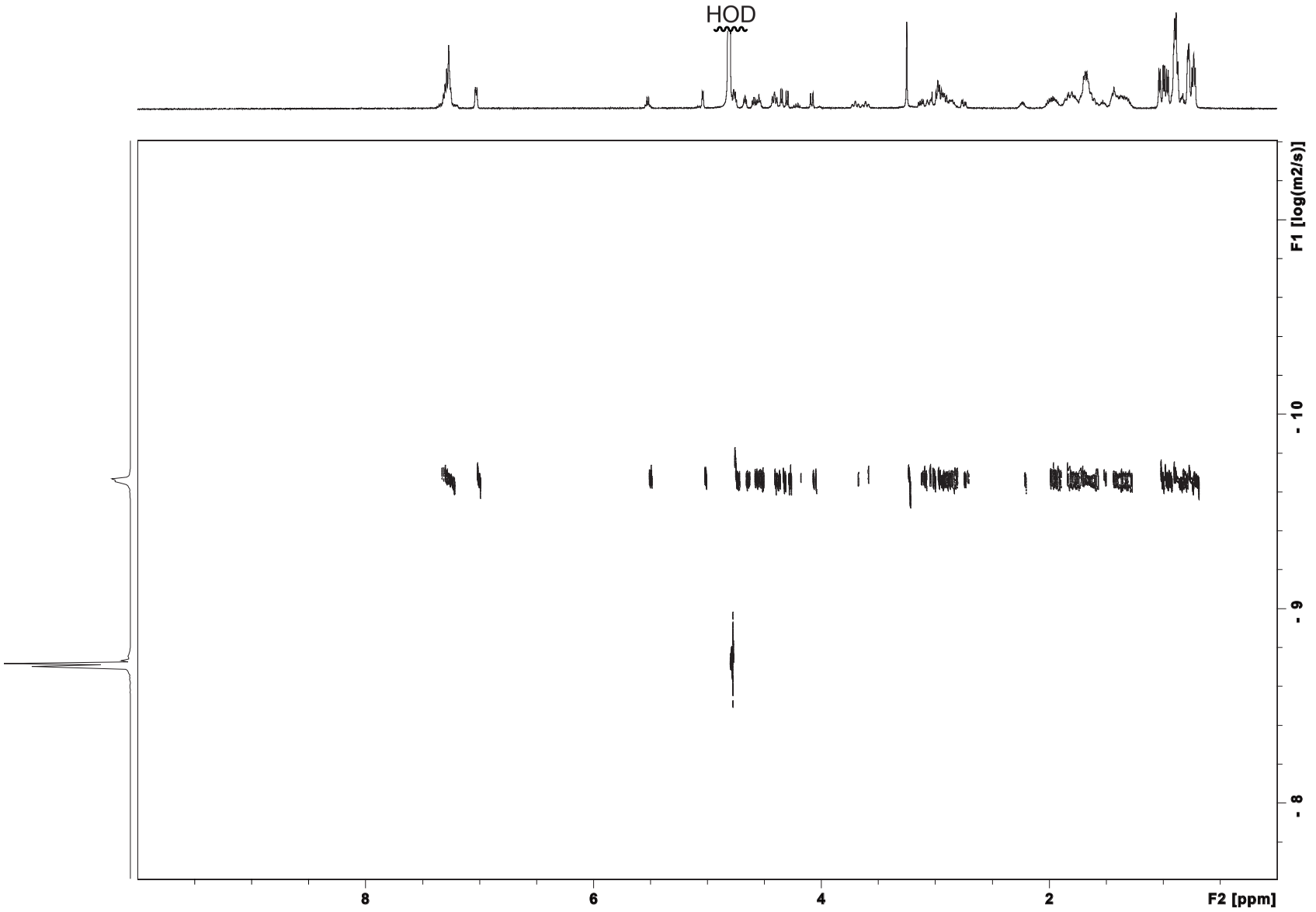
The diffusion coefficients do not vary significantly with concentrations over 1-10 mM and are comparable with those observed for similar peptide macrocycles.[3] Collectively these observations reflect that little or no self-association of macrocycle **1a** occurs at low millimolar concentrations.

[1] Wu, D.; Chen, A.; Johnson Jr., C.S. J. Magn. Reson. **1995**, *A115*, 260–264.

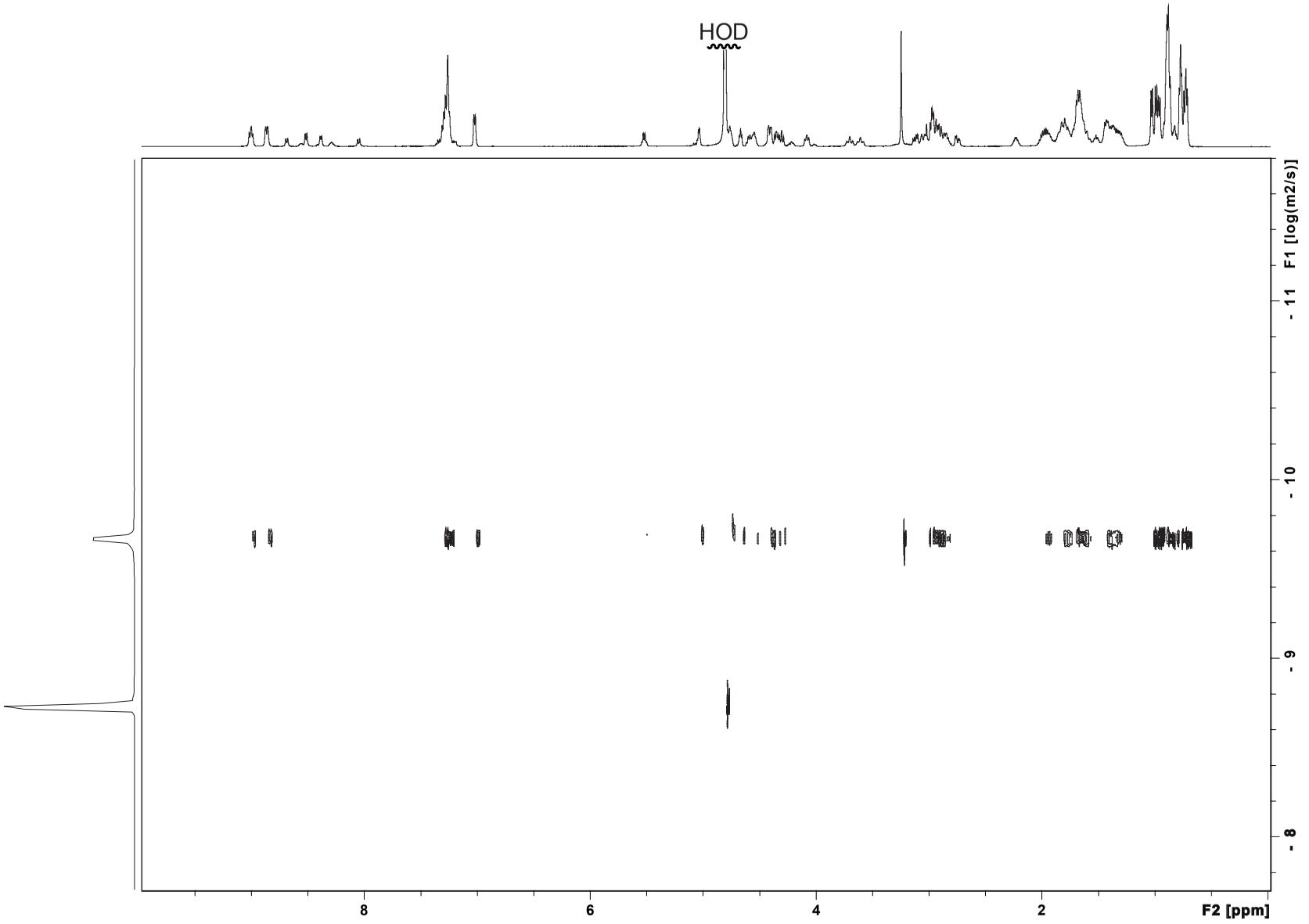
[2] Longsworth, L. G. J. Phys. Chem. **1960**, *64*, 1914–1917.

[3] Khakshoor, O.; Demeler, B.; Nowick, J. S. J. Am. Chem. Soc. **2007**, *129*, 5558-5569.

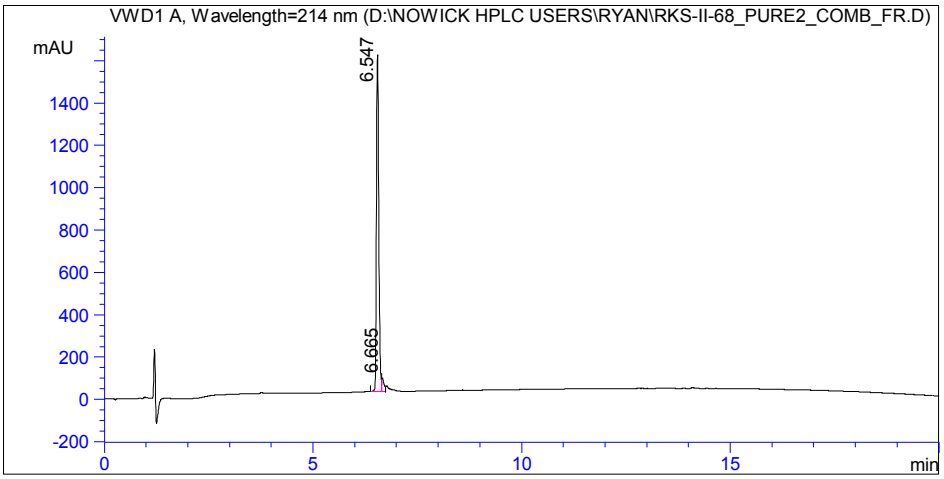
^{1}H NMR DOSY of macrocycle **1a,** 1 mM in D_2O, 500 MHz at 298 K



^{1}H NMR DOSY of macrocycle **1a,** 10 mM in D_2O, 500 MHz at 298 K

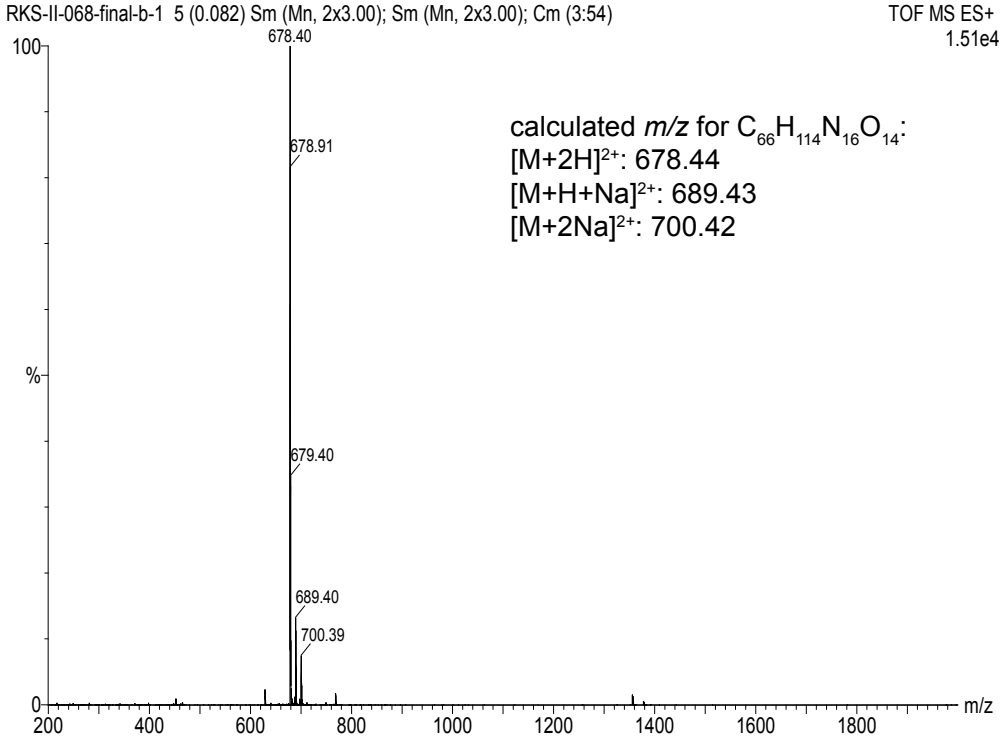


HPLC and MS ESI$^+$ TOF of macrocycle **1b**

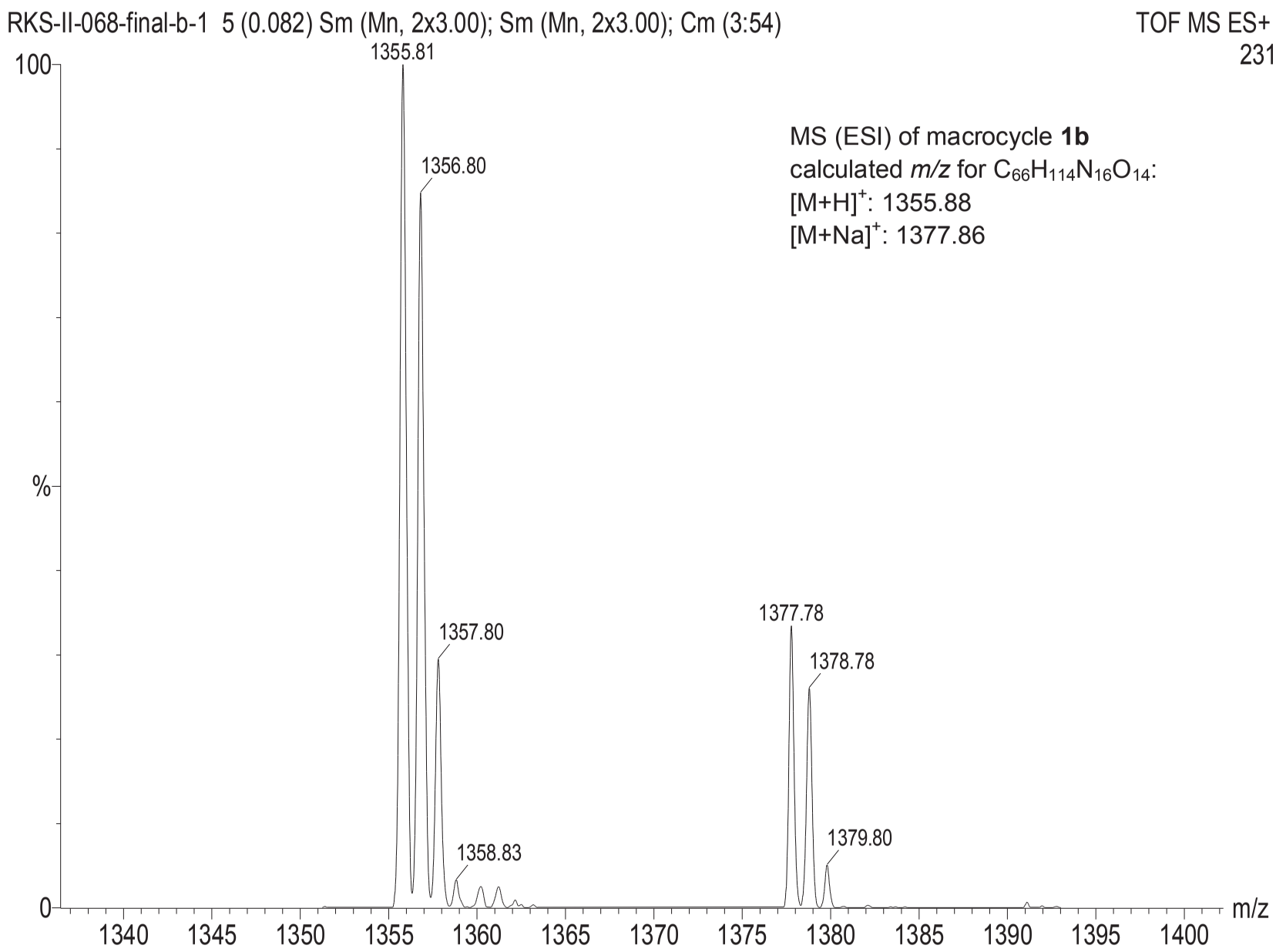


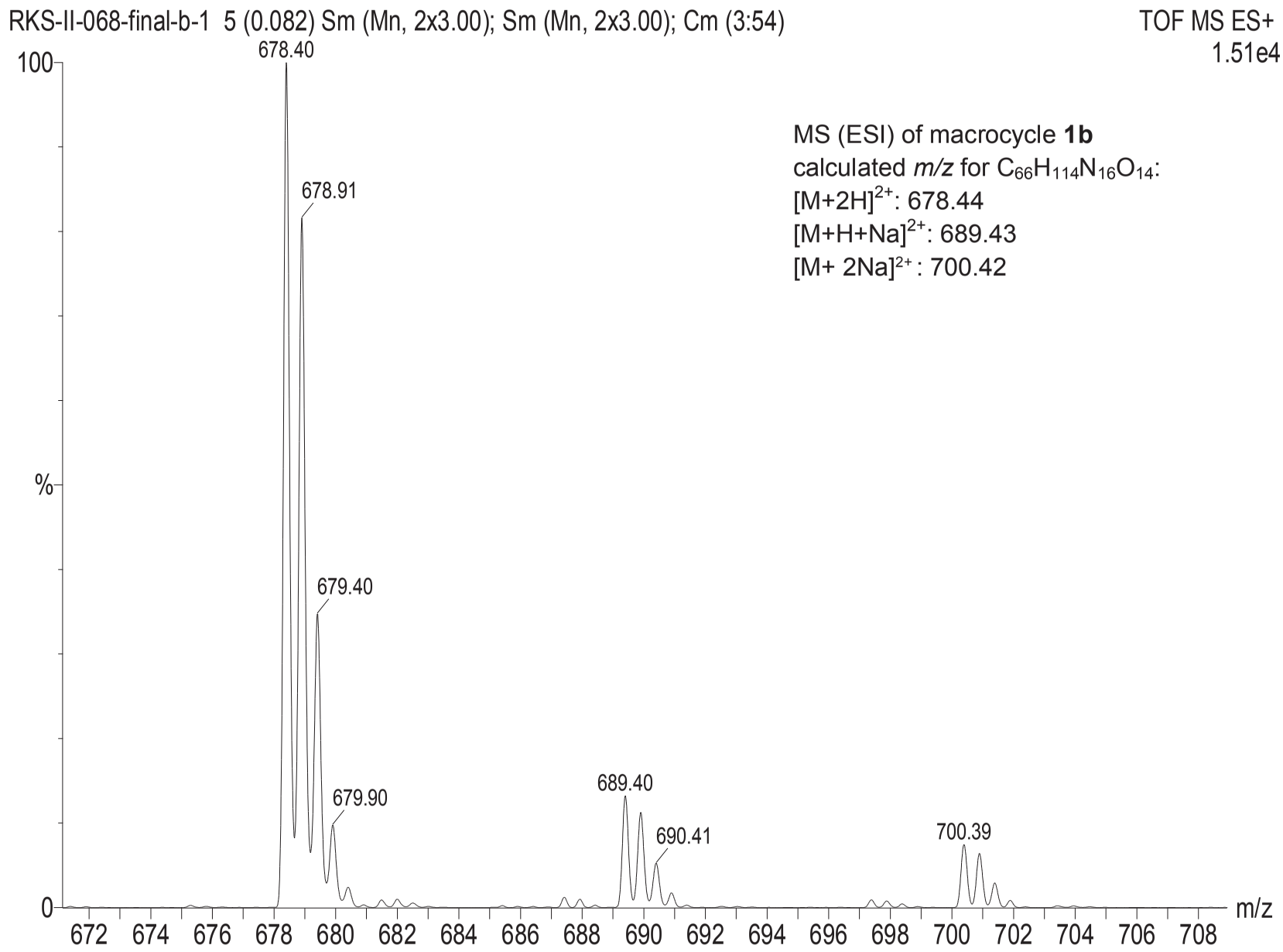
Signal 1:VWD1 A, Wavelength=214 nm

Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	6.547	BV	0.061	6176.373	96.229	96.205
2	6.665	VV	0.056	243.620	3.771	3.795

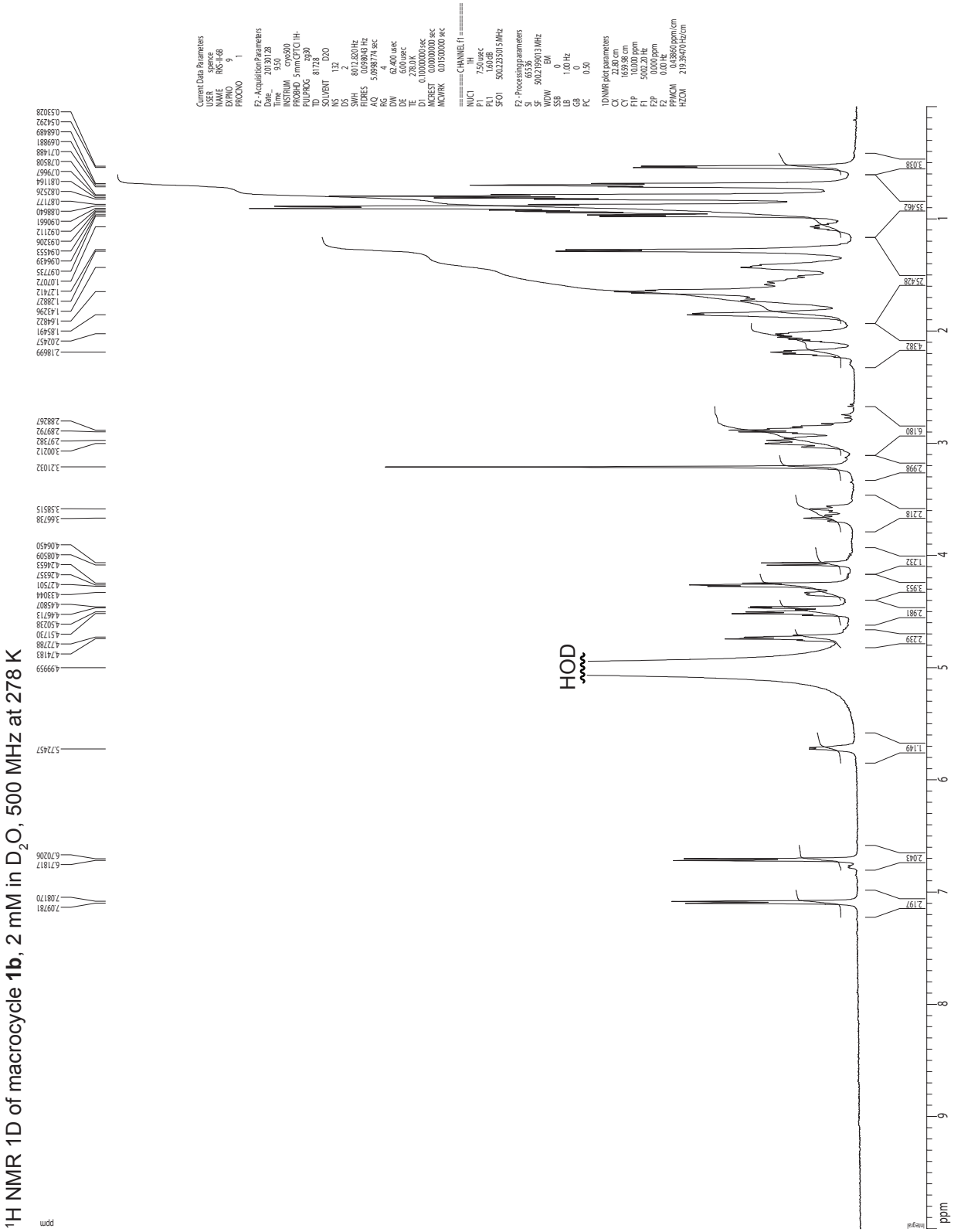


calculated *m/z* for $C_{66}H_{114}N_{16}O_{14}$:
$[M+2H]^{2+}$: 678.44
$[M+H+Na]^{2+}$: 689.43
$[M+2Na]^{2+}$: 700.42


RKS-II-068-final-b-1 5 (0.082) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:54)
TOF MS ES+
231
100
%
0
1355.81
1356.80
1357.80
1358.83
1377.78
1378.78
1379.80
MS (ESI) of macrocycle 1b
calculated m/z for C66H114N16O14:
[M+H]+: 1355.88
[M+Na]+: 1377.86
1340
1345
1350
1355
1360
1365
1370
1375
1380
1385
1390
1395
1400
m/z


RKS-II-068-final-b-1 5 (0.082) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:54)
TOF MS ES+
1.51e4
MS (ESI) of macrocycle 1b
calculated m/z for C66H114N16O14:
[M+2H]2+: 678.44
[M+H+Na]2+: 689.43
[M+ 2Na]2+: 700.42
678.40
678.91
679.40
679.90
689.40
690.41
700.39
100
%
0
m/z
672 674 676 678 680 682 684 686 688 690 692 694 696 698 700 702 704 706 708

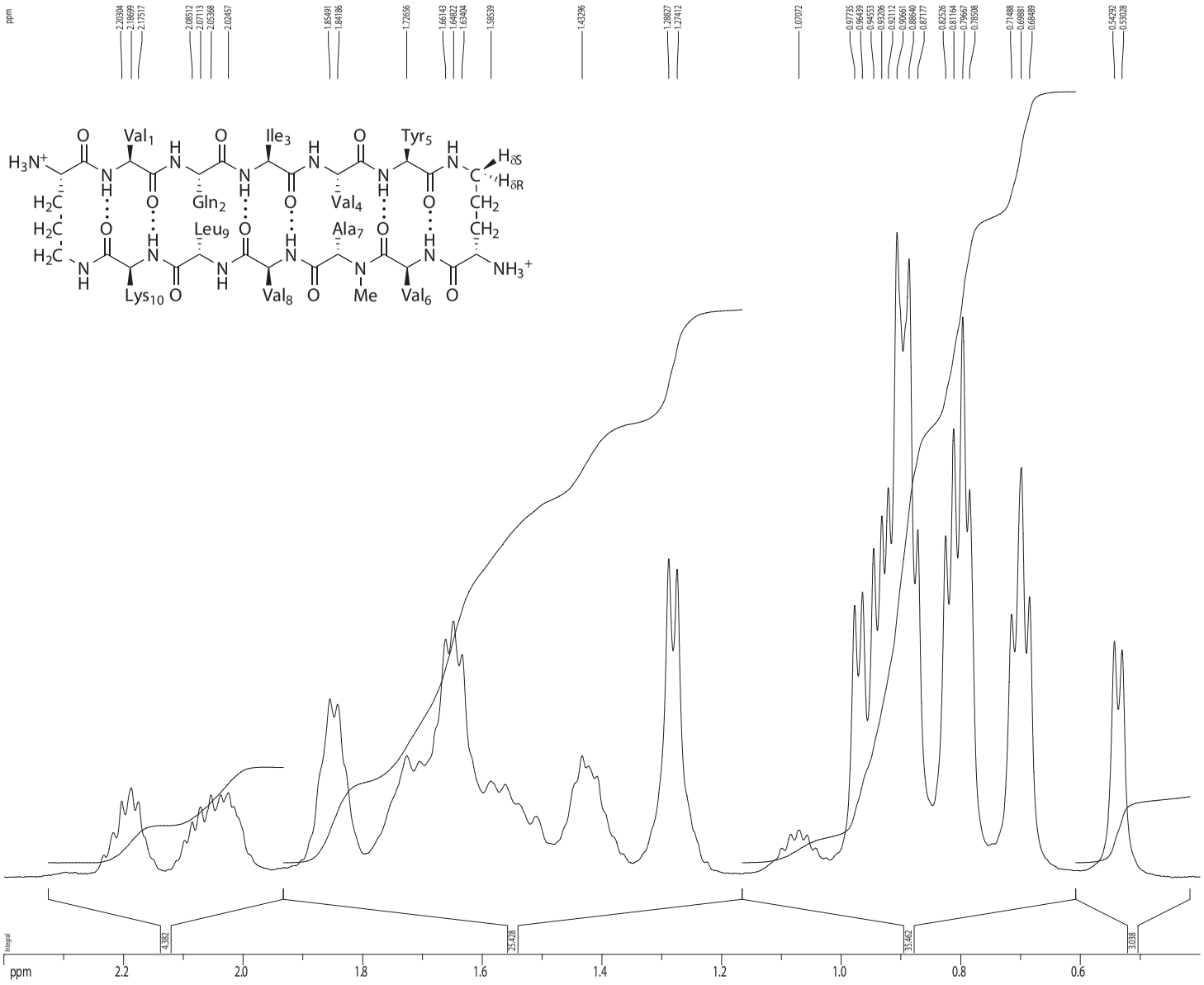
^{1}H NMR 1D of macrocycle **1b**, 2 mM in D_2O, 500 MHz at 278 K



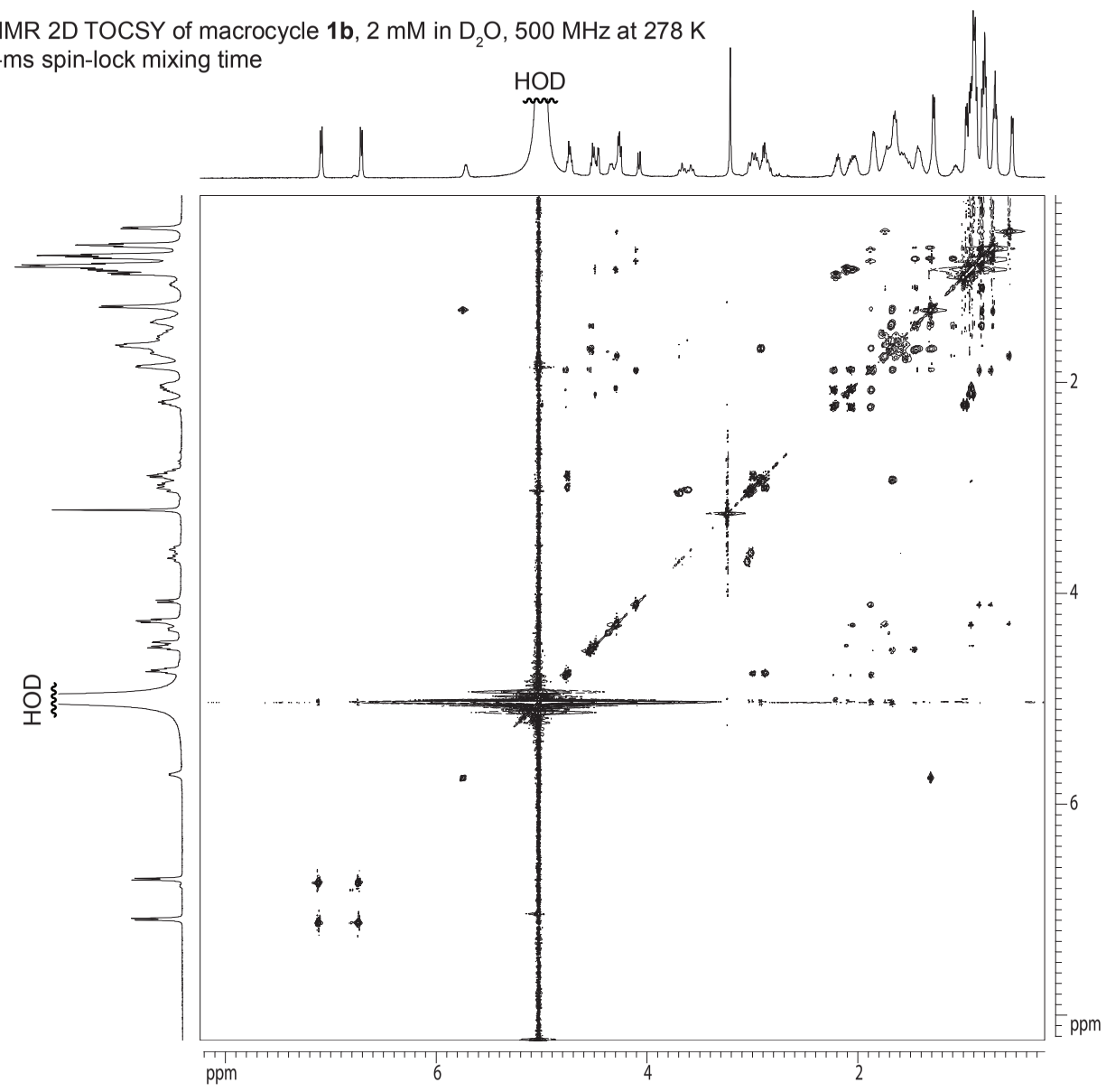
^{1}H NMR 1D of macrocycle **1b**, 2 mM in D_2O, 500 MHz at 278 K

^{1}H NMR 1D of macrocycle **1b**, 2 mM in D_2O, 500 MHz at 278 K

^{1}H NMR 1D of macrocycle **1b**, 2 mM in D_2O, 500 MHz at 278 K



^{1}H NMR 2D TOCSY of macrocycle **1b**, 2 mM in D_2O, 500 MHz at 278 K
150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME RKS-II-68
EXPNO 4
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130111
Time 21.43
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG tocsytp.wu
TD 2048
SOLVENT D2O
NS 8
DS 16
SWH 4006.410 Hz
FIDRES 1.956255 Hz
AQ 0.2556404 sec
RG 4.1
DW 124.800 usec
DE 6.00 usec
TE 278.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.02084341 sec
d12 0.00002000 sec
IN0 0.00012480 sec
L1 7

======= CHANNEL f1 =======
NUC1 1H
P1 7.50 usec
p5 23.34 usec
P6 35.00 usec
p7 70.00 usec
P17 2500.00 usec
PL1 1.60 dB
PL10 15.20 dB
SFO1 500.2220009 MHz

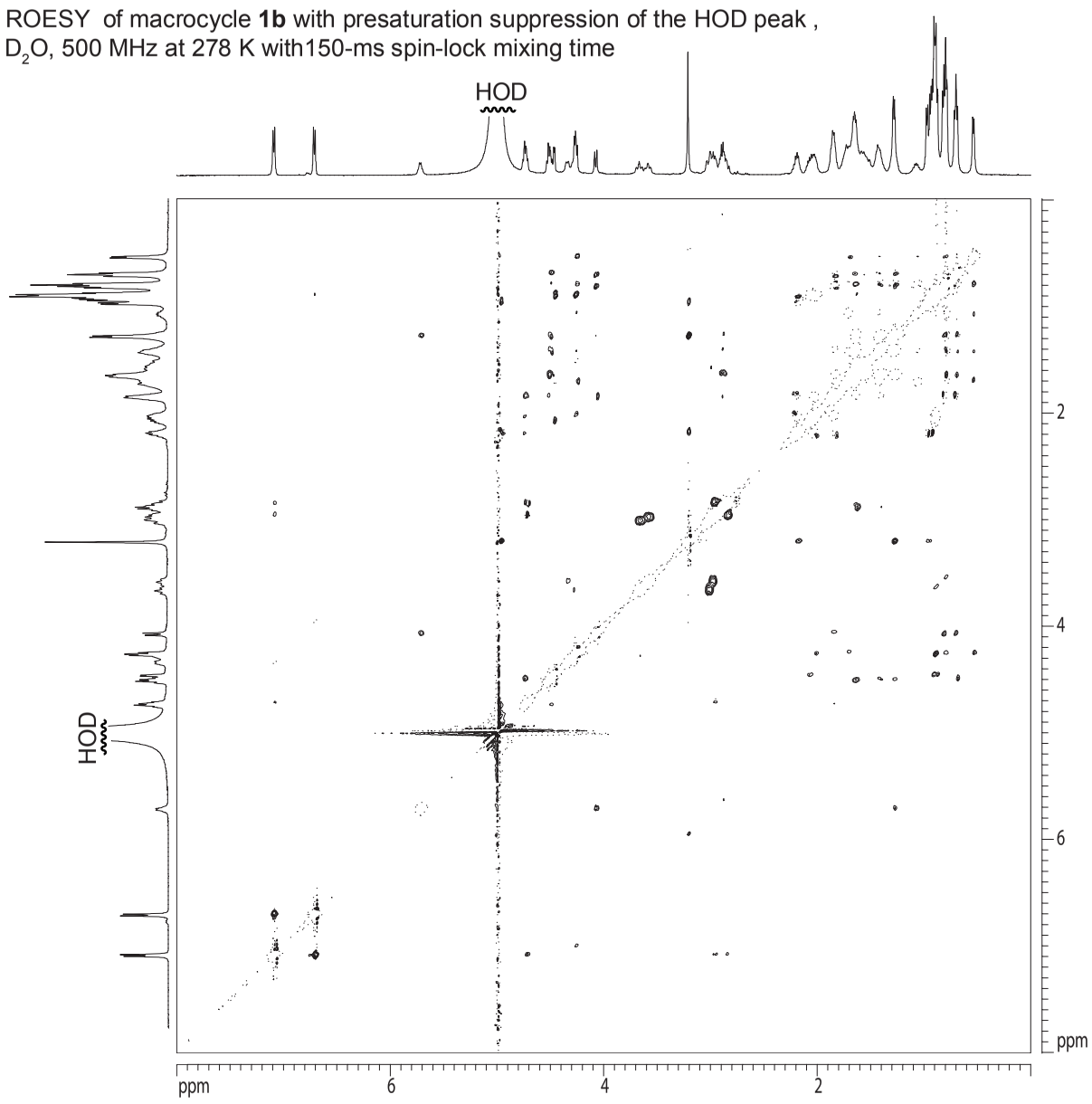
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.222 MHz
FIDRES 7.825020 Hz
SW 8.009 ppm
FnMODE undefined

F2 - Processing parameters
SI 1024
SF 500.2199013 MHz
WDW QSINE
SSB 3
LB 0.00 Hz
GB 0
PC 4.00

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2199013 MHz
WDW QSINE
SSB 3
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 8.238 ppm
F2LO 4120.97 Hz
F2PHI 0.229 ppm
F2HI 114.56 Hz
F1PLO 8.238 ppm
F1LO 4120.97 Hz
F1PHI 0.229 ppm
F1HI 114.56 Hz
F2PPMCM 0.53395 ppm/cm
F2HZCM 267.09402 Hz/cm
F1PPMCM 0.53395 ppm/cm
F1HZCM 267.09402 Hz/cm

^{1}H NMR ROESY of macrocycle **1b** with presaturation suppression of the HOD peak , 2 mM in D_2O, 500 MHz at 278 K with150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME RKS-II-68
EXPNO 5
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130112
Time 0.22
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesyprtp.2.prd
TD 2048
SOLVENT D2O
NS 8
DS 16
SWH 5482.456 Hz
FIDRES 2.676980 Hz
AQ 0.1868276 sec
RG 4.1
DW 91.200 usec
DE 6.00 usec
TE 278.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15000001 sec
d11 0.03000000 sec
d12 0.00002000 sec
d13 0.00000300 sec
IN0 0.0000912 0 sec
L4 300

======= CHANNEL f1 =======
NUC1 1H
P1 7.50 usec
P25 250.00 usec
PL1 1.60 dB
PL9 54.00 dB
PL11 26.60 dB
SFO1 500.2224049 MHz

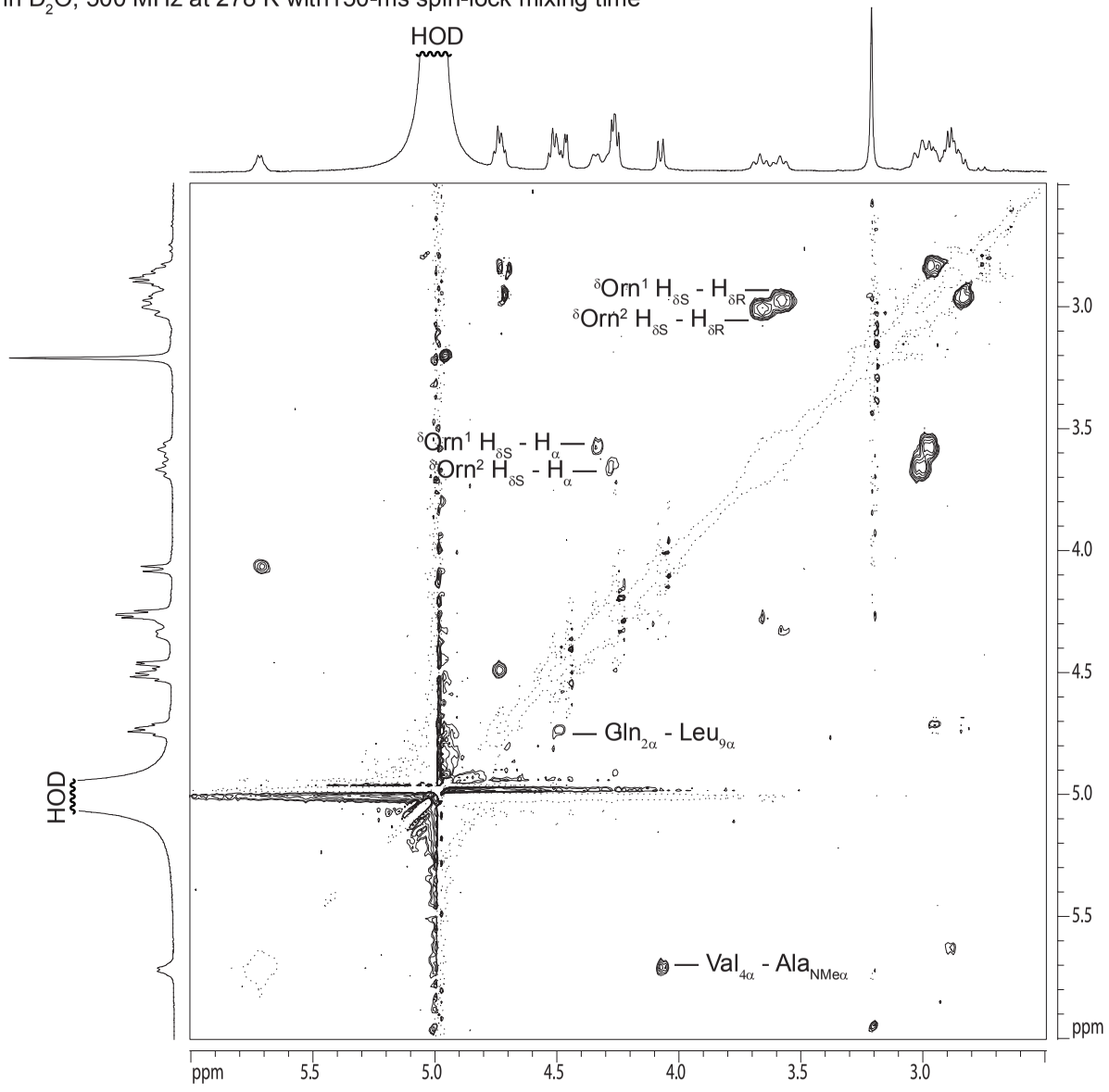
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 10.707922 Hz
SW 10.960 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2199013 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 1.40

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2199013 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

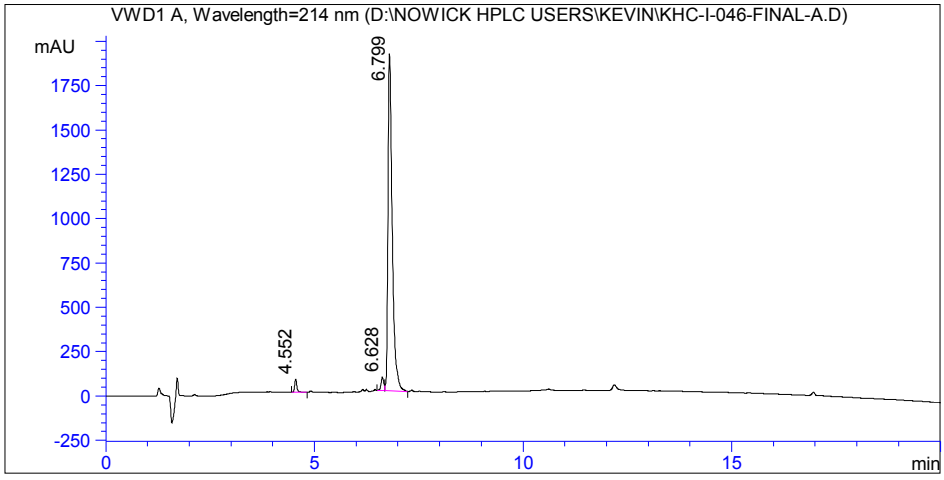
2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 8.000 ppm
F2LO 4001.76 Hz
F2PHI 0.000 ppm
F2HI 0.00 Hz
F1PLO 8.006 ppm
F1LO 4004.84 Hz
F1PHI -0.011 ppm
F1HI -5.28 Hz
F2PPMCM 0.53333 ppm/cm
F2HZCM 266.78397 Hz/cm
F1PPMCM 0.53445 ppm/cm
F1HZCM 267.34116 Hz/cm

^{1}H NMR ROESY of macrocycle **1b** with presaturation suppression of the HOD peak , 2 mM in D_2O, 500 MHz at 278 K with150-ms spin-lock mixing time



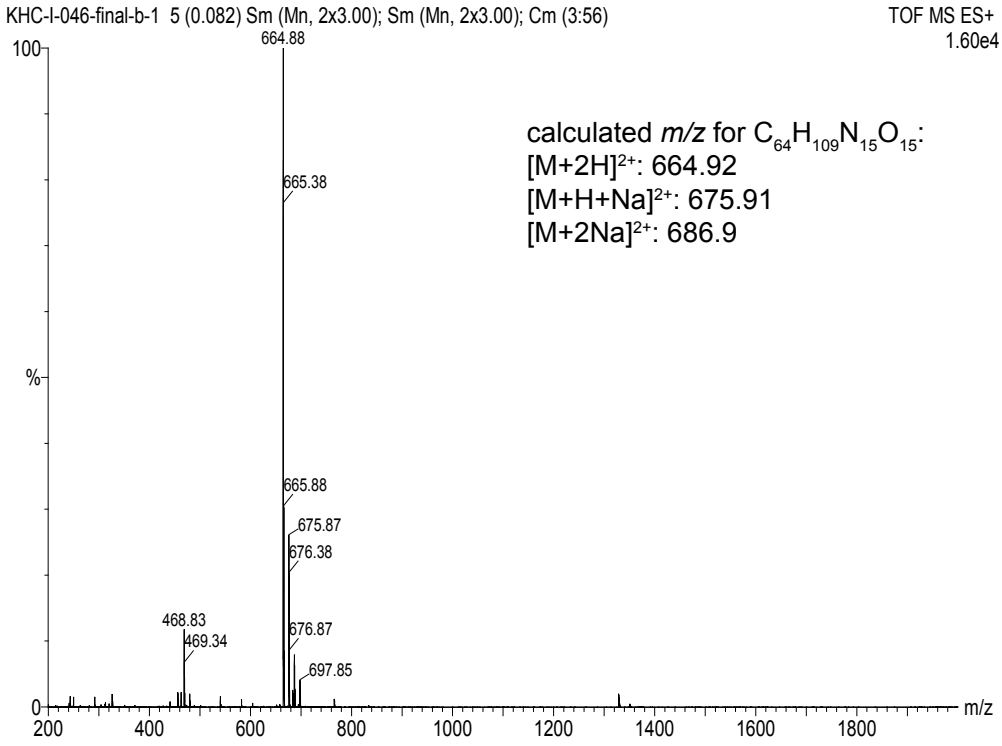

```
Current Data Parameters
USER        spence
NAME        RKS-II-68
EXPNO       5
PROCNO      1

F2 - Acquisition Parameters
Date_       20130112
Time        0.22
INSTRUM     cryo500
PROBHD  5 mm CPTCI 1H-
PULPROG  roesyprtp.2.prd
TD          2048
SOLVENT     D2O
NS          8
DS          16
SWH         5482.456 Hz
FIDRES      2.676980 Hz
AQ          0.1868276 sec
RG          4.1
DW          91.200 usec
DE          6.00 usec
TE          278.0 K
d0          0.00000300 sec
D1          2.00000000 sec
d8          0.15000001 sec
d11         0.03000000 sec
d12         0.00002000 sec
d13         0.00000300 sec
IN0         0.00009120 sec
L4          300

======= CHANNEL f1 =======
NUC1        1H
P1          7.50 usec
P25         250.00 usec
PL1         1.60 dB
PL9         54.00 dB
PL11        26.60 dB
SFO1        500.2224049 MHz

F1 - Acquisition parameters
ND0         2
TD          512
SFO1        500.2224 MHz
FIDRES      10.707922 Hz
SW          10.960 ppm
FnMODE      TPPI

F2 - Processing parameters
SI          1024
SF          500.2199013 MHz
WDW         QSINE
SSB         2
LB          0.00 Hz
GB          0
PC          1.40

F1 - Processing parameters
SI          1024
MC2         TPPI
SF          500.2199013 MHz
WDW         QSINE
SSB         2
LB          0.00 Hz
GB          0

2D NMR plot parameters
CX2         15.00 cm
CX1         15.00 cm
F2PLO       6.005 ppm
F2LO        3003.65 Hz
F2PHI       2.494 ppm
F2HI        1247.55 Hz
F1PLO       6.005 ppm
F1LO        3003.65 Hz
F1PHI       2.494 ppm
F1HI        1247.55 Hz
F2PPMCM     0.23404 ppm/cm
F2HZCM      117.07329 Hz/cm
F1PPMCM     0.23404 ppm/cm
F1HZCM      117.07329 Hz/cm
```

HPLC and MS ESI$^+$ TOF of macrocycle **1c**

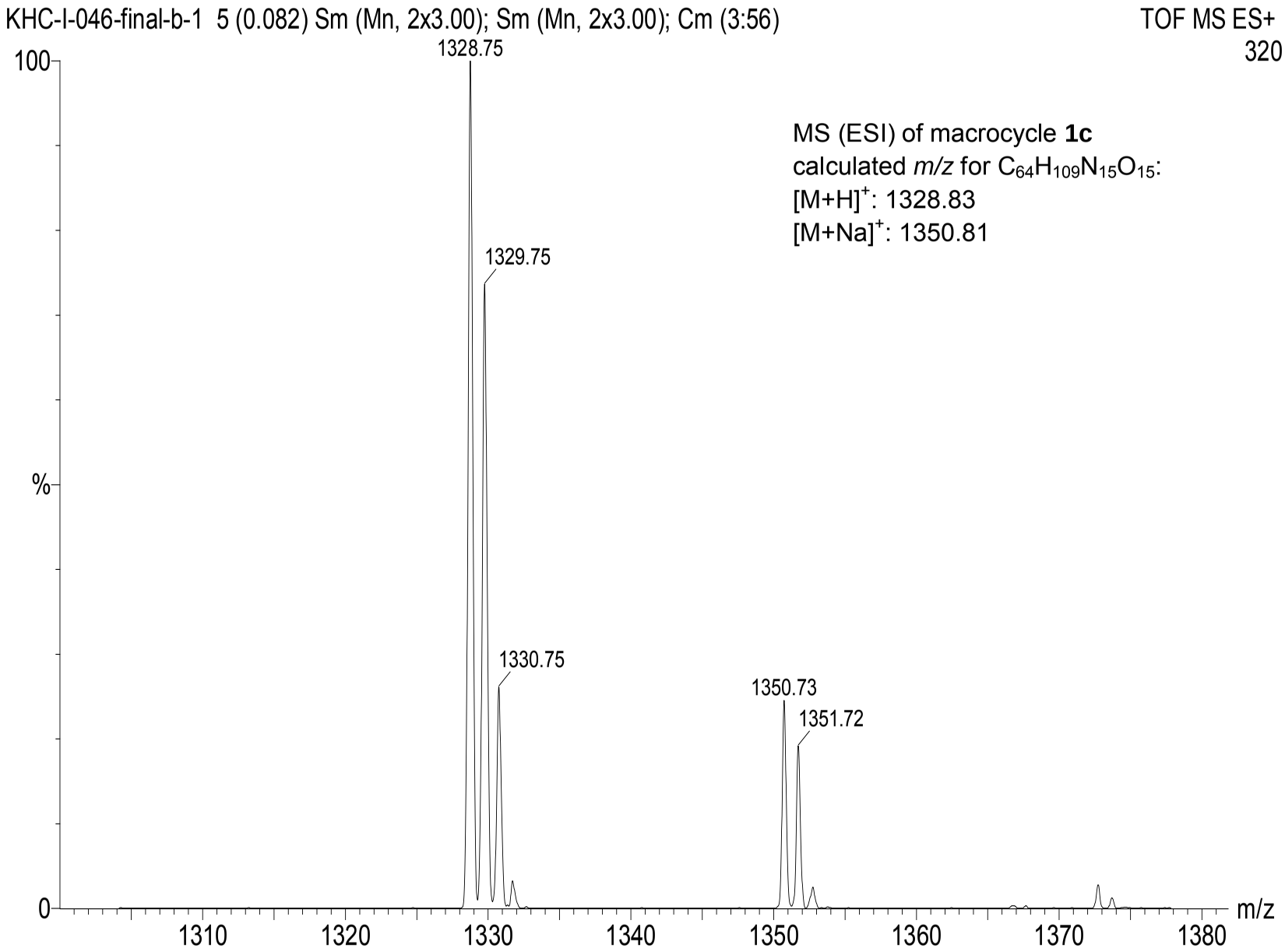


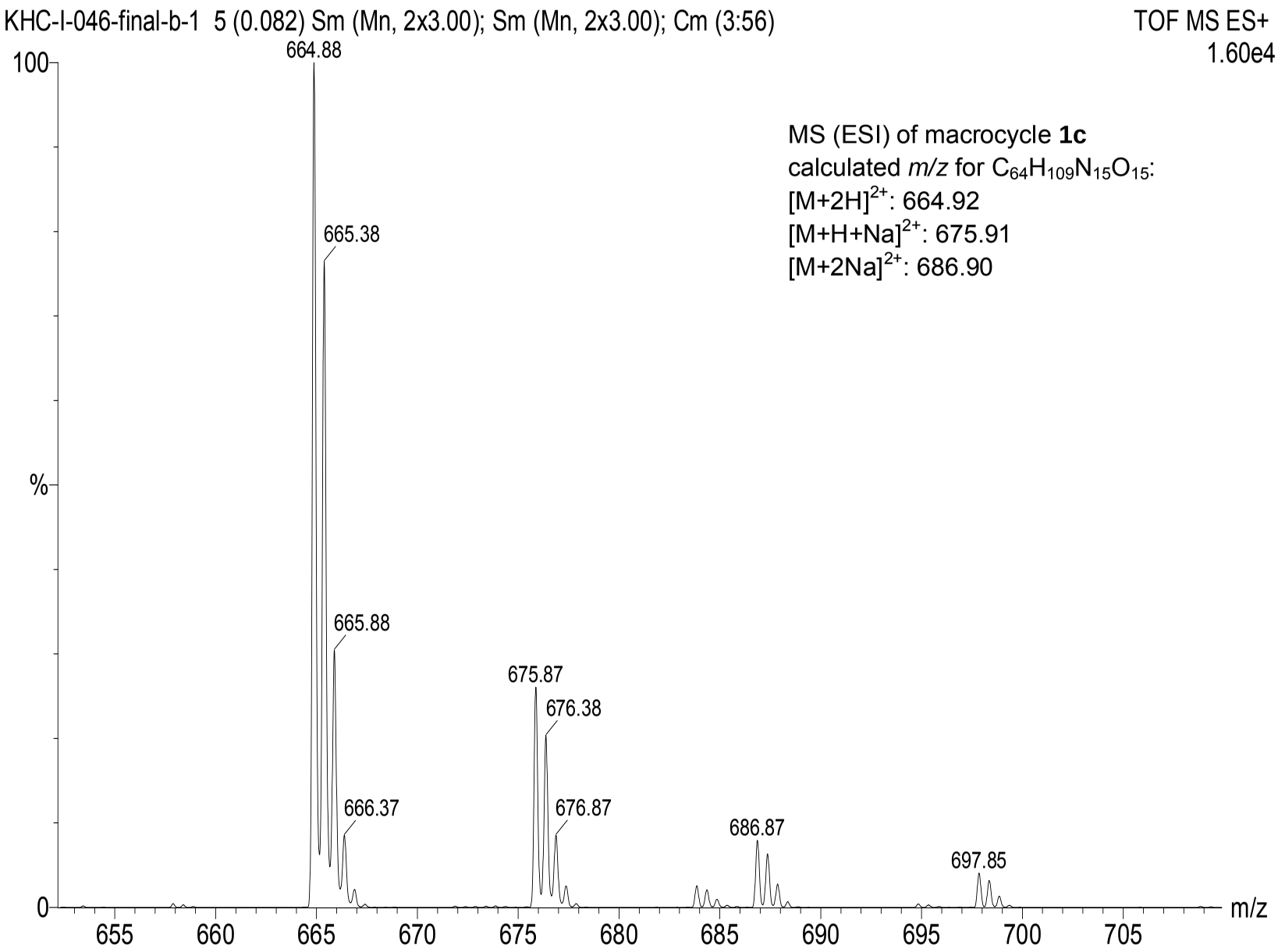
Signal 1:VWD1 A, Wavelength=214 nm

Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	4.552	BV	0.058	283.207	3.591	2.009
2	6.628	BV	0.069	345.408	3.699	2.450
3	6.799	VB	0.105	13467.859	92.710	95.541

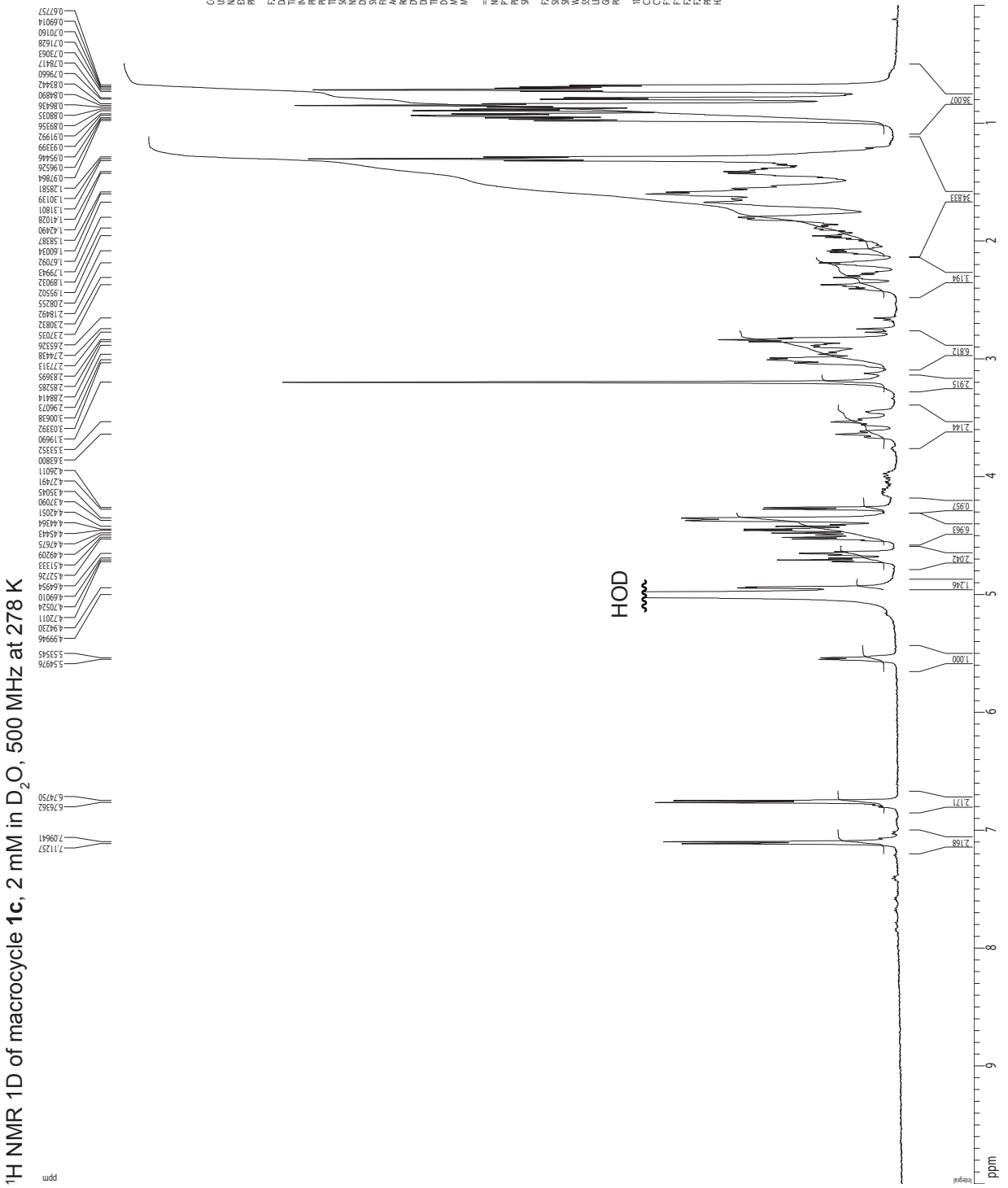


calculated *m/z* for $C_{64}H_{109}N_{15}O_{15}$:
$[M+2H]^{2+}$: 664.92
$[M+H+Na]^{2+}$: 675.91
$[M+2Na]^{2+}$: 686.9


KHC-I-046-final-b-1 5 (0.082) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:56)
TOF MS ES+
320
100
%
0
1328.75
1329.75
1330.75
1350.73
1351.72
MS (ESI) of macrocycle 1c
calculated m/z for C64H109N15O15:
[M+H]+: 1328.83
[M+Na]+: 1350.81
1310
1320
1330
1340
1350
1360
1370
1380
m/z


KHC-I-046-final-b-1 5 (0.082) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:56)
TOF MS ES+
1.60e4
100
%
0
664.88
665.38
665.88
666.37
675.87
676.38
676.87
686.87
697.85
655
660
665
670
675
680
685
690
695
700
705
m/z
MS (ESI) of macrocycle **1c**
calculated *m/z* for $C_{64}H_{109}N_{15}O_{15}$:
$[M+2H]^{2+}$: 664.92
$[M+H+Na]^{2+}$: 675.91
$[M+2Na]^{2+}$: 686.90

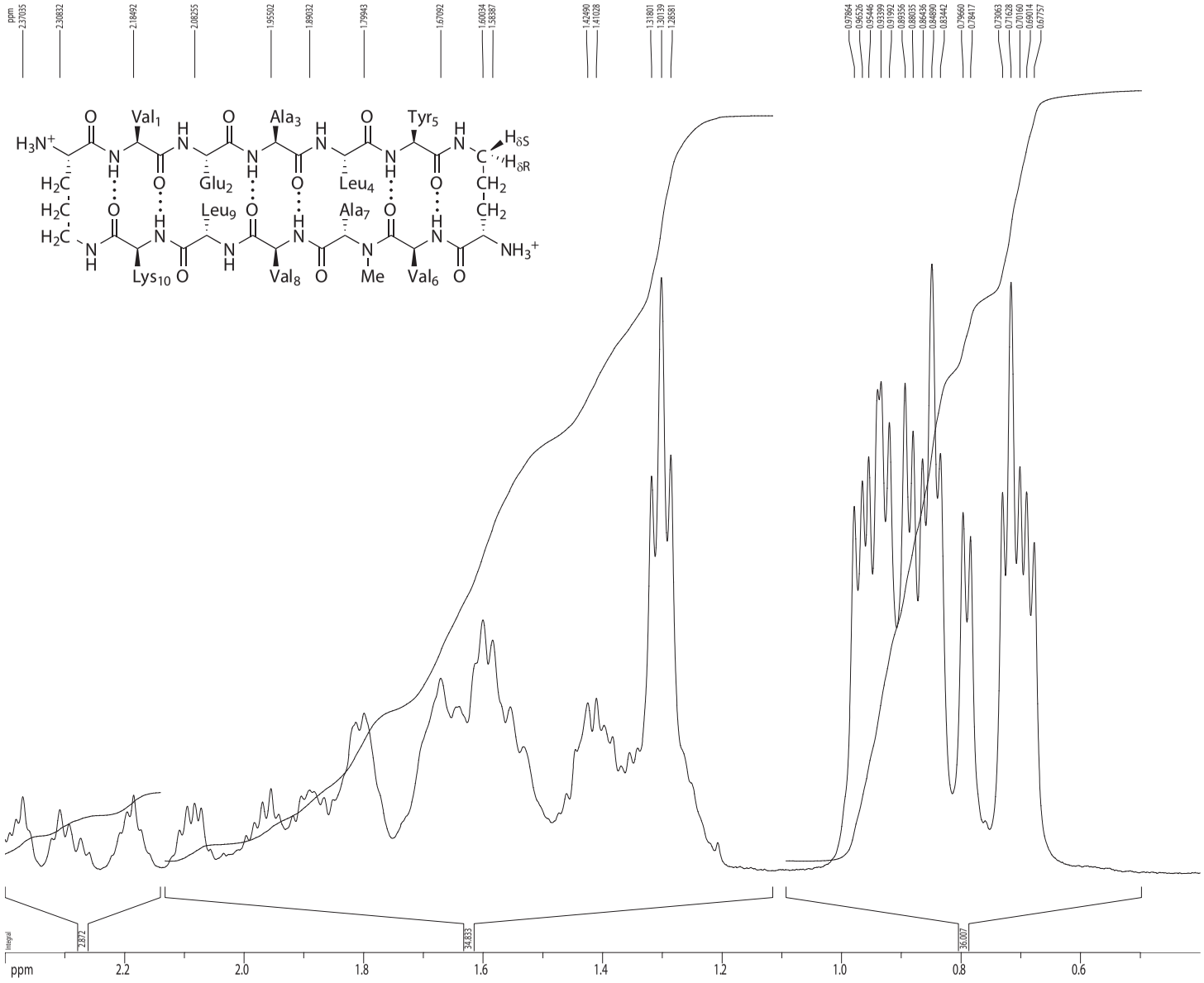
^{1}H NMR 1D of macrocycle **1c**, 2 mM in D_2O, 500 MHz at 278 K



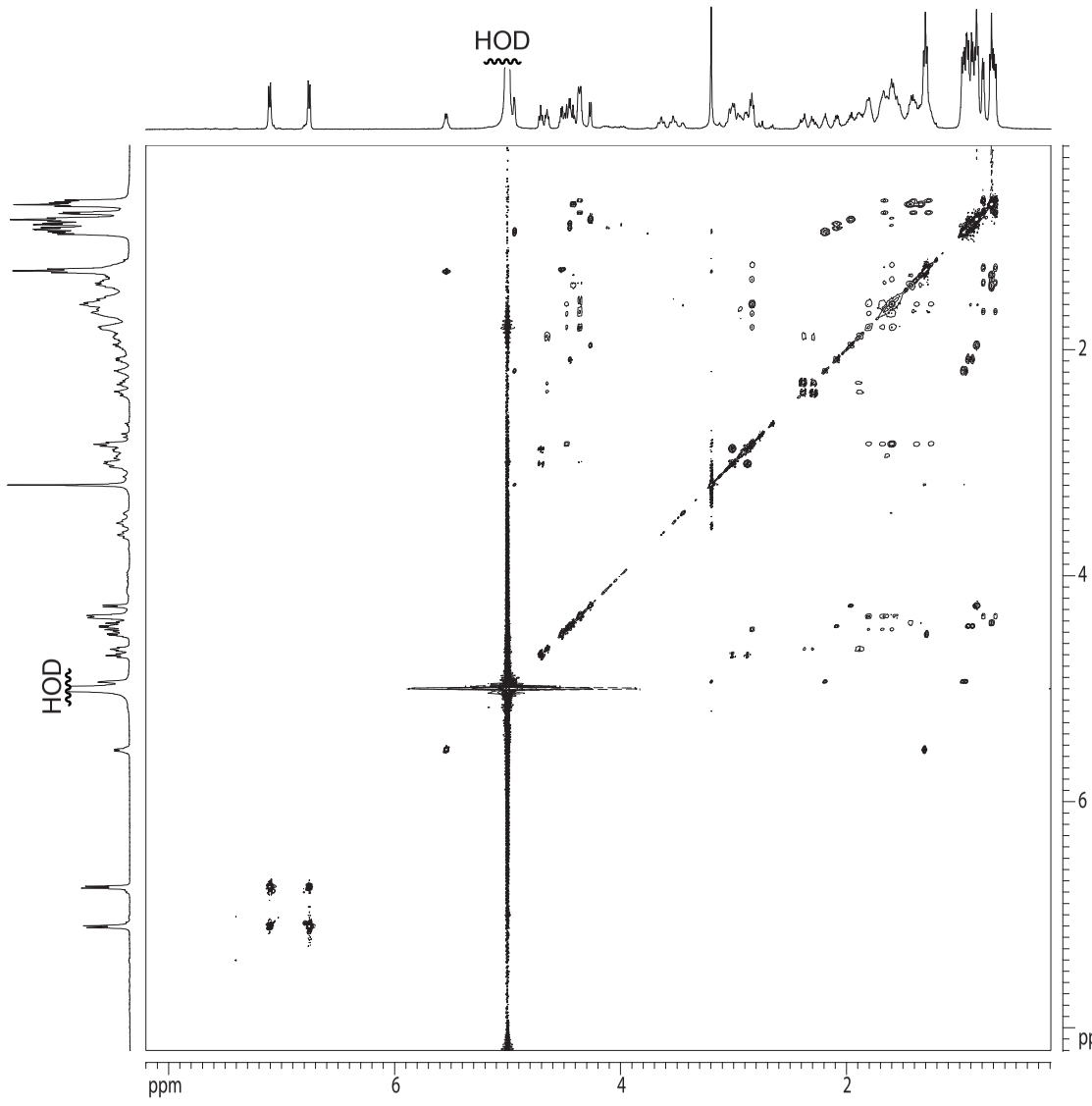
^{1}H NMR 1D of macrocycle **1c**, 2 mM in D_2O, 500 MHz at 278 K

^{1}H NMR 1D of macrocycle **1c**, 2 mM in D_2O, 500 MHz at 278 K
Ala$_{NMe}$ *N*-CH$_3$
Orn^1 H$_{\delta S}$
Orn^2 H$_{\delta S}$
Orn^1 H$_{\delta R}$
Orn^2 H$_{\delta R}$
ppm
integral

^{1}H NMR 1D of macrocycle **1c**, 2 mM in D_2O, 500 MHz at 278 K



^{1}H NMR 2D TOCSY of macrocycle **1c**, 2 mM in D_2O, 500 MHz at 278 K
150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME KHC-I-046
EXPNO 4
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130116
Time 3.18
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG tocsytp.wu
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 4006.410 Hz
FIDRES 1.956255 Hz
AQ 0.2556404 sec
RG 5.7
DW 124.800 usec
DE 6.00 usec
TE 278.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15211743 sec
d12 0.00002000 sec
IN0 0.00012480 sec
L1 65

======= CHANNEL f1 =======
NUC1 1H
P1 7.50 usec
p5 23.34 usec
P6 35.00 usec
p7 70.00 usec
P17 2500.00 usec
PL1 1.60 dB
PL10 15.20 dB
SFO1 500.2220009 MHz

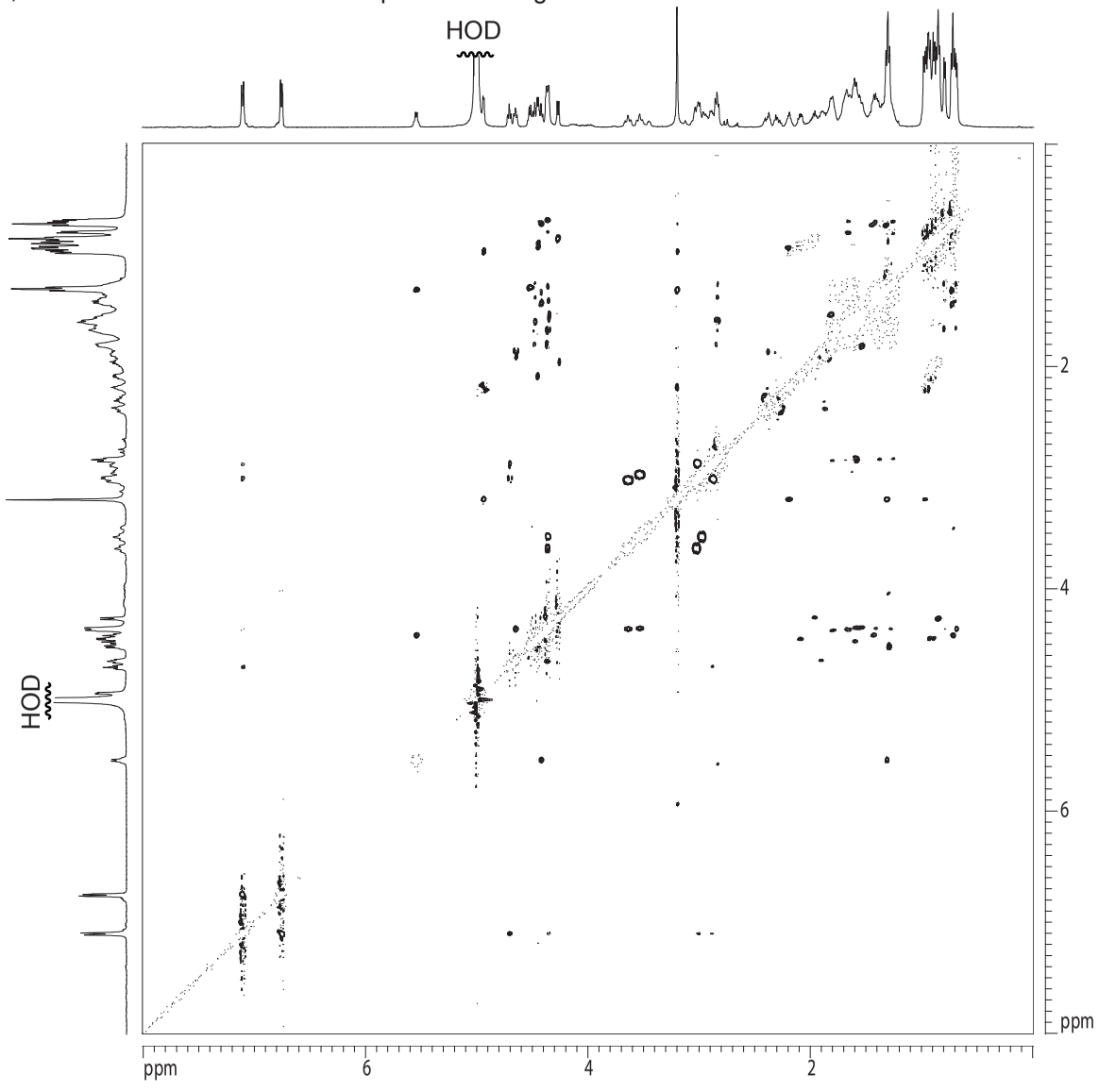
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.222 MHz
FIDRES 7.825020 Hz
SW 8.009 ppm
FnMODE undefined

F2 - Processing parameters
SI 1024
SF 500.2199016 MHz
WDW QSINE
SSB 3
LB 0.00 Hz
GB 0
PC 1.40

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2199016 MHz
WDW QSINE
SSB 3
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 8.201 ppm
F2LO 4102.45 Hz
F2PHI 0.192 ppm
F2HI 96.04 Hz
F1PLO 8.201 ppm
F1LO 4102.45 Hz
F1PHI 0.192 ppm
F1HI 96.04 Hz
F2PPMCM 0.53395 ppm/cm
F2HZCM 267.09399 Hz/cm
F1PPMCM 0.53395 ppm/cm
F1HZCM 267.09399 Hz/cm

^{1}H NMR ROESY of macrocycle **1c** with presaturation suppression of the HOD peak ,
2 mM in D_2O, 500 MHz at 278 K with150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME KHC-I-046
EXPNO 3
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130115
Time 21.50
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesyprtp.2.prd
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 5482.456 Hz
FIDRES 2.676980 Hz
AQ 0.1868276 sec
RG 456.1
DW 91.200 usec
DE 6.00 usec
TE 277.9 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15000001 sec
d11 0.03000000 sec
d12 0.00002000 sec
d13 0.00000300 sec
IN0 0.00009120 sec
L4 300

======= CHANNEL f1 =======
NUC1 1H
P1 7.50 usec
P25 250.00 usec
PL1 1.60 dB
PL9 54.00 dB
PL11 26.60 dB
SFO1 500.2224049 MHz

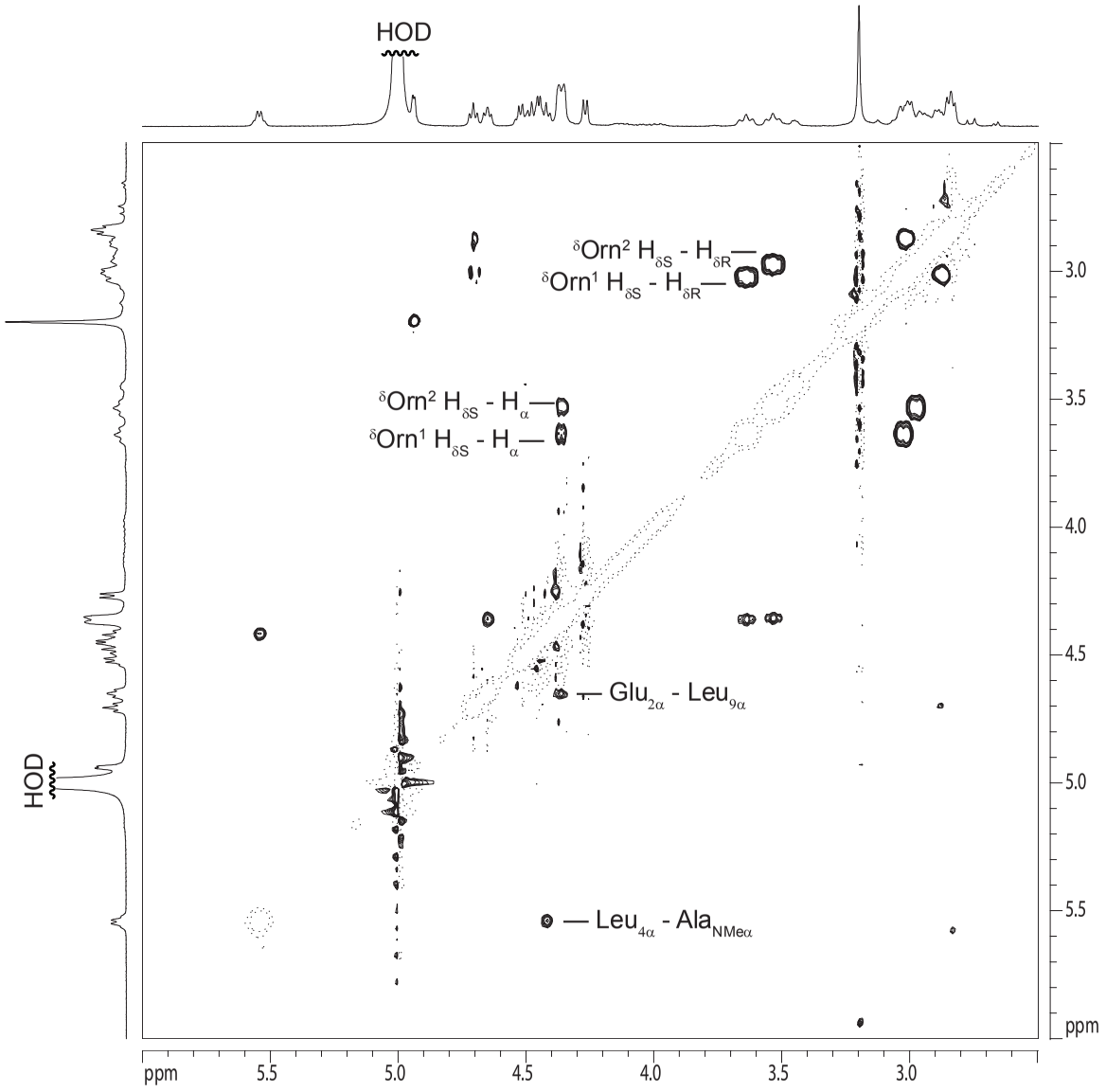
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 10.707922 Hz
SW 10.960 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2199016 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 4.00

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2199016 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 8.008 ppm
F2LO 4005.73 Hz
F2PHI -0.009 ppm
F2HI -4.38 Hz
F1PLO 8.008 ppm
F1LO 4005.73 Hz
F1PHI -0.009 ppm
F1HI -4.38 Hz
F2PPMCM 0.53445 ppm/cm
F2HZCM 267.34106 Hz/cm
F1PPMCM 0.53445 ppm/cm
F1HZCM 267.34106 Hz/cm

^{1}H NMR ROESY of macrocycle **1c** with presaturation suppression of the HOD peak , 2 mM in D_2O, 500 MHz at 278 K with150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME KHC-I-046
EXPNO 3
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130115
Time 21.50
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesyprtp.2.prd
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 5482.456 Hz
FIDRES 2.676980 Hz
AQ 0.1868276 sec
RG 456.1
DW 91.200 usec
DE 6.00 usec
TE 277.9 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15000001 sec
d11 0.03000000 sec
d12 0.00002000 sec
d13 0.00000300 sec
IN0 0.00009120 sec
L4 300

======== CHANNEL f1 ========
NUC1 1H
P1 7.50 usec
P25 250.00 usec
PL1 1.60 dB
PL9 54.00 dB
PL11 26.60 dB
SFO1 500.2224049 MHz

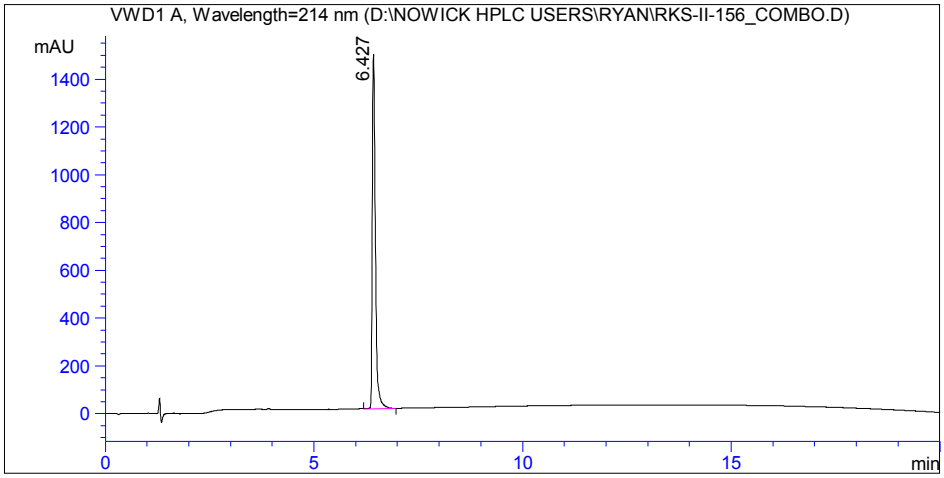
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 10.707922 Hz
SW 10.960 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2199016 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 4.00

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2199016 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

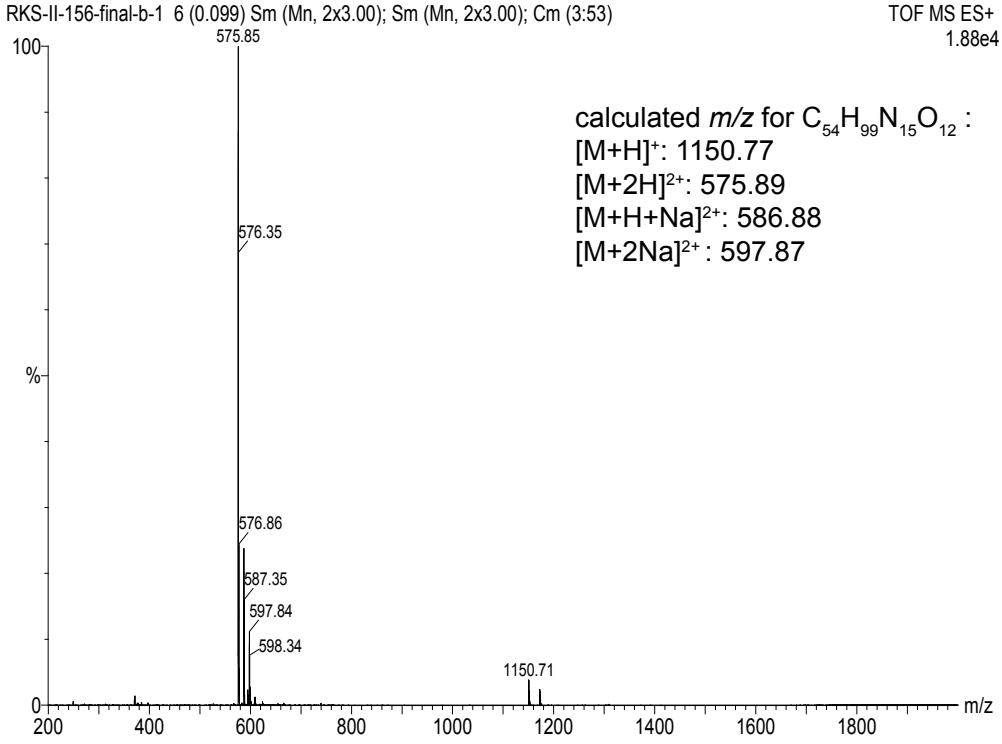
2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 6.000 ppm
F2LO 3001.32 Hz
F2PHI 2.496 ppm
F2HI 1248.44 Hz
F1PLO 6.000 ppm
F1LO 3001.32 Hz
F1PHI 2.496 ppm
F1HI 1248.44 Hz
F2PPMCM 0.23361 ppm/cm
F2HZCM 116.85833 Hz/cm
F1PPMCM 0.23361 ppm/cm
F1HZCM 116.85833 Hz/cm

HPLC and MS ESI$^+$ TOF of macrocycle **1d**

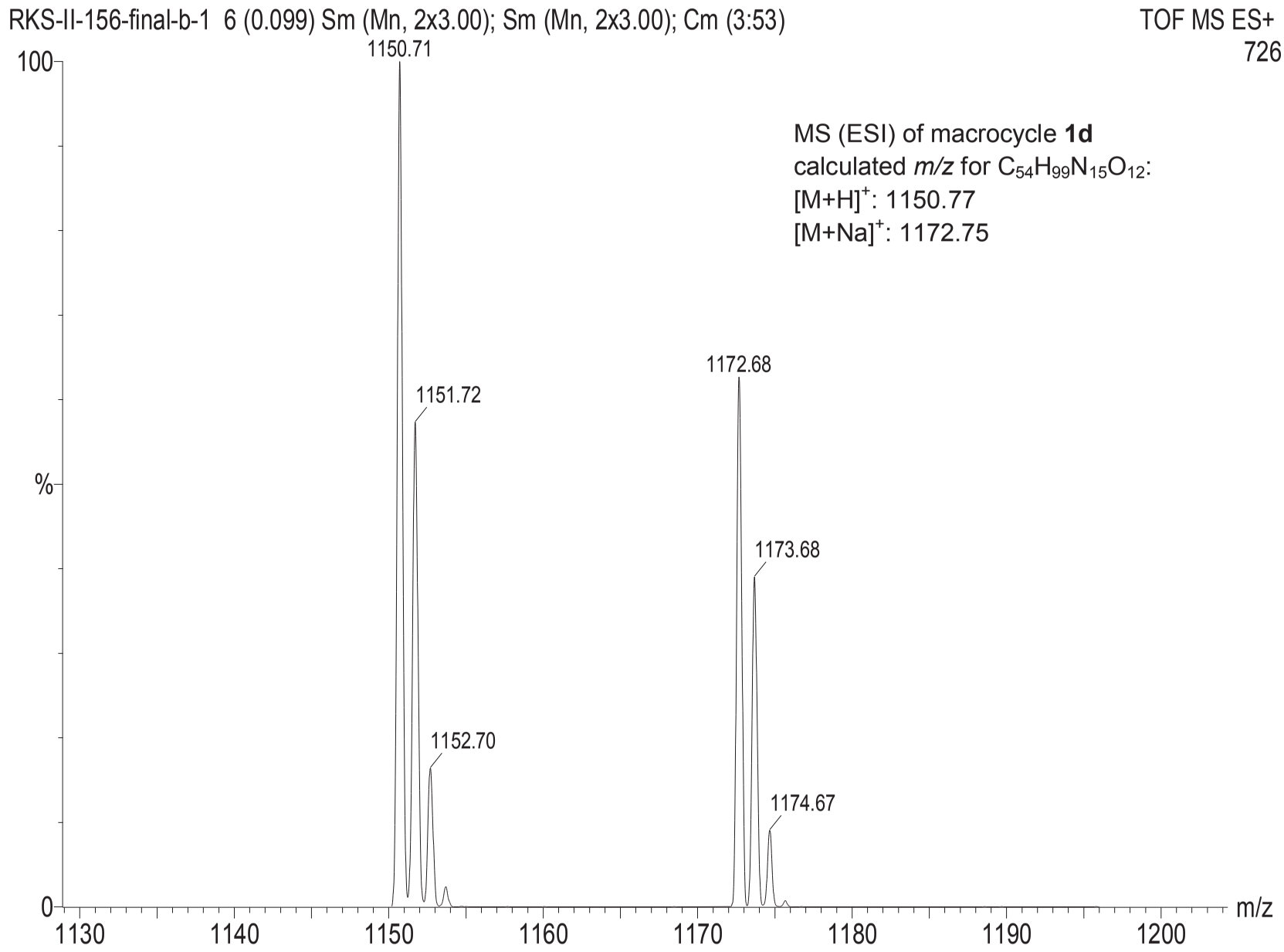


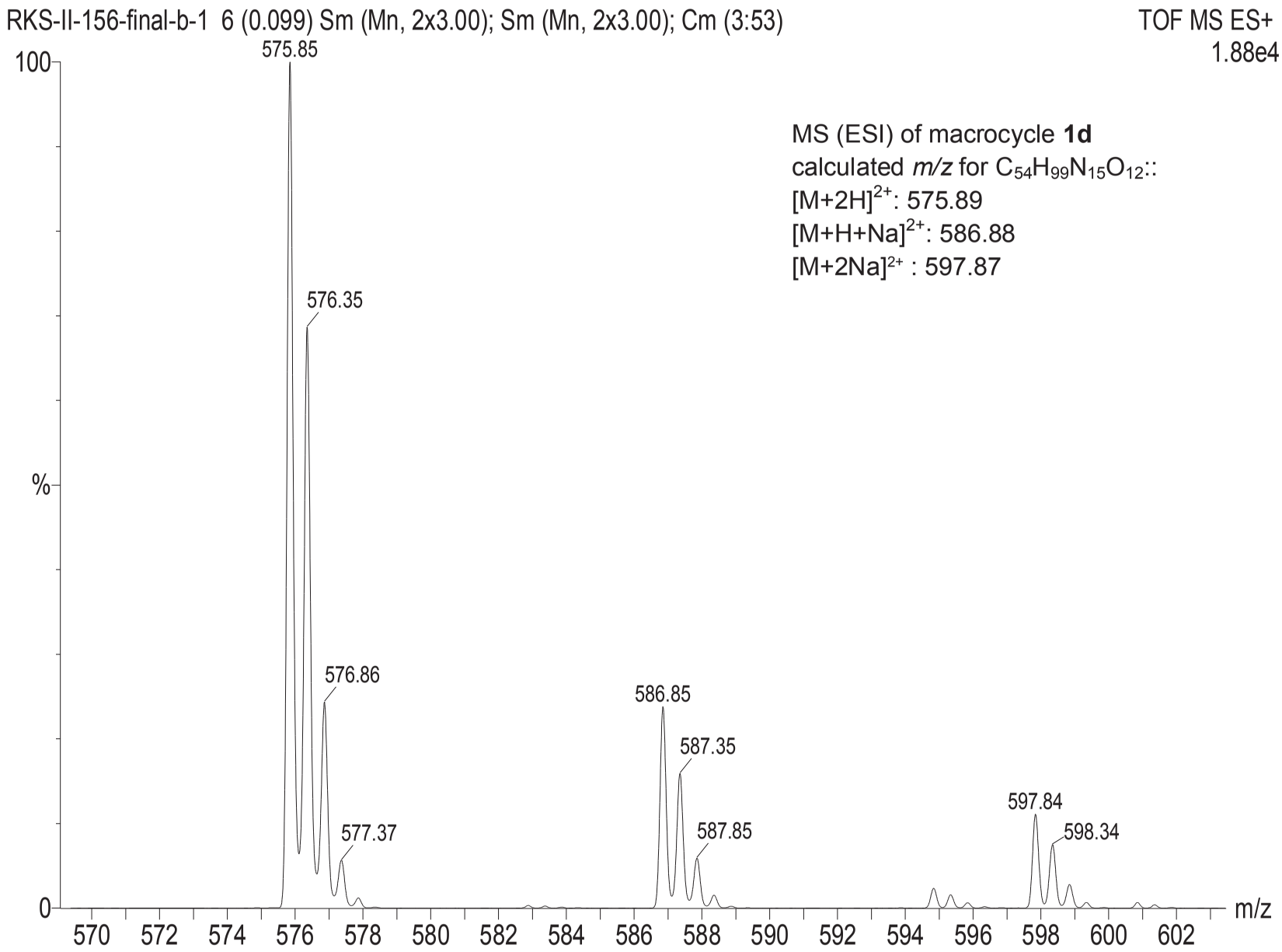
Signal 1:VWD1 A, Wavelength=214 nm

Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	6.427	VV	0.078	7672.315	100.000	100.000

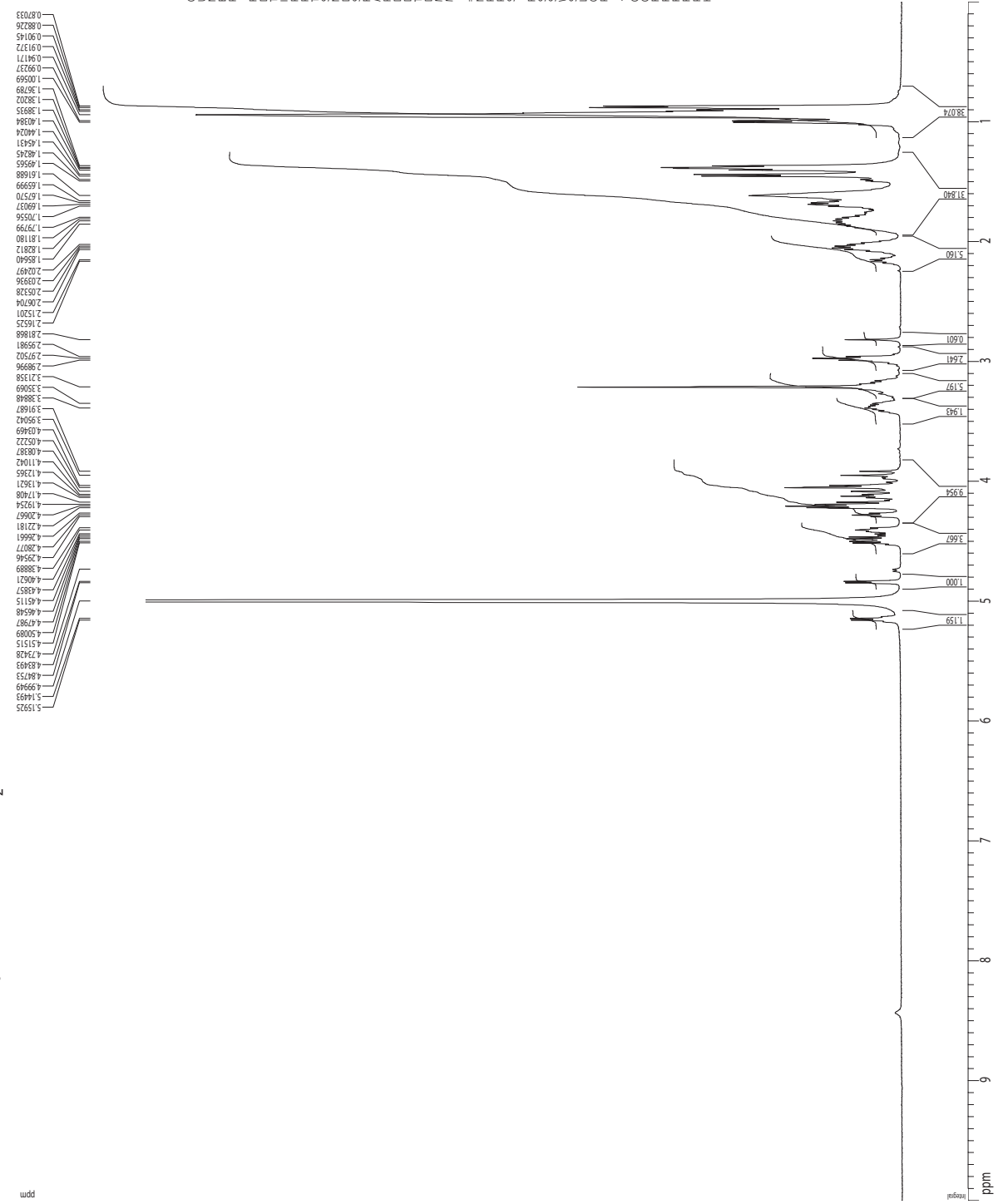


calculated *m/z* for $C_{54}H_{99}N_{15}O_{12}$:
$[M+H]^+$: 1150.77
$[M+2H]^{2+}$: 575.89
$[M+H+Na]^{2+}$: 586.88
$[M+2Na]^{2+}$: 597.87


RKS-II-156-final-b-1 6 (0.099) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:53)
TOF MS ES+
726
MS (ESI) of macrocycle 1d
calculated m/z for C54H99N15O12:
[M+H]+: 1150.77
[M+Na]+: 1172.75
1150.71
1151.72
1152.70
1172.68
1173.68
1174.67
100
%
0
m/z
1130
1140
1150
1160
1170
1180
1190
1200


RKS-II-156-final-b-1 6 (0.099) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:53)
TOF MS ES+
1.88e4
100
%
0
575.85
576.35
576.86
577.37
586.85
587.35
587.85
597.84
598.34
m/z
570 572 574 576 578 580 582 584 586 588 590 592 594 596 598 600 602
MS (ESI) of macrocycle 1d
calculated m/z for C54H99N15O12::
[M+2H]2+: 575.89
[M+H+Na]2+: 586.88
[M+2Na]2+ : 597.87

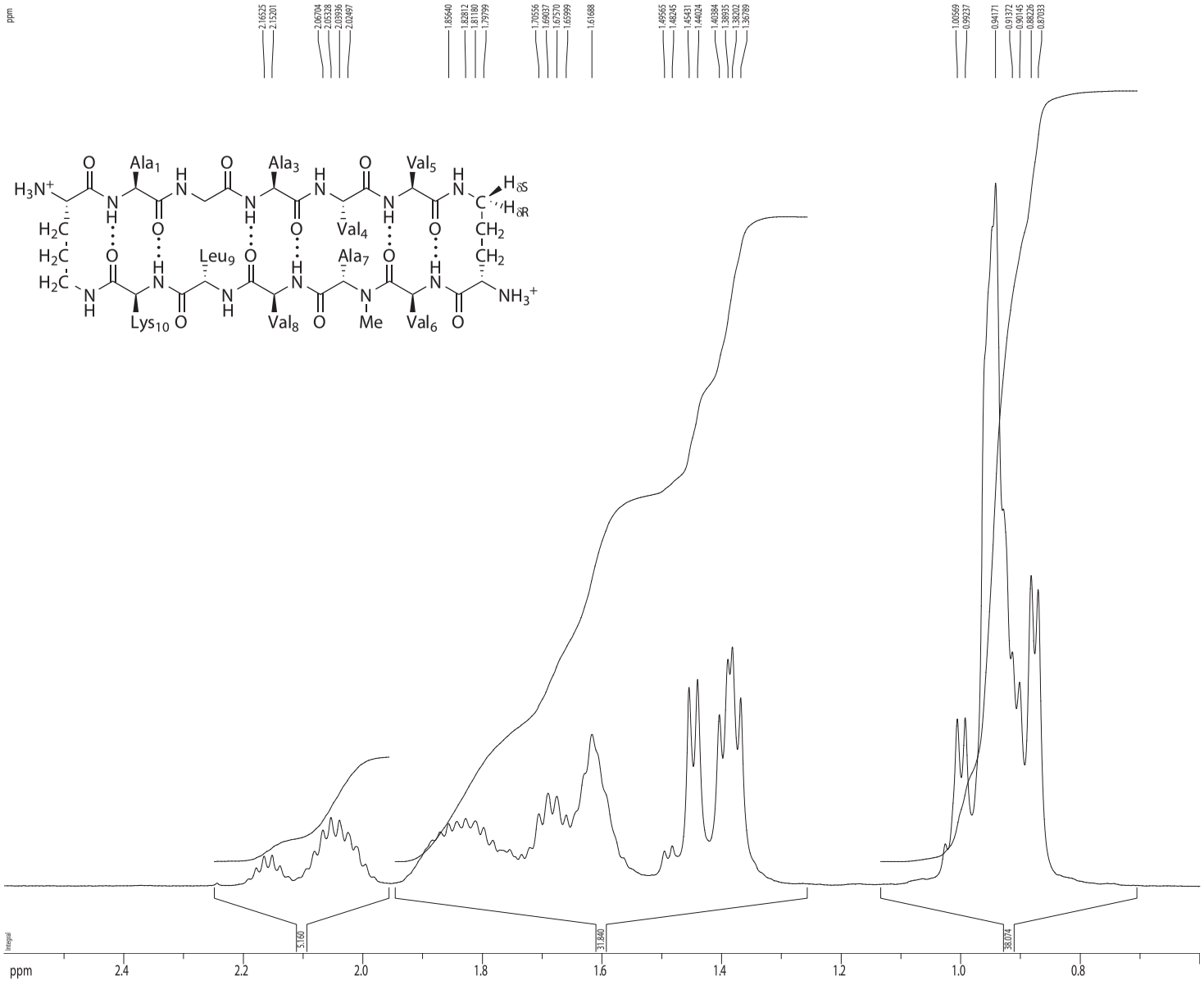
^{1}H NMR 1D of macrocycle **1d**, 2 mM in D_2O, 500 MHz at 283 K



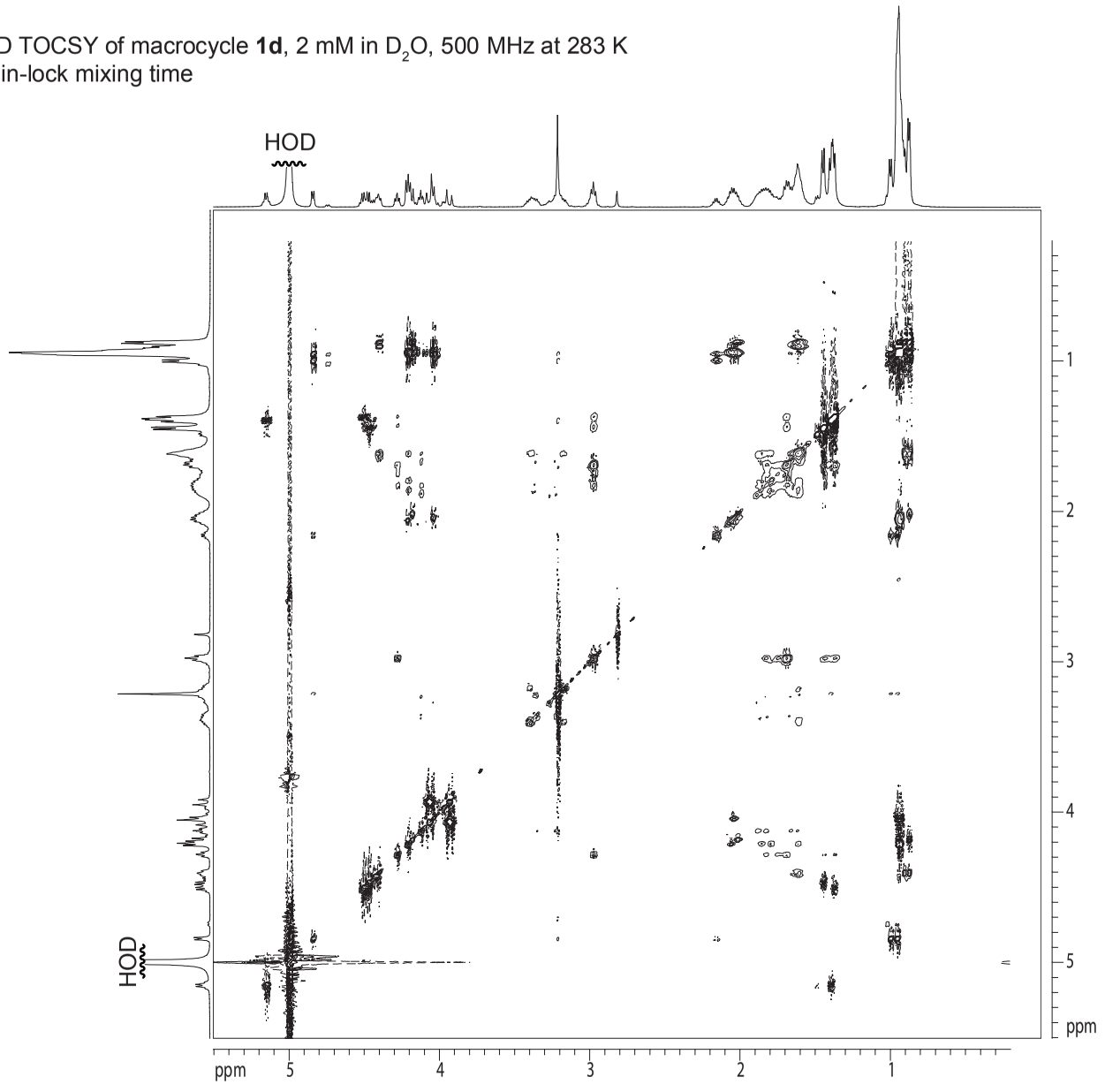
^{1}H NMR 1D of macrocycle **1d**, 2 mM in D_2O, 500 MHz at 283 K

^{1}H NMR 1D of macrocycle **1d**, 2 mM in D_2O, 500 MHz at 283 K

^{1}H NMR 1D of macrocycle **1d**, 2 mM in D_2O, 500 MHz at 283 K



^{1}H NMR 2D TOCSY of macrocycle **1d**, 2 mM in D_2O, 500 MHz at 283 K
150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME RKS-II-156
EXPNO 4
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130114
Time 3.25
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG tocsytp.wu
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 3004.808 Hz
FIDRES 1.467191 Hz
AQ 0.3408372 sec
RG 7.1
DW 166.400 usec
DE 6.00 usec
TE 283.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15211743 sec
d12 0.00002000 sec
IN0 0.00016640 sec
L1 65

======= CHANNEL f1 =======
NUC1 1H
P1 7.50 usec
p5 23.34 usec
P6 35.00 usec
p7 70.00 usec
P17 2500.00 usec
PL1 1.60 dB
PL10 15.20 dB
SFO1 500.2215007 MHz

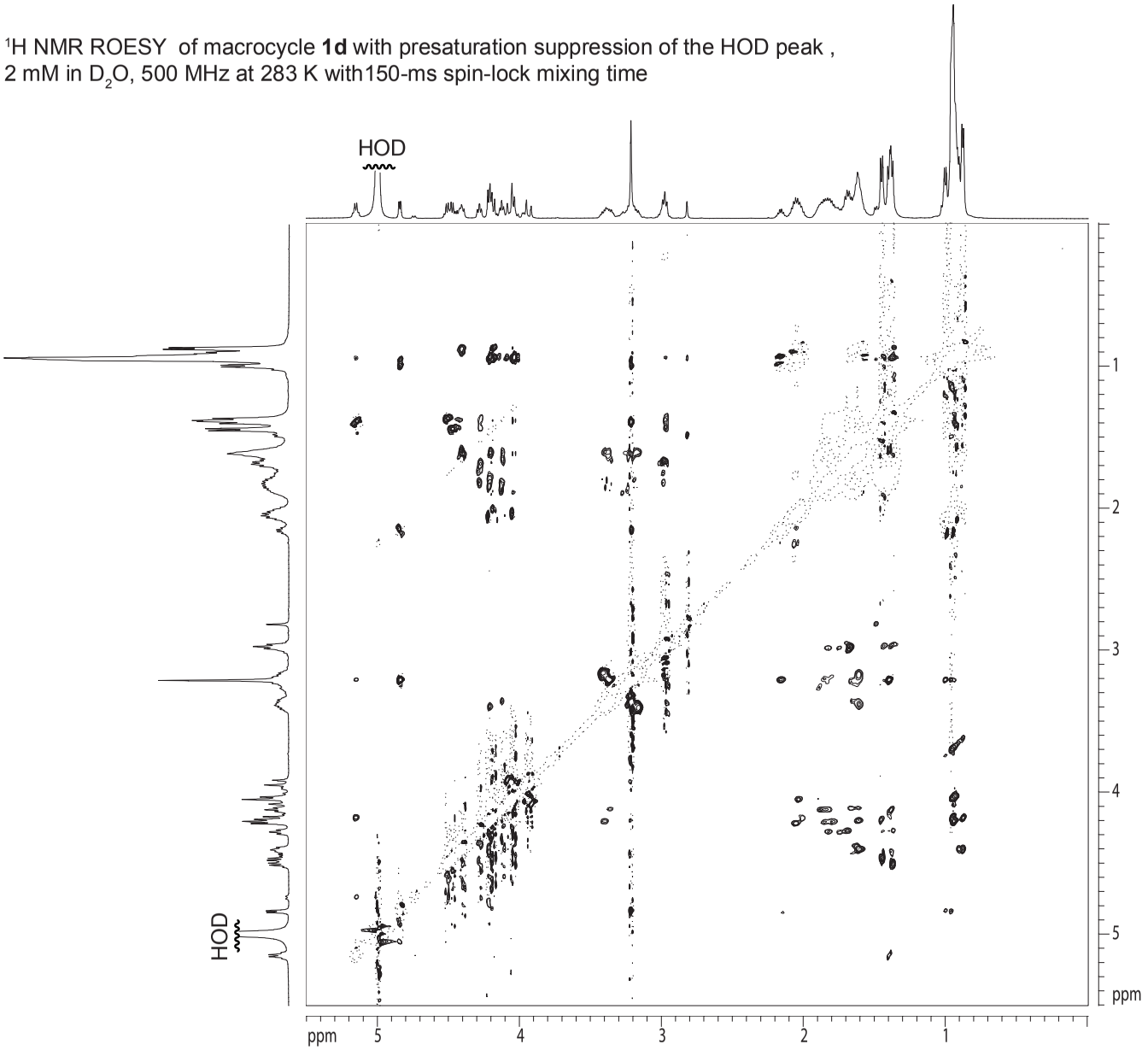
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2215 MHz
FIDRES 5.868765 Hz
SW 6.007 ppm
FnMODE undefined

F2 - Processing parameters
SI 1024
SF 500.2198981 MHz
WDW QSINE
SSB 3
LB 0.00 Hz
GB 0
PC 1.40

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2198981 MHz
WDW QSINE
SSB 3
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 5.505 ppm
F2LO 2753.64 Hz
F2PHI -0.003 ppm
F2HI -1.75 Hz
F1PLO 5.505 ppm
F1LO 2753.64 Hz
F1PHI -0.003 ppm
F1HI -1.75 Hz
F2PPMCM 0.36722 ppm/cm
F2HZCM 183.69231 Hz/cm
F1PPMCM 0.36722 ppm/cm
F1HZCM 183.69231 Hz/cm

^{1}H NMR ROESY of macrocycle **1d** with presaturation suppression of the HOD peak ,
2 mM in D_2O, 500 MHz at 283 K with150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME RKS-II-156
EXPNO 3
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130113
Time 21.57
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesyprtp.2.prd
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 5482.456 Hz
FIDRES 2.676980 Hz
AQ 0.1868276 sec
RG 362
DW 91.200 usec
DE 6.00 usec
TE 283.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15000001 sec
d11 0.03000000 sec
d12 0.00002000 sec
d13 0.00000300 sec
IN0 0.00009120 sec
L4 300

======== CHANNEL f1 =======
NUC1 1H
P1 7.50 usec
P25 250.00 usec
PL1 1.60 dB
PL9 54.00 dB
PL11 26.60 dB
SFO1 500.2224049 MHz

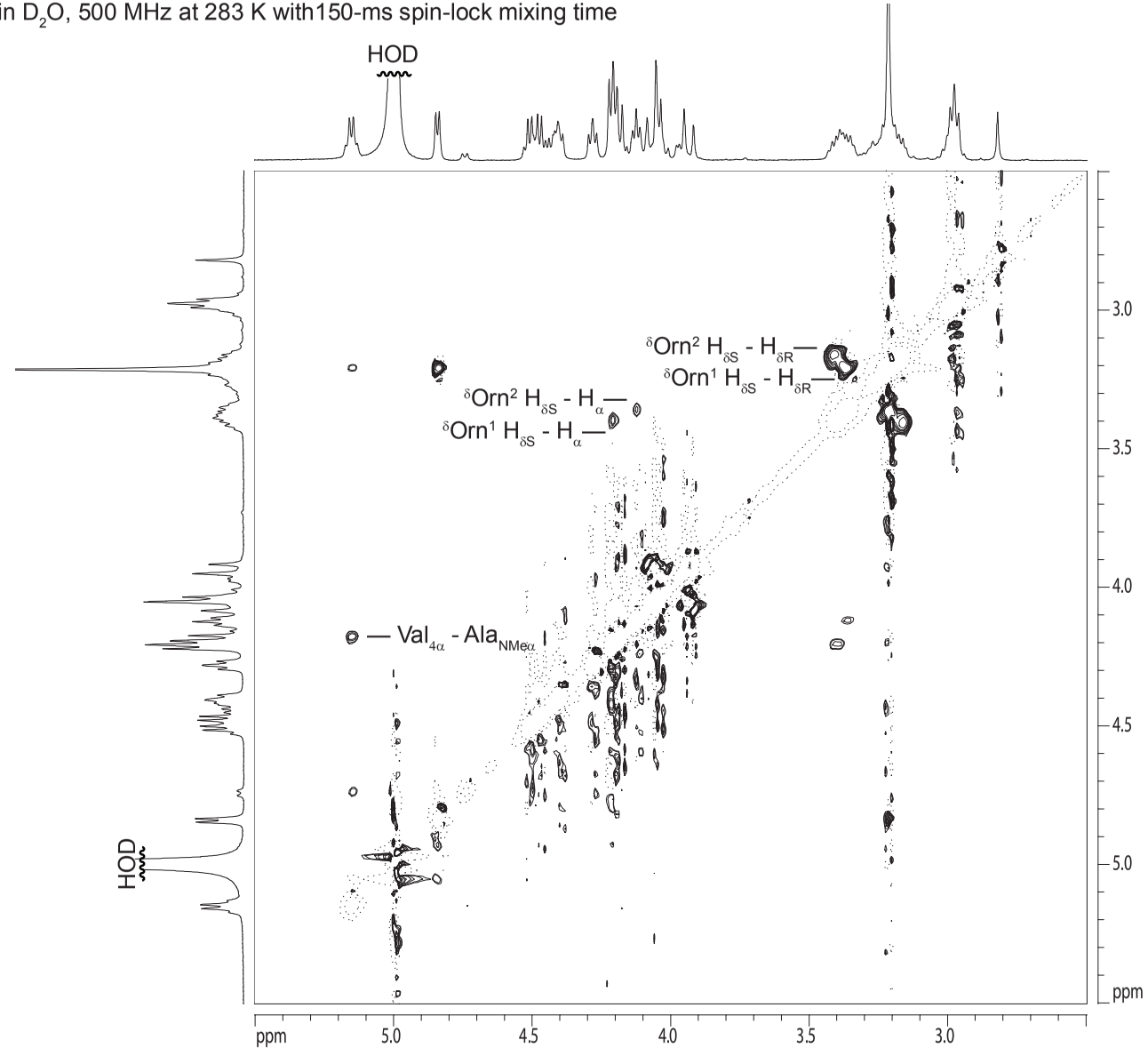
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 10.707922 Hz
SW 10.960 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2198981 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 4.00

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2198981 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 5.504 ppm
F2LO 2753.10 Hz
F2PHI -0.008 ppm
F2HI -4.19 Hz
F1PLO 5.504 ppm
F1LO 2753.10 Hz
F1PHI -0.008 ppm
F1HI -4.19 Hz
F2PPMCM 0.36748 ppm/cm
F2HZCM 183.81934 Hz/cm
F1PPMCM 0.36748 ppm/cm
F1HZCM 183.81934 Hz/cm

^{1}H NMR ROESY of macrocycle **1d** with presaturation suppression of the HOD peak , 2 mM in D_2O, 500 MHz at 283 K with150-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME RKS-II-156
EXPNO 3
PROCNO 1

F2 - Acquisition Parameters
Date_ 20130113
Time 21.57
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesyprtp.2.prd
TD 2048
SOLVENT D2O
NS 16
DS 16
SWH 5482.456 Hz
FIDRES 2.676980 Hz
AQ 0.1868276 sec
RG 362
DW 91.200 usec
DE 6.00 usec
TE 283.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.15000001 sec
d11 0.03000000 sec
d12 0.00002000 sec
d13 0.00000300 sec
IN0 0.00009120 sec
L4 300

======== CHANNEL f1 ========
NUC1 1H
P1 7.50 usec
P25 250.00 usec
PL1 1.60 dB
PL9 54.00 dB
PL11 26.60 dB
SFO1 500.2224049 MHz

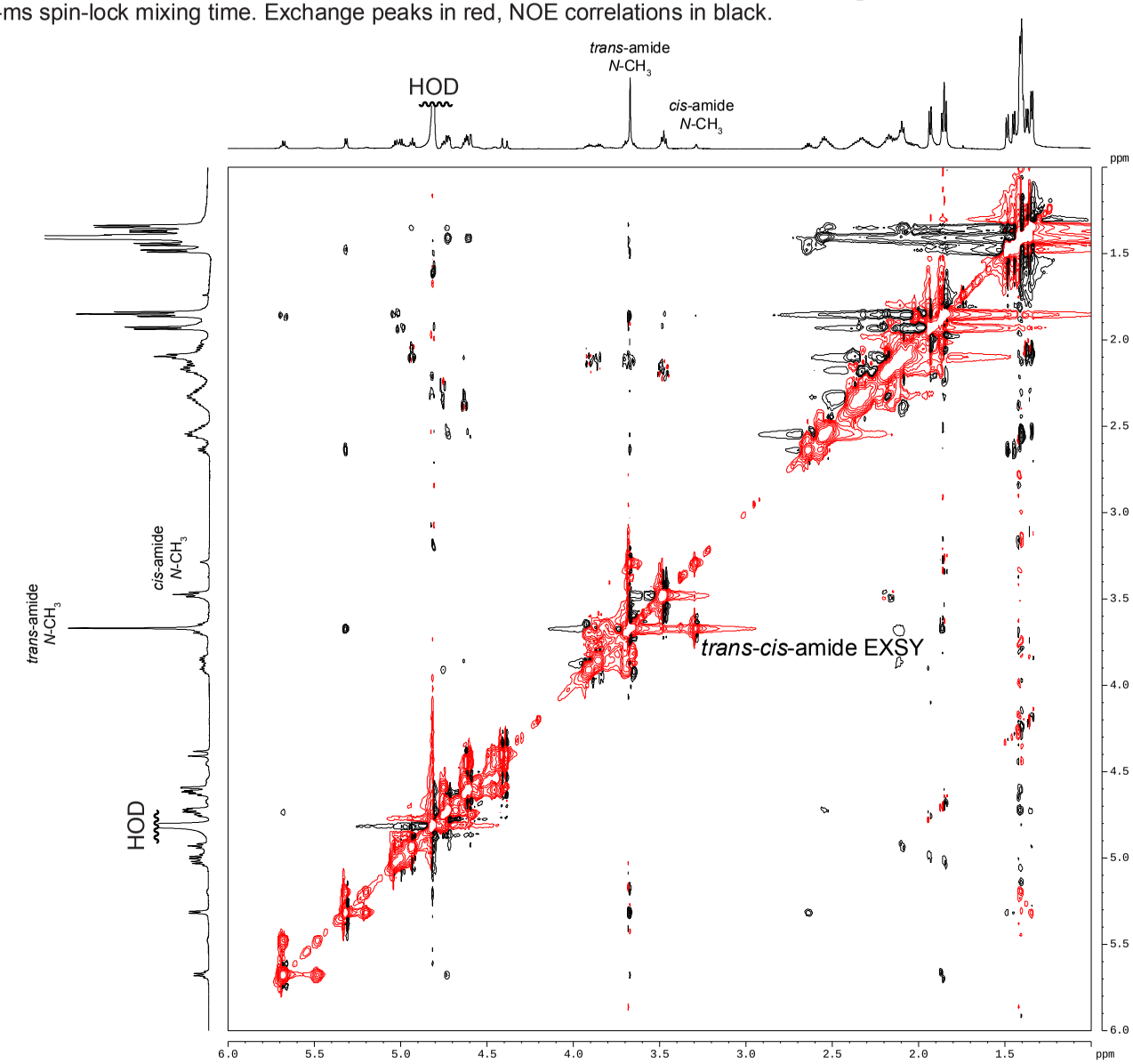
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 10.707922 Hz
SW 10.960 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2198981 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 4.00

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2198981 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

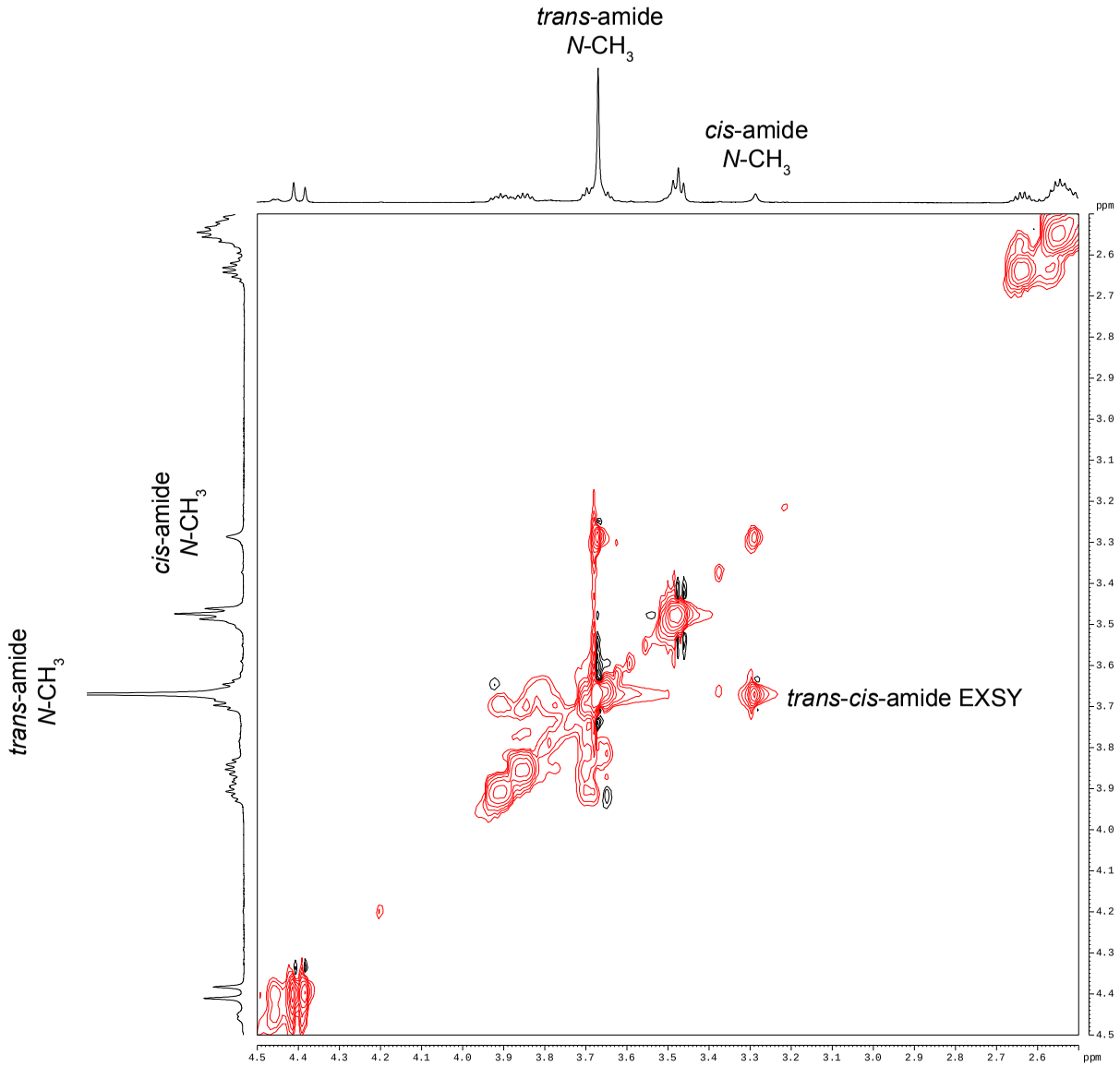
2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 5.504 ppm
F2LO 2753.10 Hz
F2PHI 2.496 ppm
F2HI 1248.64 Hz
F1PLO 5.504 ppm
F1LO 2753.10 Hz
F1PHI 2.496 ppm
F1HI 1248.64 Hz
F2PPMCM 0.20051 ppm/cm
F2HZCM 100.29754 Hz/cm
F1PPMCM 0.20051 ppm/cm
F1HZCM 100.29754 Hz/cm

^{1}H NMR 2D EXSY (ROESY showing exchange crosspeaks) of macrocycle **1d**, 2 mM in D_2O, 500 MHz at 343 K
150-ms spin-lock mixing time. Exchange peaks in red, NOE correlations in black.



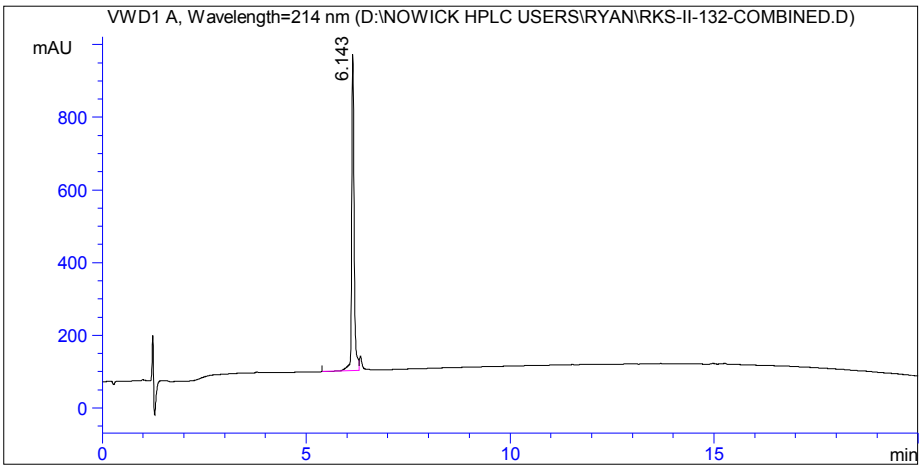

```
Current Data Parameters
NAME      RKS-II-156_600MHz
EXPNO                 3
PROCNO                1

F2 - Acquisition Parameters
Date_           20130131
Time              14.22
INSTRUM             av600
PROBHD    5 mm TBI 1H/13
PULPROG    roesyph.2.prd
TD                 2048
SOLVENT             D2O
NS                    8
DS                   16
SWH            4807.692 Hz
FIDRES         2.347596 Hz
AQ             0.2130420 sec
RG                  322
DW              104.000 usec
DE                6.00 usec
TE                342.8 K
D0           0.00009491 sec
D1           2.00000000 sec
D8           0.15000001 sec
D12          0.00002000 sec
IN0          0.00020800 sec
L4                  300
P15          150000.00 usec

======== CHANNEL f1 ========
NUC1                1H
P1                8.00 usec
P25             250.00 usec
PLW1        23.01441956 W
PLW27        0.09332543 W
SFO1        600.1324020 MHz

F1 - Acquisition parameters
TD                  256
SFO1         600.1324 MHz
FIDRES       18.780048 Hz
SW             8.011 ppm
FnMODE       States-TPPI

F2 - Processing parameters
SI                 1024
SF           600.1300000 MHz
WDW                QSINE
SSB                   2
LB         0 Hz
GB         0
PC                 1.00

F1 - Processing parameters
SI                 1024
MC2          States-TPPI
SF           600.1300000 MHz
WDW          States-TPPI
SSB                   2
LB         0 Hz
GB         0
```

^{1}H NMR 2D EXSY (ROESY showing exchange crosspeaks) of macrocycle **1d**, 2 mM in D_2O, 500 MHz at 343 K
150-ms spin-lock mixing time. Exchange peaks in red, NOE correlations in black.




```
Current Data Parameters
NAME      RKS-II-156_600MHz
EXPNO              3
PROCNO             1

F2 - Acquisition Parameters
Date_          20130131
Time             14.22
INSTRUM          av600
PROBHD   5 mm TBI 1H/13
PULPROG   roesyph.2.prd
TD               2048
SOLVENT           D2O
NS                  8
DS                 16
SWH          4807.692 Hz
FIDRES       2.347506 Hz
AQ           0.2130420 sec
RG                322
DW             104.000 usec
DE               6.00 usec
TE              342.8 K
D0          0.00009491 sec
D1          2.00000000 sec
D8          0.15000001 sec
D12         0.00002000 sec
IN0         0.00020800 sec
L4                300
P15          150000.00 usec

======== CHANNEL f1 ========
NUC1               1H
P1              8.00 usec
P25           250.00 usec
PLW1       23.01441956 W
PLW27       0.09332543 W
SFO1       600.1324020 MHz

F1 - Acquisition parameters
TD                256
SFO1       600.1324 MHz
FIDRES      18.780048 Hz
SW             8.011 ppm
FnMODE      States-TPPI

F2 - Processing parameters
SI               1024
SF         600.1300000 MHz
WDW              QSINE
SSB                 2
LB        0 Hz
GB        0
PC               1.00

F1 - Processing parameters
SI               1024
MC2         States-TPPI
SF         600.1300000 MHz
WDW         States-TPPI
SSB                 2
LB        0 Hz
GB        0
```

HPLC and MS ESI$^+$ TOF of macrocycle **1e**

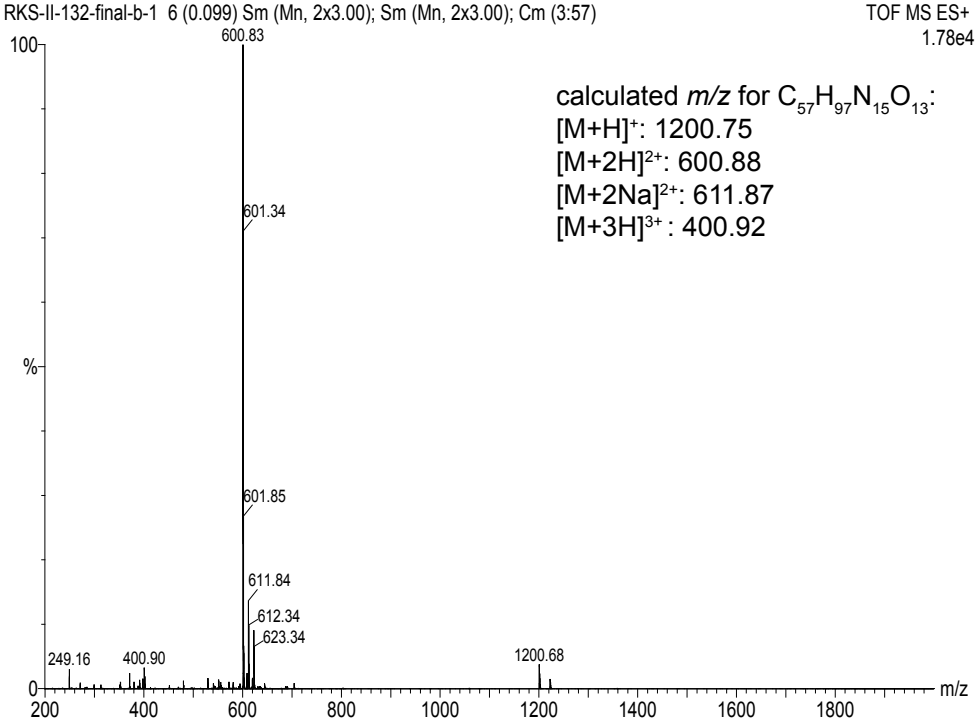


Signal 1:VWD1 A, Wavelength=214 nm

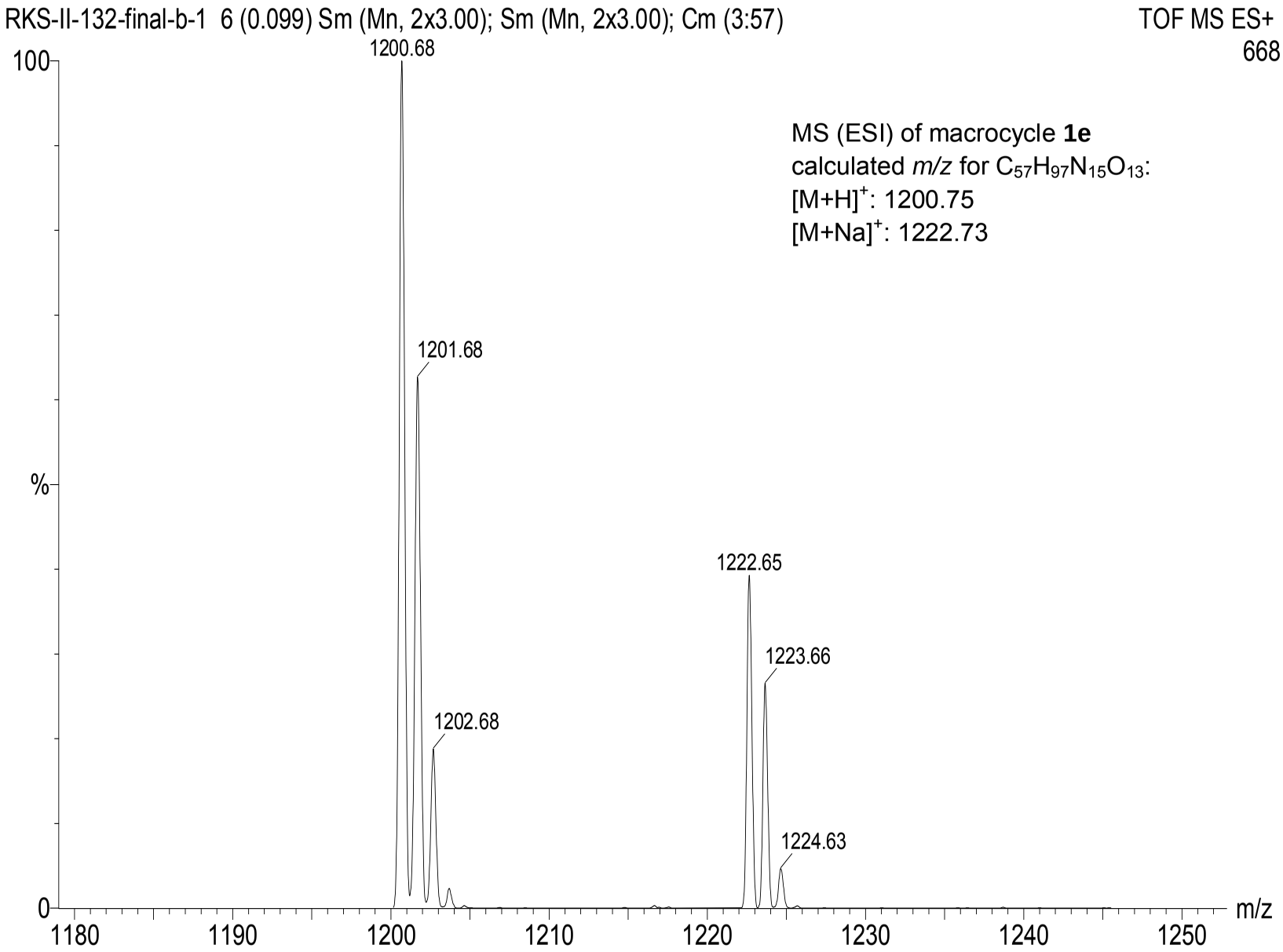
Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	6.143	BV	0.063	3636.210	100.000	100.000

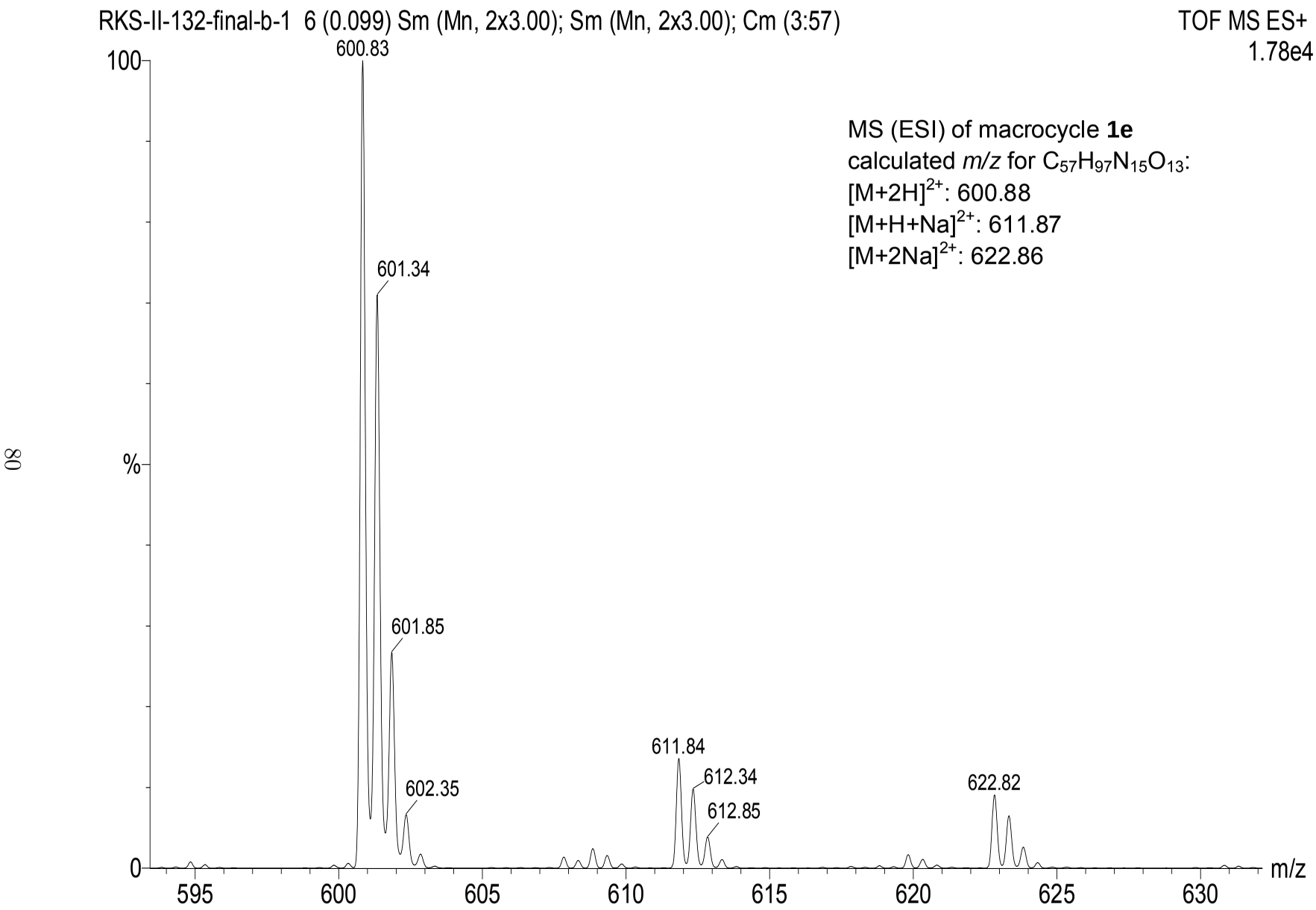
RKS-II-132-final-b-1 6 (0.099) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:57)

TOF MS ES+
1.78e4

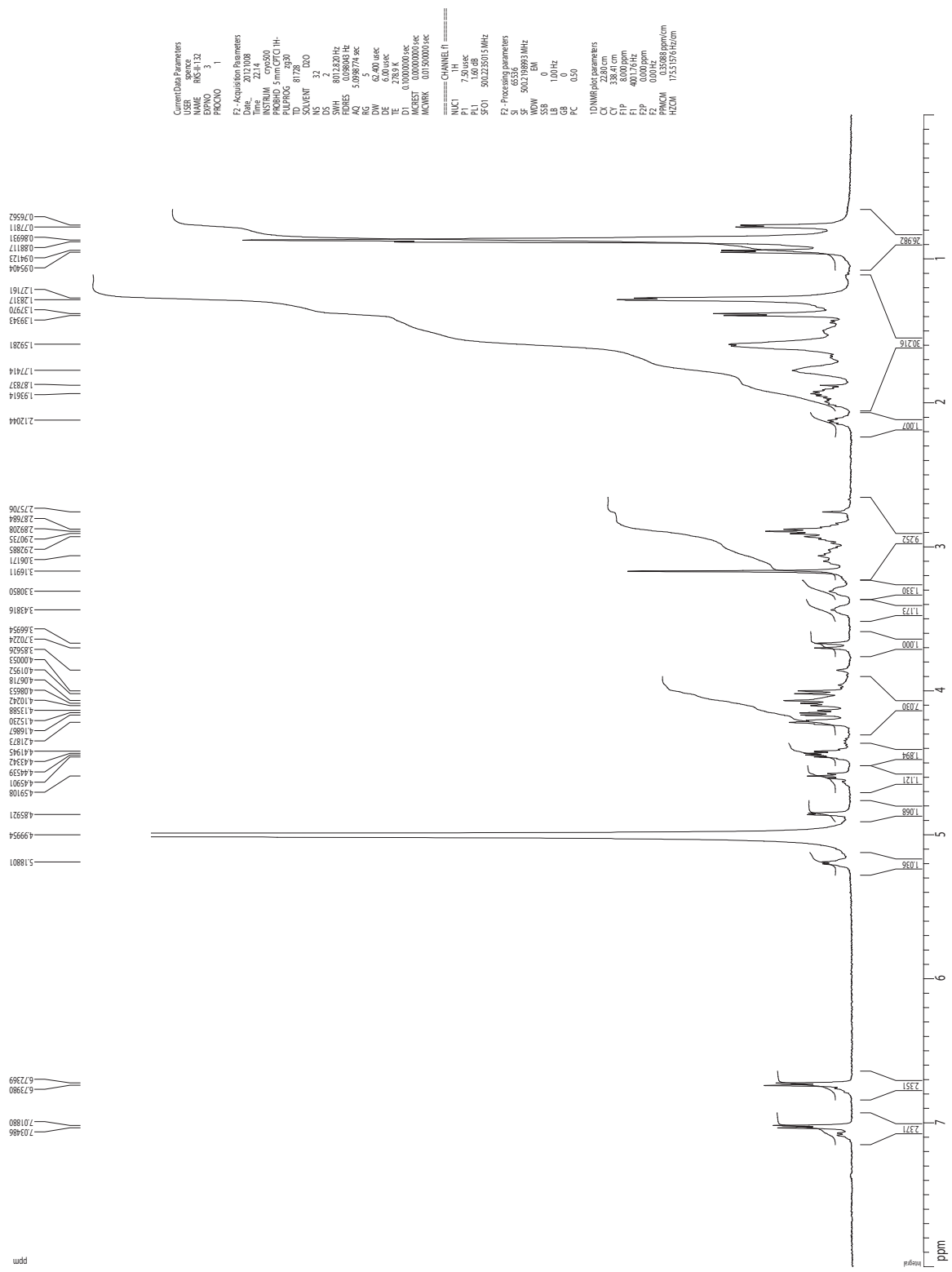


calculated *m/z* for $C_{57}H_{97}N_{15}O_{13}$:
$[M+H]^+$: 1200.75
$[M+2H]^{2+}$: 600.88
$[M+2Na]^{2+}$: 611.87
$[M+3H]^{3+}$: 400.92


RKS-II-132-final-b-1 6 (0.099) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:57)
TOF MS ES+
668
MS (ESI) of macrocycle 1e
calculated m/z for C57H97N15O13:
[M+H]+: 1200.75
[M+Na]+: 1222.73
1200.68
1201.68
1202.68
1222.65
1223.66
1224.63
100
%
0
m/z
1180
1190
1200
1210
1220
1230
1240
1250


RKS-II-132-final-b-1 6 (0.099) Sm (Mn, 2x3.00); Sm (Mn, 2x3.00); Cm (3:57)
TOF MS ES+
1.78e4
100
%
0
600.83
601.34
601.85
602.35
611.84
612.34
612.85
622.82
595
600
605
610
615
620
625
630
m/z
MS (ESI) of macrocycle 1e
calculated m/z for C57H97N15O13:
[M+2H]2+: 600.88
[M+H+Na]2+: 611.87
[M+2Na]2+: 622.86

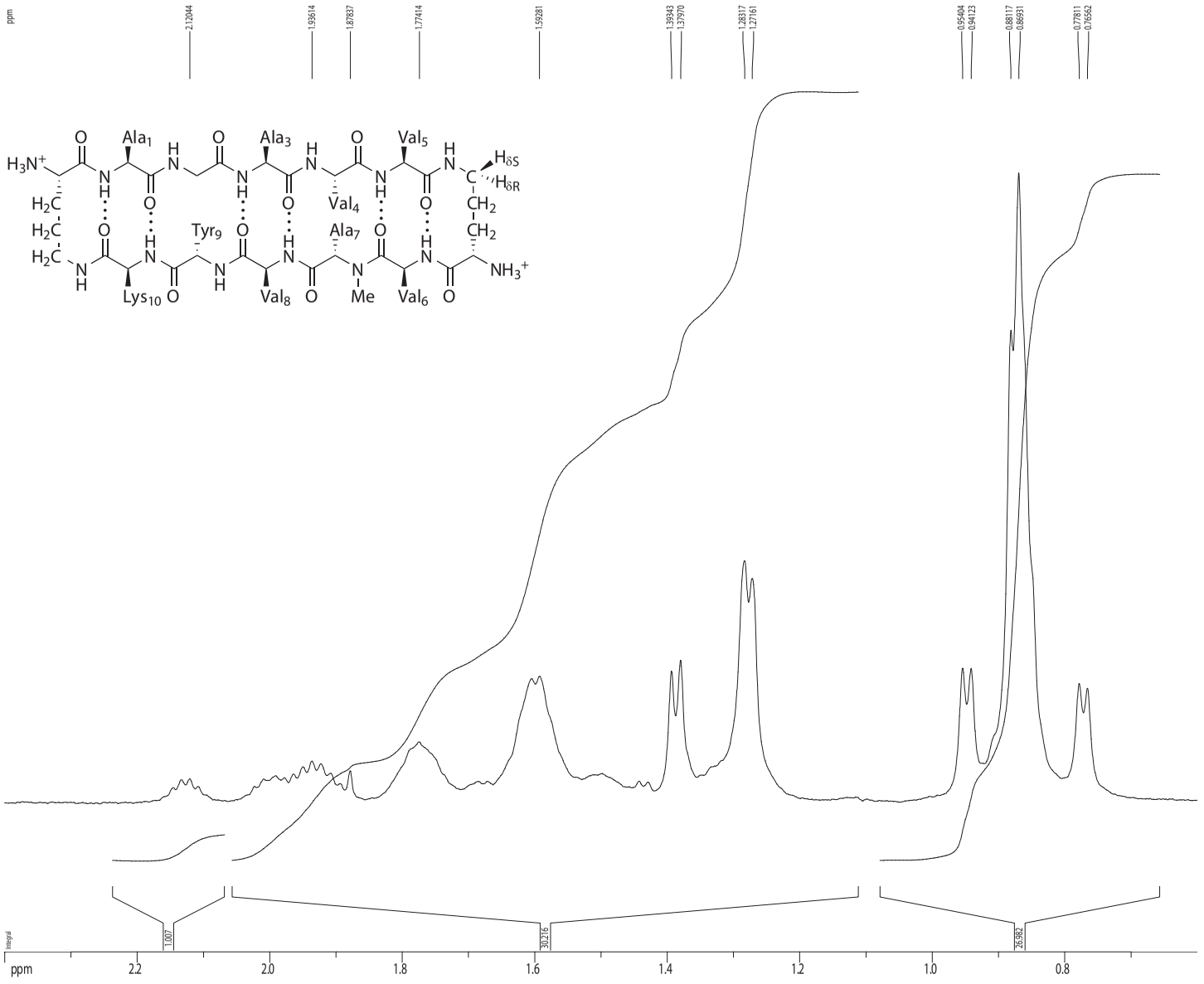
^{1}H NMR 1D of macrocycle **1e**, 2 mM in D_2O, 500 MHz at 278 K



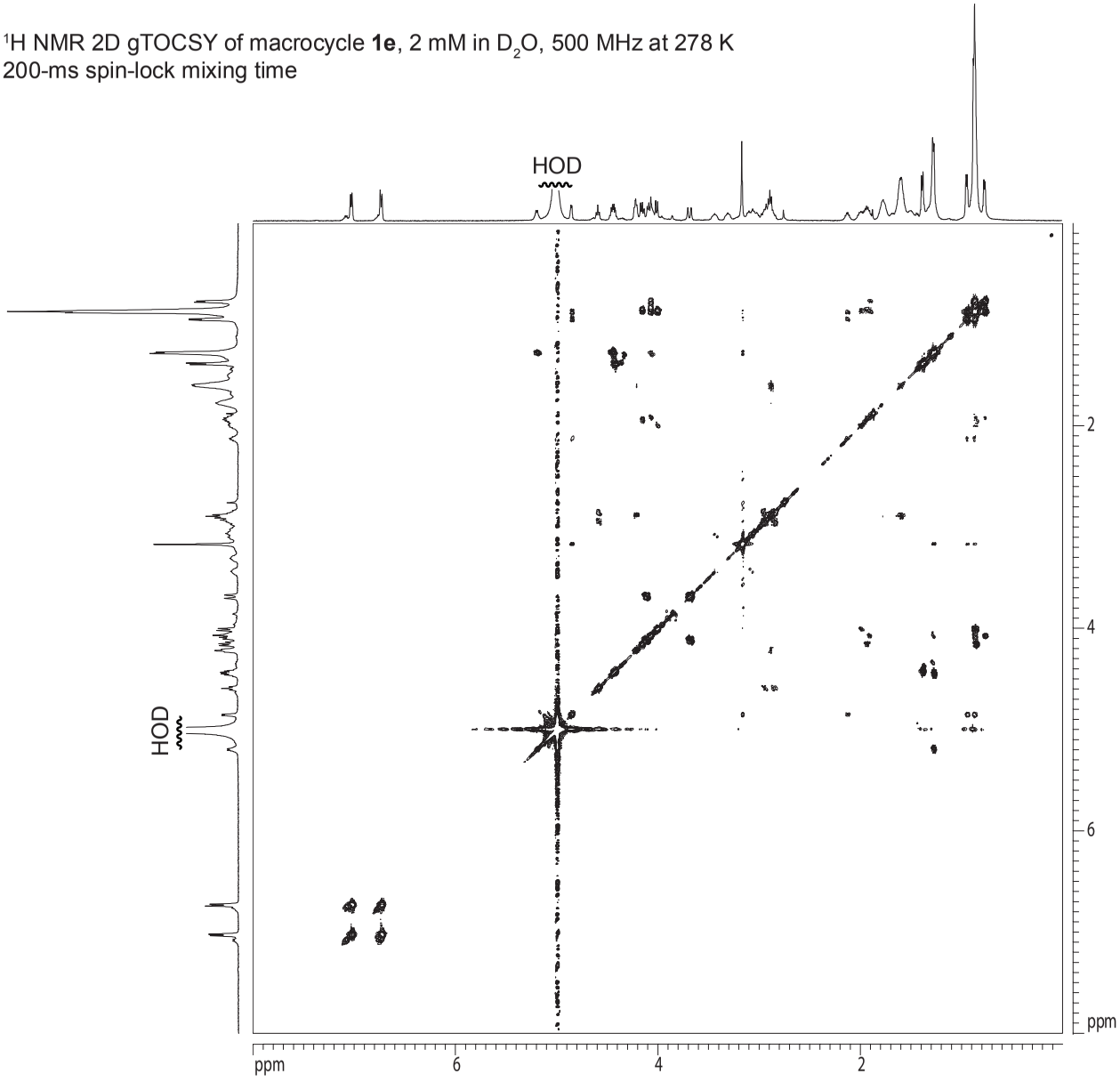
^{1}H NMR 1D of macrocycle **1e**, 2 mM in D_2O, 500 MHz at 278 K
HOD
Ala$_{NMe}$ Hα
Val$_6$ Hα
Tyr$_9$ Hα
Ala$_{1/3}$ Hα
Ala$_{1/3}$ Hα
Orn^1 Hα
Gly Hα, Val$_{5/8}$ Hα
Lys$_{10}$ Hα
Val$_4$ Hα, Orn^2 Hα
Val$_{5/8}$ Hα
Gly Hα
ppm
5.4
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
1.0358
1.0683
1.1207
1.8944
7.0297
1.0000
0.0399
5.188O1
4.9954
4.8592
4.5910
4.4590
4.4454
4.4334
4.4194
4.2187
4.1686
4.1523
4.1358
4.1024
4.0865
4.0671
4.0195
4.0005
3.8562
3.7022
3.6695
H_3N^+
H_2C
H_2C
H_2C
Ala$_1$
Lys$_{10}$
Tyr$_9$
Ala$_3$
Val$_8$
Val$_4$
Ala$_7$
Me
Val$_5$
Val$_6$
CH_2
CH_2
NH_3^+
$H_{\delta S}$
$H_{\delta R}$

^{1}H NMR 1D of macrocycle **1e**, 2 mM in D_2O, 500 MHz at 278 K

^{1}H NMR 1D of macrocycle **1e**, 2 mM in D_2O, 500 MHz at 278 K



^{1}H NMR 2D gTOCSY of macrocycle **1e**, 2 mM in D_2O, 500 MHz at 278 K
200-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME RKS-II-132
EXPNO 5
PROCNO 1

F2 - Acquisition Parameters
Date_ 20121008
Time 22.49
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG mlevgp_mo
TD 2048
SOLVENT D2O
NS 8
DS 16
SWH 5000.000 Hz
FIDRES 2.441406 Hz
AQ 0.2048500 sec
RG 57
DW 100.000 usec
DE 6.00 usec
TE 278.0 K
d0 0.00003300 sec
D1 2.00000000 sec
D9 0.20000000 sec
d12 0.00002000 sec
D16 0.00020000 sec
FACTOR1 14
IN0 0.00020000 sec
I1 84
SCALEF 6

======== CHANNEL f1 ========
NUC1 1H
P1 7.50 usec
p5 23.34 usec
P6 35.00 usec
p7 70.00 usec
P17 2500.00 usec
PL1 1.60 dB
PL10 15.20 dB
SFO1 500.2224009 MHz

====== GRADIENT CHANNEL ====
GPNAM1 sine.100
GPNAM2 sine.100
GPX1 0.00 %
GPX2 0.00 %
GPY1 0.00 %
GPY2 0.00 %
GPZ1 10.00 %
GPZ2 10.00 %
P16 1000.00 usec

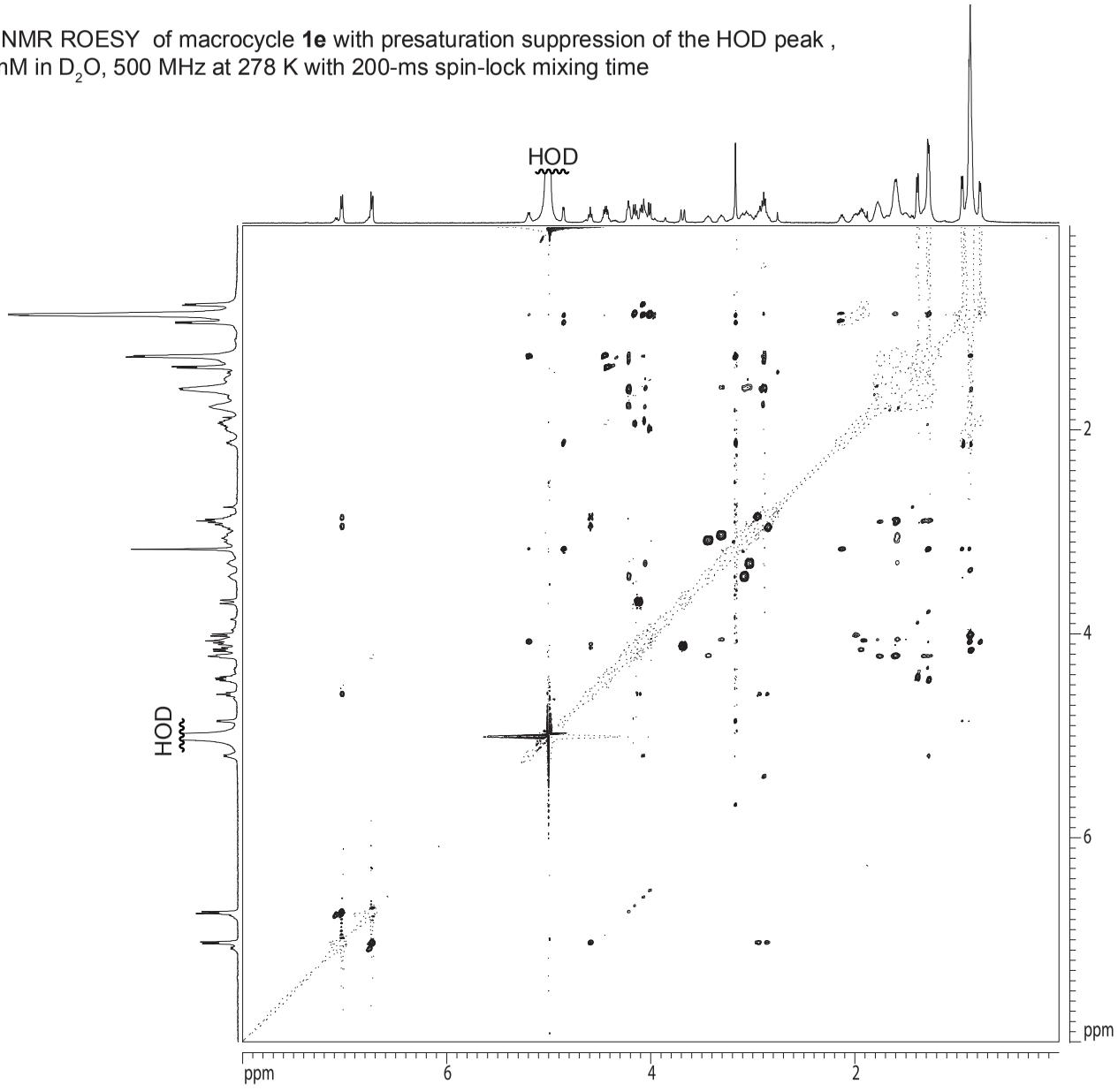
F1 - Acquisition parameters
ND0 1
TD 512
SFO1 500.2224 MHz
FIDRES 9.765625 Hz
SW 9.996 ppm
FnMODE undefined

F2 - Processing parameters
SI 1024
SF 500.2198993 MHz
WDW SINE
SSB 0
LB 0.00 Hz
GB 0
PC 4.00

F1 - Processing parameters
SI 1024
MC2 QF
SF 500.2198993 MHz
WDW SINE
SSB 0
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 8.000 ppm
F2LO 4001.76 Hz
F2PHI 0.000 ppm
F2HI 0.00 Hz
F1PLO 8.000 ppm
F1LO 4001.76 Hz
F1PHI 0.000 ppm
F1HI 0.00 Hz
F2PPMCM 0.53333 ppm/cm
F2HZCM 266.78397 Hz/cm
F1PPMCM 0.53333 ppm/cm
F1HZCM 266.78397 Hz/cm

^{1}H NMR ROESY of macrocycle **1e** with presaturation suppression of the HOD peak , 2 mM in D_2O, 500 MHz at 278 K with 200-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME RKS-II-132
EXPNO 8
PROCNO 2

F2 - Acquisition Parameters
Date_ 20121009
Time 7.09
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesyprtp.2.prd
TD 2048
SOLVENT D2O
NS 8
DS 16
SWH 5000.000 Hz
FIDRES 2.441406 Hz
AQ 0.2048500 sec
RG 35.9
DW 100.000 usec
DE 6.00 usec
TE 278.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.20000000 sec
d11 0.03000000 sec
d12 0.00002000 sec
d13 0.00000300 sec
IN0 0.00010000 sec
L4 400

======== CHANNEL f1 ========
NUC1 1H
P1 7.50 usec
P25 250.00 usec
PL1 1.60 dB
PL9 54.00 dB
PL11 26.60 dB
SFO1 500.2224009 MHz

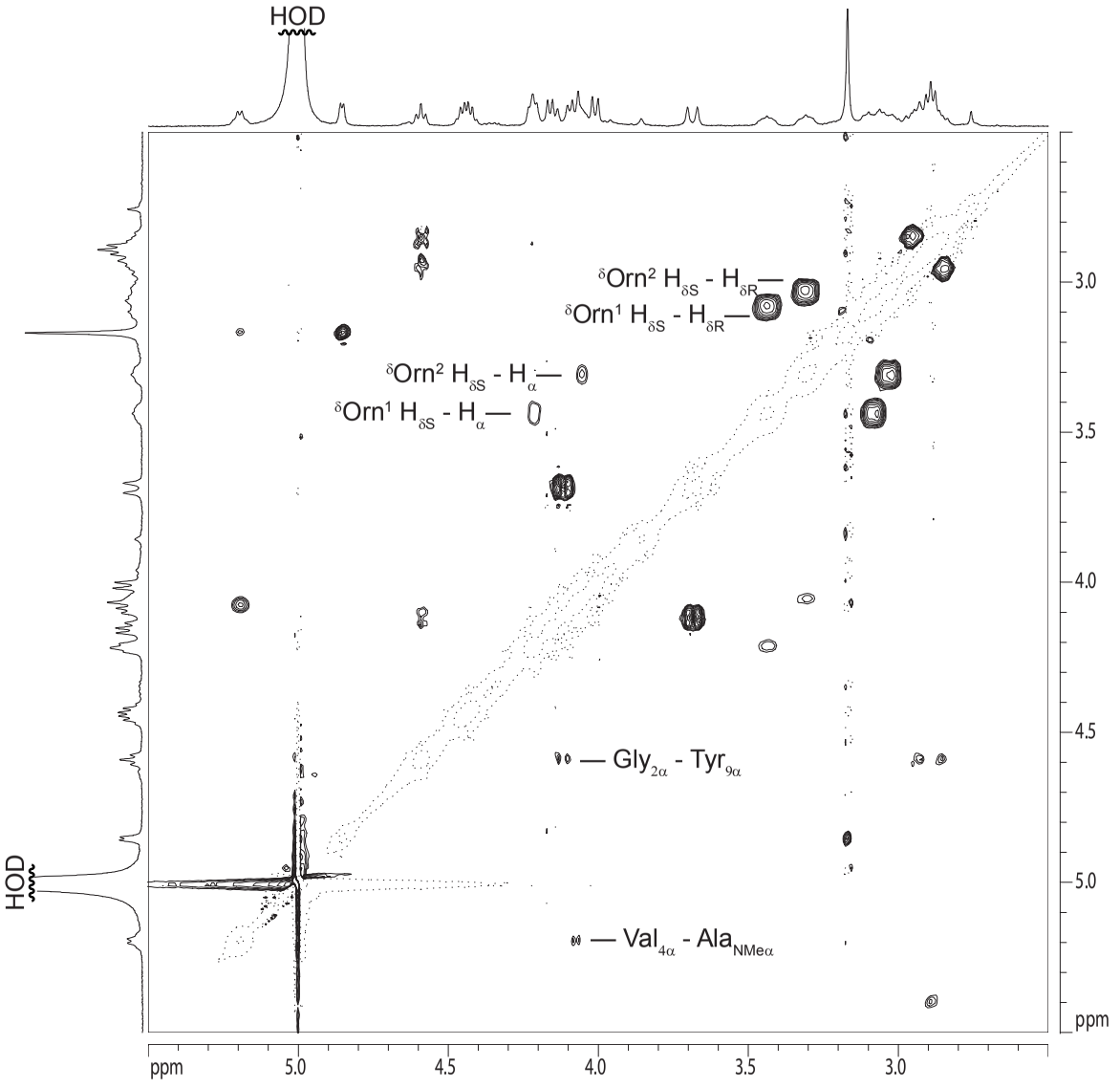
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 9.765625 Hz
SW 9.996 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2198993 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 1.40

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2198993 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 8.000 ppm
F2LO 4001.76 Hz
F2PHI 0.000 ppm
F2HI 0.00 Hz
F1PLO 8.000 ppm
F1LO 4001.76 Hz
F1PHI 0.000 ppm
F1HI 0.00 Hz
F2PPMCM 0.53333 ppm/cm
F2HZCM 266.78397 Hz/cm
F1PPMCM 0.53333 ppm/cm
F1HZCM 266.78397 Hz/cm

^{1}H NMR ROESY of macrocycle **1e** with presaturation suppression of the HOD peak ,
2 mM in D_2O, 500 MHz at 278 K with 200-ms spin-lock mixing time



Current Data Parameters
USER spence
NAME RKS-II-132
EXPNO 8
PROCNO 2

F2 - Acquisition Parameters
Date_ 20121009
Time 7.09
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG roesyprtp.2.prd
TD 2048
SOLVENT D2O
NS 8
DS 16
SWH 5000.000 Hz
FIDRES 2.441406 Hz
AQ 0.2048500 sec
RG 35.9
DW 100.000 usec
DE 6.00 usec
TE 278.0 K
d0 0.00000300 sec
D1 2.00000000 sec
d8 0.20000000 sec
d11 0.03000000 sec
d12 0.00002000 sec
d13 0.00000300 sec
IN0 0.00010000 sec
L4 400

======== CHANNEL f1 ========
NUC1 1H
P1 7.50 usec
P25 250.00 usec
PL1 1.60 dB
PL9 54.00 dB
PL11 26.60 dB
SFO1 500.2224009 MHz

F1 - Acquisition parameters
ND0 2
TD 512
SFO1 500.2224 MHz
FIDRES 9.765625 Hz
SW 9.996 ppm
FnMODE TPPI

F2 - Processing parameters
SI 1024
SF 500.2198993 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0
PC 1.40

F1 - Processing parameters
SI 1024
MC2 TPPI
SF 500.2198993 MHz
WDW QSINE
SSB 2
LB 0.00 Hz
GB 0

2D NMR plot parameters
CX2 15.00 cm
CX1 15.00 cm
F2PLO 5.500 ppm
F2LO 2751.21 Hz
F2PHI 2.500 ppm
F2HI 1250.55 Hz
F1PLO 5.500 ppm
F1LO 2751.21 Hz
F1PHI 2.500 ppm
F1HI 1250.55 Hz
F2PPMCM 0.20000 ppm/cm
F2HZCM 100.04398 Hz/cm
F1PPMCM 0.20000 ppm/cm
F1HZCM 100.04398 Hz/cm

Chapter 3

X-ray Crystallographic Structures of Trimers and Higher-Order Oligomeric Assemblies of a Peptide Derived from $A\beta_{17–36}$.[a]

Introduction

Here I report the X-ray crystallographic structures of trimers and higher-order oligomeric assemblies of a peptide derived from the β-amyloid peptide (Aβ). Oligomers of Aβ are now thought to play a central role in neurodegeneration in Alzheimers disease.[1–12] Selkoe and coworkers found that small Aβ oligomers disrupt longterm potentiation (LTP), with trimers showing the highest disruption.[7] Ashe and coworkers found that a 56 kDa Aβ oligomer, termed Aβ*56, impairs memory.[6] The oligomer appears to be a dodecamer composed of four

[a]Reproduced with permission from Spencer, R. K.; Li, H.; Nowick, J. S. *J. Am. Chem. Soc.* **2014**, *136*, 5595–5598. Copyright 2014 American Chemical Society.

trimers. Understanding the structures of these oligomers is essential to understanding their mechanism of action. Trimers are particularly enigmatic, because their structure cannot be explained by simply pairing monomers.

The hydrophobic central and C-terminal regions of Aβ are known to participate in aggregation to form fibrils and are likely involved in the aggregation of oligomers.[13–18] Although many molecular details of the aggregation processes have yet to be elucidated, the formation of β-sheets appears to be involved. While the structures of the fibrils are relatively well understood, the structures of trimers and higher-order oligomers are not known.[19–21]

Peptide fragments derived from amyloidogenic peptides and proteins are valuable tools for studying the structures of amyloid fibrils and oligomers.[22–29] In the current study, I set out to use a peptide fragment derived from both the central and C-terminal regions of Aβ to elucidate the structures of Aβ oligomers.[15,16] With the assistance of undergraduate student, Hao Li, I designed peptide **1a** as a mimic of $A\beta_{17–36}$ in which residues 17–23 (LVFFAED) and 30–36 (AIIGLMV) form the β-strands of a β-hairpin (Figure 3.1). In this structure, the δ-linked ornithine β-turn mimic connecting Asp_{23} and Ala_{30} promotes a β-hairpin structure and replaces residues 24–29 (VGSNKG). The δ-linked ornithine connecting Leu_{17} and Val_{36} forms a macrocycle to further enforce the β-hairpin structure. I incorporated an *N*-methyl group on Gly_{33} to help prevent uncontrolled aggregation through edge-to-edge hydrogen bonding between β-sheets.[30] I replaced Met_{35} with the hydrophilic isostere ornithine (α-linked) to enhance solubility and further prevent uncontrolled aggregation. I designed an analogue of peptide **1a** containing *p*-iodophenylalanine at the Phe_{19} position (peptide **1b**) for crystallographic phase determination.

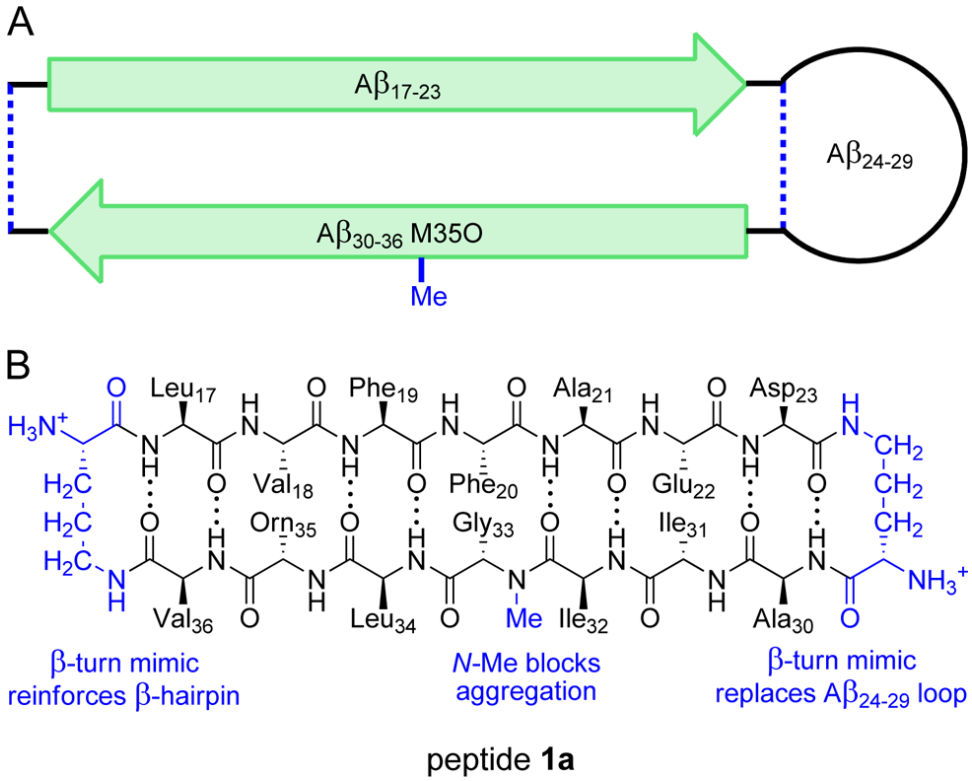


Figure 3.1: Peptide **1a**. (A) Cartoon illustrating the design of peptide **1a** and the envisioned relationship to $A\beta_{17–36}$. (B) Chemical structure of peptide **1a** illustrating $A\beta_{17–23}$, $A\beta_{30–36}$ M35O, the *N*-methyl group and the δ-linked ornithine β-turn mimic.

Results and Discussion

The synthesis and crystallization of peptides **1a** and **1b** was straightforward. The peptides were synthesized using Fmoc-based solid-phase peptide synthesis (SPPS), solution-phase cyclization, and RP-HPLC purification. Initial crystallization conditions for peptide **1a** were identified using the Hampton Research crystallization kits: Crystal Screen, Index, and PEG/Ion (288 experiments). Conditions with HEPES buffer and Jeffamine M-600 were selected and further optimized (0.1 M HEPES at pH 6.75 with 30% Jeffamine M-600) to give rapid (< 24 hours) formation of good crystals. Crystal diffraction data were collected at the Advanced Light Source at Law-rence Berkeley National Laboratory with a synchrotron source at 1.0 Å wavelength. Diffraction data were collected to 1.70 Å resolution. Data were scaled and merged using XDS and phases were determined by isomorphous replacement

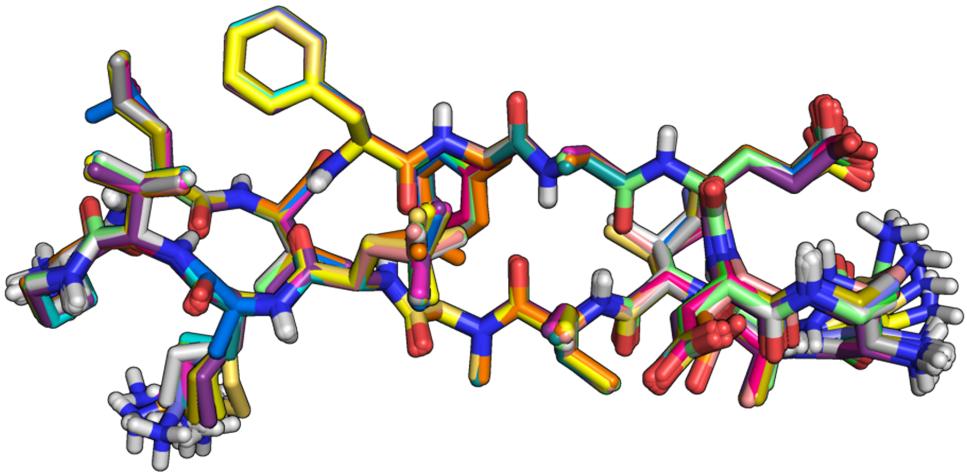


Figure 3.2: Overlay of the 16 β-hairpins in the asymmetric unit of the X-ray crystallographic structure of peptide **1a** (RMSD ≈ 0.2 Å)

of the Phe_{19} *p*-iodophenylalanine derivative **1b**.[31] The structure of peptide **1a** was solved and refined in the *R*3 space group. The asymmetric unit contains sixteen nearly identical monomers (RMSD ≈ 0.2 Å, Figure 3.2). Coordinates for hydrogens were generated by phenix.refine during refinement.[32]

The X-ray crystallographic structure reveals that peptide **1a** folds to form a β-hairpin comprising two heptapeptide β-strands. Eight residues (Leu_{17}, Phe_{19}, Ala_{21}, Asp_{23}, Ala_{30}, Ile_{32}, Leu_{34}, and Val_{36}) make up one surface of the β-hairpin (the LFA face), and six residues (Val_{18}, Phe_{20}, Glu_{22}, Ile_{31}, Gly_{33}, and Orn_{35}) make up the other surface (the VF face). The β-hairpin has a strong twist in which each residue rotates in a right-handed fashion approximately 25° along the β-strand axis. Contacts between the side chains of residues located diagonally across the β-sheet help stabilize the twisted β-hairpin (Phe_{19}-Val_{36}, Ala_{21}-Leu_{34}, and Asp_{23}-Ile_{32}; Figure 3.3).

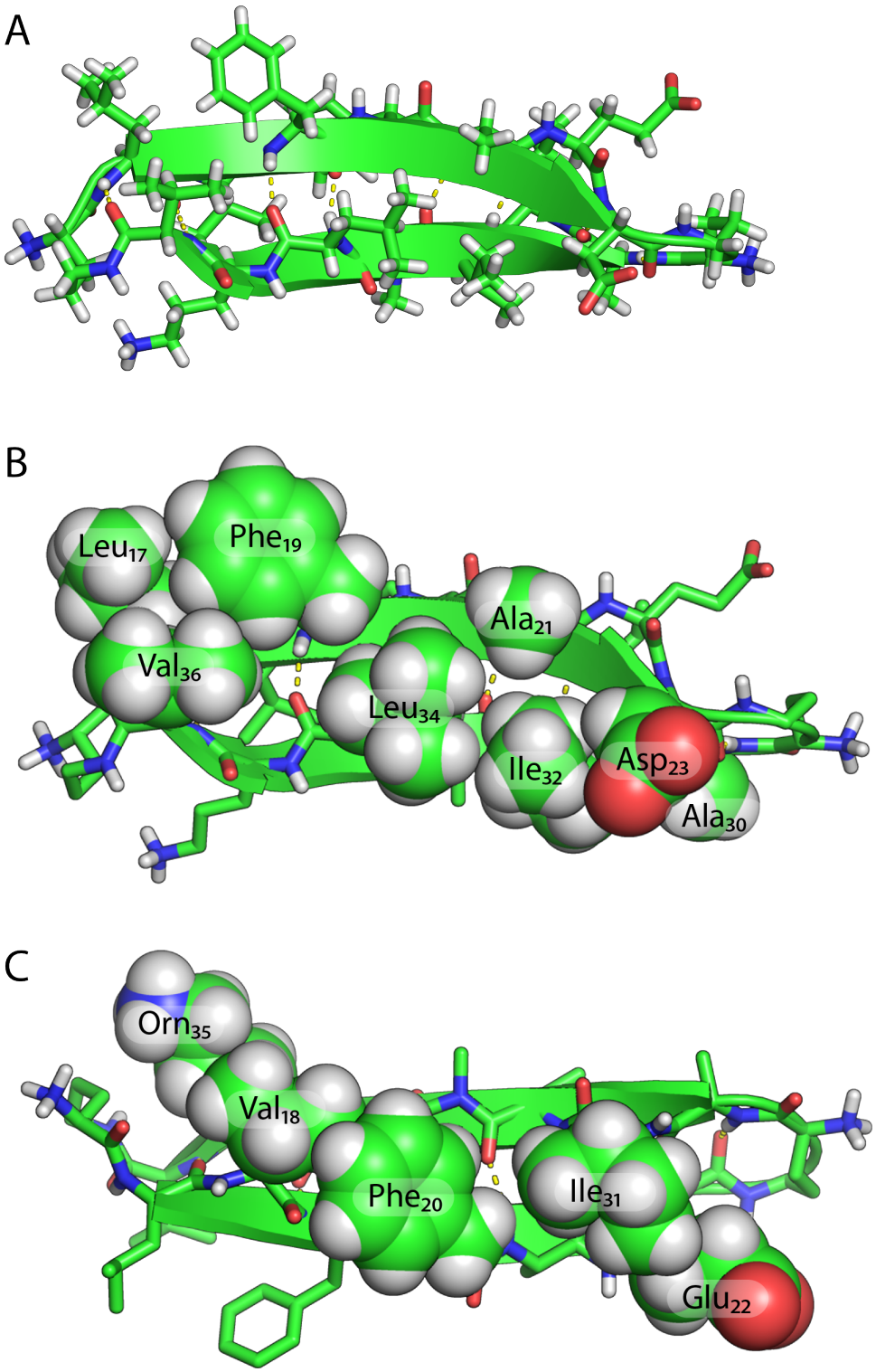


Figure 3.3: X-ray crystallographic structure of peptide **1a**. (A) β-Hairpin monomer. Contacts among side chains on the LFA (B) and VF (C) faces of peptide **1a**.

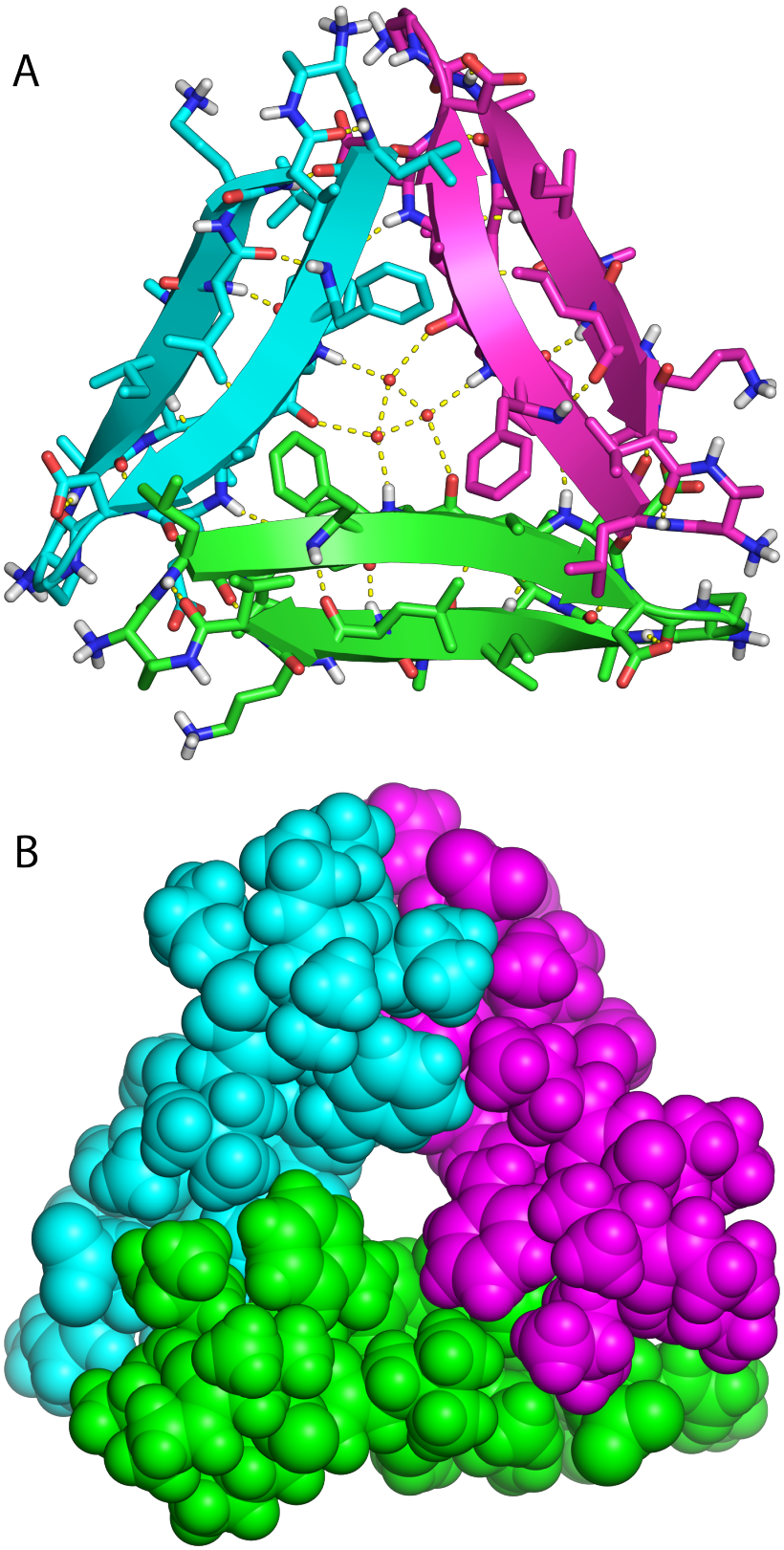


Figure 3.4: (A) Trimer, cartoon and stick representation with ordered water. (B) Trimer, space-filling representation.

Three β-hairpins assemble in a triangular fashion and interlock to form a trimer, with each β-hairpin making up one edge of the equilateral triangle (Figure 3.4). The central Aβ strand (17–23) constitutes the inner edges of the equilateral triangle, and the C-terminal strand (30–36) constitutes the outer edges. A hole runs through the center of the triangle, and the NH and C=O groups of Phe_{20} line the hole. The side chains of Phe_{19} surround the front of the hole and the side chains of Phe_{20} surround the back of the hole. Three ordered water molecules hydrogen bond to the NH and C=O groups of Phe_{20} and fill the hole in the center of the triangle (Figure 3.4A). The three ordered waters are in unusually close contact, with an average distance of ca. 2.2 Å between oxygen atoms. The β-hairpins come together at the corners of the triangle and stabilize the trimer through both polar and non-polar interactions. The main chains of Val_{18} and Glu_{22} hydrogen bond at each corner to create a four-stranded β-sheet. The side chains of Leu_{17}, Phe_{19}, Ala_{21}, Asp_{23}, Ile_{32}, Leu_{34}, and Val_{36} form an extensive hydrophobic cluster at the corners where the β-hairpins meet. The side chain of Leu_{17} makes extensive contacts with the side chains of Ala_{21}, Asp_{23}, Ile_{32}, and Leu_{34} of the adjacent β-hairpin and is buttressed by the side chains of Phe_{19} and Val_{36} (Figure 3.5).

The trimers assemble loosely in the crystal lattice to form hexamers and dodecamers (Figure 3.6). The hexamer consists of two trimers clasped through the VF faces. In the hexamer, the Ile_{31} residues of the two trimers hydrogen bond through bridging waters. Loose contacts between Phe_{20} and Ile_{31} appear to further stabilize the hexamer. The dodecamer consists of four trimers in a tetrahedral arrangement around a central cavity, with the LFA faces lining the cavity. Hydrophobic contacts occur between the Leu_{17} residues of the trimers. Salt bridges between the Asp_{23} residues and the δ-linked ornithines preceding Leu_{17} further stabilize the dodecamer. The contacts stabilizing the hexamers and dodecamers appear to be in opposition, preventing either oligomer from packing tightly.

To test whether trimer formation results from *N*-methylation of Gly_{33}, I prepared and studied peptide **2a**, in which Phe_{20} is *N*-methylated and Gly_{33} is not. Peptide **2a** crystallizes

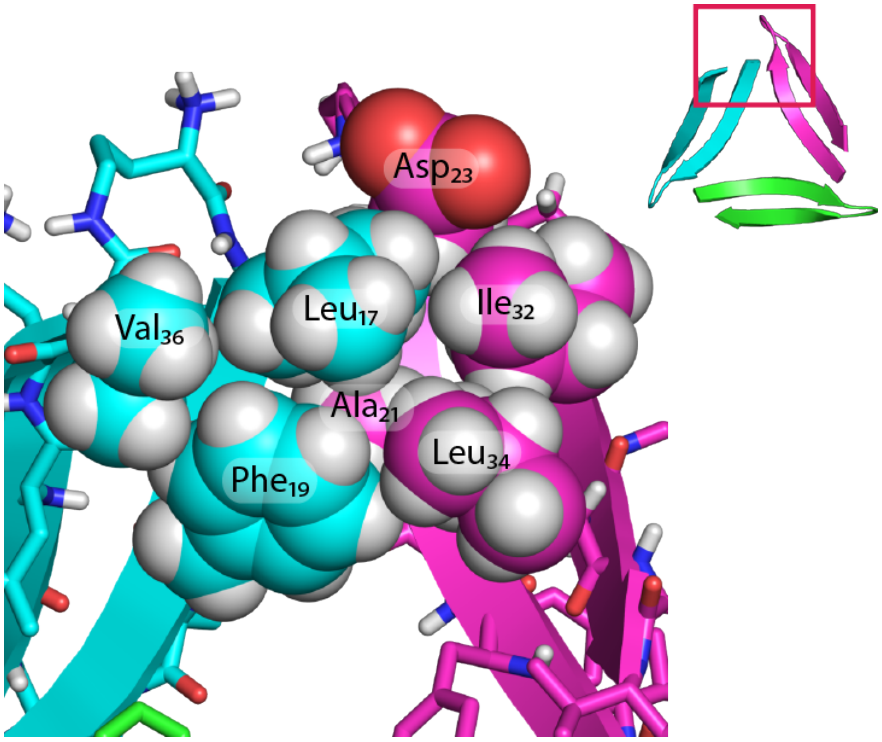
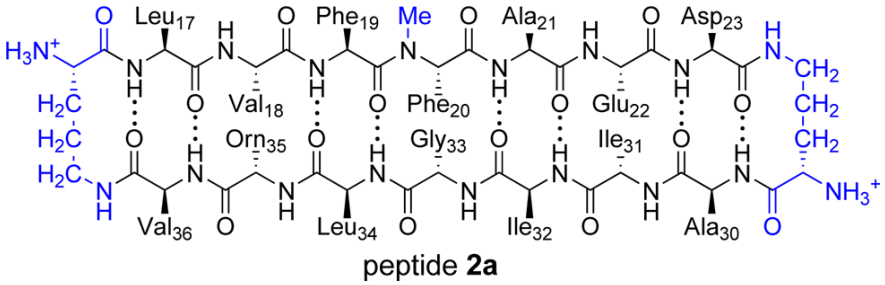


Figure 3.5: Detail of trimer interface illustrating the hydrophobic cluster formed by Leu_{17}, Phe_{19}, Ala_{21}, Asp_{23}, Ile_{32}, Leu_{34}, and Val_{36}.

in conditions similar to peptide **1a**. Phases were determined by isomorphous replacement of the Phe_{19} *p*-iodophenylalanine derivative **2b**. The structure of peptide **2a** was solved and refined in the $P3_221$ space group and contains twelve nearly identical monomers (RMSD $\approx$ 0.3 Å, Figure 3.7).



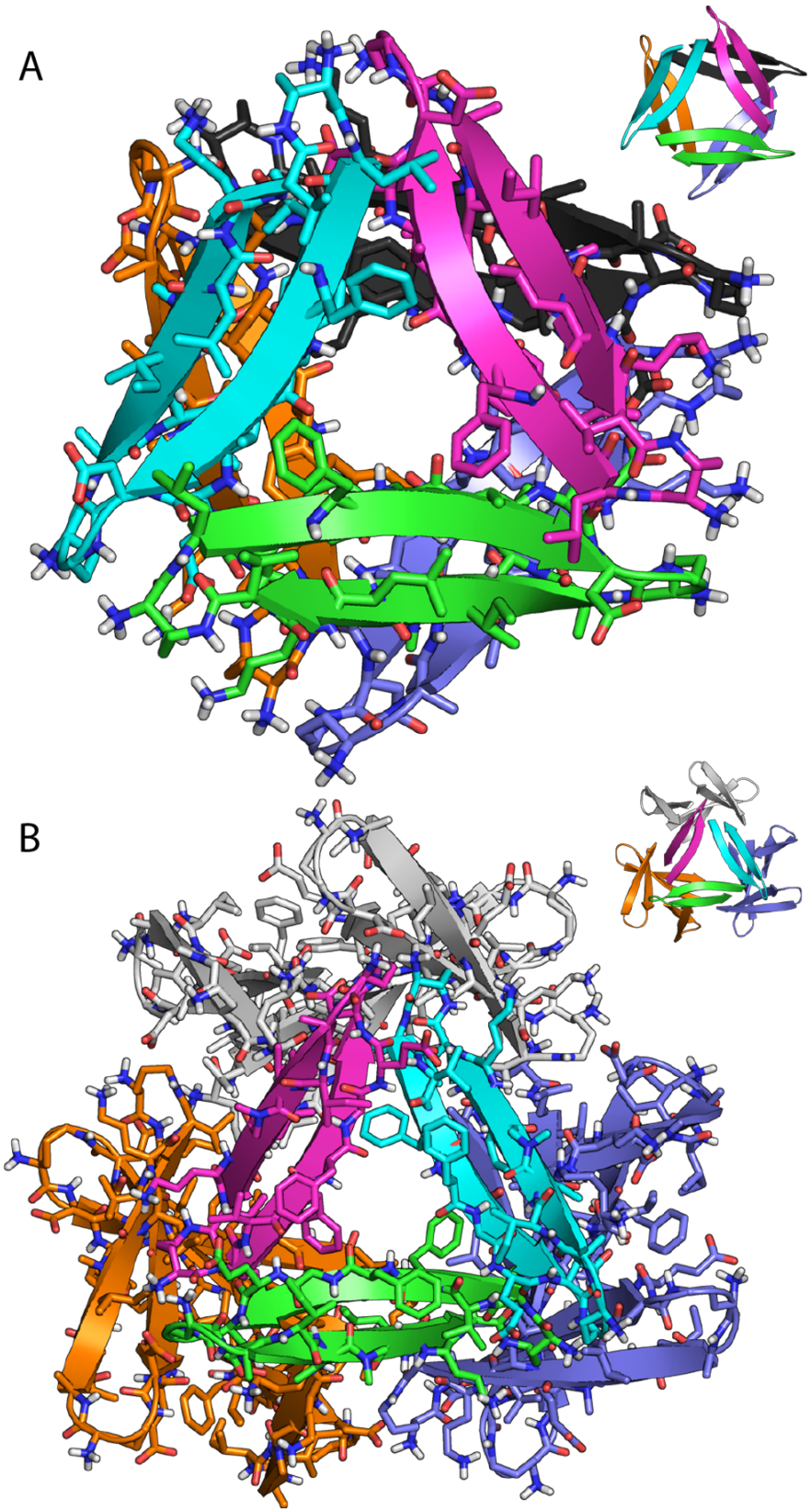


Figure 3.6: Hexamer (A) and dodecamer (B) observed in the X-ray crystallographic structure of peptide **1a**.

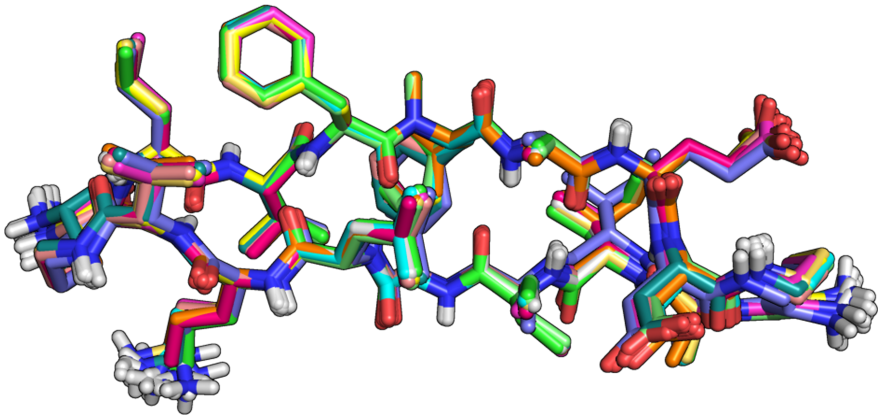


Figure 3.7: Overlay of the 12 β-hairpins in the asymmetric unit of the X-ray crystallographic structure of peptide **2a** (RMSD ≈ 0.3 Å).

The X-ray crystallographic structure of peptide **2a** is nearly identical to that of peptide **1a**. Peptide **2a** crystallizes as a β-hairpin that assembles to form trimers, which further form hexamers and dodecamers. Moving the *N*-methyl group from Gly_{33} to Phe_{20} does not significantly alter the structures of the oligomers. The central Aβ strand (17–23) still forms the inner edges of the trimer, and the C-terminal strand (30–36) still forms the outer edges. In the X-ray crystallographic structure of peptide **2a**, the *N*-methyl groups from Phe_{20} replace the three ordered waters that fill the hole in the center of the triangle (Figure 3.8). While *N*-methylation of either Phe_{20} or Gly_{33} is necessary to prevent aggregation, it does not dictate the formation of the trimer, hexamer, or dodecamer.

The trimer formed by peptides **1a** and **2a** is unlike the structures of fibrils or oligomers previously observed for amyloidogenic peptides or proteins. A similar triangular assembly of three β-hairpins occurs in actinohivin, a 114-amino-acid lectin that binds HIV gp120.[33,34] Actinohivin contains three nearly identical sequences in tandem that form the three β-hairpins and fold into a triangular assembly (Figure 3.9, PDB 3A07). Each β-hairpin comprises two pentapeptide β-strands connected by a turn of three residues. Although these β-strands are smaller than those of peptide **1a**, they occupy similar positions to residues 17–23 and 30–36 of **1a** in the trimer. Actinohivin binds to mannose-containing glycans on gp120 with high affinity through trivalent interactions of the concave surfaces formed by

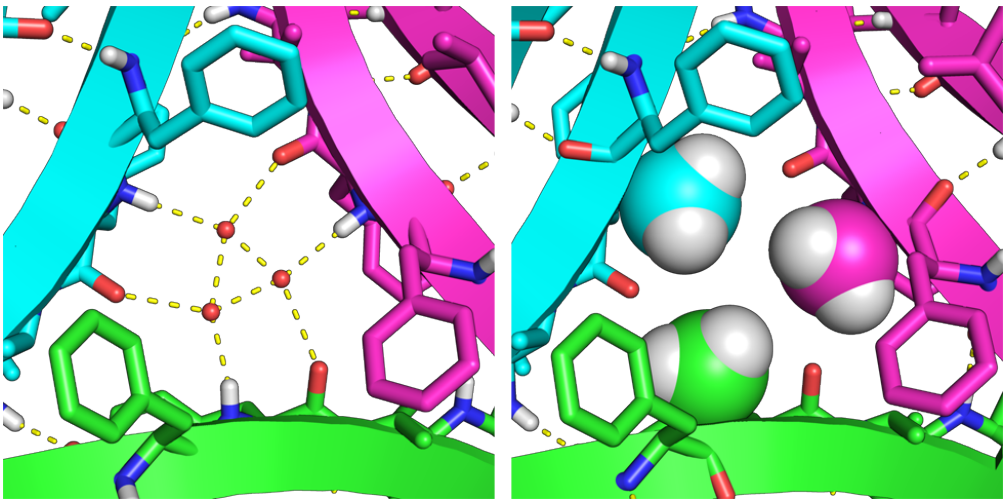


Figure 3.8: Detail of the X-ray crystallographic structure of the trimers formed by peptides **1a** (left) and **2a** (right). The *N*-methyl groups of **2a** take the place of the ordered water molecules in **1a**.

the three near-repeats. This mode of interaction suggests that trimers of Aβ may bind to molecules on the surface of neurons through trivalent interactions of three β-hairpins formed by $A\beta_{17–36}$.

I modeled a trimer of Ac-$A\beta_{17–36}$-NHMe β-hairpins to generate a working model of a trimer of Aβ. I used the crystallographic coordinates of peptide **1a** to generate residues 17–23 (LVFFAED) and 30–36 (AIIGLMV) of the trimer and added loops comprising residues 24–29 (VGSNKG). I performed replica-exchange molecular dynamics (REMD) in NAMD using the CHARMM22 force field with generalized Born implicit solvent (GBIS) to generate realistic conformations of the loops.[35,36] Figure 3.10 illustrates 20 low-energy structures from the simulation. These structures provide a working model in which residues 24–29 act as a loop connecting the 17–23 and 30–36 β-strands. The concave surfaces formed by the loops and the twisted β-strands might serve as binding sites in multivalent biological interactions.

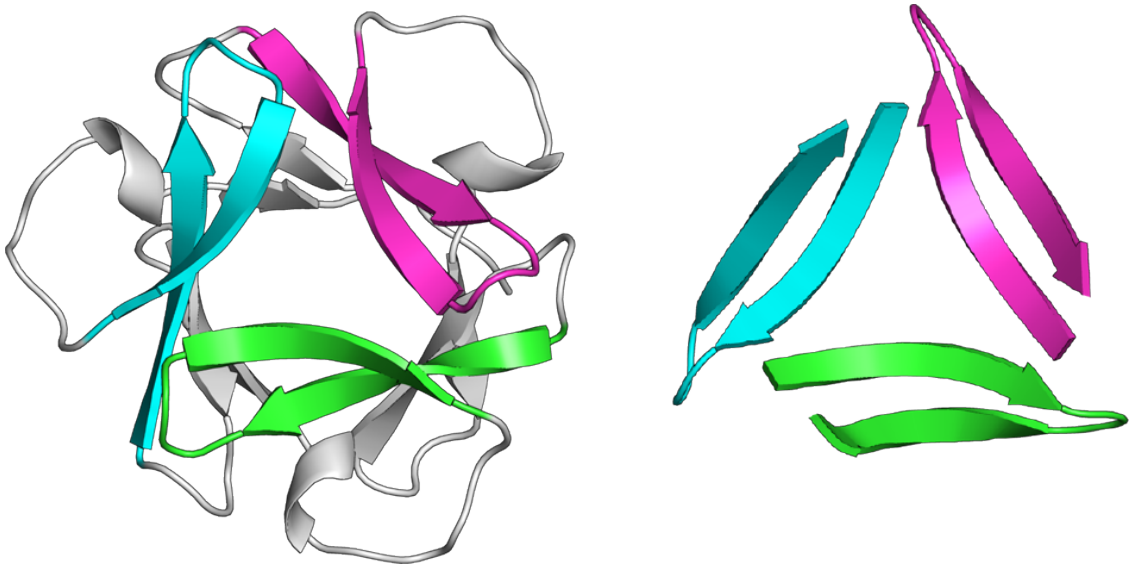


Figure 3.9: Cartoon representation showing the structural similarities of actinohivin (left) and peptide **1a** trimer (right).

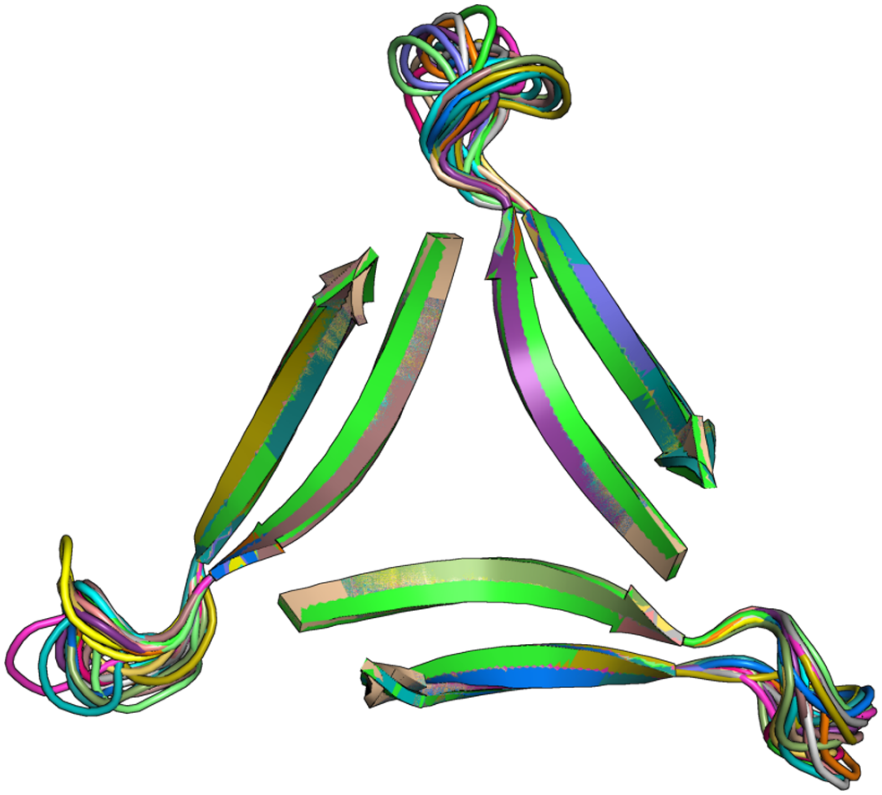


Figure 3.10: Cartoon representation of 20 low-energy structures for Ac-Aβ_{17-36}-NHMe generated by REMD.

Conclusion

The X-ray crystallographic structures of the trimers and higher-order oligomeric assemblies formed by peptides **1a** and **2a** provide working models for the structures of the trimers and higher-order oligomers of Aβ that are important in neurodegeneration. In this model, Aβ associates to form trimers comprising three β-hairpins in a triangular arrangement. The trimers are stabilized through interactions among the central region of $A\beta_{17–23}$ (LVFFAED). These interactions are buttressed by hydrophobic interactions with the hydrophobic C-terminal region of Aβ — either $A\beta_{30–36}$ (AIIGLMV) or another segment of the extensive hydrophobic C-terminal sequence. Two trimers can stack to form hexamers. While the hexamers formed by peptides **1a** and **2a** associate along the VF face, there is little distinction between the residues of the VF face and the LFA face, and stacking might occur through interactions among either face. Although the dodecamer in Figure 3.6B might explain the structure of Aβ⋆56, an assembly consisting of a stack of four trimers is also possible and may provide greater contact and stabilization. These structures and ideas serve as a starting point for developing and testing hypotheses about the structures and mechanism of action of amyloid oligomers.

References

1. Lambert, M. P.; Barlow, A. K.; Chromy, B. A.; Edwards, C.; Freed, R.; Liosatos, M.; Morgan, T. E.; Rozovsky, I.; Trommer, B.; Viola, K. L.; Wals, P.; Zhang, C.; Finch, C. E.; Krafft, G. A.; Klein, W. L. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 6448–6453.

2. Selkoe, D. J. *Science* **2002**, *298*, 789–791.

3. Nagele, R. G.; D'Andrea, M. R.; Anderson, W. J.; Wang, H. Y. *Neuroscience* **2002**, *110*, 199–211.

4. Walsh, D. M.; Klyubin, I.; Fadeeva, J. V.; Cullen, W. K.; Anwyl, R.; Wolfe, M. S.; Rowan, M. J.; Selkoe, D. J. *Nature* **2002**, *416*, 535–539.

5. Kayed, R.; Head, E.; Thompson, J. L.; McIntire, T. M.; Milton, S. C.; Cotman, C. W.; Glabe, C. G. *Science* **2003**, *300*, 486–489.

6. Lesné, S.; Koh, M. T.; Kotilinek, L.; Kayed, R.; Glabe, C. G.; Yang, A.; Gallagher, M.; Ashe, K. H. *Nature* **2006**, *440*, 352–357.

7. Townsend, M.; Shankar, G. M.; Mehta, T.; Walsh, D. M.; Selkoe, D. J. *J. Physiol.* **2006**, *572*, 477–492.

8. Haass, C.; Selkoe, D. J. *Nat. Rev. Mol. Cell. Biol.* **2007**, *8*, 101–112.

9. Shankar, G. M.; Li, S. M.; Mehta, T. H.; Garcia-Munoz, A.; Shepardson, N. E.; Smith, I.; Brett, F. M.; Farrell, M. A.; Rowan, M. J.; Lemere, C. A.; Regan, C. M.; Walsh, D. M.; Sabatini, B. L.; Selkoe, D. J. *Nature Medicine* **2008**, *14*, 837–842.

10. Zhao, W.-Q.; De Felice, F. G.; Fernandez, S.; Chen, H.; Lambert, M. P.; Quon, M. J.; Krafft, G. A.; Klein, W. L. *FASEB J* **2008**, *22*, 246–260.

11. Querfurth, H. W.; LaFerla, F. M. *N. Engl. J. Med.* **2010**, *362*, 329–344.

12. Fändrich, M. *J. Mol. Biol.* **2012**, *421*, 427–440.

13. Thirumalai, D.; Klimov, D. K.; Dima, R. I. *Curr. Opin. Struct. Biol.* **2003**, *13*, 146–159.

14. Liu, R.; McAllister, C.; Lyubchenko, Y.; Sierks, M. R. *J. Neurosci. Res.* **2004**, *75*, 162–171.

15. Hoyer, W.; Gronwall, C.; Jonsson, A.; Stahl, S.; Hrd, T. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 5099–5104.

16. Sandberg, A.; Luheshi, L. M.; Sllvander, S.; Pereira de Barros, T.; Macao, B.; Knowles, T. P.; Biverstl, H.; Lendel, C.; Ekholm-Petterson, F.; Dubnovitsky, A.; Lannfelt, L.; Dobson, C. M.; Hrd, T. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 15595–15600.

17. Yu, L. *et al. Biochemistry* **2009**, *48*, 1870–1877.

18. Cerf, E.; Sarroukh, R.; Tamamizu-Kato, S.; Breydo, L.; Derclaye, S.; Dufrne, Y. F.; Narayanaswami, V.; Goormaghtigh, E.; Ruysschaert, J. M.; Raussens, V. *Biochem. J.* **2009**, *421*, 415–423.

19. Lührs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Dbeli, H.; Schubert, D.; Riek, R. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17342–17347.

20. Petkova, A. T.; Yau, W.-M.; Tycko, R. *Biochemistry* **2006**, *45*, 498–512.

21. Lu, J.-X.; Qiang, W.; Yau, W.-M.; Schwieters, C. D.; Meredith, S. C.; Tycko, R. *Cell* **2013**, *154*, 1257–1268.

22. Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. .; Riekel, C.; Grothe, R.; Eisenberg, D. *Nature* **2005**, *435*, 773–778.

23. Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J.; McFarlane, H. T.; Madsen, A. .; Riekel, C.; Eisenberg, D. *Nature* **2007**, *447*, 453–457.

24. Laganowsky, A.; Liu, C.; Sawaya, M. R.; Whitelegge, J. P.; Park, J.; Zhao, M.; Pensalfini, A.; Soriaga, A. B.; Landau, M.; Teng, P. K.; Cascio, D.; Glabe, C.; Eisenberg, D. *Science* **2012**, *335*, 1228–1231.

25. Liu, C.; Sawaya, M. R.; Cheng, P.-N.; Zheng, J.; Nowick, J. S.; Eisenberg, D. *J. Am. Chem. Soc.* **2011**, *133*, 6736–6744.

26. Liu, C.; Zhao, M.; Jiang, L.; Cheng, P.-N.; Park, J.; Sawaya, M. R.; Pensalfini, A.; Gou, D.; Berk, A. J.; Glabe, C. G.; Nowick, J. S.; Eisenberg, D. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20913–20918.

27. Cheng, P.-N.; Pham, J. D.; Nowick, J. S. *J. Am. Chem. Soc.* **2013**, *135*, 5477–5492.

28. Pham, J. D.; Chim, N.; Goulding, C. W.; Nowick, J. S. *J. Am. Chem. Soc.* **2013**, *135*, 12460–12467.

29. Apostol, M. I.; Perry, K.; Surewicz, W. K. *J. Am. Chem. Soc.* **2013**, *135*, 10202–10205.

30. Spencer, R.; Chen, K. H.; Manuel, G.; Nowick, J. S. *Eur. J. Org. Chem.* **2013**, 3523–3528.

31. Kabsch, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125–132.

32. Adams, P. D. *et al. Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213–221.

33. Tanaka, H.; Chiba, H.; Inokoshi, J.; Kuno, A.; Sugai, T.; Takahashi, A.; Ito, Y.; Tsunoda, M.; Suzuki, K.; Taknaka, A.; Sekiguchi, T.; Umeyama, H.; Hirabayashi, J.; mura, S. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 15633–15638.

34. Suzuki, K.; Tsunoda, M.; Hoque, M. M.; Zhang, F.; Jiang, J.; Zhang, X.; Ohbayashi, N.; Tanaka, H.; Taknaka, A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1818–25.

35. Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

36. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kal, L.; Schulten, K. *J. Comp. Chem.* **2005**, *26*, 1781–1802.

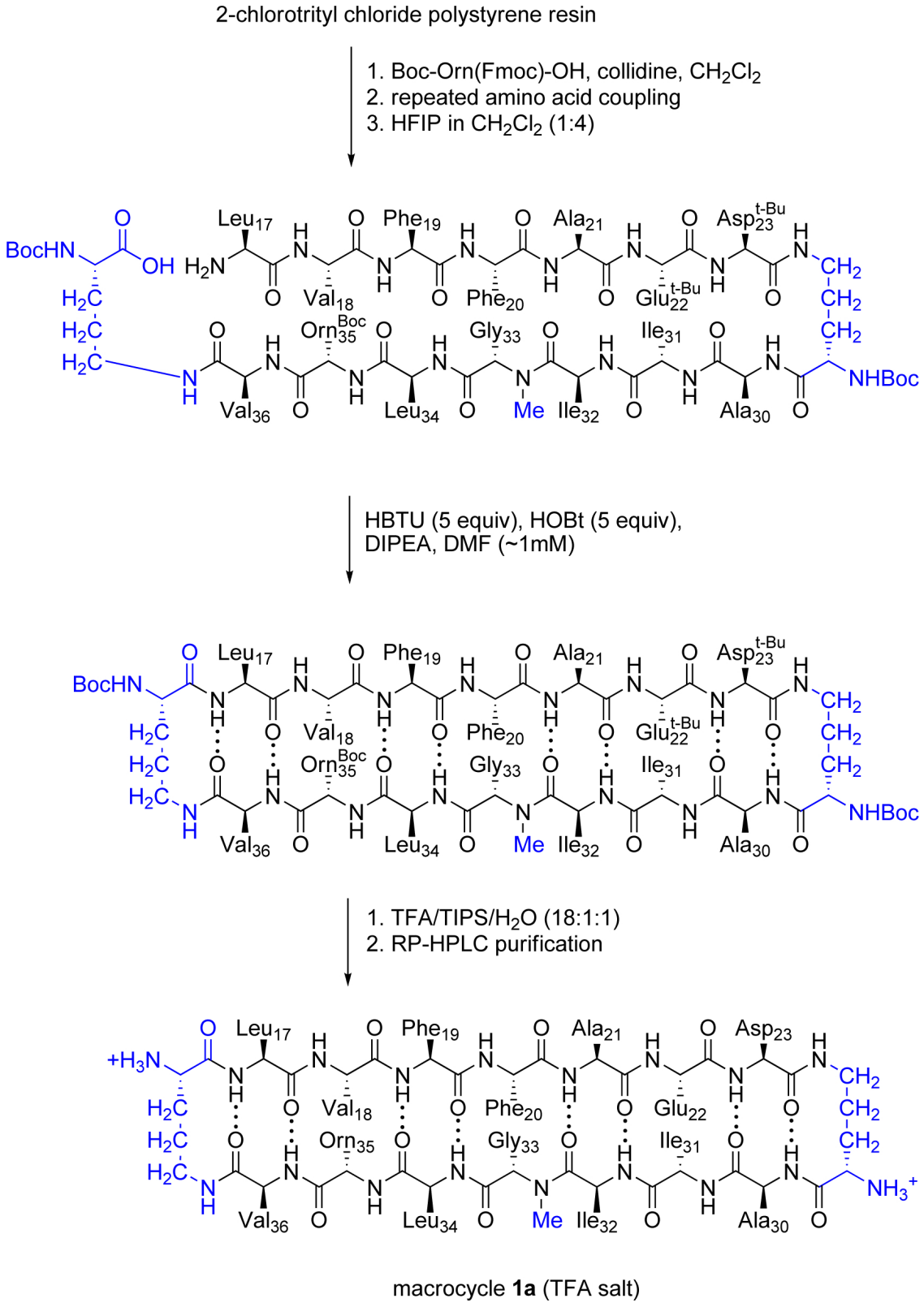
Materials and Methods

General Methods

Synthesis of peptides 1 and 2. Representative synthesis of peptide 1a.

Loading of the resin: 2-Chlorotrityl chloride resin (300 mg, 1.2 mmol/g) was added to a Bio-Rad Poly-Prep chromatography column (10 mL, 0.8×4.0 cm). The resin was suspended in dry CH_2Cl_2 (10 mL) and allowed to swell for 30 min. The solution was drained from the resin and a solution of Boc-Orn(Fmoc)-OH (0.50 equiv., 82 mg, 0.18 mmol) in 20% 2,4,6-collidine in dry CH_2Cl_2 (5 mL) was added immediately and the mixture was gently agitated for 12 h. The solution was then drained and a mixture of CH_2Cl_2/MeOH/DIPEA (17:2:1, 10 mL) was added immediately. The mixture was gently agitated for 1 h to cap the unreacted 2-chlorotrityl chloride sites. The resin was then washed with dry CH_2Cl_2 (2×) and dried by passing nitrogen through the vessel. In the synthesis of peptide **1a**, the resin loading was determined to be 0.14 mmol [0.46 mmol/g, 77% based on Boc-Orn(Fmoc)-OH] by UV analysis of the Fmoc cleavage product. Loadings of 0.120.15 mmol [70–80%, based on Boc-Orn(Fmoc)-OH] were typically observed in various repetitions of this procedure associated with the syntheses of peptides **1** and **2**.

Peptide Coupling: The PS-2-chlorotrityl-Orn(Fmoc)-Boc generated from the previous step was transferred to a solid-phase peptide synthesizer reaction vessel and submitted to cycles of automated peptide coupling with Fmoc-protected amino acid building blocks. The linear peptide was synthesized from the C-terminus to the N-terminus. Each coupling consisted of *i.* Fmoc-deprotection with 20% piperidine in DMF for 3 min, *ii.* washing with DMF (3x), *iii.* coupling of the amino acid (0.56 mmol, 4 equiv.) in the presence of HCTU (224 mg,



Scheme 3.1: Synthesis of Peptide **1a**.

0.56 mmol, 4 equiv.), and *iv.* washing with DMF (6x). Each amino acid coupling step took 20 min for all the residues of peptide **1a**. For peptides **1b** and **2b**, the phenylalanine and *p*-iodophenylalanine residues after the *N*-methyl-L-phenylalanine were double coupled (0.56 mmol, 4 equiv.) and allowed to react for 1 h per coupling with HATU (4 equiv.) and HOAt (4 equiv.). Other residues of peptides **1b** and **2b** were coupled as described previously. After coupling of the last amino acid, the terminal Fmoc group was removed with 20% piperidine in DMF. The resin was transferred from the reaction vessel of the peptide synthesizer to a Bio-Rad Poly-Prep chromatography column.

Cleavage of the Peptide from the Resin: The linear peptide was cleaved from the resin by agitating the resin for 1 hr with a solution of hexafluoroisopropanol (HFIP) in CH_2Cl_2 (1:4, 5 mL). The suspension was filtered and the filtrate was collected in a 250 mL round-bottomed flask. The resin was washed with additional HFIP in CH_2Cl_2 (1:4, 5 mL) and then with CH_2Cl_2 (2×10 mL). The combined filtrates were concentrated by rotary evaporation to give a white solid. The white solid was further dried by vacuum pump to afford the crude protected linear peptide, which was cyclized without further purification.

Cyclization of the Linear Peptide: Crude protected linear peptide was dissolved in dry DMF (125 mL). HOBt (95 mg, 0.70 mmol, 5 equiv.) and HBTU (264 mg, 0.70 mmol, 5 equiv.) were added to the solution. The reaction mixture was then stirred under nitrogen for 20 min. DIPEA (0.3 mL, 1.7 mmol, 12 equiv.) was added to the solution and the mixture was stirred under nitrogen for 24 h. The mixture was concentrated under reduced pressure to afford crude protected cyclic peptide.

Global Deprotection and Purification of the Cyclic Peptide: Protected cyclic peptide was dissolved in TFA/triisopropylsilane (TIPS)/H_2O (18:1:1, 10 mL) in a 250 mL round-bottomed flask equipped with a nitrogen-inlet adaptor. The solution was stirred for 1.5 h. The reaction mixture was then concentrated by rotary evaporation under reduced pressure to afford the deprotected cyclic peptide as a yellow oil. The oil was dissolved in H_2O and acetonitrile

(4:1, 5 mL) and the solution was filtered through a 0.20 µm syringe filter and purified by reversed-phase HPLC (gradient elution with 20–50% CH_3CN over 40 min). The pure fractions were lyophilized to afford 53 mg of the cyclic deprotected peptide 1a. The syntheses of peptides **1b**, **2a**, and **2b** afforded 64, 67, and 120 mg respectively.

Crystallization procedure

Initial crystallization conditions were determined using the hanging-drop vapor-diffusion method. Crystallization was performed in a 96-well format, with each well containing 100 µL of a solution from a Hampton 96-well screening kit. Three kits were used — Crystal Screen, Index, and Peg/Ion — for a total of 288 experiments (three 96-well plates). Hanging-drops were made by combining 300 nL of peptide **1b** (solution of 10 mg/mL in18 MΩ water) and 300 nL of the well solution using a TTP LabTech Mosquito nanodisperse instrument. Crystal grew rapidly (< 24 h) in a solution of 0.1 M 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES) buffer at pH 7.0 and Jeffamine M-600 at pH 7.0 (30% v/v).

Crystallization conditions were optimized using a 4×6 matrix Hampton VDX 24-well plate. The HEPES buffer pH was varied in each row by ±0.5 pH units (6.5, 7.0, 7.5, and 8.0) and the pH 7.0 Jeffamine M-600 concentration in each column by ±2% (24%, 26%, 28%, 30%, 32%, and 34% v/v). For the first well in the 4x6 matrix we combined 100 µL of 1 M HEPES pH 6.5, 480 µL of a 50% solution (v/v) of pH 7.0 Jeffamine M-600, and 420 µL of 18 MΩ water. [The 50% pH 7.0 Jeffamine M-600 solution was prepared by combining 200 mL of pH 10.0 Jeffamine M-600 and 200 mL of 18 MΩ water, titrating with hydrochloric acid to pH 7.0, and filtering through a 0.2 m syringe filter.] The other wells were prepared in analogous fashion, by combining 100 µL of HEPES buffer, pH 7.0 Jeffamine M-600, and 18 MΩ water for a total volume of 1 mL.

Three hanging-drops were prepared per borosilicate glass slide by combining a solution of peptides **1** or **2** (1 µL, 10 mg/mL) and the well solution (1 µL) in a ratio of 1:1, 2:1, and 1:2. Slides were inverted and pressed firmly against the silicone grease surrounding each well. Large crystals, (0.3 – 0.4 mm) grew in under 24 hours. Crystallization conditions were further optimized using smaller variations in HEPES buffer pH (±0.25 pH units) and Jeffamine M-600 concentrations ($pm1\%$). Crystal were harvested with a nylon loop attached to a copper or steel pin and flash frozen in liquid nitrogen prior to data collection. The optimized crystallization conditions for peptides **1a**, **1b**, **2a**, and **2b** are summarized in Table 3.1.

X-ray diffraction data collection and processing.

Diffraction data for peptides **1a** and **2a** were collected at Lawrence Berkeley National Laboratory (Berkeley, California) on synchrotron beamline 8.2.2 at 1.0 Åwavelength with 0.5° rotation and a detector distance of 220 mm. Diffraction data were scaled and merged using XDS. Electron density maps were generated by isomorphous replacement of coordinates from peptide **1b** using Phaser in software suite Phenix 1.8.4. Molecular manipulations of the models were performed with Coot. Coordinates were refined with phenix.refine. Diffraction data for peptides **1b** and **2b** were collected on a Rigaku Micromax-007HF X-ray diffractometer with a rotating copper anode at 1.54 Å wavelength with 0.5° rotation. Diffraction data were collected using CrystalClear. Diffraction data were scaled and merged using XDS. Coordinates for the anomalous signals were determined by HySS in the Phenix software suite. Electron density maps were generated using anomalous coordinates determined by HySS as initial positions in Autosol. Molecular manipulations of the models were performed with Coot. Coordinates were refined with phenix.refine.

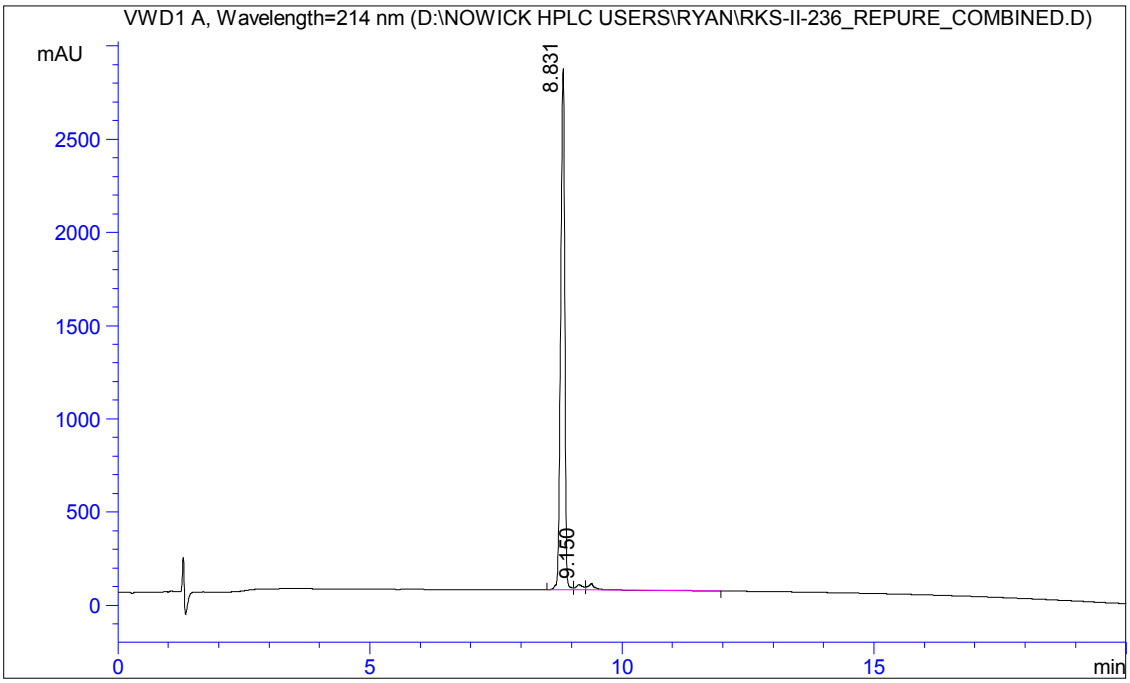
Table 3.1: Crystallographic Properties, Crystallization Conditions, and Data Collection and Model Refinement Statistics for Peptides **1a**, **1b**, **2a**, and **2b**.

peptide	1a	1b	2a	2b
PDB ID	4NTR	4NTP	4NW9	4NW8
space group	$R3$	$R3$	$P3_221$	$P3_221$
a, b, c (Å)	68 68 169.26	68.18 68.28 170.43	68.22 68.22 92.99	68.68 68.68 93.84
α, β, γ (°)	90 90 120	90 90 120	90 90 120	90 90 120
peptide per asymmetric unit	12	12	16	16
crystallization conditions	0.1 M HEPES, pH 6.75; 31% Jeffamine M-600	0.1 M HEPES, pH 6.50; 25% Jeffamine M-600	0.1 M HEPES, pH 6.50; 24% Jeffamine M-600	0.1 M HEPES, pH 7.50; 29% Jeffamine M-600
		Data Collection		
wavelength (Å)	1.00	1.54	1.00	1.54
resolution (Å)	34.36–1.70 (1.76–1.70)	34.09–1.99 (2.06–1.99)	36.54–1.65 (1.71–1.65)	29.74–2.02 (2.09–2.02)
total reflections	98412 (9014)	472254 (13643)	173645 (14559)	121088 (4284)
unique reflections	31708 (3062)	20197 (1852)	29970 (2640)	16770 (1268)
multiplicity	3.1	23.4	5.8	7.2
completeness (%)	98.74 (94.80)	99.11 (22.24)	98.42 (88.35)	97.38 (75.21)
mean I/σ	12.63 (2.25)	22.24 (3.42)	13.78 (1.02)	28.82 (3.41)
Wilson B-factor	24.24	23.13	31.33	25.70
R_{merge}	0.05 (0.55)	0.14 (0.49)	0.05 (1.48)	0.12 (0.32)
$R_{measure}$	0.06	0.14	0.06	0.13
$CC_{1/2}$	0.99 (0.81)	1.00 (0.89)	0.99 (0.53)	0.99 (0.91)
CC^*	0.99 (0.95)	1.00 (0.97)	1.00 (0.83)	1.00 (0.97)
		Refinement		
R_{work}	20.02	20.68	18.86	19.52
R_{free}	22.62	24.61	20.95	23.72
number of non-hydrogen atoms	2239	2294	1740	1703
RMS_{bonds}	0.008	0.008	0.012	0.007
RMS_{angles}	1.20	1.27	1.63	1.14
Ramachandran favored (%)	100	100	100	100
outliers (%)	0	0	0	0
clashscore	0.71	2.62	5.72	4.14
average B-factor	34.90	29.90	41.50	28.30

Modeling of Ac-A$\beta_{17–36}$-NHMe trimer using replica-exchange molecular dynamics (REMD).

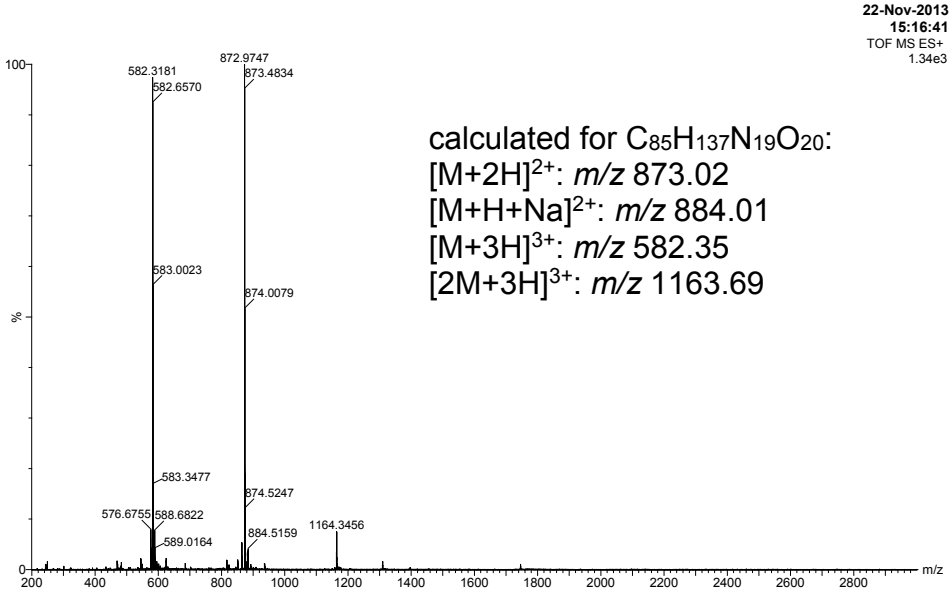
Coordinates for REMD were generated from the X-ray crystallographic coordinates of peptide **1a**. The trimer was edited in PyMOL as follows: The δ-linked ornithine turn units were removed. Aβ residues Val_{24}, Gly_{25}, Ser_{26}, Gln_{27}, Lys_{28}, and Gly_{29} were added to link Asp_{23} and Ala_{30}. Orn_{35} was mutated to Met_{35}. The N-terminus was patched as an acetylated amide (ACE) and the C-terminus was patched as a methylamide (CT3) in VMD. The requisite .psf file was generated using the autopsf plugin in VMD. The coordinates for the main chains of residues Leu_{17}–Asp_{23} and Ala_{30}–Val_{36} were frozen during the simulation. Residues Val_{24}–Gly_{29} and all side chains were allowed to move freely. REMD simulations were run in NAMD with the CHARMM22 force field and generalized Born implicit solvent (GBIS) on 32 replicas with a temperature range of 300K-800K for 20 ns. The coordinates for the 20 lowest energy conformations were selected.

HPLC and MS ESI+ TOF of peptide **1a**

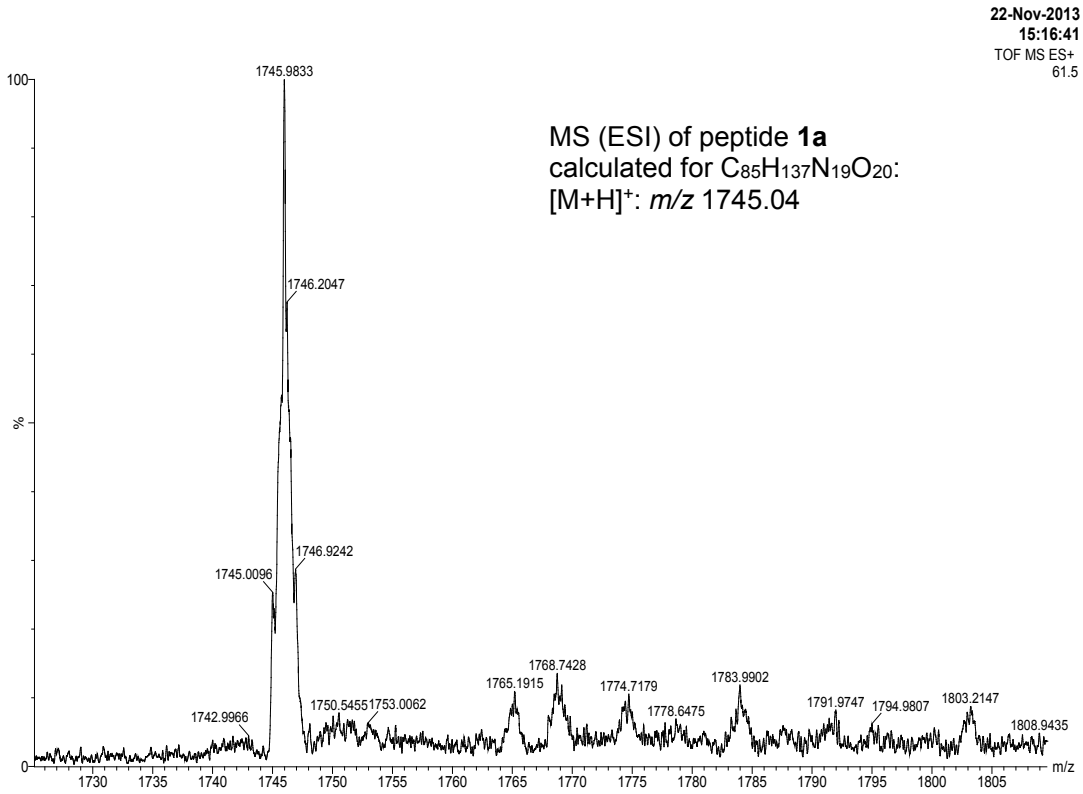


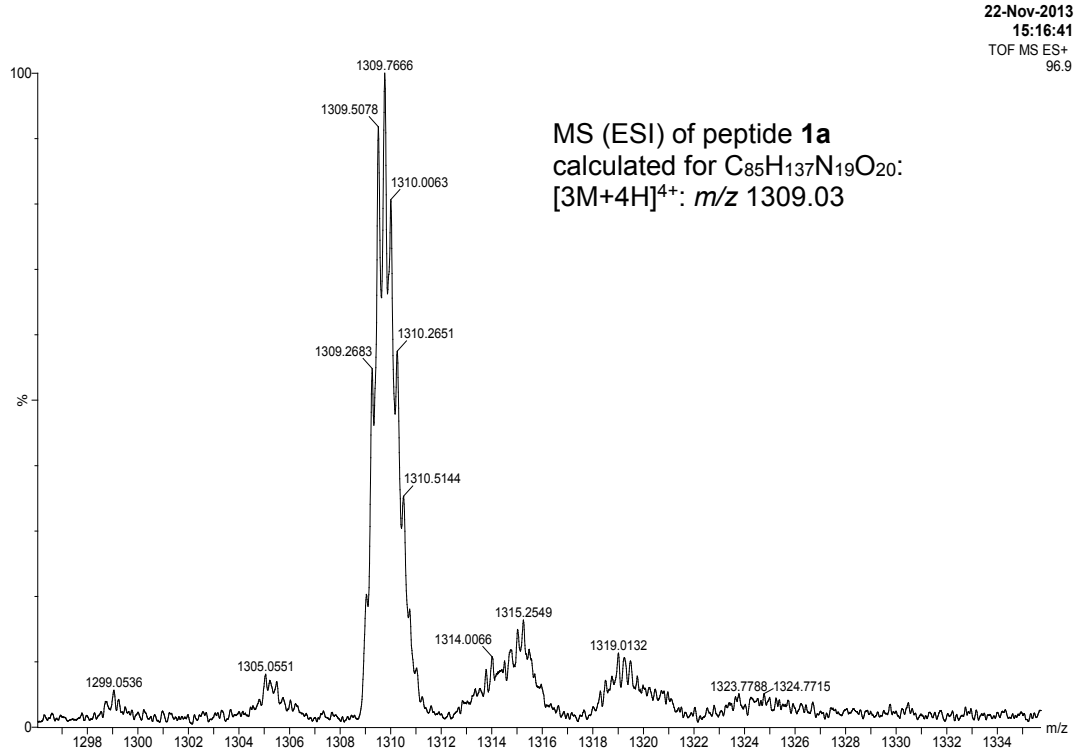
Signal 1:VWD1 A, Wavelength=214 nm

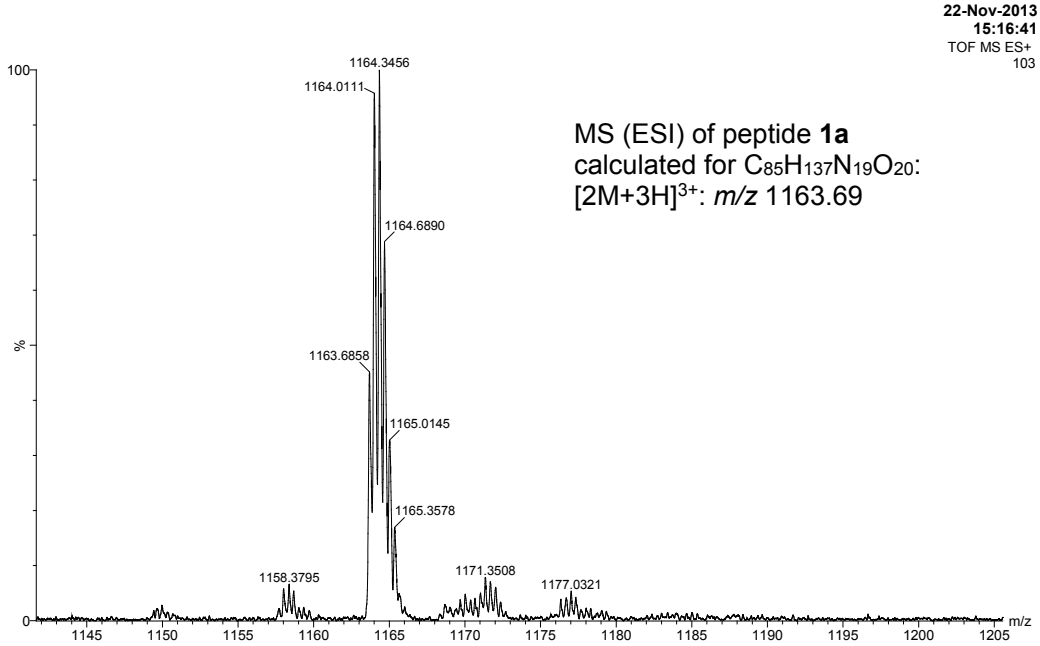
Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	8.831	BV	0.089	15767.279	97.834	95.734
2	9.150	VV	0.139	268.146	0.966	1.628
3	9.394	VB	0.165	434.398	1.201	2.638



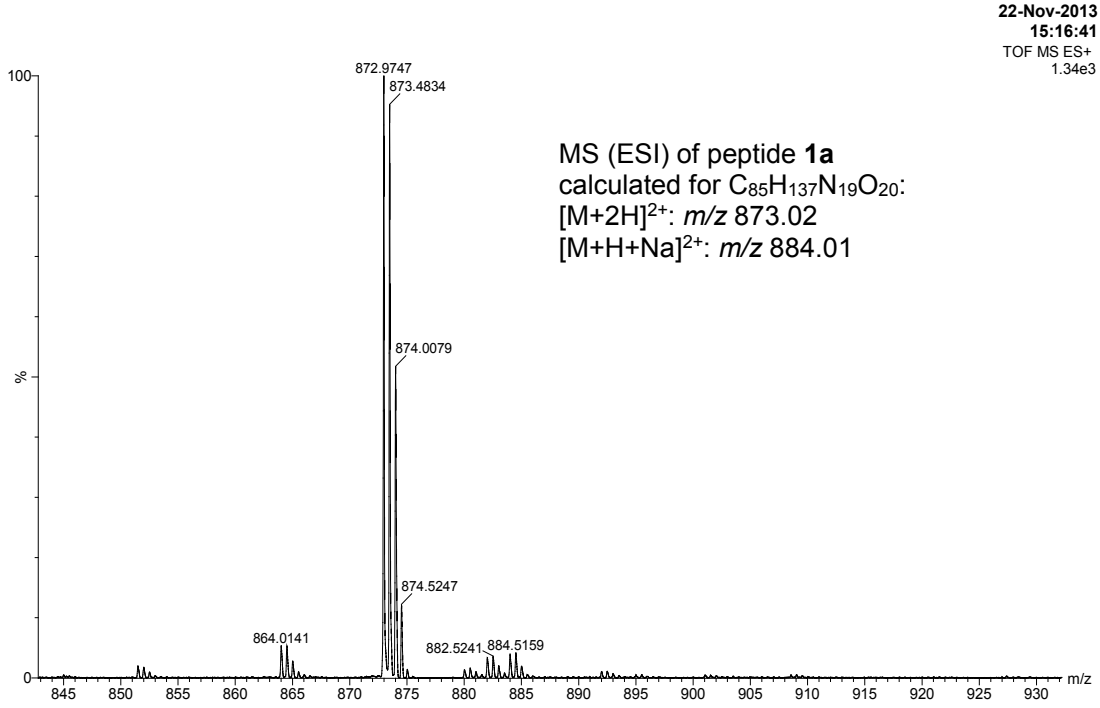
calculated for $C_{85}H_{137}N_{19}O_{20}$:
$[M+2H]^{2+}$: *m/z* 873.02
$[M+H+Na]^{2+}$: *m/z* 884.01
$[M+3H]^{3+}$: *m/z* 582.35
$[2M+3H]^{3+}$: *m/z* 1163.69


22-Nov-2013
15:16:41
TOF MS ES+
61.5
MS (ESI) of peptide **1a**
calculated for $C_{85}H_{137}N_{19}O_{20}$:
$[M+H]^+$: *m/z* 1745.04
1745.9833
1746.2047
1746.9242
1745.0096
1742.9966
1750.5455
1753.0062
1765.1915
1768.7428
1774.7179
1778.6475
1783.9902
1791.9747
1794.9807
1803.2147
1808.9435
100
%
0
m/z
1730
1735
1740
1745
1750
1755
1760
1765
1770
1775
1780
1785
1790
1795
1800
1805

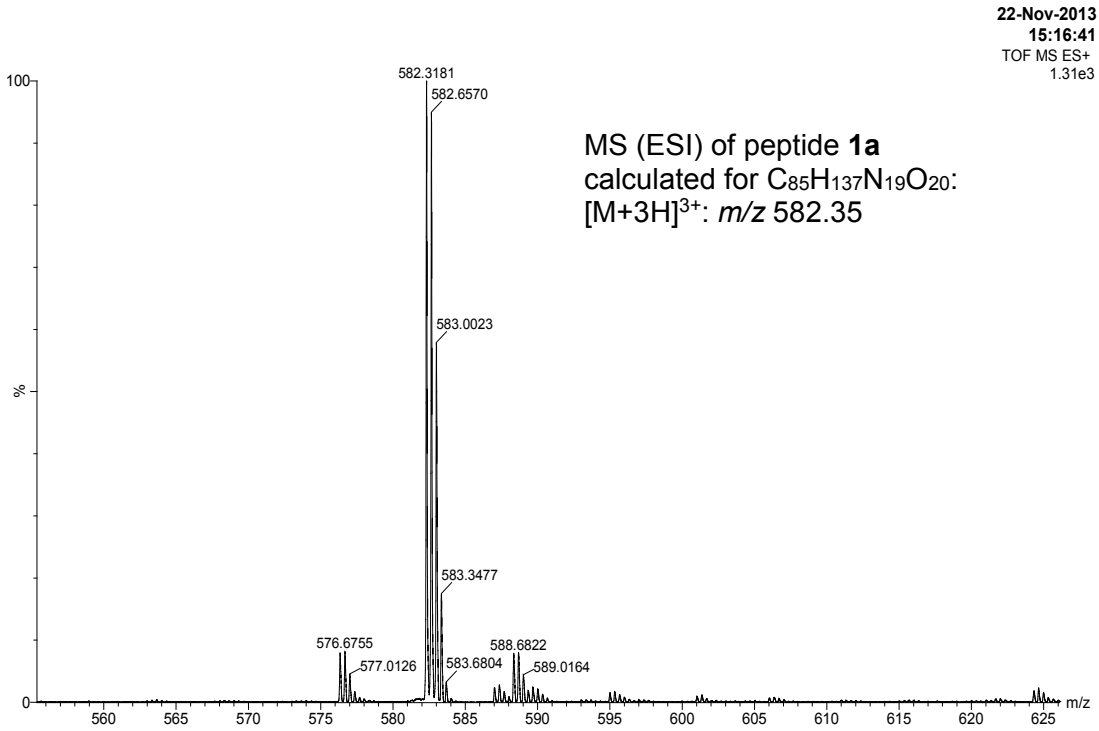

22-Nov-2013
15:16:41
TOF MS ES+
96.9
MS (ESI) of peptide **1a**
calculated for $C_{85}H_{137}N_{19}O_{20}$:
$[3M+4H]^{4+}$: *m/z* 1309.03
1309.7666
1309.5078
1310.0063
1310.2651
1309.2683
1310.5144
1315.2549
1314.0066
1319.0132
1305.0551
1299.0536
1323.7788
1324.7715
100
%
0
m/z
1298
1300
1302
1304
1306
1308
1310
1312
1314
1316
1318
1320
1322
1324
1326
1328
1330
1332
1334



MS (ESI) of peptide **1a**
calculated for $C_{85}H_{137}N_{19}O_{20}$:
$[2M+3H]^{3+}$: *m/z* 1163.69

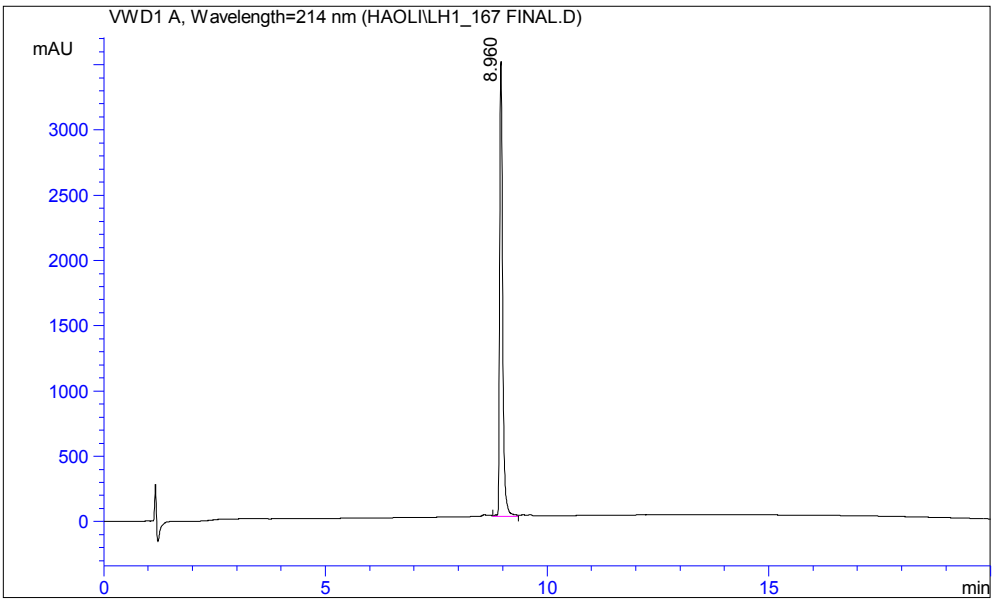


MS (ESI) of peptide **1a**
calculated for $C_{85}H_{137}N_{19}O_{20}$:
$[M+2H]^{2+}$: *m/z* 873.02
$[M+H+Na]^{2+}$: *m/z* 884.01



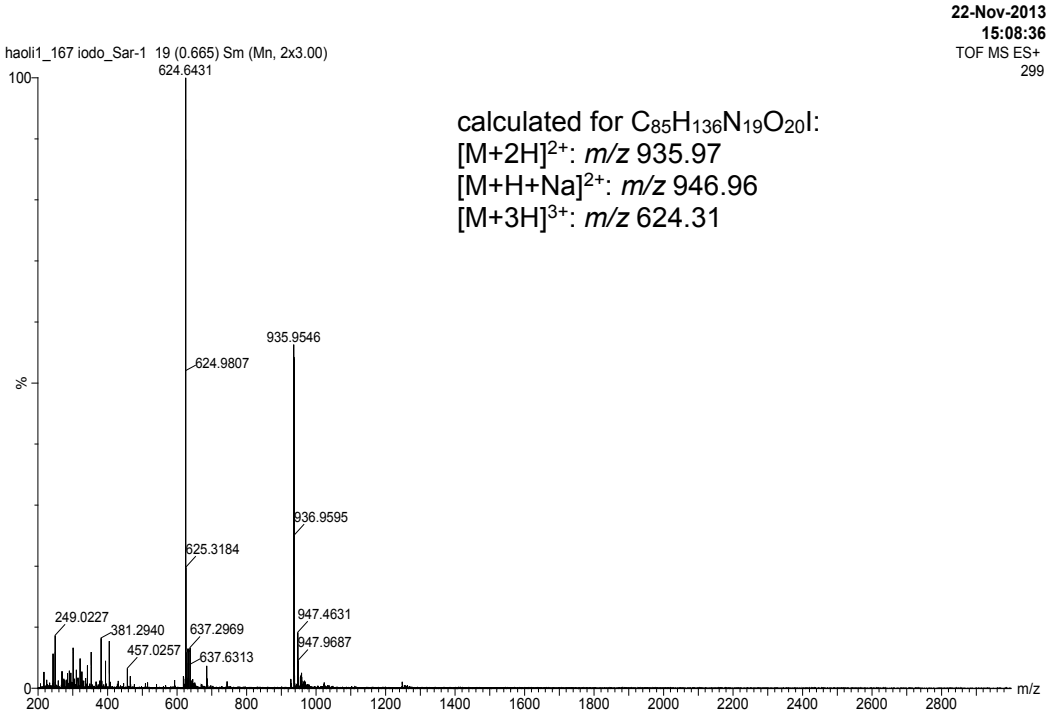
MS (ESI) of peptide **1a** calculated for $C_{85}H_{137}N_{19}O_{20}$: $[M+3H]^{3+}$: *m/z* 582.35

HPLC and MS ESI+ TOF of peptide **1b**

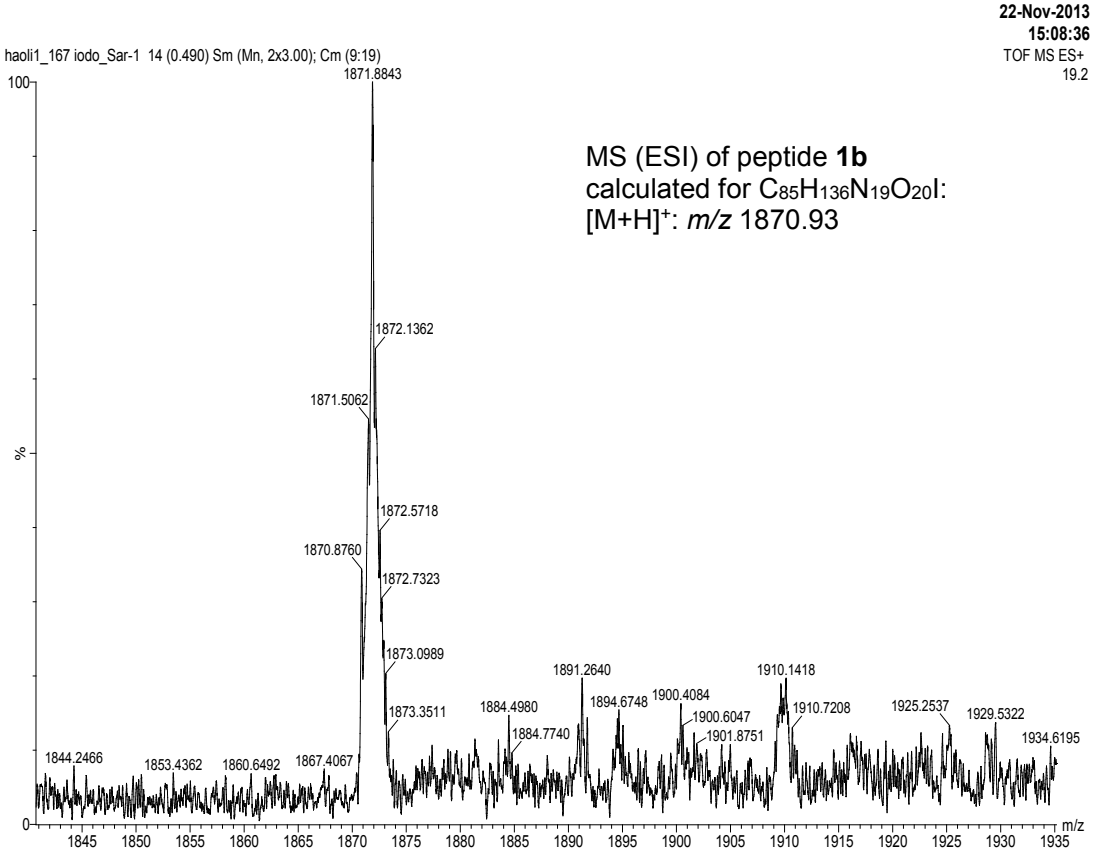


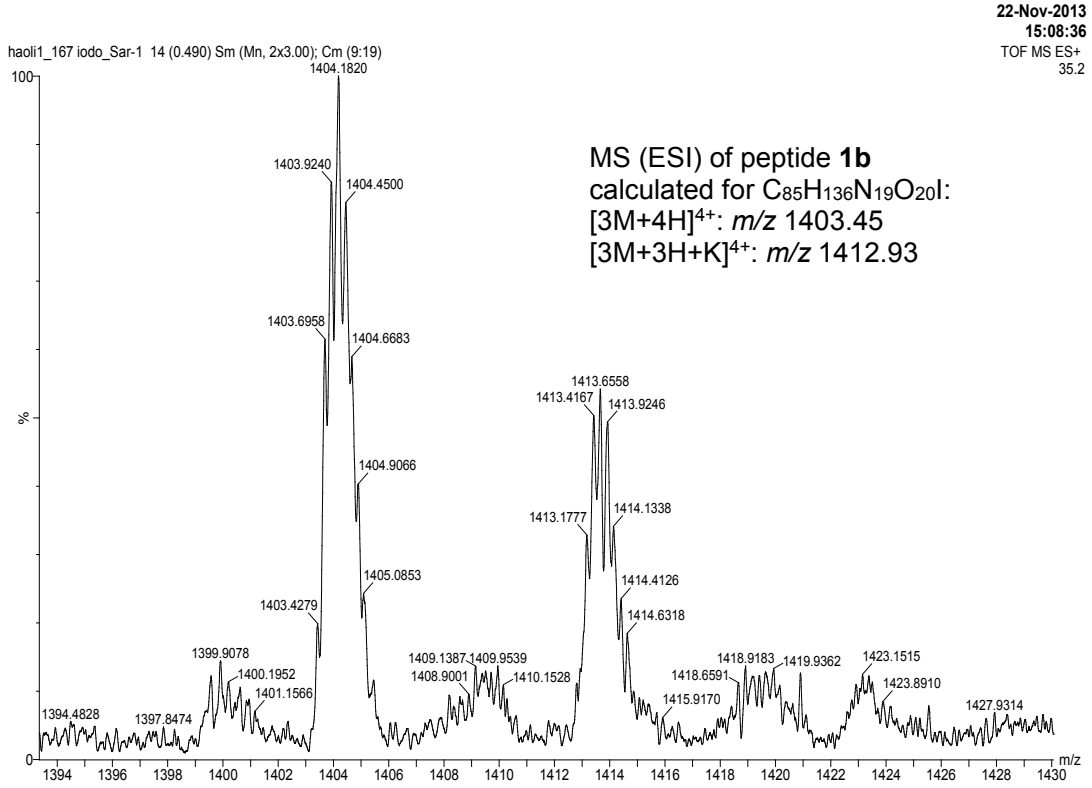
Signal 1:VWD1 A, Wavelength=214 nm

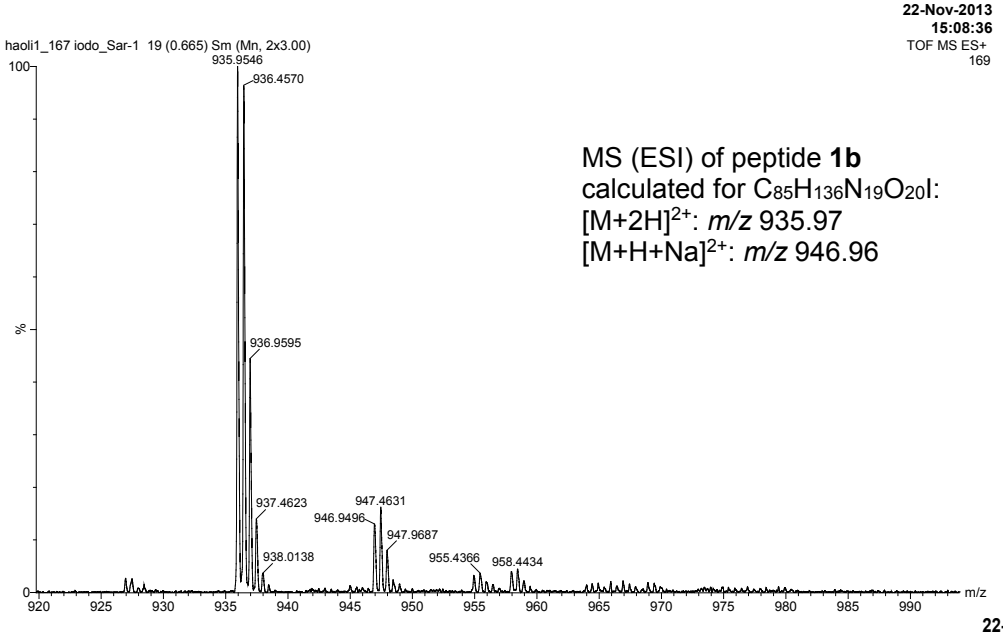
Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	8.960	VV	0.072	16299.187	100.000	100.000



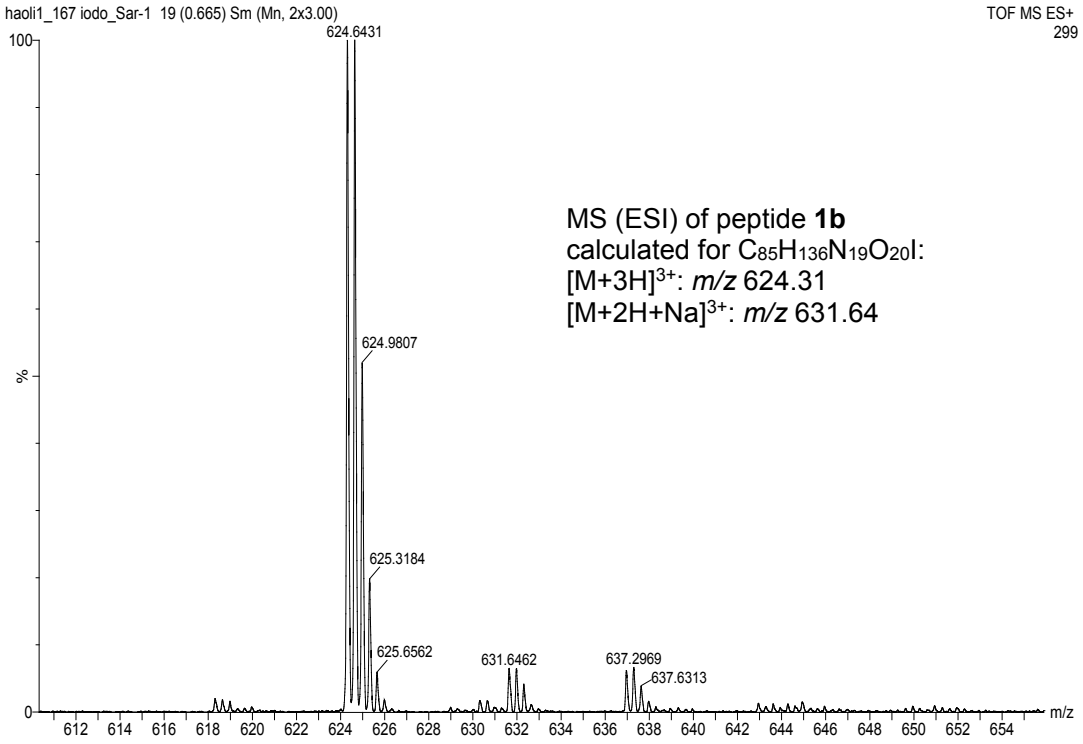
calculated for $C_{85}H_{136}N_{19}O_{20}I$:
$[M+2H]^{2+}$: *m/z* 935.97
$[M+H+Na]^{2+}$: *m/z* 946.96
$[M+3H]^{3+}$: *m/z* 624.31


22-Nov-2013
15:08:36
TOF MS ES+
19.2
haoli1_167 iodo_Sar-1 14 (0.490) Sm (Mn, 2x3.00); Cm (9:19)
MS (ESI) of peptide 1b
calculated for C85H136N19O20I:
[M+H]+: m/z 1870.93
1871.8843
1872.1362
1871.5062
1872.5718
1870.8760
1872.7323
1873.0989
1873.3511
1891.2640
1910.1418
1884.4980
1894.6748
1900.4084
1910.7208
1925.2537
1929.5322
1900.6047
1884.7740
1901.8751
1934.6195
1844.2466
1853.4362
1860.6492
1867.4067
%
m/z


22-Nov-2013
15:08:36
TOF MS ES+
35.2
haoli1_167 iodo_Sar-1 14 (0.490) Sm (Mn, 2x3.00); Cm (9:19)
MS (ESI) of peptide 1b
calculated for C85H136N19O20I:
[3M+4H]4+: m/z 1403.45
[3M+3H+K]4+: m/z 1412.93
1404.1820
1403.9240
1404.4500
1403.6958
1404.6683
1413.6558
1413.4167
1413.9246
1404.9066
1413.1777
1414.1338
1405.0853
1414.4126
1403.4279
1414.6318
1399.9078
1409.1387
1409.9539
1418.9183
1419.9362
1423.1515
1400.1952
1408.9001
1410.1528
1418.6591
1423.8910
1401.1566
1415.9170
1427.9314
1394.4828
1397.8474
%
m/z

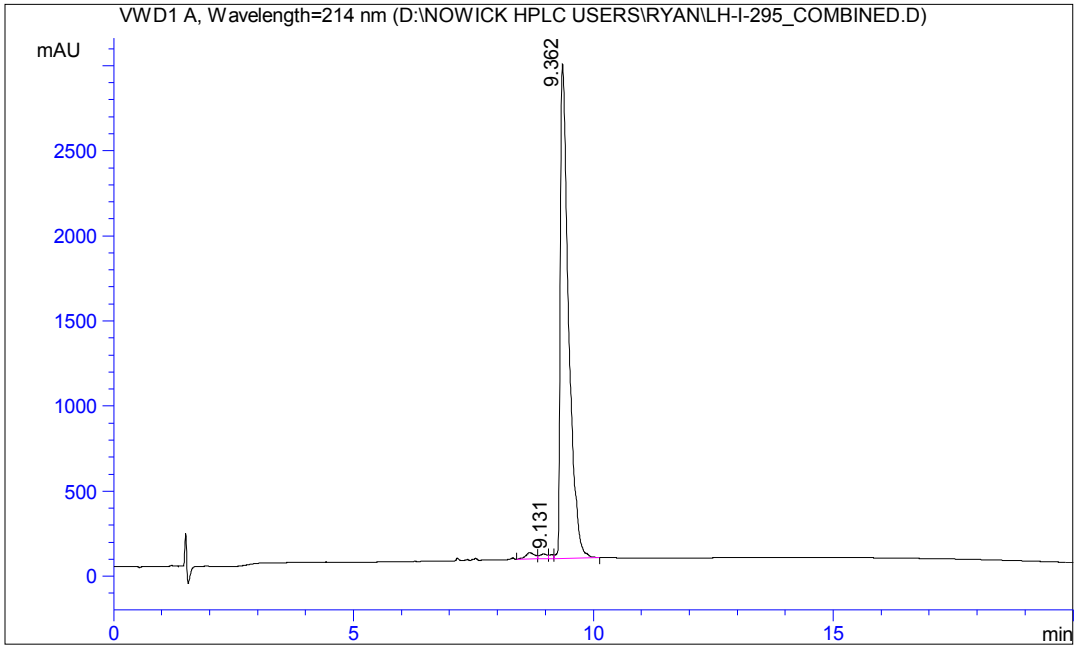


MS (ESI) of peptide **1b**
calculated for $C_{85}H_{136}N_{19}O_{20}I$:
$[M+2H]^{2+}$: *m/z* 935.97
$[M+H+Na]^{2+}$: *m/z* 946.96



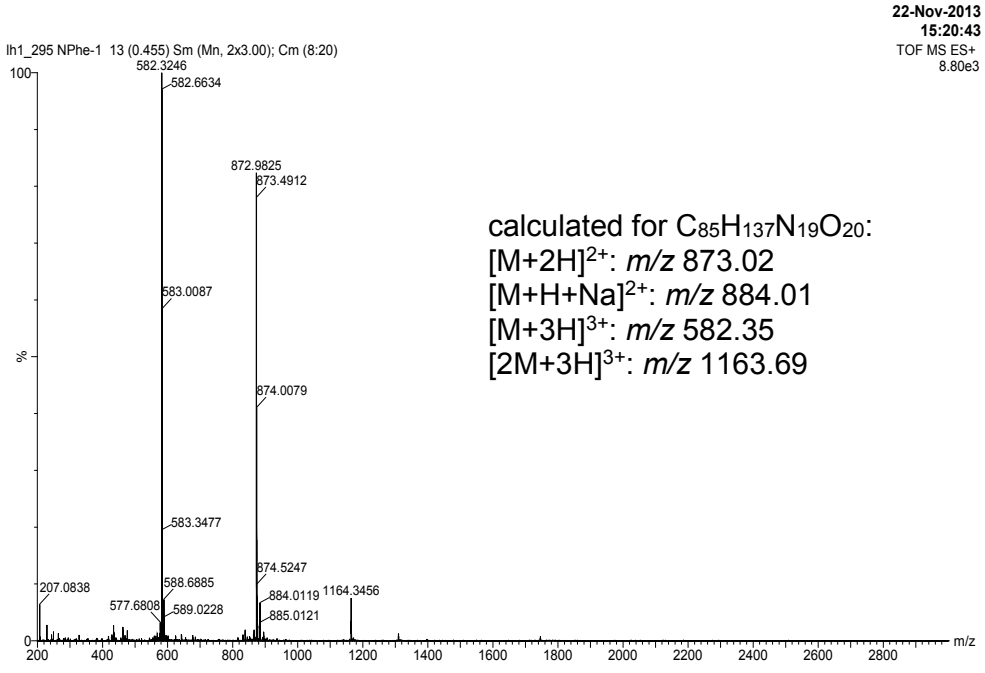
MS (ESI) of peptide **1b**
calculated for $C_{85}H_{136}N_{19}O_{20}I$:
$[M+3H]^{3+}$: *m/z* 624.31
$[M+2H+Na]^{3+}$: *m/z* 631.64

HPLC and MS ESI+ TOF of peptide **2a**

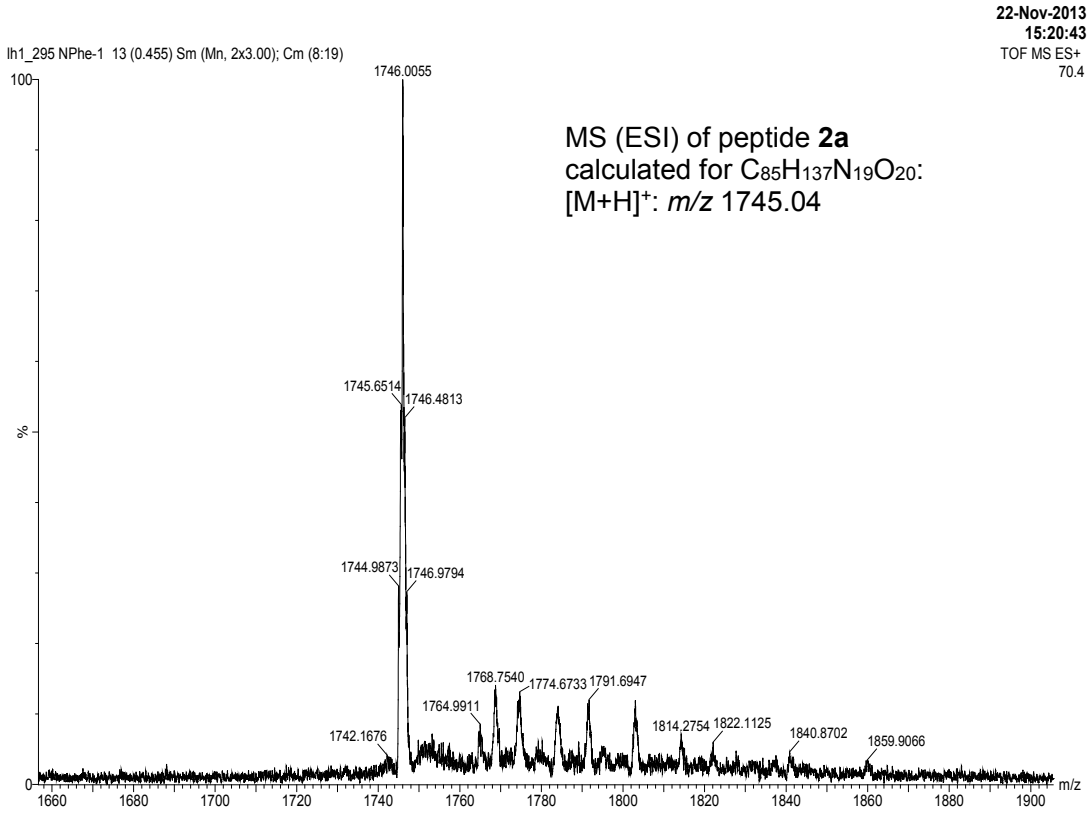


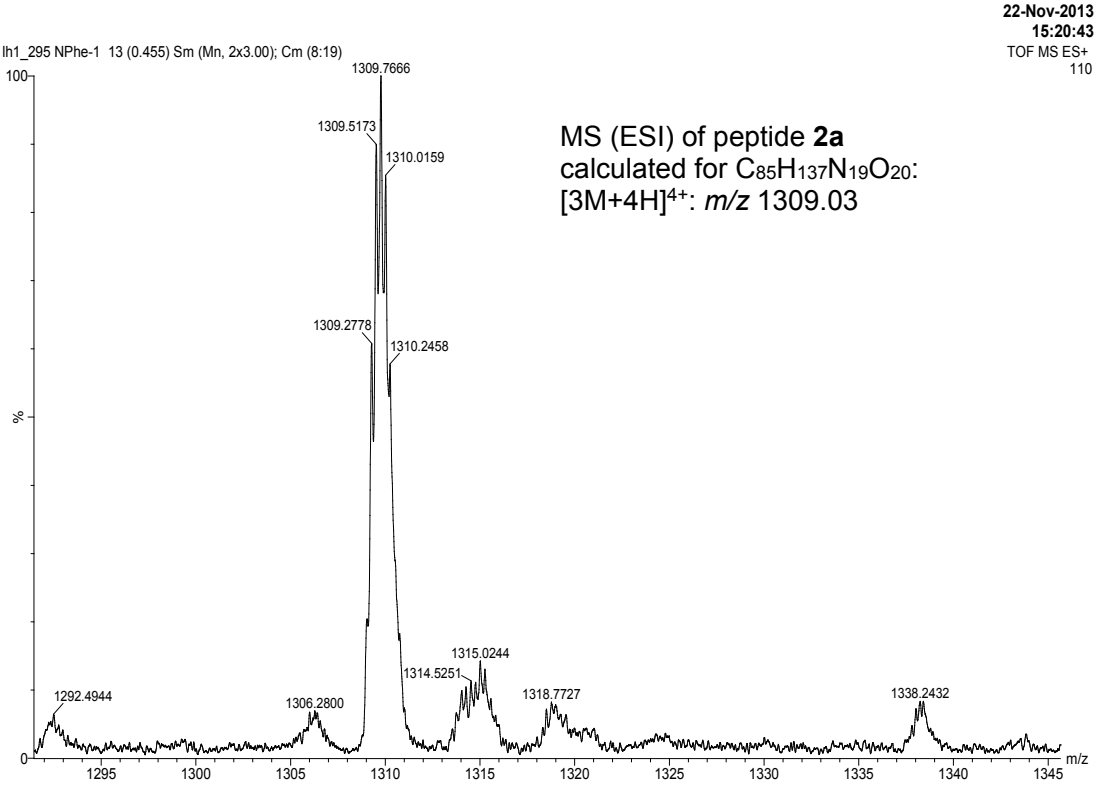
Signal 1:VWD1 A, Wavelength=214 nm

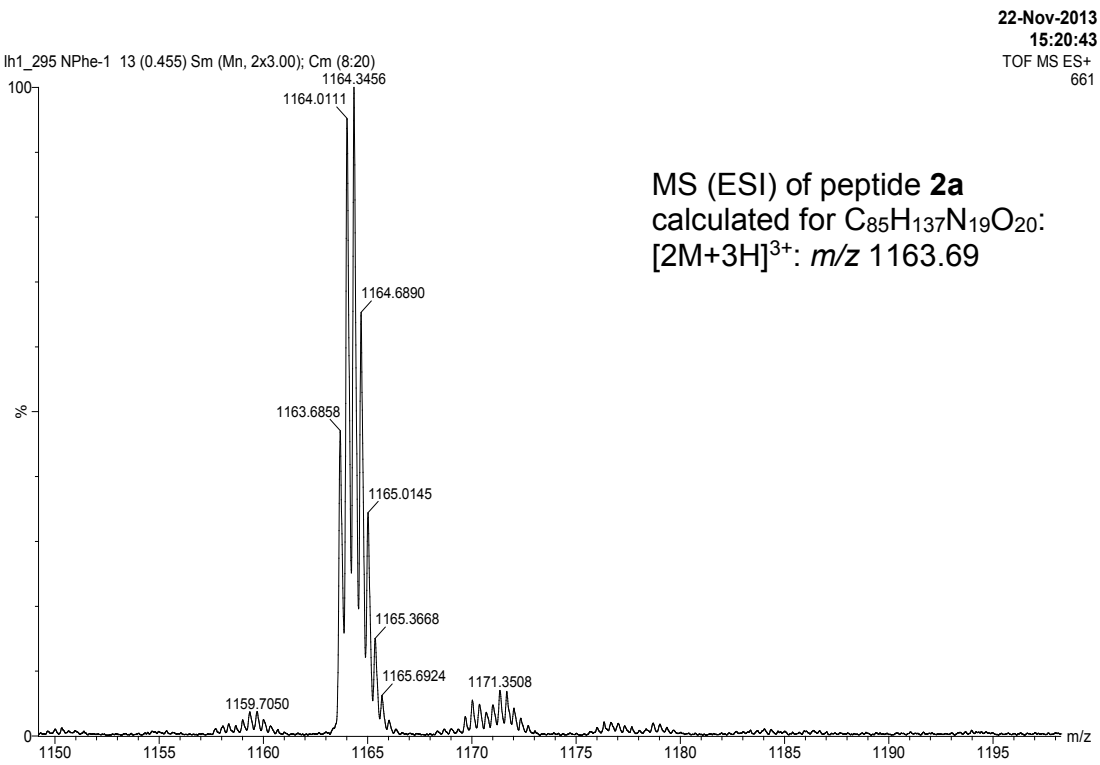
Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	8.673	BV	0.189	516.040	1.280	1.440
2	8.966	VV	0.148	310.217	0.965	0.865
3	9.131	VV	0.091	146.465	0.773	0.409
4	9.362	VB	0.178	34871.512	96.982	97.286



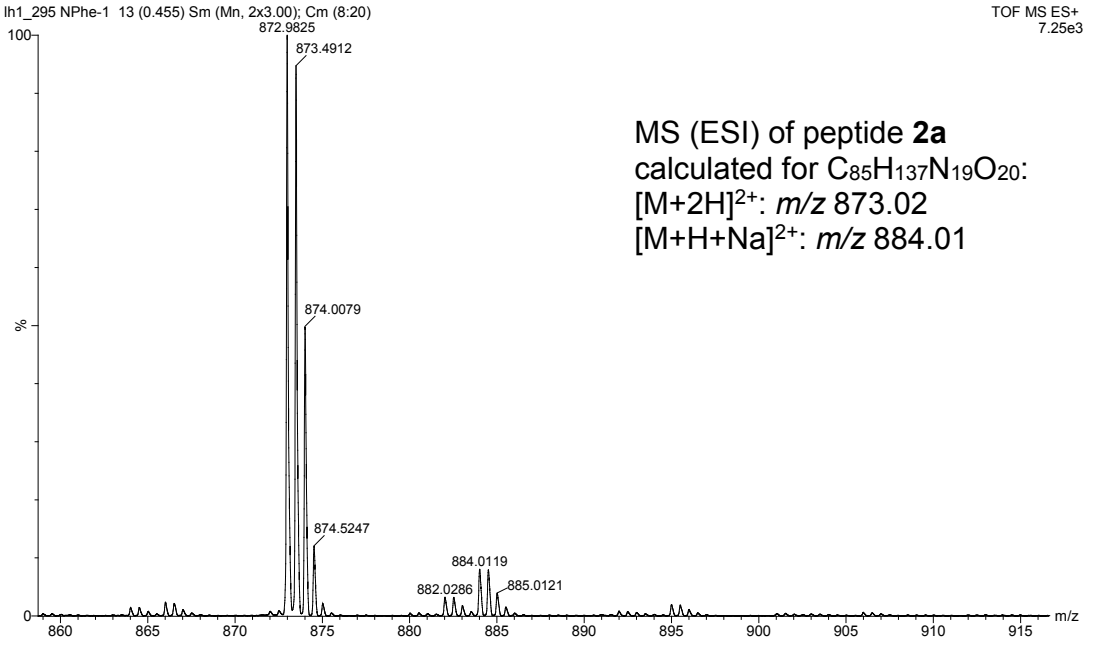
calculated for $C_{85}H_{137}N_{19}O_{20}$:
$[M+2H]^{2+}$: *m/z* 873.02
$[M+H+Na]^{2+}$: *m/z* 884.01
$[M+3H]^{3+}$: *m/z* 582.35
$[2M+3H]^{3+}$: *m/z* 1163.69


22-Nov-2013
15:20:43
TOF MS ES+
70.4
lh1_295 NPhe-1 13 (0.455) Sm (Mn, 2x3.00); Cm (8:19)
MS (ESI) of peptide 2a
calculated for C85H137N19O20:
[M+H]+: m/z 1745.04
1746.0055
1745.6514
1746.4813
1744.9873
1746.9794
1768.7540
1774.6733
1791.6947
1764.9911
1742.1676
1814.2754
1822.1125
1840.8702
1859.9066
%
m/z

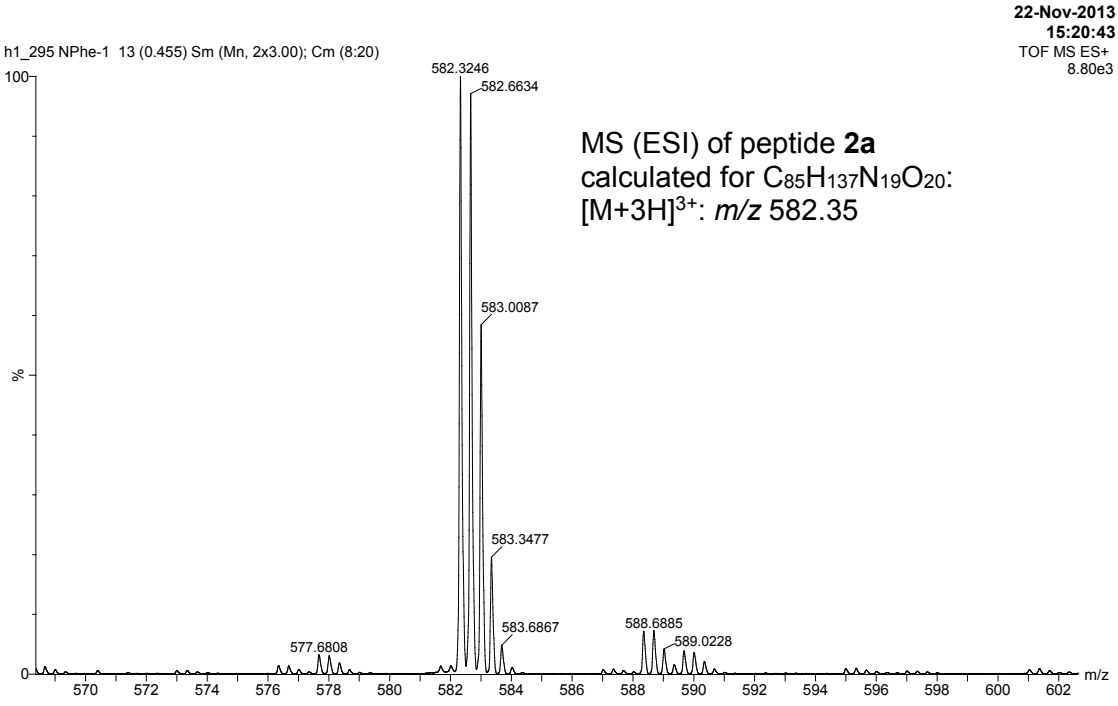

22-Nov-2013
15:20:43
TOF MS ES+
110
lh1_295 NPhe-1 13 (0.455) Sm (Mn, 2x3.00); Cm (8:19)
MS (ESI) of peptide 2a
calculated for C85H137N19O20:
[3M+4H]4+: m/z 1309.03
1309.7666
1309.5173
1310.0159
1309.2778
1310.2458
1315.0244
1314.5251
1292.4944
1306.2800
1318.7727
1338.2432
%
m/z



MS (ESI) of peptide **2a**
calculated for $C_{85}H_{137}N_{19}O_{20}$:
$[2M+3H]^{3+}$: *m/z* 1163.69

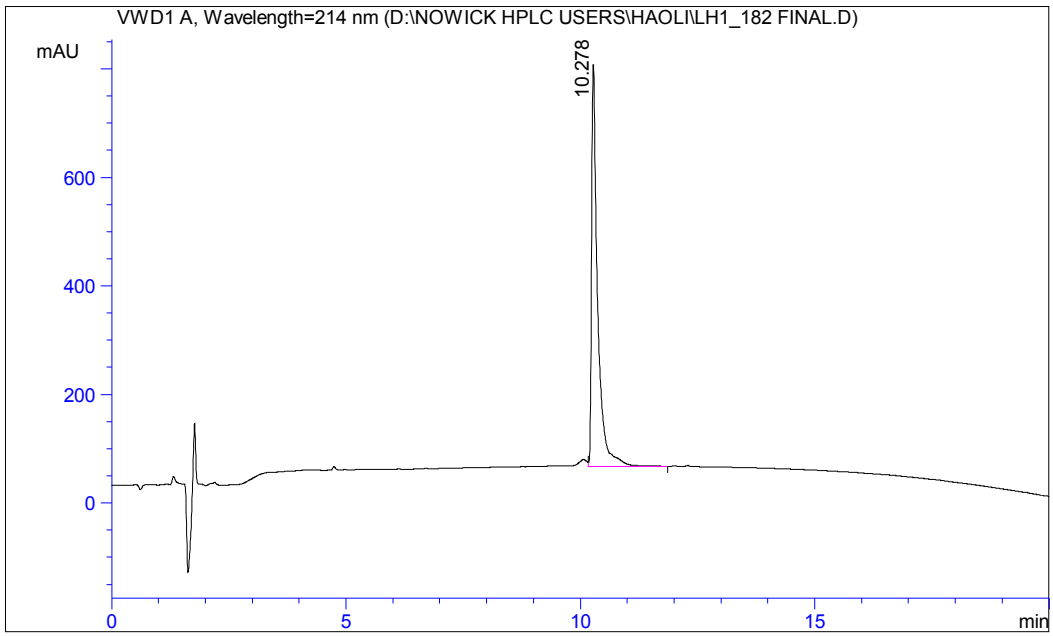


MS (ESI) of peptide **2a**
calculated for $C_{85}H_{137}N_{19}O_{20}$:
$[M+2H]^{2+}$: *m/z* 873.02
$[M+H+Na]^{2+}$: *m/z* 884.01



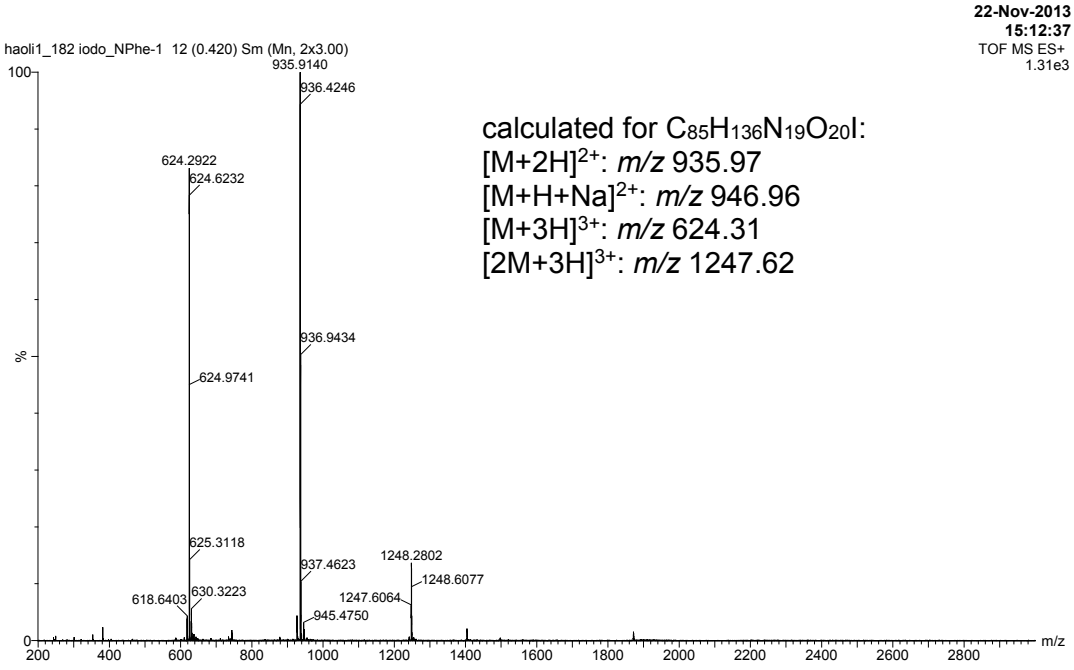
MS (ESI) of peptide **2a** calculated for $C_{85}H_{137}N_{19}O_{20}$: $[M+3H]^{3+}$: *m/z* 582.35

HPLC and MS ESI+ TOF of peptide **2b**

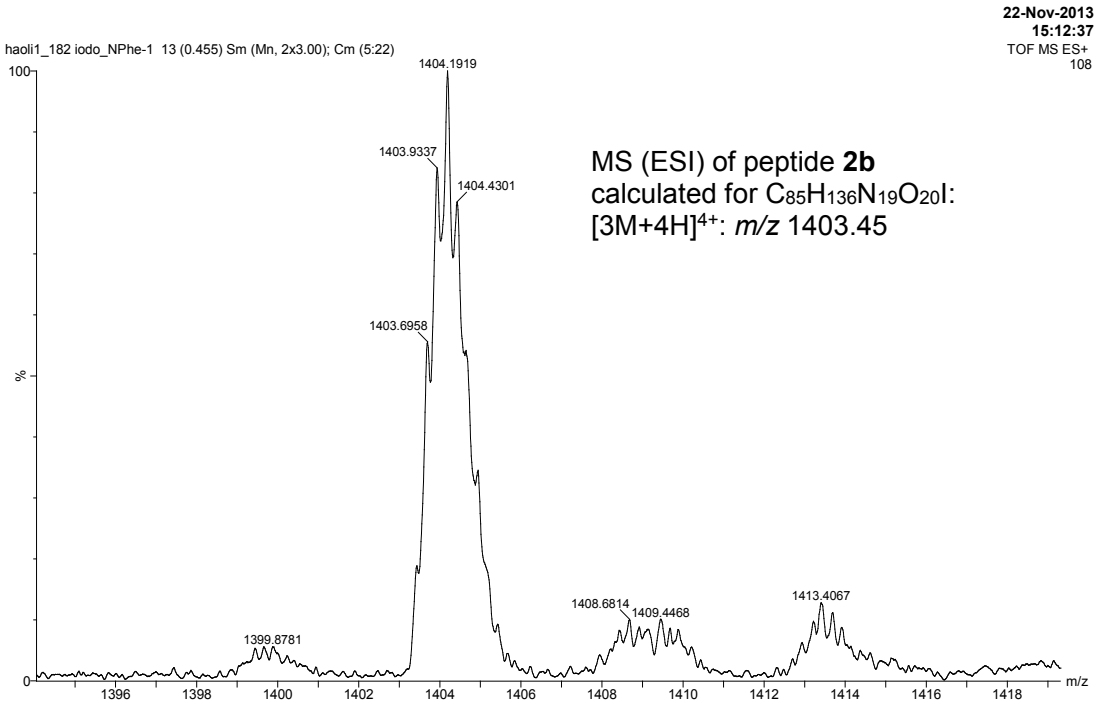


Signal 1:VWD1 A, Wavelength=214 nm

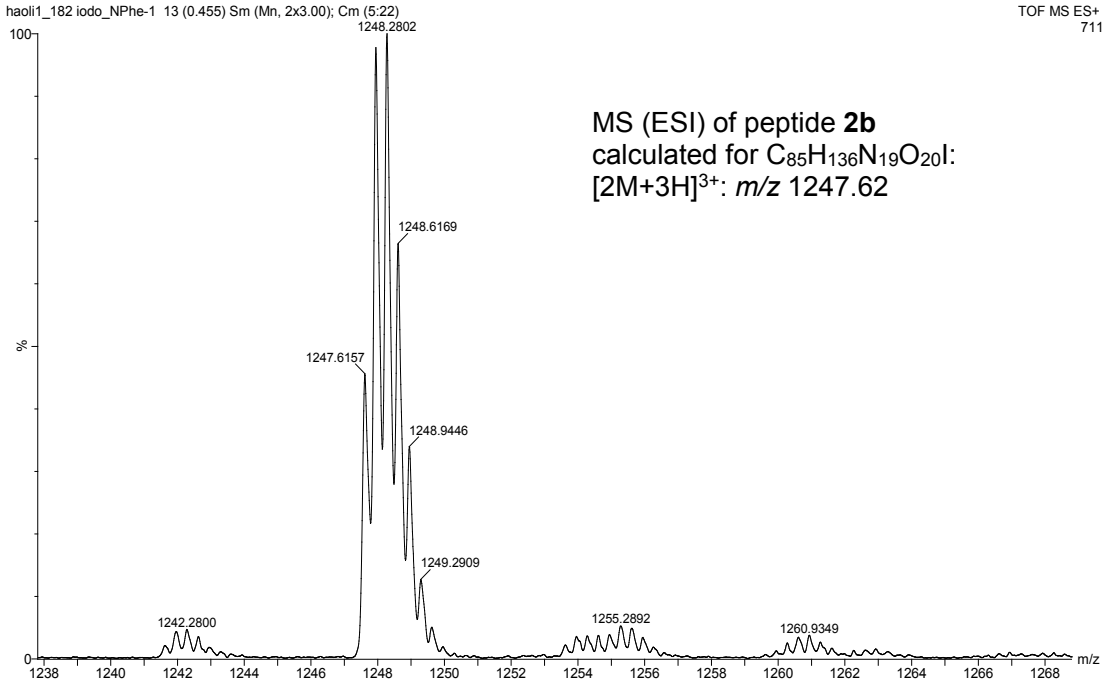
Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	10.278	VV	0.128	6660.014	100.000	100.000



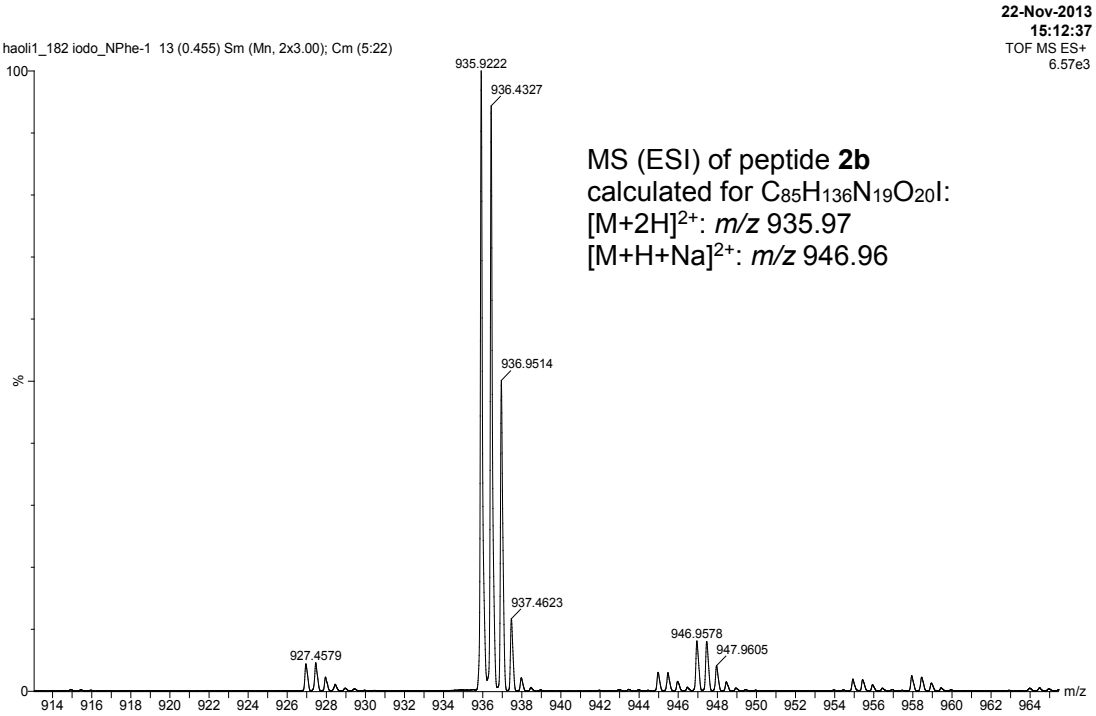
calculated for $C_{85}H_{136}N_{19}O_{20}I$:
$[M+2H]^{2+}$: *m/z* 935.97
$[M+H+Na]^{2+}$: *m/z* 946.96
$[M+3H]^{3+}$: *m/z* 624.31
$[2M+3H]^{3+}$: *m/z* 1247.62

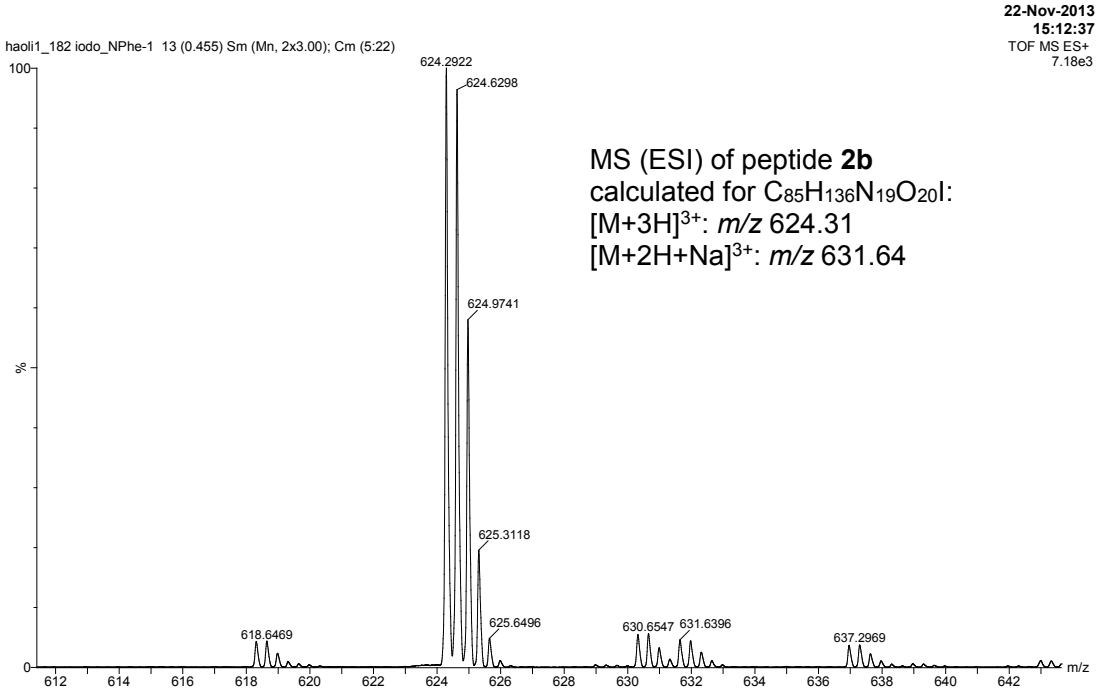


MS (ESI) of peptide **2b** calculated for $C_{85}H_{136}N_{19}O_{20}I$: $[3M+4H]^{4+}$: *m/z* 1403.45



MS (ESI) of peptide **2b** calculated for $C_{85}H_{136}N_{19}O_{20}I$: $[2M+3H]^{3+}$: *m/z* 1247.62


22-Nov-2013
15:12:37
TOF MS ES+
6.57e3
haoli1_182 iodo_NPhe-1 13 (0.455) Sm (Mn, 2x3.00); Cm (5:22)
935.9222
936.4327
936.9514
937.4623
927.4579
946.9578
947.9605
MS (ESI) of peptide **2b**
calculated for $C_{85}H_{136}N_{19}O_{20}I$:
$[M+2H]^{2+}$: *m/z* 935.97
$[M+H+Na]^{2+}$: *m/z* 946.96
%
m/z


22-Nov-2013
15:12:37
TOF MS ES+
7.18e3
haoli1_182 iodo_NPhe-1 13 (0.455) Sm (Mn, 2x3.00); Cm (5:22)
624.2922
624.6298
624.9741
625.3118
625.6496
618.6469
630.6547
631.6396
637.2969
MS (ESI) of peptide **2b**
calculated for $C_{85}H_{136}N_{19}O_{20}I$:
$[M+3H]^{3+}$: *m/z* 624.31
$[M+2H+Na]^{3+}$: *m/z* 631.64
%
m/z

Chapter 4

A Newcomer's Guide to Peptide Crystallography.[a]

Introduction

X-ray crystallography is a powerful tool for studying the structure and supramolecular assembly of small molecules and biological macromolecules.[1] Crystallography is arguably the most widely used technique for studying the structures of proteins and nucleic acids at atomic resolution. Over 90,000 biomolecular crystal structures have been deposited in the Protein Data Bank (PDB) since its inception four decades ago, with the number of deposits increasing dramatically over the past few years. While X-ray crystallography has been a boon for the structural biology of proteins, it has been underutilized for peptides. Peptide crystallography offers the promise of understanding the structures of peptides and their supramolecular assemblies.[2]

[a]Reproduced with permission from *Isr. J. Chem.*, in press. Unpublished work copyright 2015 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Advances in the tools of protein crystallography have made peptide crystallography easy for the newcomer. Increases in computer processing power, advances in charge-coupled device (CCD) area detectors, and access to synchrotron X-ray radiation sources now make it possible to rapidly determine crystal structures. Commercially available crystal screening kits and high-throughput screening robotics have greatly reduced the time required to discover appropriate crystal growing conditions while reducing the quantities of material needed for growing crystals. Precise diffraction data collection and powerful crystallographic software have allowed the development of multiple phasing techniques, such as anomalous diffraction and molecular replacement, to rapidly determine crystallographic structures within days. For protein crystallography, the computer software is so advanced that determining a crystal structure after collecting diffraction data has become almost entirely automated. Many of the techniques for diffraction data collection and computer programs developed to determine protein structures can now be applied to determining peptide crystal structures.

Over the past few years, our laboratory has begun using X-ray crystallography as a routine technique to study the structure and supramolecular assembly of macrocylic β-sheet peptides. Although we started these studies in collaboration with research groups that routinely do X-ray crystallography, I have adopted the techniques used by these laboratories to determine the crystallographic structures of over a dozen peptides within our laboratory. In this Chapter, I would like to share what I have learned thus far, to allow other researchers to use peptide crystallography for their own problems in determining peptide structures and supramolecular assembly.

Peptide Crystallography

The techniques required for peptide crystallography mirror the techniques developed for protein crystallography. Peptide crystallography involves three main stages: crystallization;

data collection and analysis; and generating an electron density map and crystallographic structure. Each of these stages involves a series of steps, starting with the purified peptide and concluding with the deposition of a crystallographic structure into the PDB (Figure 4.1).

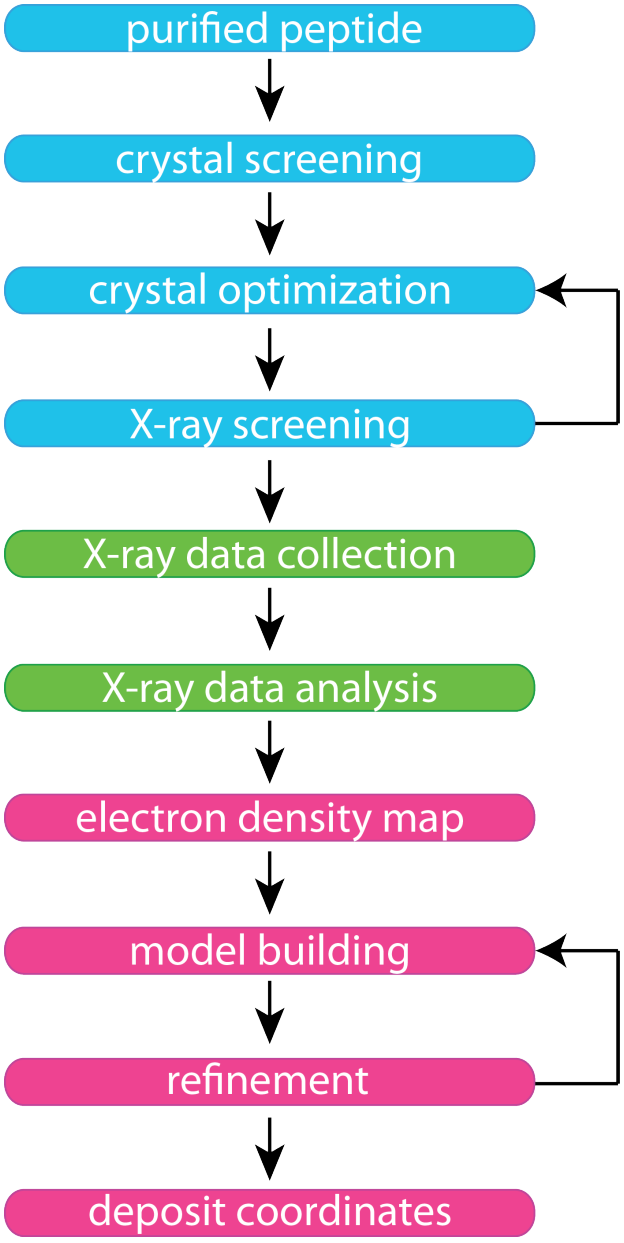


Figure 4.1: Overview of the steps involved in peptide crystallography.

The Peptide

Peptides used for crystallography experiments must be pure and soluble, and most likely will need to contain a heavy atom. Impurities, such as peptide fragments generated during peptide synthesis, totalling more than a few percent, may inhibit or prevent crystal growth. The peptide must be water soluble for the techniques described here, because the peptide is screened in various aqueous solutions containing buffers, salts, and cryogenic protectants.

The incorporation of a heavy atom into the peptide is often necessary for determining the crystallographic structure. A heavy atom is required for anomalous diffraction techniques to determine the phases required for calculating the electron density map and thus the crystal structure. Heavy atoms, such as I, Br, Se, Fe, Co, and Zn, are useful for anomalous diffraction techniques.[3] The heavy atom that is selected will dictate the experimental parameters required for the anomalous diffraction experiments. We initially used *p*-bromophenylalanine to incorporate a Br atom within our β-sheet macrocycles;[4–7] recently I have switched to *p*-iodophenylalanine to quickly solve our peptide structures using Cu radiation on an in-house X-ray diffractometer.[8] Although both bromine and iodine are suitable for anomalous diffraction data collection on a synchrotron, only the iodine permits the collection of anomalous diffraction data on an X-ray diffractometer with a Cu anode.[9–11] The choice of an appropriate heavy atom will become more apparent during the discussion of the generation of an electron density map.

Crystallization of the Peptide

The most difficult step in peptide crystallography is the formation of crystals that diffract to a high-resolution. Peptide crystals can be grown much like protein crystals, in an aqueous solution with various buffers, pH ranges, salts, and cryogenic protectants. Finding the correct combination of buffer type, pH, salt concentration and cryogenic concentrations to facilitate

crystallization can be difficult. Like proteins, the growth of peptide crystals is sensitive to changes in buffer salts, pH, additional salts, temperature, and cryogenic protectants. It is also necessary to determine the concentration of the peptide for crystal screening.

I determine the peptide concentration for crystal screening using the Hampton pre-crystallization test (PCT). I typically perform the PCT according to the instructions in the Hampton PCT kit, in a hanging-drop format, at three different concentrations (5, 10, and 20 mg/mL), using the two different crystallization solutions in the kit. After 30 minutes the drops are examined to determine the most suitable peptide concentration to perform the crystal screens. Typically, I find a peptide concentration of 10 mg/mL suitable to perform crystal screening.

Crystal screening is now routine, and many kits can be purchased commercially to do crystal screening. Crystal screening is typically performed as a hanging-drop vapour-diffusion or sitting-drop vapor-diffusion format. In hanging-drop vapor-diffusion experiments a droplet of the peptide solution is mixed with a droplet of the crystal growing solution on a plastic or glass slide. The slide is inverted and sealed over a well containing the crystal growing solution. Sitting-drop vapor-diffusion experiments are performed in a similar fashion, but the droplets are mixed in a small reservoir well next to the crystal growing solution. Screens are performed in either a 96-well or 24-well format. I typically use three different crystallization kits (Peg-Ion, Crystal Screen, and Index) in a 96-well hanging-drop format to examine 288 crystal growing conditions concurrently.

I examine each experiment in the screens under a microscope to look for the formation of crystals. Typically peptide crystals grow rapidly, within 24-72 hours, but crystals may also take longer to grow. Screens should be examined after a few hours, then a day, and then a few days later. Crystals formed during crystal screening are not typically good enough to collect good diffraction data and further optimization of the crystal growing conditions is often required. Optimization of crystal conditions is invaluable and can take a substantial amount of time.

Optimization of crystal growing conditions is performed in a 24-well format in a 4x6 matrix. Typically, two conditions (e.g., buffer pH and cryogenic protectant concentration) are examined with one condition being varied across the rows and the other condition being varied down the columns. Crystals grown during the optimization screens are examined for crystal quality (clarity, size, shape, etc.) and are also examined using an X-ray diffractometer. Crystals that appear well shaped often diffract well but sometimes diffract poorly. Conversely, crystals that do not appear well-formed sometimes diffract well. It is often best to screen all crystals within any given well to determine which crystals provide the highest quality diffraction data. The conditions that yield the best diffracting crystals are optimized further with smaller variations of the crystal growing conditions between wells. If a third component, such as a salt concentration, is present in the crystal growing solution then variations of the concentration during further optimization screens may help crystal growth. Figure 4.2 shows a single droplet of an optimized screen experiment containing multiple peptide crystals.

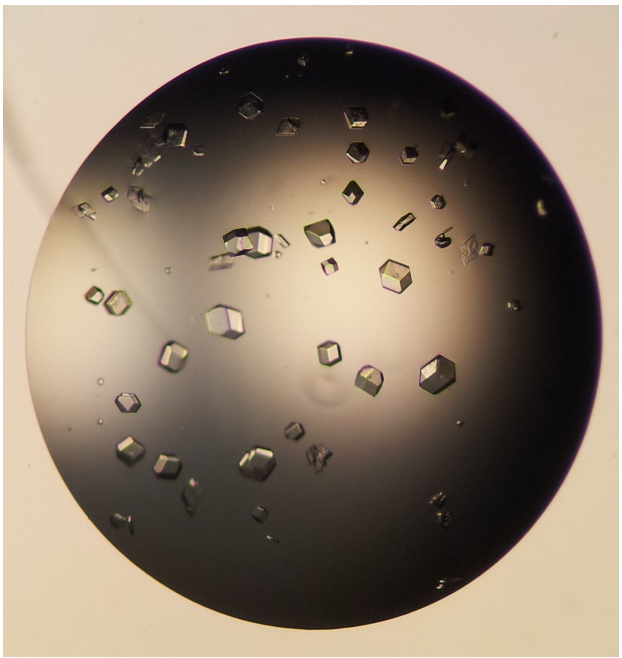


Figure 4.2: A single droplet solution of an optimized screen experiment containing multiple peptide crystals.

Harvesting the Crystals

The second most difficult step in peptide crystallography is harvesting the crystals. Peptide crystals are very fragile, and harvesting can often damage a crystal, which may limit the resolution at which the crystal diffracts. Crystals are typically harvested with a nylon loop attached to a metal pin to scoop the crystal into the center of the loop. It is often best to choose a nylon loop size that is slightly smaller than the crystal. Upon harvesting, the crystal is flash frozen in either a jet of liquid nitrogen vapor or a liquid nitrogen bath. The process of cryogenically freezing the crystal may also damage the crystal. To minimize this damage, crystals are either grown in a cryogenic protectant (glycerol, PEG, etc.) or dipped into a cryogenic protectant before freezing.[12] Cryogenic freezing is important, because it reduces radiation damage to the crystal during the X-ray diffraction experiments.

Assessing Crystal Quality

The quality of a crystal can quickly be assessed by examining images collected with an X-ray diffractometer. A typical X-ray diffractometer contains three important parts: an X-ray source, a goniometer, and a detector. The crystal is placed on a goniometer and centered relative to the incident X-ray beam. The X-ray beam passes through the crystal, is diffracted, and strikes the detector to produce a diffraction image. A single diffraction image constitutes a piece of the X-ray crystallographic data set obtained at a particular angle. In obtaining a complete diffraction data set, the crystal is rotated about a single axis and dozens or hundreds of diffraction images are collected at a series of angles.

A single diffraction image is often enough to determine the quality of the crystal. A single diffraction image is collected by rotating the crystal by a small angle, such as 0.5 or 1.0°, while exposing it to the X-ray beam. The diffraction pattern of well diffracting crystals exhibit ordered, round, well-defined diffraction spots with low mosaicity (below 1°)

and diffract to a high resolution. Diffraction images where diffraction spots are smeared or contain multiple spots around a single diffraction spot are not suitable for data collection. Figure 4.3 illustrates a good diffraction pattern with well-defined spots. Figure 4.4 illustrates a poor diffraction pattern with smeared diffraction spots and contain multiple spots around a single spot.

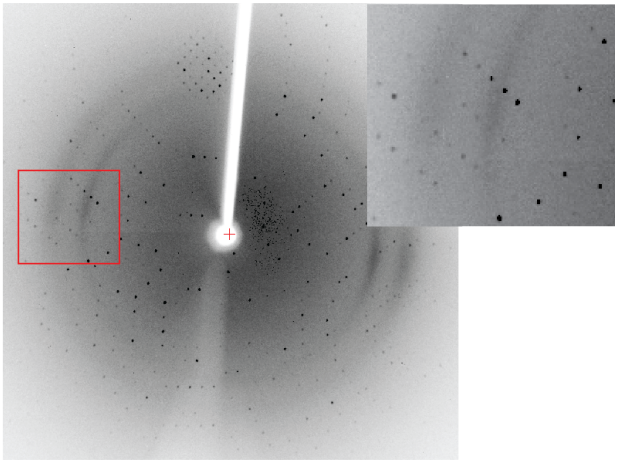


Figure 4.3: A good diffraction image, with well-defined diffraction spots. The white bar running from the top of the image to the center is from the beamstop on the diffractometer.

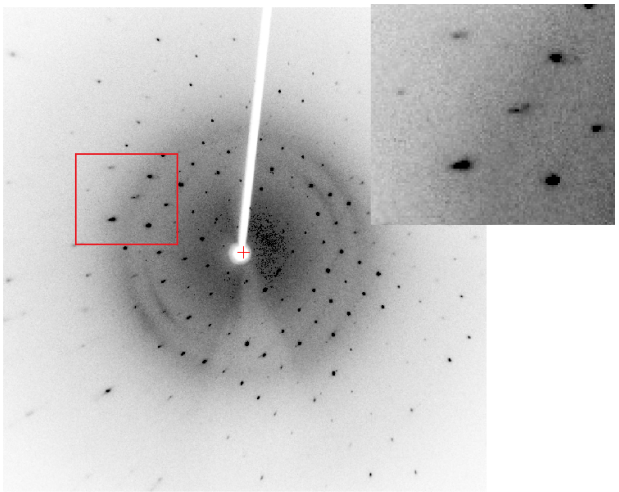


Figure 4.4: A poor diffraction image, with smeared diffraction spots and spots comprising multiple diffraction spots.

I typically collect two diffraction images to determine the crystal cell dimensions, cell angles, and a possible space group. This process is called indexing. This is often done by

analysing two diffraction images that are collected 90 degrees apart from one another. I have observed that well diffracting peptide crystals often index to space groups with high symmetry and poor diffracting crystals do not index beyond $P1$, the primitive space group for all crystals. I collect diffraction data sets on crystals that index to space groups higher than $P1$, diffract to a high resolution, and have well defined spots.

Many programs, including iMosflm,[13] HKL2000,[14] d*TREK,[15] and XDS,[16] can be used to determine the location and intensities of the spots and to calculate a probable space group. The same programs can also be used to create a strategy to collect a complete diffraction data set and to process the data set.

Collecting X-ray Diffraction Data

Three important parameters must be adjusted before collecting a diffraction data set: detector distance, exposure time, and the number of diffraction images or degrees to be collected. The parameters will depend on both the instrument (an X-ray diffractometer or a synchrotron) and the crystal being studied. I will discuss the parameters that need to be optimized for X-ray diffractometers and synchrotrons, and the advantages and disadvantages of each instrument.

The X-ray Diffractometer

X-ray diffractometers are convenient because they allow the immediate screening of crystals and collection of diffraction data in-house. The collection of diffraction data on an X-ray diffractometer, however, can be slow, requiring a few hours to a few days to collect a full data set. Data collection takes a lot of time because most CCD area detectors are small, limiting the amount of high-resolution diffraction data that can be collected per image and

requiring the collection of many images.

X-ray diffractometers are often limited to a single X-ray wavelength. A copper anode is typically used for proteins and produces X-rays at 1.54 Å wavelength. X-ray sources at other wavelengths are available from anodes such as Mo, Cr, or Co, but are generally less suited to peptide and protein crystallography.

Setting the correct distance and X-ray exposure time are key to a successful diffraction experiment on a diffractometer. The detector distance the distance from the crystal to the CCD area detector should be proportional to the longest unit cell dimension determined during indexing. If the unit cell is 45x45x135 Å, then the detector distance should be set to a minimum of 135 mm if collecting binned data. If collecting unbinned data, then the detector may be set to half this distance (e.g., 66.5 mm for a 45x45x135 Å unit cell). Moving the detector away from the crystal will limit the amount of high resolution data that can be collected. Having the detector too close to the crystal will result in the merging of low-resolution diffraction spots which will prevent the correct indexing of the crystal and subsequently prevent the generation of an interpretable electron density map. Collecting binned data allows shorter exposure times but requires the detector to be further from the crystal. Binned data collection is often preferred for crystals having unit cells in which the longest dimension is less than ca. 60 Å.

Many X-ray diffractometers offer the option of moving the CCD area detector angularly to collect high resolution data. Moving the detector to an angle (termed 2ϑ) allows for the detection of a small portion of the high-resolution data but also substantially increases the number of images and the collection time.

Before collecting a complete data set, a group of diffraction images should be taken at multiple exposure times (e.g., 15, 30, 45, 60, and 90 seconds) to determine the optimal exposure time. Optimization of the exposure time is important, because the long exposure

times required to collect high-resolution data may saturate the detector and result in loss of low-resolution data which are critical in generating an electron density map. Conversely, collection of data with too short an exposure time may result in a lower resolution structure. Once the correct detector distance and exposure time are chosen, a full diffraction data set can be collected.

The Synchrotron

Synchrotrons offer many advantages for diffraction data collection over X-ray diffractometers. Synchrotrons generate very bright X-ray radiation, at up to 10^8 times greater flux than that of an X-ray diffractometer. Synchrotrons also have tuneable wavelengths, allowing various heavy atoms to be used in single-wavelength anomalous diffraction (SAD) and multiple-wavelength anomalous diffraction (MAD) experiments. Diffraction data collection is often easier and faster because most synchrotron facilities contain much larger CCD area detectors than those on an X-ray diffractometer.

The high flux of the synchrotron radiation allows lower exposure times (0.5 – 2 seconds) and rapid collection of a full data set. Diffraction data sets that take a day or more to collect on a diffractometer take well under an hour on a synchrotron. The disadvantage of the synchrotron is that the X-ray flux is so high that it may destroy the crystal with radiation damage before it is possible to collect a complete data set. Radiation damage can be seen during data collection as a loss of resolution or shrinking of the diffraction pattern. It can also be seen as a yellowing of the crystals at the point that the beam is centered on the crystal. With large crystals it may be possible to complete a data set even after radiation damage has occurred by moving the beam to an unaffected part of the crystal and resuming collection.

Synchrotron facilities have large CCD area detectors — up to nine times larger than that

of an X-ray diffractometer — allowing the collection of high-resolution diffraction in a single image rather than multiple images. The detector distance typically does not need to be adjusted based on the longest unit cell dimension, because the large size of the CCD area detector typically permits the detector to be left at 200 mm, which is well above that which is needed for typical unit cells. If the crystals diffract at very high resolution, then the distance may be shortened to acquire the highest resolution data.

I typically try to minimize the number of images collected to minimize radiation damage to the crystals. I use software programs, such as WebIce[17,18] and iMosflm, to predict the total number of images required for a complete dataset. Fewer diffraction images are needed for crystals that have high symmetry space groups and collecting 360 degrees of the crystal is often not required.[19]

Processing Diffraction Data

Many software packages are available to process diffraction images and the choice of a particular software program is determined by the researcher. Programs such as iMosflm, HKL2000, and d*TREK provide an easy-to-follow graphical user interface for looking at diffraction images, indexing diffraction spots, and processing data sets. Programs such as XDS are script based and may be difficult for a novice. Regardless of the program, the steps for processing the images are similar for each of the software packages.

The first step in processing diffraction data is indexing the data to find the unit cell dimensions and the primitive space group. Indexing determines the unit cell dimensions and the probability of the diffraction pattern belonging to one of the 14 Bravais crystal lattices. Once a lattice is selected, either by the program or by the researcher, the remaining diffraction images are integrated and scaled. The integrated data set is scaled and merged to make a single reflection file. I typically use the program Pointless[20] to find the most

probable space group and the program Aimless[21] to merge the diffraction data and analyse the quality of the data.

Assessing Diffraction Data Quality

The scaling program Aimless provides an easy-to-follow summary of the data processing statistics. There are a few important values to note when examining the final statistics of the scaling and merging steps. Aimless lists the space group, cell dimensions, and number of reflections measured during the diffraction experiment. In addition, Aimless also lists a set of statistics involving the merging of redundant diffraction spots and centrosymmetric diffraction data (R_{merge}, R_{sym}, and R_{pim}). These values should be low (<10%) for low resolution shells and will increase substantially in the highest resolution shell. The R_{merge} value has been widely used to determine how far the resolution could be extended, but should be supplanted by the $CC_{1/2}$ value.[22,23]

Another important value that should be noted is completeness. Data sets that have completeness lower than 90% will be difficult to solve. Data sets collected on a synchrotron often easily achieve 100% completeness because of the large CCD area detectors, while data sets collected on a diffractometer typically require collecting many diffraction images to achieve completeness due to the limitations of the instrument to collect high-resolution diffraction. Additional diffraction data should be collected if completeness is below 90% or the diffraction data should be processed to a lower resolution.

In addition to listing the statistical values, Aimless suggests a high-resolution cut-off based on either the $CC_{1/2}$ value or a ratio of spot intensity to background noise ($I/\sigma(I)$) of greater than 2.0. These values are important when attempting to process data to the highest possible resolution without introducing excessive noise in the electron density map.

The Electron Density Map

The diffracted X-rays contain two pieces of information that are key to generating the electron density map: the amplitude and the phase. The amplitude and phase contain information about the location and magnitude of electron density within a crystal lattice. The amplitude comes from the intensities measured during the diffraction experiment. The phase information, however, cannot be measured directly and is lost during the diffraction experiment. As a result, the X-ray diffraction and data collection processes alone do not generally provide all of the information that is needed to generate the electron density map and thus determine the crystallographic structure of a peptide or protein. The inability to directly measure the phase information during a diffraction experiment is commonly referred to as the phase problem in X-ray crystallography.

There are three main ways to obtain the missing phase information of a diffraction experiment: direct phasing, anomalous diffraction, and molecular replacement. Direct phasing, commonly referred to as direct methods, is routinely used in small molecule crystallography and requires higher resolution diffraction data than most peptides or proteins provide, typically better than 1.2 Å. Anomalous diffraction experiments, such as multi-wavelength anomalous diffraction (MAD) and single-wavelength anomalous diffraction (SAD), require an ordered heavy atom within the crystal lattice. The specific heavy element will dictate the wavelength chosen during the diffraction experiment. I routinely use SAD techniques for determining the crystallographic phases and discuss this technique in-depth, below.

Molecular replacement is routinely used in protein crystallography to obtain phase information but is often much more difficult to use for peptide crystals. Molecular replacement relies on homologous structures to generate phases.[24] The homologous structure is used as a search model to find the location of the protein within the asymmetric unit (ASU), the smallest repeating unit within a unit cell. The success of molecular replacement depends on

the similarity between the search model used and the structure within the crystal lattice. Molecular replacement generally requires a low number of monomers within the ASU, typically one or two, to be successful. Additional monomers make molecular replacement very difficult. I have typically observed three or more monomeric peptide units in the ASU and have not generally been able to use molecular replacement to determine the crystallographic phases.

Single-Wavelength Anomalous Diffraction Phasing

I routinely use SAD experiments to accurately determine the phases and generate electron density maps and crystallographic structures of peptides. SAD experiments involve locating the heavy atom by measuring differences in diffraction data intensity that occur when the wavelength used is near the absorption edge of the heavy atom. The location of the heavy atom provides information of the phases and can then be used to help determine the phases for the entire density map.

The strength of the anomalous signal, and hence its utility, depends on the heavy atom chosen and the wavelength used for the diffraction experiment. Anomalous scattering is a fundamental property of all atoms, however only a few atoms have appreciable anomalous scattering in the wavelength range typically used for X-ray crystallography (2.47 – 0.77 Å, 5 – 16 keV). The strength of the anomalous signal depends on how close the wavelength is to the absorption edge of the heavy atom. For a given atom, the real (f') and imaginary (f'') components of the scattering factor (f) vary as a function of wavelength. Figure 4.5 shows the theoretical plot of f' and f'' for Se. The absorption edge for Se is at 0.98 Å (12.6 keV), with values of f' = -8.3 and f'' = 3.8. The large difference between the f' and f'' at 0.98 Å produces detectable differences in diffraction spot intensities. These differences can then be used to calculate the position of the Se atom in the crystal lattice and the phases of the crystal. At longer wavelengths, like that of Cu (1.54 Å), the difference between f' and f'' is

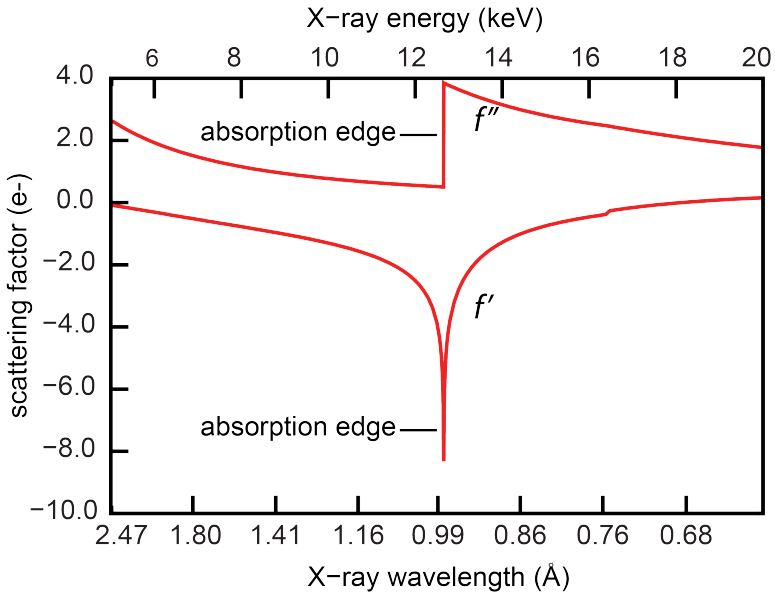


Figure 4.5: Anomalous scattering calculated for Se between 2.47 – 0.67 Å (20 keV – 5 keV). The absorption edge is at 0.98 Å (12.6 keV).[b]

substantially smaller, and measuring the difference in diffraction intensities is not generally practical.

A diffractometer with a Cu anode is suitable for anomalous phasing with I (iodine) and a number of transition metals but it not suitable for Br and Se. I routinely incorporate iodine into our peptides in the form of p-iodophenylalanine to collect anomalous diffraction data. Although, iodine does not have an edge within the typical range for X-ray diffraction, the anomalous signal at 1.54 Å is large enough for SAD phasing (f' = -0.6, f'' = 6.8). I have routinely used iodine to determine the phases and phases and generate electron density maps of our crystal structures.

The strength of the anomalous signal also depends on the localization of the heavy atom within the crystal. Heavy atoms that adopt well defined positions and have low movement within the crystal lattice promote a strong anomalous signal. Heavy atoms that are not localized to single positions within the lattice have poor anomalous signals. This is often seen

[b]Theoretical anomalous plot generated from data available at http://skuld.bmsc.washington.edu/scatter/

with heavy atoms substituents on amino acids that can adopt multiple conformations within the lattice, such as the Se heavy atom in selenomethionine or the salts of heavy atoms. In *p*-iodophenylalanine, the iodine is attached to the aromatic ring of the phenylalanine amino acid and is constrained to very few rotamers. I have noticed that the *p*-iodophenylalanine often adopts a single rotamer within each macrocycle in the crystal lattice and promotes a strong anomalous signal.

Generating the Electron Density Map from SAD Data

I routinely determine the crystallographic phases and generate the electron density map using the Phenix software package.[25] The steps I described here are general for determining the phases and generating an electron density map but the programs are specific to the Phenix software suite. Other software suites, such as CCP4i, provide similar programs for determining the phases and generating an electron density map. The steps involved in generating an electron density map are outlined in Figure 4.6.

I typically begin with an analysis of the data using the program Xtriage. Xtriage provides information about the quality of the diffraction data, similar to the information calculated by Aimless during the initial processing of diffraction data. Xtriage provides a straightforward explanation of the data quality and suggestions for improving data processing. Xtriage provides information on the total number of residues or peptides in the ASU according to Matthews coefficient, a measure of protein density in the ASU.[26] Most protein and peptide crystals contain only about 50% protein with the rest being water or other solvent. Very high or low solvent content (e.g., 70% or 30%) is unlikely. If the solvent content is calculated to be far from 50% after data processing, then the data was probably processed in the wrong space group.

Xtriage also provides information on the presence of twinning within the crystal. Twin-

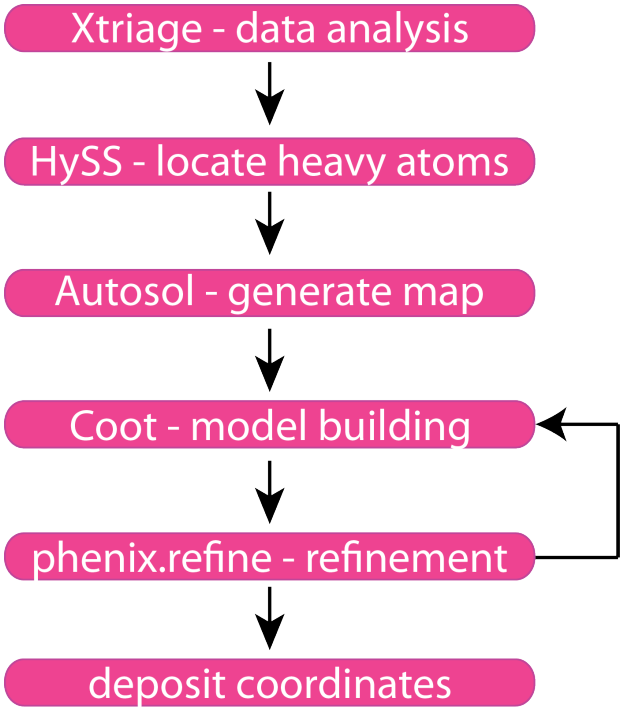


Figure 4.6: Overview of the programs and steps involved in generating an electron density map, building a molecular model, refining the structure, and depositing the coordinates.

ning is common in protein and peptide crystals. Non-merohedral twinning often leads to crystals that do not index easily and diffraction data that are difficult to process and solve. Merohedral and pseudo-merohedral twinning can lead to processing of the data in a space group with too high symmetry. Xtriage provides a multivariate Z score L-test for assessing the presence of twinning. Values greater than 3.5 often indicate twinning.

The next step in generating the electron density map is locating the heavy atoms within the ASU. We use the program HySS (hybrid substructure search) to locate the heavy atoms.[27,28] HySS requires the identity of the heavy atom, the number of heavy atoms in the ASU, and the wavelength used during diffraction experiment. Since I typically use *p*-iodophenylalanine, I know that the number of heavy atoms within the ASU is the same as the number of peptide molecules within the ASU that was determined by Xtriage.

HySS generates a set of coordinates for the heavy atoms and provides a correlation coefficient for assessing the likelihood of correct placement of the heavy atoms within the ASU.

Correlation coefficients greater than 0.4 reflect a high probability of correct placement of the heavy atoms, while values less than 0.3 often reflect incorrect placement. A low correlation coefficient value often indicates an incorrect space group or poor diffraction data quality, or that the anomalous signal is too weak to provide definitive information about the location of the heavy atoms.

HySS also provides information on the occupancy of the heavy atoms within the ASU. Normal occupancies of heavy atoms are typically 1. Occupancies above 1 are suspect and probably indicate that the data were processed incorrectly. Heavy atoms with occupancies below 0.2 are suspect and can often be ignored or removed before generating an electron density map.

The positions of the heavy atoms generated by HySS are used directly in the program Autosol to determine the phases of generate an electron density map. Autosol combines the phasing and model building operations into a single software program.[27,29–31] The sites found in HySS are used as the initial heavy atom locations. These locations serve as the starting point to determine the phases and generate the electron density map. Autosol runs through multiple algorithms, including density modification, an automatic building feature that places the amino acids directly into the electron density map, and refinement of the newly built structure.

The success of the phasing and model building steps in Autosol can be assessed by the figure of merit (FOM) and Bayes-CC values. A high FOM and a high Bayes-CC score (e.g., 0.6 and 40) characterize a correctly phased electron density map. FOM values less than 0.3 or Bayes-CC values below 30 usually indicate a poor solution, while even lower values reflect worse solutions. Figure 4.7 shows an electron density map generated by Autosol with high FOM and Bayes-CC values and an electron density map with low FOM and Bayes-CC values. The electron density map with low values appears as uninterpretable blobs of density.

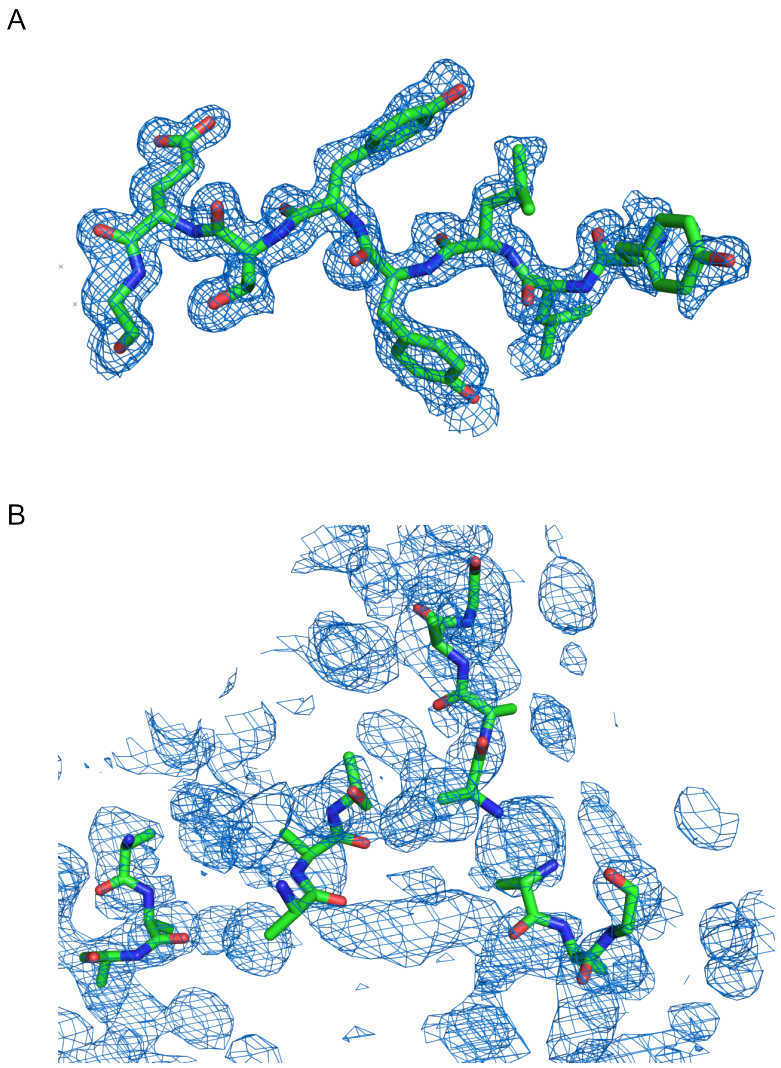


Figure 4.7: (A) A good electron density map and starting model generated by Autosol, with high FOM and Bayes-CC values. (B) A poor electron density map and model generated by Autosol, with low FOM and Bayes-CC values.

The automatic amino acid building feature in Autosol provides a good starting point for the initial model of the crystallographic structure. Unnatural amino acids and unnatural linkages are not handled well by Autosol and must be modified by the researcher. The next section will cover modifications of the initial model and refinement of the molecular model.

Model Building and Structure Refinement

Refining the crystallographic structure is iterative, requiring multiple rounds of model building and generation of the electron density map. The initial model generated by Autosol does not typically contain all of the amino acids of the peptide, but rather contains a few amino acids that can serve as a starting point for structure refinement.

The goal of model building and structure refinement is to more accurately determine the phases and generate a complete electron density map for the ASU. The initial electron density map generated by Autosol is often incomplete and may be missing electron density associated with additional amino acids. As more residues are correctly placed into the density, the accuracy of the phases increases; further iterations of refinement generate a more complete electron density map, which in turn allows further model building.

I typically start the model building with the density-modified map and the overall best placed pdb file generated by Autosol. I use the program Coot to manipulate this initial model by adding amino acids and side chains to fit the electron density map.[32] I use the program phenix.refine to iteratively refine the molecular model and electron density map.

Coot contains a library of many natural and unnatural amino acids and ligands (solvents, ions, additives, etc.) for building models within the electron density map. During modelling, Coot attempts to fit the amino acids and ligands into the electron density map by altering the conformations of the main chains and side chains. The conformational properties of each amino acid and ligand are described by the crystallographic information file (cif), which specifies atom attachments, bond lengths, chirality, torsion angles, and planes. Amino acids and ligands within the library often contain a correct cif file. For unnatural amino acids and ligands not in the library, it is necessary to generate a cif file. For uncommon amino acids and ligands in the library it may be necessary to generate a cif file if the bond lengths and bond angles are incorrect.

The program Elbow in the Phenix software suite can be used to create a cif for an unnatural amino acid or ligand not in the library.[33] Elbow takes a coordinate file or SMILES string describing the structure and generates pdb and cif files suitable for use in Coot. The pdb and cif should be examined for accuracy and edited if necessary using a text editor.

The First Refinement

Each round of refinement involves comparing the placement of the model to the diffraction data. The first round of refinement is never the last round of refinement. The first round takes the unrefined structure and offers insight into the most troublesome parts. The model and diffraction data are each converted to a set of structure factors (F). The deviation of the structure factors of the model from those of the diffraction data is calculated, and the normalized value of the deviation is termed R_{work}. The residual value R_{work} provides an assessment of how well the model explains the diffraction data. An ideal model that precisely explains all of the diffraction data would have a R_{work} of zero. A model randomly placed within the electron density map would have an R_{work} of 0.63 (63%).[34] A typical R_{work} after complete refinement is about 0.20 (20%).

A second residual value, R_{free}, provides an additional assessment of the accuracy of the model and helps prevent model bias, which often occurs in molecular replacement.[35] The residual R_{free} is similar to R_{work}, but excludes a small subset of diffraction data from refinement. This subset of data typically comprises 5 – 10% of the total diffraction data. R_{free} should always be slightly larger than R_{work} and is typically 0.02 – 0.05 (2 – 5%) higher. If R_{work} is 0.20, than R_{free} should be around 0.22 – 0.25. If R_{free} is much higher than R_{work}, then the electron density map may be flawed, suffering from model bias, unaccounted twinning, or poor quality diffraction data. If R_{free} is lower than R_{work}, the R_{free} subset was likely chosen incorrectly and a new subset should be generated. The first refinement will typically give high R_{work} and R_{free} values, typically in the 30's (0.3 – 0.4). As more residues

and ligands are placed correctly into the electron density map, R_{work} and R_{free} should drop dramatically.

The first refinement also reveals some of the problems in the model, such as poor geometries, steric clashing, and high B-factors. The molecular model and electron density map should be compared to determine if these problems should be corrected before the next round of refinement.

I use the program phenix.refine to perform the first refinement and each subsequent refinement.[36] Phenix.refine uses the initial model, the reflection file generated by Aimless, and any additional required cif files to perform the refinement. Phenix.refine offers many different refinement strategies. The default strategy refines the XYZ coordinates, isotropic B values, and occupancies of the atoms in the model. Phenix.refine also does real space refinement in which the residues are placed more precisely within the electron density map. Improper bond lengths and angles generated during the modelling step are corrected. Additional options such as TLS parameters and automatic water picking are available.

The default refinement strategy in phenix.refine typically works well for the initial refinement, with the following minor adjustments: When the structure contains heavy atoms with anomalous signals, then the f' and f'' values should be refined during the first refinement and in subsequent refinements. Phenix.refine provides an option to permit refinement of these values, which should be selected. Hydrogens should be added to the model either before refinement or using the add hydrogens feature within phenix.refine. The positions of the hydrogen atoms are typically calculated, rather than determined experimentally from the electron density map. These riding hydrogens are useful in avoiding poor geometries and steric clashes within the model during refinement.

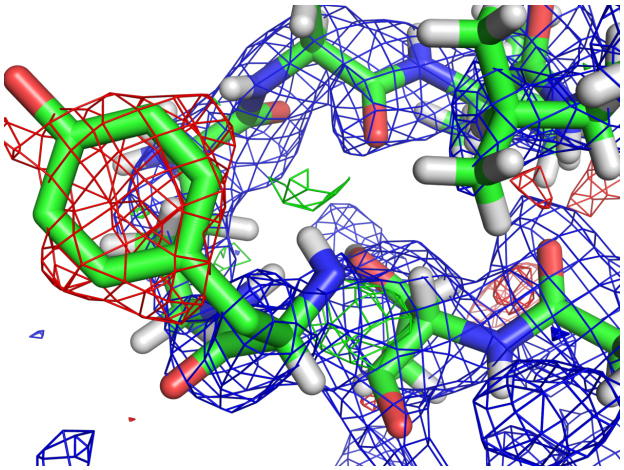


Figure 4.8: A difference electron density map overlaid on the electron density map. The electron density map is shown in blue. The difference electron density map is shown in red and green. Red areas of the difference electron density map correspond to incorrectly placed atoms or residue side chains; green areas correspond to locations that should contain additional atoms or bonds.

Further Refinement

The model is now subjected to subsequent rounds of refinement. After each round of refinement a new model and a new electron density map are generated. In phenix.refine, a copy of the original reflection file is also generated. The newly refined electron density map contains both an electron density map and a difference electron density map. The difference electron density map shows regions of the surplus electron density that should contain atoms (i.e., more electrons) and regions of the electron deficiency that should not contain atoms (i.e., fewer electrons). The regions of surplus electron density may reflect a need to add side chains, residues, ligands, and water in order to improve the model. The regions of electron deficiency reflect a need to alter the model in these regions. Figure 4.8 shows an electron density map showing electron density that should contain additional atoms (green) and a region that should not contain additional atoms (red). As the refinement progresses, waters and additional ligands should be added to the model to further lower the R_{work} and R_{free} values.

The Final Refinement

Subsequent iterations of the refinement process eventually produce diminishing improvements in the model. When the researcher finally concludes that no further significant improvements in the fit of the model to the electron density map and the R_{free} and R_{work} values can be achieved, refinement is complete. At this point, the R_{work} value should be comparable to or lower than the resolution of the data set: A data set with 2.0 Å resolution should produce an Rwork value of 20% or lower and an R_{free} value 2–5% higher.[37] Large deviations between the resolution of the structure and the R_{work} and R_{free} values indicates a poorly placed or incorrect model, the unaccounted presence of twinning, or poor quality diffraction data.

The final model should also have low RMS (root mean square) bond angle deviations and bond length deviations, and few rotamer outliers. Phenix.refine provides a summary of these values as well as a detailed description of which bond lengths and bond angles deviate. The phi (φ) and psi (ψ) dihedral angles of the amino acids within peptide are calculated and compared to the Ramachandran plot. Large deviations in these values are of particular concern and the offending parts of the model should be examined and adjusted.

Phenix.refine uses the program MolProbity to analyse steric clashes between atoms within the model.[38] MolProbity determines the overlap between atoms within the model and provides information about offending atoms on each residue or ligand. Individual steric clashes are tallied to generate a total clashscore value. An ideal model will have a clashscore of zero. A final model should have a clashscore below 10 and as close to zero as possible.

Phenix.refine also provides information about the thermal motion of atoms within the crystal, termed the B-factor or atomic displacement parameter. The atoms along the amino acid backbone typically have B-factors of 30 or lower. The B-factors of the atoms in of the side chains are slightly higher than those of the backbone. The atoms on flexible side

chains, like lysine, can have a B-factor in the 40's. Solvents and ligands typically have higher B-factors than the side chains and are typically above 30. Phenix.refine generates a B-factor histogram and flags atoms with unusually high B-factor values. High B-factors can be indications of poorly placed side chains or incorrect atom types.

The accuracy of the final molecular model ultimately depends on the researcher. When the researcher has addressed all of the issues with the model raised during refinement (poor geometry, high B-factors, high clashscore, etc.) and the structure refines to R_{work} and R_{free} values in the 20's or lower then the final molecular model becomes the X-ray crystallographic structure and represents the most accurate description of the contents of the crystal.

Depositing the Crystallographic Structure into the PDB

The dissemination of the crystallographic structure is an important step in peptide and protein crystallography, because it allows others to learn from and build upon the structure. Disseminating the crystallographic structure is typically achieved by depositing the coordinates in the Protein Data Bank, a free and open-access archive of biomolecular structures.[39] The PDB stores and makes publicly available the crystallographic coordinates and reflection files generated during the refinement process. The PDB also offers additional validations tools for assessing the quality of the model prior to deposition. The deposited coordinates are reformatted to PDB standards, run through a validation program, and assigned a four character PDB identification code.

The PDB provides an easy-to-follow web-based interface for depositing the crystallographic coordinates. Information about the crystallization experiment, the diffraction data collection experiment, data processing, and refinement are required for depositing the structure. Users are able to place a hold on releasing the coordinates to the public prior to publication. Structures are typically processed within a week or two after deposition. Severe errors in

the structure, such as the overlap of atoms or molecules, are flagged for review. These errors should then be considered by the researcher and corrected where appropriate. Any changes done to the deposited model should be rerun through a round of refinement and the new model should be submitted with the appropriate corrections.

Case Study

This section illustrates the principles and procedures in the previous section, by describing how I solved the X-ray crystallographic structures of peptides containing sequences from the β-amyloid peptide (Aβ) and observed their supramolecular assembly to form oligomers.[8] We designed macrocyclic peptides **1** and **2** to fold into a β-sheet that incorporates two heptapeptide sequences from Aβ the central region $A\beta_{17\text{-}23}$ (LVFFAED) and the C-terminal region $A\beta_{30\text{-}36}$ (AIIGLMV). We made three modifications to the heptapeptides to facilitate these studies: We changed the Met_{35} residue to the isostere ornithine to increase solubility. We changed the Gly_{33} residue to sarcosine (*N*-methyl Gly) to help prevent fibril formation, promote oligomer formation, and facilitate crystallization. In peptide **2**, we changed the Phe_{19} residue to *p*-iodophenylalanine to determine the X-ray crystallographic phases. I then used the model of peptide **2** to determine the phases of peptide **1**, which lacks an iodine atom (Figure 4.9).

Crystallization, Diffraction Data Collection, and Structure Determination of Peptide 2

I determined the crystallization conditions for peptide **2** using the techniques described previously. We screened peptide **2** at 10 mg/mL in three crystal screens (Index, PEG/Ion, and Crystal Screen) in 96-well plates using a Mosquito crystallization robot. Crystal growth

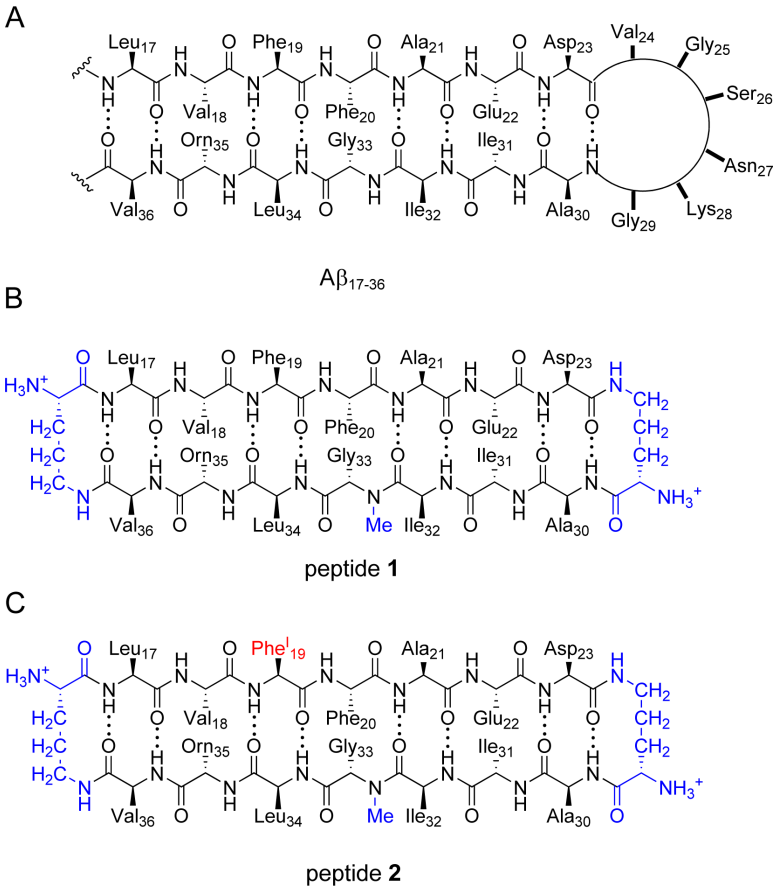


Figure 4.9: (A) Chemical structure of a β-hairpin formed by $A\beta_{17\text{-}36}$. (B) Chemical structure of peptide **1**, illustrating a macrocycle containing $A\beta_{17\text{-}23}$, $A\beta_{30\text{-}36}$ M35O, and sarcosine (*N*-methyl Gly_{33}). (C) Chemical structure of peptide **2**, containing *p*-iodophenylalanine (Phe^{I}_{19}).

occurred after 24 hours in several different conditions. I selected conditions containing 0.1 M HEPES at pH 7.5 with 20% Jeffamine M-600 at pH 7.0 for further optimization in 24-well plates. I harvested crystals suitable for X-ray crystallography from a well containing 0.1 M HEPES at pH 6.5 with 25% Jeffamine M-600 at pH 7.0.

I used a Rigaku MicroMax-007 HF diffractometer equipped with a Cu rotating anode to collect a diffraction data set of peptide **2**. Peptide **2** diffracted in the *R*3 space group with cell dimensions of 68x68x170 Å. Diffraction data were collected with a detector distance of 78 mm and 2ϑ of 15°. I used the program CrystalClear to determine the optimal strategy for data collection. The diffraction data were collected unbinned and to a resolution of 1.99 Å. The diffraction images were scaled and merged with XDS. (I now prefer Aimless instead

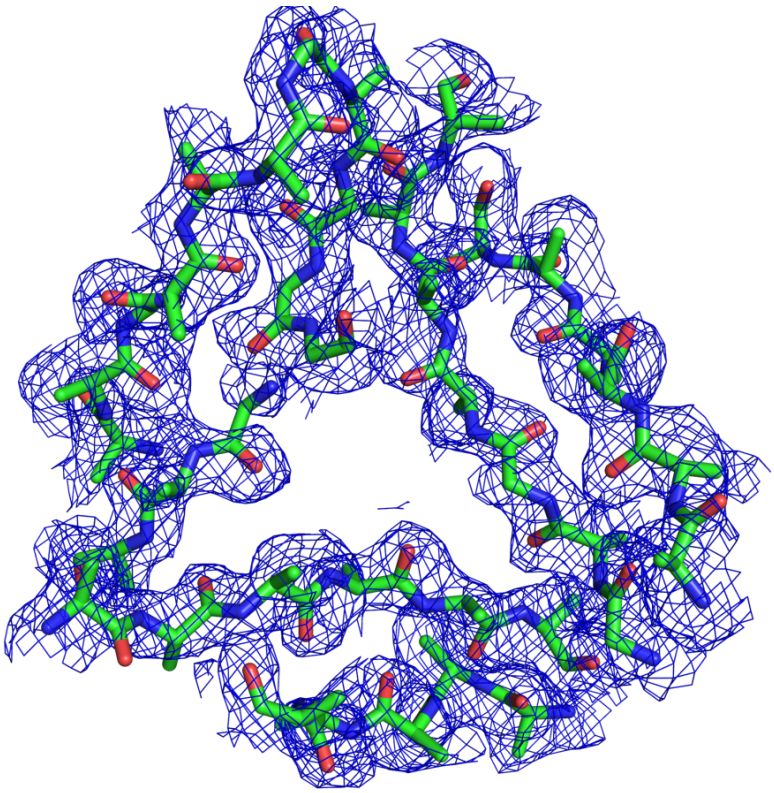


Figure 4.10: The electron density map of peptide **2** generated by Autosol with the starting model placed within the electron density map.

of XDS to merge diffraction data sets.)

Analysis of the diffraction data in Xtriage revealed that the unit cell contained 16 macrocycles with about 53% solvent content. I then searched for the iodine locations using HySS. HySS found 16 iodine sites with a correlation coefficient of 0.57. The heavy atom locations determined in HySS were used directly in Autosol to generate an electron density map with a FOM of 0.45 and a Bayes-CC of 58. Figure 4.10 shows the initial electron density map and model generated by Autosol for peptide **2**.

I built peptide **2** into the electron density map using Coot. The cif files for the ornithine and sarcosine residues were generated in Elbow, and bond lengths and angles were modified to more accurately describe each residue. A single monomer of peptide **2** was built and placed within the electron density map. The monomer was refined using phenix.refine with

hydrogens to generate a starting model that could be duplicated and placed into the electron density map an additional 15 times. The model and electron density map were refined after each new peptide was added. After all 16 peptides were added to the map, waters were added using the option within phenix.refine. I used the default refinement strategy in phenix.refine and selected refinement of the anomalous signal and TLS parameters. We achieved anR_{work} of 20.7, an R_{free} of 24.6, and a clashscore of 2 in the final refinement. I deposited the final structure and electron density map at the Protein Data Bank (PDB ID 4NTP).

Crystallization, Diffraction Data Collection, and Structure Determination of Peptide 1

I used the structure of peptide **2** to determine the phases and the structure of peptide **1**. This method is called isomorphous replacement and is possible when homologous structures have unit cells that are identical or nearly identical. It is often applied to determine the phases in protein crystallography by incorporating a heavy atom into the lattice.

We crystallized peptide **1** in similar conditions to those of peptide **2** (0.1 M HEPES at pH 6.75 and 31% Jeffamine M-600 at pH 7.0). I collected the diffraction data set on the synchrotron beamline 8.2.2 at the Advanced Light Source at Lawrence Berkeley National Laboratory, Berkeley, CA. I used iMosflm to determine the number of images to collect. I collected 200 diffraction data images over 100° of the crystal with 0.5° rotation per image. The diffraction data were collected at 1.00 Å wavelength with a detector distance of 220 mm. The diffraction data were processed to 1.70 Å resolution. Peptide **1** diffracted in the $R3$ space group with cell dimensions of 68x68x169 Å, nearly identical to the cell dimension found for peptide **2**.

I used the program Phaser in the Phenix software suite to determine the phases and

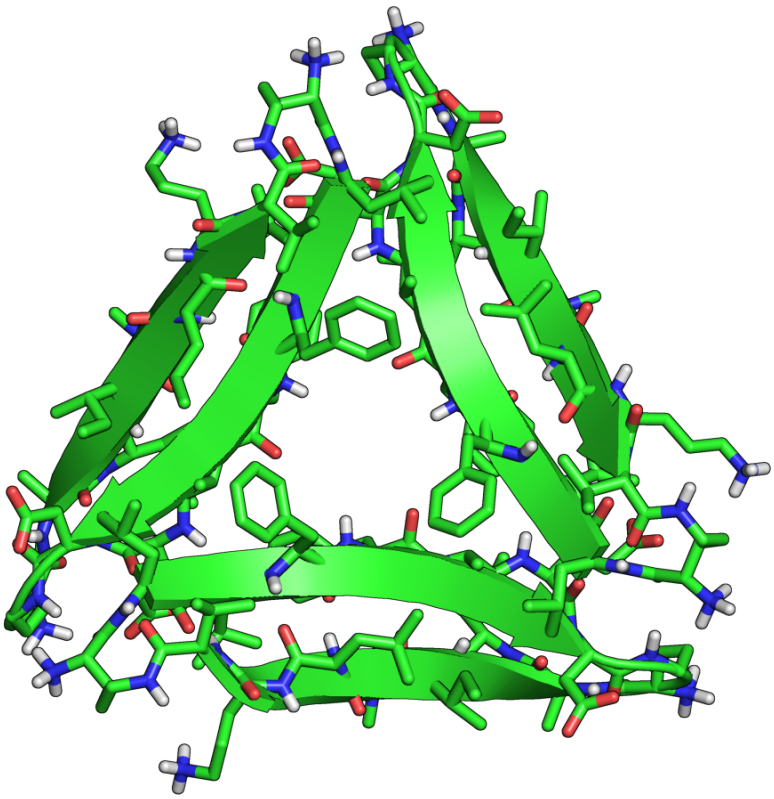


Figure 4.11: X-ray crystallographic structure of the triangular trimer formed by peptide **1** (PDB ID 4NTR).

generate an electron density map of peptide **1** using the structure determined for peptide **2**.[30] I used the starting phases generated by Phaser directly in Autosol to generate the electron density map and the starting model for peptide **1**. The model building and refinement of peptide **1** was performed by the same procedures used for peptide **2**. I achieved an R_{work} of 20.0, an R_{free} of 22.6, and a clashscore of zero in the final refinement. I deposited the final structure and electron density map at the Protein Data Bank (PDB ID 4NTR).Figure 4.11 illustrates the X-ray crystallographic structure of peptide **1** and the supramolecular assembly to form a trimer comprising three of the macrocylic β-sheets.[8]

Summary

Peptide crystallography is a powerful tool for studying the structure and supramolecular assembly of peptides. Facile peptide synthesis and the availability of amino acids containing heavy atoms permit the creation of peptides suitable for X-ray crystallographic structure determination. The availability of high-throughput crystal screening kits and automated crystallization robots have greatly reduced the time and amount of peptide needed to discover crystallization conditions and grow crystals. Ready access to X-ray diffractometers and synchrotrons, advances in X-ray diffraction technology, and the availability of free, user-friendly data processing software offer the promise of making peptide crystallography as routine as protein crystallography. The techniques described in this guide should allow other newcomers to partake of this powerful tool and apply it to their research problems.

References

1. Rupp, B. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*; Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology; Garland Science: New York, 2010; pp 1–809.

2. Benedetti, E. *Biopolymers* **1996**, *40*, 3–44.

3. Garman, E.; Murray, J. W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 1903–1913.

4. Khakshoor, O.; Lin, A. J.; Korman, T. P.; Sawaya, M. R.; Tsai, S. C.; Eisenberg, D.; Nowick, J. S. *J. Am. Chem. Soc.* **2010**, *132*, 11622–11628.

5. Liu, C.; Sawaya, M. R.; Cheng, P.-N.; Zheng, J.; Nowick, J. S.; Eisenberg, D. *J. Am. Chem. Soc.* **2011**, *133*, 6736–6744.

6. Liu, C.; Zhao, M.; Jiang, L.; Cheng, P.-N.; Park, J.; Sawaya, M. R.; Pensalfini, A.; Gou, D.; Berk, A. J.; Glabe, C. G.; Nowick, J. S.; Eisenberg, D. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 20913–20918.

7. Pham, J. D.; Chim, N.; Goulding, C. W.; Nowick, J. S. *J. Am. Chem. Soc.* **2013**, *135*, 12460–12467.

8. Spencer, R. K.; Li, H.; Nowick, J. S. *J. Am. Chem. Soc.* **2014**, *136*, 5595–5598.

9. Dauter, Z.; Dauter, M.; Rajashankar, K. R. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000**, *56*, 232–237.

10. Xie, J.; Wang, L.; Wu, N.; Brock, A.; Spraggon, G.; Schultz, P. G. *Nat. Biotechnol.* **2004**, *22*, 1297–1301.

11. Abendroth, J.; Gardberg, A. S.; Robinson, J. I.; Christensen, J. S.; Staker, B. L.; Myler, P. J.; Stewart, L. J.; Edwards, T. E. *J. Struct. Funct. Genomics* **2011**, *12*, 83–95.

12. Pflugrath, J. W. *Methods* **2004**, *34*, 415–423.

13. Leslie, A. G. W.; Powell, H. R. In *Evolving Methods for Macromolecular Crystallography*; Read, R., Sussman, J. L., Eds.; NATO: Erice, 2007; Vol. 245; pp 41–51.

14. Otwinowski, Z.; Minor, W. In *Macromolecular Crystallography, Pt A*; Carter, C., Sweet, R. M., Eds.; Academic Press: New York, 1997; Vol. 276; pp 307–326.

15. Pflugrath, J. W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 1718–1725.

16. Kabsch, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125–132.

17. McPhillips, T. M.; McPhillips, S. E.; Chiu, H. J.; Cohen, A. E.; Deacon, A. M.; Ellis, P. J.; Garman, E.; Gonzalez, A.; Sauter, N. K.; Phizackerley, R. P.; Soltis, S. M.; Kuhn, P. *J. Synchrotron Radiat.* **2002**, *9*, 401–406.

18. Gonzalez, A.; Moorhead, P.; McPhillips, S. E.; Song, J.; Sharp, K.; Taylor, J. R.; Adams, P. D.; Sauter, N. K.; Soltis, S. M. *Journal of Applied Crystallography* **2008**, *41*, 176–184.

19. Dauter, Z. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 1703–1717.

20. Evans, P. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 72–82.

21. Evans, P. R.; Murshudov, G. N. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1204–1214.

22. Karplus, P. A.; Diederichs, K. *Science* **2012**, *336*, 1030–1033.

23. Diederichs, K.; Karplus, P. A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1215–1222.

24. Evans, P.; Mccoy, A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2008**, *64*, 1–10.

25. Adams, P. D. *et al. Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213–221.

26. Matthews, B. W. *J. Mol. Biol.* **1968**, *33*, 491–497.

27. Grosse-Kunstleve, R. W.; Adams, P. D. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 1966–1973.

28. McCoy, A. J.; Storoni, L. C.; Read, R. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1220–1228.

29. Terwilliger, T. C. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000**, *56*, 965–972.

30. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

31. Terwilliger, T. C.; Adams, P. D.; Read, R. J.; McCoy, A. J.; Moriarty, N. W.; Grosse-Kunstleve, R. W.; Afonine, P. V.; Zwart, P. H.; Hung, L. W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *65*, 582–601.

32. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486–501.

33. Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *65*, 1074–1080.

34. Kleywegt, G. J.; Jones, T. A. *Methods Enzymol* **1997**, *277*, 208–230.

35. Brunger, A. T. *Nature* **1992**, *355*, 472–475.

36. Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 352–367.

37. Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. *FEBS J* **2008**, *275*, 1–21.

38. Chen, V. B.; Arendall, r., W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12–21.

39. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.

Chapter 5

X-ray Crystallographic Structures of Oligomers of Peptides Derived from β_2-Microglobulin.

Introduction

Oligomers from amyloidogenic peptides and proteins are critical in many amyloid diseases. Although the amyloidogenic peptides and proteins differ among these diseases, as do the locations within the brain and the body, the oligomers that form appear to share common features of being toxic and causing cell damage and death. Much of the understanding about amyloid oligomers has come from the β-amyloid peptide (Aβ) and Alzheimer's disease.[1–6] Other amyloidogenic peptides such as α-synuclein,[7,8] islet amyloid polypeptide (IAPP),[9–11] and β_2-microglobulin (β_2m)[12–14] are thought to form harmful oligomers in Parkinson's disease, diabetes mellitus type II, and hemodialysis-related amyloidosis. Little is known about the structures of amyloid oligomers, and there is a desperate need for atomic-resolution struc-

tures. Atomic-resolution structures of the toxic oligomers are essential to understanding the mechanisms by which they cause cell damage and death and developing effective therapies for amyloid diseases.

Many of the tools for studying the structures of amyloid oligomers have provided the molecularity and morphology of amyloid oligomers but not the structures at atomic resolution. Atomic force microscopy (AFM), transmission electron microscopy (TEM), size-exclusion chromatography (SEC), gel electrophoresis, and ion mobility mass spectrometry techniques have provided low resolution structural information about the oligomers formed by Aβ. Compact spheroids have been observed by AFM and TEM.[15,16] Dimers, trimers, and an apparent dodecamer, isolated from the brain tissue of transgenic mice, have been observed by gel electrophoresis.[4] Dimers, tetramers, hexamers, and dodecamers have also been observed by ion mobility mass spectrometry.[17] Infrared spectroscopic studies suggest that antiparallel β-sheets are involved in oligomer formation.[18–21] While NMR and X-ray diffraction studies have provided detailed structural information about amyloid fibrils,[22–27] these tools appear less suitable for studying oligomers of full-length amyloidogenic peptides and proteins.

Full-length amyloidogenic peptides and proteins are difficult to study because they often form a heterogeneous collection of soluble oligomers and insoluble fibrils. Small hydrophobic regions of these peptides and proteins, consisting of three or more hydrophobic residues, are often responsible for the aggregation.[28,29] Understanding the structures formed by these smaller regions can help elucidate the structures of full-length amyloidogenic peptides and proteins. Peptide fragments containing these regions are easier to study because they can more easily form homogeneous assemblies. Eisenberg and coworkers determined the structure of an oligomer formed from the amyloidogenic region of αB crystallin by X-ray crystallography.[30] In this structure, six β-strands associate to form a six-stranded hydrogen-bonded antiparallel β-sheet that rolls up to form a cylindrical oligomer in which hydrophobic residues

comprise the inner core of the structure. Apostol, Perry, and Surewicz determined the X-ray crystallographic structure of an oligomer formed from fragments of human prion protein (PrP).[31] In this structure, two PrP β-strands are linked through a disulfide bond to form a hydrogen-bonded dimer. Six dimers associate along the hydrogen-bonding edges to form a cylindrical hexamer with hydrophobic residues comprising the inner core of the structure. Our laboratory has determined the X-ray crystallographic structure of an oligomer formed by a macrocyclic peptide derived from Aβ. In this structure, the peptide folds into an antiparallel β-sheet and assembles to form triangular trimers and higher-order oligomers.[32]

In the current study, I set out to use X-ray crystallography to explore the range of oligomers that a single amyloidogenic peptide sequence can form. I designed macrocyclic peptides **1** and **2** to incorporate the amyloidogenic heptapeptide sequence YLLYYTE ($\beta_2m_{63–69}$) from the aggregation-prone E chain of β_2m and fold into an antiparallel β-sheet (Figure 5.1).[33–35] The peptide contains a second heptapeptide sequence with an *N*-methyl amino acid, as a template strand to block uncontrolled aggregation.[36] The two heptapeptides are connected by two δ-linked ornithine turn units, which act as β-turn mimics and reinforce β-sheet formation.[37] I replaced Tyr_{66} of $\beta_2m_{63–69}$ with a *p*-iodophenylalanine to determine the X-ray crystallographic phases. I used Lys residues at the R_1 and R_7 positions of the template β-strand to enhance solubility, Val residues at the R_2 and R_6 positions to enhance β-sheet formation, and an *N*-methyl amino acid at the R_4 position to prevent fibril formation and promote oligomer formation.

I kept the $\beta_2m_{63–69}$ peptide strand constant and varied residues R_3, R_4, and R_5 to explore the effects of residue size and hydrophobicity on oligomer formation. Peptides **1** and **2** present two surfaces: a major surface that displays the side chains of eight amino acids and a minor surface that displays the side chains of six amino acids (shown by the blue side chains and red side chains in Figure 5.1). The major surface displays Tyr_{63}, Leu_{65}, Tyr_{67}, and Glu_{69} of $\beta_2m_{63–69}$, while the minor surface displays Leu_{64}, Phe^{I}_{66}, and Thr_{68}. The

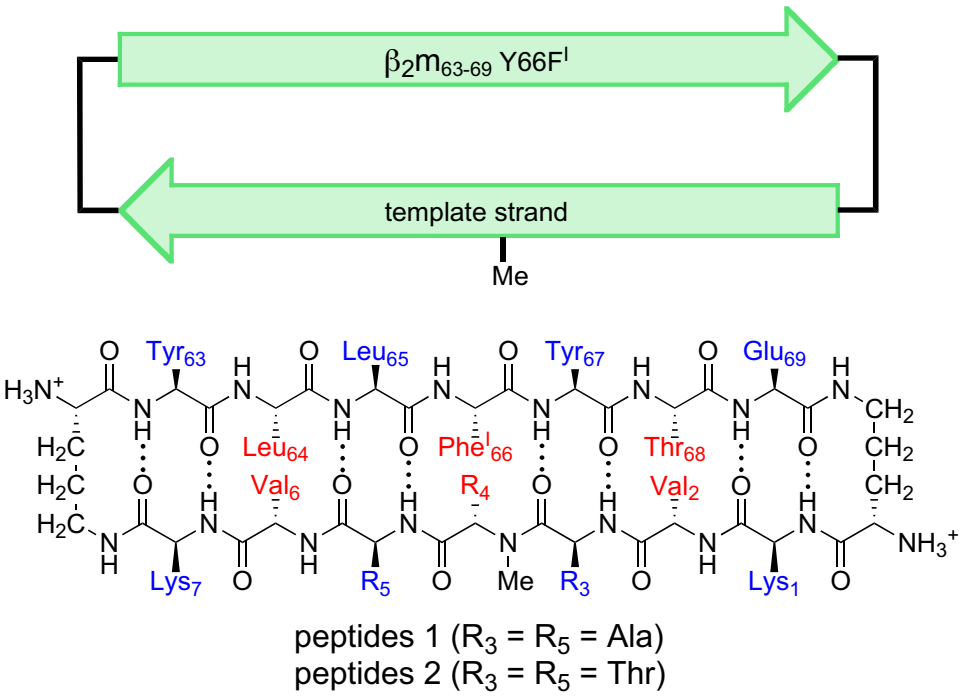


Figure 5.1: Cartoon and chemical structures of peptides **1** and **2**.

major surface also displays Lys_1, R_3, R_5, and Lys_7 of the template strand, while the minor surface displays Val_2, R_4, and Val_6. I initially synthesized and studied ten peptides. In five I incorporated alanine at positions R_3 and R_5 (**1a**–**1e**); in five we incorporated threonine at positions R_3 and R_5 (**2a**–**2e**). In each series, I varied the *N*-methylated residue R_4, to incorporate *N*-methylated alanine, valine, leucine, isoleucine, and norleucine (Nle). Table Table 5.1 summarizes the peptides we synthesized and the oligomers we observed by X-ray crystallography.

Results

Five of the ten peptides afforded crystals suitable for X-ray crystallography. Their X-ray crystallographic structures reveal three families of oligomers: hexamers, octamers, and dodecamers (Figure 5.4, Figure 5.5, and Figure 5.6). Peptides **1a** and **2a** form hexamers, peptide **2b** forms an octamer, and peptides **1b** and **1c** form dodecamers. These oligomers

Table 5.1: Peptides **1** and **2** and Oligomers Observed Crystallographically.

peptide	R_3	*N*-Me R_4	R_5	oligomer	resolution (Å)
1a	Ala	Ala	Ala	hexamer	1.97
1b	Ala	Val	Ala	dodecamer	1.50
1c	Ala	Leu	Ala	dodecamer	1.90
1d	Ala	Ile	Ala	–	
1e	Ala	Nle	Ala	–	
2a	Thr	Ala	Thr	hexamer	1.51
2b	Thr	Val	Thr	octamer	1.31
2c	Thr	Leu	Thr	–	
2d	Thr	Ile	Thr	–	
2e	Thr	Nle	Thr	–	
$\mathbf{1a_{T68V}}$	Ala	Ala	Ala	hexamer	2.02
1a + 1c	Ala	Ala/Leu	Ala	dodecamer	1.91

are composed of monomer subunits with common structures that assemble in different ways. In each of the monomer subunits, the $\beta_2 m_{63–69}$ and template strands hydrogen bond together to form a β-sheet. The β-sheets have a strong right-handed twist of about 13–22 degrees per residue along the β-strand axis, and thus mimic twisted β-hairpins.

The β-hairpins are fully hydrogen bonded, except between Glu_{69} and Lys_1, in which the hydroxyl group of Thr_{68} can disrupt the hydrogen bonding between these two residues (Figure 5.2A–D). To probe the effect of the hydroxyl group on β-hairpin structure and oligomer formation, I prepared a homologue of peptide **1a** with Val in place of Thr_{68} (peptide $\mathbf{1a_{T68V}}$). The X-ray crystallographic structure of this homologue shows a fully hydrogen-bonded β-hairpin (Figure 5.2E and F) and no appreciable difference in the structure of the oligomers that form, which are hexamers in both cases (Figure 5.3).

Hexamer

Peptide **1a** crystallizes from 0.1 M Tris buffer at pH 8.0 with 0.3 M Li_2SO_4 and 45% PEG 400, in the $P4_222$ space group, with three nearly identical β-hairpin monomers in the asymmetric unit (ASU). Expanding the ASU to generate the lattice shows hexamers composed of six

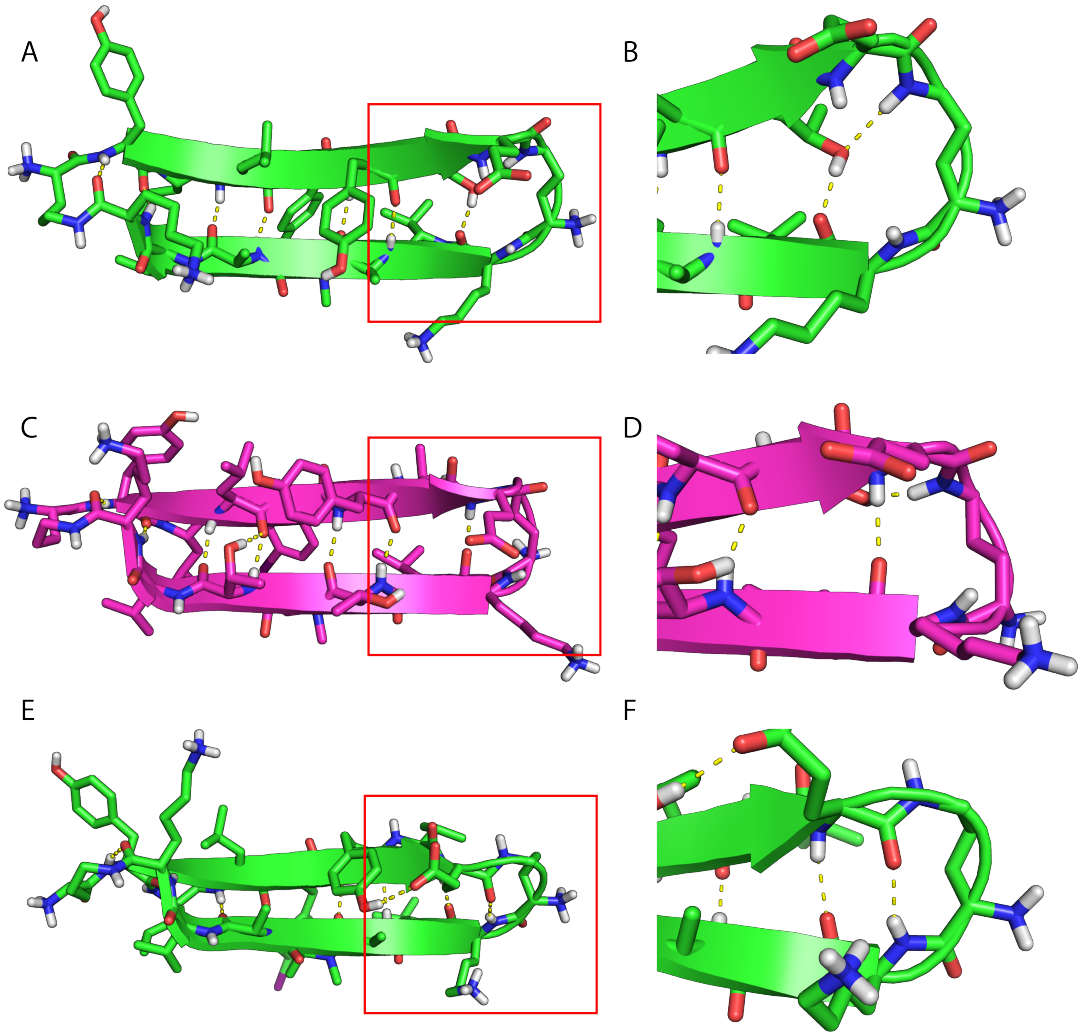


Figure 5.2: X-ray crystallographic structure of β-hairpins formed by peptides **1a**, **2b** and $\mathbf{1a_{T68V}}$. (A) β-Hairpin formed by peptide **1a**. (B) Detail showing the hydroxyl group of Thr_{68} hydrogen bonding with the carbonyl of the adjacent Lys_1 residue. (C) β-Hairpin formed by peptide **2b**. (D) Detail showing the hydroxyl group of Thr_{68} hydrogen bonding with the NH of ornithine. (E) β-Hairpin formed by peptide $\mathbf{1a_{T68V}}$. (F) Detail showing the hydrogen bonding between residues Glu_{68} and Lys_1.

β-hairpins assembled as a trimer of dimers (Figure 5.4). In each dimer, two β-hairpins come together through edge-to-edge interactions between the β_2m strands to form a four-stranded antiparallel β-sheet (Figure 5.4B). Residues Leu_{64}, Phe^{I}_{66}, and Thr_{68} of one monomer form hydrogen-bonded pairs with residues Thr_{68}, Phe^{I}_{66}, and Leu_{64} of the other monomer. Hydrophobic contacts between the side chains of residues Tyr_{63}, Leu_{65}, and Glu_{69}, appear to further stabilize the dimer.

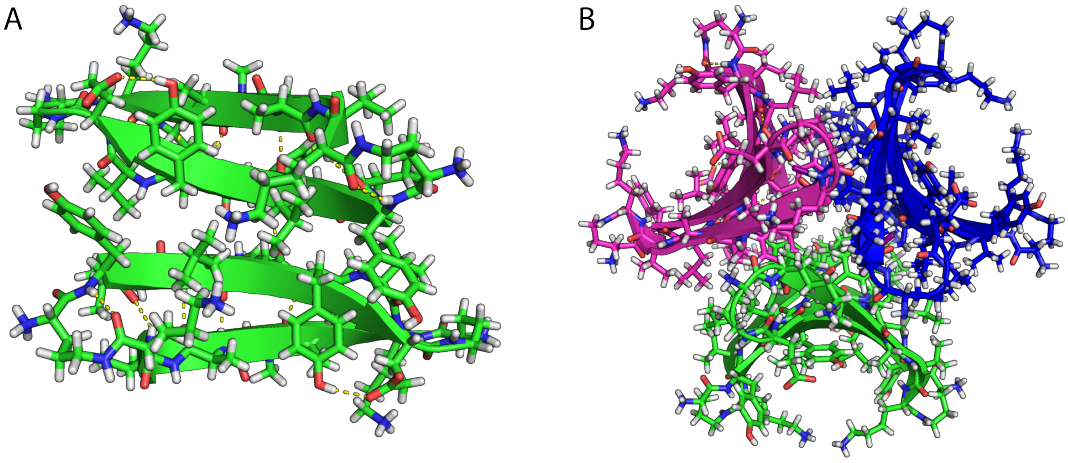


Figure 5.3: X-ray crystallographic structure of peptide $\mathbf{1a_{T68V}}$. (A) Antiparallel β-sheet dimer. (B) Hexamer, top view.

Three antiparallel β-sheet dimers come together through face-to-face interactions around a central 3-fold axis to form the hexamer (Figure 5.4C and Figure 5.4D). The minor surfaces of the β-hairpins face inward to form the hydrophobic core of the hexamer, while the major surfaces face outward and are exposed to solvent within the lattice. The six hydrophobic $\text{Phe}^{\text{I}}_{66}$ residues comprise the center of the hydrophobic core, stacking in pairs, and forming additional hydrophobic contacts among the edges of the aromatic rings. Residues Leu_{64}, Val_2, *N*-Me Ala_4, and Val_6 of the minor faces surround the iodophenyl groups and complete the hydrophobic core (Figure 5.4E and Figure 5.4F).

Peptides **2a** and $\mathbf{1a_{T68V}}$ also crystallize as hexamers from conditions similar to peptide **1a**, but in the *R*32 space group. The ASU of peptide **2a** contains seven β-hairpin monomers; the ASU of peptide $\mathbf{1a_{T68V}}$ contains only one. The hexamers formed by peptides **2a** and $\mathbf{1a_{T68V}}$ are nearly identical to those formed by peptide **1a**.

Octamer

Peptide **2b** crystallizes from 0.1 M SPG (succinic acid-phosphate-glycine) buffer at pH 10.0 and 35% PEG 1500, in the $P4_32_12$ space group, with 12 β-hairpin monomers in the ASU.

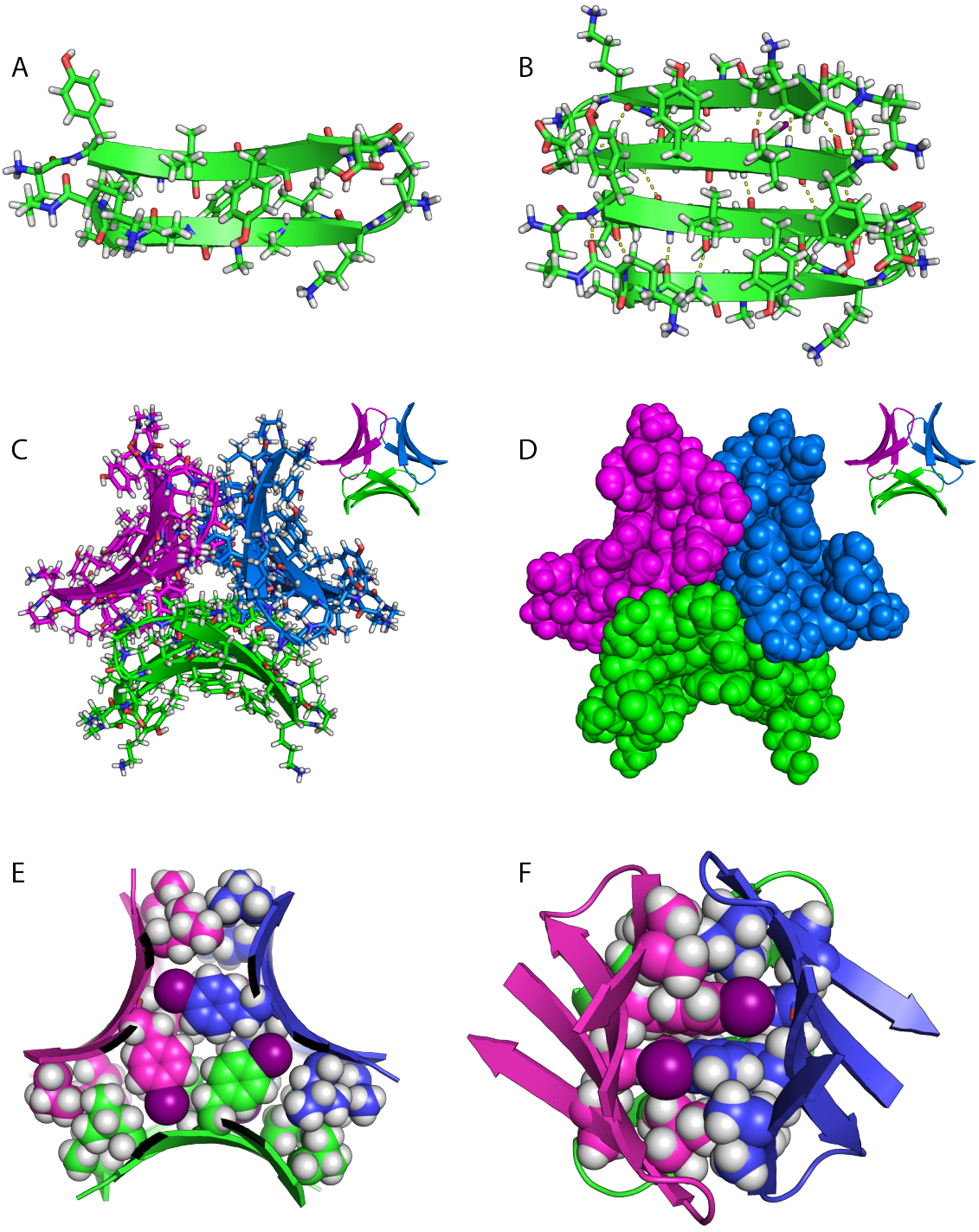


Figure 5.4: X-ray crystallographic structure of peptide **1a** (hexamer). (A) β-Hairpin monomer. (B) Antiparallel β-sheet dimer. (C) Hexamer, top view (cartoon and sticks). (D) Hexamer, top view (spheres). (E) Hydrophobic core, top view (Val_2 and *N*-Me Ala_4 omitted). (F) Hydrophobic core, side view (Val_6 omitted).

Expanding the ASU to generate the lattice shows octamers composed of eight β-hairpins assembled as a tetramer of dimers (Figure 5.5). In each dimer, two β-hairpins associate along the residues of the β_2m strand and interact through face-to-face contacts to form a facial dimer (Figure 5.5B). The two β-hairpins are oriented in an antiparallel fashion, like those in the hexamer, but interact through hydrophobic contacts among the side chains of Leu_{64}, Phe^{I}_{66}, Val_2, and *N*-Me Val_4, rather than through hydrogen bonding between the main chains of the β-hairpins.

Four facial dimers associate around a 4-fold axis to form the octamer. The minor surfaces of the β-hairpins face inward to form the hydrophobic core of the octamer, while the major surfaces face outward and are exposed to solvent within the lattice (Figure 5.5C and Figure 5.5D). The Phe^{I}_{66} pairs of the facial dimers comprise the center of the octamer. Residues Leu_{64}, Val_2, *N*-Me Val_4, and Val_6 of the minor surfaces make up the rest of the hydrophobic core, with Val_2 and *N*-Me Val_4 surrounding the iodophenyl groups and residues Leu_{64} and Val_6 packing in layers above and below the iodophenyl groups (Figure 5.5E and Figure 5.5F). Salt-bridges between Lys_1 and Glu_{69} residues and a network of hydrogen bonds between the edges of the β-hairpins of the four dimer subunits further stabilize the octamer.

Dodecamer

Peptide **1b** crystallizes from 0.1 M Tris buffer at pH 8.0 and 1.5 M $(NH_4)_2SO_4$, in the $P3_121$ space group, with 12 β-hairpins in the ASU. Expanding the ASU to generate the crystal lattice shows dodecamers composed of twelve β-hairpins. The dodecamer contains a dimer of trimers, which makes a central hexamer, and three additional pairs of β-hairpins, which surround the central hexamer (Figure 5.6).

In the trimer, three β-hairpins associate in a triangular fashion through the edges of the β_2m strands (Figure 5.6B). The β_2m strands hydrogen bond together at the corners of the

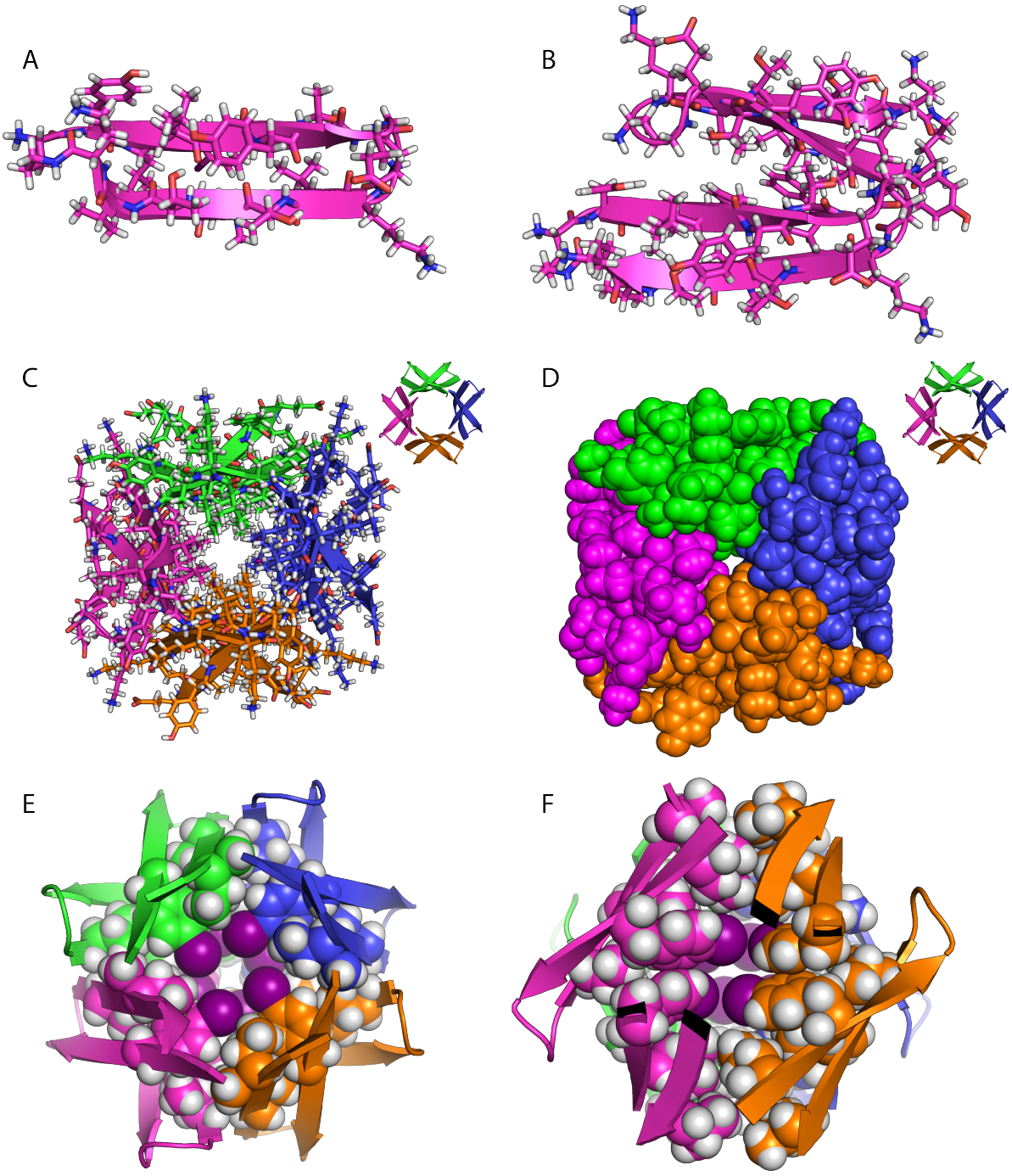


Figure 5.5: X-ray crystallographic structure of peptide **2b** (octamer). (A) β-Hairpin monomer. (B) Facial dimer. (C) Octamer, top view (cartoon and sticks). (D) Octamer, top view (spheres). (E) Hydrophobic core, top view (Leu_{64} and Val_6 omitted). (F) Hydrophobic core, side view (Val_2 omitted).

triangles, with Leu_{64} and the proximal ornithine hydrogen bonding with Phe^{I}_{66} and Thr_{68} at each corner. Three ordered waters fill the hole in the center of the triangle and form additional hydrogen bonds with Leu_{64} and Phe^{I}_{66}, thus creating a network that satisfies all of the hydrogen-bonding valences of the β_2m strands. Hydrophobic contacts among the side-chains of the β-hairpins further stabilize the trimer structure.

Two triangular trimers come together through face-to-face interactions to form the central hexamer within the dodecamer. The minor surfaces of the β-hairpins face inward and contribute to the hydrophobic core of the dodecamer, while the major surfaces face outward and are exposed to solvent within the lattice. The Phe^{I}_{66} residues of the opposing trimers stack in the center of the hydrophobic core, and the hydrophobic side chains of Leu_{64}, Val_2, *N*-Me Val_4, and Val_6 surround the iodophenyl groups. The minor surfaces of the three pairs of β-hairpins that surround the central hexamer face inward and extend the hydrophobic core through additional hydrophobic contacts. These pairs of β-hairpins do not hydrogen bond to each other, but are stabilized by hydrophobic contacts with the hexamer through the residues of the minor surfaces (Figure 5.6E and Figure 5.6F).

Peptide **1c** also crystallizes as dodecamers, but from 0.1 M Tris at pH 7.5 with 0.2 M Li_2SO_4 and 25% PEG 400, in the $P4_122$ space group, with 12 β-hairpins forming a dodecamer in the ASU. The dodecamers formed by peptide **1c** are nearly identical to those formed by peptide **1b**.

A Mixed Dodecamer

I cocrystallized peptides **1a** and **1c** to ask what would happen when peptides that formed two different oligomers (hexamers and dodecamers) were allowed to crystallize from a 1:1 mixture. Much to my surprise, the two peptides cocrystallized as a dodecamer similar to that of peptide **1c**, but with peptide **1c** forming the central hexamer and peptide **1a** forming

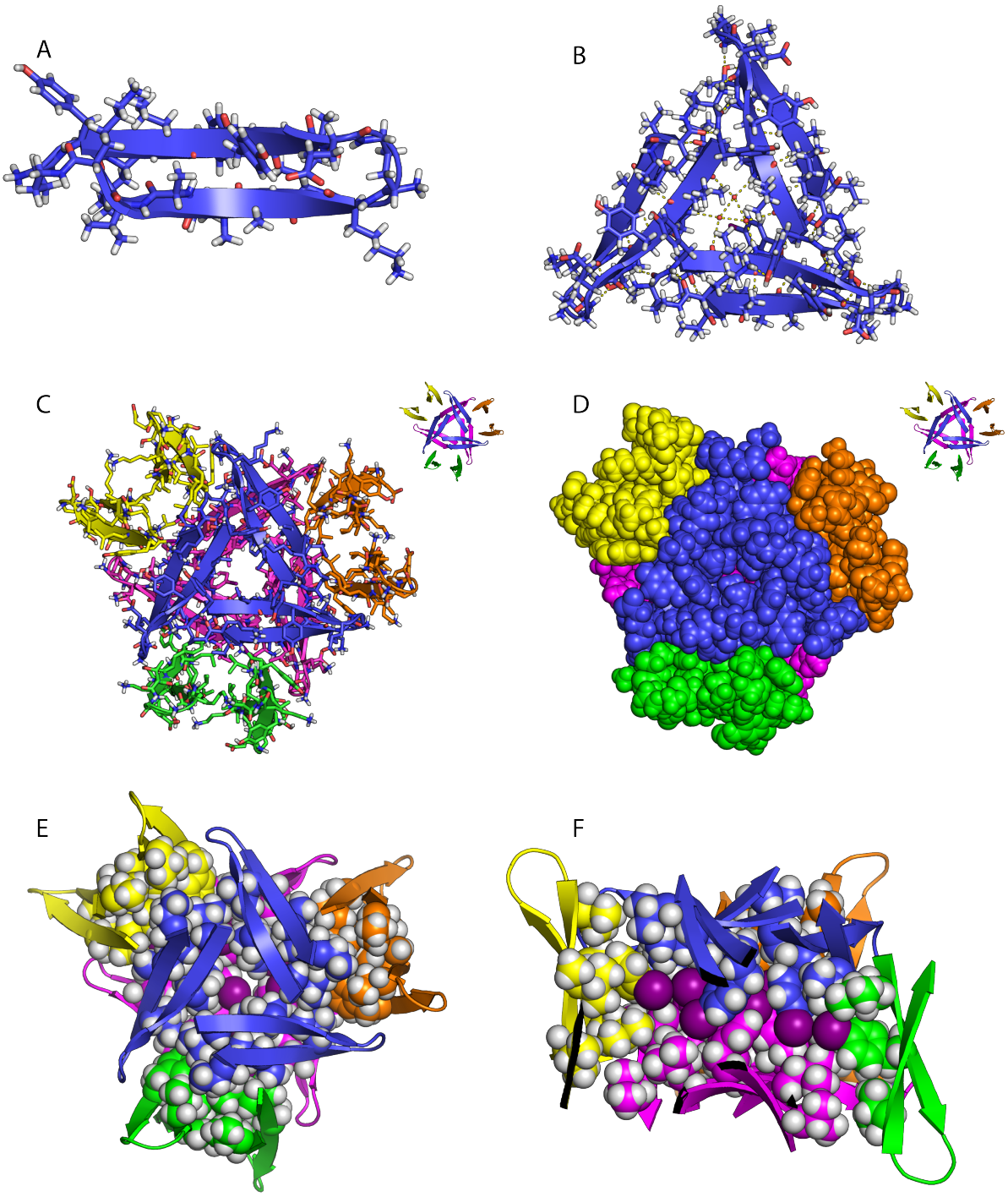


Figure 5.6: X-ray crystallographic structure of peptide **1b** (dodecamer). (A) β-Hairpin monomer. (B) Triangular trimer. (C) Dodecamer, top view (cartoon and sticks). (D) Dodecamer, top view (spheres). (E) Hydrophobic core, top view. (F) Hydrophobic core, side view (cutaway).

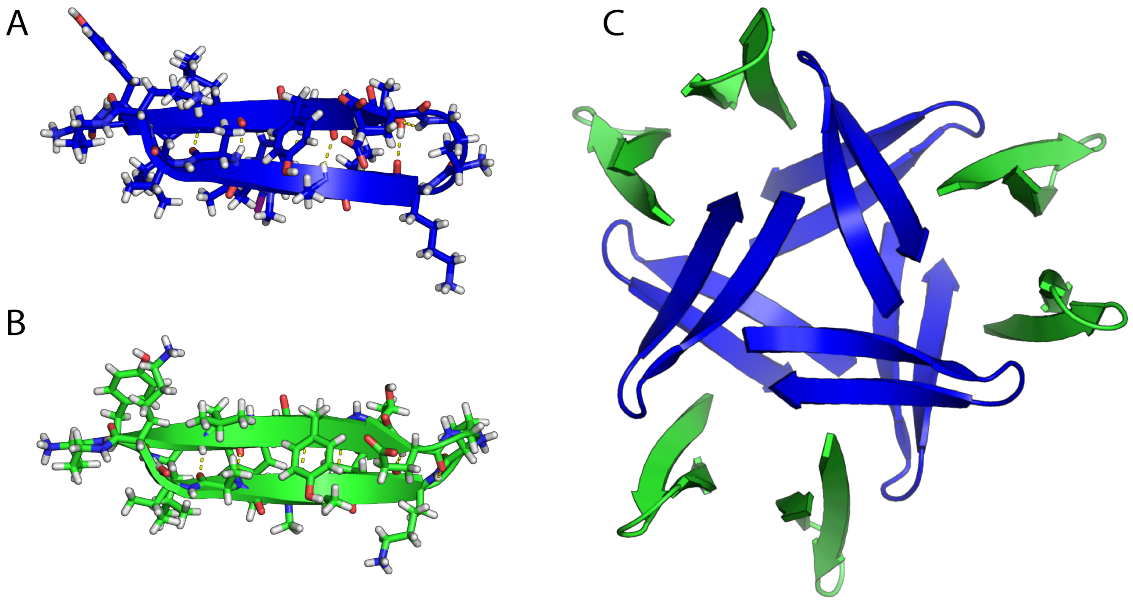


Figure 5.7: X-ray crystallographic structure of a mixed dodecamer formed by peptides **1a** and **1c**. (A) β-Hairpin formed by peptide **1c**. (B) β-Hairpin formed by peptide **1a**. (C) Mixed dodecamer with six β-hairpins of peptide **1c** (blue) forming the central hexamer (dimer of trimers) and three pairs of β-hairpins of peptide **1a** (green) surrounding the central hexamer.

the three pairs of β-hairpins surrounding the hexamer (Figure 5.7). Peptides **1a** and **1c** cocrystallize under conditions similar to those from which **1a** and **1c** crystallize individually: 0.1 M Tris buffer at pH 7.5 with 0.2 M Li_2SO_4 and 30% PEG 400, in the $P4_122$ space group, with 12 β-hairpins forming a dodecamer in the asymmetric unit (ASU). The formation of a mixed dodecamer from peptides with propensities to form different oligomers demonstrates that the oligomers formed by one peptide may alter the oligomerization of another peptide.

Size-Exclusion Chromatography of Peptides 1 and 2.

Size-exclusion chromatography (SEC) studies indicate that peptides **1** and **2** form oligomers in solution. SEC was performed on 1 mM solutions of peptides **1** and **2** in 0.1 M phosphate buffer at pH 7.4 with a Superdex 200 column. The elution profiles were compared to those of size standards vitamin B12, ribonuclease A, and chymotrypsinogen. These 1.3, 13.7, and 25.6 kDa size standards eluted at 20.3, 17.3, and 16.5 mL, respectively. Peptides **1b**–**1e** and

2a–**2e** elute between 17.3 to 18.0 mL (Table 5.2; Figure 5.10B–Figure 5.19B). (Peptide **1a** precipitates from phosphate buffer; SEC of the supernatant gives a weak signal and slightly larger elution volume.) These volumes are substantially lower than would be expected for the corresponding 2.0 kDa monomers. The elution volumes of peptides **1b**–**1e** and **2a**–**2e** are similar to that of ribonuclease A. These volumes are consistent with oligomers in the hexamer to octamer size range, for both the peptides that crystallize and those that do not. The peak shapes of the peptides are slightly broader than those of the size standards and the peaks tail slightly, reflecting an oligomer-monomer equilibrium in which the oligomer predominates. Peptides **1b** and **1c** do not appear to elute as dodecamers, suggesting that the central hexamer elutes without the three peripheral pairs of β-hairpins observed in the crystal lattice. Figure 5.8 shows the SEC chromatograms of peptides **2a** (hexamer), **2b** (octamer), and **1b** (dodecamer).

Table 5.2: SEC Elution Volumes, MTT Conversion (%), and LDH Release (%) of Peptides **1** and **2**.

peptide	SEC (mL)	MTT (%)	LDH(%)	crystallographic oligomer
1a	18.8[a]	53±5	37±4	hexamer
1b	18.0	85±8	22±5	dodecamer
1c	17.6	54±15	45±6	dodecamer
1d	17.7	110±8	18±2	–
1e	17.6	42±5	44±7	–
2a	17.4	101±5	5±1	hexamer
2b	17.3	125±9	14±2	octamer
2c	17.8	34±3	38±5	–
2d	17.5	111±8	20±2	–
2e	17.3	49±4	36±4	–
Vitamin B12	20.3			
ribonuclease A	17.3			
chymotrypsinogen	16.5			

[a]Peptide **1a** precipitated from phosphate buffer solution. SEC was performed on the supernatent after removal of the precipitate by centrifugation. The SEC data for **1a** thus can not be compared directly to the SEC data of the other peptides.

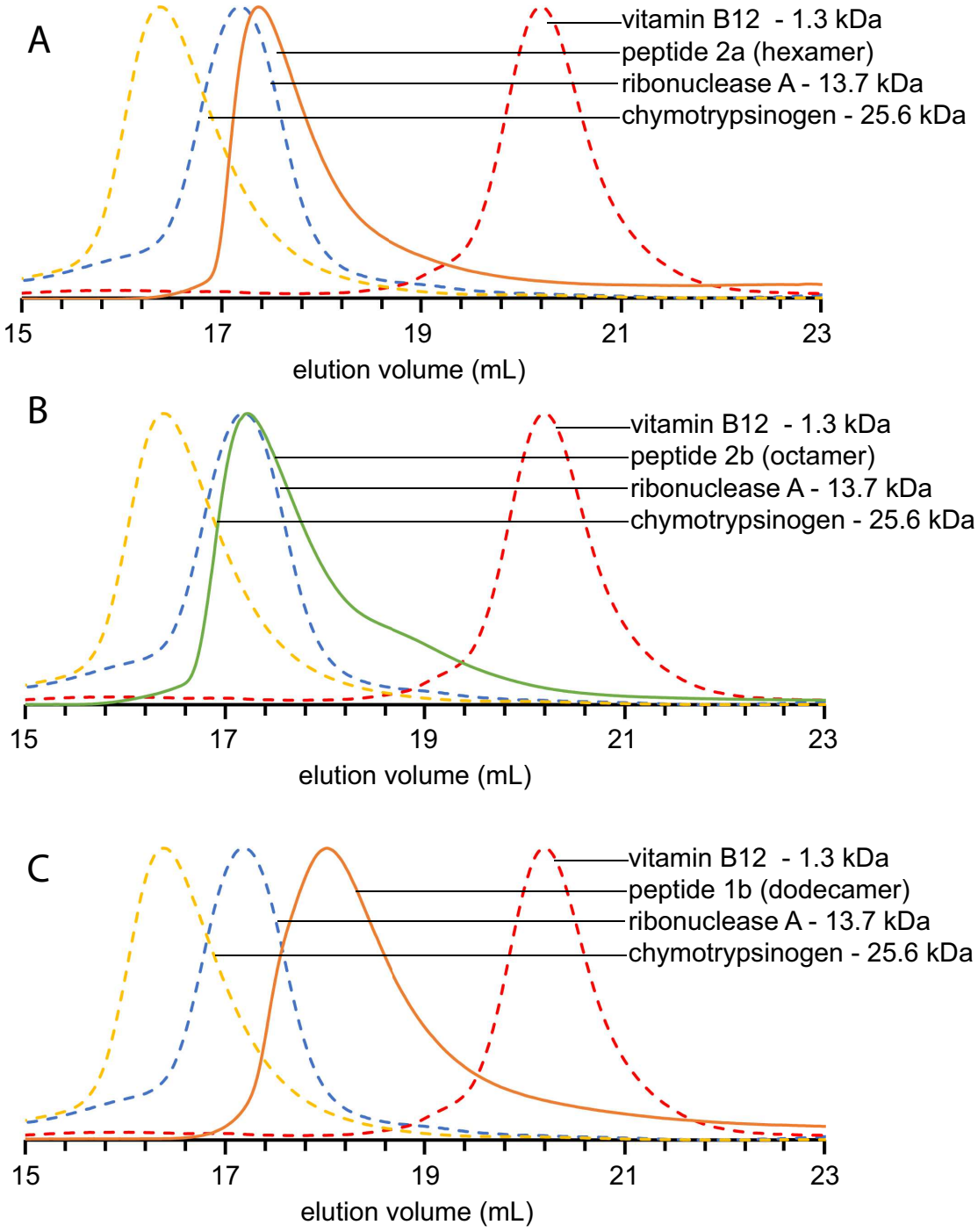


Figure 5.8: (A) SEC chromatograms of peptide **2a** (orange), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (B) SEC chromatograms of peptide **2b** (green), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (C) SEC chromatograms of peptide **1c** (orange), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red).

Cytotoxicity of Peptides 1 and 2.

In collaboration with graduate student Adam Kreutzer, I studied the cytotoxicity of peptides **1** and **2** using MTT conversion and lactate dehydrogenase (LDH) release assays in the neuroblastoma cell line SH-SY5Y. At 25 μM, peptides **1** and **2** showed a range of toxicities, with peptides **1a**, **1c**, **1e**, **2c**, and **2e** being more toxic and peptides **1b**, **1d**, **2a**, **2b**, and **2d** being less toxic (Table 5.2; Figure 5.9). These differences do not correlate with the crystallographic observation of oligomers. Two of the peptides that crystallize are more toxic (**1a** and **1c**), while three that crystallize are less toxic (**1b**, **2a**, and **2b**). Three of the peptides that do not crystallize are more toxic (**1e**, **2c**, and **2e**), while two that do not crystallize are less toxic (**1d** and **2d**). The differences in toxicity do not correlate with oligomer structure. One peptide that crystallizes as a hexamer (**1a**) is more toxic, while the other peptide that crystallizes as a hexamer (**2a**) is less toxic. One peptide that crystallizes as a dodecamer (**1c**) is more toxic, while the other peptide that crystallizes as a dodecamer (**1b**) is less toxic. The peptide that crystallizes as an octamer (**2b**) is less toxic. The differences also do not correlate with the hydrophobicity of the peptides. In the series with Ala at positions R_3 and R_5, the less hydrophobic peptide **1a** (R_4 = Ala) and more hydrophobic peptides **1c** and **1e** (R_4 = Leu and Nle) are more toxic. Peptide **1d** (R_4 = Ile) is comparable in hydrophobicity to **1c** and **1e** but is less toxic. Similar differences are observed in the series with Thr at positions R_3 and R_5, with **2c** and **2e** (R_4 = Leu and Nle) being more toxic, and **2a**, **2b**, and **2d** (R_4 = Ala, Val, and Ile) being less toxic. The differences in the observed toxicities may reflect differences in propensities of the peptides to form oligomers at concentrations 1-2 orders of magnitude lower than those used for crystallization and SEC studies.

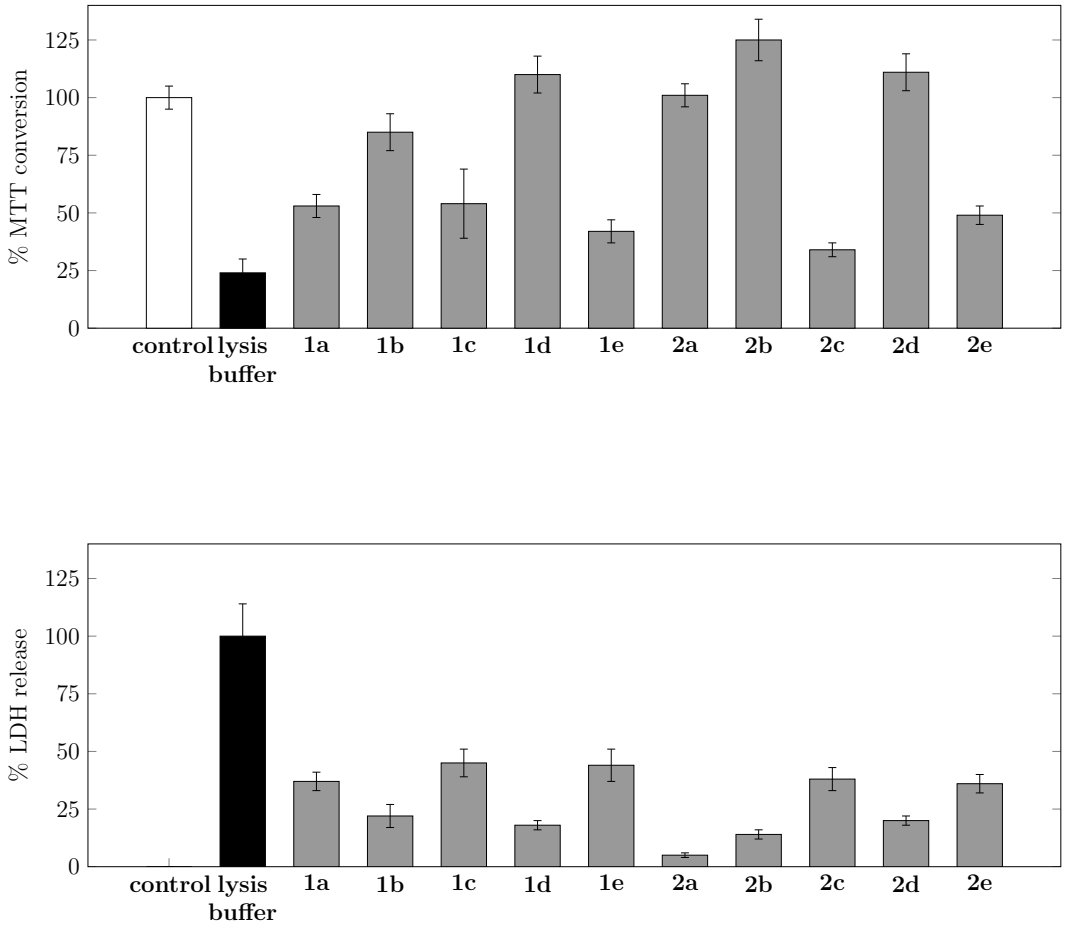


Figure 5.9: MTT and LDH assays of peptides **1** and **2** with SH-SY5Y cells. Cells were treated with 25 µM of peptides **1** or **2** for 24 hours. The percentage MTT conversion relative to untreated cells was determined spectrophotometrically (top). The percentage of LDH release was determined as described in the text of the Supporting Information (bottom).

Discussion

The seven X-ray crystallographic structures presented here illustrate a range of oligomers that a single amyloidogenic peptide sequence can form. Although the hexamers, octamers, and dodecamers differ in size and topology, they share a common theme of a hydrophobic core formed by hydrophobic side chains of β-hairpins, which form dimer and trimer subunits within the oligomers. The hexamers and octamers formed by peptides **1a**, $\mathbf{1a_{T68V}}$, **2a**, and **2b** comprise three or four dimers packed around a central hydrophobic core and thus resemble the hexamer derived from PrP.[31] The dodecamers formed by peptides **1b** and **1c**

and by a mixture of **1a** and **1c** differ substantially from the hexamers and octamers, because they are based on a pair of triangular trimers surrounded by three pairs of β-hairpins. I have previously observed similar triangular trimers in the X-ray crystallographic structures of peptides derived from $A\beta_{17\text{-}36}$.[32] Based on these observations, I now believe that the formation of higher-order oligomers from these types of dimer and trimer building blocks is a common feature of many amyloidogenic peptides and proteins.

Small differences in peptide sequence can lead to large differences in oligomer structure. The β-hairpins formed by peptides **1** and **2** are nearly identical to each other, yet they arrange in various alignments to form three families of oligomers — hexamers, octamers, and dodecamers. There is little obvious relationship between the hydrophobicity and size of residues R_3–R_5 and the oligomers that form. Peptides with both hydrophobic (Ala) and hydrophilic (Thr) residues at R_3 and R_5 (**1a** and **2a**) permit hexamer formation. Either dodecamer or octamer (**1b** or **2b**) form when R_4 is increased in size (Ala to Val). Further increasing the size of R_4 (Val to Leu; **1b** to **1c**) does not alter dodecamer formation. Other changes in R_3 and R_5 (**1c** to **2c**) or R_4 (Ile, Nle; **1d**, **1e**, **2d**, **2e**) give peptides that do not crystallize. Although I do not yet understand the relationship between the residue hydrophobicities and sizes and the oligomer structures, it is clear that the R_4 residue is important for oligomer formation. Increasing the size of the side chain at the R_4 position may change the packing of the hydrophobic core and thus change which oligomer forms.

The changes from hexamer to octamer to dodecamer that occur upon mutating a single residue may provide insights into the effects of familial mutations in amyloid diseases. Changing an alanine in peptide **1a** to valine or leucine in peptides **1b** and **1c** changes a hexamer to a dodecamer; changing an alanine in peptide **2a** to valine in peptide **2b** changes a hexamer to an octamer. These changes are similar to the point mutations that dictate early onset in Alzheimer's disease and in the synucleinopathies related to Parkinson's disease.[38,39] It is quite possible that the mutant Aβ peptides and α-synuclein protein associated

with these heritable diseases also form different oligomers than those formed by the non-mutant wild types, and that these differences in oligomer structure may alter the toxicity of the oligomers. The formation of the a mixed dodecamer from peptides **1a** and **1c** is especially intriguing, because it demonstrates that a mutant peptide or protein can dictate the structure of the oligomers that form. Similar effects may occur in individuals with a single allele for a familial mutation, and the resulting mutant peptide or protein may recruit the wild-type peptide or protein to form mixed oligomers with different oligomerization states and more toxic structures.

Conclusion

Macrocyclic peptides that mimic β-hairpins and contain an amyloidogenic peptide sequence and an *N*-methyl amino acid are valuable for exploring the structure and assembly of amyloid oligomers. These peptides are easy to synthesize and are often easy to crystallize. X-ray crystallography readily reveals the structures of the oligomers to consist of dimer and trimer subunits that assemble to create a hydrophobic core. These common structural features should also occur in the oligomers formed by full-length amyloidogenic peptides and proteins. Although the studies described here use an amyloidogenic peptide sequence from β_2m, the modes of oligomer assembly observed likely transcend individual peptide sequences and represent some of the structural diversity among amyloid oligomers.

The three families of oligomers observed — hexamers, octamers, and dodecamers — illustrate some of the polymorphism of amyloid oligomers and highlight the impact that a single mutation can have on oligomer structure. The formation of the mixed dodecamer illustrates the potential of one amyloidogenic peptide or protein to dictate oligomer formation by another. This observation may have important implications for the role of heritable mutations in familial amyloid diseases and may also be relevant to interactions among different

amyloidogenic peptides and proteins, such as Aβ, tau, α-synuclein, and IAPP, in amyloid diseases.

References

1. Walsh, D. M.; Klyubin, I.; Fadeeva, J. V.; Cullen, W. K.; Anwyl, R.; Wolfe, M. S.; Rowan, M. J.; Selkoe, D. J. *Nature* **2002**, *416*, 535–539.

2. Cleary, J. P.; Walsh, D. M.; Hofmeister, J. J.; Shankar, G. M.; Kuskowski, M. A.; Selkoe, D. J.; Ashe, K. H. *Nat. Neurosci.* **2005**, *8*, 79–84.

3. Townsend, M.; Shankar, G. M.; Mehta, T.; Walsh, D. M.; Selkoe, D. J. *J. Physiol.* **2006**, *572*, 477–492.

4. Lesn, S.; Koh, M. T.; Kotilinek, L.; Kayed, R.; Glabe, C. G.; Yang, A.; Gallagher, M.; Ashe, K. H. *Nature* **2006**, *440*, 352–357.

5. Zhao, W.-Q.; De Felice, F. G.; Fernandez, S.; Chen, H.; Lambert, M. P.; Quon, M. J.; Krafft, G. A.; Klein, W. L. *FASEB J* **2008**, *22*, 246–260.

6. Haass, C.; Selkoe, D. J. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 101–112.

7. Caughey, B.; Lansbury, P. T. *Annu. Rev. Neurosci.* **2003**, *26*, 267–298.

8. Winner, B. *et al. Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 4194–4199.

9. Lin, C. Y.; Gurlo, T.; Kayed, R.; Butler, A. E.; Haataja, L.; Glabe, C. G.; Butler, P. C. *Diabetes* **2007**, *56*, 1324–1332.

10. Haataja, L.; Gurlo, T.; Huang, C. J.; Butler, P. C. *Endocr. Rev.* **2008**, *29*, 303–316.

11. Gurlo, T.; Ryazantsev, S.; Huang, C. J.; Yeh, M. W.; Reber, H. A.; Hines, O. J.; O'Brien, T. D.; Glabe, C. G.; Butler, P. C. *Am. J. Pathol.* **2010**, *176*, 861–869.

12. Eakin, C. M.; Attenello, F. J.; Morgan, C. J.; Miranker, A. D. *Biochemistry* **2004**, *43*, 7808–7815.

13. Mustata, M.; Capone, R.; Jang, H.; Arce, F. T.; Ramachandran, S.; Lal, R.; Nussinov, R. *J. Am. Chem. Soc.* **2009**, *131*, 14938–14945.

14. Smith, D. P.; Radford, S. E.; Ashcroft, A. E. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 6794–6798.

15. Chromy, B. A.; Nowak, R. J.; Lambert, M. P.; Viola, K. L.; Chang, L.; Velasco, P. T.; Jones, B. W.; Fernandez, S. J.; Lacor, P. N.; Horowitz, P.; Finch, C. E.; Krafft, G. A.; Klein, W. L. *Biochemistry* **2003**, *42*, 12749–12760.

16. Ahmed, M.; Davis, J.; Aucoin, D.; Sato, T.; Ahuja, S.; Aimoto, S.; Elliott, J. I.; Van Nostrand, W. E.; Smith, S. O. *Nat. Struct. Mol. Biol.* **2010**, *17*, 561–567.

17. Bernstein, S. L.; Dupuis, N. F.; Lazo, N. D.; Wyttenbach, T.; Condron, M. M.; Bitan, G.; Teplow, D. B.; Shea, J. E.; Ruotolo, B. T.; Robinson, C. V.; Bowers, M. T. *Nat. Chem.* **2009**, *1*, 326–331.

18. Cerf, E.; Sarroukh, R.; Tamamizu-Kato, S.; Breydo, L.; Derclaye, S.; Dufrne, Y. F.; Narayanaswami, V.; Goormaghtigh, E.; Ruysschaert, J. M.; Raussens, V. *Biochem. J.* **2009**, *421*, 415–423.

19. Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T. *Nat. Chem.* **2011**, *3*, 172–177.

20. Celej, M. S.; Sarroukh, R.; Goormaghtigh, E.; Fidelio, G. D.; Ruysschaert, J. M.; Raussens, V. *Biochem. J.* **2012**, *443*, 719–726.

21. Buchanan, L. E.; Dunkelberger, E. B.; Tran, H. Q.; Cheng, P. N.; Chiu, C. C.; Cao, P.; Raleigh, D. P.; de Pablo, J. J.; Nowick, J. S.; Zanni, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 19285–19290.

22. Petkova, A. T.; Yau, W.-M.; Tycko, R. *Biochemistry* **2006**, *45*, 498–512.

23. Paravastu, A. K.; Leapman, R. D.; Yau, W. M.; Tycko, R. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 18349–18354.

24. Lu, J.-X.; Qiang, W.; Yau, W.-M.; Schwieters, C. D.; Meredith, S. C.; Tycko, R. *Cell* **2013**, *154*, 1257–1268.

25. Lührs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Dbeli, H.; Schubert, D.; Riek, R. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17342–17347.

26. Iwata, K.; Fujiwara, T.; Matsuki, Y.; Akutsu, H.; Takahashi, S.; Naiki, H.; Goto, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 18119–18124.

27. Jahn, T. R.; Makin, O. S.; Morris, K. L.; Marshall, K. E.; Tian, P.; Sikorski, P.; Serpell, L. C. *J. Mol. Biol.* **2010**, *395*, 717–727.

28. Pawar, A. P.; Dubay, K. F.; Zurdo, J.; Chiti, F.; Vendruscolo, M.; Dobson, C. M. *J. Mol. Biol.* **2005**, *350*, 379–392.

29. Tzotzos, S.; Doig, A. J. *Protein Science* **2010**, *19*, 327–348.

30. Laganowsky, A.; Liu, C.; Sawaya, M. R.; Whitelegge, J. P.; Park, J.; Zhao, M.; Pensalfini, A.; Soriaga, A. B.; Landau, M.; Teng, P. K.; Cascio, D.; Glabe, C.; Eisenberg, D. *Science* **2012**, *335*, 1228–1231.

31. Apostol, M. I.; Perry, K.; Surewicz, W. K. *J. Am. Chem. Soc.* **2013**, *135*, 10202–10205.

32. Spencer, R. K.; Li, H.; Nowick, J. S. *J. Am. Chem. Soc.* **2014**, *136*, 5595–5598.

33. Trinh, C. H.; Smith, D. P.; Kalverda, A. P.; Phillips, S. E.; Radford, S. E. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9771–9776.

34. Jones, S.; Manning, J.; Kad, N. M.; Radford, S. E. *J. Mol. Biol.* **2003**, *325*, 249–257.

35. Ivanova, M. I.; Thompson, M. J.; Eisenberg, D. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 4079–4082.

36. Spencer, R.; Chen, K. H.; Manuel, G.; Nowick, J. S. *Eur. J. Org. Chem.* **2013**, *2013*, 3523–3528.

37. Nowick, J. S.; Brower, J. O. *J. Am. Chem. Soc.* **2003**, *125*, 876–877.

38. Karran, E.; Mercken, M.; De Strooper, B. *Nat Rev Drug Discov* **2011**, *10*, 698–712.

39. Bendor, J. T.; Logan, T. P.; Edwards, R. H. *Neuron* **2013**, *79*, 1044–1066.

Materials and Methods

Synthesis of peptides 1 and 2.

Peptides **1** and **2** were prepared by procedures similar to those that we have described previously and were isolated as the corresponding trifluoroacetate (TFA) salts.[1,2] Peptides **1** and **2** were synthesized by Fmoc-based solid-phase peptide synthesis, solution-phase cyclization, deprotection, and purification by RP-HPLC.

Representative synthesis of peptide 1a

Loading of the resin. 2-Chlorotrityl chloride resin (300 mg, 1.2 mmol/g) was added to a Bio-Rad Poly-Prep chromatography column (10 mL). The resin was suspended in dry CH_2Cl_2 (10 mL) and allowed to swell for 30 min. The solution was drained from the resin and a solution of Boc-Orn(Fmoc)-OH (0.50 equiv., 82 mg, 0.18 mmol) in 20% 2,4,6-collidine in dry CH_2Cl_2 (5 mL) was added immediately and the mixture was gently agitated for 12 h. The solution was then drained and a mixture of CH_2Cl_2/MeOH/DIPEA (17:2:1, 10 mL) was added immediately. The mixture was gently agitated for 1 h to cap the unreacted 2-chlorotrityl chloride sites. The resin was then washed with dry CH_2Cl_2 (2×) and dried by passing nitrogen through the vessel. In the synthesis of peptide **1a**, the resin loading was determined to be 0.14 mmol [0.45 mmol/g, 75% based on Boc-Orn(Fmoc)-OH] by UV analysis of the Fmoc cleavage product. Loadings of 0.12–0.15 mmol [65–80%, based on Boc-Orn(Fmoc)-OH] were typically observed in various repetitions of this procedure associated with the syntheses of peptides **1** and **2**.

Peptide coupling. The PS-2-chlorotrityl-Orn(Fmoc)-Boc generated from the previous step was transferred to a solid-phase peptide synthesizer reaction vessel and submitted to cycles

of automated peptide coupling with Fmoc-protected amino acid building blocks. The linear peptide was synthesized from the C-terminus to the N-terminus starting at Lys_7. Each coupling consisted of: *i.* Fmoc-deprotection with 20% piperidine in DMF for 3 min, *ii.* washing with DMF (3×), *iii.* coupling of the amino acid (0.56 mmol, 4 equiv.) in the presence of HCTU (224 mg, 0.56 mmol, 4 equiv.), and *iv.* washing with DMF (6×). Each amino acid coupling step took 20 min. Special coupling conditions were used for the amino acids (R_3) that followed the *N*-methyl amino acids (R_4): The amino acids after the *N*-methyl-L-alanine, *N*-methyl-L-leucine, and *N*-methyl-L-norleucine were double coupled (0.56 mmol, 4 equiv.) and allowed to react for 1 h per coupling with HATU (4 equiv.) and HOAt (4 equiv.); the amino acids after *N*-methyl-L-valine and *N*-methyl-L-isoleucine were double coupled (0.56 mmol, 4 equiv.) and allowed to react for 6 h per coupling with HATU (4 equiv.) and HOAt (4 equiv.). The remaining residues were coupled as described in steps *i*–*iv*. After the last amino acid was coupled (Tyr_{63}), the terminal Fmoc group was removed with 20% piperidine in DMF. The resin was transferred from the reaction vessel of the peptide synthesizer to a Bio-Rad Poly-Prep chromatography column.

Cleavage of the peptide from the resin. The linear peptide was cleaved from the resin by agitating the resin for 1 hr with a 1:4 mixture of hexafluoroisopropanol (HFIP) and CH_2Cl_2 (5 mL). The suspension was filtered and the filtrate was collected in a 250 mL round-bottomed flask. The resin was washed with additional HFIP-CH_2Cl_2 mixture (5 mL) and then with CH_2Cl_2 (2 × 10 mL). The combined filtrates were concentrated by rotary evaporation to give a white solid. The white solid was further dried under vacuum ($\leq$ 0.1 mmHg) to afford the crude protected linear peptide, which was cyclized without further purification.

Cyclization of the linear peptide. The crude protected linear peptide was dissolved in dry DMF (125 mL). HOBt (95 mg, 0.70 mmol, 5 equiv.) and HBTU (264 mg, 0.70 mmol, 5 equiv.) were added to the solution. The reaction mixture was then stirred under nitrogen for 20 min. DIPEA (0.3 mL, 1.7 mmol, 12 equiv.) was added to the solution and the mixture

was stirred under nitrogen for 24 h. The mixture was concentrated under reduced pressure to afford the crude protected cyclic peptide.

Global deprotection and purification of the cyclic peptide. The protected cyclic peptide was dissolved in TFA/triisopropylsilane (TIPS)/H_2O (18:1:1, 10 mL) in a 250 mL round-bottomed flask equipped with a nitrogen-inlet adaptor. The solution was stirred for 1.5 h. The reaction mixture was then concentrated by rotary evaporation under reduced pressure to afford the deprotected cyclic peptide as a yellow oil. The oil was dissolved in 4:1 mixture of H_2O and acetonitrile (5 mL) and the solution was filtered through a 0.20 µm syringe filter and purified by reversed-phase HPLC (gradient elution with 20–50% CH_3CN over 40 min). Pure fractions were combined and lyophilized. Peptide **1a** was isolated as 52 mg of a white powder. Peptides **1b**–**1e**, **2a**–**2e**, and $\mathbf{1a_{T68V}}$ were prepared and isolated by similar procedures as white powders: **1b**, 14 mg; **1c** 45 mg; **1d**, 56 mg, **1e**, 44 mg; **2a**, 35 mg; **2b**, 22 mg; **2c**, 60 mg; **2d**, 20 mg; **2e**, 43 mg; $\mathbf{1a_{T68V}}$, 93 mg.

Crystallization procedure for peptides 1 and 2

Standard protein crystallography techniques were used to grow crystals of peptides **1** and **2**. Optimal peptide concentrations for crystallization experiments were determined using the Hampton Pre-Crystallization Test (PCT) kit. Crystallization conditions were screened for peptides **1a**, **1b**, **1d**, **1e**, and **2b**-**2e** using five crystallization kits in a 96-well-plate format (Hampton Index, PEG/Ion, and Crystal Screen; Qiagen JCSG+ and PACT; 480 experiments). Peptides **1a**, **1b**, and **2b** readily formed crystals under a variety of conditions. Peptides **1c** and **2a** grew crystals under conditions similar to peptide **1a** and thus were not subjected to additional screening. Crystallization conditions were further optimized by varying pH and additives in a 24-well-plate format. The crystals grew rapidly, typically in less than 48 hours. Optimized crystal growing conditions for peptides **1a**, **1b**, **1c**, **2a**,

and **2b** are summarized in Table 5.3; growing conditions for peptide $\mathbf{1a_{T68V}}$ and the mixed dodecamer (**1a** + **1c**) are summarized in Table 5.4. Peptides **1d**, **1e**, **2c**, **2d**, and **2e** failed to crystallize under any of the conditions studied.

Representative crystallization of peptide 1a

Crystallization conditions were determined using the hanging-drop vapor-diffusion method. Crystallization was performed in a 96-well format, with each well containing 100 µL of a solution from a 96-well screening kit. Five kits were used (Hampton Index, PEG/Ion, and Crystal Screen; Qiagen JCSG+ and PACT) for a total of 480 experiments (five 96-well plates). Hanging drops were made by combining 300 nL of peptide **1a** solution (10 mg/mL in 18 MΩ deionized water) and 300 nL of the well solution using a TTP LabTech Mosquito crystal screening robot. Crystals grew rapidly (< 48 h) in multiple wells. A solution of 0.2 M lithium sulfate, 0.1 M Tris pH 8.5 and 40% PEG 400 was selected from the JCSG+ crystal screen for further optimization.

Crystallization conditions were optimized using a 4×6 matrix Hampton VDX 24-well plate. The PEG concentration was varied in increments of 2% across the rows and the pH of the Tris buffer was varied in increments of 0.5 pH units down the columns. Three hanging-drops were suspended on each borosilicate glass slide. The drops were prepared by combining 1- or 2-µL aliquots of a solution of peptide **1a** (10 mg/mL) and the well solution in 1:1, 2:1, and 1:2 ratios. Slides were inverted and pressed firmly against the silicone grease surrounding each well. Additional plates were optimized further by varying the lithium sulfate concentration in increments of 0.05 M. Crystals suitable for X-ray crystallography were selected from a well containing 0.1 M Tris at pH 8.0 with 0.3 M Li_2SO_4 and 45% PEG.

Crystallization conditions for peptides **1a**, **1b**, **1c**, **2a**, and **2b** are shown in Table 5.3. Crystal growing conditions for the mixed dodecamer and $\mathbf{1a_{T68V}}$ are shown in Table 5.4.

X-ray diffraction data collection, data processing, and structure determination for peptides 1a, 1b, 1c, 2a, 2b, and 1a + 1c

X-ray diffraction data sets for peptides **1a**, **1b**, **1c**, **2a**, and **2b** were collected at the Advanced Light Source (ALS) at the Lawrence Berkeley National Laboratory (Berkeley, California) on synchrotron beamlines 8.2.1 (peptides **1a**, **1b**, and **2b**) and 8.2.2 (peptides **1c** and **2a**) at 1.0 Å wavelength with 0.5° rotation per image (Table 5.3). X-ray diffraction data for the mixed dodecamer (peptides **1a + 1c**) were collected at the Stanford Synchrotron Radiation Lightsource (SSRL) on synchrotron beamline 11.1 at 0.98 Å wavelength with 0.5° rotation per image (Table 5.4). Diffraction data were scaled using XDS,[3] analyzed using Pointless,[4] and merged using Aimless.[5] Coordinates for the iodine anomalous signal were determined using the program HySS in the Phenix software suite.[6] Electron density maps were generated using anomalous coordinates determined by HySS as initial positions in Autosol. Molecular manipulations of the models were performed with Coot[7]. Coordinates were refined with phenix.refine. Models were refined with riding hydrogen atoms. Table 5.3 and Table 5.4 show the statistics of data collection and model refinement for peptides **1a**, **1b**, **1c**, **2a**, **2b**, and **1a + 1c** (the mixed dodecamer).

X-ray diffraction data collection, data processing, and structure determination for peptide $\mathbf{1a_{T68V}}$

The X-ray diffraction data set for peptide $\mathbf{1a_{T68V}}$ was collected on a Rigaku Micromax-007 HF diffractometer with a Cu rotating anode at 1.54 Å wavelength. We used the program CrystalClear to determine the number of diffraction images to collect. Diffraction data sets were collected with a 0.5° rotation and at a detector distance of 50 mm. Diffraction data were scaled using XDS, analyzed using Pointless, and merged using Aimless. Coordinates for the iodine anomalous signal were determined using the program HySS in the Phenix software

suite. Electron density maps were generated using anomalous coordinates determined by HySS as initial positions in Autosol. Molecular manipulations of the models were performed with Coot. Coordinates were refined with phenix.refine. Models were refined with riding hydrogen atoms. Table 5.4 shows the statistics of data collection and model refinement for peptide $\mathbf{1a_{T68V}}$.

Size-Exclusion chromatography of peptides 1 and 2

The oligomerization of peptides **1** and **2** was studied by size-exclusion chromatography (SEC) at 5 °C in 100 mM potassium phosphate buffer at pH 7.4 as follows: Each peptide was dissolved in phosphate buffer to a concentration of 1 mM. The peptide solutions were centrifuged for 30 seconds at 14,000 rpm and the supernatant was loaded onto a GE Superdex 200 10/300 GL column at 0.2 mL/min over 5 min. After loading, samples were run with phosphate buffer at 0.3 mL/min. Chromatograms were recorded at 280 nm and normalized to the highest absorbance value. Peptides **1** and **2** eluted between 17 and 19 mL (Table 5.2). SEC chromatograms for peptides **1** and **2** are shown in Figure 5.10B–Figure 5.19B. Standards (chymotrypsinogen, ribonuclease A, and vitamin B12) were run in a similar fashion and are plotted along with the peptides.

Cytotoxicity assays of peptides 1 and 2.

The toxicity of peptides **1** and **2** toward SH-SY5Y cells was assessed by LDH release and MTT conversion assays. Cells were incubated in the presence or absence of 25 µM peptide for 24 hours in 96-well plates, and the toxicity of peptides **1** and **2** toward the cells was then assessed by standard LDH and MTT assay techniques. The LDH assay was performed using the Pierce LDH Cytotoxicity Assay Kit from Thermo Scientific. Experiments were performed with 10 peptides (**1a**–**1e** and **2a**–**2e**) in 5 replicates in each plate (50 wells) and an additional

10 wells were used for controls. Cells were cultured in the inner 60 wells (rows B–G, columns 2–11) of the 96-well plate. DMEM:F12 media (100 µL) was added to the outer wells (rows A and H, columns 1 and 12), in order to ensure the greatest reproducibility of data generated in the inner wells. Experiments were replicated across three plates; representative data of a single plate is presented in Table 5.2 and Figure 5.9.

*Preparation of stock solutions of peptides **1** and **2**.* A solution of the peptide was prepared by dissolving 1 mg in 100 µL of 18 MΩ deionized water that was either filtered through a 0.2 µm syringe filter or autoclaved. The concentration was determined spectrophotometrically at 280 nm using a calculated value of ε_{280}= 2820 $M^{-1}cm^{-1}$ (Tyr = 1280, Phe^{I} = 260).[8] The solution was diluted to 250 µM. Separate freshly prepared 250 µM stock solutions were used in the MTT and LDH assays.

Preparation of SH-SY5Y cells for LDH and MTT assays. SH-SY5Y cells were plated in three separate 96-well plates at 15,000 cells per well. Cells were incubated in 100 µL of a 1:1 mixture of Dulbelcco's Modified Eagle Medium and Nutrient Mixture F12 Ham's (DMEM:F12) media supplemented with 10% fetal bovine serum, 100 U/mL penicillin, and 100 µg/mL streptomycin at 37 °C in a 5% CO_2 atmosphere for 24 hours.

*Treatment of SH-SY5Y cells with peptides **1** and **2**.* After 24 hours, the culture media was removed and replaced with 90 µL of serum-free DMEM:F12 media. A 10 µL aliquot of peptide **1** or **2** stock solution was added to each well, for a total of 25 µM of peptides **1** or **2** per well. Experiments were run in quintuplicate, with 50 wells used for peptides. An additional 5 wells were used as controls and received a 10 µL aliquot of water. Another 5 wells were left untreated, to be subsequently used as controls with lysis buffer. Cells were incubated at 37 °C in a 5% CO_2 atmosphere for 23 hours.

LDH release assay. After 23 hours, 10 µL of 10× lysis buffer was added to the five untreated wells. The cells were incubated for an additional hour. A 50 µL aliquot from each

well was transfered to a new 96-well plate and 50 µL of LDH substrate solution was added to each well. The treated plates were stored in the dark for 30 min, then 100 µL of stop solution was added to each well. The absorbance of each well was measured at 490 and 680 nm (A_{490} and A_{680}). Data were analyzed by calculating the differential absorbance for each well (A_{490} - A_{680}) and comparing those values to those of the lysis buffer controls and the untreated controls:

$$\% \text{ LDH release} = \frac{(A_{490}-A_{680})_{peptide}-(A_{490}-A_{680})_{untreated}}{(A_{490}-A_{680})_{lysis}-(A_{490}-A_{680})_{untreated}} \times 100$$

Representative data of a single plate are presented in Table 5.2 and Figure 5.9.

MTT conversion assay. The remaining cell media was removed and replaced with 100 µL of serum-free, phenol-red-free DMEM:F12 media containing 1.2 mM MTT. The cells were incubated in the presence of MTT at 37 °C in a 5% CO_2 atmosphere for 12 hours. After 12 hours, 100 µL of a 10% SDS in 10 mM HCl solution was added to each well and the cells were incubated for an additional 4 hours. After 4 hours, the absorbance of each well was measured at 570 nm. Data were analyzed by comparing absorbance for each well (A_{570}) to that of the untreated controls:

$$\% \text{ MTT conversion} = \frac{(A_{570})_{peptide}}{(A_{570})_{untreated}} \times 100$$

Representative data of a single plate are presented in Table 5.2 and Figure 5.9.

1. Spencer, R.; Chen, K. H.; Manuel, G.; Nowick, J. S. *Eur. J. Org. Chem.* **2013**,*2013*, 3523–3528.
2. Spencer, R. K.; Li, H.; Nowick, J. S. *J. Am. Chem. Soc.* **2014**, *136*, 5595–5598.

3. Kabsch, W *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125–132.
4. Evans, P. R. Acta Crystallogr., Sect. D: Biol. Crystallogr. **2011**, *67*, 282–292.
5. Evans, P. R.; Murshudov, G. N. Acta Crystallogr., Sect. D: Biol. Crystallogr. **2013**, *69*, 1204–1214.
6. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213–221.
7. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486–501.
8. De Filippis, V.; Colombo, G.; Russo, I.; Spadari, B.; Fontana, A. *Biochemistry*, **2002**, *41*, 13556–13569.

Table 5.3: Crystallographic Properties, Crystallization Conditions, and Data Collection and Model Refinement Statistics for Peptides **1a**, **1b**, **1c**, **2a**, and **2b**.

peptide	**1a**	**1b**	**1c**	**2a**	**2b**
PDB ID	4P4V	4P4W	4P4X	4P4Y	4P4Z
space group	$P4_222$	$P3_121$	$P4_122$	$R32$	$P4_32_12$
a, b, c (Å)	62.03 62.03 62.03	58.19 58.19 121.28	61.34 61.34 113.96	66.26 66.26 154.63	63.31 63.31 126.71
α, β, γ (°)	90 90 90	90 90 120	90 90 90	90 90 120	90 90 90
peptide per asymmetric unit	3	12	12	7	12
crystallization conditions	0.1 M Tris, pH 8.0; 0.3 M Li_2SO_4; 45% PEG 400	0.1 M Tris, pH 8.0; 1.5 M $(NH_4)_2SO_4$	0.1 M Tris, pH 7.5; 0.2 M Li_2SO_4; 25% PEG 400	0.1 M Tris, pH 7.5; 0.2 M Li_2SO_4; 26% PEG 400	0.1 M SPG, pH 10.0; 35% PEG 1500
Data Collection[a]					
wavelength (Å)	1.00	1.00	1.00	1.00	1.00
resolution (Å)	31.3–1.97 (2.04–1.97)	31.53–1.50 (1.55–1.50)	32.3–1.90 (1.97–1.90)	33.13–1.51 (1.56–1.51)	35.13–1.31 (1.36–1.31)
total reflections	45138(4263)	77828 (7458)	35831 (3478)	104626 (9221)	438334 (43398)
unique reflections	6556 (621)	38934 (3748)	17924 (1744)	20672 (1972)	62669 (6159)
multiplicity	6.9	2.0	2.0	5.1	7.0
completeness (%)	99.88 (99.20)	99.79 (97.94)	99.91 (99.09)	98.72 (95.54)	99.96 (100.00)
mean I/σ	13.79 (1.64)	26.23 (4.36)	12.49 (2.02)	18.26 (3.17)	16.09 (1.76)
Wilson B-factor	37.08	15.36	30.49	19.50	11.96
R_{merge}	0.11 (1.21)	0.03 (0.19)	0.03 (0.38)	0.05 (0.41)	0.8 (1.11)
$R_{measure}$	0.12	0.04	0.04	0.06	0.08
$CC_{1/2}$	0.99 (0.63)	0.99 (0.88)	0.99 (0.62)	0.99 (0.89)	0.99 (0.64)
CC^*	0.99 (0.88)	0.99 (0.97)	1.00 (0.88)	0.99 (0.97)	1.00 (0.88)
Refinement					
R_{work}	19.46	17.45	20.87	15.73	15.09
R_{free}	22.18	20.88	23.50	17.47	17.45
number of non-hydrogen atoms	477	1942	1732	1123	2162
RMS_{bonds}	0.008	0.040	0.010	0.014	0.014
RMS_{angles}	1.28	2.42	1.45	1.53	1.60
Ramachandran favored (%)	100	100	100	100	100
outliers (%)	0	0	0	0	0
clashscore	2.12	12.98	6.64	4.66	5.26
average B-factor	46.60	26.10	47.60	31.00	18.00

Table 5.4: Crystallographic Properties, Crystallization Conditions, and Data Collection and Model Refinement Statistics for Peptides **1a + 1c** and $\mathbf{1a_{T68V}}$.

peptide	**1a + 1c**	$\mathbf{1a_{T68V}}$
PDB ID	4WC8	4X0S
space group	$P4_122$	$R32$
a, b, c (Å)	58.05 58.05 112.95	51.06 51.06 31.42
α, β, γ (°)	90 90 90	90 90 120
peptide per asymmetric unit	12	1
crystallization conditions	0.1 M Tris, pH 7.5, 0.2 M Li_2SO_4, 30% PEG 400	0.1 M Tris, pH 8.5, 0.2 M Li_2SO_4, 30% PEG 400
	Data Collection[a]	
wavelength (Å)	0.98	1.54
resolution (Å)	31.59–1.91 (1.98–1.91)	25.61–2.03 (2.10–2.03)
total reflections	85498 (7250)	7604 (387)
unique reflections	15594 (1482)	1072 (98)
multiplicity	5.5	7.1
completeness (%)	99.20 (96.93)	98.89 (93.33)
mean I/σ	12.15 (2.61)	49.07 (16.40)
Wilson B-factor	26.16	23.01
R_{merge}	0.08 (0.53)	0.06 (0.10)
$R_{measure}$	0.09	0.07
$CC_{1/2}$	0.99 (0.87)	0.99 (0.99)
CC*	1.00 (0.96)	1.00 (0.99)
	Refinement	
R_{work}	16.87	16.49
R_{free}	21.86	20.21
number of non-hydrogen atoms	1739	147
RMS_{bonds}	0.008	0.010
RMS_{angles}	1.18	1.19
Ramachandran favored (%)	99	100
outliers (%)	0	0
clashscore	5.01	0
average B-factor	37.90	39.40

[a]Values for the highest resolution shell are show in parentheses.

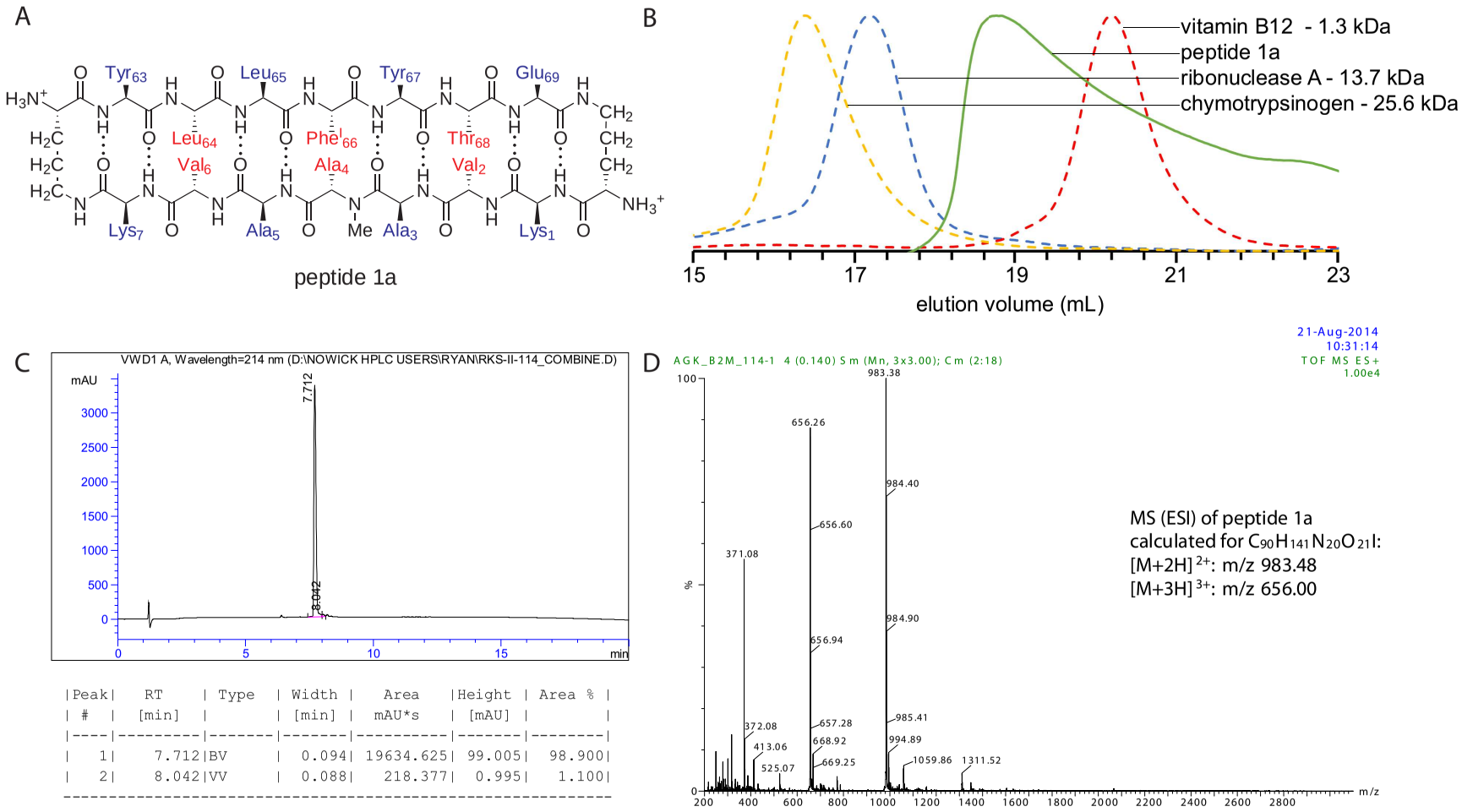


Figure 5.10: (A) Chemical structure of peptide **1a**. (B) SEC chromatograms of the supernatant of peptide **1a** (green), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). Peptide **1a** precipitated from phosphate buffer solution. SEC was performed on the supernatent after removal of the precipitate by centrifugation. (C) HPLC chromatogram of peptide **1a** after synthesis and isolation. (D) Mass spectrum of peptide **1a**.

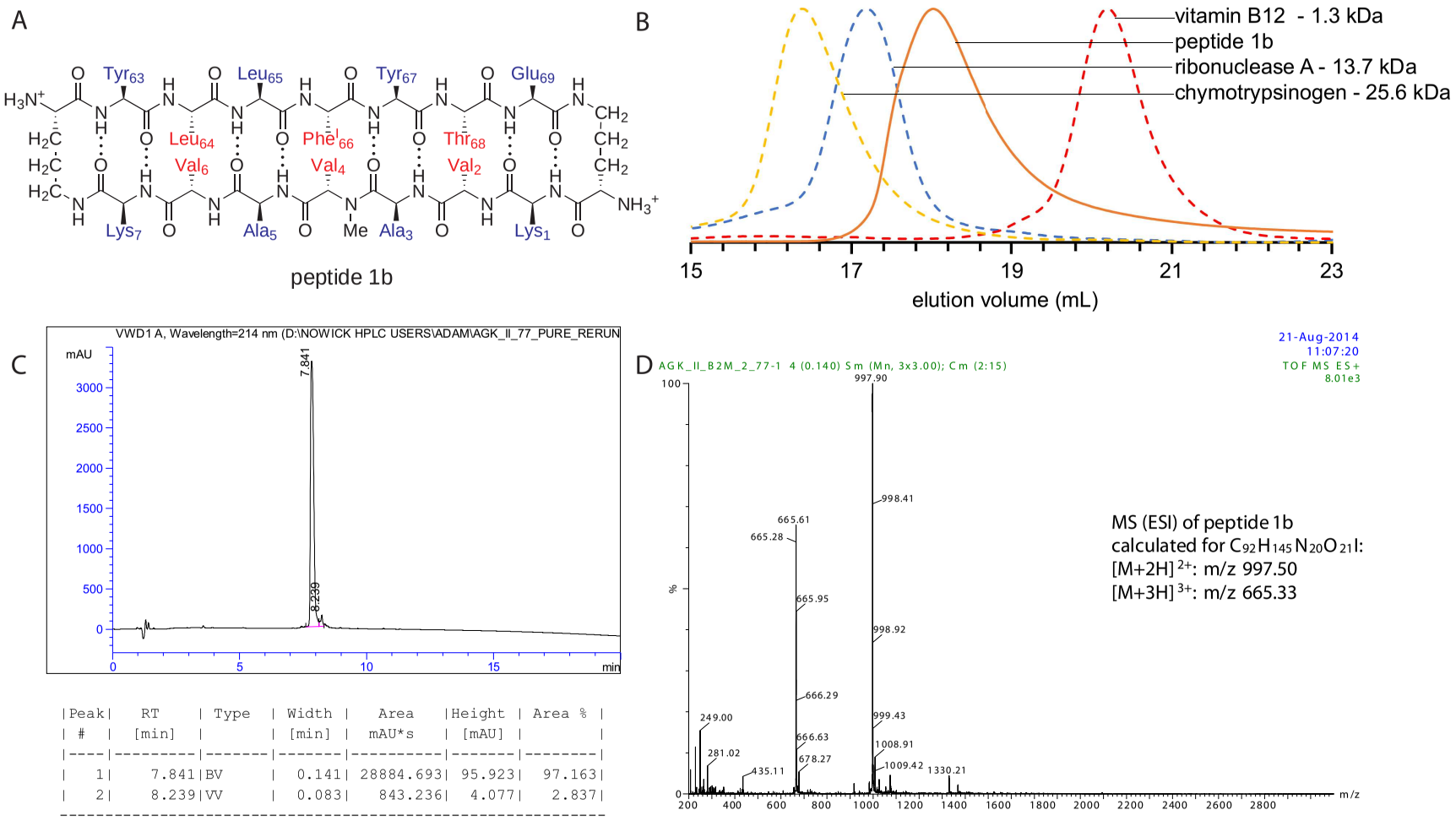


Figure 5.11: (A) Chemical structure of peptide **1b**. (B) SEC chromatograms of peptide **1b** (orange), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (C) HPLC chromatogram of peptide **1b** after synthesis and isolation. (D) Mass spectrum of peptide **1b**.

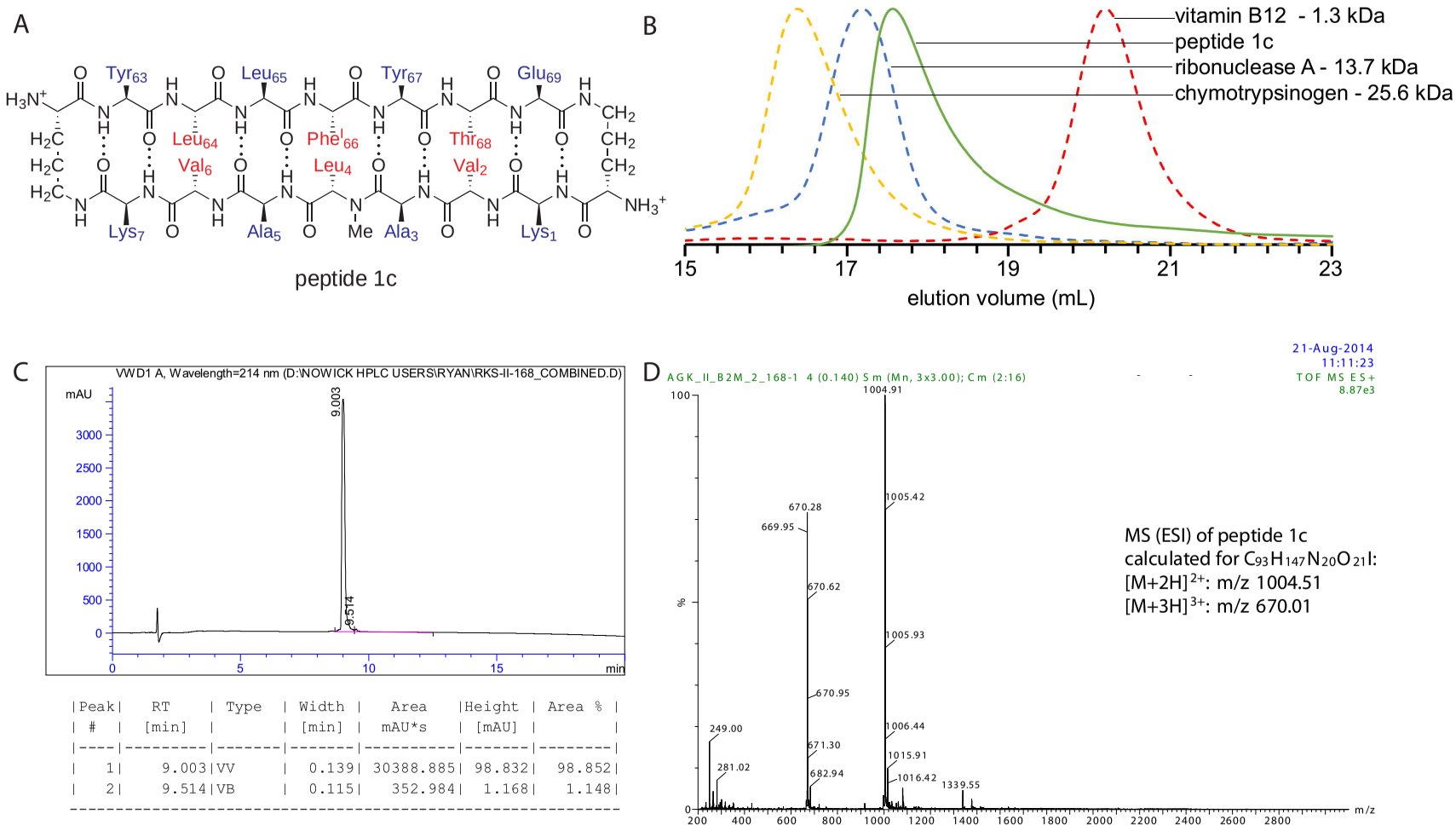


Figure 5.12: (A) Chemical structure of peptide **1c**. (B) SEC chromatograms of peptide **1c** (green), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (C) HPLC chromatogram of peptide **1c** after synthesis and isolation. (D) Mass spectrum of peptide **1c**.

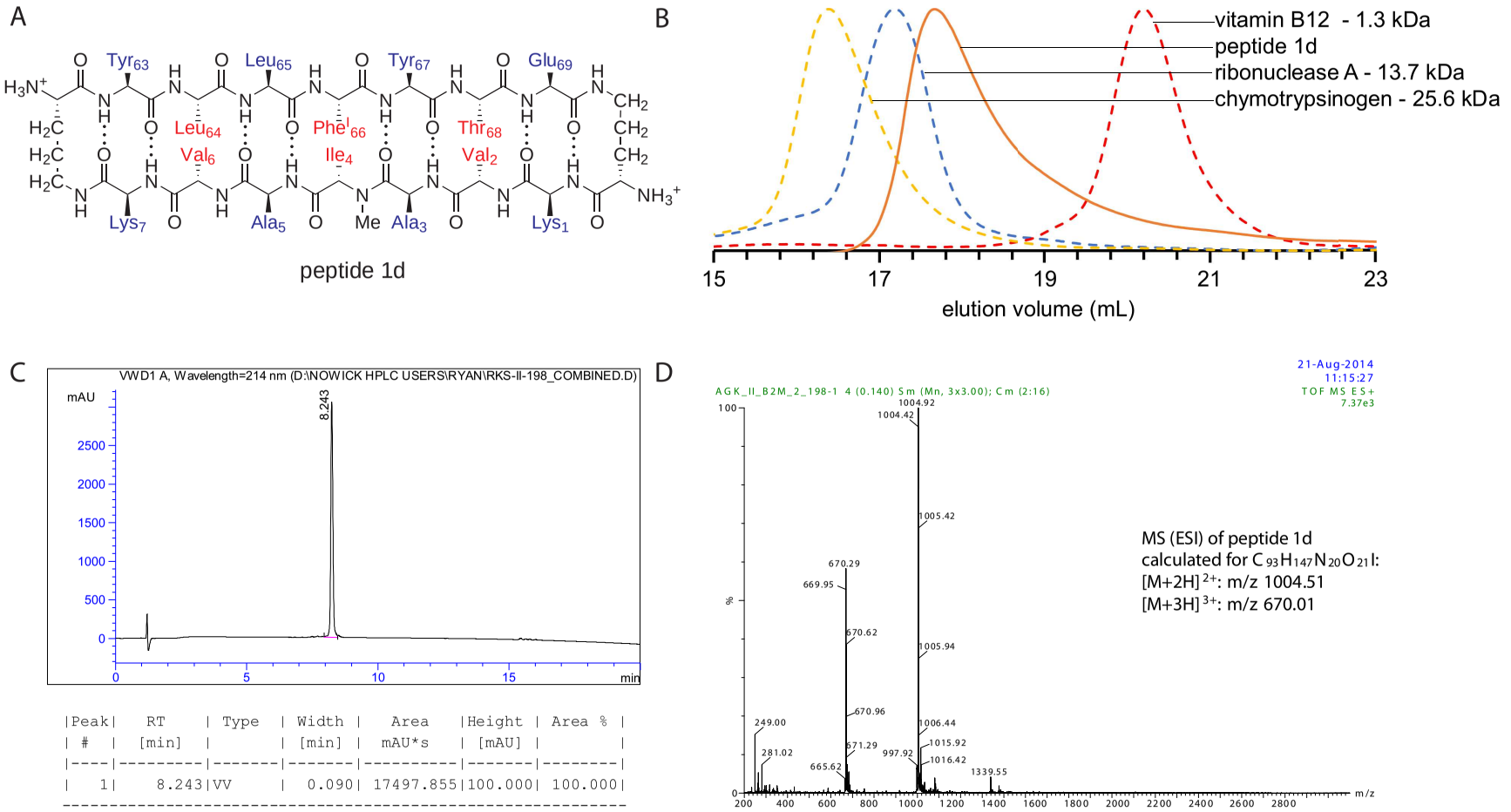


Figure 5.13: (A) Chemical structure of peptide **1d**. (B) SEC chromatograms of peptide **1d** (orange), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (C) HPLC chromatogram of peptide **1d** after synthesis and isolation. (D) Mass spectrum of peptide **1d**.

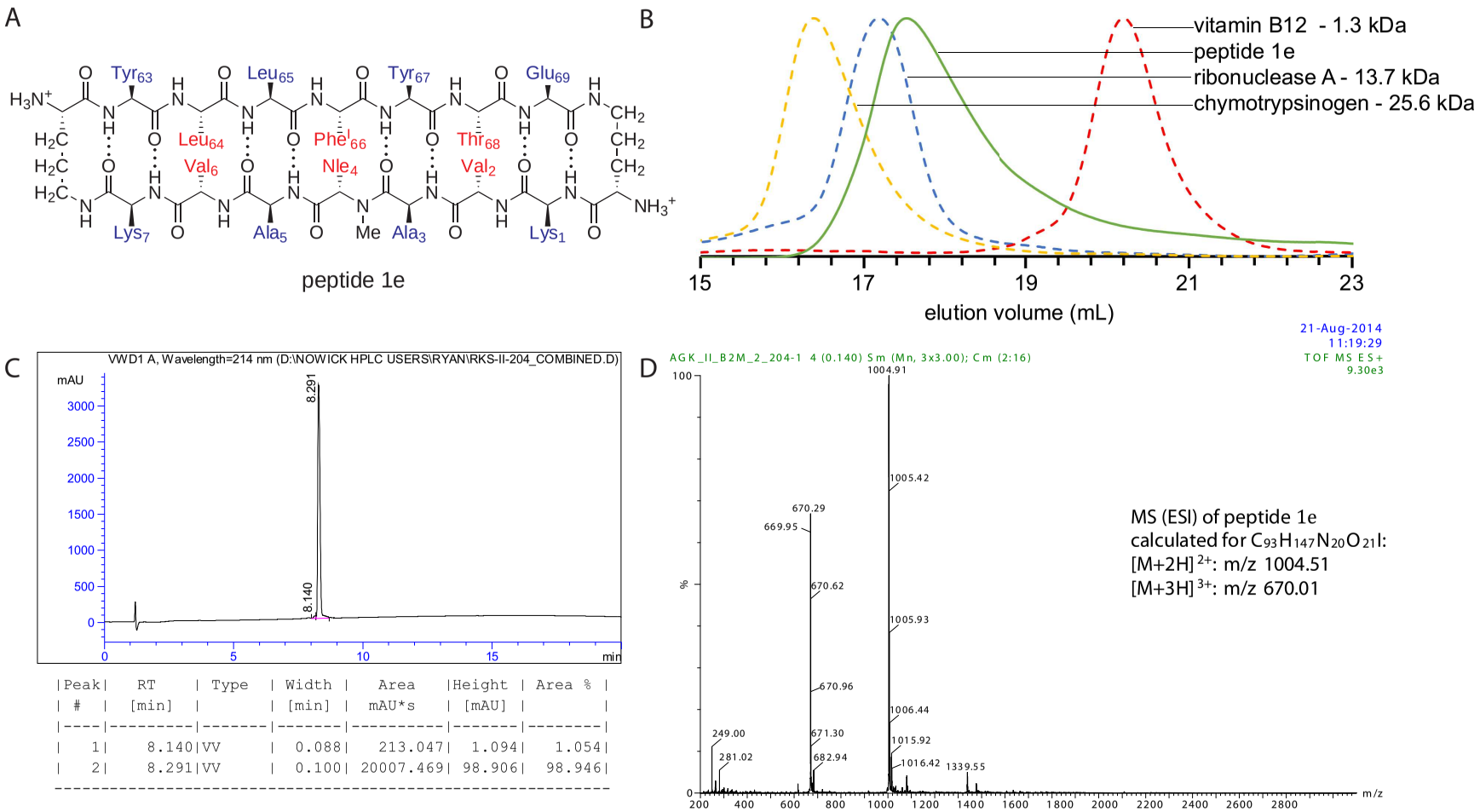


Figure 5.14: (A) Chemical structure of peptide **1e**. (B) SEC chromatograms of peptide **1e** (green), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (C) HPLC chromatogram of peptide **1e** after synthesis and isolation. (D) Mass spectrum of peptide **1e**.

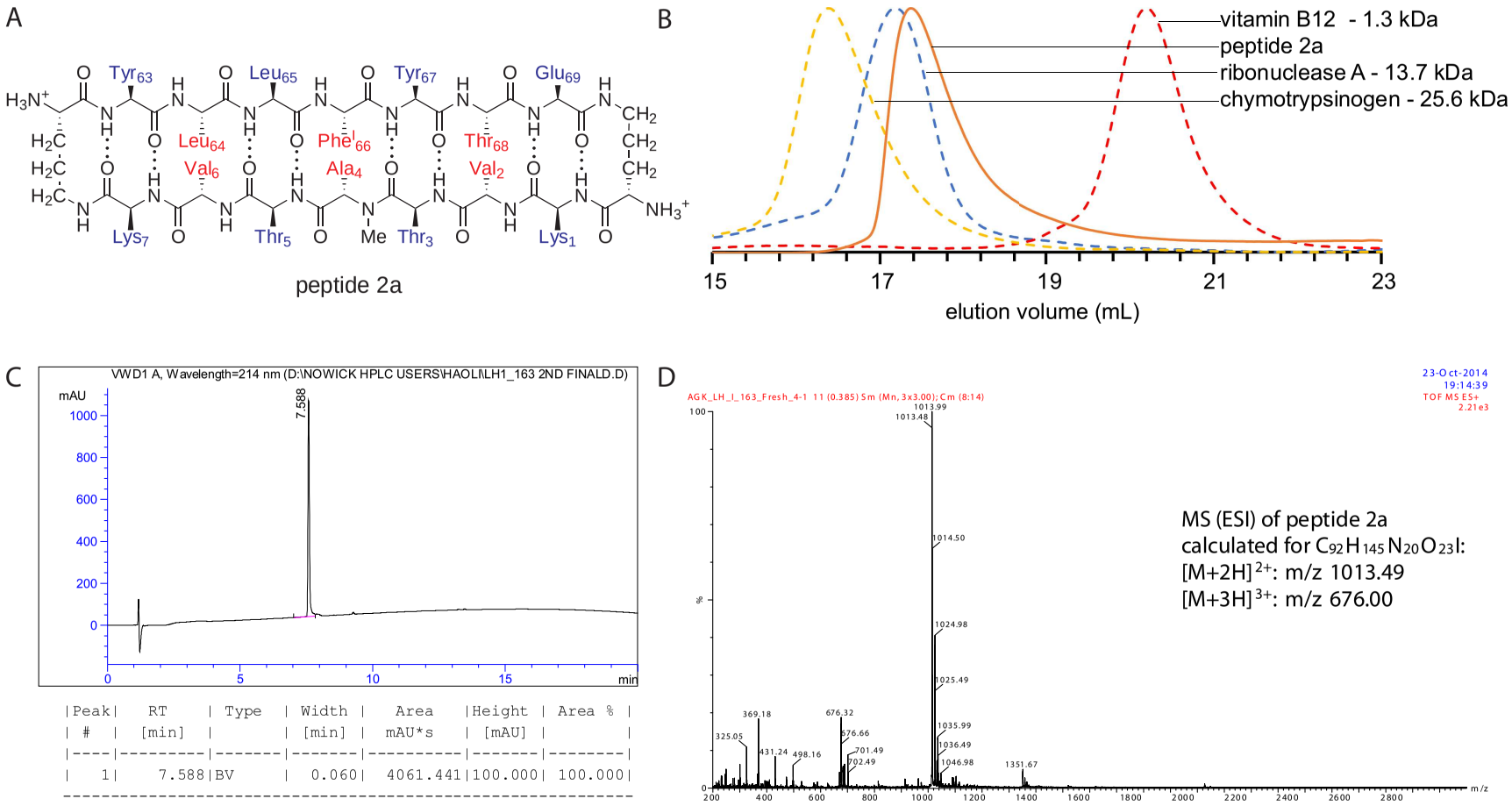
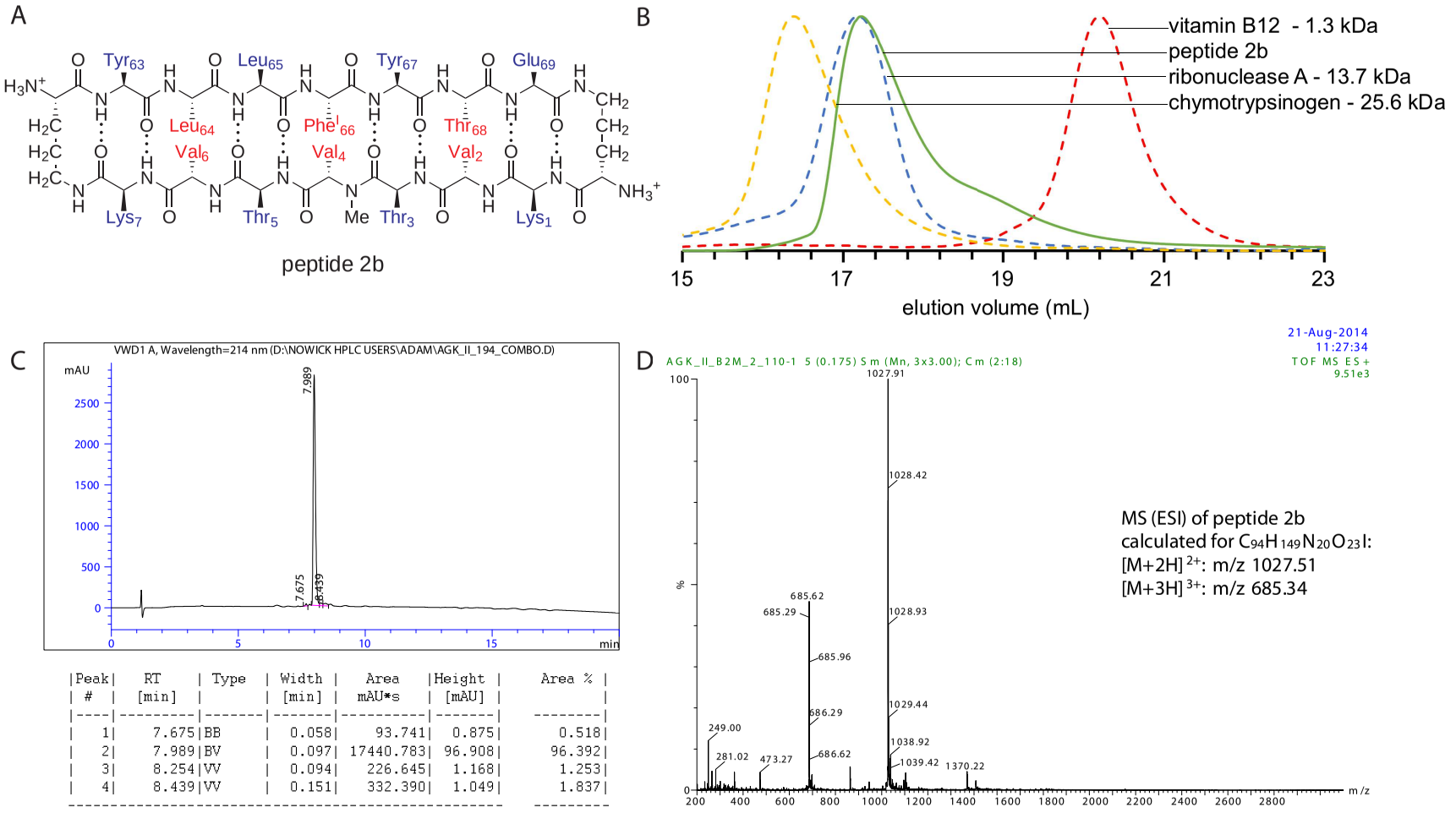


Figure 5.15: (A) Chemical structure of peptide **2a**. (B) SEC chromatograms of peptide **2a** (orange), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (C) HPLC chromatogram of peptide **2a** after synthesis and isolation. (D) Mass spectrum of peptide **2a**.



Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	7.675	BB	0.058	93.741	0.875	0.518
2	7.989	BV	0.097	17440.783	96.908	96.392
3	8.254	VV	0.094	226.645	1.168	1.253
4	8.439	VV	0.151	332.390	1.049	1.837

Figure 5.16: (A) Chemical structure of peptide **2b**. (B) SEC chromatograms of peptide **2b** (green), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (C) HPLC chromatogram of peptide **2b** after synthesis and isolation. (D) Mass spectrum of peptide **2b**.

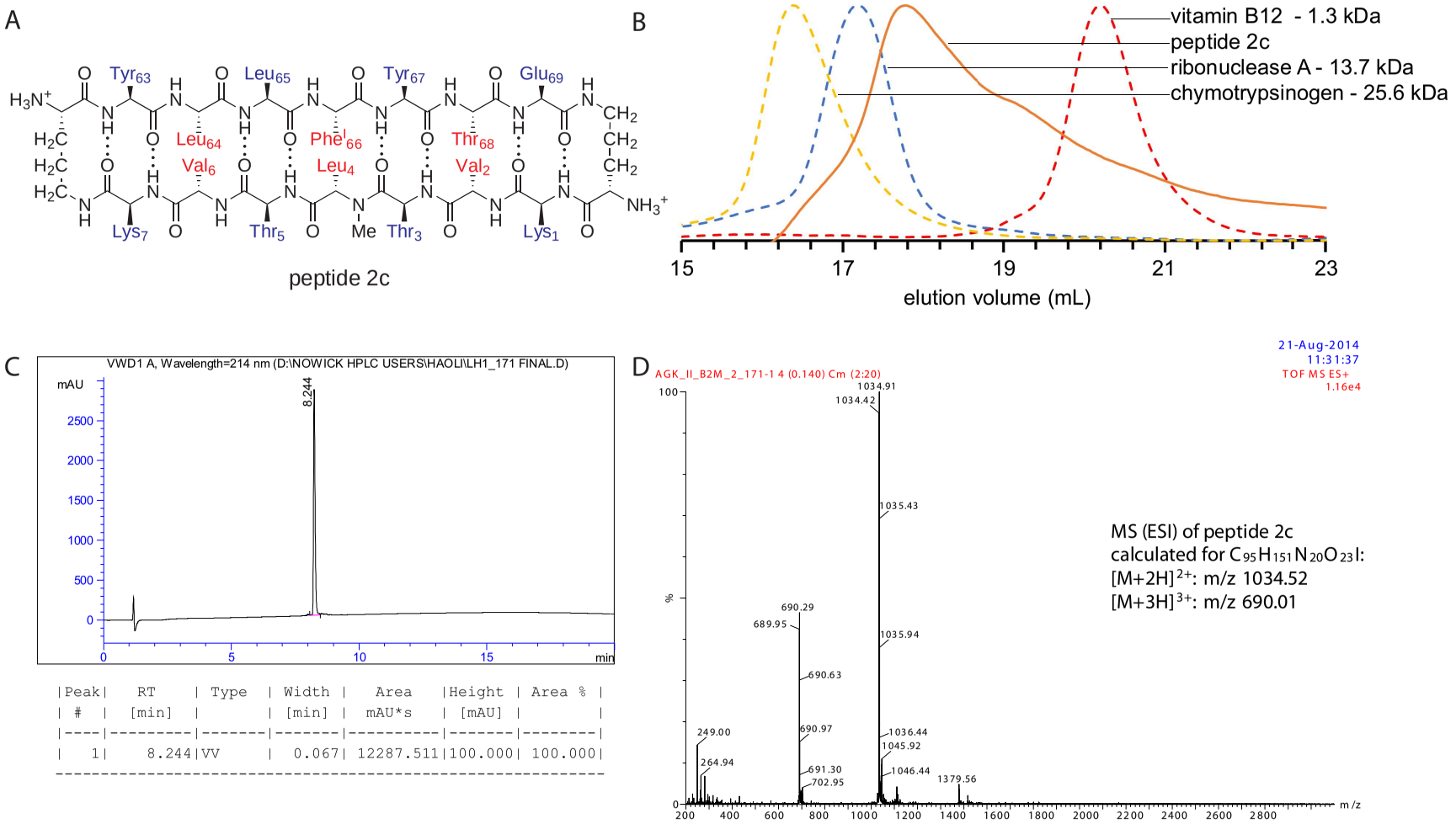


Figure 5.17: (A) Chemical structure of peptide **2c**. (B) SEC chromatograms of peptide **2c** (orange), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (C) HPLC chromatogram of peptide **2c** after synthesis and isolation. (D) Mass spectrum of peptide **2c**.

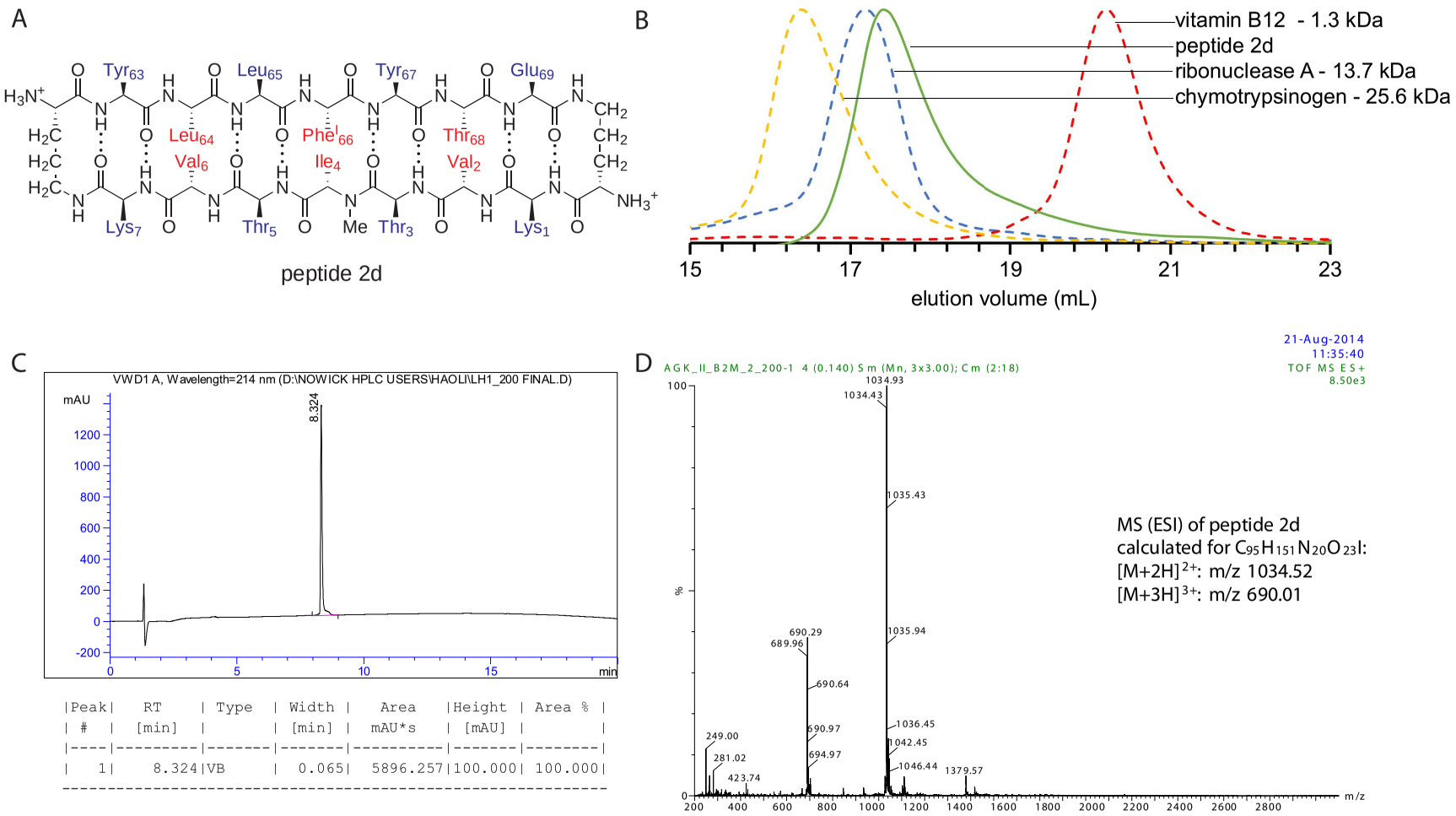
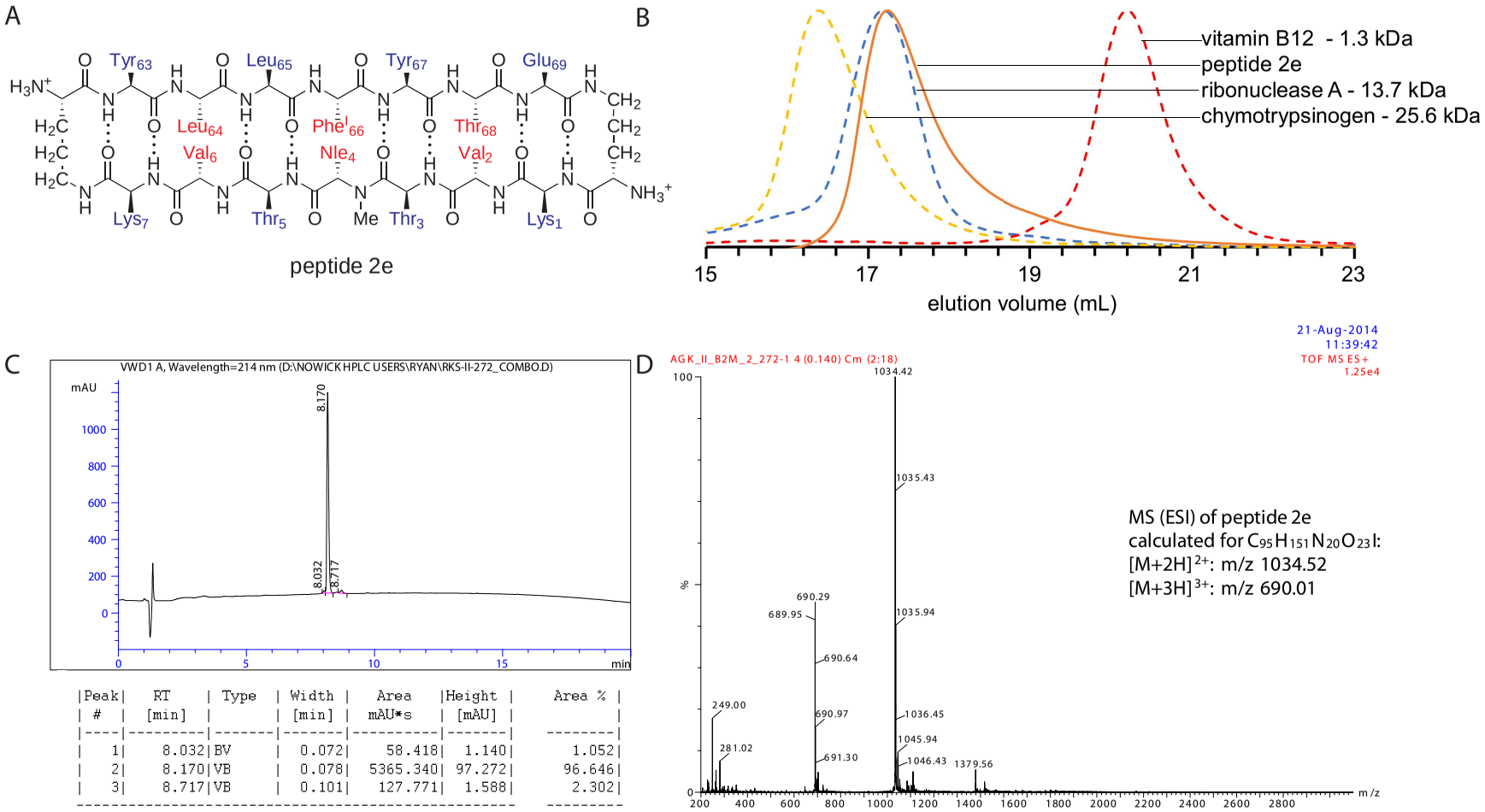


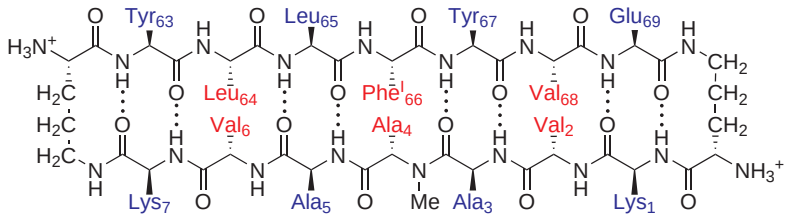
Figure 5.18: (A) Chemical structure of peptide **2d**. (B) SEC chromatograms of peptide **2d** (green), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (C) HPLC chromatogram of peptide **2d** after synthesis and isolation. (D) Mass spectrum of peptide **2d**.



Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	8.032	BV	0.072	58.418	1.140	1.052
2	8.170	VB	0.078	5365.340	97.272	96.646
3	8.717	VB	0.101	127.771	1.588	2.302

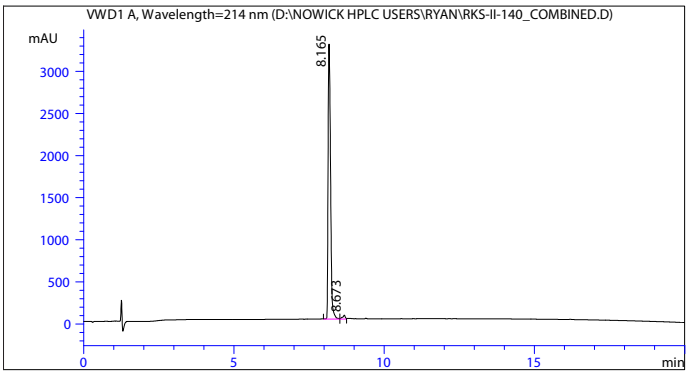
Figure 5.19: (A) Chemical structure of peptide **2e**. (B) SEC chromatograms of peptide **2e** (orange), chymotrypsinogen (yellow), ribonuclease A (blue), and vitamin B12 (red). (C) HPLC chromatogram of peptide **2e** after synthesis and isolation. (D) Mass spectrum of peptide **2e**.

A



peptide $1a_{T68V}$

B



Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	8.165	MM	0.096	18914.260	98.555	98.440
2	8.673	MM	0.104	299.743	1.445	1.560

C

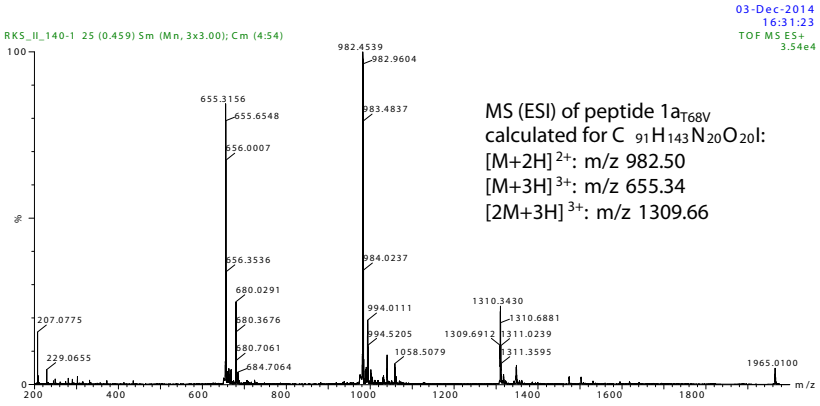


Figure 5.20: (A) Chemical structure of peptide $\mathbf{1a_{T68V}}$. (B) HPLC chromatogram of peptide $\mathbf{1a_{T68V}}$ after synthesis and isolation. (C) Mass spectrum of peptide $\mathbf{1a_{T68V}}$.