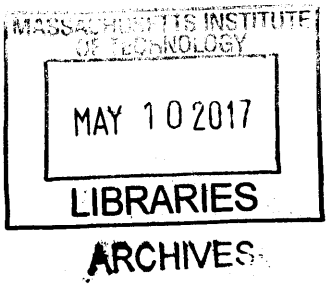
Fast Flow Biopolymer Synthesis



by

Mark David Simon

B.S. Chemistry
University of Florida, 2011

Submitted to the Department of Chemistry
In Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

at the

Massachusetts Institute of Technology

February 2017

© Massachusetts Institute of Technology
All rights reserved

Signature of Author: Signature redacted
Department of Chemistry
January 13, 2017

Certified by: Signature redacted
Bradley L. Pentelute
Pfizer-Laubach Career Development Professor of Chemistry
Thesis Supervisor

Accepted by: Signature redacted
Robert W. Field
Haslam and Dewey Professor of Chemistry
Chairman, Departmental Committee for Graduate Students

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Signature redacted

Alexander M. Klibanov
Novartis Professor of Chemistry and Bioengineering
Thesis Committee Chair

Signature redacted

Bradley L. Pentelute
Pfizer-Laubach Career Development Professor of Chemistry
Thesis Supervisor

Signature redacted

Jeremiah A. Johnson
Firmenich Career Development Professor of Chemistry

Fast Flow Biopolymer Synthesis

by

Mark David Simon

Submitted to the Department of Chemistry
on January 13, 2017 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Chemistry

Abstract

This thesis describes the development and application of fast flow solid phase synthesis for the preparation of peptides and phosphorodiamidate morpholino oligomers (PMOs), as well as the application of fast, reliable peptide synthesis to study non-natural protein folding and function.

In the first chapter, solid supported peptide synthesis was accelerated using flow by continuously delivering preheated solvents and reagents to the solid support at high flow rate, thereby maintaining maximal concentrations, quickly exchanging reagents, and eliminating the need to heat reagents after they were added to the vessel.

In the second chapter, these chemical principles were expanded upon and mechanical challenges particular to accelerated solid phase synthesis were overcome to build a fully automated fast flow peptide synthesizer than incorporates amino acids in as little as 40 seconds each. First, mechanical systems were developed to rapidly switch between the many reagents needed for peptide synthesis while maintaining the proper stoichiometry of all reaction components at all times. Second, conditions under which reagents did not appreciably degrade during storage or synthesis were found. Finally, synthetic outcomes were substantially improved by increasing temperature without degrading the protected, resin bound peptide.

The third chapter describes the expansion of fast flow synthesis to PMOs. A 10-fold acceleration of PMO synthesis was realized using mechanical systems adapted from chapter 1, increasing the reaction temperature to 90°C, and introducing a Lewis acid catalyst. The acidity of the deprotection reagent was reduced to prevent cleavage of the backbone during 3' detritylation.

In the final chapter, a "D-scan" of two small proteins, the disulfide-rich *Ecballium elaterium* trypsin inhibitor II (EETI-II) and a minimized Z domain of protein A (Z33), is reported. For each protein, the chirality of one amino acid at a time was inverted to generate a series of diastereomers, and study the critical stereocenters of EETI-II and Z33. Twelve out of 30 EETI-II analogs folded and were high-affinity trypsin inhibitors, but most active analogs were less stable to reduction than EETI-II. Similarly, twelve Z33 analogs retained high binding affinity to IgG, but most were substantially less stable than WT-Z33.

Thesis Supervisor: Bradley L. Pentelute
Pfizer-Laubach Career Development Professor of Chemistry

List of figures and tables

Acknowledgements

I would like to thank my advisor, Brad Pentelute, for fostering a creative, collaborative environment. I have worked with many wonderful people during my time here, and the following efforts would not have been possible without them. In particular, I would like to thank Alex Vinogradov, Chi Zhang, Alex Mijalis, Rocco Policarpo, Ted Li, Yuta Maki, Gunnar Hanson, and Kyle Totaro for their invaluable contributions.

I would also like to thank Sam Colgate for spending many patient hours in the shop and classroom with me, and encouraging my application to MIT.

This work was generously funded by MIT startup funds for Brad Pentelute, the MIT Deshpande Center, Dr Reddy's Laboratories, Sarepta Therapeutics, and the Defense Advanced Research Projects Agency.

Introduction

The total chemical synthesis of biomimetic polymers promises to enable the next generation of therapeutics. Synthetic peptide drugs now treat HIV[1], diabetes[2], bacterial infections[3], aggressive and refractory cancers[4,5], and a variety of other conditions[6,7]. Phosphorodiamidate morpholino oligomers, DNA analogs with uncharged backbones, are undergoing trials to treat Duchene's muscular dystrophy[8], dengue[9], and hemorrhagic fevers[10]. Modified RNA is used to manage spinal muscular atrophy[11] and genetic hypercholesterolemia[12], and natural RNA conjugated to suitable delivery vehicles is under investigation for a wide variety of conditions from pancreatic cancer to dry eyes[13].

Today, biopolymers are often prepared chemically, through solid phase synthesis. Supported by synthetic methodology, traditional medicinal chemistry tools are employed to create the best molecule without the constraints of a biological expression system[14–17]. The size and complexity of these molecules, however, makes synthesis challenging.

In solid phase synthesis, a terminal protected monomer is bound to a macroscopic solid support, and additional monomers are added by iteratively partially deprotecting the previous monomer and coupling a new protected monomer[18]. Intermediates remain bound to the solid support, reactions are performed with a large excess of reagent, and excess reagent and reaction byproducts are mechanically washed away after each step. The solid support solubilizes otherwise insoluble protected intermediates, and purifications are performed with mechanical washing rather than chromatographic columns[19,20]. Since Bruce Merrifield's 1963 report of solid phase peptide synthesis, solid phase methods have been significantly extended to become standard for preparing peptides[21–23], carbohydrates[24], oligonucleotides[25], and certain small molecules[26], and are used to access advanced intermediates in the total synthesis of proteins[27,28], genes[29], and even full length chromosomes[30]and genomes[31].

With these techniques, the assembly of large synthetic biopolymers became feasible, but was still challenging and time consuming[32,33]. Most efforts at accelerating solid supported synthesis, including our own, have focused on peptides[34–37]. As functional natural products, peptides are attractive research tools and drug leads, but their synthesis often proves difficult.

Inspired by the success of continuous flow chemistry[38–42], successful automation of solid phase carbohydrate synthesis[43,44], and earlier work in flow peptide synthesis [45,46], we set out to accelerate biopolymer synthesis in flow, starting with peptides. Peptide synthesis was a natural starting point because strategies have been under continuous development for over a century, leading to a broad, well documented set of synthetic chemistries and potential problems. The rich literature[47] and general interest around peptide synthesis allowed our first effort, presented in Chapters 2 and 3, to focus on challenges

particular to rapid flow based synthesis, and interpret synthetic failures in the context of existing literature.

In the course of establishing a robust, automated fast flow peptide synthesis method, we articulated critical aspects of rapid solid phase synthesis and developed the apparatus to support it. In Chapter 4, this knowledge and hardware were leveraged to accelerate phosphorodiamidate morpholino oligomer synthesis. Free to focus on the chemical aspects of fast PMO synthesis at elevated temperature, we realized a ten fold acceleration of synthesis over patent procedures[48].

Finally, Chapter 5 describes the application of the automated peptide synthesizer to synthesis of two series of protein diastereomers, and the investigation of their structure and function. The effect of inverting the chirality of a single amino acid in EETI-II, a disulfide rich trypsin inhibitor, and Z33, a minimized form of the IgG binding Z domain of Protein A was studied. Fast peptide synthesis made it feasible to chemically prepare more than 60 proteins for detailed biological study, and, for the first time, we have a map of how the chirality of each residue affects the stability and function of two structurally distinct proteins.

Manual fast flow peptide synthesis

The work presented in this chapter was published in the following manuscript, and Appendix 1 is derived from the supporting information

Mark D. Simon, Patrick L. Heider, Andrea Adamo, Alexander A. Vinogradov, Surin K. Mong, Xiyuan Li, Tatiana Berger, Rocco L. Policarpo, Chi Zhang, Yekui Zou, Xiaoli Liao, Alexander M. Spokoyny, Klavs F. Jensen, Bradley L. Pentelute. Rapid Flow-Based Peptide Synthesis *Chembiochem* 2014, 15, 713–720.

2.1 Introduction

The total chemical synthesis of peptides has been of great interest for over a century[49]. Chemical synthesis of peptides and proteins enables incorporation of non-proteogenic functionalities, without restriction on their location or number[15,16,50]. With the introduction of solid phase peptide synthesis in 1963, the total synthesis of short peptides became routine[18]. Subsequent advances extended the ability to assemble long polypeptides[51–53], and the advent of native chemical ligation enabled the preparation of polypeptides of theoretically unlimited length from smaller fragments[28].

Despite these advances, the time required to assemble polypeptides, either as final targets or as fragments of a larger molecule, is a major limitation on studies employing such synthetic material. Most peptides are synthesized with Fmoc protocols, rather than faster and higher yielding Boc procedures, because highly toxic hydrogen fluoride is not required, and only small amounts of costly trifluoroacetic acid are used[51]. Standard Fmoc solid phase peptide synthesis methods require 60 to 100 minutes to incorporate each amino acid residue[32,33], and some specialized procedures use complex microwave systems to reduce this to 5-20 minutes per residue[35–37]. In this chapter, we describe the development of a conventionally heated, manually operated flow based peptide synthesizer that incorporates an amino acid residue every 3 min. In Chapter 3, the principles elucidated here are used to construct a fully automated instrument that incorporates an amino acid every 40 seconds.

2.2 Apparatus design

To perform such rapid synthesis, we began with an analysis of existing kinetic data. It is known that at room temperature amide bond formation is 99.9% complete in less than 2 minutes with 0.5 M solutions of protected proteogenic amino acids and HATU activation[54], and removal of N-terminal Fmoc protection is effected in 4–6 minutes with 20% (v/v) piperidine in DMF[55]. Our own model studies corroborated these data, showing a reaction half-life of 4.6±0.6 seconds for the formation of an amide bond between 0.3 M HBTU activated leucine and resin bound phenylalanine at room temperature (See Appendix 1). Standard procedures allow these steps to proceed for much longer in an effort to improve the quality of difficult sequences, although this strategy is often of marginal utility[56].

Assuming the reaction rate for these processes doubles for every 10 °C increase in temperature, at 60 °C amide bond formation should be complete in less than 10 seconds and Fmoc removal in less than 20 seconds. Based on this, we believed that robust Fmoc based peptide synthesis could be carried out in substantially less than five minutes per residue at 60 °C, using conventional heaters.

To study peptide synthesis on these time scales, we sought to build a device capable of delivering preheated reagents and solvents to a synthesis vessel, continuously monitoring the process, and rapidly switching between reagents. The second criteria pushed us to revisit continuous flow peptide synthesis, as several previous systems were able to effectively track the progress of peptide synthesis by monitoring the UV absorbance of the reaction effluent[45,46]. Furthermore, with a flow-based system, solvents and reagents can be rapidly preheated by pumping them through a high efficiency heat exchanger with low residence time[57]. This arrangement eliminates the time required to heat reagents after they are delivered to the synthesis vessel and prevents degradation from prolonged storage of reagents at elevated temperature[58].

Previous flow-based peptide synthesis systems had two major drawbacks, however. First, washes were very slow. All of these systems used a long, packed column containing the solid support, similar to an HPLC column. As with equilibrating an HPLC column, several column volumes of solvent were required to effectively wash them. At low flow rate, this required tens of minutes to hours. As flow rate increased, the solid support collapsed and backpressure rose rapidly. Eventually, the support extruded through the frits used to confine it[45]. Although several solutions to this problem were proposed, none were ultimately successful, and commercial variants of these synthesizers are not currently available[45,46,59,60]. Second, these systems recirculated low concentration reagents rather than continuously replenishing high concentration reagents. This conserved activated amino acids, but resulted in slow amide bond formation. To overcome these problems, we designed a low volume, low pressure reaction vessel that was shorter and wider than common laboratory HPLC columns. This vessel reduced the volume of wash solvent required and enabled delivery of solvent and reagent at high flow rates, allowing reagents to be maintained at maximal concentration and removed rapidly.

To deliver the reagents required for peptide synthesis without a complex fluidic manifold, the apparatus shown in Figure 2.2.1A was developed. An HPLC pump was used to deliver either DMF or 50% (v/v) piperidine in DMF for the common washing and deblocking steps, and a syringe pump was used to deliver coupling reagents. The HPLC pump solvent was selected with a manually actuated 3-way valve, and the HPLC pump outlet was attached to the reaction vessel via a luer-lock quick connect. For the coupling step, the quick connect was manually moved to a syringe pump that delivered a solution of activated amino acid. The effluent was passed through a UV detector to continuously monitor the absorbance at 304 nm, a region where Fmoc amino acids and the dibenzofulvene-piperidine deprotection adduct absorb strongly[46].

The first generation reaction vessel, shown in Figure 2.2.1B, was designed to be simple and easy to construct, while still giving low back pressure. A ¼" (6.4 mm) inner diameter × 3" (76 mm) long perfluoroalkoxy tube with Swagelok reducing unions as the inlet and outlet was used. A frit was

positioned in the outlet using a short piece of tubing with a ¼" outer diameter. Installation of the outlet fitting and concurrent compression of the ferrule and tube sealed the frit in place. The total volume of the vessel was ≈2.5 mL. This vessel was assembled without machine or glass shop support, and allowed us to conduct model studies with up to 100 mg of resin.

To verify the feasibility of Fmoc SPPS with this system, we synthesized the model peptide Fmoc-ALFALFA-CONH2 on a 0.1 mmol scale (100 mg of resin). Based on an initial estimate, we chose to start with a 2 minute DMF wash at 10 mL/min, a 2 minute Fmoc deprotection at 6 mL/min, another 2 minute DMF wash, and a 6 minute room temperature coupling with 2 mmol of activated amino acid delivered at 1 mL/min. This sequence yielded highly pure material, enabling peptide synthesis in 12 minutes per residue. To achieve a maximal concentration of activated amino acid and rate of amide bond formation, coupling solutions were prepared by dissolving amino acids in one equivalent of 0.4 M HBTU in DMF. The activating base was added immediately before use, giving a final concentration of activated amino acid of about 0.3 M. This concentration of coupling reagent was used for all experiments, including those with HATU activation.

Based on our initial investigations and prior reports, we decided to carry out all subsequent studies at 60 °C to minimize the cycle time without significantly increasing formation of side products[61,62]. To consistently and quickly bring reagents to 60 °C, a heat exchanger was placed between the synthesis vessel and the luer-lock quick connect. A 5' (1.6 m) coil of open tubing with a 1/16" (1.59 mm) outer diameter and 0.030" (0.76 mm) inner diameter was used. This preheat loop was immersed with the reaction vessel in a water bath maintained at 60 °C and effectively increased the temperature of solvent from 18 °C (RT) to 59 °C, as measured by a thermocouple inserted into the outlet of the loop. PFA tubing was effective at flow rates up to 20 mL/min; stainless steel was used for all experiments at higher flow rates.

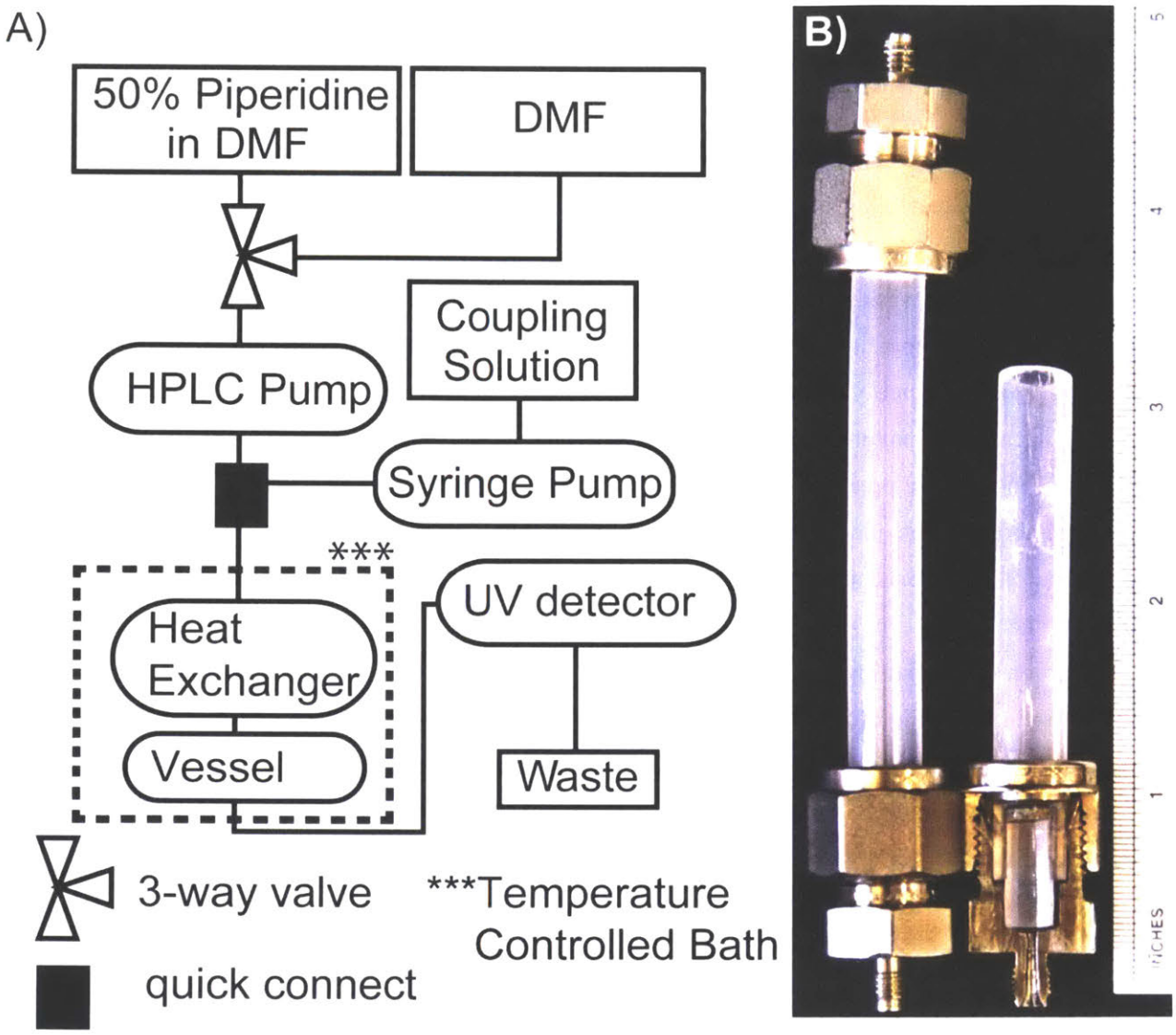


Figure 2.2.1 Manual flow platform for Fmoc SPPS. A) Schematic of the synthesizer. The synthesis vessel can be placed in a temperature-controlled bath. **B)** Photograph of the assembled reaction vessel, left, and a cutaway showing the down-stream components, right. Photograph has been contrast enhanced.

2.3 Determination of minimum cycle time

To determine minimal cycle times, each step of peptide synthesis was studied. First, the time required to wash amino acid solution out of the reaction vessel was investigated as a function of flow-rate via the UV absorbance of the effluent. At 1 mL/min, about 16 minutes (16 mL) were required to remove 99% of the amino acid precursor. As flow rate increased, however, the amount of solvent required decreased, with only 1 minute (10 mL) required at 10 mL/min. To guarantee an effective wash, 20 mL of DMF were used. These 20 mL were delivered over two minutes at 10 mL/min, as the first generation reaction vessel could not reliably accommodate higher flow rates. Analysis of the crude peptides did not show double incorporation of amino acids, which might result from an inadequate wash. Increasing the wash volume did not improve the crude peptide quality.

We then investigated the rate of Fmoc removal by monitoring formation of the UV active dibenzofulvene-piperidine adduct. We chose to use 50% (v/v) piperidine in DMF over the more common 20% (v/v) because our preliminary work indicated it removes Fmoc more rapidly. To decouple the time spent removing the Fmoc protecting group and the time required to wash the byproduct from the resin, deprotection reagent was delivered briefly, the resin was washed, and more deprotection reagent was delivered. A second UV absorbance peak indicated the formation of additional dibenzofulvene-piperidine, and an incomplete initial deprotection. The initial deprotection was performed at 10 mL/min for 60, 30, 15, or 6 seconds. A second peak was observed after a 6 second deprotection, but not a 15, 30, or 60 second deprotection. The rate of Fmoc removal is reported to be sequence dependent[55], so, to ensure robust deblocking, the deprotection reagent was delivered for 20 seconds at 10 mL/min.

With wash and deblocking conditions, the time required for robust amide bond formation was determined by synthesizing two model peptides, LYRAG-CONH2 and Fmoc-ALF-CONH2. Each peptide was synthesized five times, and, for each synthesis, 0.3 M amino acid solutions were coupled for a nominal time of 90, 45, 30, 15, or 7 seconds (Figure 2.3.1). For syntheses with 90, 45, and 30 second couplings, 2 mmol of each amino acid were used. Since the syringe pump could not infuse 6 mL (2 mmol, 0.3 M) of amino acid solution in less than 30 seconds, for the synthesis with 15 second couplings, 3 mL (1 mmol) was used, and 1.2 mL (0.4 mmol) was used for the synthesis with 7 second couplings. For Fmoc-ALF-CONH2, we found no significant difference in the quality of the crude product as a function of coupling time. For LYRAG-CONH2 we observed a significant increase in the Arg deletion peptide when all residues were coupled for 7 seconds. Based on these results, a 30 second coupling time was selected. The final synthetic timeline used with the first generation reaction vessel is shown in Figure 2.3.1C.

To further explore our approach under a variety of synthetic conditions, we studied the synthesis of ACP(65-74). This peptide serves as a model to validate new peptide synthesis protocols, as it is considered difficult to prepare[37,61,63,64]. The main synthetic impurity is a chromatographically resolved Val deletion. The LCMS data for the synthesis of ACP(65-74) with our methodology, as well as two controls, is shown in Figure 2.3.2. Using our protocol and the HATU coupling agent, a minor Val deletion product was observed. When using HBTU, more Val deletion was observed; this is consistent with prior reports[61]. ACP(65-74) synthesized with our protocol, but at room temperature, showed large Val and Gln deletions, confirming that reaction temperature is important. No major differences between the product composition from this flow-based room temperature synthesis and an analogous room temperature batch synthesis following the same synthetic timeline were observed.

Next, two additional "difficult" peptides were prepared: a conotoxin variant and a fragment of the HIV-1 protease[54]. The initial syntheses of these peptides yielded several products of equal molecular weight, which were determined to result from racemization of cysteine during activation. Model studies using the peptide GCF which has a resolved diastereomer, GcF, identified several conditions that produced less than 1% diastereomer and maintained the same cycle time (See Appendix 1, Figure A1.5). This level of racemization is consistent with literature protocols for Fmoc synthesis[65]. Using modified activation conditions for Cys, the peptides shown in Figure 2.3.3 were prepared on a 0.1 mmol scale using the first generation synthesis vessel and cycle.

To demonstrate the suitability of flow-based peptide synthesis for ligation based protein synthesis, it was applied to the synthesis of a 58 residue tri-helical protein based on the Z domain of protein A (referred to as the affibody). The synthetic strategy uses peptide-hydrazides as thioester precursors. Peptide hydrazides can be oxidized with $NaNO_2$ to form a C-terminal acyl-azide, which will react with a thiol to form a peptide thioester suitable for use in native chemical ligation[66,67]. The LCMS data for the crude synthetic peptides are shown in Figure 2.3.4. We purified each peptide, synthesized the affibody according to the strategy in Figure 2.3.4A, and isolated highly pure, full-length affibody after purification (Figure 2.3.4E). The fragments were produced and cleaved from the resin in one day. In contrast, production of similar fragments with optimized Boc in situ neutralization protocols required more than three days, and yielded crude peptides of similar quality (Figure A1.7).

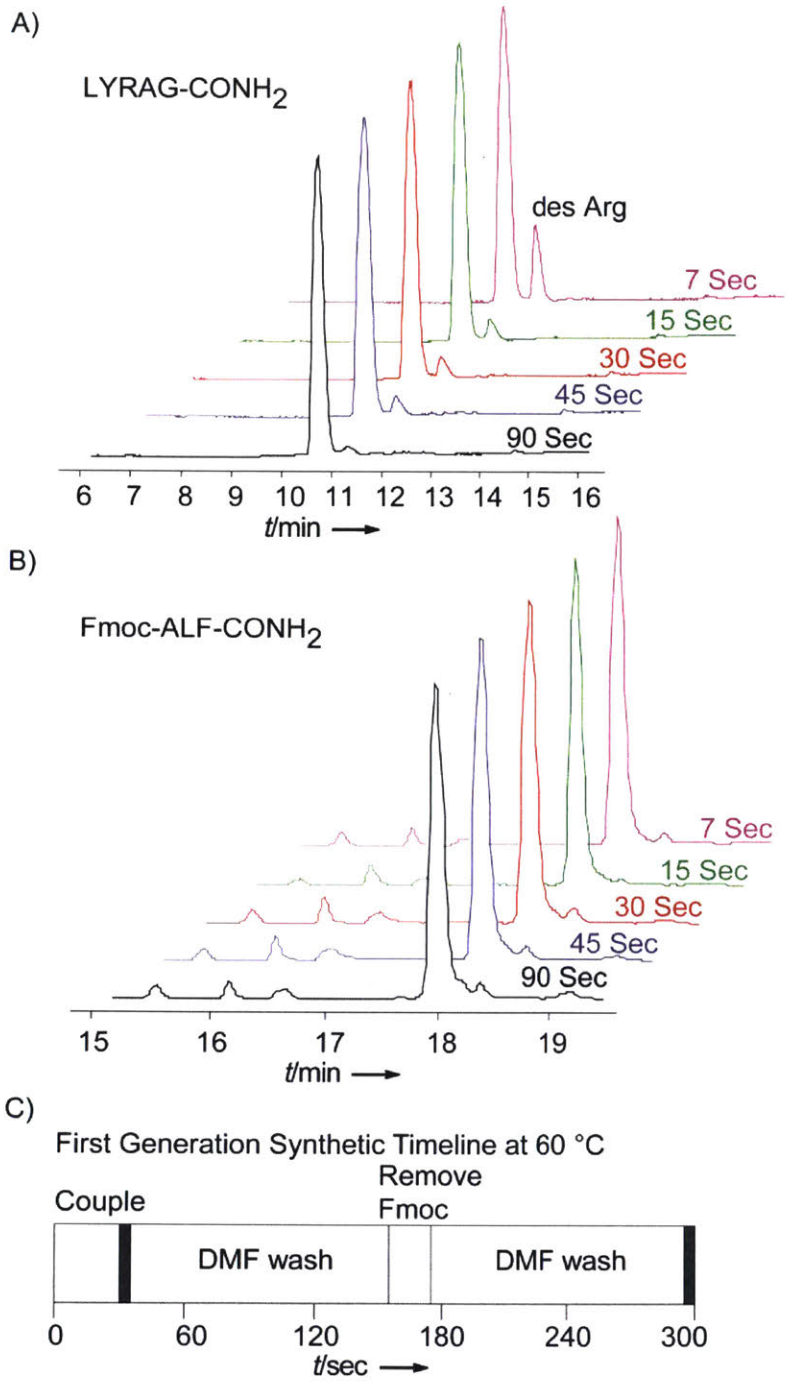


Figure 2.3.1 Investigation of coupling time. A) LC data for crude LYRAG-CONH2 produced by coupling every amino acid for a nominal 90, 45, 30, 15, or 7 seconds. The amount of Arg deletion peptide was greatest for 7 second coupling. **B)** LC data for the synthesis of Fmoc-ALF-CONH2 under the same conditions, showing no change in peptide quality with reduction in coupling time. **C)** The final synthetic timeline used with the first generation reaction vessel, where grey bars represent the time required to move the quick connect. An amino acid residue is incorporated every 300 seconds. The total ion current is displayed in each chromatogram.

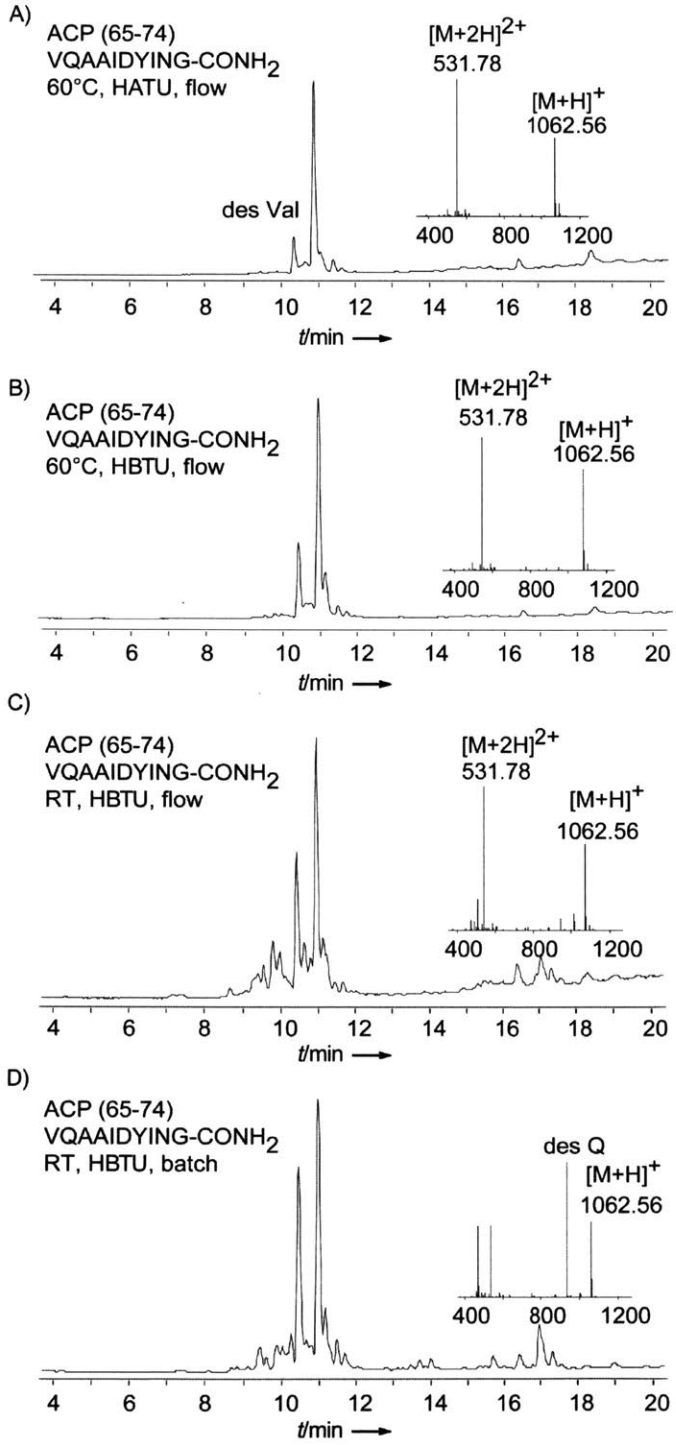


Figure 2.3.2 ACP(65-74) model studies. The synthetic timeline in Figure 2.3.1C was used. Crude LCMS chromatograms for ACP(65-74) synthesized at **A)** 60 °C using HATU as an activator, **B)** 60 °C using HBTU as an activator, **C)** RT using HBTU as an activator, and **D)** RT using a manual batch method following the same timeline. The total ion current is displayed in each chromatogram.

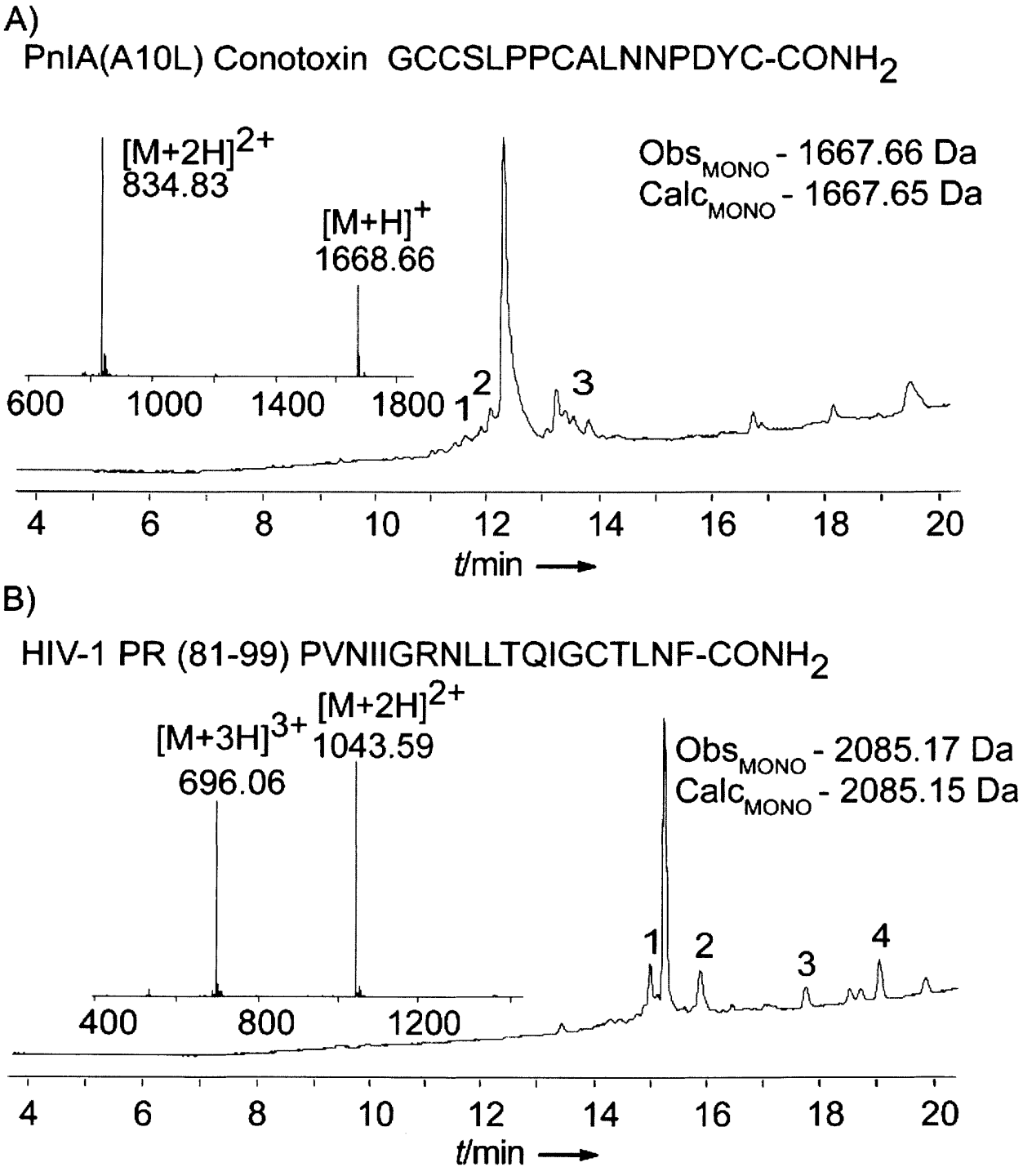


Figure 2.3.3 Synthesis of difficult peptides. LCMS data for the crude peptides: **A)** PnIA(A10L) (1 = Cys deletion, 2 = Cys deletion, 3 = incomplete side-chain protecting group removal) and **B)** HIV-1 PR (81-99) (1 = peptide truncation at Arg and 2,3,4 = incomplete side-chain protecting group removal). The total ion current is displayed.

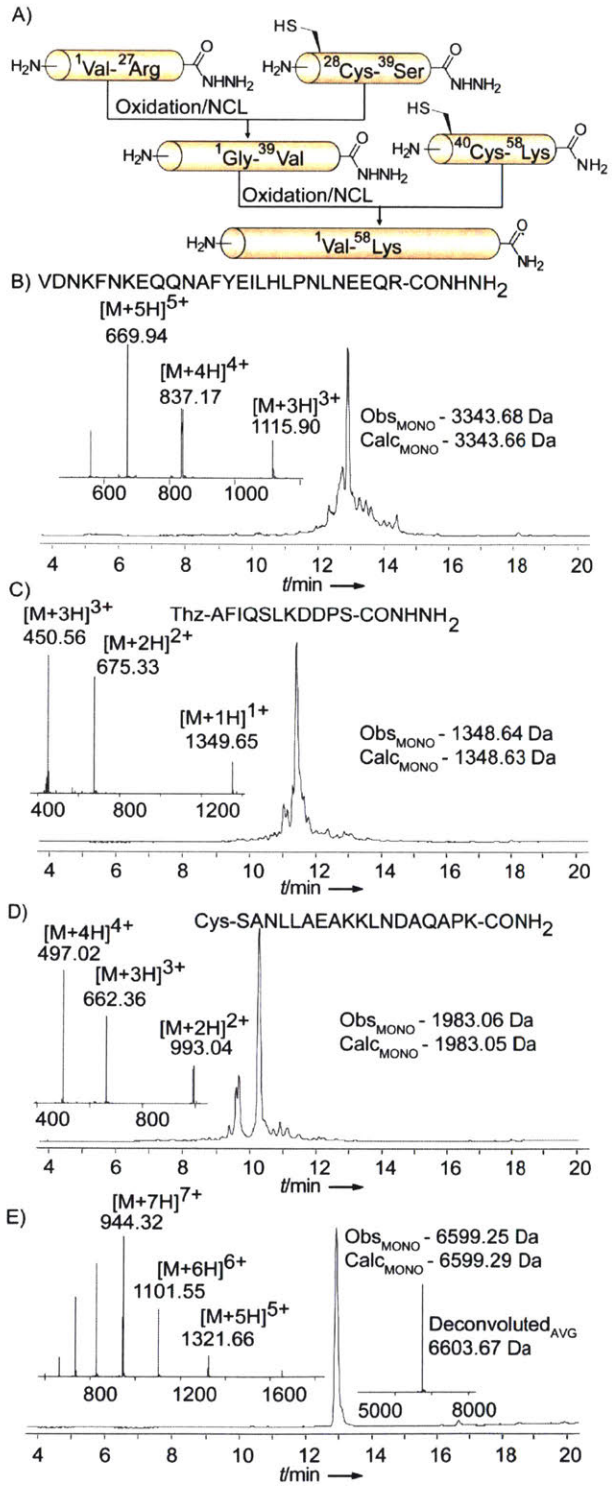


Figure 2.3.4 Synthesis of the affibody. A) Synthetic scheme used to produce the affibody **B)** Crude fragment [1-27]-CONHNH2 **C)** Crude fragment Thz-[28-39]-CONHNH2 **D)** Crude fragment Cys-[40-58]-CONH2 **E)** Purified full-length affibody. Listed observed and calculated masses are monoisotopic. The total ion current is displayed in each chromatogram.

2.4 Acceleration and scale-up

After the completion of these model studies, we sought to increase synthetic scale and decrease cycle time. All attempts to increase the flow rate or add more resin to the first generation reaction vessel were thwarted by rapidly increasing backpressure. A high pressure stainless steel reaction vessel was constructed to study the effect of simply providing more pressure to maintain a high flow rate, but this resulted in extrusion of the resin through the frit, as has been previously observed[45].

Therefore, the second generation synthesis vessel shown in Figure 2.4.1 was constructed. The diameter is twice that of the first generation and volume limiting inserts restrict the volume to 2 mL, about the same as the first generation. Increasing the diameter of the vessel drastically reduced the backpressure, and maintaining a comparable volume allowed the same volumes of solvents and reagents to be used. This vessel accommodated up to 200 mg of resin, and flow rates up to 100 mL/min. More resin should not be used, because the resin swells as the peptide is elongated and the volume limiting inserts restrict the swollen volume to 2 mL. Long peptides may result in reactor failure if more than 200 mg of resin is used.

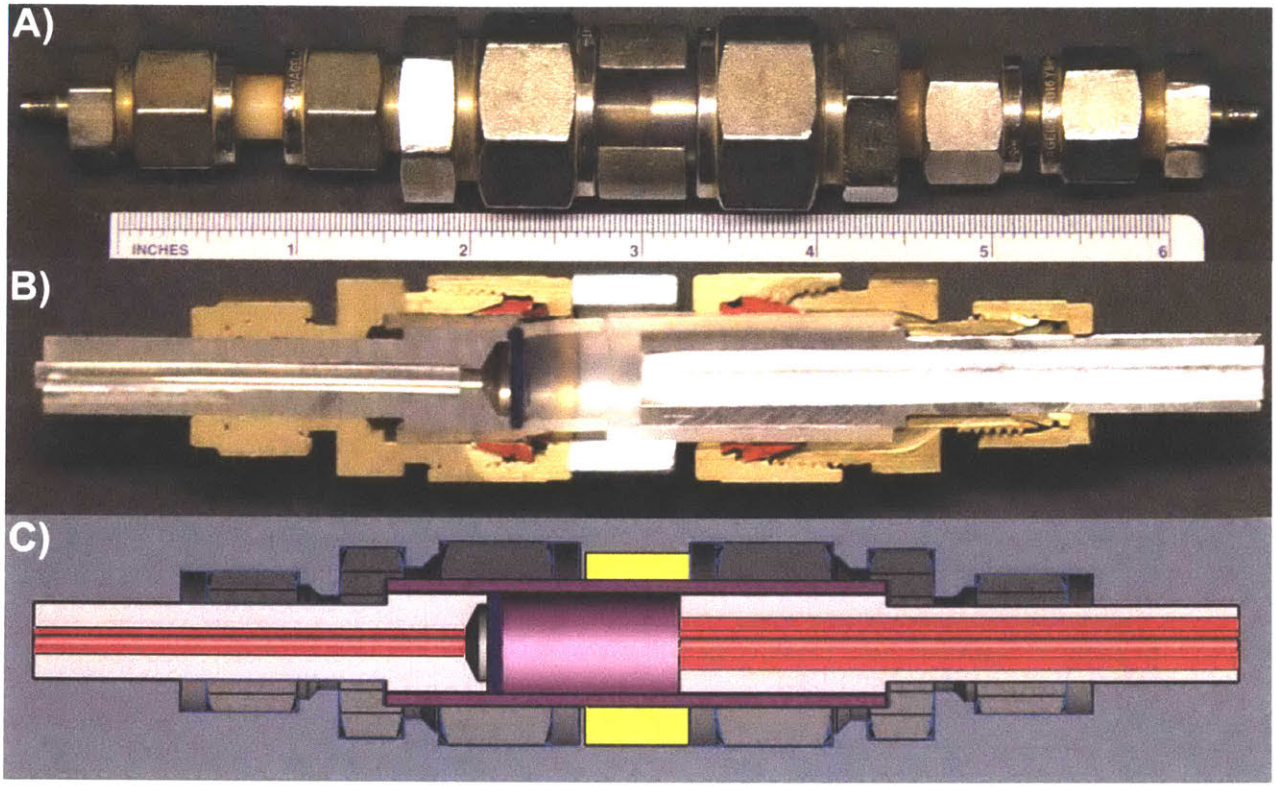
With this new reaction vessel in hand, cycle time was reduced by washing at a higher flow rate. As expected, the required wash time continued to decrease with increasing flow rate, with 99% of the amino acid removed in 36 seconds at 20 mL/min and 20 seconds at 40 mL/min. To allow operators adequate time to prepare each subsequent amino acid, one minute washes at 20 mL/min were used. After the instrument became available for general laboratory use, users often reduced the wash time to 30 seconds at 20 ml/min with no adverse effect on synthesis.

To verify that the performance of the second generation synthetic protocol and vessel, using faster cycles and fewer equivalents of amino acids, was comparable to that of the first generation synthetic procedure, the syntheses of ACP(65-74), the conotoxin, the HIV-1 protease fragment, and the affibody fragments were repeated on 200mg of resin with the synthetic timeline in Figure 2.4.1D. The peptides were of comparable crude quality (Figures A1.9 to A1.11).

Dozens of unique peptides were made for diverse applications, and almost all were of high crude purity. Several representative case studies are included here and in Appendix 1 (Figures 2.4.2, A1.12 to A1.14). This method is sufficiently robust that all of these peptides were synthesized without UV monitoring of the reactor effluent. Figure 2.4.2 shows a library of model cysteine containing peptides, including some of low crude purity. Appendix 1 shows additional, longer analogues (Figure A1.13), a library of 10 model glutathione-like peptides prepared in a single day (Figure A1.12), and two biotinylated protease recognition sites (Figure A1.14) all of which are of high crude purity. In all cases,

peptides were produced on a 0.2 mmol scale, the major peak is the desired product, and crude material was successfully purified in one preparative RP-HPLC step.

This method was also used to produce peptide fragments of two additional proteins, the 113 residue B. a. RNAse and the 130 residue pE59 DARPin. Following rapid optimization of the synthetic protocol, the full length proteins were obtained. For a detailed account of the synthesis of these targets, including effective and ineffective methods of suppressing aspartamide formation, deletions, methionine oxidation, and premature termination of the peptide chain, refer to the report by Mong et al[68].



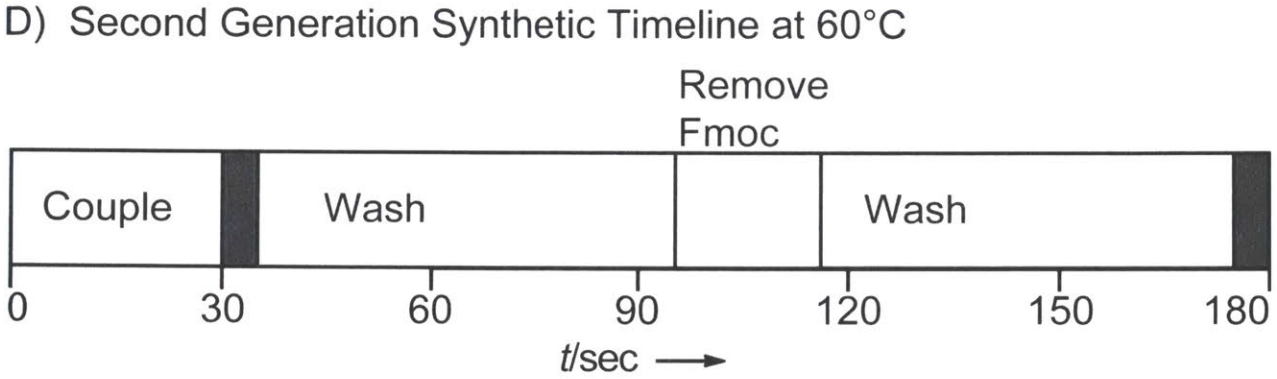


Figure 2.4.1 Second generation reaction vessel A) Assembled unit after hundereds of syntheses **B)** Cutaway with non-wetted parts shown in brass, the frit in blue, and large ferrules in red. The image has been color enhanced and background objects removed. **C)** False color drawing of the cutaway with non-wetted parts shown in dark gray and the frit in blue **D)** The final synthetic timeline used with the second generation reaction vessel; grey bars represent the time required to move the quick connect. An amino acid residue is incorporated every 180 seconds.

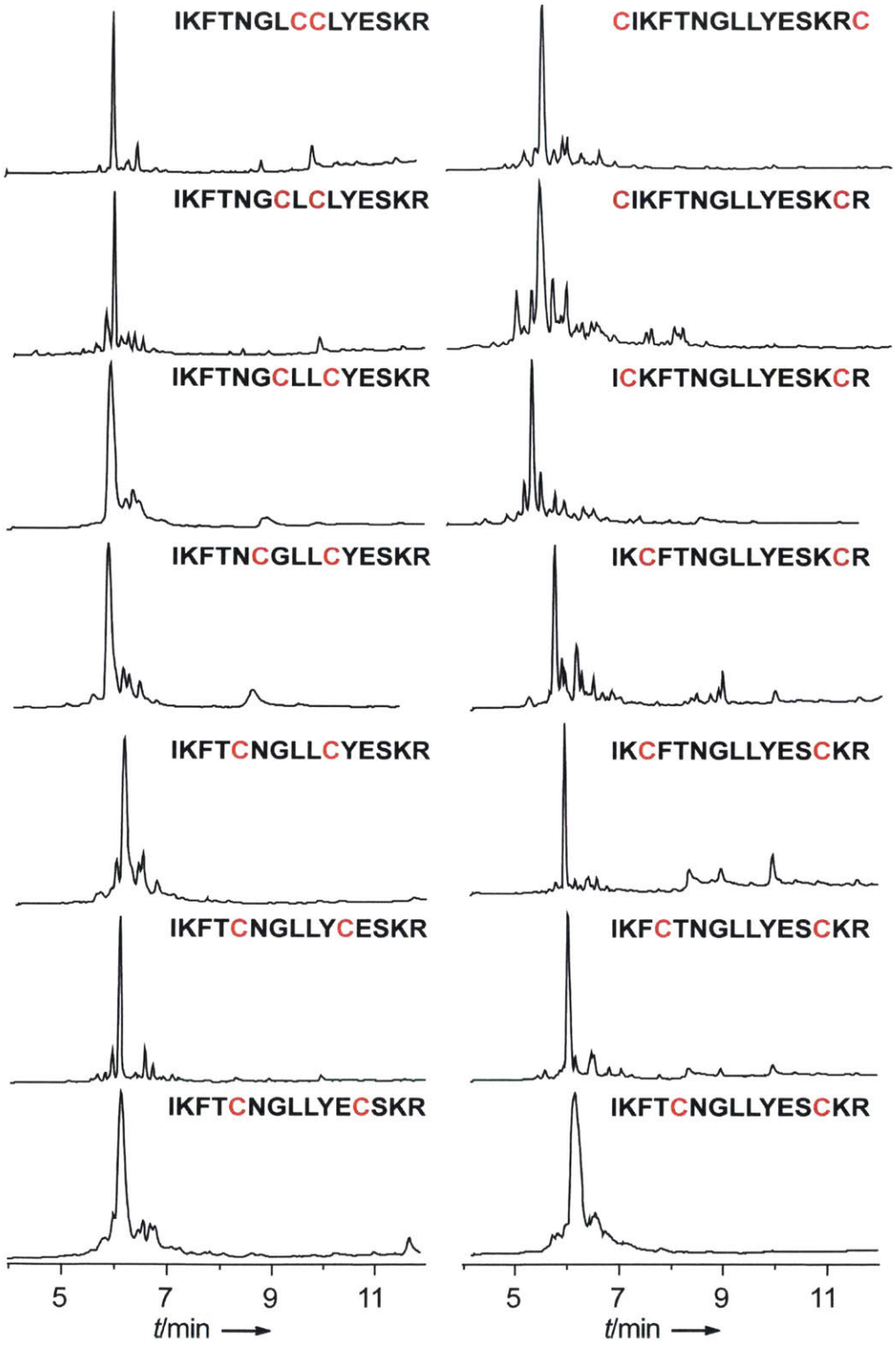


Figure 2.4.2 Model 15-mers for cysteine macrocyclization research synthesized in 45 minutes each. Cysteine is highlighted in red and total ion chromatograms are shown. All peptides were produced as C-terminal amides, and, in all cases, the main peak was the desired mass (1772.9 Da observed and calculated monoisotopic).

2.5 Conclusions

The manual fast flow peptide synthesizer described here rapidly generated high quality peptides and provided a strong demonstration of the scientific principles used to accelerate peptide synthesis in flow. In all cases, the desired product was obtained and readily isolable by RP-HPLC. Based on published syntheses of ACP(65-74), the conotoxin, and the HIV protease fragment[54], our experiments with manual Boc syntheses of the affibody fragments, and the synthesis of ACP(65-74) at room temperature in batch and under flow, we believe that crude material obtained with this method is nearly identical to material obtained from traditional batch processes.

Extensive work has shown that SPPS chemistry can be fast and efficient at elevated temperatures[36]. The manual fast flow method described here further accelerates SPPS chemistry, far beyond what is currently possible with microwave assisted or other rapid peptide synthesizers, by leveraging a flow based approach. In addition to constantly supplying high concentration reagents, the flow-based platform overcomes a number of significant obstacles that hinder standard and microwave-assisted approaches. First, the completely sealed reaction vessel and heat exchanger can be immersed in a temperature controlled bath which allows solvents and reagents to be heated in a consistent and controlled manner immediately before reaching the resin bed. Rapid preheating is crucial to avoid thermal degradation of reagents[58] while quickly reaching the desired temperature, but this is extremely difficult in a batch system. Second, the flow platform can be scaled without increasing the cycle times. As demonstrated in the transition from the first to second generation reaction vessels, increasing the diameter and flow rate effectively increases the maximum scale, without slowing the synthesis. Third, stirring is not required to effect adequate mass transfer, eliminating failure-prone moving parts and facilitating scale up. Finally, high quality peptides can be obtained quickly without double coupling, double deprotection, or colorimetric tests of coupling efficiency. During our studies with ACP(65-74), we observed no decrease in the Val deletion peptide after double coupling Val and double deprotecting the preceding Gln, and these results are consistent with our experience optimizing the fragments of B. a. RNAse and the DARPin[68]. Such additional steps are often employed in batch syntheses, hindering synthetic progress.

Fast flow synthesis enables the efficient, rapid Fmoc synthesis of polypeptides and provides a reproducible and systematic study of flow-based Fmoc SPPS chemistry at elevated temperature. To prove the synthetic utility and reliability of this system, we produced dozens of peptides suitable for various applications, including three fragments of an affibody that were successfully ligated to produce a full length protein.

By reanalyzing flow-based SPPS[45,46,59,60] and carefully designing a new system, we were able to overcome two longstanding challenges preventing rapid peptide synthesis. First, we were able to reduce

the wash time from several minutes to one minute or less by minimizing the volume of the system. Although it is often not discussed, washing the resin requires a significant investment of time and solvent in most SPPS systems. With this method, wash times and solvent usage are significantly less than previous systems, but washing remains a key challenge in further accelerating the cycle.

Second, we eliminated the extremely high capital and maintenance costs of a microwave heating system by employing a simple, effective heat exchanger in a water bath. As a UV detector is not essential, this method can be reproduced for about $1000 with a used pump. Furthermore, all components of the system can be serviced by a non-expert, drastically reducing instrument downtime. We believe lowering the cost and complexity of rapid peptide synthesis is a major step towards its general adoption by chemists. Similarly, the reported system uses the cheapest, most common peptide synthesis chemistry. All reagents are commercially available, and no extraordinary hazards are associated with chain assembly or cleavage.

This platform was used to produce of dozens of peptides, and support the total synthesis of the 113 residue B. a. RNAse and a 130 residue DARPin[68]. Importantly, during the synthesis of the fragments of these proteins, most of the major side reactions in peptide synthesis were encountered. The cited work[68] provides an extensive description of procedures to overcome these side reactions, and constitutes a detailed tutorial for sequence specific optimization of long and complex peptides using our flow-based synthesizer.

In conclusion, we developed a simple, rapid, highly robust peptide synthesis platform. The system can be easily and cheaply assembled, then leveraged to generate high quality peptides. Ultimately, this provides a guide for chemists inexpert in peptide and protein synthesis to quickly and independently carry out total syntheses of these complex biomolecules without the need for sophisticated tools, reagents, or equipment. Moreover, the scientific principles of fast flow peptide synthesis were well validated, encouraging automation and further acceleration of peptide synthesis.

Automated fast flow peptide synthesis

The work presented in this chapter is currently under review, and Appendix 2 is derived from the supporting information

Mark D. Simon, Alex Mijalis, Kyle A. Totaro, Daniel Dunkelmann, Alexander A. Vinogradov, Chi Zhang, Yuta Maki, Justin Wolfe, and Bradley L. Pentelute. Automated fast flow peptide synthesis. *(Under Review)*

3.1 Introduction

In this chapter, the scientific principles validated with the manual fast flow peptide synthesizer were used to create an automated fast flow peptide synthesizer (AFPS) that was highly flexible in the choice of activation and deprotection for each amino acid, continuously charged with stably stored reagents, and incorporated amino acid residues in as little as 40 seconds at temperatures up to 120°C. The instrument was controlled through a graphical user interface, and operated across a variety of scales. The results of this work highlight the interplay of chemical, mechanical, and control variables in the construction of an automated instrument to perform rapid flow based chemistry on a solid support (Fig 3.1.1).

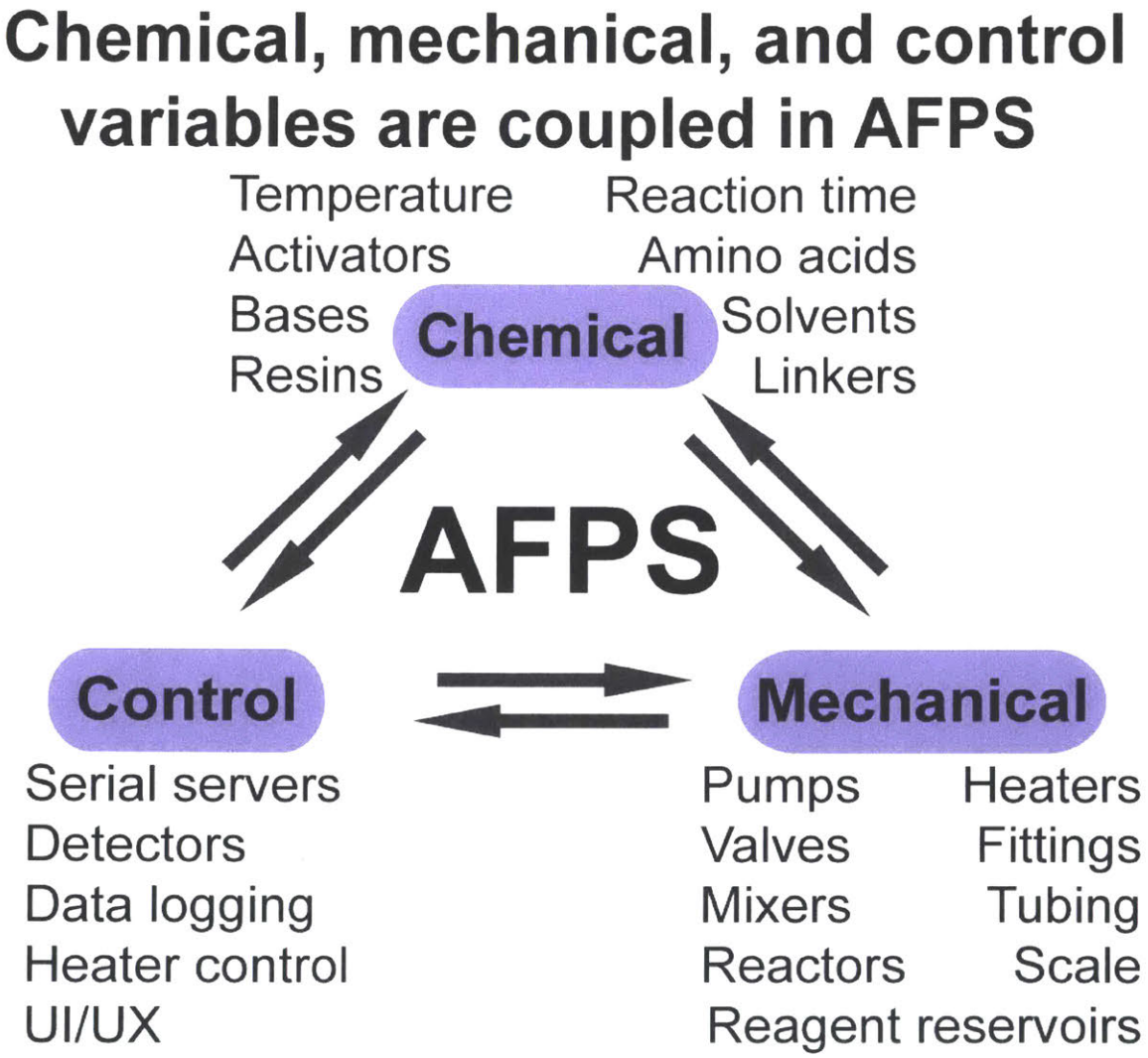


Figure 3.1.1 Automated fast flow peptide synthesis was enabled by careful consideration of the interplay between chemical, mechanical, and control variables.

3.2 Summary

In the course of designing and testing four generations of automated fast flow peptide synthesizers (AFPS) over five years, we found that the mechanical and chemical problems encountered were closely coupled. As solutions to chemical problems became more sophisticated, hardware and software was needed to match that sophistication. With each generation of AFPS, we learned important chemical and mechanical lessons that eventually culminated in the fourth generation AFPS. Along the way, we elucidated several critical principles for fast flow solid supported synthesis.

3.2.1 Mechanical principles

The mechanical transition from one reagent to the next through a wash solvent must be extremely rapid since these transitions and washes dominate the synthetic timeline. In increasing the speed of transitions between reagents, however, it is import to maintain correct stoichiometry of reagents, both during the transitions and in a steady state if it is ever established. In contrast to true continuous flow synthesis, reaction byproducts generated prior to mechanical equilibrium are bound to the solid support and carried through the entire synthesis, reducing yield and contaminating the final product. It is therefore essential to maintain proper stoichiometry at all times.

3.2.2 Chemical principles

Reagents must be stable in storage for much longer than in a typical manual synthesis setting. Solutions of Fmoc protected amino acids codissolved in DMF with HBTU degraded to give tetramethyl urea in solution and truncated peptides during synthesis. Without HBTU, N-α unprotected amino acids slowly precipitated as Fmoc protection was lost. These problems were overcome by first separating amino acids and activators, then dissolving amino acids in NMP or DMI instead of DMF to preserve the Fmoc protection. In both steps, careful attention to mechanical detail was required to ensure that stoichiometry was maintained throughout synthesis.

Further, reagents, including resin bound intermediates, must be stable during use. We found Fmoc-His(Trt)-OH rapidly epimerized at elevated temperature to yield the racemic monomer, as has been reported[62,69,70]. This was suppressed by Boc protecting the imidazole ring, inductively deactivating it. Surprisingly, the peptidyl resin appeared completely stable at elevated temperature, and synthetic outcomes improved monotonically as temperature was increased up to 120°C, the maximum investigated.

3.2.3 User interface principles

To fully leverage the flexibility of the AFPS and prevent operator errors, the user interface was carefully designed. Reagents were arranged so they were easy to access and clearly labeled, and software was written with a convenient graphical interface. This proved critical to driving adoption of the instrument within our laboratory, and enabled non-expert users to effectively make peptides.

3.2.4 Data analysis method

Over the course of five years of development, the various versions of the AFPS were used to make many different peptides in support of diverse projects in our laboratory. To usefully compare one version of the AFPS to another based on these large data sets of distinct peptides, an automated, peptide agnostic processing technique was used. This method was based on molecular feature extraction (MFE) from analytical LC/MS data, and captured three major features of crude synthetic peptides: their LC/MS purity, the impurities of comparable molecular weight to the desired compound, and impurities less than 90% of the molecular weight of the desired compound.

The LC/MS purity of peptides was the primary metric of success, called the *estimated purity* below. Impurities of comparable molecular weight included commonly reported peptide synthesis byproducts [22,71–75]that are frequently the target of peptide specific parameter optimization, such as cleavage adducts, deletions, and aspartimide derived products. Impurities with molecular weight 90% or less of the desired peptide were largely peptidic truncates, with or without a capping moiety – products that reflect on the fundamental performance of the instrument design, and are difficult to suppress by optimizing parameters. These products are collectively called *truncates* below.

With dozens to hundreds of syntheses performed on each generation of AFPS, the average, aggregate performance was well characterized with this analytical technique, and could be used to compare generations of the AFPS, even when specific compounds and impurities could not.

3.2.5 Outcome

Taken together, these insights have resulted in an extremely rapid, fully automated peptide synthesizer that we believe can be adapted to accelerate solid phase synthesis of many compounds.

3.3 First generation automated fast flow peptide synthesis

3.3.1 Key findings

The basic design in Figure 3.3.1 was validated, but it became clear that mechanical components, especially valves and pumps, were needed that have low volume and arrange amino acids symmetrically about the HPLC pump. Without careful attention to mechanical detail, transitions from one reagent to another dominated the synthetic timeline, and asymmetrically arranged amino acids led to variable stoichiometry of reagents across amino acids.

3.3.2 Design of first generation AFPS

The first generation automated fast flow peptide synthesizer (AFPS), first reported with our manual method[34], attempted to directly translate the success of manual fast flow synthesis into an automated format. To this end, the instrument shown schematically in Figure 3.3.1A was assembled. Rather than replace a human operator with an anthropomorphic robot, two HPLC pumps were used to deliver reagents to the solid support contained in the resin vessel. One delivered DMF wash solvent, piperidine for Fmoc removal, or amino acids and activators codissolved in DMF for the coupling. The other HPLC pump infused DIEA before the combined stream was passed through a static mixer to the resin vessel via a preheat loop. The preheat loop and resin vessel were heated in a water bath to rapidly bring solvents and reagents to the desired temperature, and accelerate amide bond formation and Fmoc removal. A UV spectrometer provided real time information on synthetic quality before the reaction vessel effluent was passed to waste.

This design was directly analogous to a user preparing amino acids codissolved with activators before starting a manual synthesis, and mixing in DIEA immediately before each coupling. The UV detector, preheat loop and resin vessel from the manual method were reused without modification[34].

The arrangement in Figure 3.3.1 had the advantage of relying on valves and pumps, which are robust and easy to control, rather than a complex robotic arm. Furthermore, by drawing from large stocks of amino acids, individual aliquots of reagents did not have to be prepared for each residue coupled, which saved significant labor.

3.3.3 Characterization of first generation AFPS

Although we had some initial success using the synthetic timeline shown in Figure 3.3.1B, several major problems were identified, and few peptides were made beyond very simple model peptides. Most critically, transitions between reagents in this automated instrument were slow. The custom valve

manifolds supplied by Controls for Automation (CFA) had large internal volume, and there was substantial non-displaced volume in the heads of the Knauer HPLC pumps used. Ultimately, 40mL of solvent were required to wash the manifold, pump head, and resin, but only 10mL were required to wash the resin alone. The problem was so severe that 84% of the total synthesis time was spent washing (90 seconds out of every 107 second cycle).

Additionally, the valves were arranged asymmetrically about the outlet of the manifold, giving amino acids further from the outlet broader, wider concentration profiles once selected (Figs 3.3.1C, 3.3.1D, 3.3.2). The large volume within the manifold made the asymmetry quite significant, and caused differences in the stoichiometry of DIEA used to activate each amino acid.

Our initial success with this method, and the limitations of the selected pumps and valves, motivated us to retain the same basic design while solving mechanical problems.

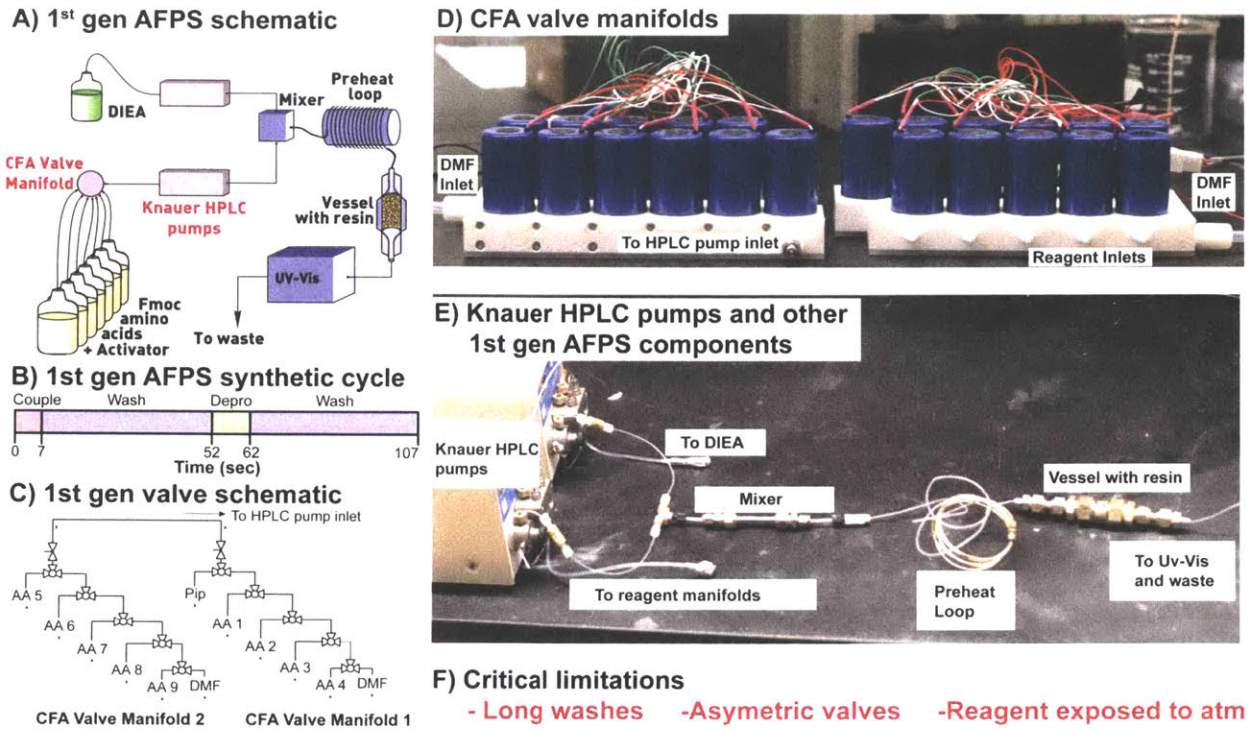


Figure 3.3.1 First generation AFPS A) Schematic representation of the first generation automated fast flow peptide synthesizer (AFPS). Fmoc protected amino acids codissolved with activators were pumped through an HPLC pump to a mixer where a second HPLC pump infused DIEA. The mixed fluid then traveled through a preheat loop that served as an activation reactor to form the active ester. Upon exiting the preheat loop, the newly formed active ester moved through the vessel containing the solid phase and effected amide bond formation. The effluent was passed through a UV-Vis spectrometer and to waste. The resin vessel and preheat loop were heated with a water bath. **B)** First generation AFPS synthetic cycle, run at 50mL/min. **C)** Schematic of the first generation valves. Several custom valve manifolds supplied by CFA were used. **D)** Photograph of CFA valves **E)** Photograph of the first generation AFPS with select components labeled. **F)** The first generation AFPS suffered from long washout times, asymmetric reagent positions within the valves, and exposure of reagents to atmosphere during storage.

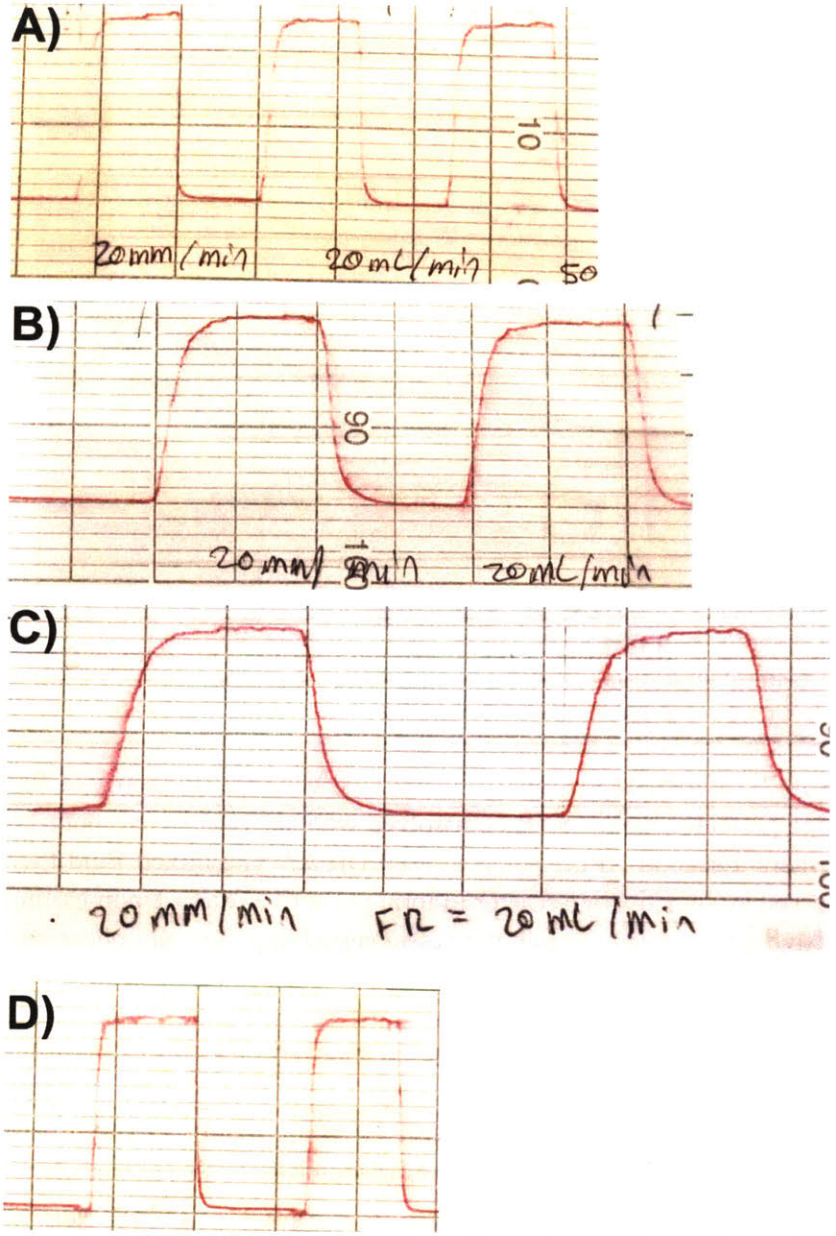


Figure 3.3.2 Asymmetry of valves These traces above show the effect of the asymmetry of the first generation valve design (CFA manifolds). The traces were acquired by pumping 20mL/min of a UV dye (4mM Fmoc-Gly-OH in DMF) through either the first position **A)**, third position **B)**, or fifth position **C)** of the manifold. These traces are from testing of an improved, lower volume variant of the CFA manifolds, but still illustrate the problem. The improved CFA manifolds were never used because the asymmetry was still significant, as illustrated in panels A-C, and severe manufacturing delays became unacceptable. **D)** shows the second generation valve design (Valco manifold) tested under the same conditions (all positions were identical). The Valco manifold was slightly better than the best position on the improved CFA manifolds, and substantially better than the worst position. The strip chart was run at 20mm/min and the distance between vertical lines is 10mm (30 seconds, 10mL of solvent). Photographs of strip charts are shown, and have been brightness and contrast enhanced.

3.4 Second generation automated fast flow peptide synthesis

3.4.1 Key findings

Selection of improved pumps and valves overcame the mechanical issues with the first generation AFPS, but several chemical issues were encountered. Fmoc protected amino acids were not stable codissolved with HBTU in DMF, even in the absence of an exogenous base such as DIEA. Synthesized peptides were contaminated with truncated byproducts, and over time, tetramethyl urea was observed in solution. N-im Trityl protected histidine was particularly unstable when codissolved with HBTU and when used to incorporate histidine at high temperature; these problems were overcome with N-im Boc protection.

3.4.2 Design of second generation AFPS

The second generation automated fast flow peptide synthesizer, shown schematically in Figure 3.4.1A, overcame the mechanical problems of the first generation AFPS with three major design changes. The valves were replaced with symmetric, low volume alternatives from Valco, the pumps were replaced with the Varian 210s used in the manual system, and the reagent reservoirs were maintained under nitrogen to protect them from the atmosphere. Together, these changes allowed us to directly translate the manual fast flow synthesis method into an automated instrument.

3.4.2.1 Valves

After evaluating a number of alternative valving schemes, Valco ChemInert 10 position selector valves with 0.060" flow channels arranged as shown in Figure 3.4.1C were used. In the arrangement shown, all of the amino acid reservoirs were completely symmetric. The length of the tubing between the reagent reservoirs and the first valve was not important, as that volume was an extension of the reservoir, but the volume of the other tubing was minimized by positioning the valves as physically close to the master valve as possible, and positioning the master valve as close to the amino acid pump head as possible (Fig 3.4.1D). The valves had negligible internal volume, and the internal volume of the 1/16" ID tubing used was 50 microliters per inch, resulting in a total swept volume of several hundred microliters on each channel. In contrast, the CFA manifolds had an internal volume of several milliliters on the longest channel. Further, the Valco valves were inexpensive, commercial, robust, field serviceable, and physically much smaller than the CFA manifolds, an important consideration in a crowded fume hood.

In practice, the Valco valves required less than 1mL of solvent to wash (<1s at 50ml/min), compared to 15mL for the CFA manifolds (18 seconds at 50ml/min), which was more than an order of magnitude improvement in the washout time (Figure 3.3.2).

3.4.2.1 Pumps

It was deemed impractical to re-engineer the Knauer pumps used in the first generation automated synthesizer, so we reverted to the Varian 200 series HPLC pumps used in the manual synthesizer, despite the pulsatile flow from their single piston pump heads. The pumps have very low non-displaced volume in the pump head, and generally achieve 99% of maximum concentration of a new solvent or reagent within 4 pump strokes (2 seconds at max flow rate; 1.7mL for 50mL head). This performance was an order of magnitude better than the Knauer pumps, and with a similar improvement in the valves enabled rapid transitions between reagents in the second generation AFPS.

3.4.2.2 Amino acid stocks

The AFPS used large stocks of reagents that were left on the instrument for days or weeks, rather than using freshly prepared reagents for each synthesis. To protect the hygroscopic amino acid solutions from atmospheric moisture, the media bottles used as reservoirs were placed under a positive pressure of dry nitrogen. The nitrogen pressure also suppressed degassing in the manifold and cavitation in the pump heads. Commercial tops we evaluated explicitly warned against building up a positive pressure of inert gas, so the custom insert shown in Figure 3.4.1F was designed for GL45 threaded tops with open caps, and similar inserts were later used with GL32 and 38-400 threaded tops.

3.4.3 Characterization and use of second generation AFPS

With these three improvements, we assembled an automated peptide synthesizer that was conceptually similar to the proof of concept first generation AFPS, but with substantially improved mechanical properties. To validate this design, model peptides ALFALFA and ACP(65-74) were synthesized as follows on the second generation AFPS. Amide bond formation was effected by flowing 2ml/min of DIEA and 20mL/min of ~0.4M HBTU and protected amino acids codissolved in DMF for 14 seconds. Fmoc protecting groups were removed with 20% (v/v) piperidine in DMF at a flow rate of 20ml/min for 20 seconds, and washes were with 20ml/min of DMF for 30 seconds, for a total cycle time of 94 seconds per residue (Fig 3.4.1B). Both ACP(65-74) and ALFALFA were synthesized on 180-190 mg of chlorotrityl hydrazide functionalized 1% divinylbenzene crosslinked polystyrene resin (0.7mmol/g) and were of high quality (Fig 3.4.1G).

The instrument was then made available for general lab use to identify remaining problems before iterating on the design. Two major synthetic campaigns were undertaken with this instrument: synthesis of about 80 cell penetrating peptides ranging in length from 6 residues to 40, and synthesis of 30 diastereomers of the small protein EETI-II (Fig 3.4.2, A2.15.10 to A2.15.40, and A2.15.69 to A2.15.144). The cell penetrating peptides were synthesized on 250- 300mg of amino methyl polystyrene resin (1.1mmol/g) using the Rink amide linker, and had an average estimated purity of 36% (range 5% to 70%). EETI-II diastereomers were prepared on 140-170mg of ChemMatrix resin (0.5mmol/g; Rink amide linker) with an average estimated purity of 26% (range 13% to 36%). Most peptides were successfully isolated by RP-HPLC for further studies, but several syntheses of CPPs failed and had to be repeated with freshly prepared reagents. Overall, three major problems were identified during these efforts.

First, substantial low molecular weight peptidic products were observed: truncates were an average of 28% (range 20%-45%) of the assigned ions from EETI-II analog syntheses and 33% (range 14% to 84%) of the assigned ions from CPP syntheses. Second, Fmoc-His(Trt)-OH co-dissolved with HBTU in DMF degraded rapidly, turning dark brown overnight and entirely failing to incorporate histidine after several days. Finally, it was difficult to refill the regent reservoirs, with frequent spills, and occasional synthetic failures due to insufficient reagent. Each of these problems was studied in detail to inform a third generation AFPS.

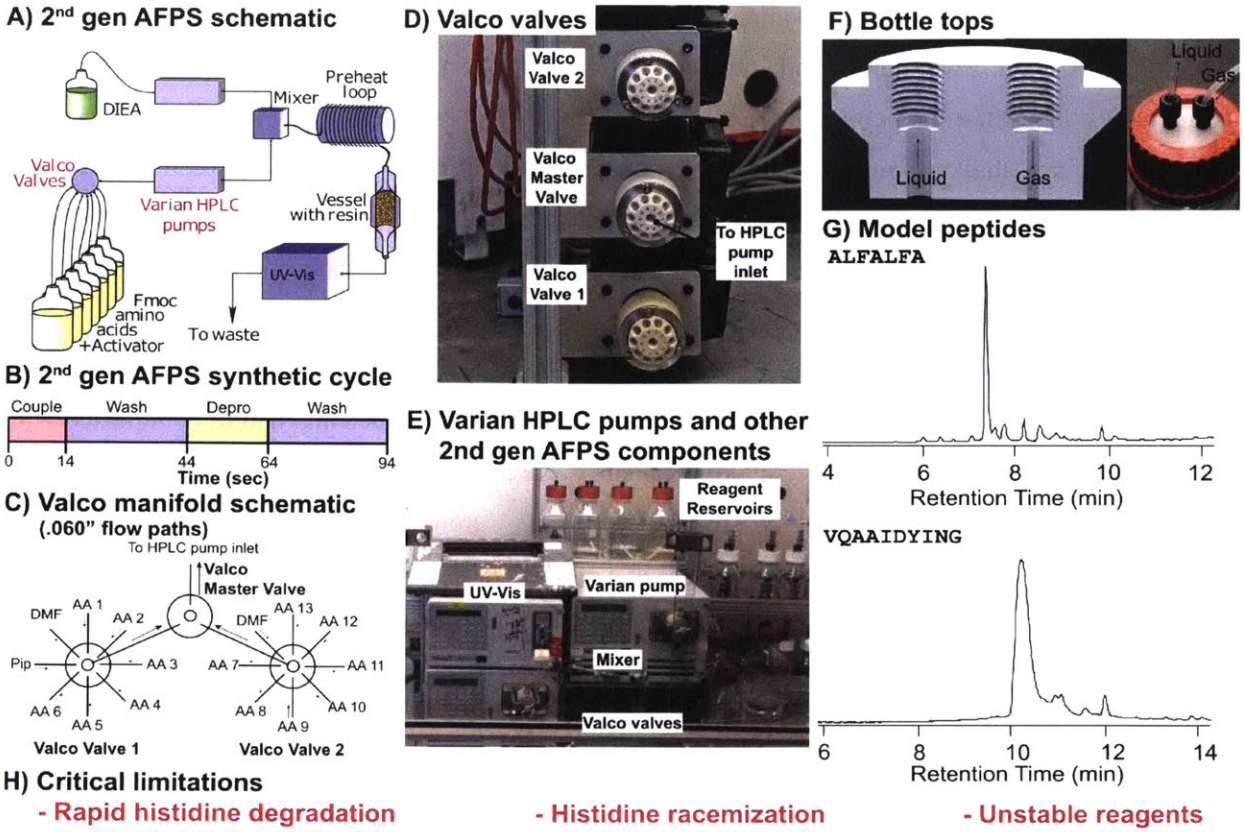


Figure 3.4.1 Second generation AFPS A) Schematic representation of the second generation automated fast flow peptide synthesizer (AFPS). Fmoc protected amino acids codissolved with activators were pumped through an HPLC pump to a mixer where a second HPLC pump infused DIEA. The mixed fluid then traveled through a preheat loop that served as an activation reactor to form the active ester. Upon exiting the preheat loop, the newly formed active ester moved through the vessel containing the solid phase and effected amide bond formation. The effluent was passed through a UV-Vis spectrometer and to waste. The resin vessel and preheat loop were heated with a water bath. The first and second generation AFPS differed primarily in the valves, pumps, and reagent reservoirs. **B)** Second generation AFPS synthetic cycle, run at 20mL/min. **C)** Schematic of the second generation valves. Several 10 position selectors were used, but two valves with 8 positions each are drawn for simplicity. The master valve was also a 10 position selector, with unused ports plugged. All reagent reservoirs were positioned symmetrically, and the volume was minimal. **D)** Photograph of Valco valves **E)** Photograph of the second generation AFPS with select components labeled. **F)** Left: Rendering of a sectioned GL45 bottle top insert showing a liquid (reagent) withdraw port and gas supply port. The liquid tube is inserted to the bottom of the bottle through the 1/8" hole and the gas supply tube seats against the bottom of the gas port. Both ports use standard ¼-28 flat bottom fittings. Right: Photograph of the insert installed with a PFA liquid withdraw tube and a PE gas supply tube. The red open top compresses the insert against the glass, forming a seal. **G)** Model peptides ALFALFA and ACP(65-74) synthesized on the second generation AFPS under default conditions on 180-190mg chlorotrityl hydrazide resin (0.7mmol/g). **H)** The second generation AFPS suffered from degradation of the reagents generally, and a very rapid degradation of histidine. Histidine also racemized at elevated temperature.

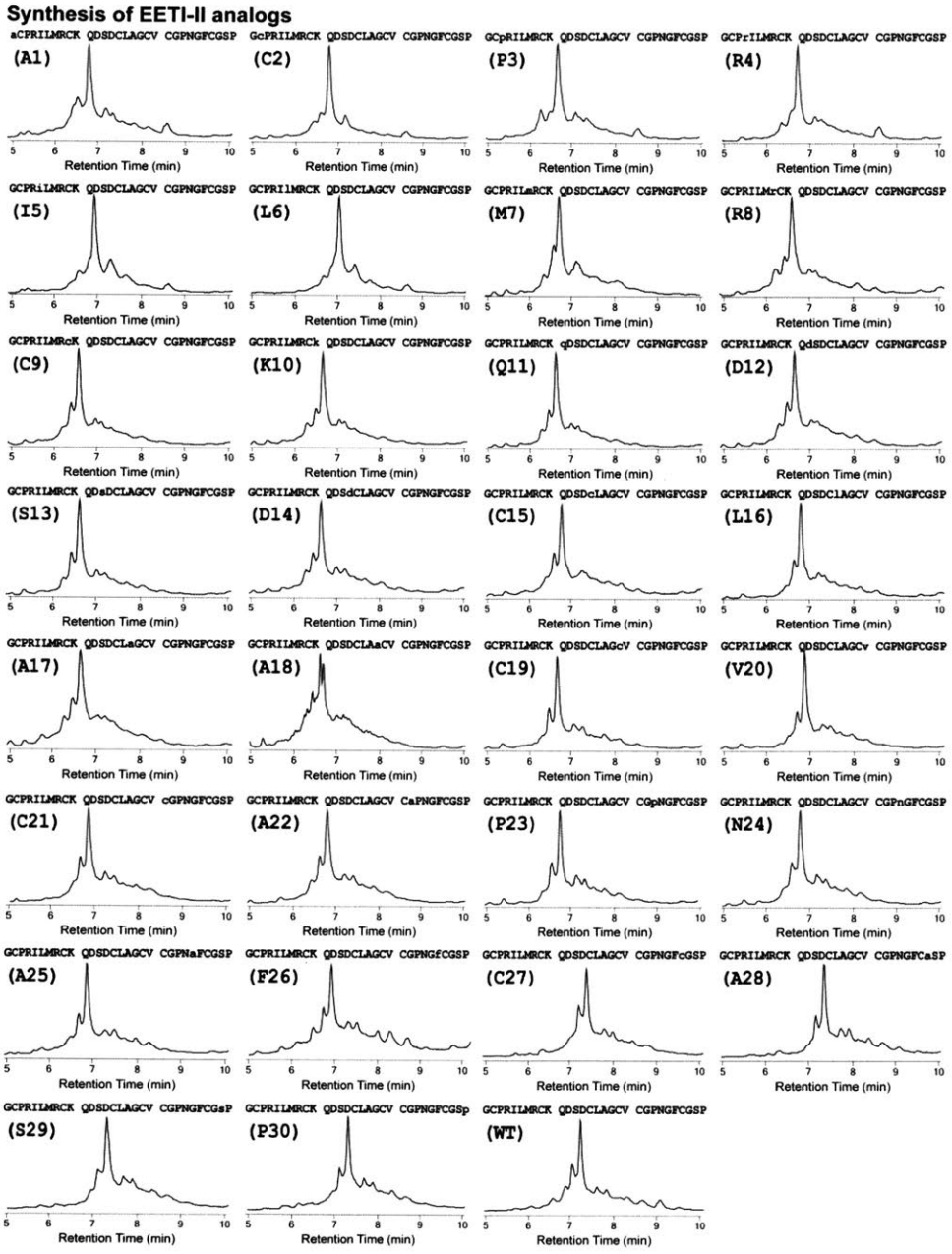


Figure 3.4.2 30 EETI-II analogs synthesized on the second generation AFPS. Capital letters represent L amino acids, and lower case letters represent D amino acids. In all cases, the major peak was the desired product, and sufficient material was isolated by preparative RP-HPLC for further studies (LC/MS average estimated purity 26%, range 13% to 36%). A18 shows a large, late eluting alanine deletion, and all syntheses have an early eluting Cys-Gly double deletion, and late eluting t-butyl adducts. The broad peak bases consist of a complex mixture of truncates (average 28%, range 20%-45%). For a detailed view of impurities, assignments of the ten most prevalent compounds identified by MFE for each synthesis are in Appendix 2. All syntheses were on 140-170mg of ChemMatrix Rink amide resin (0.5mmol/g) at 70°C under default conditions on the second generation AFPS, and total ion chromatograms are shown. See Appendix 2 for chromatographic conditions, mass spectral data, and large format chromatograms.

3.4.3.1 Truncations

To compare the prevalence of low molecular weight products in this design to our manual flow method, GHRH (1-29) was prepared under the default conditions described above. This 29 residue hormone is synthesized effectively under manual control conditions, and serves as a benchmark for synthesis of long peptides. Notably, the major peak was the desired product, but numerous low intensity truncates constituted a prominent, highly heterogeneous baseline in the automated synthesis (36-37% vs 20% for the manual control), consistent with observations from the synthesis of EETI-II diastereomers and CPPs.

Annotation of compounds identified by MFE confirmed many observed minor products were truncates of the desired peptide, often with a modification that led to a truncate 42 or 26 Daltons more massive than the unmodified peptide (see Appendix 2 for annotation of the 50 most prevalent compounds identified with MFE for each synthesis). The high baseline of truncates was unaffected by increasing coupling time, increasing deprotection time, or removing the static mixer, which indicated they were from a species formed in the reagent reservoirs (Fig 3.4.3). Further evidence for the degradation of reagents in the reservoirs was found during the NMR studies, below.

3.4.3.2 Stability of Fmoc-His(Trt)-OH and HBTU in DMF

1H NMR analysis of an HBTU/His(Trt) solution clearly showed the time dependent formation of tetramethyl urea, a byproduct of activation with HBTU that has a distinctive peak well resolved from all starting components (Figs A2.14.3-1 to A2.14.6). Interestingly, a solution of Fmoc-Ala-OH and HBTU in DMF analyzed as a negative control also showed formation of tetramethyl urea, but at levels 70 fold lower, suggesting that long term storage of amino acids in solution with activators was not viable (Figs A2.14.7 and A2.14.9).

Based on the NMR findings and the observed brown color, we hypothesized that the Trityl protected imidazole deprotonated the carboxylic acid, initiating formation of the OBt ester, followed by degradation to the carboxylate or an amide with trace water or dimethyl amine in the DMF (Scheme A2.1). To improve the stability of histidine in DMF with activator, N-im Boc protection was used to inductively deactivate the unprotected nitrogen, and NMR analysis confirmed a 7 fold slower degradation of HBTU (Figs A2.14.7 and A2.14.8).

The functional stability of solutions of protected histidine derivatives codissolved with HBTU in DMF was studied by synthesizing CPP#55 under several conditions (ZGSPWGLQHHPPRT; Z = 4-Pentynoic acid). CPP#55 was a cell penetrating peptide that showed prominent des-His and des-His-His

products following automated synthesis. It was selected for histidine stability model studies because it has two possible identical histidine deletion products, enhancing observed effects.

All syntheses were performed using the manual peptide synthesizer and the previously reported two minute synthesis protocol (30 second amide bond formation at 6ml/min with ~0.4M amino acid activated with HBTU, 20 second deprotection with 20% (v/v) piperidine in DMF at20ml/min, and two 30 second DMF washes at 20ml/min) [68]. 18 experiments were performed: synthesis was conducted with either Fmoc-His(Trt)-OH, Fmoc-His(Trt)-OH + 1eq HOBt, Fmoc-His(Boc)-OH, or Fmoc-His(Boc)-OH-CHA that was dissolved in DMF with equimolar HBTU immediately before synthesis, one day before synthesis, or 4 days before synthesis. For syntheses with Fmoc-His(Trt)-OH and Fmoc-His(Boc)-OH-CHA, one set of syntheses was performed with DIEA added as usual immediately before use (190uL, 1.1eq) and one was performed without DIEA. For syntheses with Fmoc-His(Boc)-OH and Fmoc-His(Trt)-OH + 1eq HOBt, only the trials with the usual 1.1eq of DIEA were performed.

The results of these trials (Fig 3.4.4) support the hypothesis that Trityl protection of the imidazole ring allows solutions of protected histidine and HBTU in DMF to degrade through an OBt ester intermediate, and that the degradation is initiated by deprotonation of the carboxylate by the imidazole ring.

Formation of the desired product when using Fmoc-His(Trt)-OH without DIEA shows the Trityl protected imidazole acts as a base to deprotonate the carboxylic acid. Further, synthetic outcomes with Fmoc-His(Boc)-OH-CHA were insensitive to addition of DIEA, and were always worse than the Fmoc-His(Trt)-OH controls, showing degradation is initiated by deprotonation of the carboxylic acid. When HOBt was added to the solution of Fmoc-His(Trt)-OH, synthetic outcomes were markedly worse, consistent with an OBt ester intermediate.

The success of Fmoc-His(Boc)-OH at all three time points confirms that Boc protection of the imidazole is sufficiently deactivating to largely prevent deprotonation of the carboxylic acid and initiation of degradation, but the minor deletion product in the 4 day trial and tetramethyl guanidinium observed by NMR indicate that this protection is not perfect.

In addition to improving the stability of HBTU in solution with histidine, boc protection and inductive deactivation of the imidazole ring prevented epimerization of histidine at high temperature, which is a severe side reaction with Fmoc-His(Trt)-OH[62,69,70]. To confirm that N-im Boc protection is sufficiently deactivating to prevent direct deprotonation of C-α by the basic imidazole moiety, a series of tri peptides, FHL, with a well resolved diastereomer, FhL, were prepared. Either Fmoc-His(Boc)-OH or Fmoc-His(Trt)-OH was activated with either 1.0eq or 2.9eq DIEA and incorporated at either RT or 90°C. In addition to these 8 trials, the authentic diastereomer was prepared using Fmoc-D-His(Trt)-OH activated with 1.0eq DIEA and incorporated at 60°C. Results are in Figure 3.4.5, and clearly show that Fmoc-

His(Trt)-OH substantially racemizes under all conditions, but Fmoc-His(Boc)-OH minimally racemizes, even at 90°C. Racemization appears independent of DIEA concentration, confirming that the imidazole moiety directly deprotonates C-α.

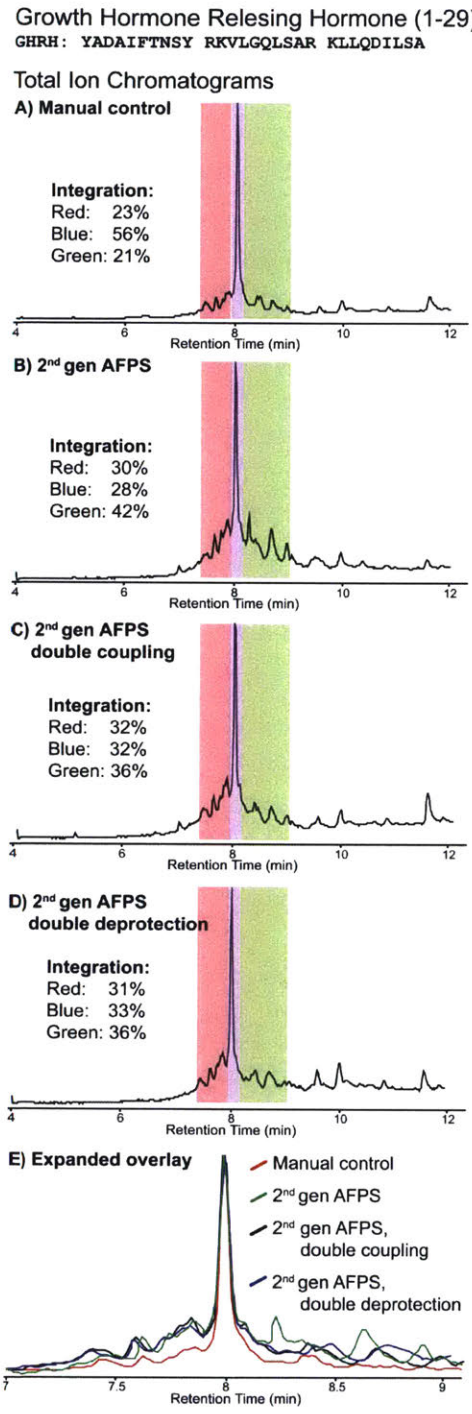


Figure 3.4.3 GHRH synthesized with the second generation AFPS under several conditions: **A)** manual fast flow synthesizer, **B)** second generation AFPS, **C)** second generation AFPS with doubled coupling time and volume, and **D)** second generation AFPS with doubled deprotection time and volume. **E)** Overlay of the four chromatograms shown in A-D, showing all three automated syntheses to be worse than the manual synthesis. Total ion chromatograms are shown and integrated. See Appendix 2 for chromatographic conditions, mass spectral data, and assignment of the 50 most prevalent compounds identified with MFE in each synthesis

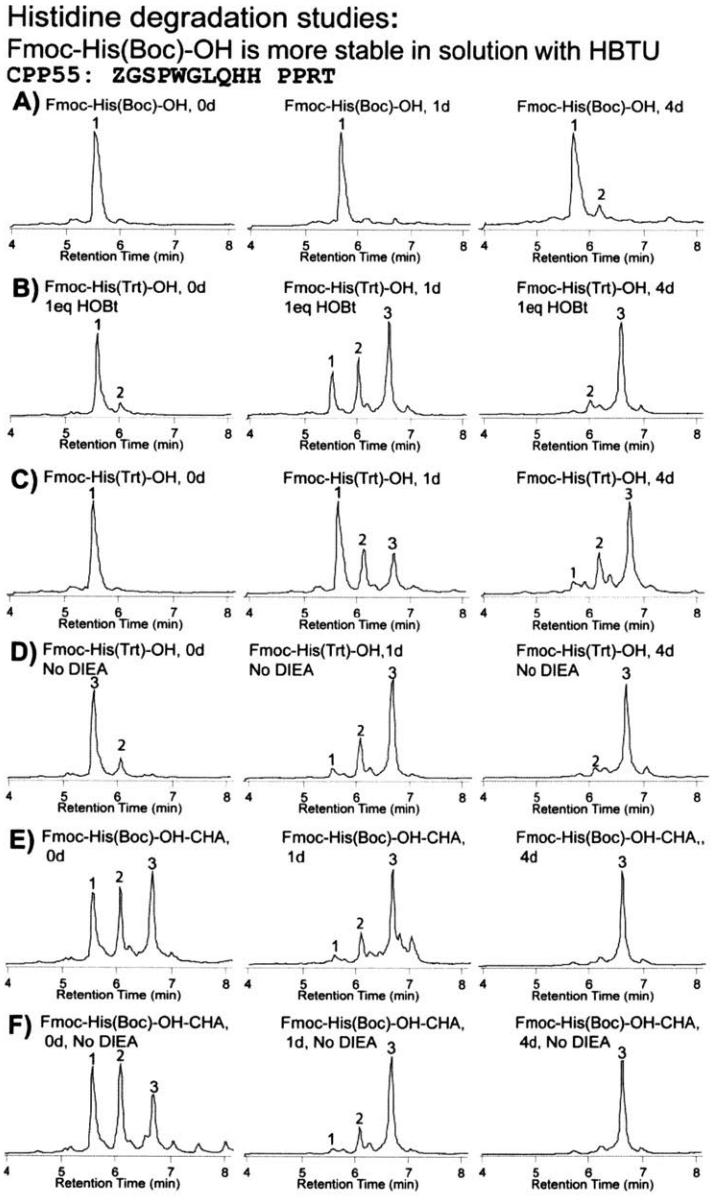


Figure 3.4.4 Histidine degradation studies on ZGSPWGLQHHPPRT. The left column shows syntheses performed with histidine monomers freshly codissolved with HBTU, the middle column shows syntheses with reagents that had been dissolved for one day, and the right column shows syntheses with reagents that had been dissolved for four days. Only the Fmoc-His(Boc)-OH condition gave appreciable desired product in the 4 day trial. **A)** Fmoc (Boc)-OH performs well even after 4 days codissolved with HBTU. **B)** Fmoc-His(Trt)-OH degrades faster with HOBt than in the comparable trial without it in panel C, supporting an OBt ester intermediate. **C)** Fmoc-His(Trt)-OH performs poorly after 1 day codissolved with HBTU, and fails to incorporate histidine after 4 days. **D)** Fmoc-His(Trt)-OH performs comparably without DIEA, showing that the imidazole ring is sufficiently basic to deprotonate the carboxylate. **E) and F)** A cyclohexylammonium salt of Fmoc-His(Boc)-OH shows rapid degradation with (E) or without (F) DIEA, indicating that deprotonation of the carboxylate alone is sufficient to degrade the solution, and DIEA added immediately before use has little influence. 1 = desired product; 2 = des-His; 3 = des-His-His. See Appendix 2 for chromatographic conditions.

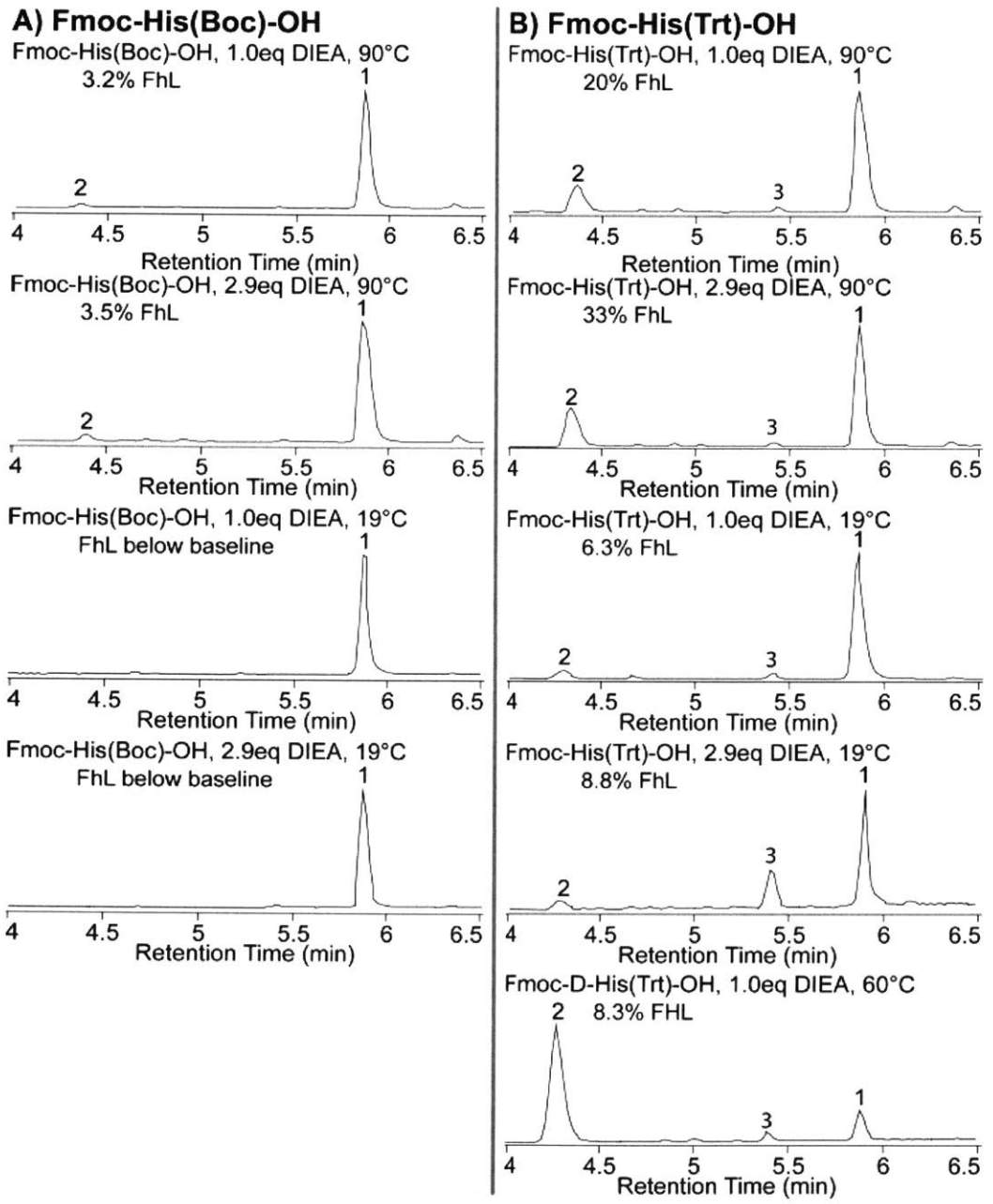


Figure 3.4.5 Histidine racemization studies on FHL. **A)** The left column shows syntheses performed with Fmoc-His(Boc)-OH. **B)** The right column shows syntheses performed with Fmoc-His(Trt)-OH, including synthesis of authentic FhL with Fmoc-D-His(Trt)-OH. The Fmoc-His(Boc)-OH conditions all gave minimal racemization, even at 90°C, and the Fmoc-His(Trt)-OH conditions all resulted in some diastereomer formation. The proportion of diastereomer was determined by integration of absorbance at 214nm. 2 = FhL; 3 = des-His; 1 = FHL. Total ion chromatograms are shown. See Appendix 2 for chromatographic conditions.

3.5 Third generation automated fast flow peptide synthesis

3.5.1 Key Findings

The regent degradation observed with the second generation AFPS was overcome by separation of amino acids and activators. Mechanical systems were designed and tested to properly blend amino acids and activators during rapid transitions so that excess activator was never present. Furthermore, a scalable, electrically heated reaction vessel was implemented, and used to show that synthetic outcomes improved monotonically up to 120°C, its maximum operational temperature. Finally, a new user interface was created that substantially reduced synthetic failures due to human factors.

After extensive use of the third generation AFPS, it was observed that Fmoc protected amino acids are not stable in DMF, losing the Fmoc protecting group and precipitating as the N-α unprotected monomer. Using NMP or DMI as solvents substantially improved stability of the N-α Fmoc group.

3.5.2 Design of third generation AFPS

The third generation automated fast flow peptide synthesizer shown in Figure 3.5.1 incorporated several major design changes that improved the synthesis quality and user interface substantially, and facilitated scale up. First and foremost, amino acids and activators were stored separately, rather than codissolved in DMF, to improve their stability, dramatically improving synthetic outcomes. Further, the water bath, preheat loop, and resin vessel from the manual synthesizer were replaced with an electrically heated resin vessel and several independently heated tubular activation reactors, which enabled activation studies at various temperatures and residence times, as well as synthesis up to 120°C. Finally, the user interface was improved with enhanced storage vessels that were easier to refill and clean, and software with a graphical user interface that automatically recorded data and synthetic parameters.

3.5.2.1 Separation of amino acid and activator

To eliminate the truncates that were characteristic of peptides produced with the second generation synthesizer, the third generation AFPS stored amino acids without activator co-dissolved. The amino acids were connected to the amino acid pump as before, and a dedicated activator pump was used to infuse an activator dissolved in DMF at the same point DIEA was infused. The mixed amino acid, activator, and DIEA solution was passed through a static mixer and to the resin vessel (Fig 3.5.1A). In addition to improving the stability of reagents, this design enables trivially selecting one of several activators with a valve attached to the activator pump. In principle, a different activator could be used for each coupling in a synthesis, although this capability was not utilized in the present work.

In the third generation AFPS it was essential that the stoichiometry of the amino acid and activator was always matched; if at any point there was an excess of activating agent, the peptide chain was terminated with a tetramethyl guanidinium cap. To achieve matched stoichiometry with the cyclic, single piston Varian 210 pumps we retained for their low non-displaced volume, the pumps were started and run absolutely synchronously. Initiating operation simultaneously and using the same pump heads, flow rate, withdraw time and compressibility settings caused the pump heads to cycle in exactly the same way, and the stoichiometry of reagents was therefore always correct.

To verify the stoichiometry was always matched, dyed solvents in amino acid and activator reservoirs were used, and the flow during a coupling cycle was observed with a high speed camera (Fig 3.5.2).

3.5.2.2 New reactor design

Next, the resin vessel, preheat loop, and water bath retained from the manual synthesizer were replaced with an electrically heated resin vessel and several independently heated activation reactors. The new resin vessel design offered a number of advantages. Most importantly, synthesis could be performed at temperatures over 100°C. Further, the effect of residence time and temperature on a particular activation step could be studied independently of the time and temperature for amide bond formation, Fmoc removal, and other activation steps.

Additionally, resin was handled in disposable fritted polypropylene syringes, which facilitated downstream operations and reduced handling losses. It also allowed us to provision the reactor platform with a robotic arm to reliably load syringes of resin without air pockets, and automatically remove finished resin from the heated zone.

Finally, the reactor platform was designed to accommodate three scales with minimal changes to the hardware and software, enabling facile scale up of synthesis from the standard 0.08mmol scale to an 0.4mmol scale (at 0.5mmol/g resin loading; higher loadings enabled larger scales).

In the new reactor design, shown in Figures 3.5.1H-3.5.1K, incoming fluid was passed through one of several independently thermostated tubular activation reactors selected with a Valco column selector valve, then into a heated resin vessel. Much like the preheat loop used in the first and second generation AFPS, the tubular activation reactors consisted of a coil of heated stainless steel tubing (0.030" ID; 1/16" OD). After exiting the heated zone, however, solvents and reagents were passed back through the Valco column selector valve, through a pressure sensor, and through about 10 inches on uninsulated PEEK tubing before entering the heated resin vessel inlet. This spatial separation allowed the fluid to cool, and effectively decoupled the activation temperature from the amide bond formation and Fmoc removal temperature.

The resin vessel consisted of a heated aluminum cylinder that accommodated a fritted syringe containing the solid phase. The syringe was sealed at the inlet with an O-ring seal in the barrel and at the outlet with a female luer fitting. A robotic arm loaded syringes of resin without air pockets by positioning the syringe on a fork under the resin vessel inlet, waiting for the pumps to fill the syringe with DMF, then lowering the resin vessel inlet into the syringe, displacing excess DMF into a drip tray. The fork withdrew, and the resin vessel inlet and syringe were lowered, sealing the luer slip syringe into the mating luer fitting in the resin vessel outlet.

To increase the scale, the resin vessel inlet, resin vessel body, and fork were changed to accommodate 12mL and 24mL fritted syringes, as shown in Figures 3.5.1J and 3.5.1K. Heater positions, thermocouple positions, and mechanical attachment points were the same for all scales, and the alignment was retained when changing scales. The scale selection in the software controlled the vertical movement of the inlet to properly close each reactor.

3.5.2.3 User interface - software

The software for the first and second generation AFPS was based on an Arduino microcontroller[34,76]. Although this software was sufficiently flexible to enable any and all experiments with the hardware, the interface was command line only, and no data was stored electronically. To improve the user experience, RM Beaumont Corporation (RBC) was contracted to develop improved control hardware and software.

In writing new software and overhauling the control hardware, a LabVIEW software package with a graphical user interface was created that allowed the user to select the synthetic scale and easily change any synthetic parameter for any step. UV and pressure data were displayed in real time to allow users to detect problems with the synthesis, and alarms were set for low reagent levels and low heater temperatures. For every synthesis, a data folder was created that recorded the instrument configuration parameters, synthetic conditions, and data collected during the synthesis. Digital recording of UV data enabled informative automated post processing [55] (See Appendix 2), and in principle, real time analysis could be integrated into the LabVIEW instrument controller to modify synthetic parameters during synthesis. Details of the LabVIEW instrument can be found in Appendix 2, and several key user facing features are shown in Figure 3.5.3.

3.5.2.4 User interface - reagent reservoirs

The media bottles used as reagent reservoirs in the second generation AFPS were positioned in racks that were far back in the fume hood and difficult to reach and see, especially for people less than 5'6" tall. Furthermore, the bottles had a single GL45 threaded opening that was used for both filling the

bottle and withdrawing reagents. This required the seal between the bottle top and the insert be broken and remade every time reagents were added, which was inconvenient, required significant force, and often caused nitrogen leaks.

To overcome this user interface issue, the rack and reservoir system was significantly redesigned for the third generation AFPS (Fig 3.5.1C-3.5.1E). Reagents were stored in three neck flasks with one neck capped permanently, one neck used for nitrogen supply and reagent withdraw, and the third neck used exclusively as a fill port. Additionally, the rack that held the bottles was moved forward in the hood, and redesigned so it was easy to see the fill level of all bottles. The rack confined the bottles vertically so that it was easier to remove the side arm fill cap, and rotameters were added to detect nitrogen leaks that evaporated DMF over time. In total, 45 amino acid reservoirs were employed (9 on each of five 10-position valves), which held the common L and D amino acids, as well as five specialty locations. 9 reservoirs for activators and deprotection reagents were used, all on a single activator valve. Labels were engraved and mounted above each reservoir so it was always clear which compound was in each vessel; variable locations had the character used to call that location when setting up a synthesis.

Together, these changes improved the user interface with the reagent reservoirs dramatically. Reagents were never run dry, large nitrogen leaks were detected and corrected immediately, and filling was much more convenient.

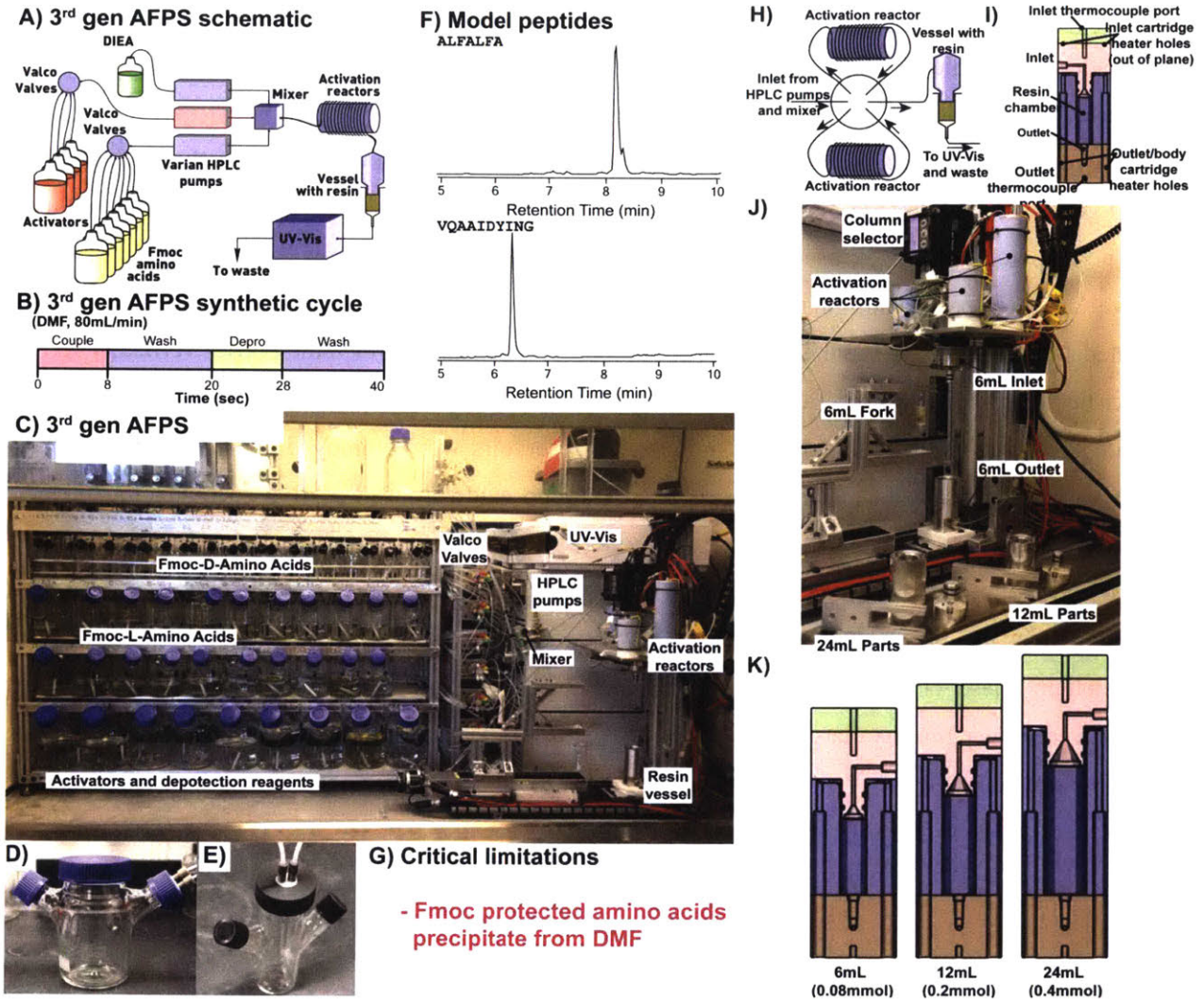


Figure 3.5.1 Third generations AFPS A) Schematic of the third generation AFPS. Fmoc protected amino acids are pumped through an HPLC pump to a mixer where a second HPLC pump infuses an activating agent, and a third HPLC pump infuses DIEA. The mixed fluid then travels through a tubular activation reactor to form the active ester. Upon exiting the activation reactor, the newly formed active ester moves through an independently thermostated vessel containing the solid phase, and effects amide bond formation. The effluent is passed through a UV-Vis spectrometer and to waste. The third generation AFPS is distinguished from the second generation by separation of the activator and amino acids, as well as a different reagent bottle arrangement and resin vessel design. Separation of amino acids and activators is supported by a third pump, highlighted in red. **B)** Third generation AFPS synthetic cycle, run at 80mL/min. **C)** Photograph of the third generation AFPS with select components labeled. Reagent reservoirs are filled through the necks facing the camera; the nitrogen supply/reagent withdraw is in the back. **D)** 250mL (500mL total capacity) three neck spinner flask used for proteogenic amino acids, showing a gas/reagent adaptor installed in a GL32 threaded side arm. The 500mL (1L capacity) spinner flasks for activators and deprotection reagents are similar, but have GL45 threaded arms. **E)** 25mL (50mL capacity) three neck flask used for non-proteogenic amino acids, showing a gas/reagent adaptor installed in the 38-400 threaded top port. **F)** Model peptides ALFALFA and ACP(65-74) synthesized with the third generation AFPS. **G)** Amino acids on the third generation AFPS were not stable, precipitating as N-α unprotected material after days to weeks. **H)** Detailed schematic of the electrically heated reactor module. An incoming mixture of DIEA, amino acid, and activator is passed through a Valco column selector valve

into one of several individually thermostated activation reactors of various lengths (two are shown, up to six are supported). The mixture, which now contains the active ester of the amino acid, is then passed back through the column changer and into the independently thermostated resin vessel. **I)** Sectioned rendering of the 6mL resin vessel showing the arrangement of the inlet, resin chamber, and outlet. Mechanical connections are not shown **J)** Photograph of the reactor module with the 6mL resin vessel and three activation reactors installed. A fork positions syringes for a bubble free fill and removes the resin from the heated zone when synthesis is finished. In the foreground, hardware for larger scales can be seen. **K)** Sectioned rendering of the three resin vessel scales tested here. The small gap between the inlet (light red) and rector body (blue) accommodates a disposable syringe (not shown). Note that the inlet mounting plate (light green) and outlet (orange) are universal, allowing facile conversion from one scale to another.

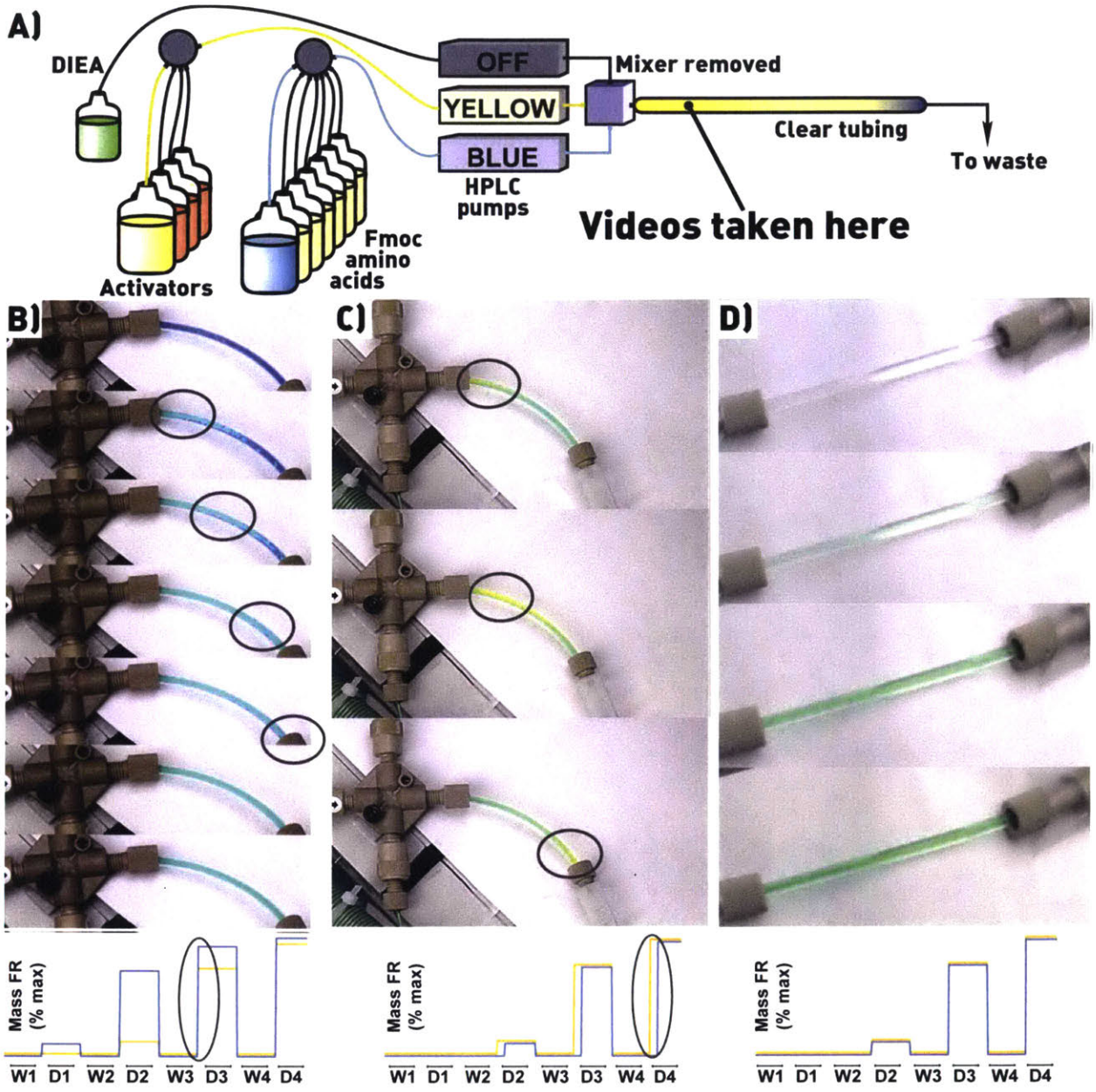


Figure 3.5.2 Pump synchronicity studies using visible dye, digitally color enhanced. **A)** The amino acid pump was charged with blue dye, the activator pump was charged with yellow dye, the DIEA pump was deactivated, and the exit of the merging region was observed through a clear piece of tubing during a coupling cycle. The mixer was removed. The mass flow rate of each dye vs time is plotted below the photographs in panels B-D. During pump withdraw (marked with a W), the mass flow rate is zero, and during delivery (marked with a D), the mass flow rate depends on the concentration in the pump head, which does not change during delivery. Concentration steps are caused by laminar flow in the reagent manifold supplying the pump heads. **B)** A leading edge of blue followed by a steady stream of green indicated that the blue reagent arrived too early. During delivery 1 and withdraw 2 (D1 and W2), the tube appears faint blue or colorless (not shown). During D2 and W3, the tube appears intensely blue (top photograph), and during D3 it transitions to green (bottom photograph). The junction between green and blue is circled in all panels. **C)** Regular pulses of yellow indicate that the pumps are running asynchronously or the blue pump head is cavitating. In all panels, a pulse of yellow is circled as it moves

through a stream of green. The detrimental effects of this phenomenon are shown in Figure 3.6.2. **D)** Proper formation of a mixed stream of blue and yellow showing a gradual change from colorless to green with no blue or yellow regions. Note the step change in green intensity shown is actually observed and expected because mixing in the pump heads discretizes the concentration changes of each reagent. All photographs are derived from individual frames of high speed video and have been contrast and brightness enhanced and cropped. Additionally, panel A has had blue enhanced and panel B has had yellow enhanced to more clearly show the transitions described. The original contrast and brightness enhanced image with no hue specific manipulation is in Appendix 2 (Figure A2.4).

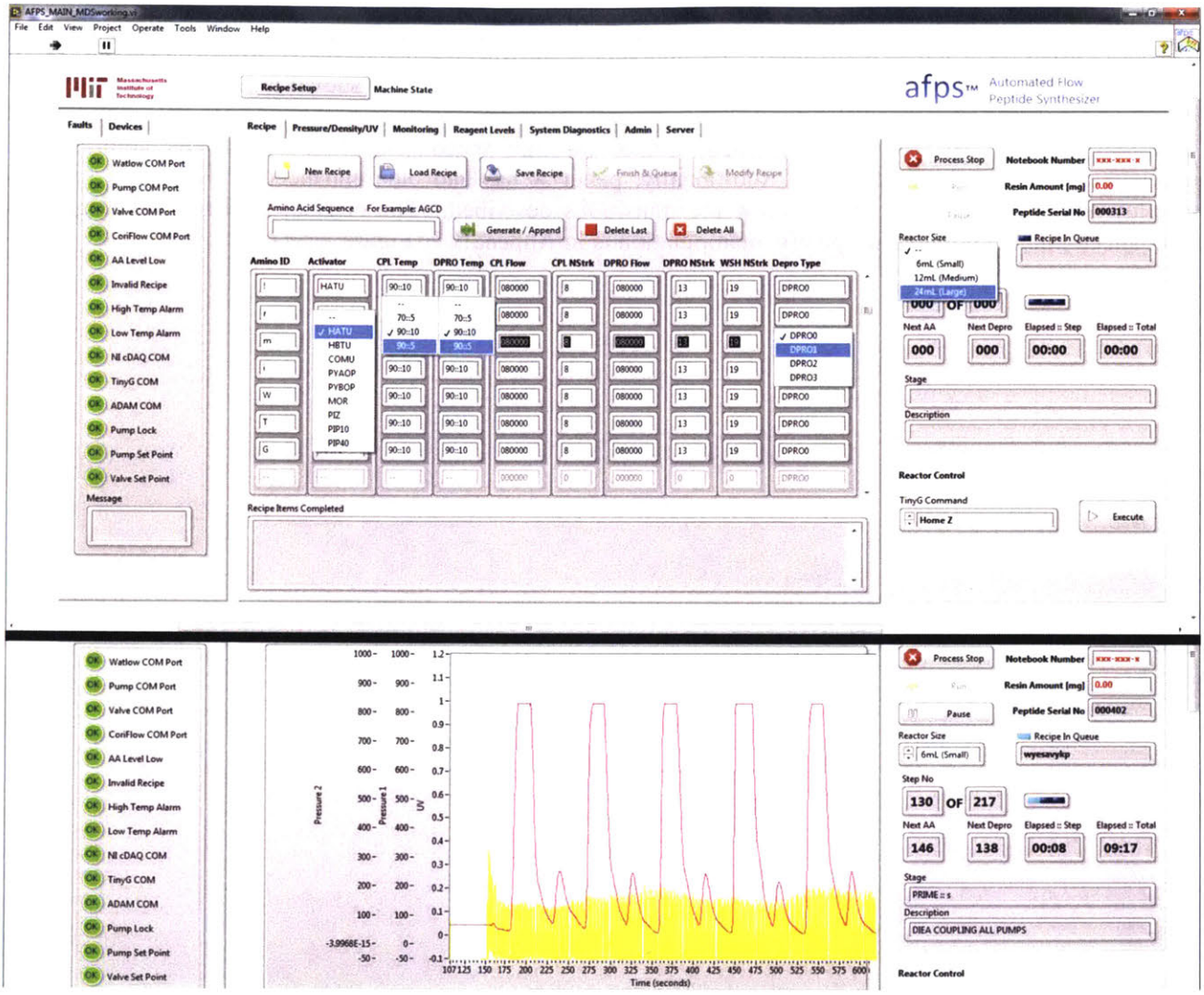


Figure 3.5.3 LabVIEW software Top: Overlaid screen shots of the LabVIEW VI that controlled the instrument showing options for selecting the activator, coupling and deprotection temperatures, coupling and deprotection flow rates, deprotection protocol, and scale. Bottom: Pressure and UV data are displayed in real time. This partial screen shot shows the first five couplings and deprotections of a 9 residue peptide synthesized on the fourth generation AFPS (80 second cycle, see below). The sequence can be seen on the right, along with the current step.

3.5.3 Characterization of third generation AFPS

3.5.3.1 Model peptides

To validate the third generation automated peptide synthesizer, model peptides ALFALFA and ACP(65-74) were synthesized in the 6mL resin vessel at 90°C with a 10ft activation reactor, also at 90°C. Each synthetic cycle consisted of an 8 second coupling, an 8 second deprotection, and two 12 second washes. For the wash steps, both the amino acid and activator pump were used to deliver DMF at 40ml/min each. During the coupling step, the amino acid pump delivered 0.4M amino acid in DMF at 40ml/min, the activator pump delivered 0.38M HATU in DMF at 40ml/min, and the DIEA pump delivered 4mL/min of neat DIEA. During the deprotection step, the amino acid pump delivered 40ml/min of DMF, and the activator pump delivered 40ml/min of 40% (v/v) piperidine in DMF. The pumps were run synchronously at all times; activator and amino acid pumps had 50ml/min pump heads, and the DIEA pump had a 5ml/min pump head. For each synthesis, 150-180mg of ChemMatrix Rink amide resin was used (0.5mmol/g) in the 6mL resin vessel, and the peptides were found to be high quality (Fig 3.5.1F). These default conditions were used for all subsequent syntheses with the third generation AFPS unless noted.

Additionally, the 12mL and 24mL resin vessels were used to synthesize ALFALFA and ACP(65-74) at larger scale. For each peptide, two syntheses were performed in the 12mL resin vessel on about 350mg of resin, and 4 syntheses were performed in the 24mL resin vessel on about 750mg of resin. In one synthesis in the 12mL resin vessel, the default conditions described above were used; in the other, all volumes (and times) were doubled with no other changes. In the 24mL resin vessel, one synthesis was performed with default conditions, one with only the wash volume doubled, one will all volumes doubled, and one with all volumes quadrupled. Controls were repeated under the default conditions above on 150mg of resin in the 6mL resin vessel.

Syntheses were of high quality, and yields of peptidyl resin and crude peptide are in Table 3.5.1. Most strikingly, no peptidic products were obtained from synthesis in the 24mL resin vessel with default volumes, but doubling the volume of wash solvent resulted in effective synthesis with the default reagent volumes: only 5 equivalents of amino acid were used, in contrast to 25 equivalents in the control synthesis. This result demonstrates that fast flow synthesis can be effective with few equivalents of reagents, and suggests there is substantial room to optimize the equivalency of smaller scale synthesis by changing the volume or concentration of the coupling reagents.

3.5.3.2 Temperature studies

With an electrically heated resin vessel and activation reactor, peptide synthesis was studied above the 90°C limit of the water bath. The 6mL resin vessel and default synthetic conditions described above were used without modification, except the temperature of the 10ft activation reactor and resin vessel were both set at 60°C, 70°C, 80°C, 90°C, 100°C, 110°C, or 120°C. Temperatures above 120°C were not studied because the polypropylene syringes used to contain the resin inside the resin vessel melt at about 125°C.

The JR-10mer was selected for temperature studies because it shows persistent deletions at lower temperature which ionize much more efficiently than the desired compound during LC/MS analysis, and are thus easy to detect at low levels. As can be seen in Figure 3.5.4, the synthesis gradually improved with temperature, with the 110°C and 120°C experiments showing minimal deletion products, even in the total ion chromatogram where they are over represented.

With the completion of these model studies, a synthetic coiled coil fragment was synthesized at 90°C, 100°C, 110°C, and 120°C after the UV peak heights following deprotection declined dramatically during synthesis at 90°C. Figure 3.5.5 shows that although the synthesis did not fail at 90°C, it gradually improved with increasing temperature. Declining UV peak heights during deprotection indicated that Fmoc removal was proceeding more slowly, but did not necessarily indicate that synthesis had failed. The default coupling and deprotection times (volumes and flow rates) are much greater than typically necessary so that syntheses with unexpectedly slow amide bond formation or Fmoc removal will yield useable peptidyl resin, as was the case with the 90°C synthesis of the coiled coil fragment.

3.5.3.3 General lab use

With initial validation complete, we undertook the synthesis of about sixty mixed chirality proteins in support of other projects in the lab, and to detect remaining problems with the third generation AFPS.

All of the peptides were synthesized under the default synthetic conditions above at 90°C, using 140-170mg of ChemMatrix Rink amide resin (0.5mmol/g) in the 6mL resin vessel. Chromatograms from select syntheses are shown in Figure 3.5.6 and Section A2.15.3, and the remainder are shown in Section A4.17.

The synthesis went smoothly and was completed over the course of about a week. Notably, the severe truncates that plagued synthesis with the second generation instrument were much less prevalent, and histidine was effectively incorporated throughout the project, signaling that the major chemical problems observed with the second generation synthesizer had been overcome.

Shortly after the syntheses above were completed, however, a precipitate formed in several amino acid reservoirs. Although synthesis wasn't impacted in these experiments, insoluble particulates were

concerning. Mechanical failures from blocked selector valves, stuck check valves, and destroyed piston seals were expected, as were chemical failures from a reduced concentration of amino acid and an increased concentration of the poorly soluble degradation product.

3.5.3.4 Reagent stability study

Ultimately, our goal was to build an instrument that was always filled with reagents and ready to use, much like an HPLC is left with solvents, ready to run. Following the above syntheses, it became clear that the only major barrier to this objective was the stability of the amino acid stocks.

Therefore, a reagent stability study with all the Fmoc protected proteogenic amino acids was conducted. We hypothesized that precipitates were N-α unprotected amino acids formed when DMF disproportionates, liberating dimethyl amine and removing Fmoc. In contrast, NMP was not expected to substantially disproportionate, so 0.4M solutions of each amino acid were prepared in DMF and NMP, and left for one month. The solvents used were dry and amine free. Solutions were prepared on the bench, and stored in airtight media bottles.

After one month, 13 out of 20 amino acids stored in DMF had formed precipitates. In contrast, only proline and glycine precipitated from NMP. 1H and 13C NMR confirmed each precipitate to be the N-α unprotected amino acids with side chain protecting groups intact, and Figure 3.5.7 illustrates the striking difference in proline stability under these conditions. To determine whether Fmoc protected amino acids were more stable in NMP, or N-α unprotected degradation products were simply more soluble, the supernatants of all the solutions were assayed for amines with the quantitative ninhydrin test[77]. NMR was not used due to high background from DMF, fully protected amino acids, and dibenzofulvene fromed during Fmoc removal. In all cases, the concentration of primary amine in NMP solution was less than or equal to the concentration in DMF solution, showing that Fmoc protected amino acids are much more stable in NMP than DMF (Table 5.3.2).

The study was allowed to continue for an additional 10 weeks, and all of the amino acids except Fmoc-Arg(Pbf)-OH and Fmoc-Asn(Trt)-OH precipitated from DMF to form gelatinous bricks. The bottles were found to be under positive pressure when opened, consistent with liberation of one equivalent of CO_2 during Fmoc removal. In contrast, only four amino acids (His, Pro, Gly, Met) had significantly precipitated from NMP after 14 weeks. An additional five (Val, Thr, Ile, Leu, Lys, Ala) showed very slight precipitation.

In an effort to extend the shelf life of proline and glycine, which precipitated from NMP after only four weeks, their stability was assessed in dimethylimidazolidone (DMI), which was hypothesized to be even less likely than NMP to form amines and remove Fmoc. A slight precipitate was first noticed in the

glycine solution after 10 weeks of standing, and the ninhydrin test showed only 9 mM primary amine in the supernatant. Proline did not precipitate in this time, and cannot be detected with the ninhydrin test.

Table 3.5.1 Yields of ACP(65-74) and ALFALFA across three reactor scales under various conditions.

Peptide	Reactor	Condition	Resin Loaded	Peptidyl Resin	Crude Peptide
ALFALFA	6mL	Default	141mg	**166mg (86%)**	**12mg (23%)**
	24mL	Default	352mg	**406mg (84%)**	**28mg (21%)**
		2x all volumes	364mg	**469mg (94%)**	**35mg (26%)**
	24mL	Default	759mg	**753mg (72%)**	**14mg (5%)**
		2x wash	756mg	**825mg (79%)**	**62mg (22%)**
		2x all volumes	745mg	**833mg (81%)**	**62mg (22%)**
		4x all volumes	731mg	**937mg (93%)**	**100mg (36%)**
ACP(65-74) VQAAIDYING	6mL	Default	174mg	**251mg (79%)**	**63mg (75%)**
	12mL	Default	364mg	**498mg (79%)**	**94mg (51%)**
		2x all volumes	340mg	**529mg (85%)**	**78mg (43%)**
	24mL	Default	N/A	N/A	N/A
		2x wash	757mg	**992mg (72%)**	**203mg (51%)**
		2x all volumes	40mg	**1091mg (81%)**	**186mg (48%)**
		4x all volumes	768mg	**1254mg (89%)**	**146mg (36%)**

Table 5.3.2 Stability of Fmoc protected amino acids in DMF and NMP after 4 weeks

Amino Acid	Precipitation[a]		Amine conc. (mM) [b]	
	DMF	NMP	DMF	NMP
Fmoc-Ala-OH	**YES**	NO	**9 ± 6**	**5.1 ± 2.0**
Fmoc-Cys(Trt)-OH	NO	NO	**8.8 ± 1.7**	0.6 ± 0.7
Fmoc-Asp(tBu)-OH	NO	NO	1.7 ± 0.8	0.4 ± 0.6
Fmoc-Glu(tBu)-OH	**YES**	NO	**17.9 ± 1.8**	3.1 ± 0.7
Fmoc-Phe-OH	**YES**	NO	**5.7 ± 1.1**	1.5 ± 0.6
Fmoc-Gly-OH	**YES**	**YES**	**5.2 ± 2.0**	3.9 ± 0.7
Fmoc-His(Boc)-OH	NO	NO	**5.3 ± 1.5**	1.9 ± 0.6
Fmoc-Ile-OH	**YES**	NO	**7.5 ± 1.0**	3.1 ± 0.8
Fmoc-Lys(Boc)-OH	**YES**	NO	**6.2 ± 1.5**	2.8 ± 0.6
Fmoc-Leu-OH	**YES**	NO	4.8 ± 2.1	2.6 ± 0.7
Fmoc-Met-OH	**YES**	NO	**5.3 ± 0.8**	3.2 ± 0.7
Fmoc-Asn(Trt)-OH	NO	NO	1.9 ± 0.8	0.7 ± 0.6
Fmoc-Pro-OH	**YES**	**YES**	n/a[c]	n/a[c]
Fmoc-Gln(Trt)-OH	NO	NO	**30 ± 8**	**5.8 ± 0.8**
Fmoc-Arg(Pbf)-OH	NO	NO	**5.3 ± 2.0**	1.5 ± 0.6
Fmoc-Ser(tBu)-OH	**YES**	NO	**12.7 ± 1.7**	3.2 ± 0.7
Fmoc-Thr(tBu)-OH	**YES**	NO	**13.1 ± 2.2**	**10.8 ± 1.0**
Fmoc-Val-OH	**YES**	NO	**6.3 ± 2.3**	3.5 ± 0.7
Fmoc-Trp(Boc)-OH	NO	NO	4.2 ± 1.0	1.5 ± 0.6
Fmoc-Tyr(tBu)-OH	**YES**	NO	**6.8 ± 1.6**	0.8 ± 0.6

[a] Conditions that caused precipitation are highlighted in red [b] Amine concentration in the supernatant, determined with the ninhydrin test. Standard errors are given, amine concentrations over 5mM are bolded, and amine concentrations over 10mM are bold and red. [c] Proline contains a secondary amine, and is unreactive in the ninhydrin assay.

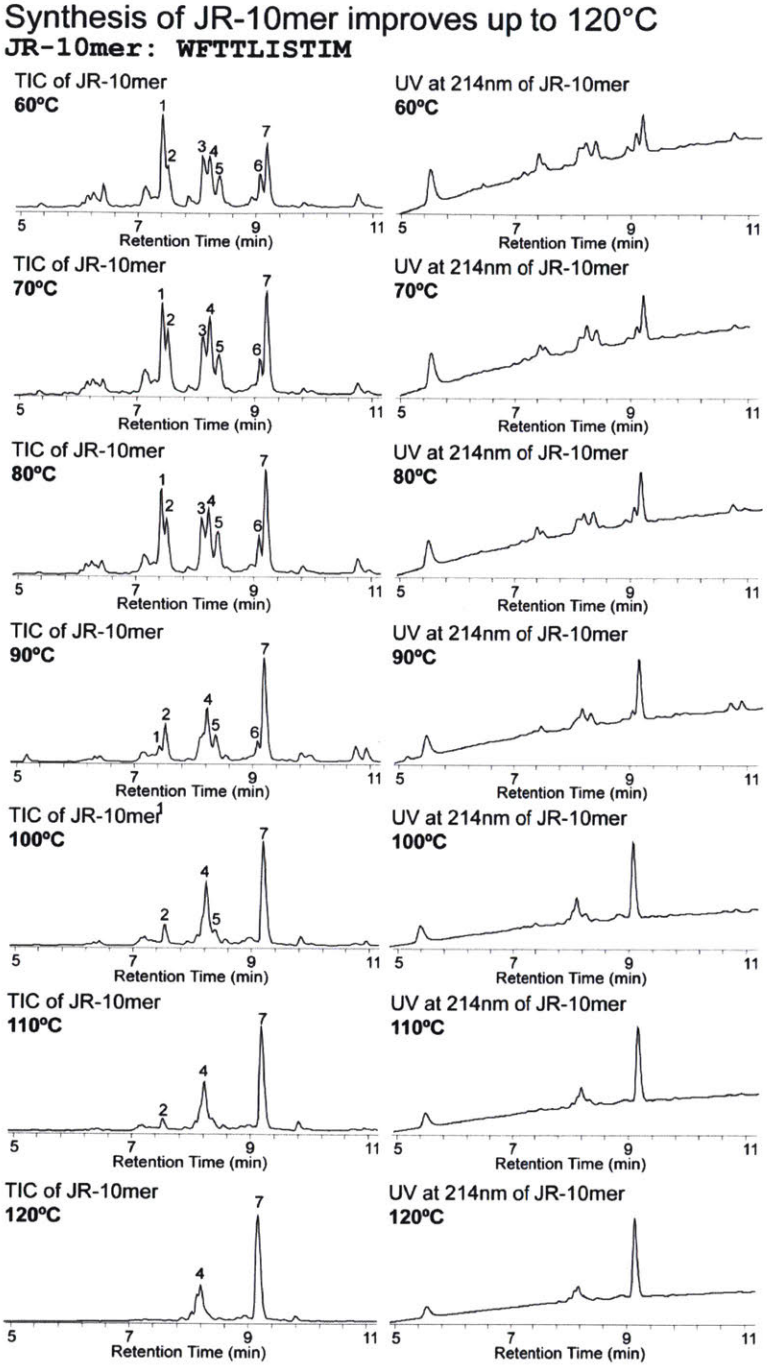


Figure 3.5.4 Monotonic improvement in JR-10mer synthesis from 60°C to 120°C, the maximum reactor service temperature. Each row shows chromatograms from a synthesis performed 10 degrees hotter than the previous row, with the 60°C synthesis on top and the 120°C synthesis on the bottom. The left column shows total ion chromatograms, which more clearly illustrate differences in syntheses, and the right column shows absorbance at 214nm which more accurately convey the true synthetic quality. 1=des-Trp, Phe, Thr, 2=des-Trp-Phe,3=des-Trp-Thr, 4=des-Trp, 5= des-Phe, and 6=Des-Thr. See Appendix 2 for chromatographic conditions.

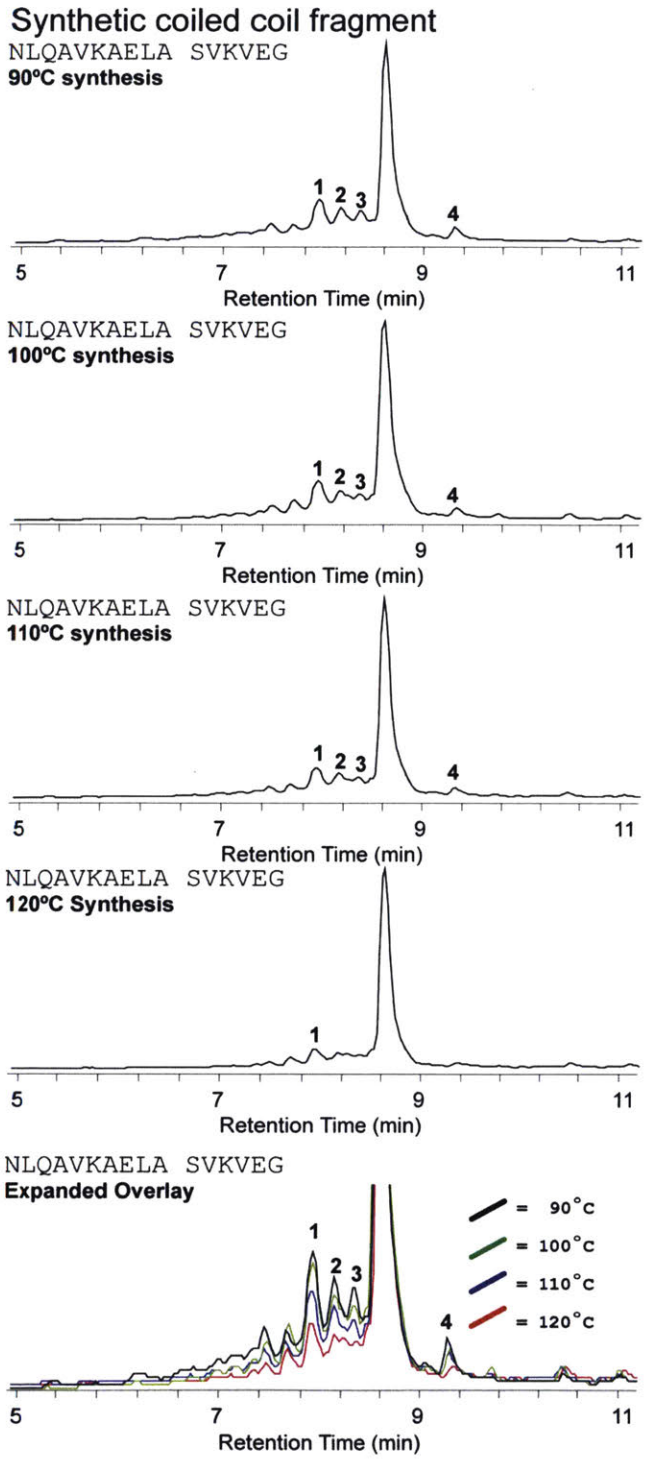


Figure 3.5.5 Synthetic coiled coil fragment with the best synthesis at 120°C. All syntheses were on 140-170mg of ChemMatrix Rink amide resin (0.5mmol/g) under default conditions in the 6mL resin vessel, and total ion chromatograms are shown. 1= des-Asn and des-Ans-Gln, 2=des-Gln, 3= des-Ala and des-Glu, and 4=des-Lys. See Appendix 2 for chromatographic conditions.

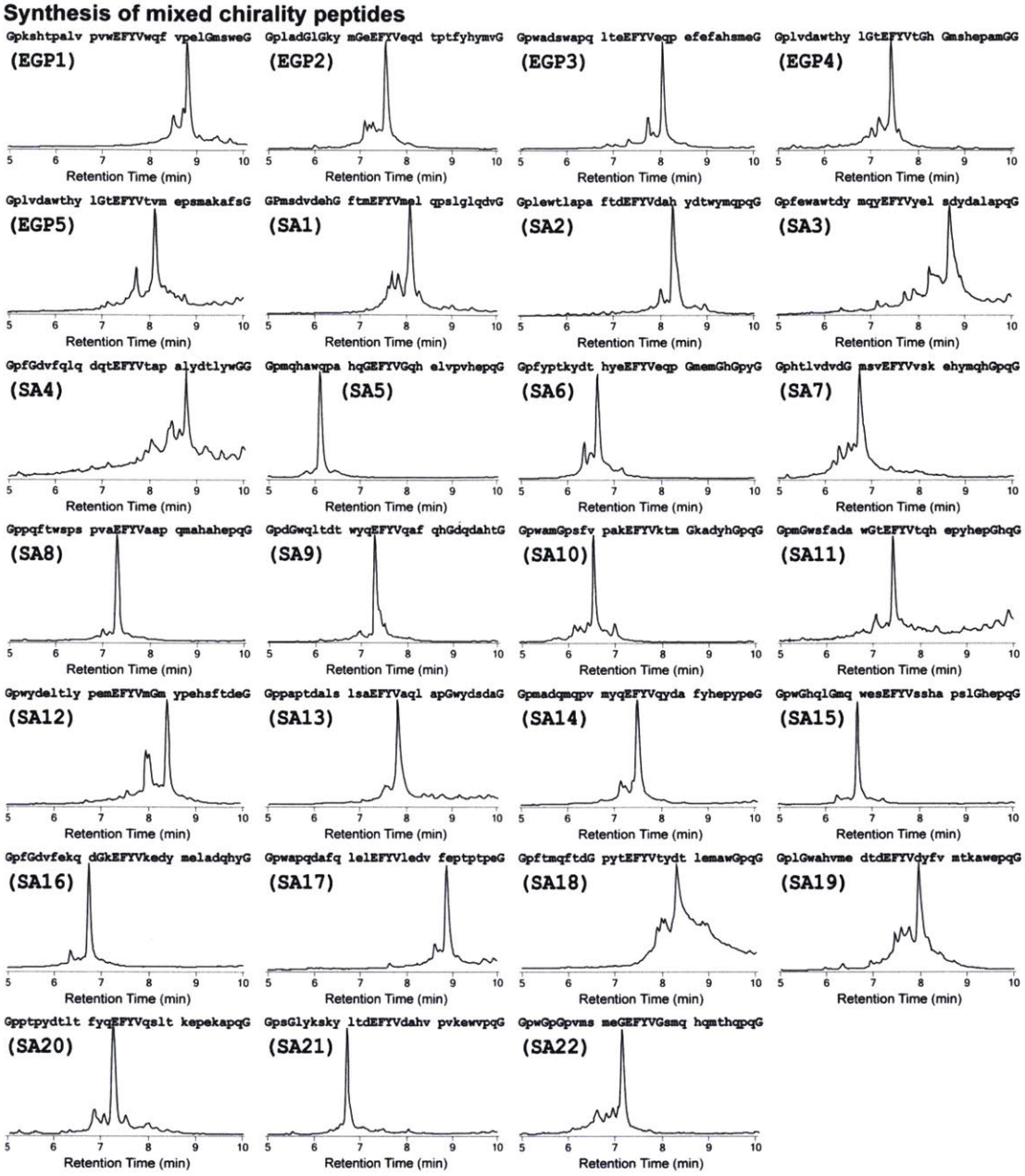


Figure 3.5.6 30 mixed chirality 30-mer peptides synthesized on the third generation AFPS. Capital letters represent L amino acids, and lower case letters represent D amino acids. In all cases, the major peak was the desired product, and sufficient material was isolated by preparative RP-HPLC for further studies (average estimated purity 46%, range 22%-73%). In contrast to synthesis with the second generation AFPS, the impurities in the peptides shown here were predominantly deletions, cleavage adducts, and aspartimide related products. Truncations were still present, but at dramatically lower levels (average 11%, range 0%-33%, vs average 28%, range 20%-45% previously). All syntheses were on 140-170mg of ChemMatrix Rink amide resin (0.5mmol/g) at 90°C under default conditions in the 6mL resin vessel, and total ion chromatograms are shown. See Appendix 2 for chromatographic conditions, mass spectral data, large format chromatograms, and assignment of the ten most prevalent compounds identified with MFE in each synthesis.

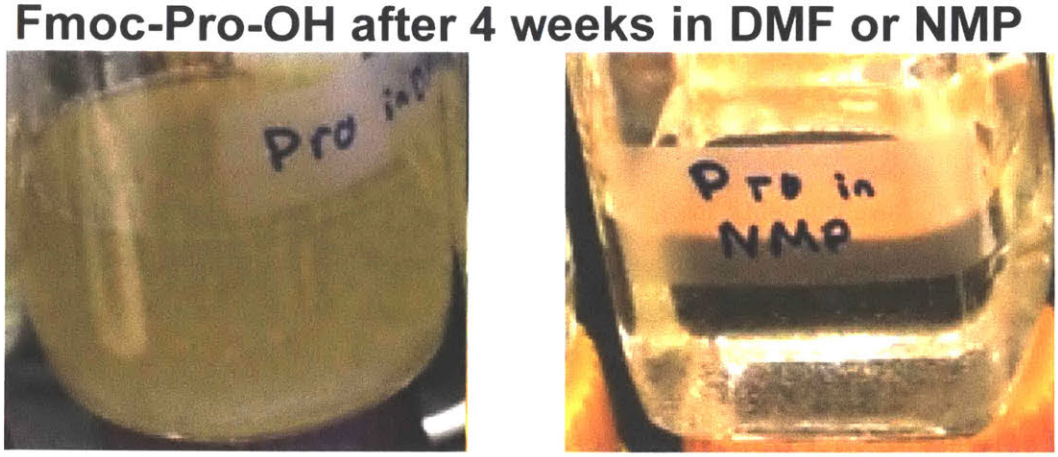


Figure3.5.7 Photographs of 0.4M solutions of Fmoc-Pro-OH in DMF (left) or NMP (right) after 4 weeks. The DMF solution turned into an opaque sludge, and there is a slight precipitate in the NMP solution. Proline did not precipitate from DMI, even after 10 weeks.

3.6 Fourth generation AFPS

3.6.1 Key findings

Use of amino acids dissolved in NMP in AFPS required that pump withdraw times be increased substantially to prevent cavitation, and consequently maximum flow rates and synthetic speed were reduced. Synthesis with NMP solutions was successful, however, and reagents have been used for up to six weeks with no decrease in synthetic quality, and no mechanical problems.

3.6.2 Effect of solvent on fast flow synthesis

Prior to designing a fourth generation AFPS, the effects of substituting NMP or DMI for DMF on fast flow peptide synthesis were evaluated. For these experiments, the manual fast flow peptide synthesizer was used to avoid preparing large numbers of stock solutions. The previously reported two minute cycle [68]was used to make ACP(65-74) and the JR-10mer at 90°C using DMF, NMP, or DMI as the solvent. To accentuate the effect of changing the solvent, all reagents and the wash solvent were prepared with the solvent under investigation.

As shown in Figure 3.6.1, the three syntheses of ACP yielded identical crude material. The three syntheses of JR-10mer, however, showed that DMI is strictly worse than the other two solvents. Using NMP lead to a different impurity profile, but was not better than DMF. There was a reduction in the Phe deletion and Phe-Trp double deletion, but a large increase in the Trp deletion.

More significantly, these model studies show that all three solvents are suitable for fast flow peptide synthesis. Therefore, the solvent a reagent is most stable in can be used, with a minimal effect on synthesis. Based on this information, a fourth generation AFPS was created with most amino acid stocks prepared in NMP, and proline and glycine stocks prepared in DMI. Piperidine and HATU continued to be dissolved in DMF, and DMF was still used for the wash solvent.

3.6.3 Design and characterization of fourth generation AFPS

The fourth generation AFPS used the same hardware at the third generation (Fig 3.5.1); it was marked by a substantial change in the software settings and synthetic protocol (Fig 3.6.2).

The denser and more viscous NMP solutions initially caused the amino acid pump to cavitate as in Figure 3.5.2C. The pump cavitation caused sub stoichiometric amino acid to be delivered and resulted in tetramethyl guanidinylated truncates during benchmark GHRH synthesis. The truncates were eliminated with longer pump refill times, which eliminated cavitation (Fig 3.6.2B). The longer pump refill times reduced the maximum achievable flow rate to ~40% of the rated maximum of each pump head, and the

cycle time was doubled to deliver the same quantities of reagents. Additionally, since the flow rate was lowered, the 10ft activation reactor temperature was reduced from 90°C to 70°C, yielding slightly cleaner GHRH, comparable to the manual control. Even without extensive optimization of these conditions, a number of peptides were successfully synthesized, including the difficult LL-37 human lysosomal antibacterial peptide[3,78] (Fig 3.6.3), and a number of all D peptides.

Importantly, no precipitates have yet been observed in amino acid reservoirs, and no new problems have arisen during these efforts.

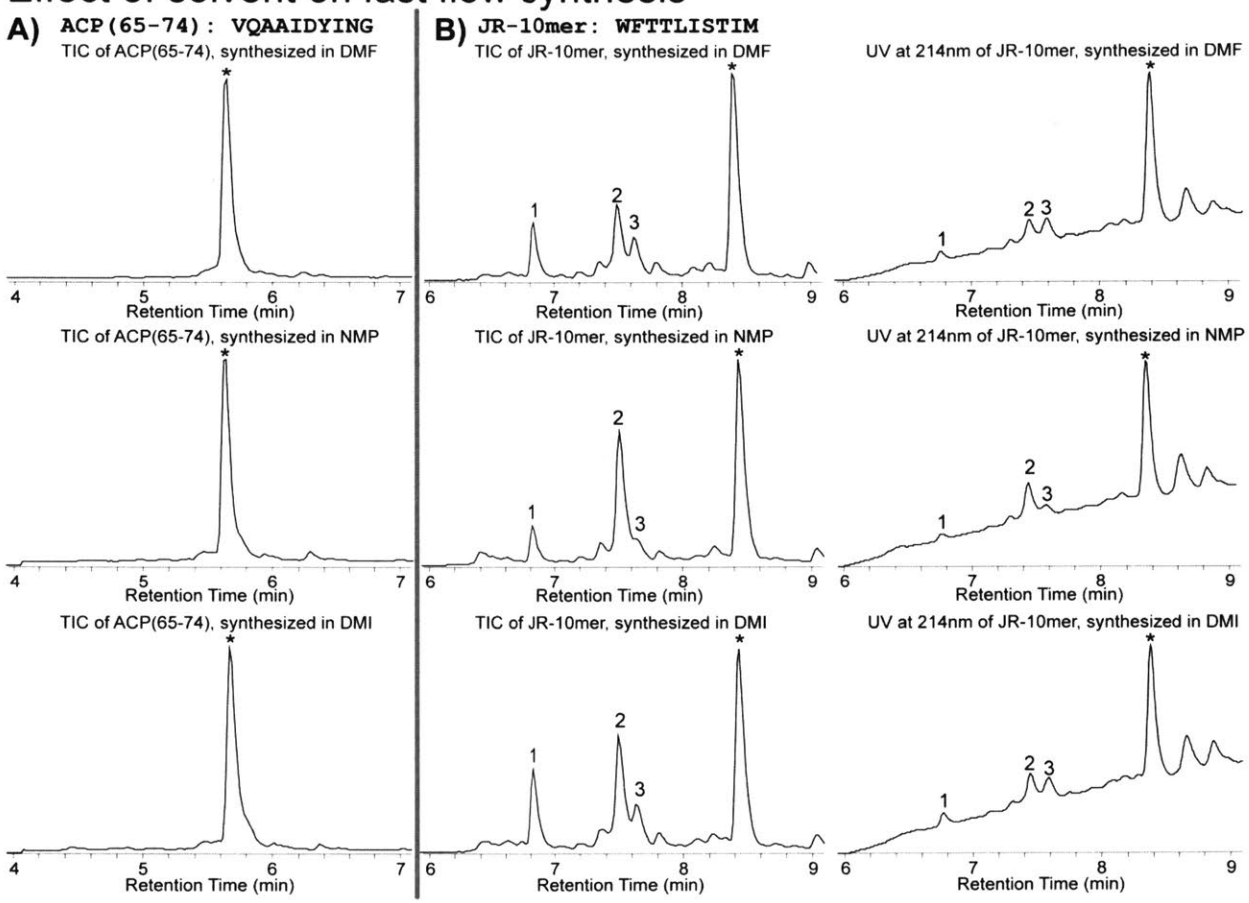


Figure 3.6.1 Effect of solvent on fast flow synthesis A) Syntheses of ACP(65-74) in DMF (top), NMP (middle), and DMI (bottom) showing that identical products are obtain with all three solvents. TICs are shown in all cases. **B)** Syntheses of JR-10mer in DMF (top), NMP (middle), and DMI (bottom) showing that comparable, but not identical, products are obtain with the three solvents. TICs are shown on the left to more clearly illustrate synthetic differences, and UV absorbance at 214nm is shown on the right to more accurately convey synthetic quality. All peptides were synthesized at 90°C on ChemMatrix Rink amide resin (0.5mmol/g) using two minute cycles on the manual fast flow synthesizer. * = Desired product; 3 = des-Phe; 2 = des-Trp 1 = des-Phe-Trp. See Appendix 2 for chromatographic conditions.

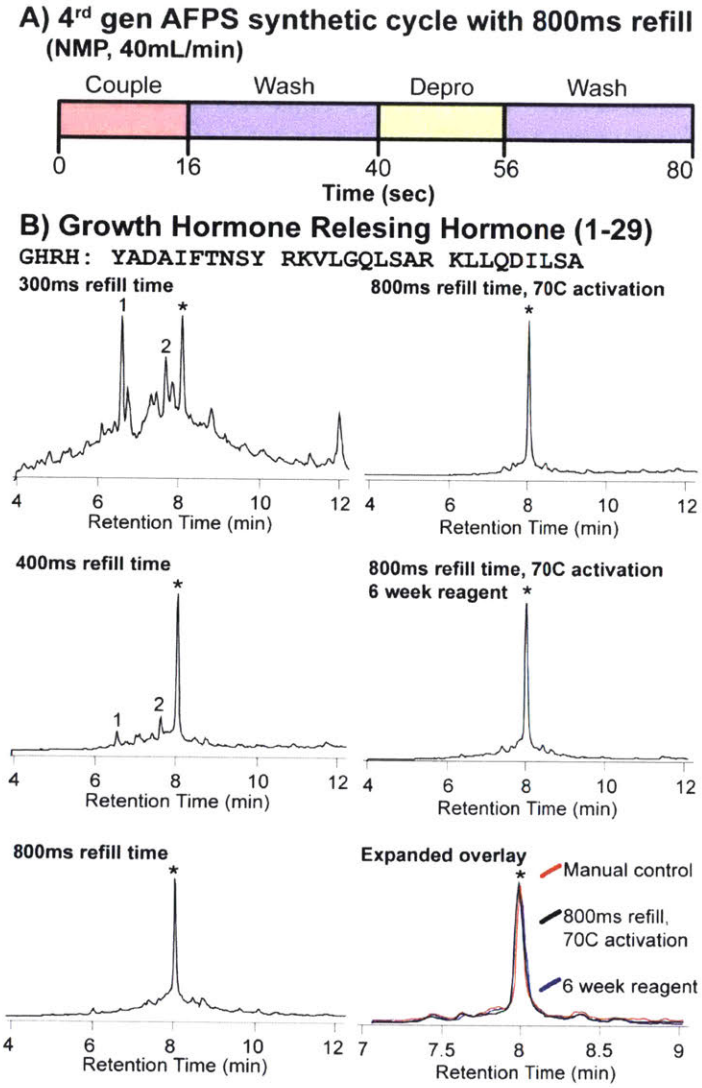


Figure 3.6.2 Fourth generation AFPS **A)** Fourth generation synthetic cycle using most amino acid stocks dissolved in NMP, and Pro and Gly dissolved in DMI. Run at 40ml/min. **B)** GHRH synthesized using most amino acids dissolved in NMP, and Pro and Gly dissolved in DMI. With inadequate pump refill times, the amino acid pump cavitates, but the activator pump does not, resulting in truncations. Under the best conditions, GHRH is similar to the manual control (67% estimated purity with 9% truncates vs 53% estimated purity with 20% truncates in the control). The synthesis was comparable using reagents that were on the instrument for 6 weeks (61% estimated purity, 10% truncates), but low levels of double incorporation were identified by MFE, consistent with slow loss of Fmoc in the amino acid reservoirs. * = GHRH; 1 = truncation with tetramethyl guanidinium cap before Arg-20, after coupling Lys-21; 2 = truncation with tetramethyl guanidinium cap before Arg-11, after coupling Lys-12. Total ion chromatograms are shown. See Appendix 2 for chromatographic conditions, mass spectral data, and annotated molecular feature extractions.

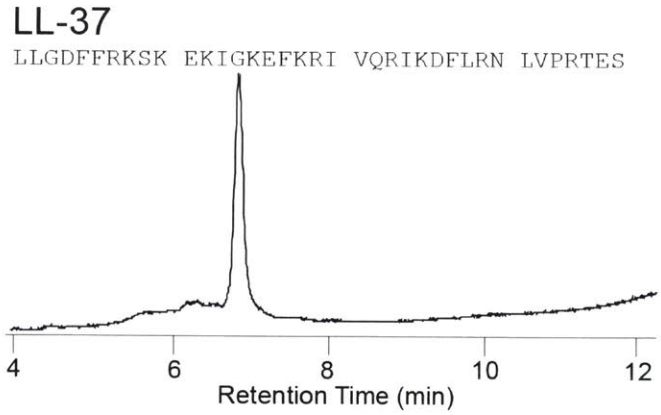


Figure 3.6.3 LL-37 synthesized on the fourth generation AFPS using amino acids dissolved in NMP, 800ms pump refill time, and 70°C activation on 170mg of ChemMatrix resin (HMPB linker, 0.5mmol/g). Integration of the TIC shown suggests a purity of 65% (MFE failed due to the high molecular weight). See Appendix 2 for chromatographic conditions and mass spectral data.

3.5 Conclusions

In conclusion, we have demonstrated an extremely rapid, robust, fully automated peptide synthesizer that incorporates amino acid residues five times faster than the fastest commercial offering[37]. Building on our previous work in which we accelerated peptide synthesis in a manually operated flow based instrument, we have further doubled the synthesis speed and robustly automated the process. This is accomplished without microwave heating, stirring, double coupling, double deprotecting, or capping by taking advantage of enhanced heat and mass transfer in flow. Four critical advances enabled robust automation and further acceleration of synthesis, and we believe these advances offer a template for accelerating a wide variety of solid supported syntheses with an automated flow instrument.

First, transitions between reagents were very rapid with stoichiometry that was always correct. In a true continuous flow process, the transitions between reagents are negligible. The time required to reach equilibrium is brief compared to the total run time, and poor quality product from the pre-equilibrium period can be discarded. In contrast, in solid supported flow based synthesis, the transitions dominate the synthesis. Incorporation of each amino acid residue required four transitions between solvents and reagents, and a steady state was rarely, if ever, achieved under our operational conditions. Furthermore, if byproducts were formed at any point in the synthesis, including during the dozens of transitions, that product was bound to the resin as an impurity. It could not simply be mechanically discarded: it had to be separated by RP-HPLC, and drastically reduced the yield.

To ensure transitions were rapid and that the stoichiometry was always and everywhere correct, a low volume reagent delivery manifold and HPLC pumps with low non-displaced volume were used, and HPLC pumps were run synchronously. The stoichiometry was verified by high speed videography of the process running with dyed solvent instead of reagent.

Next, reagents were stable both in the storage reservoirs, and as they were used at temperature. Three mechanisms by which peptide synthesis reagents degrade were identified. First, the Trityl protected histidine monomer epimerized, and this problem was particularly pronounced at elevated temperature. Using Boc protection instead of Trityl overcame this issue. Second, Fmoc protected amino acids were not stable when co-dissolved with uronium activators. This was overcome by storing amino acids and activators separately, and mixing them immediately before use. Third, the N-α Fmoc protecting group was found to be unstable in DMF. This problem was mitigated by using NMP as the primary solvent for amino acids.

Further, the resin bound intermediates were stable under the conditions used. We expected to find a maximum temperature for SPPS during the temperature studies, but there was no evidence that the peptides we explored degraded at temperatures up to 120°C, the maximum service temperature of the

reactor module. Surprisingly, no additional side products were observed, and the crude purity of the peptides continued to improve as the temperature was increased.

Single deletion products were expected to become less pronounced as temperature increased not only because acylation is inherently faster, but also because peptides are less likely to aggregate. Deletion products arise during peptide synthesis because intramolecular hydrogen bond formation within the peptidyl resin sterically occludes the reactive N-terminus of the growing peptide chain, giving rise to "difficult sequences"[79,80]. At higher temperature, however, formation of hydrogen bonds becomes thermodynamically less favorable as hydrogen bond formation is an enthalpically driven process with an entropic barrier[81]. Therefore, increasing the temperature of peptide synthesis should be highly effective in disrupting aggregation and overcoming difficult couplings. In this work, we observed a substantial increase in synthetic quality with an increase in temperature, and temperatures above 120°C may be of further benefit in peptide synthesis.

Finally, a flexible, modular control package was created. The hardware and software used provides exquisite control over the process, and can readily be repurposed to perform similar reactions with more or fewer pumps, valves, and sensors. Similarly, the graphical user interface is intuitive and easy to use for peptide synthesis, and can be modified to provide a high quality user experience for many other types of iterative solid supported chemistry.

These advances at the interface of chemistry and engineering overcame numerous coupled chemical and mechanical problems to enable rapid peptide synthesis, and we believe they can be leveraged to enhance other solid phase syntheses.

Fast flow PMO synthesis

The work presented in this chapter is currently under review, and Appendix 3 is derived from the supporting information

Mark D. Simon*, Kyle A. Totaro*, Ming Zhou, Hong Zong, Gunnar J. Hannson, and Bradley L. Pentelute. Fast Flow Phosphorodiamidate Morpholino Oligomer Synthesis. *(Under Review)*

* Equally contributing authors

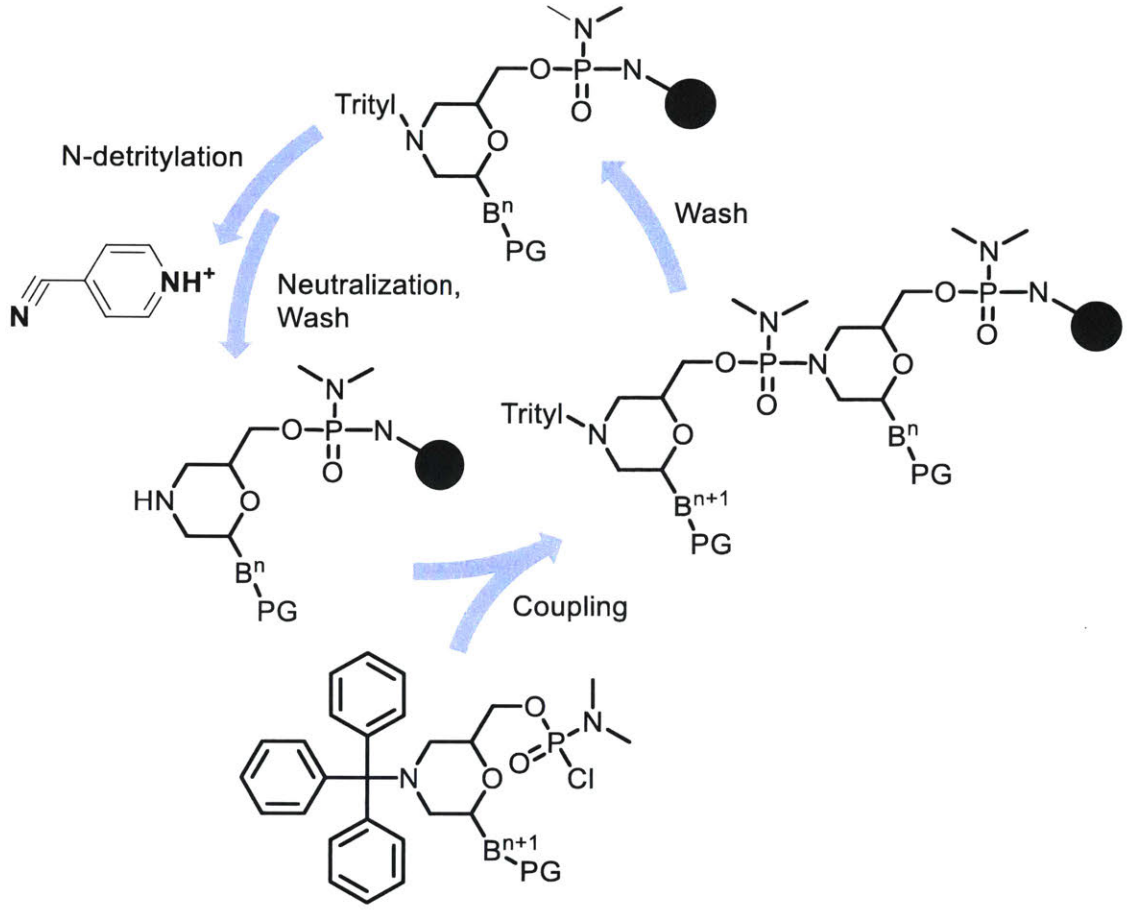
4.1 Introduction

Antisense therapy is an attractive route for treating difficult genetic diseases and viral infections. Antisense therapeutics are synthetic RNA mimetics that bind to mRNA via complementary base-paring interactions and cause alternative splicing or inhibit translation, reducing expression or yielding altered proteins[8,82]. In this way, critical viral proteins can be deactivated, exons with lethal frame shift mutations can be excised to yield a functional truncated protein, or oncogenic gene expression can be suppressed. Phosphorodiamidate morpholino oligomers (PMOs) are a class of antisense theraputics in which the 5-member ribose ring of RNA has been replaced with a 6-member morpholine ring, and phosphate linkages have been replaced with uncharged posphorodiamidates[83]. These modifications make PMOs resistant to nucleases[84] and substantially more cell permeable than naked RNA[85]. Several PMOs, including second-generation structural analogs and peptide-PMO conjugates, are currently in the clinic with a first in class Duchenne muscular dystrophy drug recently approved[86] and trials underway for the treatment of Dengue[9], Marburg[10], Ebola[10], and influenza viral infections[87].

Despite the breadth of potential disease targets, lengthy preparation of PMOs limits their exploration and application. To alleviate this problem, and enable efficient identification and production of lead compounds, we developed a fast, flow based PMO synthesis procedure based on knowledge presented in Chapters 2 and 3.

Unlike synthetic DNA and RNA, which can be amplified with biological catalysts from tiny amounts of chemically synthesized material[88], PMOs are not biologically accessible and must be chemically synthesized at the desired scale. Typically, a solid phase synthetic approach is used in which PMOs are synthesized from the 5' end to the 3' end on a crosslinked polystyrene support with relatively high loading (0.1-1mmol/g)[48,83,89,90]. Nucleotides are incorporated using 3' Trityl protected chlorophosphoramidate monomers in solution with base, and 3' Trityl protection is removed in dilute acid to regenerate a reactive 3' amine for the next coupling (Scheme 4.1.1). Monomer coupling dominates the synthetic time, and was the major focus of this accelerated protocol. Previous PMO synthesis required several hours to form the phosphorodiamidate linkage, and this was recently reduced to 5 minutes using phosphoramidite chemistry[91]. In this work, we have accelerated direct phosphorodiamidate formation from chlorophosphoramidates in a fast flow platform, requiring only 2 minutes for complete coupling, and are optimistic that this approach can further accelerate the phosphoramidite approach as well.

Solid Phase PMO Synthesis Scheme



Scheme 4.1.1 PMO flow synthesis scheme. Each nucleotide is incorporated as a Trityl protected phosphorchloride. Detritylation and neutralization regenerate a reactive 3' amine for the next coupling reaction.

4.2 Results and Discussion

PMO synthesis was translated into a fast flow format following the template established in our earlier work on peptide synthesis, presented in Chapter 3. Mechanically robust hardware previously developed to enable rapid solid supported chemistry at high temperature was reused with minimal modification, allowing the work in this chapter to focus on chemical aspects of accelerating PMO synthesis. To accelerate synthesis, we focused on increasing the coupling temperature and reactivity of the chlorophosphoramidate monomers, while maintaining the stability of the protected, resin bound oligomer during synthesis.

First, high boiling solvents were explored for PMO synthesis at elevated temperature. Second, the stability of activated nucleotides and resin bound oligomers were investigated under various conditions to establish limits on reaction temperature. Third, the time scale of coupling and deprotection at elevated temperature was studied. Finally, these insights were used to develop an accelerated, flow-based PMO synthesis protocol that incorporates a nucleotide every 13 minutes.

4.2.1 High boiling solvents for PMO synthesis.

Increasing the reaction temperature was our most powerful tool to accelerate biopolymer synthesis, so we began our studies by investigating high boiling solvents to replace dichoromethane (DCM) during 3' detritylation. In previous protocols[83,89,90], DCM was the preferred solvent for washes and detritylation, while either 1,3-dimethyl-2-imidazolidinone (DMI) or *N*-methyl-2-pyrrolidone (NMP) were preferred coupling solvents. The latter solvents were appropriate for flow synthesis at high temperatures, but the atmospheric pressure boiling point of DCM is only 42 °C. Although it was possible to use solvents above their atmospheric pressure boiling point, and this was eventually done, we sought to improve process safety and reliability by substituting a higher boiling solvent for DCM during detritylation.

To screen high boiling solvents for PMO synthesis, a streamlined, small-scale (4μmol) room temperature batch synthesis protocol was used with 15 minute couplings of 25 eq (0.1mmol) of monomer (see Appendix 3). This abbreviated batch protocol afforded the model tetramer 5'-Tail-ACGT-Trt-3' in good yield. The presence of minor deletion products, however, indicated it would be sensitive to decreases in reaction performance, an important consideration when using small model systems to make conclusions applicable to long and difficult sequences.

Alternative solvents were investigated by repeated synthesis of the above model tetramer, 5'-Tail-ACGT-Trt-3'. To examine the effect of detritylation solvents, each detritylation was performed with 100mM (~2% w/v) 4-cyanopyridine trifluoroacetate, 1% (v/v) ethanol, and 20% (v/v) trifluoroethanol

dissolved in dioxane, diglyme, toluene, NMP, dimethylformamide (DMF), dichloroethane (DCE), acetonitrile, or DCM (Control). All couplings in these experiments were performed with 0.2M chlorophosphoramidate monomer and 0.4M N-ethylmorpholine (NEM) in DMI. The products from each of these eight syntheses were examined by LC/MS and DCM, DCE, and toluene were found to be suitable detritylation solvents, with DCE slightly outperforming the DCM control (Table 4.2.1).

To select between DCE and toluene as the detritylation solvent, these were evaluated as coupling solvents along with DMI (control), NMP, and DMF. In each of 5 experiments, every coupling in the model tetramer was carried out with 0.2M monomer and 0.4M NEM in one of these solvents (Table 4.2.1). DCE performed the best during the coupling trials, indicating residual DCE would have minimal effect on coupling, so it was selected as the detritylation solvent.

Table 4.2.1 High boiling solvents for PMO synthesis

Coupling solvent	Detritylation Solvent	Product[a]	G-Deletion	C-Deletion
DMI	DCM	88	2.0	0.6
DMI	**DCE**	**91**	**0.40**	**0.08**
DMI	Toluene	84	0.43	0.19
DMI	Dioxane		Failed	
DMI	Acetonitrile		Failed	
DMI	Diglyme		Failed	
DMI	NMP		Failed	
DMI	DMF		Failed	
DCE	DCE	86	3.1	0.17
Toluene	DCE	47	34	1.5
NMP	DCE	87	0.31	0.12
DMF	DCE	87	0.98	0.72

[a] Values extracted from LC-MS data. 3' Trityl protection is lost in the ESI source, so abundances of the indicated species and their detritylated conterparts were summed following identification with Agilent's automated molecular feature extraction utility. A and T deletions were very low intensity or unidentified; the last 10-20% of assigned ions were unknown species.

4.2.2 PMO reagent stability

We investigated the stability of both morpholino monomers and resin bound protected morpholino oligomers at various temperatures. First, the stability of the morpholino subunits was evaluated. In sealed vials, coupling solutions of 0.2 M moA and 0.4 M DIEA in DMI were heated for 5 minutes at temperatures from 90°C to 150°C in 10°C steps with or without 0.21M LiBr, a previously reported catalyst for phosphorodiamidate formation[90]. After 5 minutes, coupling solutions were quenched with a 10% (v/v) solution of piperidine in NMP and analyzed by LC-MS. The moA monomer was stable at 90°C, but degraded to unknown products at higher temperature, with LiBr accelerating degradation (Fig 4.2.1). The other monomers, moC, moG, and moT, were then evaluated under the same conditions at 90°C and found to be stable (Fig A3.5). These studies established 90°C as a conservative upper temperature limit for PMO synthesis, although in principle reagents could be rapidly brought to higher temperatures in flow and used before they substantially degrade.

Next, the stability of the resin-bound PMO was determined in much the same way as monomer stability. Resin-bound model PMO tetramer with either a Trityl protected or free 3'-end was placed in sealed vials with each solvent or reagent to be used during synthesis, and incubated at 90°C. After 4 hours, representative of the exposure to each reagent during a long synthesis, the resin was isolated, washed, cleaved, and analyzed by LC-MS. Tritylated and detritylated resin-bound PMO was not stable in the presence of the 4-cyanopyridine trifluoroacetate detritylating solution, showing degradation to low molecular weight products via acidolysis of the phophorodiamidate linkages. Furthermore, detritylated resin was unstable in most solvents and reagents, but stability improved when the detritylated resin was neutralized with a 5% (v/v) solution of DIEA in 3:1 DCE and isopropanol prior to stability studies (Fig A3.4). This demonstrated the utility of neutralizing the resin during synthesis, which was unclear from the literature.

Finally, we performed the same series of experiments on a longer PMO sequence, to amplify any deleterious effects. A 10 residue PMO, Eteplirsen(1-10), was synthesized in batch, according to literature procedures, and incubated at 90°C with the reagents and solvents used in the previous experiments. Additionally, trials were conducted with NEM replacing DIEA and DCM replacing DCE. Trials with DCM were conducted at 70°C to reduce pressure in the glass vials. After 4 hours, the resin was isolated, washed, cleaved, and analyzed by LC-MS. Interestingly, at 90°C, DCE and DIEA both promote degradation of the resin-bound 10mer PMO. DCM and NEM however were not found to degrade the morpholinyl resin (Fig 4.2.2).

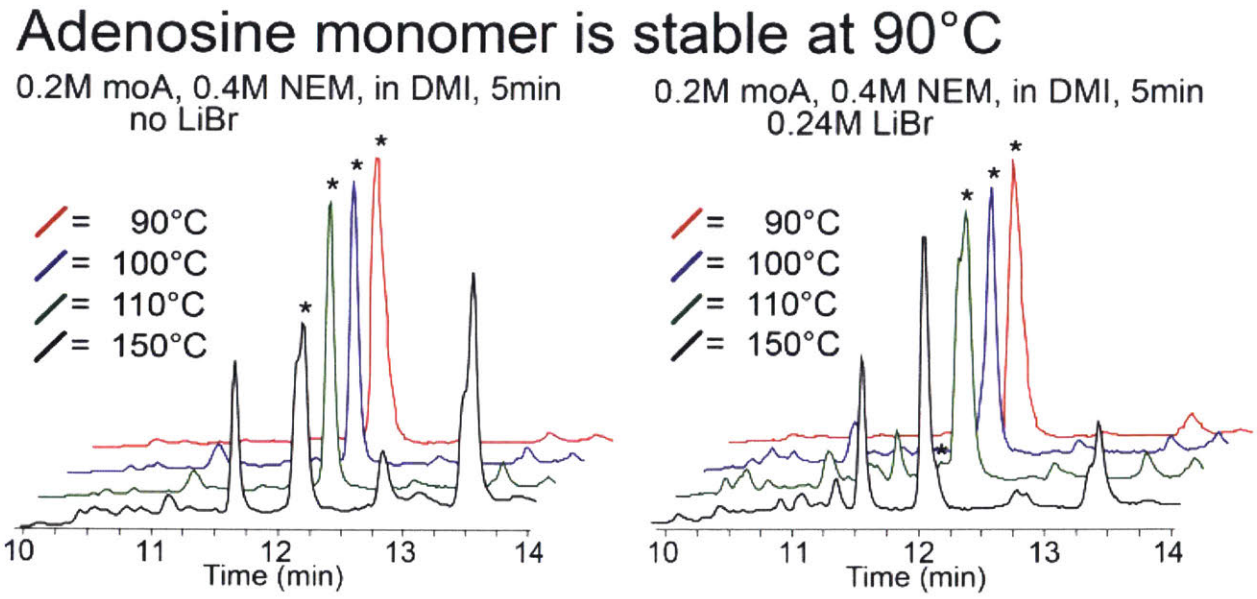


Figure 4.2.1 moA is stable at 90°C for 5 minutes, but degrades to a product 156 daltons less massive at higher temperatures. * = desired quenched product; the primary degradation product is an early eluting shoulder.

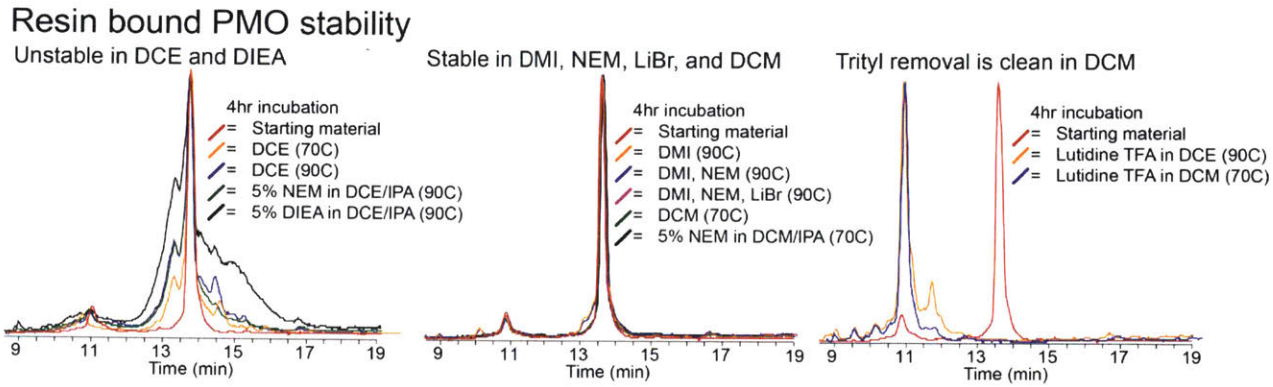


Figure 4.2.2 Resin bound PMO stability. Eteplirsen(1-10) is stable with DMI, DCM, LiBr, and NEM, but not with DCE or DIEA. Trityl is cleanly removed with DCM as the detritylation solvent.

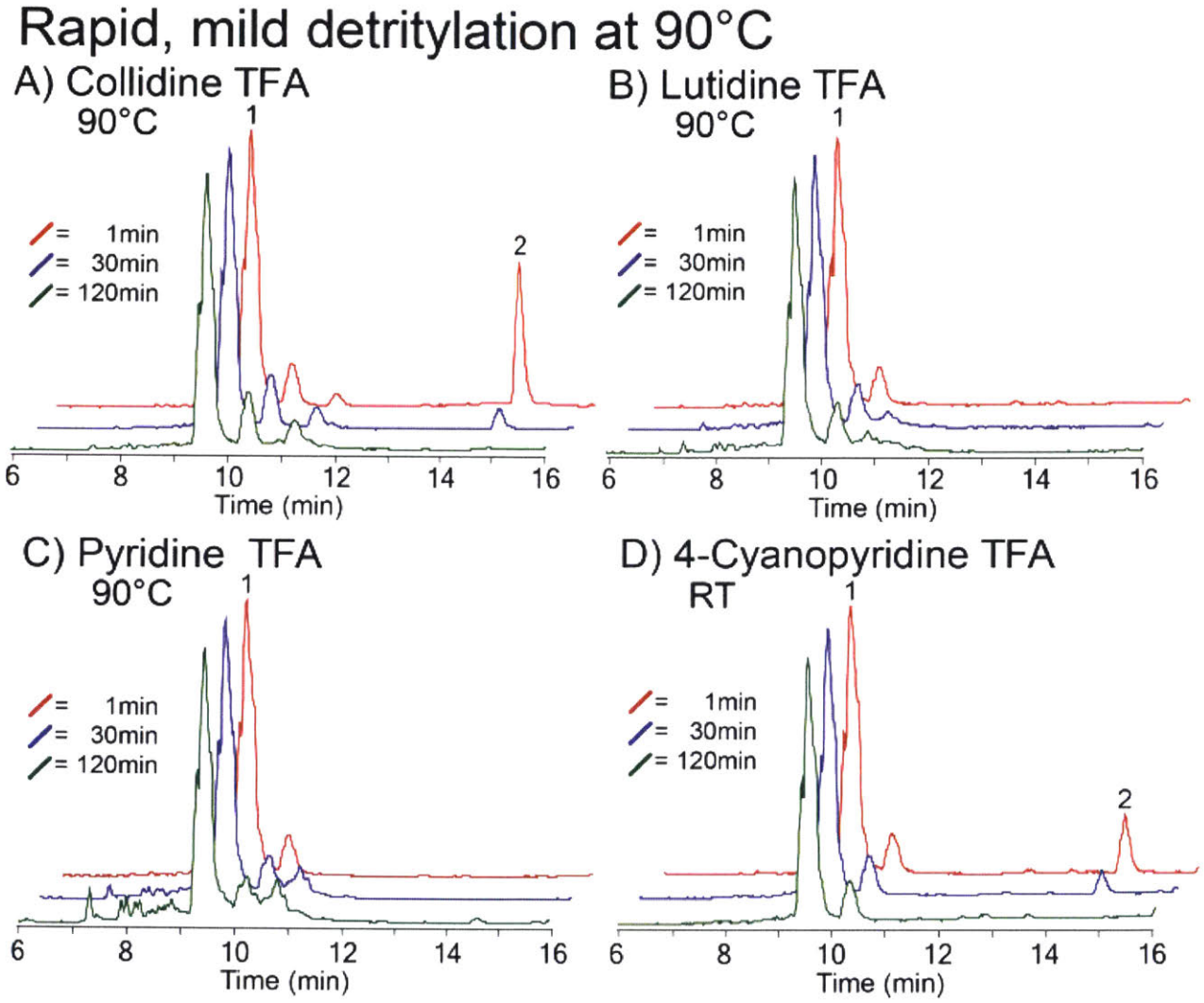


Figure 4.2.3 Rapid, mild detritylation at 90°C Lutidine trifluoroacetate is an ideal Trityl removal reagent at 90°C **(B)**. Removing Trityl more rapidly than collidine TFA **(A)** , but without the degradation caused by pyridine TFA **(C)**. Cyanopyridine TFA cleanly detritylates at room temperature **(D)**.

4.2.3 Detritylation reactivity

100mM 4-cyanopyridine trifluoroacetate detritylating solution causes severe acidolysis at high temperature, so we evaluated less acidic pyridine triflouroacetates[48]. 3' tritylated model PMO tetramer was incubated with trifloroacetate salts of 4-cyanopyridine (control, pK_a = 1.9), ethyl nicotinate (pK_a = 4.5), pyridine (pK_a = 5.5), 3,5-lutidine (pK_a = 6.2), and 2,4,6-collidine (pK_a = 7.4) at room temperature or 90 °C. Time points were taken at 1, 30, and 120 minutes. As expected, we observed an increase in detritylation time as pK_a increased (Fig 4.2.3). At 90°C, 3,5-lutidine trifluoroacetate solution detritylated the model PMO completely in under 1 minute with minimal degredation, even after 2 hours. In contrast, 2,4,6-collidine trifluoroacetate required longer than 30 minutes for detritylation at 90°C. Pyridine trifluoroacetate and more acidic reagents extensively degraded the PMO after 2 hours.

4.2.4 Coupling reactivity

The coupling of nucleodsides to resin is the longest step in PMO synthesis, and increasing the coupling rate was our primary objective in these studies. To assess the time scale of coupling, model PMO tetramer was prepared at room temperature with each coupling step performed for 2, 5, or 15 minutes. Coupling was complete in 15 minutes, but not 5 minutes. Assuming the reaction rate doubles for every 10°C increase in temperature, coupling was expected to be complete in 30 seconds or less at 70°C. The same experiment was performed, but the coupling steps were performed in a water bath at 70°C for 1, 5, or 15 minutes. Remarkably, coupling was complete after 5 minutes, but not after 1 minute, showing less than the expected rate enhancement.

To further enhance the coupling rate, catalysts were evaluated to make the chlorophosphoramidate monomers more reactive[90]. On a 0.1 mmol scale with only 5eq of monomer, the streamlined small scale room temperature timeline used to screen solvents yielded model PMO tetramer that was substantially contaminated with deletion products, making this a convenient model system to study additives for rate enhancement. In each of 7 experiments, every nucleotide was coupled with a solution of 0.2M monomer, 0.4M NEM, and 0.24M of one of the following additives in DMI: lithium bromide, lithium iodide, lithium triflate, tetrabutyl ammonium bromide (TBAB), dimethyl amino pyridine (DMAP), or N-methyl imidazole (NMI).

These experiments show reactivity is primarily enhanced through lithium catalysis, not halogen exchange. Adding TBAB to the coupling solution had little effect on synthesis, but all four lithium salts performed well. Interestingly, synthesis with lithium triflate and lithium iodide had more prevalent cytosine deletion products than the synthesis with lithium bromide, indicating exchange of the chloride for a better leaving group is detrimental to moC coupling (Table 4.2.2).

Similarly, synthesis with DMAP and NMI resulted in no high molecular weight products, suggesting that the activated phosphoroamides are too reactive to use in manual PMO synthesis. Analysis of the monomer solutions showed the formation of unreactive phosphoroanhydrides, corroborating this conclusion. Other hard, small Lewis acids such as Mg^{2+} and first row transition metal ions may offer enhanced reactivity without the rapid degradation of a more activated phosphorous center, but these studies remain for future work. Here it was decided to retain the previously reported LiBr catalyst.

Table 4.2.2 PMO Coupling catalysts [a]

Catalyst	Product	G-Deletion	C-Deletion
None	64	10	5.8
TBAB	66	7.1	4.7
LiCl	75	3.4	2.5
LiBr	72	3.3	3.4
LiI	74	2.2	5.0
LiOTf	71	2.6	6.5

[a] Values extracted from LC-MS data. 3' Trityl protection is lost in the ESI source, so abundances of the indicated species and their detritylated conterparts were summed following identification with Agilent's automated molecular feature extraction utility. About 20% of assigned ions were unknown species.

4.2.5 Design of flow synthesizer

To leverage the above findings for rapid PMO synthesis, the apparatus shown in figure 4.2.4 was assembled, based on our previously reported manual fast flow peptide synthesizer[34]. The instrument was designed to deliver preheated solvents and reagents, rapidly switch between reagents, maintain high concentrations of reactive components by continuously delivering fresh reagent, and continuously monitor the synthesis via UV absorbance of the effluent.

An HPLC pump was used to deliver DCE and DMI wash solvents, detritylation reagent, and neutralization reagent for the common washing, neutralization, and deprotection steps, and a syringe pump was used to deliver small volumes of freshly prepared coupling reagents. The HPLC pump solvent was selected with a manually actuated four position selector valve, and the HPLC pump outlet was connected to the preheat loop and reactor with a luer lock quick connect. During the coupling, the quick connect was disconnected and moved to a syringe with coupling reagent.

Solvents and reagents were passed through a preheat loop, and into the reactor with resin, then through a UV detector to waste. The reactor and preheat loop were submerged in a heated water bath for steps above room temperature, and in an unheated water bath for steps at room temperature. When DCM was used above its atmospheric pressure boiling point, a 250psi back pressure regulator was placed between the reactor outlet and UV detector, and bypassed for coupling steps (the syringe pump could not generate 250psi and DMI was never used above its boiling point). A 20psi back pressure regulator on the UV detector outlet prevented still warm DCM from boiling in the detector (Fig 4.2.4).

The detritylating solution was designed to quench the trityl cation, forming products with absorbance overlapping the pyridine trifluoroacetates. To regenerate the intensely yellow Trityl cation with well separated visible absorbance, a second HPLC pump was used to infuse methane sulfonic acid (20% v/v in DCM) into the reactor effluent before it was passed through the UV detector (Fig A3.6).

The second-generation reaction vessel described previously for fast-flow peptide synthesis was used for PMO synthesis for its chemical compatibility, mechanical reliability, and optical access[34]. It was equipped with finer 2μM frits to better retain the finer PMO synthesis resin, and otherwise unmodified.

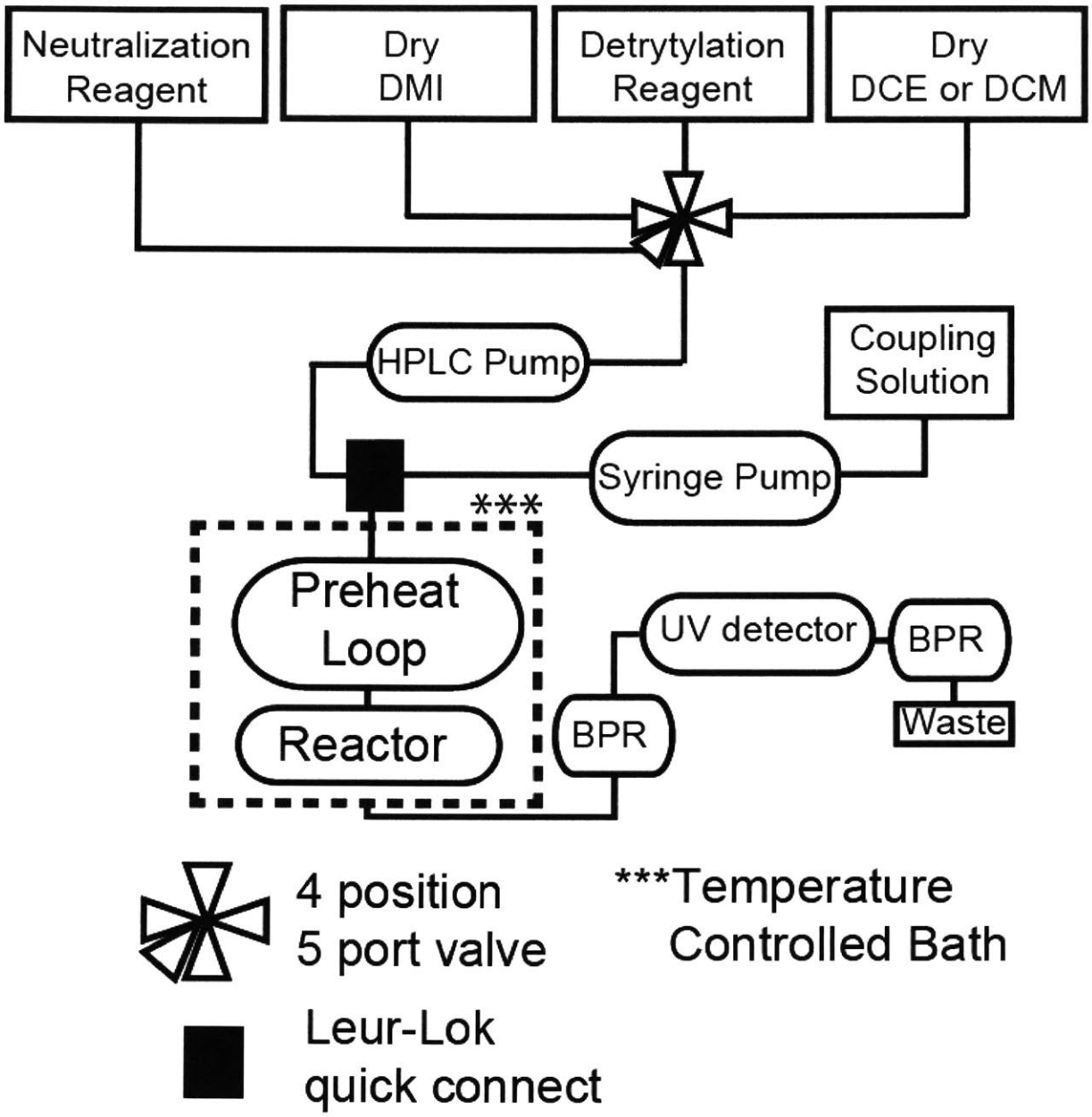


Figure 4.2.4 Fast flow PMO synthesizer schematic.

4.2.6 Leveraging stability and reactivity data

Based on reactivity and stability information from the above batch studies, the following flow protocol was used with 200mg of tritylated, tail loaded resin presowllen in DCE for 5 minutes. Despite the finding that prolonged exposure to DCE degraded the resin bound PMO, it was believed that brief exposure during fast flow synthesis may be acceptable, and DCE was initially used for its higher boiling point. First, 10ml/min of 100mM 4-cyanopyridine trifluoroacetate in 4:1 DCE:TFE with 1% ethanol was used to detritylate the resin for two minutes. The resin was then washed with DCE for 30 seconds at 10ml/min, and neutralized with 5% DIEA in 3:1 DCE:IPA for one minute at 10ml/min. After a second DCE wash, the resin was washed with DMI (10ml/min, 30 seconds) in preparation for the coupling step. The coupling solution consisted of 0.5mmol (5eq) of morpholino subunit dissolved to 0.2 M in DMI with 0.24 M LiBr and 0.4 M DIEA. This solution was delivered via syringe pump at 3 mL/min for approximately 1 minute. DMI washed the morpholino subunit through the reactor at the same flow rate for 3 minutes, to allow for complete coupling at high concentration. A final 30 second DCE wash was performed at 10 mL/min before repeating the procedure to install the next nucleotide. Isolated resin was cleaved in a mixture of 28% aqueous ammonia:40% (aq) methylamine (1:1) for 15 minutes at 65C.

The model PMO tetramer was synthesized four times with this flow protocol at different temperatures. In all cases, the product was left 3' tritylated to help distinguish deletion products, which retain the 3' Trityl protection, from truncations, which do not. At room temperature, the desired product was heavily contaminated with deletion products, as expected. In contrast, at 70 °C, the synthesis was substantially improved, though deletion products were still present and PMO truncates became prominent. To limit PMO degradation in 4-cyanopyridine trifluoroacetate, detritylations were then performed at room temperature, with couplings at 70°C or 90°C. Both syntheses gave substantially improved products, and the 90°C synthesis was slightly cleaner.

Next, to eliminate the numerous mechanically inconvenient transitions from room temperature to 90°C, a synthesis was performed with all steps at 90 °C and 3,5-lutidine replacing 4-cyanopyridine in the detritylation reagent. Finally, DIEA in the coupling and neutralization solutions was replaced with NEM, eliminating an early eluting impurity 1Da less massive than the product, and resulting in a very clean synthesis at a single temperature (Fig 4.2.5).

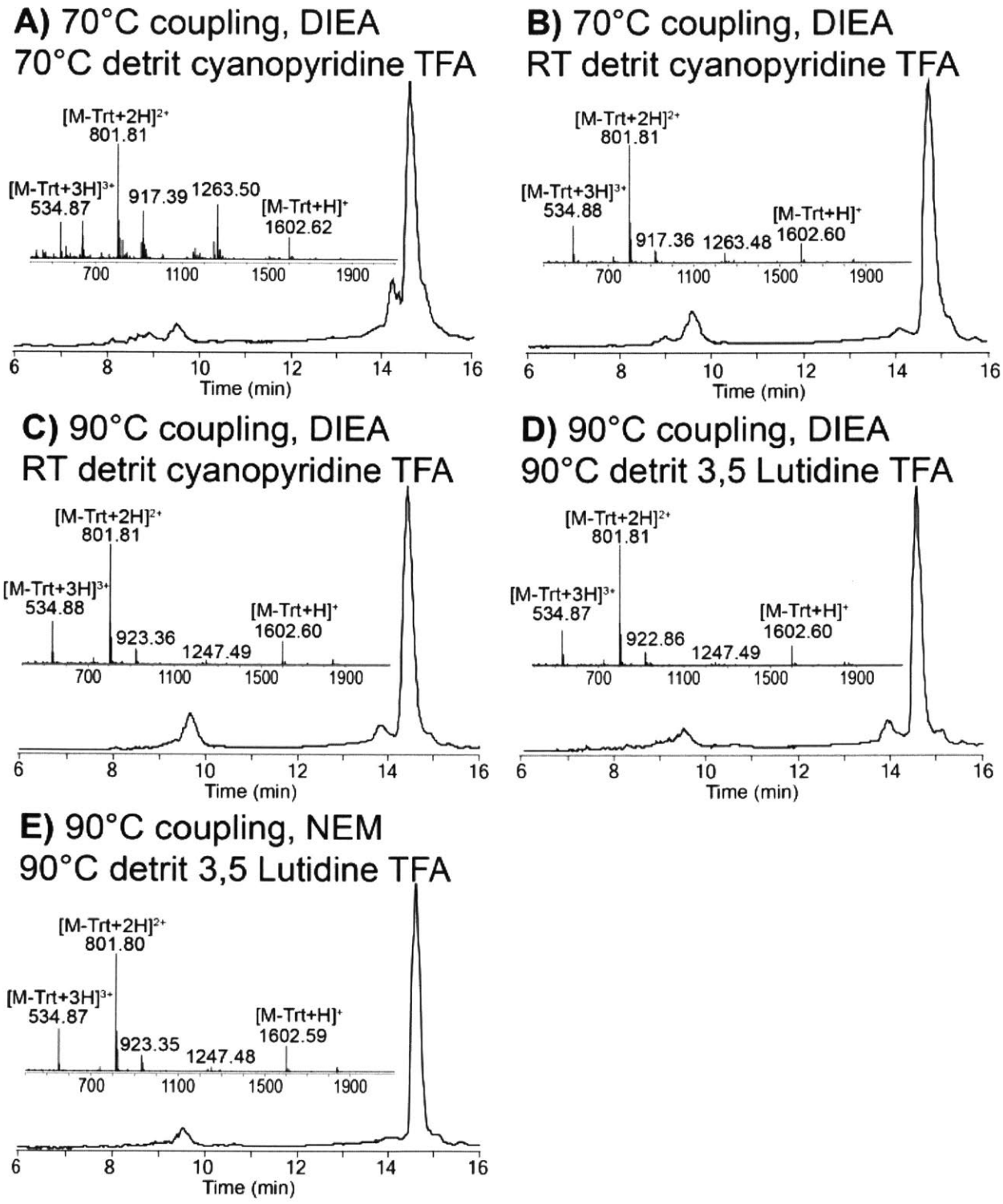


Figure 4.2.5 Synthesis of the model tetramer 5'-Tail-ACGT-Trt-3' under several conditions. In each panel, the total ion chromatogram is shown with the mass spectrum integrated from 6 to 16 minutes inset. The cleanest synthesis used NEM as the coupling base and lutidine trifluoroacetate as the detritylation acid.

4.2.7 Synthesis of Eteplirsen(1-10)

With these model studies complete, the 10 residue 5' end of Eteplirsen was synthesized using the two best tetramer synthesis conditions: 90 °C couplings with either room temperature detritylations using 4-cyanopyridine trifluoroacetate or 90 °C detritylations using 3,5-lutidine trifluoroacetate. In all couplings the amount of morpholino subunit delivered to the reactor was doubled to 1mmol (10eq) and delivered over 2 minutes. This change reflects reported routine double couplings in longer PMOs[48]. Cleavage was in aqueous ammonia and ethanol (1:4) for 16 hours at 55°C.

Both protocols afforded the desired 10 residue PMO in good yield by LC-MS, but the 90°C detritylations resulted in more low molecular weight byproducts, likely due to prolonged DCE exposure at high temperature. MALDI-TOF analysis confirmed this observation, showing the room temperature detritylation protocol to be comparable to the batch control, which was quite pure.

To reduce truncations while maintaining a single temperature throughout synthesis, DCE was replaced with DCM, despite the challenges of operating well above the atmospheric pressure boiling point. Synthesis was repeated with 3,5-lutidine trifluoroacetate based detritylations, and all steps at 70°C or 90°C. Synthesis at 90°C afforded the highest quality crude PMO, which was comparable to synthesis with 90°C couplings and room temperature 4-cyanopyridine trifluoroacetate detritylations (Fig 4.2.6).

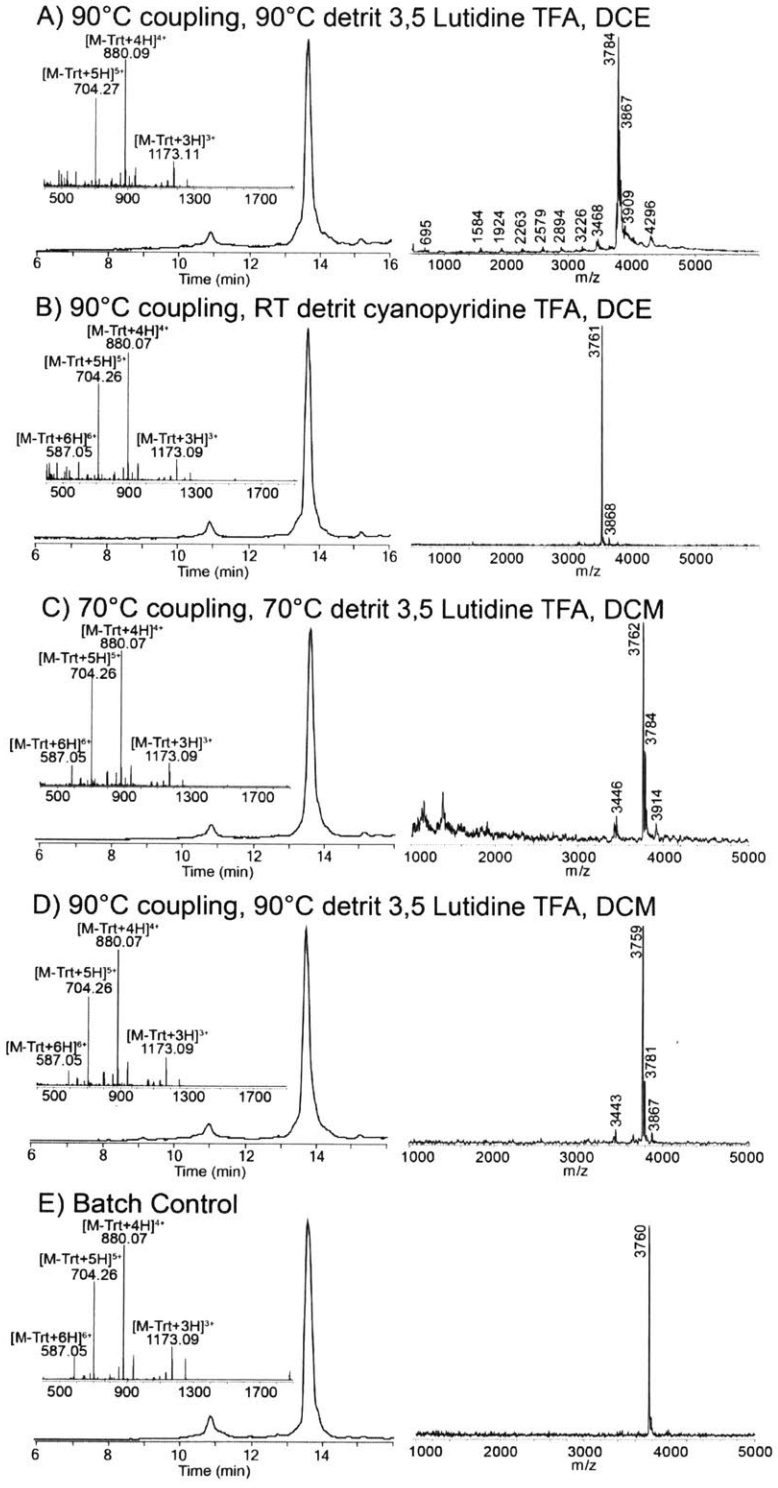


Figure 4.2.6 Synthesis of Eteplirsen(1-10) under several conditions. In each panel, the total ion chromatogram is shown on the left with the mass spectrum integrated from 6 to 16 minutes inset. On the right, MALDI-TOF mass spectra are shown.

4.3 Conclusions

Fast flow synthesis of PMOs has been demonstrated. With approval of a first in class PMO for the treatment of muscular dystrophy, the therapeutic potential of PMOs has been validated. As PMO therapeutics are expanded to cover more muscular dystrophy genotypes, and other orphan diseases, the demand for rapid manufacturing methods to treat small cohorts will intensify. In this work, we have accelerated PMO synthesis by applying the lessons learned in fast flow peptide synthesis (Chapters 2 and 3).

First, high boiling solvents were sought for PMO synthesis to improve process safety. By operating below or near the atmospheric pressure boiling point of the solvents used, flow synthesis was performed at lower pressure and less mechanical strain was placed on components. Ultimately, we reverted to low boiling dichloromethane, however, for its better chemical compatibility with the resin bound PMO. Operating at the required pressure did not cause any mechanical failures during these experiments.

Second, the stability of the reagents used was investigated. It was determined that the activated nucleotides were unstable above 100°C, rapidly forming the unreactive phosphoroanhydride and several unknown products. Based on this result, we established 90°C as a conservative upper limit for PMO synthesis. Similarly, the morpholinyl resin was found to be unstable in DIEA, DCE, and detritylation solution at 90°C, and these reagents were avoided in later syntheses.

Third, the reactivity of the pyridinium salt used for 3' detritylation was reduced to protect the PMO backbone from acidolysis, while maintaining rapid detritylation at 90°C. Similarly, the reactivity of the activated nucleotide was enhanced by using lithium bromide as a Lewis acid catalyst.

Finally, these insights were used to implement a fast flow PMO synthesis using repurposed fast flow peptide synthesis hardware[34]. This hardware enabled rapid preheating of reagents and high heat and mass transfer within the resin bed on a scalable, mechanically robust platform. Reusing reliable rapid peptide synthesis hardware allowed up to focus on the unique chemical problems of accelerated PMO synthesis and ultimately accelerate synthesis by about an order of magnitude, from two hours per nucleotide to about 13 minutes.

In conclusion, we have demonstrated the challenges and opportunity to use fast flow technology to accelerate the synthesis of PMOs. We find unique value in this methodological approach in accelerating solid phase synthesis, and we believe the lessons learned here will prove useful in accelerating other biopolymer synthesis chemistries, such as the recently reported phosphoramidate based PMO synthesis[91].

D-amino acid scan of two small proteins

The work presented in this chapter was published in the following manuscript, and Appendix 4 is derived from the supporting information

Mark D. Simon, Yuta Maki, Alexander A. Vinogradov, Chi Zhang, Hongtao Yu, Yu-Shan Lin, Yasuhiro Kajihara, Bradley L. Pentelute. D-Amino Acid San of Two Small Proteins. *J. Am. Chem. Soc.*, **2016**, 138 (37), pp 12099–12111 DOI: 10.1021/jacs.6b03765

5.1 Introduction

Nature is chiral. Biopolymers from DNA to lipids are handed, as are many small molecule natural products. But while Nature uses multiple isomers of some products, such as carbohydrates and steroids, natural proteins are homochiral polymers, composed exclusively of L-amino acids[92–94]. There are post-translational amino acid isomerases and non-ribosomal synthases that produce short peptides with D-amino acids, but such enzymes are uncommon[95–97]. Why doesn't nature make extensive use of genetically encoded D-amino acids or post-translational epimerases? In principle, using D-amino acids in a protein would allow access to a broader folding and structural space[98]. Research has shown rational substitution of D-amino acids for glycine in a favorable confirmation can improve thermal stability[98,99], and targeted insertions elsewhere can improve stability toward proteases[100]. Further, D scans of several bioactive peptides have shown them to be tolerant of D-amino acid substitutions, with some constructs having enhanced biological properties[101–105].

To investigate the effect of introducing D-amino acids into existing proteins, we selected two proteins, *Ecballium elaterium* trypsin inhibitor II (EETI-II) and a minimized Z domain of protein A (Z33)[106], and prepared two series of diastereomers by inverting a single amino acid in each analog (Fig 5.1.1). For achiral glycine residues, we substituted D-alanine, as it is hypothesized that glycine is Nature's placeholder for D amino acids[98,99].Like alanine scans used to determine which side chains are important for function and folding, this "D scan" identified the important chiral centers for folding and activity of EETI-II and Z33. Strikingly, most variants folded, and about a third retained binding affinity within an order of magnitude of the wild type. EETI-II is a three-disulfide cysteine knot, while Z33 is composed of two antiparallel 3.6_{13}-helices bridged by a short loop. These two proteins were selected for several reasons.

First, both EETI-II and Z33 are highly representative of small protein scaffolds that are of interest in developing novel protein functionality. EETI-II-like cysteine knots have high thermal and proteolytic stability, tolerance to mutation, oral bioavailability, and low immunogenicity, making them attractive targets for therapeutic protein engineering. Several successful engineering efforts have yielded constructs that are FDA approved and used clinically[107–115]. Minimal helical scaffolds such as Z33 have also been extensively used to engineer novel biological functions[116–119]. Since α-helix interfaces are commonplace in protein-protein binding interactions, such scaffolds are routinely evolved to bind proteins, and thus to elicit biologically relevant functions[120–122].

Second, they are chemically accessible. Although there exist methods to genetically encode non-proteogenic amino acids, chemical synthesis offers a more general and robust route to proteins incorporating many different D amino acids[123,124]. Both EETI-II and Z33 can be assembled in good yield

by fast flow stepwise solid phase peptide synthesis without the need for fragment condensation or native chemical ligation[125]. This vastly simplified synthesis and improved yields, an important consideration when undertaking the total synthesis of over sixty proteins.

Third, the folding of EETI-II and Z33 can be readily monitored. The folding reaction of EETI-II can be monitored via LC/MS. Native EETI-II folds by forming three disulfide bonds[126,127]. The first and second form rapidly, giving a stable, early eluting two disulfide intermediate. The third disulfide then forms, locking the protein into its final knotted conformation, and yielding a chromatograph-ically resolved, early eluting final product[24]. The nature of the intermediate and final products can be confirmed by the loss of 2Da per disulfide bond. Similarly, WT Z33 forms two helices, which yield a canonically helical far-UV CD spectrum[106]. Perturbations to the helical structure are observed as a loss in helicity, and an increase in random coil content.

Fourth, there is a high-resolution crystal structure of EETI-II bound to trypsin[128], a high resolution crystal structure of the Z-domain of protein A bound to a human IgG[129], and an NMR structure of Z33 with five additional N-terminal residues[130]. Observed changes in the folding, structure, and function of EETI-II and Z33 can be correlated to the structure of the native proteins and their interaction with their binding partners, enabling a richer analysis of our findings. Furthermore, structural data supported molecular dynamics (MD) simulations to provide additional insight into the effect of D-amino acid substitutions.

Finally, there are facile assays available to calculate the binding affinity of EETI-II and Z33 to their corresponding targets. A chromogenic assay can be used to calculate the binding affinity of non-covalent trypsin inhibitors[131,132]. In this assay, trypsin hydrolyzes Nα-benzoyl-Arg-(4-nitro)anilide, releasing the intensely yellow nitro aniline. The rate of increase in absorbance is proportional to the trypsin activity, and the binding affinity is determined by calculating trypsin activity in the presence of various concentrations of EETI-II[132]. As a minimized Z-domain of protein A, Z33 binds to hIgG with a nanomolar affinity, and this interaction can be monitored using Bio-layer Interferometry methods.

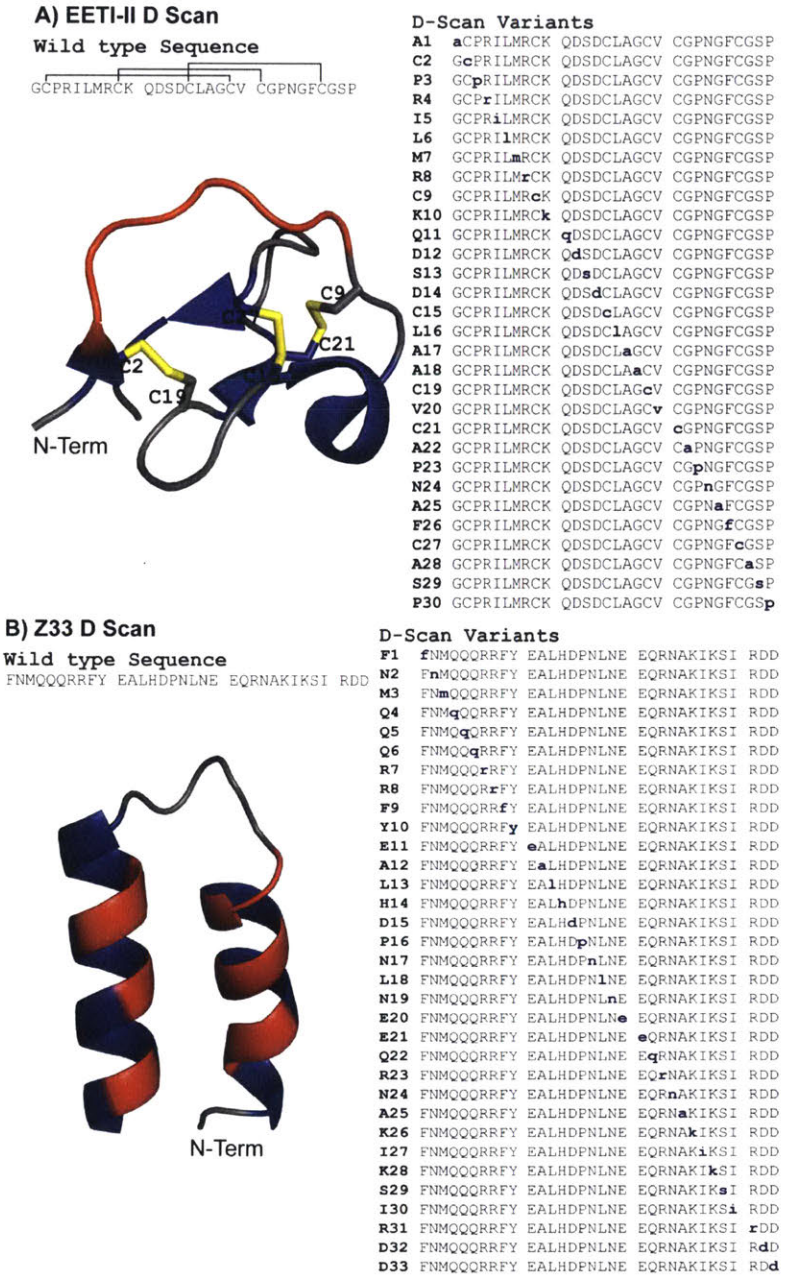


Figure 5.1.1 Illustration of the D-Scan of EETI-II and Z33. Amino acids represented by black capital letters are L-amino acids, and amino acids represented by blue lower case letters are D-amino acids. The name of each analog is the substituted D-amino acid and its position. (**A**) D-scan of EETI-II. In the wild type sequence of EETI-II, disulfide connectivity is shown with black lines. In the cartoon representation of EETI-II, constructed in PyMOL from crystallographic data reported by Kraetzner et al[128], the disulfide bridges are yellow, the trypsin binding loop is red, secondary structure annotated in the PDB is blue, and the remaining regions are grey. (**B**) D-scan of Z33. The cartoon representation of Z33 was constructed from NMR structural data reported by Starovasnik et al[133], amino acid residues at the binding interface with IgG are red, secondary structure annotated in the PDB is blue, and the remaining regions are grey. Binding residues were determined from the co-crystal structure of intact Z-domain and an IgG fragment[129].

5.2 D-scan of EETI-II

5.2.1 Synthesis and folding of EETI-II analogs

Linear peptides were assembled using an improved and optimized Fast Flow peptide synthesis protocol (0.1mmol scale)[34,68], affording about a hundred milligrams of each crude analog. Stepwise solid phase peptide synthesis was performed on ChemMatrix Rink amide resin at 70°C using HATU activation and N-terminal Fmoc protection. Amide bond formation was allowed to proceed for 30 seconds, and Fmoc groups were removed in 20 seconds with 20% (v/v) piperidine in DMF. The total time required to incorporate an amino acid was 2 minutes, enabling synthesis of five analogs per day. Representative LC/MS analyses of crude peptides are shown in Appendix 4.

We next investigated folding conditions for wild type EETI-II. The reduced polypeptide required 10 minutes to fold in a glutathione based redox buffer[132], and was insensitive to the concentrations of oxidized and reduced glutathione in the range we explored. In cysteine/cystine redox buffer or aerated phosphate buffer, WT EETI-II folded in several hours. In an effort to drive analogs towards properly folded thermodynamic products, the most reducing glutathione based folding buffer we investigated was selected (2mM GSSH, 10mM GSH, 25mM Tris, pH=7.8).

After establishing folding conditions for wild type EETI-II, we undertook analytical folding and trypsin inhibition experiments with the analogs. First, we purified several milligrams of each linear peptide via RP-HPLC (Fig 5.2.1). The purified linear peptides were then folded by dissolution to 3.3mM in dissolving buffer (6M guanidinium chloride, 0.2M phosphate, pH=6.9) and diluted to 0.33mM with folding buffer. Folding reactions were monitored by LC/MS, with time points at two minutes, one hour, and three hours (Fig 5.2.2).

Under these folding conditions, 21 out of 30 analogs formed products that appeared as a single, well-defined, early eluting peak 6 Da less massive than the starting material during LC/MS analysis. This material was deemed "folded," although it is possible that in some cases there was a mixture of co-eluting three disulfide products. If folding was disrupted, however, an inhomogeneous mixture of products was seen. Folded three disulfide products were isolated by RP-HPLC to afford pure material for the trypsin inhibition assay (Fig 5.2.2).

The nine analogs that failed to fold in the glutathione redox buffer were subjected to cysteine/cystine folding buffer, and oxidation with atmospheric oxygen in phosphate buffer (100mM phosphate, pH=8). None folded well in the cysteine based buffer, but three gave a single, early eluting three disulfide product upon air oxidation. These folded products were then purified by RP-HPLC for the trypsin inhibition assay (Fig A4.4).

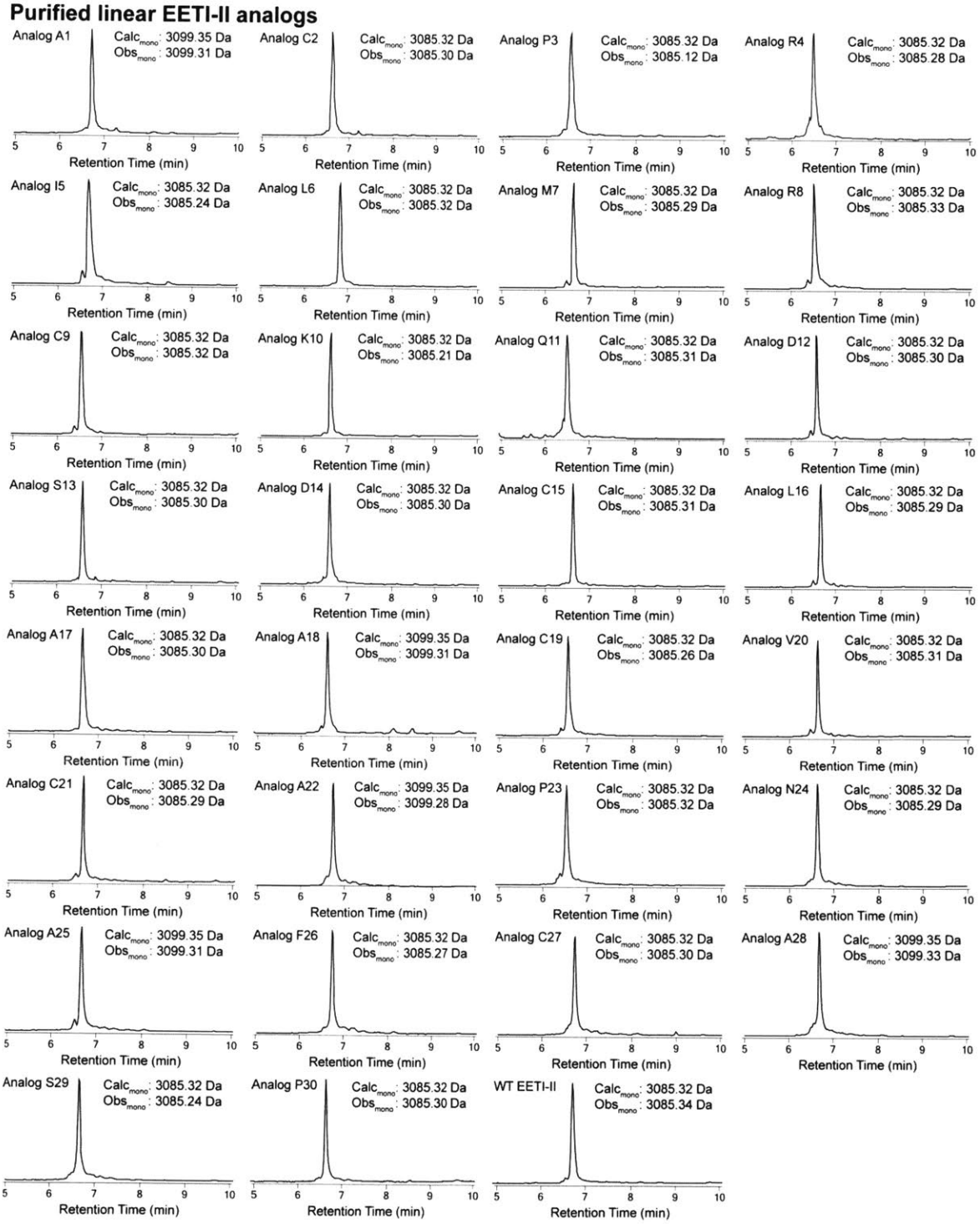


Figure 5.2.1 Purified linear EETI-II analogs. Each panel shows LC/MS analysis of pooled fractions from the RP-HPLC purifications of a linear, unfolded analog of EETI-II. In all cases, calculated and observed masses are monoisotopic, and total ion chromatograms are shown. See Appendix 4 for chromatographic conditions.

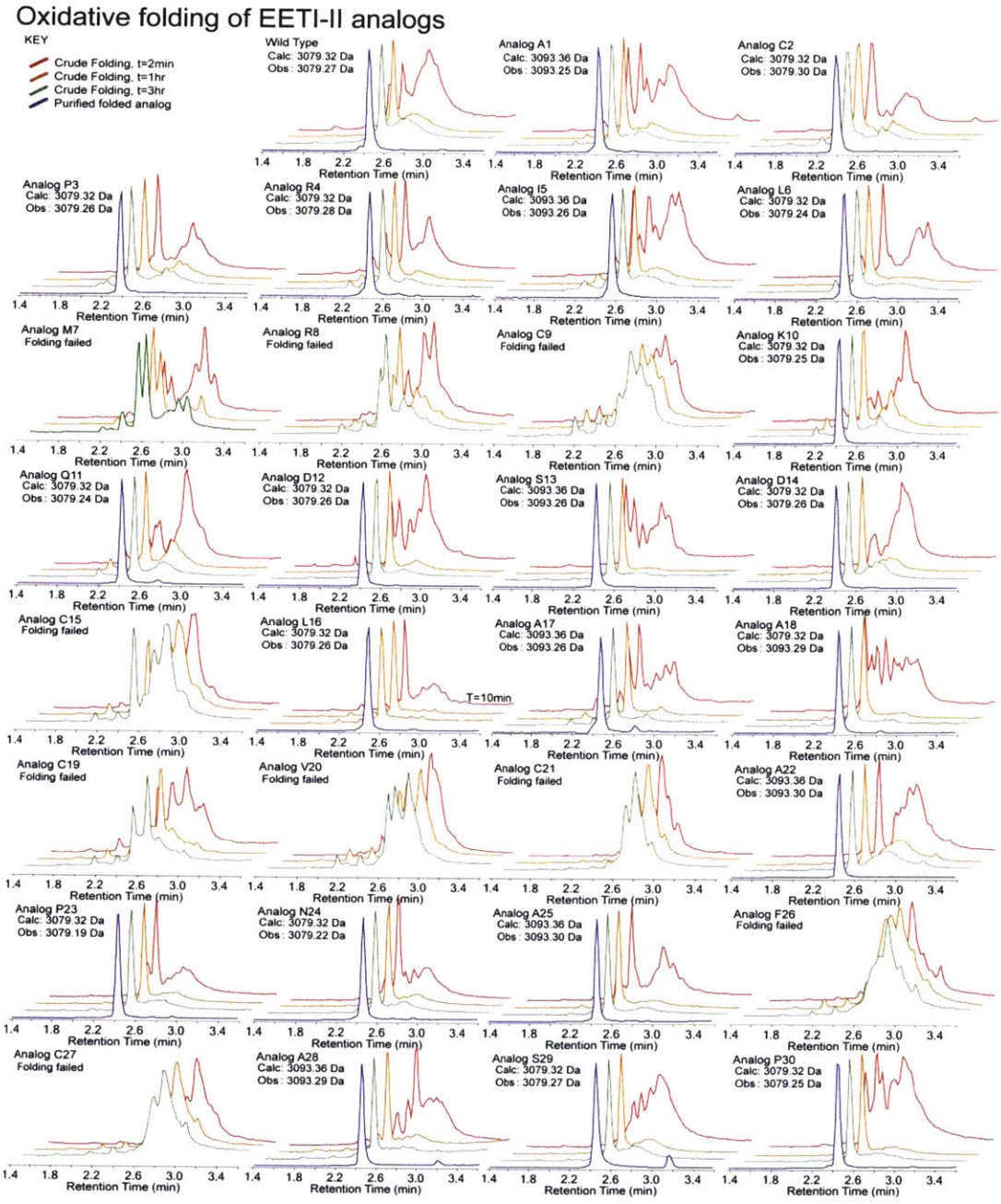


Figure 5.2.2 Time coursed folding reaction for each EETI-II analog. Each EETI-II analog was dissolved in a glutathione based redox buffer (2mM GSSH, 10mM GSH, pH=7.8) and the folding reaction was monitored by LC/MS, with aliquots quenched and analyzed after two minutes, one hour, and three hours. The formation of a product six Daltons less massive than the starting material that elutes as a single peak significantly earlier the starting material is characteristic of a successful folding reaction. Unsuccessful folding reactions generally produce a heterogeneous mixture of products that elute as a broad peak somewhere between the desired, folded product and the starting material. The desired product was isolated form all successful folding reactions by preparative RP-HPLC and the LC/MS analysis of pooled fractions is shown in blue, where applicable. In all cases, calculated and observed masses are monoisotopic, and total ion chromatograms are shown. See Appendix 4 for chromatographic conditions.

5.2.2 Trypsin inhibition activity of EETI-II analogs

With 24 folded, purified EETI-II analogs in hand, we assayed their activity against a commercial sample of bovine pancreatic trypsin[131,132]. EETI-II was diluted with 2mM Nα-benzoyl-Arg-(4-nitro)anilide in trypsin buffer (60mM Tris, 20mM $CaCl_2$, pH=7.8) and 80nM trypsin in 1mM HCl was added to give a final trypsin concentration of 4nM. Final EETI-II concentrations of 100nM, 32nM, 16nM, 8nM, 4nM, 2nM, 1nM, 500pM, 250pM, and 0pM were used to obtain measurements to accurately calculate a broad range of binding affinities. The absorbance at 410nm was monitored to track the formation of the yellow nitro aniline liberated by catalytic hydrolysis of Nα-benzoyl-Arg-(4-nitro)anilide. The rate of change in absorbance became linear in less than 30 minutes, and the reaction was monitored for a total of 3 hours. Long monitoring times were necessary because a low concentration of trypsin was used to improve the sensitivity of the assay. We found binding affinities more than fivefold lower than the trypsin concentration cannot be quantified. The rate of change in absorbance was taken as a direct measure of residual trypsin activity and used to calculate the binding affinity[132]. For each analog, the assay was performed in quadruplicate in a 96 well plate (2 analogs per plate). Representative data for four analogs of varying activity are shown in Figure 5.2.3, and the complete results are tabulated in Table 5.2.1. Of 24 analogs assayed, 12 were high affinity trypsin inhibitors (K_D < 15nM), and three were as active as WT EETI-II.

5.2.3 Thermodynamic stability of EETI-II analogs

To investigate the thermodynamic stability of the 12 active EETI-II analogs, the synthesis was scaled up to 0.2mmol and each of the active variants was folded on a larger scale. The crude linear peptide was dissolved to a concentration of 3.3mM in dissolving buffer, then slowly diluted with folding buffer to a final concentration of 0.29mM. After one hour, each folded analog was isolated by preparative RP-HPLC, affording 10-20% of highly pure, folded material. This procedure was found to be higher yielding than first purifying the linear peptide, folding, and then repurifying the final product.

We then attempted to establish a melting temperature for EETI-II. The far UV circular dichroism spectrum of wild type EETI-II showed no change with temperature or urea concentration, and thermally induced conformational changes were not detected by differential scanning calorimetry. These results are not surprising because EETI-II is stabilized primarily by three disulfide bonds, with little canonical secondary structure, and similar results have been reported for another cysteine knot protein[134]. Thus, to determine the relative stability of each active EETI-II analog, we investigated the reduction potential of the disulfide bonds using a procedure adapted from Nolan and co-workers[135].

To evaluate the reduction potential, each active variant was incubated in a series of dithiothreitol based redox buffers with potentials ranging from -240mV to -390mV. After incubation at 4°C for two weeks, aliquots of each sample were quenched with acid and analyzed by RP-HPLC. For each analog, it was observed that one disulfide bond was readily reduced, and then the second two were reduced together. The one disulfide intermediate was generally not prevalent and poorly resolved from the fully reduced product. Furthermore, we observed that the one disulfide intermediate can form from the linear peptide even in acidic media, so observed one disulfide products were integrated with the fully reduced product.

Separate midpoint potentials were determined for the two disulfide core, and the third disulfide. Eight of thirteen samples showed Nernstian behavior, while the remaining five (WT, A1, A17, N24, and A25) had not yet reached equilibrium under these conditions. After incubation at 4°C for an additional 4 weeks, these five analogs also exhibited Nernstian behavior. Raw chromatograms for WT EETI-II and analog D12 are shown in Figure 5.2.4, and fitted data for four analogs of varying stability is shown in Figure 5.2.5. The remaining data is in Appendix 4. Complete results are summarized in Table 5.2.2.

Most of the analogs investigated were substantially destabilized. The third disulfide bond in A17, for example, was destabilized by 25mV (1.2kcal/mol), and the two disulfide cores of K10, Q11, and D14 were destabilized by almost 50mV (4.6kcal/mol) relative to wild type EETI-II. As can be seen in Table 5.2.2, the only three active analogs with both a two disulfide core and the third disulfide comparably stable to WT EETI-II were A1, N24 and A25. Despite large differences in the stability of the analogs, however, we did not observe a clear relationship between stability and trypsin binding affinity.

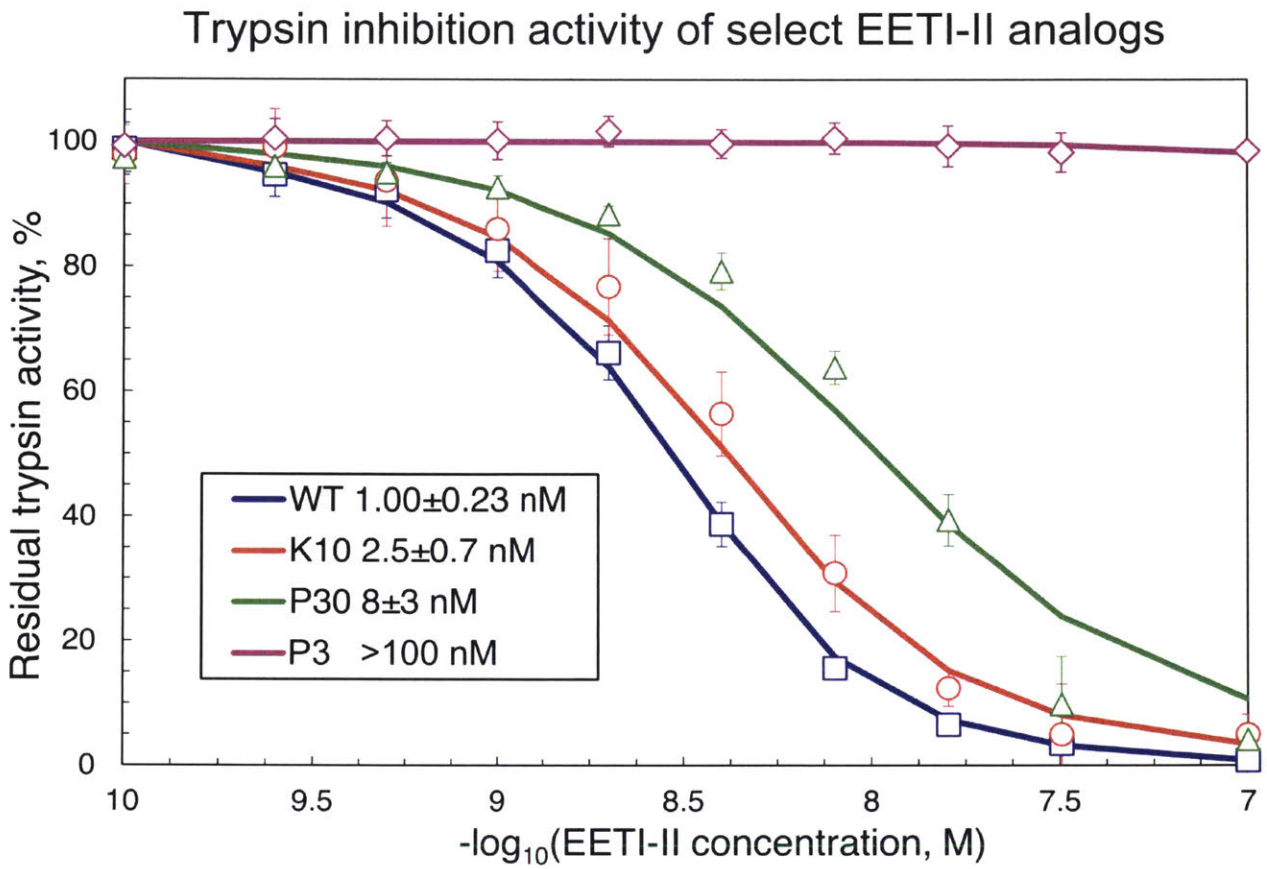


Figure 5.2.3 Representative trypsin inhibition assay results. Binding affinity was determined by incubating trypsin with various concentrations of each analog and measuring the residual catalytic hydrolysis of Nα-benzoyl-Arg-(4-nitro)anilide, a chromogenic trypsin substrate, via the absorbance at 410nm. This implementation of the assay is sensitive enough to differentiate a 2.5nM binder (K10) from a 1nM binder (WT). Errors are given as 95% confidence intervals. See Appendix 4 for details of the assay and fitting procedure.

Table 5.2.1 Trypsin affinity of EETI-II analogs

Analog	Trypsin Binding (nM)[a]	Analog	Trypsin Binding (nM)[a]
WT	1.00 ±0.23	L16	>100
A1	2.01±0.27	A17	3.2±0.8
C2	>100	A18	>100
P3	>100	C19	N/A
R4	>100	V20	N/A
I5	>100	C21	N/A
L6	>100	A22	>100
M7	>100	P23	0.92±0.15
R8	>100	N24	0.92±0.16
C9	N/A	A25	0.89±0.13
K10	2.5±0.7	F26	N/A
Q11	13±5	C27	N/A
D12	3.0±2.0	A28	72±28
S13	3.1±1.0	S29	7.3±1.8
D14	11±4	P30	8±3
C15	>100		

[a] The trypsin binding affinity of each analog calculated via the chromogenic trypsin inhibition assay is tabulated, with errors reported as 95% confidence intervals. Trypsin binding affinities >100nM could not be detected, and binding affinities below 0.8nM could not be accurately quantified. N/A indicates an analog that failed to fold.

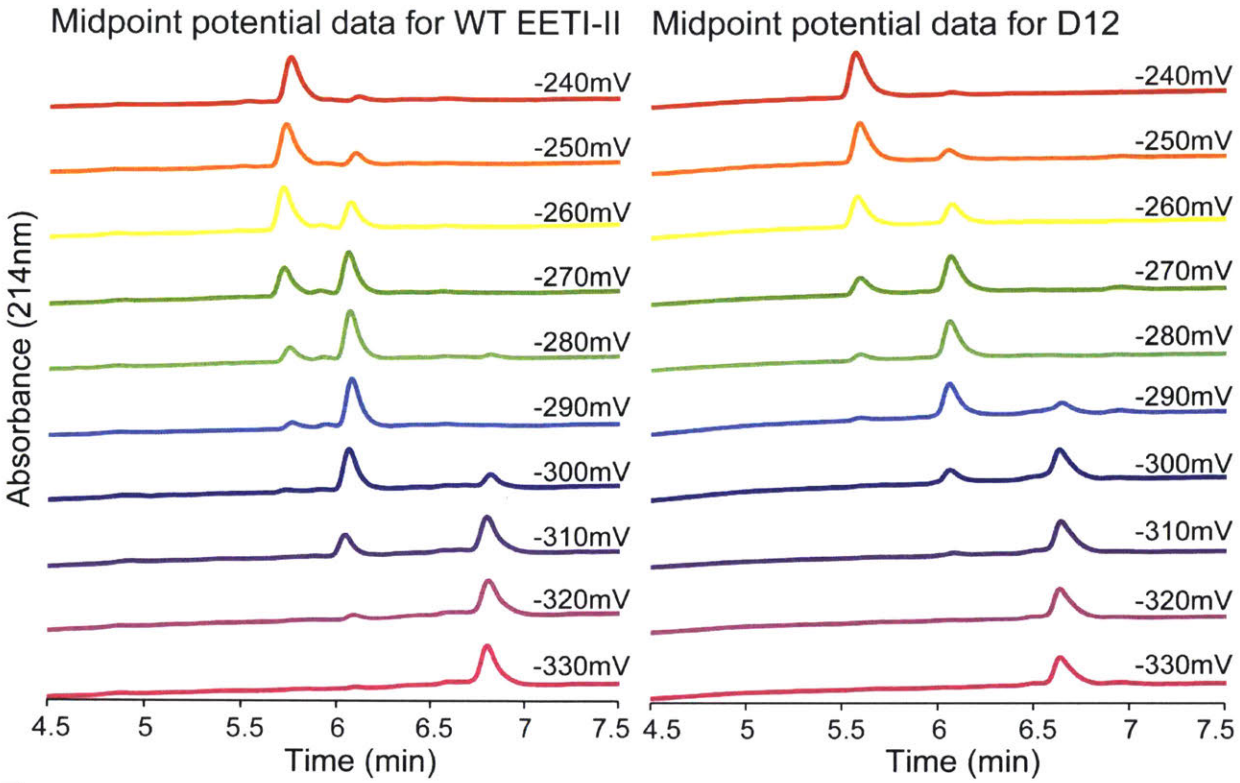


Figure 5.2.4 Raw reduction potential data from WT EETI-II and analog D12. Raw chromatograms of WT EETI-II and analog D12 after incubation in redox buffers of various potentials are shown. The topmost (red) trace is after incubation in a-240mV buffer and shows that the product is almost completely folded in both cases. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom (purple) trace showing complete reduction to the linear peptide after incubation in a -330mV buffer. Upon inspection of the -300mV and -310mV results, D12 is seen to be less stable than WT, even before quantitative data processing.

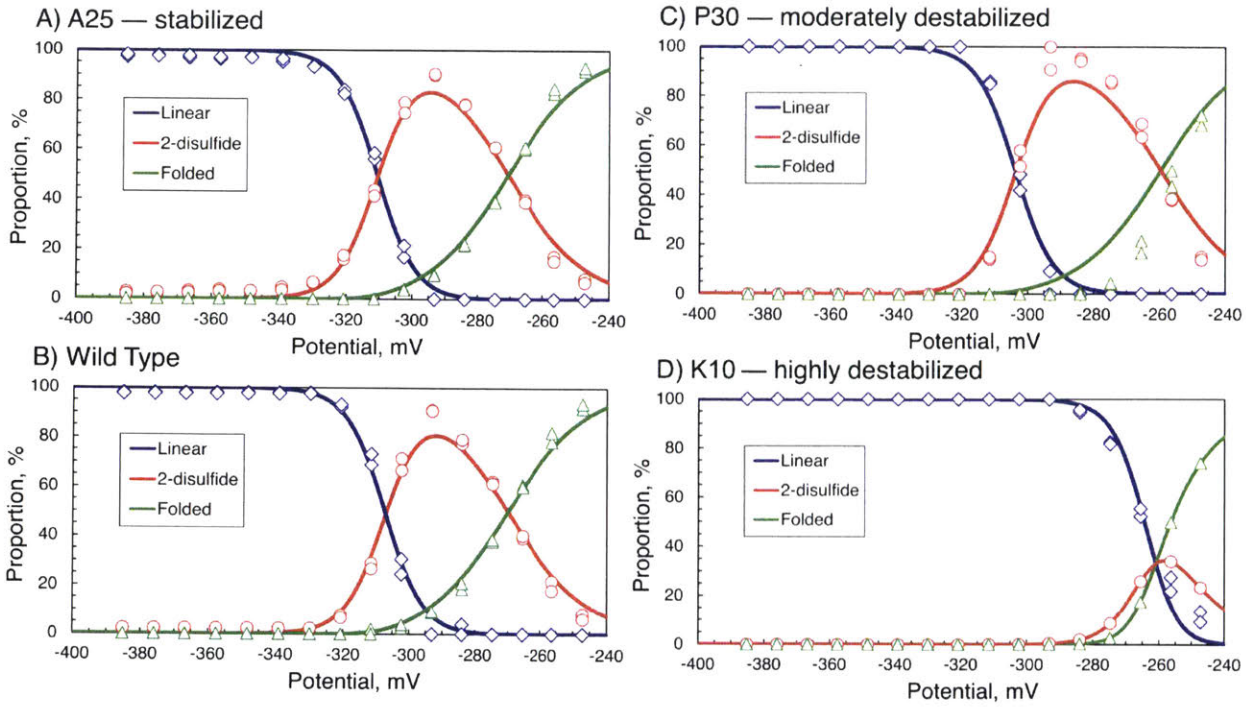


Figure 5.2.5 Representative thermodynamic data from several EETI-II analogs. The stability of each active analog of EETI-II was determined by incubating the analog in a series of redox buffers with potentials from -390mV to -250mV until equilibrium was achieved. Reduction potentials were calculated by fitting the equilibrium proportions of each species to the Nernst equation. In all cases, it was observed that one disulfide bond was reduced first, and then two more were reduced together. **A)** through **D)** show the equilibrium proportion of the linear peptide, the two disulfide intermediate, and the folded protein at various potentials for A25, WT, P30, and K10, respectively. **A)** A25 is slightly more stable than WT **B)** WT **C)** P30 is less stable than WT **D)** K10 is significantly less stable than WT.

Table 5.2.2 **Reduction potential of EETI-II analogs**

EETI-II Analog	Two disulfide core reduction potential (mV)[a]	Third disulfide reduction potential (mV)[a]
WT	-307.2±1.1	-270.0±1.5
A1	-304.3±0.4	-266.3±0.9
K10	-260.9±1.5	-260.9±1.5
Q11	-264.1±1.4	-267.1±1.3
D12	-296.5±1.4	-269.2±1.8
S13	-293.7±1.9	-273.8±2.0
D14	-261.7±2.5	-275.4±1.6
A17	-305.6±0.6	-244.7±1.2
P23	-293.2±1.6	-270.0±2.0
N24	-304.1±0.8	-266.5±1.1
A25	-310.5±0.9	-270.7±1.3
S29	-305.1±1.4	-252.9±2.2
P30	-303.7±1.7	-259.7±2.5

[a] Reduction potentials of the two disulfide core and third disulfide of each active EETI-II analog were calculated by incubation in redox buffers of varying potential, and fitting the equilibrium proportions of folded protein, two disulfide intermediate, and linear peptide to the Nernst equation. Results are tabulated here, and reported as 95% confidence intervals.

5.2.4 Molecular dynamics simulations of select EETI-II analogs

Molecular dynamics (MD) simulations of WT EETI-II, and analgos A18, P23, N24, and A25 were performed to provide further insight into the effects of D-amino acid substitutions at these positions (See Appendix 4 for detailed methods)[136–139]. It was observed that in three out the four MD runs performed on A18, the G18a substitution resulted in significant RMSD from the PDB structure (run 1 shown in Fig 5.2.6 and all runs in Fig A4.5). In wild type EETI-II, a small beta sheet is formed by residues 20–22 and residues 26–28 (Fig 5.2.6A and 5.2.6C, left column), and the C-terminus of the peptide (residues 29 and 30) had close contacts with residues 18–19 (Fig 5.2.6D, left column). In the A18 analog, the G18a mutation pushed away residues 29 and 30 toward residues 3-8 in the N-terminus (Fig 5.2.6D, right column). In runs 1 and 2 of the A18 (G18a) simulations, the original beta sheet interactions between residues 20–22 and 26–28 were significantly perturbed, and new beta sheet interactions formed between residues 26–27 and 7–8 (Fig 5.2.6 and A4.5). On the other hand, minimal structural perturbation was observed in A25 (Fig A4.55). The results of the P23 and N24 analogs (three runs for each) are shown in Figures 5.2.7, A4.6, and A4.7. Neither mutation seemed to disturb the structures significantly. We found that in the WT, residues 23–24 formed a type I beta turn. When P23 was mutated to p23, residues 23–24 instead adopted a type II' beta turn; when N24 was mutated to n24, residues 23–24 formed a type II beta turn.

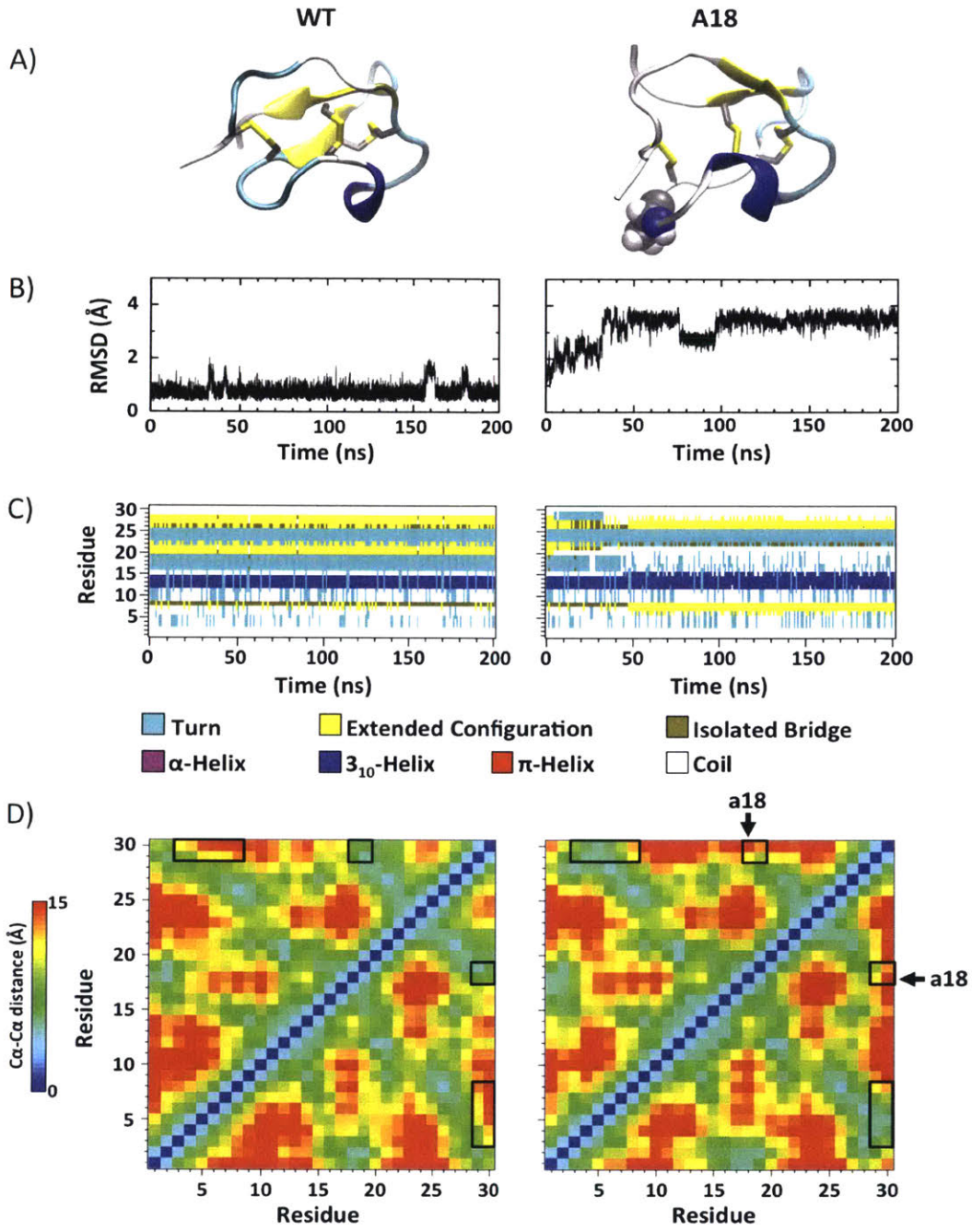


Figure 5.2.6 Molecular dynamic simulations of WT EETI-II, and analog A18. A) Cartoon representations of the peptide structures after 200ns MD simulations. The structures are colored according to the secondary structure calculated using STRIDE (turn: cyan; extended conformation: yellow; isolated bridge: brown; α-helix: magenta; 3_{10}-helix: blue; π-helix: red; coil: white). The disulfide bonds are represented by sticks and the D-amino acids are highlighted with spheres. **B)** Peptide backbone RMSD trajectories with respect to the PDB structure (PDB ID=1H9H). **C)** Secondary structure evolutions of each residue during the 200 ns simulations. The same color code is used as in **A)**. **D)** Cα-Cα distance map calculated using 100–200ns of each trajectory. Arrows denote the locations of the D-amino acids.

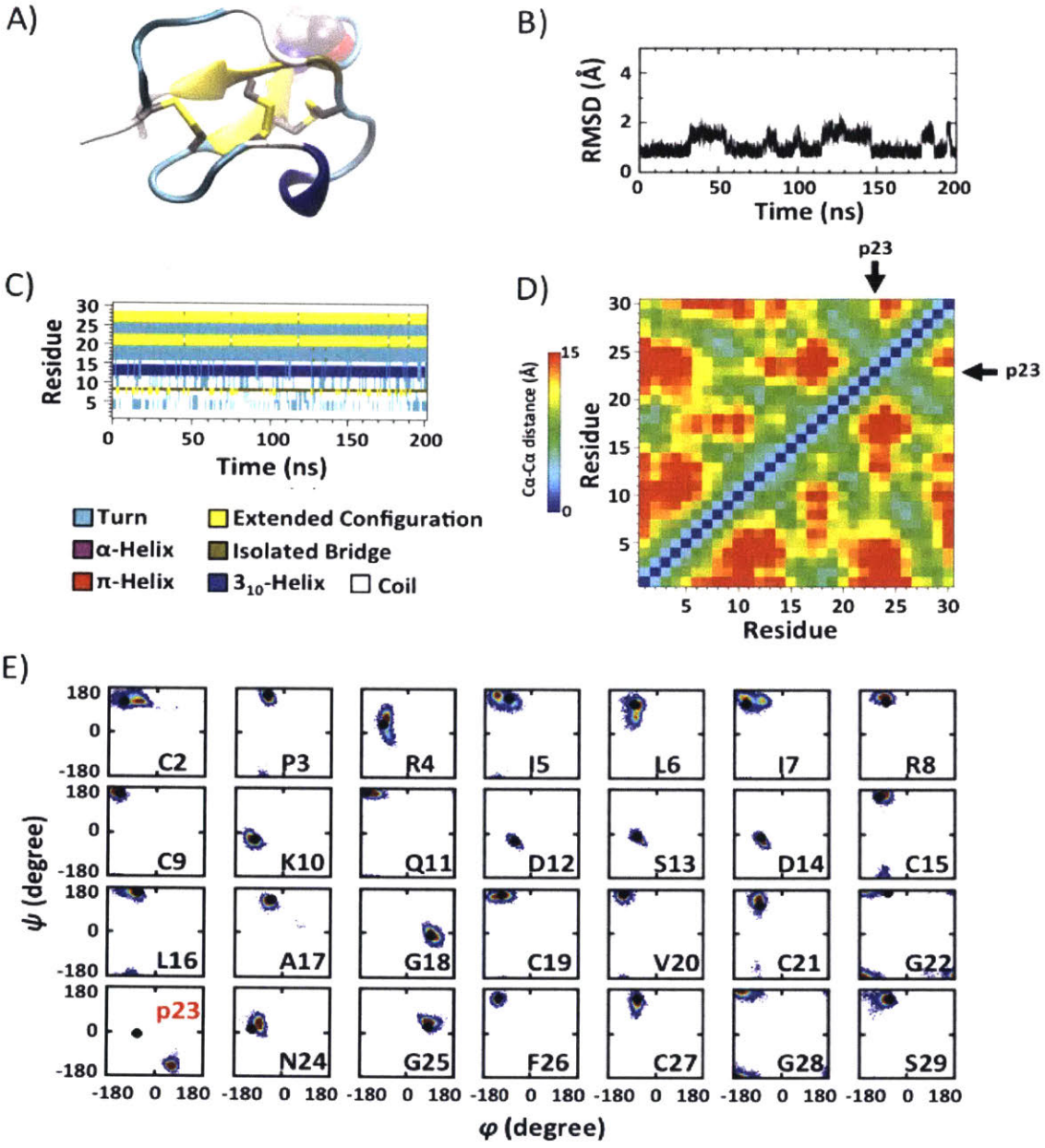


Figure 5.2.7 Molecular dynamic simulations of P23. A) Cartoon representation of the P23 analog structure after 200ns MD simulations. The structure is colored according to the secondary structure calculated using STRIDE (turn: cyan; extended conformation: yellow; isolated bridge: brown; α-helix: magenta; 3_{10}-helix: blue; π-helix: red; coil: white). The disulfide bonds are represented by sticks and the D-amino acids are highlighted with spheres. **B)** Peptide backbone RMSD trajectory with respect to the PDB structure. **C)** Secondary structure evolution of each residue during the 200 ns simulation. The same color code is used as in **A)**. **D)** Cα-Cα contact mas calculated using 100–200ns of the trajectory. Arrows denote the location of the D-amino acid. **E)** Ramachandran plot for each residue calculated using 100–200ns of the trajectory. The D-amino acid is labeled with a red lower-case letter. φ/ψ angles of the WT are shown as black dots.

5.3 D-scan of Z33

5.3.1 Synthesis and purification of Z33 analogs

Peptides were assembled by stepwise solid phase peptide synthesis on 0.1mmol ChemMatrix rink amide resin using a further improved Fast Flow peptide synthesis protocol (0.1mmol scale)[34,68], affording about 150mg of each crude analog. Synthesis was performed at 90°C using HATU activation and N-terminal Fmoc protection. Amide bond formation was allowed to proceed for 8 seconds, and Fmoc groups were removed in 8 seconds with 20% (v/v) piperidine in DMF. The total time required to incorporate an amino acid was 0.7 minutes, enabling synthesis of twelve analogs per day. WT Z33 and all 33 analogs were purified by RP-HPLC to yield highly pure material (Fig. 5.3.1)

5.3.2 Binding of Z33 analogs to IgG

The IgG binding affinity of the Z33 analogs was determined via Bio-layer Interferometry with biotinylated trastuzumab antibody immobilized on the surface of streptavidin biosensors. Sensograms for each Z33 analog were obtained at a series of concentrations from 156nM to 5,000nM at 25°C and fitted to obtain the binding constant (K_D) following the manufacturer's protocol (see Appendix 4).

As summarized in Table 5.3.1, the binding constant of every Z33 analog was greater than wild type Z33 (K_D = 24±1 nM). Ten analogs were still potent IgG binders (K_D < 250 nM), and twelve others retained some IgG binding (K_D between 250 nM and 2.5 μM). The rest of the Z33 analogs showed little to no binding to trastuzumab. Representative data for four analogs of varying binding affinity are in Figure 5.3.2, and the remaining sensograms are in Appendix 4.

5.3.3 Thermodynamic stability of Z33 analogs

To determine how the secondary structure of Z33 is affected by a single amino acid inversion, far-UV circular dichroism (CD) spectroscopy experiments were performed for each analog. First, 195-260 nm CD spectra were obtained. Purified peptides were dissolved in PBS containing 30% (v/v) glycerol to a final peptide concentration of 0.2 mg/mL, and spectra were acquired at −10°C. Glycerol prevents the buffer from freezing, and enables spectroscopy to be performed at low temperature to differentiate analogs of compromised stability. Spectra were then deconvoluted using the K2D3 algorithm to determine the α-helical content[140]. A D-amino acid in a secondary structure element may alter the spectral signature, but this was not taken into account and the α-helicity is likely slightly underestimated for all analogs. The theoretical helicity for the WT Z33 was calculated based on the solution NMR structure reported by Wells and co-workers[133].

As summarized in Table 5.3.2, we found that every analog was less helical than WT Z33. Five analogs (F1, M3, Q4, N17, and R31) retained more than 90% of the wild type's helical character, while 8 other analogs (R7, F9, L13, H14, P16, L18, R23, and K26) showed less than 35% α-helical character (less than half of the WT). The remaining 20 proteins, although significantly compromised, were still substantially helical. Selected spectra from analogs with varying helical character are shown in Figure 5.3.3a, and the complete data set is in Appendix 4.

Next, CD-aided urea denaturation of Z33 and its analogs was performed. The proteins were dissolved to 0.2 mg/mL in PBS containing 30% (v/v) glycerol and varying amounts of urea (0 to 7.5M), and the ellipticity was measured at 220 nm and −10°C. In 13 cases, unfolding curves appeared to extend to the regions of fully folded protein. In these cases, the resulting unfolding curves were processed assuming a two-state transition to obtain $\Delta G^u(H_2O, 263K)$ and $[urea]_{1/2}$, (see Appendix 4). Otherwise, thermodynamic parameters were not obtained; such analogs were significantly less stable than WT Z33. Figure 5.3.3b shows urea denaturations of several analogs of varying stability, and results are presented in Table 5.3.2. Complete data sets for all analogs are in Appendix 4.

In general, urea denaturation experiments agreed with the results of spectral deconvolution, with analogs that showed little helical character melting quickly and analogs with substantial helical character melting more slowly. Of 12 proteins with complete unfolding curves, only Q4 was as stable as WT, while the other 11 analogs were less stable, both in terms of $\Delta G^u(H_2O, 263K)$ and $[urea]_{1/2}$.

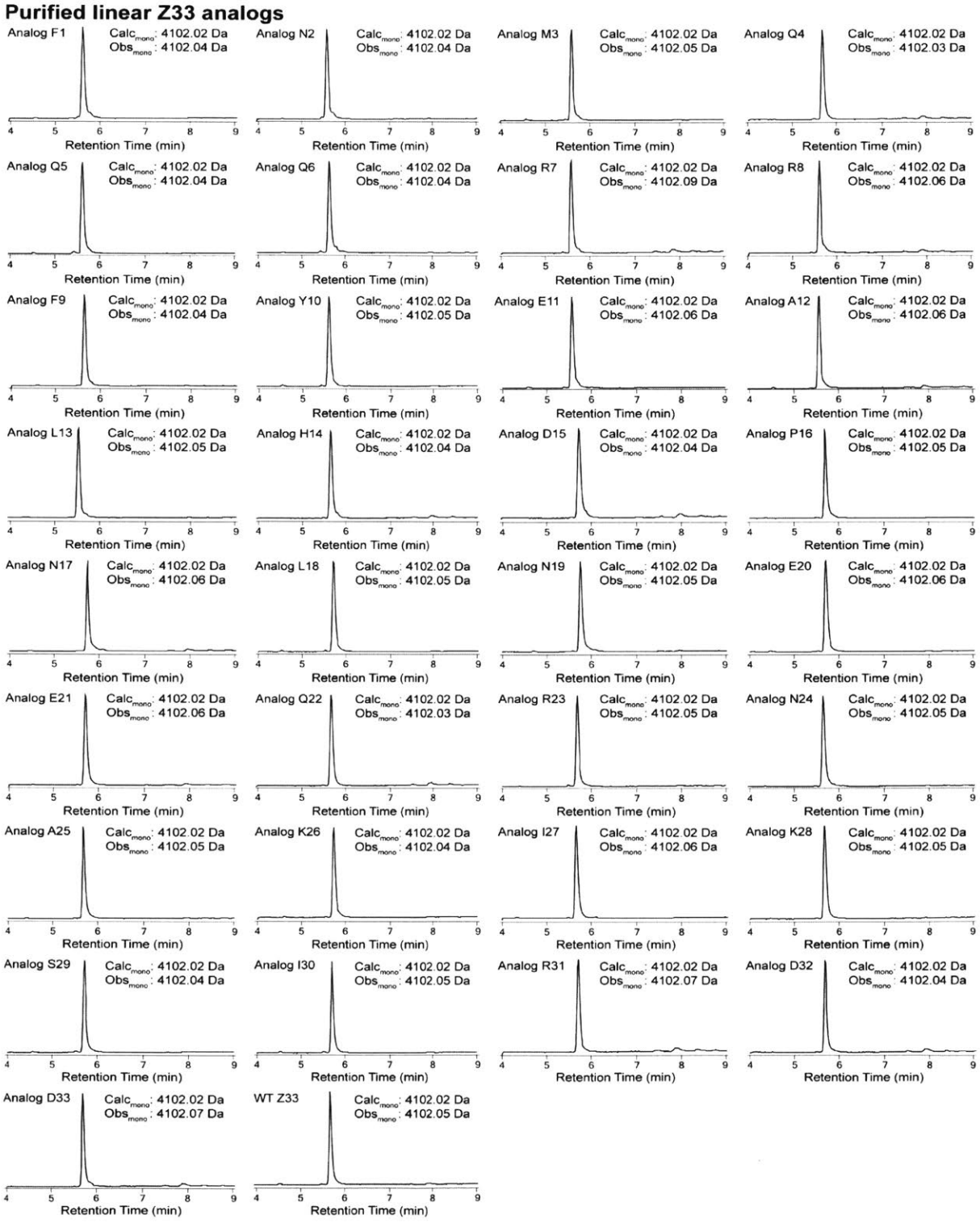


Figure 5.3.1 Purified linear Z33 analogs. Each panel shows LC/MS analysis of pooled fractions from the RP-HPLC purifications of a linear, unfolded analog of Z33. In all cases, calculated and observed masses are monoisotopic, and total ion chromatograms are shown. See Appendix 4 for chromatographic conditions.

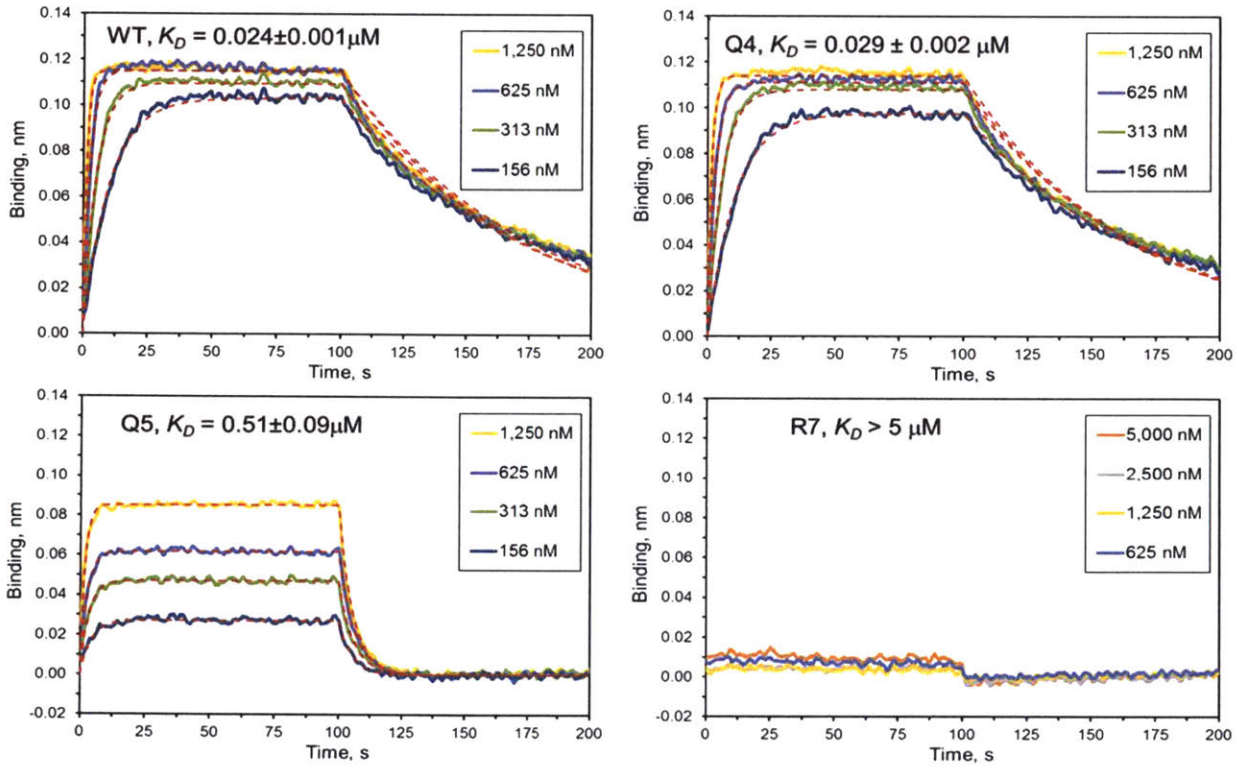


Figure 5.3.2 Representative sensograms from Z33 and several analogs. The IgG binding affinity of each analog of Z33 was determined using Bio-layer Interfereometry. Streptavidin biosensors loaded with biotinylated trastuzumab antibody were sampled against Z33 variants in solution. The concentrations of Z33 variants used are shown in the legend of each sensorgram. Wild Type Z33 had a K_D of 24±1nM, Q4 had a comparable K_D of 29±2 nM, Q5 had a decreased K_D of 0.51±0.09 μM, and R7 didn't associate with the biosensor even at a concentration of 5 μM.

Table 5.3.1 Binding constants of Z33 analogs

Z33 Analog	K_D, μM [a]	Z33 Analog	K_D, μM [a]
WT	0.024±0.001	N17	0.10±0.01
F1	0.125±0.005	L18	1.2±0.5
N2	0.28±0.03	N19	1.5±0.6
M3	0.088±0.003	E20	0.060±0.002
Q4	0.029±0.002	E21	0.083±0.005
Q5	0.51±0.09	Q22	0.19±0.5
Q6	4±2	R23	>5
R7	>5	N24	5±2
R8	0.25±0.04	A25	0.25±0.07
F9	>5	K26	1.2±0.2
Y10	>5	I27	4±1
E11	0.7±0.3	K28	1.2±0.2
A12	1.8±0.5	S29	0.056±0.001
L13	>5	I30	>5
H14	2.4±0.2	R31	2.2±0.8
D15	1.79±0.09	D32	0.036±0.002
P16	3±2	D33	0.030±0.002

[a] Standard errors are given

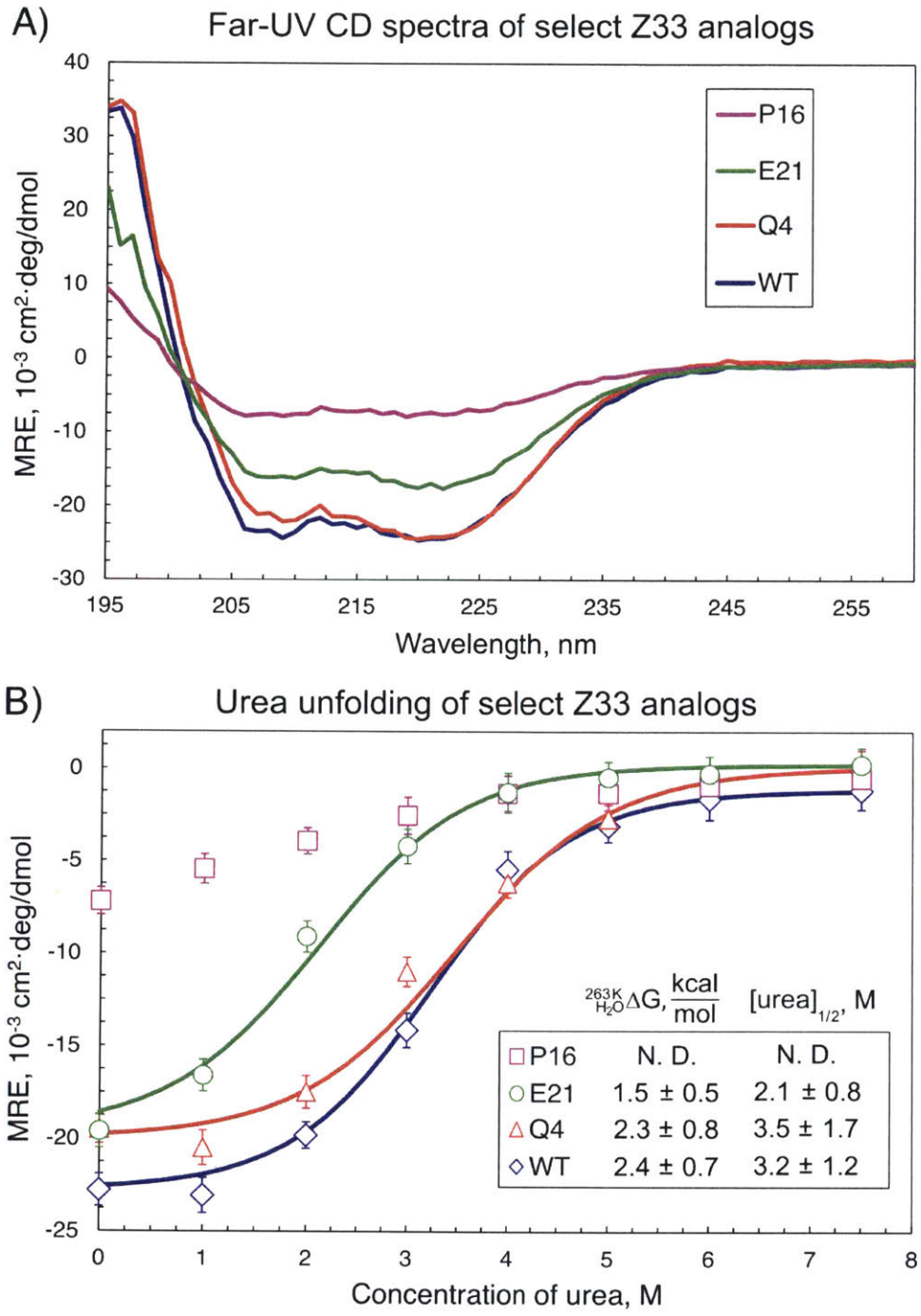


Figure 5.3.3 CD spectra and urea denaturations of select Z33 analogs. A) 195-260 nm CD spectra were obtained for 0.2 mg/mL protein solutions in PBS with 30% (v/v) glycerol at -10 °C. B) Mean molar ellipticity was monitored at 220 nm for 0.2 mg/mL protein solutions in PBS with 30% (v/v) glycerol and variable amounts of urea at -10 °C. Points on the plot are data points, lines are fit curves. The data were fit as described in detail in Appendix 4, and 95% confidence intervals for $\Delta G^u(H_2O, 263K)$ and $[urea]_{1/2}$ were calculated based on the fit parameters.

Table 5.3.2 Thermodynamic stability of Z33 analogs

Analog	$^{267K}\Delta G(H_2O)$, kcal/mol	$[Urea]_{1/2}$, M	Helicity, %
WT	2.4±0.7[a]	3.2±1.2	67
F1	1.7±0.6	2.8±1.3	62
N2	1.1±0.4	2.7±1.2	44
M3	1.8±0.6	2.6±1.0	65
Q4	2.3±0.8	3.5±1.7	64
Q5	1.5±0.4	2.6±0.9	41
Q6	N.D.[b]	N.D.	43
R7	N.D.	N.D.	33
R8	1.4±0.4	2.4±1.0	44
F9	N.D.	N.D.	32
Y10	N.D.	N.D.	53
E11	N.D.	N.D.	36
A12	N.D.	N.D.	37
L13	N.D.	N.D.	31
H14	N.D.	N.D.	28
D15	N.D.	N.D.	38
P16	N.D.	N.D.	11
N17	1.9±0.8	2.4±1.4	61
L18	N.D.	N.D.	31
N19	N.D.	N.D.	38
E20	1.5±0.3	2.5±0.6	57
E21	1.5±0.5	2.1±0.8	46
Q22	N.D.	N.D.	45
R23	N.D.	N.D.	19
N24	2.3±0.4	2.6±0.7	47
A25	N.D.	N.D.	46
K26	N.D.	N.D.	22
I27	N.D.	N.D.	38
K28	N.D.	N.D.	42
S29	1.4±0.4	2.6±1.1	54
I30	N.D.	N.D.	53
R31	1.5±0.2	2.5±0.4	61
D32	N.D.	N.D.	52
D33	1.9±0.6	3.1±1.2	49

[a] Data are 95% confidence intervals; [b] N.D.: not determined

5.4 Discussion

5.4.1 D-scan of EETI-II

The D-scan of *Ecballium elaterium* trypsin inhibitor II showed that the protein is remarkably resistant to changes in the chirality of the backbone (Fig 5.4.1). 24 out of 30 analogs oxidized to give defined compounds 6 Da less massive than the starting material and significantly earlier eluting on RP-HPLC, products we consider folded. Of the 24 analogs that folded, 12 were high affinity trypsin binders, and 12 showed little to no activity. Of the twelve that were highly active, three were as stable as wild type, and nine were moderately to severely destabilized.

Only six of the thirty synthesized analogs failed to fold: C9, C19, V20, C21, F26, and C27. Four of these are cysteine inversions, which would intuitively be expected to directly disrupt folding by changing the position of the sulfur atoms that form disulfides. Further, F26 and V20 are adjacent to cysteine. What is surprising, however, is that all of the other analogs folded, including C2 and C15, as well as several analogs with D-amino acids adjacent to cysteine. Even analogs inverting rigid proline folded efficiently. In wild type EETI-II, C2 and C15 are disulfide bonded with C19 and C27, respectively, indicating that there is significant flexibility in certain regions. Our MD simulation results of P23 and N24 suggest that although one might have expected large effects from these D amino acid substitutions, the peptide may simply change the turn type in this region to accommodate the replacements.

All 6 of the mutations in the binding loop (P3-R8) abolished the inhibitory activity of EETI-II, as well as the C2 mutation adjacent to it. This result was expected based on the co-crystal structure of EETI-II and trypsin, which shows a tight interaction extending from C2 to C9[128]. What was unexpected, however, was that changes far from the binding site would abolish activity, even though the analogs folded well. C15 is disulfide bonded to the binding loop, and L16 is adjacent to it, but three substitutions of D-alanine for glycine - A18, A22, and A28 - also abolished activity. In these three cases, substitution of Gly with D-Ala far from the binding loop prevented efficient trypsin binding but did not disrupt folding. Substitution of D-alanine for glycine-18 is particularly interesting because it has a positive Ramachandran φ angle, which is hypothesized to indicate tolerance to substitution of D-amino acids[98]. The only other amino acid with a positive Ramachandran φ angle, A25, folded well, was a high affinity trypsin inhibitor, and was slightly more stable than wild type EETI-II.

In all cases, investigated analogs show the formation of an early eluting two disulfide product followed by the slower formation of a still earlier eluting three disulfide product. This is consistent with the known folding pathway of WT EETI-II where the C9-C21 and C15-C27 disulfides form first to generate a stable 2-disulfide core, then the C2-C19 disulfide bond forms, locking the trypsin binding loop

in place. This observation does not, however, support the notion that all analogs have the native disulfide bond configuration.

Out of the analogs investigated, A1, N24, and A25 were as stable as WT and the rest were significantly destabilized. A25 has a positive φ angle, and A1 is on the N-terminus, suggesting that substitution of D-amino acids may not disrupt the conformation of the backbone. A18, the only other substitution at a residue with a positive Ramachandran φ angle, was completely inactive. Our MD simulation results suggest that the G18a mutation in A18 pushed away residues 29 and 30 toward residues 3–8 in the N-terminus, resulting in perturbation of the native beta sheet interactions between residues 20–22 and 26–28 and formation of new beta sheet interactions between residues 26–27 and 7–8. Since residues 2–8 participate in the binding to trypsin, perturbation of these residues likely affects the binding greatly. In contrast, the G25a substitution in A25 led to minimal structural perturbation. Therefore, although both G18 and G25 have a positive φ angle and may be considered potential placeholders for D amino acids, the small size of glycine at G18 might also play an important role in the structure and function of EETI-II.

Inversion of prolines 23 and 30 resulted in folded, active diastereomers of EETI-II, suggesting EETI-II is much more tolerant of disruption of proline than most proteins[141]. Surprisingly both were only moderately destabilized. The 2 disulfide core of P23 is about 15mV less stable than WT, with the third disulfide unaffected presumably because P23 is in the core of the protein. Interestingly, however, inversion of P30 has a minimal effect on the core of the protein, but destabilizes the third disulfide by 10mV. The third disulfide, putatively between C2 and C19, would seem less likely than the core to be effected by changes to the C-terminus, suggesting that the disulfide bond configuration may be different in P30 than WT-EETI-II. Inversion of S29 shows the same trend with even greater (18mV) destabilization of the third disulfide. The last analog with a significantly destabilized third disulfide, A17, is the active mutation closest to C19. As there is not currently a reliable, high throughput method to map disulfide bonds in proteins, in depth structural analysis will be the subject of future work[126,128,142,143].

5.4.2 D-scan of Z33

The D-scan of Z33 showed that it also had significant tolerance to D amino acid substitution. Although no analogs as active as WT Z33 were found, 22 of the 33 analogs still bound to IgG, and of these, ten were high affinity IgG binders with K_D less than 250 nM.

Of the 11 analogs with abolished binding, 10 (Q6, R7, F9, Y10, L13, P16, R23, N24, I27, I30, R31) were located either on the binding face of the helix or adjacent to it, confirming the importance of these residues for the binding of Z33 to IgG (Fig 5.4.2). The eleventh inactive analog is P16.,As expected,

inversion of the rigid proline in the loop connecting the two helices abolished activity, possibly by disrupting their intramolecular organization, which is known to be essential[106,141].

The helical structure of Z33 analogs was found to be correlated to IgG binding (Fig 5.4.3). The 8 analogs that showed significantly decreased α-helicity (R7, F9, L13, H14, P16, L18, R23, and K26) also had significantly decreased binding affinity. Four of these analogs (R7, F9, L13, and R23) did not bind to IgG, and the rest (H14, P16, L18, and K26) had binding affinities more than 40 fold lower than wild type. Analogs that retained binding affinity comparable to WT (K_D < 250 nM) (F1, M3, Q4, N17, E20, E21, Q22, S29, D32, and D33) also largely retained WT helicity, with the least helical analog being Q22 (45% helicity, 67% retained vs WT). Three analogs, Y10, I30, and R31, were highly helical but poor IgG binders. All of these mutations are located directly on the binding face of the helix, and the positions of side chains at these sites are likely essential for Z33 to bind IgG.

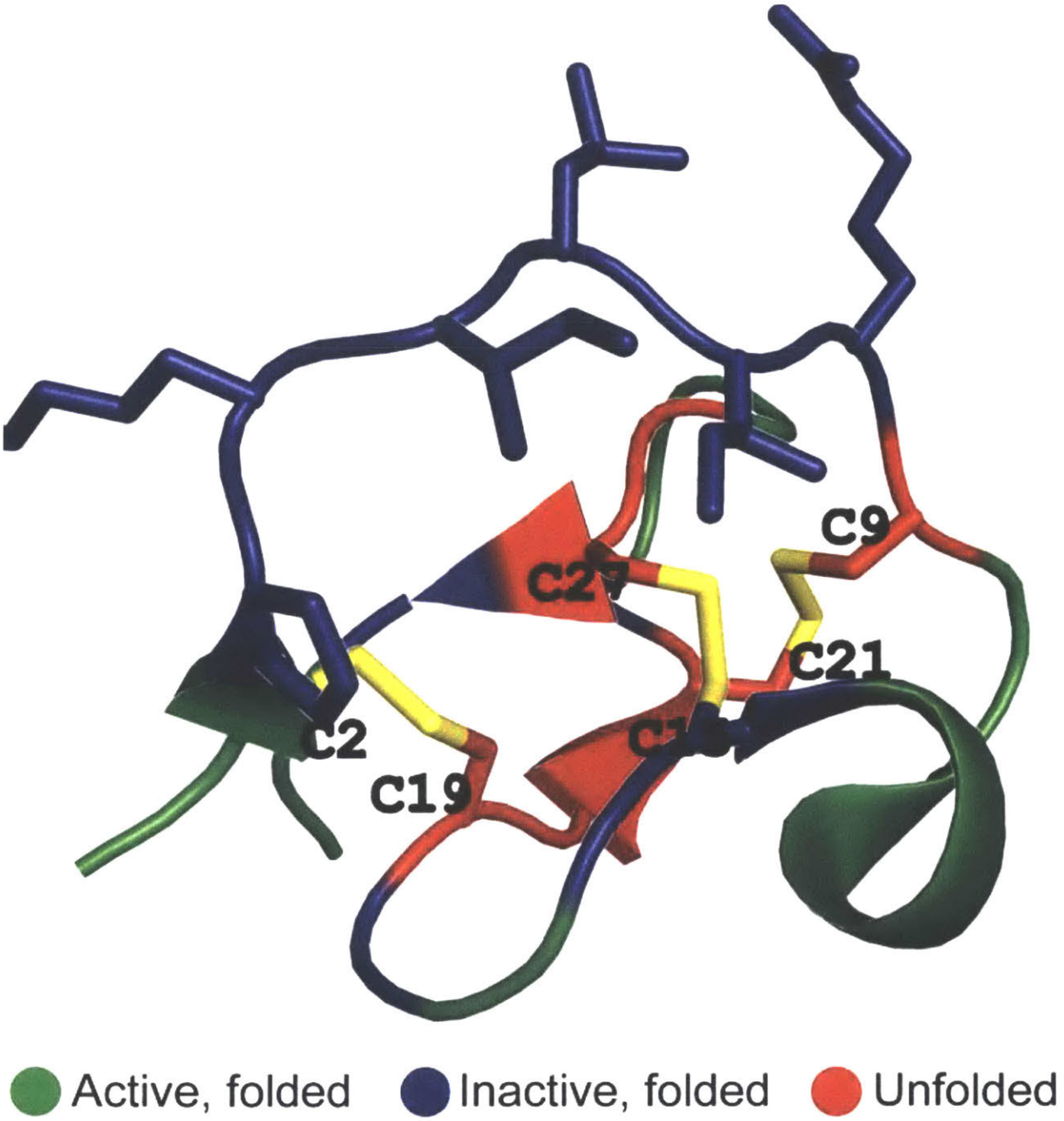


Figure 5.4.1 Graphical summary of folding and activity of EETI-II analogs. Highlighted regions show where inversion of stereochemistry resulted in analogs that did not fold (red), folded but were inactive (blue), and folded with retention of trypsin inhibition activity (green). Disulfide bonds are shown in yellow. Note that most analogs that did not fold (red regions) are mutated near disulfide bonds and most analogs that folded but did not bind (blue regions) are in the binding loop. The graphic was created with PyMOL using structural data for WT EETI-II reported by Kraetzner et al[128].

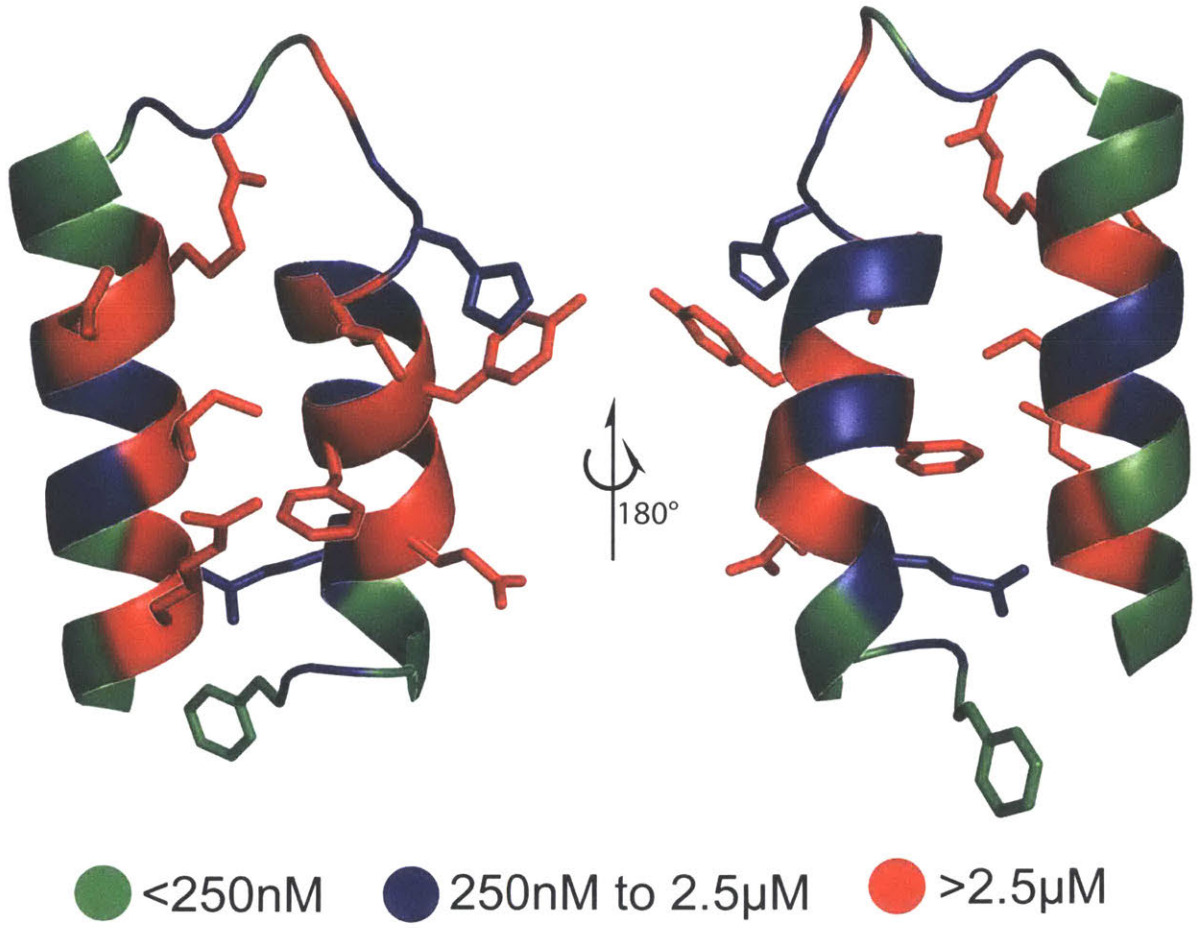


Figure 5.4.2 Graphical summary of folding and activity of Z33 analogs. Highlighted regions show where inversion of stereochemistry resulted in analogs that did not bind IgG (red, KD > 2.5μM), bound IgG weakly (blue, 250nM> KD > 2.5μM), and retained high affinity IgG binding (green, KD < 250nM). Side chains are displayed for residues at the binding interface, as determined from the co-crystal structure of intact Z-domain and an IgG fragment[129]. Note that most analogs that did not bind (red regions) are mutated at or near the binding face. The graphic was created with PyMOL using reported structural data[133].

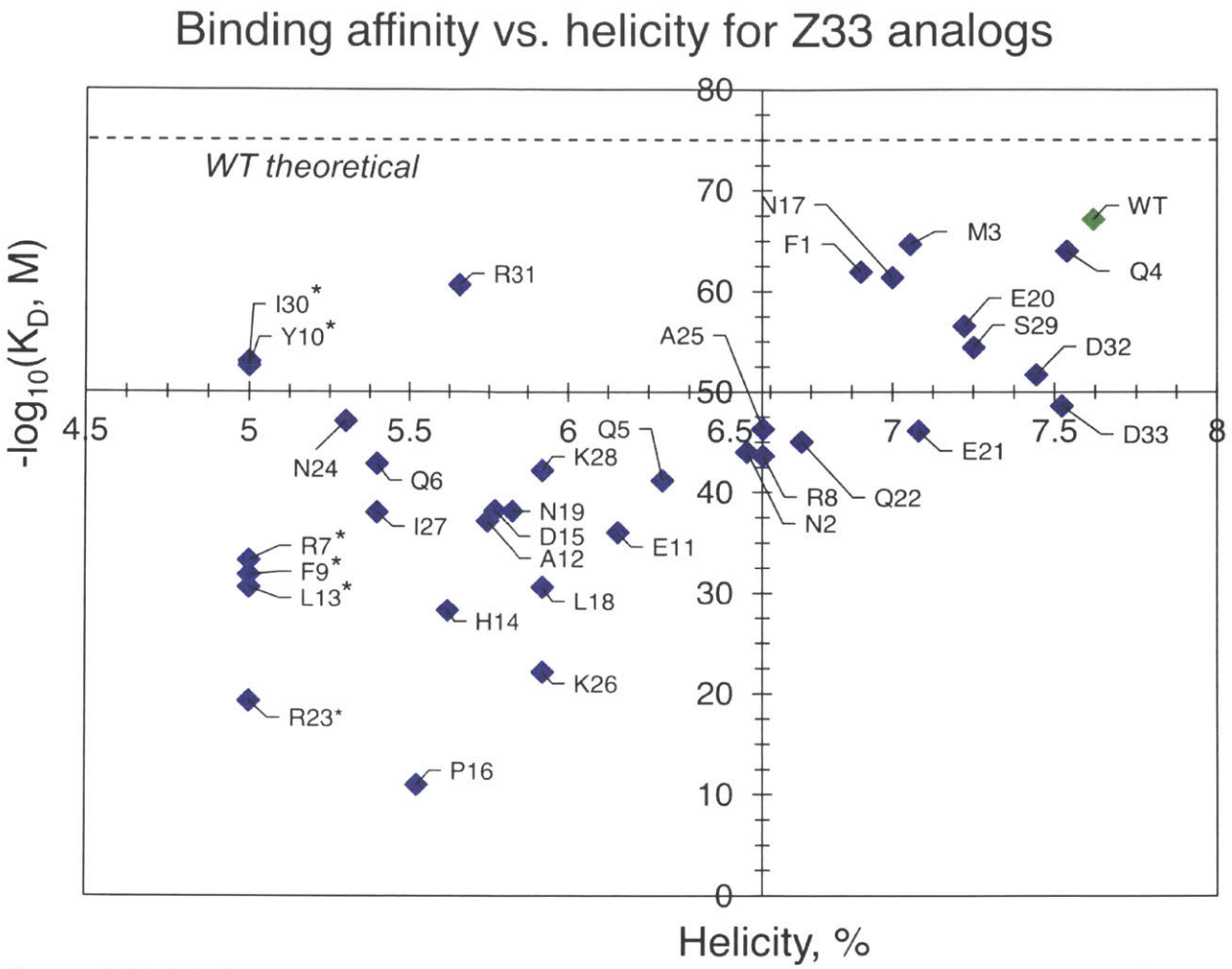


Figure 5.4.3 Binding affinity vs. helicity of Z33 analogs. The helical character of the analogs was estimated via K2D3 deconvolution[140] of their 195-240 nm CD spectra, and the values for K_D are those obtained from bilayer experiments. Analogs marked with a "*" did not bind IgG under the experiment conditions; such analogs are arbitrarily assigned high K_D values of 10^{-5}M to demonstrate the overall trend. The theoretical WT helicity character is calculated from the corresponding NMR structure[133].

5.5 Conclusion

In conclusion, this work paves the way for detailed investigation of the role of chirality in protein folding and function. For the first time, the effect of inversion of each chiral center in a protein has been systematically studied, showing both EETI-II, a disulfide cross-linked cysteine knot, and Z33, two canonical α-helices, to be remarkably tolerant to changes in stereochemistry. In both proteins, mutations at the binding interfaces abolished activity, but many other analogs retained appreciable binding affinity.

We did not, however, find evidence to support the hypothesis that glycine acts as an achiral placeholder for D-amino acids. Substitution of D-alanine for glycine abolished activity or reduced stability in most cases. However, the only case in which substitution of a D-amino acid improved stability and retained native activity was substitution of D-alanine at glycine-25 in EETI-II. This indicates that the effect of D-amino acid substitution in proteins is more complex than previously hypothesized, and D-scans of additional proteins are needed to further elucidate the effect of D-amino acids substitutions in proteins.

So why doesn't Nature make extensive use of D-amino acids in higher organisms? Taken together, these results begin to offer some insight. We found that single point inversions, like other mutations, are generally harmful or of no consequence. No analogs were identified that were significantly more active than the WT proteins, but about 35% retained high activity. In contrast to mutations from one L-amino acid to another, which arise randomly and may be accumulated until an improved multipoint mutant arises, inversion of an amino acid requires either a posttranslational epimerase or non-ribosomal synthase. This complex machinery is difficult to evolve and unlikely to be retained without conferring a significant benefit. If advantageous single point inversions are in fact rare, it may be extremely difficult for potentially more beneficial proteins with multiple inversions to evolve from homochiral precursors.

Protein engineering efforts, however, can incorporate an unlimited number of D-amino acids into known or novel proteins to generate mutants with several inversions. Targeted insertions of D-amino acids could stabilize proteins against proteolysis, abolish undesired functions, and change the metabolites of degraded proteins. Such constructs are likely to be particularly useful for knottins and other highly stable proteins that can be somewhat destabilized without compromising the overall structure and function.

We believe the D-scan can be leveraged to identify cases where a protein can accept a D-amino acid without greatly affecting its function or folding, and inform efforts to engineer mixed chirality proteins with novel properties.

Concluding Remarks

In this work, we have made great progress towards true on-demand peptide synthesis, but we have just begun to explore the scope and application of rapid, solid supported synthesis at elevated temperature. With a firm understanding of the underlying scientific principles, and robust mechanical apparatus, fast flow synthesis can now be extended to improve the speed and fidelity of diverse precision polymer syntheses.

The D-Scan presented in chapter five shows the remarkable power of rapid chemical synthesis of proteins. With turnaround times of hours instead of weeks, it is now feasible to design a functional mixed chirality protein through iterative D-Scans of a biological precursor. Analogous investigations of functional RNA may yield nucleic acids with a mixture of morpholine, ribose, and deoxyribose backbones and various phosphate linkages with novel properties in vivo.

Similarly, rapid access to novel compounds is poised to support a new era of personalized therapeutics with peptide or oligonucleotide drugs produced on demand. Personal cancer immunotherapies now in the clinic rely on a mixture of patient and tumor specific peptide antigens that must be produced on demand. Fast synthesis of these drugs will be necessary for the broad adoption of this therapy, and now, for the first time, it is possible. Likewise, fast synthesis is a cornerstone of the vision that every patient will receive oligonucleotide therapeutics tailored to his or her individual genotype. Although other technical challenges with delivery and stability remain for these approaches, the chemical techniques necessary to support discovery and clinical programs is now available.

With many of the fundamental challenges of fast, solid supported chemistry solved, we can begin to use this technology as a tool to explore new frontiers in science and medicine.

References

[1] J. P. Lalezari, K. Henry, M. O'Hearn, J. S. G. Montaner, P. J. Piliero, B. Trottier, S. Walmsley, C. Cohen, D. R. Kuritzkes, J. J. Eron, et al., *N. Engl. J. Med.* **2003**, *348*, 2175–2185.

[2] R. E. Amori, J. Lau, A. G. Pittas, *J. Am. Med. Assoc.* **2007**, *298*, 194–206.

[3] R. E. W. Hancock, H.-G. Sahl, *Nat. Biotechnol.* **2006**, *24*, 1551–1557.

[4] M. Yadav, S. Jhunjhunwala, Q. T. Phung, P. Lupardus, J. Tanguay, S. Bumbaca, C. Franci, T. K. Cheung, J. Fritsche, T. Weinschenk, et al., *Nature* **2014**, *515*, 572–6.

[5] M. Noguchi, F. Moriya, N. Koga, S. Matsueda, T. Sasada, A. Yamada, T. Kakuma, K. Itoh, *Cancer Immunol. Immunother.* **2016**, *65*, 151–160.

[6] D. J. Craik, D. P. Fairlie, S. Liras, D. Price, *Chem. Biol. Drug Des.* **2013**, *81*, 136–147.

[7] K. Fosgerau, T. Hoffmann, *Drug Discov. Today* **2015**, *20*, 122–128.

[8] S. Cirak, V. Arechavala-Gomeza, M. Guglieri, L. Feng, S. Torelli, K. Anthony, S. Abbs, M. E. Garralda, J. Bourke, D. J. Wells, et al., *Lancet* **2011**, *378*, 595–605.

[9] K. L. Holden, D. A. Stein, T. C. Pierson, A. A. Ahmed, K. Clyde, P. L. Iversen, E. Harris, *Virology* **2006**, *344*, 439–452.

[10] P. L. Iversen, T. K. Warren, J. B. Wells, N. L. Garza, D. V Mourich, L. S. Welch, R. G. Panchal, S. Bavari, *Viruses* **2012**, *4*, 2806–30.

[11] C. Zanetta, M. Nizzardo, C. Simone, E. Monguzzi, N. Bresolin, G. P. Comi, S. Corti, S. Ogino, D. G. Leonard, H. Rennert, et al., *Clin. Ther.* **2014**, *36*, 128–40.

[12] F. J. Raal, R. D. Santos, D. J. Blom, A. D. Marais, M.-J. Charng, W. C. Cromwell, R. H. Lachmann, D. Gaudet, J. L. Tan, S. Chasan-Taber, et al., *Lancet* **2010**, *375*, 998–1006.

[13] R. Kanasty, J. R. Dorkin, A. Vegas, D. Anderson, *Nat. Mater.* **2013**, *12*, 967–77.

[14] J. Kurreck, *Eur. J. Biochem.* **2003**, *270*, 1628–1644.

[15] B. Aussedat, B. Fasching, E. Johnston, N. Sane, P. Nagorny, S. J. Danishefsky, *J. Am. Chem. Soc.* **2012**, *134*, 3532–3541.

[16] B. L. Pentelute, Z. P. Gates, V. Tereshko, J. L. Dashnau, J. M. Vanderkooi, A. A. Kossiakoff, S. B. H. Kent, *J. Am. Chem. Soc.* **2008**, *130*, 9695–9701.

[17] M. Liu, C. Li, M. Pazgier, C. Li, Y. Mao, Y. Lv, B. Gu, G. Wei, W. Yuan, C. Zhan, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 14321–6.

[18] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149.

[19] S. Sakakibara, *Biopolymers* **1995**, *37*, 17–28.

[20] S. B. Kent, *Chemical Synthesis of Peptides and Proteins*, **1988**.

[21] R. Sheppard, *J. Pept. Sci.* **2003**, *9*, 545–552.

[22] M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pept. Protein Res.* **1992**, *40*, 180–193.

[23] L. a Carpino, G. Y. Han, *J Org Chem* **1972**, *37*, 3404–3409.

[24] P. H. Seeberger, W. C. Haase, *Chem. Rev.* **2000**, *100*, 4349–4393.

[25] S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* **1981**, *22*, 1859–1862.

[26] P. H. H. Hermkens, H. C. J. Ottenheijm, D. Rees, *Tetrahedron* **1996**, *52*, 4527–4554.

[27] S. Kent, *J. Pept. Sci.* **2003**, *9*, 574–593.

[28] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science (80-.).* **1994**, *266*, 776–9.

[29] D. M. Hoover, J. Lubkowski, *Nucleic Acids Res.* **2002**, *30*, e43.

[30] N. Annaluru, H. Muller, L. A. Mitchell, S. Ramalingam, G. Stracquadanio, S. M. Richardson, J. S. Dymond, Z. Kuang, L. Z. Scheifele, E. M. Cooper, et al., *Science (80-.).* **2014**, *344*.

[31] C. A. Hutchison, R.-Y. R.-Y. Chuang, V. N. Noskov, N. Assad-Garcia, T. J. Deerinck, M. H. Ellisman, J. Gill, K. Kannan, B. J. Karas, L. Ma, et al., *Science (80-.).* **2016**, *351*, aad6253-aad6253.

[32] I. Coin, M. Beyermann, M. Bienert, *Nat. Protoc.* **2007**, *2*, 3247–3256.

[33] W. C. Chan, P. D. White, *Fmoc Solid Phase Peptide Synthesis : A Practical Approach*, Oxford University Press, **2000**.
[34] M. D. Simon, P. L. Heider, A. Adamo, A. A. Vinogradov, S. K. Mong, X. Li, T. Berger, R. L. Policarpo, C. Zhang, Y. Zou, et al., *Chembiochem* **2014**, *15*, 713–720.
[35] B. Bacsa, C. O. Kappe, *Nat. Protoc.* **2007**, *2*, 2222–7.
[36] S. L. Pedersen, A. P. Tofteng, L. Malik, K. J. Jensen, *Chem. Soc. Rev.* **2012**, *41*, 1826–44.
[37] J. M. Collins, K. A. Porter, S. K. Singh, G. S. Vanier, *Org. Lett.* **2014**, *16*, 940–943.
[38] A. Adamo, R. L. Beingessner, M. Behnam, J. Chen, T. F. Jamison, K. F. Jensen, J.-C. M. Monbaliu, A. S. Myerson, E. M. Revalor, D. R. Snead, et al., *Science (80-.).* **2016**, *352*.
[39] A. Nagaki, M. Togai, S. Suga, N. Aoki, K. Mae, J. Yoshida, *J. Am. Chem. Soc.* **2005**, *127*, 11666–11675.
[40] D. Webb, T. F. Jamison, *Chem. Sci.* **2010**, *1*, 675.
[41] H. Wakami, J. Yoshida, *Org. Process Res. Dev.* **2005**, *9*, 787–791.
[42] T. Noël, S. L. Buchwald, *Chem. Soc. Rev.* **2011**, *40*, 5010.
[43] O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science (80-.).* **2001**, *291*, 1523–1527.
[44] H. S. Hahm, C.-F. Liang, C.-H. Lai, R. J. Fair, F. Schuhmacher, P. H. Seeberger, *J. Org. Chem.* **2016**, *81*, 5866–5877.
[45] T. J. Lukas, M. B. Prystowsky, B. W. Erickson, *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78*, 2791–2795.
[46] A. Dryland, R. C. Sheppard, *J. Chem. Soc. Perkin Trans. 1 Org. Bio-Organic Chem.* **1986**, 125–137.
[47] A. Felix, L. Moroder, C. Toniolo, M. Goodman, *Methods of Organic Chemistry: Synthesis of Peptides and Peptidomimetics*, Thieme Medical Publishers, New Yory, NY, **2004**.
[48] C. M. J. Fox, D. D. Weller, *US Pat.* **2008**, *US8299206*.
[49] T. Kimmerlin, D. Seebach, "'100 Years of Peptide Synthesis': Ligation Methods for Peptide and Protein Synthesis With Applications To B-Peptide Assemblies," **2005**.
[50] P. Nagorny, N. Sane, B. Fasching, B. Aussedat, S. J. Danishefsky, *Angew. Chemie - Int. Ed.* **2012**, *51*, 975–979.
[51] M. Amblard, J.-A. Fehrentz, J. Martinez, G. Subra, *Mol. Biotechnol.* **2006**, *33*, 239–254.
[52] G. B. Fields, R. L. Noble, *Int. J. Pept. Protein Res.* **1990**, *35*, 161–214.
[53] S. B. H. Kent, *Annu. Rev. Biochem.* **1988**, *57*, 957–989.
[54] L. P. Miranda, P. F. Alewood, *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 1181–1186.
[55] M. Quibell, W. Turnell, T. Johnson, *J. Org. Chem.* **1994**, *59*, 1745–1750.
[56] I. Clark-Lewis, R. Aebersold, H. Ziltener, J. W. Schrader, L. E. Hood, S. B. Kent, *Science (80-.).* **1986**, *231*, 134–139.
[57] S. Kakac, H. Liu, *CRC Press* **2002**, *2nd Editio*, DOI 10.1016/0378-3820(89)90046-5.
[58] A. Pernille Tofteng, S. L. Pedersen, D. Staerk, K. J. Jensen, *Chem. - A Eur. J.* **2012**, *18*, 9024–9031.
[59] E. Bayer, G. Jung, I. Halász, I. Sebestian, *Tetrahedron Lett.* **1970**, *11*, 4503–4505.
[60] R. P. W. Scott, K. K. Chan, P. Kucera, S. Zolty, *J. Chromatogr. Sci.* **1971**, *9*, 577–591.
[61] L. M. Varanda, M. T. Miranda, *J. Pept. Res.* **1997**, *50*, 102–108.
[62] M. P. Souza, M. F. M. Tavares, M. T. M. Miranda, *Tetrahedron* **2004**, *60*, 4671–4681.
[63] P. Alewood, D. Alewood, L. Miranda, S. Love, W. Meutermans, D. Wilson, *Methods Enzymol.* **1997**, *289*, 14–29.
[64] C. Hyde, T. Johnson, D. Owen, M. Quibell, R. C. Sheppard, *Int. J. Pept. Protein Res.* **1994**, *43*, 431–440.
[65] Y. Han, F. Albericio, G. Barany, *J. Org. Chem.* **1997**, *62*, 4307–4312.
[66] P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
[67] G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. Bin Li, Y. Lin, H. K. Cui, L. Liu, *Angew. Chemie - Int. Ed.* **2011**, *50*, 7645–7649.
[68] S. K. Mong, A. A. Vinogradov, M. D. Simon, B. L. Pentelute, *ChemBioChem* **2014**, *15*, 721–733.

[69] L. A. Carpino, A. El-Faham, F. Albericio, *Tetrahedron Lett.* **1994**, *35*, 2279–2282.
[70] S. A. Palasek, Z. J. Cox, J. M. Collins, *J. Pept. Sci.* **2007**, *13*, 143–148.
[71] H. Huang, D. L. Rabenstein, *J. Pept. Res.* **1999**, *53*, 548–553.
[72] E. Nicolás, E. Pedroso, E. Girald, *Tetrahedron Lett.* **1989**, *30*, 497–500.
[73] M. Quibell, D. Owen, L. C. Packman, T. Johnson, I. Schon, T. Szirtes, A. Rill, G. Balogh, Z. Vadasz, J. Seprodi, et al., *J. Chem. Soc. Chem. Commun.* **1994**, *12*, 2343.
[74] F. Albericio, N. Kneib-Cordonier, S. Biancalana, L. Gera, R. I. Masada, D. Hudson, G. Barany, *J. Org. Chem* **1990**, *55*, 3730–3743.
[75] I. Coin, R. Dölling, E. Krause, M. Bienert, M. Beyermann, C. D. Sferdean, L. A. Carpino, *J. Org. Chem.* **2006**, *71*, 6171–6177.
[76] A. D'Ausilio, *Behav. Res. Methods* **2012**, *44*, 305–13.
[77] V. K. Sarin, S. B. H. Kent, J. P. Tam, R. B. Merrifield, *Anal. Biochem.* **1981**, *117*, 147–157.
[78] J. Dutta, S. Ramesh, S. M. Radebe, A. M. Somboro, B. G. de la Torre, H. G. Kruger, S. Y. Essack, F. Albericio, T. Govender, *Int. J. Pept. Res. Ther.* **2015**, *21*, 13–20.
[79] R. C. de L. Milton, S. C. F. Milton, P. A. Adams, *J. Am. Chem. Soc.* **1990**, *112*, 6039–6046.
[80] C. M. Deber, M. K. Lutek, E. P. Heimer, A. M. Felix, *Pept. Res.* **1989**, *2*, 184–8.
[81] D. van der Spoel, P. J. van Maaren, P. Larsson, N. Tîmneanu, *J. Phys. Chem. B* **2006**, *110*, 4393–4398.
[82] J. H. Chan, S. Lim, W. F. Wong, *Clin. Exp. Pharmacol. Physiol.* **2006**, *33*, 533–540.
[83] J. Summerton, D. Weller, *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 187–195.
[84] R. M. Hudziak, E. Barofsky, D. F. Barofsky, D. L. Weller, S.-B. Huang, D. D. Weller, *Antisense Nucleic Acid Drug Dev.* **1996**, *6*, 267–272.
[85] V. Arora, D. C. Knapp, M. T. Reddy, D. D. Weller, P. L. Iversen, *J. Pharm. Sci.* **2002**, *91*, 1009–1018.
[86] C. A. Stein, *Mol. Ther.* **2016**, *24*, 1884–1885.
[87] Q. Ge, M. Pastey, D. Kobasa, P. Puthavathana, C. Lupfer, R. K. Bestwick, P. L. Iversen, J. Chen, D. A. Stein, *Antimicrob. Agents Chemother.* **2006**, *50*, 3724–33.
[88] H. Erlich, D. Gelfand, J. Sninsky, *Science (80-.).* **1991**, *252*, 1645–1651.
[89] J. E. Summerton, D. D. Weller, *US Pat.* **1991**, *US5185444*.
[90] T. Harakawa, H. Tsunoda, A. Ohkubo, K. Seio, M. Sekine, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 1445–1447.
[91] S. Paul, M. H. Caruthers, *J. Am. Chem. Soc.* **2016**, *138*, 15663–15672.
[92] M. Szumski, B. Buszewski, *J. Sep. Sci.* **2004**, *27*, 837–842.
[93] P. Dwivedi, B. Bendiak, B. H. Clowers, H. H. Hill, *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1163–1175.
[94] K. Simons, D. Toomre, *Nat. Rev. Mol. Cell Biol.* **2000**, *1*, 31–39.
[95] G. Kreil, *Annu. Rev. Biochem.* **1997**, *66*, 337–45.
[96] D. Soyez, J. Y. Toullec, C. Ollivaux, G. Géraud, *J. Biol. Chem.* **2000**, *275*, 37870–37875.
[97] L. Luo, R. M. Kohli, M. Onishi, U. Linne, M. A. Marahiel, C. T. Walsh, *Biochemistry* **2002**, *41*, 9184–9196.
[98] B. Anil, B. Song, Y. Tang, D. P. Raleigh, *J. Am. Chem. Soc.* **2004**, *126*, 13194–13195.
[99] D. Bang, G. I. Makhatadze, V. Tereshko, A. A. Kossiakoff, S. B. Kent, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 3852–3856.
[100] M. Werle, A. Bernkop-Schn??rch, *Amino Acids* **2006**, *30*, 351–367.
[101] T. L. Peeters, M. J. Macielag, I. Depoortere, Z. D. Konteatis, J. R. Florance, R. A. Lessor, A. Galdes, *Peptides* **1992**, *13*, 1103–1107.
[102] O. Roda, M. L. Valero, S. Peiro, D. Andreu, F. X. Real, P. Navarro, *J. Biol. Chem.* **2003**, *278*, 5702–5709.
[103] ‡,‖ Paolo Grieco, ‡,† Preeti M. Balse, § David Weinberg, § and Tanya MacNeil, ‡ Victor J. Hruby*, **2000**.
[104] A. L. Roth, E. Marzola, A. Rizzi, M. Arduin, C. Trapella, C. Corti, R. Vergura, P. Martinelli, S.

Salvadori, D. Regoli, et al., *J. Biol. Chem.* **2006**, *281*, 20809–20816.
[105] K. Ramalingam, S. R. Eaton, W. L. Cody, G. H. Lu, R. L. Panek, L. A. Waite, S. J. Decker, J. A. Keiser, A. M. Doherty, *Bioorg. Med. Chem.* **1995**, *3*, 1263–1272.
[106] A. C. Braisted, J. A. Wells, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 5688–5692.
[107] M. Werle, T. Schmitz, H.-L. Huang, A. Wentzel, H. Kolmar, A. Bernkop-Schnürch, *J. Drug Target.* **2006**, *14*, 137–46.
[108] H. Kolmar, *Curr. Opin. Pharmacol.* **2009**, *9*, 608–614.
[109] A. Wentzel, A. Christmann, R. Krätzner, H. Kolmar, *J. Biol. Chem.* **1999**, *274*, 21037–21043.
[110] J. L. Lahti, A. P. Silverman, J. R. Cochran, *PLoS Comput. Biol.* **2009**, *5*, e1000499.
[111] R. H. Kimura, D. S. Jones, L. Jiang, Z. Miao, Z. Cheng, J. R. Cochran, *PLoS One* **2011**, *6*, e16112.
[112] S. J. Moore, J. R. Cochran, *Methods Enzymol.* **2012**, *503*, 223–51.
[113] S. J. Moore, M. G. Hayden Gephart, J. M. Bergen, Y. S. Su, H. Rayburn, M. P. Scott, J. R. Cochran, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 14598–603.
[114] J. A. Williams, M. Day, J. E. Heavner, *Expert Opin. Pharmacother.* **2008**, *9*, 1575–83.
[115] Y. Guo, D.-M. Sun, F.-L. Wang, Y. He, L. Liu, C.-L. Tian, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 14276–81.
[116] B. Imperiali, J. J. Ottesen, *J. Pept. Res.* **1999**, *54*, 177–184.
[117] C. Grönwall, S. Ståhl, *J. Biotechnol.* **2009**, *140*, 254–269.
[118] Y. Fezoui, D. L. Weaver, J. J. Osterhout, *Proc Natl Acad Sci USA* **1994**, *91*, 3675–3679.
[119] H. K. Binz, P. Amstutz, A. Pluckthun, *Nat. Biotechnol.* **2005**, *23*, 1257–1268.
[120] J. W. Chin, A. Schepartz, *Angew. Chemie - Int. Ed.* **2001**, *40*, 3806–3809.
[121] N. P. Nord Karin, Gunneriusson E, Ringdahl J, Ståhl S, Uhlén M, *Nat. Biotechnol.* **1997**, *15*, 772–777.
[122] B. Nilsson, T. Moks, B. Jansson, L. Abrahmsen, A. Elmblad, E. Holmgren, C. Henrichson, T. A. Jones, M. Uhlen, *Protein Eng.* **1987**, *1*, 107–113.
[123] T. Fujino, Y. Goto, H. Suga, H. Murakami, *J. Am. Chem. Soc.* **2013**, *135*, 1830–7.
[124] L. M. Dedkova, N. E. Fahmi, S. Y. Golovine, S. M. Hecht, *J. Am. Chem. Soc.* **2003**, *125*, 6616–7.
[125] D. Le-Nguyen, D. Nalis, B. Castro, *Int. J. Pept. Protein Res.* **1989**, *34*, 492–497.
[126] K. J. Nielsen, D. Alewood, J. Andrews, S. B. Kent, D. J. Craik, *Protein Sci.* **1994**, *3*, 291–302.
[127] A. Walewska, A. Ja??kiewicz, G. Bulaj, K. Rolka, *Chem. Biol. Drug Des.* **2011**, *77*, 93–97.
[128] R. Krätzner, J. E. Debreczeni, T. Pape, T. R. Schneider, A. Wentzel, H. Kolmar, G. M. Sheldrick, I. Uson, *Acta Crystallogr. D. Biol. Crystallogr.* **2005**, *61*, 1255–62.
[129] J. Deisenhofer, *Biochemistry* **1981**, *20*, 2361–2370.
[130] M. A. Starovasnik, A. C. Braisted, J. A. Wells, *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 10080–5.
[131] B. F. Erlanger, N. Kokowsky, W. Cohen, *Arch. Biochem. Biophys.* **1961**, *95*, 271–278.
[132] T. Durek, J. Zhang, C. He, S. B. H. Kent, *Org. Lett.* **2007**, *9*, 5497–5500.
[133] M. A. Starovasnik, A. C. Braisted, J. A. Wells, *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 10080–10085.
[134] M. L. Colgrave, D. J. Craik, *Biochemistry* **2004**, *43*, 5965–75.
[135] Y. Zhang, F. B. L. Cougnon, Y. A. Wanniarachchi, J. A. Hayden, E. M. Nolan, *ACS Chem. Biol.* **2013**, *8*, 1907–1911.
[136] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.
[137] B. Hess*, C. Kutzner, D. van der Spoel, E. Lindahl, **2008**.
[138] F. Jiang, Y.-D. Wu, **2014**.
[139] F. Jiang, C.-Y. Zhou, Y.-D. Wu, **2014**.
[140] C. Louis-Jeune, M. A. Andrade-Navarro, C. Perez-Iratxeta, *Proteins Struct. Funct. Bioinforma.* **2012**, *80*, 374–381.
[141] M. W. MacArthur, J. M. Thornton, *J. Mol. Biol.* **1991**, *218*, 397–412.
[142] M. Reinwarth, O. Avrutina, S. Fabritz, H. Kolmar, *PLoS One* **2014**, *9*, DOI 10.1371/journal.pone.0108626.

[143] K. Gupta, M. Kumar, P. Balaram, *Anal. Chem.* **2010**, *82*, 8313–9.

Appendices

In this section, additional experimental details are provided to support reproduction of the work in this thesis, and supplementary figures and tables are given that would be too cumbersome to include in the main text. Each appendix is derived from the supporting information of the referenced work.

The sections on automated fast flow synthesis reference additional electronic materials. These were submitted with the paper copy of this thesis, but may not be adequately preserved or readily accessible for future use. The referenced files have largely been reproduced within Appendix 2, however it was not practical to completely describe the LabVIEW software or mechanical assembly within Appendix 2.

A1: Appendix 1 - Manual fast flow synthesis

This section contains additional experimental details and supplementary figures in support of material in Chapter 2.

A1.1 Common Material

A1.1.1 Materials

2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 2-(7-Aza-1Hbenzotriazole-1-yl)-1,1,3,3-tetramethyluroniumhexafluorophosphate (HATU), hydroxybenzotriazole (HOBt), and Nα-Fmoc protected amino acids were from Chem-Impex International, IL, NovaBioChem, Darmstadt, Germany, and Peptide Institute, Japan. 4-methylbenzhydrylamine functionalized crosslinked polystyrene (MBHA resin) and p-Benzyloxybenzyl alcohol functionalized crosslinked polystyrene (Wang resin) were from Anaspec, CA. Chlorotrytyl chloride resin was from Chem Impex International, IL. All resins were 200-400 mesh, with 100-200 mesh found to be unsuitable. N,N-Dimethylformamide (DMF), dichloromethane (DCM), diethyl ether, methanol (MeOH) and HPLC-grade acetonitrile were from VWR, PA. Triisopropyl silane (TIPS) and 1,2-ethanedithiol were from Alfa Aeser, MA. Solvents for LC-MS were purchased from TJ Baker and Fluka. All other reagents were purchased from Sigma-Aldrich, MO. Unless noted, solvents were not anhydrous, and all material was used without further purification. ***Common solvent mixtures used throughout these experiments are: 0.1% (v/v) TFA in water (A), 0.1% (v/v) Formic acid in water (A'), 0.1% (v/v) TFA in acetonitrile (B), and 0.1% (v/v) formic acid in acetonitrile (B').***

A1.1.2 LC-MS Analysis of Peptides

All peptides were analyzed on an Agilent 6520 Accurate Mass Q-TOF LC-MS under one of four conditions, as indicated below. **Condition 1:** An Agilent C3 Zorbax SB column (2.1 x 150 mm, 5 μm packing) was used with a flow rate of 0.4mL/min of the following gradient: A' with 1% B' for 3 minutes, 1-61% B' ramping linearly over 15 min, and 61% B' for 4 minutes. **Condition 2:** an Agilent C18 Zorbax SB column (2.1 x 250 mm, 5 μm packing) was used. The flow rate was 0.4mL/min of the following gradient: A' with 1% B' for 5 minutes, 1-61% B' ramping linearly over 15 min, and 61% B' for 4 minutes. **Condition 3:** An Agilent C3 Zorbax SB column (2.1 x 150 mm, 5μm packing) was used with a flow rate of 0.8mL/min of the following gradient: A' with 5% B' for 3 minutes, 5-65% B' ramping linearly over 9 min, and 65% B' for 1 minute. **Condition 4:** An Agilent C3 Zorbax SB column (2.1 x 150 mm, 5 μm packing) was used with a flow rate of 0.4mL/min of the following gradient: A' with 5% B' for 3 minutes, 5-95% B' ramping linearly over 15 min, and 95% B' for 4 minutes.

Unless noted, peptides were analyzed under condition 1. LYRAG used in the coupling time study and GCF used in the cysteine activation studies were analyzed under condition 2. Peptides in Figure 2.4.2, A1.12a-b, A1.13, and A1.14 were analyzed under condition 3. Peptides in Figure A1.12c were analyzed under condition 4.

Total ion current is displayed in all chromatograms, both in Chapter 2 and this appendix.

A1.1.3 Manual Synthesizer Design

Overview:

Figure A1.1 is a schematic of the synthesizer. An HPLC pump was used to deliver either methanol purge solvent for washing the pump heads, reaction vessel, and UV detector after use; DMF wash solvent for removal of reagents and byproducts during synthesis; or 50% (v/v) piperidine in DMF for deprotection of the N-terminus. The positions of valves 1 and 2 determined which fluid was delivered. A back pressure regulator on the outlet of the HPLC pump prevents siphoning of the solvents.

A syringe pump was used to deliver coupling solution. To switch between the syringe pump and HPLC pump, a quick connect was manually moved from the outlet of the HPLC pump to the syringe on the syringe pump. A valve would be ineffective here because the line between the syringe pump and valve would retain coupling solution, causing incorrect incorporation in the next cycle.

The reaction vessel and a heat exchanger were submerged in a water bath to maintain a constant 60°C.

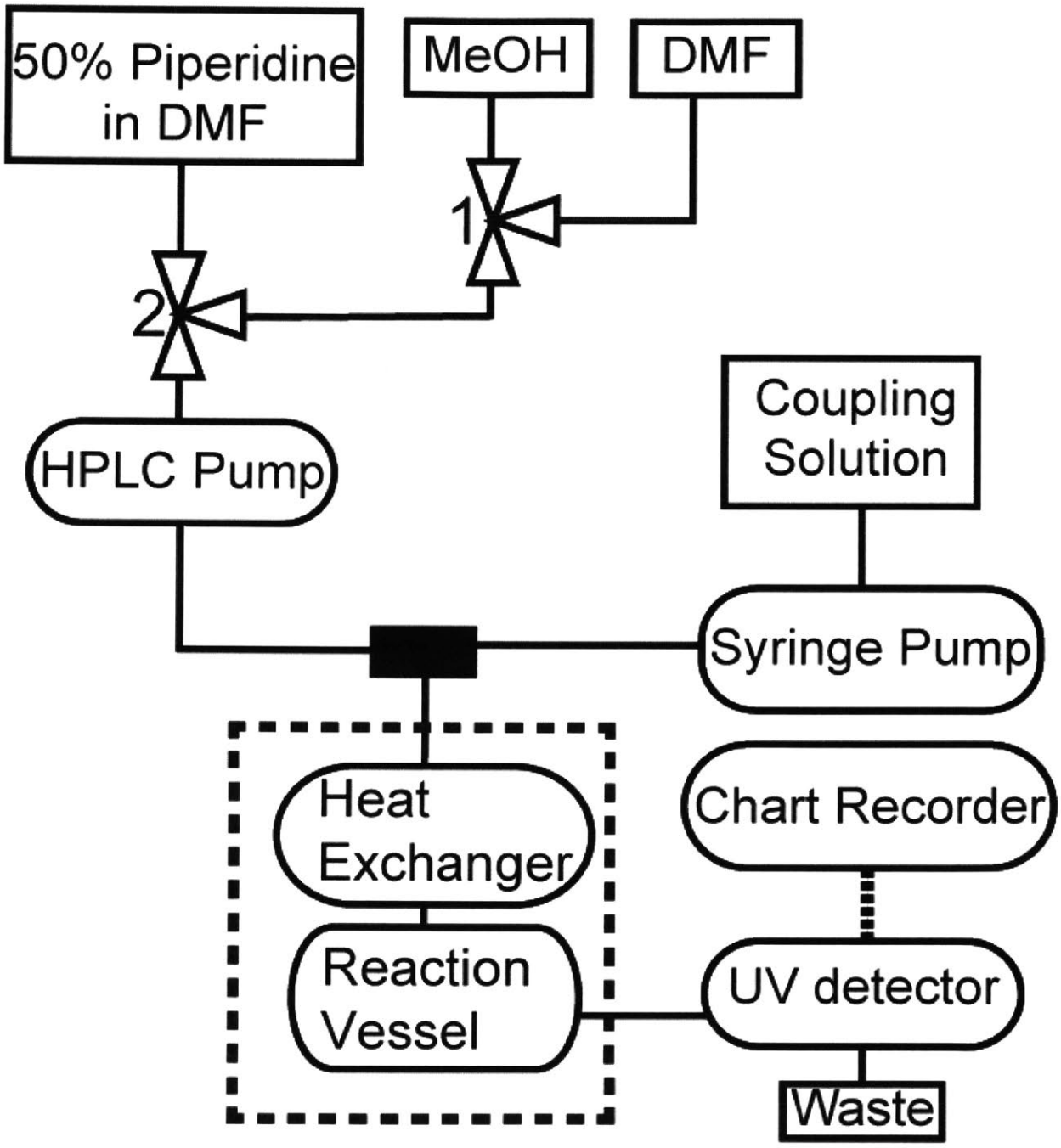


Figure A1.1 Schematic of the Manual Synthesizer. Fluid paths are illustrated. The dashed box enclosing the heat exchanger and reaction vessel represents a water bath. The dashed line from the UV Detector to the Chart recorder is a data link.

Detailed Description:

A Varian Prostar 210 HPLC pump, KD Scientific KDS200, Harvard Apparatus Model 22 or Harvard Apparatus PHD2000 Infusion syringe pump, Varian Prostar 320 UV detector set to 304nm, Amersham Pharmacia Biotech chart recorder, and VWR 39032-214 water bath were used. The HPLC pump delivered about 95% of the nominal flow rate. Disposable 10mL syringes (BD 309604) were used to deliver coupling solutions. Valves 1 and 2 were Swagelok 1/8" 3-way valve (SS- 41GXS2). The methanol, DMF, and 50% (v/v) piperidine lines were 1/8" outer diameter (OD), 1/16" inner diameter (ID) FEP or PFA (Idex 1521 or McMaster 5773K11). All other lines were 1/16" OD, 0.030" ID PFA (Idex 1514L). All lengths were minimal, except the tubing between the quick connect and the reaction vessel. This included a 1.6m coil which was submerged along with the reaction vessel and served as a heat exchanger to ensure that reactants were at 60°C before reaching the reaction vessel. For experiments with flow rates above 20ml/min, this section of tubing was replaced with stainless steel tubing (McMaster 51755K35). Whenever tubing had to be joined, Swagelok 1/16" unions were used (SS-100-6). These were used to attach the heat exchanger and join the outlet of the synthesis vessel to the UV detector. The manually changed quick connect was a female luer to 10-32 female HPLC fitting (Idex P-659). This connected directly to the syringes on the syringe pump or to a mating male luer to 10-32 female fitting on HPLC pump outlet (Idex P- 656). The connection between the UV detector and chart recorder was a data link (three 18ga insulated copper wires).

A1.2 Synthesis with First Generation Reaction Vessel

A1.2.1 Conditions for Peptides Synthesis with the First Generation Vessel

All peptides in this section, except the ACP(65-74) batch in Figure 2.3.2D, were synthesized on the reported flow platform. All peptides except ACP(65-74) batch and ACP(65-74) flow RT in Figures 2.3.2D and 2.3.2C were synthesized at 60 °C, with reagents heated immediately before use via a coil of tubing in the water bath (see synthesizer design). One synthetic cycle consisted of amide bond formation (coupling), removal of coupling reagent (wash), N_α fmoc removal (deprotection), and removal of the deprotection reagent and the reaction product, piperidine-dibenzofulvene (piperidine-DBF) (wash).

Unless noted, coupling was performed by delivering the following coupling solution at 12 mL/min (for approximately 30 seconds). The coupling solution consisted of 2 mmol of N_α-Fmoc and side chain protected amino acid dissolved in 5 mL of 0.4 M HBTU in DMF and 1 mL of DIEA[11]. Cysteine was dissolved in 5 mL 0.4 M HBTU in DMF, 0.687 mL neat DMF, and 0.313 mL DIEA. In both cases, amino acids were dissolved in HBTU solution up to several hours before use, and DIEA was added within two

minutes of use. Volumetric measurements were made at RT (18-20 °C). The ACP(65-74) shown in Figure 2.3.2A was synthesized by substituting HATU for HBTU in the above solution.

Next, the coupling solution was removed with 20 mL of DMF delivered at 10 mL/min over 2 minutes, and then the N_{α}-Fmoc protecting group was removed with 3.3 mL of 50% (v/v) piperidine in DMF delivered at 10 mL/min over 20 seconds. Excess piperidine and piperidine-DBF were removed with 20 mL of DMF delivered at 10 mL/min over 2 minutes to complete one cycle.

All peptides were synthesized on 100 mg of 1% divinyl benzene crosslinked polystyrene resin. To produce C-terminal carboxamide peptides, MBHA functionalized resin with a loading of 1 mmol per gram was used, and the TFA labile Rink linker was coupled as the first amino acid[2]. To produce CS8 terminal hydrazide peptides for ligation, Wang resin, functionalized as below, was used. The loading was 0.6 mmol/g (0.06 mmol scale).

Non-cysteine containing carboxamide peptides were cleaved from the resin and side-chain deprotected by treatment with 2.5% (v/v) water and 2.5% (v/v) TIPS in TFA for two hours. Cysteine containing carboxamide peptides were cleaved from the resin and side chain deprotected with 2.5% (v/v) EDT, 2.5% (v/v) TIPS, and 1% (v/v) water in TFA for two hours. Hydrazide peptides were cleaved with 5% (v/v) EDT, 5% (v/v) TIPS, and 2.5% (v/v) water in TFA for two hours. In all cases, the resin was removed and compressed air was used to evaporate the cleavage solution to dryness at RT. The resulting solids were washed three times with cold diethyl ether, dissolved in 50% A/50% B (v/v), and lyophilized. Side chain protection was as follows: Arg(Pbf), Tyr(tBu), Lys(Boc), Asp(OtBu), Gln(Trt), Ser(tBu), His(Trt), Asn(Trt), Trp(Boc), Glu(OtBu), Thr(tBu), Cys(Trt).

The Wang resin was functionalized as follows:[3] 5.47 g Wang resin was added to a 500 mL round bottom flask and suspended in 98 mL of DCM and 1.12 mL of N-Methyl morpholine. This was stirred in an ice bath for 5 min and 2.03 g p-nitrophenol chloroformate was added as a powder. This mixture was stirred for 8.5 hours. The ice in the bath was not replenished, which allowed the reaction to slowly reach RT. The mixture was filtered and the solids washed with DCM, DMF, MeOH, and DCM to give a white resin. The resulting resin was placed in a clean 500 mL round bottom flask in an ice bath, and suspended in a prepared mixture of 210 mL DMF, 54 mL DCM, and 1.1 mL hydrazine monohydrate prechilled to 0°C. This yielded a bright yellow solution. The reaction proceeded for 18 hours in an ice bath that was allowed to melt. The mixture was then filtered, and the solids washed as before to give hydrazine-functionalized Wang resin. **WARNING**: This resin promotes severe double incorporation of glycine, regardless of the position of glycine in the polypeptide chain. The chlorotrytyl hydrazide resin used for studies with the second generation reaction vessel (and described there) is superior. This section is included only for completeness.

A1.2.2 Reaction Vessel Assembly

Figure A1.2 shows the first generation reaction vessel. The reaction vessel consisted of a tube with standard compression fittings on each end (3/8" to 1/16" reducing unions). On the downstream end there was also a frit. This was positioned by a support designed to fit inside the reaction vessel and seat against the bottom of the fitting on the outlet. Various frit porosities were used (see High Pressure section). The part number below is for a 20 micron frit, the most commonly used.

The body (blue) was a 3.5 inch segment of PFA tubing with outer diameter 3/8" and inner diameter ¼". The frit (red) was a ¼" sintered stainless steel disk 1/16" thick. The frit support (white) was a 0.5" length of ¼" OD PTFE tubing. As the fittings are tightened, the nut (dark grey) compresses the ferrule (yellow) against the fitting body (light grey), sealing the reaction vessel body to the fitting body. This also compresses the reaction vessel body against the frit, forming an internal seal against the frit.

The reaction vessel body and frit were purchased from McMaster-Carr as part numbers 51805K73 and 9446T314, respectively. The nut, ferrule, and fitting body are available as a set along with the 1/16" nut and ferrule from Swagelok as part number SS-600-6-1. Replacement ferrules are available as SS-600-SET.

The reaction vessel was assembled by first cutting the body and frit support to length, ensuring the ends were square. We used a sharp razor blade and steady hand for these operations. Next, the outlet (downstream) end was assembled. The frit was placed on a solid, clean surface and the reaction vessel body was pressed onto it. We verified the frit was square and flush with the end of the reaction vessel, then carefully pushed the frit support in slightly, pushing the frit up towards its final position. Firmly seating the reaction vessel body in the fitting body forced the frit to its final position. We verified that the frit was square and properly positioned under the ferrule, then installed the fitting according to the manufacturer's instructions. Once sealed, the frit could not be removed and reseated. Finally, the inlet fitting was installed according to the manufacturer's instructions. A high pressure reaction vessel with a stainless steel body was also built. In this case, the downstream fitting had to be tightened well beyond specification to effect a seal with the frit. The reaction vessel was typically replaced every 3-8 syntheses. When replacing the reaction vessel, the ferrules, frit and reaction vessel body were not reused. All other parts were reused. The nuts were recovered by cutting the reaction vessel body in half.

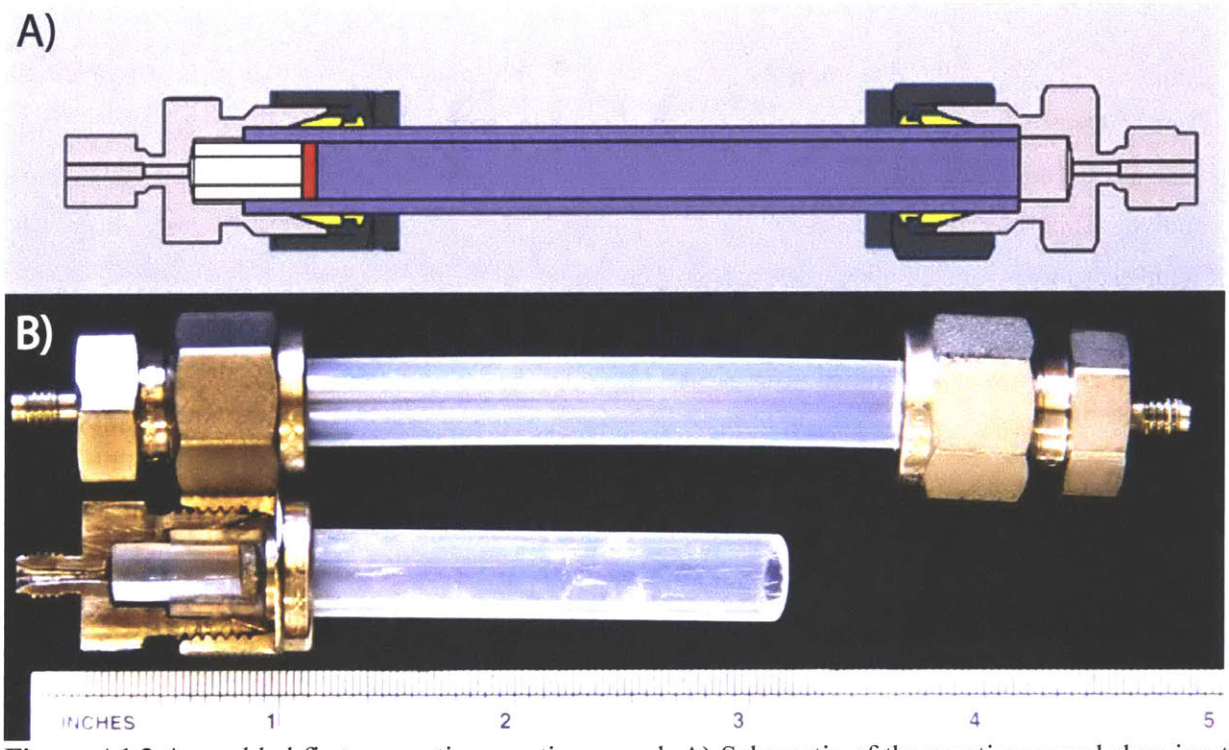


Figure A1.2 Assembled first generation reaction vessel. **A)** Schematic of the reaction vessel showing the reaction vessel body (blue), fitting bodies (grey), nuts (dark grey), ferrules (yellow), frit (red), and frit support (white). **B)** Photograph of an assembled reaction vessel (top) and cutaway showing the internal components of the fritted end (bottom). The 1/16" nuts have been removed from the fitting bodies.

A1.2.3 Resin Loading Procedure

To load the reaction vessel with resin, the upstream fitting body was removed and a slurry of resin in methanol was pipetted in. The reaction vessel was completely filled with methanol, and the fitting body was reinstalled. The inlet line and heat exchanger were filled with solvent by attaching them to the quick connect and running the HPLC pump before attaching them to the reaction vessel. The reaction vessel was then kept upright in the water bath so that any small bubble would move to the top and not interfere with synthesis. Before the first coupling, the resin was washed for two minutes with DMF at 10 mL/min.

A1.2.4 High Pressure

A low pressure polymer reaction vessel was used, so an overpressure alarm on the HPLC pump was set to shut off the pump at 240 psi. This alarm was occasionally triggered, most often after coupling Rink or Asn(Trt). When the alarm was triggered, the system was allowed to rest for 30 seconds, and the pumps were restarted without further incident. During this resting phase, the resin visibly expanded. By observing the HPLC pump pressure, we concluded that if too much pressure was applied to these beads, they begin to collapse. This increases the pressure drop across the bed, and the rate of collapse. This quickly triggers the over pressure alarm. Similar 1% divinyl benzene crosslinked polystyrene resin

available for gel permeation chromatography from Bio-Rad is recommended for gravity driven separations only, because it is very soft once swollen.

When the reaction vessel was disassembled immediately after such an event, the resin looked like a solid block and, when probed with a pipette tip, felt like a hard mass. It was impossible to immediately pipette it out. After a few tens of seconds, the resin relaxed and could be pipetted out. A high pressure stainless steel reaction vessel was built and tested, but the very high pressure necessary to maintain a high flow through a compacted bed (>1000 psi) resulted in extrusion of the resin through the frit.

To overcome these problems and scale up, the second generation reaction vessel, described below, was constructed.

A1.2.5 Quantification of UV detector response

To understand the UV traces produced and the wash efficiencies they represent, the response of the UV detector was quantified. To determine the approximate concentration of amino acid in the UV traces, a serial dilution of Fmoc-Ala-OH coupling solution was prepared. The initial concentration of amino acid was about 0.3 M (2 mmol in 6.5 mL total volume). 10x, 100x, 1000x, 10,000x and 100,000x dilution standards were prepared and injected directly into the UV detector. The 100x dilution standard ($3x10^{-3}$ M) was just below saturation. The 10,000x dilution ($3x10^{-5}$M) standard was just above baseline, about 1% of scale, as expected. The 100,000x dilution was below the detection limit. The highly reproducible wash-out traces (Figure A1.3) show that this is representative of all amino acids (qualitatively different traces between cycles would be expected if the absorbance was vastly different).

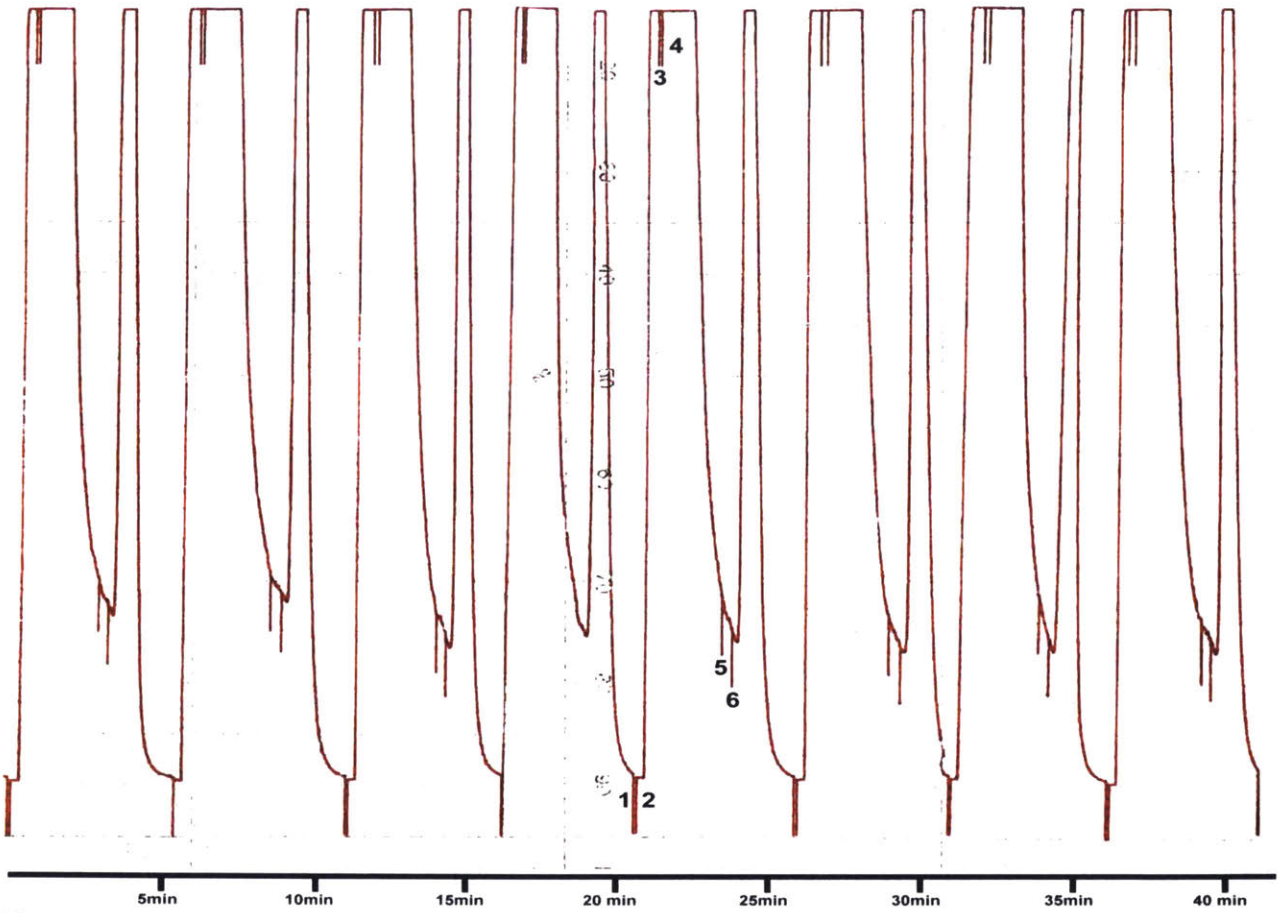


Figure A1.3 UV record of the incorporation of the final eight residues during the conotoxin synthesis. Negative marks represent manual actions. The scanned trace has been color enhanced and a timeline added, taking zero to be the beginning of the trace. The marks of one cycle have been annotated with 1 indicating the end of the previous wash, 2 indicating the beginning of coupling, 3 indicating the end of coupling, 4 indicating the start of the first wash, 5 indicating the end of the first wash and start of the deprotection, and 6 indicating the end of the deprotection and start of the second wash. The quick connect is moved between 1 and 2, and between 3 and 4. Inconsistencies in cycle time and missing marks are due to human error.

A1.2.6 N^{α} Deprotection Optimization

In developing an Nα deprotection protocol, piperidine in DMF was selected as the standard deblocking reagent. A concentration of 50% (v/v) in DMF was selected over the more common 20% (v/v) in DMF,[2] because the deprotection solution is diluted as it enters the reaction vessel. A higher concentration is therefore desirable. The flow rate was set at 10 mL/min (the maximum reliably obtainable with the first generation reaction vessel) to reach an effective concentration in the minimum time. To determine the length of the deprotection step, ALF was synthesized with a double deprotection of every residue, and the UV absorbance of the effluent was monitored at 304 nm. Piperidine and DMF do not absorb well at this wavelength, but piperidine-DBF, the deprotection product, does. Therefore, the presence of a second peak after the second deprotection indicated that the initial deprotection was inadequate. No second peak was observed after 60, 30, and 15 second initial deprotections, but a small peak was observed after a 6 second initial deprotection. In all cases, the first deprotection was at 10 mL/min, and the second was for one minute at 10 mL/min. Fmoc removal is reported to be sequence dependent, so the final deprotection time was set at 20 seconds. Further optimization is likely possible, but was not considered necessary at this time.

The double deprotection protocol must be used to determine deprotection time because it takes significantly longer to wash the piperidine-DBF adduct out of the resin than to remove the Nα Fmoc group. If the effluent is simply monitored until the absorbance returns to near-baseline, most of the "deprotection" time will be spent washing the resin with deprotection reagent after deblocking is complete.

A1.2.7 Wash Optimization:

Visual observation of the reaction vessel during wash cycles showed recirculation and mixing of DMF wash solvent and coupling solution. Other solvent exchanges showed the same behavior. Differences in color and refractive indices allowed the direct observation of all exchanges. Based on this observation, it was expected that the wash efficiency was primarily dependent on the volume of solvent used, as predicted by a continuous dilution model. To test this, the UV absorbance of the reaction vessel effluent was monitored at 304 nm during the triplicate synthesis of ACP(65-74).

During each synthesis, two consecutive residues were washed at 10 mL/min, two at 7 mL/min, two at 4 mL/min, two at 2 mL/min, and the final two at 1 mL/min. Wash rates were randomly assigned to blocks of amino acids, ensuring that no block was washed at the same rate twice. The time required for the detector to desaturate was measured for each residue. Desaturation represents approximately 99% reduction in amino acid concentration. This wash efficiency was selected because the wash is essentially complete, but air and particulate contamination in the detector are less significant than at lower signal

levels. The data are shown in Figure A1.4. It is interesting that the exponent (parameter b) is significantly below negative one. This means that less solvent is required to desaturate the detector at higher flow rates. This is not consistent with the proposed continuous dilution model, and we have not yet determined the source of this discrepancy. Further, we cannot explain the large variance in wash times at low flow rates.

Based on these results, the maximum reliable flow rate (10 mL/min) was selected for the wash. The wash time was set at two minutes, which consistently reduced the final concentration of coupling solution to 0.2% of the initial concentration. No double incorporation, the expected outcome of an inadequate wash, was observed. The apparatus used did not provide a direct way to monitor the removal of piperidine (the UV absorbance is similar to DMF at accessible wavelengths), so the same wash cycle was used for the second wash. If it is assumed that piperidine is removed at the same rate as piperidine-DVB, this is an overestimate of the necessary wash time (see UV trace, Figure A1.3).

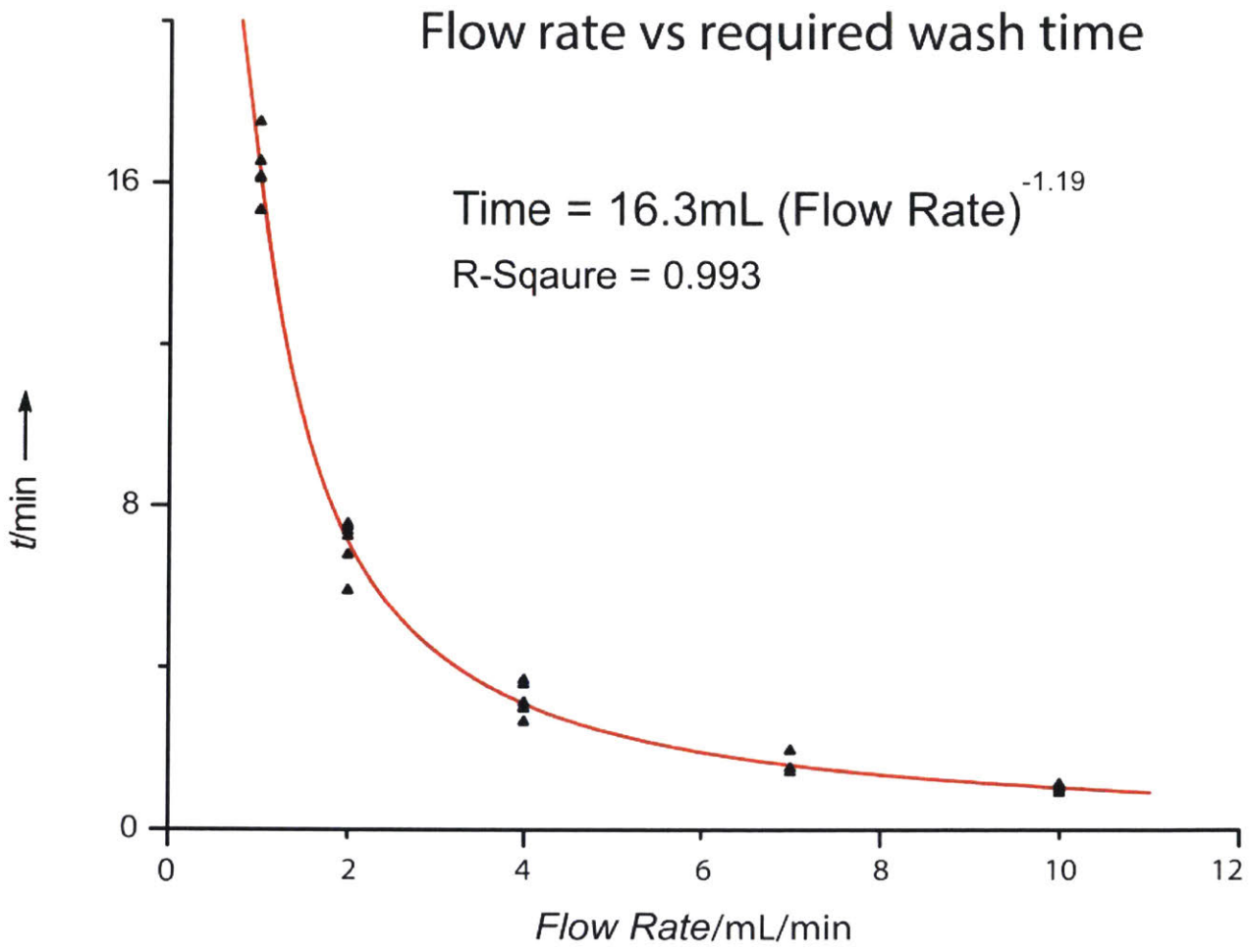


Figure A1.4 Graph of flow rate vs required wash time, showing the exponential decline in wash time with increasing flow rate.

A1.2.8 Determination of Minimum Coupling Time:

LYRAG was selected as a model peptide to determine the minimum coupling time, because there is an arginine deletion that can be monitored. For 90, 45, and 30 second nominal coupling times, the coupling solution was delivered at 4, 8, and 12 mL/min, respectively. This allowed the delivery of a full 2 mmol of amino acid. Flow rates above 12 mL/min were not reliably obtainable, so for the 15 second trial, half of the coupling solution was used (1 mmol AA in 2.5 mL 0.4 M HBTU in DMF, with 0.5mL DIEA). At a coupling time of 7 seconds, the time spent manually moving the quick connect (5-6 seconds) is very significant, so 1.2 mL of coupling solution was delivered. This was the volume of the heat exchanger and inlet line, so the coupling solution did not reach the reaction vessel until it was cleared from the lines by the DMF wash. The wash, at 10 mL/min, takes 7.2 seconds to clear 1.2 mL, giving a 7 second coupling time. In the other runs, the 5 seconds to move the quick connect is added to the nominal coupling time, as is about a 10% increase over the nominal volume of coupling solution. The difference in the time taken for the DMF wash solvent and coupling solution to clear the inlet line is subtracted. Thus, more accurate coupling times are 93, 53, 39, 23, and 7 seconds. This does not include the time required to wash the coupling solution from the reaction vessel (see Wash Optimization). The seven second coupling showed increased Arginine deletion, so the 30 second protocol was selected as a conservative estimate for the general case. Fmoc-ALF was produced with the same procedure, and showed no change in peptide quality with reduction in coupling time. Data are presented in Figure 2.3.1.

A1.2.9 Minimization of Cysteine Racemization

In the initial syntheses of PnIA (A10L) conotoxin and HIV-1 PR(81-99), cysteine was activated like all other amino acids (1 eq HBTU, 2.9 eq DIEA), and significant diastereotopic impurities were observed in the products. These were determined to be the result of cysteine racemization. To investigate conditions that reduce racemization, a model system, GCF, was selected because the diasteromer formed upon racemization is resolved by RP-HPLC.[4] The standard synthetic procedure was used, except coupling time was increased to one minute for 60 °C runs and 6 minutes for RT runs. Rink, Gly, and Phe were all activated according to the standard procedure, with one equivalent of HBTU (5 mL, 0.4 M) and 2.9 equivalents of DIEA (1 mL). For cysteine, eight different activation procedures were tried. The conditions for each are described in Table A1.1, and below. Additionally, an authentic diasteromer was produced using Fmoc-D-Cys(Trt)-OH and the successful activation procedure of trial 5. TIC traces are shown for trials 4, 5, 6, 7 and the authentic diastereomer (Figure A1.5). Runs not shown are similar to 4. For procedures that used less than 1 mL of DIEA, DMF was used to replace this volume. In all cases, the activator, additive, and 2 mmol amino acid were dissolved in 5mL DMF, and additional DMF was added,

if needed. The base was added immediately before use. Trial 8 employed an isolated C-terminal pentafluorophenyl (Pfp) ester (Fmoc-Cys(Trt)- OPfp) without additional activators, additives, or base.

Table A1.1 summarizes the results, with racemization quantified by integration of the extracted ion current. This enables quantification of racemization below the TIC baseline. The results obtained are consistent with previous reports.[8]

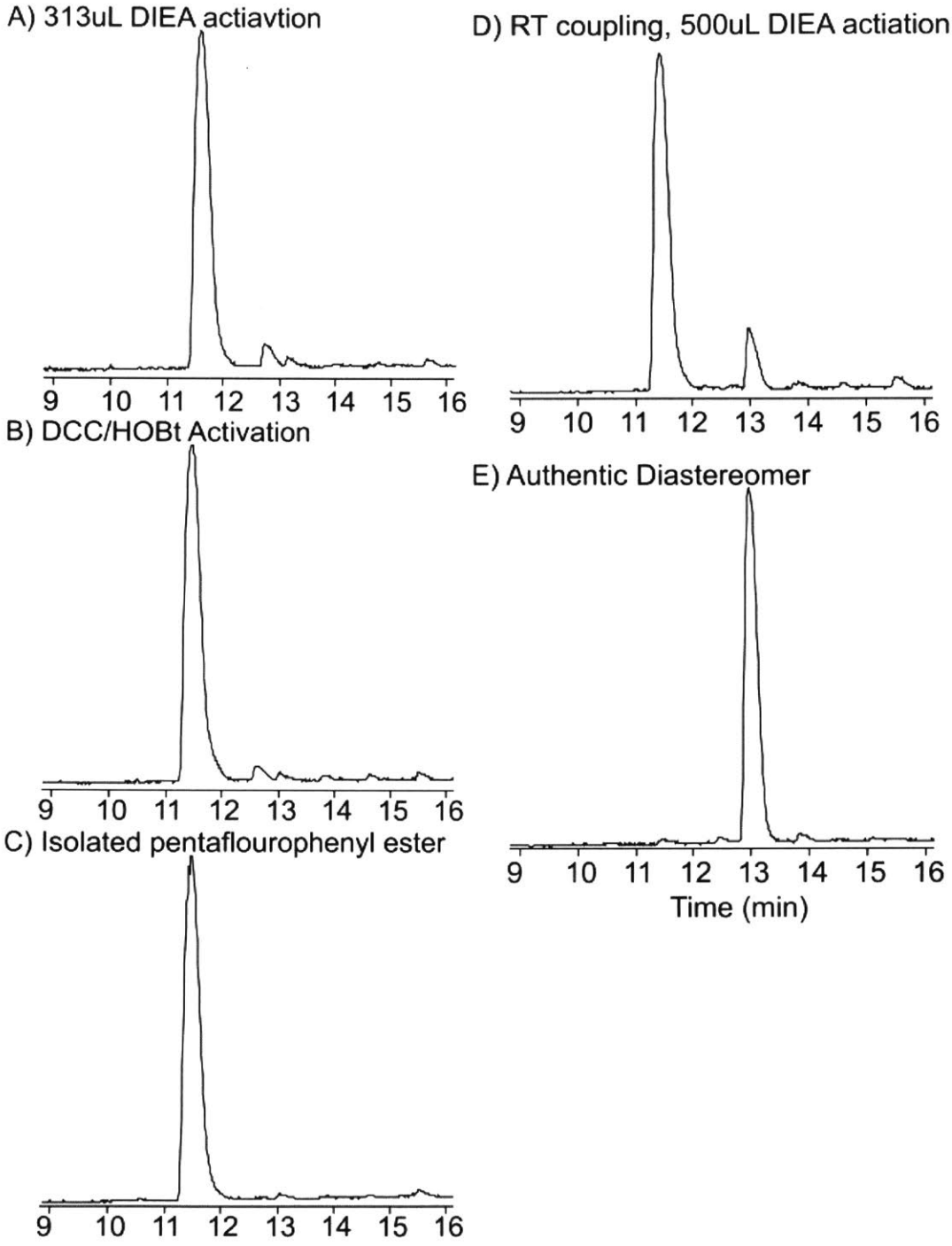


Figure A1.5 GCF produced with various cysteine activation schemes. The peak eluting between the desired product and diastereomer in A and B is hydrolysis of the C-terminal carboxamide. The diastereomer is barely visible. Conditions are listed in Table A1.1. **A)** Condition 5: 0.9 eq DIEA, 1eq HBTU **B)** Condition 7: 0.9 eq DCC, 1.1 eq HOBT **C)** Condition 8: isolated Pfp ester **D)** Condition 4: 2.9 eq DIEA, 1 eq HBTU, RT coupling **E)** Authentic Gly-D-Cys-L-Phe. The total ion current is displayed in each chromatogram.

Table A1.1 Activation procedures explored to reduce cysteine racemization; 5-8 were successful.

Trial	Activator	Additive	Base	Temp	Racemization
1	**HBTU (1 eq)**	**None**	**DIEA (2.9eq)**	**60°C**	**10%**
2	HBTU (1 eq)	HOBt (1eq)	DIEA (2.9eq)	60°C	18%
3	**HBTU (1 eq)**	**None**	**DIEA (2.9eq)**	**Cys RT G,F 60°C**	**11%**
4	HBTU (1 eq)	None	DIEA (2.9eq)	RT	10%
5	**HBTU (1 eq)**	**None**	**DIEA (0.9eq)**	**60°C**	**1%**
6	HBTU (1 eq)	HOBT (1eq)	DIEA (0.9eq)	60°C	1%
7	**DCC (0.9eq)**	**HOBt (1.1eq)**	**None**	**60°C**	**1%**
8	OPfp	None	None	60°C	1%

A1.2.10 Affibody Synthesis

Throughout this section, ligation buffer refers to a 6 M GnHCl, 0.2 M sodium phosphate buffer at the specified pH. Buffer P is a 20 mM Tris, 150 mM NaCl solution at pH 7.5.

Oxidation chemistry reported by Fang *et al.*[3] was used to ligate three fragments into a synthetic, 58 residue protein. Thiozolidine was found to be unstable in the conditions used (see Stability of Thiozolidine Against $NaNO_2$ Oxidation), so 9.6 mg (7 μmol) of fragment Thz-[28-39]-$CONHNH_2$ was converted to a free N-terminal cysteine by treatment with 83 mg (1 mmol) methoxyamine hydrochloride in 5 mL of ligation buffer at pH 4 overnight.[5] Quantitative conversion was observed. The N-to-C assembly shown in Figure 2.3.4A was employed instead of the C-to-N synthesis used when thioesters are accessed directly. Fragment [1-27]-$CONHNH_2$ was oxidized to the C-terminal azide by dropwise addition of 0.1 mL of 200mM aqueous $NaNO_2$ to a solution of 11 mg (3 μmol) purified fragment [1-27]-$CONHNH_2$ in 1 mL ligation buffer at pH 3 and 0°C. The reaction proceeded for 20 minutes at 0°C, and was then quenched by the addition of 172 mg 4-mercaptophenylacetic acid (MPAA) and 34 mg tris(2-carboxyethyl)phosphine·HCl (TCEP·HCl) dissolved in 4.4 mL ligation buffer to give a final concentration of 200mM MPAA and 20mM TCEP (pH 7, RT). To the resulting thioester, 3.4 mL of the solution of crude methoxyamine treated fragment Thz-[28-39]-$CONHNH_2$ were added (4.7 μmol of the peptide). After a two hour RT ligation, one half of the crude reaction mixture was purified, with 2 mg (0.4 μmol) of highly pure material recovered. Of this, 0.6 mg (0.1 μmol) were oxidized by dissolution in 0.1 mL of ligation buffer and dropwise addition of 0.01 mL of 200 mM aqueous $NaNO_2$ at 0° C. The reaction proceeded for 25 minutes, and was then quenched by addition of 3.1 mg MPAA and 0.78 mg TCEP·HCl dissolved in 0.1 mL ligation buffer (pH 7, RT) to give a final concentration of 100mM MPAA and 20mM TCEP. The pH was adjusted to 7 and 0.3 mg (0.15 μmol) of fragment Cys-[40-58]-$CONH_2$ were added to the reaction mixture. After a two hour RT ligation, the mixture was diluted with 0.21 mL buffer P, then a further 0.63 mL buffer P to fold the resulting affibody. The crude mixture was concentrated over a 3 kDa membrane to a final volume of 0.075mL. The crude, folded protein was diluted with 36mg TCEP·HCl in 3mL buffer P (60mM TCEP) and purified to homogeneity (Figure A1.6).

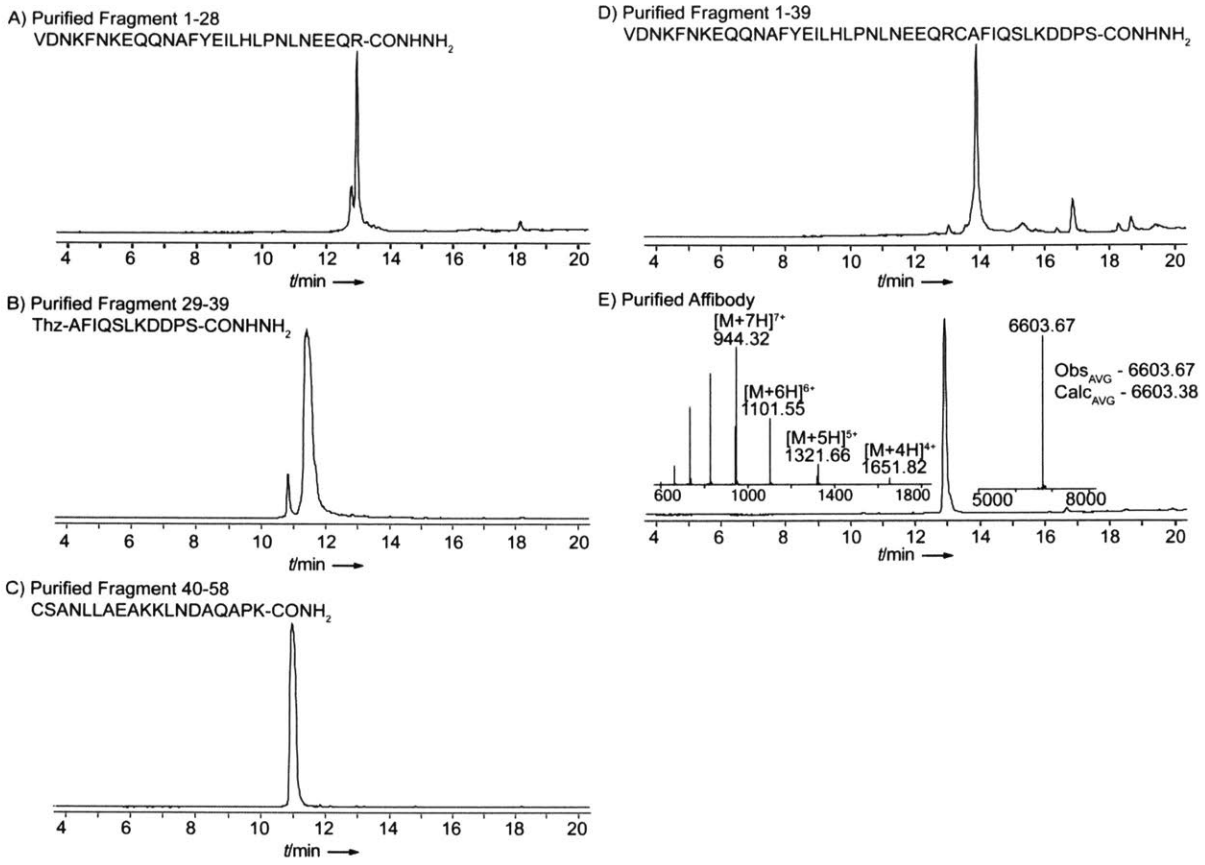


Figure A1.6 Purified affibody synthesis intermediates and final product. **A)** Fragment [1- 27]-CONHNH2. The minor peak is 18Da below desired mass, possibly from formation of aspartamide. **B)** Fragment Thz-[28-39]- CONHNH2. The minor peak is 12Da below the desired mass, likely from minor removal of Thz protection. **C)** Fragment Cys-[40-58]-CONH2 **D)** Ligation fragment [1-39]-CONHNH2. All minor peaks are column contamination. **E)** Final affibody.

A1.2.11 Purification

Crude affibody peptides were dissolved in 95% A/5% B (v/v) and purified on a Waters preparative HPLC with an Agilent Zorbax SB C18 column (21.2 × 250 mm, 7 μm packing), a linear gradient from 5%–45% B in A over 80 min, and a flow rate of 10 mL/min. The crude affibody Fragment 1-39 ligation product was purified on a Beckman System Gold semi-preparative HPLC with a Zorbax C_{18} column (9.4 × 250 mm, 5μm packing), a linear gradient from 10% to 55% B in A over 90 minutes, and a flow rate of 5 mL/min. The final affibody was purified on the same system with the same gradient, using a Jupiter C_{18} column (4.6 x 250mm, 5 μm packing) and a flow rate of 2.3 mL/min.

For all purifications, one minute fractions were collected and screened for the correct mass on a PerSpective Biosystems Voyager-DE MALDI-TOF using 2 μL of the fraction co-crystallized with 2 μL of 50% A'/50% B' (v/v) saturated with alpha-cyano-4-hydroxycinnamic acid matrix. The purity of pooled fractions was confirmed by LC-MS, as above.

A1.2.12 Stability of Thiozolidine Against $NaNO_2$ Oxidation

Thiozolidine was found to be unstable in the oxidation conditions used. After oxidation according the protocol of Fang *et al.*,[3] followed by a 100 mM MPAA and 10 mM TCEP·HCl quench, quantitative conversion to a product 29 Da more massive than expected was observed. This product successfully ligated with an N-terminal cysteine, but showed no change in mass upon exposure to methoxyamine·HCl. No reaction was observed between the mexthoxyamine·HCl treated ligation product and a C-terminal thioester. Fang *et al.* have since reported this limitation, and an analogous protecting group that overcomes it.[6]

A1.2.13 Comparison to Boc in situ Neutralization Synthesis

We synthesized comparable affibody fragments using manual Boc in situ neutralization methods.[1] LC-MS data for these crude peptides is shown in Figure A1.7, next to the crude peptides synthesized on the reported platform. In all cases, the quality was comparable.

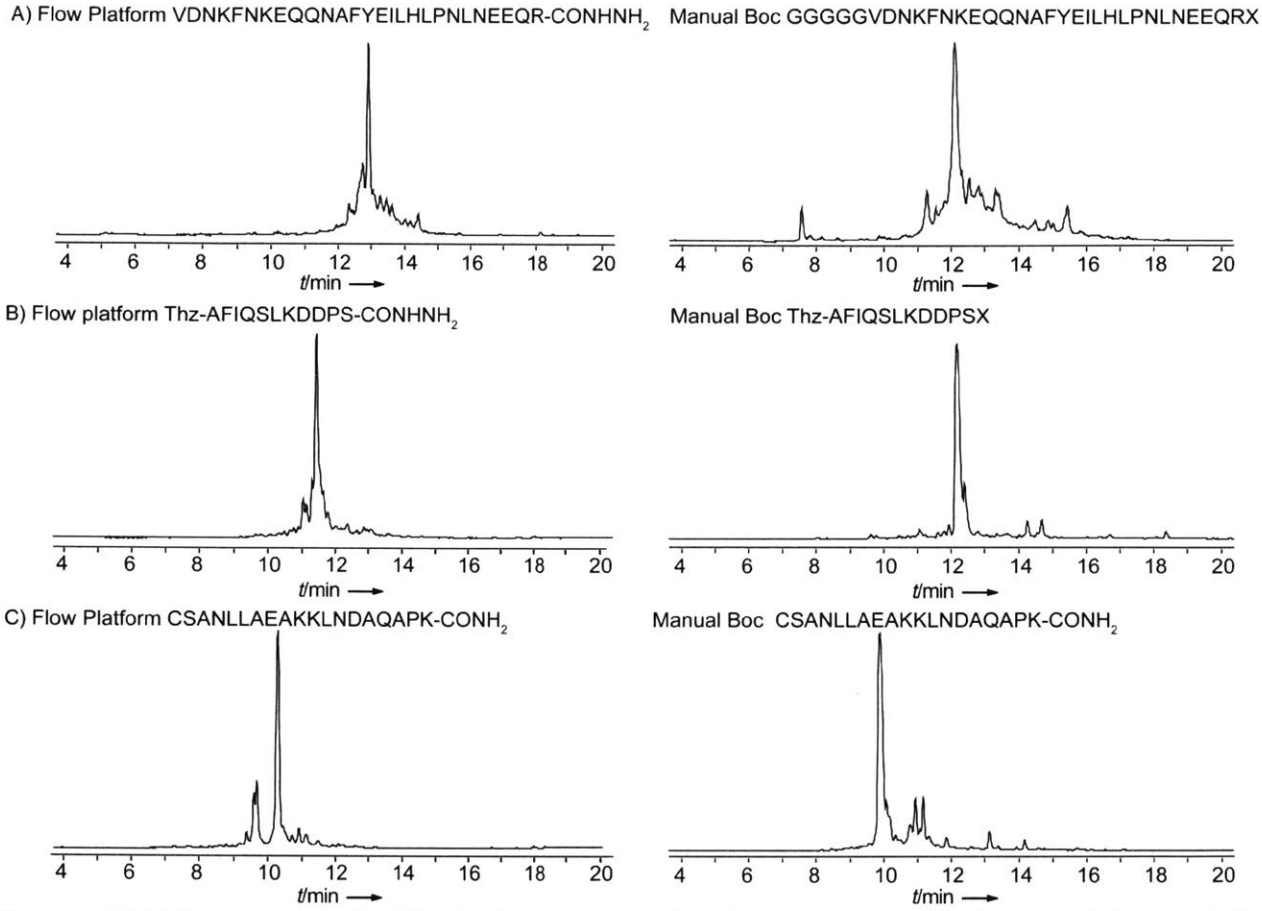


Figure A1.7 Comparison of affibody fragments produced on the reported platform (left) to similar fragments produced with a manual Boc in-situ neutralization method (right). In the Boc peptide sequences, X represents a thioester.[7] (A) N-terminal fragment (B) Middle fragment (C) C-terminal fragment. Crude products from Boc synthesis were analyzed under condition 4.

A1.3 Synthesis with Second Generation Reaction Vessel

A1.3.1 General Conditions for Peptides Synthesis with the Second Generation Vessel:

All peptides in this section, except ACP(65-74) in Figure A1.9c-d were synthesized on the reported flow platform at 60 °C, with reagents heated immediately before use via a coil of tubing in a water bath (see synthesizer design, above). ACP(65-74) shown in Figure A1.9c was synthesized at room temperature on the reported flow platform, and ACP(65-74) shown in Figure A1.9d was synthesized in a standard glass synthesis vessel. In all cases, one synthetic cycle consisted of amide bond formation (coupling), removal of coupling reagent (wash), Nα Fmoc removal (deprotection), and removal of the deprotection reagent and reaction product, piperidine-dibenzofulvene (piperidine-DBF) (wash).

Unless noted, coupling was performed by delivering the following coupling solution at 6 mL/min (for approximately 30 seconds). The coupling solution consisted of 1 mmol of Nα-Fmoc and side chain protected amino acid dissolved in 2.5 mL of 0.4 M HBTU in DMF and 0.5 mL of DIEA.[1] Cysteine was dissolved in 2.5 mL 0.4 M HBTU in DMF and 0.157 mL DIEA. In both cases, amino acids were dissolved in HBTU solution up to several hours before use, and DIEA was added within two minutes of use. Volumetric measurements were made at RT (18-20 °C). The ACP(65-74) shown in Figure A1.9a, and the protease sites in Figure A1.14 was synthesized by substituting HATU for HBTU in the above solution.

Next, the coupling solution was removed with 20 mL of DMF delivered at 20 mL/min over 1 minute, and then the Nα-Fmoc protecting group was removed with 6.6 mL of 50% (v/v) piperidine in DMF delivered at 20 mL/min over 20 seconds. Excess piperidine and piperidine-DBF were removed with 20 mL of DMF delivered at 20 mL/min over 1 minute to complete one cycle.

Peptides were synthesized on 1% divinyl benzene crosslinked polystyrene resin. To produce C-terminal carboxamide peptides, 175 mg of MBHA functionalized resin with a stated loading of 1 mmol per gram was used, and the TFA labile Rink linker was coupled as the first amino acid.[2] To produce C-terminal hydrazide peptides, 200 mg of chlorotrytyl hydrazide (hydrazide) resin, prepared as below, was used.

Non-cysteine containing peptides were cleaved from the resin and side-chain deprotected by treatment with 2.5% (v/v) water and 2.5% (v/v) TIPS in TFA for two hours at RT. Cysteine containing peptides were cleaved from the resin and side chain deprotected with 2.5% (v/v) EDT, 2.5% (v/v) water, and 1% (v/v) TIPS in TFA for two hours at RT. In all cases, the resin was removed and nitrogen was used to evaporate the cleavage solution to dryness at RT. The resulting solids were washed three times with cold diethyl ether, dissolved in 50% A/50% B (v/v), and lyophilized. Side chain protection was as follows: Arg(Pbf), Tyr(tBu), Lys(Boc), Asp(OtBu), Gln(Trt), Ser(tBu), His(Trt), Asn(Trt), Trp(Boc), Glu(OtBu), Thr(tBu), Cys(Trt).

For studies conducted with the second generation synthesis vessel, a different resin was used to generate C-terminal hydrazides, because the Wang resin, used above, was found to promote significant double incorporation of glycine. We were unable to solve this problem, so changed resins. The hydrazide resin was produced as follows: 16 grams of chlorotrytyl chloride resin with a stated loading of 1.2 mmol/gram were suspended in 150 mL of dry, amine free DMF and stirred for 15 minutes. To this, a suspension of 25 mL of DIEA, 50 mL of DMF, and 10 mL of anhydrous hydrazine was added dropwise. During addition, two layers formed and the bottom was added first. After addition was complete, the mixture was allowed to stir for one hour, and was then quenched with 50 mL of methanol. The resin was removed and washed with five 100 mL portions of each of DMF, water, DMF, methanol, and diethyl ether (for a total wash volume of 2.5 L). The resin was then dried for three hours at reduced pressure ($\approx$ 5 torr), resulting in a free flowing powder with lumps. The lumps were gently broken before use, taking care not to generate fines that could clog the reaction vessel's frit.

A1.3.2 Second Generation Synthesis Vessel Construction:

Figure A1.8 shows the larger reaction vessel. The design principles of the first generation translate directly to larger scales. Two problems were encountered, however, when scaling up to a 5/8" OD, ½" ID tube. First, there are no standard 5/8" to 1/16" compression fittings. Second, the minimum distance between 5/8" fittings is quite large, meaning there is a large minimum volume. To overcome the first problem, a Swagelok 5/8" to 3/8" (SS-1010-6-6) fitting followed by a Swagelok 3/8" to 1/16" (SS-600-6-1))fitting was used, but this necessitated a joining length of 3/8" tubing that greatly increased the already large volume of the reaction vessel.

To reduce the reaction vessel volume, a 316SS insert was machined that consisted of a nominal ½" OD segment followed by a 3/8" OD segment with a 1/4" through hole. A 5/8" to 3/8" reducing union was bored out to give a 3/8" through hole, the insert was seated, and the 3/8" ferrule swaged on. After this, the insert could not be separated from the fitting. When installed, the ½" part of this insert-fitting sat in the top of the reaction vessel and limited the volume.

This was effective, but there was still a large volume from the 1/4" through hole. This volume was reduced by inserting a 1/4" OD, 1/8" ID PFA tube and cutting it flush. To further reduce the volume, a 1/8" OD, 1/16" ID PFA tube was inserted by heating and drawing a section of tubing to a narrower diameter, threading it through, and pulling until all tubing in the insert was of the proper diameter. Both sides of the tube were cut flush and the drawn section was discarded. A 3/8" to 1/16" reducing union was installed on the open end of the 3/8" segment to interface with the rest of the system. This insert-fitting is

pictured in Figure A1.8A (left). To prevent the upstream insert fitting from becoming permanently sealed into the tube like the frit, the nominal ½" segment was machined to 0.496" and polished.

A similar piece was machined for the outlet side, with a ½" section the proper length to seat the frit under the ferrule. To prevent all of the solvent from being forced through a small central section of the frit, a 3/8" diameter step 0.05" deep was cut. The bottom of this step tapered to a 1/8" through hole at 31 degrees from horizontal (a standard drill bit taper). A 1/8" OD, 1/16" ID PFA tube was inserted to further limit the volume. The ½" section of the outlet insert positioned the frit and sat largely below the ferrule, so a standard finish was adequate. The one pictured in Figure A1.8A is PTFE, and was installed in a bored through 5/8" to 3/8" reducing union in exactly the same way as the upstream insert. Subsequent reaction vessels, including the one in the cutaway, used stainless steel outlet inserts. A 3/8" to 1/16" reducing union was installed on the open end of the 3/8" segment to interface with the rest of the system.

Tubing was used to limit the internal volume of the inserts, rather than directly making inserts with small holes, to simplify fabrication.

To make reaction vessels with consistent volumes, aluminum spacers were used (Figure A1.8B). A vertical window maintains adequate optical access. The spacer sets the internal volume to 2 mL and enables reproducible assembly of reaction vessels. Furthermore, the spacer helps remove the inlet insert-fitting after synthesis. It prevents the nut from moving down, and instead ejects the insertfitting when the nut it turned. Figure A1.8D shows the assembled reaction vessel, A1.8E shows a cutaway, and A1.8F shows a drawing of the cutaway.

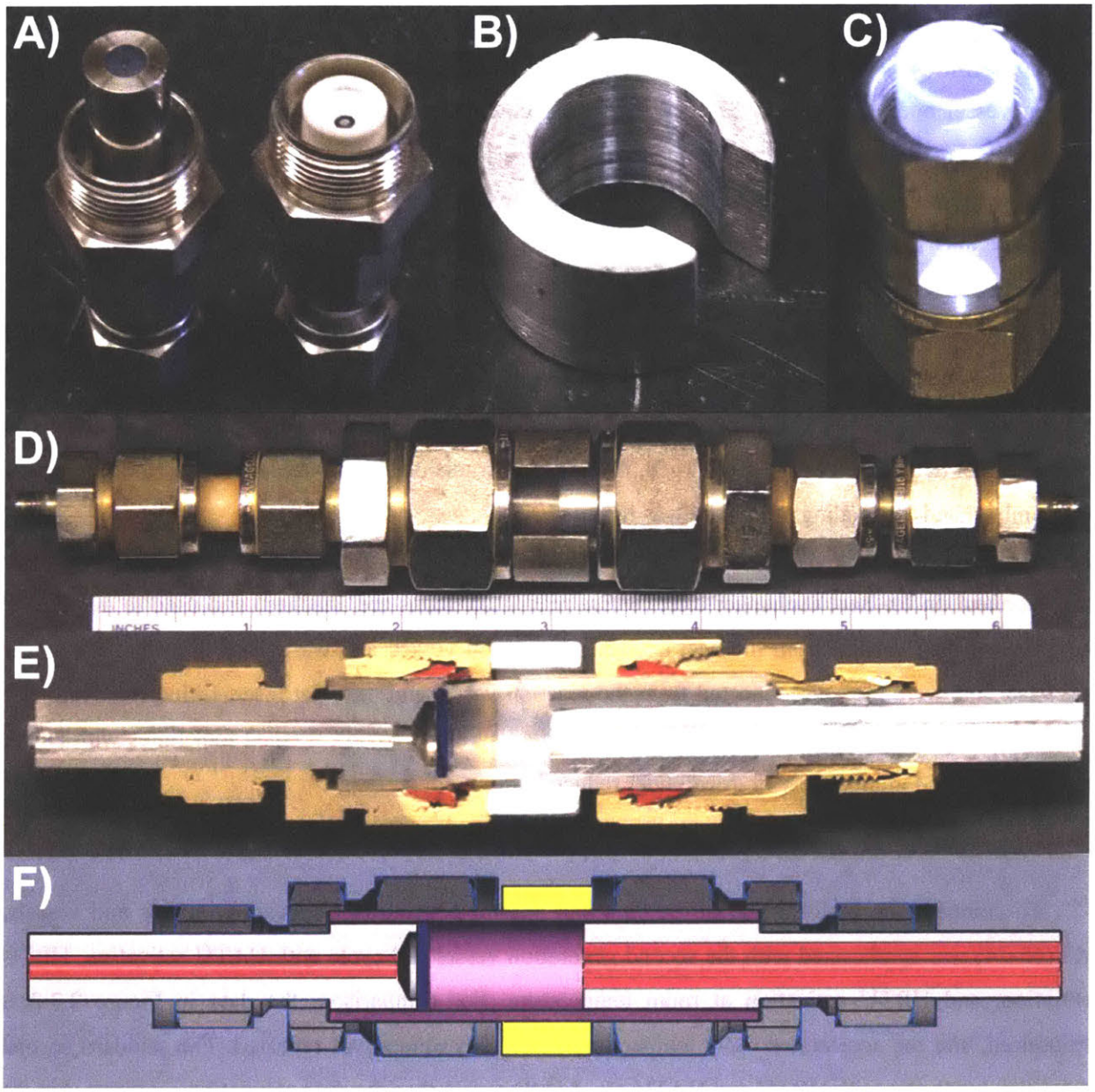


Figure A1.8 Second generation synthesis vessel. **A)** Photograph of the inlet (left) and outlet (right) insert-fittings **B)** Photograph of the spacer **C)** Photograph of the reaction vessel body unit showing the frit as a bright disk in the bottom **D)** Photograph of the assembled reaction vessel **E)** Color enhanced photograph of a sectioned reaction vessel with non-wetted parts in brass, the frit painted blue, the 5/8" ferrules painted red and a single piece of PTFE tubing inserted into the upper insert-fitting rather than the usual two pieces of PFA. The terminal 3/8" nuts are not shown, and background items positioning the cutaway have been digitally removed. **F)** Schematic of the reaction vessel showing the reaction vessel body (purple), frit (blue), spacer (yellow), machined inserts (light grey), and tubing inserted to limit the volume of the inserts (red). The components of the fittings are not individually portrayed and the terminal 3/8" nuts are not shown.

A1.3.3 Reaction Vessel Assembly and Loading

To assemble the reaction vessel, the frit was squarely pressed into one end of the ½" ID PFA reaction vessel body (Mcmaster, 5773K17) and the downstream insert-fitting pressed in behind it to completely seat the frit. The downstream insert-fitting was then installed as a regular fitting. Next, the spacer was placed on top of the outlet insert-fitting, the inlet nut placed on top of that, and the tube cut to length. The inlet insert-fitting was then installed as a regular fitting. To load the reaction vessel, the inlet insert-fitting was removed, the resin was added dry, and the vessel filled with methanol. The inlet insert-fitting was then attached to the heat exchanger and purged with methanol. Without removing it from the heat exchanger, the purged inlet insert-fitting was installed, causing excess methanol to exit the vessel through the outlet insert-fitting. If there was a large bubble, the reaction vessel was turned upside down (with the inlet below the outlet) and purged. If this failed to dislodge the bubble, the reaction vessel was disassembled and the loading procedure was repeated.

To remove the resin after synthesis, a syringe filled with 10 mL of air was attached to the luer-lock quick connect on the heat exchanger inlet (where reagent syringes are attached) and used to deliver the air. This removed solvent from the heat exchanger, reaction vessel, and waste line. The heat exchanger and waste line were then disconnected from the reaction vessel and the inlet insert-fitting removed. The resin was suspended in DCM and decanted into a fritted syringe (Toviq), washed four times with DCM, and either cleaved immediately or dried under reduced pressure for storage.

A1.3.4 Synthesis of ACP(65-74)

To compare the performance of the first and second generation synthesis protocols and vessels, ACP(65-74) was synthesized with the second generation vessel and cycle with HATU activation, HBTU activation, and HBTU activation at room temperature. For comparison, the data in Figure 2.3.2 is reproduced, and the accelerated room temperature batch experiment was repeated. The standard second generation syntheses with HBTU and HATU are slightly worse than the first generation synthesis, but the second generation RT control has significantly reduced co-eluting Gln deletion product. The performance of the two protocols is different, but comparable. Overall, the second generation system is preferred for its reduced cycle time and vastly improved mechanical reliability.

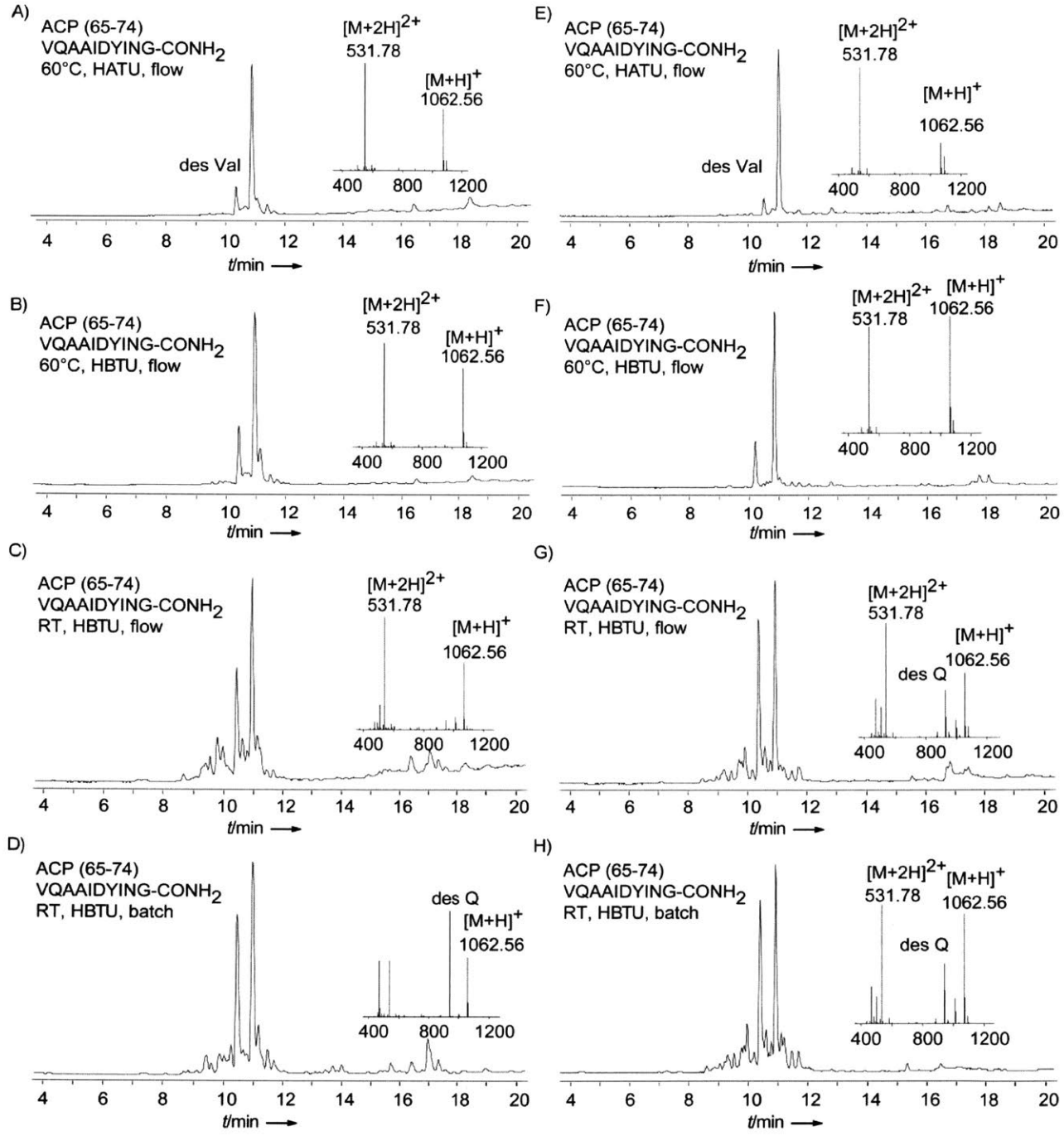


Figure A1.9 LCMS data for ACP(65-74) model studies. Crude LCMS chromatograms for ACP(65-74) synthesized with the second generation protocol at **A)** 60 °C using HATU as an activator, **B)** 60 °C using HBTU as an activator, and **C)** RT using HBTU as an activator. **D)** RT synthesis using a comparable manual batch method. For comparison, ACP(65-74) synthesized with the first generation protocol at **E)** 60 °C using HATU as an activator, **F)** 60 °C using HBTU as an activator, **G)** RT using HBTU as an activator, and **H)** RT using a comparable manual batch method, are shown. The total ion current is displayed in each chromatogram.

A1.3.5 Synthesis of HIV-1 PR (81-99) and PnI(A10L) Conotoxin

To further compare the performance of the first and second generation synthesis protocols and vessels, the syntheses of the HIV-1 protease fragment and the PnI(A10L) Conotoxin were repeated. The data is shown in Figure A1.10, along with the data from Figure 2.3.3 for comparison. The performance of the two protocols is comparable.

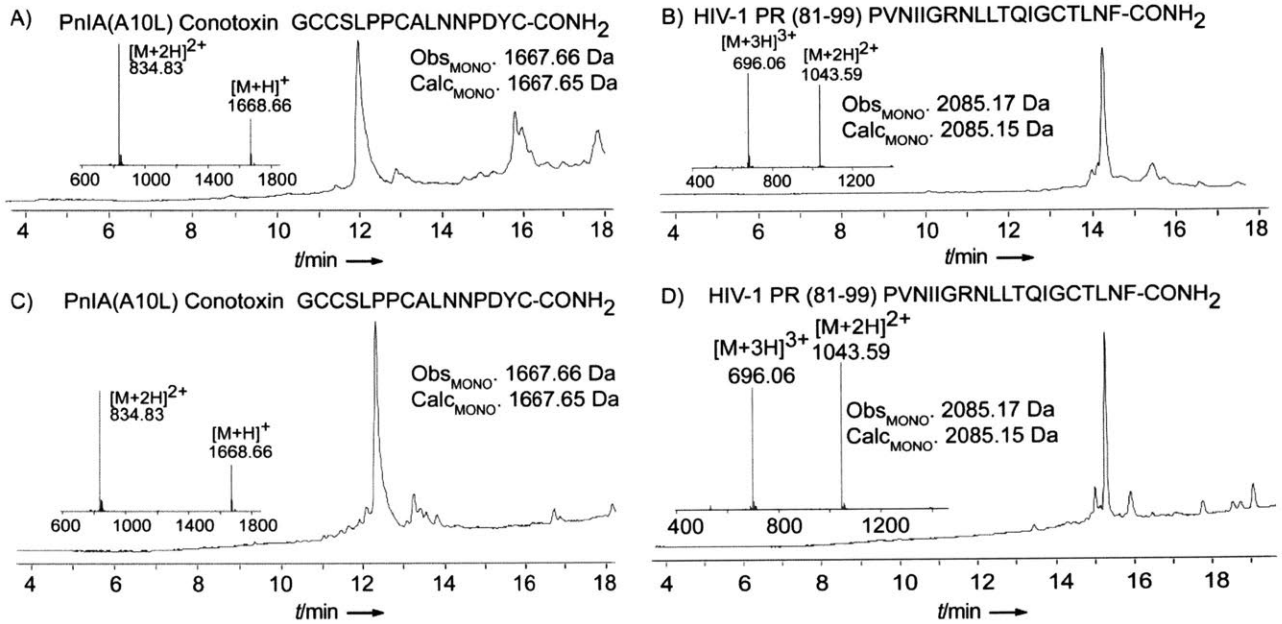


Figure A1.10 Resynthesis of HIV-1 protease fragment and PnI(A10L) Conotoxin. **A)** Conotoxin synthesized with the second generation vessel and conditions. Material eluting after 15 minutes is non-peptidic. **B)** HIV-1 PR (81-99) synthesized with the second generation vessel and conditions. **C)** Conotoxin synthesized with the first generation vessel and conditions. **D)** HIV-1 PR (81-99) synthesized with the first generation vessel and conditions. Late eluting material is side chain protected products. Retention times differ slightly because the LC/MS column was replaced between first and second generation reactor studies.

A1.3.6 Synthesis of Affibody Fragments

To validate the performance of the chlorotrytyl hydrazide resin and further explore synthesis with the second generation reaction vessel and cycle, the synthesis of the affibody fragments was repeated. The data is shown in Figure A1.11, along with the data from Figure 2.3.4 for comparison. The performance of the two linkers is comparable when synthesizing fragments without glycine. The modified Wang resin promotes significant double incorporation of glycine, which motivated our move to hydrazide resin. In both cases, pure material suitable for ligation was obtained after preparative chromatography.

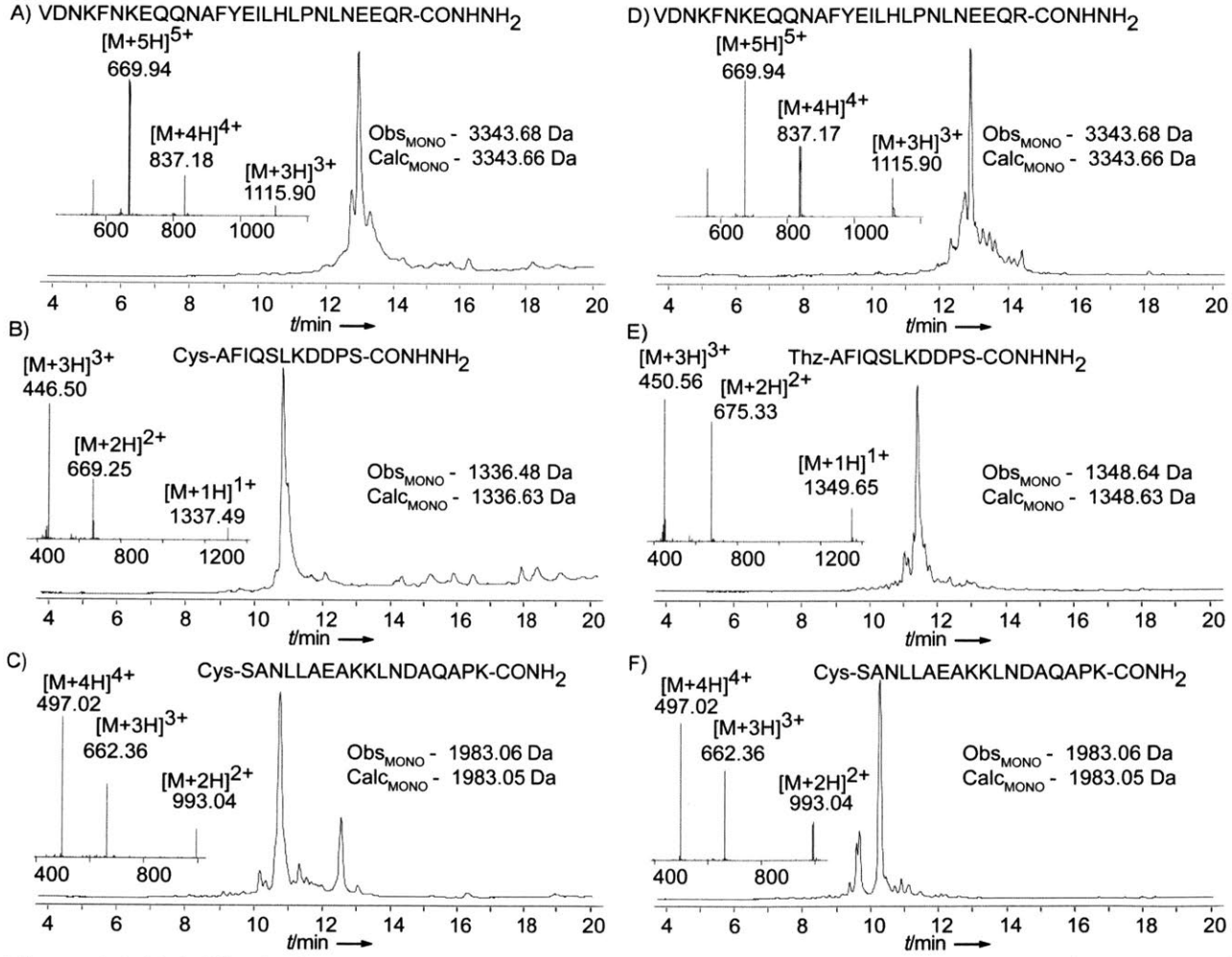


Figure A1.11 Affibody fragments. **A-C:** Fragments synthesized on chlorotrytyl hydrazide functionalized polystyrene with the second generation protocol. **D-F:** Fragments synthesized on modified Wang resin with the first generation protocol.

A1.3.7 Synthesis of Glutathione Analogue Library

In support of other work in our group, the library of glutathione analogues in Figure A1.12 was synthesized. These were all produced in one day, and afforded highly pure material after preparative chromatography (purified material is not shown).

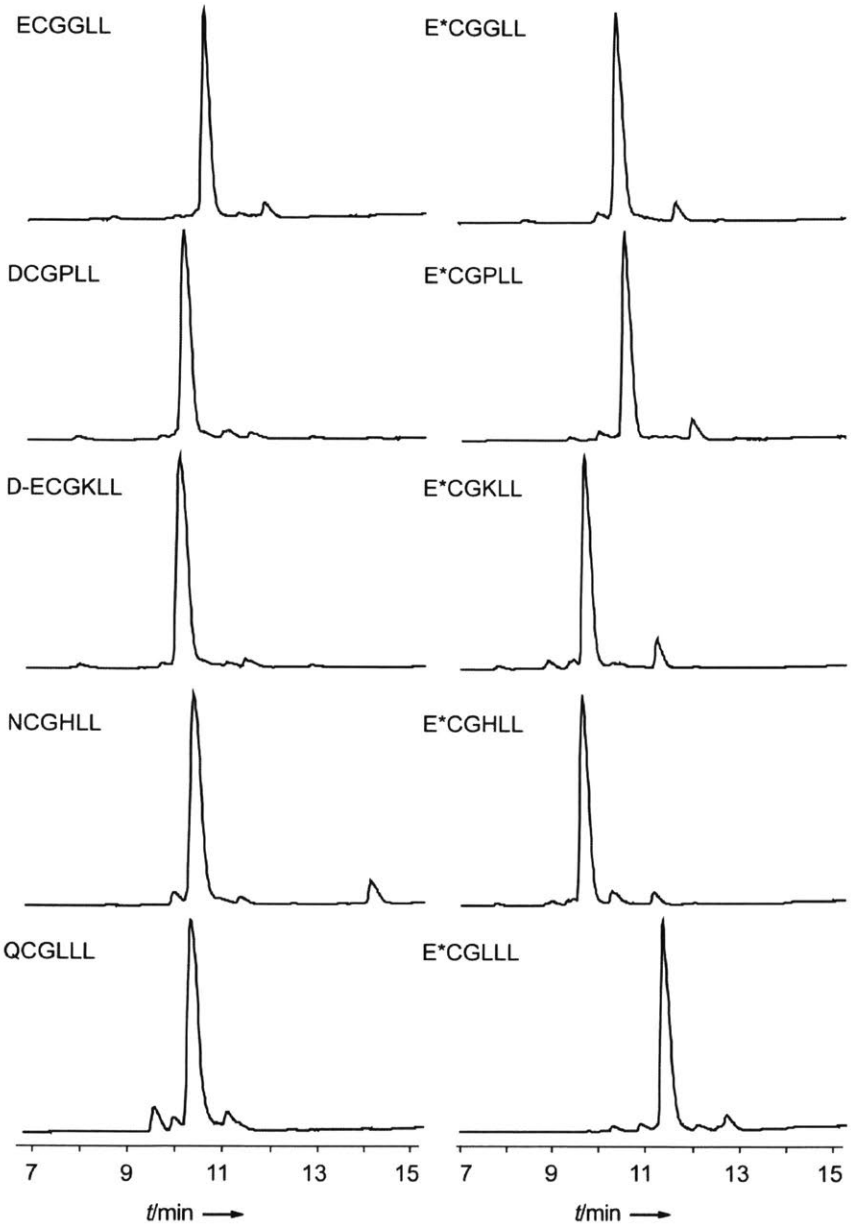


Figure A1.12 A library of glutathione analogues synthesized in 18 minutes each and used as non-natural substrates for Glutathione-S-Transferase.[9] E* represents glutamic acid coupled to the polypeptide through the side chain carboxylate. All peptides were analyzed under condition 3, and the main peak has the desired mass.

A1.3.8 Synthesis of Cysteine Rich Peptides

In support of other work in our group, the cysteine rich peptides shown in Figure A1.13 were synthesized. All cysteine was acetamidomethyl (Acm) protected, to prevent side reactions during cleavage and side chain deprotection. Cysteine in these peptides was activated with 0.190 mL of DIEA rather than 0.157 mL. Highly pure material was isolated after preparative chromatography (purified material is not shown).

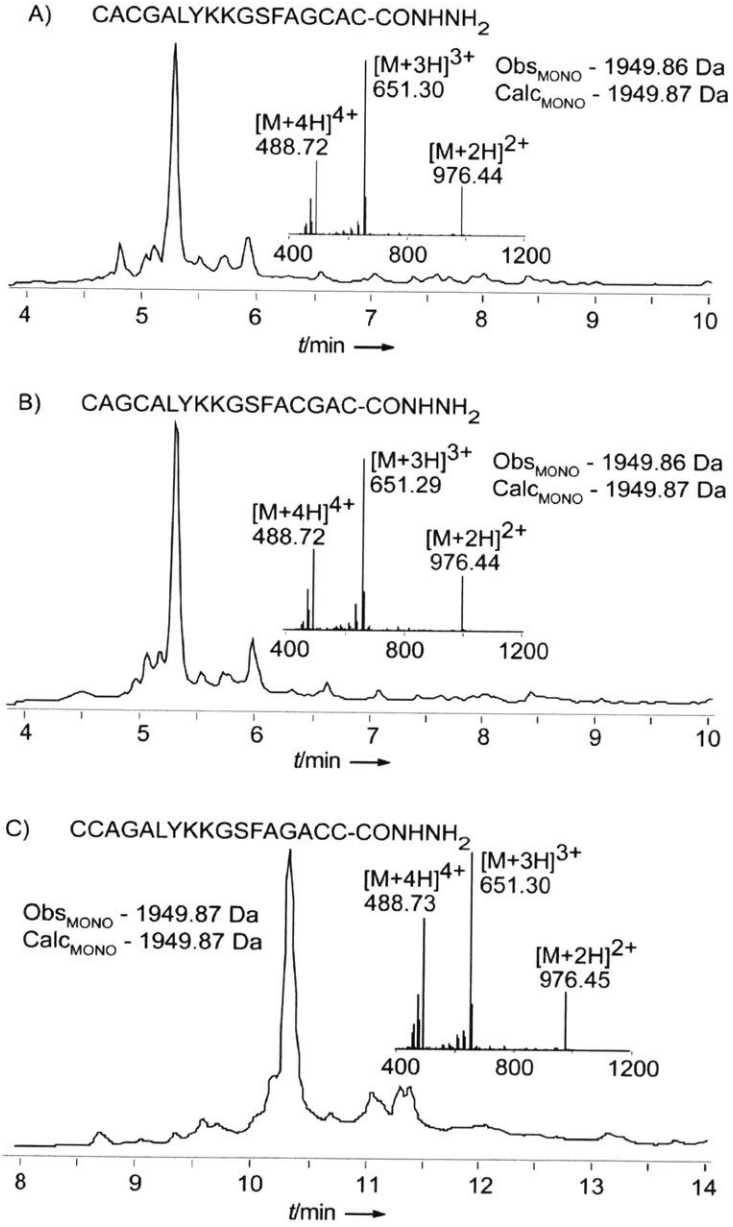


Figure A1.13 Several cysteine rich peptides. All cysteine was Acm protected. **A)** and **B)** were analyzed under condition 3; **C)** was analyzed under condition 4.

A1.3.9 Synthesis of Biotinylated Peptides Sites

In support of other work in our group, the biotinylated peptides shown in Figure A1.14 were synthesized. Highly pure material was isolated after preparative chromatography (purified material is not shown).

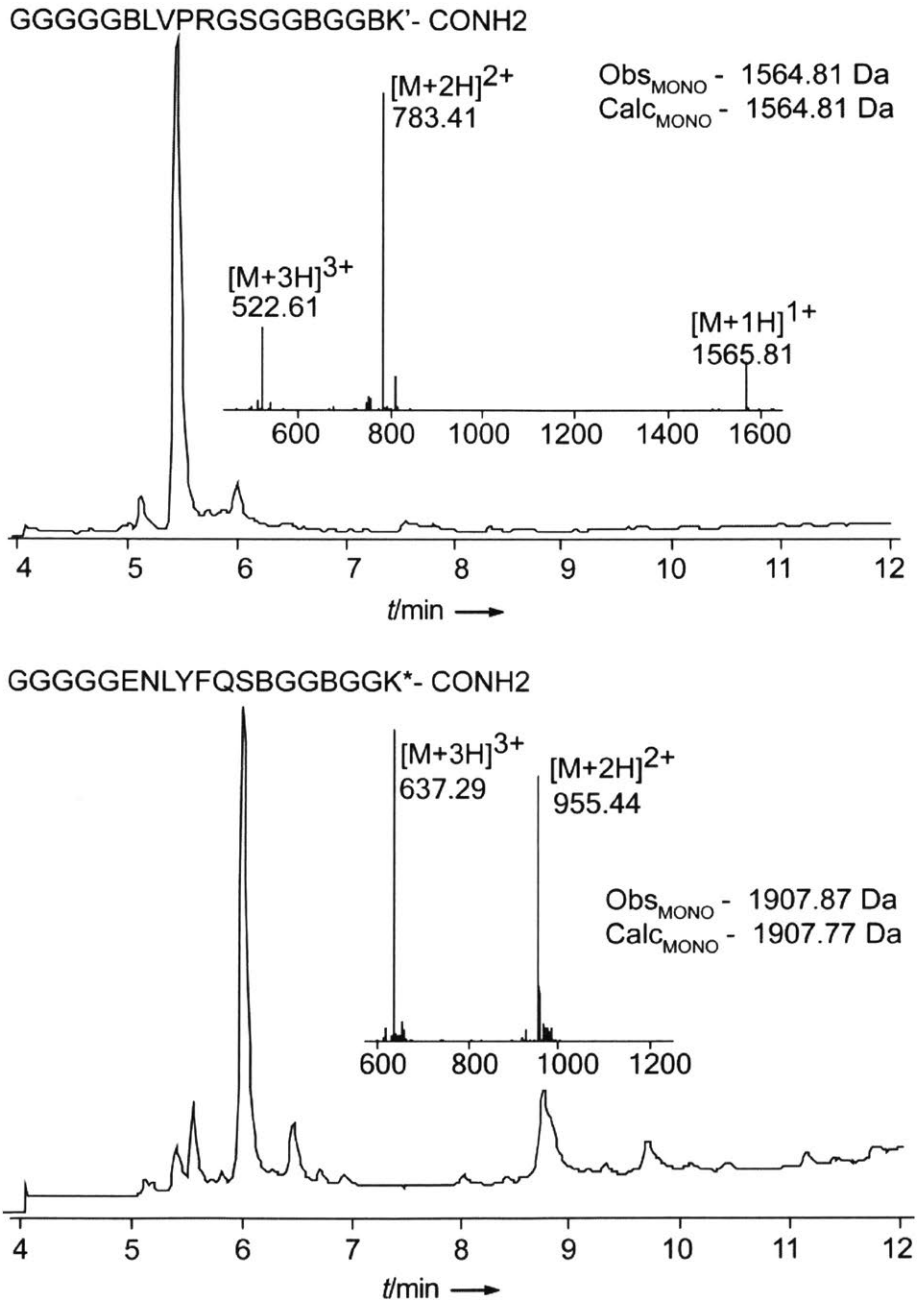


Figure A1.14 Biotinylated protease recognition sites. B is β-alanine, K* is biotinylated lysine, and K' is alloc protected lysine. The alloc protected peptide was hydrogenated in batch on-resin to give free lysine, and biotin was coupled in batch. Material eluting after 8 minutes is non-peptidic.

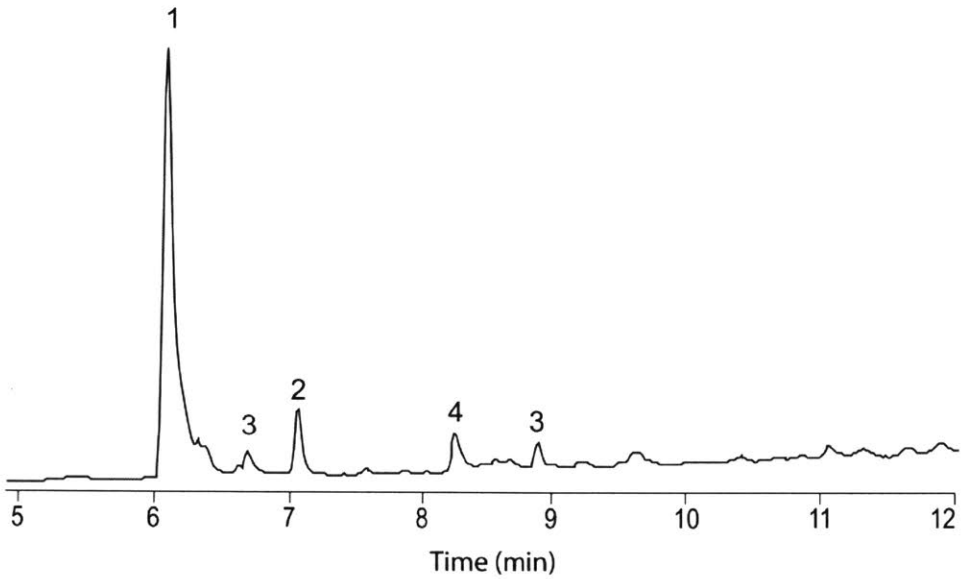
A1.4 Estimation of Timescale of Amide Bond Formation

We studied the timescale of amide bond formation between Fmoc-Leucine and resin bound FALFA using 0.3 M amino acids and HBTU activation. To estimate the timescale of amide bond formation, FALFA bound to chlorotrytyl hydrazide resin was prepared and dried under vacuum. A test cleavage and LC/MS analysis under condition 3 revealed the FALFA-$CONHNH_2$ to be of high purity. Fifteen 5 mg portions of this resin in Ependorff tubes were pre-swollen for > 15 minutes in 1 mL of DMF, centrifuged, and the excess DMF decanted. Next, a coupling solution consisting of 0.2 mmol (40 eq) of Fmoc-Leu dissolved in 500 μL of 0.4 M HBTU in DMF and preactivated with 100 μL of DIEA for 2 minutes was added to the resin in the Ependorff tubes. The reaction was allowed to proceed for 5, 10, 15, 25, or 35 seconds, and was then quenched with 200 μL of anhydrous hydrazine. A 45 second control showed the reaction to be essentially complete, and a pre-quenched control, with hydrazine added to the coupling solution before it was added to the resin, showed no product (Figure A1.16). After each trial, the resin was removed from the Ependorff tube, washed once with DMF, three times with DCM, and dried under vacuum. Each time trial was performed in triplicate. Each trial was cleaved for two hours at room temperature with a mixture of 95% TFA, 2.5% H_2O and 2.5% TIPS. Hydrazine was selected to quench the reaction because it is a strong nucleophile but a weak base.

The crude cleavages were analyzed on an Agilent 1260 analytical HPLC under chromatographic condition 3, and the absorbance at 214 nm was integrated for the product and starting material. To determine the mass of FALFA and Fmoc-LFALFA represented by the observed integrations, standard curves were prepared. Two samples each of 0.5 mg/mL purified FALFA and Fmoc- LFALFA were prepared. Each sample was injected 5 times, at 4, 12, 20, 30, and 40 uL injection volume, to generate a standard curve.

For 10, 15, 25, and 35 second time points, FALFA levels were too low to quantify from the same HPLC run as Fmoc-LFALFA, so a second run, with ten times more material injected, was performed to quantify the amount of FALFA (Fmoc-LFALFA was well above the linear range in the second run). The data were taken to be pseudo-first order and fit to the curve yield = $1\text{-}e^{-At}$. The 95% confidence interval for parameter A was calculated to be 0.147 ± 0.019 s^{-1}. This corresponds to a reaction half-life of 4.7±0.6 seconds. Therefore, this reaction is expected to be 99.9% complete in less than one minute at RT. When performed at 60 °C, under otherwise identical conditions, the reaction was found to be complete in 10 seconds (Figure A1.18), as expected.

A) Total Ion Chromatogram of Crude FALFA



B) Total Ion Chromatogram of 45 Second Reaction

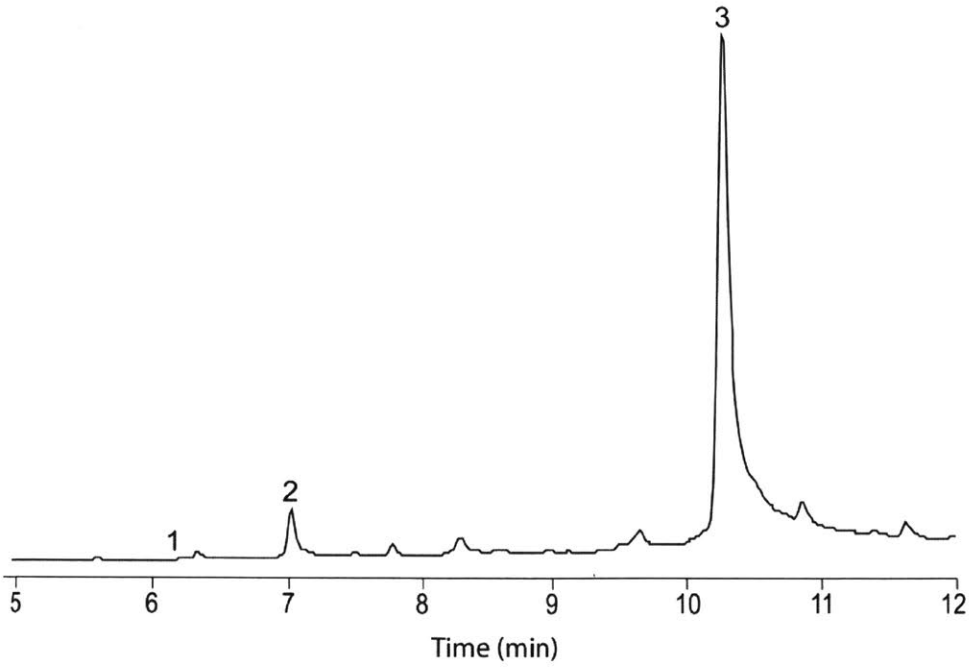


Figure A1.16 A) FALFA-CONHNH2 starting material. 1 = FALFA 2 = Trifloroacetylated FALFACONHNH2 3 = unknown, possibly column contamination 4 = non-peptidic. **B)** Reaction products after 45 second coupling of Fmoc-Leu at RT. Starting material, 1, is below the TIC baseline. 2 is minor LFALFA-CONHNH2 (des-Fmoc). 3 is desired product, Fmoc-LFALFA-CONHNH2.

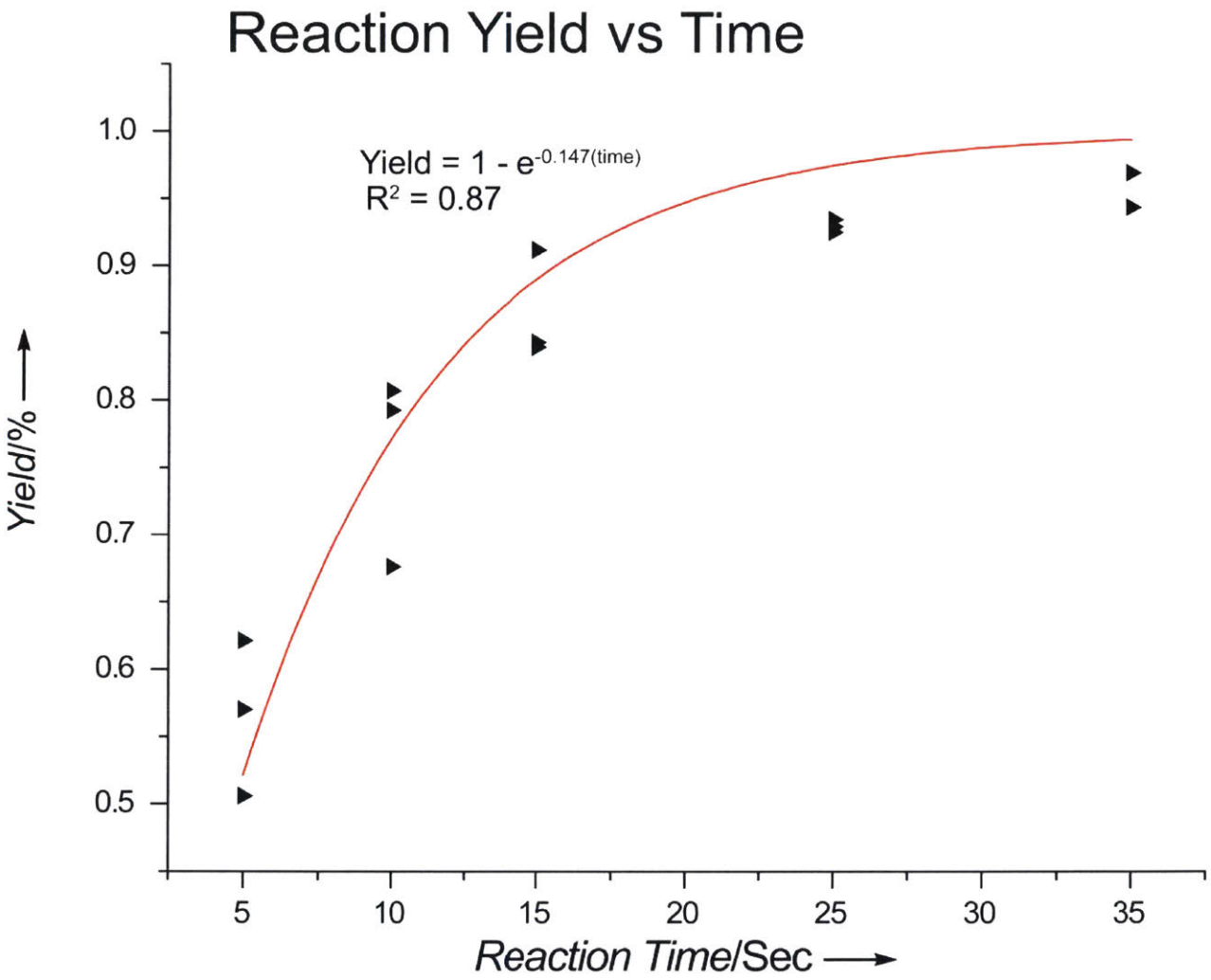


Figure A1.17 Reaction yield vs. time for the batch mode coupling of Fmoc-leu to resin bound FALFA. The curve was fit with Origin software.

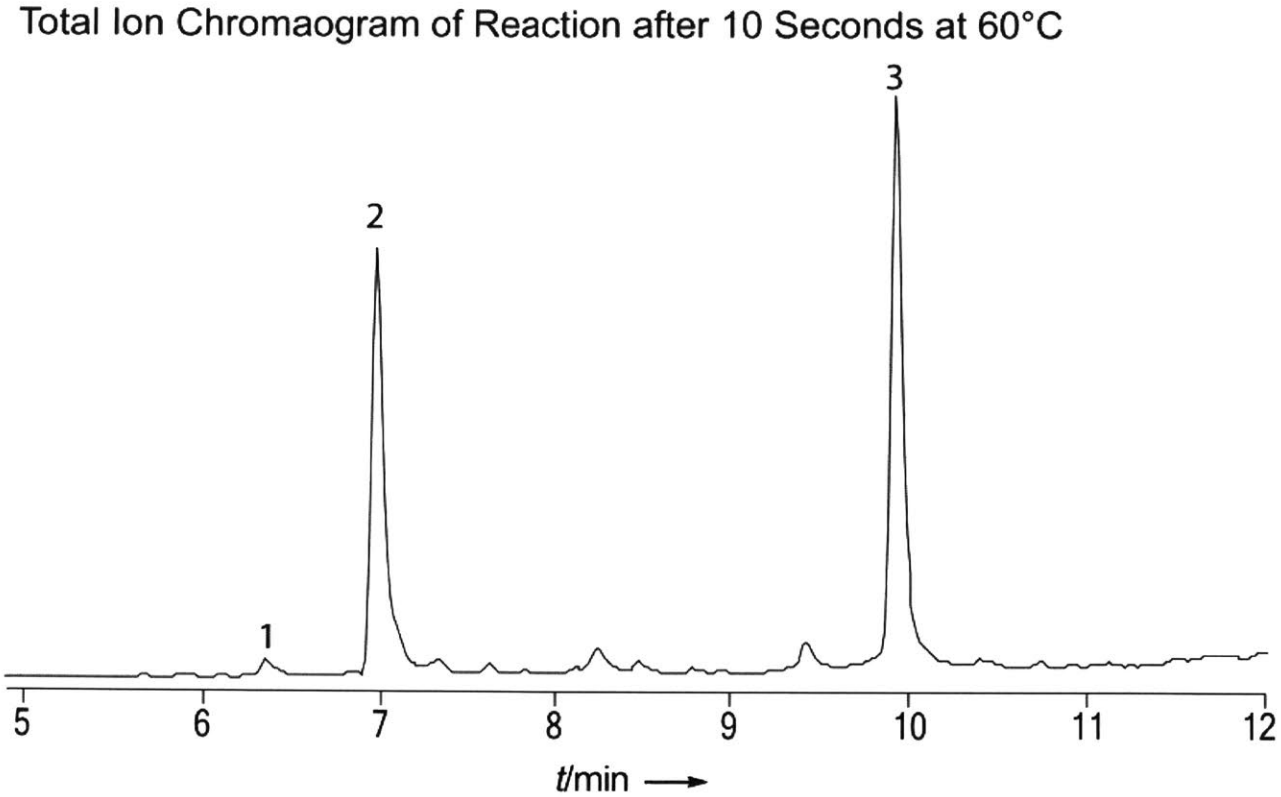


Figure A1.18 Total ion chromatogram of the batch mode coupling of Fmoc-Leu to resin bound FALFA after 10 seconds at 60 °C. Starting material, 1, is just above the TIC baseline. 2 is LFALFA-$CONHNH_2$, with Fmoc removed during the hydrazine quench, and 3 is the desired Fmoc-LFALFA-$CONHNH_2$.

A1.5 References for Appendix 1

(1) M. Schnolzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pept. Res. Ther.* 2007, *13*, 31-44.

(2) W.C. Chan, P.D. White, *Fmoc solid phase peptide synthesis : a practical approach*, Oxford University Press, New York, 2000.

(3) G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li, Y. Lin, H. K. Cui, L. Liu, *Angew. Chem.* 2011, 123, 7787-7791; *Angew. Chem. Int. Ed.* 2011, 50, 7645-7649.

(4) Y. Han, F. Albericio, G. Barany, *J. Org. Chem.* 1997, *62*, 4307-4312.

(5) D. Bang, S. B. Kent, *Angew. Chem.* 2004, *116,* 2588-2592; *Angew. Chem. Int. Ed.* 2004, *43*, 2534-2538.

(6) G. M. Fang, J. X. Wang, L. Liu, *Angew. Chem.* 2012, *41,* 10493-10496; *Angew. Chem. Int. Ed.* **2012**, *51*, 10347-10350.

(7) T. M. Hackeng, J. H, Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068-10073.

(8) Y. X. Han, F. Albericio, G. Barany, *J. Org. Chem.* **1997**, 62, 4307.

(9) C. Zhang, A. M. Spokoyny, Y. Zou, M. D. Simon, B. L. Pentelute, *Angew. Chem.* 2013, *52,*14251-14255; *Angew. Chem.Int. Ed.* 2013, *52*, 14001-14005

Appendix 2: Automated fast flow synthesis

A2.1 Materials

2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 2-(7-Aza-1Hbenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU), and Nα-Fmoc and side chain protected D-amino acids were from Chem-Impex International, IL. Nα-Fmoc and side chain protected L-amino acids were from Advanced ChemTech, KY. Side chain protection was as follows: Arg(Pbf), Asn(Trt), Asp(OtBu), Cys(Trt), Gln(Trt), Glu(OtBu), His(Trt) or His(Boc), Lys(Boc), Ser(tBu), Thr(tBu), Trp(Boc), Tyr(tBu).

Rink amide-ChemMatrix and HMPB-ChemMatrix polyethylene glycol resins were purchased from Matrix Innovation (35-100 mesh, 0.49mmol/g). Chlorotrityl hydrazide and aminomethyl polystyrene resins were prepared as described previously by Mong et al[1].

N,N-Dimethylformamide (DMF), dichloromethane (DCM), diethyl ether, and HPLC-grade acetonitrile were from EMD Millipore. Acetonitrile for LC-MS were purchased from Sigma-Aldrich. Water for LC-MS was purified with a Millipore water purifier. All other reagents were purchased from Sigma-Aldrich. N-Methyl pyrolidone (NMP) and N,N' dimethylimidazolidinone (DMI) were purchased anhydrous from Sigma-Aldrich and repurified to remove water and amines on a Pure Process solvent purification system, as was DMF used to dissolve amino acids and activators for use with the third generation AFPS. Other solvents were not anhydrous, and all other material was used without further purification.

Common solvent mixtures used throughout these experiments are: 0.1% (v/v) TFA in water (A), 0.1% (v/v) Formic acid in water (A'), 0.1% (v/v) TFA in acetonitrile (B), and 0.1% (v/v) formic acid in acetonitrile (B').

A2.2 LC-MS Analysis

All peptides and proteins were analyzed on an Agilent 6520 Accurate Mass Q-TOF LC-MS under the following condition, unless noted below: An Agilent C3 Zorbax 300SB column (2.1 x 150 mm, 5 μm packing) was used with a flow rate of 0.8mL/min of the following gradient: A' with 5% B' for 2 min, 5-65% B' ramping linearly over 9 min, and 65% B' for 1 min.

FHL was analyzed on an Agilent 6520 Accurate Mass Q-TOF LC-MS under the following condition: An Agilent C18 Zorbax 300SB column (2.1 x 150 mm, 5 μm packing) was used with a flow

rate of 0.8mL/min of the following gradient: A' with 1% B' for 2 min, 1-31% B' ramping linearly over 10 min, and 31% B' for 1 min.

LL-37 was analyzed on an Agilent 6550 Q-TOF with an Agilent C3 Poroshell 300SB column (1.0x75mm, 5μm packing) and a flow rate of 0.4mL/min of the following gradient: 1% B' in A' for two minutes, then 1% B' in A' ramping linearly to 61%B' in A' over 10 minutes.

ACP(65-74) shown in Figure A2.2 was analyzed on an Agilent 6520 Accurate Mass Q-TOF LC-MS under the following condition: An Agilent C3 Zorbax 300SB column (2.1 x 150 mm, 5 μm packing) was used with a flow rate of 0.4mL/min of the following gradient: A' with 1% B' for 3 minutes, 1-61% B' ramping linearly over 15 minutes, and 61% B' for 4 minutes.

CPPs 12, 20, and 84 were analyzed on an Agilent 6520 Accurate Mass Q-TOF LC-MS under the following condition: An Agilent C18 Zorbax 300SB column (2.1 x 250 mm, 5 μm packing) was used with a flow rate of 0.4mL/min of the following gradient: A' with 1% B' for 5 min, 1-61% B' ramping linearly over 17 min, and 61% B' for 3 min.

CPP 39 was analyzed on an Agilent 6520 Accurate Mass Q-TOF LC-MS under the following condition: An Agilent C3 Zorbax 300SB column (2.1 x 150 mm, 5 μm packing) was used with a flow rate of 0.8mL/min of the following gradient: A' with 5% B' for 3 min, then 5-85% B' ramping linearly over 15 min.

A2.3 Automated LC/MS data processing

Agilent's molecular feature extraction utility was used to automatically process the large volume of LC/MS data generated throughout these studies. Molecular features were extracted, and the most prevalent feature was taken to be the desired product for the purpose of the below analysis. The relative peak volume of the desired product (percentage of assigned ions assigned to the particular peak), is the "estimated purity". Any ions that were assigned to molecules less than 90% of the mass of the desired product were assumed to be truncates, and the sum of peak volumes of such features is the "truncates" parameter. That is, if a chromatogram has 15% "truncates" it means that 15% of the ions assigned by Agilent's MFE utility were assigned to products less than 90% of the mass of the most prevalent product.

This analysis is helpful to understand trends in the data, but is not strictly quantitative. The estimation of purity does not include salts, solvents, protecting groups, or other small molecules in the sample, and is therefore expected to be strictly higher than isolated yields following preparative RP-HPLC. It is in good agreement, however, with HPLC yields calculated from integration of the crude total ion chromatogram.

The MFE parameters were set as follows: To be assigned, compounds were required to have a peak height >0.1% of the greatest peak height and at least 5000 counts, and charge state assignment was with a tolerance of 0.05Da plus 7ppm. The count threshold was set to avoid picking up column background.

A2.4 Automated UV data processing

Digitally recorded UV data was processed off line to gain more insight into synthesis than visual inspection alone, and the tools to do so are included as an additional electronic supplement.

The large absorbance of HATU and released HOAt overlaps with the Fmoc absorbance and makes it impossible to meaningfully monitor the coupling reaction by UV. It is possible, however, to monitor the deprotection reaction. To prevent the detector from saturating, the UV absorbance was monitored at 304nm, away from the absorbance maximum of Fmoc. In this way, UV data like that shown in Figures 10 and S1 was collected.

There were two interesting properties of the deprotection peaks: their area, which corresponded to the total amount of Fmoc removed, and their width, which correlated with the rate of Fmoc removal. A decrease in peak area that recovered in the next steps indicated a single or multiple deletion. A decrease in peak area that did not recover but remained constant was caused by a truncation of the peptide chain. A steady decline in peak area indicated that the linker was not stable and the peptide was slowly being released from the resin. This behavior was observed, for example, with chlorotrityl chloride esters, but not with the HMPB ester used to produce LL-37 as a peptide acid.

Increases in peak width indicated that Fmoc removal, and likely amide bond formation, were slower than usual, and often anticipated deletions below the detection limit of the peak area change. This was the case for the coiled coil synthesis in Figure 12, which was erroneously assumed to have failed at 90°C based on inspection of peak heights. Such erroneous conclusions based on casual inspection of the UV data were common and motivated the production of automated tools to extract quantitative data. An integrator was written in Matlab and identified deprotection peaks, then output their peak area, peak width, and peak width at half maximum. Most people in our lab were much more comfortable using excel than Matlab, so the output was written to an excel file, and the examples below are plotted in excel, from the output file.

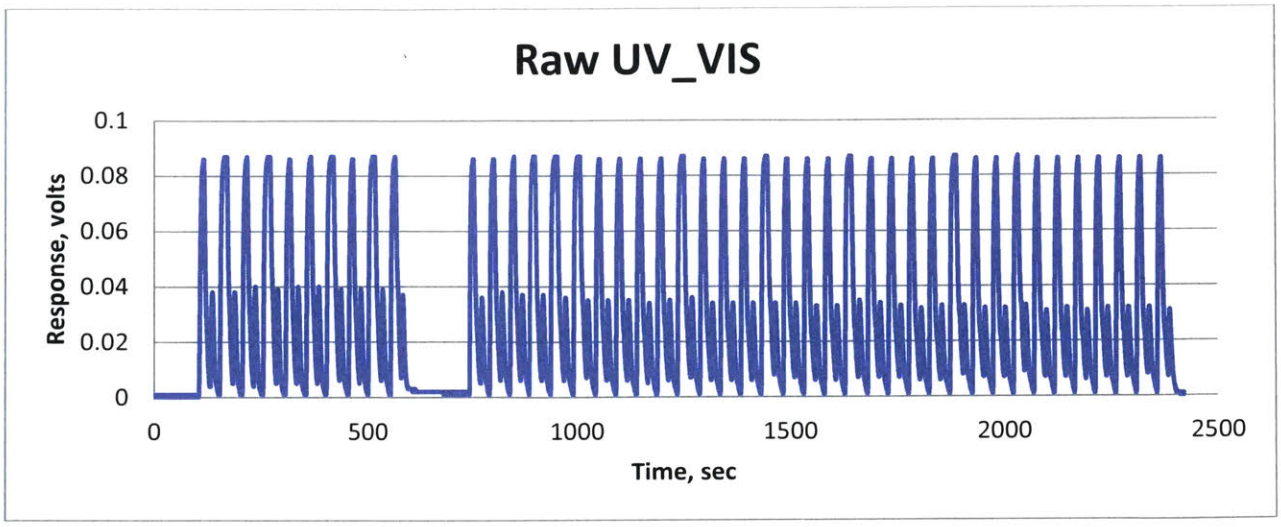


Figure A2.1: Raw UV data processed below. The synthesis was paused after 10 couplings, then resumed for another 34 couplings.

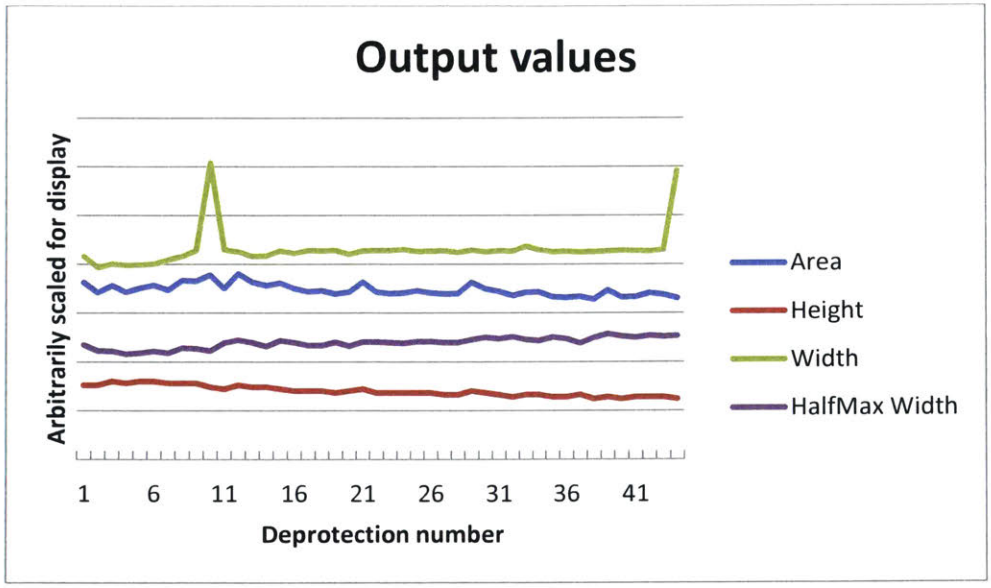
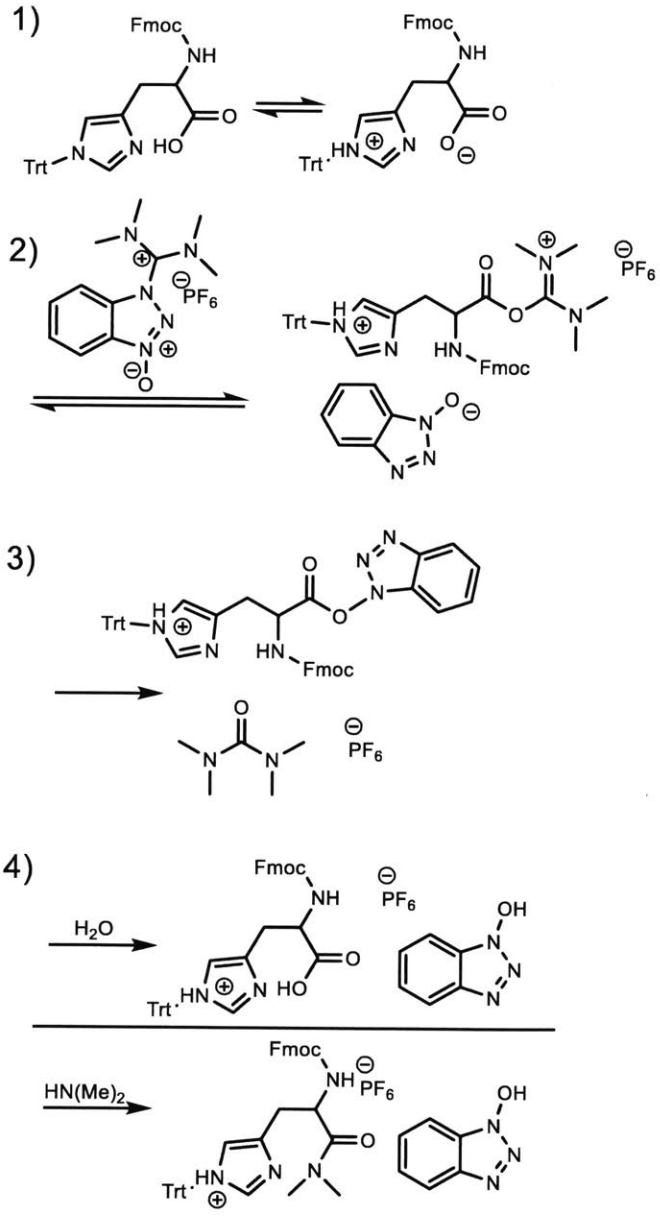


Figure A2.2: Peak area, height, width, and width at half maximum of the deprotection peaks in Figure A2.1. Notice that peak height declines significantly, but peak width increases, and peak area remains almost constant. The very wide 10^{th} and 44^{th} peaks are a failure of the program to detect the end of the peak properly because the run was paused (after the 10^{th} deprotection) or ended (after the 44^{th}). The Half maximum width, height, and area are correct, illustrating the utility these parameters.

A2.5 Histidine degradation scheme



Scheme A2.1 Proposed degradation mechanism of solutions of Fmoc His(Trt) OH and HBTU. 1) The trityl protected imidazole deprotonates the carboxylic acid 2) The carboxylate reacts with HBTU, forming an isoacyl urea and releasing HOBt. 3) HOBt attacks the isoacyl urea, irreversibly forming tetramethyl urea and an OBt ester. 4) The OBt ester degrades by reaction with trace water or dimethyl amine in DMF, liberating HOBt and accounting for the observed colour change.

A2.6 High speed videography

A Sony RX10 II was used to capture dye experiments at 960 frames per second. The resulting video was analyzed and enhanced with the Adobe Premiere software suite and VNC media player. Color, contrast, and brightness were enhanced, and individual frames were cropped.

A2.7 Photographs from figure 3.5.2 without hue specific enhancement

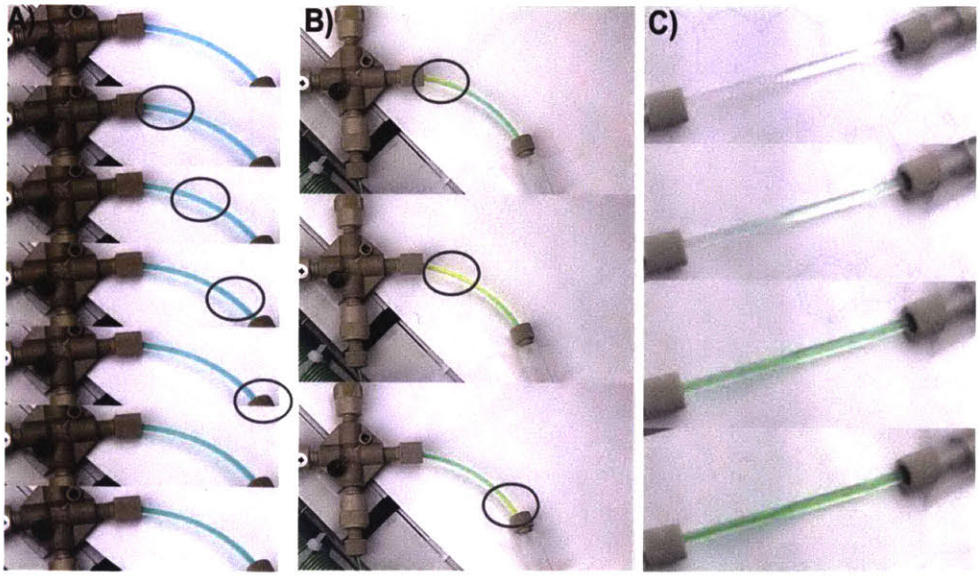


Figure A2.4: Images shown in figure 3.5.2 without hue-specific enhancement. The images here were only brightness and contrast enhanced. Panels A, B, and C correspond to Figures 3.5.2B, 3.5.2C, and 3.5.2D, respectively.

A2.8 Ninhydrin test

The ninhydrin test of protected amino acid solution supernatants was performed as follows, according to Merrifield's method[2]:

Two developing reagents were used.

Reagent 1 was prepared by mixing 10 grams of phenol dissolved in 2.5mL ethanol with 1mL of (65mg KCN in 100mL water) diluted to 25mL in pyridine, for a total volume of 35mL.

Reagent 2 was prepared by dissolution of 1.25g of Ninhydrin in 25mL of ethanol, and was stored in the dark.

13mM glycine in water was used as an amine standard. Serial dilutions were performed to obtain standards of 3.25mM, 6.5mM, and 13mM.

To assay for amines, 10μL of the sample or standard was mixed with 100μL of reagent 1 and 25μL of reagent 2. The mixture was heated to 100C for 10 minutes in a sand bath. 1mL of 60% ethanol in water was then added, and the absorbance was promptly measured at 570nM on a BioTek plate reader. The assay was performed in duplicate, and all pure solvents were negative for amines.

A2.9 Third Generation Automated Fast Flow Peptide Synthesizer – Construction

This section is divided into five subsections, each covering a different aspect of the construction of the third generation AFPS: the reagent bottles and storage system, the valve module, the reactor module, the high pressure plumbing, and the electronics. The software systems are covered in the following section.

The goal of this section is to provide enough information to successfully reproduce the third generation AFPS from the technical drawings and software included as additional supplements. Some mechanical drawings and partial drawings are included in Appendix 2 for reference while reading the text, and complete models of all parts and assemblies are attached as Solidworks files. Additionally, the electrical specifications are attached as a PDF, and the LabView virtual instrument and its required functions and libraries are included.

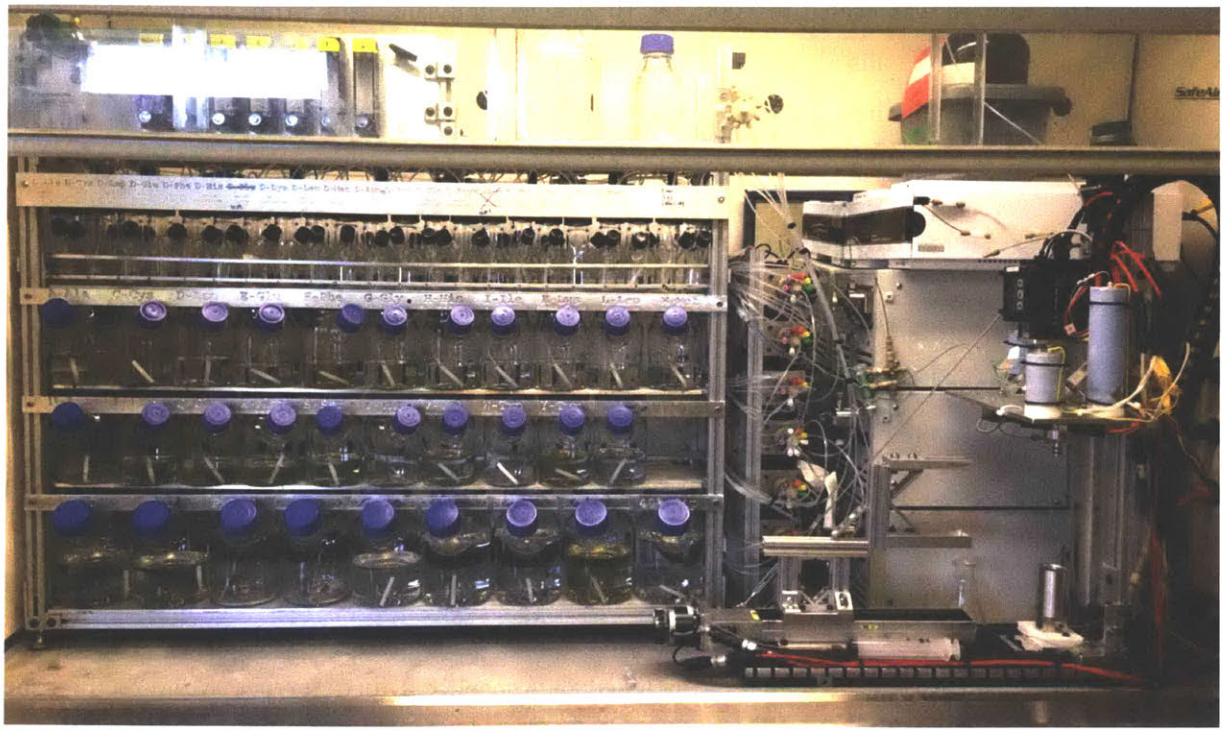


Figure A2.5: The photograph above shows the completed third generation automated peptide synthesizer.

A2.9.1 Reagent bottles and storage system

The reagent bottle and storage system is a conceptually simple module, but it is essential to a high quality user experience. The design here was created to fit in the available fume hood, and is probably close to the most compact possible design for the number and size of reservoirs used that still allows users easy access to all the fill ports.

1) In designing the reagent module, there were five key concerns:
2) The module should contain ~50 amino acid bottles, ~10 activator bottles, and wash solvent bottles
3) Bottles should be maintained under positive pressure
4) Bottles should be easy to inspect for fluid level and precipitates
5) Bottles should be easy to refill
6) Bottles must be safe.

Bottle design In order to achieve the above goals, commercial spinner flasks from Chemglass were used. These flasks have three necks: one was capped permanently, one was used as a fill port, and the last was fitted with a machined adaptor for nitrogen and liquid transfer. A dedicated fill port made it easier to refill bottles and prevented nitrogen leaks which could be introduced when the machined adaptor was

removed and reinstalled. To test the bottles, the fill neck and third neck were capped, and the adaptor for nitrogen and reagent withdraw was installed. One port on the adaptor was plugged, and the second was connected to a nitrogen supply. The bottles were submerged in water or ethanol and pressurized to 35 psig to ensure that they were safe and did not leak at operational pressure (<10psig). Numerous leaks were detected and stopped, but no bottles exploded. Initial pressurization was always conducted with a shield in place to contain glass fragments, should a bottle rupture.

All 70mm caps required silicone liners to seal. The silicone liners were likely not compatible with the chemicals used, and contact with reagents was scrupulously avoided. All GL-45 and GL-32 necks were fitted with pour rings, although there were not nessesary.

To make bottle filling easier, disposable polypropylene weighing funnels were used. These were manufactured by TWD Scientific and purchased from Sigma (Z409480-1PAK and Z409499-1PAK). The large were used to fill 500mL and 250mL flasks, and the medium were used to fill 25mL flasks.

Drawings for machined adaptors for GL45, GL32, and 38-400 bottle tops for use with commercial Chemglass tops with holes on the 500mL, 250mL, and 25mL spinner flasks, respectively, can be found in Appendix 1. Chemglass part numbers were as follows: 500mL spinner flask (CLS-1401-500), GL45 cap (CLS-1480-13SC), GL45 cap with a hole (CLS-1481-21), GL45 pour ring (CLS-1482-13), 70mm cap (CLS-1487-70), silicone liner for 70mm cap(CLS-1487-72), 250mL spinner flask (CLS-1401-250), GL32 cap (CLS-1480-11), GL32 cap with hole (CLS-1481-20), GL32 pour ring (CLS-1482-11), 25mL Spinner flask (CLS-1400-25), 18-415 side arm cap with PTFE liner (191-05PC), 38-400 cap with a hole (MIT-1511-101CLS).

Shelf Design

The primary purpose of the shelf was to hold the bottles so that they were easy to see, refill, and clean. The major constraint was that it had to fit in the hood with everything else. The bottom shelf held 9 500mL spinner flasks (1L capacity) for deprotection solutions and activators, the middle two shelves each held 11 250mL (500mL capacity) spinner flasks for proteogenic amino acids (two extra for non-standard protecting groups), and the top shelf held 22 25mL (50mL capacity) spinner flasks for D-amino acids and other unusual reagents. DIEA (500mL media bottle) and DMF wash solvent (2000mL media bottle) were placed on the roof.

The fill necks of the flasks stuck out beyond the face of the shelves for easy filling and there were no front cross bars of aluminum extrusion, so the bottles were easy to inspect and slide in and out. The engraved name plates seen in the above photograph had to be slid up before bottle could be removed, but they were non-structural and this was easy. The shelves were spaced vertically such that the 500mL and 250mL flasks fit snuggly. This prevented them from sliding around and enabled facile removal and replacement of the cap on the fill neck. Since the top neck of the 25mL spinner flasks was used for

nitrogen supply and reagent withdraw, the 25mL flasks were confined with an aluminum holder with snug fitting holes.

Engraved aluminum plates with the name of each reagent were added later. These labels were used as they could not fall off and were impervious to solvent. Unassigned bottles had the character used to reference the bottle in software on the label and space to write name with a marker as reagents were added. We found that Bic Mark-It markers were very slightly better than other brands at resisting solvent, and written labels could be covered with clear tape to protect them.

A2.9.2 Fluid Lines

After the shelf was constructed and the bottles were tested, the bottles were plumbed in. Each bottle had one nitrogen supply line and one reagent withdraw line. The liquid and gas lines were routed along the back of the shelving unit. In general, liquid lines were clustered based on the valve they served and routed along the bottom of their respective shelf and gas lines were clustered by the valve their bottles served and routed along the top of the their respective shelf. Liquid lines were confined with cable carriers, and these are color coded by valve in the CAD files.

The nitrogen lines were organized as follows: the main nitrogen supply was passed through a regulator (Swagelok KPR1DFC412A20000) to a main rotameter (McMaster 4112k34). The outlet was used to supply a manifold which in turn supplied one rotameter per reagent valve (McMaster 4112K38). The bottles for each valve were supplied by a single rotameter via a manifold. Manifolds were Idex P-190 9-port manifolds; if more ports were needed, Idex three port manifolds (Idex P-713) were used. Unused ports were plugged. The reagent withdraw lines were routed directly to their assigned valves. The DMF was pressurized to the same pressure as the reagent bottles on the shelf using a spare port on the manifold served by the main rotameter, and a three way valve (Swagelok SS-41GXS2) was used to vent the bottle during refilling (the common port was connected to the bottle, another port to nitrogen, and the third to atmosphere). DIEA was not pressurized or stored under inert atmosphere. Nitrogen lines were 1/8" OD 1/16" ID polyethylene tubing (McMaster 5181K15) and reagent lines were 1/8" OD 1/16" ID PFA tubing (McMaster 5773K11). All fittings were Idex ¼-28 flatbottom fittings with superflangeless ferrules (Idex P-359) and either standard (Idex P-331) or headless nuts (Idex P-387).

All of the lines were left long enough to remove each bottle from the shelf and access the neck with the machined insert. This made it easy to fully remove the bottles for cleaning.

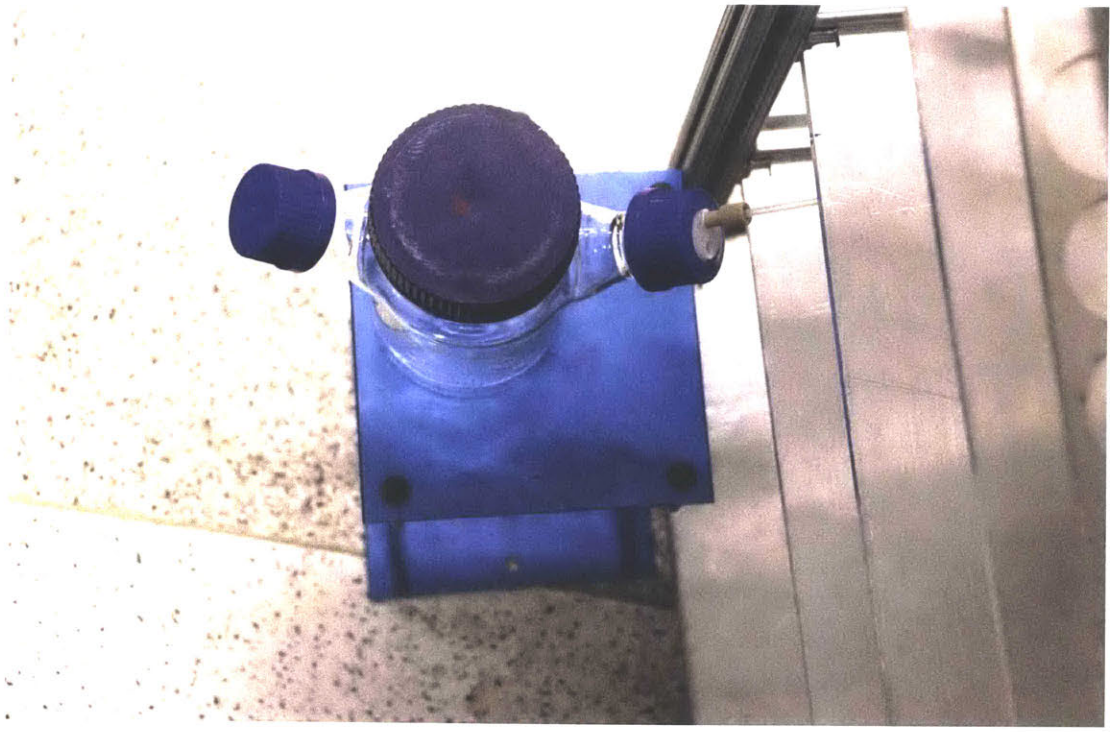


Figure A2.6: Lines were cut to length with the bottles on a jack stand off the rack to ensure they could be removed without being tied back by nitrogen or reagent lines. The nitrogen line is being cut in this photo; the reagent line is not yet installed.

Figure A2.7: Bottles to serve one valve with nitrogen lines installed. Note that reagent lines have not yet been installed in this photo.

Figure A2.8: Front view of the assembled shelf module with bottles, prior to installation of the engraved labels.

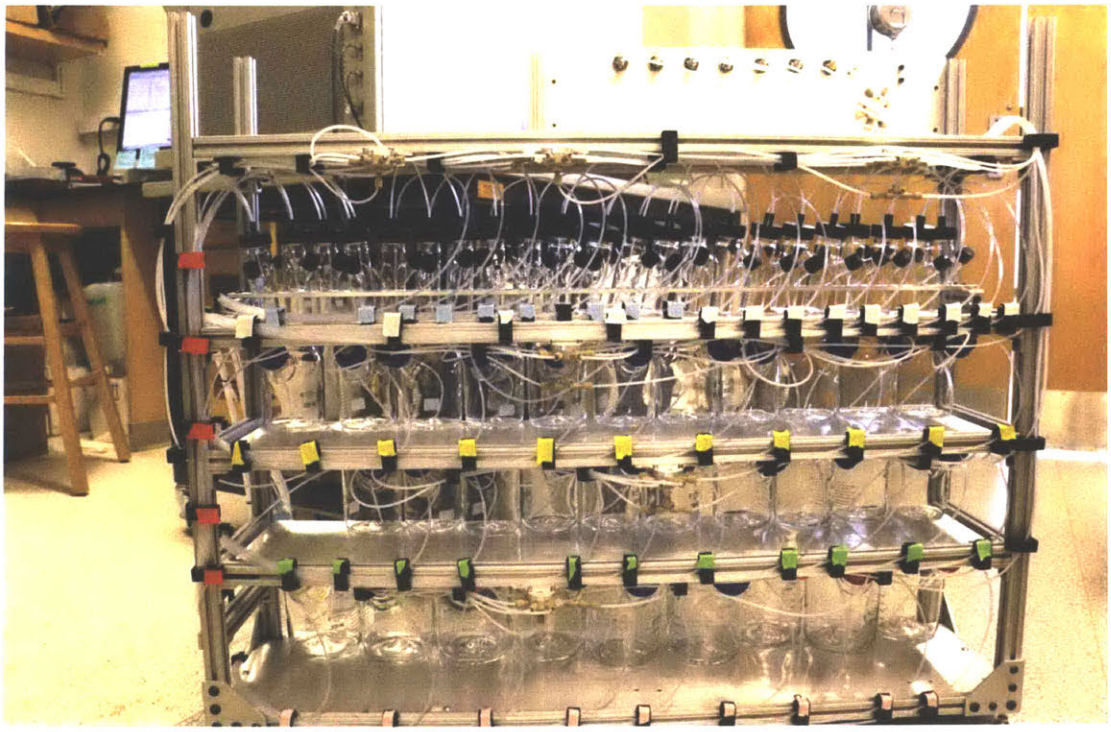


Figure A2.9: Rear view of the assembled shelf module with the bottles. Note that cable carriers used for fluid lines are color coded by valve, whereas nitrogen lines are confined by uncoded cable carriers.

Figure A2.10: Liquid lines were bundled by valve and labeled. If the labels fall off it is difficult to trace them later. Foil caps prevented dust from entering prior to connecting the valves.

A2.9.3 Valve Module

1) The valve module was built with the following design goals:
2) Maintain access to the front of the valves to facilitate service of fluid lines and valve heads
3) Protected components from DMF as much as possible
4) Maintain visual and mechanical access to the manual controllers to manually step the valves and verify their positions
5) Place valves physically close to one another and the pump
6) Support 7 Cheminert 10 position Valco C25-6180EMH selector valves (1 for activators and deprotection reagents, five valves for amino acids, and one master valve to select between the five amino acid valves)

To meet these goals and conserve horizontal space in the hood, the valves were mounted vertically, as close to one another as practical. The manual controllers were affixed to the sides of the valves and covered with adhesive backed FEP film (McMaster 85905K66) to protect them from DMF. An electrical distribution panel (McMaster 7527K51) was mounted on the reverse side of the valve mounting plate to supply 24V power, and all wires were FEP insulated. Valco cables (I-22697) were used for serial

connections, but these occasionally mechanically fail and are not resistant to DMF, so should not be used again.

A2.9.4 Low pressure plumbing

The valve module was plumbed as follows using highly transparent "extreme purity" 1/8" OD, 1/16" ID PFA tubing (McMaster 5773K11). All flatbottom fittings used superflangeless ferrules for 1/8" tubing (Idex P-359), including the fittings attaching tubing to the valves. Various ¼-28 nuts were used with these ferrules: Idex headless knurreld nuts for the connections to the 25mL flasks (Idex P-387), standard Idex nuts for the 250 and 500mL flasks (Idex P-331), and the Cheminert nuts that came with the valves for the valve connections (the cheminert ferrules were found to fail when subject to fluid hammer, so were discarded and replaced with superflangeless ferrules).

The tenth position of each valve was connected directly to DMF wash solvent (stored on the roof of the reagent shelf) via a manifold (Idex P-190). A check valve (Idex CV-3315) was used on the inlet of the manifold to eliminate the possibility of backflow into the DMF reservoir. The remaining nine inlets of valve 1 were connected to the nine bottles on the bottom shelf, which contained activators and deprotection reagents. The outlet of valve 1 was connected to the inlet of the activator pump with a machined adaptor (see Appendix 1).

The remaining 9 positions of valve 3 were connected to the nine left most bottles on the second shelf, the remaining nine positions on valve 5 were connected to the nine leftmost bottles on the third shelf, the remaining nine positions on valve 6 were connected to the nine leftmost bottles one the top shelf, the remaining nine positions of valve 7 were connected to the tenth through 18th bottles from left on the top shelf, and the remaining 9 positions of valve 2 were connected to the remaining bottles or plugged if unneeded. The outlets of each valve were connected to their respective positions on valve 4 (two to two, three to three, ect). Positions 1, 4, 8, and 9 were plugged on valve 4.

The outlet of valve four was connected to the inlet of the amino acid pump using the same kind of machined adaptor as was used with the activator pump.

DIEA was stored on the roof of the reagent shelf and connected directly to the DIEA pump without passing through a valve. No adaptor was necessary.

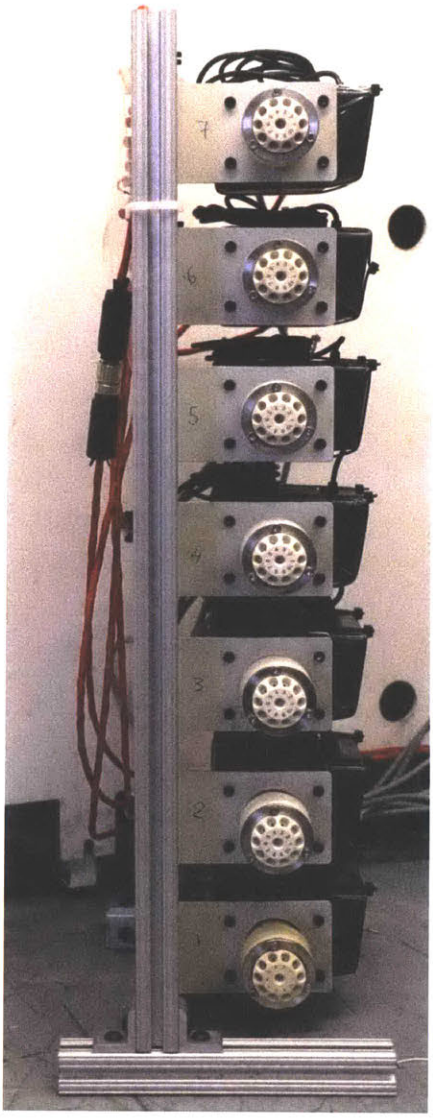


Figure A2.11: Front view of the valve module showing the spacing of valves. The physical number on each valve corresponds to its valve ID and is used to address it in software. The 24V distribution panel can be seen on the left side and is covered by an FEP sheet to prevent shorts. The bottom ports on the lowest valve are difficult to access. In the future there should be more space between the bottom valve and the floor.

Figure A2.12: Side view of the valve module. Manual controllers were affixed to each valve to facilitate troubleshooting and system verification.

Figure A2.13: Rear view of the valve module showing electrical (red), manual controller (black) and data (grey) connections.

Figure A2.14: Photograph of the plumbed valves. The union cross, check valves, and static mixer on the high pressure side are also visible.

A2.9.5 Reactor module

The reactor module was built with the following design goals:

1) Provide several independently heated activation reactors of various length to study the effect activation time and temperature on synthetic outcomes independently of the amide bond formation time and temperature.
2) Have an independently heated resin vessel to study the effect of temperature on amide bond formation
3) Load syringes of resin without bubbles
4) Remove syringes of resin from the heated zone when synthesis finished.
5) Accommodate scales 1x, 2x, and 4x larger that of the second generation manual fast flow peptide synthesis reaction vessel
6) Protect components as much as possible from DMF

The reactor module consisted of two major parts: several individually heated tubular activation reactors with a valco column changer to select between them, and a heated cylinder that confined a fritted syringe filled with the solid support for peptide synthesis. Two linear actuators loaded syringes of resin without air bubbles and removed them from the heated zone when synthesis finished.

Tubular activation reactors

The tubular activation reactors and the Valco column changer that selected between them (Valco EMTCST6UW) were mounted on the plated used to mount the resin chamber inlet and inlet pressure sensor to minimize the distance (and time) reactive species traveled before reaching the resin. This mounting plate is referred to as the "middle plate" in the accompanying CAD drawings and was itself mounted on the vertical linear actuator (Parker Daedal 404350XRMS track and Parker Compumotor Zeta 57-102-MO motor). The entire module containing the valve, activation reactors, inlet pressure sensor, and resin chamber inlet moved up and down to open and close the resin chamber.

Each tubular activation reactor consisted of a heated aluminum cylinder with a stainless steel tube wrapped around it (1/16" OD, 0.030" ID, Idex U-164). Once wrapped with the stainless steel tube, the cylinder was insulated with blue silicone insulation (Dow RTV-630), giving them the distinctive blue color seen in the photographs. Each activation reactor was mounted to the "middle plate," with a PTFE spacer to insulate it and provide clearance for the heater wires. A drawing of the aluminum cylinder and mounting spacer is in Appendix 1.

Resin Vessel

The resin vessel confined and heated a fritted syringe barrel that contained the solid support. It consisted of three parts: the inlet, the body, and the outlet. The resin vessel inlet had a cylindrical lower section that formed a double o-ring seal with the inside of the syringe barrel and had an internal channel

for reagent delivery. The inlet was heated with cartridge heaters and controlled as its own heated zone. The inlet mounted to an aluminum inlet mounting plate that confined the heaters and thermocouple and provided a consistent alignment for inlets of all scales. The inlet mounting plate in turn mounted to the "middle plate" with a PTFE spacer to insulate it and provide clearance for the heater and thermocouple wires. The mounting slots on the middle plate allowed the inlet to slide back and forth so it could be properly aligned.

The resin vessel body confined and heated the syringe barrel. The body was a hollow aluminum cylinder with the inner diameter matched to the outer diameter of the syringe. The length was slightly (~0.10") less than the overall length of the syringe, so when the resin vessel closed, the syringe was pressed tightly into the outlet. Additionally, the resin vessel body had relief for the flange and slightly larger OD of the syringe's top collar. The body also had holes for ¼" cartridge heaters and countersunk clearance holes for 10-32 screws used to secure it to the outlet.

The resin vessel outlet mated to the body and provided a female luer connection for the syringe outlet and a flow channel for waste. The outlet of the resin vessel outlet was connected to the outlet pressure sensor, and then to the UV detector. The resin vessel outlet housed the thermocouple that monitored the temperature of the resin vessel body/outlet unit and shared cartridge heaters with the resin vessel body. The outlet was mounted to an optical bread board (Thor Labs MB3060/M, cut to size) shared with the vertical stand. The mount was PTFE and allowed the outlet to move in two directions for proper alignment.

Syringes of resin were loaded and unloaded with a syringe carrying fork on a horizontal linear actuator (Parker Daedal 402006LNMP track and Parker Compumotor Zeta 57-51 motor). To effect bubble free loads, fritted syringes with resin were placed on the fork, which held them by the flange. The fork then moved towards the open resin vessel, positioning the syringe with resin directly below the inlet. DMF was pumped through the inlet to fill the syringe, and a drip tray caught excess DMF. The inlet was then lowered into the syringe and raised slightly so the syringe flange cleared the lip of the fork as the fork was withdrawn. The resin vessel inlet was then lowered to its fully closed position, sealing the syringe in the heated chamber.

After synthesis, the resin was removed from the heated zone. The resin vessel inlet and syringe were raised until the syringe was level with the fork, the fork was extended, and the inlet was fully raised, leaving the syringe on the fork. The fork was withdrawn, and the drip tray caught DMF as it drained from the syringe. Removing the syringe from the resin vessel relied on the friction created by the o-ring seal and was not completely reliable. If the syringe got stuck, it was quickly removed with a pair of pliers. It is likely that the reliability can be improved by reducing the ID of the resin vessel body, especially near the top (inlet).

By changing the inlet, body, and fork, synthesis could be performed in 6mL, 12mL, or 24mL syringes. The only change in the operation of the resin chamber closure at different scales was the final position of the inlet: it was higher at larger scales. Selecting the appropriate scale in the software instructed the vertical actuator to adjust its final, closed position. The alignment did not need to be readjusted when changing scales because the outlet and inlet mounting plate were not removed. Similarly, the fork's position was fully defined by the two bolts securing it.

Fritted syringes were prepared by punching out a circular section of 7um polyethylene frit from a 0.125" thick sheet (Interstate Specialty Products POR-9948) and stuffing it in the bottom of the syringe using the plunger. Normject syringes were used, and the punch diameters were ½" for the 6mL syringe (Air-Tite ABC5LS), 16mm for the 12mL syringe (Air-Tite ABC10LS), and 13/16" for the 24mL syringe (Air-Tite ABC20LS). The nominal capacity for ChemMatrix resin (high swelling PEG) was <200mg for the 6mL syringe, <400mg for the 12mL syringe, and <800mg for the 24mL syringe, but the syringes could be overloaded somewhat, especially the larger ones.

Figure A2.15: Photograph of the reactor module prior to plumbing and wiring. The fork is mounted to the horizontal actuator, but is not in its final configuration (see CAD drawings). The 6mL resin vessel body and universal outlet are clearly visible on their mounts, and the 6mL inlet is just barely visible under the middle plate. Three activation reactors, the column changing valve, and both pressure sensors are also clearly visible. The yellow wires are thermocouple leads and the white wires are cartridge heater leads. The grey wires are leads for pressure sensors, motors, and the limit switches used to home the motors.

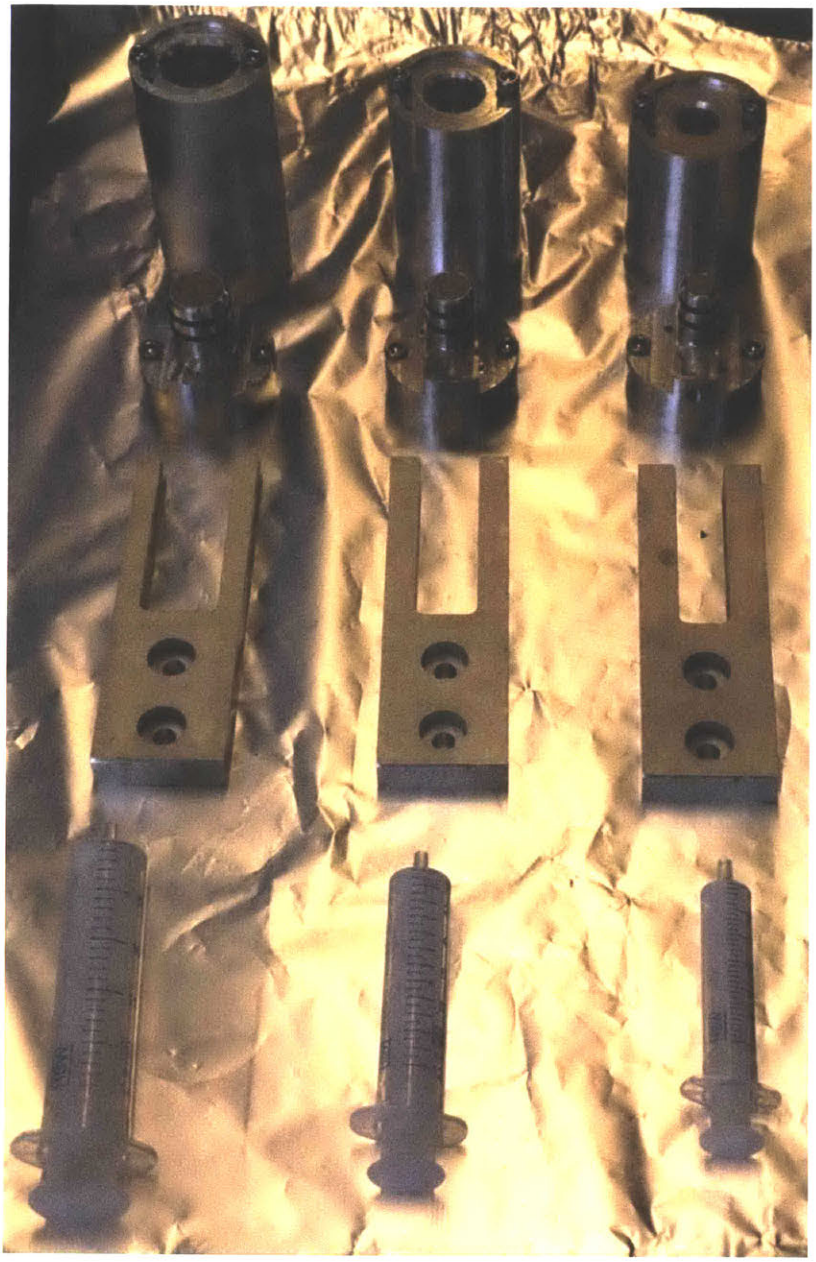


Figure A2.16: Photograph showing resin vessel inlets, resin vessel bodies, and forks for three scales with the syringes they accommodate. On the left is the 24mL scale, the 12mL scale is center, and the 6mL scale is on the right.

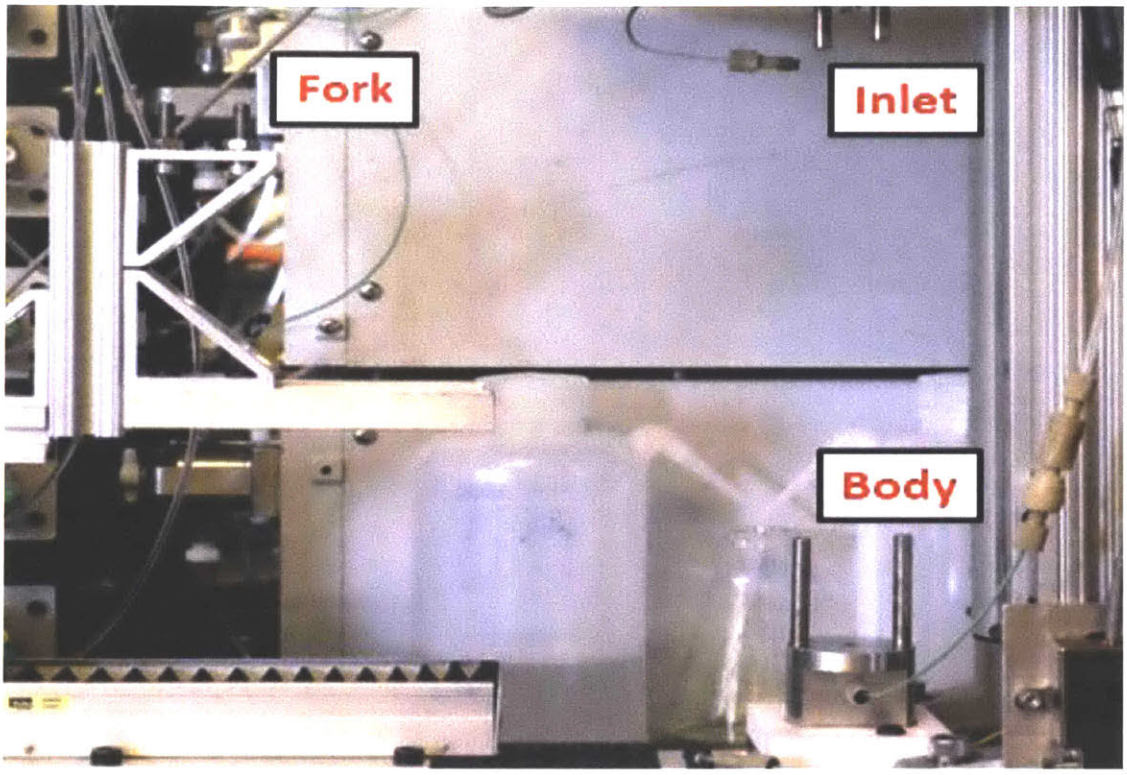


Figure A2.17: Photograph showing the reactor module with the fork, resin vessel inlet, and resin vessel body removed. Any of the three sets of hardware above could be installed to perform synthesis on any of three scales. The cartridge heaters and thermocouples are captive in the universal outlet and universal inlet mounting plate, so they do not need to be manipulated when changing scale. In this photograph, the outlet pressure sensor has been removed from the flow path.

A2.9.6 High pressure plumbing

Whenever ¼ -28 flatbottom fittings were required, Idex superflangeless fittings (Idex P-255 nuts and Idex P-250 ferrules) were used. These fittings have the highest pressure rating of any flat bottom fittings considered, and seem to be slightly more resistant to loosening in a high vibration environment. They do still have to be tightened from time to time, however. In the future, it may be better to employ Swagelok fittings which will not loosen when subject to vibration.

Wherever not specified, 1/16" OD 0.030" ID PEEK tubing was used (Idex 1533XL). This tubing has a sufficient pressure rating that the pumps will stall before the tubing ruptures, reducing the possibility of a massive mess. It is also easy to work with as it is flexible and simple to cut, unlike stainless steel tubing of comparable OD and ID.

Equal lengths of tubing were connected to the pump head outlets using the appropriate fitting (Idex F-310). These tubes were then connected to check valves (Idex CV-3301) installed on three ports of a union cross (Idex P-722). The fourth port of the union cross was connected to a curved tube static mixer consisting of 22 turns of the above PEEK tubing around a ½" diameter shaft (a 12mL syringe barrel was used).

The outlet of this was in turn connected through a union (Idex P-702) to the inlet of the Valco column selection valve on the reactor module. The inlet and outlet of each activation loop were connected to a short piece of PEEK tubing with a union (Swagelok SS-100-6) and the PEEK tubing was connected to the appropriate positions on the column selector using the fittings supplied with the valve. The outlet of the column selector was connected to the inlet of the inlet pressure sensor and the outlet of this pressure sensor was connected to the resin vessel inlet. The resin vessel outlet was connected to the outlet pressure sensor, then to the UV detector via a check valve and union (Idex CV-3301, Idex P-702). The UV detector effluent was passed directly to waste.

NOTE: The check valve on the amino acid pump inlet to the cross tended to fail, presumably from the acidic reagent pumped through it. The cracking pressure from this check valve was necessary to prevent siphoning of reagents, so it is recommended to have spares and/or put a backpressure regulator/check valve with cracking pressure between the static mixer and Valco column changer inlet.

A2.9.7 Electronics

The electronics package to control this instrument was designed and installed by RM Beaumont Corporation (RBC) in response to our requirements. The detailed schematics and specifications provided by RBC are attached as an additional supplement; this section provides a brief operational overview.

The electronics package was responsible for controlling all aspects of the AFPS instrument, and digitizing data from all of the installed sensors. The electronics package controlled the heaters, pumps,

valves, and reactor module motors. Data was read in from the UV-Vis, thermocouples, and pressure sensors. Additionally, the software and electronics were capable of logging data from resistance temperature detectors and a mass flow meter, but this hardware was not installed. Furthermore, the electronics package supplied power to the valves, pressure sensors, heaters (via solid state relays), and reactor module motors (via the TinyG CNC controller).

The temperature of each individually thermostated zone was maintained by an eight channel Watlow PID controller. Five zones were implemented: three activation loops, the resin vessel inlet, and the resin vessel body/outlet. Temperatures were measured with K type thermocouples, and the Watlow controller switched power to Omega high watt density ¼" cartridge heaters via solid state relays. The temperature setpoints were established using Watlow's EZone control software. If any temperatures were changed, the heated zone was allowed to equilibrate for about 20 minutes prior to initiating synthesis. The LabView virtual instrument was able to read temperatures measured by the Watlow, but was not able to change set points.

The pumps and valves were controlled on two separate RS-485 serial communication channels from a Vlinx serial server, and the reactor motors were controlled by a TinyG CNC controller. The TinyG was configured to keep the motors powered when idle, so that the Z-axis motor would not be back driven by the force applied by the pressure in the resin vessel. If the motor was unpowered, the resin vessel would open during synthesis. This is an increasingly sever problem as scale (and inlet area) increases. A more powerful motor is recommended. Further, the Parker Zeta Compumotors were supplied in an 8 wire configuration; they were wired in series to give the best low speed/high torque performance and make them compatible with the TinyG (Blue and yellow shorted; orange and brown shorted in the motors used)

The UV and pressure data was read in by a National Instruments ADC. Thermocouple readings were taken from the Watlow controller. The UV detector was an Agilent G1365D, and the pressure sensors were DJ Instruments MOD DF2-TI-01-6000 SQ, 5V.

Where possible, connectors and wires were insulated with perfloropolymers to resists solvent spills in the hood.

A2.10 Software

A2.10.1 Overview

The software to control this instrument was largely written by RM Beaumont Corporation (RBC) in response to our requirements. The instrument is controlled as a LabView virtual instrument. The VI and all of its dependent VI's are included as an additional electronic supplement to this manuscript. The

principle of operation is briefly described here and key features are highlighted. Throughout the software "prime" means "coupling" and "recipe" means "synthetic procedure."

Key features:

1) All synthetic parameters, including scale, could be adjusted by the user in the GUI
2) All synthetic parameters, including instrument configuration files, were saved automatically for each synthesis
3) All UV, pressure, and temperature data was automatically logged
4) Reagent and waste levels were monitored, and a warning was issued if levels were too low
5) The peptide synthesis template was sufficiently flexible that, with some creativity, any series of pump and valve actuations could be added at any point during peptide synthesis, enabling unusual chemistries to be explored (although these experiments are not presently reported)
6) Additional utilities separate from the main VI enabled independant control of the valves, pumps, and reactor motors for trouble shooting and other experiments outside of the paradigm of peptide synthesis

A2.10.2 Main VI

The main VI is organized as a flat sequence array with 11 steps; the sequence is executed once per synthesis. **In the first step**, numerous variables are declared and set to zero.

In the second step, the contents of several configuration files are loaded. Except for AFPS_CONFIG.xml, the configuration files are comma separated value files, editable with notepad or excel, but have unusual file extensions to help keep track of them.

The loaded configuration files are as follows:

HX_PARAMS.adm is a file that records the length, temperature set point, and valve position of each heat exchanger. The only parameter that affects hardware operation is the valve position; the length and temperature set point are only used to generate the name of the heat exchanger displayed to the user. Length and temperature set point are hard coded and not verified anywhere.

AMINO_PARAMS.adm is a file that records the position of each amino acid on the valve manifold, as well as default coupling and deprotection conditions. The position on the manifold consists of three parameters: the bank valve ID, the position on the bank valve, and the position on the master valve (this was always the same as the bank valve ID in hardware). The conditions recorded are: activator, flow rate for coupling, number of pump strokes for coupling, activation reactor for coupling, number of pump strokes for the washes, flow rate for deprotection, activation reactor for deprotection, and deprotection

type (see next section). The amount of reagent remaining in the reservoir is also recorded here and update during synthesis.

SYNTHESIS_SEQUENCE_BASELINE.syn is a file that configures the hardware sequence used to perform each synthetic step. A synthesis consists of four elements: the prewash sequence (everything that happens before the first coupling), and one prime (coupling) and one deprotection step for each amino acid, and the postwash sequence (everything after the last deprotection).

There are four types of deprotection steps (DEPRO0 thru DEPRO3). The user selects between them for each amino acid coupled (the default is controlled by AMINO_PARAMS.adm). These alternate deprotection conditions can be used to remove Fmoc in different ways (for example with piperizine or morpholine) or to enable unusual chemistry to be implemented. Each deprotection condition is a fully programmable synthesis step consisting of a series of hardware commands (pump and valve actuations). This feature was not used in this work, however it does enable exploring diverse solid supported chemistry with the instrument.

AFPS_CONFIG.xml is a file that records a large number of instrument parameters including the (virtual) communications ports for communication with pumps, valves, and motors, the DMF position on all the valves, the master valve ID, ect. Not all of the parameters are used. Some were programed for anticipated alterations to the hardware (for example, there was no DIEA valve implemented in hardware, but the configuration file assigns it an ID).

The third step uses the loaded sequence file to initialize a lot of variables to their correct (non-zero) values and declares a few additional zero-valued variables.

In the fourth step, there is a 5 second pause for the watlow to start.

In the fifth step, communication is established with controlled hardware and the ADC.

The sixth step changes the machine state to "Recipe Setup."

The seventh step contains the loops that setup and execute synthesis. There are three main loops in this step that execute in parallel. The first loop contains nested loops that control interaction with the user and run the synthesis. The second loop records data and the third loop updates amino acid levels, prints data to the front panel, and handles some minimal data processing tasks. The execution of the loops is asynchronous and, in general, LabView does not have constant execution rates, so where timing was essential explicit timers were used.

The operation of the second and third loops is clear, so this section will focus on the first loop. In the first loop, a bunch of variables are declared and/or re-declared. This is so they update in real time as the loop is running. There are also two nested loops, an upper one and a lower one. The upper loop handles actual execution of the synthesis once the necessary steps have been determined, and the lower

loop interacts with the user to determine the necessary steps for synthesis, among other user-operated functions.

The lower loop is an Event Structure. Whenever a value of interest is changed by the user, the loop responds. The user can directly control the reactor module motors, load saved syntheses, change synthetic parameters from the loaded or default parameters, save the synthesis, and start the run. Most of the procedures in this loop revolve around properly interacting with the user so that he or she can input chemical parameters, and then properly converting those chemical parameters into machine-level instructions for the actuation of pumps, valves, and motors. The conversion functions are largely handled by Sub-VIs which can be opened from the attached VIs and inspected; here two important features will be highlighted.

Critically, the "recipe files" created when the user saves a synthetic procedure store the chemical level data visible to the user in the GUI, so even if the hardware is reconfigured, files can be used to reproduce earlier synthesis without modification. This improves the reproducibility of synthesis, and helps prevent the use of old synthetic protocols from rolling back hardware or software changes. For example, the time allowed for the pumps to fill a syringe with DMF prior to a bubble free load was increased by hardcoding an increase in the "dwell time" for that step. This change is not rolled back by loading old recipes.

Furthermore, when the "Run" button is clicked, a data folder is created for the synthesis. In this data folder, there is the data log file that UV, pressure, and temperature data is written to as well as a copy of all the instrument configuration files, the saved recipe file, and the actual recipe file. The actual recipe file should not differ from the saved file, but there are a few circumstances in which it can. Most commonly, if a user loads an old recipe, makes changes, and does not save it again the "recipe" and "actual recipe" will be different.

The upper loop steps through the synthetic procedure generated by the lower loop from the user's input. There are several important parameters for the execution of each step: the "Dwell Time" which may be hard coded or calculated during the creation of the step, the "Device Type" which selects between pumps, valves, and motor control (TinyG), and the commands to send (Arg1, Arg2, and Arg3). For each step, the loop establishes a step start time, transmits the appropriate commands, then checks to see if enough time has elapsed to move on to the next step. If enough time hasn't passed, the program waits, updating the front panel and blinking indicator lights every second to reassure the user the program hasn't frozen. Once enough time has passed, the upper loop moves on to the next step.

The eighth step, set the machin state to shutting down and waited 500ms for the Master Stop variable to update.

The ninth step, which executed after synthesis was complete, closed the data file process data was written to, as well as ADC communications. The pumps were also stopped in this step, as a safety measure to prevent them from running forever if the "postwash" sequence was not configured properly.

The tenth step closed communications with the valves, pumps, TinyG (motors), mass flow meter (hardware not implemented), and ADAM6015 (RTD reader; hardware not implemented). The configuration file, AFPS_CONFIG.xml was also updated in this step, as were the final reagent levels.

The eleventh step reset the Master Stop variable and changed the machine state variable to Stopped.

After execution of the eleventh step, the VI stopped and had to be run again to configure and run another sequence.

A2.10.3 Additional notes on the software

One major challenge in writing this software was initiating operation of the pumps simultaneously. To accomplish this, each set of pump control serial commands was concatenated into a single command and sent with one call of the VISA write utility (LabView utility for sending serial commands). If pump commands were sent as separate serial commands with multiple, sequential calls of VISA write, the pumps ran asynchronously. See the sub VI "Varian Send Multi"

Sometimes LabView crashed after synthesis was finished, and did not close properly. In these cases, there was often a LabView process that could not be ended in the task manager. The presence of this process prevented acquisition of analog data by a second LabView process when LabView was restarted. This bug is possibly related to the global Master Stop variable updating too slowly. To recover from this problem, the electronics were switched off, the computer was shut down, the electronics were switched on, the computer was turned on, and LabView was started again.

These VIs are written in LabView 2015 and run on windows 7. An upgrade to windows 10 broke serial control of all the hardware, so was rolled back.

A2.10.4 Additional Utilities

There are utilities to directly control the pumps, valves, and motors that are very useful for trouble shooting and doing experiments outside of the context of synthesis. For example, the dye study was done using the pump and valve utilities, rather than forcing the procedure to awkwardly fit into the template for peptide synthesis. These utilities are called "Varian_Controller.VI", "Valco_Controller.VI", and "TinyG_Controller.VI", and it is clear how they work when they are opened. They are not called by the main VI.

A2.10.5 Synthetic procedure

As mentioned above, mechanical steps to complete each synthesis were generated from a combination of instrument parameters and default synthetic conditions stored in configuration files, and user-defined conditions entered while setting up the synthesis.

In the experiments presented here, the following hardware sequence was used to perform each synthesis, with synthetic parameters as indicated by the user or default values. Each row is in 1:1 correspondence with SYNTHESIS_SEQUENCE_BASELINE.syn, which is included as part of the LabView VI supplement (open it with notepad or with Excel as a comma delimited file). The first three columns are copied directly from SYNTHESIS_SEQUENCE_BASELINE.syn, and the fourth column is a longer explanation than is in the "desc" column. SYNTHESIS_SEQUENCE_BASELINE.syn has 11 columns in it with various parameters. Some values are hard coded, and other are included as variables indicated with a "#" symbol. The variables are replaced with numerical values as the hardware steps for synthesis are being created and can be system paramerters, default values, or user set values. The 11 columns are as follows. The "Step ID" and "Stage" are used to index and group the steps while building up a list of hardware commands for synthesis. "Desc" is a brief description of the step. The "Dwell Time (ms)" is the minimum time before advancing to the next step. "Device type," "Device ID," and "Action" all relate to the type of command and device it is being sent to. "Arg1" is the primary argument of any command (the flow rate for a pump, the position for a valve). "Arg2" is only used for setting Varian pump displays, and "Arg3" is used for level monitoring.

Table A2.1 Hardware sequence used for synthesis. Each synthesis consisted of four stages, the PREWASH (steps 1-15 performed once at the beginning), the PRIME (steps 16-29 performed for each coupling), the DEPRO (Depro_0, Depro_1, Depro_2, or Depro_3 are user selectable but identical at this time – Depro_0 is steps 30-37 and repeated for each deprotection step), and the POSTWASH (steps 62-66, performed once at the end of synthesis).

STEP ID	STAGE	DESC	Purpose
1	PREWASH	Lock all pumps	Pumps must be locked before accepting any other serial commands. This makes sure they are.
2	PREWASH	All pumps off	The flow rate of all pumps is set to zero, just in case they were running.
3	PREWASH	Position HX valve	The HX valve is turned to a hardcoded position known to be in use
4	PREWASH	Query TinyG (wake it up)	TinyG is queried to wake it up
5	PREWASH	Home linear actuators	TinyG is sent commands to home X and Z
6	PREWASH	Close reactor- stage 1	The X actuator positions the syringe under the inlet and the inlet is lowered a bit
7	PREWASH	Master valve to DMF	Ensures the master valve is on DMF
8	PREWASH	AA pump on (Fill syringe with DMF)	Fills the syringe with DMF using the AA pump. Can potentially overfill into the drip tray.
9	PREWASH	AA pump off	Stops the pump.
10	PREWASH	Master valve to blank	Turns the mater valve to a plugged port so it does not siphon
11	PREWASH	Close reactor- stage 2	Puts the syringe on the inlet, retracts the fork, and finishes closing the reactor
12	PREWASH	Act valve to DMF	Ensures the activator valve is on DMF
13	PREWASH	Master valve to DMF	Ensures the master valve is on DMF
14	PREWASH	AA and Act pumps on (Prewash)	Runs the activator and AA pumps for a 20sec prewash.
15	PREWASH	AA and Act pumps off	Stops both pumps
16	PRIME	Position HX valve	Positions the HX valve to select a default or user defined activation reactor for coupling
17	PRIME	Position AA bank valve	Position the correct AA bank valve to the desired AA
18	PRIME	Position master valve	Positions the master valve to the correct bank
19	PRIME	Position act valve	Positions the activator valve to select a default or user defined activating agent
20	PRIME	AA and Act pumps on (Prime)	Runs the AA and activator pumps long enough for reagent to reach the pump heads
21	PRIME	AA and Act pumps off	Stops both pumps - necessary to start all three synchronously in next step
22	PRIME	DIEA, AA, and Act pumps on (Couple)	Runs all three pumps to infuse DIEA now that reagent is coming through - continues for user defined or default time/flow rate
23	PRIME	All pumps off	Stops all pumps
24	PRIME	Act valve to DMF	Turns the activator valve to DMF

25	PRIME	AA valve to DMF	Turns the AA bank valve to DMF
26	PRIME	DIEA, AA, and Act pumps on (Wash valves)	Runs all three pumps to infuse DIEA as the valves are washed and residual reagent comes thru
27	PRIME	All pumps off	Stops all pumps
28	PRIME	AA and Act pumps on (Wash resin)	Runs the activator and AA pumps for a DMF wash of user defined or default duration/flow rate
29	PRIME	All pumps off	Stops all pumps
30	DEPRO_0	Position HX valve	Positions the HX valve to select a default or user defined activation reactor for deprotection
31	DEPRO_0	Master valve to dmf	Ensures the master valve is on DMF
32	DEPRO_0	Pip (act) valve to pip	Ensures the activator valve is on DMF
33	DEPRO_0	AA and Act pumps on (Deprotect)	Runs the activator and AA pumps for deprotection of user defined or defualt duration/flow rate - 40% pip is diluted with an equal volume of DMF
34	DEPRO_0	AA and Act pumps off	Stops both pumps
35	DEPRO_0	Pip (act) valve to dmf	Turns the activator valve to DMF
36	DEPRO_0	AA and Act pumps on (Wash)	Runs the activator and AA pumps for a DMF wash of userdefined or defualt duration/flow rate
37	DEPRO_0	AA and Act pumps off	Stops both pumps
xxxx	xxxxxxxxxx	xxxxxxxxxxxxxxxxxxxx	Depro_1, Depro_2, and Depro_3 are identical to Depro_0, but are included to be changed for custom chemistries. Lines may be added or subtracted, but Step IDs must be changed to remain sequential.
62	POSTWASH	AA and Act pumps on (Postwash)	Runs the activator and AA pumps for a 20sec postwash.
63	POSTWASH	AA and Act pumps off	Stops both pumps
64	POSTWASH	Set pump display	Sets the pump display to "Sequence finished"
65	POSTWASH	Master valve to blank	Turns the master valve to a plugged port so it does not siphon (acid quickly destroys the check valve that provides cracking pressure)
66	POSTWASH	Open reactor	Opens the reactor and extracts the syringe with the fork

A2.11 References for appendix 2:

(1) Mong, S. K.; Vinogradov, A. A.; Simon, M. D.; Pentelute, B. L. ChemBioChem 2014, 15 (5), 721.

(2) Sarin, V. K.; Kent, S. B. H.; Tam, J. P.; Merrifield, R. B. Anal. Biochem. 1981, 117 (1), 147.

A2.12 Additional electronic files supporting Chapter 3

1) Matlab scripts for automated UV data processing
2) Solidworks parts and assemblies of all mechanical components
3) Documentation of the electronics package (mechanical and electrical)
4) LabView VIs
5) Technical reference manual for the Varian 210 pumps (it is out of print and hard to get)

A2.13 Drawings of select machined parts

This appendix includes drawings of particularly useful machined parts. Complete drawings for all parts and assemblies can be created as needed from the Solidworks models that accompany this document as additional supplementary information. Commercial mechanical components such as screws and washers were purchased almost exclusively from McMaster Carr, and part numbers are generally included in the model names. Tolerances are not discussed, but were held to ±0.005". Attention must be paid to the tolerances of the resin vessel inlet O-ring grooves and the ID of the resin vessel body, or finished peptidyl resin may not be removed properly from the heated resin vessel body. All extruded T-slotted framing was McMaster 47065T101.

All dimensions are given in inches, however a few items mount on 24mm or 25mm centers resulting in unusual inch dimensions and non-interchangeable mounting parts. It is recommended to standardize all mounts that are approximately one inch to exactly one inch and use 10-32 threads to replace M5 for simplicity.

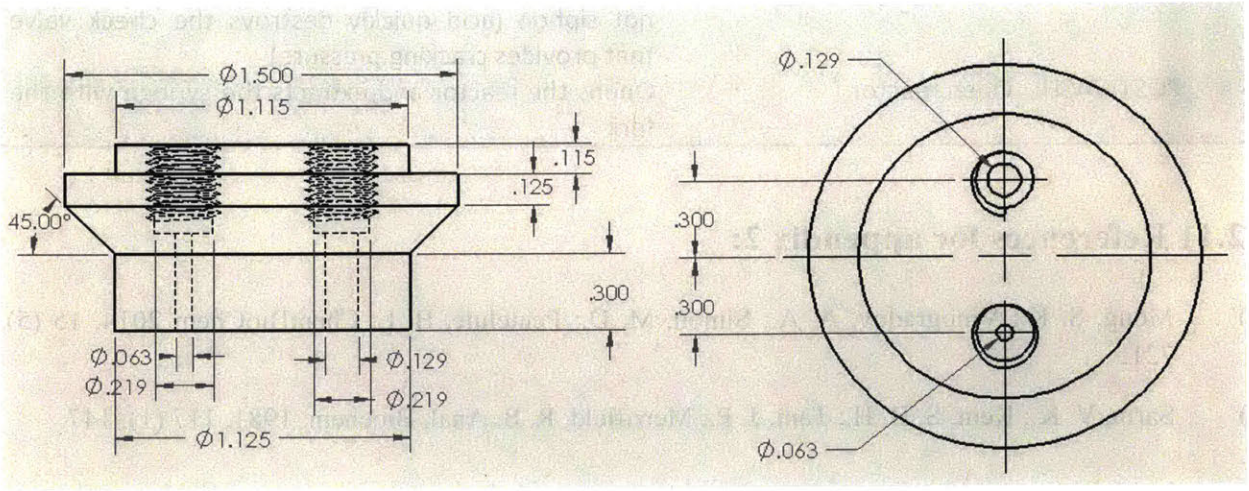


Figure A2.13.1: Drawing of the GL45 machined insert for nitrogen supply and reagent withdraw. Material was polypropylene.

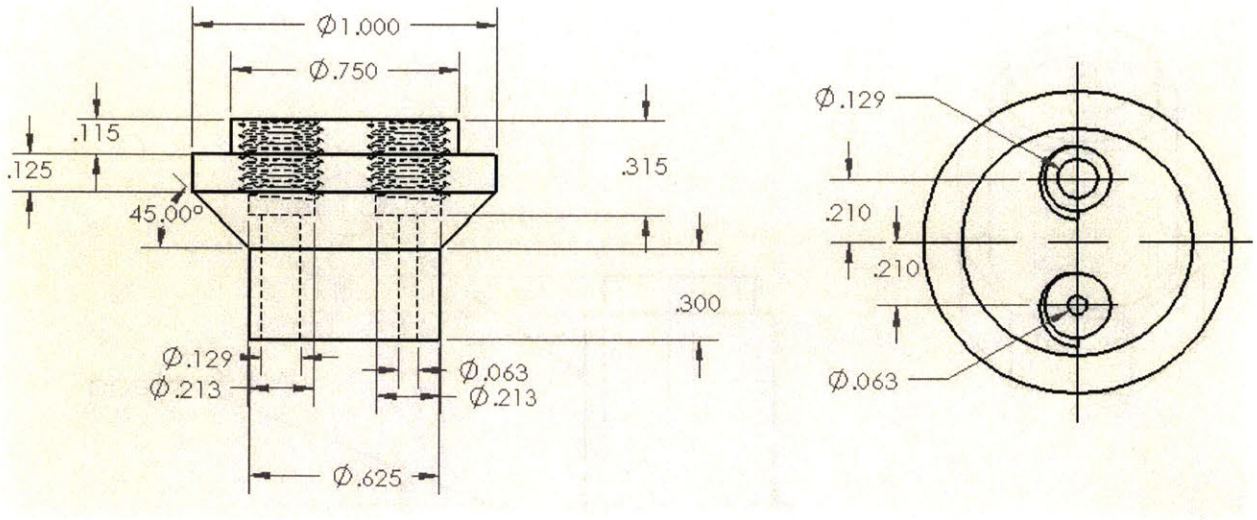


Figure A2.13.2: Drawing of the GL32 machined insert for nitrogen supply and reagent withdraw. Material was polypropylene.

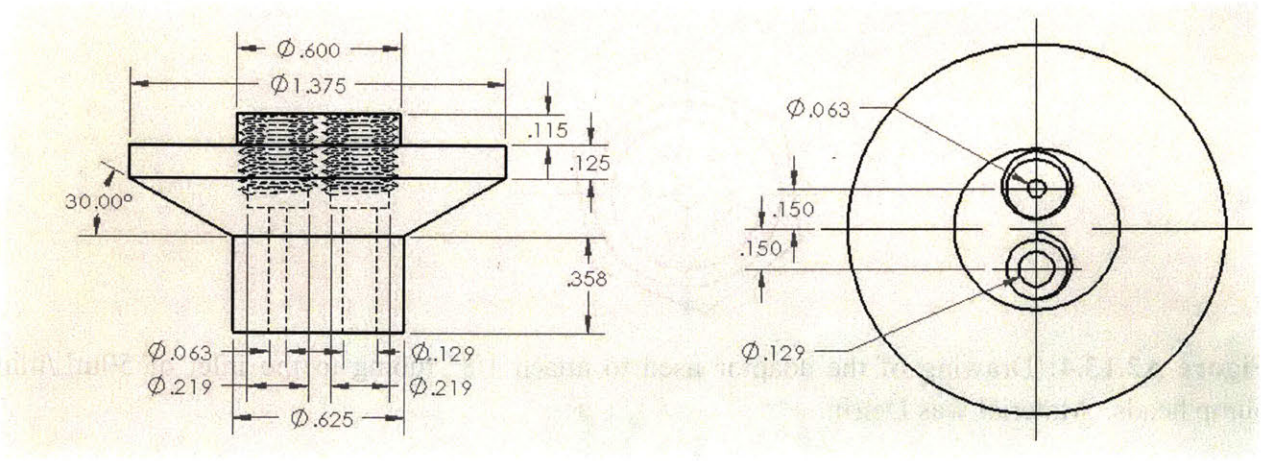


Figure A2.13.3: Drawing of the 38-400 machined insert for nitrogen supply and reagent withdraw. The angle specified as 30 degrees should be increased to at least 45 degrees; this design requires a lot of clamping force to seal. Note that headless knurreled nuts are required for the ¼-28 flatbottom fittings in this design because nuts with standard heads will collide with each other. Material was polypropylene.

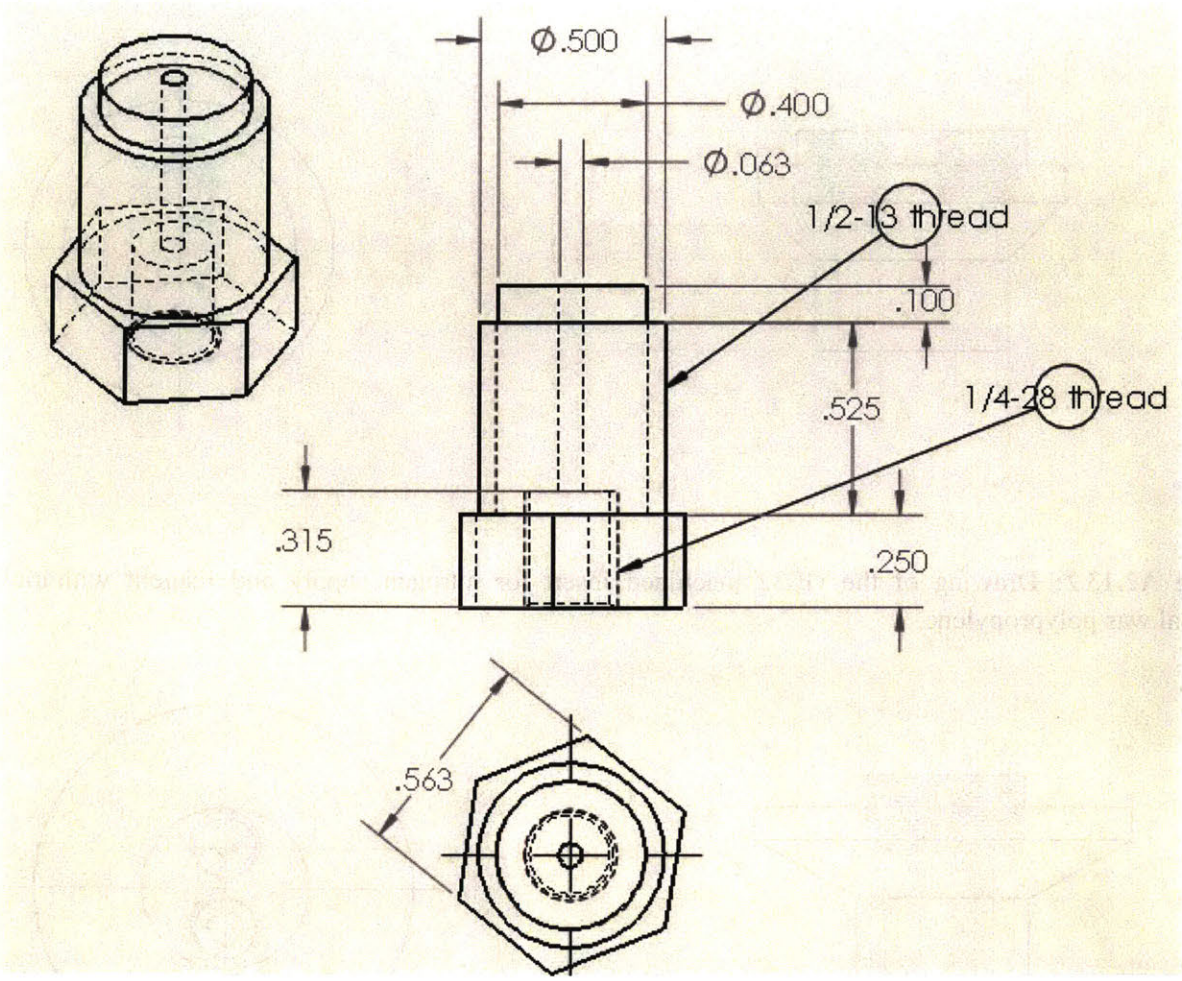


Figure A2.13.4: Drawing of the adaptor used to attach 1/8" tubing to the inlet of 50mL/min pump heads. Material was Delrin.

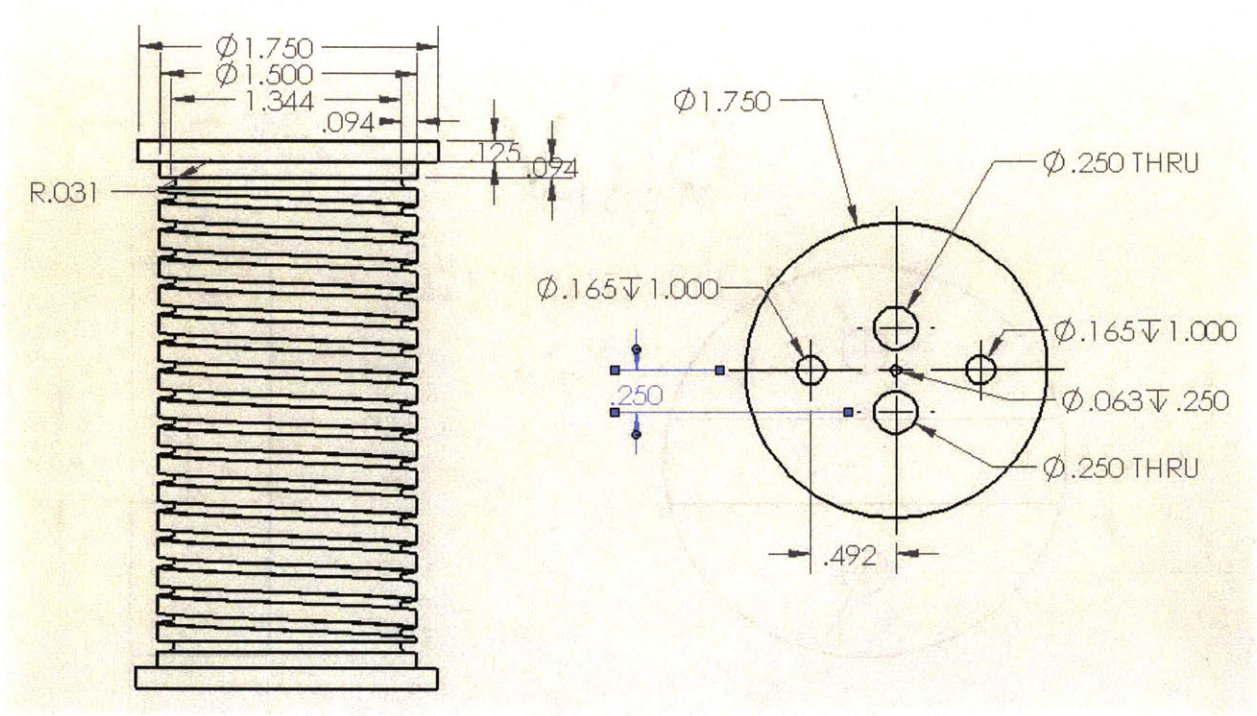


Figure A2.13.5: Drawing of an aluminum cylinder used to heat the tubular activation reactors. The pitch of the spiral and/or overall length was changed to accommodate shorter or longer stainless steel tubes. The center hole is for a thermocouple, the ¼" thru holes are for cartridge heaters, and the .165" OD holes were tapped M5 and used to mount the reactor on 25mm centers. Material was aluminum.

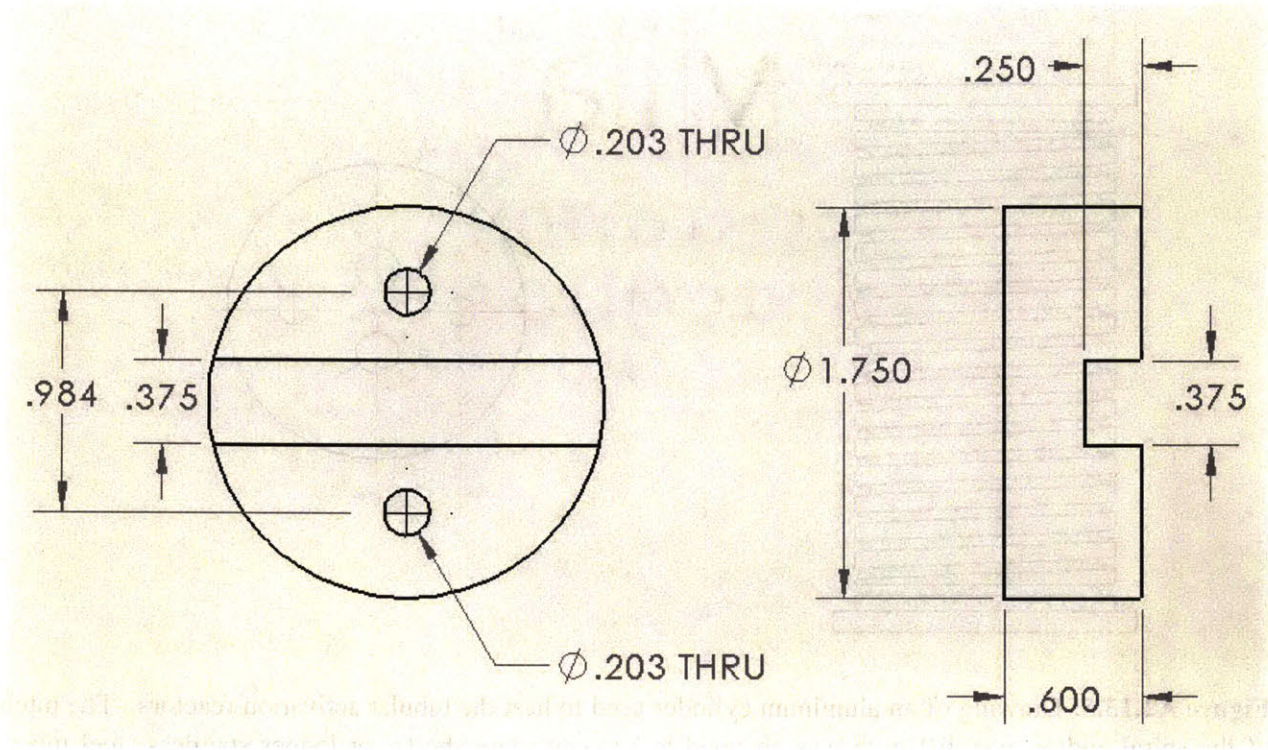


Figure A2.13.6: Mounting spacer for activation reactor. Material was PTFE.

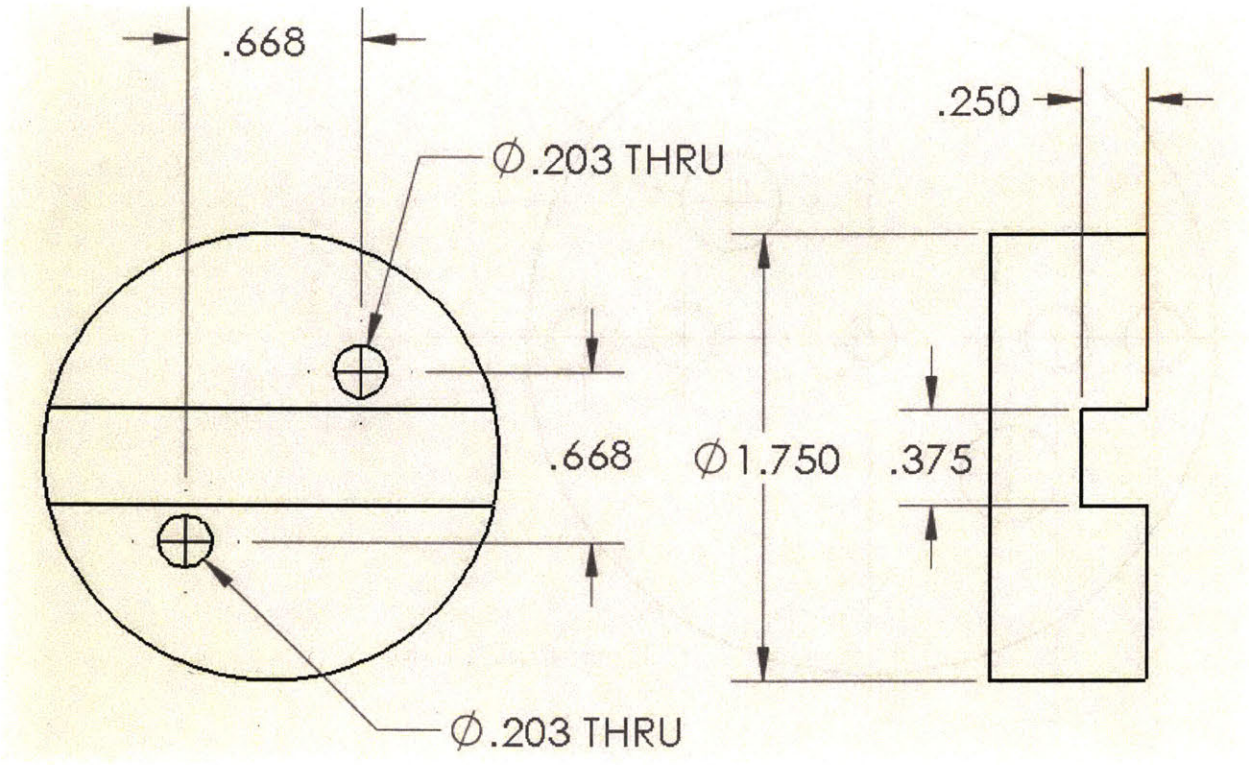


Figure A2.13.7: Mounting spacer for resin vessel inlet. Material was PTFE.

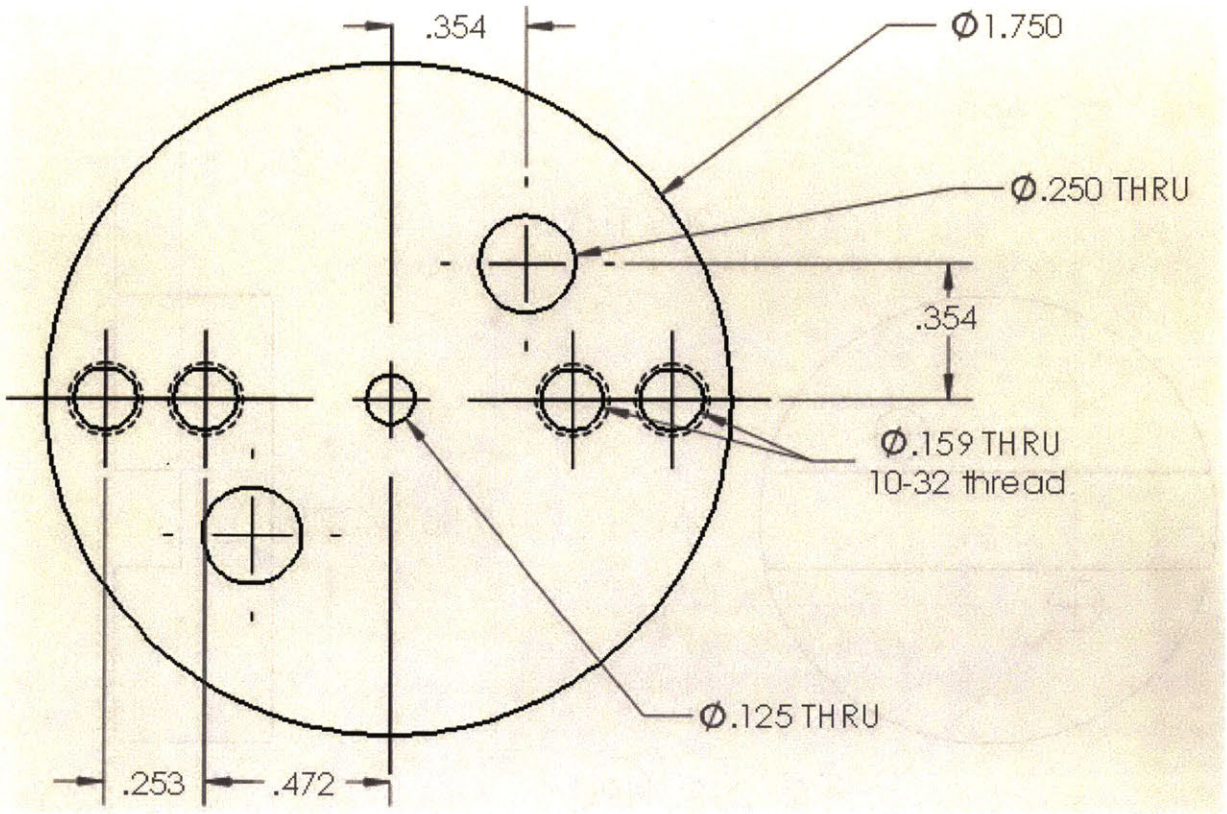


Figure A2.13.8: Universal mounting plate for resin vessel inlet. Accommodates all inlet sizes: 6mL, 12mL, and 24mL. Material was aluminum.

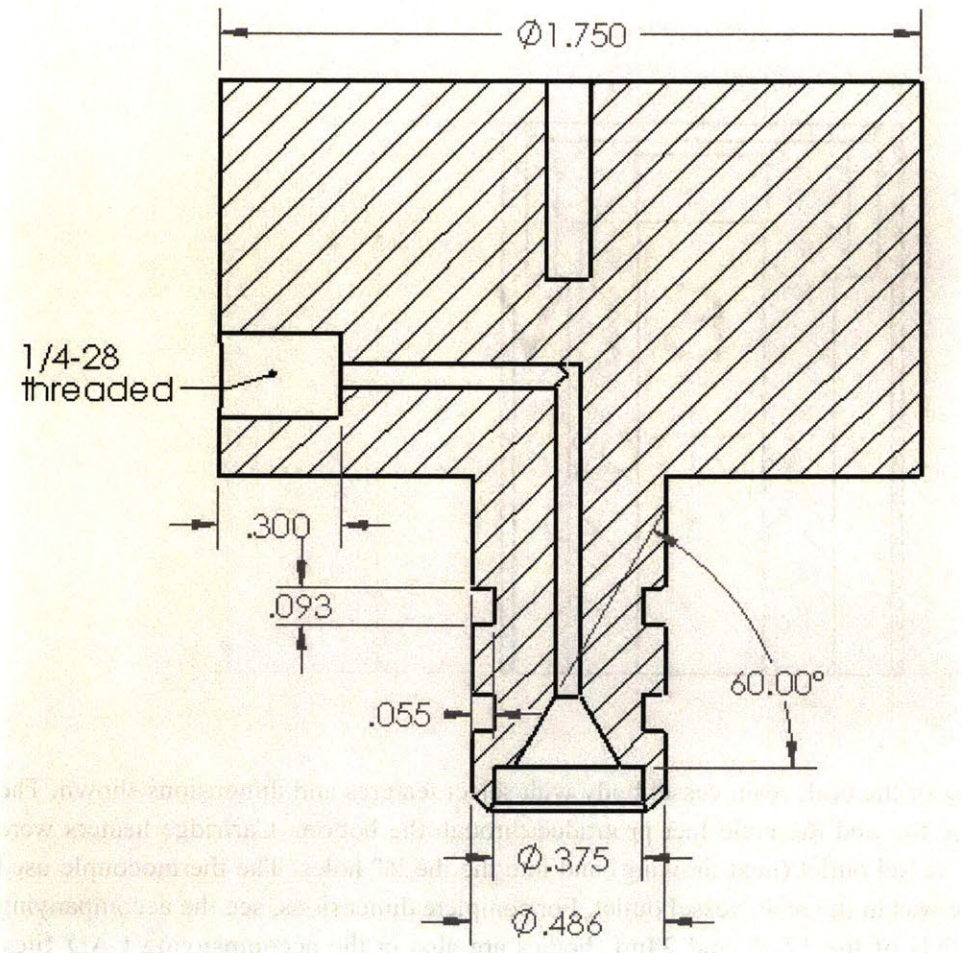


Figure A2.13.9: Drawing of the 6mL resin vessel inlet with select features and dimensions shown. For complete dimensions, see the accompanying CAD files. The inlet accepted a ¼-28 flatbottom fitting. The two o-ring glands used Kalrez O-rings (size 012) to seal against a nominal 0.5" ID 6mL syringe barrel. A 3/8" diameter stainless steel frit 1/8" thick was pressed into the outlet to help achieve plug flow in the syringe, and the 60 degree cone leading up to the frit further promoted plug flow in the syringe. The 12mL and 24mL inlets were conceptually similar, but had larger diameter cylinders on the outlet to seal larger diameter syringes with 014 and 017 o-rings, respectively. Complete models of the 12mL and 24mL inlets are in the accompanying CAD files. Material was aluminum.

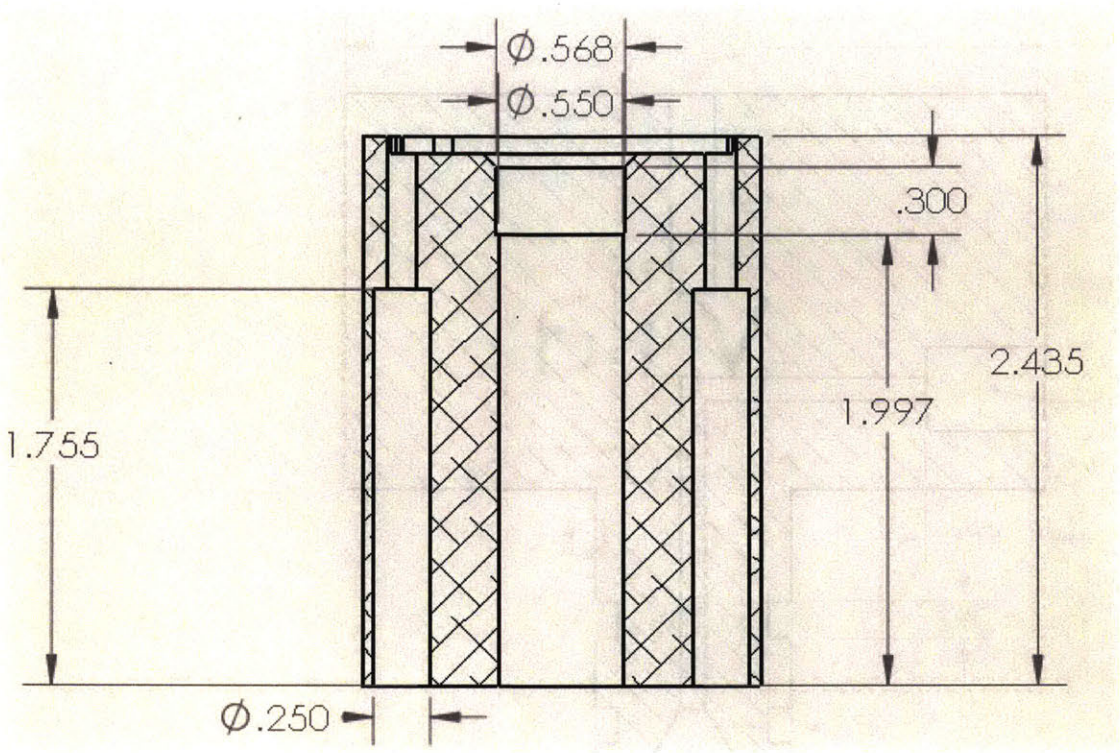


Figure A2.13.10: Drawing of the 6mL resin vessel body with select features and dimensions shown. The syringe was inserted in the top and the male luer protruded through the bottom. Cartridge heaters were inserted through the resin vessel outlet (next drawing) and into the the ¼" holes. The thermocouple used to monitor the temperature was in the resin vessel outlet. For complete dimensions, see the accompanying CAD files. Complete models of the 12mL and 24mL bodies are also in the accompanying CAD files. Material was aluminum.

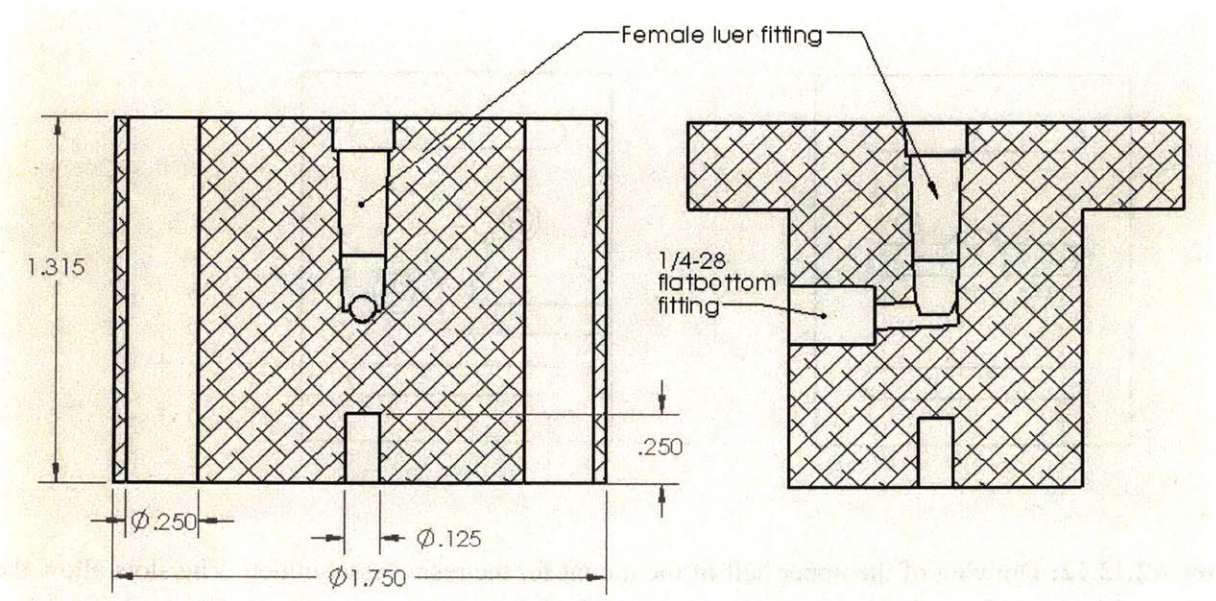


Figure A2.13.11: Drawing of the universal outlet for all resin vessel scales with select features and dimensions shown. The female luer fitting for the syringe tip was machined with a commercial luer reamer (Cockrill precision products) and may not be accurately drawn. The thermocouple to read the temperature of both the resin vessel outlet and body was inserted into the 1/8" hole on the bottom. ¼" cartridge heaters were inserted through the ¼" holes shown into the resin vessel body. No O-ring seal is present between the reactor body and outlet because the syringe fit tightly enough that residual solvent in the bottom of the resin vessel could pose a problem. It was preferred that solvent that accidentally collected in the resin vessel between syntheses be squeezed out the side as the syringe was lowered, rather than destroy the syringe or cause the resin vessel to fail to close. The outlet was connected to the waste line with a ¼-28 flatbottom fiting. The complete model can be found in the accompanying CAD drawings. Material was aluminium.

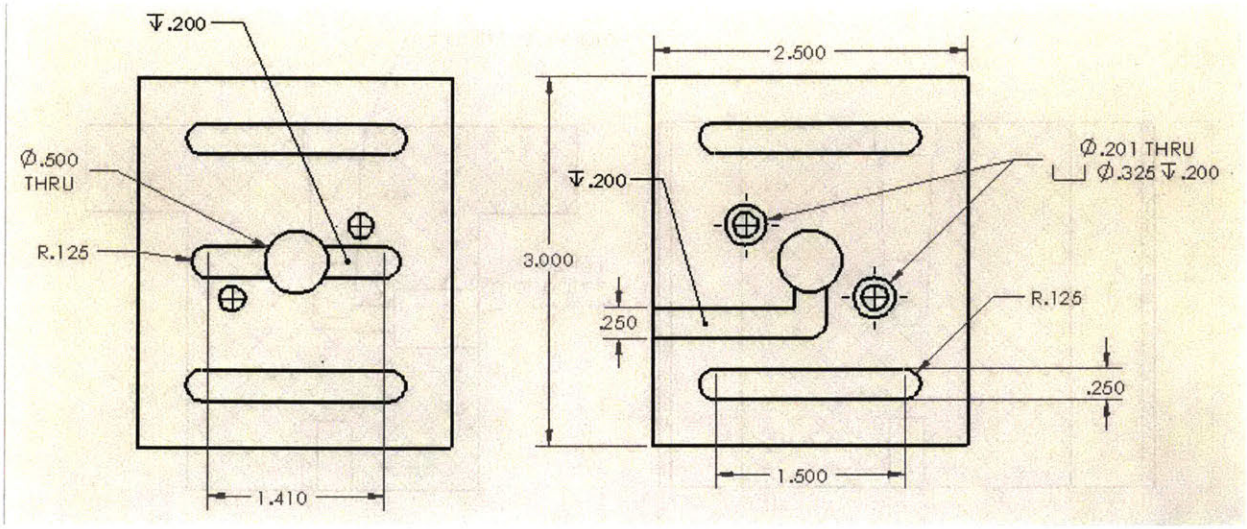


Figure A2.13.12: Drawing of the upper half of the mount for the resin vessel outlet. The slots allow the assembly with the resin vessel outlet to move in one direction. On the top side (left image) there is a clearance trench for heater and thermocouple wires. These were passed through a ½" hole and out through a clearance trench on the other side (right drawing). This pass through and exit from below was designed to protect the wires mechanically and chemically. Material was PTFE.

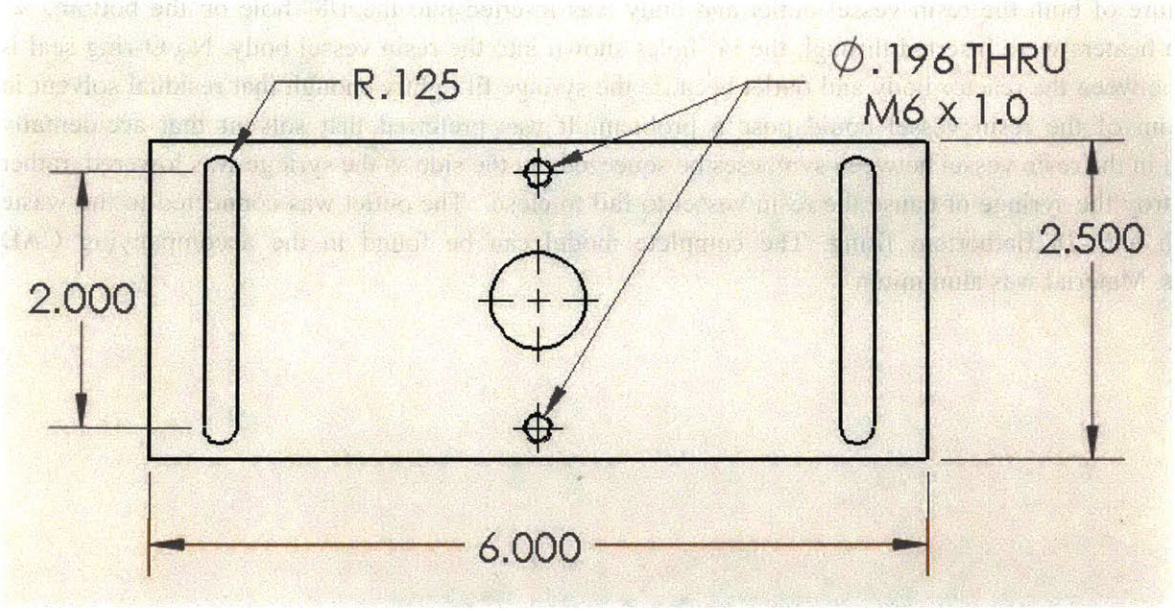


Figure A2.13.13: Drawing of the lower half of the mount for the resin vessel outlet. The slots allow the resin vessel outlet assembly to move in a second, orthogonal direction. The center hole had no purpose. Material was aluminum.

A2.14 NMR Spectra

All NMR spectra were collected on a Bruker Avance-600 NMR spectrometer with a Magnex Scientific superconducting actively-shielded magnet. Spectra were collected at room temperature and 1H chemical shifts were referenced from residual protic solvent.

All NMR experiments on the stability of reagents co-dissolved with HBTU were conducted by dissolution of commercial starting materials in per deuterated DMF for acquisition of 1H spectra. To simulate actual use conditions, commercial materials were not dried or otherwise repurified.

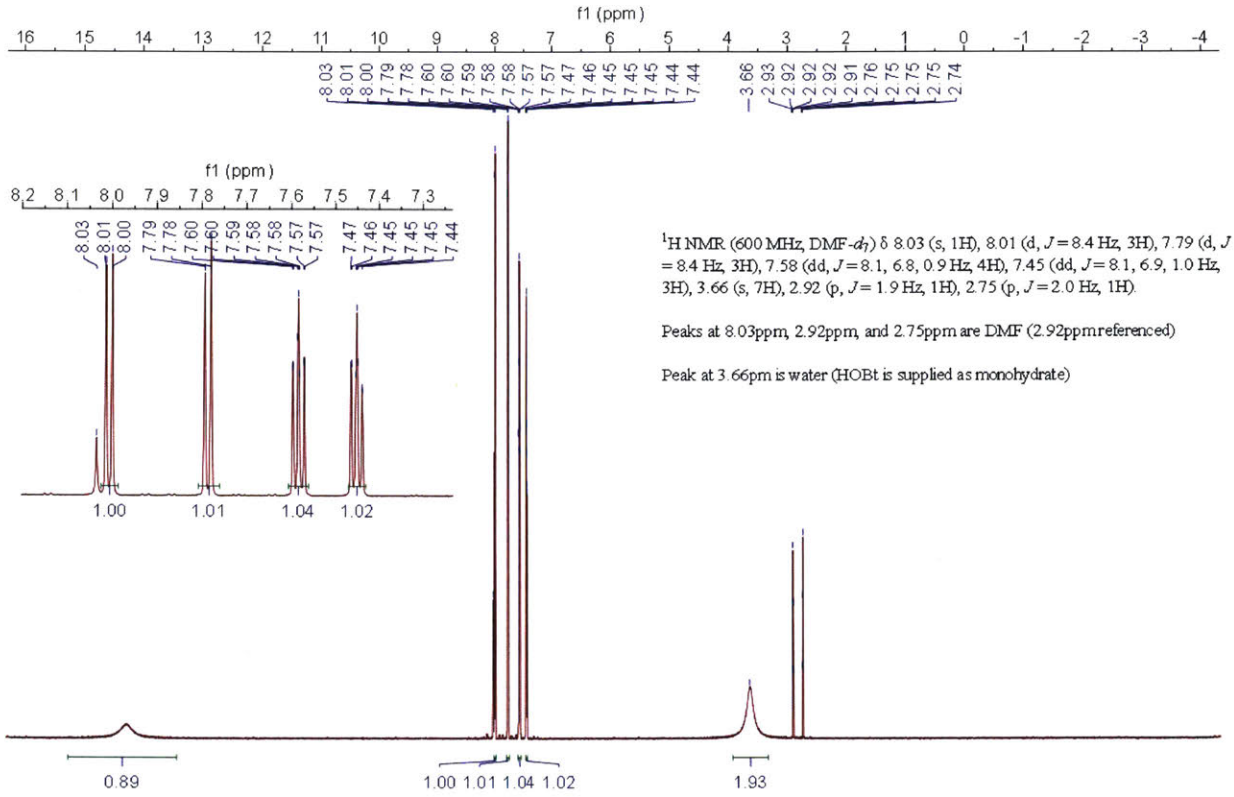
All precipitates from DMF and NMP were isolated by filtration, washed extensively with DMF and DCM, and dried under high vacuum prior to dissolution in deuterium oxide for acquisition of 1H and 13C spectra.

A2.14.1 Stability studies of amino acids co-dissolved with activators:

The stability of Fmoc-His(Trt)-OH codissolved with HBTU was studied as follows: HBTU (7mg, 0.031M) and Fmoc-His(Trt)-OH (14mg, 0.037M) were co-dissolved in 0.6mL of perdeuterated dimethylformamide. Spectra were acquired as soon as possible (about 5 minutes, recorded as t=0), after 30minutes, and after one day (Figures A2.14.4 – A2.14.6). Standards of HBTU, HOBt, and tetramethyl urea were collected for reference (Figures A2.14.1 – A2.14.3). The proportion of HBTU that had degraded was calculated by integrating the methyl protons in HBTU (3.65ppm and 3.27ppm) and comparing this signal to the methyl protons in tetramethyl urea (2.75ppm). The DMF signal at 2.76ppm overlapped with the tetramethyl urea peak, so they were integrated together and the area of the non-overlapping DMF peak (2.92ppm) was subtracted out.

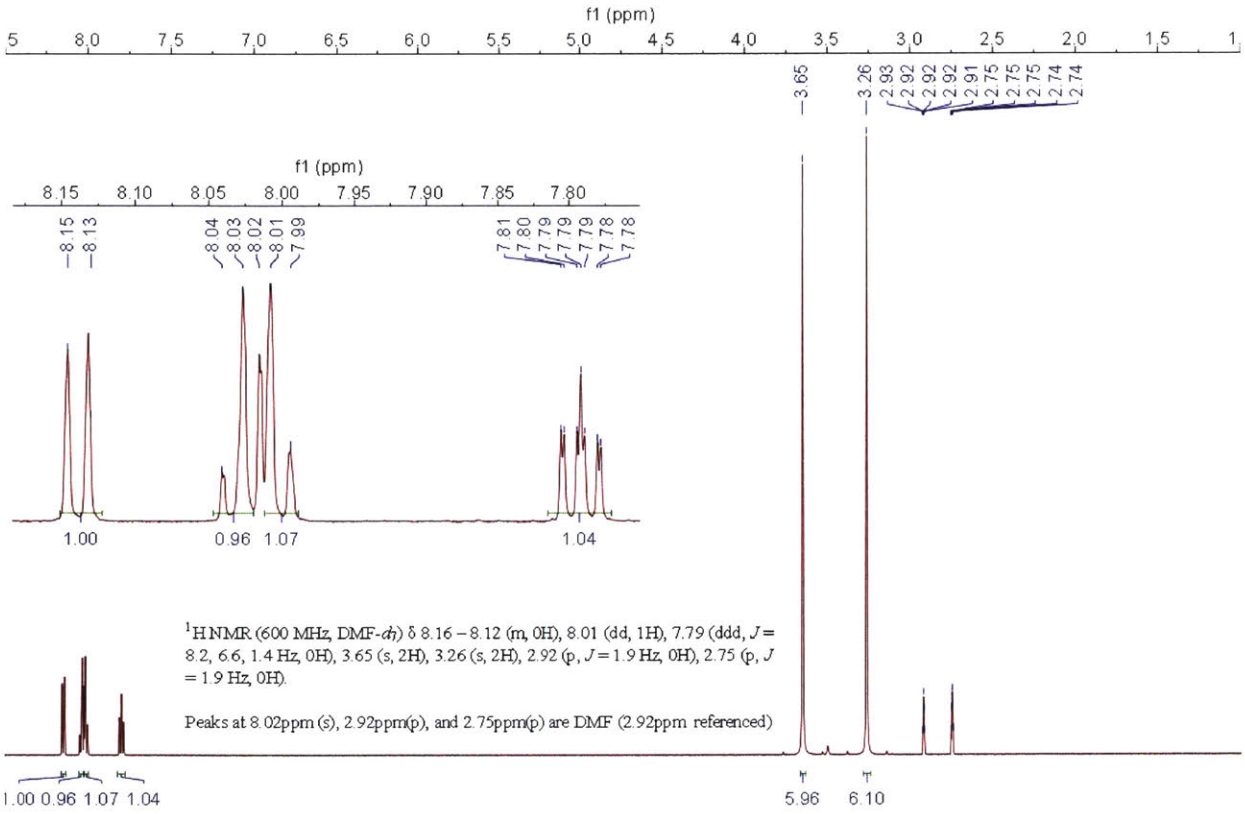
At t="0", the HBTU was 13% degraded, after 30 minutes, it was 23% degraded, and after 1 day it was 62% degraded.

To compare the relative stability of HBTU codissolved with Fmoc-His(Trt)-OH, Fmoc-His(Boc)-OH, and Fmoc-Ala-OH, HBTU (76mg, 0.33M) was codissolved with each amino acid (0.34M, 123mg His(Trt), 68mg Ala, or 95mg His(Boc)) in 0.6mL perdeuterated DMF. Each of the three samples was analyzed by NMR after 1 hour, and the proportion of HBTU that degraded was calculated as above. In the Fmoc-His(Trt)-OH sample, 23% had degraded. In the Fmoc-His(Boc)-OH sample, 3.4% had degraded, and in the Fmoc-Ala-OH sample, 0.35% of the HBTU had degraded. The data is displayed in figures A2.14.7 through A2.14.9.



049-172-HOBt.1.fid — 049-172-HOBt

Figure A2.14.1: 1H NMR spectrum of HOBt standard in DMF. HOBt is shipped as a monohydrate, giving rise to two water protons.



049-172-HBTU.1.fid — 049-172-HBTU

Figure A2.14.2: 1H NMR spectrum of HBTU standard in DMF. Note the well resolved methyl protons (3.65ppm and 3.27ppm)

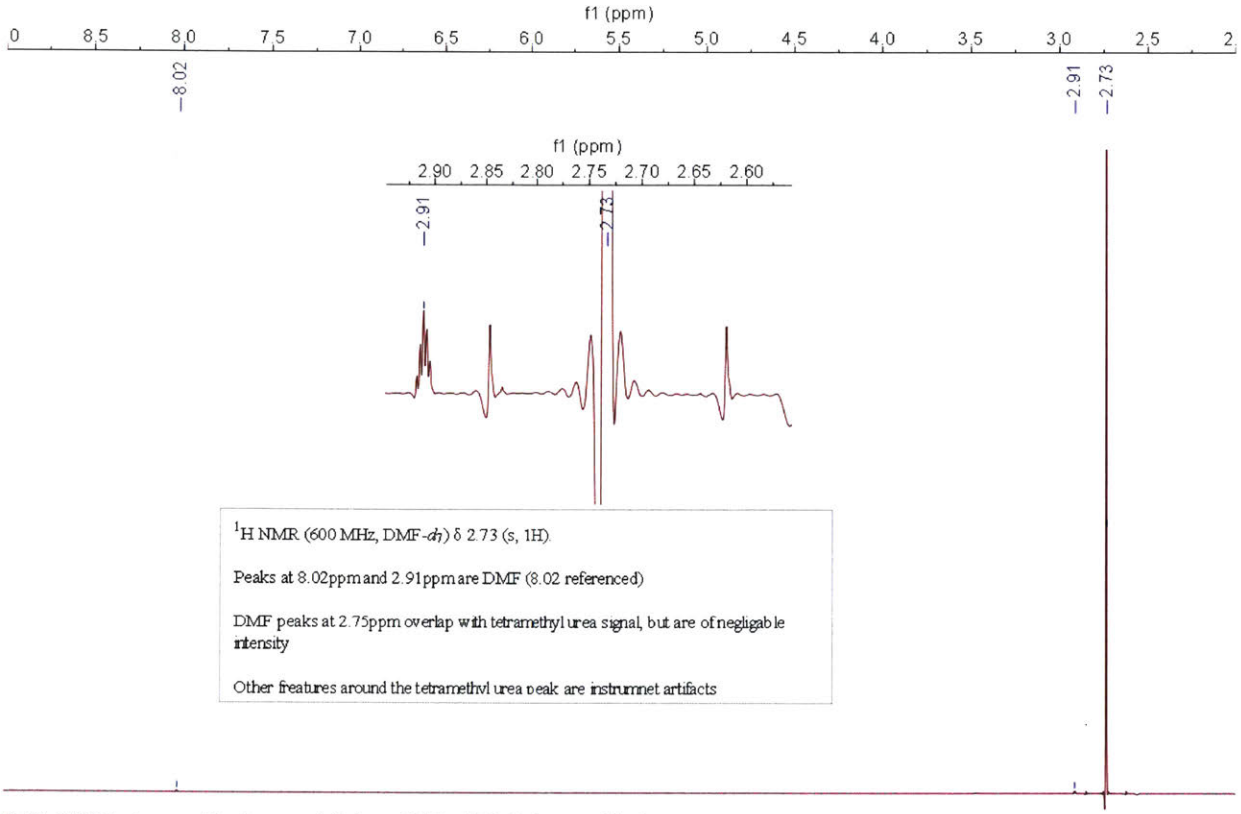


Figure A2.14.3: 1H NMR spectrum of tetramethyl urea standard in DMF. Methyl protons overlap with one DMF signal (2.75 ppm)

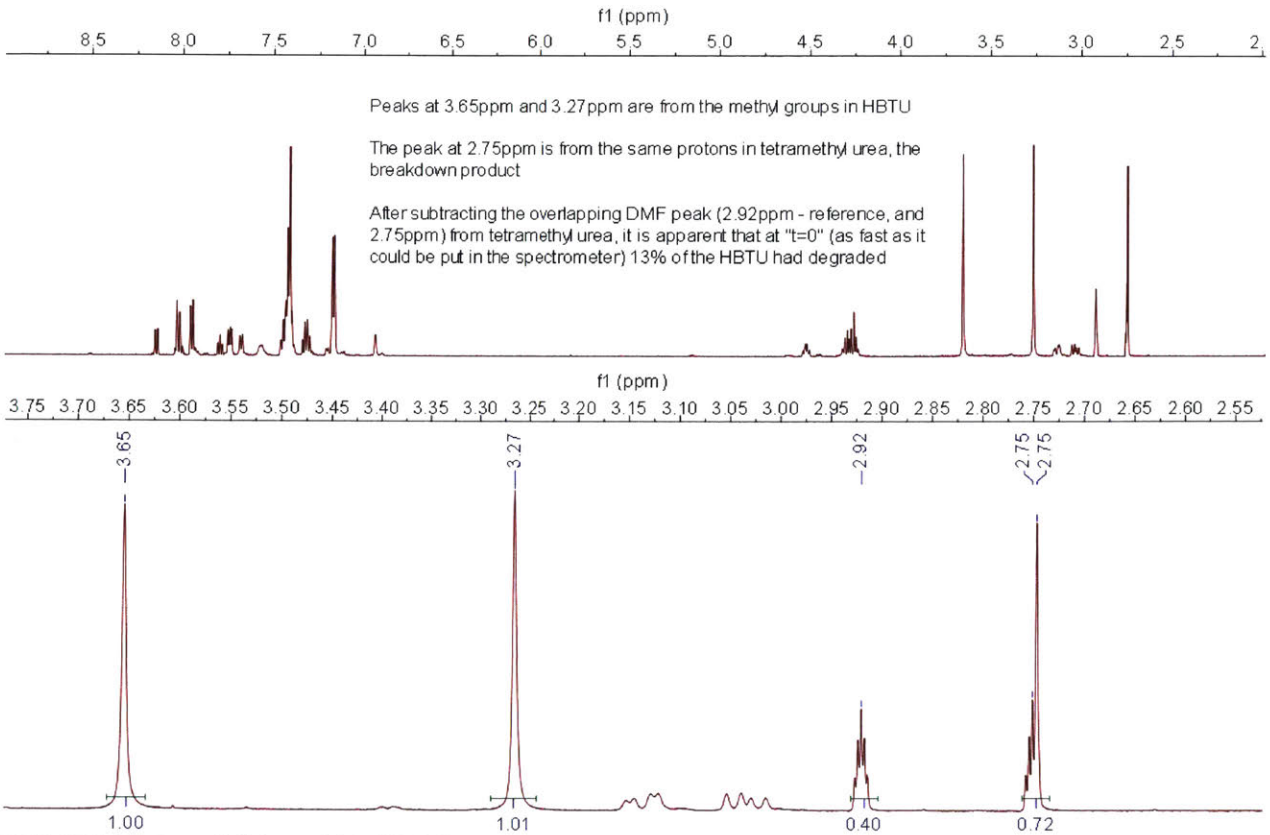


Figure A2.14.4: 1H NMR spectrum of Fmoc-His(Trt)-OH (0.037M) and HBTU (0.033M) codissolved in DMF acquired as soon as possible after dissolution. HBTU is 13% degraded.

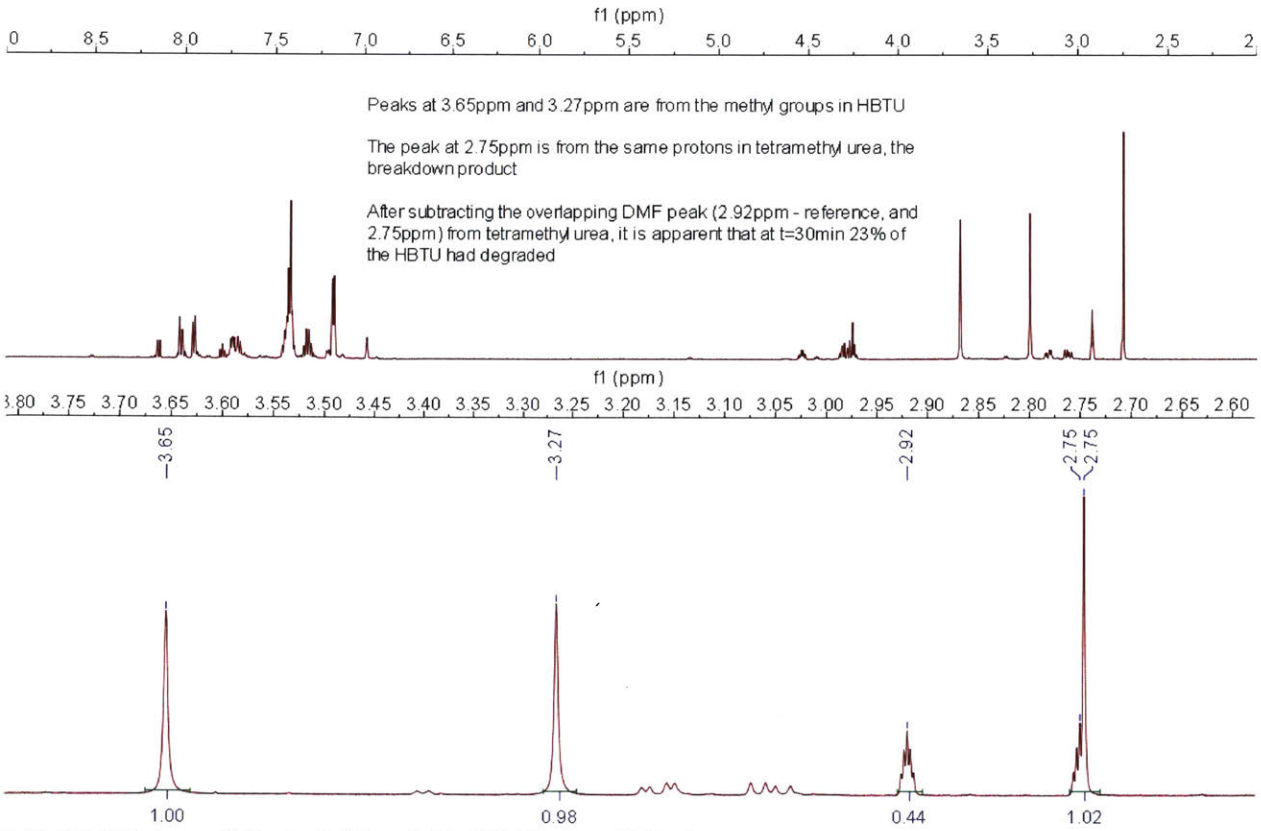


Figure A2.14.5: 1H NMR spectrum of Fmoc-His(Trt)-OH (0.037M) and HBTU (0.033M) codissolved in DMF after 30 minutes. HBTU is 23% degraded.

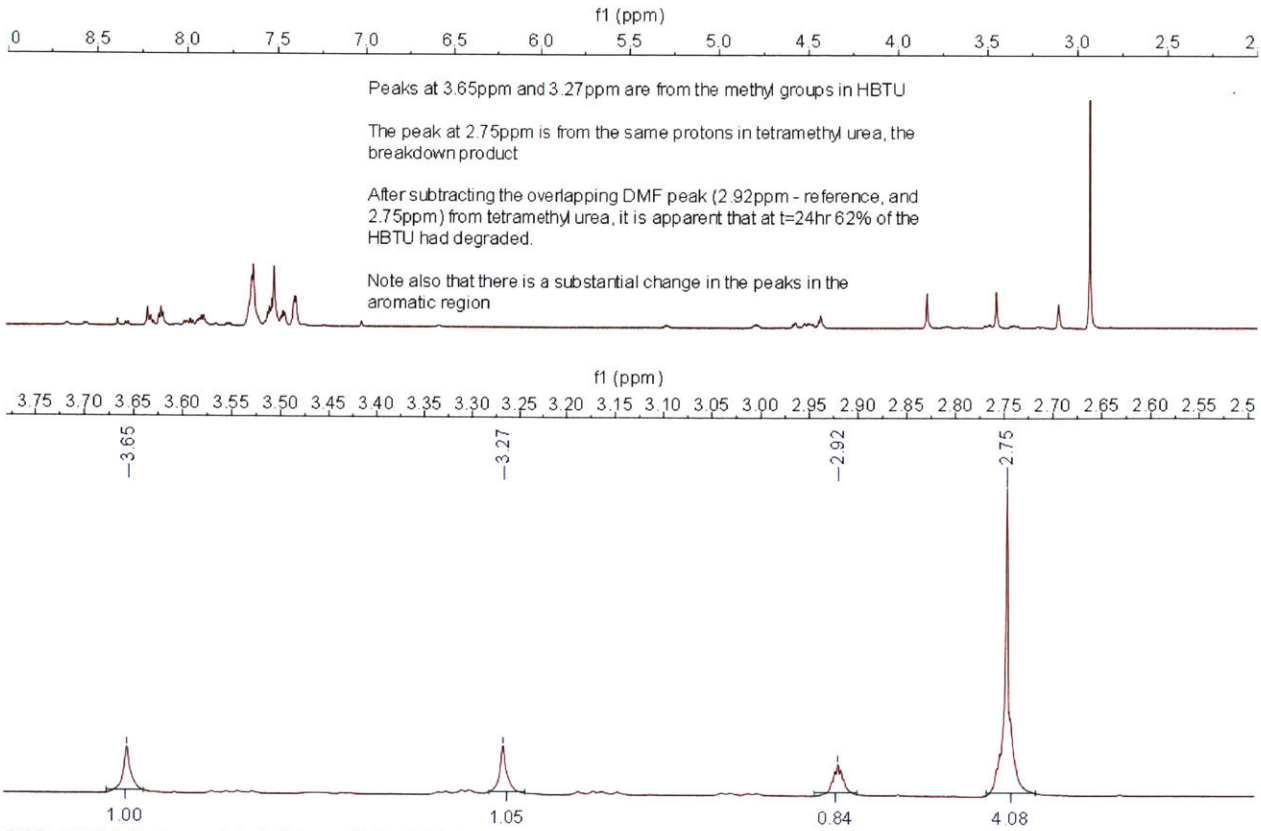
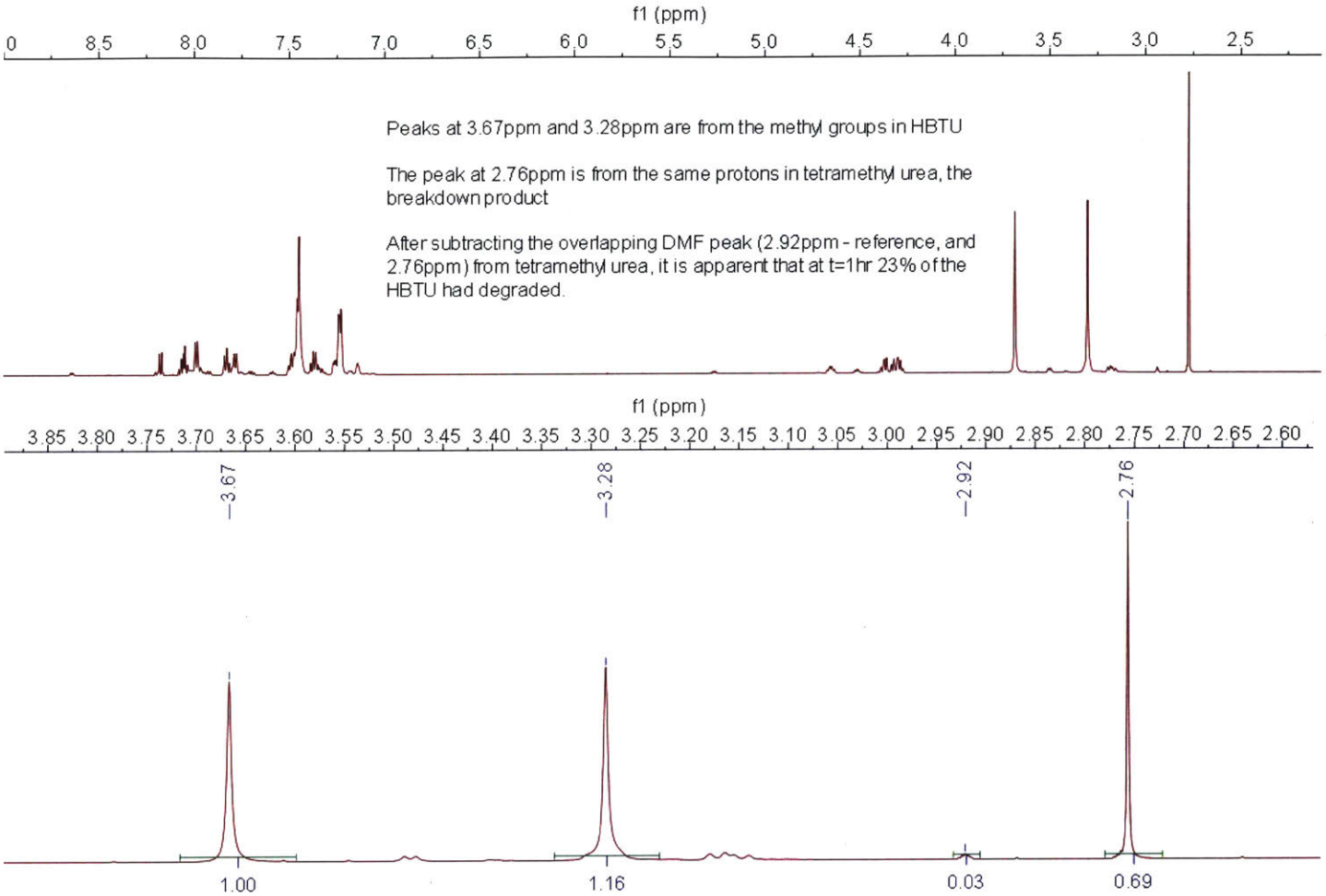
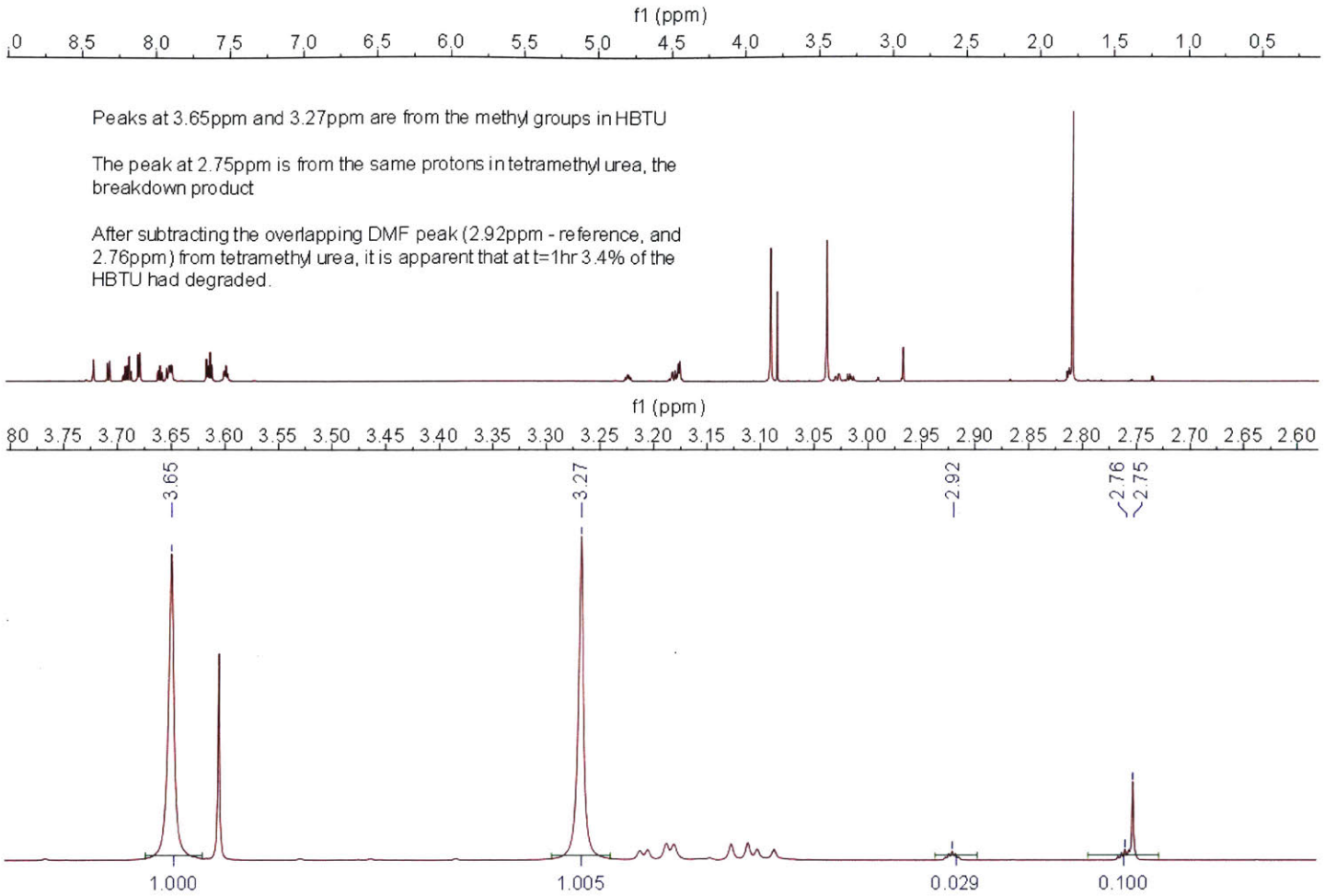


Figure A2.14.6: 1H NMR spectrum of Fmoc-His(Trt)-OH (0.037M) and HBTU (0.033M) codissolved in DMF after 24 hours. HBTU is 62% degraded.



049-172-His(Trt)-one.1.fid — 049-172-His(Trt)-one

Figure A2.14.7: 1H NMR spectrum of Fmoc-His(Trt)-OH (0.33M) and HBTU (0.34M) codissolved in DMF after 1 hour. HBTU is 23% degraded.



049-172-His(Boc)-one.1.fid — 049-172-His(Boc)-one

Figure A2.14.8: 1H NMR spectrum of Fmoc-His(Boc)-OH (0.33M) and HBTU (0.34M) codissolved in DMF after 1 hour. HBTU is 3.4% degraded.

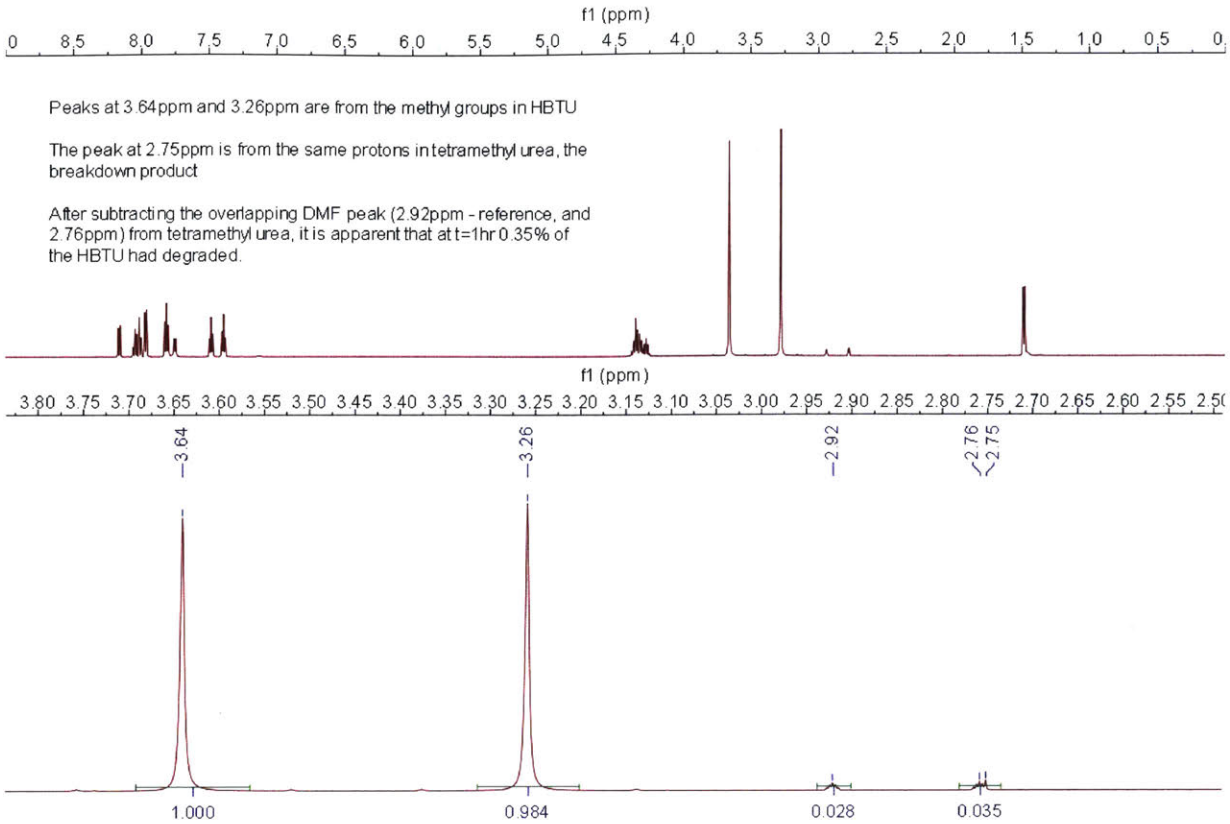


Figure A2.14.9: 1H NMR spectrum of Fmoc-Ala-OH (0.33M) and HBTU (0.34M) codissolved in DMF after 1 hour. HBTU is 0.35% degraded.

A2.14.2 Spectra of precipitates isolated from DMF:

The following spectra were obtained from precipitates isolated from solutions of Fmoc protected amino acids in DMF after four weeks at room temperature. In all cases, the Fmoc protecting group has been lost.

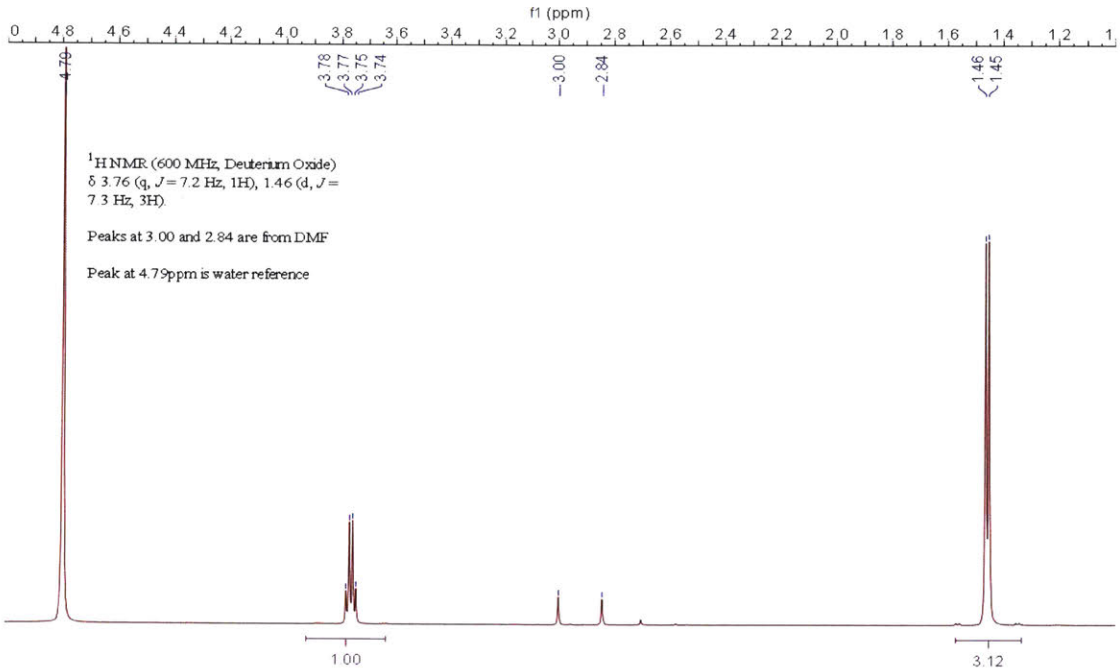


Figure A2.14.10: 1H spectrum of alanine, isolated from a solution of Fmoc-Ala-OH in DMF

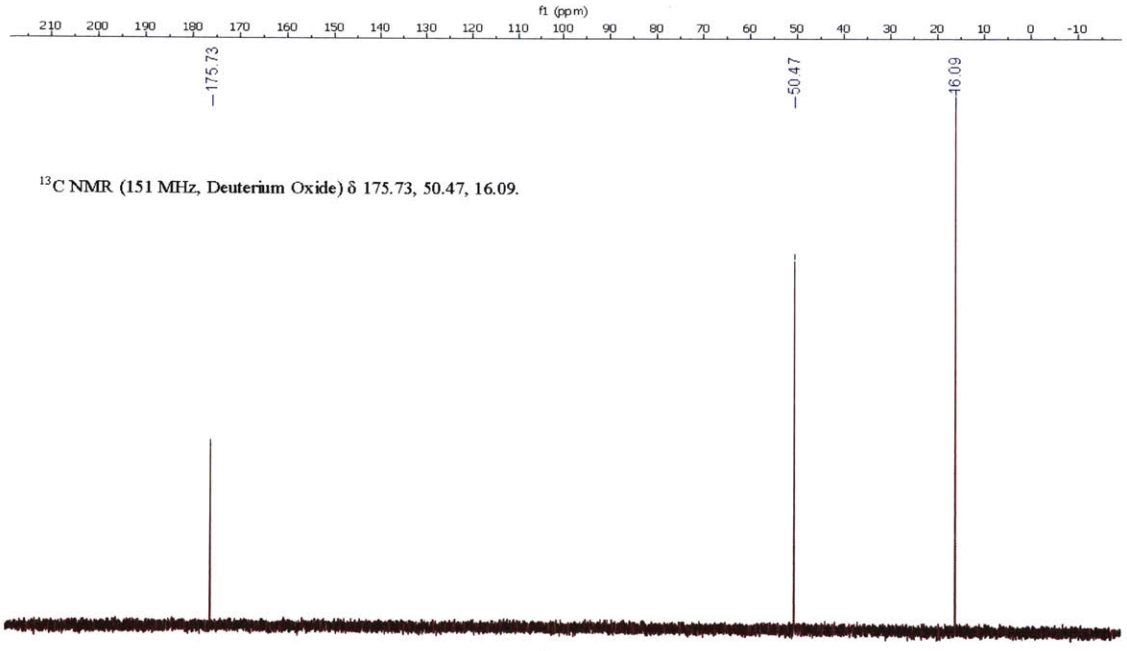
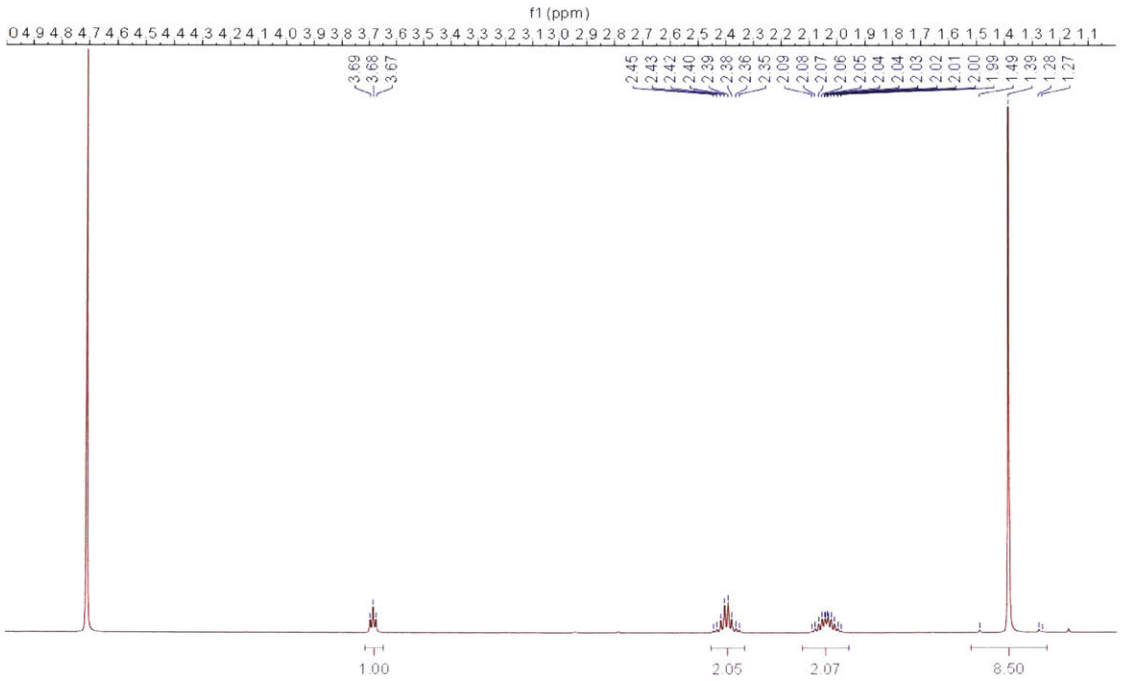
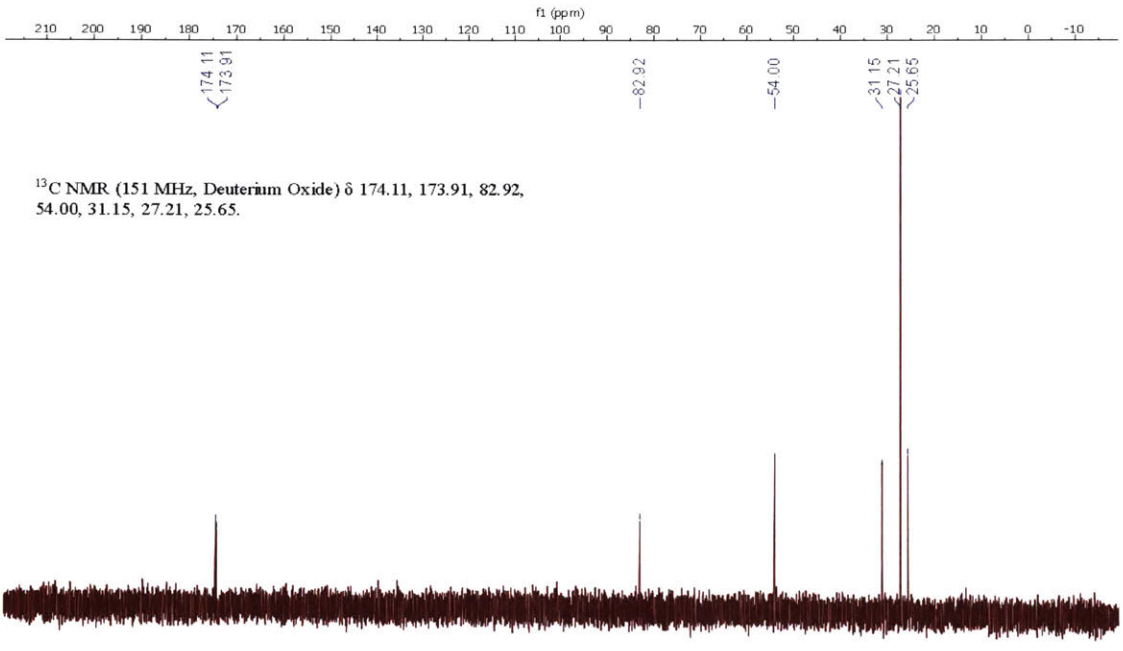


Figure A2.14.11: 13C spectrum of alanine, isolated from a solution of Fmoc-Ala-OH in DMF



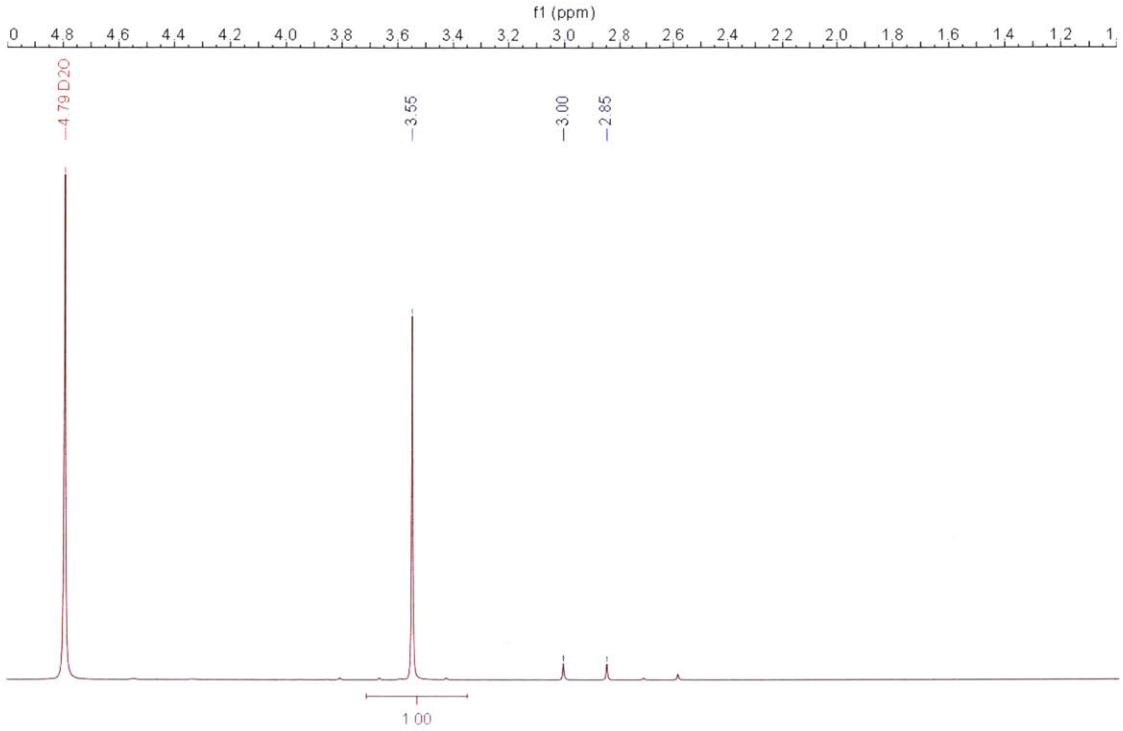
073-181 Glu DMF 1H_1.1.fid —

Figure A2.14.12: 1H spectrum of γ t-butyl ester of glutamic acid, isolated from a solution of Fmoc-Glu(tBu)-OH in DMF



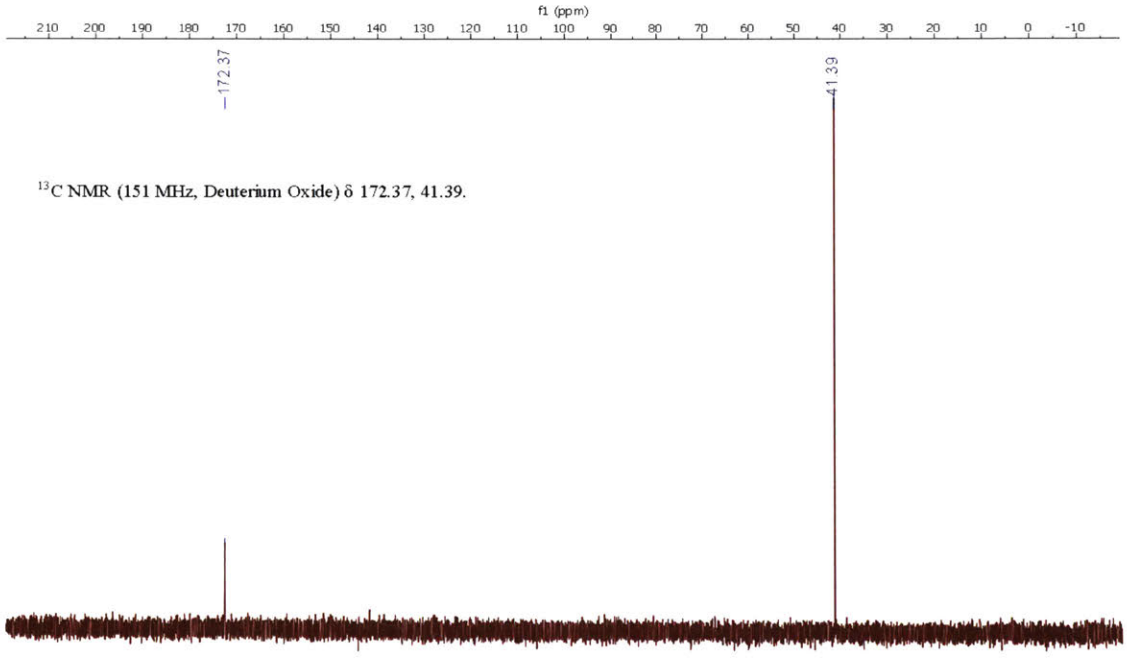
073-181 Glu DMF 13C_1.1.fid —

Figure A2.14.13: 13C spectrum of γ t-butyl ester of glutamic acid, isolated from a solution of Fmoc-Glu(tBu)-OH in DMF



073-181 Gly DMF 1H_1.1.fid —

Figure A2.14.14: 1H spectrum of glycine, isolated from a solution of Fmoc-Gly-OH in DMF



073-181 Gly DMF 13C_1.1.fid —

Figure A2.14.15: 13C spectrum of glycine, isolated from a solution of Fmoc-Gly-OH in DMF

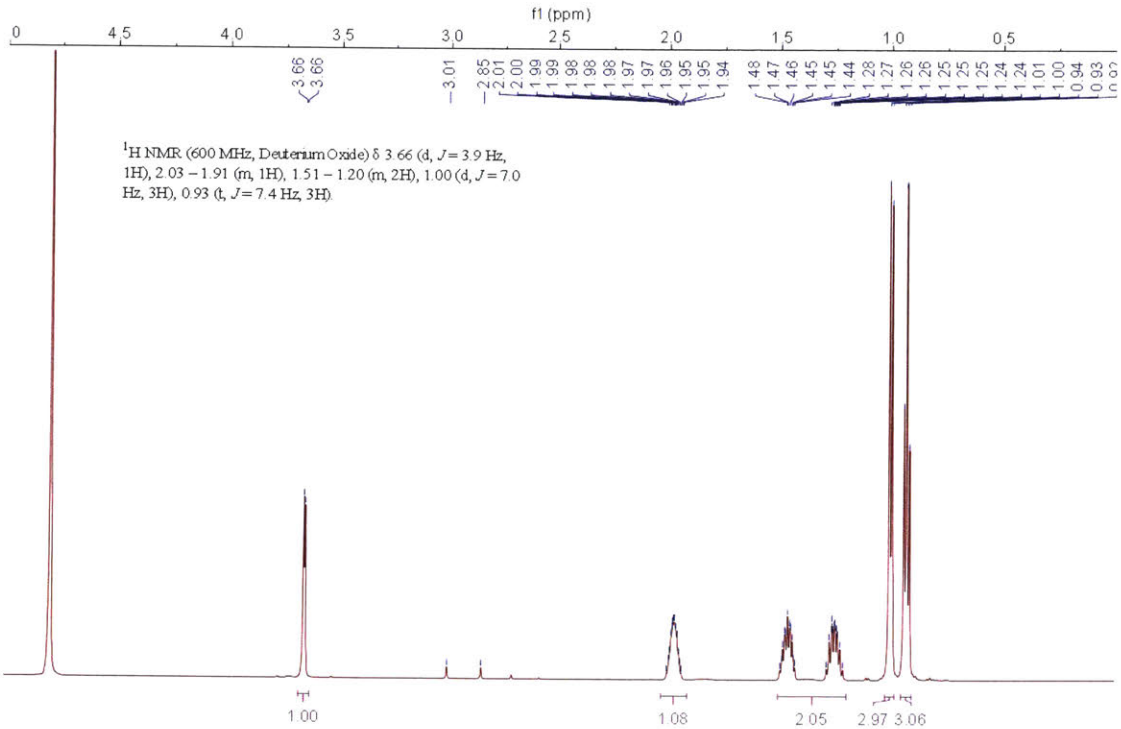


Figure A2.14.16: 1H spectrum of isoleucine, isolated from a solution of Fmoc-Ile-OH in DMF

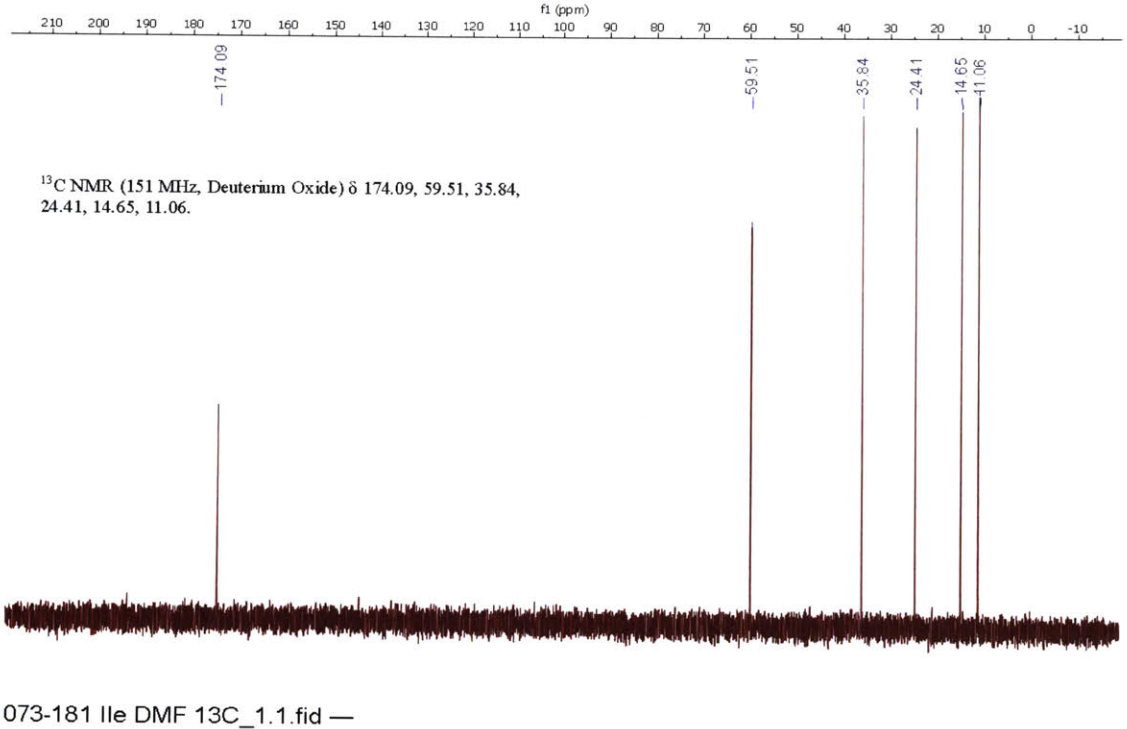
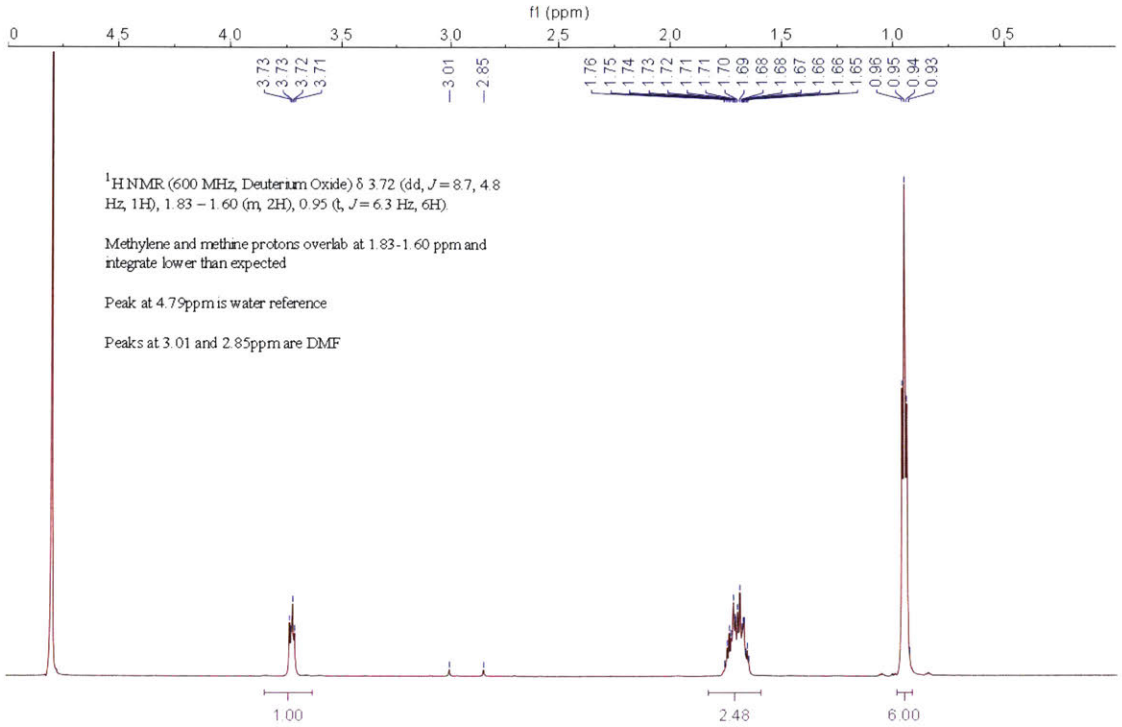
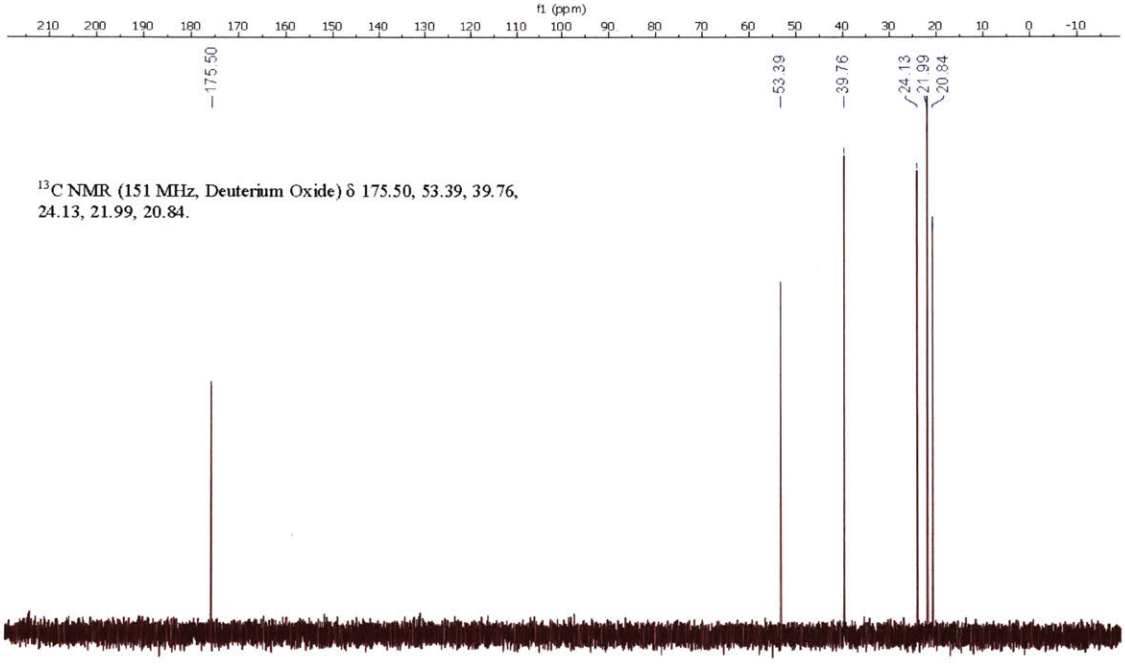


Figure A2.14.17: 13C spectrum of isoleucine, isolated from a solution of Fmoc-Ile-OH in DMF



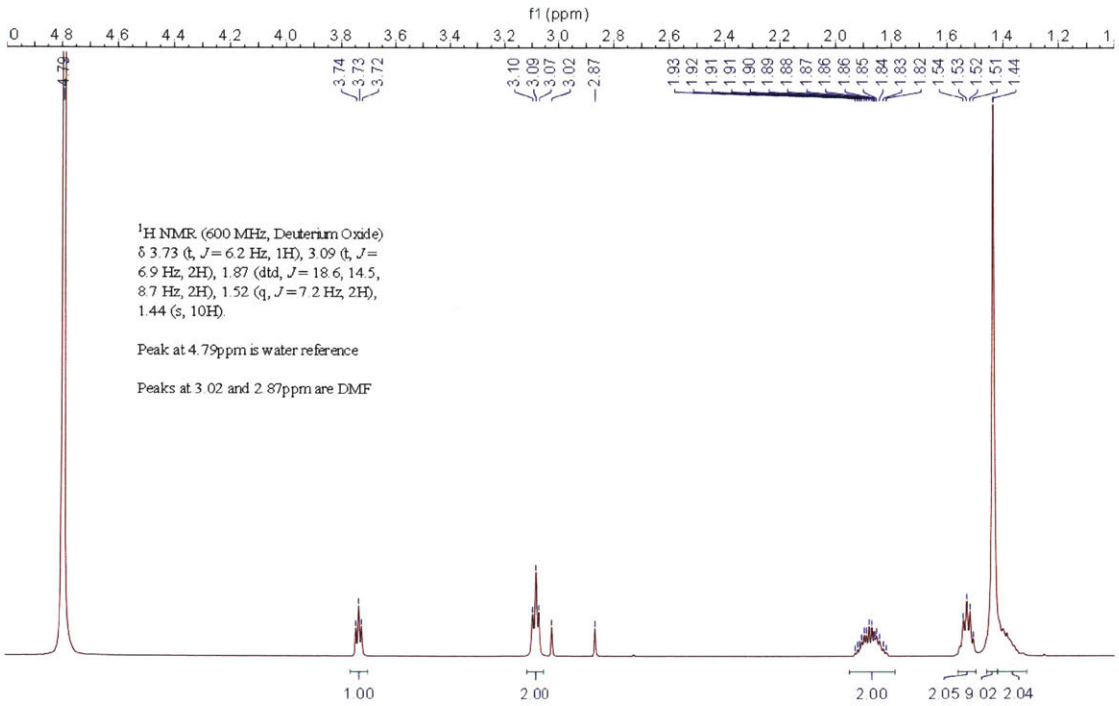
073-181 Leu DMF 1H_1.1.fid —

Figure A2.14.18: 1H spectrum of leucine, isolated from a solution of Fmoc-Leu-OH in DMF



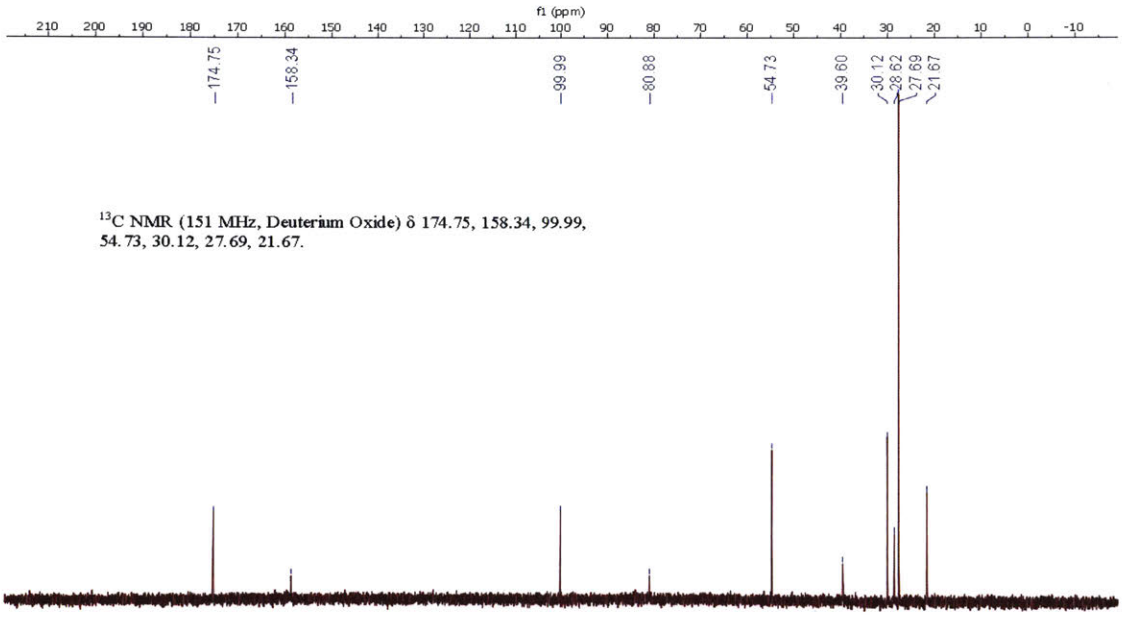
073-181 Leu DMF 13C_1.1.fid —

Figure A2.14.19: 13C spectrum of leucine, isolated from a solution of Fmoc-Leu-OH in DMF



073-181 Lys DMF 1H_1.1.fid —

Figure A2.14.20: 1H spectrum of N-ε Boc lysine, isolated from a solution of Fmoc-Lys(Boc)-OH in DMF



073-181 Lys DMF 13C_1.1.fid —

Figure A2.14.21: 13C spectrum of N-ε Boc lysine, isolated from a solution of Fmoc-Lys(Boc)-OH in DMF

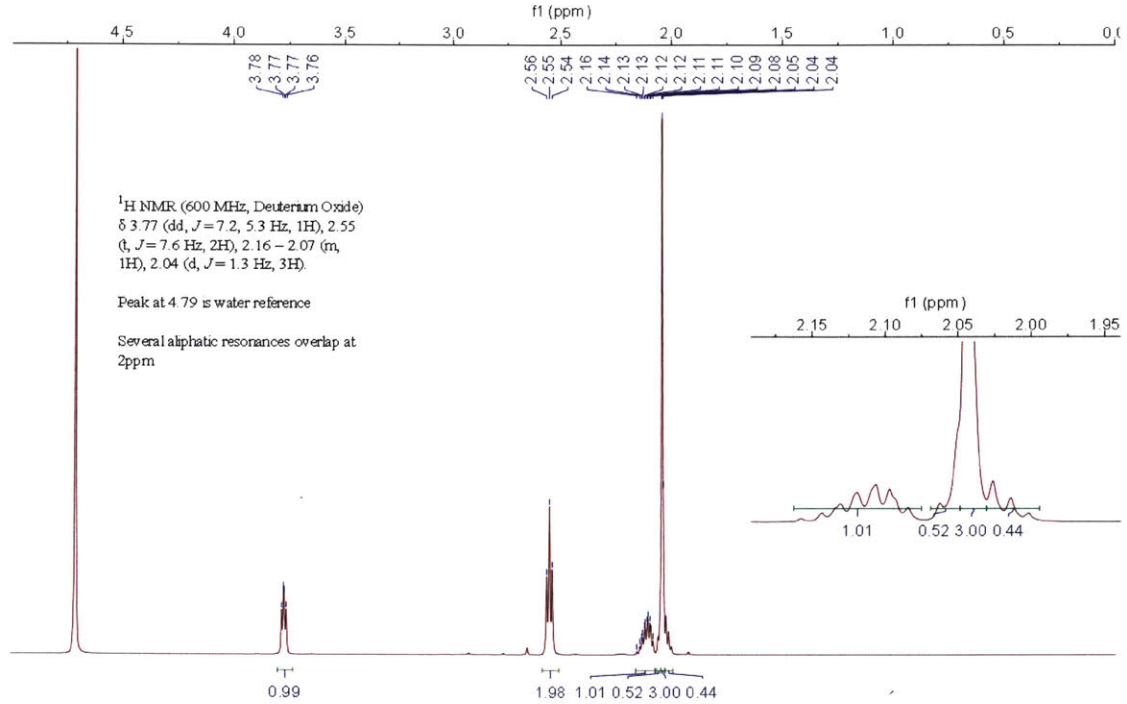


Figure A2.14.22: 1H spectrum of methionine, isolated from a solution of Fmoc-Met-OH in DMF

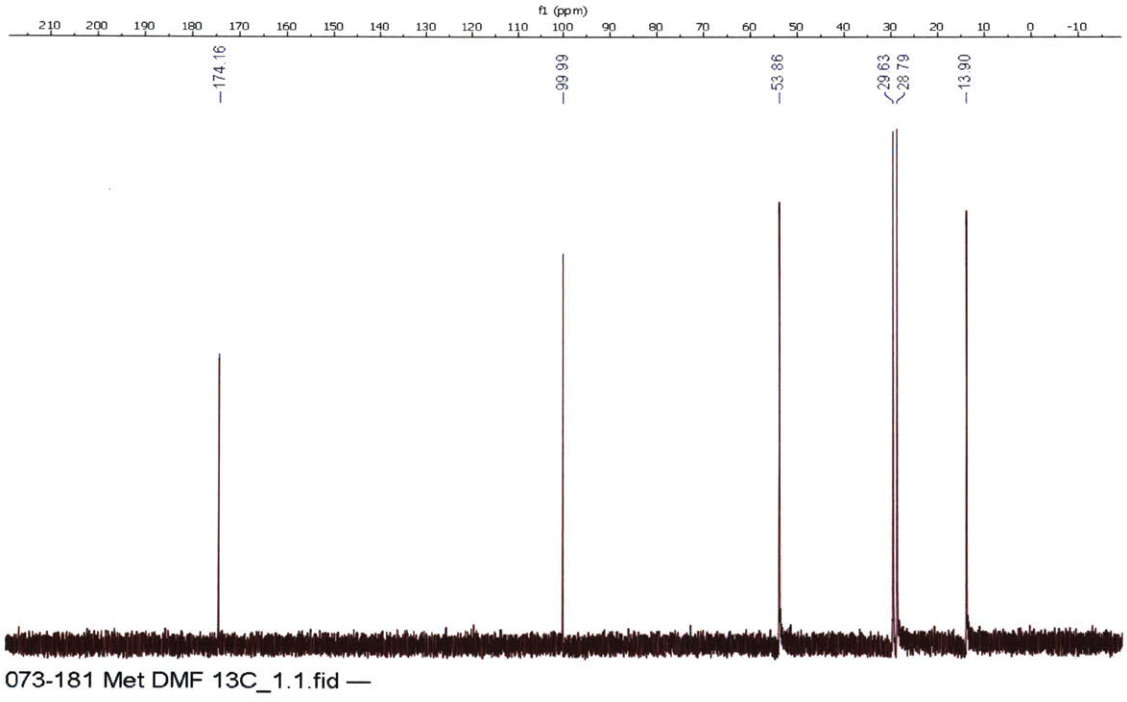


Figure A2.14.23: 13C spectrum of methionine, isolated from a solution of Fmoc-Met-OH in DMF

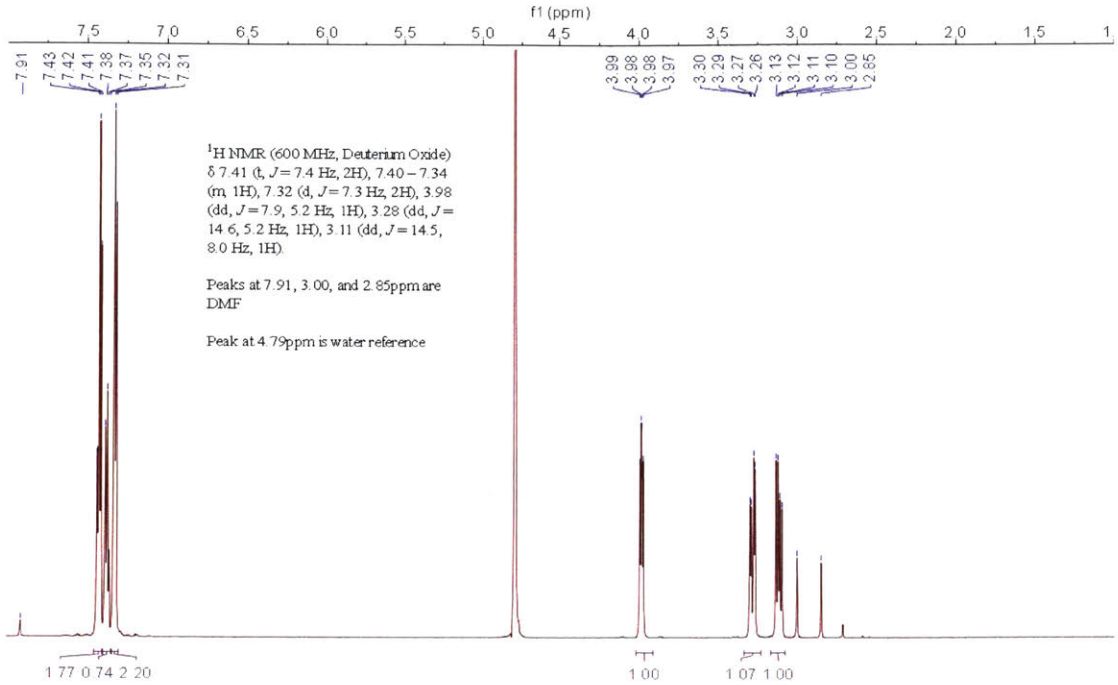


Figure A2.14.24: 1H spectrum of phenylalanine, isolated from a solution of Fmoc-Phe-OH in DMF

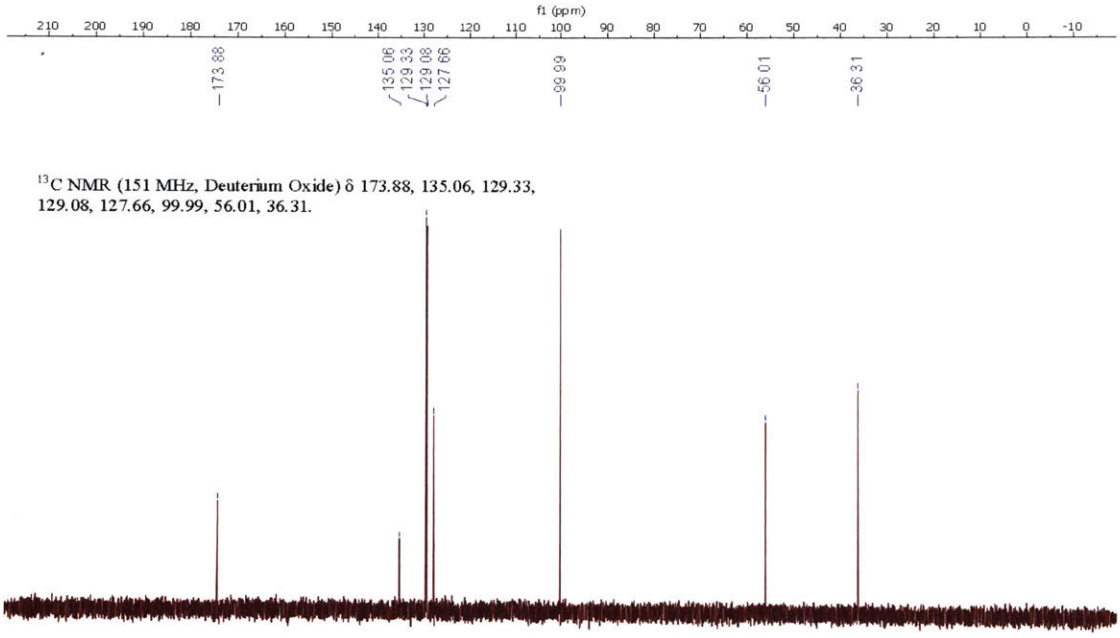
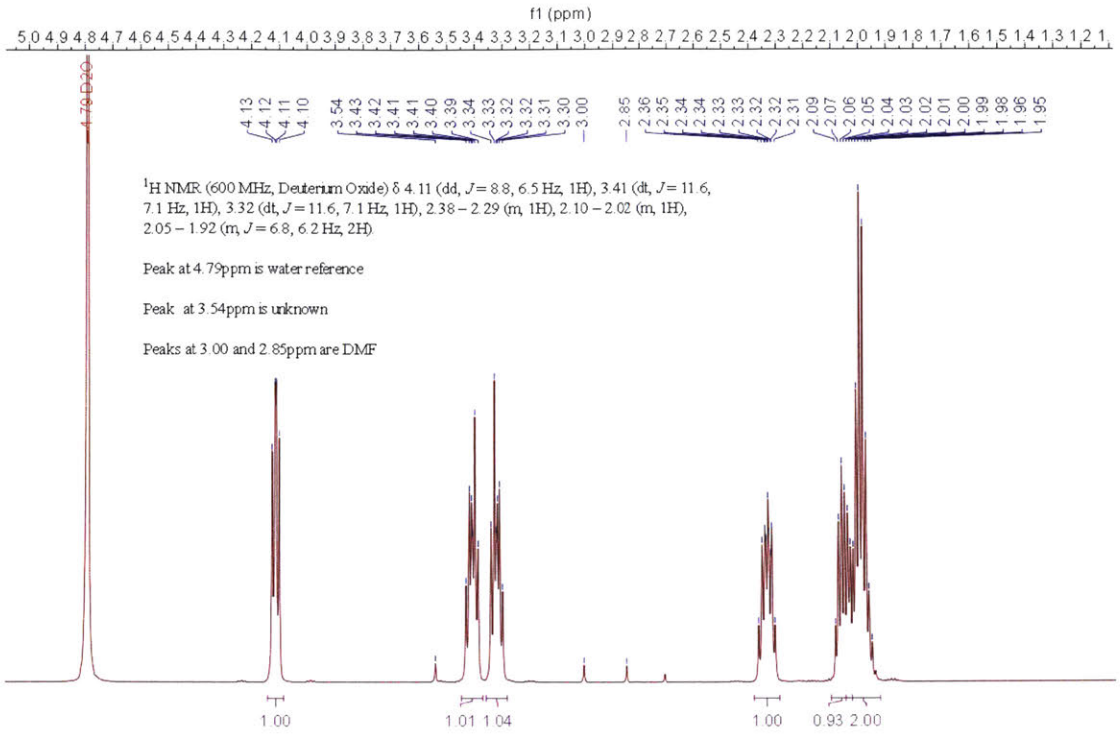
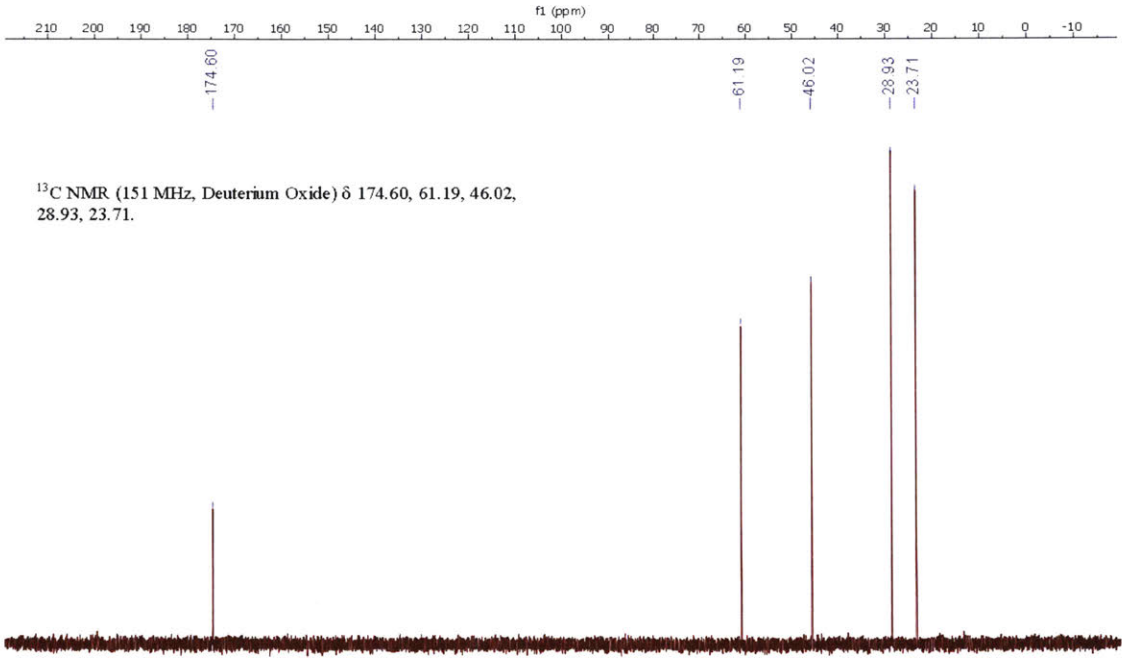


Figure A2.14.25: 13C spectrum of phenylalanine, isolated from a solution of Fmoc-Phe-OH in DMF



073-181 Pro DMF 1H_1.1.fid —

Figure A2.14.26: 1H spectrum of proline, isolated from a solution of Fmoc-Pro-OH in DMF



073-181 Pro DMF 13C_1.1.fid —

Figure A2.14.27: 13C spectrum of proline, isolated from a solution of Fmoc-Pro-OH in DMF

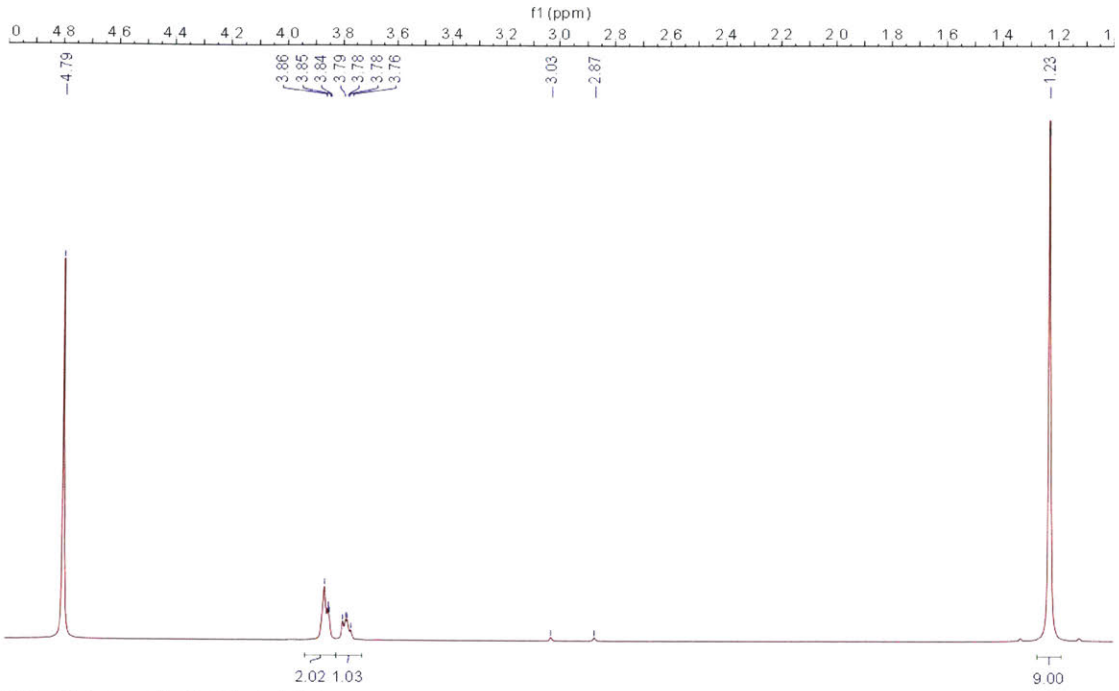


Figure A2.14.28: 1H spectrum of t-butyl ether of serine, isolated from a solution of Fmoc-Ser(tBu)-OH in DMF

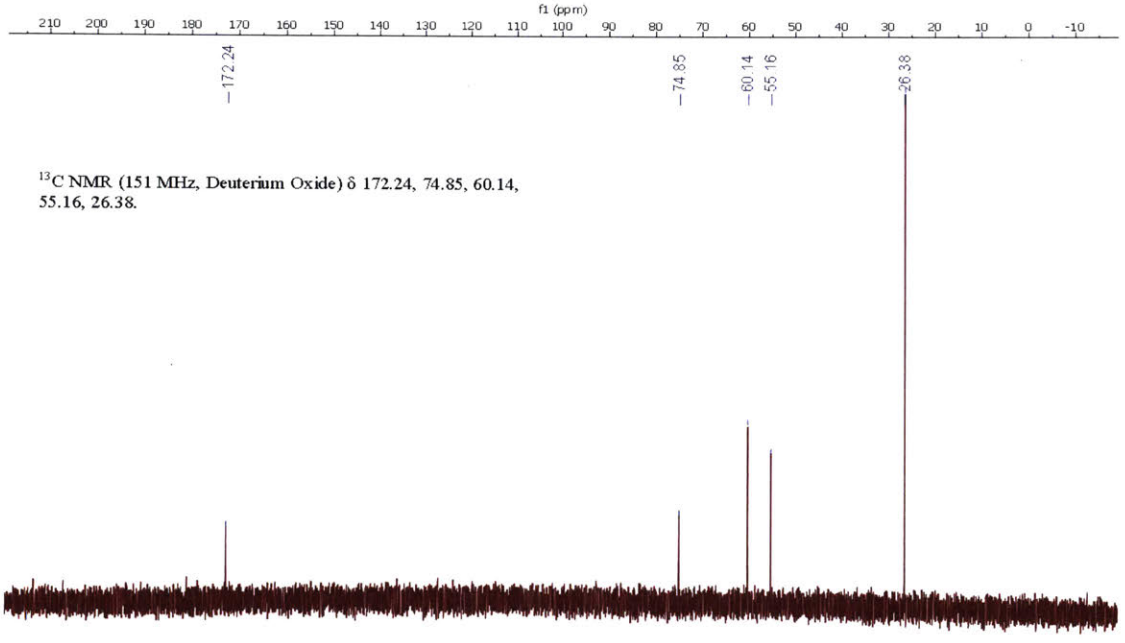
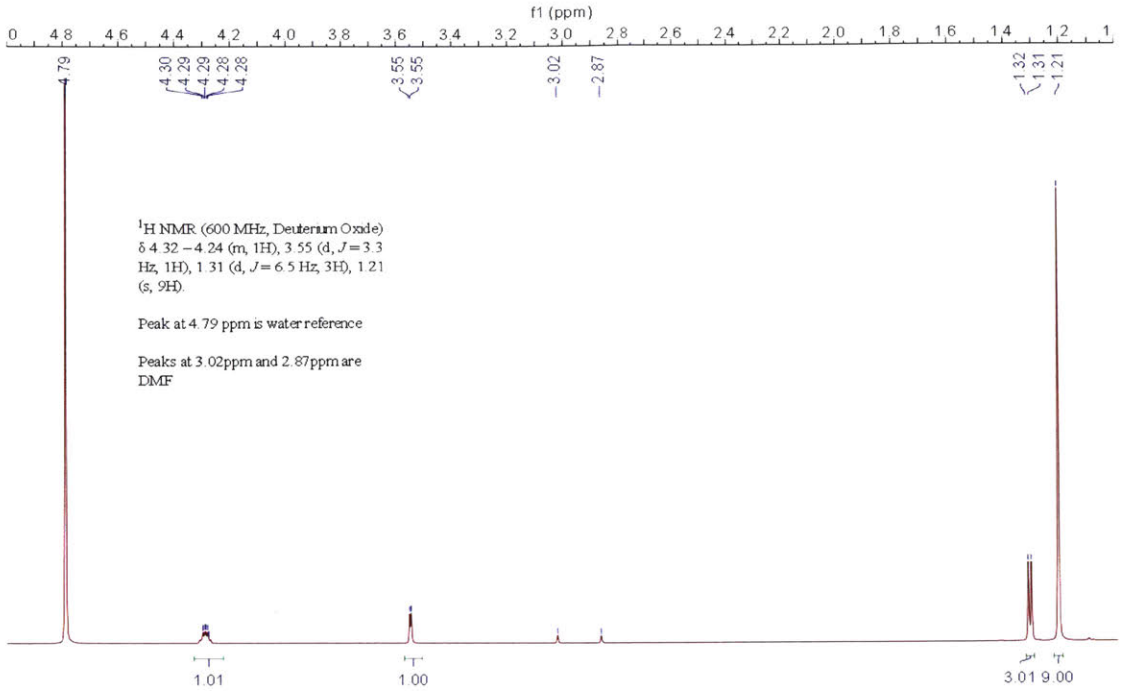
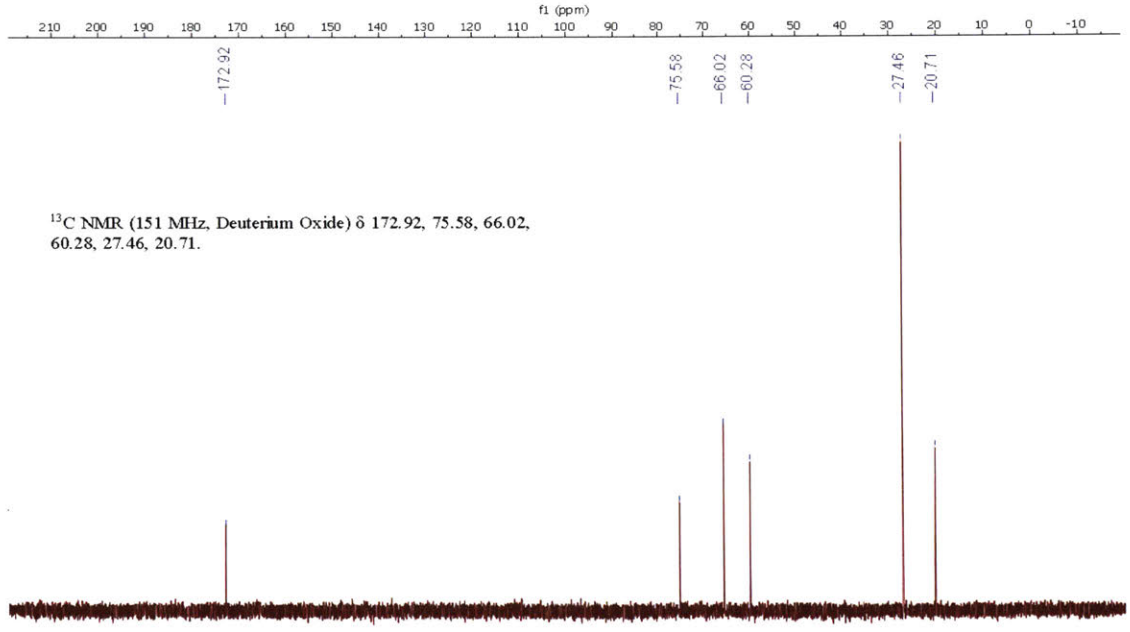


Figure A2.14.29: 13C spectrum of t-butyl ether of serine, isolated from a solution of Fmoc-Ser(tBu)-OH in DMF



073-181 Thr DMF 1H_1.1.fid —

Figure A2.14.30: 1H spectrum of t-butyl ether of threonine, isolated from a solution of Fmoc-Thr(tBu)-OH in DMF



073-181 Thr DMF 13C_1.1.fid —

Figure A2.14.31: 13C spectrum of t-butyl ether of threonine, isolated from a solution of Fmoc-Thr(tBu)-OH in DMF

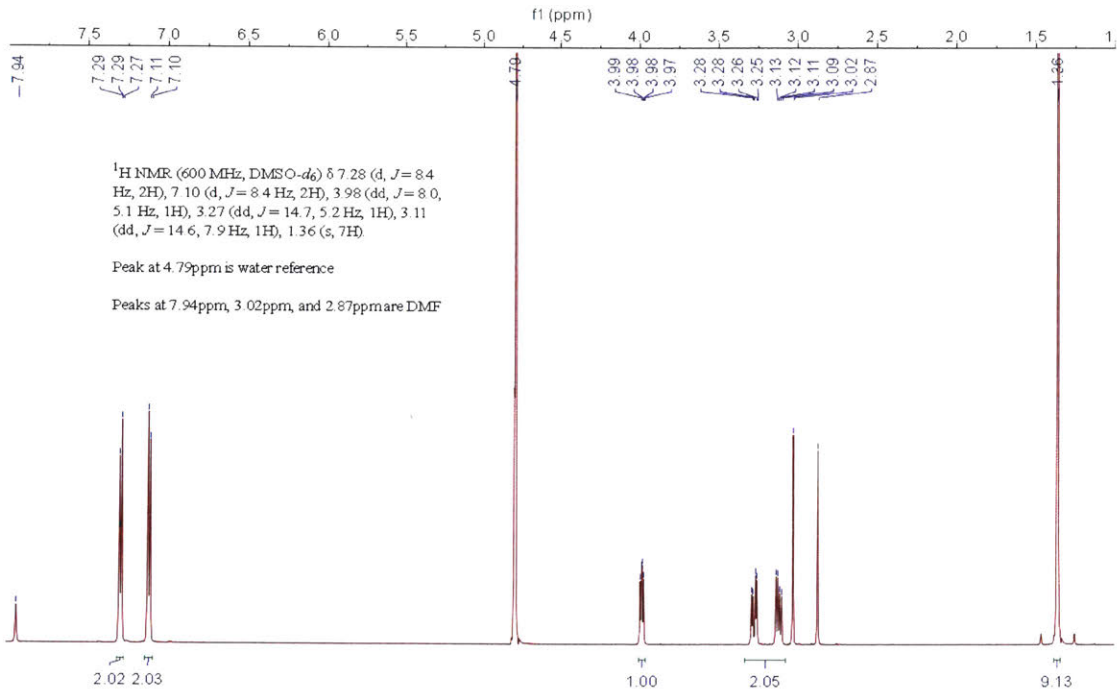


Figure A2.14.32: 1H spectrum of t-butyl ether of tyrosine, isolated from a solution of Fmoc-Try(tBu)-OH in DMF

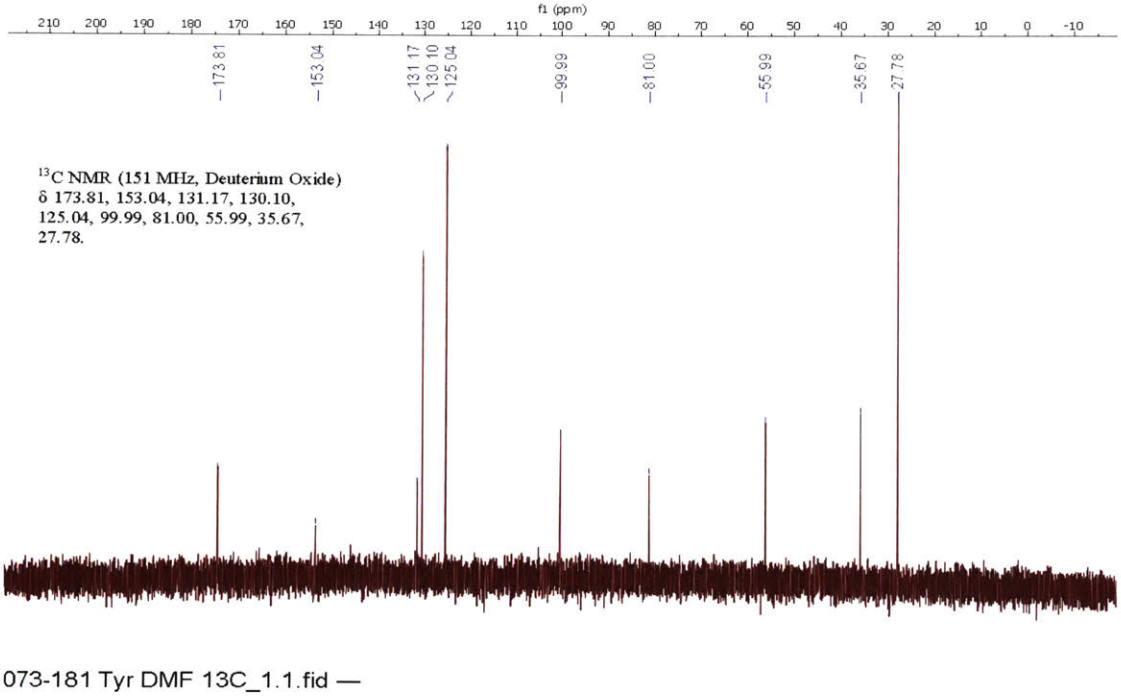


Figure A2.14.33: 13C spectrum of t-butyl ether of tyrosine, isolated from a solution of Fmoc-Try(tBu)-OH in DMF

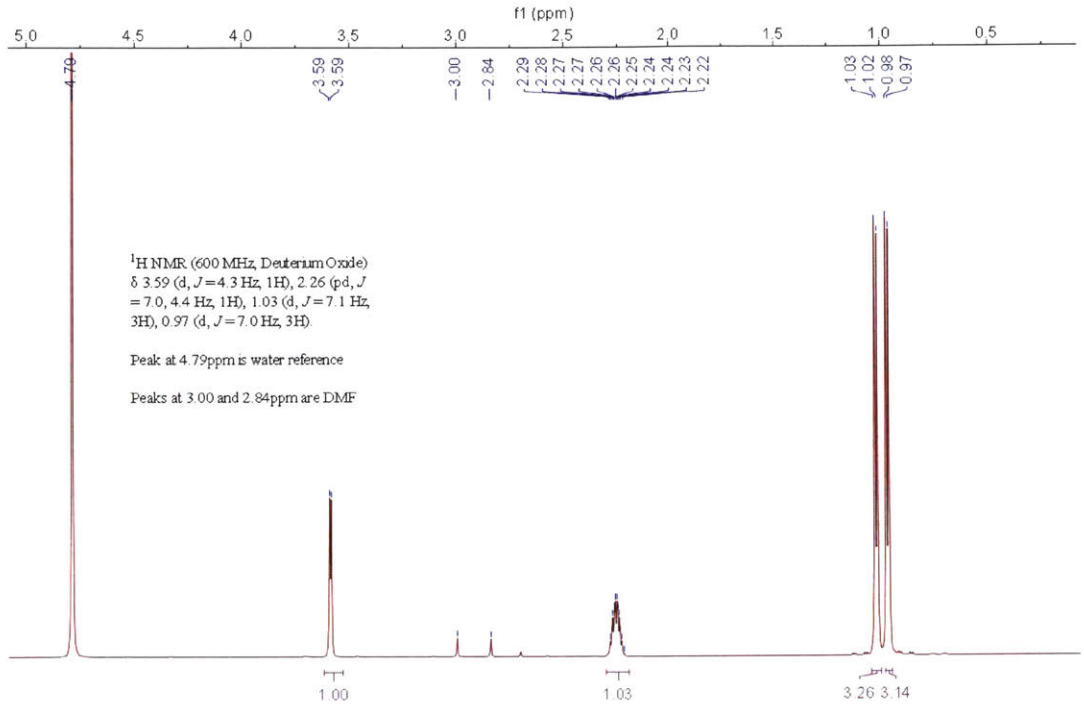


Figure A2.14.34: 1H spectrum of valine, isolated from a solution of Fmoc-Val-OH in DMF

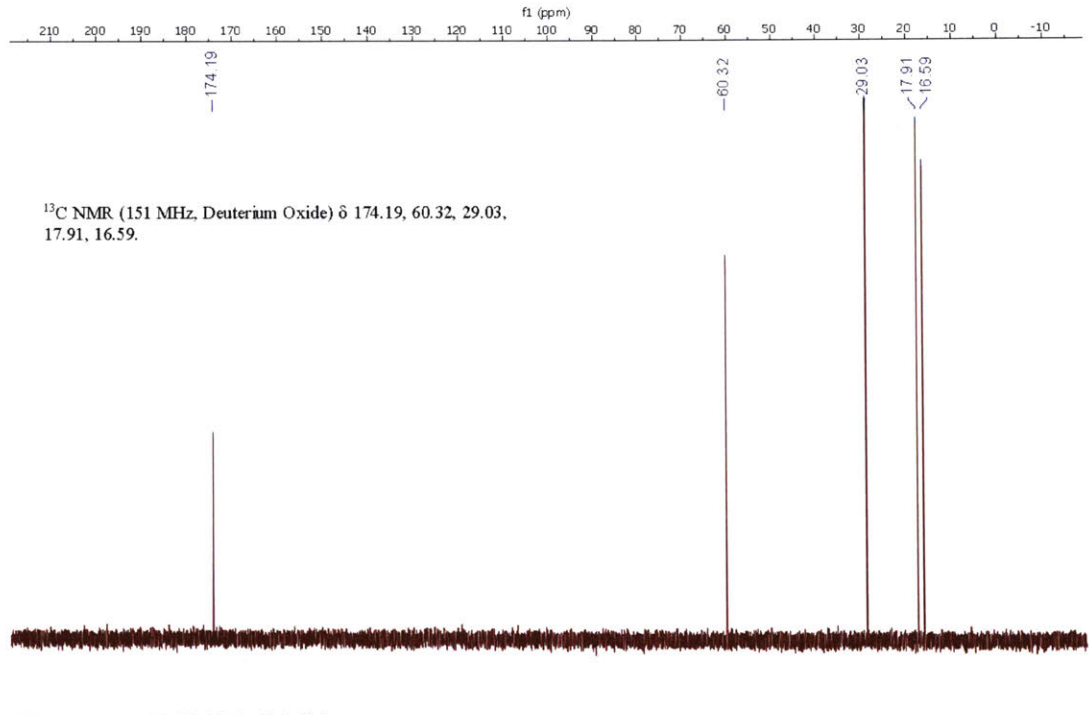


Figure A2.14.35: 13C spectrum of valine, isolated from a solution of Fmoc-Val-OH in DMF

A2.14.3 Spectra of precipitates – NMP:

The following spectra were obtained from precipitates isolated from solutions of Fmoc protected amino acids in NMP after four weeks at room temperature. In all cases, the Fmoc protecting group has been lost.

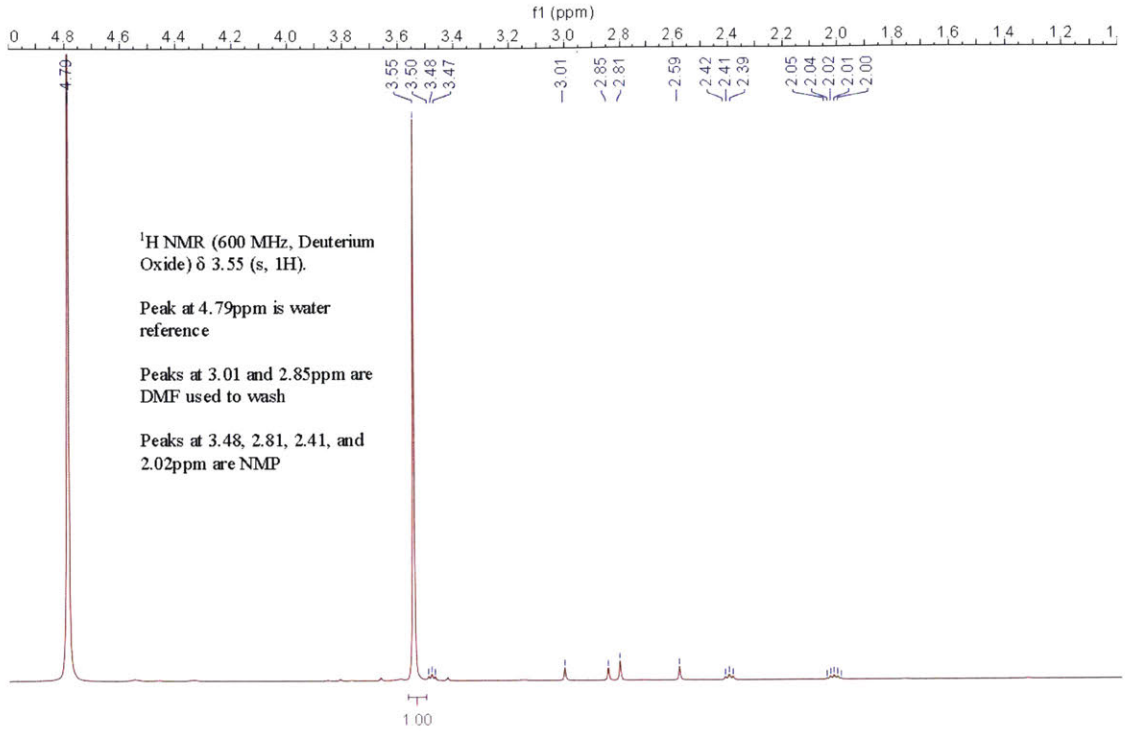


Figure A2.14.36: 1H spectrum of glycine, isolated from a solution of Fmoc-Gly-OH in NMP

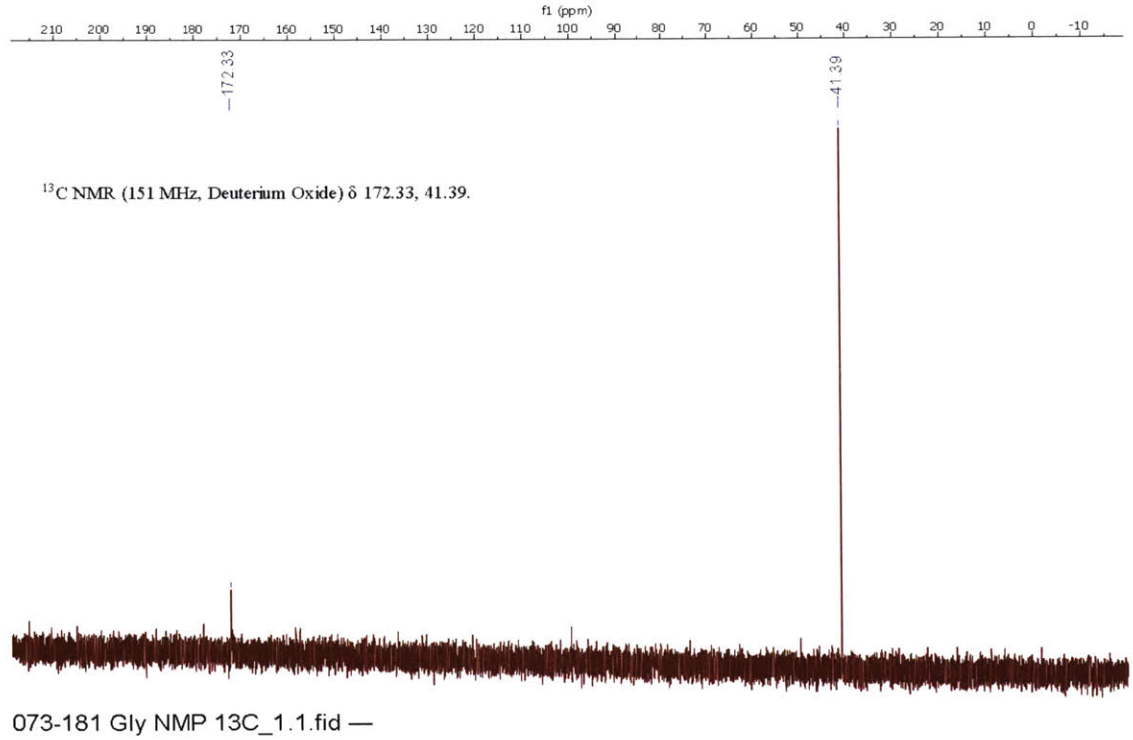


Figure A2.14.37: 13C spectrum of glycine, isolated from a solution of Fmoc-Gly-OH in NMP

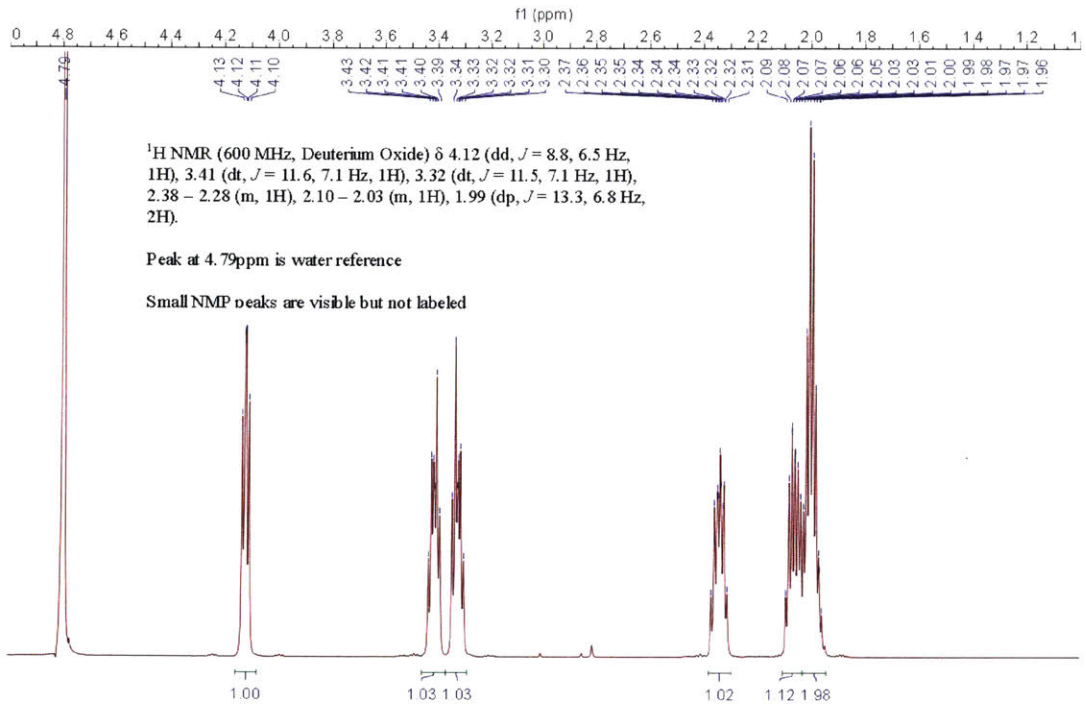


Figure A2.14.38: 1H spectrum of proline, isolated from a solution of Fmoc-Pro-OH in NMP

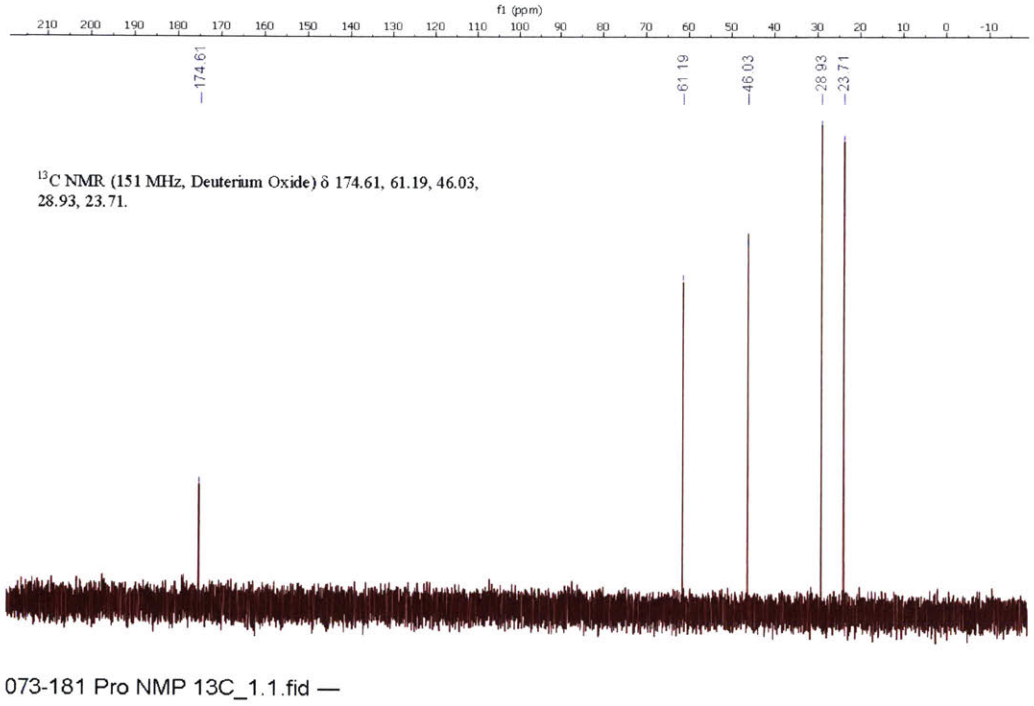


Figure A2.14.39: 13C spectrum of proline, isolated from a solution of Fmoc-Pro-OH in NMP

A2.15 Large format chromatograms, mass spectra, and annotated molecular feature extractions

A2.15.1 GHRH

The following figures and tables show expanded views of the TICs presented in Figure 5, mass spectral data from the indicated peak apex, and compounds identified with MFE.

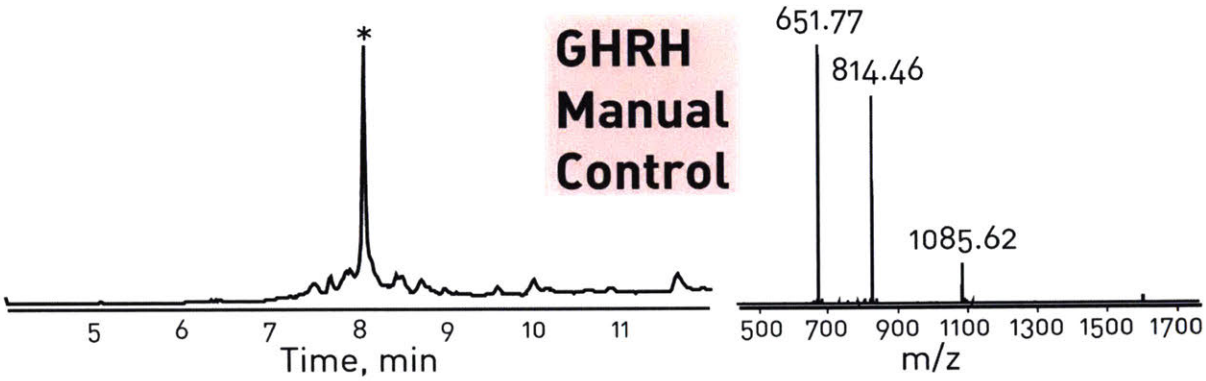


Figure A2.15.1: Large format chromatogram and MS data for GHRH shown in figure 3.4.3A synthesized with the manual fast flow peptide synthesizer. TIC is shown. Calculated monoisotopic mass = 3252.79Da, Observed monoisotopic mass = 3252.83Da.

Table A2.15.1 Assigned MFE of GHRH control synthesized with the manual flow synthesizer (Figs A2.15.1 and 3.4.3A). The 50 most abundant compounds are shown and annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3252.825	651.5725	8.075	52.62
42Da cap on Leu-17	1467.884	734.9486	8.44	2.64
Unknown	3238.805	648.7686	8.036	2.53
42Da cap on Lys-12	1993.217	665.4131	8.169	2.48
Gln deletion	3124.764	625.9602	8.138	2.28
42Da cap on Arg-11	2149.329	538.3373	7.696	1.55
Leu/Ile deletion	3139.728	628.9539	7.517	1.45
Unknown. Possible t-Butylation	3305.733	827.4413	8.075	1.42
Unknown	3238.808	648.7699	7.934	1.39
Ala deletion	3181.786	637.3644	8.033	1.26
Isomer	3252.823	651.5717	7.494	1.16
42Da cap on Ser-9	2399.406	600.859	7.879	1.09
Tyr deletion	3089.756	618.9579	7.983	1
42Da cap on Ala-2	3131.762	783.9478	8.497	0.93
Aspartimide	3235.805	809.9575	8.067	0.91
Isomer	3252.825	651.5723	7.825	0.82
Isomer	3252.819	651.5719	7.69	0.81
42Da cap on Lys-21	1040.629	521.3217	7.93	0.79
Unknown	3290.767	1097.93	8.074	0.78
Arg deletion	3096.71	775.1848	8.493	0.76
Unknown	3224.795	645.9693	7.93	0.66
Gln deletion	3124.751	625.9579	7.881	0.58
Ser deletion	3165.773	634.1618	8.017	0.58
Small molecule	399.2476	400.2549	9.956	0.52
Unknown	3278.836	820.7156	8.469	0.49
Arg deletion	3096.708	775.1844	8.705	0.48
Small molecule	420.226	421.2332	9.957	0.47
42Da cap on Asn-8	2513.456	629.3708	7.839	0.45
Thr deletion	3151.781	631.3631	8.327	0.45
42Da cap on Val-13	1865.112	622.711	8.866	0.42
Unknown truncate	2650.552	663.6458	7.523	0.4
42Da cap on Asp-25	557.3556	558.3629	7.701	0.39
Asp deletion	3137.792	785.457	7.987	0.39
42Da cap on Leu-14	1766.052	589.6912	8.66	0.39
Ala deletion	3181.79	637.3671	8.169	0.38
Asp deletion	3137.772	628.5627	7.882	0.37
Unknown. Possible Ser deletion	3166.78	792.7022	8.022	0.36
Unknown truncate	1450.853	484.6248	8.425	0.36
Unknown. Possible Ser deletion	3164.759	792.1969	8.073	0.35
Lys deletion	3124.715	782.1861	8.514	0.35
Unknown truncate	1891.127	631.3809	9.108	0.35
Unknown truncate	2627.453	657.8713	8.087	0.34
42Da cap on Tyr-10	2312.381	579.1027	7.874	0.33
Leu/Ile deletion	3139.745	628.9562	7.747	0.31
Unknown	3274.806	1092.609	8.073	0.31
Ser deletion	3165.805	634.1654	8.234	0.3
Small molecule	404.2777	427.2675	8.236	0.3
Unknown	2937.678	588.5432	7.557	0.29
Small molecule	404.2747	427.2639	7.876	0.29
Unknown. Possible Leu/Ile deletion	3140.752	629.1578	8.005	0.29

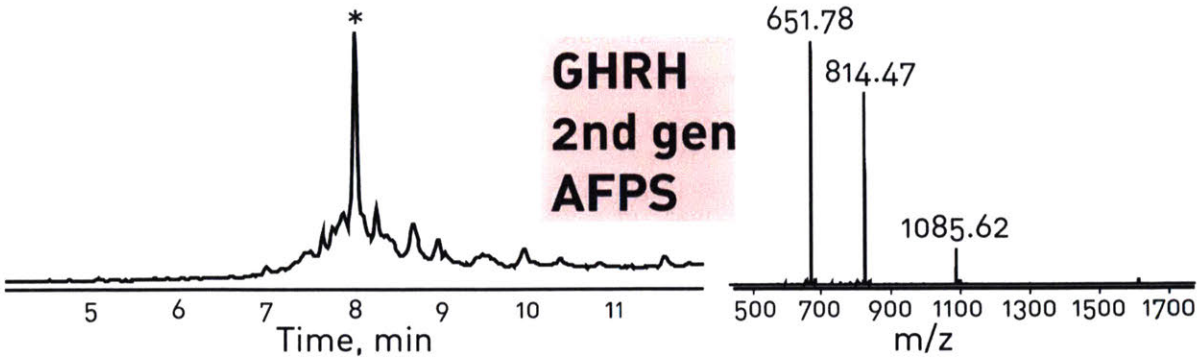


Figure A2.15.2: Large format chromatogram and MS data for GHRH shown in figure 3.4.3B synthesized with the second generation AFPS, under default conditions. TIC is shown. Calculated monoisotopic mass = 3252.79Da, Observed monoisotopic mass = 3252.86Da.

Table A2.15.2. Assigned MFE of GHRH synthesized with the second generation AFPS, under default conditions (Figs A2.15.2 and 3.4.3B). The 50 most abundant compounds are shown and annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3252.841	651.5765	8.053	37.47
Thr deletion	3151.784	631.3649	8.296	6.44
42Da cap on Leu-17	1467.886	734.95	8.413	2.38
26Da cap on Val-13	1849.092	617.3723	8.991	2.07
42Da cap on Lys-12	1993.278	665.4332	8.147	1.96
Leu/Ile deletion	3139.765	628.9606	7.991	1.94
Unknown	3238.809	648.7706	8.008	1.89
26Da cap on Tyr-10	2296.349	575.0953	7.795	1.84
Ala deletion	3181.789	637.3649	8.011	1.78
26Da cap on Leu-17	1451.859	726.9365	8.37	1.66
42Da cap on Lys-21	1040.626	521.3209	7.914	1.64
26Da cap on Arg-11	2133.29	534.3301	7.689	1.52
Small molecule	412.1658	413.173	7.046	1.49
42Da cap on Arg-11	2149.319	538.338	7.675	1.33
Val deletion	3153.753	631.7611	7.9	1.3
Unknown truncate	2940.627	981.2148	9.494	1.04
Arg deletion	3096.715	775.1862	8.662	1.02
Unknown truncate	1891.121	631.3802	9.077	0.91
26Da cap on Ser-9	2383.389	596.8558	7.825	0.9
26Da cap on Leu-14	1750.028	584.3513	8.728	0.9
Tyr deletion	3089.765	618.9602	7.949	0.86
42Da cap on Ser-9	2399.412	600.8608	7.86	0.82
42Da cap on Gln-16	1579.911	527.645	8.329	0.8
Unknown (high molecular weight)	4682.532	781.4305	7.877	0.64
Arg deletion	3096.723	775.1888	8.456	0.62
Unknown truncate	2218.376	555.6015	8.213	0.58
42Da cap on Ala-2	3131.779	783.9521	8.46	0.57
Possible Phe/Thr double deletion	3004.719	752.187	7.944	0.56
26Da cap on Lys-21	1023.601	512.8076	7.915	0.53
Small molecule	420.2266	421.2339	9.935	0.53
Isomer	3252.828	651.5728	7.471	0.51
Unknown. Possible t-Butylation	3306.782	662.3685	8.052	0.51
Leu/Ile deletion	3139.754	785.9456	7.533	0.5
42Da cap on Leu-17	1766.047	589.6901	8.624	0.48
Unknown. Possible Leu/Ile deletion	3140.746	629.1565	7.53	0.47
Unknown truncate	1832.069	611.6968	8.991	0.47
Truncation at Ser-18	1311.792	656.9033	8.746	0.46
Small molecule	404.2746	427.2638	8.219	0.45
Small molecule	404.2756	427.2648	7.863	0.43
Unknown truncate	2381.453	596.3707	8.232	0.42
Ser deletion	3165.808	634.1688	7.996	0.41
Unknown truncate	1837.105	613.3758	8.939	0.41
Unknown	3530.117	883.5373	9.133	0.4
26Da cap on Leu-17	1451.857	726.9331	8.521	0.39
42Da cap on Val-13	1865.118	622.7147	8.831	0.39
Unknown truncate	2412.449	604.1201	7.384	0.37
Double Ala incorporation	3323.857	831.9716	8.115	0.37
Unknown. Possible Ser deletion	3166.782	792.7028	8.011	0.36
Possible Ser/(Leu/Ile) double deletion	3052.722	764.187	8.098	0.36
Isomer	3252.809	651.5674	7.669	0.32

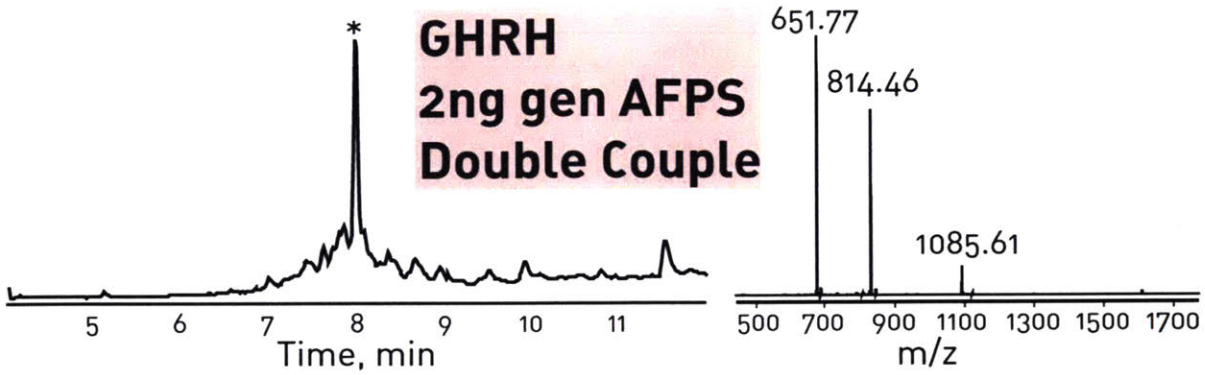


Figure A2.15.3: Large format chromatogram and MS data for GHRH shown in figure 3.4.3 synthesized with the second generation AFPS, with double coupling. TIC is shown. Calculated monoisotopic mass = 3252.79Da, Observed monoisotopic mass = 3252.82Da.

Table A2.15.3. Assigned MFE of GHRH synthesized with the second generation AFPS, with double coupling (Figs A2.15.3 and 3.4.3C). The 50 most abundant compounds are shown and annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3252.809	651.5685	8.005	43.36
42Da cap on Lys-12	1993.209	665.4101	8.098	3.32
42Da cap on Leu-17	1467.879	734.9469	8.369	3.18
42Da cap on Lys-21	1040.627	521.3209	7.865	2.15
Unknown	3238.8	648.7678	7.963	2.10
Small molecule	412.1618	413.1691	7.016	1.92
Unknown. Possibly t-Butylation	3305.735	662.1546	8.004	1.77
42Da cap on Arg-11	2149.316	538.3367	7.625	1.65
Tyr deletion	3089.754	618.9576	7.898	1.41
42Da cap on Ser-9	2399.417	600.8614	7.807	1.12
Unknown (high molecular weight)	4682.544	781.4323	7.829	1.11
26Da cap on Arg-11	2133.289	534.33	7.64	1.03
26Da cap on Val-13	1849.083	617.3674	8.958	1.02
26Da cap on Leu-17	1451.857	726.9352	8.329	0.99
Leu/Ile deletion	3139.758	628.9588	7.473	0.95
Unknown	3239.808	648.9683	7.861	0.95
Leu/Ile deletion	3139.745	628.9562	7.943	0.93
Ala deletion	3181.785	637.3643	7.96	0.93
Isomer.	3252.839	651.5751	7.421	0.92
26Da cap on Tyr-10	2296.353	575.0959	7.745	0.92
Unknown truncate	1891.143	631.3886	9.041	0.83
Unknown truncate	2650.551	663.6446	7.451	0.80
26Da cap on Ser-9	2383.385	596.8541	7.772	0.80
26Da cap on Lys-21	1023.601	512.8078	7.866	0.74
42Da cap on Ala-2	3131.762	783.9481	8.422	0.69
Unknown	607.4232	608.4305	11.539	0.65
Unknown	3224.793	645.966	7.851	0.64
Unknown (high molecular weight)	5636.111	806.1665	7.814	0.62
Unknown. Possible Leu/Ile deletion (-1Da)	3140.741	786.1926	7.942	0.62
Isomer.	3252.834	651.5759	7.75	0.60
Unknown truncate	2218.368	555.5998	8.168	0.59
Small molecule	420.2257	421.233	9.894	0.56
Small molecule	399.2473	400.2546	9.892	0.55
Arg deletion	3096.706	775.1831	8.631	0.54
Ala deletion	3181.773	637.3624	8.096	0.50
Thr deletion	3151.784	631.3638	8.255	0.47
42Da cap on Leu-17	1766.059	589.6946	8.589	0.47
42Da cap on Ala-19	1267.767	634.8914	7.647	0.46
Arg deletion	3096.709	775.1843	8.418	0.46
Unknown truncate	1750.02	584.3466	8.698	0.43
42Da cap on Val-13	1865.114	622.7121	8.798	0.43
Unknown truncate	716.5075	717.5136	10.762	0.43
Unknown truncate	2467.403	494.4892	6.867	0.40
Val deletion	3153.755	631.7592	7.847	0.40
Aspartimide	3235.795	809.956	7.995	0.40
Isomer	3252.821	651.5733	7.619	0.39
Truncation at Tyr-10	2270.375	568.6002	7.258	0.38
Unknown. Posible Ser deletion	3164.731	792.1901	8	0.37
Unknown truncate	1005.592	503.8033	7.867	0.36
42Da cap on Gln-16	1579.916	527.6471	8.281	0.36

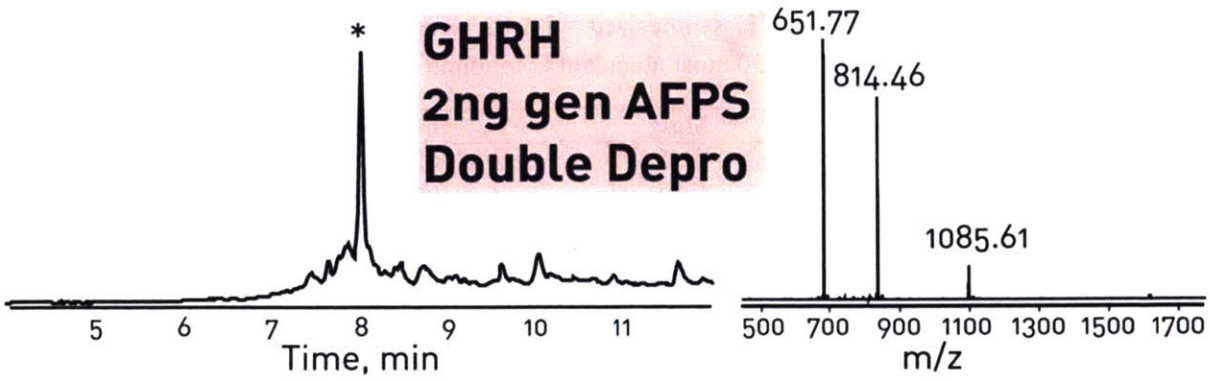


Figure A2.15.4: Large format chromatogram and MS data for GHRH shown in figure 3.4.3 synthesized with the second generation AFPS, with double deprotection. TIC is shown. Calculated monoisotopic mass = 3252.79Da, Observed monoisotopic mass = 3252.84Da.

Table A2.15.4. Assigned MFE of GHRH synthesized with the second generation AFPS, with double deprotection (Fig A2.15.4 and 3.4.3D). The 50 most abundant compounds are shown and annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3252.827	651.5735	8.068	33.93
42Da cap on Lys-12	1993.206	665.4094	8.159	3.26
42Da cap on Leu-17	1467.88	734.9472	8.441	2.96
Isomer. Likely aspartimide derived	3252.802	651.5677	8.127	2.11
42Da cap on Lys-21	1040.618	521.3161	7.921	2.02
42Da cap on Arg-11	2149.335	538.3389	7.687	1.83
Unknown truncate	2940.619	981.2131	9.618	1.82
Tyr deletion	3089.755	618.9598	7.96	1.79
Unkown (high molecular weight)	4682.532	781.4296	7.889	1.7
Probably missasignement of produt ions	3253.807	814.4593	8.131	1.57
Isomer. Likely aspartimide derived	3252.812	651.5706	7.486	1.41
Ala deletion	3181.801	637.3681	8.02	1.3
Aspartimide	3234.818	647.971	8.023	1.26
Unknown. Possible Leu/Ile deletion (-1Da)	3140.757	629.1596	8.001	1.25
Arg deletion	3096.713	775.1858	8.51	1.25
26Da cap on Ser-9	2383.383	596.8544	7.834	1.19
Arg deletion	3096.723	775.1878	8.774	1.16
Uknown	3238.818	810.7115	8.023	1.12
26Da cap on Arg-11	2133.283	534.3278	7.707	1.1
42Da cap on Ser-9	2399.409	600.8596	7.871	1.1
Small molecule	399.2467	400.2539	10.014	1.1
Unknown truncate	2650.542	663.6444	7.517	1.05
26Da cap on Leu-17	1451.852	726.9318	8.393	1.03
26Da cap on Tyr-10	2296.341	575.0932	7.807	1.02
Isomer. Likely aspartimide derived	3252.824	651.5724	7.683	0.98
Small molecule	420.225	421.2323	10.015	0.97
Unknown truncate	1891.12	631.3795	9.209	0.95
Unkown (+14Da)	3238.784	648.7645	7.923	0.9
Unknown	3239.828	648.9729	8.022	0.9
Isomer. Likely aspartimide derived	3252.815	651.5729	7.816	0.88
Truncation at Ser-18	1311.784	656.8993	8.829	0.88
Small molecule	404.2755	427.2648	7.856	0.84
Unknown	3278.831	820.7182	8.485	0.84
Small molecule	404.2755	427.2652	8.22	0.81
Unknown (high molecular weight)	5636.101	806.168	7.88	0.76
Thr deletion	3151.769	631.3601	8.319	0.74
26Da cap on Lys-21	1023.59	512.8024	7.92	0.72
42Da cap on Ala-2	3131.764	783.9483	8.516	0.69
26Da cap on Val-13	1849.088	617.3683	9.129	0.69
Small molecule	448.148	449.1553	10.211	0.67
26Da cap on Leu-17 (Isomer of above)	1451.856	726.9339	8.568	0.62
Ala deletion	3181.769	637.3603	8.156	0.6
t-Butylation	3307.75	662.5611	8.067	0.56
Ser deletion	3165.779	634.1631	8.008	0.53
42Da cap on Leu-17	1766.044	589.6897	8.7	0.53
Unknown	3239.783	810.9531	7.924	0.52
Unknown truncate	1837.123	613.3819	9.059	0.49
Unknown truncate	2467.39	494.4852	6.917	0.47
Unknown truncate	1294.76	648.3872	8.83	0.46
Val deletion	3153.743	631.7546	7.912	0.43

The following figures and tables show expanded views of the TICs presented in Figure 3.6.2, mass spectral data from the indicated peak apex, and compounds identified with MFE.

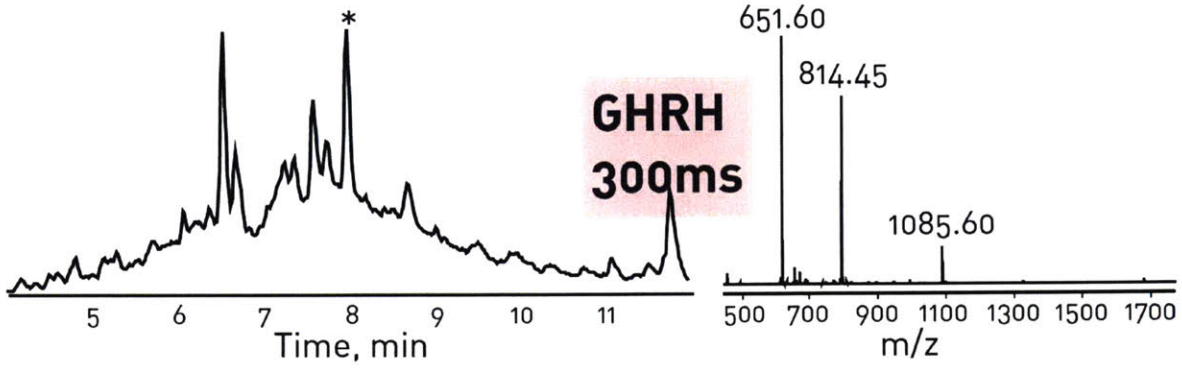


Figure A2.15.5: Large format chromatogram and MS data for GHRH shown in figure 3.6.2 synthesized with 300ms pump refill time and 90C activation. TIC is shown. Calculated monoisotopic mass = 3252.79Da, Observed monoisotopic mass = 3252.76Da. MFE was not performed.

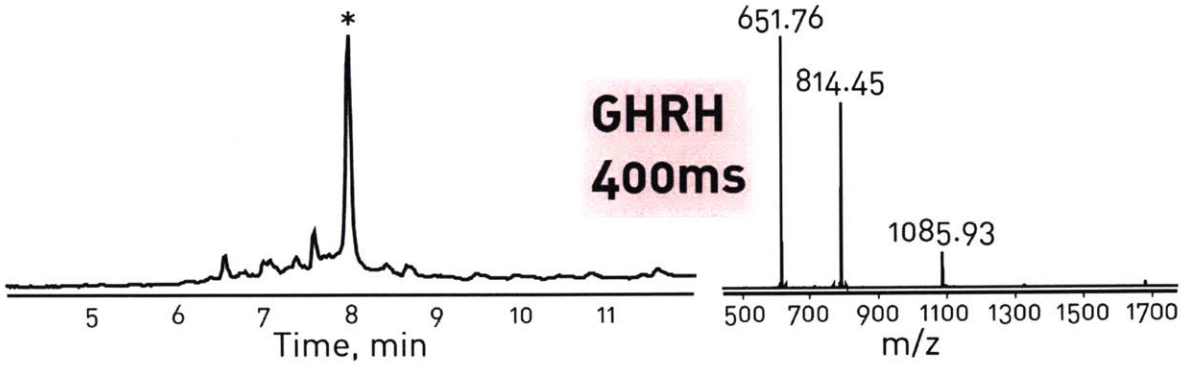


Figure A2.15.6: Large format chromatogram and MS data for GHRH shown in figure 3.6.2 synthesized with 400ms pump refill time and 90C activation. TIC is shown. Calculated monoisotopic mass = 3252.79Da, Observed monoisotopic mass = 3252.76Da. MFE was not performed.

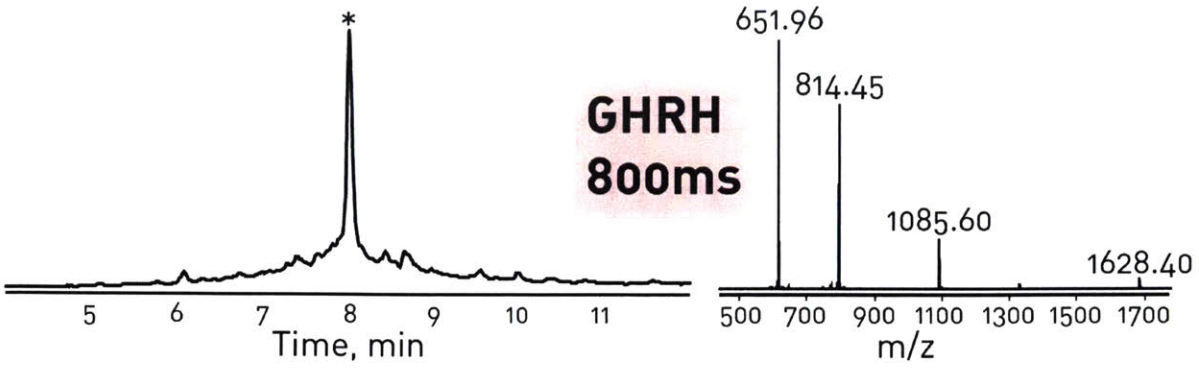


Figure A2.15.7: Large format chromatogram and MS data for GHRH shown in figure 3.6.2 synthesized with 800ms pump refill time and 90C activation. TIC is shown. Calculated monoisotopic mass = 3252.79Da, Observed monoisotopic mass = 3252.76Da. MFE was not performed.

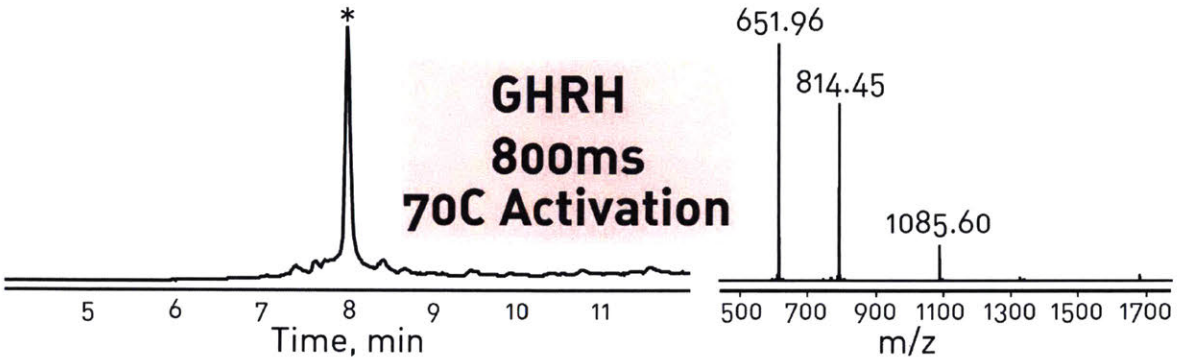


Figure A2.15.8: Large format chromatogram and MS data for GHRH shown in figure 3.6.2 synthesized with 800ms pump refill time and 70C activation. TIC is shown. Calculated monoisotopic mass = 3252.79Da, Observed monoisotopic mass = 3252.76Da

Table A2.15.5. Assigned MFE of GHRH synthesized with the fourth generation AFPS (Fig A2.15.8 and 3.6.2B). The 50 most abundant compounds are shown and annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3252.766	651.5599	7.997	67.1
Aspartimide	3234.751	647.9565	7.965	3.54
Isomer. Likely aspartimide derived	3252.702	651.546	7.415	1.73
42Da cap on Ala-2	3131.704	783.9332	8.42	1.59
42Da cap on Arg-11	2149.27	538.3244	7.629	1.43
Unknown. Possibly t-Butylation	3305.679	662.142	8.003	1.37
Isomer. Likely aspartimide derived	3252.738	651.5554	7.617	1.16
Isomer. Likely aspartimide derived	3252.707	651.5516	7.748	1.1
Leu/Ile deletion	3139.628	628.9335	7.421	1.09
Tyr deletion	3089.687	618.9415	7.899	0.98
Asp deletion	3137.737	628.5543	7.914	0.96
Unknown truncate	716.5035	717.5106	10.721	0.89
Unknown. Possibly Ser deletion	3164.721	792.1869	7.988	0.87
Arg deletion	3096.665	775.1729	8.425	0.78
Unknown	3238.743	648.7554	7.845	0.7
42Da cap on Leu-17	1467.867	734.9411	8.372	0.64
Asn deletion	3138.714	628.7505	8.225	0.63
Gln deletion	3124.697	625.9468	7.801	0.6
Truncation at Ala-4	2903.579	581.7228	7.731	0.59
Ala deletion	3181.719	637.351	7.952	0.59
Unknown truncate	702.5123	725.5009	11.302	0.59
Arg deletion	3096.664	775.1728	8.643	0.56
Pbf clavage adduct	3504.815	701.969	8.955	0.56
Ala deletion	3181.728	637.353	8.094	0.54
Leu/Ile deletion	3365.831	674.1719	7.9	0.46
42Da cap on Lys-12	1993.188	665.4029	8.106	0.43
Aspartimide	3234.726	647.9498	7.625	0.41
Thr deletion	3151.699	631.3448	8.254	0.41
Unknown	3262.318	653.4709	8.001	0.39
Unknown truncate. Possibly 42Da cap on Leu-	1778.026	593.6822	8.538	0.39
Unknown	3274.766	1092.598	7.996	0.37
Small molecule	404.2767	427.2662	8.095	0.34
Unknown	2869.499	574.9068	7.072	0.32
Unknown truncate	716.4993	717.5064	11.072	0.32
Asn deletion	3137.735	628.5559	7.8	0.31
Unknown truncate	702.5214	703.5292	11.63	0.31
Small molecule	684.499	685.5063	11.301	0.3
Small molecule	404.2718	427.2609	7.734	0.29
Unknown	3325.791	666.1655	8.078	0.29
Unknown truncate	738.4865	739.4937	10.38	0.29
Small molecule	698.4937	699.501	10.38	0.29
Unknown	3077.673	770.4252	7.997	0.28
Unknown (high molecular weight)	5709.421	816.6388	7.998	0.28
Unknown	3217.714	644.55	7.995	0.26
Unknown truncate	1479.863	740.9389	8.631	0.26
42Da cap on Asn-8	2513.375	629.3507	7.767	0.22
Unknown (high molecular weight)	4894.038	816.6803	7.996	0.22
Unknown truncate	1957.068	653.3634	8.001	0.21
Unknown truncate	2130.274	711.0986	8.435	0.21
Unknown truncate	698.494	699.5013	10.723	0.2

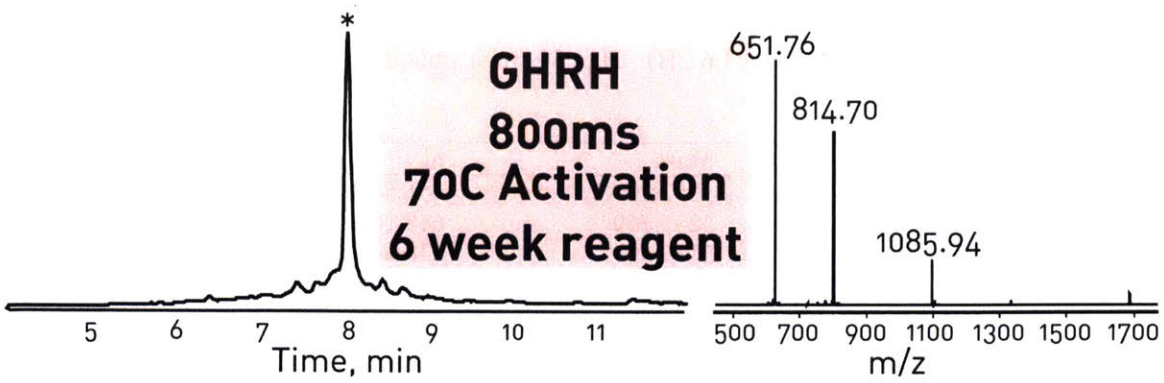


Figure A2.15.9: Large format chromatogram and MS data for GHRH shown in figure 3.6.2 synthesized with 800ms pump refill time and 70C activation six weeks after reagents were prepared. TIC is shown. Calculated monoisotopic mass = 3252.79Da, Observed monoisotopic mass = 3252.79Da

Table A2.15.6. Assigned MFE of GHRH synthesized with the fourth generation AFPS 6 weeks after reagents were prepared (Fig A2.15.9 and 3.6.2B). The 50 most abundant compounds are shown and annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3252.791	651.5648	7.976	60.84
Unknown (-17Da) Possible aspartimide	3235.769	648.1612	7.962	2.53
Leu deletion	3139.69	628.944	7.399	2.03
Unknown (-14Da)	3238.771	648.7608	7.936	1.84
42Da cap on Arg-11	2149.297	538.331	7.612	1.58
42Da cap on Ala-4	3131.731	783.9399	8.404	1.53
Arg deletion	3096.684	775.1777	8.631	1.26
Asp deletion	3137.755	628.5575	7.891	1.13
Possible 26Da cap on Leu-17 (-1Da)	1450.85	734.9459	8.36	0.92
Gln deletion	3124.722	625.9521	7.781	0.89
Tyr deletion	3089.712	618.9481	7.886	0.86
Unknown	3146.733	809.9505	7.962	0.86
Ala deletion	3181.749	637.3562	7.931	0.74
Possible Ser deletion	3164.729	792.19	7.968	0.74
Unknown cleavage adduct	3415.846	684.1772	8.262	0.74
Unknown truncate	1786.934	596.6518	6.387	0.66
Unknown	3217.77	644.5611	7.974	0.64
Possible Asp, Gly double deletion	3079.657	775.179	8.411	0.61
42Da cap on Lys-12	1993.203	665.4083	8.088	0.59
Ala deletion	3181.746	637.3574	8.074	0.57
Unknown (-14Da)	3238.76	648.7593	7.824	0.55
Unknown	3274.691	662.7562	7.983	0.52
Asn deletion	3138.73	628.7522	8.204	0.52
Double Leu coupling	3365.869	674.1803	7.875	0.5
Gln deletion	3124.723	625.9522	8.046	0.49
Ser deletion	3165.747	634.155	8.139	0.48
Asp deletion	3137.748	628.5572	7.779	0.45
Unknown	3077.698	770.4313	7.976	0.45
Thr deletion	3151.734	631.353	8.235	0.44
Unknown (+26Da)	3278.792	656.7655	7.812	0.41
Ala deletion	3181.747	796.4454	8.279	0.41
Possible Thr deletion	3150.736	788.6888	7.813	0.4
Possible Ser deletion	3166.729	634.353	7.935	0.39
Unknown	2964.606	742.1587	7.976	0.39
Unknown (-13Da)	3239.769	810.9489	7.826	0.37
Double Ala coupling	3323.814	831.9609	8.054	0.37
Unknown truncate	1670.933	844.9875	8.941	0.37
Double Gln coupling	3380.838	677.1737	7.375	0.36
Lys deletion	3124.702	782.1825	8.639	0.35
Ser deletion	3165.745	792.4433	8.685	0.35
Unknown truncate	2342.238	586.5659	6.869	0.33
Unknown truncate	2938.564	603.3316	7.286	0.33
Unknown	3114.71	1044.92	8.402	0.31
Unknown (high molecular weight)	4566.426	653.3539	7.976	0.3
t-Butylation	3308.844	662.7779	8.465	0.3
Unknown truncate	2186.146	729.7228	7.077	0.29
Unknown truncate	2115.108	706.0428	6.992	0.28
26Da cap on Lys-21	1023.598	521.3188	7.823	0.28
Unknown truncate	1761.008	593.6853	8.523	0.28
42Da cap on Ser-9	2399.397	600.8571	7.792	0.27

A2.15.2 EETI-II

The following figures show expanded views of the TICs presented in Figure 3.4.2, as well as the mass spectral data from the indicated peak apex.

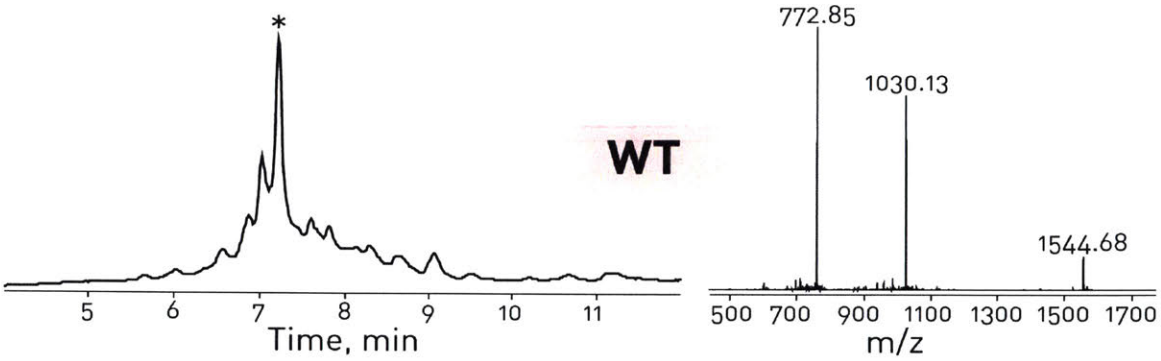


Figure A2.15.10: WT-EETI-II GCPRILMRCKQDSDCLAGCVCGPNGFCGSP TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.37Da

Table A2.15.7. Assigned MFE of WT-EETI-II synthesized with the second generation AFPS (Figs A2.15.10 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.367	772.3501	7.233	35.61
Cys, Gly double deletion	2925.329	732.3414	7.041	11.43
Truncation at Val-20	1037.505	519.7601	6.581	5.12
42Da cap on Pro-3	2967.345	990.1221	7.84	4.34
Unknown	3075.292	769.8318	7.063	3.47
t-Butylation	3141.44	786.3679	7.621	3.31
Unknown truncation	2838.299	710.5833	7.263	2.63
Gly deletion	3028.34	767.8472	7.255	2.46
Cys deletion	2982.371	746.6027	7.152	2.15
Unknown (-14Da)	3071.351	768.8466	7.164	1.99
	2585.065	862.6982	7.834	1.59
	506.2243	507.2333	6.04	1.56
	2998.337	750.5945	7.29	1.52
	2681.125	900.3913	8.308	1.46
	2938.289	735.5805	6.896	1.42
	1012.445	1035.433	7.09	1.32
	1421.596	711.807	7.727	1.31
	2895.322	724.8389	7.222	1.24
	2971.288	991.4365	7.237	1.19
	2851.259	951.4313	7.704	1.15

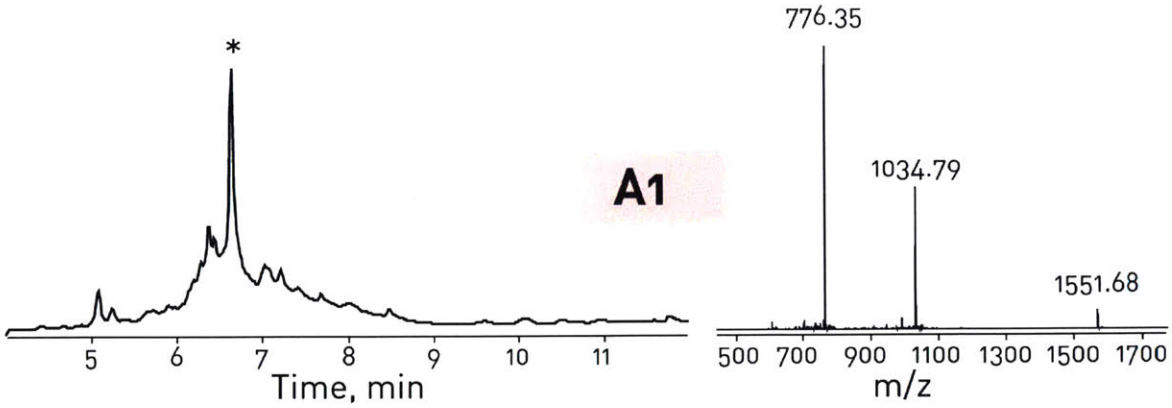


Figure A2.15.11: EETI-II Analog A1 aCPRILMRCKQDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3099.34Da, Observed monoisotopic mass = 3099.35Da

Table A2.15.8. Assigned MFE of EETI-II analog A1 synthesized with the second generation AFPS (Figs A2.15.11 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3099.281	775.8277	6.653	26.74
Cys, Ala double deletion	2925.234	732.3161	6.397	6.24
Unknown truncation	1634.723	409.6882	5.118	4.08
t-Butylation	3155.336	789.8414	7.045	2.8
42Da cap on Pro-3	2967.278	990.0998	7.225	2.45
Truncation at Cys-15	1482.569	742.291	6.462	2.03
Val deletion	3000.233	751.0656	6.464	1.7
Unknown (-17Da)	3082.241	771.5674	6.778	1.69
Unknown truncation	2437.066	610.2742	6.295	1.51
Unknown truncation	2534.098	634.5318	6.351	1.44
	1012.436	507.2256	5.261	1.18
	3271.249	818.8194	7.377	1.11
	3081.288	771.3293	6.663	1.03
	3086.284	772.5783	6.569	1.02
	2852.245	714.0686	6.684	0.99
	3042.286	761.5792	6.693	0.94
	3012.259	754.0719	6.711	0.91
	1421.544	711.7799	7.103	0.81
	3065.309	767.3347	6.726	0.76
	875.3508	876.3585	5.288	0.72

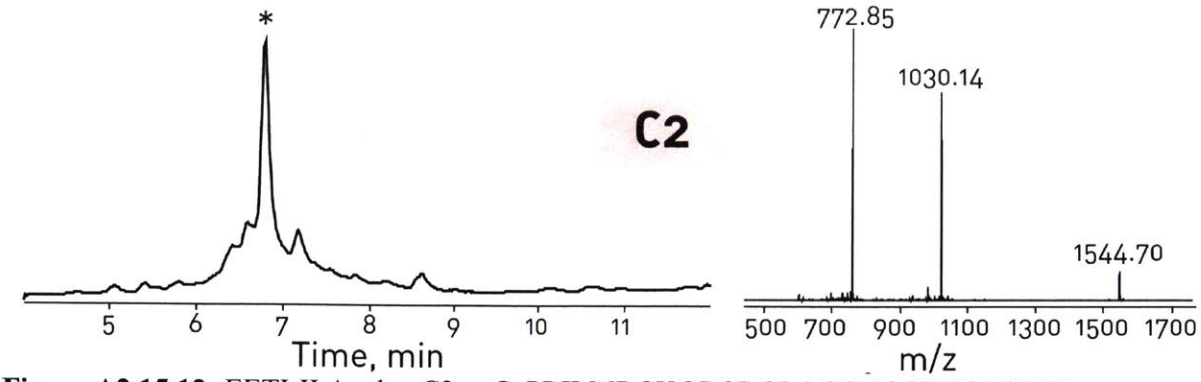


Figure A2.15.12: EETI-II Analog C2 GcPRILMRCKQDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.35Da

Table A2.15.9. Assigned MFE of EETI-II analog C2 synthesized with the second generation AFPS (Figs A2.15.12 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.31	772.3347	6.623	34.96
Unknown truncation	1620.738	406.1923	4.937	4.17
t-Butylation	3141.319	786.3365	7.025	3.74
Val deletion	2986.235	747.5663	6.424	2.87
Truncation at Cys-15	1482.564	742.2895	6.462	2.36
Unknown (-17Da)	3068.244	768.068	6.751	2.25
Unknown (-14Da)	3071.258	768.822	6.58	2.08
Cys, Gly double deletion	2925.266	732.3255	6.397	1.97
Triple t-Butylation	3257.224	815.313	7.371	1.61
Gly deletion	3028.287	758.0799	6.662	1.43
Cys deletion	2982.277	746.5787	6.549	1.22
	2724.154	682.0461	6.703	1.22
	3067.295	767.8309	6.63	1.15
	2998.241	750.5673	6.682	1.08
	2838.252	710.5701	6.654	1.05
	1460.71	487.9105	4.937	0.9
	2520.097	631.0318	6.307	0.88
	2988.209	748.0597	6.574	0.87
	2971.242	743.8178	6.671	0.78
	2423.042	606.7695	6.246	0.76

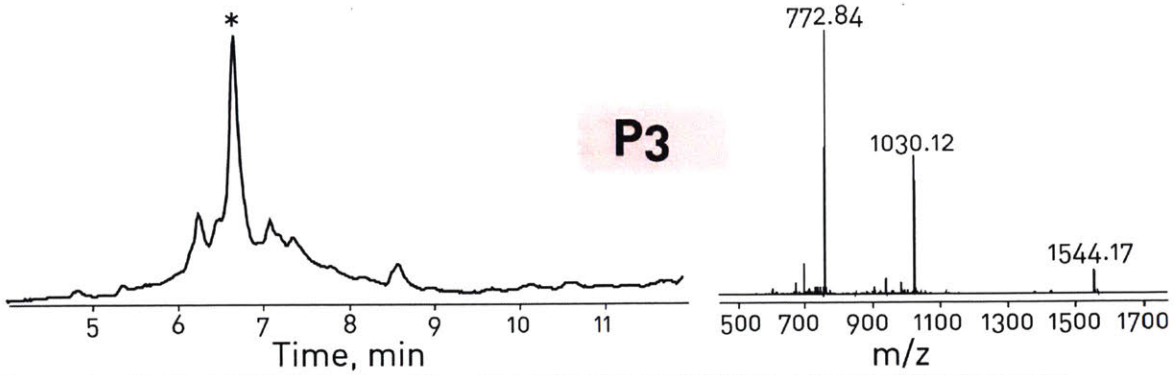


Figure A2.15.13: EETI-II Analog P3 GCpRILMRCKQDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.35Da
Table

Table A2.15.10. Assigned MFE of EETI-II analog P3 synthesized with the second generation AFPS (Figs A2.15.13 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.28308	772.32811	6.537	26.00
Unknown truncation	1620.73864	406.19226	4.763	4.34
Unknown. Possible Cys, Gly, Ser tripple deletion	2838.23336	710.56621	6.569	4.16
Ser deletion	2998.25984	750.57229	6.599	2.86
t-Butylation	3141.31185	786.33517	6.956	2.61
Unknown truncation	2691.14227	673.79303	6.145	2.56
Truncation at Cys-15	1482.56786	742.29092	6.464	2.38
Unknown (-17Da)	3068.27178	768.07528	6.661	2.07
Unknown truncation	2741.18613	686.30419	6.507	1.91
Unknown truncation	2594.09369	649.53106	6.048	1.78
	2423.03139	606.76536	6.134	1.64
	3085.23469	618.05701	6.372	1.58
	2925.23701	732.31648	6.356	1.40
	2967.28268	990.10176	7.23	1.33
	2986.19886	747.5571	6.341	1.27
	3071.28364	768.82813	6.479	1.21
	2520.1028	631.03334	6.203	1.12
	2724.17129	682.05082	6.608	1.12
	3028.27154	758.07489	6.574	1.04
	1460.70764	487.90982	4.763	0.98

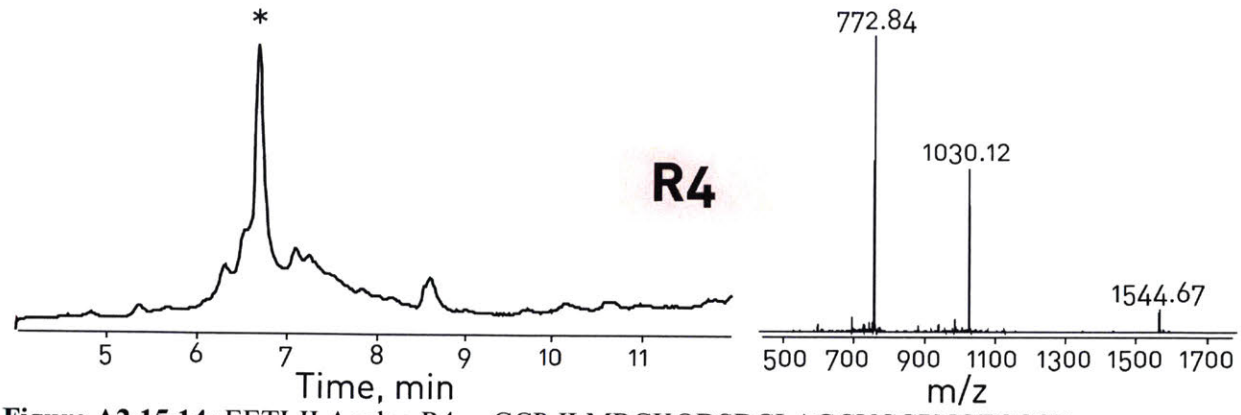


Figure A2.15.14: EETI-II Analog R4 GCPrILMRCKQDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.35Da

Table A2.15.11. Assigned MFE of EETI-II analog R4 synthesized with the second generation AFPS (Figs A2.15.14 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.301	772.3328	6.555	29.21
Unknown truncation	1620.738	406.193	4.773	4.22
Cys, Gly double deletion	2925.243	732.3182	6.399	3.48
t-butylation	3141.361	786.3479	6.964	3.11
Truncation at Cys-15	1482.558	742.2863	6.461	2.82
42Da cap on Pro-3	2967.275	990.1002	7.139	2.4
Unknown (-17Da)	3068.251	768.0705	6.684	2.02
Val deletion	2986.208	747.5594	6.352	1.83
Unknown. Possible Cys, Gly, Ser triple deletion	2838.21	710.5583	6.584	1.4
Ser deletion	2998.256	750.5718	6.613	1.3
	3257.233	815.3159	7.312	1.27
	3084.251	772.0697	6.463	1.24
	3028.243	758.0688	6.593	1.15
	3072.227	769.0639	6.47	0.94
	3067.274	767.8256	6.559	0.89
	2724.167	682.0497	6.628	0.87
	2988.197	748.0572	6.503	0.79
	1421.569	711.7916	7.101	0.79
	1012.447	507.231	5.26	0.74
	2585.018	862.6807	7.216	0.72

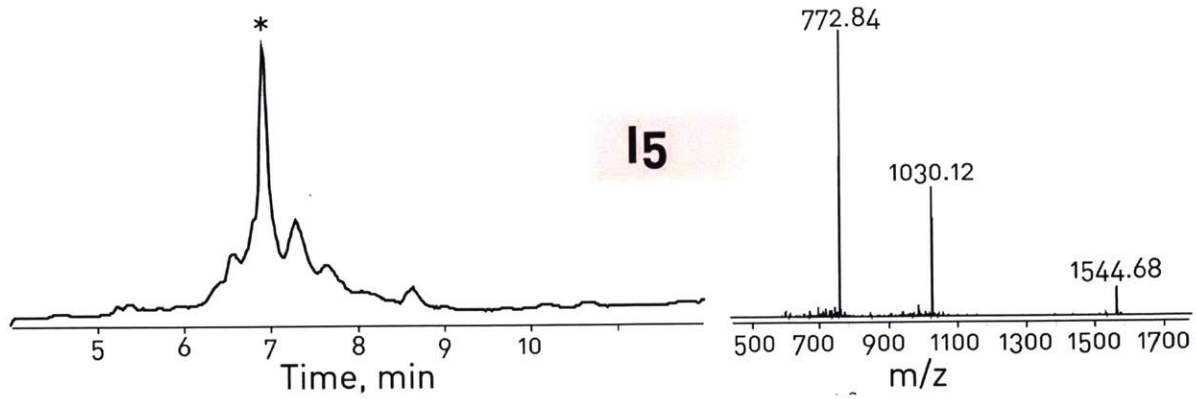


Figure A2.15.15: EETI-II Analog I5 GCPRiILMRCKQDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.35Da

Table A2.15.12. Assigned MFE of EETI-II analog I5 synthesized with the second generation AFPS (Figs A2.15.15 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.313	772.3353	6.732	28.12
Cys, Gly double deletion	2925.273	732.3256	6.625	5.17
Unknown truncation	1620.739	406.1926	5.162	4.89
Truncation at Cys-15	1482.572	742.2936	6.46	2.83
t-Butylation	3141.354	786.3462	7.136	2.77
42Da cap on Pro-3	2967.233	990.0847	7.08	2.36
Unknown (-17Da)	3068.239	768.0673	6.869	2.11
Ser deletion	2998.252	750.5705	6.795	1.92
Unknown. Possible Cys, Gly, Ser triple deletion	2838.255	710.571	6.771	1.59
Val deletion	2986.173	747.5505	6.546	1.43
Gly deletion	3028.286	758.0803	6.759	1.38
	3257.233	815.3154	7.469	1.38
	3067.287	767.8296	6.766	1.17
	3139.23	785.8133	6.728	1.08
	2988.251	748.0698	6.69	1.01
	3099.266	1034.096	7.228	0.97
	3085.294	618.0661	6.554	0.96
	2983.283	746.828	6.648	0.91
	2585.008	862.677	7.214	0.9
	2970.264	743.5735	6.704	0.84

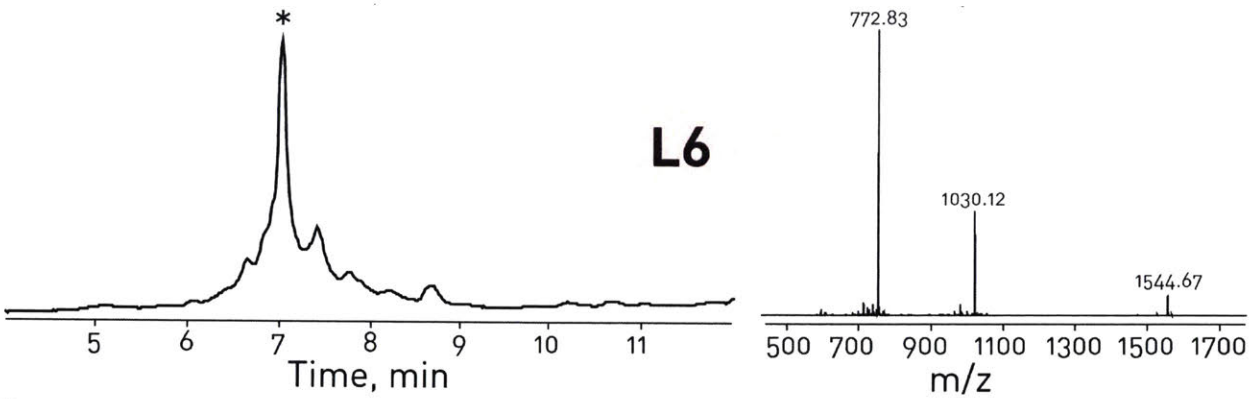


Figure A2.15.16: EETI-II Analog L6 GCPRIIMRCKQDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.32Da

Table A2.15.13. Assigned MFE of EETI-II analog L6 synthesized with the second generation AFPS (Figs A2.15.16 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.308	772.334	6.807	29.05
Cys, Gly double deletion	2925.274	732.3259	6.719	6.33
Unknown truncation	1620.744	406.1932	5.328	4.57
Truncation at Cys-15	1482.587	742.3004	6.46	2.93
t-Butylation	3141.377	786.3519	7.208	2.79
Unknown (-17Da)	3068.27	768.0749	6.944	2.64
Isomer. Unknown source	3085.302	618.0684	6.62	2.31
42Da cap on Pro-3	2967.272	990.0979	7.167	2.11
Val deletion	2986.251	747.5693	6.627	1.91
Gly deletion	3028.273	758.0756	6.834	1.85
Cys deletion	2982.297	746.5817	6.735	1.67
	3257.264	815.3226	7.537	1.61
	3071.28	768.8272	6.669	1.47
	2970.275	743.5761	6.782	1.06
	3067.249	767.8193	6.854	0.86
	2971.215	743.811	6.87	0.77
	3099.277	1034.1	7.278	0.76
	1421.577	711.7957	7.1	0.75
	1012.441	507.2276	5.261	0.73
	2988.267	748.0741	6.76	0.69

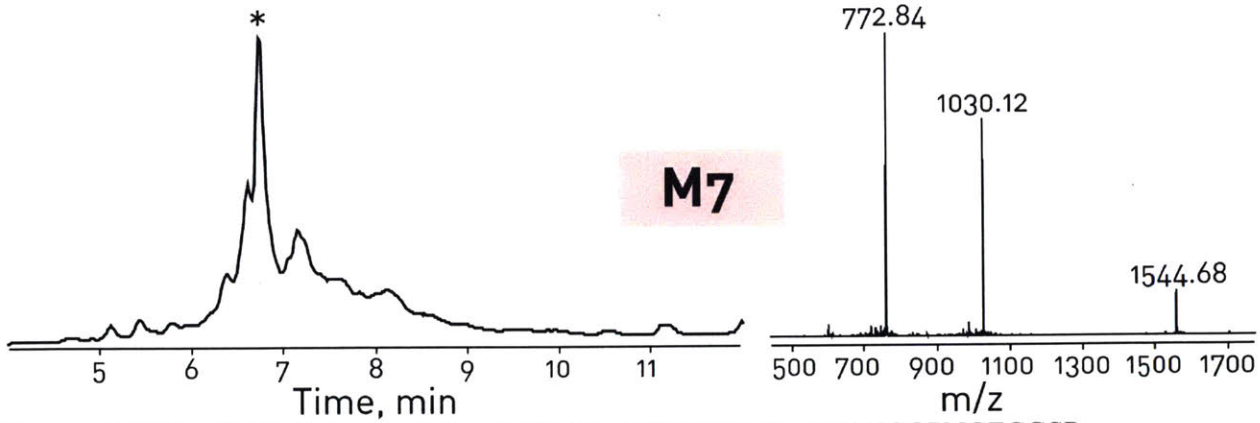


Figure A2.15.17: EETI-II Analog M7 GCPRILmRCKQDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.35Da

Table A2.15.14. Assigned MFE of EETI-II analog M7 synthesized with the second generation AFPS (Figs A2.15.17 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.327	772.3389	6.616	26.54
Cys, Gly double deletion	2925.219	732.3126	6.486	7.58
Unknown truncation	1620.735	406.1917	4.981	3.74
t-Butylation	3141.319	786.3373	7.028	2.66
Truncation at Cys-15	1482.592	742.3031	6.46	2.6
42Da cap on Pro-3	2967.238	990.0873	7.109	2.56
Unknown (-17Da)	3068.277	768.0768	6.757	1.95
Unknown (-14Da)	3071.26	768.8218	6.547	1.67
Val deletion	2986.222	747.5619	6.427	1.63
Cys deletion	2982.276	746.578	6.528	1.58
	3257.236	815.3164	7.371	1.48
	3028.281	758.077	6.654	1.4
	3099.268	1034.097	7.278	0.93
	3067.291	767.8305	6.652	0.92
	1421.554	711.7843	7.101	0.82
	2988.235	748.0661	6.565	0.75
	2981.301	746.3326	6.928	0.74
	1012.44	507.2271	5.261	0.72
	2971.157	743.7964	6.468	0.72
	3085.312	618.0697	6.458	0.7

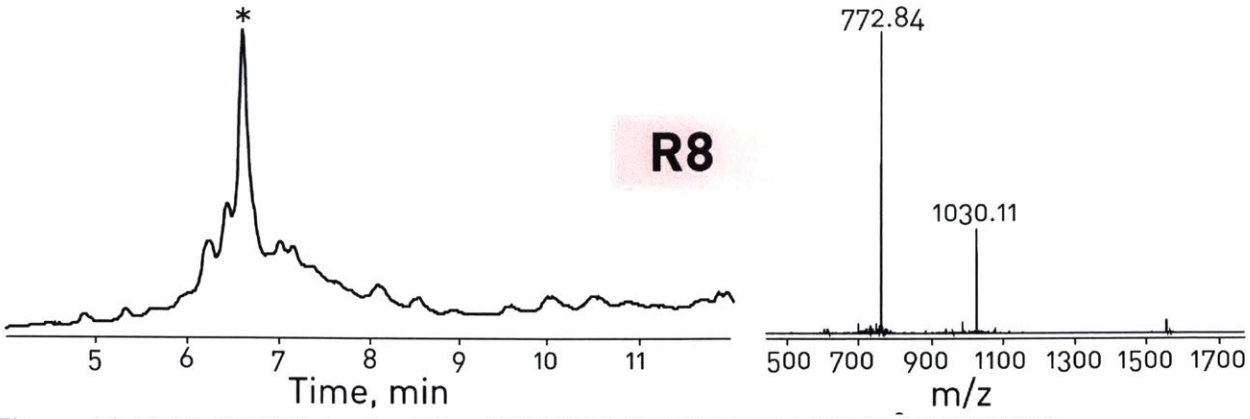


Figure A2.15.18: EETI-II Analog R8 GCPRILMrCKQDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.31Da

Table A2.15.15. Assigned MFE of EETI-II analog R8 synthesized with the second generation AFPS (Figs A2.15.18 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.297	772.3317	6.523	25.00
Cys, Gly double deletion	2925.264	732.3232	6.357	6.45
Unknown truncation	1620.723	406.1884	4.84	4.54
Truncation at Cys-15	1482.583	742.2984	6.461	2.85
t-Butylation	3141.344	786.3433	6.938	2.59
42Da cap on Pro-3	2967.262	990.0948	7.086	2.29
Unknown truncation	2520.08	631.0275	6.193	2.27
Unknown (-17Da)	3068.268	768.0741	6.657	2.15
Unknown cleavage adduct	3257.259	815.3221	7.287	1.34
Unknown truncation	2423.032	606.7655	6.123	1.28
	2982.304	746.5839	6.45	1.26
	2986.202	747.5577	6.33	1.24
	3028.294	758.0819	6.561	1.11
	3071.29	768.8295	6.46	1.05
	2838.26	710.5722	6.556	1.03
	3067.284	767.8285	6.557	1.03
	2971.24	743.821	6.547	1.00
	3083.264	771.8233	6.366	0.87
	2998.252	750.5707	6.583	0.77
	1421.573	711.794	7.101	0.76

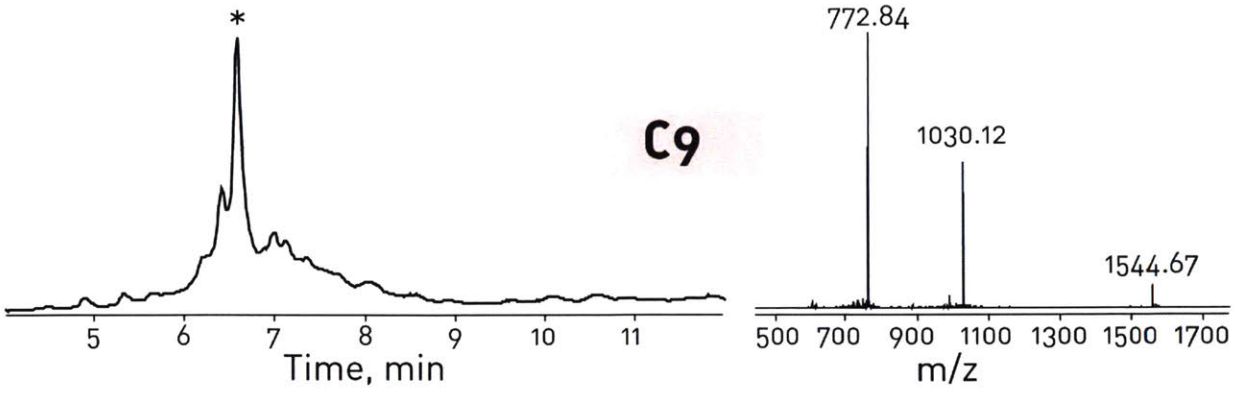


Figure A2.15.19: EETI-II Analog C9 GCPRILMRcKQDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.35Da

Table A2.15.16. Assigned MFE of EETI-II analog C9 synthesized with the second generation AFPS (Figs A2.15.19 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.331	772.3394	6.528	26.42
Cys, Gly double deletion	2925.269	732.3249	6.357	6.86
Unknown truncation	1620.748	406.1945	4.847	4.08
t-Butylation	3141.333	786.3406	6.938	2.91
Truncation at Cys-15	1482.547	742.2808	6.46	2.83
42Da cap on Pro-3	2967.22	990.0807	7.074	2.69
Unknown (-17Da)	3068.269	768.0747	6.665	2.01
Val deletion	2986.187	747.5541	6.334	1.77
Unknown cleavage adduct	3257.232	815.3152	7.286	1.6
Cys deletion	2982.234	746.5661	6.452	1.56
Gly deletion	3028.267	758.0737	6.563	1.53
	3085.284	618.0652	6.389	1.27
	3067.271	767.825	6.562	1.21
	3071.199	768.8084	6.463	0.99
	1012.439	507.2265	5.261	0.89
	2981.304	746.3332	6.807	0.76
	3085.259	618.0599	6.284	0.75
	875.3542	876.3629	5.288	0.68
	1421.564	711.7894	7.1	0.68
	1502.715	501.9124	5.572	0.6

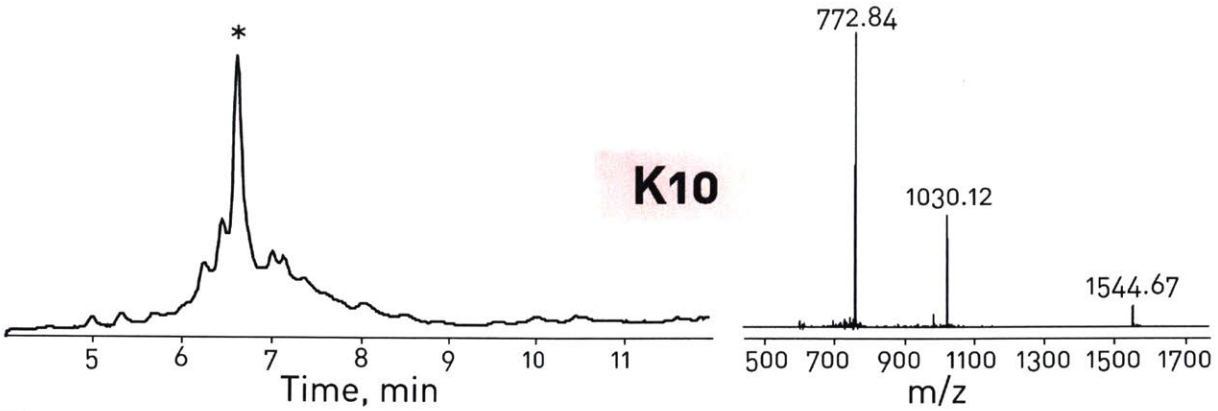


Figure A2.15.20: EETI-II Analog K10 GCPRILMRCkQDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.35Da

Table A2.15.17. Assigned MFE of EETI-II analog K10 synthesized with the second generation AFPS (Figs A2.15.20 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Misassigned product ions	3086.316	772.5862	6.559	20.92
Product	3085.318	772.3371	6.559	9.18
Cys, Gly double deletion	2925.274	732.3258	6.394	6.26
Unknown truncation	1620.753	406.1955	4.952	3.95
Unknown (-17Da)	3068.291	768.0802	6.67	3.78
t-Butylation	3141.38	786.3515	6.962	3
Truncation at Cys-15	1482.586	742.3004	6.461	2.82
42Da cap on Pro-3	2967.293	990.1055	7.095	2.4
Isomer. Unknown source	3085.301	618.0684	6.383	1.81
Unknown cleavage adduct	3257.261	815.3224	7.295	1.57
	2982.293	746.5812	6.482	1.55
	2986.242	747.5679	6.374	1.54
	2520.131	631.04	6.245	1.06
	3071.279	768.827	6.483	1.05
	2838.259	710.5725	6.594	1.03
	3028.293	758.0804	6.596	1.01
	2423.073	606.7757	6.178	0.9
	1421.585	711.7997	7.101	0.76
	2691.179	673.8022	6.183	0.67
	3098.322	786.5805	6.544	0.67

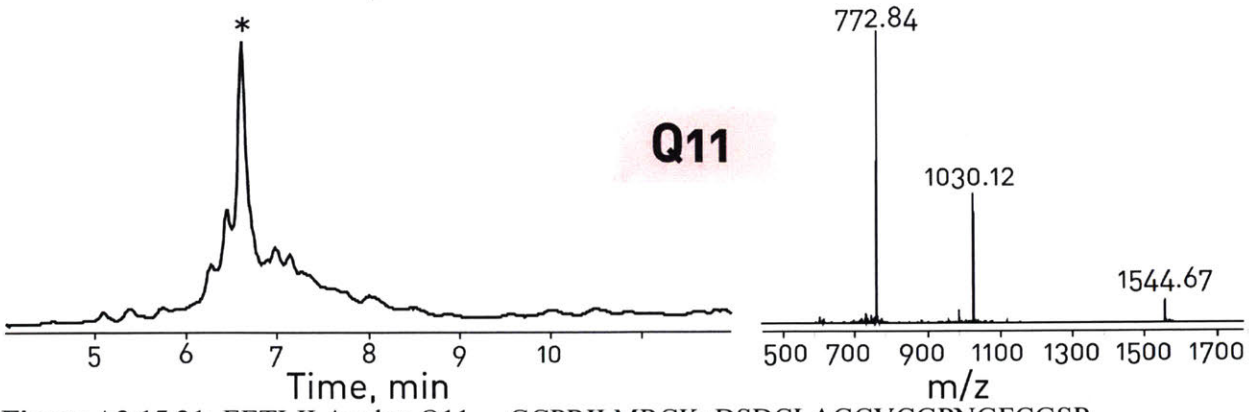


Figure A2.15.21: EETI-II Analog Q11 GCPRILMRCKqDSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.34Da

Table A2.15.18. Assigned MFE of EETI-II analog Q11 synthesized with the second generation AFPS (Figs A2.15.21 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Misassigned product ions	3086.31	772.5847	6.562	23.04
Product	3085.306	772.3339	6.561	9.68
Cys, Gly double deletion	2925.279	732.3273	6.387	6.84
Unknown truncation	1620.751	406.1949	5.008	4.54
Truncation at Cys-15	1482.582	742.2979	6.463	3.75
t-Butylation	3141.364	786.348	6.963	3.21
Unknown (-17Da)	3068.293	768.0807	6.643	3.15
42Da cap on Pro-3	2967.288	990.1041	7.134	2.87
Cys deletion	2982.304	746.5836	6.488	2.21
Val deletion	2986.233	747.5656	6.375	1.63
	2520.12	631.0374	6.248	1.57
	2925.271	732.3251	6.519	1.2
	3028.287	758.0789	6.6	1.17
	1421.582	711.7981	7.103	1.05
	3071.282	768.8275	6.483	1.02
	3051.303	785.8146	6.561	1.01
	3083.286	771.8289	6.397	0.96
	1012.442	507.2278	5.262	0.84
	3085.299	618.0671	6.373	0.84
	2724.187	682.0543	6.641	0.8

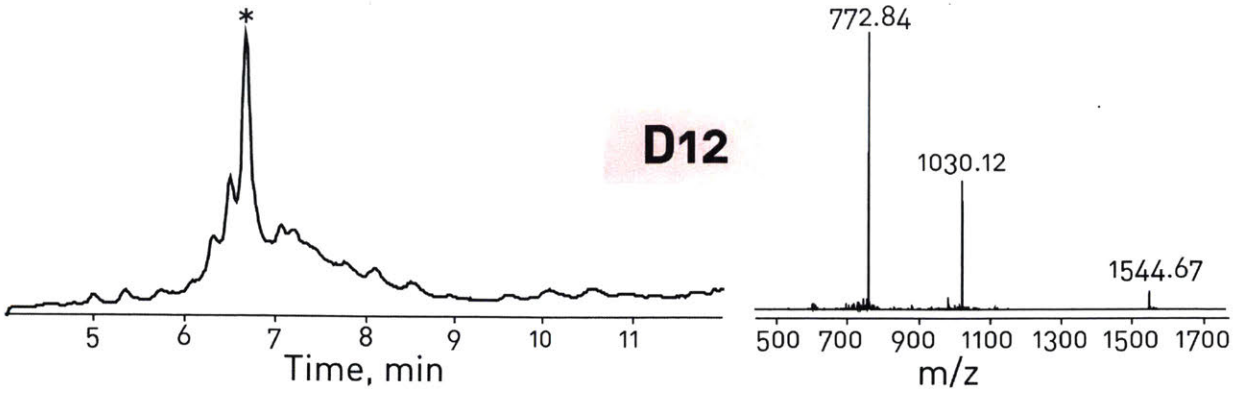


Figure A2.15.22: EETI-II Analog D12 GCPRILMRCKQdSDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.34Da

Table A2.15.19. Assigned MFE of EETI-II analog D12 synthesized with the second generation AFPS (Figs A2.15.22 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.296	772.3304	6.584	23.76
Cys, Gly double deletion	2925.258	732.3219	6.406	7.32
Unknown truncation	1620.755	406.1964	4.935	3.36
42Da cap on Pro-3	2967.287	990.1031	7.166	3.13
Truncation at Cys-15	1482.567	742.2904	6.461	2.98
t-Butylation	3141.348	786.3443	6.995	2.77
Cys deletion	2982.238	746.5673	6.518	1.78
Unknown (-14Da)	3071.253	768.8206	6.511	1.27
Unknown truncation	1012.445	507.2291	5.26	1.19
Val deletion	2986.226	747.5634	6.4	1.18
	1421.561	711.7881	7.101	1.06
	3099.244	1034.088	7.382	0.99
	3069.246	768.3186	6.724	0.97
	875.353	876.3594	5.291	0.85
	2981.27	746.325	6.859	0.78
	2970.234	743.5666	6.554	0.77
	2698.069	900.3635	7.736	0.76
	2998.223	750.5632	6.647	0.73
	2585.006	1293.515	7.27	0.67
	1077.429	1078.437	6.188	0.62

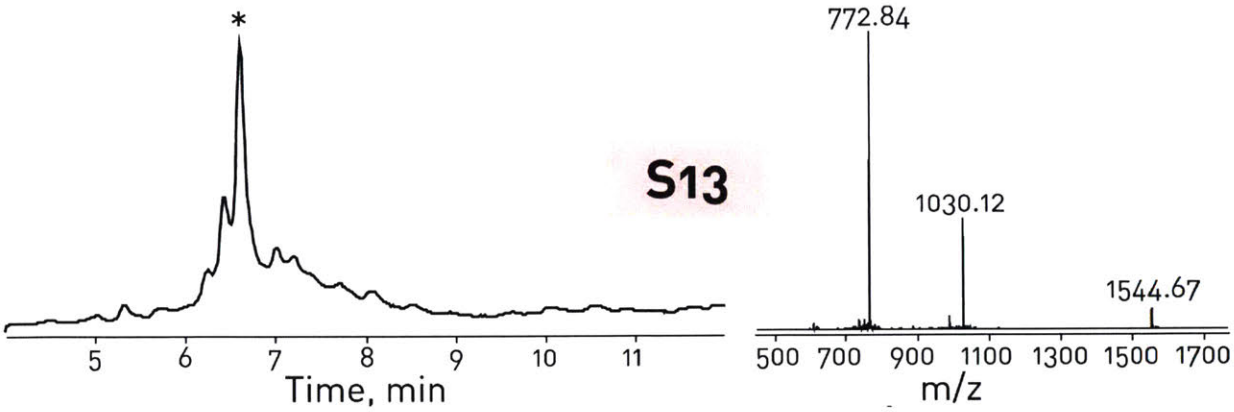


Figure A2.15.23: EETI-II Analog S13 GCPRILMRCKQDsDCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.34Da

Table A2.15.20. Assigned MFE of EETI-II analog S13 synthesized with the second generation AFPS (Figs A2.15.23 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.292	772.3297	6.556	18.44
Cys, Gly double deletion	2925.262	732.3222	6.377	7.40
Probably misassigned product ions	3086.301	1029.774	6.556	5.54
Unknown truncation	1620.754	541.2585	4.973	3.78
42Da cap on Pro-3	2967.268	990.0964	7.175	3.18
Truncation at Cys-15	1482.58	742.2961	6.461	2.98
t-Butylation	3141.275	786.3256	6.972	2.64
Unknown truncation	1012.445	507.2293	5.26	1.31
Val deletion	2986.238	747.5668	6.378	1.19
Unknown (-14Da)	3099.252	1034.091	7.376	0.99
	3083.288	1028.77	6.553	0.97
	2982.29	746.5805	6.469	0.96
	1421.57	711.7919	7.103	0.96
	3085.296	618.0664	6.337	0.83
	875.363	876.3713	5.291	0.79
	3028.268	758.0751	6.526	0.79
	3028.252	758.0703	6.608	0.75
	2998.244	750.5686	6.616	0.69
	1535.126	768.5702	6.669	0.67
	3257.235	815.3159	7.305	0.65

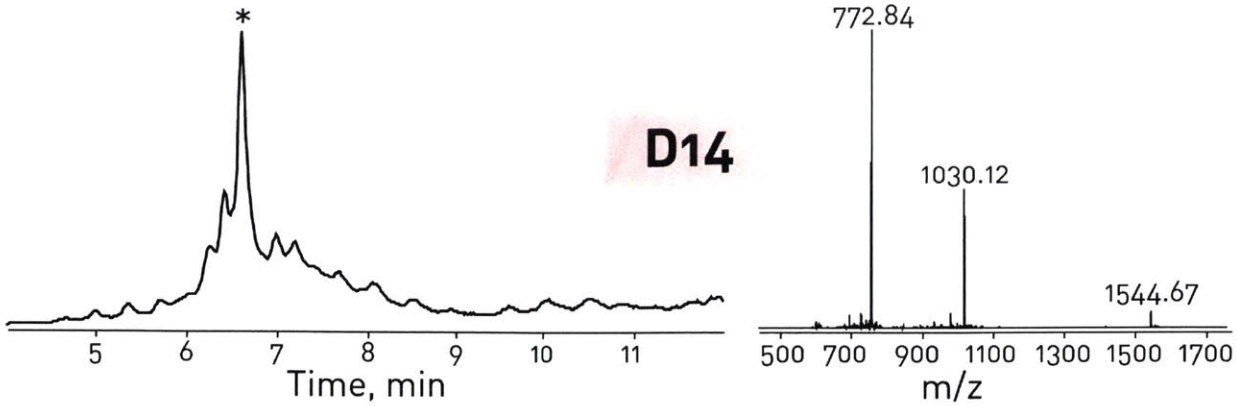


Figure A2.15.24: EETI-II Analog D14 GCPRILMRCKQDSdCLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.34Da

Table A2.15.21. Assigned MFE of EETI-II analog D14 synthesized with the second generation AFPS (Figs A2.15.24 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.257	772.3212	6.55	24.13
Cys, Gly double deletion	2925.258	732.3219	6.352	7.13
Unknown truncation	1620.749	406.1948	4.937	3.04
42Da cap on Pro-3	2967.248	990.0901	7.169	2.97
t-Butylation	3141.318	786.3365	6.948	2.95
Truncation at Cys-15	1482.588	742.301	6.462	1.85
Asp deletion	2970.228	743.5647	6.559	1.58
Cys deletion	2982.282	746.5782	6.49	1.49
Unknown truncation	1012.444	507.2281	5.261	1.15
Unknown. Possibly Cys, Gly, Ser triple deletion	2838.202	710.5579	6.585	1.14
	3071.272	768.8251	6.473	1.12
	3069.245	768.3191	6.668	1.1
	3099.238	1034.087	7.389	1.09
	2987.242	747.8177	6.372	1.01
	3028.233	758.0659	6.594	0.94
	1421.561	711.7873	7.103	0.92
	2698.072	900.3643	7.654	0.79
	875.3604	876.3686	5.292	0.75
	3139.183	785.8034	6.546	0.74
	2981.259	746.3222	6.796	0.73

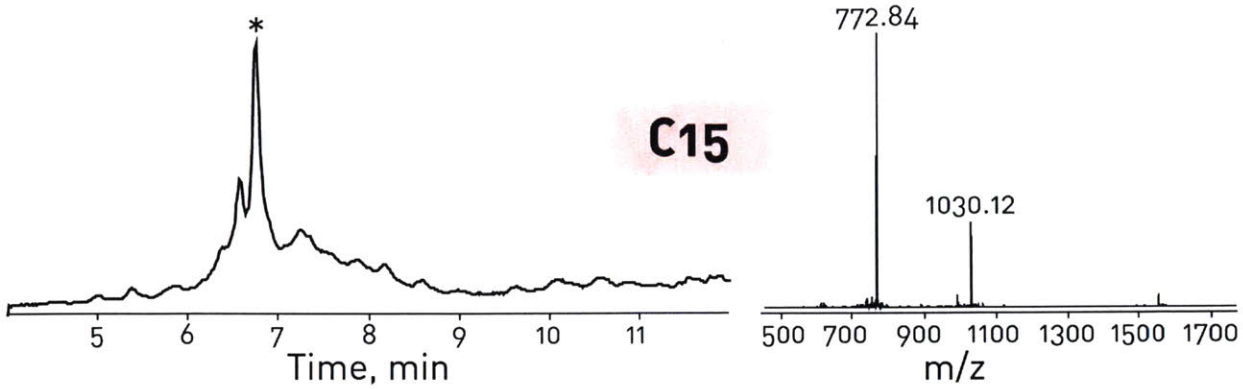


Figure A2.15.25: EETI-II Analog C15 GCPRILMRCKQDSDcLAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.34Da

Table A2.15.22. Assigned MFE of EETI-II analog C15 synthesized with the second generation AFPS (Figs A2.15.25 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.285	772.3286	6.653	22.58
Cys, Gly double deletion	2925.239	732.317	6.49	7.41
Unknown truncation	1620.754	406.196	4.933	3.42
42Da cap on Pro-3	2967.256	990.0927	7.249	3.05
Aspartimide	3067.238	767.8169	6.789	2.27
Val deletion	2986.213	747.5606	6.478	2.21
t-Butylation	3141.314	786.3357	7.152	1.94
Truncation at Cys-15	1482.569	742.291	6.694	1.72
Cys deletion	2982.291	746.5809	6.588	1.4
Gly deletion	3028.244	758.0682	6.697	1.22
	1012.442	507.2282	5.26	1.13
	3141.319	786.3369	7.062	1.13
	3099.247	1034.09	7.459	1.02
	2971.168	743.7974	6.703	0.89
	3072.284	769.0782	6.586	0.87
	3085.266	618.0608	6.465	0.85
	875.3601	876.3683	5.29	0.74
	1946.889	487.7297	5.844	0.73
	3118.377	624.6841	6.391	0.65
	3067.261	767.8231	6.636	0.65

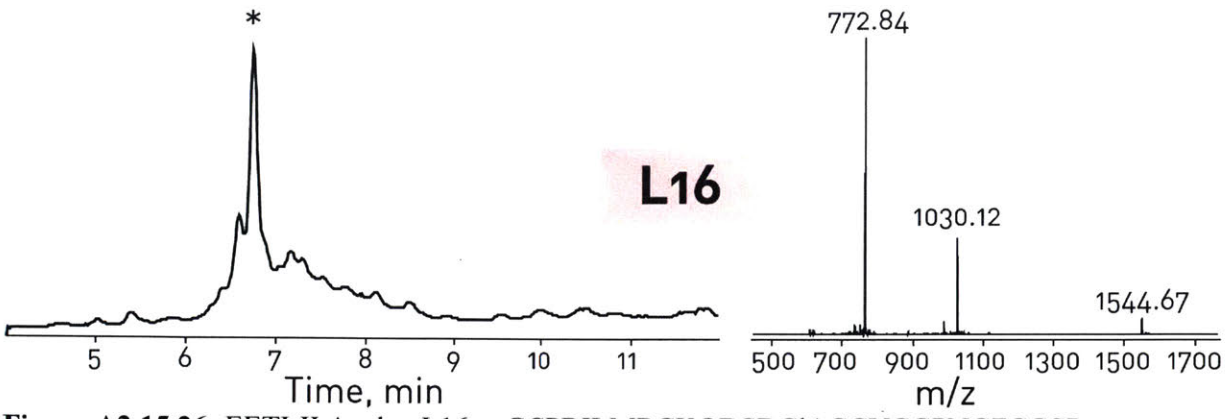


Figure A2.15.26: EETI-II Analog L16 GCPRILMRCKQDSDCIAGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.34Da

Table A2.15.23. Assigned MFE of EETI-II analog L16 synthesized with the second generation AFPS (Figs A2.15.26 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Misassigned product ions	3086.288	772.5794	6.675	13.74
Product	3085.296	1029.439	6.674	7.89
Cys, Gly double deletion	2925.266	732.3235	6.513	7.23
Unknown truncation	1620.747	406.1942	4.93	3.66
42Da cap on Pro-3	2967.26	990.0942	7.271	3.11
t-Butylation	3141.342	786.3431	7.118	3.02
Truncation at Cys-15	1482.568	742.2904	6.826	2.84
Possibly misassigned product ions	3083.273	771.8257	6.668	2.22
Aspartimide	3067.253	767.8202	6.805	2.09
Val deletion	2986.246	747.5681	6.481	1.46
Cys deletion	2982.285	746.5787	6.58	1.31
	3085.298	618.0681	6.494	1.20
	3071.279	768.8269	6.608	1.06
	1012.446	507.2293	5.259	1.00
	3099.277	1034.1	7.48	0.98
	2970.26	743.5722	6.661	0.95
	3028.27	758.0748	6.642	0.76
	2981.317	746.3366	6.997	0.75
	875.3625	876.3712	5.289	0.73
	1421.572	711.7927	7.207	0.73

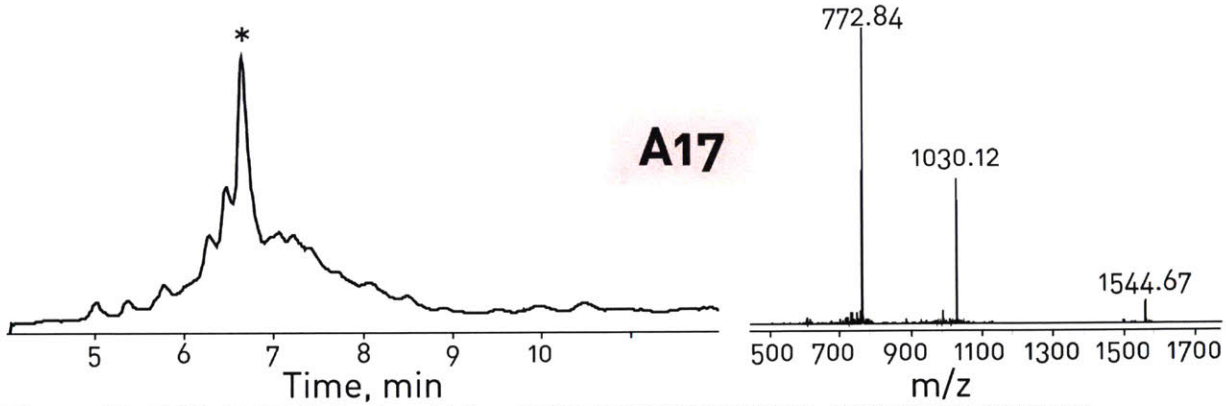


Figure A2.15.27: EETI-II Analog A17 GCPRILMRCKQDSDCLaGCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.34Da

Table A2.15.24. Assigned MFE of EETI-II analog A17 synthesized with the second generation AFPS (Figs A2.15.27 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.265	772.323	6.587	23.04
Cys, Gly double deletion	2925.262	732.3227	6.408	7.15
Unknown truncation	1620.752	406.1953	4.932	4.40
42Da cap on Pro-3	2967.256	990.0929	7.186	2.87
Truncation at Cys-15	1482.569	742.2917	6.636	2.00
Val deletion	2986.228	747.5635	6.404	1.78
Cys deletion	2982.274	746.5764	6.487	1.64
t-Butylation	3141.346	786.3438	7.006	1.44
t-Butylation	3141.292	786.3303	6.923	1.31
Aspartimide	3067.258	767.822	6.642	1.29
	1012.445	507.2295	5.26	1.11
	3099.244	1034.089	7.397	1.06
	3069.232	768.3152	6.729	1.03
	3071.273	768.8253	6.514	1.01
	3028.251	758.0697	6.579	0.96
	2998.236	750.5668	6.654	0.91
	1946.898	487.7318	5.817	0.86
	1502.73	501.9175	5.719	0.76
	1421.572	711.7934	7.156	0.73
	875.3581	876.3657	5.29	0.69

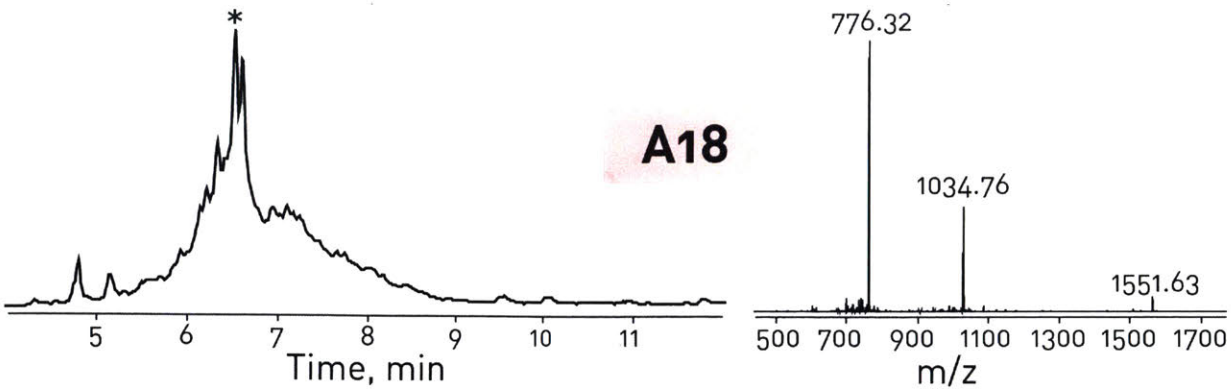


Figure A2.15.28: EETI-II Analog A18 GCPRILMRCKQDSDCLAaCVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3099.34, Observed monoisotopic mass = 3099.26

Table A2.15.25. Assigned MFE of EETI-II analog A18 synthesized with the second generation AFPS (Figs A2.15.28 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3099.273	775.8256	6.604	12.72
Ala deletion	3028.116	758.0375	6.686	11.16
Cys, Gly double deletion	2939.279	735.8271	6.413	4.1
Unknown truncation	2868.212	718.06	6.497	3.17
Unknown truncation	1620.753	406.1956	4.932	3.16
Truncation at Cys-15	1496.574	749.2942	6.566	1.64
Unknown truncation	2534.105	634.5338	6.301	1.56
Unknown truncation	1012.442	507.228	5.259	1.53
t-Butylation	3155.247	789.8195	7.004	1.44
42Da cap on Pro-3	2981.259	994.7608	7.216	1.29
	3081.121	771.2887	6.684	1.23
	2910.172	971.0642	7.296	1.17
	2852.172	714.0506	6.638	1.11
	875.3564	876.364	5.289	1.08
	3084.228	772.065	7.139	1.06
	1425.557	713.7856	6.548	0.99
	2929.196	733.3065	6.432	0.95
	2781.129	696.2908	6.717	0.87
	1435.556	718.7851	7.166	0.86
	2437.001	610.2586	6.239	0.84

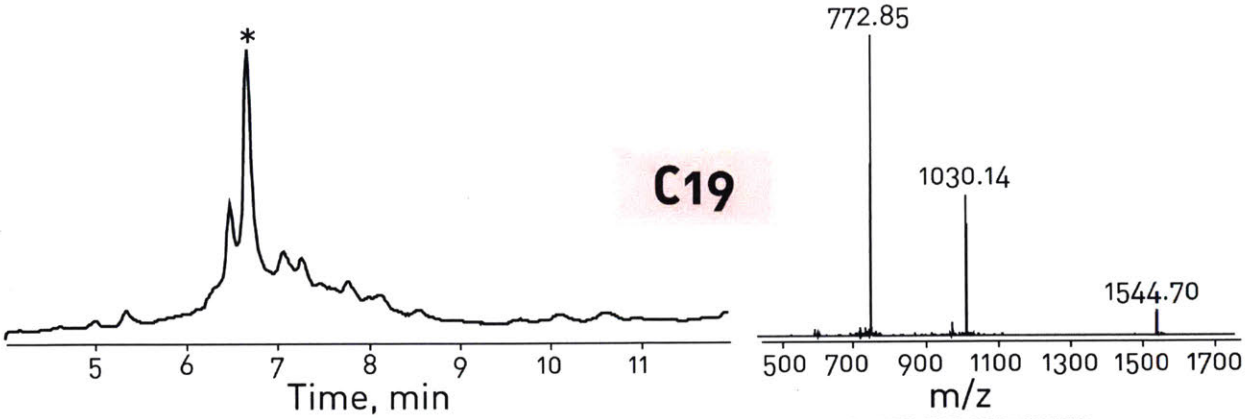


Figure A2.15.29: EETI-II Analog C19 GCPRILMRCKQDSDCLAGcVCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.39Da

Table A2.15.26. Assigned MFE of EETI-II analog C19 synthesized with the second generation AFPS (Figs A2.15.29 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.282	772.3269	6.605	24.23
Cys, Gly double deletion	2925.262	732.3223	6.414	7.95
Unknown truncation	1620.749	406.1942	4.932	3.79
42Da cap on Pro-3	2967.271	990.098	7.221	3.41
t-Butylation	3141.305	786.3334	7.001	2.96
Truncation at Cys-15	1482.581	742.2966	6.565	2.72
Val deletion	2986.229	747.5638	6.402	1.61
Cys deletion	2982.303	746.5845	6.502	1.32
Unknown truncation	1012.443	507.228	5.261	1.17
Unknown (-14Da)	3071.274	768.8257	6.529	1.17
	3099.278	1034.1	7.437	1.11
	2981.255	746.3214	6.858	0.90
	3028.252	758.0699	6.646	0.77
	875.3591	876.3669	5.291	0.76
	1421.57	711.7914	7.169	0.73
	2698.087	900.3692	7.734	0.72
	3257.245	815.3187	7.396	0.65
	1502.731	501.9179	5.72	0.61
	2827.228	707.8148	6.335	0.59
	2851.181	951.4008	7.118	0.57

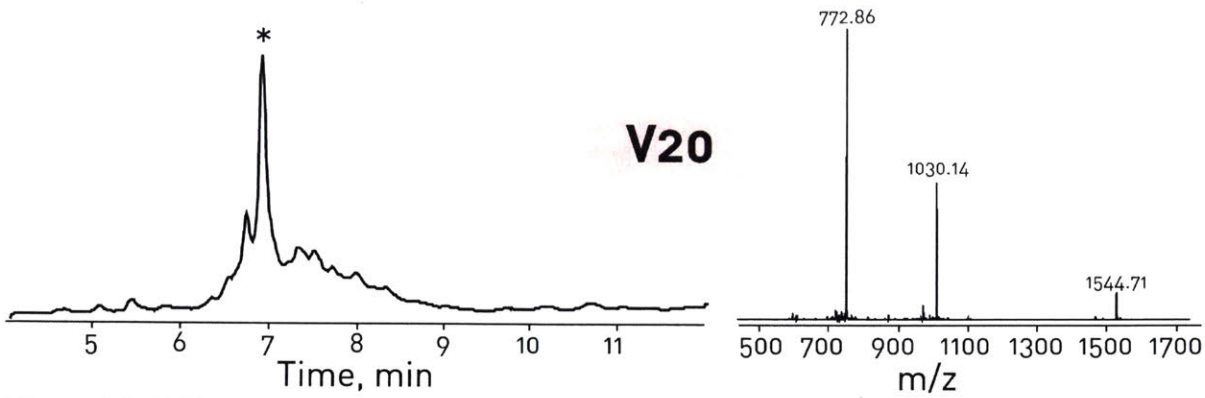


Figure A2.15.30: EETI-II Analog V20 GCPRILMRCKQDSDCLAGCvCGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.41Da

Table A2.15.27. Assigned MFE of EETI-II analog V20 synthesized with the second generation AFPS (Figs A2.15.30 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.293	772.3293	6.685	24.48
Cys, Gly double deletion	2925.288	732.329	6.512	8
42Da cap on Pro-3	2967.25	990.0907	7.289	3.36
t-Butylation	3141.339	786.3421	7.097	3.1
Unknown truncation	1620.749	406.1946	4.934	2.28
Aspartimide	3067.282	767.8279	6.755	2.05
Truncation at Cys-15	1482.583	742.2977	6.691	1.8
Unknown truncation	1012.445	507.2294	5.261	1.44
Unknown (-14Da)	3071.267	768.8238	6.614	1.33
Cys deletion	2982.289	746.5795	6.551	1.29
	875.358	876.3657	5.29	0.98
	2698.084	900.3686	7.82	0.85
	2981.31	746.336	6.966	0.77
	1421.572	711.7931	7.324	0.77
	1077.434	1078.442	6.353	0.73
	1012.445	1013.453	6.435	0.73
	3099.249	1034.091	7.501	0.72
	2988.247	748.0692	6.634	0.71
	2851.172	951.3979	7.209	0.69
	3138.234	628.6526	6.685	0.62

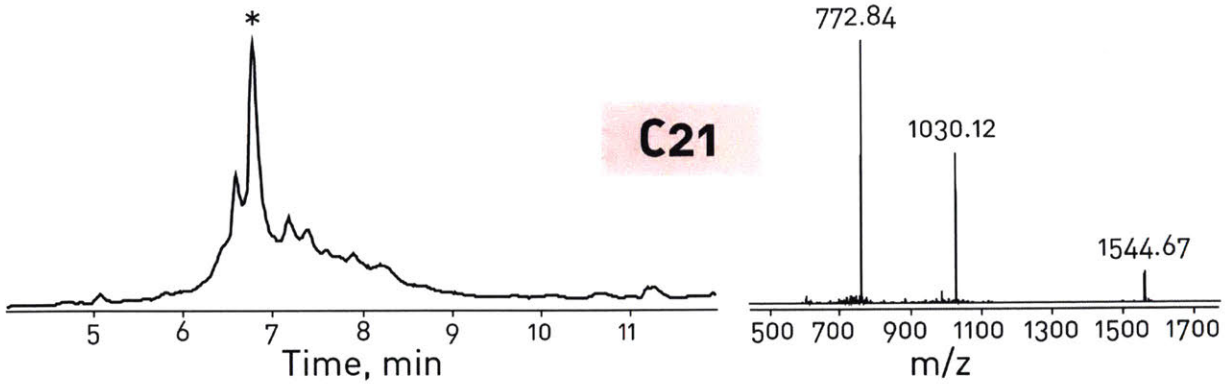


Figure A2.15.31: EETI-II Analog C21 GCPRILMRCKQDSDCLAGCVcGPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.33Da

Table A2.15.28. Assigned MFE of EETI-II analog C21 synthesized with the second generation AFPS (Figs A2.15.31 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.406	772.3602	6.744	30.29
Cys, Gly double deletion	2925.37	732.351	6.555	8.57
42Da cap on Pro-3	2967.377	990.1328	7.353	3.78
t-Butylation	3141.469	786.3759	7.14	3.08
Val deletion	2986.356	747.5975	6.548	2.60
Cys deletion	2982.406	746.6124	6.636	1.78
Unknown truncation	1620.785	406.2047	5.083	1.60
Unknown (-14Da)	3071.389	768.8557	6.672	1.60
Asn deletion	2971.334	991.4501	6.753	1.59
Ser deletion	2998.372	750.6019	6.812	1.54
	3028.377	758.1037	6.771	1.51
	2838.341	710.5942	6.778	1.42
	1482.624	742.3208	6.671	0.99
	2585.095	1293.553	7.384	0.96
	3069.397	768.3565	6.838	0.93
	2698.138	900.3884	7.849	0.89
	2981.428	746.3649	6.992	0.87
	2851.298	951.4417	7.241	0.80
	1421.622	711.8208	7.295	0.80
	1077.46	1078.464	6.446	0.78

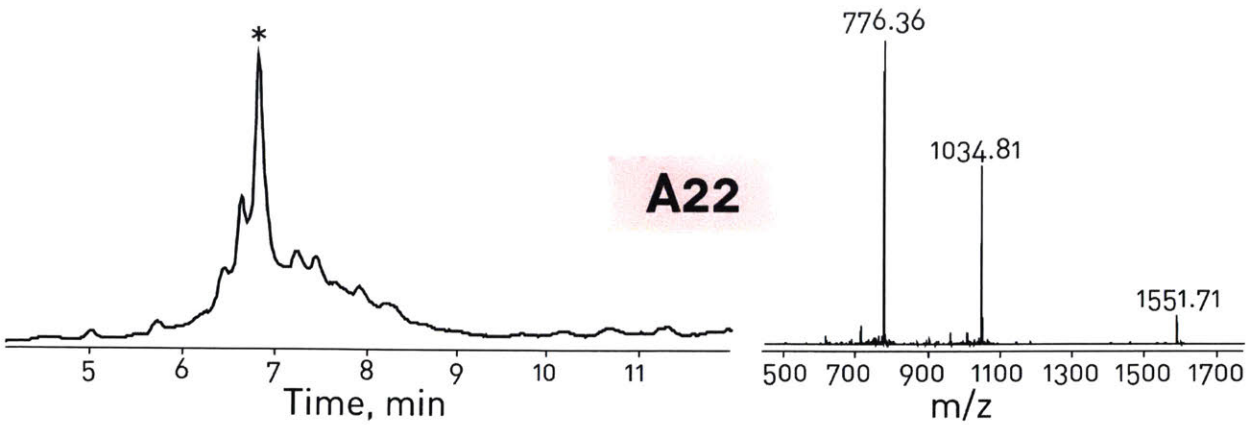


Figure A2.15.32: EETI-II Analog A22 GCPRILMRCKQDSDCLAGCVCaPNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3099.34, Observed monoisotopic mass = 3099.41

Table A2.15.29. Assigned MFE of EETI-II analog A22 synthesized with the second generation AFPS (Figs A2.15.32 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3099.427	775.8654	6.886	28.28
Cys, Gly double deletion	2939.389	735.8557	6.699	7.74
Val deletion	3000.36	751.0984	6.713	4.45
42Da cap on Pro-3	2981.39	994.8042	7.486	2.53
Unknown. +24Da cap on Arg-4	2852.358	714.0989	6.901	2.34
Unknown (-14Da)	3085.41	772.3605	6.814	2.03
t-Butylation	3155.211	789.8116	7.297	1.71
Ser deletion	3012.39	754.1062	6.945	1.64
Unknown	1620.774	406.2019	5.101	1.50
t-Butylation of truncation at G-22	889.3816	890.3877	5.763	1.13
	2840.288	711.0814	6.493	1.08
	1496.639	749.3281	6.834	1.04
	2599.086	1300.548	7.532	0.94
	3155.487	789.8806	7.249	0.92
	2712.199	905.076	8.001	0.87
	3082.417	771.6118	6.954	0.85
	2534.186	634.5551	6.551	0.83
	1435.639	718.8287	7.488	0.82
	2996.411	750.1108	6.78	0.80
	2995.443	749.8685	7.141	0.77

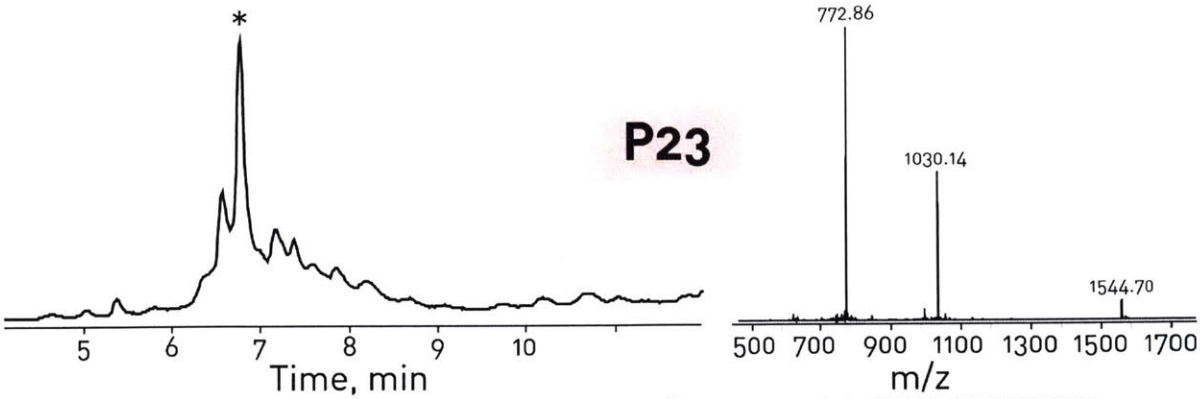


Figure A2.15.33: EETI-II Analog P23 GCPRILMRCKQDSDCLAGCVCGpNGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.40Da

Table A2.15.30. Assigned MFE of EETI-II analog P23 synthesized with the second generation AFPS (Figs A2.15.33 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.333	772.3418	6.719	30.16
Cys, Gly double deletion	2925.328	732.3401	6.517	8.87
42 Da cap on Pro-3	2967.357	990.1262	7.345	3.38
t-Butylation	3141.473	786.3766	7.12	3.03
Unknown (-14Da)	3071.316	768.8368	6.644	1.78
Val deletion	2986.291	747.5804	6.552	1.77
Unknown truncation	1012.448	507.2329	5.438	1.45
Asn deletion	2971.268	991.4255	6.749	1.34
Gly deletion	3028.302	758.0844	6.73	1.25
Unknown truncation	2698.12	900.3824	7.835	1.17
	1482.588	742.3027	6.587	1.15
	875.3629	876.3697	5.447	1.14
	1421.622	711.8187	7.214	1.12
	1620.769	406.2003	5.111	1.07
	2585.072	1293.539	7.341	1.05
	3070.308	768.5851	6.819	0.93
	2851.298	951.4434	7.203	0.89
	3166.359	1056.461	7.879	0.89
	1077.435	1078.44	6.307	0.86
	2981.358	746.3476	6.957	0.83

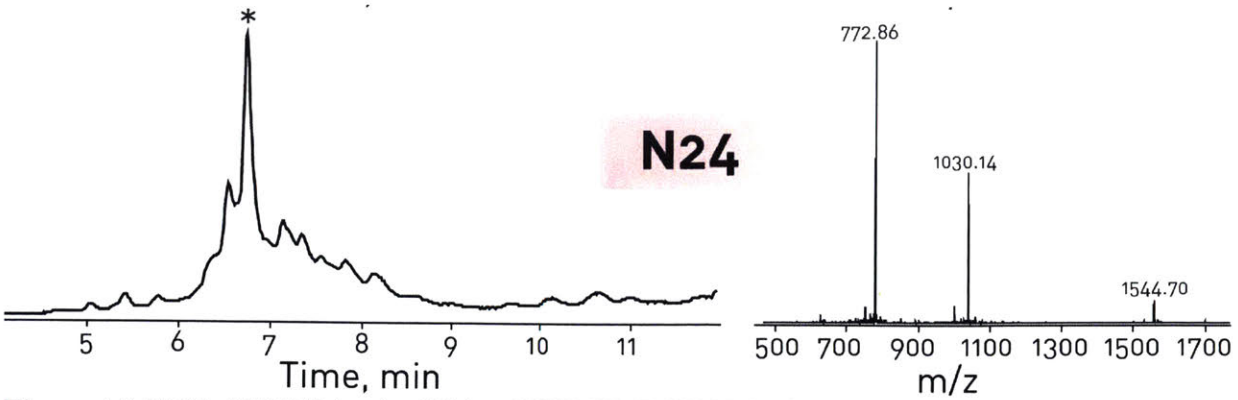


Figure A2.15.34: EETI-II Analog N24 GCPRILMRCKQDSDCLAGCVCGPnGFCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.39Da

Table A2.15.31. Assigned MFE of EETI-II analog N24 synthesized with the second generation AFPS (Figs A2.15.34 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.385	772.3549	6.795	28.42
Cys, Gly double deletion	2925.349	732.3456	6.581	7.91
Asn deletion	2971.332	743.8433	6.839	3.80
42Da cap on Pro-3	2967.357	990.126	7.42	3.68
t-Butylation	3141.446	786.3699	7.211	3.64
Val deletion	2986.324	747.5897	6.602	1.80
Aspartimide	3067.371	767.8515	6.84	1.66
Unknown (-13Da)	3072.367	769.1006	6.713	1.40
Unknown truncation	1012.456	507.2365	5.498	1.35
Gly deletion	3028.355	758.0977	6.805	1.31
	875.3682	876.3746	5.515	1.20
	2811.296	703.8341	6.622	1.18
	2698.122	900.3857	7.885	1.17
	2585.078	1293.544	7.424	1.06
	3199.423	800.8647	6.7	1.04
	1421.612	711.8154	7.302	1.02
	3166.402	1056.474	7.917	0.98
	2981.411	746.3605	7.049	0.95
	2851.278	951.4356	7.29	0.85
	1012.459	1013.463	6.637	0.83

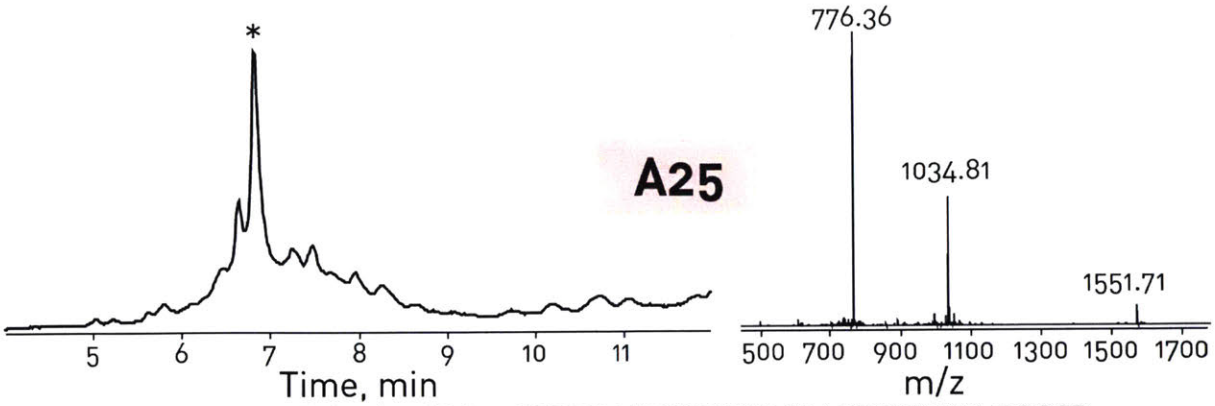


Figure A2.15.35: EETI-II Analog A25 GCPRILMRCKQDSDCLAGCVCGPNaFCGSP TIC is shown. Calculated monoisotopic mass = 3099.34Da, Observed monoisotopic mass = 3099.42Da

Table A2.15.32. Assigned MFE of EETI-II analog A25 synthesized with the second generation AFPS (Figs A2.15.35 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3099.331	775.8413	6.89	28.99
Cys, Gly double deletion	2939.333	735.842	6.697	6.80
42Da cap on Pro-3	2981.31	994.7777	7.496	3.52
t-Butylation	3155.346	789.8444	7.28	2.83
Val deletion	3000.301	751.0836	6.71	2.00
Unknown truncation	1026.463	514.2405	5.726	1.45
42Da cap on Gly-22	889.3756	890.3817	5.849	1.34
Gly deletion	3042.311	761.5866	6.871	1.24
Unknown truncation	1620.768	541.2626	5.197	1.15
Cys deletion	2996.335	750.0919	6.797	1.14
	3086.334	772.5908	6.815	1.12
	1435.6	718.8097	7.489	1.10
	2712.089	905.0393	7.986	1.08
	2423.106	606.7852	6.436	1.05
	2984.313	995.7726	6.863	1.05
	3081.31	771.336	6.91	1.02
	2599.044	1300.526	7.526	0.98
	3180.347	1061.122	8.004	0.98
	2520.145	631.0452	6.495	0.97
	2865.206	956.0765	7.368	0.97

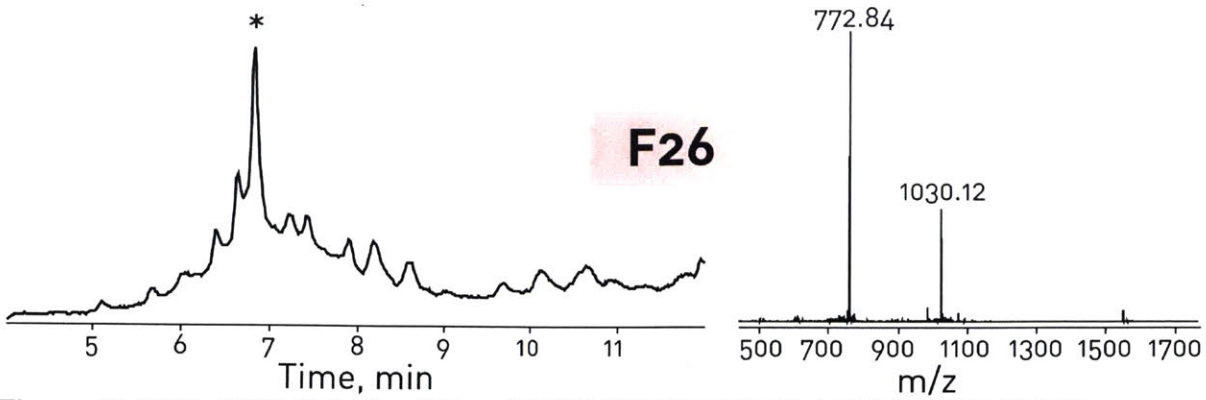


Figure A2.15.36: EETI-II Analog F26 GCPRILMRCKQDSDCLAGCVCGPNGfCGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.33

Table A2.15.33. Assigned MFE of EETI-II analog F26 synthesized with the second generation AFPS (Figs A2.15.36 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.383	772.3545	6.896	25.52
Cys, Gly double deletion	2925.349	732.3458	6.706	7.20
Phe deletion	2938.313	735.587	6.463	4.23
Val deletion	2986.334	747.5917	6.715	2.96
42Da cap on Pro-3	2967.355	990.1257	7.48	2.83
t-butylation	3141.448	786.3704	7.279	2.24
Unknow. +26Da cap on Leu 6	2585.078	1293.543	7.5	1.74
Cys deletion	2982.379	746.6049	6.757	1.72
Unknown truncation	1012.458	507.2378	5.709	1.70
Unknown truncation	1620.787	541.2687	5.114	1.50
	875.3711	876.3775	5.729	1.38
	2778.276	695.5778	6.18	1.19
	2698.135	900.3875	7.96	1.19
	1421.611	711.8148	7.448	1.15
	3067.37	767.8515	6.912	0.99
	2838.321	710.5893	6.949	0.92
	1482.612	742.3147	6.824	0.90
	3072.371	769.1001	6.825	0.89
	3028.365	758.0986	6.919	0.88
	2971.303	991.4417	6.896	0.87

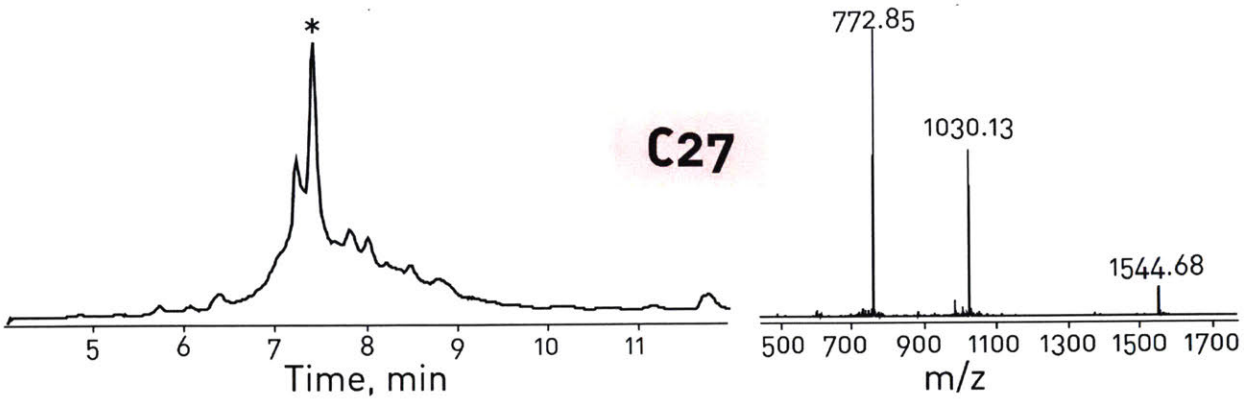


Figure A2.15.37: EETI-II Analog C27 GCPRILMRCKQDSDCLAGCVCGPNGFcGSP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.33Da

Table A2.15.34. Assigned MFE of EETI-II analog C27 synthesized with the second generation AFPS (Figs A2.15.37 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.366	772.3479	6.931	27.63
Cys, Gly double deletion	2925.346	732.3451	6.723	8.04
42Da cap on Pro-3	2967.352	990.1241	7.533	3.63
Cys deletion	2982.369	746.601	6.757	2.65
t-Butylation	3141.44	786.3684	7.341	2.64
Possibly Val deletion (-1Da)	2987.325	747.8395	6.726	1.74
Unknown. -14Da	3071.366	768.8495	6.849	1.70
Isomer. Unknown source	3085.383	772.3515	6.799	1.51
Unknown truncation	1012.457	507.2373	5.782	1.39
42Da cap on G-22	875.3697	876.376	5.835	1.27
	2698.134	900.3869	8.015	1.21
	3028.357	758.0976	6.929	1.17
	2838.318	710.5894	6.841	1.13
	2585.071	1293.539	7.563	1.12
	3166.415	1056.479	8.032	1.01
	1620.779	406.2033	5.201	0.99
	3095.441	774.8676	7.137	0.98
	1421.608	711.8134	7.523	0.95
	2971.278	991.4332	6.932	0.94
	3070.365	768.5997	7.02	0.82

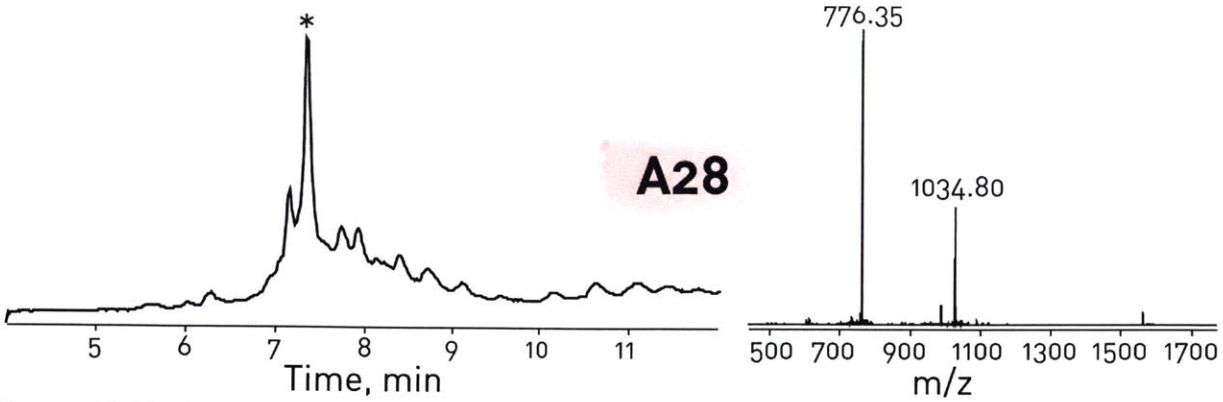


Figure A2.15.38: EETI-II Analog A28 GCPRILMRCKQDSDCLAGCVCGPNGFCaSP
TIC is shown. Calculated monoisotopic mass = 3099.34Da, Observed monoisotopic mass = 3099.37Da

Table A2.15.35. Assigned MFE of EETI-II analog A28 synthesized with the second generation AFPS (Figs A2.15.38 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3099.395	775.8573	6.858	29.89
Cys, Gly double deletion	2939.359	735.8481	6.657	8.42
42Da cap on Pro-3	2981.369	994.7978	7.482	3.42
t-Butylation	3155.194	789.8071	7.27	2.22
Val deletion	3000.327	751.0905	6.666	2.06
Asn deletion	2985.329	996.1115	6.863	1.95
Unknown. -14Da	3085.382	772.3534	6.786	1.87
Unknown truncation	1026.472	514.2442	5.691	1.39
42Da cap on G-22	889.3835	890.3906	5.734	1.25
	2712.126	1357.067	7.965	1.25
	3109.466	778.3748	7.066	1.24
	1435.625	718.8215	7.433	1.17
	3180.388	1061.135	7.993	1.04
	2599.092	1300.551	7.498	0.91
	1091.462	1092.47	6.556	0.84
	1620.771	406.2009	5.156	0.80
	3042.366	761.6007	6.806	0.78
	2995.42	749.8637	7.113	0.77
	621.2607	622.2683	5.293	0.75
	2996.377	750.1027	6.743	0.74

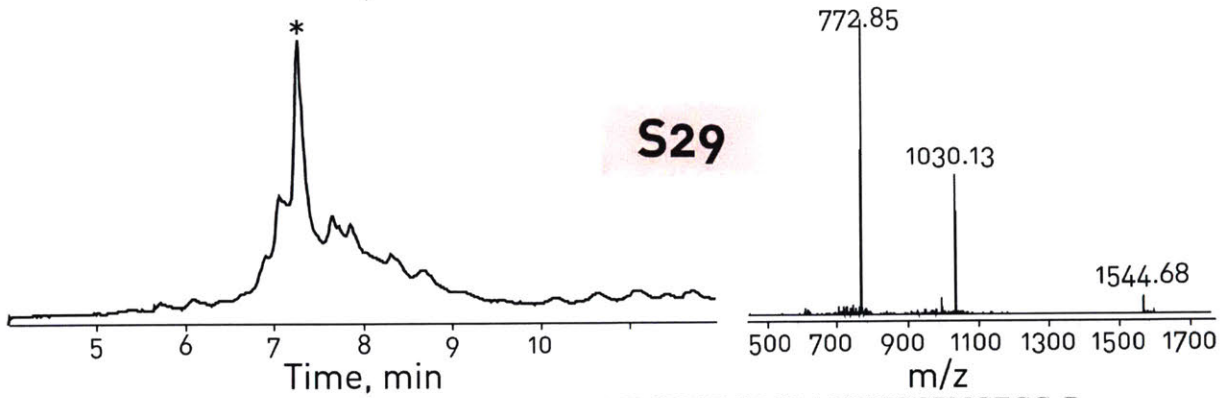


Figure A2.15.39: EETI-II Analog S29 GCPRILMRCKQDSDCLAGCVCGPNGFCGsP
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.36Da

Table A2.15.36. Assigned MFE of EETI-II analog S29 synthesized with the second generation AFPS (Figs A2.15.39 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.369	772.3518	6.821	24.83
Cys, Gly double deletion	2925.34	732.3433	6.62	7.45
Ser	2998.311	750.5865	6.882	7.25
42Da cap on Pro-3	2967.35	990.1235	7.416	3.18
t-Butylation	3141.436	786.3669	7.203	2.40
Unknown truncation	2838.301	710.5839	6.687	2.09
Unknown (-14Da)	3071.321	768.8381	6.748	1.60
Cys deletion	2982.36	746.6003	6.745	1.57
Asn deletion	2971.283	991.4309	6.824	1.48
Val deletion	2986.317	747.5877	6.639	1.44
	3095.442	774.8682	6.999	1.16
	3028.325	758.0927	6.863	1.14
	2838.289	710.5788	6.849	1.06
	2981.37	746.3503	7.036	1.06
	2698.141	900.3897	7.877	1.03
	2880.315	961.1121	7.483	0.96
	1620.774	406.2018	5.17	0.92
	875.362	876.3687	5.51	0.89
	1012.449	507.233	5.482	0.86
	1421.606	711.8118	7.312	0.85

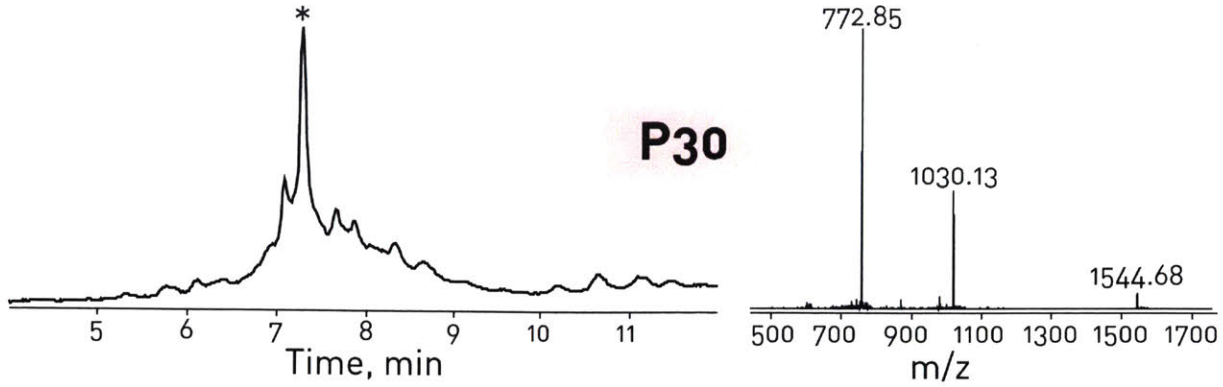


Figure A2.15.40: EETI-II Analog P30 GCPRILMRCKQDSDCLAGCVCGPNGFCGSp
TIC is shown. Calculated monoisotopic mass = 3085.32Da, Observed monoisotopic mass = 3085.36Da

Table A2.15.37. Assigned MFE of EETI-II analog P30 synthesized with the second generation AFPS (Figs A2.15.40 and 3.4.2). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3085.317	772.3378	6.799	29.42
Cys, Gly double deletion	2925.324	732.3378	6.59	8.28
42Da cap on Pro-3	2967.339	990.1203	7.423	3.60
t-Butylation	3141.432	786.3659	7.204	2.98
Val deletion	2986.302	747.5853	6.612	2.87
Unknown (-14Da)	3071.332	768.8409	6.722	1.79
Gly deletion	3028.308	758.0847	6.824	1.72
Unknown truncation	1012.445	507.2311	5.483	1.61
Cys deletion	2982.369	746.6031	6.701	1.50
26Da cap on Ile-5	2698.13	900.385	7.917	1.28
	3095.384	774.8536	6.988	1.19
	875.3595	876.3665	5.518	1.18
	1421.599	711.8084	7.308	1.09
	1620.759	406.1974	5.149	1.04
	2585.062	1293.536	7.423	1.02
	3166.393	1056.471	7.961	0.99
	3067.321	767.8374	6.838	0.92
	2851.249	951.4243	7.283	0.91
	2981.299	746.3312	7.04	0.80
	2971.247	991.4228	6.806	0.75

A2.15.3 Mixed chirality peptides:

The following figures show expanded views of the TICs presented in Figure 14, as well as the mass spectral data from the indicated peak apex.

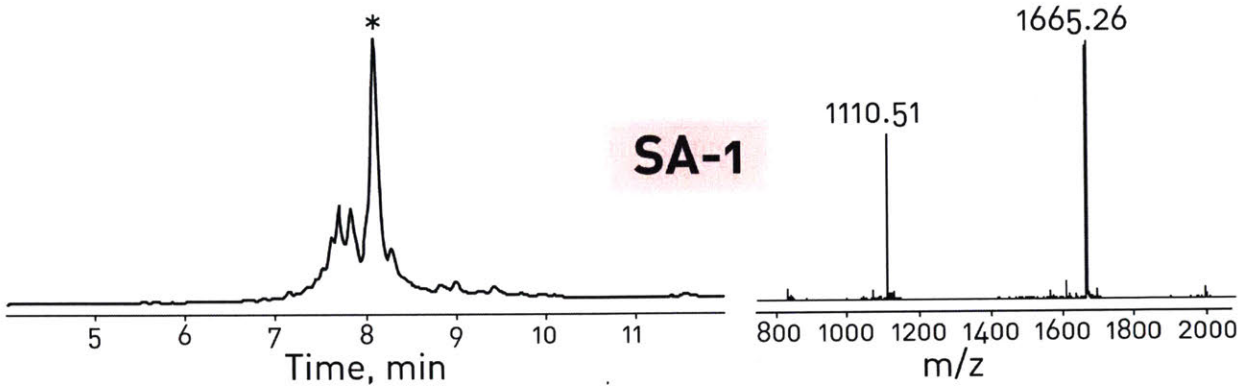


Figure A2.15.41: SA-1 GPmsdvdehGftmEFYVmelqpslglqdvG
TIC is shown. Calculated monoisotopic mass = 3326.48Da, Observed monoisotopic mass = 3326.50Da

Table A2.15.38. Assigned MFE of mixed chirality peptide SA-1 synthesized with the third generation AFPS (Figs A2.15.41 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3326.506	1664.261	8.088	28.97
t-Bu adduct	3382.56	1128.527	7.695	8.97
t-Bu adduct	3382.558	1128.527	7.835	6.43
Asp deletion	3211.466	1606.738	8.114	3.12
Ser Deletion	3239.462	1620.736	8.145	3.04
double Gly	3383.563	1128.862	7.613	1.85
Truncation after M-13	2053.988	1028.001	7.894	1.82
Unknown, possibly -Gly, +tBu	3325.54	1109.521	7.83	1.73
Unknown truncate	2233.084	1117.547	8.286	1.72
Truncation after G-10	2359.126	1180.569	8.303	1.58
	1525.784	763.9002	7.529	1.27
	3438.628	860.6648	7.44	1.08
	3195.448	1598.731	8.009	1.08
	3267.524	1090.182	7.691	1.04
	2277.067	1139.54	8.999	1.04
	3108.416	1555.213	8.027	1.03
	3364.451	1122.49	8.092	1.03
	3227.426	1614.716	8.045	1.02
	2233.08	1117.546	8.839	0.99
	3308.481	1103.836	8.163	0.96

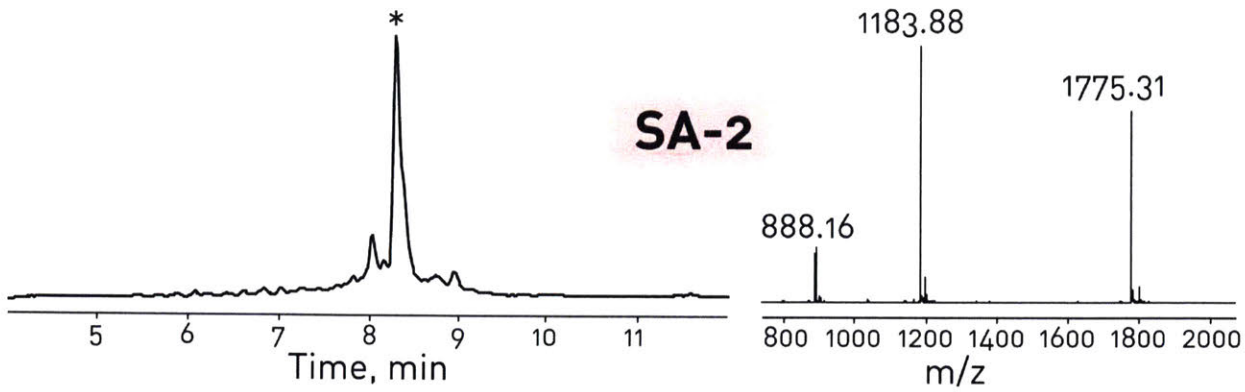


Figure A2.15.42: SA-2 GPlewtlapaftdEFYVdahydtwymqpqG
TIC is shown. Calculated monoisotopic mass = 3546.58Da, Observed monoisotopic mass = 3546.59Da

Table A2.15.39. Assigned MFE of mixed chirality peptide SA-2 synthesized with the third generation AFPS (Figs A2.15.42 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3546.602	1183.208	8.271	41.05
Carboxylation of Trp (converts to product)	3590.581	1197.868	8.365	7.66
t-butylation	3602.654	1201.891	8.008	5.12
Aspartimide	3528.579	1177.2	8.347	3.97
Isomer. Likely aspartimide derived	3546.6	1183.205	8.003	2.98
Unknown	3569.596	1198.202	8.259	2.22
Unknown	3643.051	1215.356	8.923	2.06
Met oxidation	3562.585	1188.534	8.135	1.75
Unknown +14Da	3560.623	1202.54	8.283	1.67
Unknown 2x +14Da	3584.545	1195.854	8.275	1.66
	3546.589	1183.202	7.796	1.28
	3590.57	1796.292	8.258	1.14
	3602.649	1201.889	8.746	1.09
	2442.029	815.0191	6.988	0.85
	3642.572	1215.196	8.969	0.84
	2486.015	1244.012	7.746	0.80
	3600.519	901.1371	8.275	0.80
	3418.538	1140.521	8.348	0.76
	3634.565	1212.53	8.333	0.70
	3642.57	1215.199	8.706	0.70

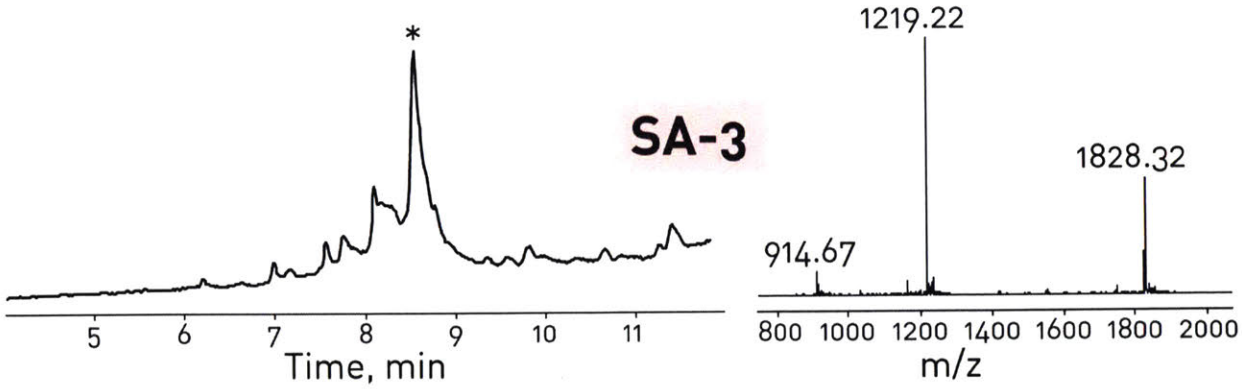


Figure A2.15.43: SA-3 GpfewawtdymqyEFYVyelsdydalapqG
TIC is shown. Calculated monoisotopic mass = 3652.57Da, Observed monoisotopic mass = 3652.62Da

Table A2.15.40. Assigned MFE of mixed chirality peptide SA-3 synthesized with the third generation AFPS (Figs A2.15.43 and 3.5.6). All identified compounds are shown, and the ten most abundant are annotated. It is likely that minor components were too low intensity to pass the 5000count filter and would be identified if more material was injected.

Putative Identity	Mass	m/z	RT	Vol %
Product	3652.639	1218.554	8.677	50.54
t-Butylation	3708.694	1237.24	8.251	11.72
Gln deletion	3524.584	1175.871	8.823	5.99
42Da cap on Phe-13	1890.901	946.4589	7.895	5.05
Aspartimide	3634.62	1212.55	8.76	4.89
Unknown	3360.518	1121.183	8.742	4.73
42Da cap on Glu-14	2019.939	1010.977	7.716	3.84
Unknown	3706.546	1236.523	8.679	3.42
Tyr deletion	3489.576	1164.202	8.658	3.34
42Da cap on V-17	1580.763	791.3899	7.129	3.08
	3525.576	1763.795	8.806	2.12
	1318.626	660.322	6.348	1.29

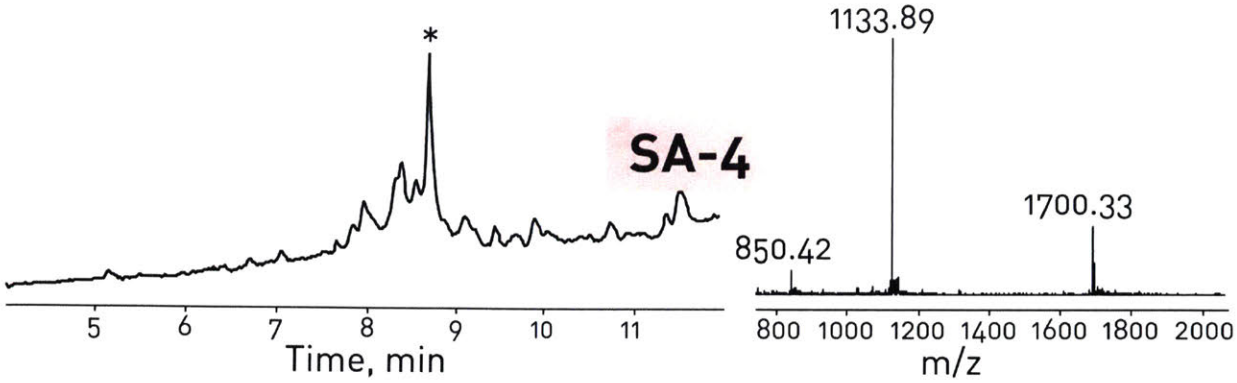


Figure A2.15.44: SA-4 GpfGdvfqlqdqtEFYVtapalydtlywGG
TIC is shown. Calculated monoisotopic mass = 3396.57Da, Observed monoisotopic mass = 3396.63Da

Table A2.15.41. Assigned MFE of mixed chirality peptide SA-4 synthesized with the third generation AFPS (Figs A2.15.44 and 3.5.6). All identified compounds are shown and annotated. It is likely that minor components were too low intensity to pass the 5000 count filter and would be identified if more material was injected.

Putative Identity	Mass	m/z	RT	Vol %
Product	3396.655	1133.226	8.734	45.78
Isomer. Possibly carboxylation of trp fragementing in source	3396.652	1133.225	8.403	21.00
122Da cap on Thr-25	816.3913	817.3978	7.884	8.31
122Da cap on Thr-18	1547.751	774.884	8.596	6.70
122Da cap on Thr-13	2187.053	1094.533	9.154	4.38
Unknown	772.3997	773.4065	7.723	3.24
42Da cap on Val-15	1566.796	784.4074	8.006	3.11
Truncation at Leu-5	593.3027	594.31	5.204	2.78
Truncation at Ala-13	1324.662	663.3397	7.095	2.58
Unknown	772.3985	773.4057	6.753	2.12

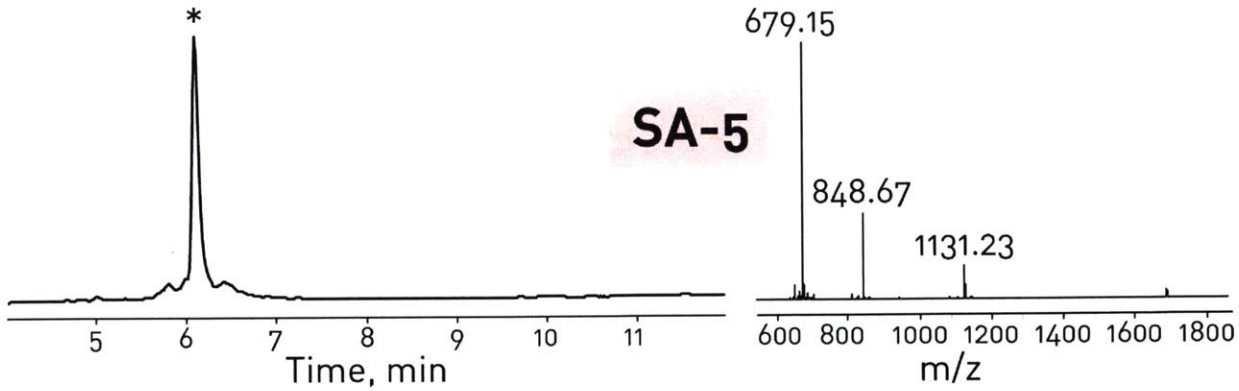


Figure A2.15.45: SA-5 GpmqhawqpahqGEFYVGqhelvpvhepqG
TIC is shown. Calculated monoisotopic mass = 3388.58Da, Observed monoisotopic mass = 3388.64Da

Table A2.15.42. Assigned MFE of mixed chirality peptide SA-5 synthesized with the third generation AFPS (Figs A2.15.45 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3388.695	678.7485	6.118	64.14
Gln deletion	3260.618	653.132	6.161	5.97
His deletion	3251.605	813.9088	6.444	1.67
Gly deletion	3330.651	667.1374	6.13	1.66
Truncation at Ala-6	2838.424	710.6141	6.156	1.62
Unknown +14Da	3402.649	681.5368	6.133	1.59
Double glycine	3445.686	690.1445	6.118	1.32
t-butylation	3444.732	689.954	5.833	0.98
Truncation at Gln-18	2148.084	717.0354	6.017	0.82
Isomer. Possibly carboxylation of trp fragementing in source	3388.689	678.745	5.83	0.81
	1102.598	552.3066	5.039	0.78
	3404.678	681.9445	5.997	0.74
	3131.575	627.3232	6.189	0.71
	3331.641	833.9182	6.133	0.69
	3459.699	692.9475	6.13	0.65
	3486.73	698.3533	6.188	0.62
	3292.617	659.5313	6.049	0.6
	1565.825	783.9191	5.69	0.59
	3516.723	704.3531	6.102	0.54
	3163.56	633.7203	6.076	0.53

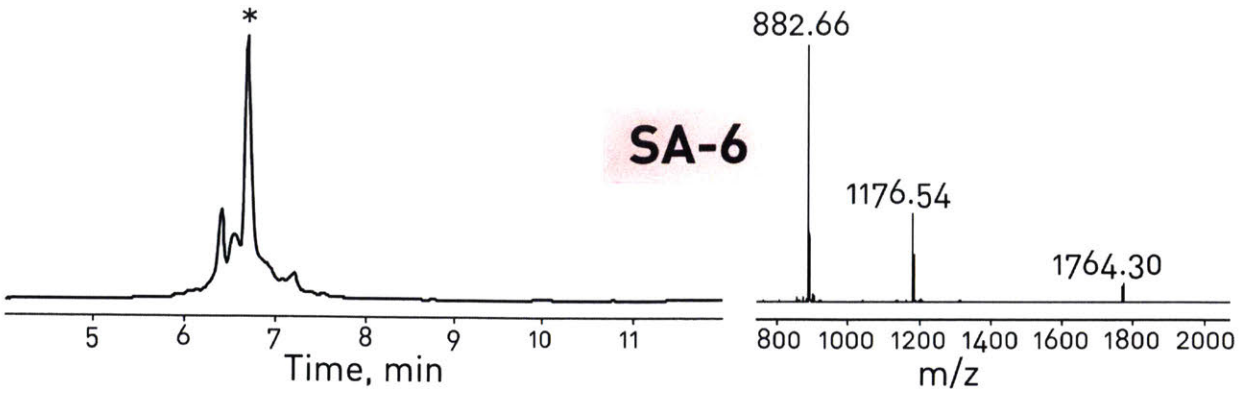


Figure A2.15.46: SA-6 GpfyptkydthyeEFYVeqpGmemGhGpyG
TIC is shown. Calculated monoisotopic mass = 3524.50Da, Observed monoisotopic mass = 3524.57Da

Table A2.15.43. Assigned MFE of mixed chirality peptide SA-6 synthesized with the third generation AFPS (Figs A2.15.46 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Misassignment of product ions.	3525.701	882.4326	6.69	30.59
Product	3524.621	882.1648	6.691	15.19
t-Butylation	3580.659	717.1405	6.402	9.75
Oxidation of Met	3540.584	886.1538	6.526	2.91
Isomer. Unknown source	3524.596	705.9274	6.55	2.73
Isomer. Unknown source	3524.611	705.9295	6.403	2.00
Tyr deletion	3361.519	841.3878	6.575	1.91
Gly deletion	3467.572	867.9009	6.731	1.57
122Da cap on Thr-10	2578.111	1290.06	7.169	1.49
122Da cap on Thr-6	3085.355	1029.459	6.937	1.45
t-butylation	3580.657	896.1724	7.219	1.39
Unknown +14Da	3538.555	885.6463	6.71	1.04
	3581.603	896.4084	6.67	0.91
	2217.956	1109.984	6.593	0.84
	3102.377	1035.134	6.619	0.84
	3621.639	906.4183	6.75	0.82
	3532.546	884.1438	6.695	0.75
	3395.543	849.8933	6.705	0.72
	2355.021	786.0157	6.357	0.66
	1528.671	765.3442	6.037	0.59

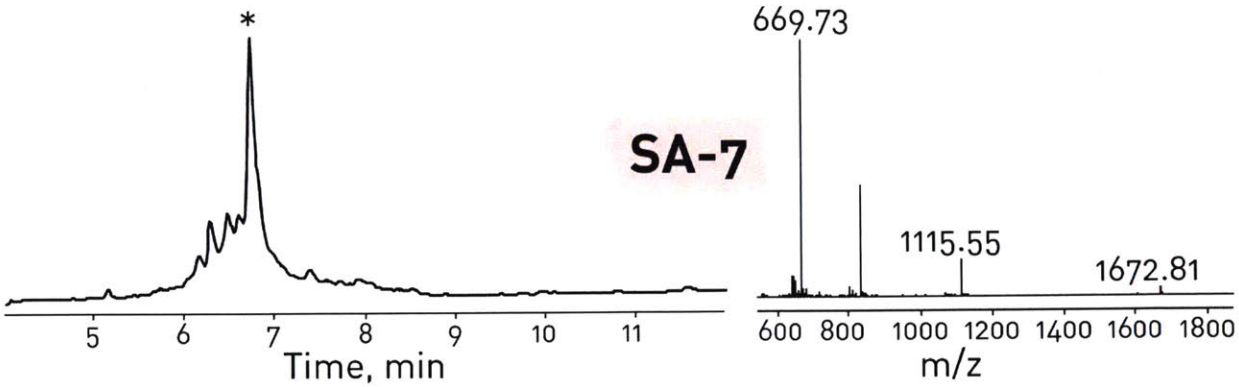


Figure A2.15.47: SA-7 GphtlvdvdGmsvEFYVvskehymqhGpqG
TIC is shown. Calculated monoisotopic mass = 3341.55Da, Observed monoisotopic mass = 3341.60Da

Table A2.15.44. Assigned MFE of mixed chirality peptide SA-7 synthesized with the third generation AFPS (Figs A2.15.47 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3341.649	669.3383	6.71	32.85
Ser deletion	3254.608	651.93	6.764	7.99
Aspartimide	3323.635	665.7355	6.82	5.04
Val deletion	3242.58	649.5251	6.586	4.18
Ala, Phe double deletion	3123.57	625.7222	6.53	3.52
t-Butylation	3397.72	680.5521	6.282	3.1
Gln deletion	3213.585	643.7257	6.741	2.63
Truncation at Leu-9	2949.445	738.3689	6.81	2.48
Val deletion	3242.577	649.5232	6.467	1.98
Isomer. Likely from aspartimide	3341.649	669.3371	6.424	1.58
	2175.072	726.0321	6.47	1.44
	3123.461	1042.161	7.375	1.42
	3024.498	605.908	6.257	1.39
	3240.588	649.1256	6.679	1.37
	2201.09	734.7043	6.793	1.31
	2075.999	693.0075	6.142	1.29
	3155.541	632.1159	6.591	1.21
	3398.661	680.741	6.687	1.21
	3050.503	763.6331	6.896	1.14
	3341.655	669.3381	6.281	1.01

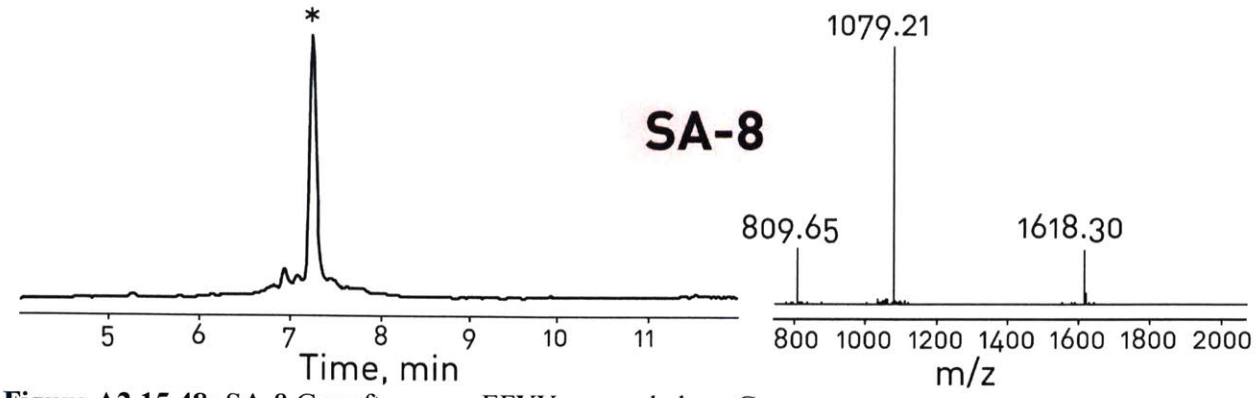


Figure A2.15.48: SA-8 GppqftwspspvaEFYVaapqmahahepqG
TIC is shown. Calculated monoisotopic mass = 3232.51Da, Observed monoisotopic mass = 3232.58Da

Table A2.15.45. Assigned MFE of mixed chirality peptide SA-8 synthesized with the third generation AFPS (Figs A2.15.48 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3232.605	1078.543	7.242	51.07
Probable missasignement of product ions	3232.295	1078.439	7.298	13.1
Probable missasignement of product ions	3232.6	809.1578	6.951	3.97
Gln deletion	3104.245	1035.755	7.308	2.97
Gln deletion	3104.53	1035.851	7.267	1.59
Met oxidation	3248.584	1083.869	7.104	1.45
t-Butylation	3288.646	823.1698	6.949	1.44
Ala deletion	3161.549	1054.857	7.232	1.06
Unknown	3285.519	1096.177	7.24	1.03
42Da cap on V-17	1483.732	742.8726	5.285	0.85
Gly deletion	3175.566	1059.529	7.249	0.85
	3095.516	1032.848	7.471	0.82
	3246.545	1083.188	7.252	0.72
	2828.35	943.7917	7.778	0.71
	3270.533	818.6409	7.241	0.7
	3085.51	1029.509	6.85	0.69
	3329.629	1110.885	7.246	0.61
	3233.587	1078.869	7.412	0.49
	3303.608	1102.209	7.248	0.48
	3254.558	1085.862	7.24	0.47

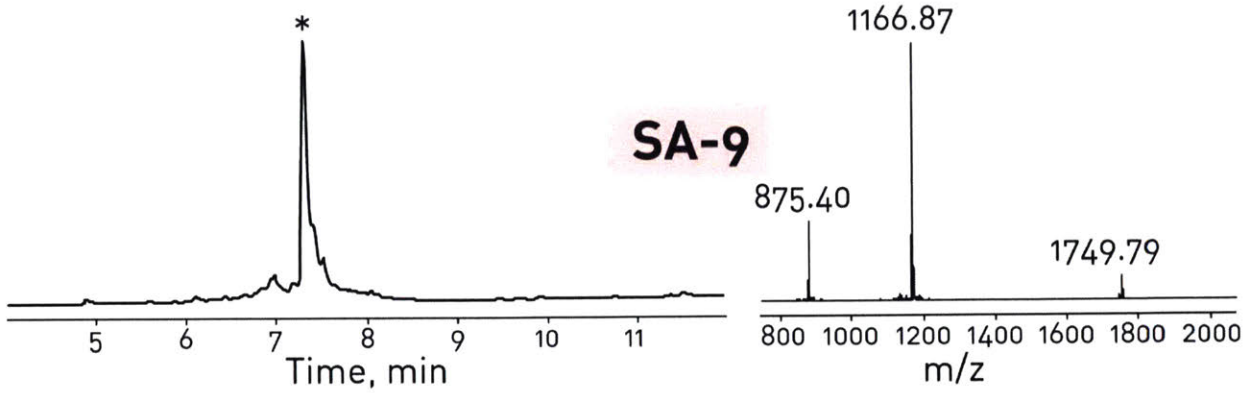


Figure A2.15.49: SA-9 GpdGwqltdtwyqEFYVqafqhGdqdahtG
TIC is shown. Calculated monoisotopic mass = 3495.51Da, Observed monoisotopic mass = 3495.56Da

Table A2.15.46. Assigned MFE of mixed chirality peptide SA-9 synthesized with the third generation AFPS (Figs A2.15.49 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3495.598	1166.206	7.35	45.55
Aspartimide	3477.578	1160.2	7.343	10
Aspartimide	3477.576	1160.198	7.449	7.86
Isomer. Probably aspartimdie derived	3495.595	874.9066	7.005	3.99
Unknown. 122Da cap on Thr-10	2648.189	883.7388	7.561	3.52
Unknown	3459.558	1154.192	7.446	1.96
Unknown truncation	2864.264	955.7635	7.467	1.92
Unknown	3501.503	1168.175	7.351	1.61
Thr deletion	3394.543	1132.522	7.353	1.3
Unknown	3548.518	1183.852	7.353	1.21
	3438.571	1147.197	7.358	1.21
	3477.59	870.4049	7.014	1.17
	3509.562	1170.858	7.352	1.11
	3533.541	884.3937	7.355	1.03
	3366.527	1123.183	7.441	1.02
	2604.203	869.0745	6.863	0.94
	3380.552	1127.858	7.349	0.82
	2630.17	877.7324	7.558	0.82
	2715.27	906.0987	8.089	0.73
	3551.646	1184.887	7.951	0.69

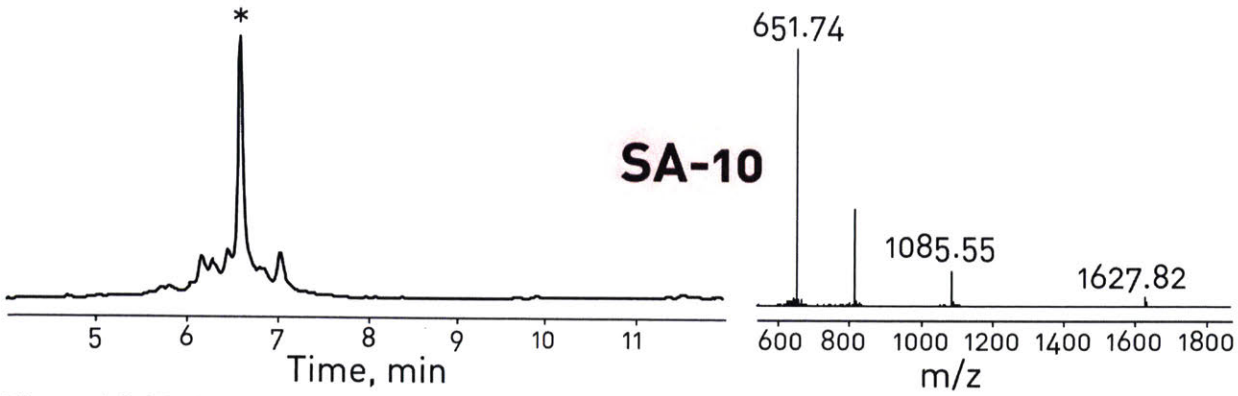


Figure A2.15.50: SA-10 GpwamGpsfvpakEFYVktmGkadyhGpqG
TIC is shown. Calculated monoisotopic mass = 3251.56Da, Observed monoisotopic mass = 3251.61Da

Table A2.15.47. Assigned MFE of mixed chirality peptide SA-10 synthesized with the third generation AFPS (Figs A2.15.50 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3251.658	651.34	6.611	48.52
Lysine deletion	3123.546	781.8944	7.057	5.31
Val deletion	3152.582	631.525	6.486	4.93
Isomer. Possibly carboxylation of trp fragementing in source	3251.661	651.3407	6.207	3.03
Lysine deletion	3123.589	625.7257	6.651	2.74
t-Butylation	3307.72	662.5503	6.198	2.69
Ala deletion	3180.606	637.1294	6.687	2.69
t-Butylation	3307.717	662.5502	6.312	2.37
Isomer. Possibly carboxylation of trp fragementing in source	3251.654	651.3385	6.33	1.89
Glycine deletion	3194.635	639.9345	6.608	1.81
	3265.659	654.1396	6.625	1.56
	3024.474	757.1264	6.899	1.21
	1381.639	691.8259	5.761	1.05
	3308.67	662.7412	6.599	1.03
	3233.645	647.7368	6.618	0.83
	1967.968	656.9969	5.849	0.72
	3305.573	827.3998	6.61	0.64
	3267.65	654.54	6.489	0.59
	3081.541	617.3168	6.539	0.56
	3179.608	636.9288	6.593	0.56

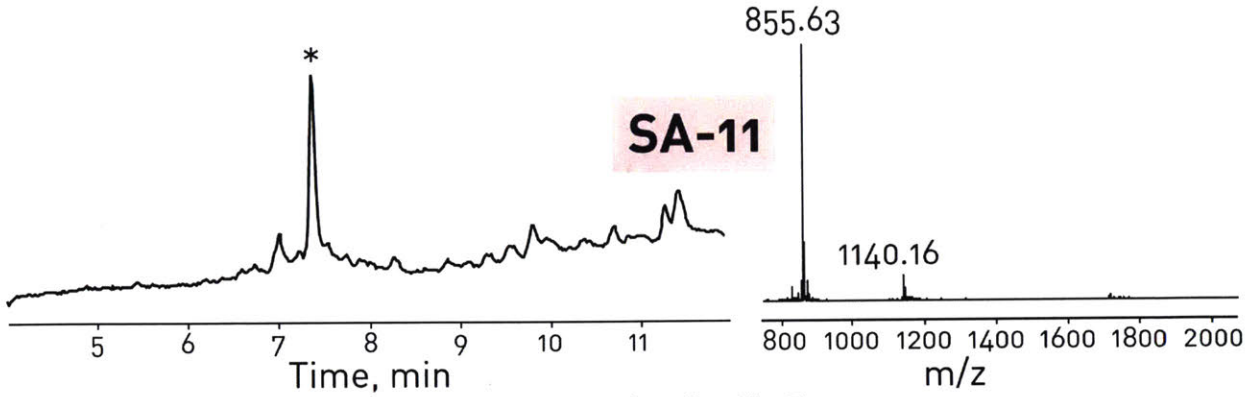


Figure A2.15.51: SA-11 GpmGwsfadawGtEFYVtqhepyhepGhqG
TIC is shown. Calculated monoisotopic mass = 3416.46Da, Observed monoisotopic mass = 3416.46Da

Table A2.15.48. Assigned MFE of mixed chirality peptide SA-11 synthesized with the third generation AFPS (Figs A2.15.51 and 3.5.6). All identified compounds are shown and annotated. It is likely that minor components were too low intensity to pass the 5000 count filter and would be identified if more material was injected.

Putative Identity	Mass	m/z	RT	Vol %
Product	3416.495	855.1318	7.432	73.46
Isomer. Likely aspartimide related	3416.497	684.306	7.072	15.19
Aspartimide	3398.517	850.6346	7.432	11.35

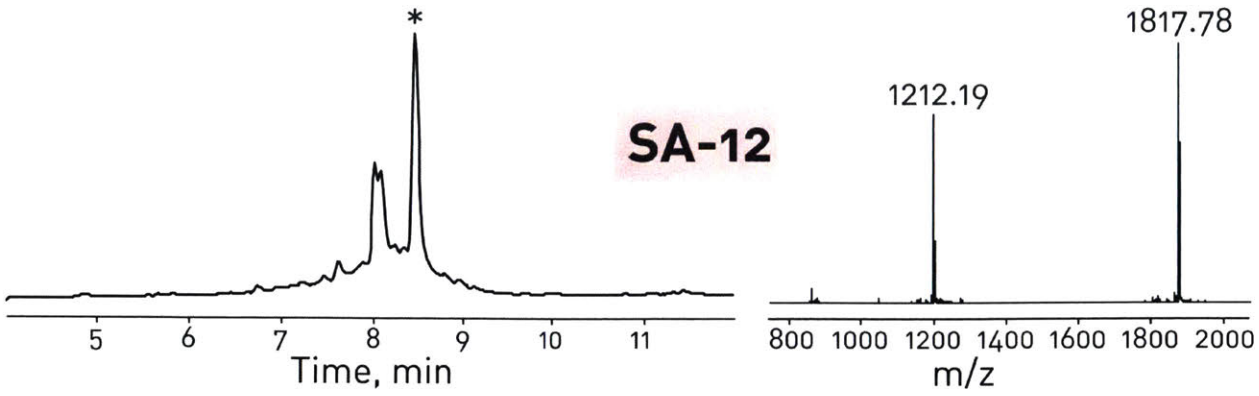


Figure A2.15.52: SA-12 GpwydeltlypemEFYVmGmypehsftdeG
TIC is shown. Calculated monoisotopic mass = 3631.52Da, Observed monoisotopic mass = 3631.55Da

Table A2.15.49. Assigned MFE of mixed chirality peptide SA-12 synthesized with the third generation AFPS (Figs A2.15.52 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3631.555	1816.786	8.468	34.69
t-Butylation	3687.606	1230.209	8.051	22.72
Isomer. Likely aspartimide derived	3631.545	1211.522	8.112	4.08
Aspartimide	3613.523	1807.766	8.503	3.82
Unknown. 2x +14Da	3669.582	1224.203	8.071	2.55
2x t-Butylation	3743.664	936.9241	7.62	1.43
Glu deletion	3502.49	1752.25	8.503	1.25
Met oxidation	3647.532	1216.851	8.23	1.15
Thr deletion	3530.487	1766.248	8.49	1.14
t-Butylation	3687.609	922.9097	7.619	0.94
	3558.542	1187.188	8.074	0.85
	10896.3	2180.268	8.463	0.79
	3687.589	1230.205	8.978	0.78
	3574.516	1788.265	8.499	0.76
	3688.509	1230.509	8.444	0.71
	3516.501	1759.257	8.499	0.71
	3685.466	922.3757	8.469	0.69
	3743.664	936.9236	7.462	0.67
	2435.978	1218.993	7.894	0.66
	3387.467	1694.735	8.516	0.65

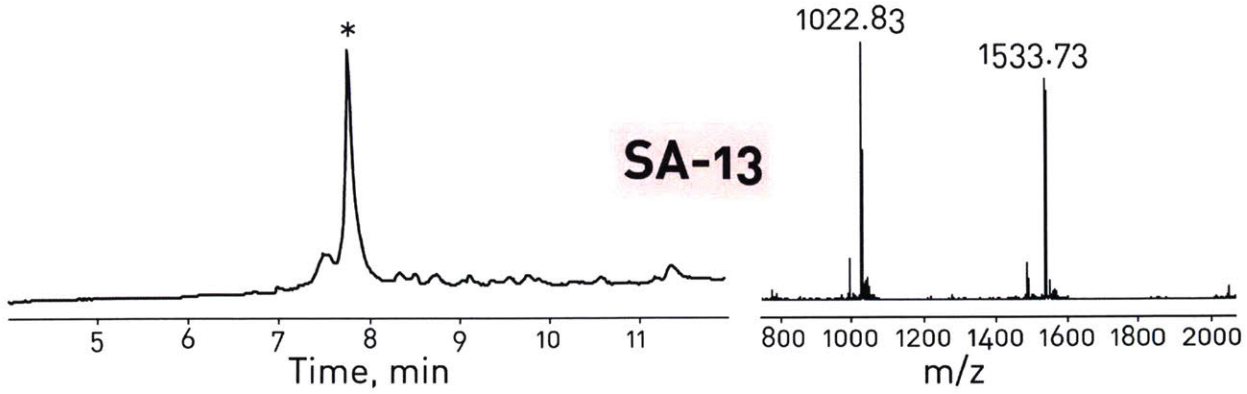


Figure A2.15.53: SA-13 GppaptdalslsaEFYVaqlapGwydsdaG
TIC is shown. Calculated monoisotopic mass = 3064.43Da, Observed monoisotopic mass = 3064.45Da

Table A2.15.50. Assigned MFE of mixed chirality peptide SA-13 synthesized with the third generation AFPS (Figs A2.15.53 and 3.5.6). All identified compounds are shown and annotated. It is likely that minor components were too low intensity to pass the 5000 count filter and would be identified if more material was injected.

Putative Identity	Mass	m/z	RT	Vol %
Product	3064.455	1022.495	7.811	63.94
Thr deletion	2963.405	988.8118	7.799	9.28
Unknown	3117.373	1040.13	7.81	6.46
Aspartimide	3046.44	1016.489	7.88	5.41
Isomer. Likely aspartimide derived	3064.456	1022.494	7.603	4.45
Unknown. 2x +14Da	3102.4	1035.14	7.815	3.66
Asp deletion	2949.43	984.154	7.844	2.97
Truncation at Asp-7	2544.19	1273.099	7.737	2.84
42Da cap on Val-17	1489.685	745.8513	7.043	0.99

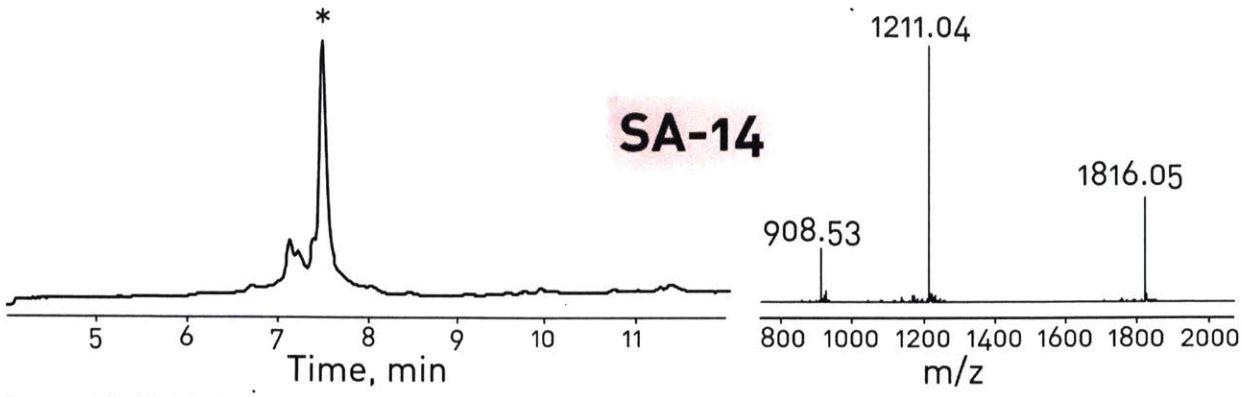


Figure A2.15.54: SA-14 GpmadqmqpvmyqEFYVqydafyhepypeG
TIC is shown. Calculated monoisotopic mass = 3628.53Da, Observed monoisotopic mass = 3628.09Da

Table A2.15.51. Assigned MFE of mixed chirality peptide SA-14 synthesized with the third generation AFPS (Figs A2.15.54 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3628.107	1210.377	7.437	54.73
t-Butylation	3684.613	1229.211	7.085	7.26
Aspartimide	3610.095	1204.372	7.485	5.92
Glu deletion	3499.077	1167.367	7.478	4.44
t-Butylation	3684.608	1229.21	7.175	3.76
Unknown	3681.004	1228.009	7.438	1.75
Unknown	3683.029	921.7646	7.439	1.73
Unknown truncation	3239.385	1080.803	7.355	1.61
Tyr deletion	3465.485	1156.169	7.343	1.52
Unknown	3635.022	1212.682	7.439	1.51
	3650.068	1217.696	7.443	1.49
	3644.07	917.5194	7.446	1.17
	3666.02	1223.014	7.447	1.08
	3684.615	1229.211	7.996	1.05
	3643.108	1215.378	7.449	1.00
	3402.015	1135.013	7.456	0.78
	3644.531	1215.851	7.227	0.71
	1754.763	878.389	6.656	0.68
	3571.095	1191.372	7.465	0.63
	3725.13	1242.717	7.477	0.58

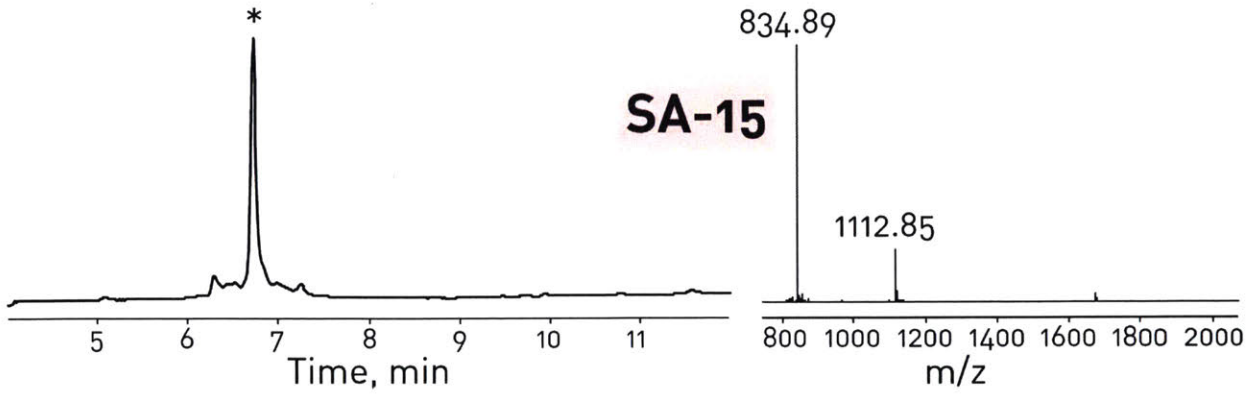


Figure A2.15.55: SA-15 GpwGhqlGmqwesEFYVsshapslGhepqG
TIC is shown. Calculated monoisotopic mass = 3333.49Da, Observed monoisotopic mass = 3333.51Da

Table A2.15.52. Assigned MFE of mixed chirality peptide SA-15 synthesized with the third generation AFPS (Figs A2.15.55 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3333.539	834.3929	6.705	67.93
Gln deletion	3205.476	802.3762	6.806	3.59
t-Butylation	3389.599	678.9274	6.282	3.32
Unknown	3386.46	678.3005	6.708	2.49
Unknown. +14Da	3347.524	837.8882	6.712	2.05
Unknown	3391.542	848.8928	6.698	1.84
Possible Gly deletion, mass 1Da off	3275.519	819.887	6.716	1.75
Ser deletion	3246.498	812.6326	6.73	1.22
Unknown	3381.513	846.3864	6.712	0.96
Isomer	3333.528	667.7139	6.28	0.93
	3389.575	848.4013	7.239	0.92
	2799.289	934.1044	6.621	0.89
	3371.486	675.3056	6.708	0.89
	3349.519	838.3875	6.466	0.88
	3430.573	858.6505	6.735	0.72
	3333.561	667.7195	6.423	0.65
	3404.562	852.148	6.697	0.59
	1442.697	722.3537	5.081	0.58
	3461.564	866.3983	6.674	0.56
	2713.227	905.4171	7.232	0.55

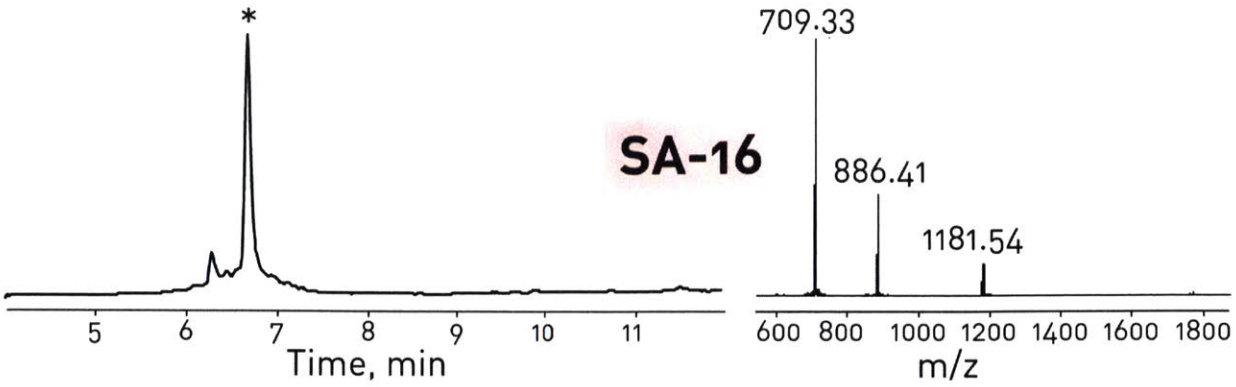


Figure A2.15.56: SA-16 GpfGdvfekqdGkEFYVkedymeladqhyG
TIC is shown. Calculated monoisotopic mass = 3539.59Da, Observed monoisotopic mass = 3539.60Da

Table A2.15.53. Assigned MFE of mixed chirality peptide SA-16 synthesized with the third generation AFPS (Figs A2.15.56 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3539.637	708.9353	6.712	37.05
Aspartimide	3521.618	705.3318	6.71	23.00
Probably misassigned product ions	3540.628	886.1643	6.712	10.63
t-Butylation	3595.701	720.147	6.333	3.66
Glu deletion	3410.594	683.1271	6.788	2.02
Unknown. 2x +14Da	3577.688	716.5448	6.335	1.56
Isomer. Likely from aspartimide	3539.647	590.9492	6.333	1.52
Possibly misassigned product ions	3537.622	885.4143	6.713	1.47
Unknown	3503.615	701.7299	6.735	1.22
Unknown	3592.559	719.5224	6.713	1.21
	3555.63	712.1347	6.717	1.12
	3376.562	676.3211	6.585	0.93
	3606.692	722.3465	6.854	0.93
	3597.631	720.5334	6.709	0.89
	3482.626	697.5323	6.727	0.88
	3555.622	712.1319	6.494	0.83
	3595.68	720.1432	7.187	0.75
	3521.638	587.949	6.335	0.74
	3579.601	716.9295	6.709	0.72
	3393.589	679.7274	6.776	0.66

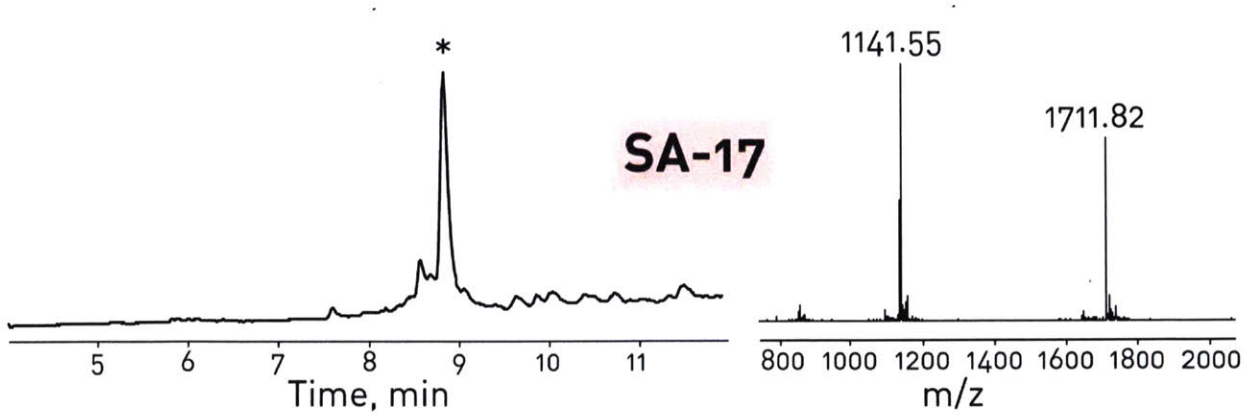


Figure A2.15.57: SA-17 GpwapqdafqlelEFYVledvfeptptpeG
TIC is shown. Calculated monoisotopic mass = 3419.61Da, Observed monoisotopic mass = 3419.62Da

Table A2.15.54. Assigned MFE of mixed chirality peptide SA-17 synthesized with the third generation AFPS (Figs A2.15.57 and 3.5.6). All identified compounds are shown, and the ten most abundant are annotated. It is likely that minor components were too low intensity to pass the 5000count filter and would be identified if more material was injected.

Putative Identity	Mass	m/z	RT	Vol %
Product	3419.631	1140.886	8.876	64.72
Leu deletion	3306.541	1103.189	8.625	7.27
Unknown	3472.552	1158.527	8.875	6.66
Glu deletion	3290.581	1097.867	8.911	5.90
Possibly non-peptidic	716.5	739.4881	10.77	3.19
Isomer. Likely aspartimide derived	3419.629	1140.882	8.77	2.17
Aspartimide	3401.605	1134.877	8.909	1.96
Unknown	3474.536	1738.275	8.875	1.86
Possibly non-peptidic	716.4987	739.4874	10.434	1.73
Unknown	3457.591	1153.538	8.878	1.64
	1569.761	785.8895	7.647	1.40
	698.4899	699.4972	10.43	0.84
	590.3078	591.3154	5.89	0.66

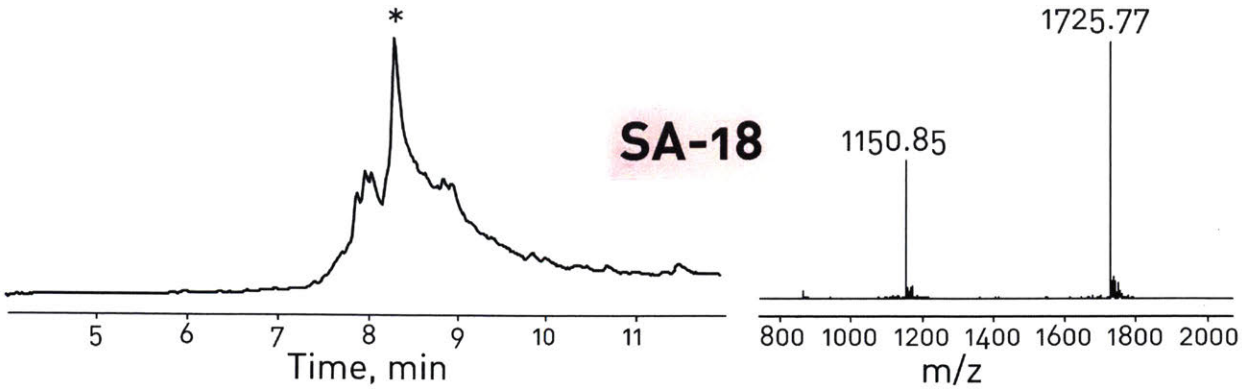


Figure A2.15.58: SA-18 GpftmqftdGpytEFYVtydtlemawGpqG
TIC is shown. Calculated monoisotopic mass = 3447.50Da, Observed monoisotopic mass = 3447.52Da

Table A2.15.55. Assigned MFE of mixed chirality peptide SA-18 synthesized with the third generation AFPS (Figs A2.15.58 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated. Product ions were likely misassigned because the main peak is not approximately Gaussian.

Putative Identity	Mass	m/z	RT	Vol %
Product	3447.519	1724.766	8.36	26.61
Isomer. Probably misasigned product ions	3447.514	1150.181	8.636	10.01
Isomer. Probably misasigned product ions	3447.282	1724.647	8.484	9.89
Isomer. Probably misasigned product ions	3447.516	1150.18	8.109	8
t-Butylation	3503.564	1168.865	8.029	7.16
Triflouroacetylation	3543.501	1772.755	8.993	6.5
Carboxylation of trp	3491.502	1164.844	8.278	5.28
t-Butylation	3503.578	1168.867	7.934	4.19
t-Butylation	3503.547	1168.857	8.865	1.93
Met oxidation	3463.507	1155.51	8.122	1.77
	2227.982	1114.997	8.689	1.43
	6897.887	2300.303	8.366	1.23
	3457.477	1751.732	8.35	1.16
	3268.406	1635.207	8.845	1.02
	3547.561	1183.529	7.883	1
	1209.548	1210.553	7.829	0.94
	3599.551	1200.857	8.62	0.88
	3543.501	1182.176	8.73	0.87
	716.5037	739.4934	10.769	0.86
	2828.266	1415.138	8.724	0.85

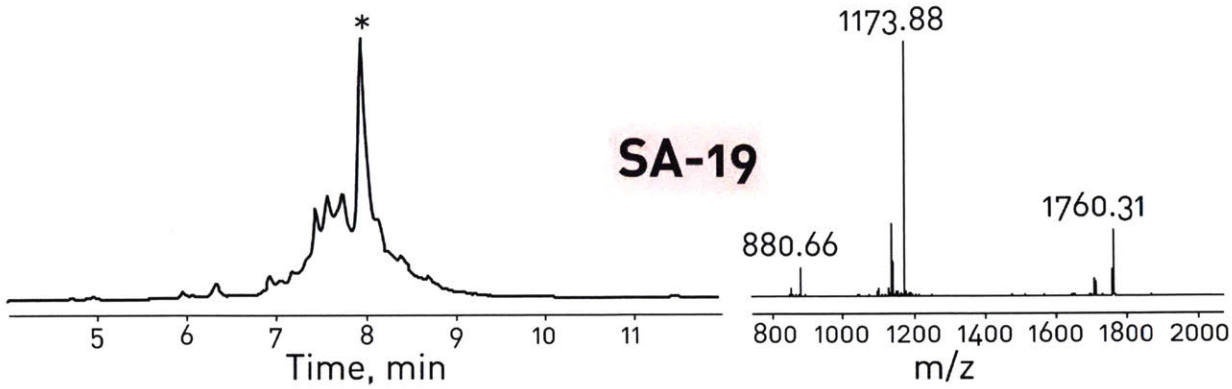


Figure A2.15.59: SA-19 GplGwahvmedtdEFYVdyfvmtkawepqG
TIC is shown. Calculated monoisotopic mass = 3516.57Da, Observed monoisotopic mass = 3516.60Da

Table A2.15.56. Assigned MFE of mixed chirality peptide SA-19 synthesized with the third generation AFPS (Figs A2.15.59 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3516.632	1173.22	7.977	21.72
Thr deletion	3415.568	1139.531	7.956	7.1
t-butylation	3572.671	894.1758	7.466	3.3
Val deletion	3417.541	1140.188	7.78	3.15
t-Butylation	3572.669	894.1759	7.601	3.05
Glu deletion	3387.566	1130.197	8.051	2.92
Isomer. Likely aspartimide derived	3516.613	880.1609	7.609	2.01
Aspartimide	3498.597	1167.206	8.111	1.66
Asp deletion	3401.577	1134.867	8.158	1.63
Unknown	3397.555	1133.526	8.045	1.56
	3300.534	1101.186	8.004	1.49
	3286.518	1096.514	8.022	1.36
	2446.1	1224.057	8.411	1.34
	3516.249	880.0703	7.704	1.32
	1710.817	856.4168	7.556	1.21
	1103.584	552.8012	6.958	1.2
	3171.488	1058.17	8.084	1.19
	3522.529	1175.184	7.978	1.16
	1036.503	1037.509	6.364	1.13
	3280.472	1641.241	8.186	1.07

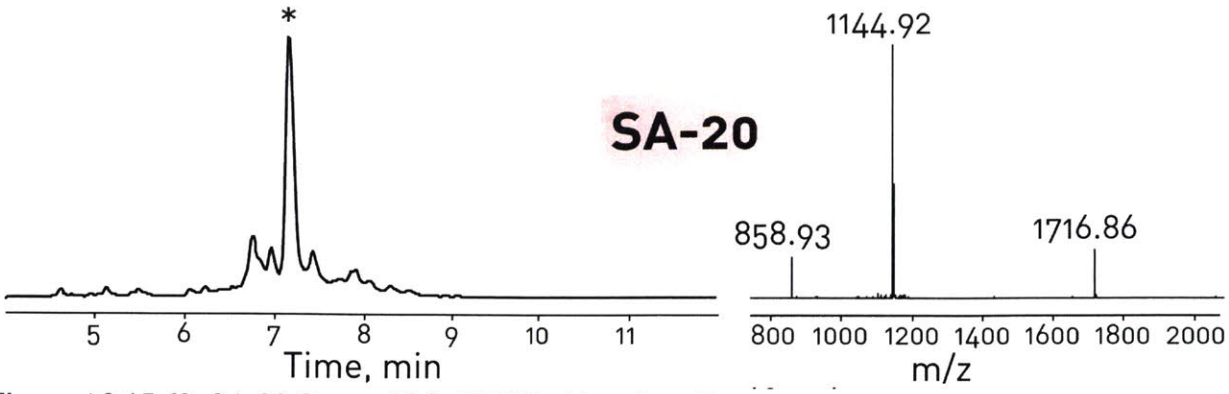


Figure A2.15.60: SA-20 GpptpydtltfyqEFYVqsltkepekapqG
TIC is shown. Calculated monoisotopic mass = 3429.67Da, Observed monoisotopic mass = 3429.70Da

Table A2.15.57. Assigned MFE of mixed chirality peptide SA-20 synthesized with the third generation AFPS (Figs A2.15.60 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3429.751	1144.26	7.188	37.3
Isomer. Unknown source	3429.72	858.4382	6.794	12.58
Lys deletion	3301.663	1101.563	7.235	4.14
Deamidation (in source fragmentation)	3412.678	1145.894	7.188	3.43
Unknown truncation	2610.295	1306.153	7.462	2.84
122Da cap on Thr-8	2824.425	942.4837	7.948	2.3
Truncation at Pro-5	3077.532	1026.851	7.001	1.3
Unknown truncation	2566.303	856.4416	6.718	1.07
Lys deletion	3301.659	826.4228	6.827	1.02
42Da cap on Val-17	1551.847	776.9297	5.171	0.99
	2601.34	868.1211	6.707	0.83
	3526.765	1176.596	7.205	0.83
	3485.776	1162.932	7.854	0.78
	2119.08	1060.544	6.101	0.66
	2780.437	927.8196	7.129	0.65
	3331.657	1111.56	7.145	0.64
	3371.702	1124.908	7.203	0.62
	1289.717	645.8655	5.54	0.61
	3367.693	1123.573	8.332	0.61
	2387.222	796.7481	6.256	0.58

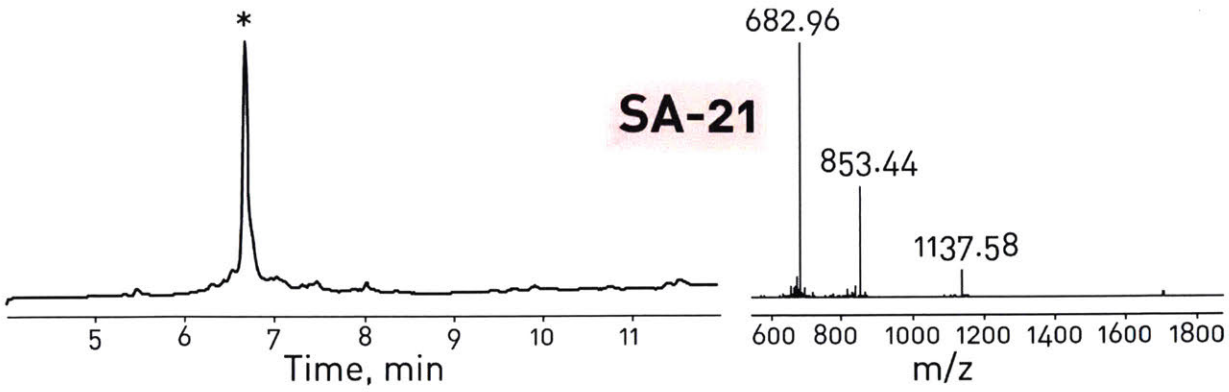


Figure A2.15.61: SA-21 GpsGlykskyltdEFYVdahvpvkewvpqG
TIC is shown. Calculated monoisotopic mass = 3407.71Da, Observed monoisotopic mass = 3407.72Da

Table A2.15.58. Assigned MFE of mixed chirality peptide SA-21 synthesized with the third generation AFPS (Figs A2.15.61 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3407.754	682.5595	6.71	54.96
Gly deletion	3350.734	671.1558	6.721	5.75
Lys deletion	3279.69	656.9455	6.785	5.02
Aspartimide	3389.744	678.9565	6.79	4.71
Gly, Ser double deletion	3263.701	653.7501	6.711	2.61
Unknown	3331.684	833.9284	8.062	2.3
Ser deletion	3320.715	665.1532	6.717	1.91
122Da cap on Thr-12	2336.119	1169.065	7.503	1.75
Truncation at Val-21	1136.642	569.3288	5.52	1.59
Unknown	3465.767	694.1607	6.712	1.59
	3407.765	682.5591	6.566	1.45
	3166.649	634.3382	6.678	1.33
	3151.622	788.9134	7.071	0.98
	3310.701	663.1476	6.681	0.95
	3463.805	693.769	7.155	0.9
	3460.683	693.1425	6.711	0.89
	3464.73	867.1898	6.713	0.7
	1600.849	801.4291	6.347	0.68
	3293.729	659.753	6.751	0.65
	2292.139	765.0565	7.363	0.64

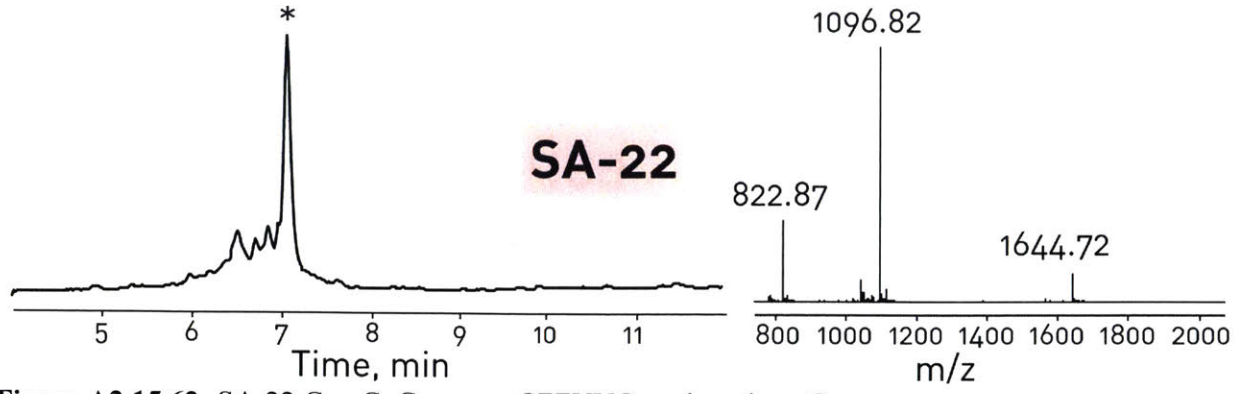


Figure A2.15.62: SA-22 GpwGpGpvmsmeGEFYVGsmqhqmthqpqG
TIC is shown. Calculated monoisotopic mass = 3285.42Da, Observed monoisotopic mass = 3285.42Da

Table A2.15.59. Assigned MFE of mixed chirality peptide SA-22 synthesized with the third generation AFPS (Figs A2.15.62 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3285.448	1096.157	7.09	41.1
Gln deletion	3157.389	1053.47	7.128	4.55
Pro deletion	3188.387	1063.803	7	4.45
t-Butylation	3341.506	836.3845	6.535	4.38
Arg deletion	3131.367	1044.795	7.094	3.7
Gly deletion	3228.415	1077.147	7.135	3.65
Isomer. Unknown source	3285.453	836.3807	6.742	3.53
Val deletion	3186.376	1063.131	6.885	3.02
Unknown	3299.478	836.87	7.072	2.99
Unknown	3343.453	1115.492	7.065	2.01
	3316.395	1113.798	7.088	1.38
	3285.448	822.3696	6.889	1.33
	2791.208	931.4096	6.433	1.1
	3089.314	1030.783	6.781	1.06
	3154.407	1052.475	6.796	1.02
	3212.446	822.1233	6.541	0.86
	3245.452	812.3704	6.464	0.78
	3382.478	1128.502	7.136	0.78
	3186.397	797.6066	6.512	0.71
	3032.294	1011.774	6.889	0.71

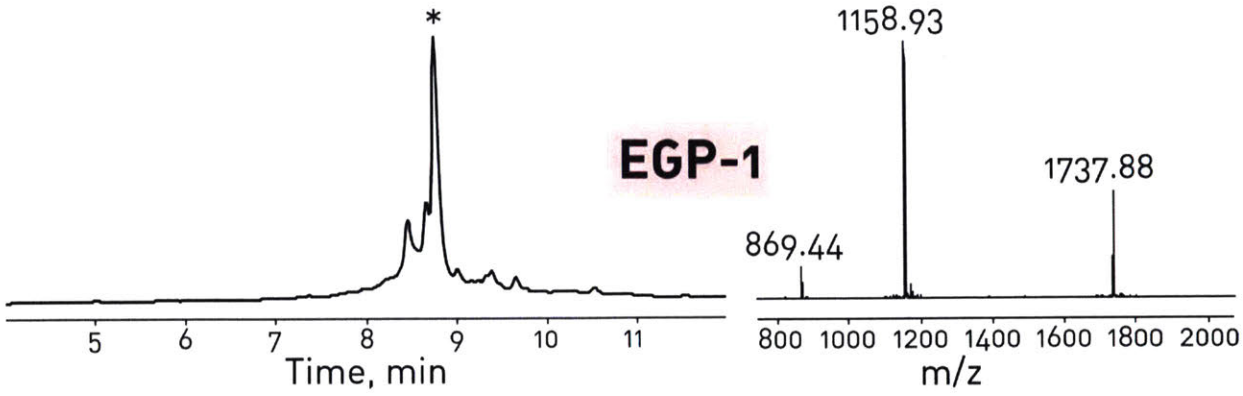


Figure A2.15.63: EGP-1 GpkshtpalvpvwEFYVwqfvpelGmsweG
TIC is shown. Calculated monoisotopic mass = 3471.70Da, Observed monoisotopic mass = 3471.74Da

Table A2.15.60. Assigned MFE of mixed chirality peptide EGP-1 synthesized with the third generation AFPS (Figs A2.15.63 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3471.769	1158.265	8.75	38.83
Carboxylation of Trp (converts to product)	3515.745	1172.922	8.657	8.47
t-Butylation	3527.818	882.9632	8.44	6.85
Isomer. Unknown source	3471.759	868.948	8.466	3.97
Missasignment of product ions	3476.665	1159.896	8.75	3.4
Val deletion	3372.676	1125.232	8.637	2.91
Truncation at Pro-7	2864.432	1433.221	9.38	2.24
98Da cap on Pro-7	2962.516	988.5142	9.643	2.17
Met oxidation	3487.735	1163.587	8.581	1.67
Glu deletion	3342.701	1115.242	8.83	1.49
Trp deletion	3285.662	1096.229	8.505	1.43
	3156.622	1053.216	8.532	1.26
	3516.742	1173.255	8.783	1.02
	3528.746	1177.258	8.722	0.97
	2965.482	1483.744	9.325	0.88
	3484.676	1162.562	8.749	0.73
	3384.698	1129.241	8.831	0.68
	3414.727	1139.25	8.797	0.67
	3527.803	1176.942	9.219	0.64
	3571.801	893.9575	8.373	0.59

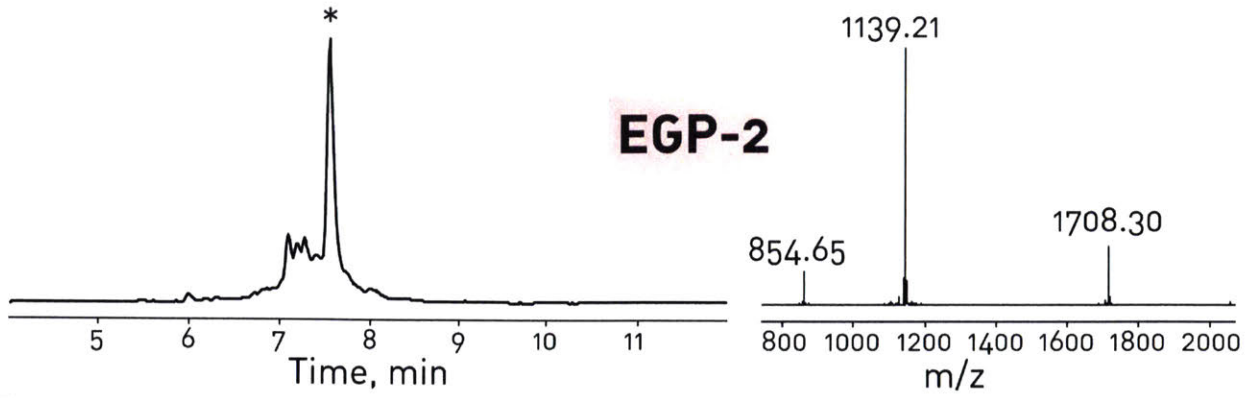


Figure A2.15.64: EGP-2 GpladGlGkymGeEFYVeqdtptfyhymvG
TIC is shown. Calculated monoisotopic mass = 3412.53Da, Observed monoisotopic mass = 3412.57Da

Table A2.15.61. Assigned MFE of mixed chirality peptide EGP-2 synthesized with the third generation AFPS (Figs A2.15.64 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3412.623	1138.549	7.557	38.34
Aspartimide	3394.588	1132.538	7.583	8.43
t-Butylation	3468.657	868.1726	7.107	5.17
Isomer. Likely aspartimide derived	3412.59	854.1566	7.28	4.83
Met oxidation	3428.58	1143.868	7.323	2.24
Probably t-Butylation (-1Da)	3469.662	868.4229	7.19	1.89
Isomer. Likely aspartimide derived	3412.6	854.1575	7.19	1.85
Gly deletion	3355.569	1119.53	7.534	1.7
Aspartimide	3394.579	849.6529	7.267	1.6
Unknown. 2x +14Da	3450.638	863.6681	7.115	1.38
Try deletion	3249.516	1084.179	7.448	1.38
	1112.532	557.2748	6.021	1.25
	2092.999	1047.506	7.401	1.07
	3281.538	1094.854	7.382	0.75
	3468.56	1157.194	7.5	0.72
	3468.631	1157.22	7.19	0.69
	1726.795	864.4083	7.042	0.66
	3315.535	1106.186	7.599	0.63
	3265.51	1089.512	7.249	0.61
	3285.545	1096.189	7.621	0.59

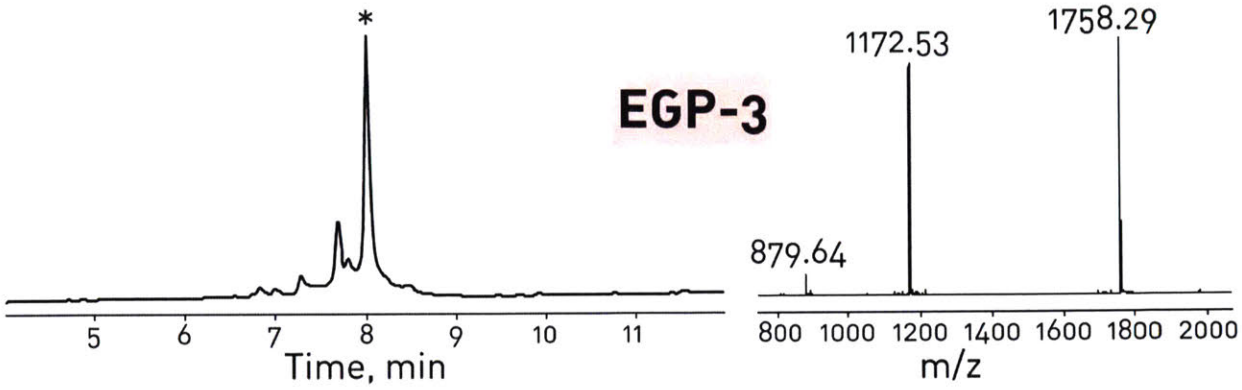


Figure A2.15.65: EGP-3 GpwadswapqlteEFYVeqpefefahsmeG
TIC is shown. Calculated monoisotopic mass = 3512.52Da, Observed monoisotopic mass = 3512.52Da

Table A2.15.62. Assigned MFE of mixed chirality peptide EGP-3 synthesized with the third generation AFPS (Figs A2.15.65 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3512.591	1757.304	8.011	54.85
t-Butylation	3568.628	1190.551	7.694	7.62
Isomer. Unknown source	3512.574	1171.865	7.693	6.37
98Da cap on Phe-15	2042.949	1022.479	7.299	3.27
Met oxidation	3528.56	1177.195	7.808	2.11
Unknown	3518.478	1173.833	8.013	1.35
Missasigned product ions	10539.6	2108.928	8.01	1.22
t-Butylation	3568.622	1190.55	8.498	1.09
Missasigned product ions	7025.059	2342.693	8.008	1
42Da cap on Val-17	1676.732	839.3751	6.84	0.96
	2426.03	1214.02	8.018	0.96
	3526.537	1176.522	8.027	0.79
	3494.546	1165.861	8.052	0.77
	2382.042	1192.025	7.269	0.76
	3565.485	1189.502	8.013	0.75
	3383.511	1692.763	8.043	0.73
	3566.495	892.6311	8.018	0.7
	3455.561	1152.861	8.008	0.62
	2382.041	1192.025	7.824	0.61
	3165.444	1056.158	7.892	0.6

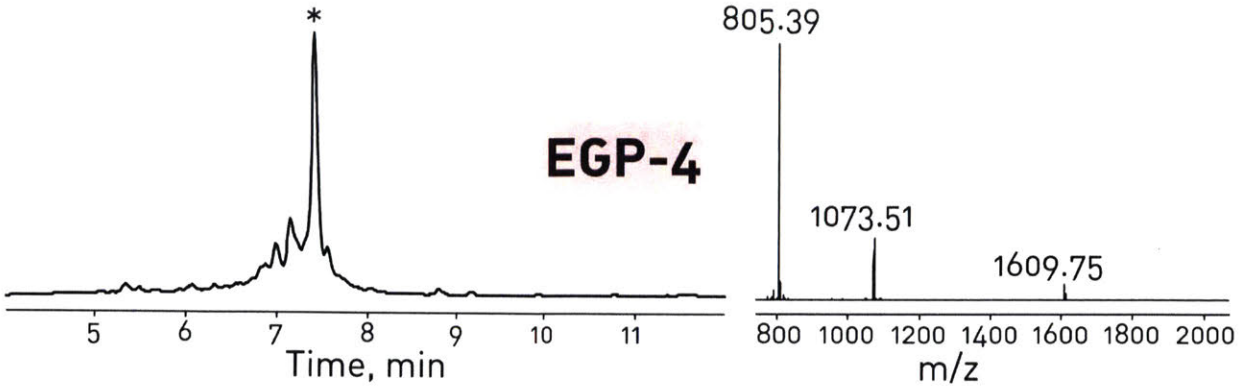


Figure A2.15.66: EGP-4 GplvdawthylGtEFYVtGhGmshepamGG
TIC is shown. Calculated monoisotopic mass = 3215.45Da, Observed monoisotopic mass = 3215.49Da

Table A2.15.63. Assigned MFE of mixed chirality peptide EGP-4 synthesized with the third generation AFPS (Figs A2.15.66 and 3.5.6). The 20 most abundant compounds are shown, and the ten most abundant are annotated.

Putative Identity	Mass	m/z	RT	Vol %
Product	3215.54	804.89442	7.398	44.42
t-Butylation	3271.576	655.32368	7.149	4.71
Isomer. Unknown source.	3215.519	644.11238	6.984	3.19
Met oxidation	3231.51	808.88505	7.301	2.81
Isomer. Unknown source.	3215.521	644.11196	7.149	2.8
Pro deletion	3118.452	780.62075	7.551	2.59
Leu deletion	3102.422	776.61406	7.217	2.4
Gly deletion	3158.49	790.63024	7.414	2.18
Possible Gly deletion (-1Da)	3159.491	790.8799	7.344	1.77
122Da cap on Thr-18	1388.582	695.29897	5.343	1.37
	3061.43	766.36546	7.109	1.31
	3272.52	819.13717	7.47	1.22
	3215.518	644.11146	6.808	0.85
	2239.019	747.34762	6.321	0.79
	1344.592	673.30318	5.481	0.7
	3170.399	1057.80764	8.805	0.7
	3005.37	752.34992	7.017	0.68
	3272.523	819.13801	7.364	0.66
	2948.346	738.09413	6.888	0.65
	2027.875	1014.94365	7.113	0.63

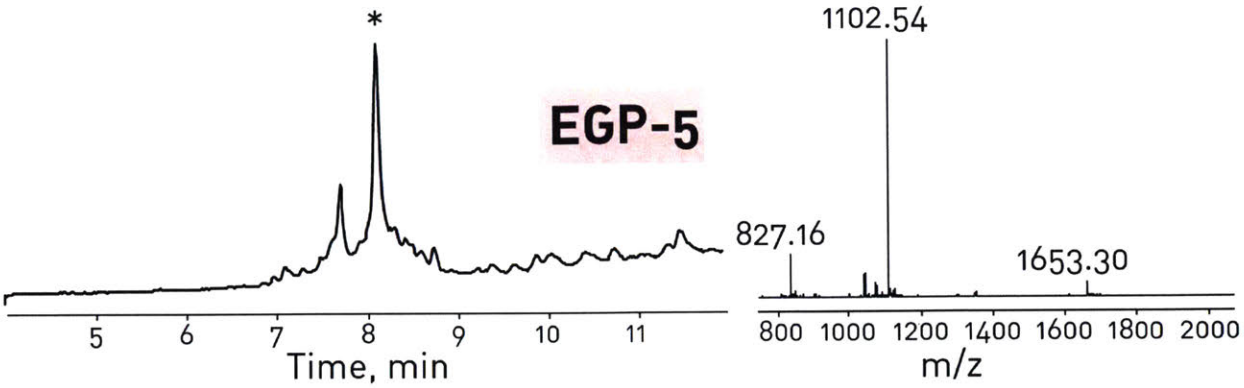


Figure A2.15.67: EGP-5 GplvdawthylGtEFYVtvmepsmakafsG
TIC is shown. Calculated monoisotopic mass = 3302.57Da, Observed monoisotopic mass = 3302.58Da

Table A2.15.64. Assigned MFE of mixed chirality peptide EGP-5 synthesized with the third generation AFPS (Figs A2.15.67 and 3.5.6). All identified compounds are shown, and the ten most abundant are annotated. It is likely that minor components were too low intensity to pass the 5000count filter and would be identified if more material was injected.

Putative Identity	Mass	m/z	RT	Vol %
Product	3302.619	1101.88	8.156	49.56
Isomer. Unknown source.	3302.624	826.6633	7.764	9.99
t-Butylation	3358.686	840.6798	7.751	9.19
Thr deletion	3201.561	1068.194	8.178	4.38
Unknown truncation	2071.006	1036.51	8.14	4.31
Unknown truncation	2200.083	1101.045	8.785	4.03
122Da cap on Thr-13	2114.987	1058.501	8.349	3.32
Unknown truncation	1542.776	772.3967	8.482	2.61
Unknown truncation	1475.691	738.8557	7.676	2.53
Possibly non-peptidic	716.4921	717.497	10.775	2.24
	1891.904	946.9592	7.145	2.23
	2182.061	1092.041	8.672	1.95
	1560.784	781.4009	8.563	1.75
	1494.741	748.3791	7.026	1.05
	3303.646	826.9187	7.645	0.86

A2.15.4 LL37:

LL37 was synthesized using the 800ms pump refill and 70C activation described in the text using Ser loaded ChemMatrix HMPB resin (0.5mmol/g). The resin was prepared by reacting 10eq of Fmoc-Ser-OH, 5eq of diisopropylcarbodiimide (DCC), 1eq of resin, and 0.1eq of diaminopyridine (DMAP) for 15 hours at RT. The reaction was performed on a 1mmol scale using 2g of resin. Fmoc-Ser-OH and DCC were dissolved in 30mL of dry, amine free DMF and allowed to stand for 10 minutes. The resin was then added, and DMAP was added about 2 minutes later.

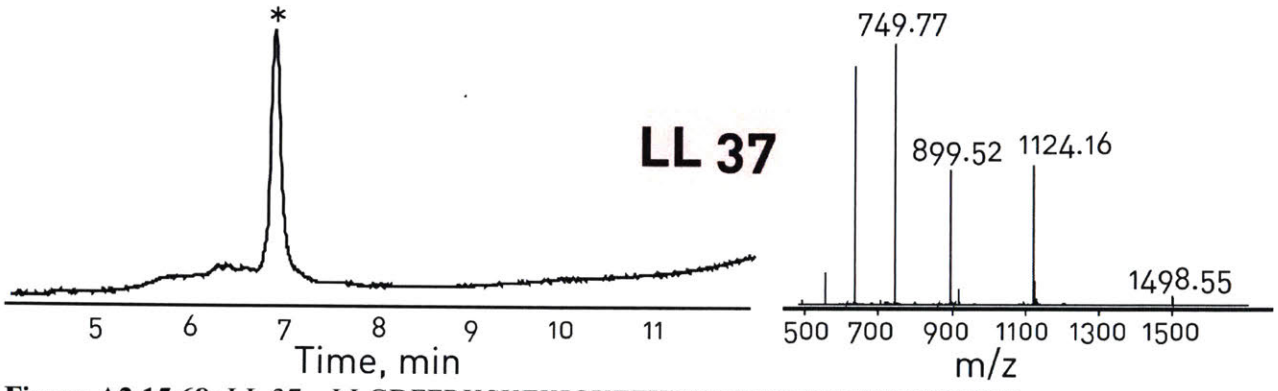


Figure A2.15.68: LL 37 LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES
This crude LL-37 is of much higher quality than the LL-37 analog synthesized with the second generation AFPS on polystyrene resin shown in Figure A2.15.86 (CPP18). TIC is shown. Calculated monoisotopic mass = 4490.57Da, Observed monoisotopic mass = 4490.56Da. MFE failed.

A2.15.5 CPPs

The following CPPs were synthesized with the second generation AFPS instrument under the default conditions described in the manuscript. When problems with histidine incorporation became obvious, several histidine containing CPPs were initially skipped, and are not presented here, thus the numbering of CPPs is not strictly sequential. The skipped CPPs were later synthesized either with freshly prepared reagents on the second generation AFPS, Boc protected histidine on the second generation AFPS, on the third generation AFPS, or using the manual instrument. In all cases, Z = 4-pentynoic acid, introduced as a handle for azide-alkyne click chemistry.

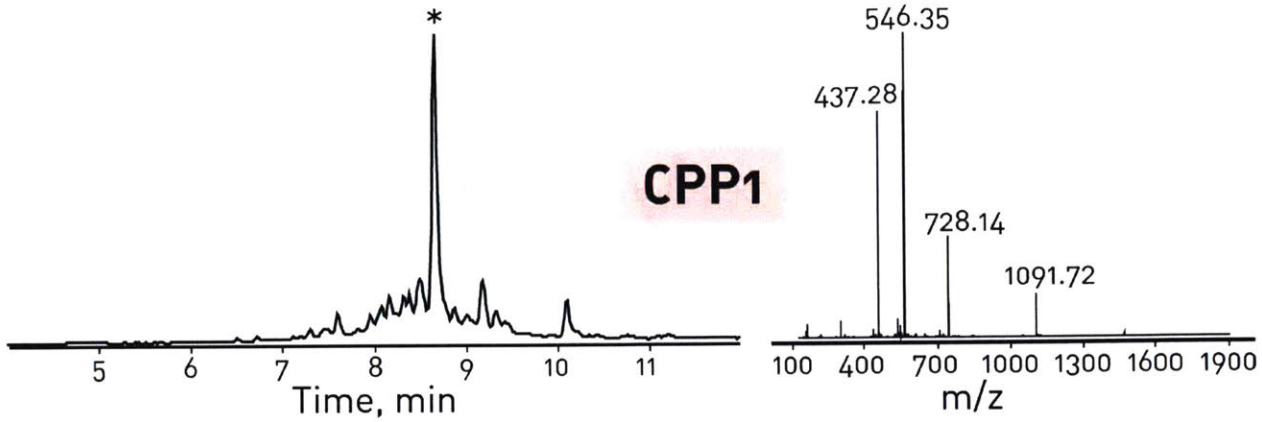


Figure A2.15.69: CPP1 AGYLLGKINLKALAALAKKIL
TIC is shown. Calculated monoisotopic mass = 2180.40Da, Observed monoisotopic mass =2180.39Da

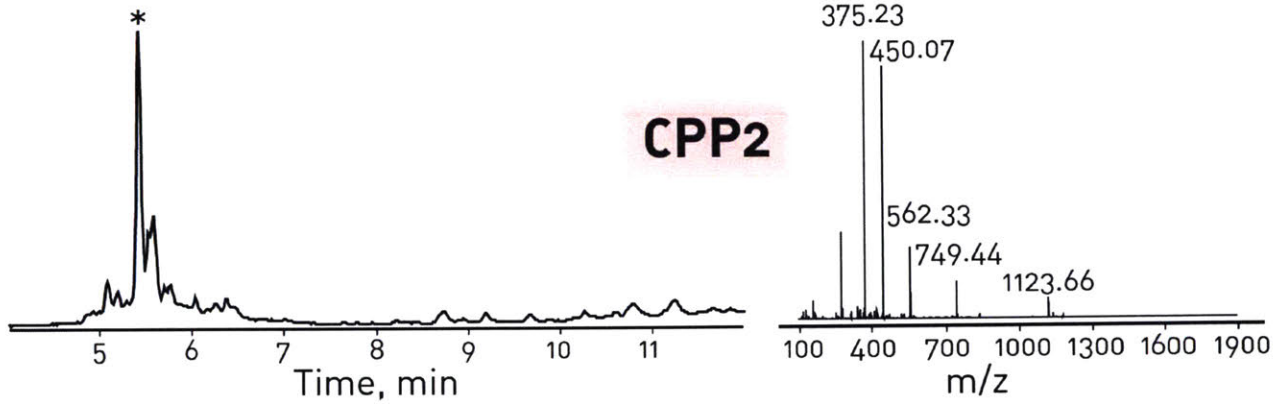


Figure A2.15.70: CPP2 RQIKIWFQNRRMKWKK
TIC is shown. Calculated monoisotopic mass = 2244.30Da, Observed monoisotopic mass =2243.88Da

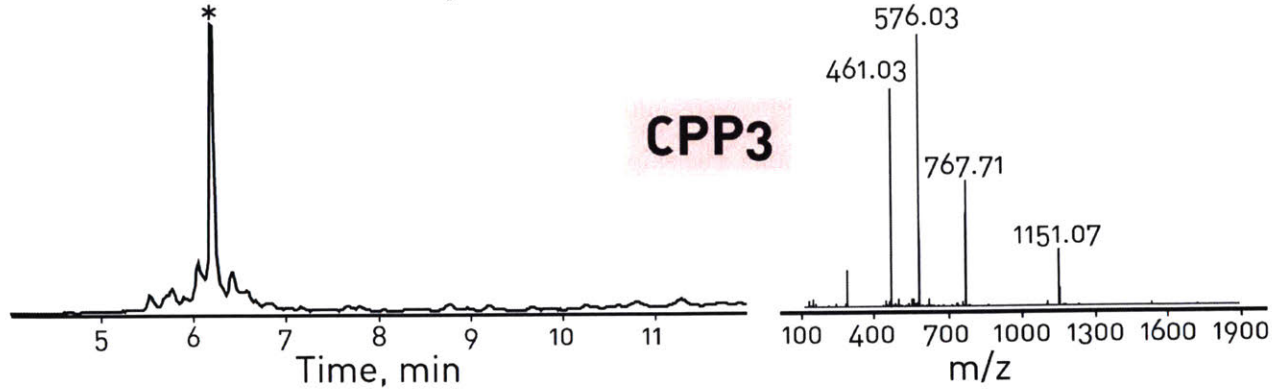


Figure A2.15.71: CPP3 NGYTSRFEREYGKFNKFGT
TIC is shown. Calculated monoisotopic mass = 2299.10Da, Observed monoisotopic mass =2299.11Da

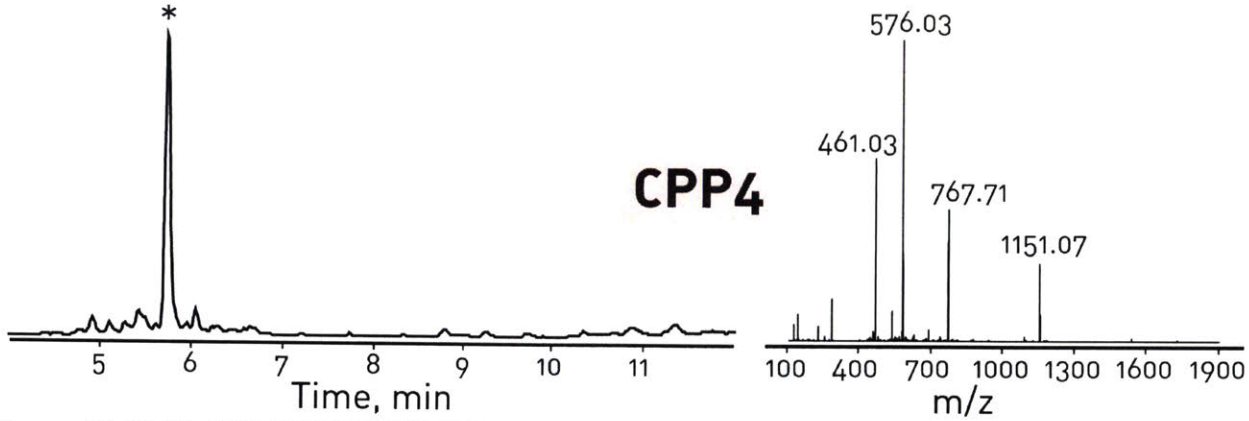


Figure A2.15.72: CPP4 TFFYGGSRGKRNNFKTEEY
TIC is shown. Calculated monoisotopic mass = 2299.10Da, Observed monoisotopic mass =2299.12Da

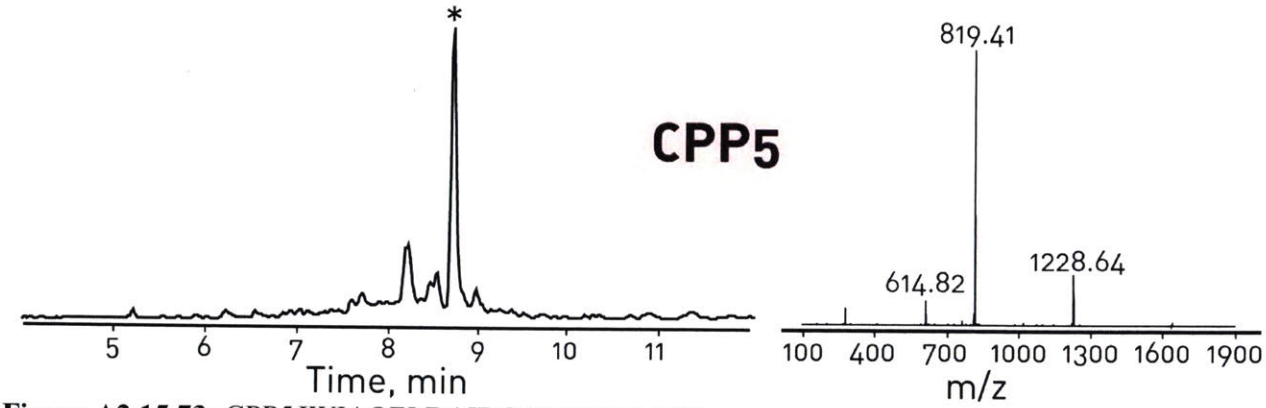


Figure A2.15.73: CPP5 IWIAQELDAIDGAFNAYYARR
TIC is shown. Calculated monoisotopic mass = 2454.23Da, Observed monoisotopic mass =2454.22

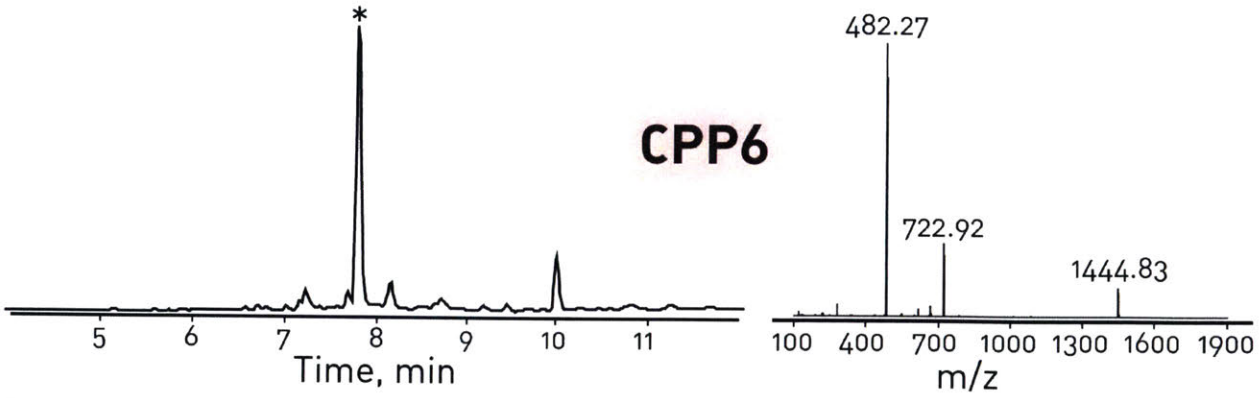


Figure A2.15.74: CPP6 ITFADLLAYYKK
TIC is shown. Calculated monoisotopic mass = 1443.80Da, Observed monoisotopic mass =1443.79Da

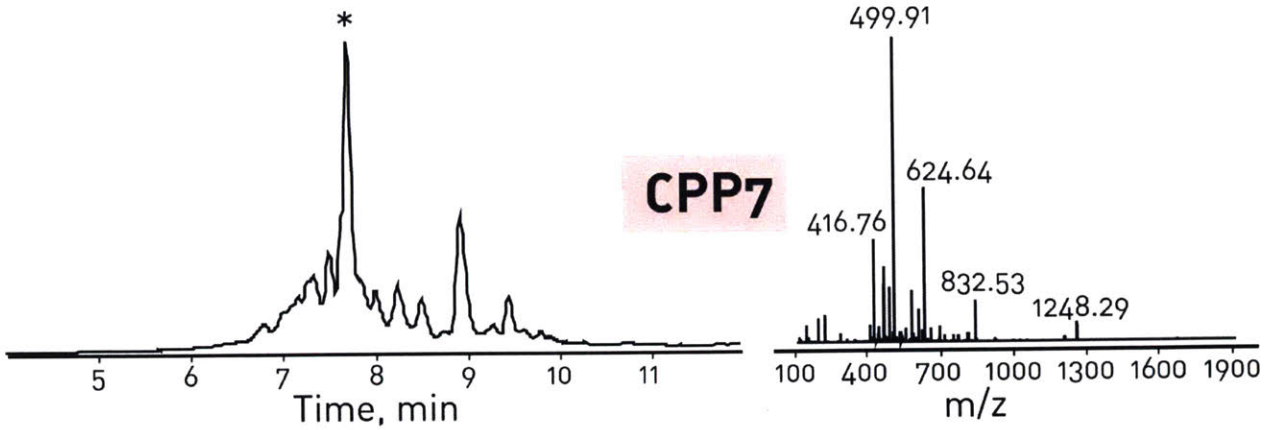


Figure A2.15.75: CPP7 ZMGLGLHLLVLAAALQGAKKKRKV
TIC is shown. Calculated monoisotopic mass = 2493.51Da, Observed monoisotopic mass =2493.51Da

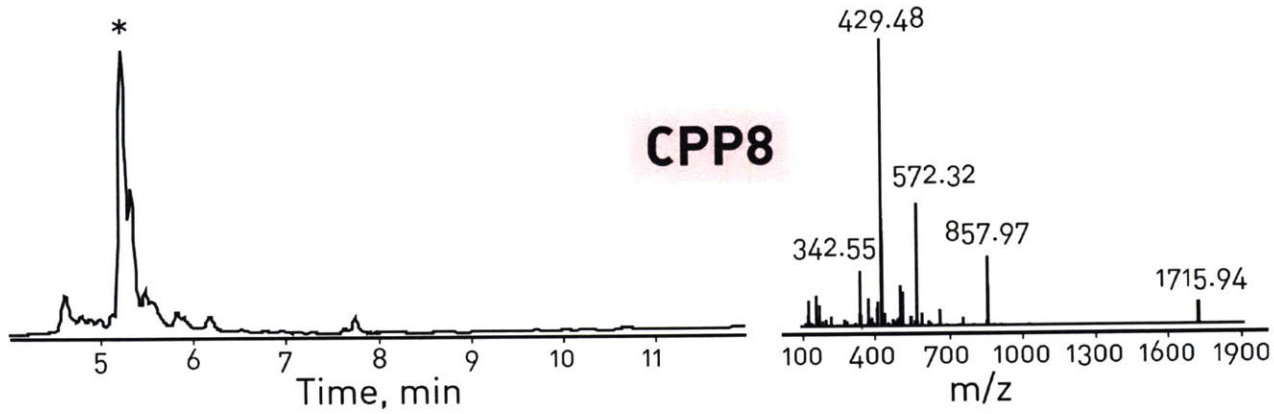


Figure A2.15.76: CPP8 ZMVKSKIGSWILVLFVAMWSDVGLCKKRPKP
TIC is shown. Calculated monoisotopic mass = 3494.91Da, Observed monoisotopic mass =1713.89Da (Failed – truncation without cap after coupling Ala-17)

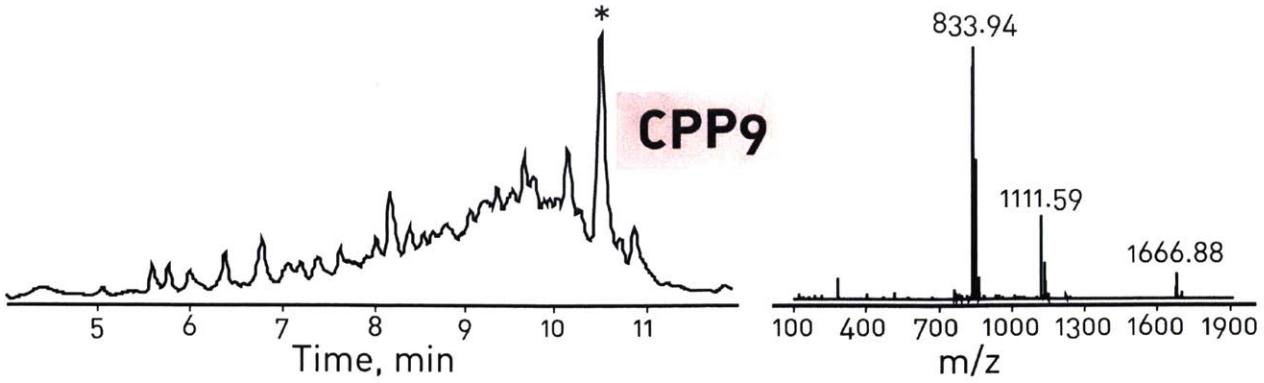


Figure A2.15.77: CPP9 ZMANLGYWLLALFVTMWTDVGLCKKRPKP
TIC is shown. Calculated monoisotopic mass = 3329.73Da, Observed monoisotopic mass =3329.74Da

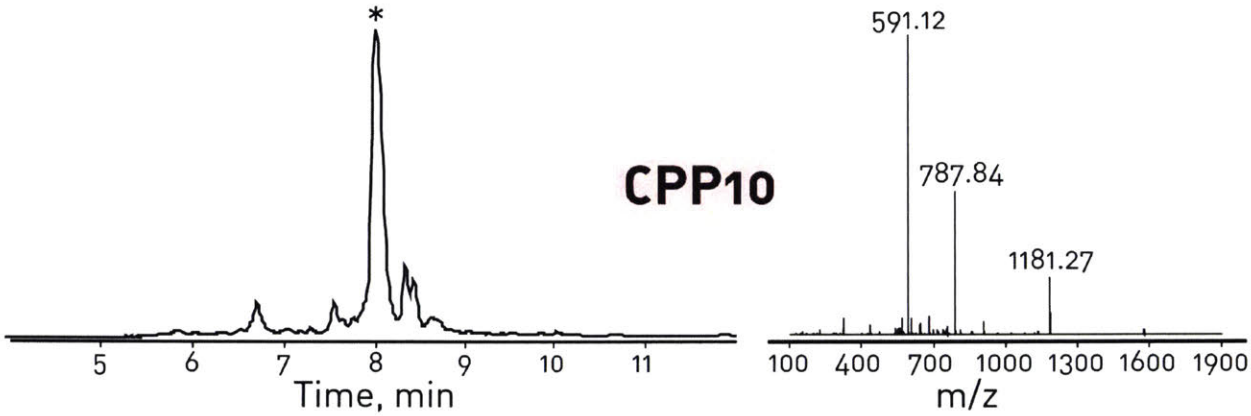


Figure A2.15.78: CPP10 ZAAVLLPVLLAAPVQRKRQKLP
TIC is shown. Calculated monoisotopic mass = 2359.46Da, Observed monoisotopic mass =2359.47Da

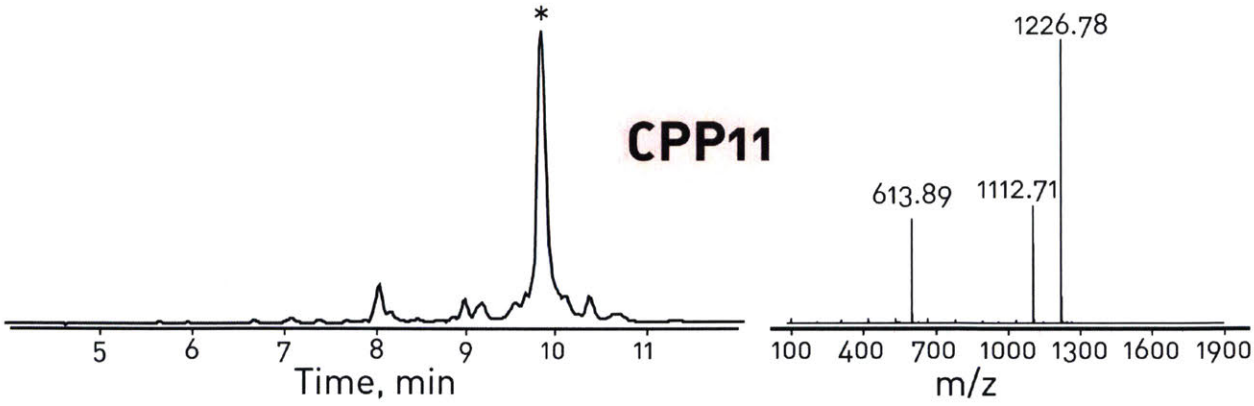


Figure A2.15.79: CPP11 ZAAVLLPVLLAAP
TIC is shown. Calculated monoisotopic mass = 1225.74Da, Observed monoisotopic mass =1225.78Da

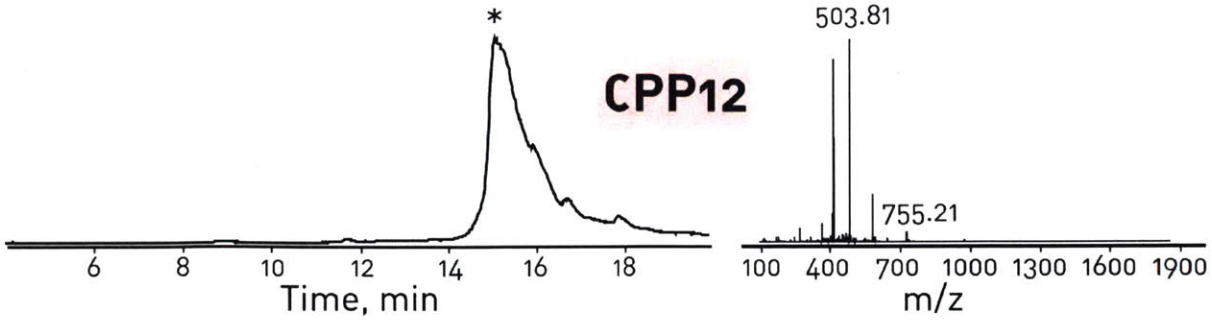


Figure A2.15.80: CPP12 ZRRIRPRPPRLPRPRPRPLPFPRPG
TIC is shown. Calculated monoisotopic mass = 3015.80Da, Observed monoisotopic mass =3015.81Da

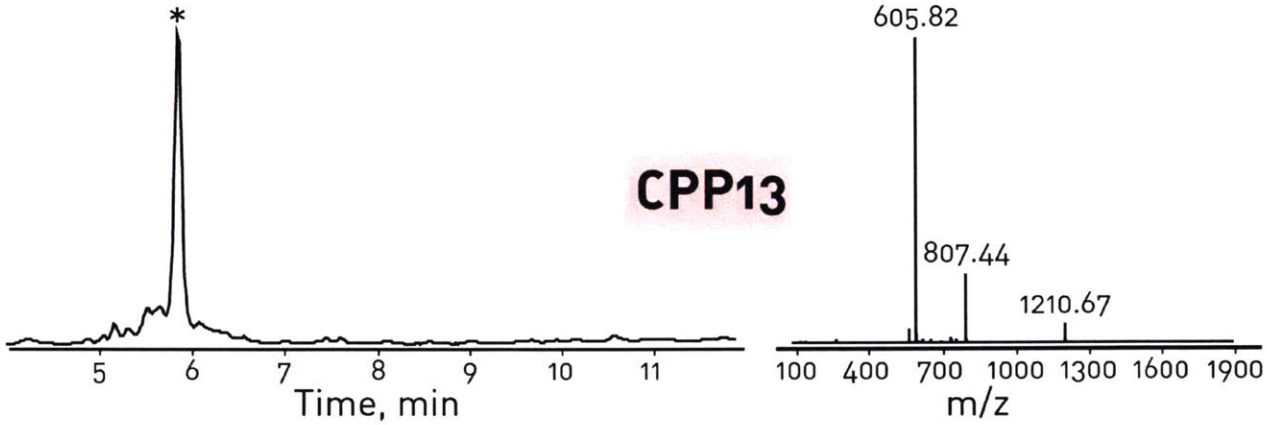


Figure A2.15.81: CPP13 ZVDKGSYLPRPTPPRPIYNRN
TIC is shown. Calculated monoisotopic mass = 2418.26Da, Observed monoisotopic mass =2418.27Da

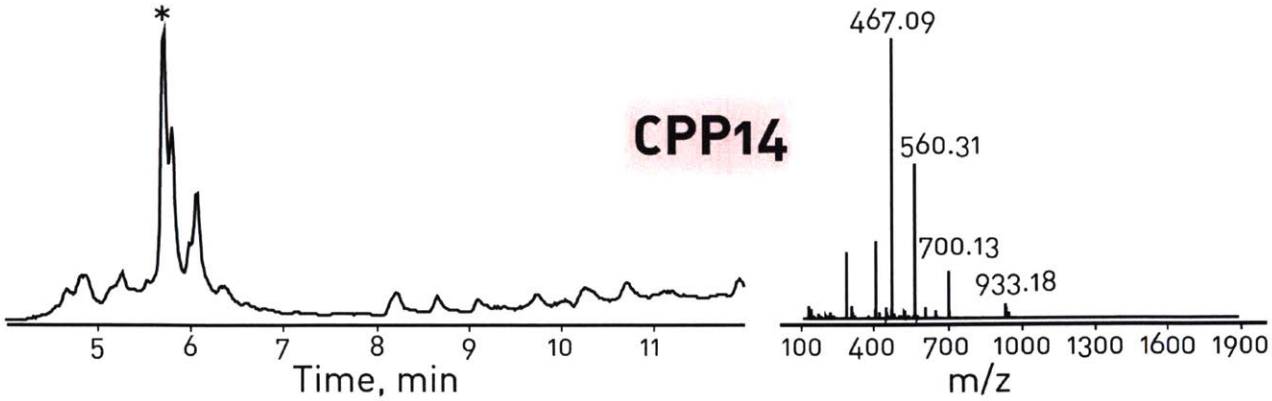


Figure A2.15.82: CPP14 ZKCFQWQRNMRKVRGPPVSCIKR
TIC is shown. Calculated monoisotopic mass = 2795.45Da, Observed monoisotopic mass = 2795.51Da

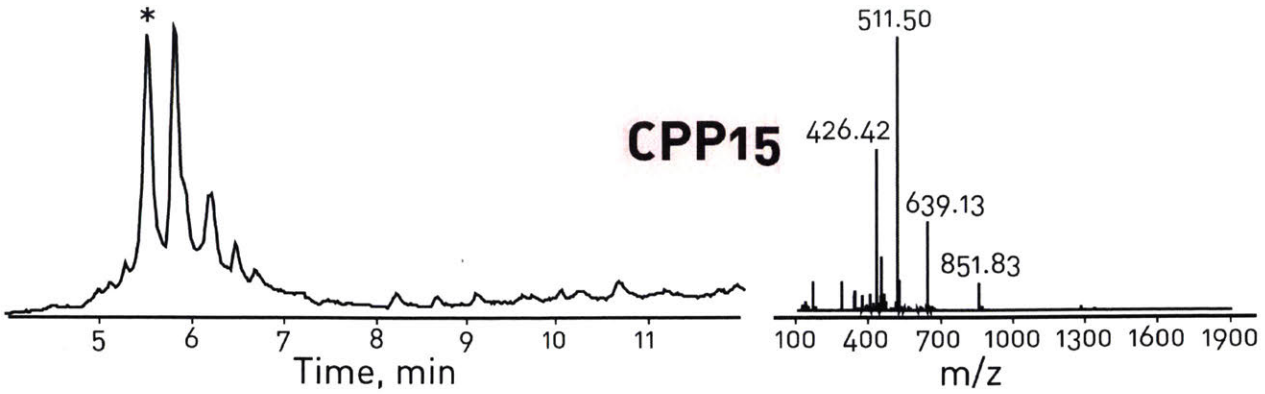


Figure A2.15.83: CPP15 ZTRSSRAGLQWPVGRVHRLLRK
TIC is shown. Calculated monoisotopic mass = 2551.45Da, Observed monoisotopic mass = 2551.47Da

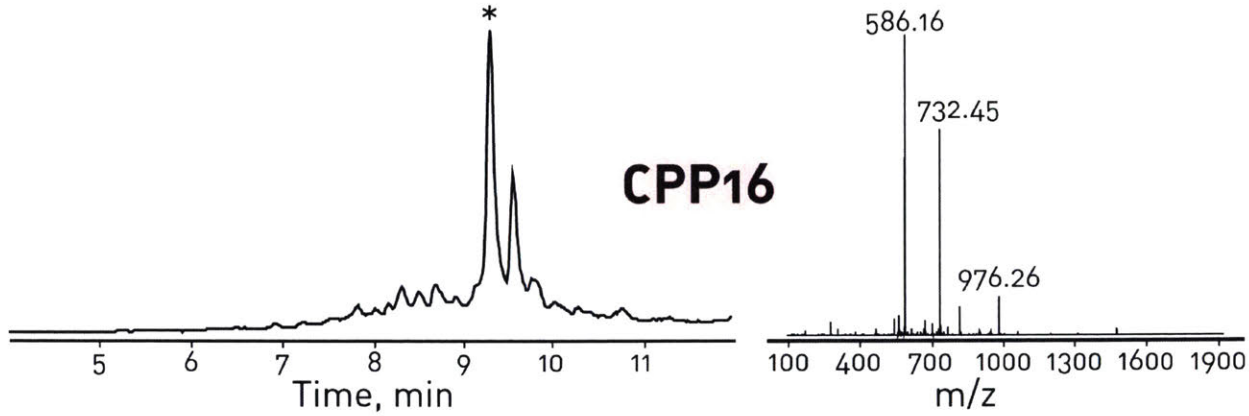


Figure A2.15.84: CPP16 ZGIGAVLKVLTTGLPALISWIKRKRQQ
TIC is shown. Calculated monoisotopic mass = 2924.74Da, Observed monoisotopic mass = 2924.77Da

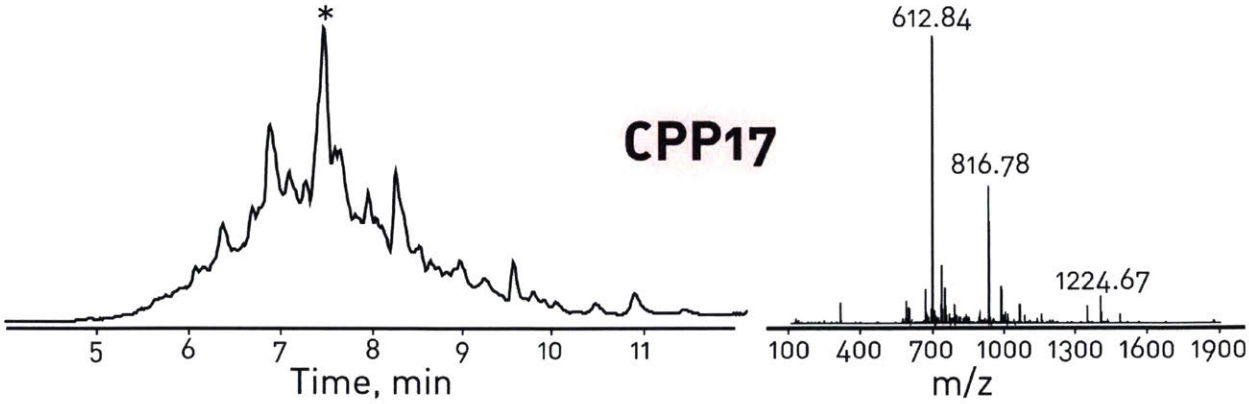


Figure A2.15.85: CPP17 ZGIGKWLHSAKKFGKAFVGEIMNS
TIC is shown. Calculated monoisotopic mass = 2583.34Da, Observed monoisotopic mass = 2446.31Da
(Most prominent product is des-His; desired product is the peak eluting at 6.9min)

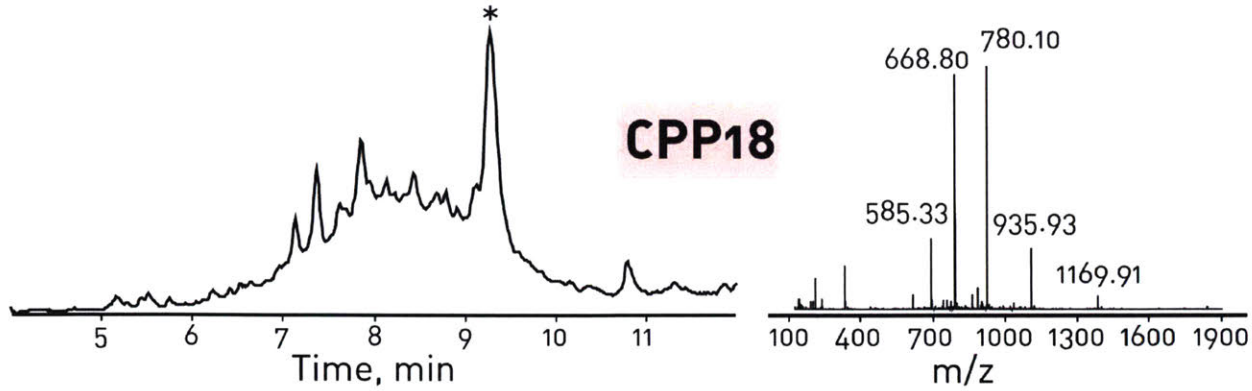


Figure A2.15.86: CPP18 ZLLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTESC
This CPP is an LL-37 with an N-terminal alkyne handle and a C-terminal cysteine. The native LL-37, produced with the third generation AFPS on ChemMatrix PEG resin is of much higher quality (Figure A2.15.68). TIC is shown. Calculated monoisotopic mass = 4672.59Da, Observed monoisotopic mass = 4672.59Da

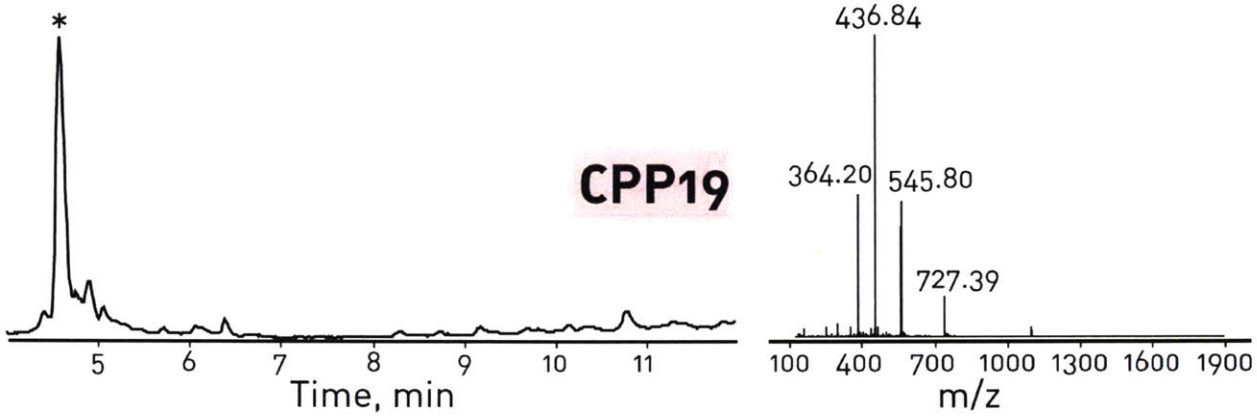


Figure A2.15.87: CPP19 ZRGGRLSYSRRRFSTSTGR
TIC is shown. Calculated monoisotopic mass = 2178.13Da, Observed monoisotopic mass =2178.16Da

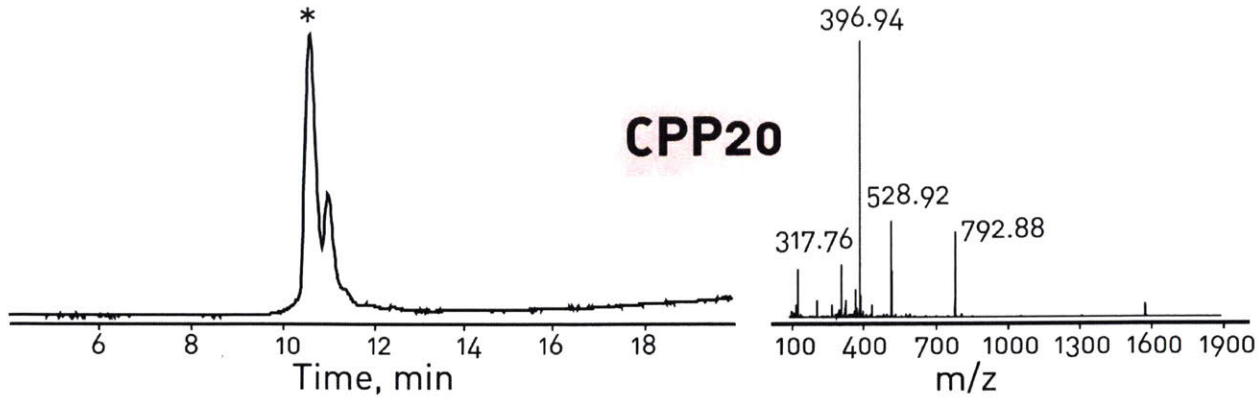


Figure A2.15.88: CPP20 ZYKQCHKKGGKKGSG
TIC is shown. Calculated monoisotopic mass = 1583.80Da, Observed monoisotopic mass =1583.74Da

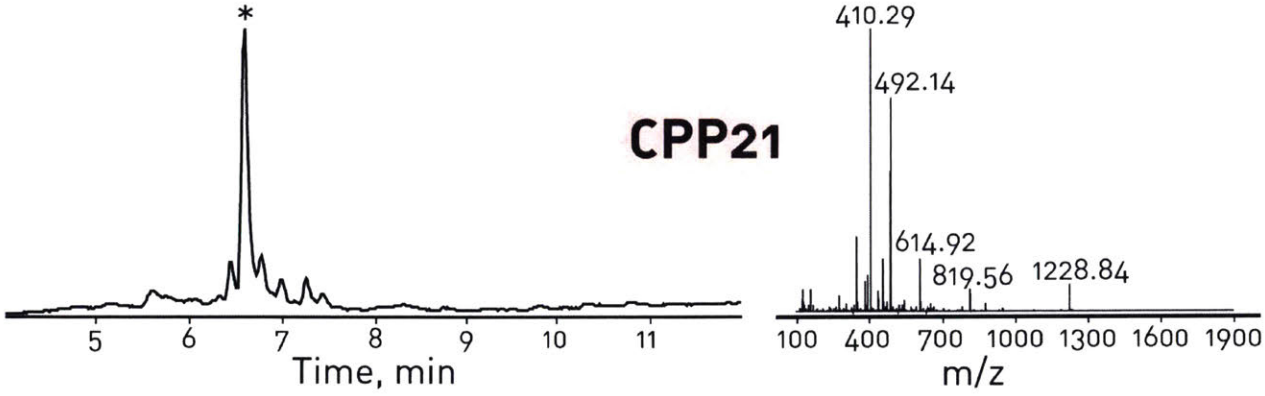


Figure A2.15.89: CPP21 ZALWKTLLKKVLKAPKKKRKV
TIC is shown. Calculated monoisotopic mass = 2454.60Da, Observed monoisotopic mass =2454.67Da

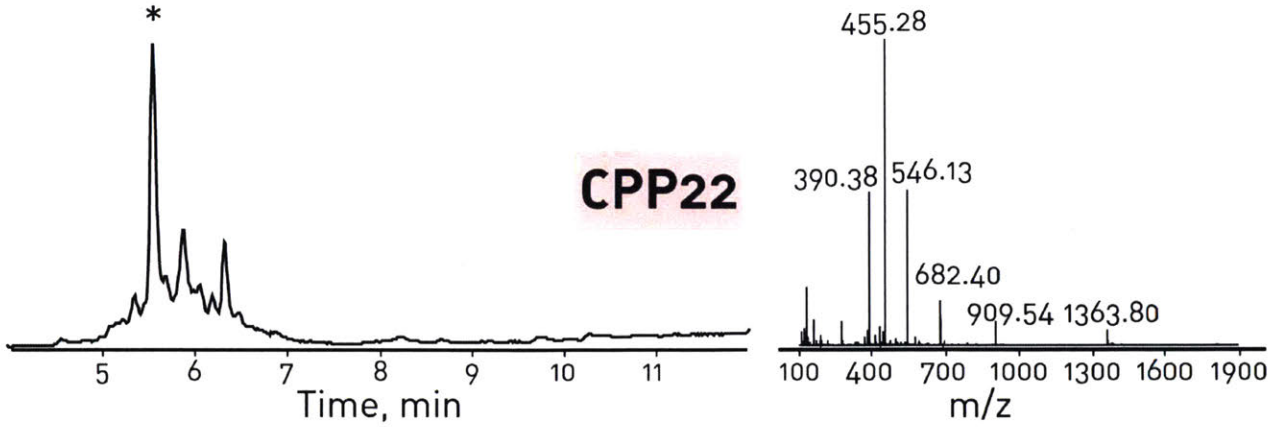


Figure A2.15.90: CPP22 ZHARIKPTFRRLKWKYKGKFW
TIC is shown. Calculated monoisotopic mass = 2724.54Da, Observed monoisotopic mass =2724.60Da

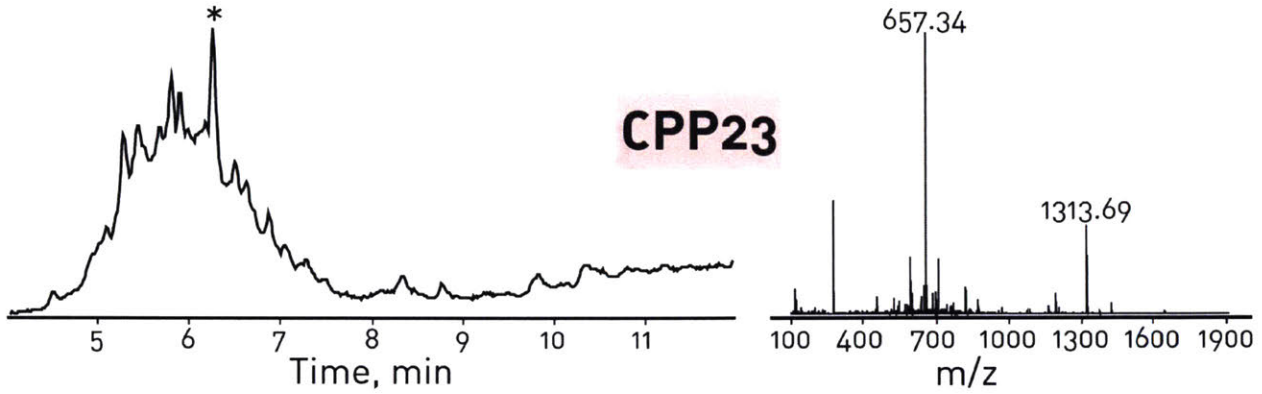


Figure A2.15.91: CPP23 ZTKRRITPKDVIDVRSVTTEINT
TIC is shown. Calculated monoisotopic mass = 2620.43Da, Observed monoisotopic mass =1312.67Da (Failed – desired product not found amongst low molecular weight peptidic products)

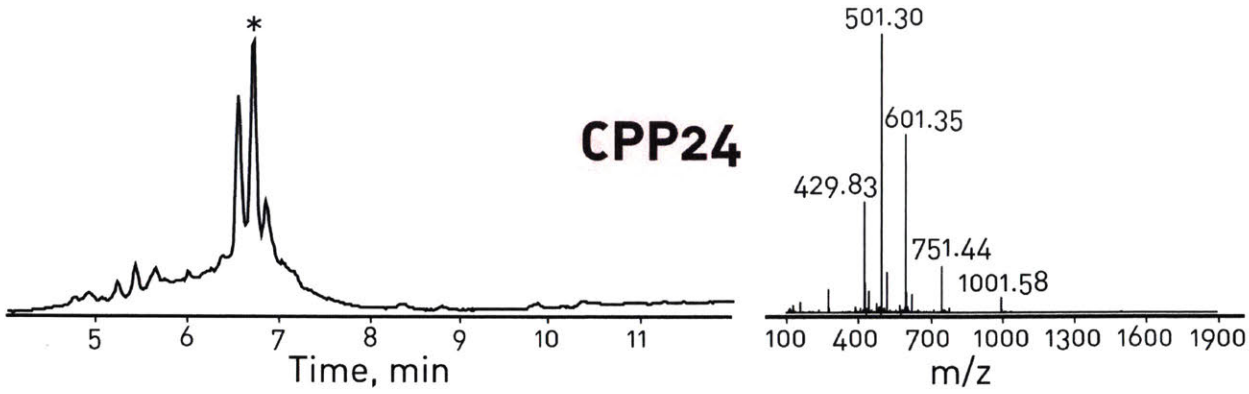


Figure A2.15.92: CPP24 ZRQGAARVTSWLGRQLRIAGKRLEGRSK
TIC is shown. Calculated monoisotopic mass = 3128.77Da, Observed monoisotopic mass =3000.73Da (Des-Gln was the primary product. The desired product was the prominent earlier eluting peak at 6.5min)

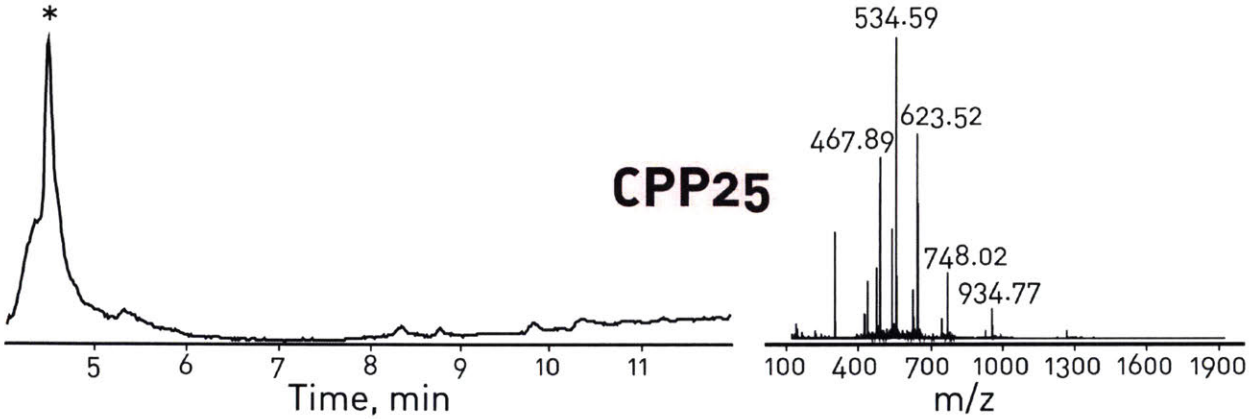


Figure A2.15.93: CPP25 ZNAATATRGRSAASRPTQRPRAPARSASRPRRPVQ
TIC is shown. Calculated monoisotopic mass = 3733.01Da, Observed monoisotopic mass =3733.07Da

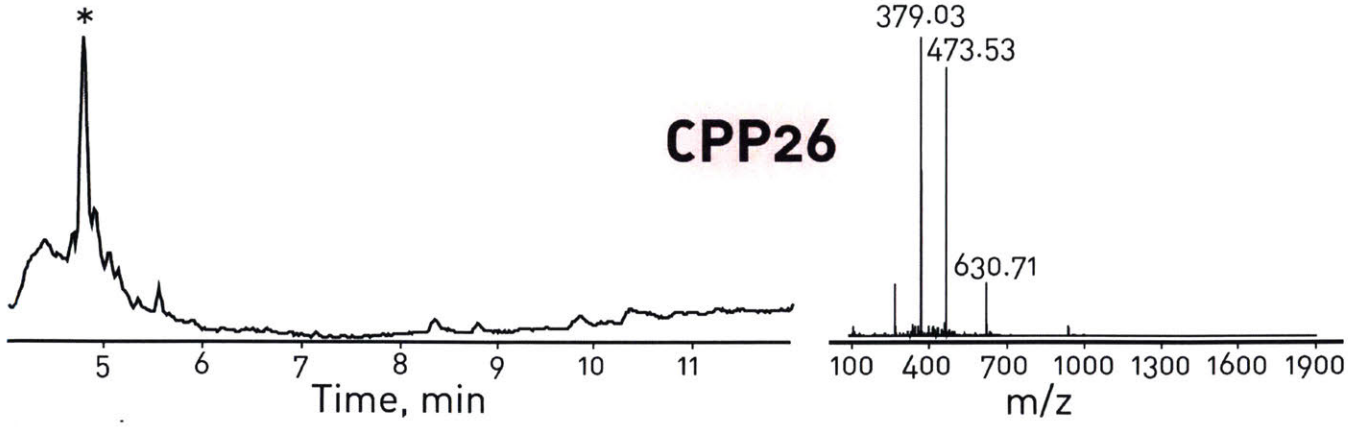


Figure A2.15.94: CPP26 ZRHSRIGIIQQRRTRNG
TIC is shown. Calculated monoisotopic mass = 2026.11Da, Observed monoisotopic mass =1889.11Da (Failed – major product is des-His)

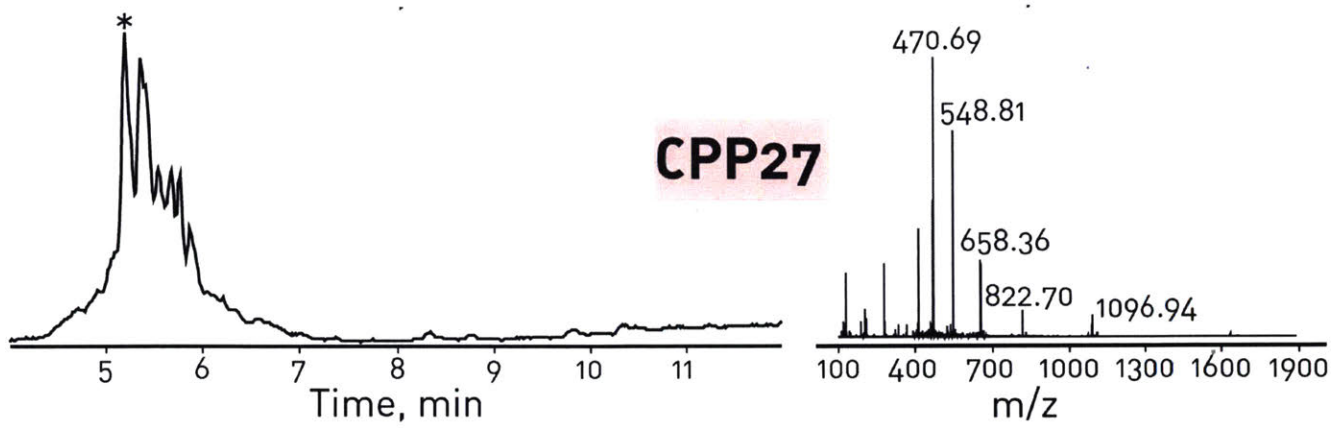


Figure A2.15.95: CPP27 ZKLIKGRTPIKFGKADCDRPPKHSQNGMGK
TIC is shown. Calculated monoisotopic mass = 3285.73Da, Observed monoisotopic mass =3285.79Da

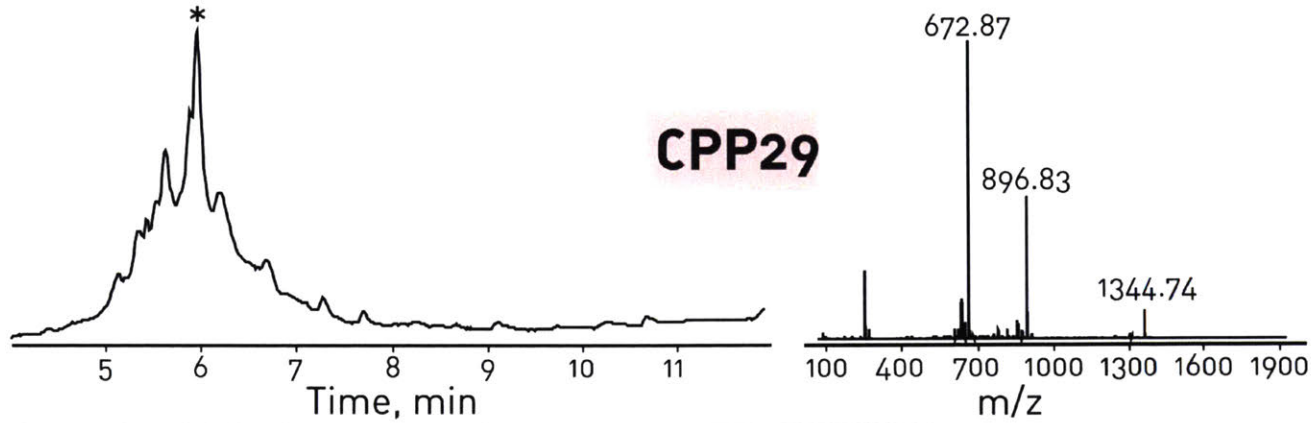


Figure A2.15.96: CPP29 ZDPKGDPKGVTVTVTVTVTGKGDPKPD
TIC is shown. Calculated monoisotopic mass = 2686.38Da, Observed monoisotopic mass =2686.45Da

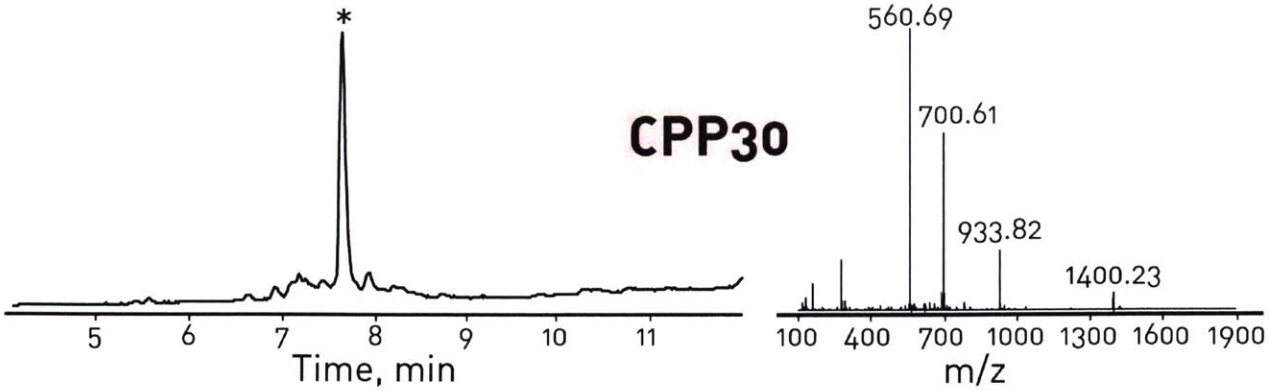


Figure A2.15.97: CPP30 ZKETWWETWWTEWSQPKKRKV
TIC is shown. Calculated monoisotopic mass = 2797.38Da, Observed monoisotopic mass =2797.42Da

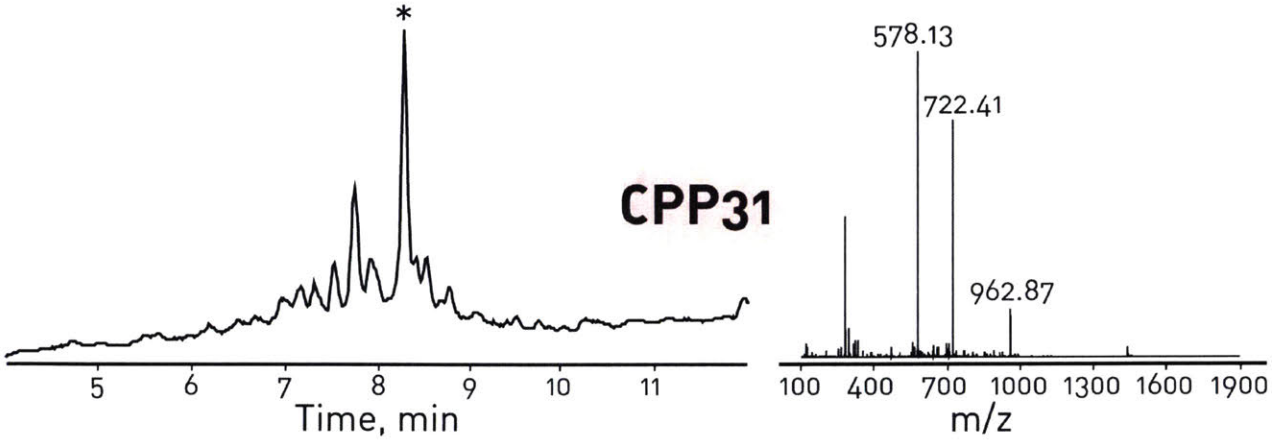


Figure A2.15.98: CPP31 ZGLAFLGFLGAAGSTMGAWSQPKKKRKV
TIC is shown. Calculated monoisotopic mass = 2884.55Da, Observed monoisotopic mass =2884.59Da

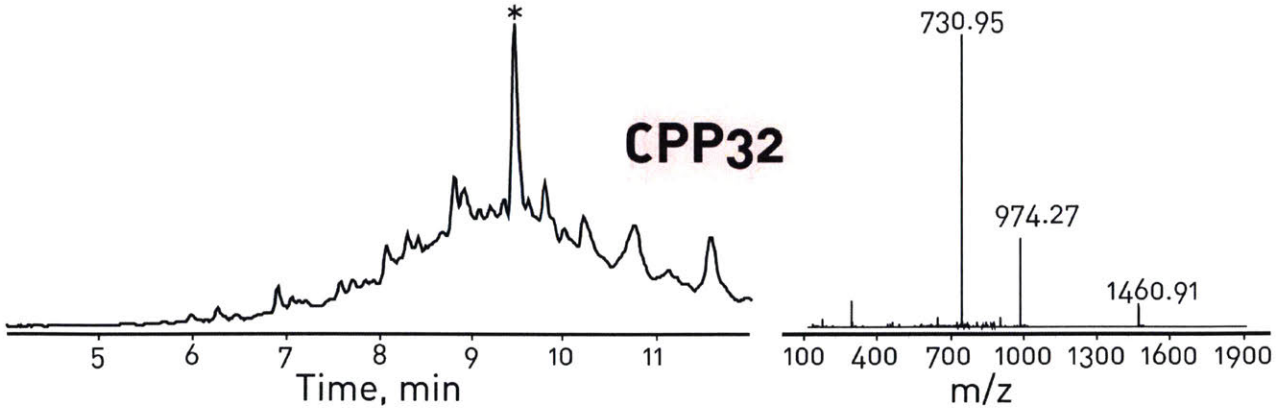


Figure A2.15.99: CPP32 ZGWTLNSAGYLLGKINLKALAALAKKIL
TIC is shown. Calculated monoisotopic mass = Da, Observed monoisotopic mass =2918.79Da

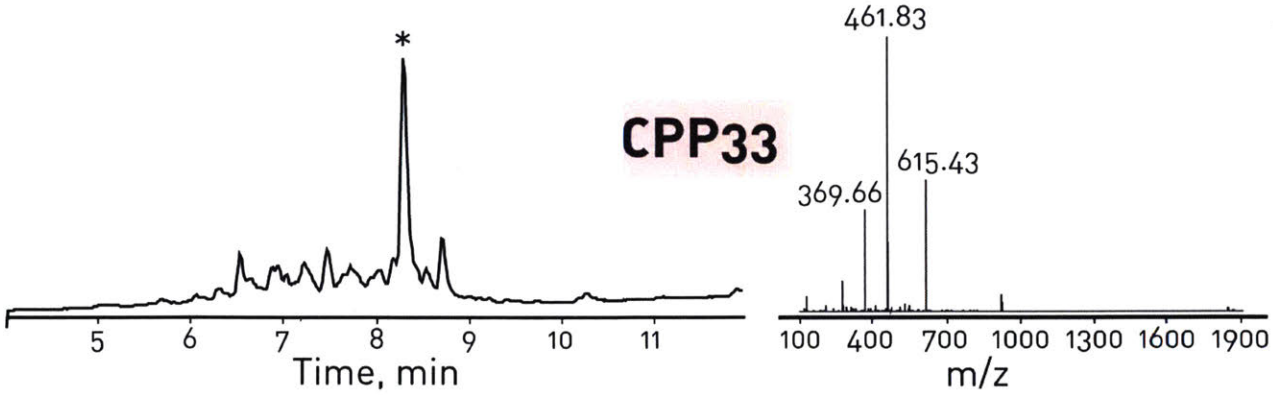


Figure A2.15.100: CPP33 ZKLALKALKALKAALKLA
TIC is shown. Calculated monoisotopic mass = 1842.22Da, Observed monoisotopic mass =1842.27Da

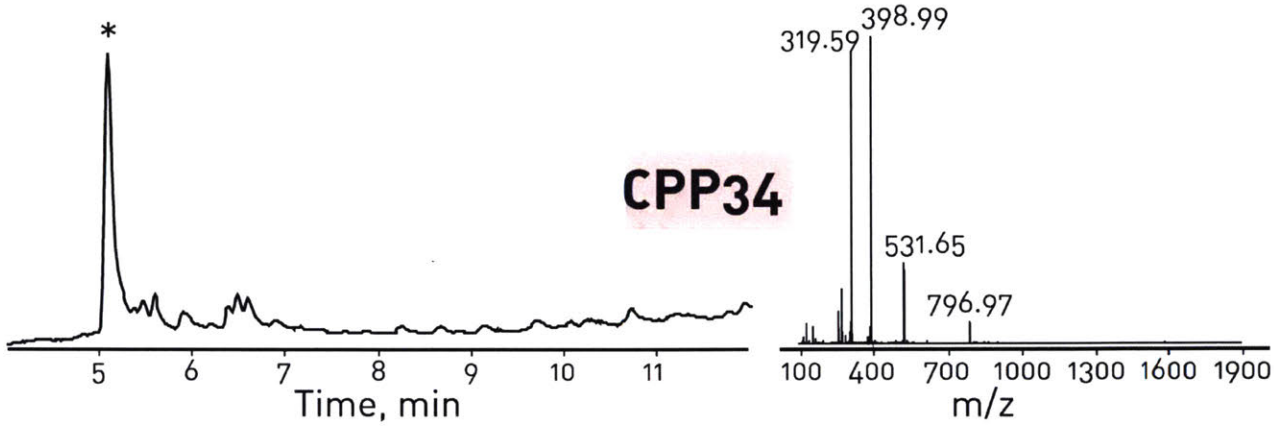


Figure A2.15.101: CPP34 ZRRWWRRWRR
TIC is shown. Calculated monoisotopic mass = 1591.86Da, Observed monoisotopic mass =1591.92Da

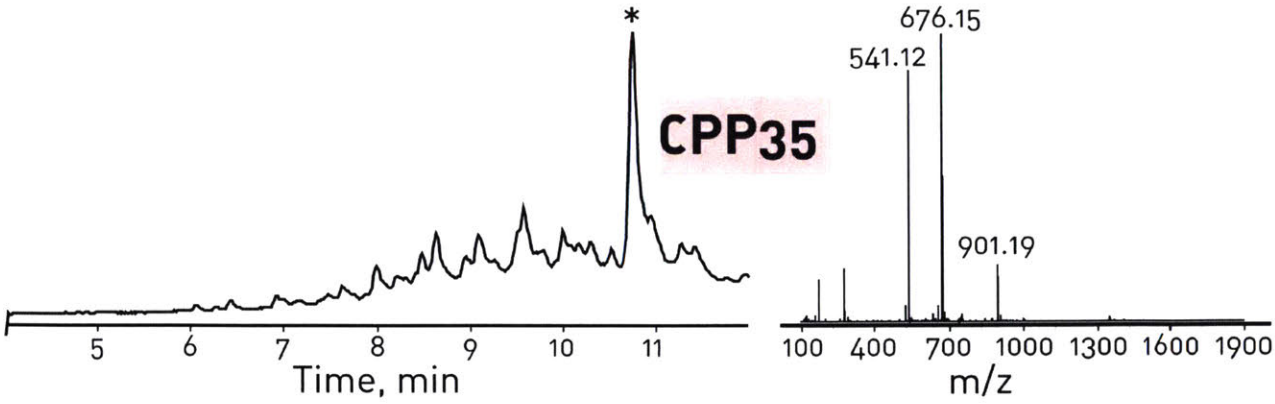


Figure A2.15.102: CPP35 ZGLWRALWRLLRSLWRLLWRA
TIC is shown. Calculated monoisotopic mass = 2699.56Da, Observed monoisotopic mass =2699.55Da

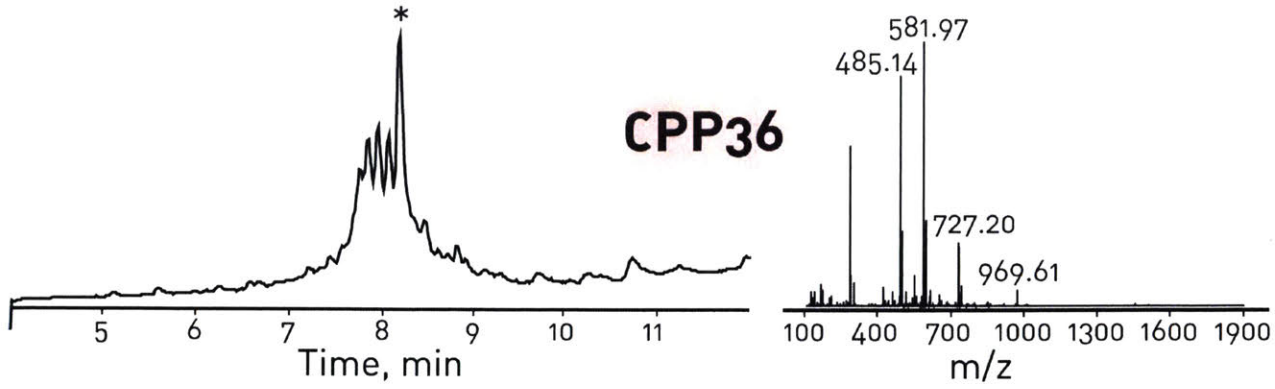


Figure A2.15.103: CPP36 ZLIRLWSHLIHIWFQNRRLKWKKK
TIC is shown. Calculated monoisotopic mass = 3177.85Da, Observed monoisotopic mass =2903.78Da (Failed – primary product was double histidine deletion)

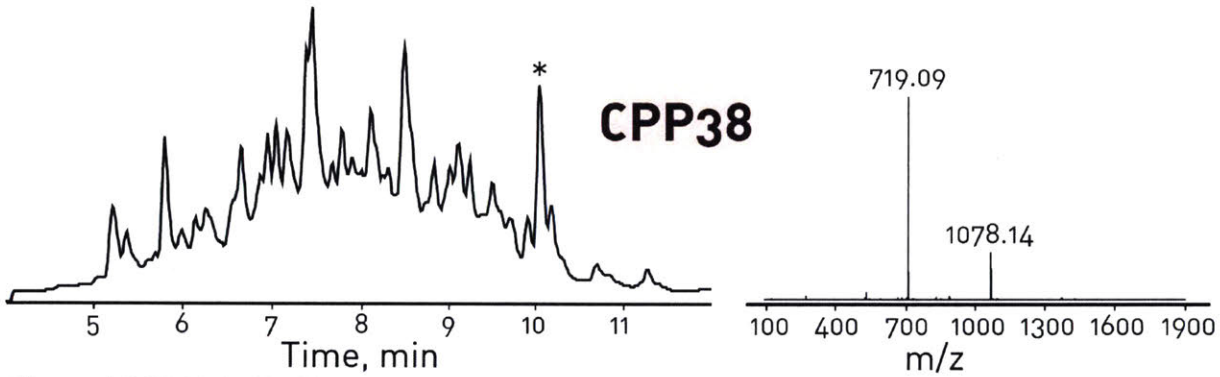


Figure A2.15.104: CPP38 ZLKTLTETLKELTKTLTEL
TIC is shown. Calculated monoisotopic mass = 2153.22Da, Observed monoisotopic mass =2153.24Da

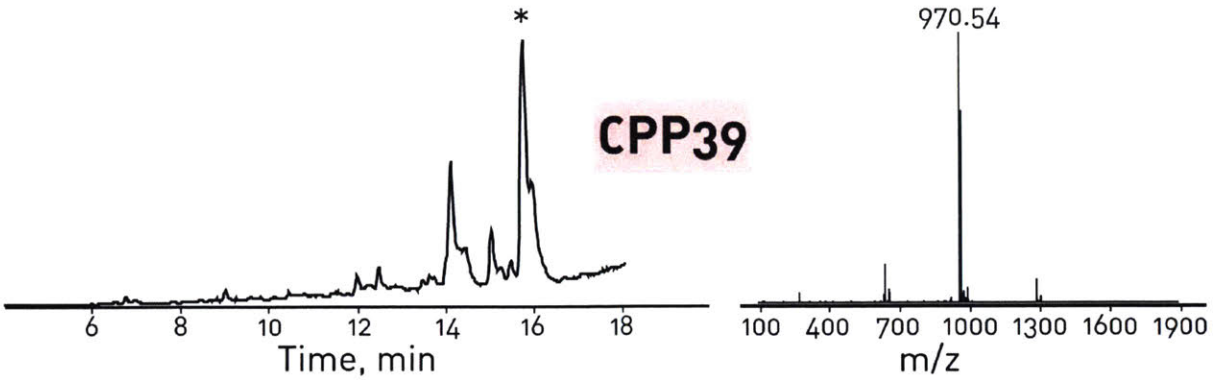


Figure A2.15.105: CPP39 ZQLALQLALQALQAALQLA
TIC is shown. Calculated monoisotopic mass = 1955.12Da, Observed monoisotopic mass =1955.11Da. The most prominent peak in the MS is $[M+H^{+}-NH_2^{-}]^{2+}$ but $[M+2H^{+}]^{2+}$ is also observed, as are other species. This peptide has unusual behavior in the ESI source because it has no charge carriers. It also elutes in ~80% acetonitrile.

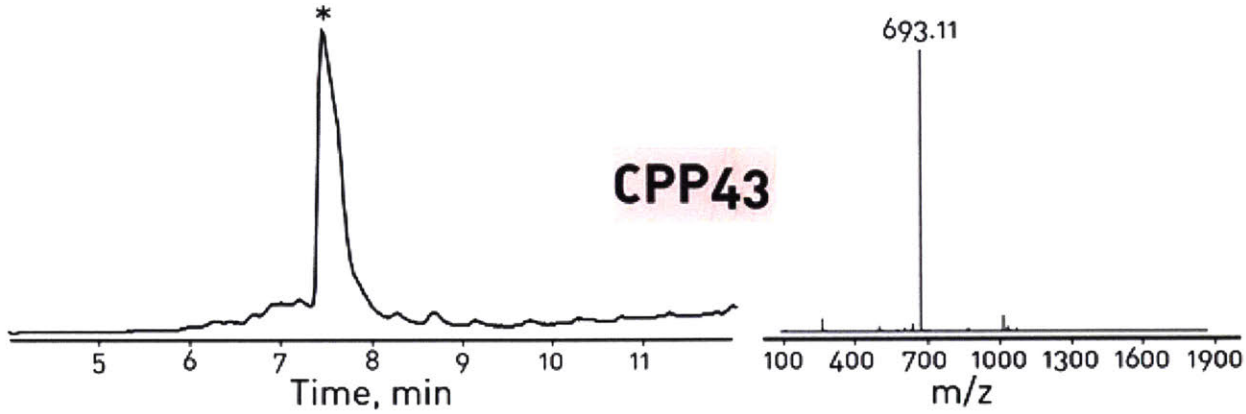


Figure A2.15.106: CPP43 ZVRLPPPVRLPPPVRLPPP
TIC is shown. Calculated monoisotopic mass = 2075.25Da, Observed monoisotopic mass = 2075.30Da

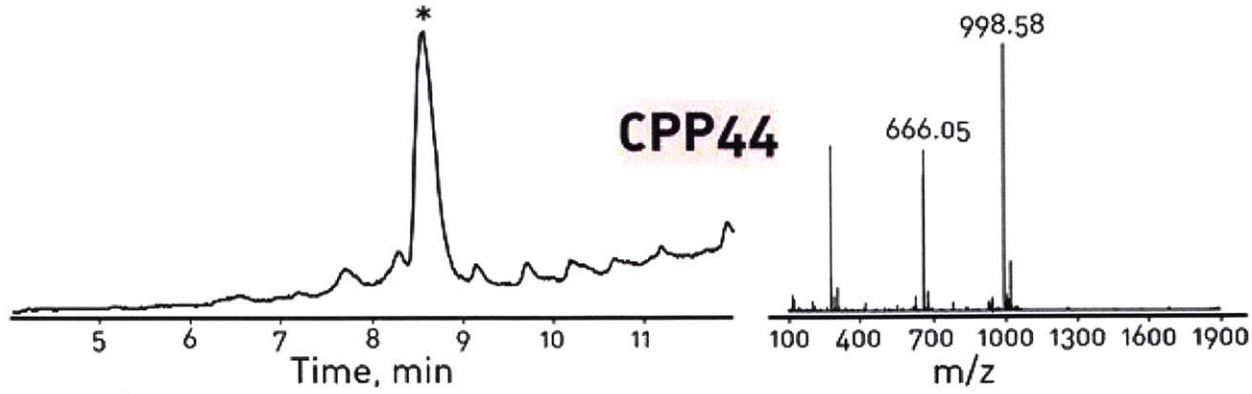


Figure A2.15.107: CPP44
TIC is shown. Calculated monoisotopic mass = 1994.08Da, Observed monoisotopic mass = 1994.14Da

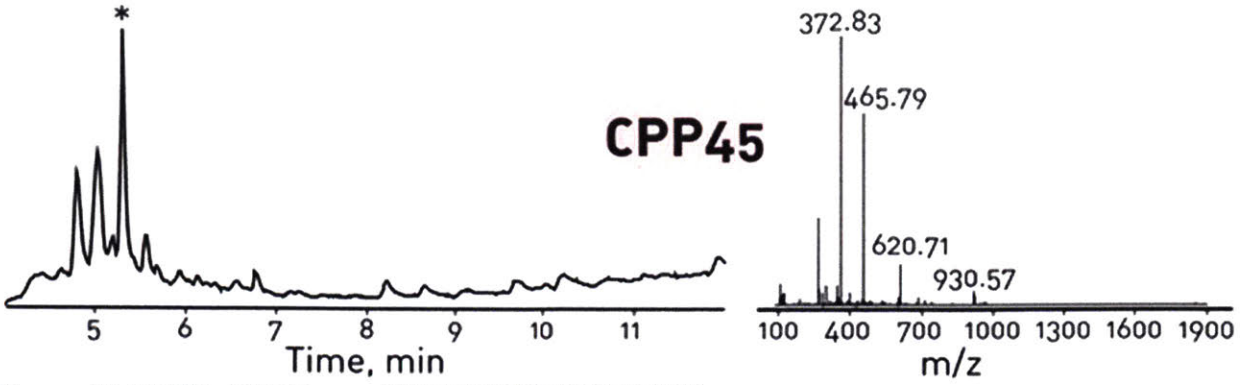


Figure A2.15.108: CPP45 ZFKIYDKKVRTRVVKH
TIC is shown. Calculated monoisotopic mass = 1995.15Da, Observed monoisotopic mass = 1858.12Da (Failed – primary product is des-His)

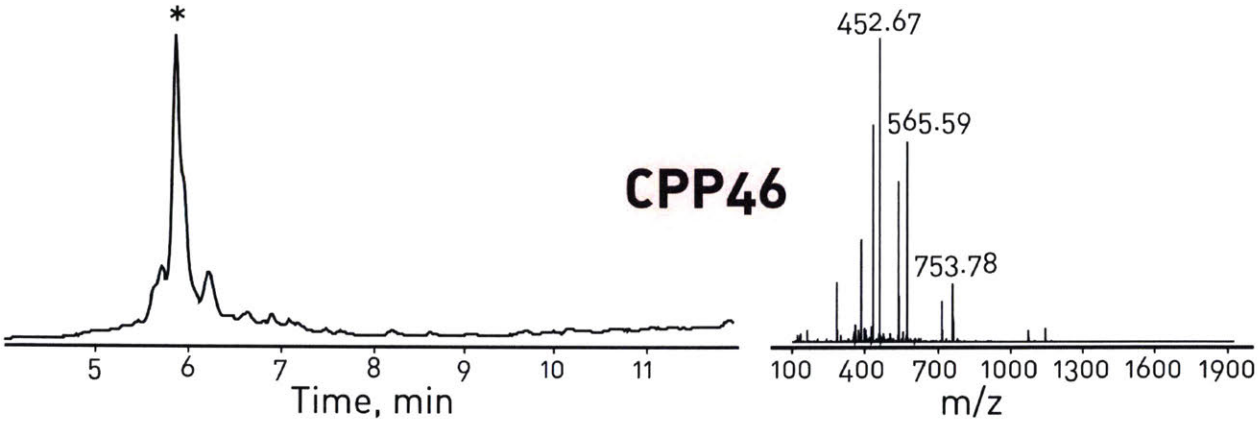


Figure A2.15.109: CPP46 ZRASKRDGSWVKKLHRILE
TIC is shown. Calculated monoisotopic mass =2257.25 Da, Observed monoisotopic mass = 2257.33Da

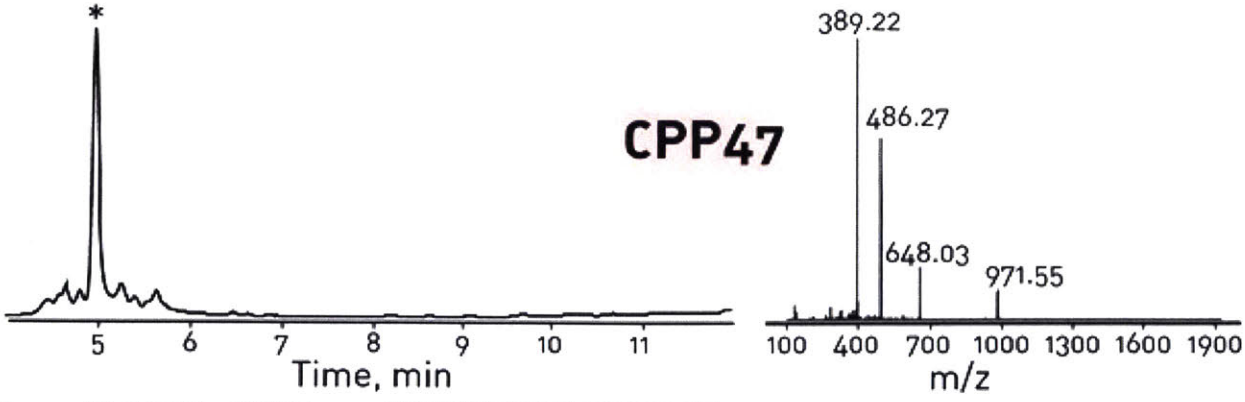


Figure A2.15.110: CPP47 ZKGTYKKKLMRIPLKGT
TIC is shown. Calculated monoisotopic mass = 1940.14Da, Observed monoisotopic mass = 1940.04Da

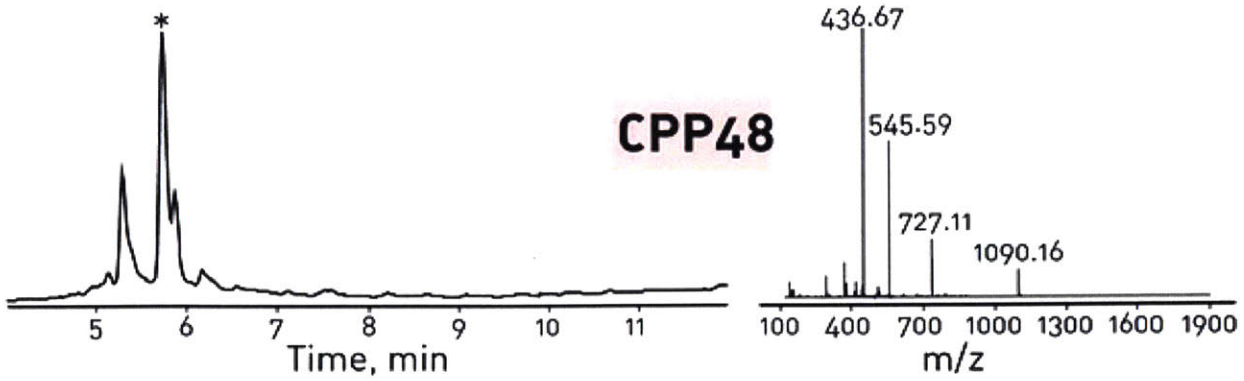


Figure A2.15.111: CPP48 ZLYKKGPAKKGRPPLRGWFH
TIC is shown. Calculated monoisotopic mass = 2314.30Da, Observed monoisotopic mass =2177.30Da (Failed – Primary product is des-His)

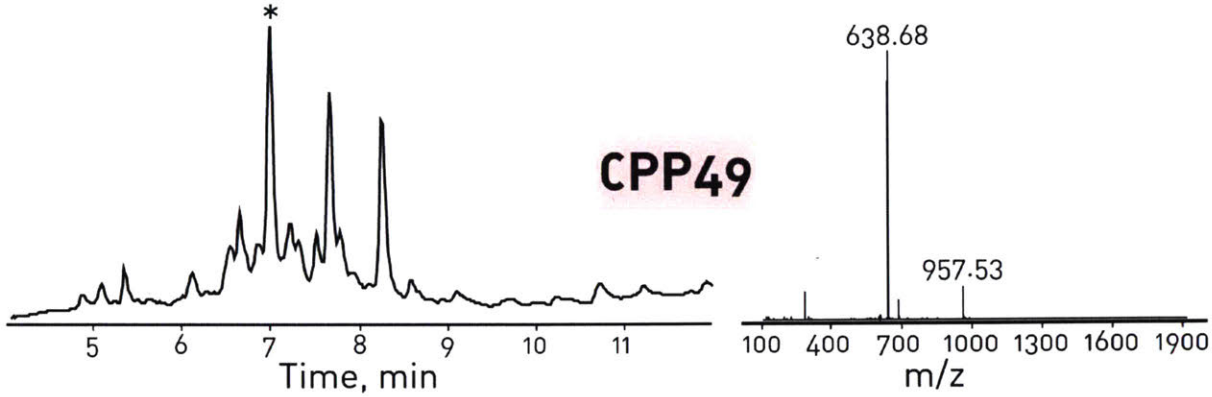


Figure A2.15.112: CPP49 ZHSPIIPLGTRFVCHGVT
TIC is shown. Calculated monoisotopic mass = 1911.98Da, Observed monoisotopic mass = 1912.04Da Single and double histidine deletions are observed at 7.6min and 8.3min, respectively.

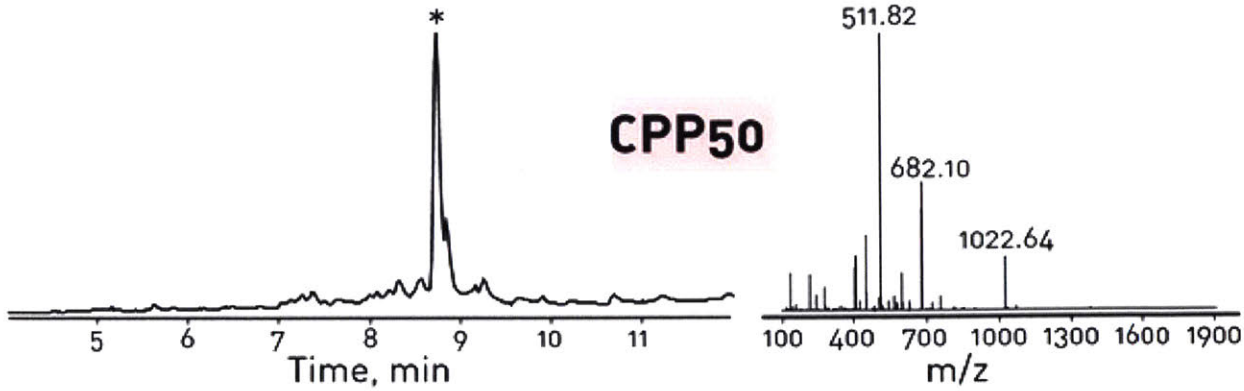


Figure A2.15.113: CPP50 ZYTAIAWVKAFIRKLRK
TIC is shown. Calculated monoisotopic mass = 2042.19Da, Observed monoisotopic mass = 2042.26Da

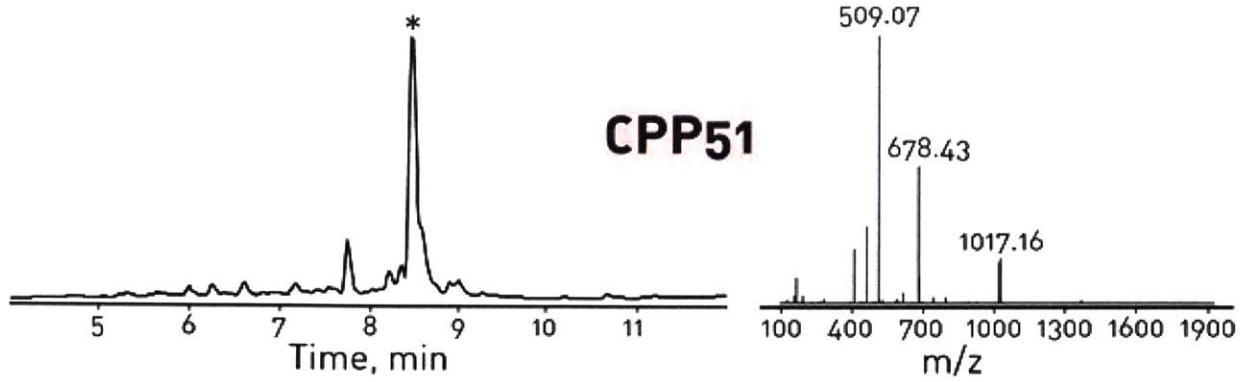


Figure A2.15.114: CPP51 ZIAWVKAFIRKLRKGPLG
TIC is shown. Calculated monoisotopic mass = 2031.23Da, Observed monoisotopic mass = 2031.26 Da

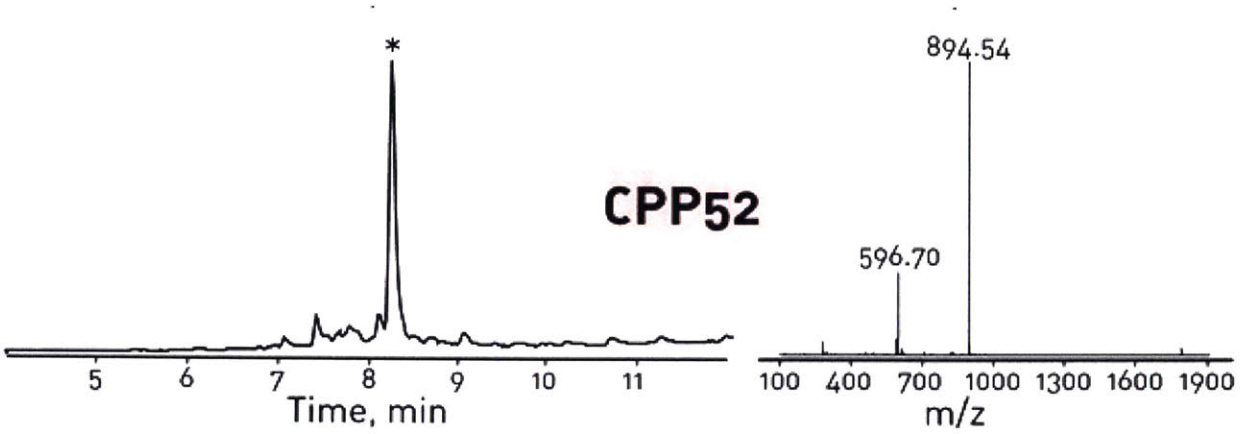


Figure A2.15.115: CPP52 ZRLSGMNEVLSFRWL
TIC is shown. Calculated monoisotopic mass = 1785.90 Da, Observed monoisotopic mass = 1786.05Da

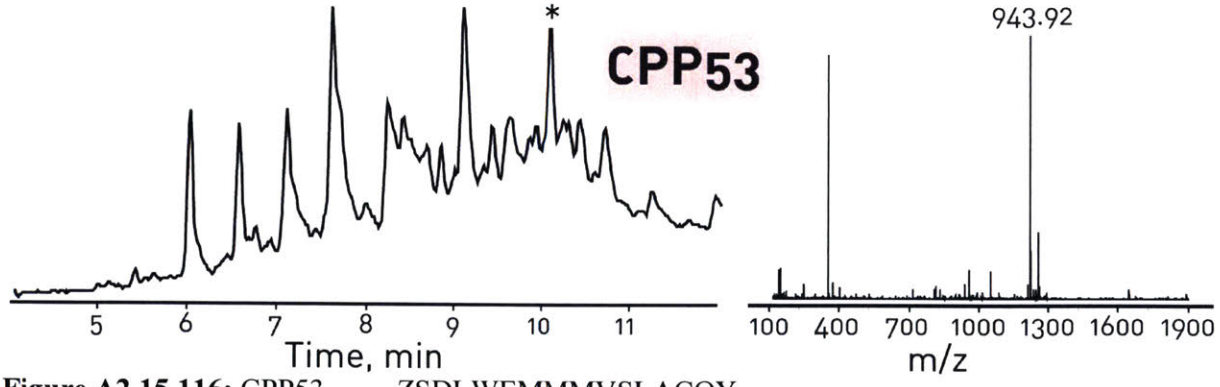


Figure A2.15.116: CPP53 ZSDLWEMMMVSLACQY
TIC is shown. Calculated monoisotopic mass = 1884.76Da, Observed monoisotopic mass = 1884.82Da

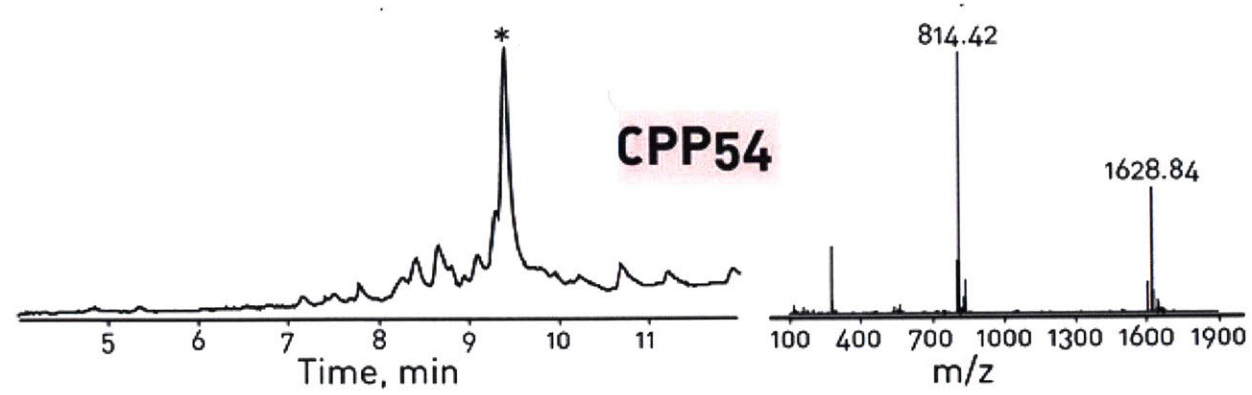


Figure A2.15.117: CPP54 ZVTWTPQAWFQWV
TIC is shown. Calculated monoisotopic mass = 1626.76Da, Observed monoisotopic mass = 1626.82Da

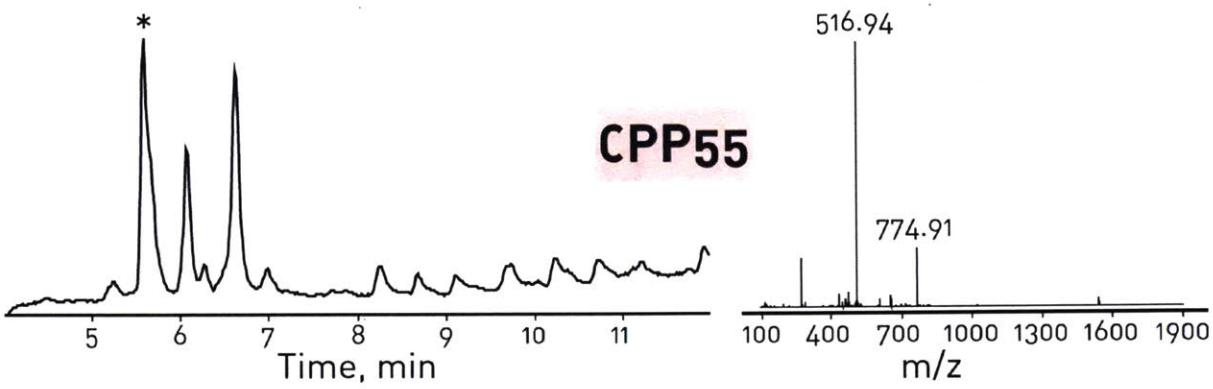


Figure A2.15.118: CPP55 ZGSPWGLQHHPPRT
TIC is shown. Calculated monoisotopic mass = 1547.74Da, Observed monoisotopic mass = 1547.80Da. Single and double histidine deletions are also observed at 6.1min and 6.6min, respectively.

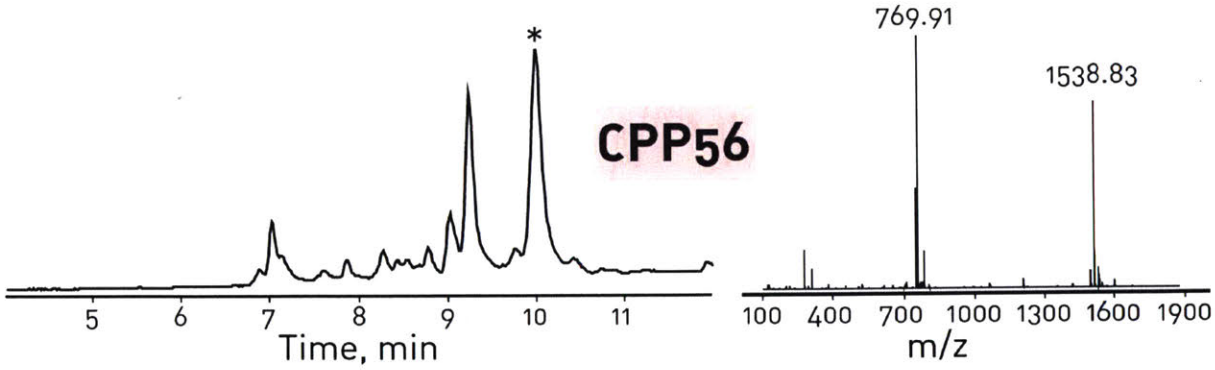


Figure A2.15.119: CPP56 ZGPFHFYQFLFPPV
TIC is shown. Calculated monoisotopic mass = 1673.80Da, Observed monoisotopic mass = 1536.81Da (Failed – the primary product is the histidine deletion product). The desired product is observed at 9.2min.

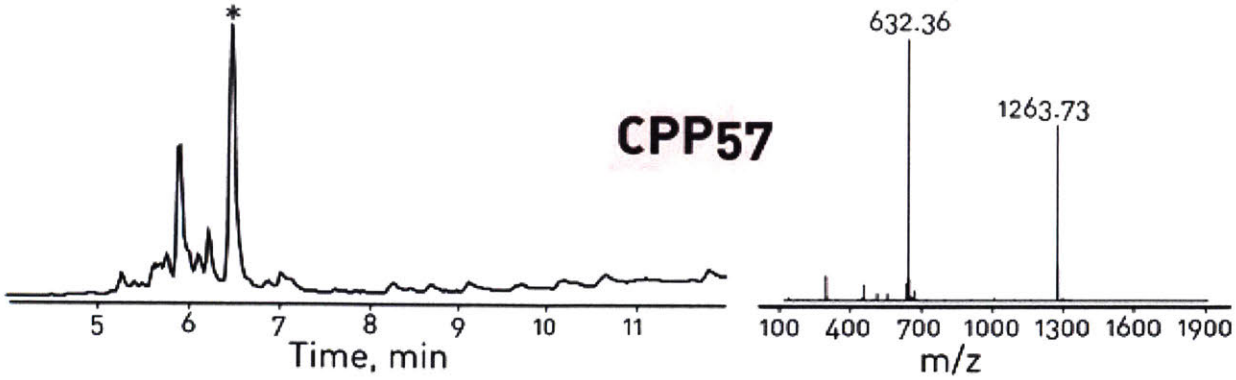


Figure A2.15.120: CPP57
TIC is shown. Calculated monoisotopic mass = 1399.72Da, Observed monoisotopic mass = 1262.73Da (Failed – the primary product is the histidine deletion product). The desired product is observed at 5.9min.

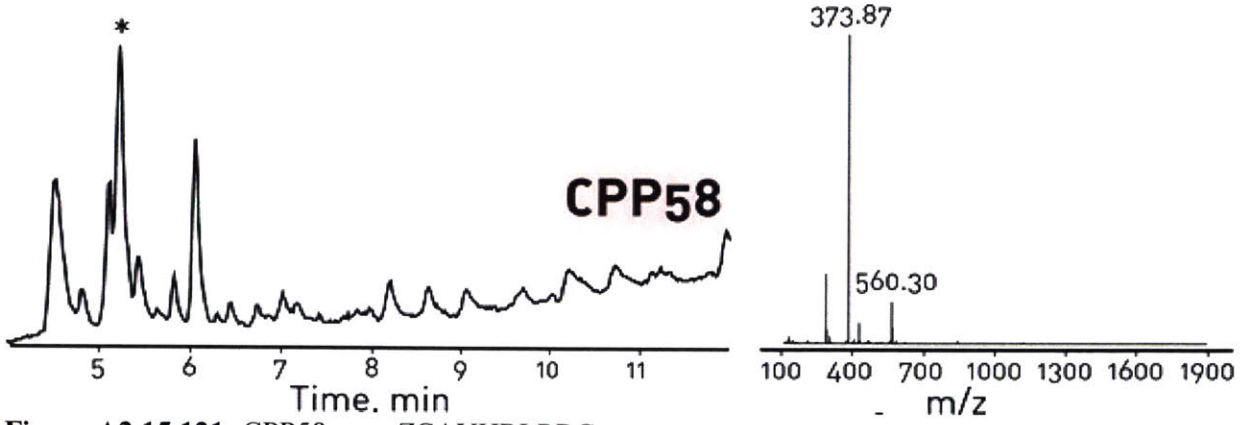


Figure A2.15.121: CPP58 ZCAYHRLRRC
TIC is shown. Calculated monoisotopic mass = 1255.58Da, Observed monoisotopic mass = 1118.57Da (Failed – the primary product is the histidine deletion product). The desired product is observed at 4.6min.

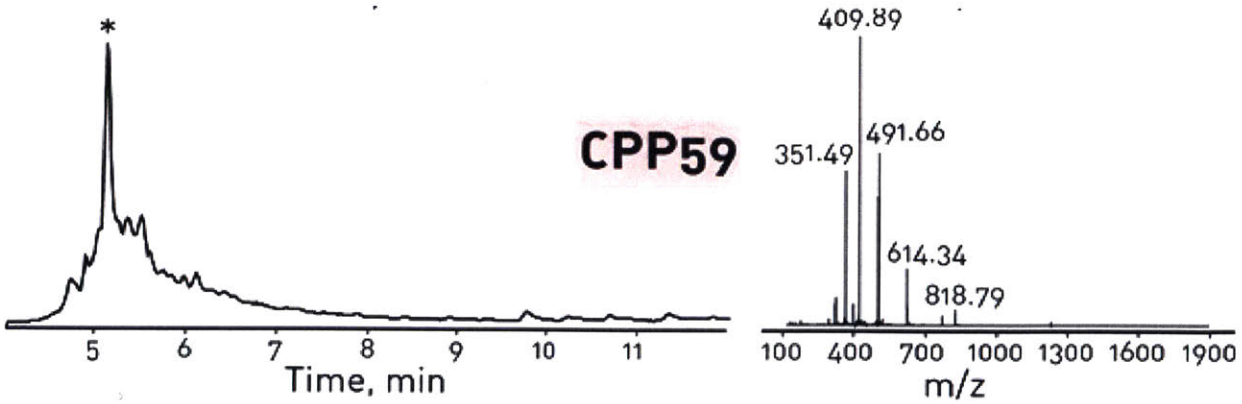


Figure A2.15.122: CPP59 ZRCGRASRCRVRWMRRRRI
TIC is shown. Calculated monoisotopic mass = 2452.31Da, Observed monoisotopic mass = 2452.34Da

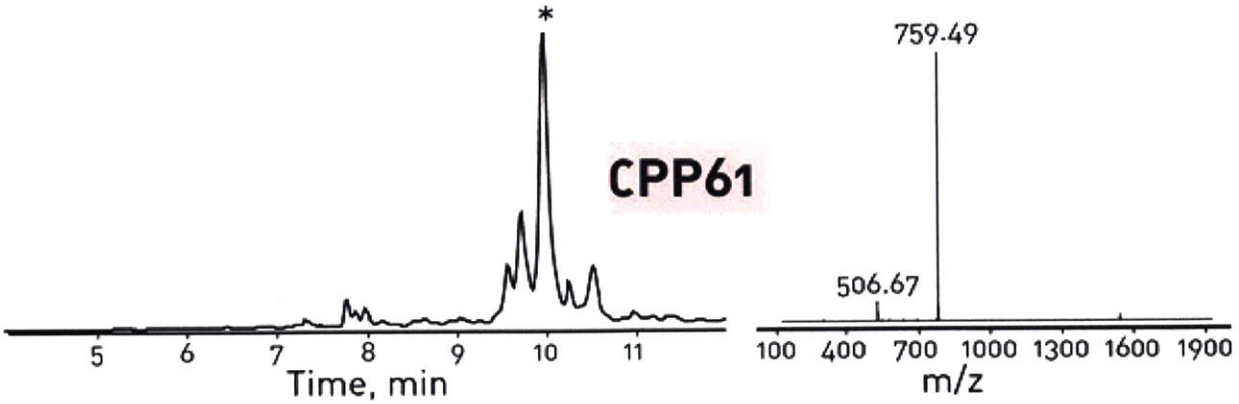


Figure A2.15.123: CPP61 ZPLILLRLLRGQF
TIC is shown. Calculated monoisotopic mass = 1516.92Da, Observed monoisotopic mass = 1516.96Da

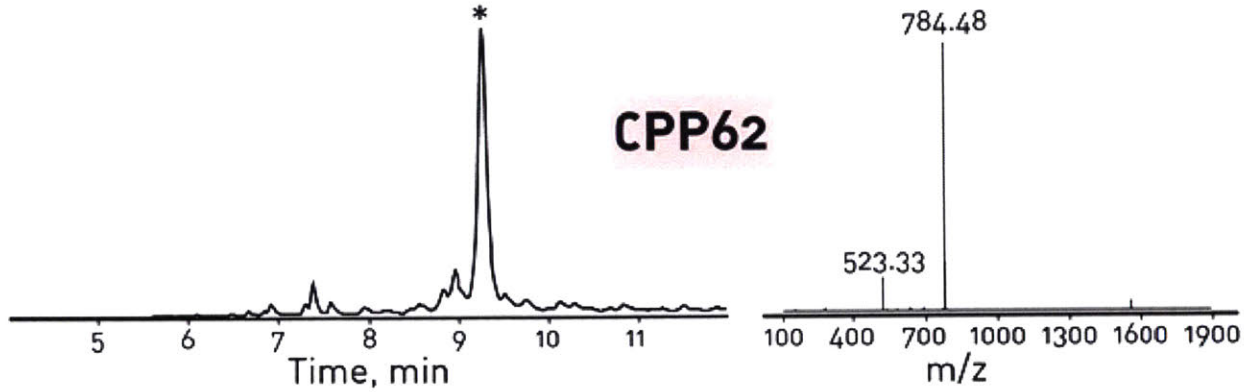


Figure A2.15.124: CPP62 ZPLIYLRLLRGQF
TIC is shown. Calculated monoisotopic mass = 1566.90Da, Observed monoisotopic mass = 1566.94Da

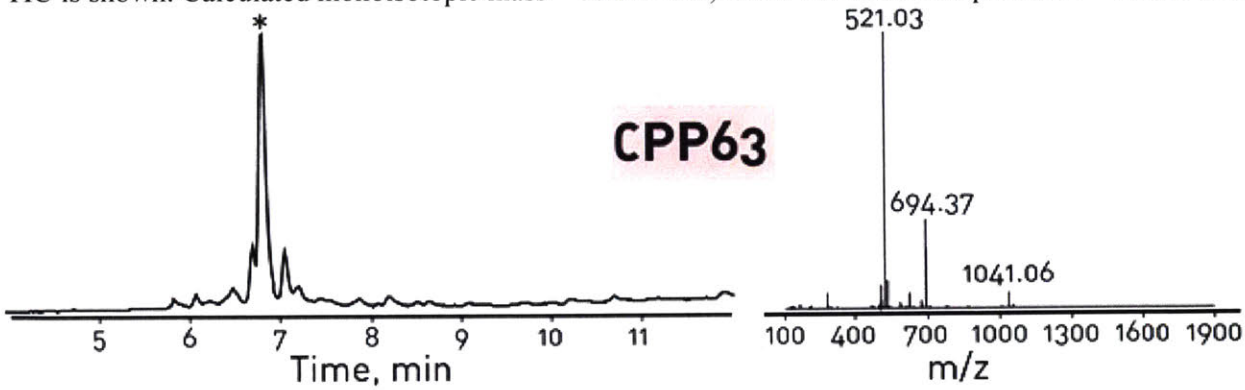


Figure A2.15.125: CPP63 ZKLWMRWYSPTTRRYG
TIC is shown. Calculated monoisotopic mass = 2079.02Da, Observed monoisotopic mass = 2079.07Da

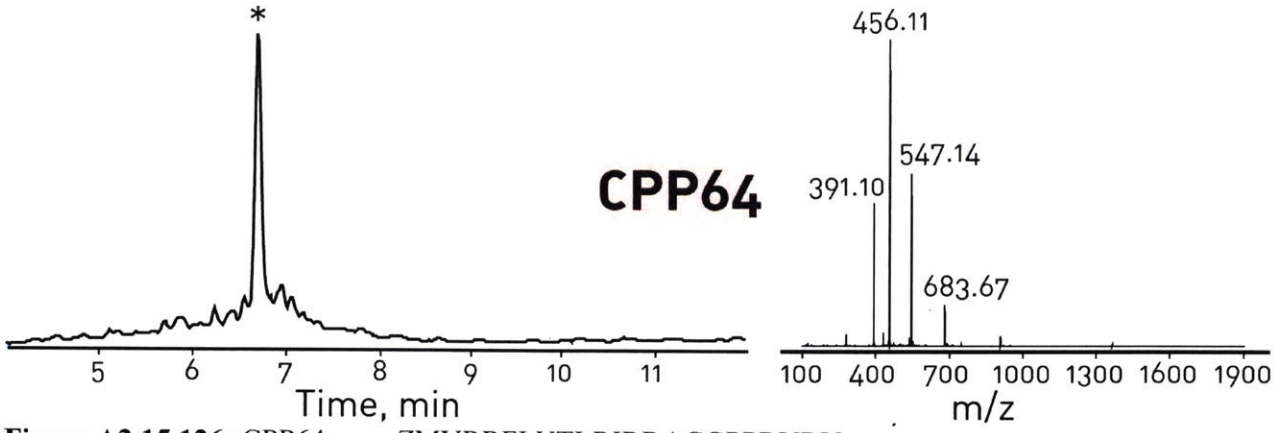


Figure A2.15.126: CPP64 ZMVRRFLVTLRIRRACGPPRVRV
TIC is shown. Calculated monoisotopic mass = 2729.58Da, Observed monoisotopic mass = 2729.64Da

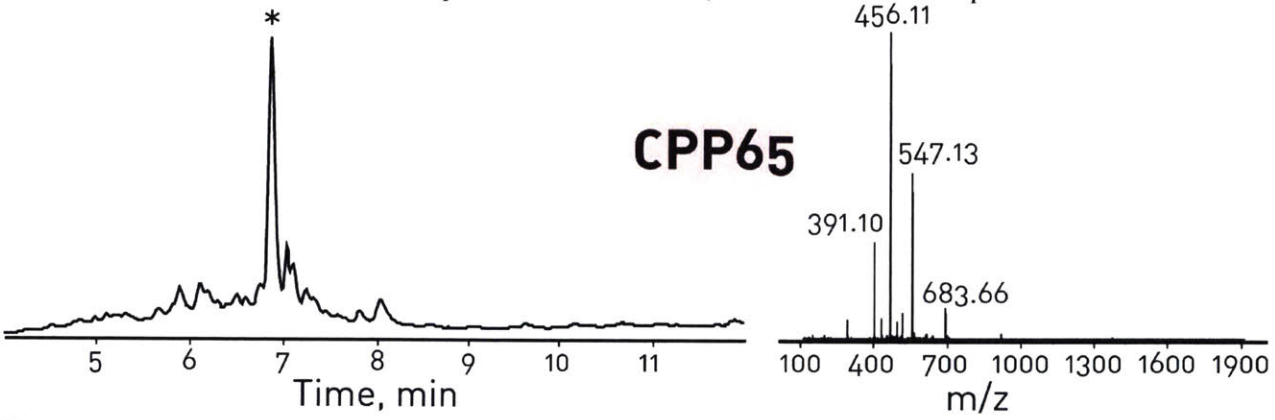


Figure A2.15.127: CPP65 ZMVTVLFRRLRIRRACGPPRVRV
TIC is shown. Calculated monoisotopic mass = 2729.58Da, Observed monoisotopic mass = 2729.61Da

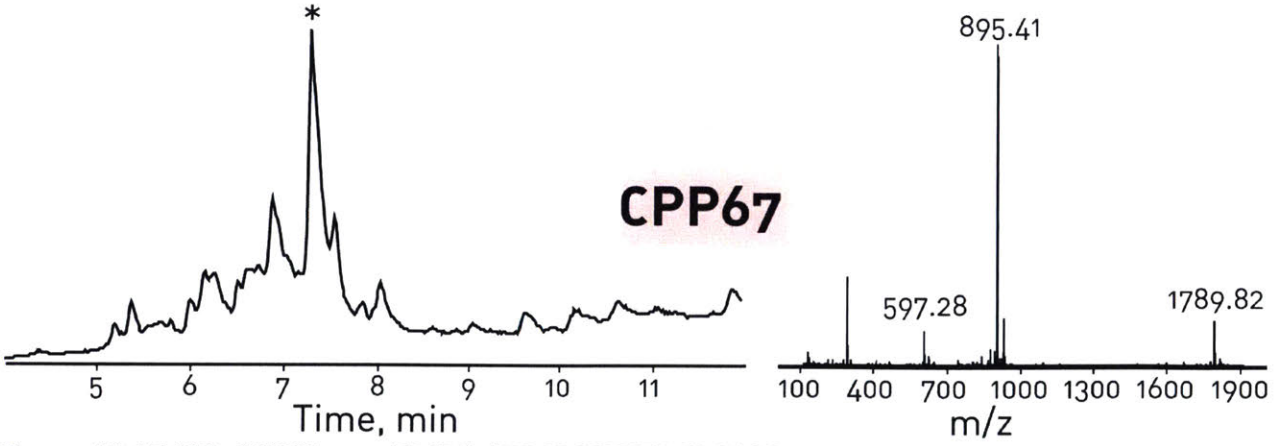


Figure A2.15.128: CPP67 ZLSTAADMQGVVTDGMASG
TIC is shown. Calculated monoisotopic mass = 1788.77Da, Observed monoisotopic mass = 1788.80Da

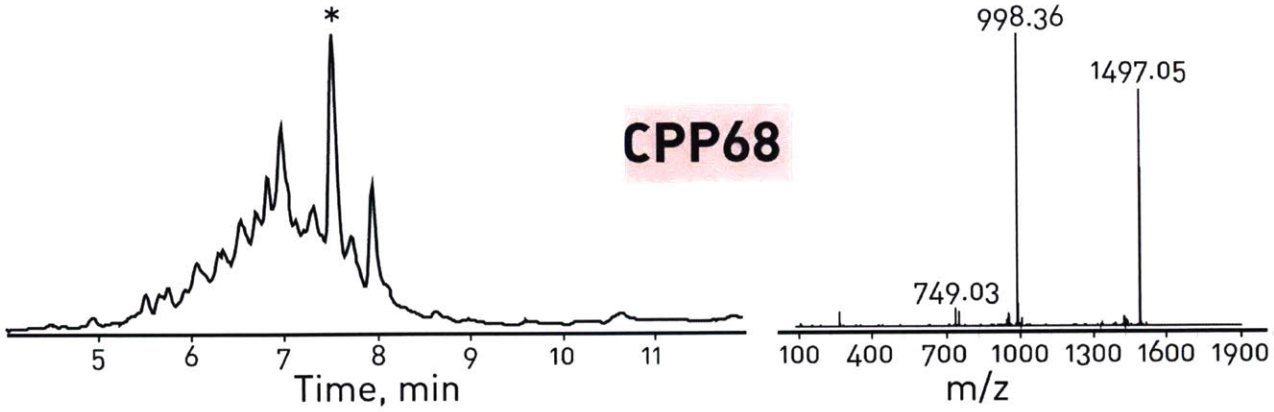


Figure A2.15.129: CPP68 ZLSTAADMQGVVTDGMASGLDKDYLKPDD
TIC is shown. Calculated monoisotopic mass = 2991.35Da, Observed monoisotopic mass = 2991.09Da

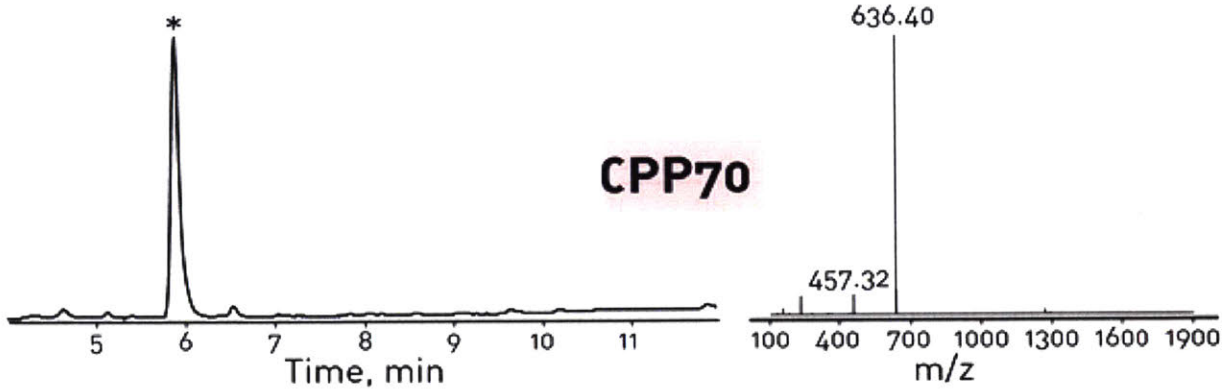


Figure A2.15.130: CPP70 ZVPTLK
TIC is shown. Calculated monoisotopic mass = 635.34Da, Observed monoisotopic mass = 635.39Da

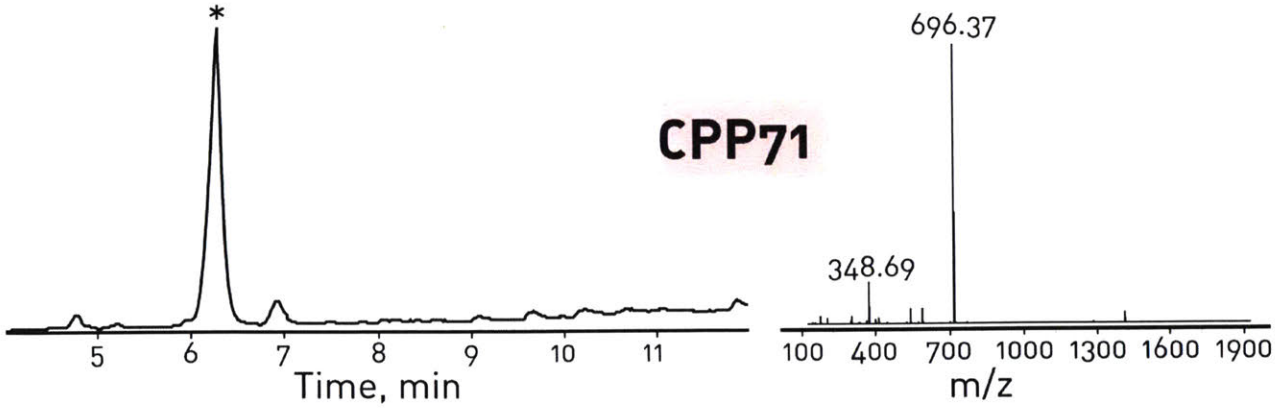


Figure A2.15.131: CPP71 ZPMLKE
TIC is shown. Calculated monoisotopic mass = 695.33Da, Observed monoisotopic mass = 695.36Da

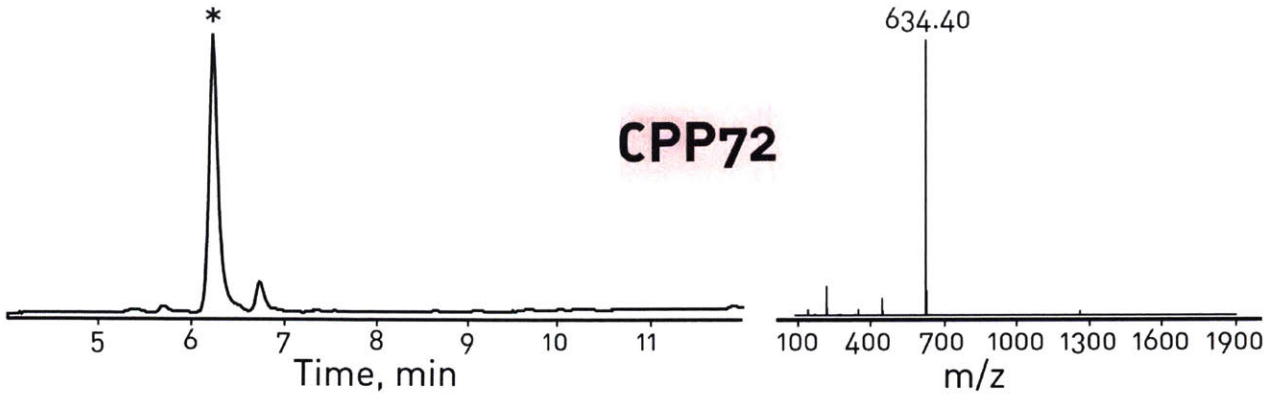


Figure A2.15.132: CPP72 ZVPALR
TIC is shown. Calculated monoisotopic mass = 633.36Da, Observed monoisotopic mass = 633.39Da

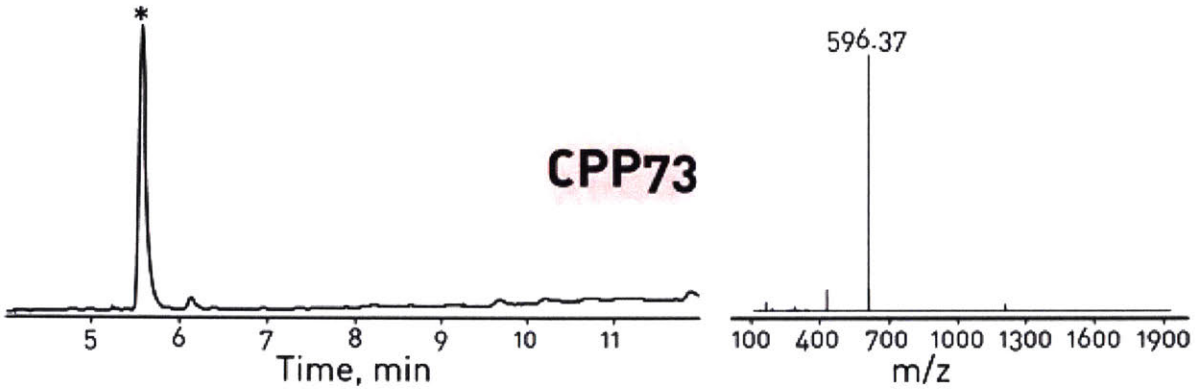


Figure A2.15.133: CPP73 ZVSALK
TIC is shown. Calculated monoisotopic mass = 595.33Da, Observed monoisotopic mass = 595.36Da

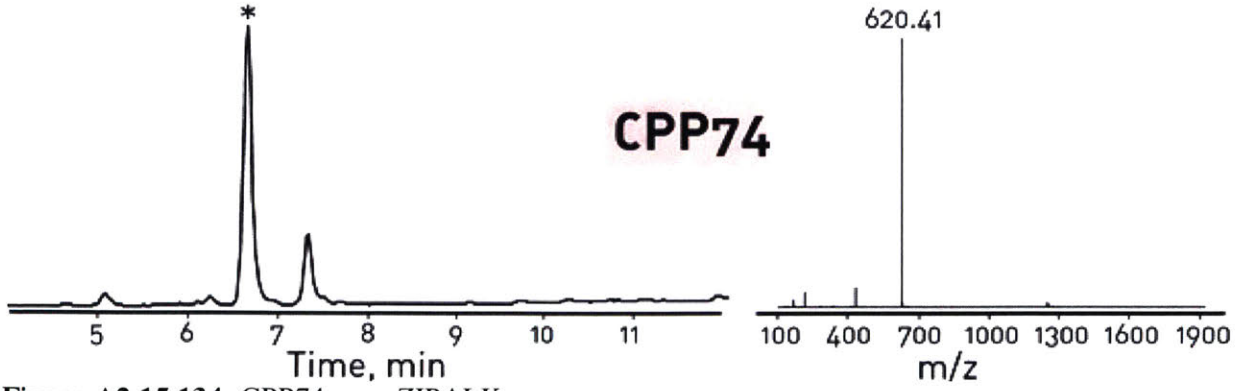


Figure A2.15.134: CPP74 ZIPALK
TIC is shown. Calculated monoisotopic mass = 619.37Da, Observed monoisotopic mass = 619.40Da

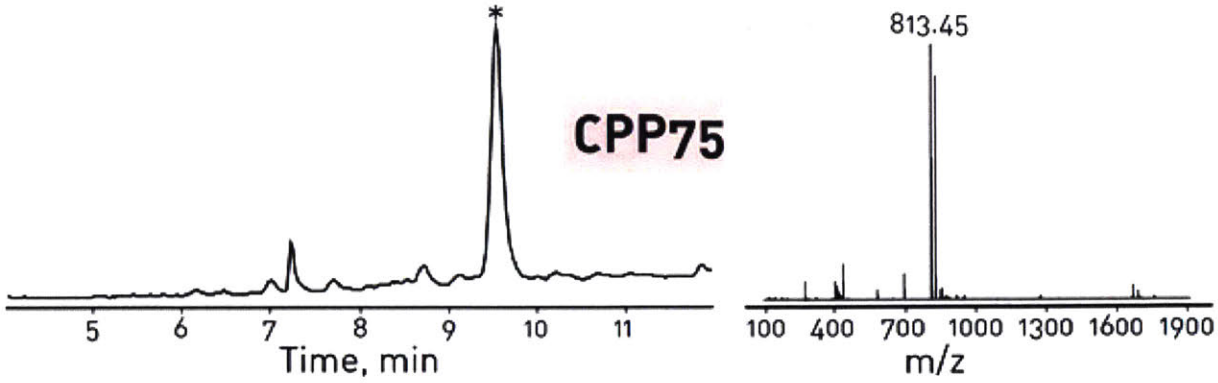


Figure A2.15.135: CPP75 ZPFVYLI
TIC is shown. Calculated monoisotopic mass = 829.44Da, Observed monoisotopic mass = 829.47Da (The most prominent peak in the MS is $[M+H^+-NH_3]^+$)

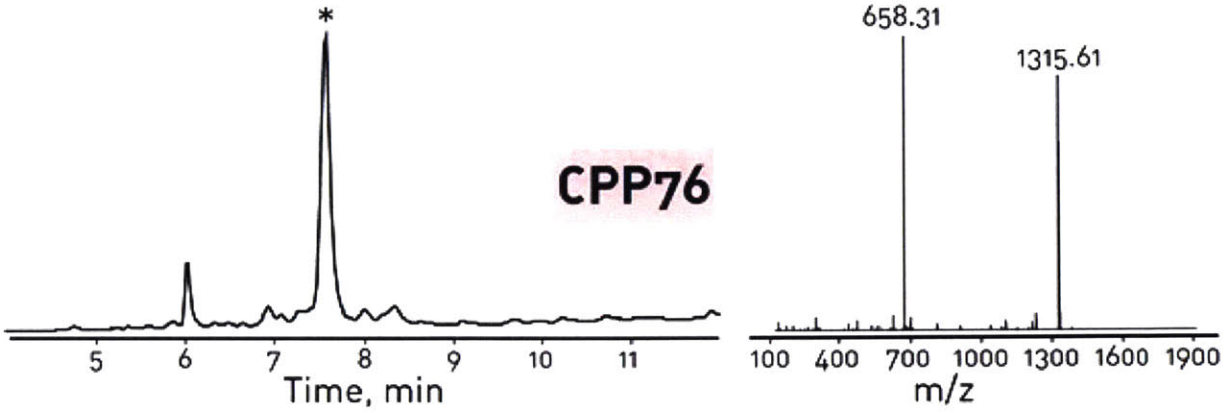


Figure A2.15.136: CPP76 ZPIEVCMYREP
TIC is shown. Calculated monoisotopic mass = 1314.57Da, Observed monoisotopic mass = 1314.60Da

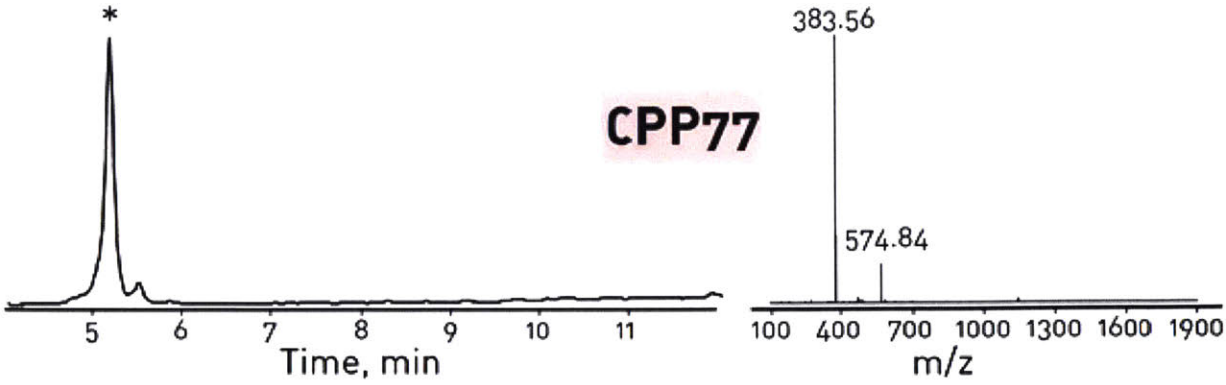


Figure A2.15.137: CPP77 ZPPRPPRPPR
TIC is shown. Calculated monoisotopic mass = 1147.64Da, Observed monoisotopic mass = 1147.65Da

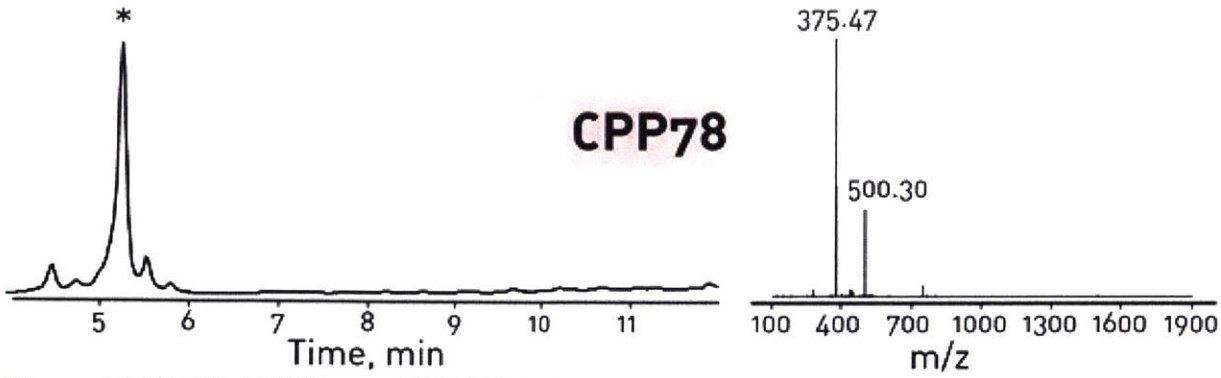


Figure A2.15.138: CPP78 ZPPRPPRPPRPPR
TIC is shown. Calculated monoisotopic mass = 1497.84Da, Observed monoisotopic mass = 1497.86Da

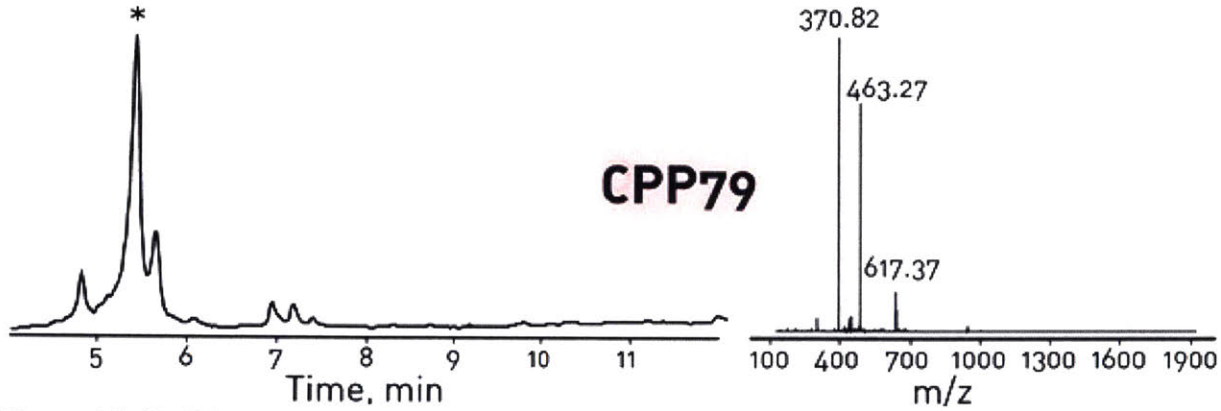


Figure A2.15.139: CPP79 ZPPRPPRPPRPPRPPR
TIC is shown. Calculated monoisotopic mass = 1848.05Da, Observed monoisotopic mass = 1848.09Da

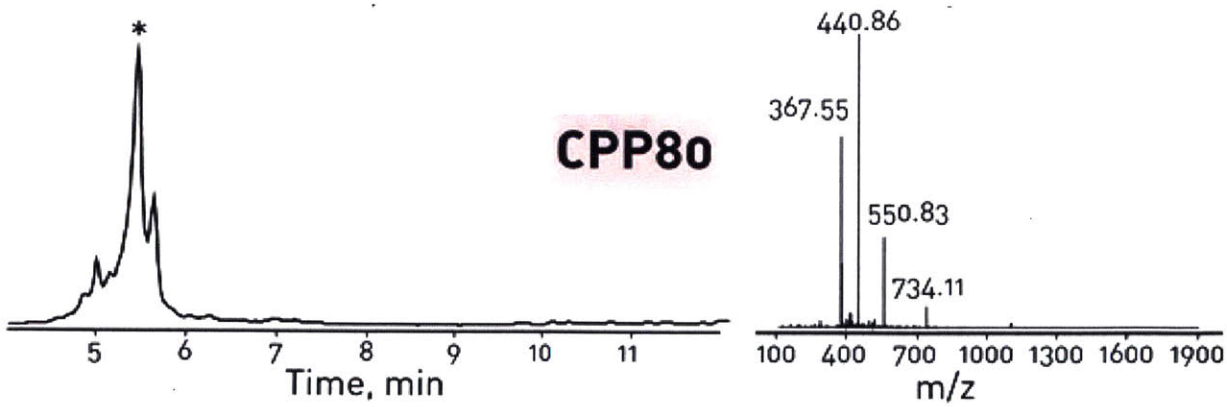


Figure A2.15.140: CPP80 ZPPRPPRPPRPPRPPRPPR
TIC is shown. Calculated monoisotopic mass = 2198.26Da, Observed monoisotopic mass = 2198.25Da

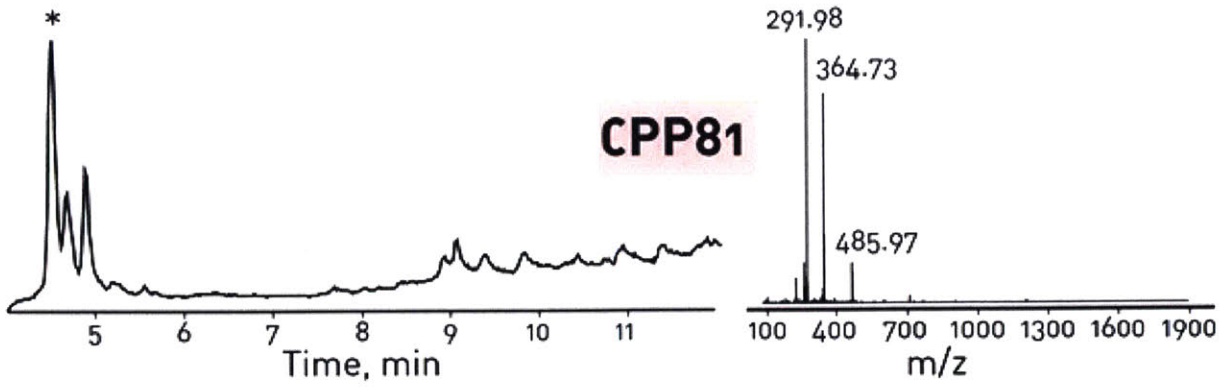


Figure A2.15.141: CPP81 ZRRIPNRRPRR
TIC is shown. Calculated monoisotopic mass = 1454.86Da, Observed monoisotopic mass = 1454.88Da

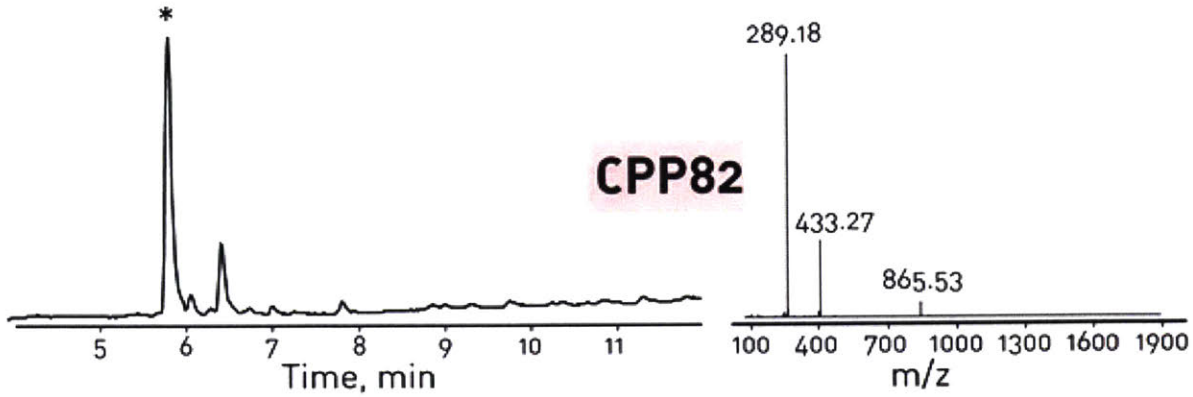


Figure A2.15.142: CPP82 ZRLRWR
TIC is shown. Calculated monoisotopic mass = 864.48Da, Observed monoisotopic mass = 864.50Da

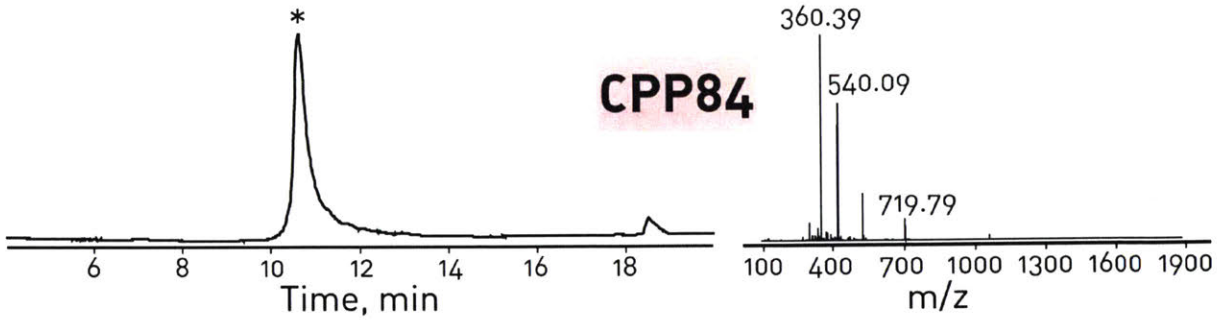


Figure A2.15.143: CPP84 ZGRPRESGKKRKRKRLKP
TIC is shown. Calculated monoisotopic mass = 2155.30Da, Observed monoisotopic mass = 2155.33Da

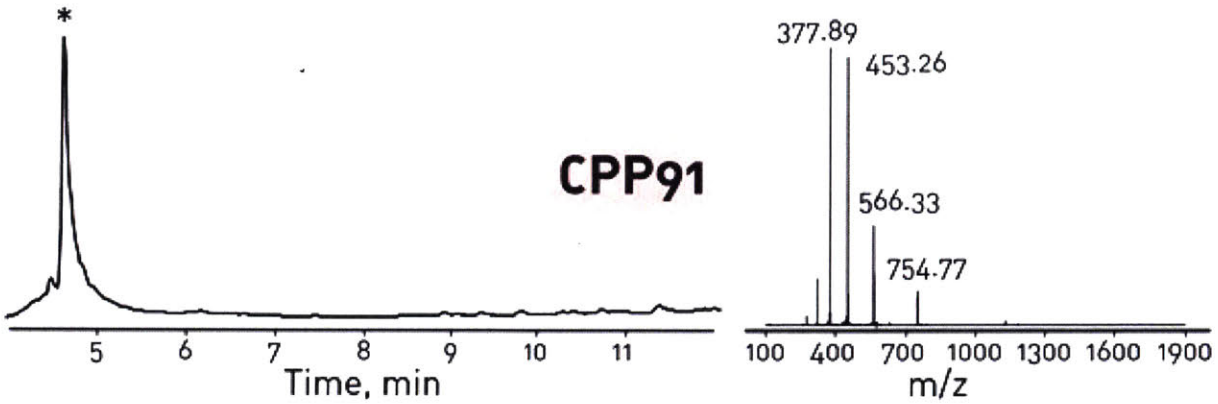


Figure A2.15.144: CPP91 ZLRRERQSRLRRERQSR
TIC is shown. Calculated monoisotopic mass = 2260.30Da, Observed monoisotopic mass = 2260.29Da

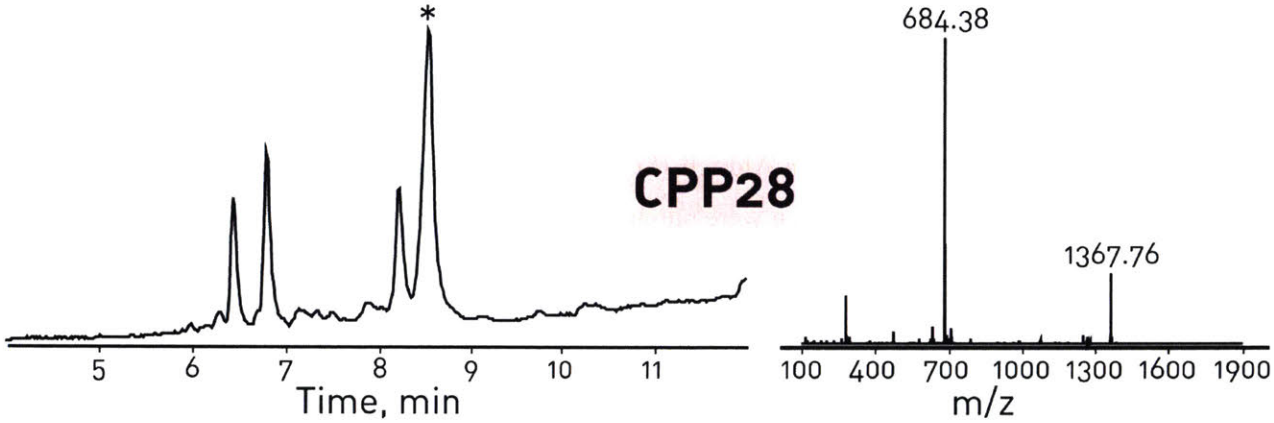
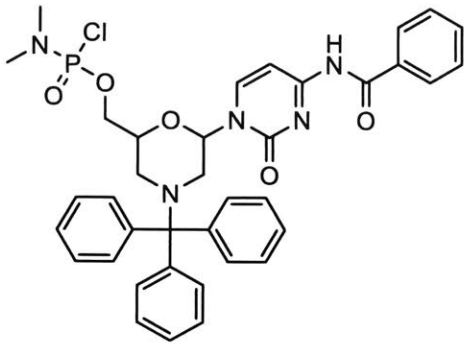


Figure A2.15.145: CPP28 ZPLSSIFSRIGDP
TIC is shown. Calculated monoisotopic mass = 1366.69Da, Observed monoisotopic mass =1366.74Da

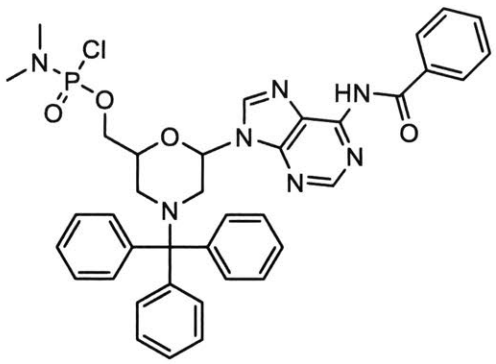
Appendix 3: Fast flow PMO synthesis

A3.1 Materials

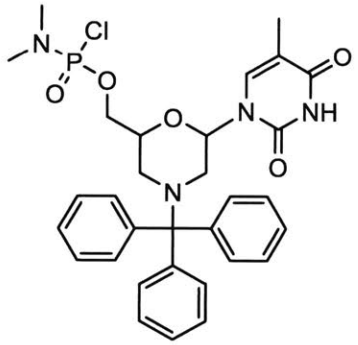
Activated morpholino subunits (moA, moC, moG, and moT - Figure A3.1), functionalized aminomethyl polystyrene resin (Figure A3.2), and functionalized polyethylene glycol linker ("Tail," Figure A3.3) were all provided by Sarepta Therapeutics. Aminomethyl polystyrene resin used in all studies had a loading of 0.39-0.43 mmol/g and was functionalized with the "Tail" PEG linker bearing a piperazine terminus, according to literature. 1,3-dimethyl-2-imidazolidinone (DMI), dichloroethane (DCE), dimethylformamide (DMF), N-methylpyrolidone (NMP), dicholormethane (DCM), and diisopropylamine (DIEA) were obtained anhydrous from Sigma-Aldrich (St. Louis, MO) and repurified on a Pure Process Technology (Nashua, NH) solvent purification system to remove water and amines. HPLC-grade Acetonitrile was purchased from VWR International (Philadelphia, PA) and LC-MS grade acetonitrile was purchased from Sigma-Aldrich (St. Louis, MO). Econo-Pac chromatography columns and accessories for solid-phase extraction were purchased from Bio-Rad (Hercules, CA). Water for HPLC was purified to 18.2MΩ resistivity on a Millipore Milli-Q system. All other reagents and solvents were purchased from Sigma-Aldrich (St. Louis, MO) as the purest anhydrous grades available, and used without further purification.



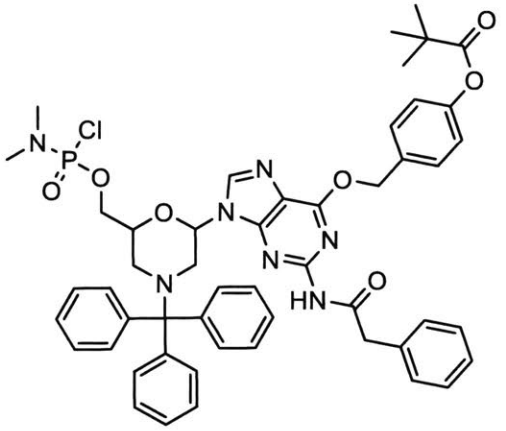
Activated, protected cytosine monomer
"moC"



Activated, protected adenine monomer
"moA"

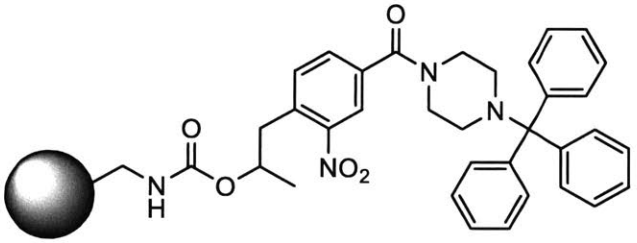


Activated, protected thymine monomer
"moT"



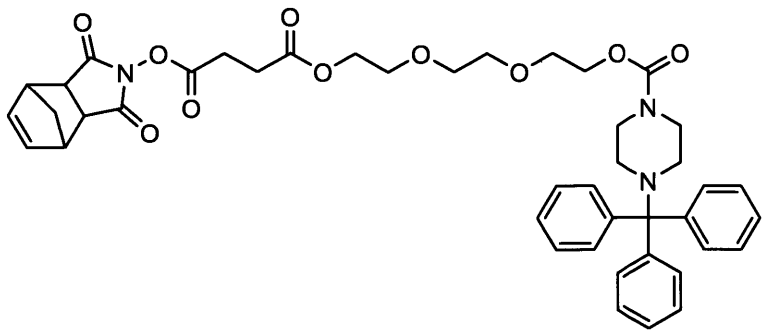
Activated, protected guanine monomer
"moG"

Figure A3.1: Structures of morpholino subunits



Functionalized, protected polystyrene resin
0.39-0.43 mmol/g

Figure A3.2: Structure of functionalized polystyrene resin.



Activated, protected polyethylene glycol linker
"Tail"

Figure A3.3: Structure of the activated polyethylene glycol "Tail"

A3.2 LC-MS analysis

All PMOs were analyzed on an Agilent 1260 Infinity Quaternary HPLC coupled to an Agilent 6520 ESI-Q-TOF mass spectrometer. The solvent mixtures used for LC-MS chromatography are: A = water + 0.1% formic acid (LC-MS-grade), B = acetonitrile + 0.1% formic acid (LC-MS-grade). The following conditions were used for PMO analysis. **Condition 1:** Column: Zorbax 300-SB C_3 (5 μm, 2.1 x 150 mm); Flow Rate: 0.8 mL/min; Gradient: 0-2 min 1% B, 2-8 min 1-30% B, 8-13 min 30-95% B, 13-14 min 95% B. **Condition 2:** Column: Zorbax 300-SB C_3 (5 μm, 2.1 x 150 mm); Flow Rate: 0.8 mL/min; Gradient: 0-3 min 5% B, 3-17 min 5-95% B, 17-18 min 95% B. **Condition 3:** Column: Zorbax 300-SB C_3 (5 μm, 2.1 x 150 mm); Flow Rate: 0.8 mL/min; Gradient: 0-2 min 1% B, 2-35 min 1-100% B. **Condition 4:** Column: Zorbax 300-SB C_3 (5 μm, 2.1 x 150 mm); Flow Rate: 0.8 mL/min; Gradient: 0-2 min 1% B, 2-20 min 1-55% B, 20-22 min 55% B. Total ion current (TIC) is displayed for all spectra shown in both the manuscript and supporting information.

A3.3 Quantification of tetramer synthesis

Tetramer synthesis products were poorly resolved on RP-HPLC, so they were quantified from the mass spectra using Agilent's Molecular Feature Extraction utility. For each species quantified, the abundances of the following identified compounds were summed: [M], [M+14], [M-Trt], and [M-Trt+14]. 3' Trityl protection was lost in the ESI source, and cleavage in methylamine resulted in ammonia to methylamine exchange in cytosine (+14Da, not present in C-deletion product).

A3.4 PMO synthesis reagents

The following is a list of solution identities and their composition, which were used throughout the investigation of PMO synthesis:

Detritylating solution

100mM <pyridine> trifluoroacetate in 4:1 (v/v) <solvent>/trifluoroethanol (TFE) + 1% (v/v) ethanol. The solution was made by first mixing the <solvent> and TFE, followed by sequential addition of <pyridine>, trifluoroacetic acid (TFA), and ethanol. After addition, the reagent solution was stirred for 30 minutes to dissolve solid pyridine derivatives.

<pyridine> = 4-cyanopyridine, ethyl nicotinate, pyridine, 3,5-lutidine, or 2,4,6-collidine

<solvent> = dichloroethane (DCE) or dichloromethane (DCM)

Neutralization solution

5% <base> in 3:1 <solvent>/isopropanol (*i*PrOH)

<base> = diisopropylethylamine (DIEA) or *N*-ethylmorpholine (NEM)

<solvent> = DCE or DCM

Coupling solution

0.2 M morpholino subunit, 0.4 M <base>, and 0.21 M LiBr (or other additive, when applicable) in dry DMI. The solution was made by first dissolving the subunit and LiBr in DMI. Base was added immediately prior to coupling. In all experiments prior to preparation of the 10 residue PMO, 0.5 mmol of monomer were used. For the 10 residue PMO experiments, 1.0 mmol of monomer was used. Note: Sonication can accelerate slow dissolution of anhydrous lithium bromide pellets in DMI.

<base> = DIEA or NEM

A3.5 Small scale batch synthesis for solvent screening

High boiling solvents were screened using the following synthetic cycle at room temperature on 10 mg of tail loaded resin. Resin was swollen in DCM for 15 minutes. 4-Cyanopyridinium trifluoroacetate was used to detritylate the resin for two minutes, repeated three times (6 minutes total). For coupling solvent studies, DCE was used as the detritylation solvent. The resin was then washed three times with DCM and the next monomer was coupled for 15 minutes. The coupling solution was prepared with 0.1 mmol of monomer and no additive, and the coupling solvent was always DMI in detritylation solvent studies. The resin was again washed three times with DCM in preparation for the next detritylation. The product was left Trityl protected.

A3.6 Batch synthesis for additive screening

Additives to increase the reactivity were screened using the following synthetic cycle at room temperature on 200 mg of tail loaded resin. Resin was swollen in DCE for 5 minutes. Cyanopyridinium trifluoroacetate was used to detritylate the resin for two minutes. DCE was used as the detritylation solvent. The resin was then washed three times with DCE and two times with DMI, and the next monomer was coupled for 15 minutes. The coupling solution was prepared with 0.5 mmol of monomer, and the coupling solvent was DMI. The resin was again washed three times with DCE in preparation for the next detritylation. The product was left Trityl protected.

A3.7 Eteplirsen(1-10) batch synthesis

The batch control Eteplirsen(1-10) was synthesized as follows according to reported procedures. 250 mg (0.1 mmol) of tail loaded resin was swollen in NMP for 15 minutes. The resin was then washed twice with DCM and detritylated five times for two minutes each time (10 minutes total) using 4-cyanopyridine and DCM based detritylation reagent. Following detritylation, the resin was washed twice with DCM and neutralized with a DIEA/DCM neutralization solution. Two more DCM washes and two DMI washes prepared the resin for a 90 minute coupling with 0.5 mmol (5 eq.) of monomer with LiBr and DIEA. Two DCM washes completed the synthetic cycle prior to the next detritylation.

A3.8 Cleavage

Completed resin bound PMOs were washed with DCM and dried under vacuum prior to cleavage. Early cleavages were performed in a 1:1 mixture of ammonia (sat. aq.) and methylamine (sat. aq.) at 65°C for 15 minutes. This cleavage procedure resulted in substantial substitution of methyl amine for ammonia

in cytosine residues, and later cleavages were performed in a 4:1 mixture of ethanol:ammonia (sat. aq.) at 55°C for 16 hours.

A3.9 Tetramer stability

Resin bound model tetramer, 5'-Tail-ACGT-Trt-3' was found to be unstable in acid and when detritylated but not neutralized (i.e. with protonated 3' morpholine). When neutralized, the detritylated resin bound tetramer was stable. (Fig. A3.4)

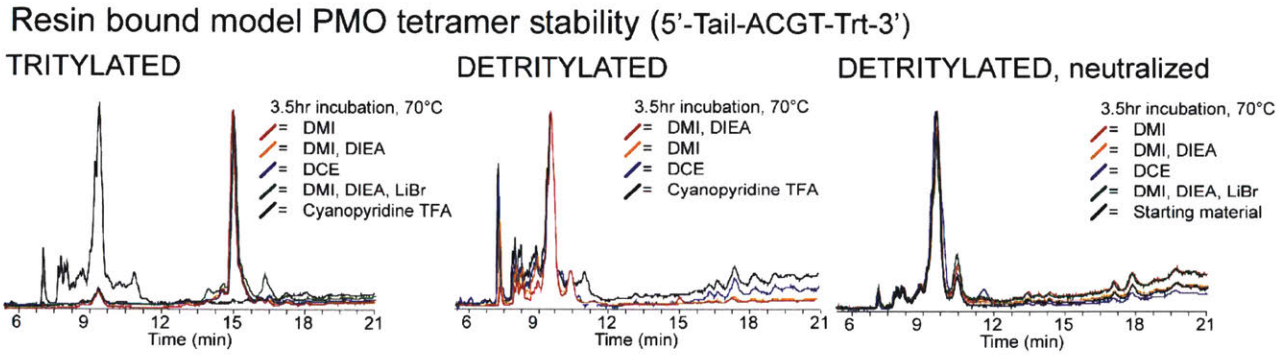


Figure A3.4: Stability of the resin bound tetramer 5'-Tail-ACGT-3' when tritylated (Left), detritylated with a protonated 3' amine (Center), and when detritylated with a neutralized 3' amine (Right). Note the drastic improvement in stability upon neutralization.

A3.10 Monomer stability at 90°C

Monomer stability was examined at 90°C by heating coupling solutions of 0.2 M moA and 0.4 M DIEA in DMI in sealed vials for 5 minutes with or without LiBr. After 5 minutes, coupling solutions were quenched with a 10% (v/v) solution of piperidine in NMP and analyzed by LC-MS (Fig. A3.5)

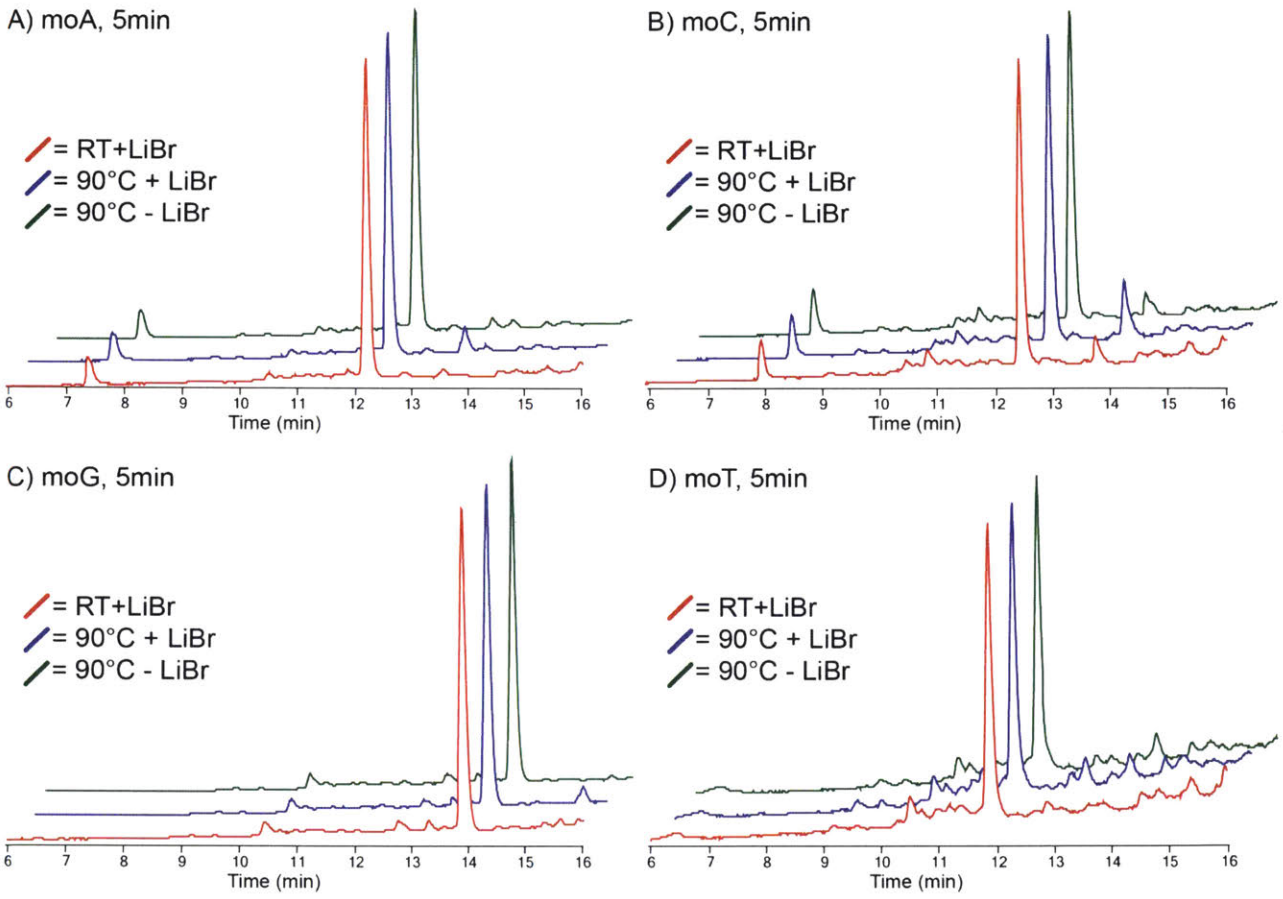


Figure A3.5: moA, moC, moG, and moT were all stable for 5 minutes at 90°C in coupling solution with or without LiBr.

A3.11 Design of Flow Synthesizer

The flow PMO synthesizer is shown in Figure A3.6. Reagent reservoirs were GL45 threaded media bottles equipped with a top to maintain a positive pressure of dry nitrogen and allow anhydrous transfer of solvents (Fig A3.7). Each reagent reservoir was connected to one of four selectable ports on a four positon Swagelok SS-43ZFS2 manual selector valve. The common port was connected to the HPLC pump via a Swagelok (SS-QM2-B-200KR and SS-QM2-S-200) shutoff quick connect which could be disconnected when not in use to prevent siphoning of the pressurized solvents. The HPLC pump was a Varian 210 with a 25ml/min stainless steel pump head. Low pressure (inlet side) tubing was 1/8" OD, 1/16" ID PFA. The HPLC pump outlet was connected to a 40psi back pressure regulator (P-785) and male leur lock quick connect (Idex P-655) using 1/16" OD, 0.030" ID PFA tubing. When the HPLC pump was in use, this quick connect was mated to the female leur lock quick connect on the reactor inlet line. When the HPLC pump was not in use, the reactor inlet line was attached to a syringe of coupling reagent on the syringe pump (Harvard Apparatus PhD 3000).

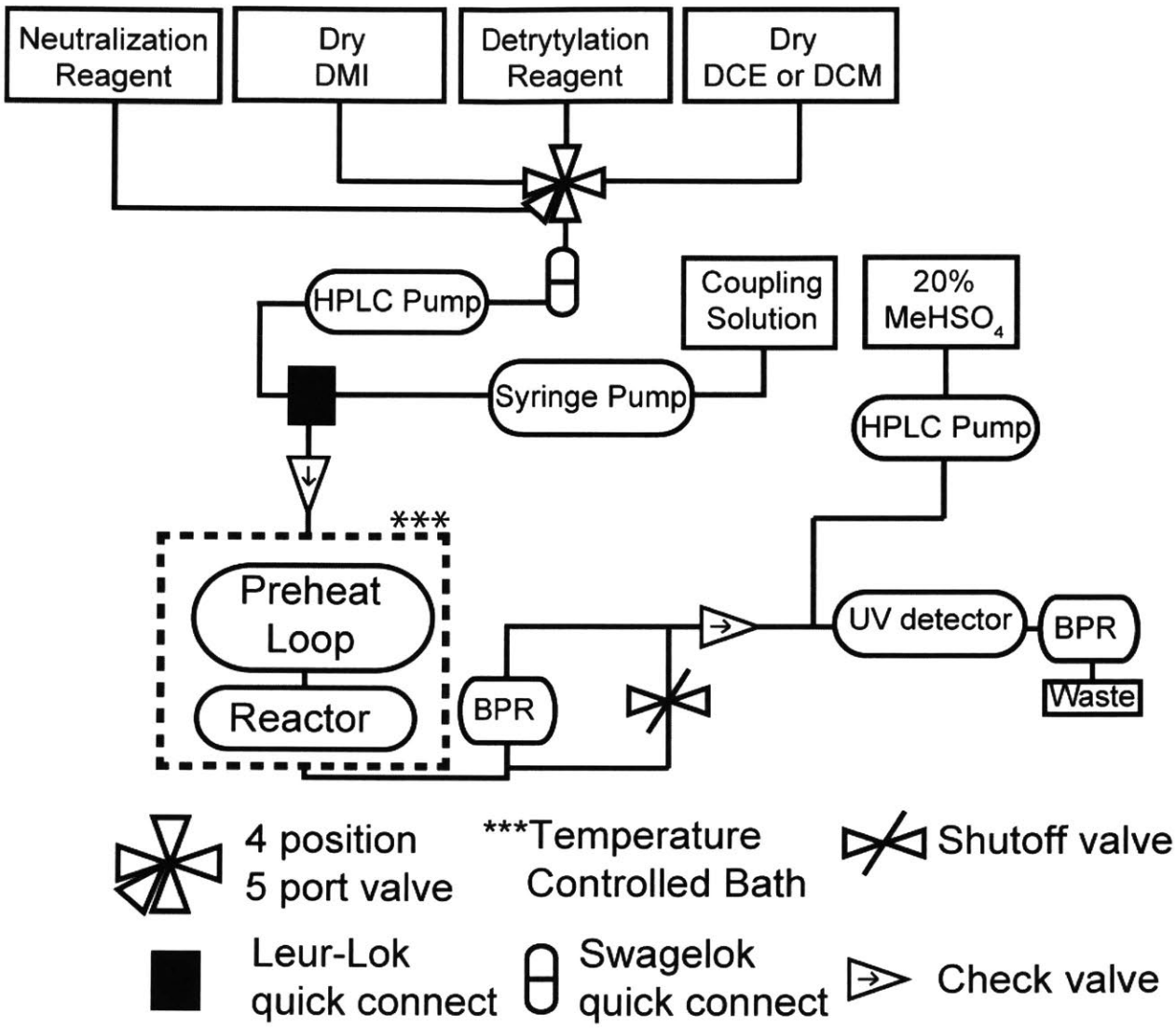


Figure A3.6: Complete schematic of the flow PMO synthesizer.

The reactor inlet line consisted of a female luer lock quick connect (Idex P-658) and check valve (Idex CV-3316) joined with about 18 inches of 1/16" OD PFA tubing to a 5 foot stainless steel preheat loop (1/16" OD, 0.030" ID, Idex U-107) via a Swagelok union (SS-100-6). This preheat loop was connect to a reactor identical to the one described previously for peptide synthesis. The reactor outlet was connected to a 250psi back pressure regulator (Idex P-788), which could be bypassed by opening a bypass valve (Idex P-733). The bypass fluid path rejoined the outlet of the back pressure regulator in a T (P-632), passed through a check valve (Idex P-788), and six inches of tubing before arriving at the methane sulfonic acid T (Swagelok SS-100-3). At this point, 20% MeHSO4 in DCM was optionally infused with a Knauer smartline HPLC pump with a 50mL/min titanium head to regenerate the Trityl cation for UV monitoring of deprotection. The mixed fluid was then passed through a UV detector (Agilent G1315D), a 20psi back pressure regulator (Idex P-791), and to waste.

Where not described, tubing downstream of the HPLC pump and syringe was 1/16" OD, 0.030" ID PFA. All ¼-28 flat bottom fittings were Idex super flangeless (XP-131 and XP-141).

WARNING: $MeHSO_4$ is extremely corrosive. It will degrade PEEK, giving a yellow compound with absorbance overlapping the Trityl cation. Further, the Knauer pump seals were eventually destroyed, either by DCM or the acid (or both).

Reagent reservoirs

Reagent reservoirs were provisioned with machined adaptors for use with GL-45 caps with holes (Chemglass). These adaptors had three ¼-28 threaded ports. Two ports had 1/8" thru holes; the third had a 1/16" thru hole. The first hole was used for a reagent withdraw line. The second was used as a fill port, and the third was used for the nitrogen gas supply (4psi). The reagent withdraw line was a 1/8" line inserted into the bottom of the reagent reservoir and sealed in place with a super flangeless fitting (Idex XP-131). Nitrogen gas was supplied by a 1/8" line seated against the 1/16" thru hole and sealed in with a super flangeless fitting. The fill port consisted of a thin wall 1/8" OD, 0.1" ID, stainless steel tube inserted to just below the bottom of the machined adaptor and sealed in place with a super flangeless fitting. The free side was fitted with a shut off quick connect (Swagelok SS-QM2-B-200KR) that was sealed when not in use but could be used to fill the reservoir from a second reservoir of anhydrous solvent under slightly higher pressure. The second reservoirs were of a similar design and 10psi argon was used to transfer dry solvent. During filling, the nitrogen system was allowed to vent through a 5psi back pressure regulator and oil bubbler.

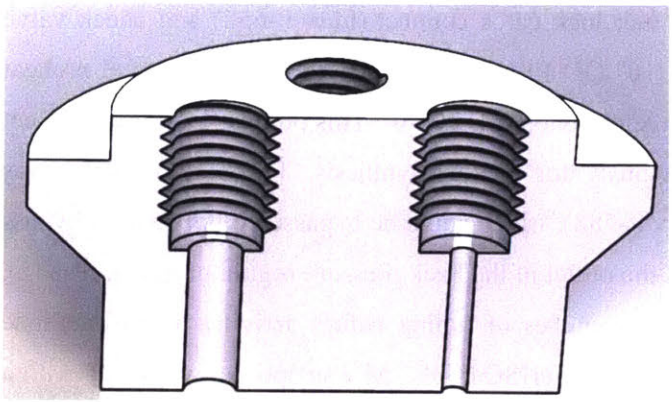


Figure A3.7: Rendering of sectioned reagent reservoir top showing three ports (two are sectioned, one is in the back).

A3.12 General procedure for flow synthesis

The following procedure was used for flow synthesis. Resin was loaded into the reactor, the reactor was connected to the HPLC pump, and halogenated wash solvent was delivered at 10 mL/min to remove air. The flow was stopped and the resin was allowed to swell for 10 minutes. The flow protocol was initiated with an initial halogenated solvent wash at 10 mL/min for 60 seconds. Detritylation was performed with 100 mM of a pyridine trifluoroacetate for 120 seconds at the same flow rate. After a 30 second halogenated solvent wash, neutralization was performed with 5% DIEA or NEM for 60 seconds. The resin was then washed with the halogenated solvent and DMI for 60 seconds each at 10 mL/min. The HPLC pump was then halted for the coupling step. Coupling solution (0.2 M subunit, 0.4 M DIEA or NEM, and 0.21 M LiBr in dry DMI) was placed in a 10 mL syringe and delivered via syringe pump at 3 mL/min over 1 minute (0.5 mmol monomer) or 2 minutes (1 mmol monomer). When all of the solution was delivered, the HPLC pump delivered DMI at 3 mL/min for three minutes. This protocol was repeated for each residue until synthesis was complete. The finished resin was removed from the reactor, washed 5 times with DCM in a fritted syringe (Torviq), and dried under vacuum. Cleavage was performed as above.

For steps at elevated temperature, the reactor and preheat loop were placed in a thermostated water bath.

For HPLC purification, resin was cleaved under mild conditions (4:1 ethanol/ammonium hydroxide) and then isolated by solid-phase extraction (SPE). Initially, the cleaved resin was filtered and washed 4 times with methanol. The cleavage solution and washes were collected and concentrated to dryness using a rotary evaporator. The residue was dissolved in 10 mL of Milli-Q water prior to SPE. Separately, SPE columns were prepared and conditioned using the following procedure. 20 mL Econo-Pac columns from Bio-Rad were charged with 3-4 mL of Amberchrome CG-300M resin and sealed with a frit. Then 8 mL of the following solutions were added to the column, in order, and drained before adding the next solution: 80% ACN in 1% NH_4OH, 0.5 M NaOH in 20% EtOH, Milli-Q water, 50 mM H_3PO_4 in 80% ACN, Milli-Q water, 0.5 M NaOH in 20% EtOH, Milli-Q water, 1% NH_4OH. When conditioning was complete, the column was stored in 8 mL of 1% NH_4OH at room temperature until used. The column was rinsed two times with 12 mL of Milli-Q water before loading the PMO onto the column. Then, the column was rinsed once with 3 mL of 1 M NaCl, followed by three rinses with 12 mL of Milli-Q water, and once with 3mL of 10% acetonitrile in water. The PMO was then eluted with two 3mL rinses of 50% acetonitrile in water. The eluent from the 50% acetonitrile wash was collected into a pre-weighed 50 mL conical centrifuge tube and lyophilized to afford the crude PMO as a white powder suitable for LC/MS analysis and/or preparative HPLC purification.

Several PMOs were purified on a mass-directed purification system consisting of an Agilent 1260 Infinity Quaternary HPLC coupled to an Agilent 6130 single quadrupole mass spectrometer. The solvent mixtures used for purification were as follows: A = 5 mM NH4OAc (pH =8), B = 90% acetonitrile + 10% 5 mM NH4OAc (pH =8). Purification was performed using the following conditions: Column: Zorbax 300-SB C3 (5 μm, 21.2 x 100 mm); Flow Rate: 20 mL/min; Gradient: 0-2 min 2% B, 2-60 min 2-60% B, 60-70 min 75% B.

Appendix 4: D-Scan of two small proteins

A4.1 Materials:

2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 2-(7-Aza-1Hbenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU), and Nα-Fmoc and side chain protected D-amino acids were from Chem-Impex International, IL. Nα-Fmoc and side chain protected L-amino acids were from Advanced ChemTech, KY. Side chain protection was as follows: Arg(Pbf), Asn(Trt), Asp(OtBu), Cys(Trt), Gln(Trt), Glu(OtBu), His(Trt), Lys(Boc), Ser(tBu), Thr(tBu), Trp(Boc), Tyr(tBu).

Rink amide-ChemMatrix polyethylene glycol resin was purchased from Matrix Innovation (35-100 mesh, 0.49mmol/g). Aminomethyl polystyrene resin was prepared from BioRad SX-1 beads as described previously (Mong 2014).

N,N-Dimethylformamide (DMF), dichloromethane (DCM), diethyl ether,

and HPLC-grade acetonitrile were from EMD Millipore. Acetonitrile for LC-MS were purchased from Fluka. Water for LC-MS was purified with a Millipore water purifier. All other reagents were purchased from Sigma-Aldrich, MO. Solvents were not anhydrous, and all material was used without further purification.

Common solvent mixtures used throughout these experiments are: 0.1% (v/v) TFA in water (A), 0.1% (v/v) Formic acid in water (A'), 0.1% (v/v) TFA in acetonitrile (B), and 0.1% (v/v) formic acid in acetonitrile (B').

A4.2 LC-MS Analysis:

All peptides and proteins were analyzed on an Agilent 6520 Accurate Mass Q-TOF LC-MS under one of three conditions.

Condition 1: An Agilent C3 Zorbax SB column (2.1 x 150 mm, 5 μm packing) was used with a flow rate of 0.8mL/min of the following gradient: A' with 5% B' for 2 min, 5-61% B' ramping linearly over 9 min, and 65% B' for 1 min. This method was used to assess the purity of pooled fractions and crude peptides.

Condition 2: An Agilent C3 Zorbax SB column (2.1 x 150 mm, 5 μm packing) was used with a flow rate of 0.8mL/min of the following gradient: 5-22% B' in A' ramping linearly over 0.7 min, then ramping linearly to 45% B' over 3.5 min. This method was used to monitor folding reactions.

Condition 3: An Agilent C3 Zorbax SB column (2.1 x 150 mm, 5 μm packing) was used with a flow rate of 0.8mL/min of the following gradient: 5-28% B' in A' ramping linearly over 0.7 min, then ramping linearly to 45% B' over 2.8 min. This method was used to analyze individual fractions (data not shown).

A4.3 MALDI analysis:

MALDI-TOF mass spectra were acquired on a PerSeptive Biosystems Voyager-DE Pro Biospectrometry workstation. The matrix was saturated α-cyano-4-hydroxycinnamic acid in 50% A' / 50% B'.

A4.5 Peptide Synthesis:

To facilitate production and purification of many analogs of EETI-II, synthesis was optimized using our previously reported fast flow peptide synthesizer (Simon 2014). We synthesized wild type EETI-II under 8 conditions: at 70C or 90C, on PEG or PS resin, and with HBTU or HATU activation. Other conditions were as previously reported[1]. The synthesis at 70C on PEG resin with HATU activation was determined to be the best, and all analogs were synthesized under these conditions. Representative crude traces of WT-EETI-II and analogs D12 and A17 are shown below in Figure A4.1. In all cases, the major peak is the desired product. The prominent late eluting products have masses consistent with single t-butylations of the peptide.

Z33 and its analogs were synthesized at 90C on PEG resin with HATU activation using an improved version of our previously reported[1] fast flow peptide synthesizer (manuscript in preparation). Amide bond formation was effected in 8 seconds, and Fmoc groups were removed in 8 seconds with 20%(v/v) piperidine in DMF. Overall cycle times were about 40 seconds. Crude and purified peptides are shown in section A4.17 (Sub-Appendix 5).

Cleavage from resin and global side chain deprotection was effected by treatment with 25mL of 94% trifluoroacetic acid, 2.5% water, 2.5% ethane dithiol, and 1% triisopropyl silane per gram of peptidyl resin for 7 minutes at 60°C. The crude cleavage reaction was then diluted with 100mL of cold (-80C) diethyl ether per 25mL of cleavage cocktail to precipitate the crude peptide. The crude peptide was isolated by centrifugation, washed two times with additional cold ether, then dissolved in 50%B in A and lyophilized prior to purification or folding.

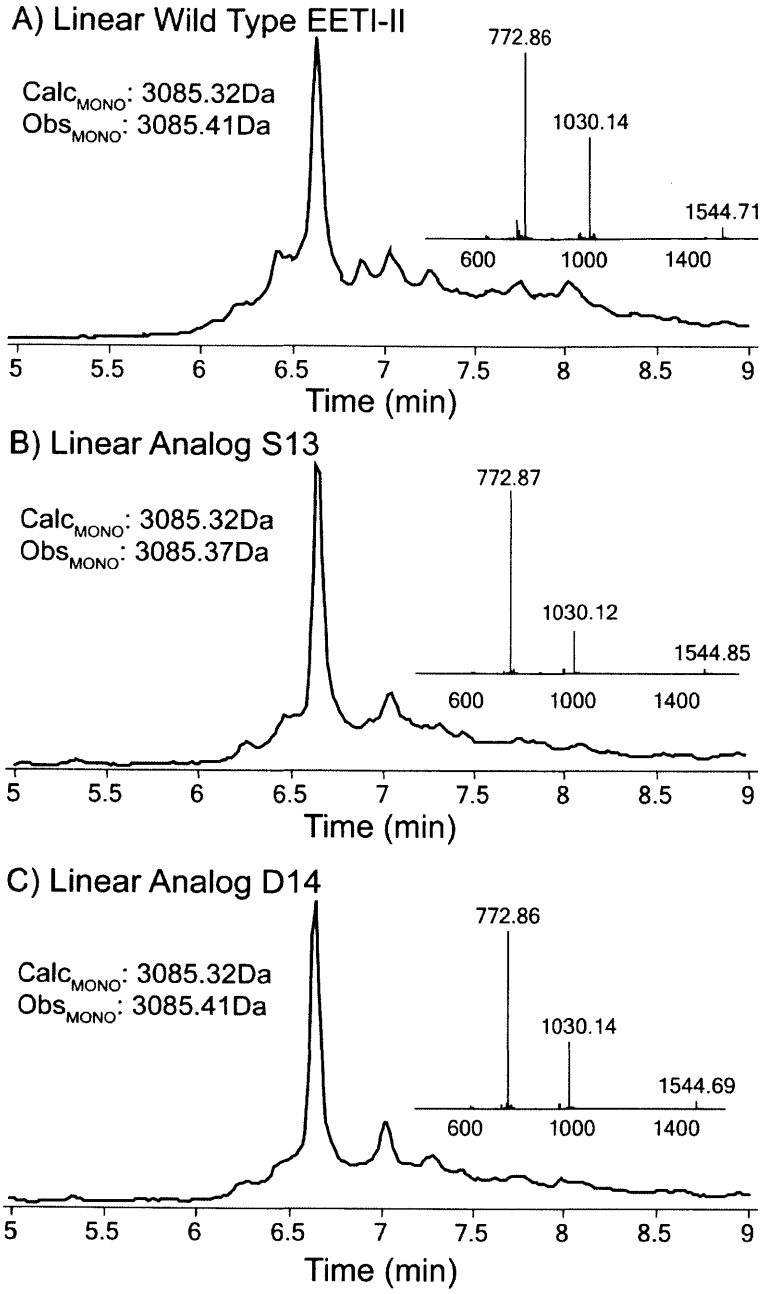


Figure A4.1: LCMS analysis of the crude linear peptides: A) Wild Type EETI-II, B) Analog D12, and C) Analog S13. Total ion current is displayed, and masses are monoisotopic.

A4.6 Purification:

Analytical Linear EETI-II

20mg of each linear peptide was dissolved in 5mL of 10%B in A and purified on a Waters 600 HPLC. An Agilent C3 Zorbax SB column (9.4 x 250 mm, 5μm packing) was used with a flow rate of 5mL/min of the following gradient: 10% B ramping linearly to 15% over 5 minutes, then a linear ramp to 40% B over 75 minutes. Fractions were analyzed by MALDI and pooled. The purity of the pooled fractions was confirmed by analysis under LC-MS condition 1.

Analytical Folded EETI-II

Analytical folding reactions were purified using a Waters 600 HPLC. An Agilent C3 Zorbax SB column (4.6 x 150 mm, 3.5μm packing) was used with a flow rate of 1.2mL/min of the following gradient: 5% B ramping linearly to 10% over 5 minutes, then a linear ramp to 35% B over 75 minutes. Fractions were analyzed by MALDI and pooled. The purity of the pooled fractions was confirmed by analysis under LC-MS condition 1.

Large Scale EETI-II

Purification of large scale folding reactions was performed on an Agilent mass directed purification system (1260 infinity LC and 6130 single quad MS) with an Agilent C3 Zorbax SB column (21 x 250 mm, 7μm packing) and a flow rate of 35mL/min of the following gradient: 5% B ramping linearly to 10% B over 5 min, then a linear ramp to 15% B over 10 min, then a linear ramp to 20%B over 25 min, then a linear ramp to 25% B over 5 minutes, and finally a flush at 70% B for 5 min. Fractions were pooled based on in-line mass spectral analysis and the purity of the pooled fractions was confirmed by analysis under LC-MS condition 1.

Large Scale Z33

Purification of Z33 variants was performed on an Agilent mass directed purification system (1260 infinity LC and 6130 single quad MS) with an Agilent C3 Zorbax SB column (9.4 x 250 mm, 5μm packing) and a flow rate of 4 mL/min of the following gradient: 5% mobile phase C (acetonitrile with 0.1% formic acid) and 95% mobile phase D (water with 0.1% formic acid) for 2 min, and then ramping linearly to 15% C over 40 min. Fractions were pooled based on in-line mass spectral analysis and the purity of the pooled fractions was confirmed by analysis under LC-MS condition 1.

A4.7 Folding:

Folding of EETI-II analogs was attempted under four conditions. Small scale folding reactions were performed in glutathione based redox buffer, cysteine based redox buffer, or in phosphate buffer exposed to atmospheric oxygen.

Glutathione: Folding of all analogs was attempted in glutathione based redox buffer. 1-2 mg of purified analog was dissolved to a concentration of 10mg/mL (3.3mM) in dissolving buffer (6M Guan HCl, 0.2M phosphate, pH=6.9) and diluted to a concentration of 1mg/mL with folding buffer (25mM Tris, 10mM GSH, 2mM GSSH, pH=7.8). After at least three hours, each analog that folded was purified as described elsewhere in this Appendix. Analogs that did not fold were discarded. No attempt was made to reduce and recover misfolded material.

Cysteine: For each analog that did not fold under the above conditions, the folding reactions were additionally attempted in a cysteine based redox buffer. 1-2mg of each analog was dissolved to 1mg/mL in cysteine redox buffer (100mM NH_4OAc, 2mM EDTA, 1M GnHCl, 4mM cysteine, 0.2mM cystine, pH=7.8) and allowed to stand overnight (22hrs). All analogs that failed to fold in the glutathione buffer also failed to fold in the cysteine buffer. Figure A4.2 shows WT EETI-II folding in this redox buffer over several hours.

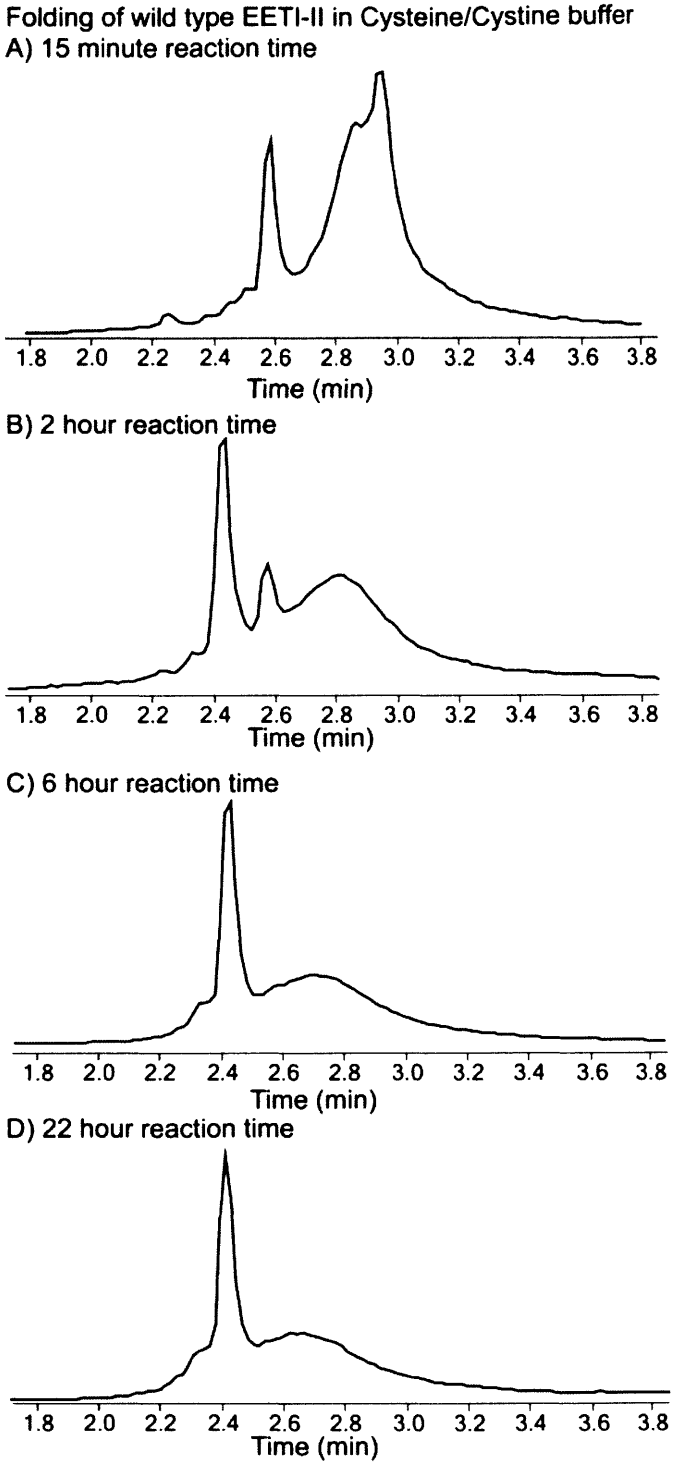


Figure A4.2: Folding of wild type EETI-II in 100mM NH_4OAc, 2mM EDTA, 1M GnHCl, 4mM cysteine, and 0.2mM cystine at pH=7.8. Folding was complete in less than 6 hours. Total ion current is displayed in each chromatogram.

Air: For each analog that did not fold under either of the above conditions, the folding reaction was additionally attempted in phosphate buffer open to the air. 1-2mg of each analog was dissolved to a concentration of 1mg/mL in 100mM phosphate buffer (pH=8) and allowed to stand for 2-3 days. Analogs that folded successfully were purified as described above. Figure A4.3 shows wild type EETI-II folding in oxygenated phosphate buffer, and Figure A4.4 shows the folding reaction time courses for all of the analogs that were subjected to oxygenated phosphate buffer.

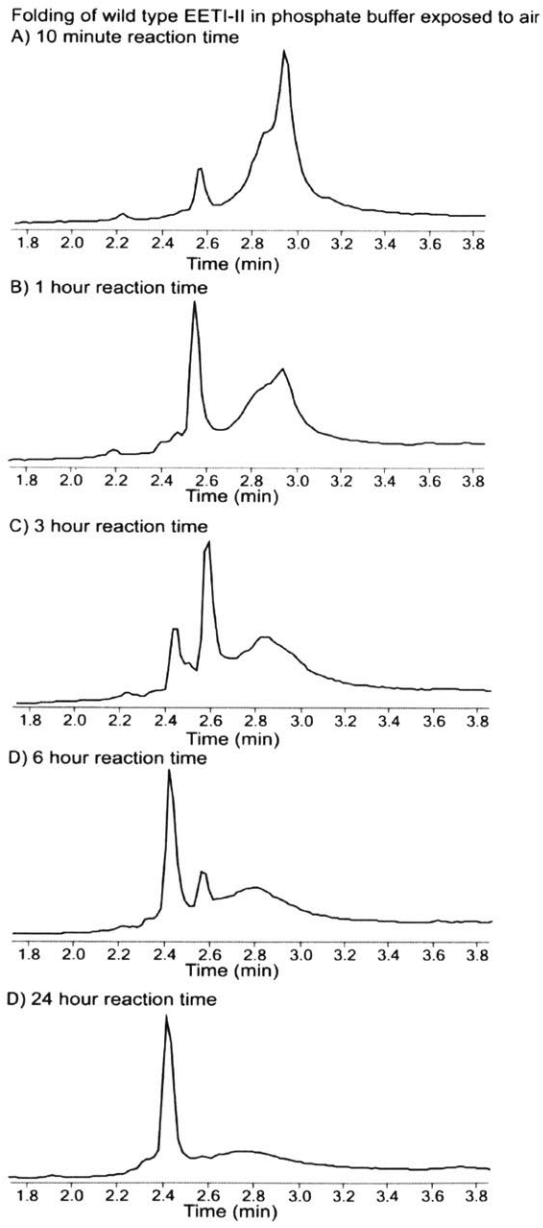


Figure A4.3: Folding of wild type EETI-II in oxygenated phosphate buffer (100mM, pH=8). The reaction is complete within 24 hours. Total ion current is displayed in each chromatogram.

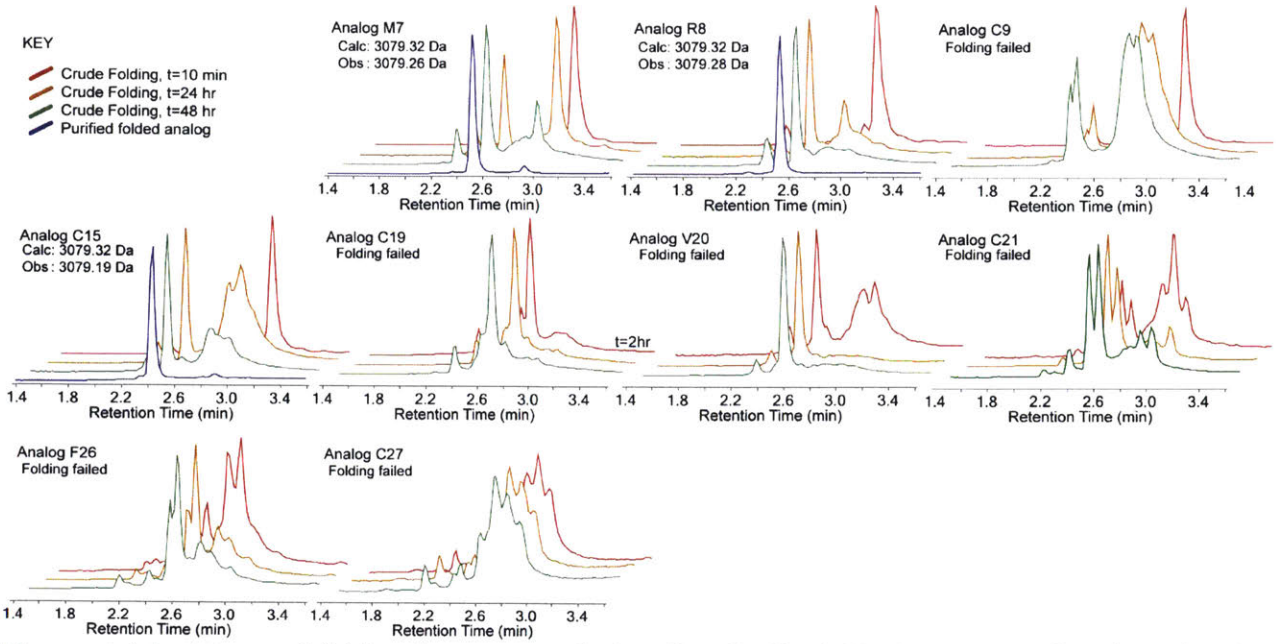


Figure A4.4: Attempted folding in oxygenated phosphate buffer (100mM, pH=8) of each analog that failed to fold in glutathione and cysteine redox buffers. Three analogs folded only under these conditions. Total ion current is displayed in each chromatogram, and all masses are monoisotopic.

Large Scale: Large scale folding reactions were performed as follows. 100mg of each crude analog that was selected for large scale synthesis was dissolved in 10mL of dissolving buffer and diluted with 100mL of folding buffer at a rate of about 200mL/min with continuous stirring. After at least two hours, the crude reaction mixture was passed through a 0.2μm filter and purified as described above. Yields, shown in table A4.1, were generally between 10 and 20% from crude.

Analog	Yield	Analog	Yield
WT	7.2%	A17	13.0%
A1	11.1%	P23	15.4%
K10	15.0%	N24	13.7%
Q11	16.5%	A25	19.2%
D12	18.7%	S29	10.5%
S13	17.3%	P30	17.1%
D14	10.7%		

Table A4.1: Yields of each scaled up analog after folding and purifying 100mg of crude linear peptide. Folding from crude and running a shallow preparative RP-HPLC gradient proved higher yielding than first purifying the linear analog, folding, then purifying the folded analog. In general, impurities that partially fold or do not fold do not have as marked a retention time shift as the properly folded analog, making them easier to remove after folding.

A4.8 Trypsin Inhibition Assay:

The trypsin inhibition assay was performed as follows. Each well in a row of a 96-well plate was filled with 180uL of 2mM Nα-benzoyl-Arg-(4-nitro)anilide in trypsin buffer (60mM Tris, 20mM $CaCl_2$, pH=7.8). 10uL of the appropriate concentration of EETI-II analog in trypsin buffer was then added. Finally, 10uL of 80nM trypsin in 1mM HCl was added and the absorbance at 410nM was monitored on a BioTek Epoch Plate reader for three hours. Concentrations of EETI-II were as follows in each row: 100nM, 32nM, 16nM, 8nM, 4nM, 2nM, 1nM, 0.5nM, 0.25nM, 0nM (Uninhibited 1), 0nM (Uninhibited 2), 0nM and no trypsin (substrate only blank). Each experiment was performed in quadruplicate, with two analogs assayed in each plate.

To determine the trypsin binding affinity of each analog, the data was fit to the following equation, adapted from Durek et al.[2], using unweighted non-linear least squares regression.

$$O = (M - U)\left(\frac{P_0 - I_0 - K_d + \sqrt{K_d^2 + 2K_d(P_0 + I_0) + (P_0 - I_0)^2}}{2P_0}\right) + U$$

Where O is the observed rate, M is the maximum rate, U is the background rate in the absence of trypsin, P_0 is the concentration of trypsin, I_0 is the concentration of inhibitor, and K_d is the dissociation constant.

For each set of data, M, U, and K_d were fit. P_0 and I_0 were assumed to be errorless. Sub-Appendix 1 contains plots of the raw data and fits for each analog.

Fits were performed using all 48 data points (quadruplicate runs of 12 points each). U, the background rate which accounts for drift in the spectrometer or degradation of the chromophore, was close to zero in all cases, as expected. (Minimum = -1.0, Maximum = 0.3 Average = -0.2, Standard Deviation = 0.4) Maximum rates, which are expected to be identical across samples, were not as consistent as hoped, but the complete data sets in Sub-Appendix 1 clearly show that this inconsistency is a real effect, not a statistical aberration (Min = 7.5, Max =13.5, Average = 10.7 Standard Deviation = 1.7). The maximum rate is determined by the activity of the trypsin, which may vary based on handling procedures, even taking care to prevent self-cleavage.

A4.9 Midpoint Potential Determination:

Midpoint potentials of each active analog were determined by incubating the analog in buffers of defined redox potential in an anaerobic glovebox at 4°C. The experiments were performed at 4°C because we did not have a suitable glove box at RT. Each experiment was performed in duplicate.

Oxidized DTT; reduced DTT; purified, lyophilized EETI-II analogs; argon purged, degassed 10%TFA in water; and argon purged, degassed buffer (75mM HEPES, pH=7.0) were brought into the

glove box. Solutions of 100mM DTT, 100mM oxidized DTT, and 0.67mM EETI-II analogs were prepared. A series of redox buffers was prepared by mixing oxidized and reduced DTT solutions. The redox potential of each buffer was calculated using the Nernst equation:

$$E_{Buffer} = E^{\circ}_{(DTTred:DTTox)} - \frac{RT}{nF}\ln\left(\frac{[DTTred]}{[DTTox]}\right)$$

Where E_{Buffer} is the redox potential of the buffer, $E^{\circ}_{(DTTred:DTTox)}$ is the midpoint potential of reduced and oxidized DTT (-0.330V), R is the universal gas constant (8.314 $VCmol^{-1}K^{-1}$), T is the temperature (277K), n is the number of electrons transferred (2 per disulfide bond), and F is the faraday constant (96,485 $Cmol^{-1}$).

Buffers were prepared with redox potentials from -390mV to -240mV in 10mV increments. For each analog, a sample of 500uLof 13uM EETI-II analog was prepared in each buffer. After two weeks, 75uL of each sample was quenched with 75uL of 10% TFA in water. Quenched samples were analyzed on an Agilent 1260 HPLC using the chromatographic conditions described in LC-MS analysis condition 1. The proportions of folded product, two disulfide product, and linear product were determined by integration of the 214nm UV chromatogram. One disulfide intermediates were generally not prevalent and poorly resolved, and therefore integrated with the linear product. Each experiment was performed in duplicate, and midpoint potentials were calculated via unweighted least squares fit to the Nernst equation (eq 2):

$$E_{Buffer} = E^{\circ}_{(Analog)} - \frac{RT}{nF}\ln\left(\frac{[ANALOGred]}{[ANALOGox]}\right)$$

Where E_{Buffer} is the redox potential of the buffer, $E^{\circ}_{(Analog)}$ is the midpoint potential of the analog, R is the universal gas constant (8.314 $VCmol^{-1}K^{-1}$), T is the temperature (277K), n is the number of electrons transferred (2 per disulfide bond, 2 for the two-three disulfide midpoint and 4 for the zero-two disulfide midpoint), and F is the faraday constant (96,485 $Cmol^{-1}$).

In Sub-Appendix 2, one set of raw HPLC traces is shown for each analog. Experiments were performed in duplicate, but only one set of chromatograms is shown for each analog.

Sub-Appendix 3 contains fitted integration data for each analog. In every case, blue diamonds are the linear peptide, red squares are the two disulfide intermediate, and green triangles are the folded analog. The black curves are the best fit curves.

In cases where the fit was poor, an additional aliquot was quenched after 4 more weeks and analyzed again. In all cases the fit was good after the additional incubation time, as seen in Sub-Appendix 3.

A4.10 Molecular Dynamics Simulation:

Multiple molecular dynamics (MD) simulations were performed for the WT EETI-II protein and four D-amino acid analogs (A18, P23, N24, and A25). The initial structure for the WT protein is obtained from the X-ray structure (PDB ID 1H9H)[3]. The initial structures for the four analogs were built by replacing the residue with the D-amino acid at the corresponding position using the UCSF Chimera package[4]. The NH_3^+ group was used for the N-terminus, while the C-terminus was capped by an amide. All MD simulations in this study were implemented using the GROMACS 4.6.7 package[5] with the RSFF1 force field[6,7]. For each protein, the starting structure was first energy minimized for 1000 steps and then solvated in a cubic TIP4P-Ew water box of size 51×51×51 $Å^3$.[8] Two chloride ions were added to neutralize the system. Before any MD simulation, the solvated system was further minimized for 5000 steps using the steepest descent algorithm. With all heavy atoms of the protein restrained, each system was heated from 5 K to 300 K within 20 ps and relaxed for 30 ps. Before production, an additional 500 ps equilibration was also performed with only the protein backbone restrained.

The 200 ns production runs were performed in an isobaric-isothermal (NPT) ensemble at 300 K/1.0 bar. The temperature of the system was maintained using the Nosé-Hoover algorithm with a coupling time constant of 1.0 ps[9,10]. To avoid the "hot solvent-cold solute" artifact[11,12], two separate thermostats were applied to the solvent and the protein. The pressure was regulated using the Berendsen barostat with a coupling time of 2.0 ps and a compressibility of 4.5×10^{-5} bar^{-1}.[13] All bonds were constrained using the LINCS algorithm[14], allowing the use of 2.0 fs time step with the leapfrog integrator. The non-bonded interactions were truncated at 8.0 Å. Long-range Coulomb interactions were calculated using the Particle Mesh Ewald (PME) method with a Fourier spacing of 1.2 Å and an interpolation order of 4.[15] A long-range analytic dispersion correction was applied to both the energy and pressure to account for the truncation of Lennard-Jones interactions[19].

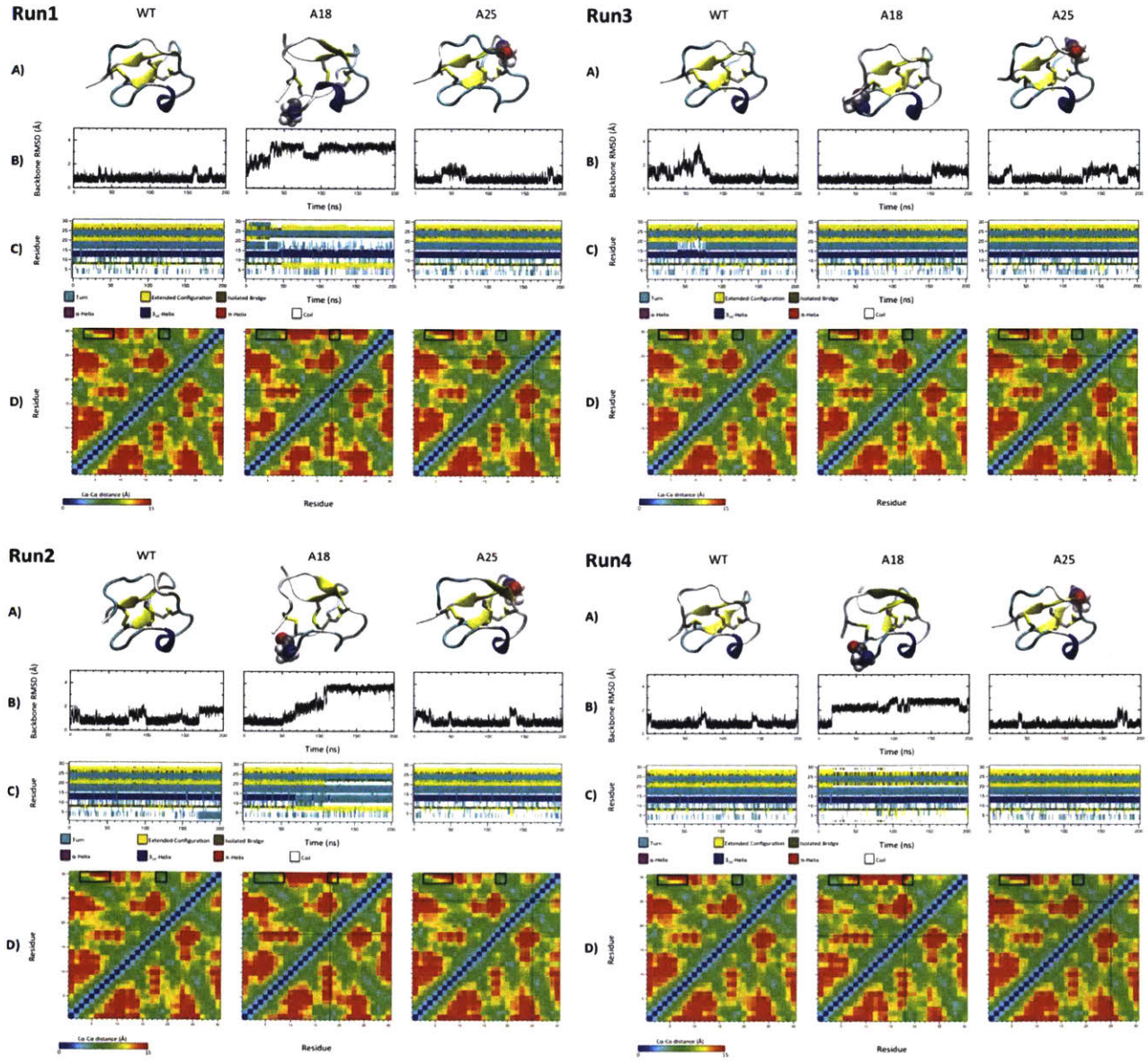


Figure A4.5. Molecular dynamic simulations of WT, A18 and A25. (A) Cartoon representations of the peptide structures after 200 ns MD simulations. The structures are colored according to the secondary structure calculated using STRIDE (turn: cyan; extended conformation: yellow; isolated bridge: brown; α -helix: magenta; 3_{10}-helix: blue; π-helix: red; coil: white). The disulfide bonds are represented by sticks and the D-amino acids are highlighted with spheres. **(B)** Peptide backbone RMSD trajectories with respect to the PDB structure (PDB ID=1H9H). **(C)** Secondary structure evolutions of each residue during the 200 ns simulations. The same color code is used as in (A). **(D)** Cα-Cα distance map calculated using 100–200 ns of each trajectory. The locations of the D-amino acids are denoted by black lines. Four independent MD runs were performed. Run 1 was also provided in Chapter 5 as Figure 5.2.6.

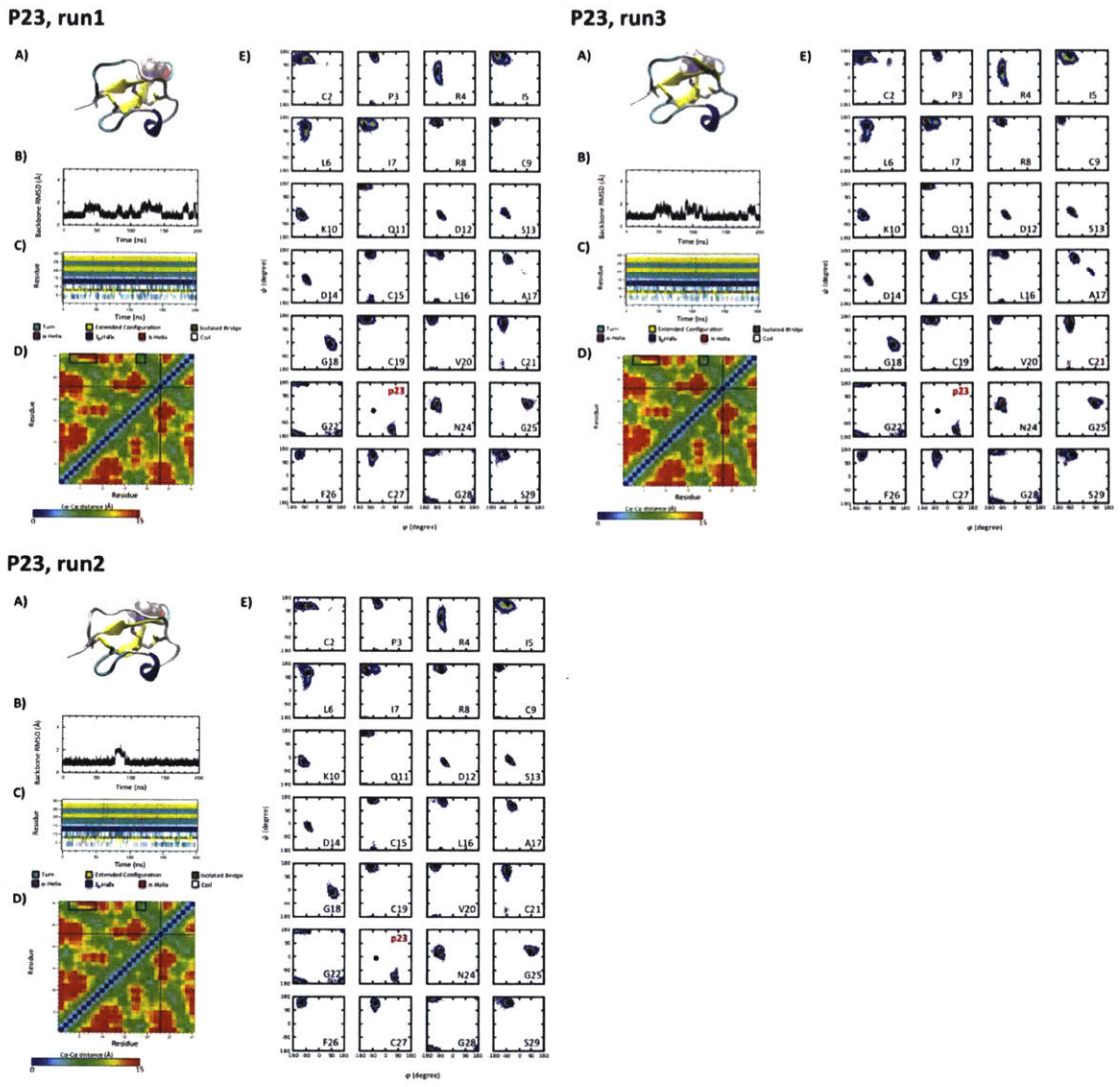


Figure A4.6. Molecular dynamic simulations of P23. (A) Cartoon representation of the P23 analog structure after 200 ns MD simulations. The structure is colored according to the secondary structure calculated using STRIDE (turn: cyan; extended conformation: yellow; isolated bridge: brown; α-helix: magenta; 3_{10}-helix: blue; π-helix: red; coil: white). The disulfide bonds are represented by sticks and the D-amino acids are highlighted with spheres. **(B)** Peptide backbone RMSD trajectory with respect to the PDB structure. **(C)** Secondary structure evolution of each residue during the 200 ns simulation. The same color code is used as in (A). **(D)** Cα-Cα contact mas calculated using 100–200 ns of the trajectory. The location of the D-amino acid is denoted by black lines. **(E)** Ramachandran plot for each residue calculated using 100–200 ns of the trajectory. The D-amino acid is labeled with a red lower-case letter. φ/ψ angles of the WT are shown as black dots. Three independent MD runs were performed. Run 1 was also shown in Chapter 5 as Figure 5.2.7.

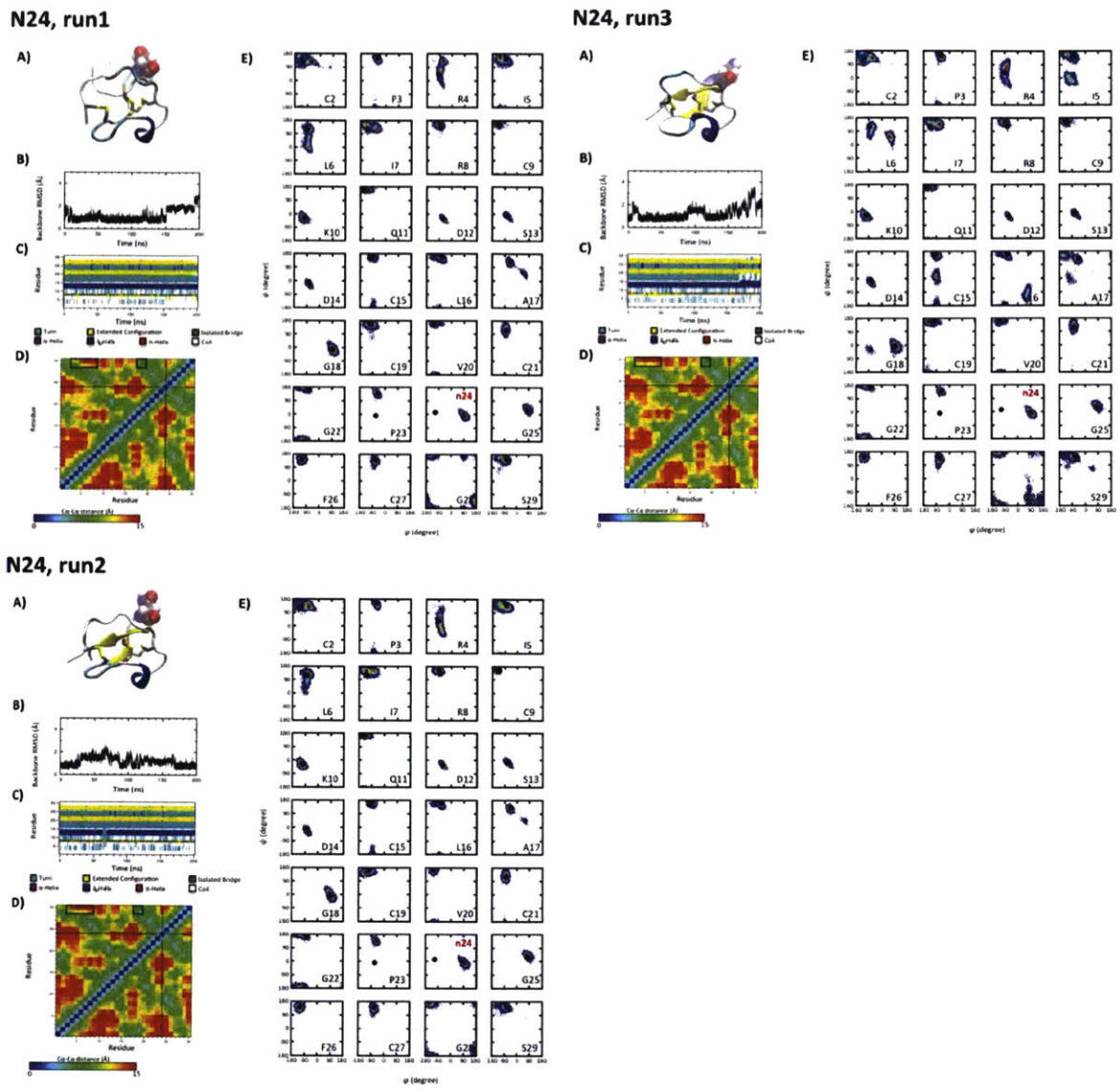


Figure A4.7. Molecular dynamic simulations of N24. (A) Cartoon representation of the N24 analog structure after 200 ns MD simulations. The structure is colored according to the secondary structure calculated using STRIDE (turn: cyan; extended conformation: yellow; isolated bridge: brown; α-helix: magenta; 3_{10}-helix: blue; π-helix: red; coil: white). The disulfide bonds are represented by sticks and the D-amino acids are highlighted with spheres. **(B)** Peptide backbone RMSD trajectory with respect to the PDB structure. **(C)** Secondary structure evolution of each residue during the 200 ns simulation. The same color code is used as in (A). **(D)** Cα-Cα contact mas calculated using 100–200 ns of the trajectory. The location of the D-amino acid is denoted by black lines. **(E)** Ramachandran plot for each residue calculated using 100–200 ns of the trajectory. The D-amino acid is labeled with a red lower-case letter. φ/ψ angles of the WT are shown as black dots. Three independent MD runs were performed.

A4.11 BioLayer Interferometry Binding Assay:

Binding kinetics for the Z33/IgG interaction were measured on an Octet RED96 instrument from ForteBio. All measurements were performed at 25° C in a 96-well plate shaking at 1000 rpm. Biotinylated trastuzumab was prepared by reacting trastuzumab with biotin-$(PEG)_4$-NHS following a previously described protocol[16]. Streptavidin-coated tips (from ForteBio) were dipped in water for 10 minutes to remove the sucrose coating. Experiments consisted of five steps:

1) Base line reading. Tips were dipped in a well containing fresh binding buffer (1x phosphate buffered saline, 1% w/v bovine serum albumin, 0.02% v/v tween-20, pH 7.4) for 60 seconds.
2) Ligand loading. ~50 nM solutions of biotinylated trastuzumab antibody in binding buffer were associated onto the tips for 60 seconds.
3) Base line reading #2. Tips were dipped in a well containing fresh binding buffer for 60 seconds.
4) Association. Trastuzumab loaded tips were dipped into well containing variable concentrations of Z33 variants (5 μM – 0.156 μM) in binding buffer for 100 seconds. A reference well with binding buffer was included in each measurement.
5) Dissociation. The tips were dipped back into wells containing binding buffer for 100 seconds.

Data processing was done using the ForteBio data analysis software package. First, the reference well association was subtracted from all sensorgrams. Then, each association and dissociation curve was locally fitted to obtain the association constant k_{on} and dissociation constant k_{off} following the manufactuer's protocol. The binding constant K_D was calculated as k_{off}/k_{on}.

For all variants except His14, K_D was calculated as the average of binding constants from measurements at 4 different concentrations, as shown in the sensorgrams in Sub-appendix 4. Errors were reported as the standard deviation of these four K_D values.

For His14, a steady state fitting was used to obtain the K_D value. First, equilibrium response Req was obtained from the software by averaging the response at equilibrium. Then, concentration of His14 in each measurement was plotted against Req. The concentration-Req curve was fitted using the Fortebio software to obtain the K_D value, error was reported as the error from fitting.

A4.12 CD Analysis of the Thermodynamic stability of Z33 Analogs:

Far-UV CD spectra for Z33 and its analogs were acquired on an Aviv 202 Circular Dichroism Spectrometer at –10 °C in a 1 mm quartz cuvette from 260 to 195 nm with six seconds averaging times at

each wavelength. 0.2 mg/mL protein solutions were prepared by dissolving solid samples in PBS containing 30% (v/v) glycerol (CD buffer). Data processing included solvent background correction (subtraction) and adjustment for pathlength and concentration (mean residue ellipticity, $MRE = [\theta]\lambda = \frac{\theta_{obs}}{10lcn}$, where θ_{obs} is measured ellipticity, θ — mean residual ellipticity in deg·cm^2·dmol^{-1}, l — pathlength in cm; c — concentration of peptide in M; n — # of amino acids). K2D3 algorithm[17] was used to perform spectral deconvolution in the range of 195-240 nm, assuming 33 amino acids for every analog. In every case the input spectrum agreed well with the deconvolution result. Deconvolution values obtained in this way (percentages of α-helix and β-strand) were directly used for further analysis.

Urea unfolding experiments were performed using the same setup, except 15 second acquisition times were used. Protein unfolding was monitored at 220 nm, −10 °C and 0.2 mg/mL protein concentration. For each peptides, eight data points corresponding to different urea concentrations (0, 1, 2, 3, 4, 5, 6, and 7.5M) were collected. The analytes were prepared volumetrically from 3.2 mg/mL peptide stock solutions in CD buffer, CD buffer, and 8M urea in the same buffer. In cases where the regions of fully folded and unfolded protein were observed, the data were processed in the following way.

Measured ellipticity values were converted to MRE for each data point as explained above. MRE at 0M urea concentration was generally assumed to correspond to a fully folded protein, y_f, while MRE at 7.5M urea concentration was generally assumed to correspond to a fully unfolded protein, y_u. Under these assumptions, the values for ΔG of unfolding were calculated from the following expressions:

$$^{263}_{\Delta}G^{K}[urea] = -RTln\frac{y_f - y}{y - y_u},$$

where T = 263 K, R = 1.987 cal·K^{-1}·M^{-1}, y — MRE at concentration of urea *[urea]*. $^{263}_{\Delta}G^{K}[urea]$ values corresponding to the transition region rather than the complete unfolding curve were used to performed linear least-squares regression to fit $^{263}_{\Delta}G^{K}[urea]=0$ and m, where m is the slope of the fit[18]:

$$^{263}_{\Delta}G^{K}[urea] = {}^{263}_{\Delta}G^{K}[urea]=0 + m \cdot [urea]$$

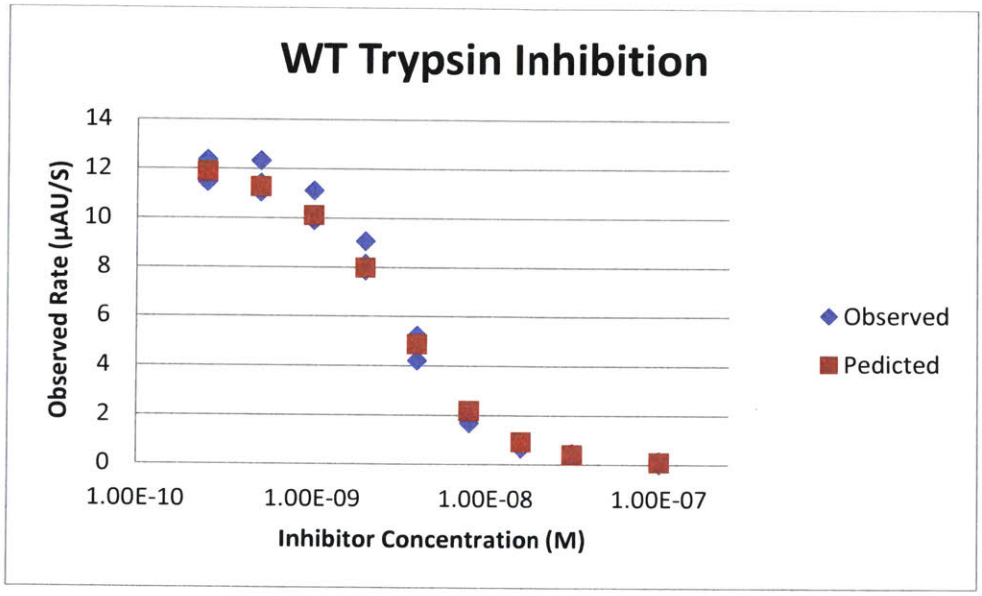
From these values, values for $[urea]_{1/2}$were calculated as $[urea]_{1/2} = -\frac{^{263}_{\Delta}G^{K}[urea]=0}{m}$. Both $^{263}_{\Delta}G^{K}[urea]=0$ and $[urea]_{1/2}$ values were reported as 95% confidence intervals. Regression lines in Fig. 10b were plotted as functions of fit parameters:

$$MRE_{[urea]} = \frac{y_f + y_u \cdot e^{-\frac{^{263}_{\Delta}G^{K}[urea]=0 + m\cdot[urea]}{RT}}}{e^{-\frac{^{263}_{\Delta}G^{K}[urea]=0 + m\cdot[urea]}{RT}} + 1}.$$

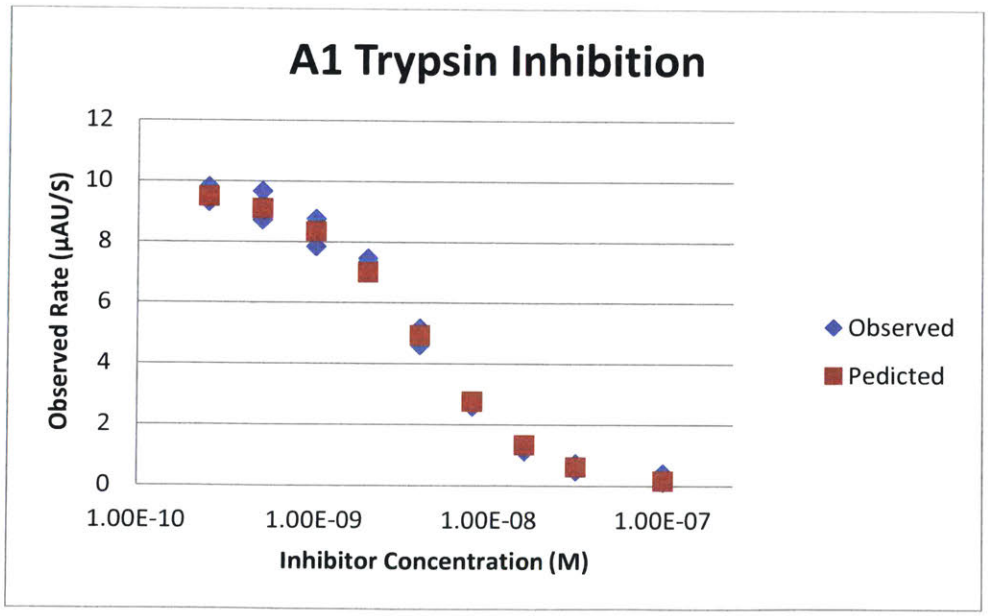
Analysis and fitting of the data missing either y_u or y_f regions was not performed.

A4.13 Inhibition Assay Data

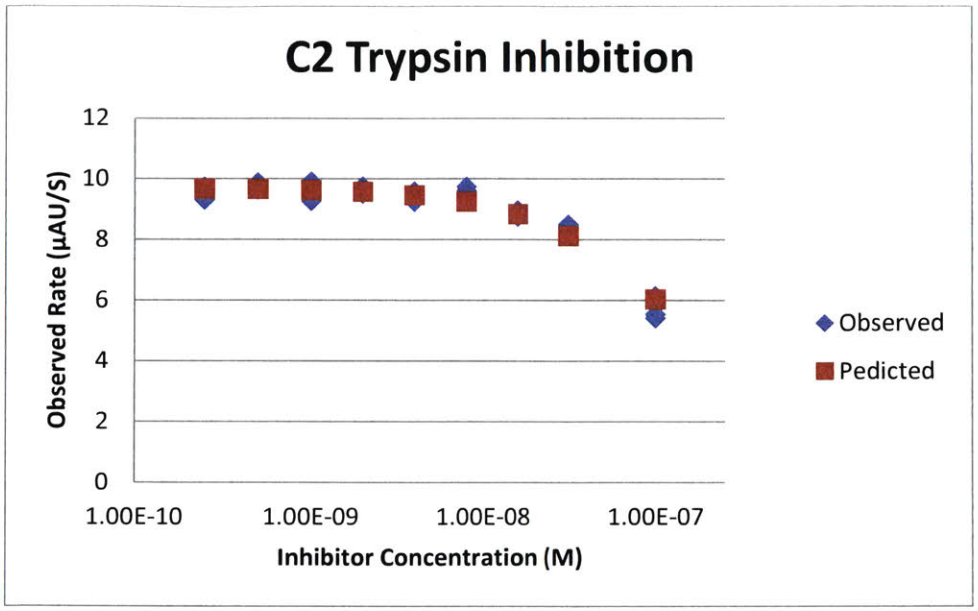
Complete data sets for the trypsin inhibition assays for all analogs are shown below. Each plot shows the measured relative trypsin activity at various concentrations of inhibitor (blue diamonds), and the relative trypsin activity predicted by equation 1 with parameters determined by a least squares fit, if applicable (red squares). For very low affinity analogs, a fit is not reported because relative errors were too high.



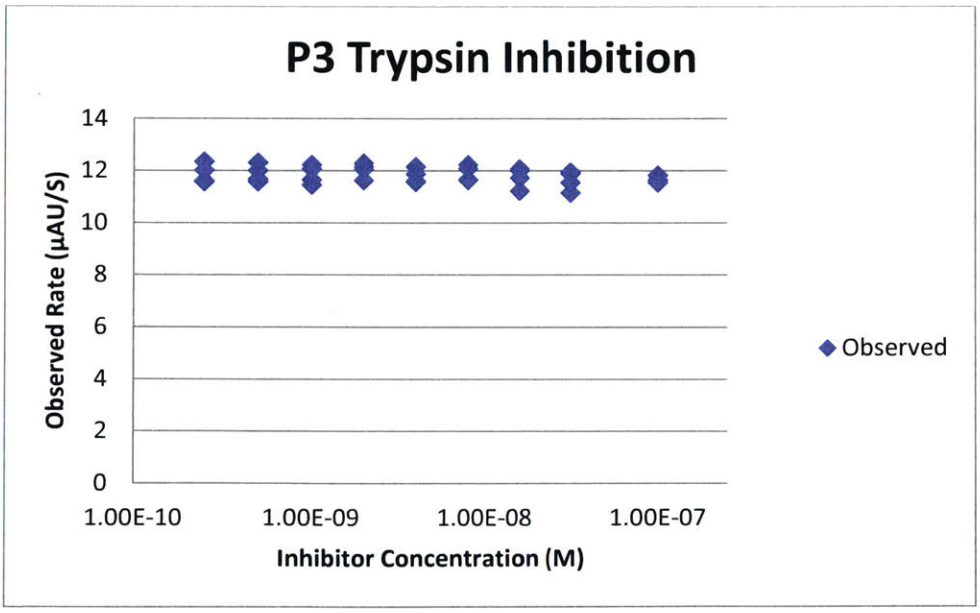
Trypsin binding affinity of WT EETI-II.



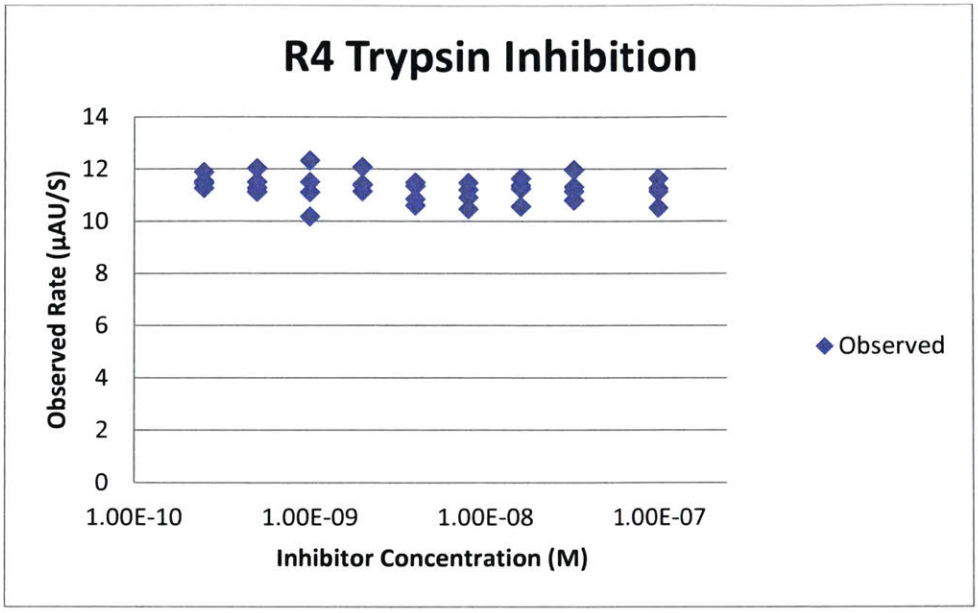
Trypsin binding affinity of analog A1.



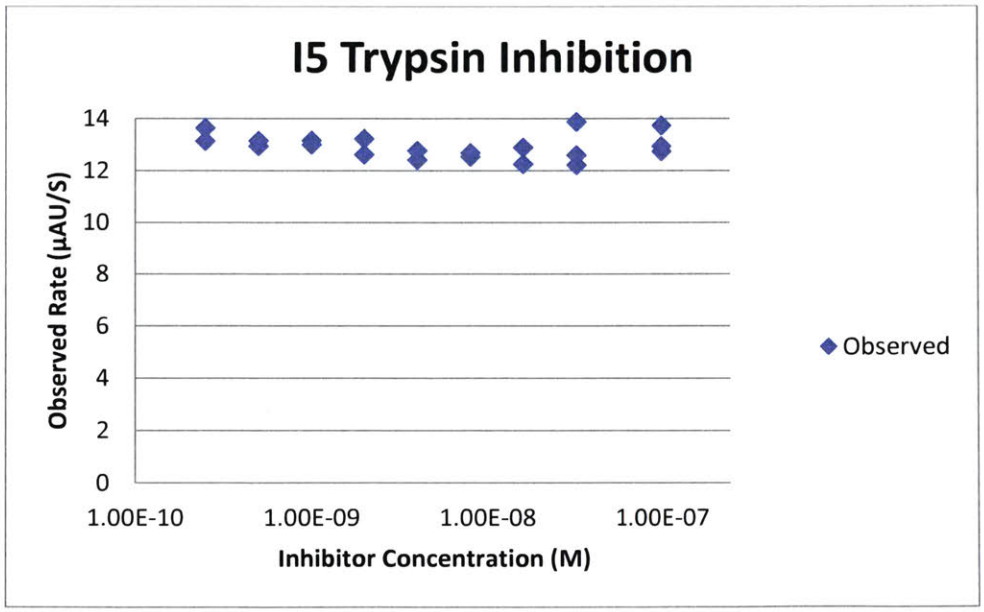
Trypsin binding affinity of analog C2.



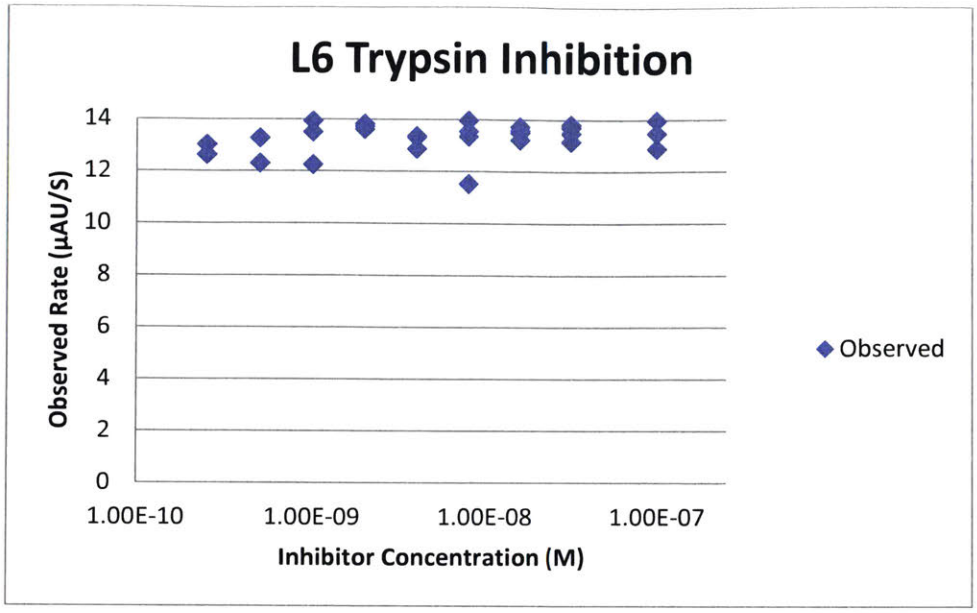
Trypsin binding affinity of analog P3.



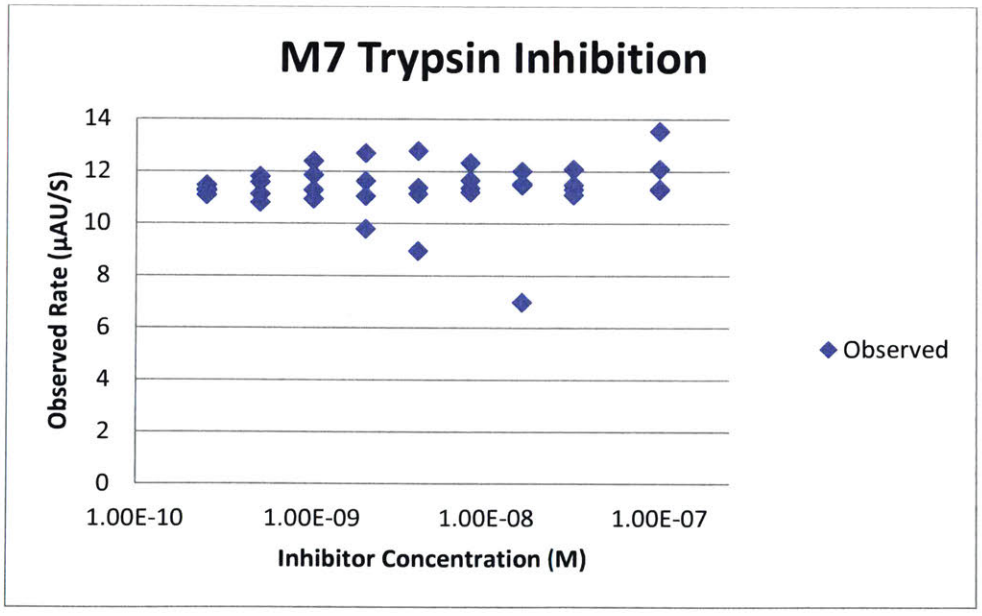
Trypsin binding affinity of analog R4.



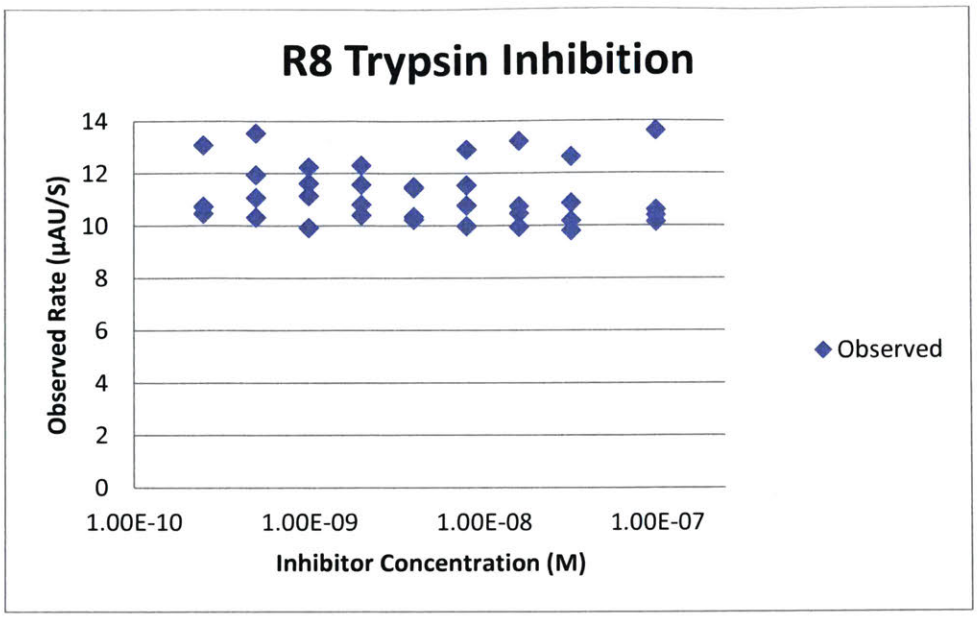
Trypsin binding affinity of analog I5.



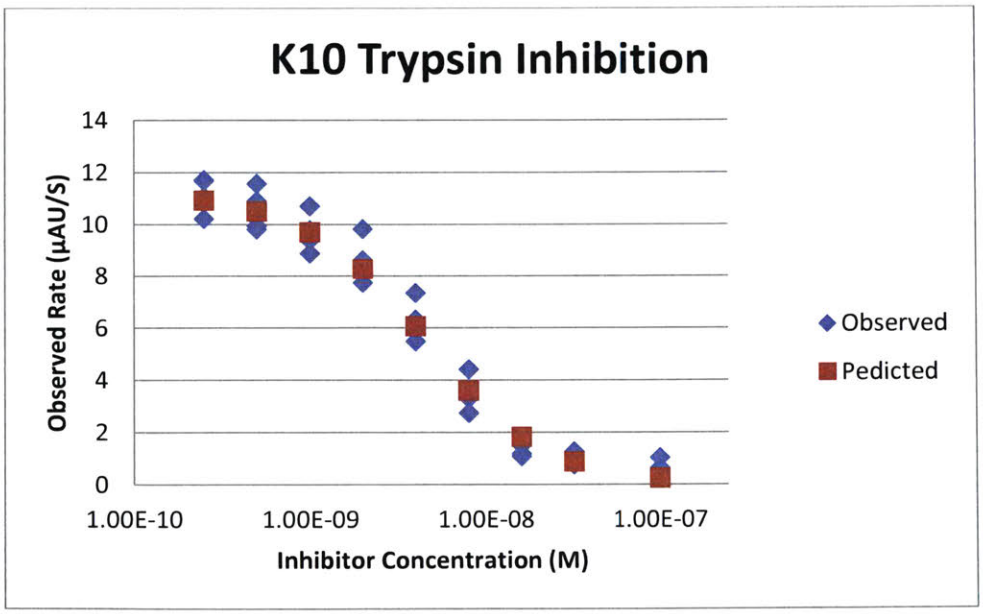
Trypsin binding affinity of analog L6.



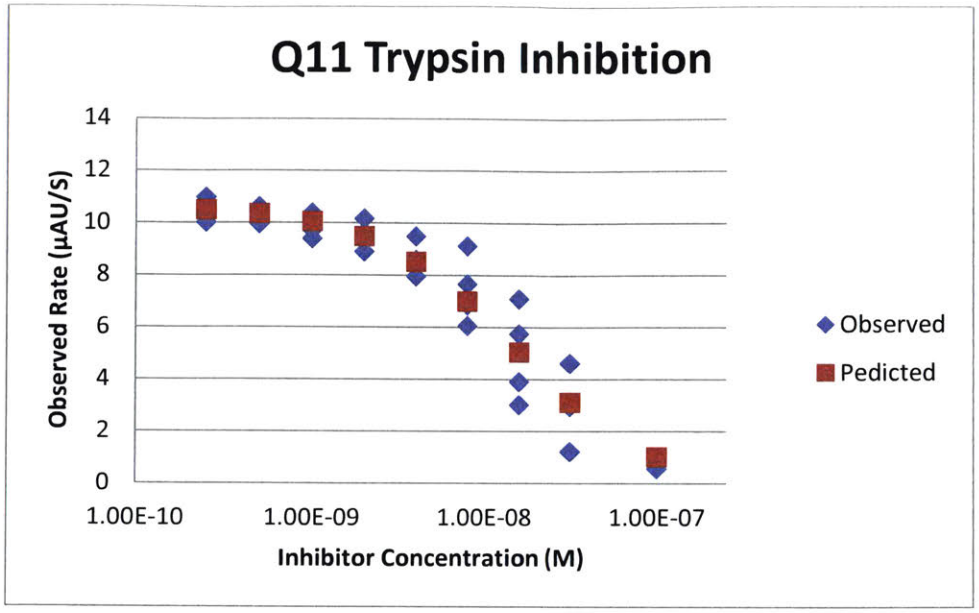
Trypsin binding affinity of analog M7.



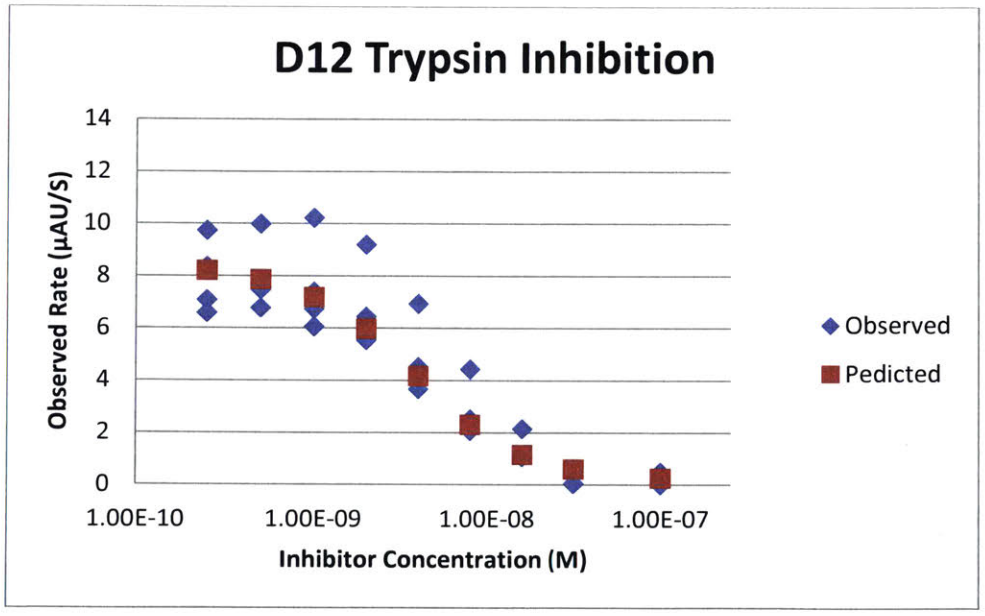
Trypsin binding affinity of analog R8.



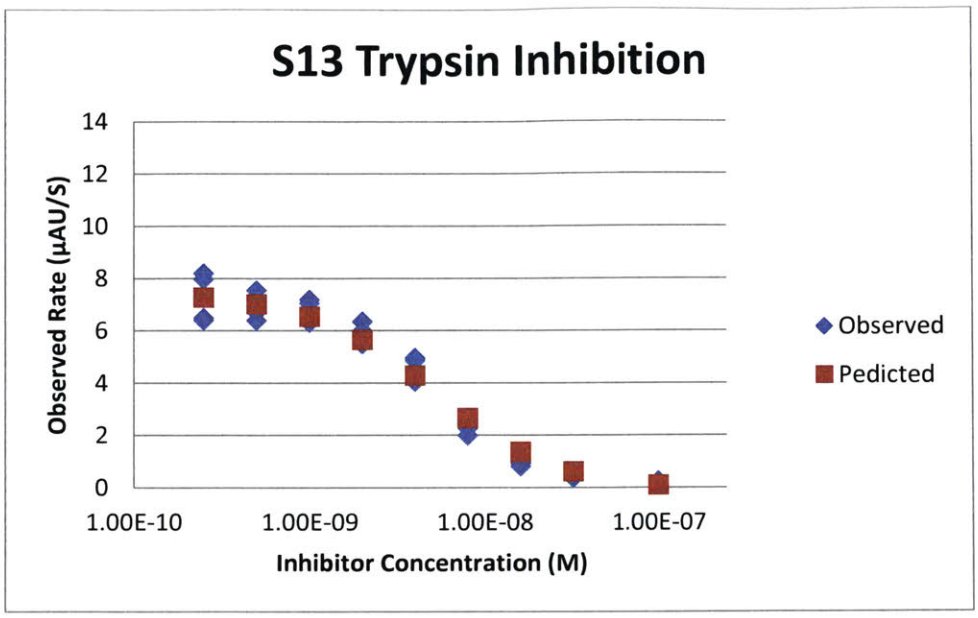
Trypsin binding affinity of analog K10.



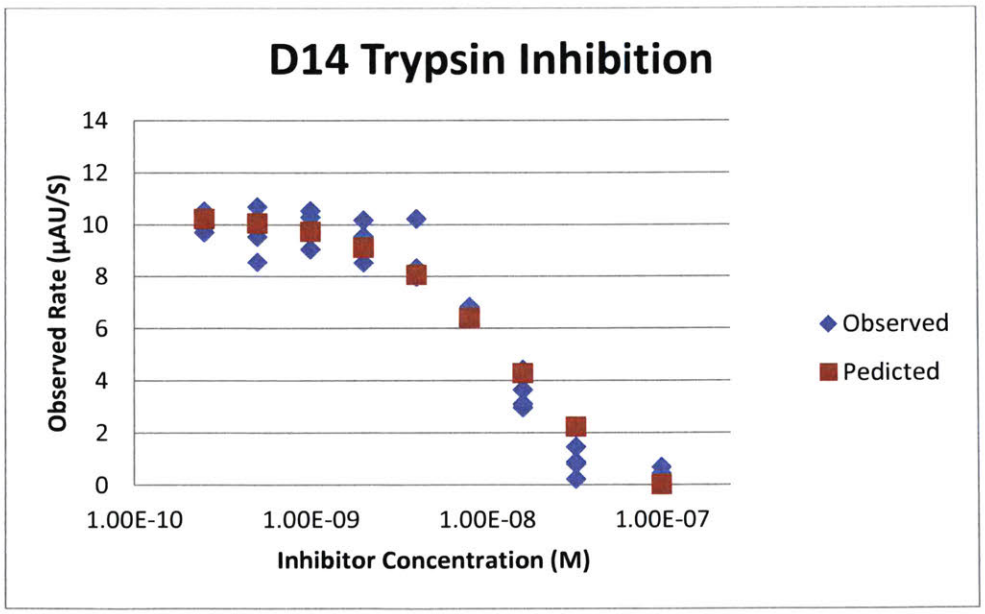
Trypsin binding affinity of analog Q11.



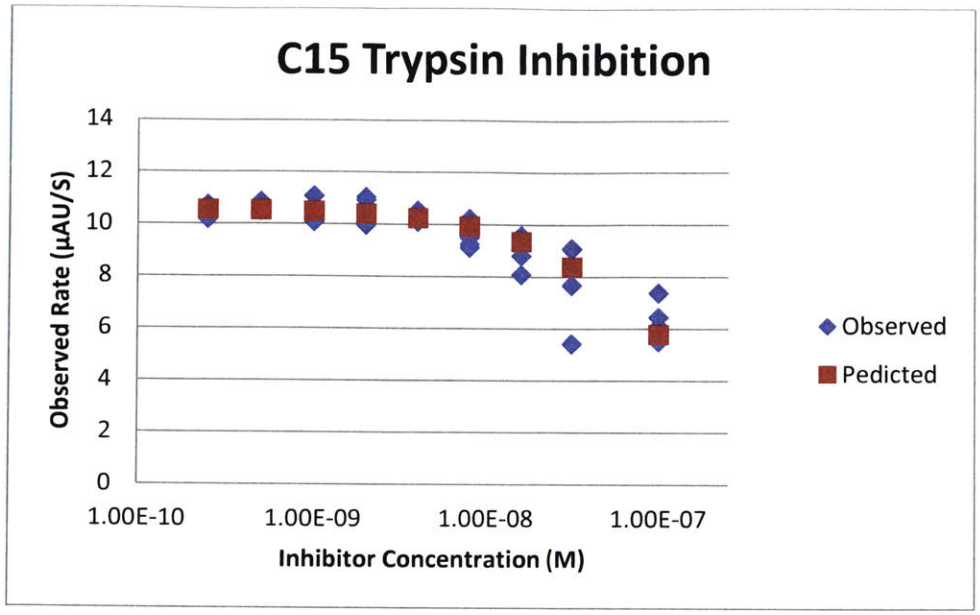
Trypsin binding affinity of analog D12.



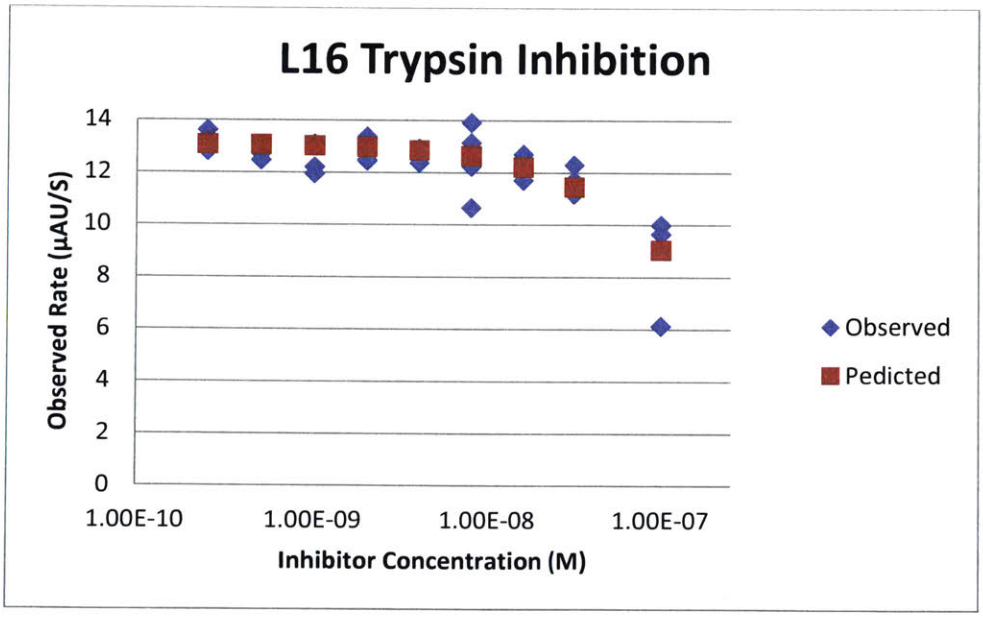
Trypsin binding affinity of analog S13.



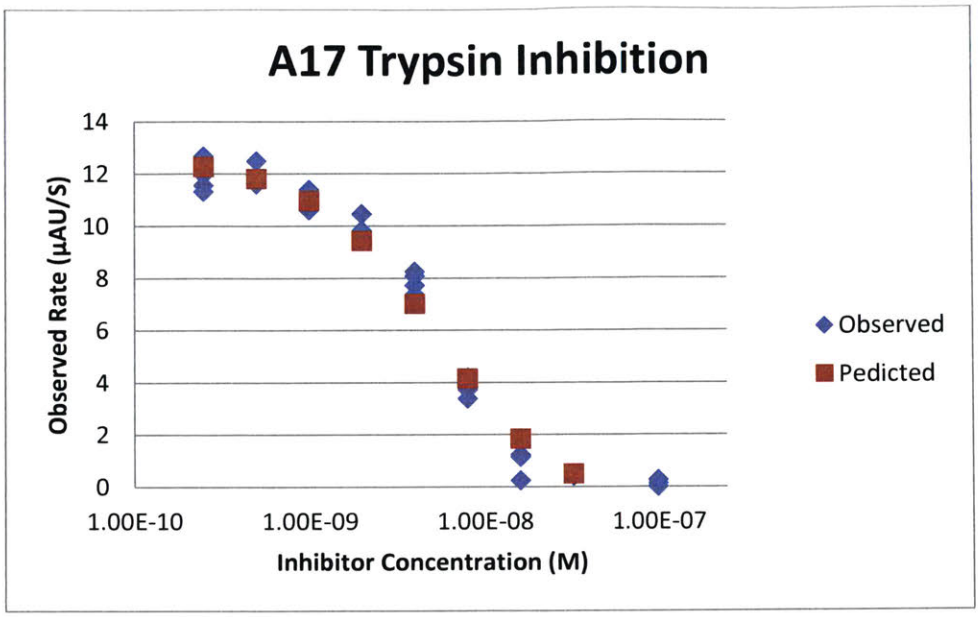
Trypsin binding affinity of analog D14.



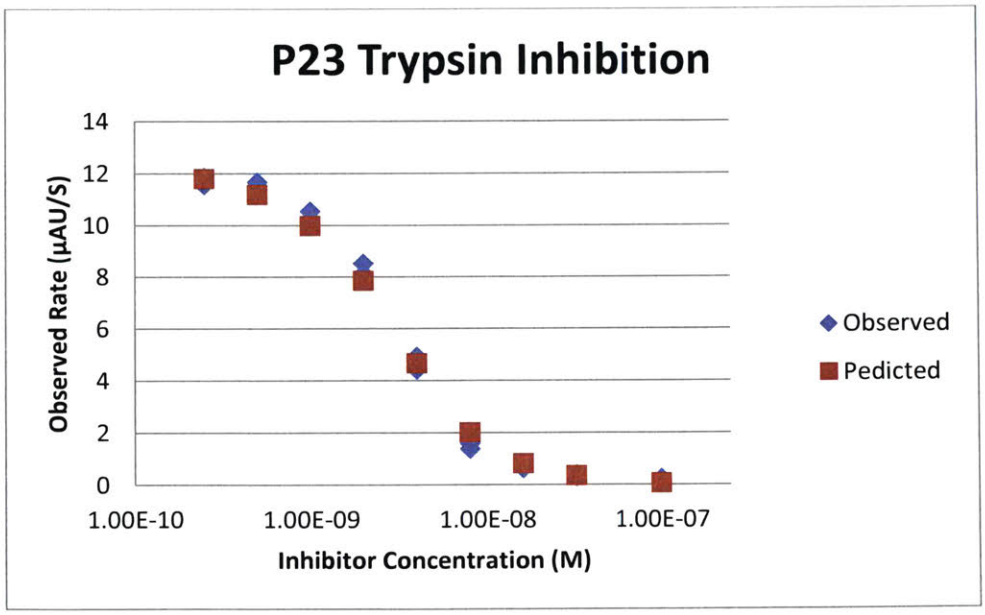
Trypsin binding affinity of analog C15.



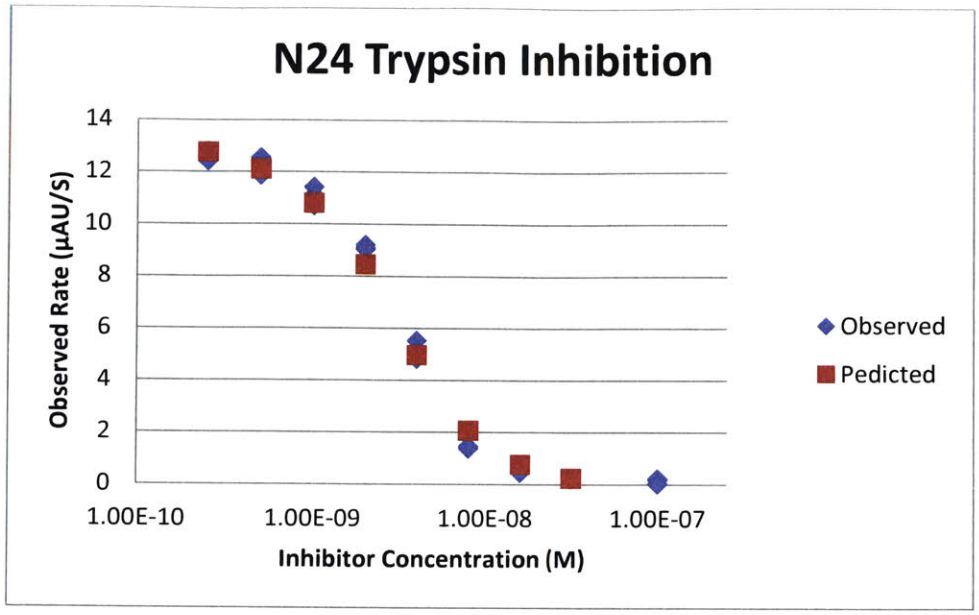
Trypsin binding affinity of analog L16.



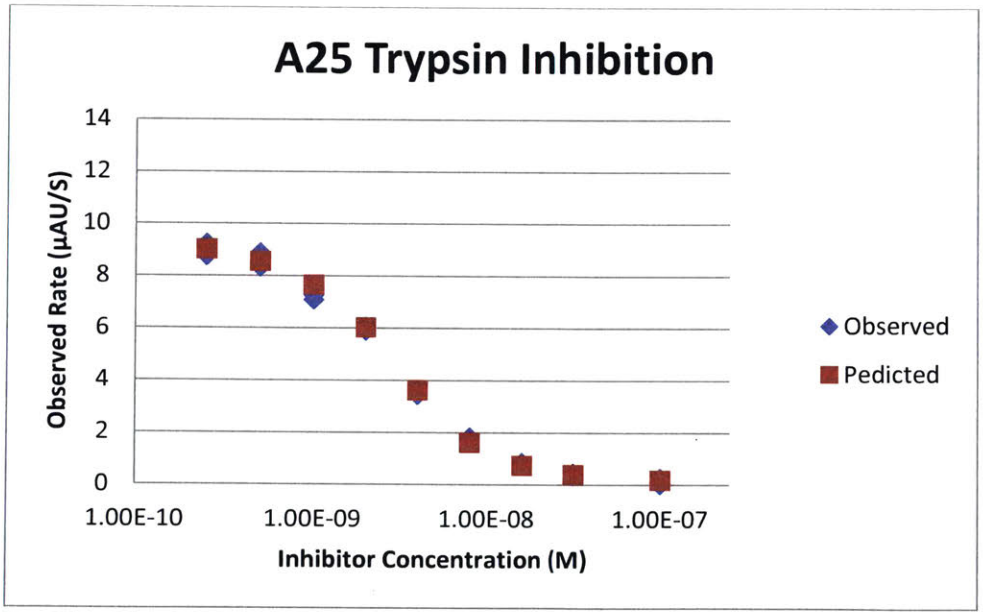
Trypsin binding affinity of analog A17.



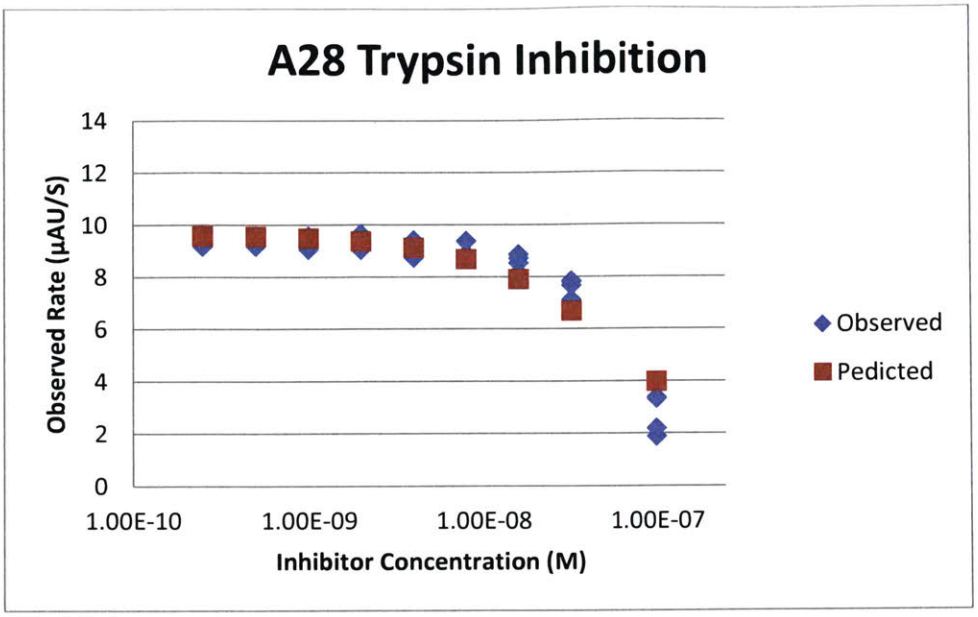
Trypsin binding affinity of analog P23.



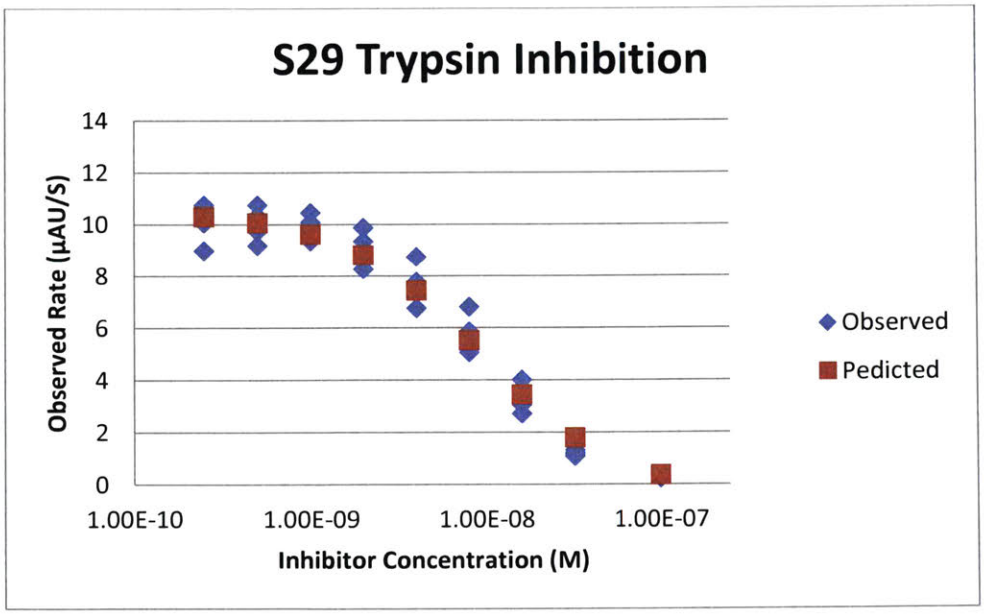
Trypsin binding affinity of analog N24.



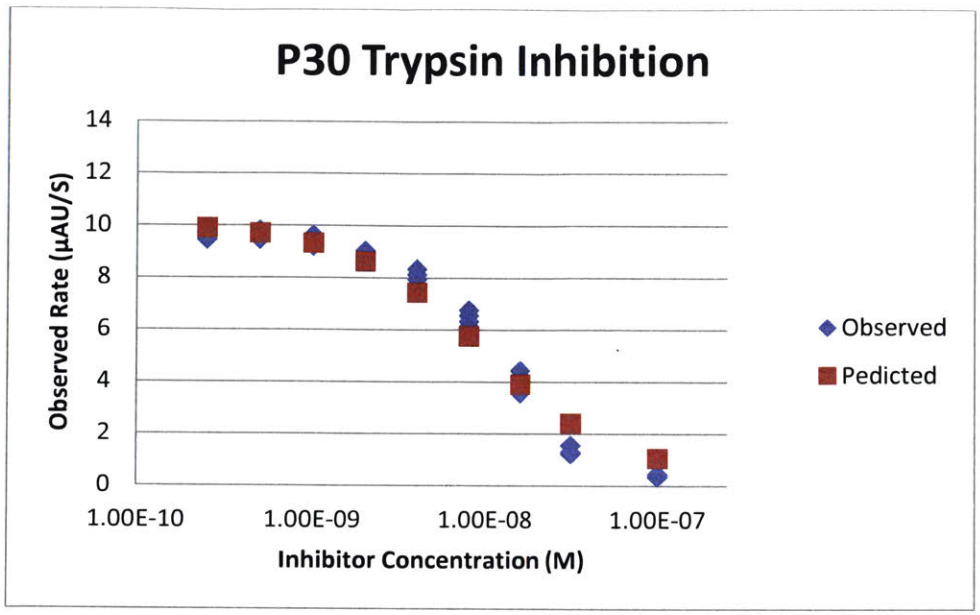
Trypsin binding affinity of analog A25.



Trypsin binding affinity of analog A28.



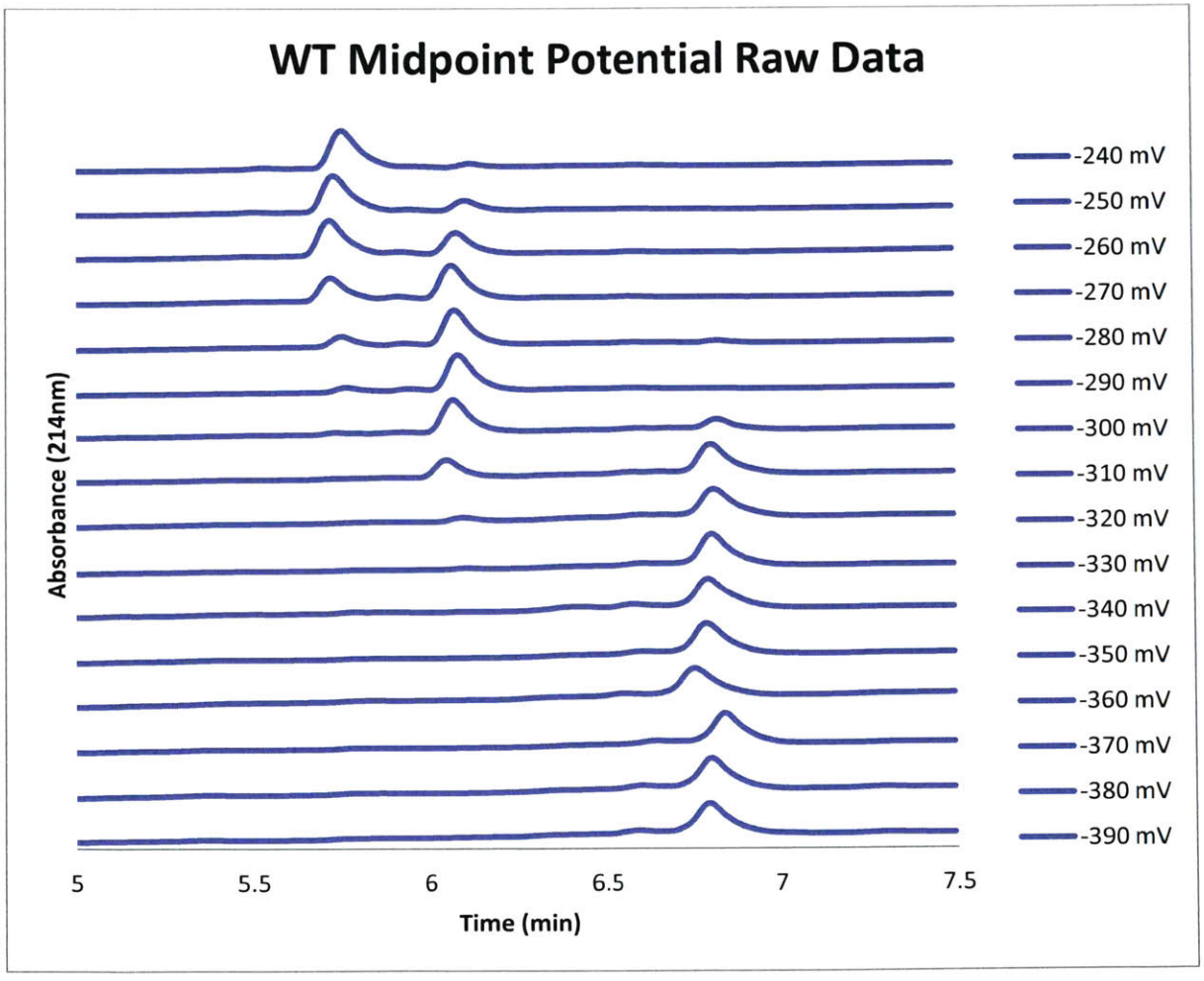
Trypsin binding affinity of analog S29.



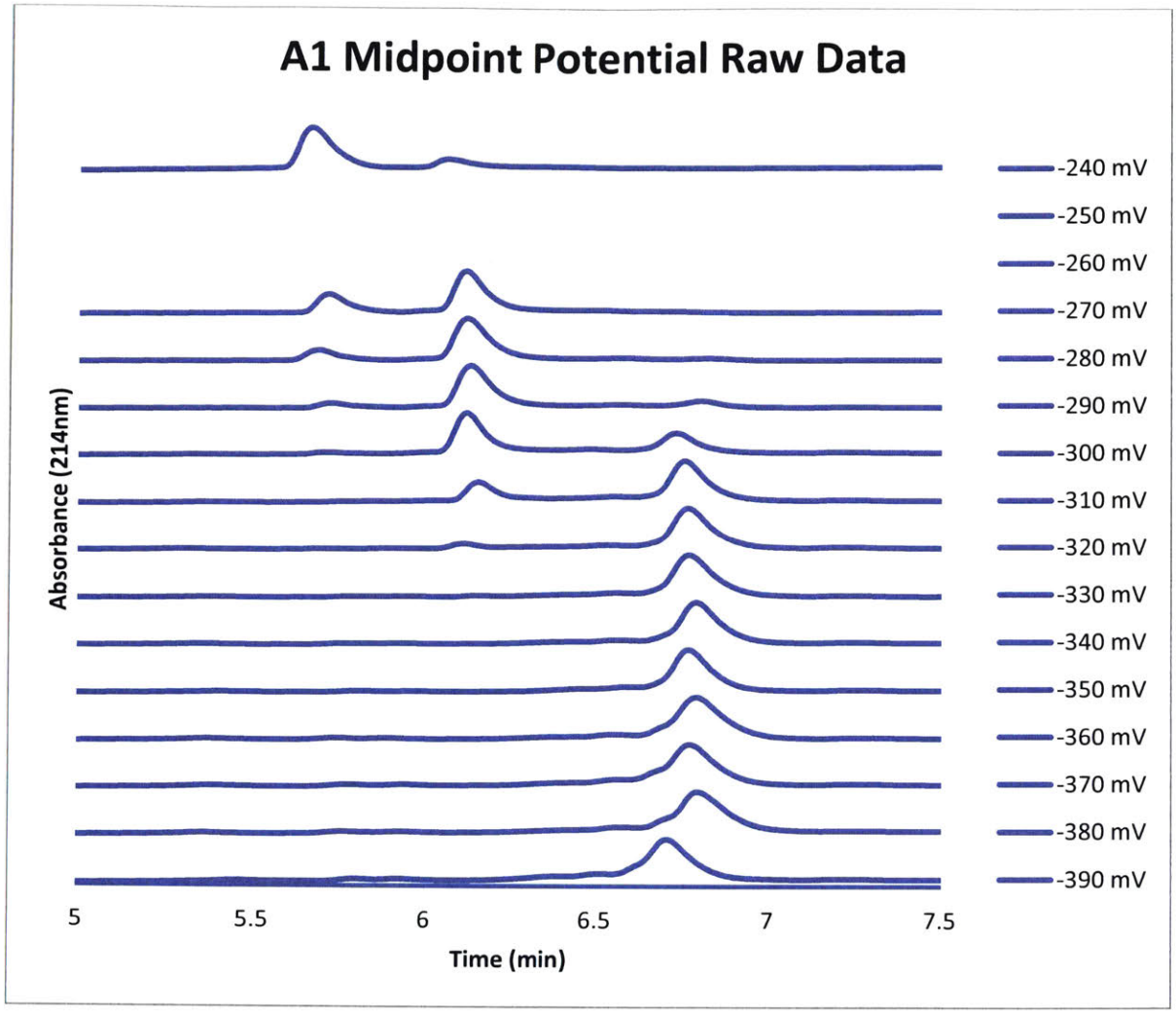
Trypsin binding affinity of analog P30.

A4.14 Raw midpoint potential data.

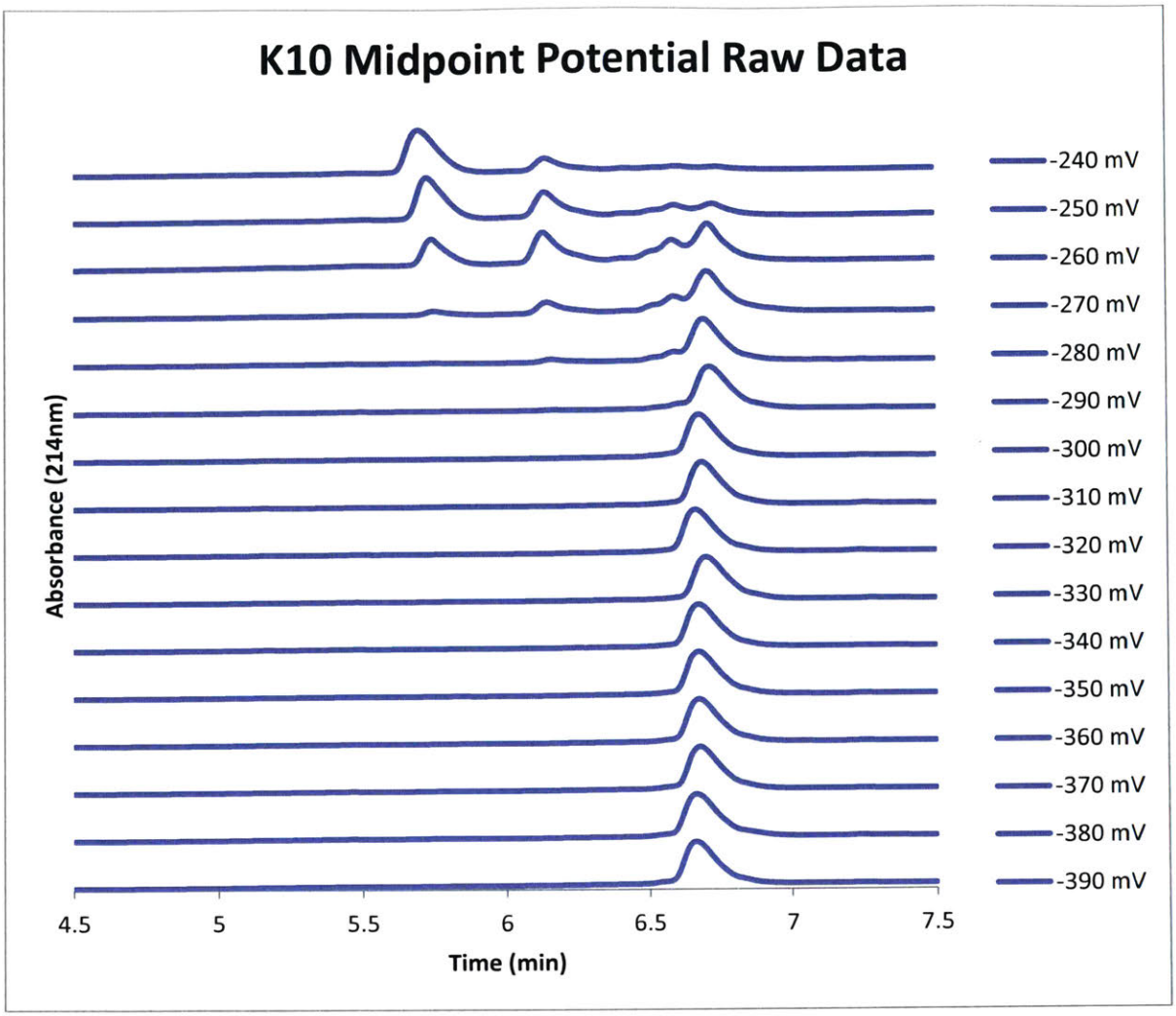
Each Figure A4.hows a series of chromatograms obtained after incubation in redox buffers of various potentials. In all cases, the earliest eluting peak is the folded, three disulfide product, the large, well resolved central peak is the two disulfide intermediate, and the latest eluting peak is the linear peptide. In some cases, a fourth peak can be observed, poorly resolved from the linear peptide. This is a one disulfide intermediate and is known to form in acid media for WT-EETI-II (for WT EETI-II below, it is an unresolved left shoulder under the chromatographic conditions used). The one disulfide product was integrated with the linear peptide.



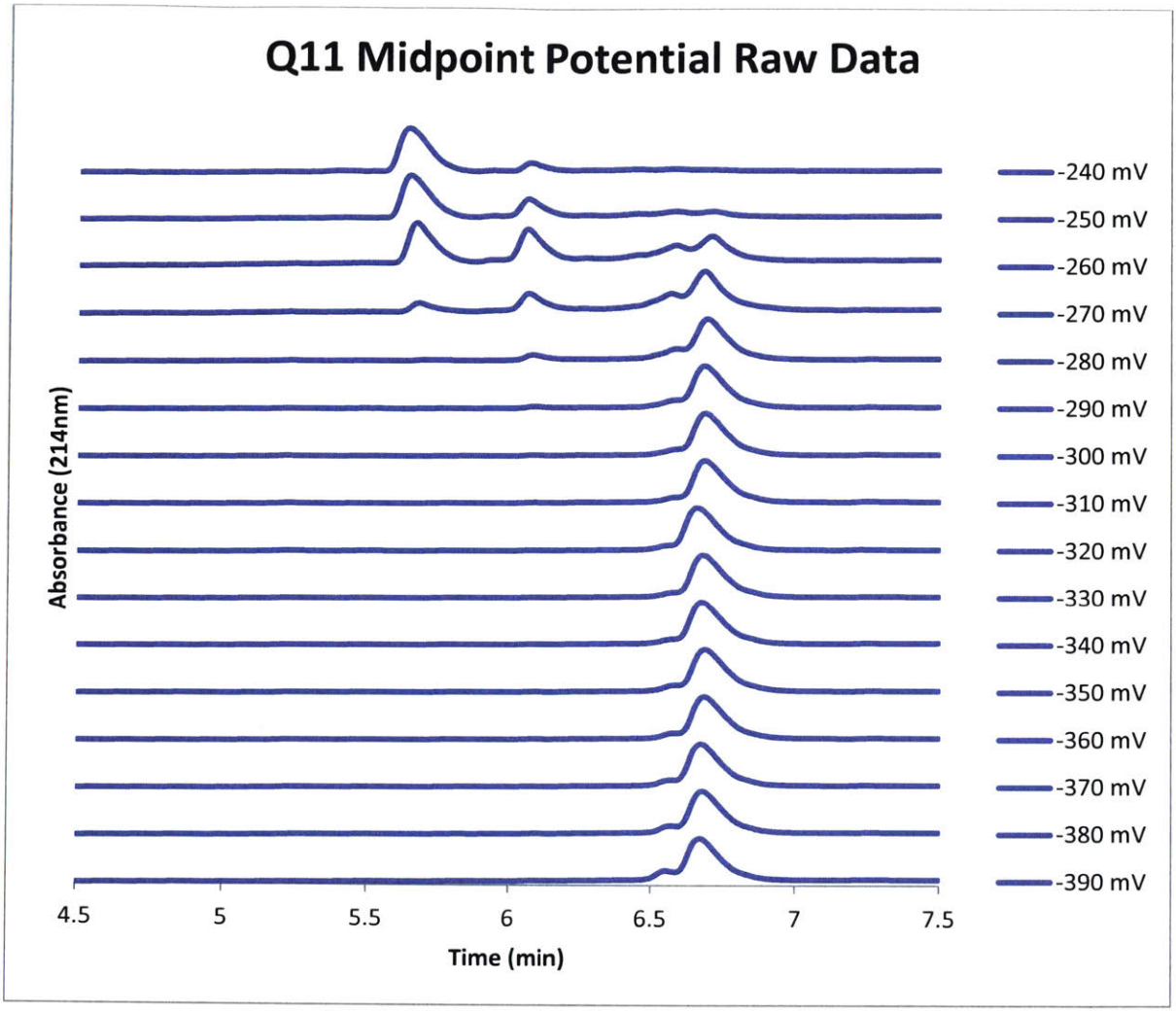
Raw chromatograms of WT EETI-II in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



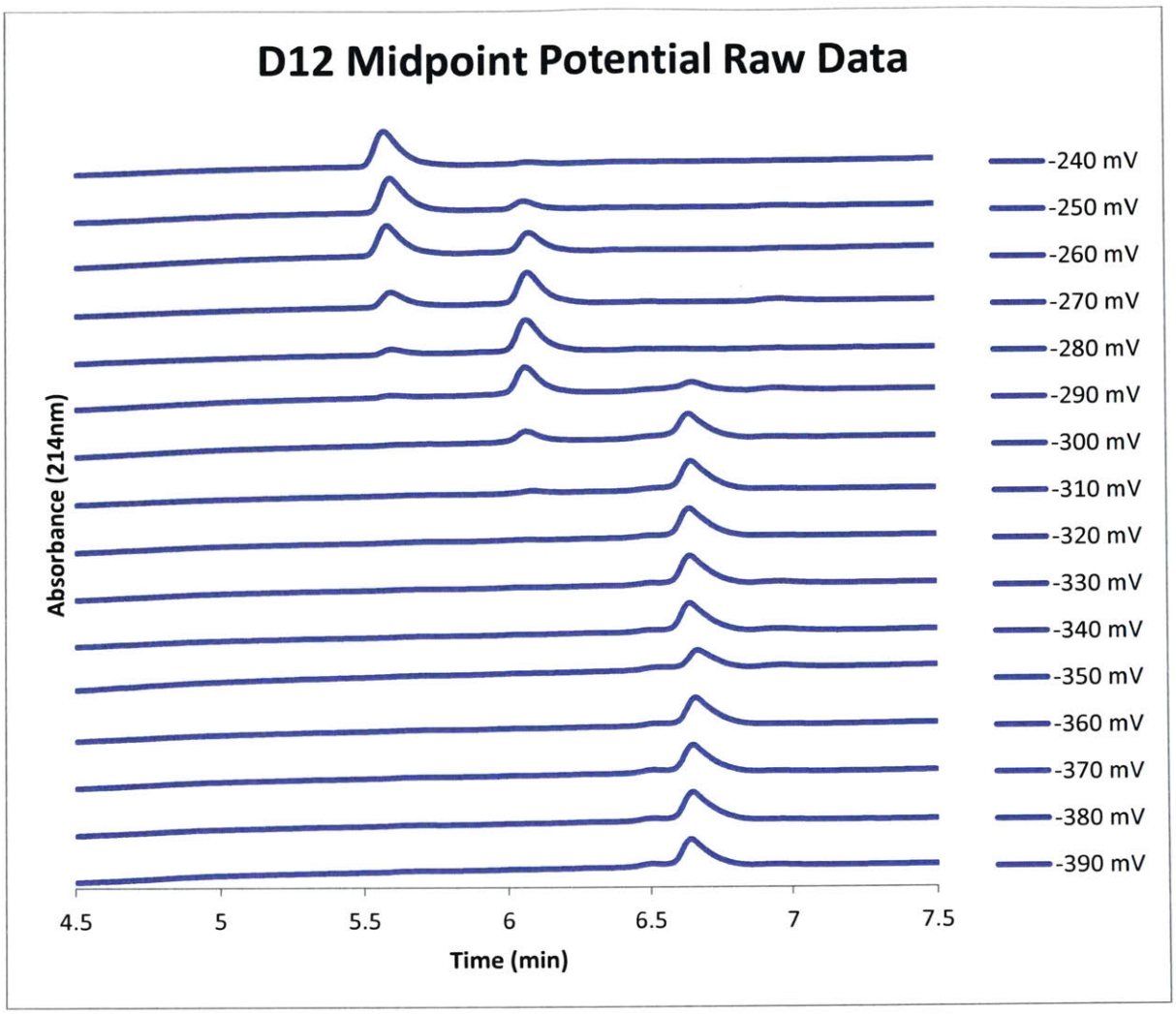
Raw chromatograms of EETI-II analog A1 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer. -250mV and -260mV trials were not run for analog A1.



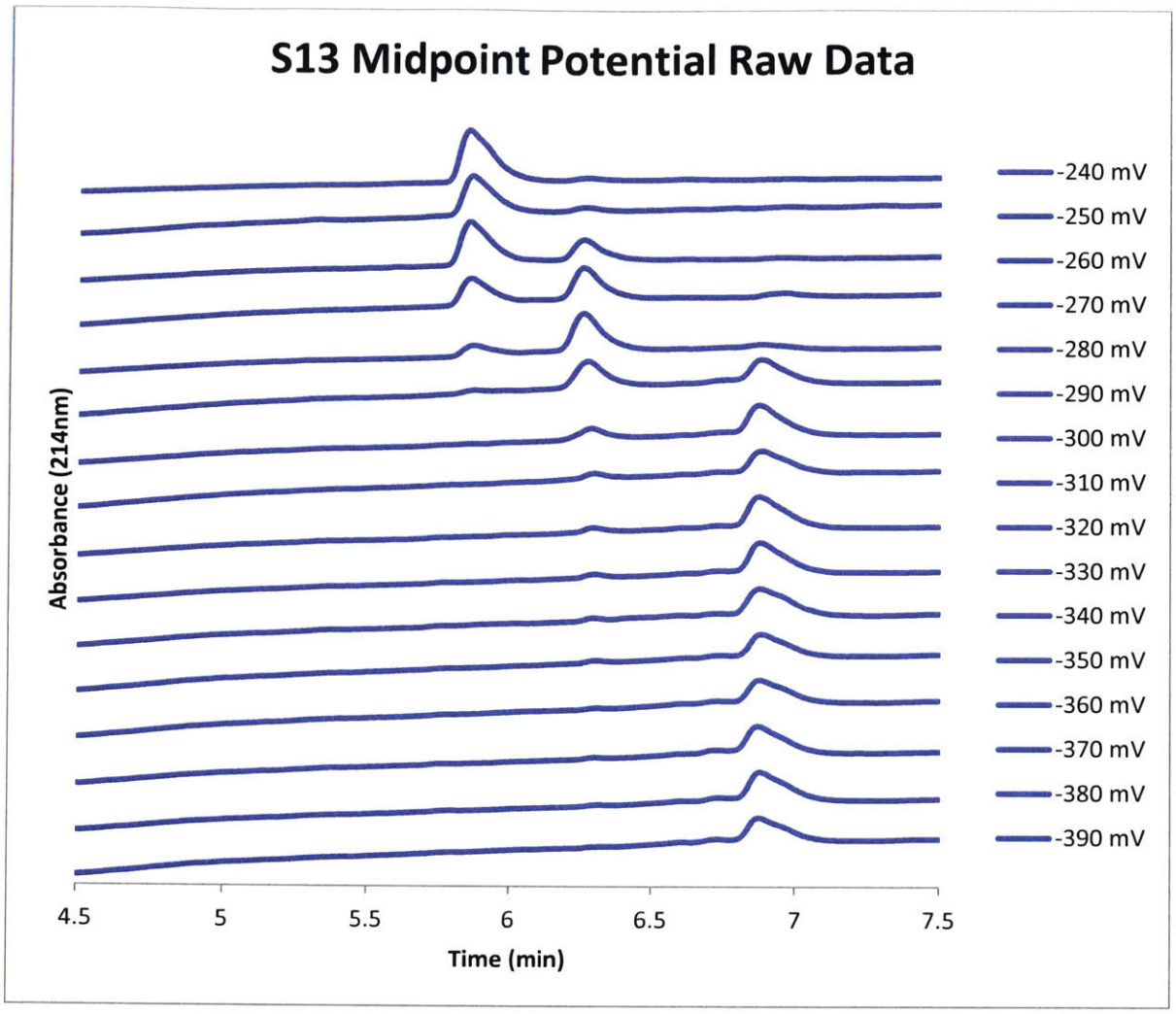
Raw chromatograms of EETI-II analog K10 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



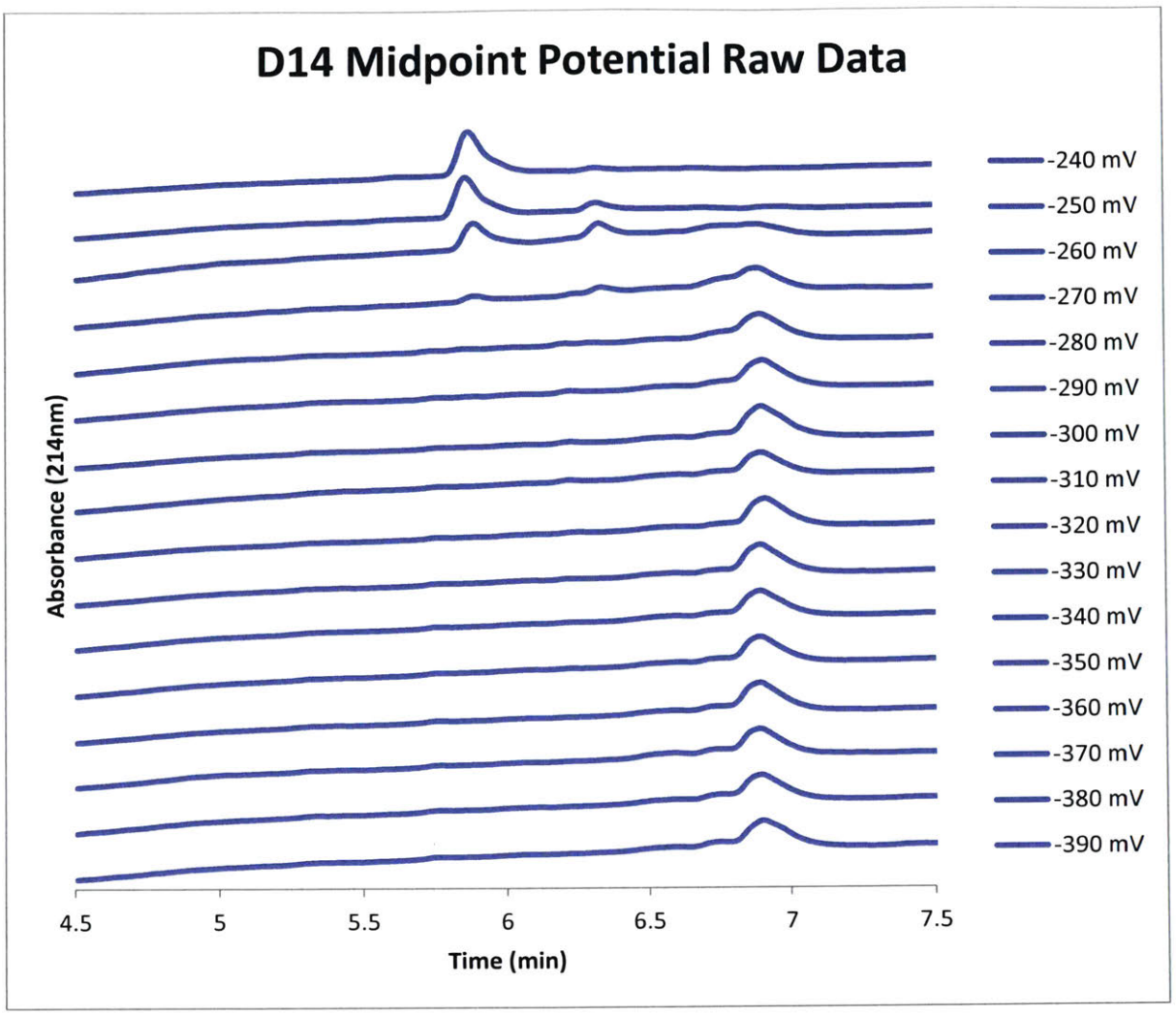
Raw chromatograms of EETI-II analog Q11 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



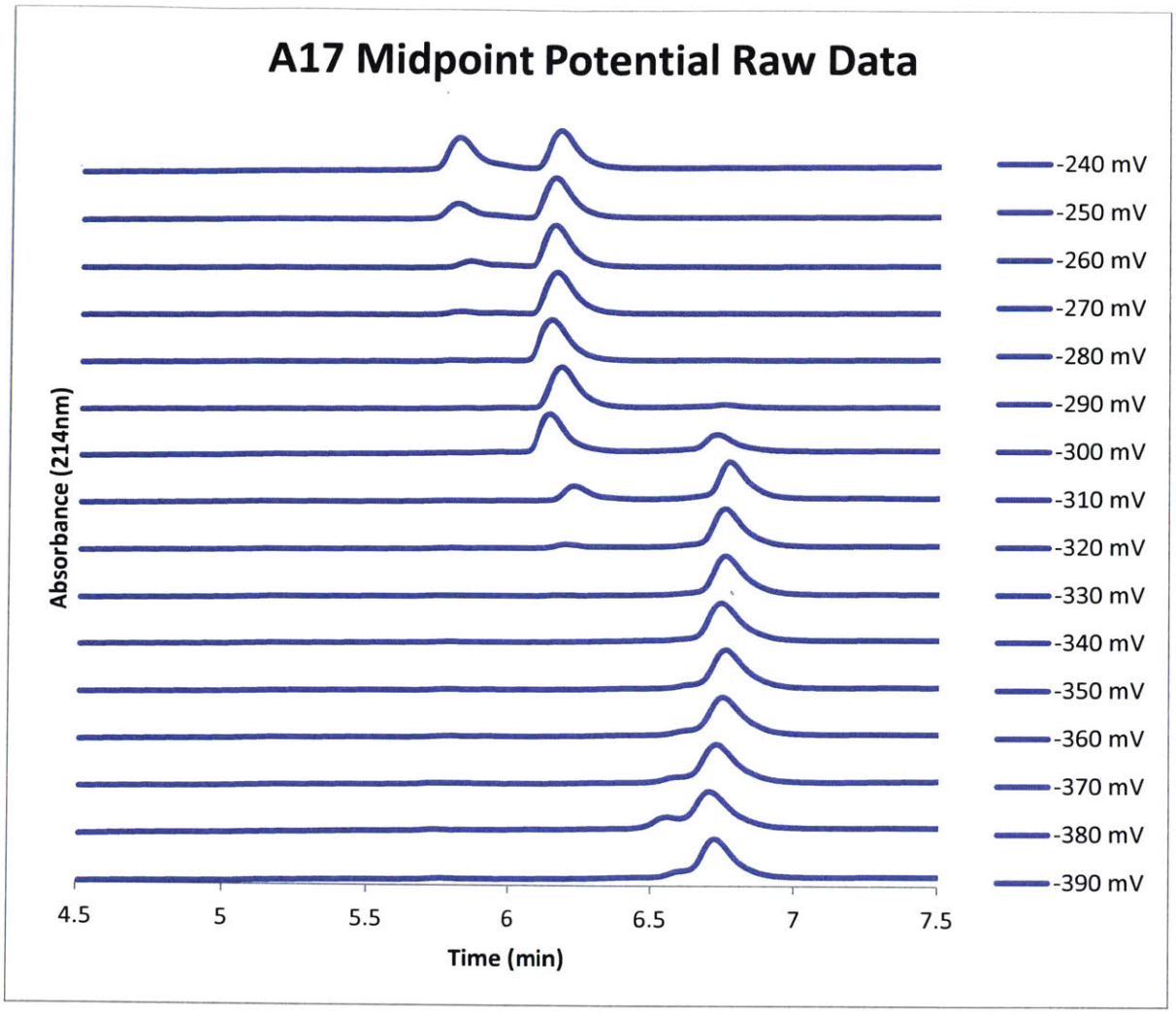
Raw chromatograms of EETI-II analog D12 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



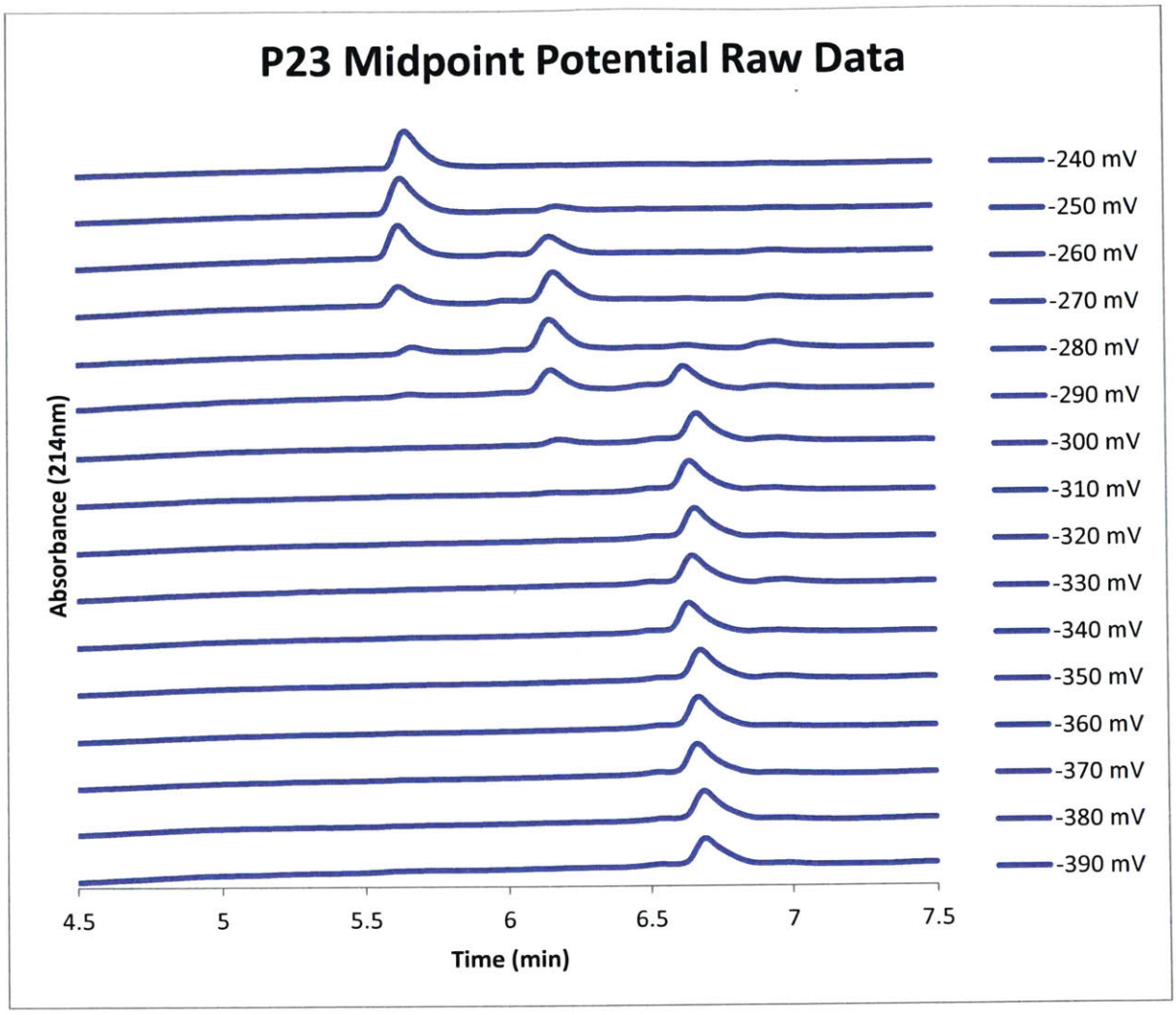
Raw chromatograms of EETI-II analog S13 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



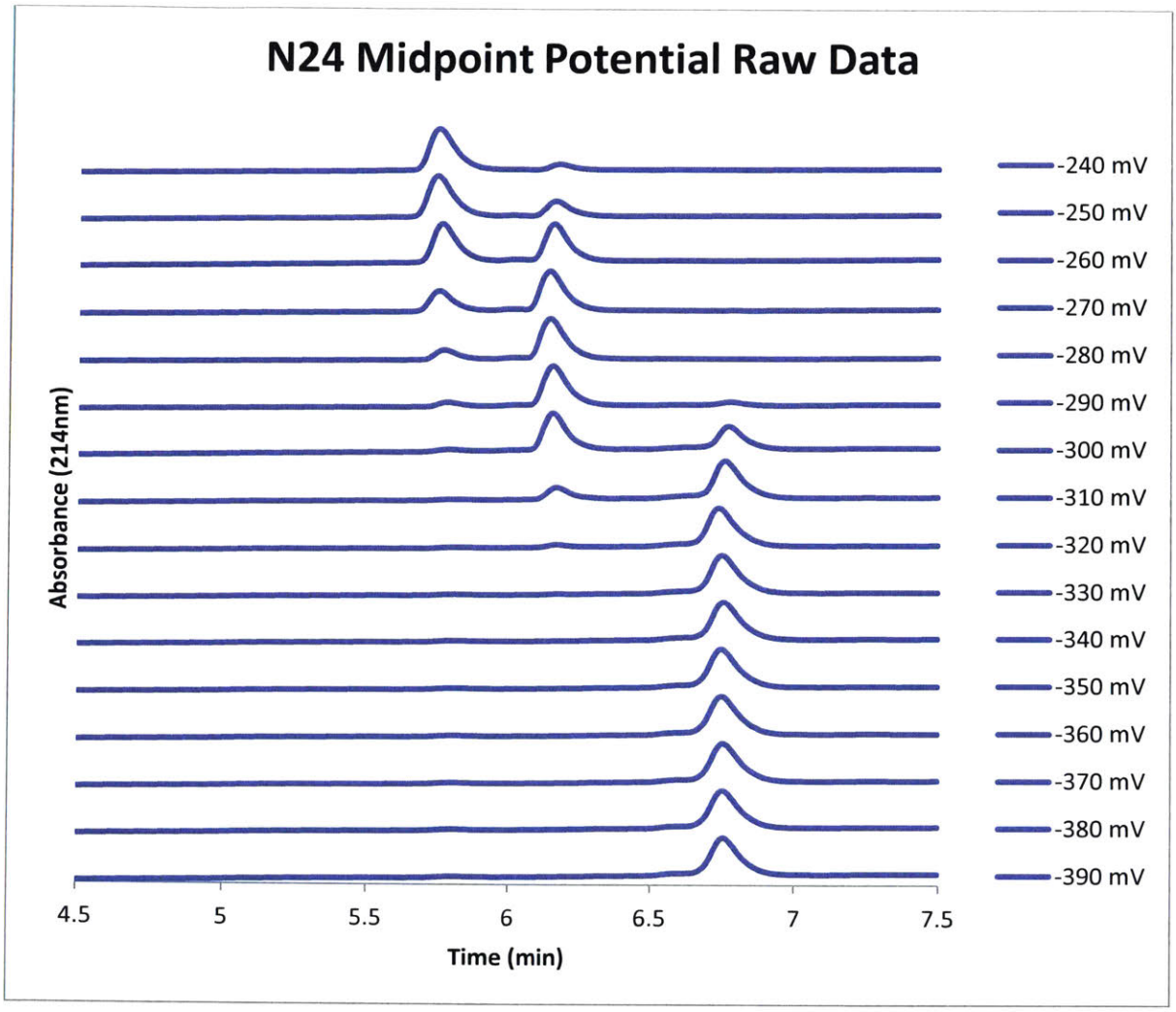
Raw chromatograms of EETI-II analog D14 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



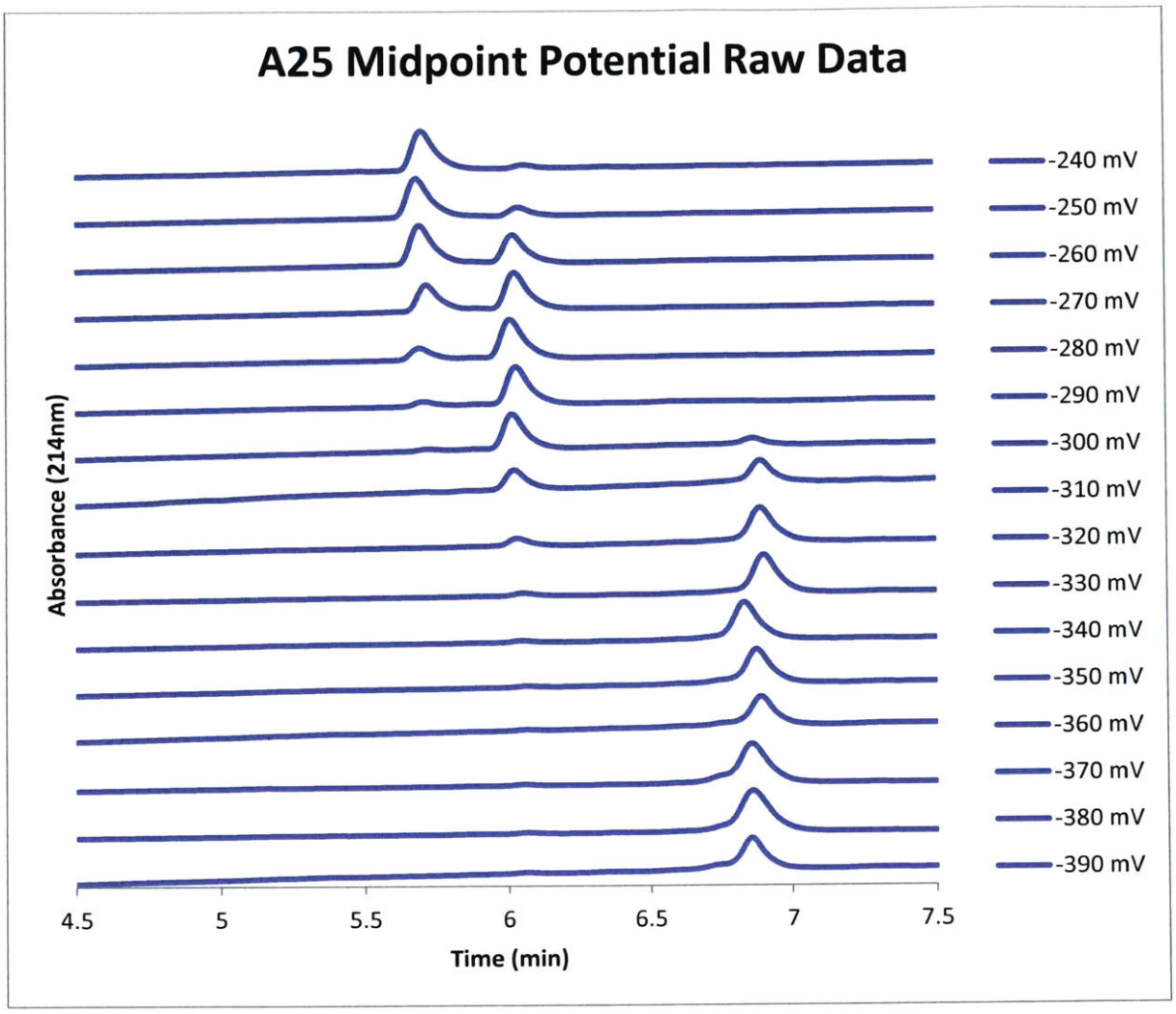
Raw chromatograms of EETI-II analog A17 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



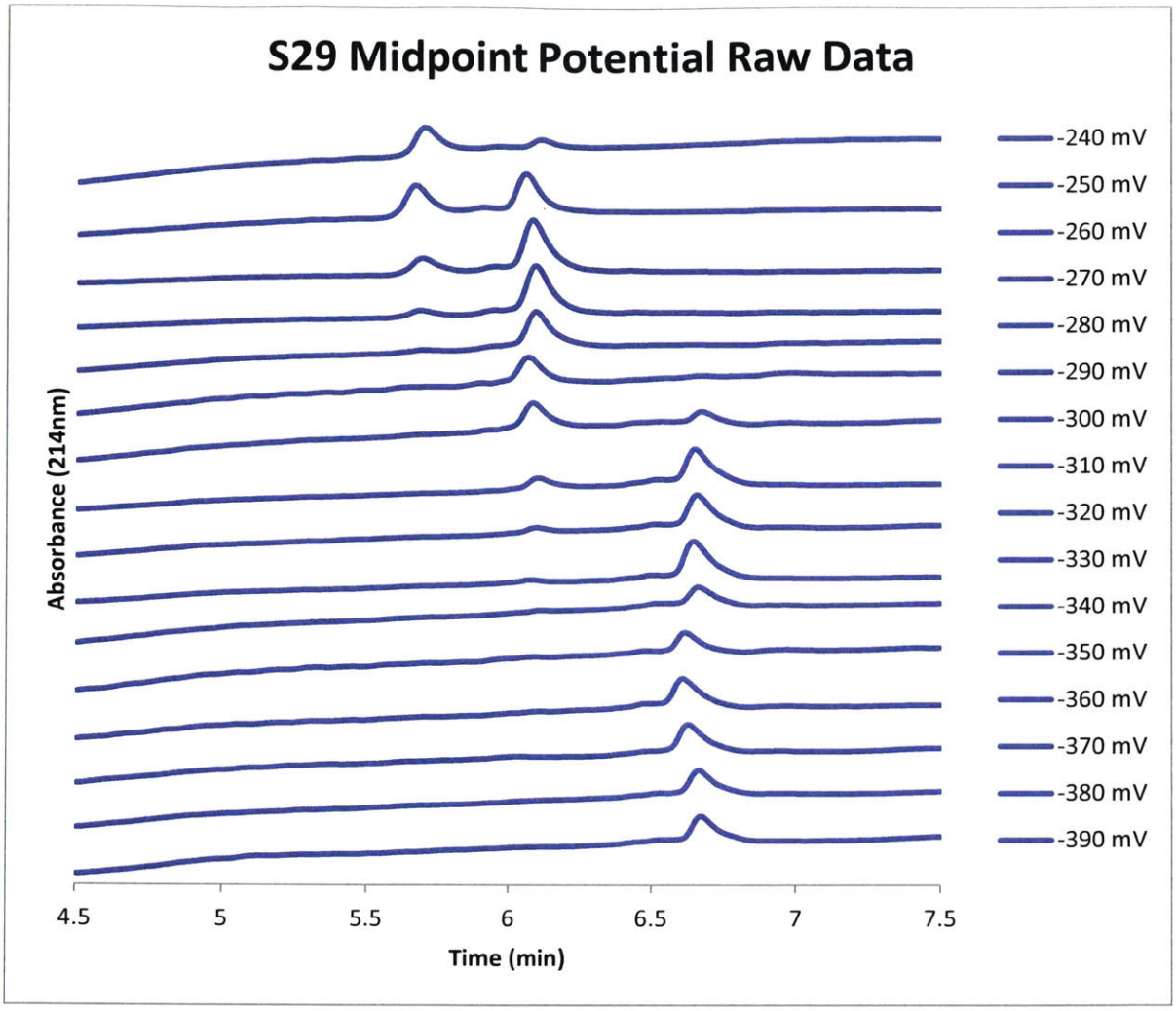
Raw chromatograms of EETI-II analog P23 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



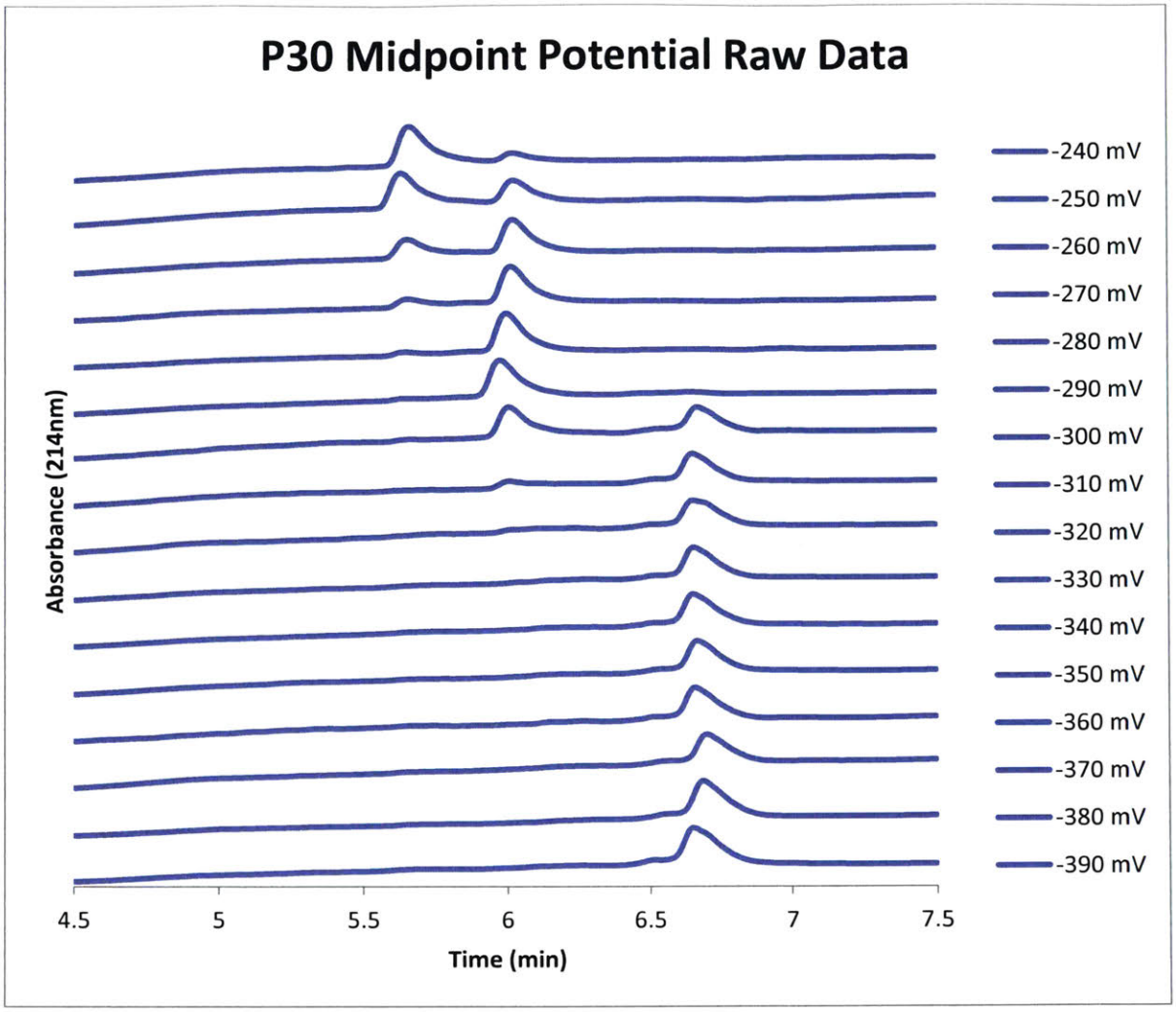
Raw chromatograms of EETI-II analog N24 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



Raw chromatograms of EETI-II analog A25 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



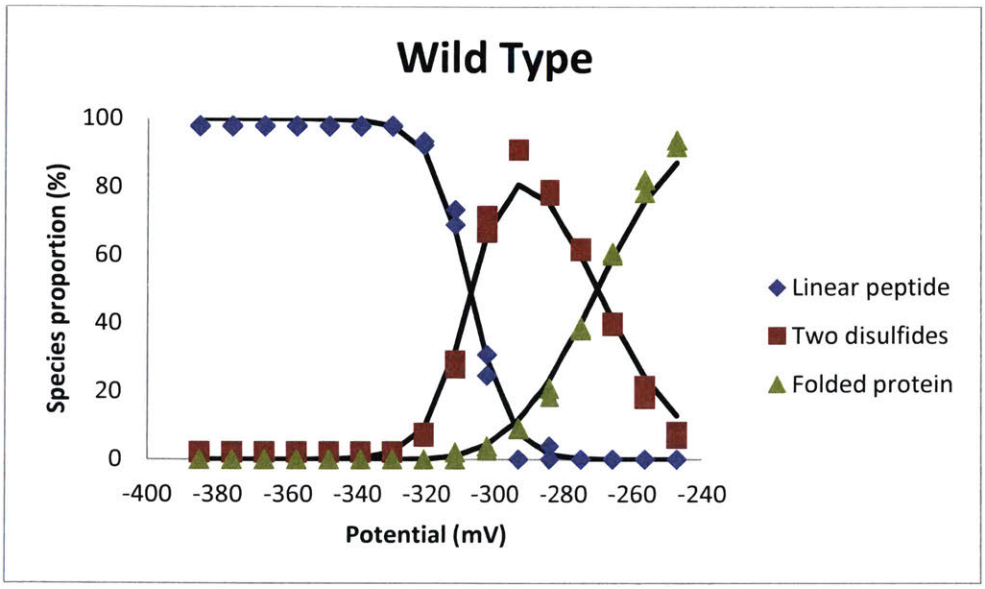
Raw chromatograms of EETI-II analog S29 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.



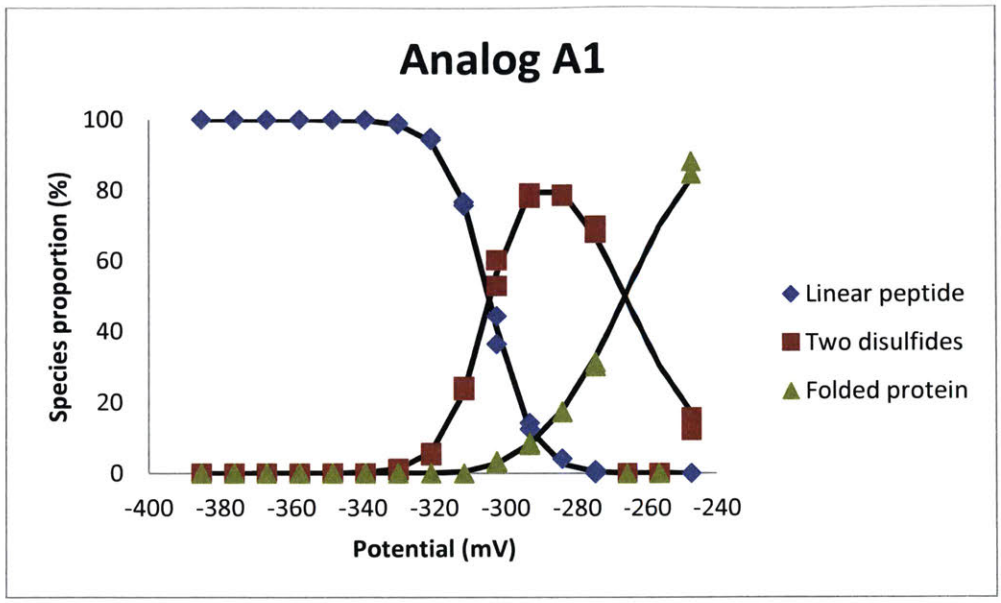
Raw chromatograms of EETI-II analog P30 in redox buffers of various potentials. The topmost trace is after incubation in a-240mV buffer and shows that the product is almost completely folded. The traces below are after incubation in buffers that are each 10mV more reducing than the previous buffer, with the bottom trace showing complete reduction to the linear peptide after incubation in a -390mV buffer.

A4.15 Fitted midpoint potential data

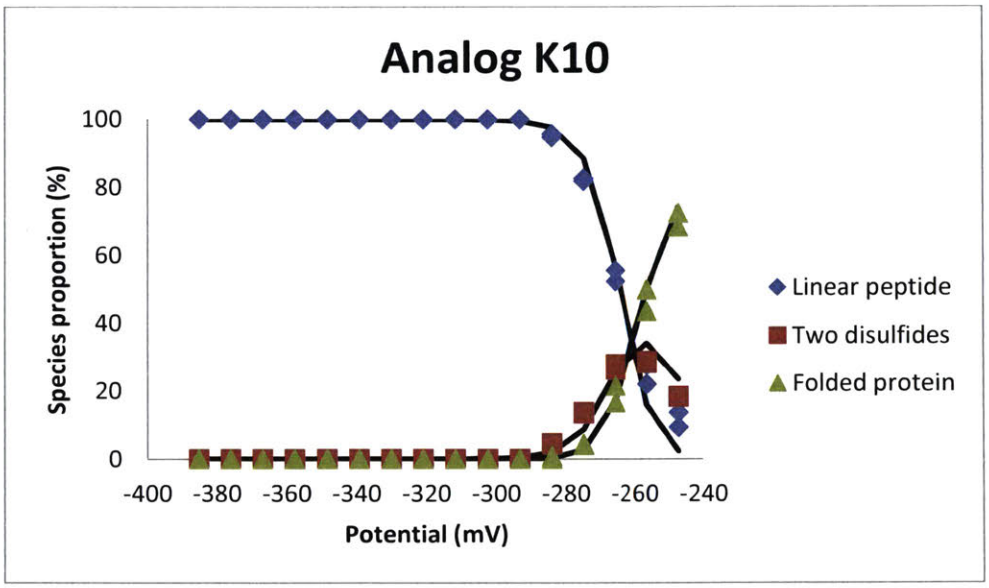
Complete data sets for the midpoint potential determination of the twelve active analogs are shown below. Each plot shows the observed proportions of folded protein (green triangles), two disulfide intermediate (red squares), and linear peptide (blue diamonds) at various potentials. Predicted values based on the fit of the Nernst equation are shown as black lines.



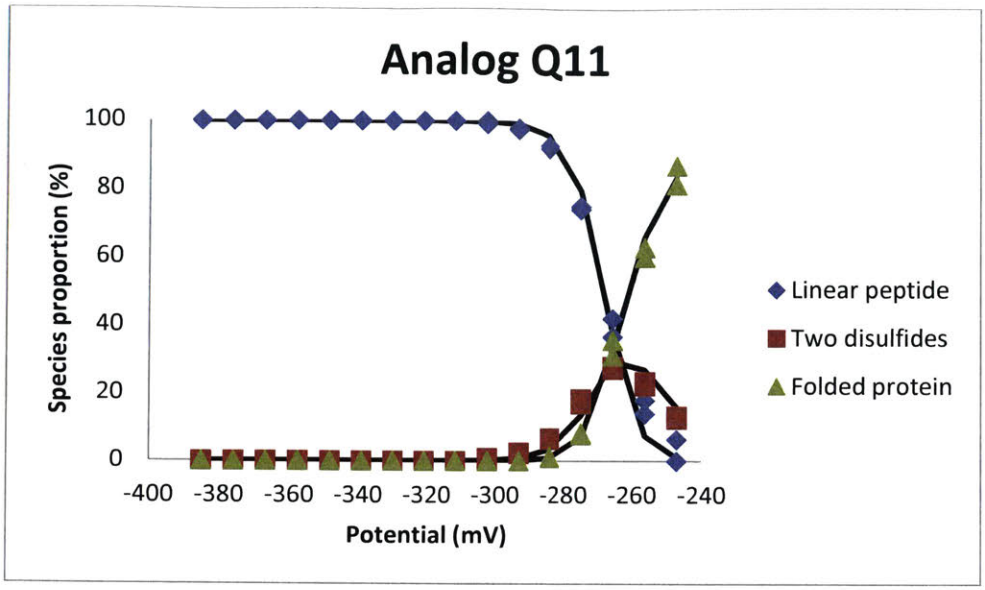
Midpoint potential determination for WT EETI-II. Black lines are fits.



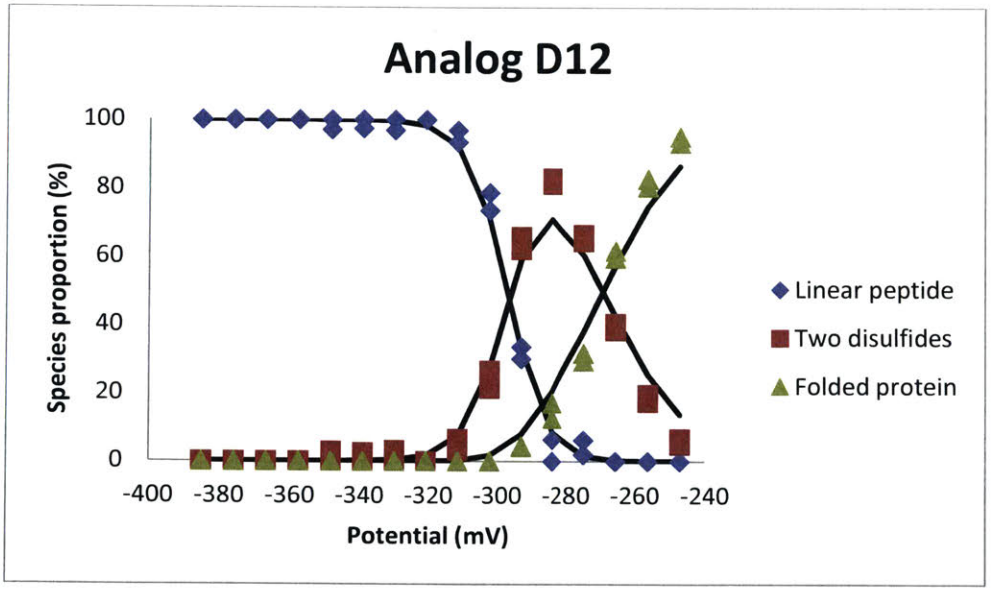
Midpoint potential determination for analog A1. Black lines are fits.



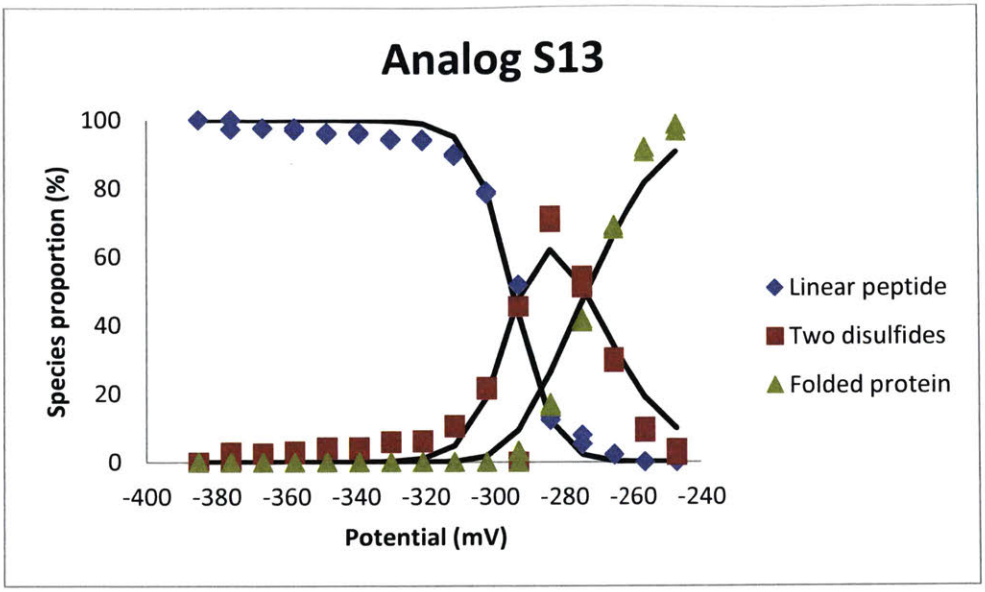
Midpoint potential determination for analog K10. Black lines are fits.



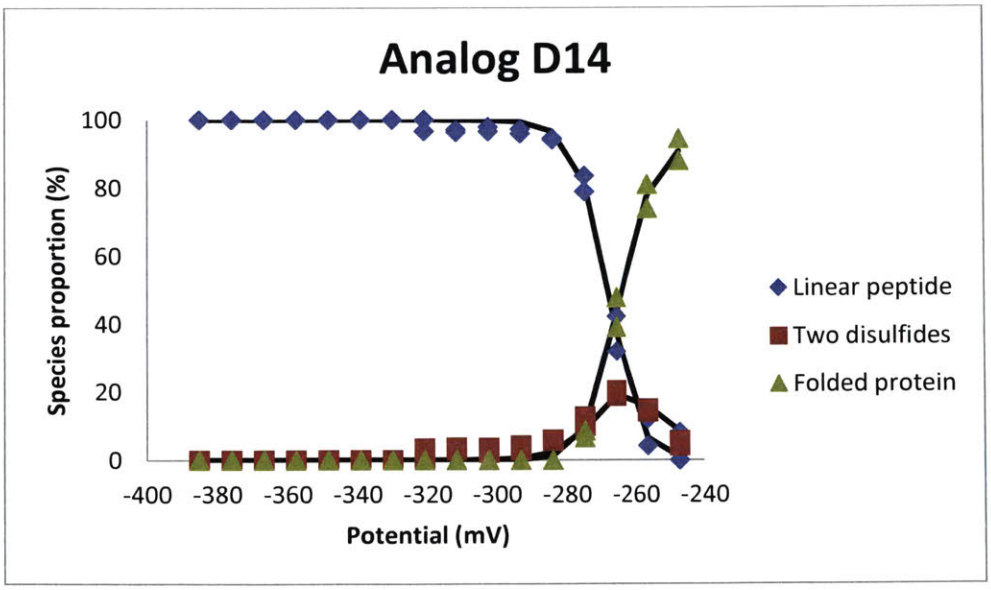
Midpoint potential determination for analog Q11. Black lines are fits.



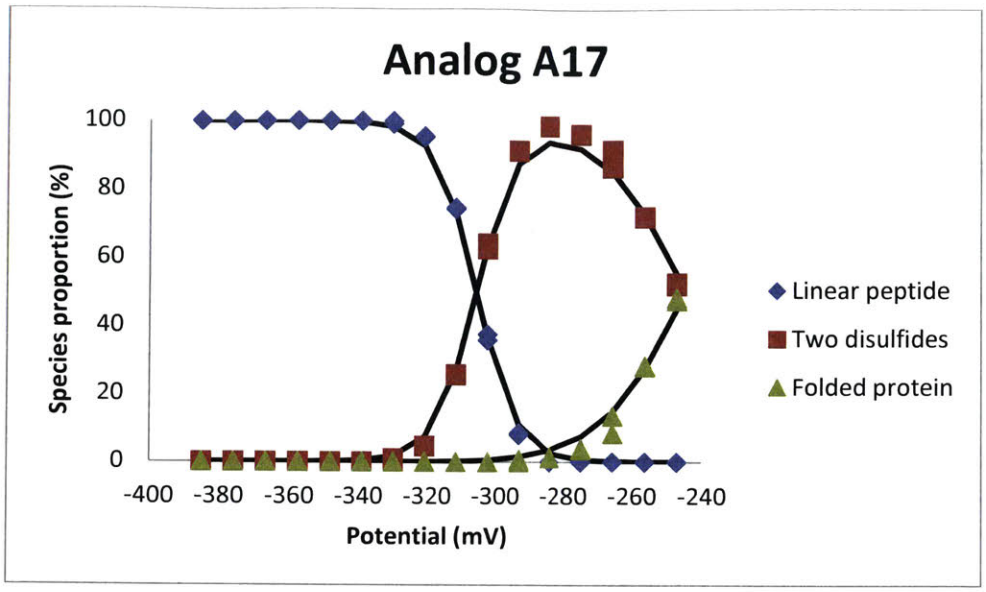
Midpoint potential determination for analog D12. Black lines are fits.



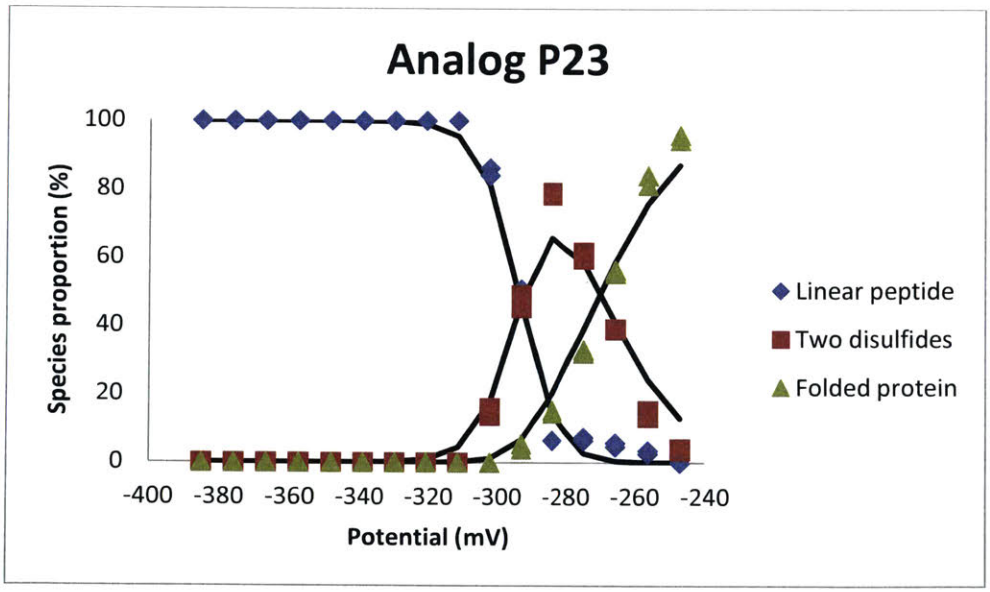
Midpoint potential determination for analog S13. Black lines are fits.



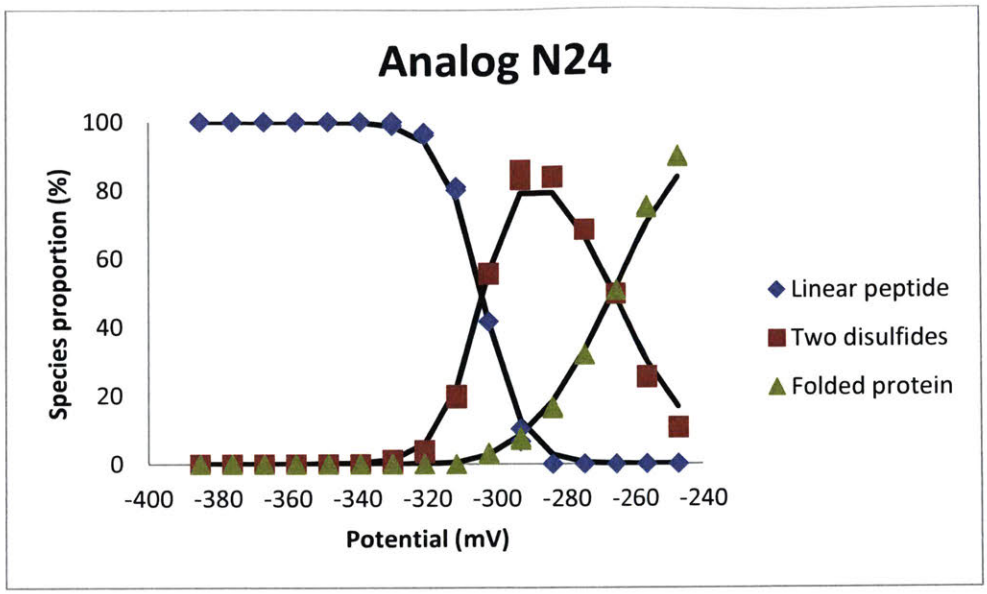
Midpoint potential determination for analog D14. Black lines are fits.



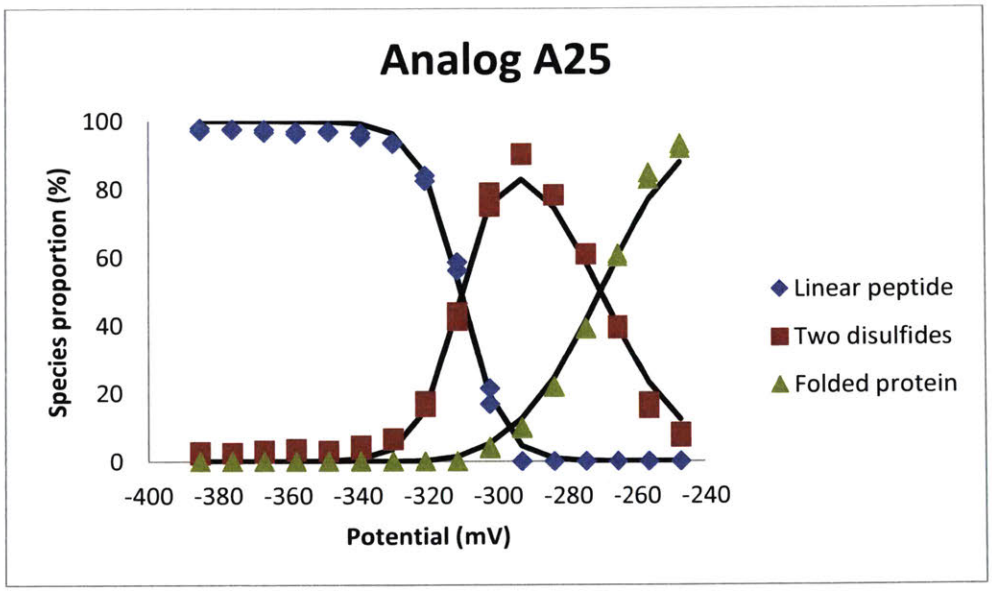
Midpoint potential determination for analog A17. Black lines are fits.



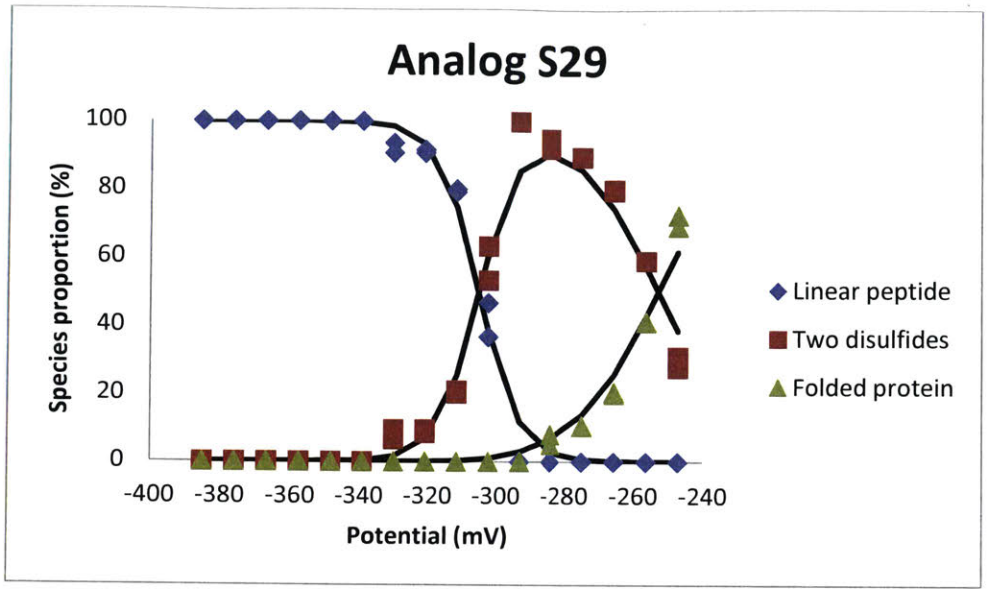
Midpoint potential determination for analog P23. Black lines are fits.



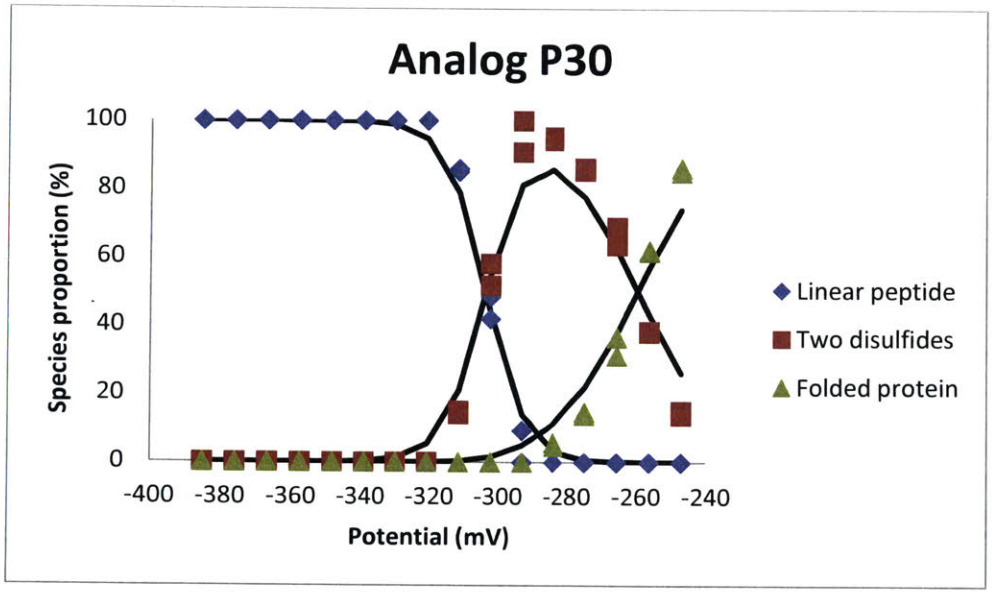
Midpoint potential determination for analog N24. Black lines are fits.



Midpoint potential determination for analog A25. Black lines are fits.



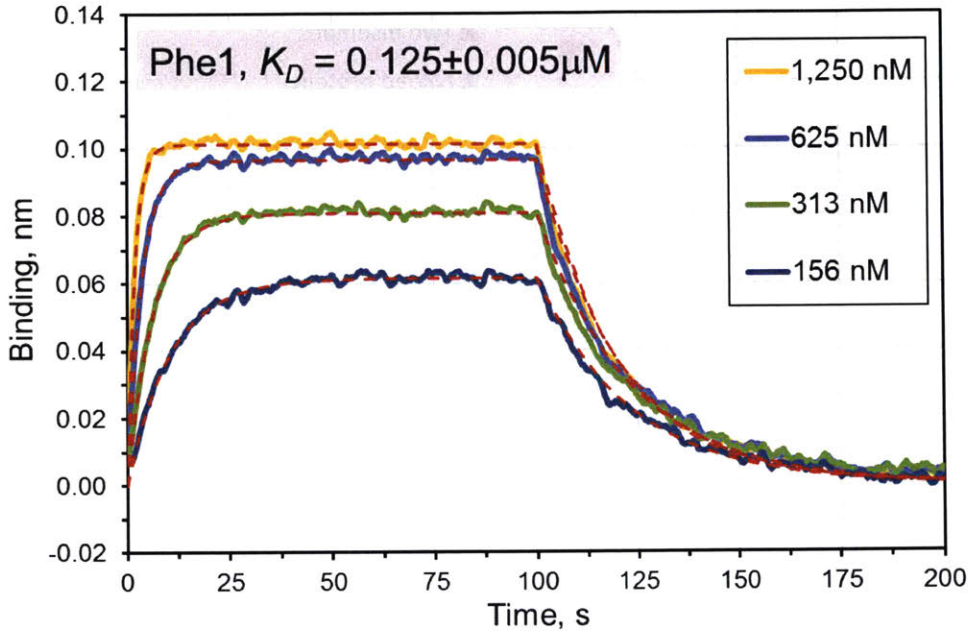
Midpoint potential determination for analog S29. Black lines are fits.

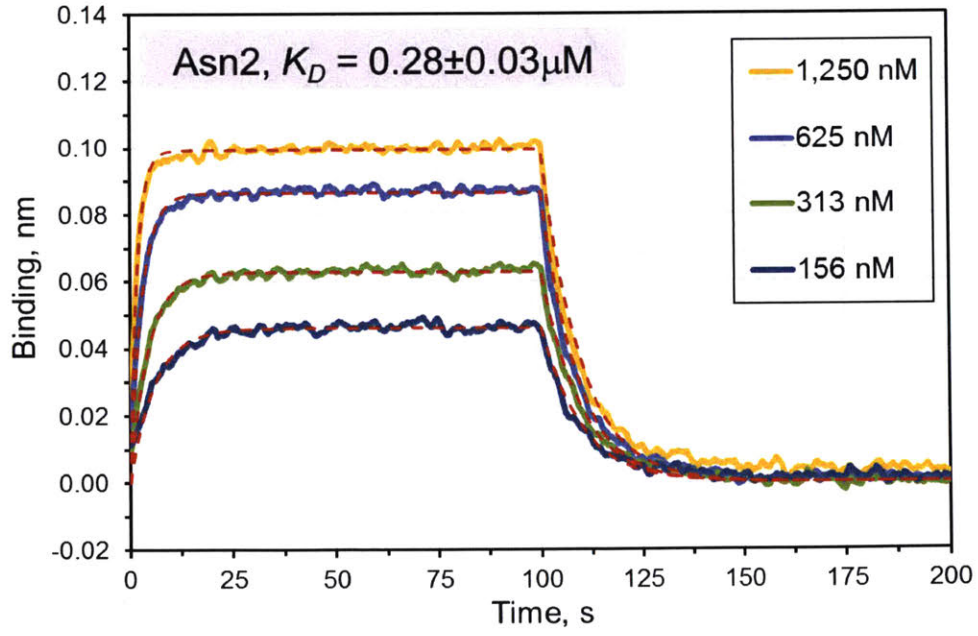


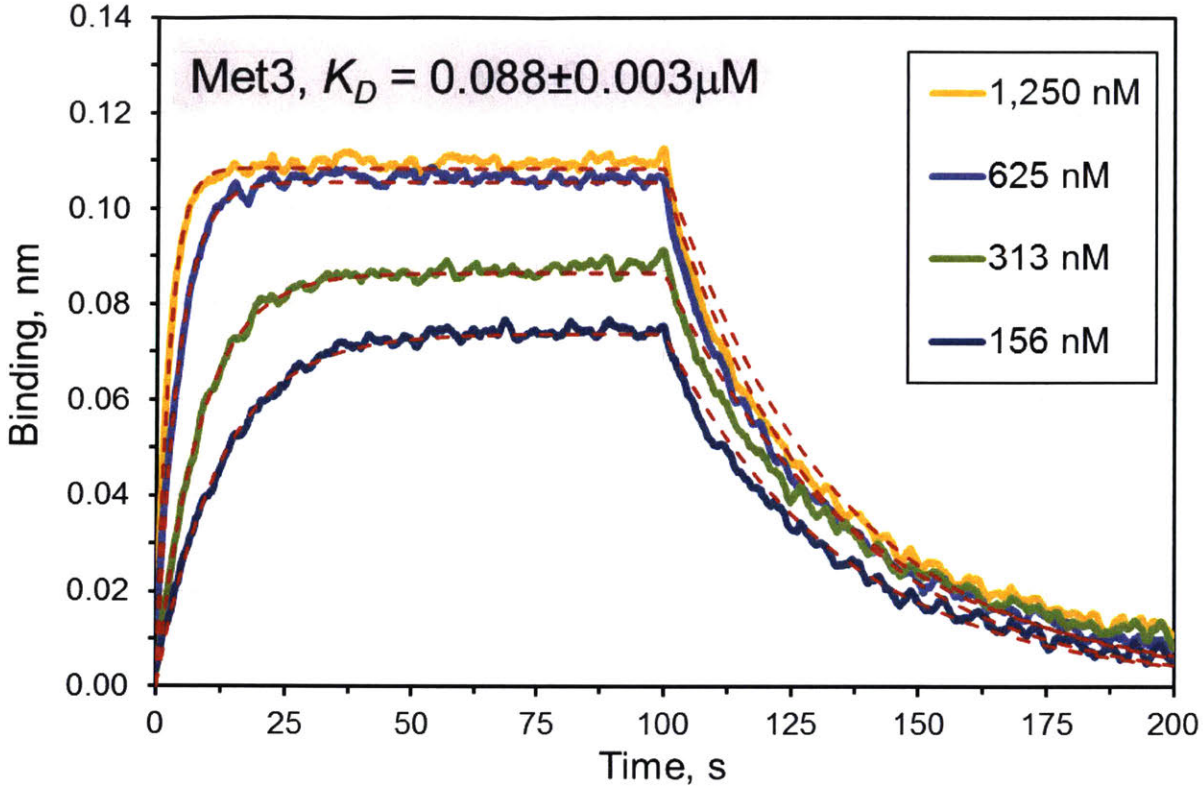
Midpoint potential determination for analog P30. Black lines are fits.

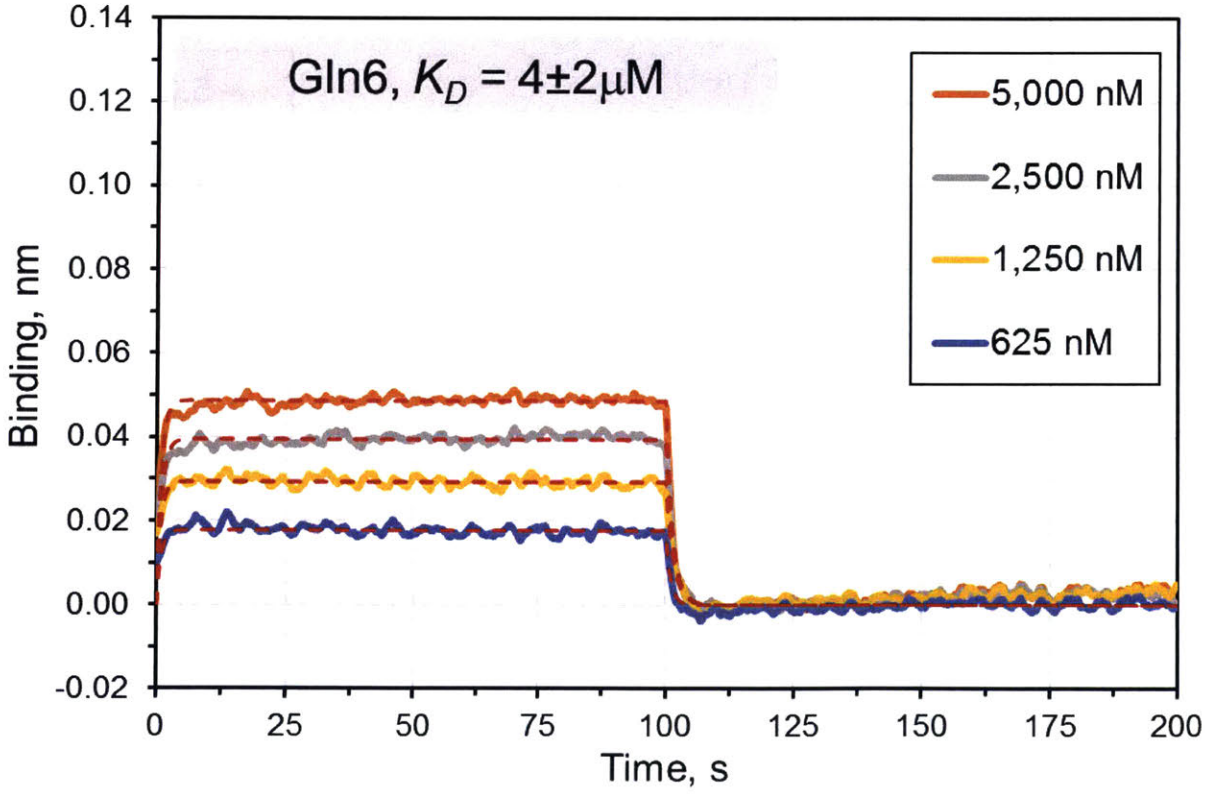
A4.16 Fitted Bilayer Interferometry Sensograms

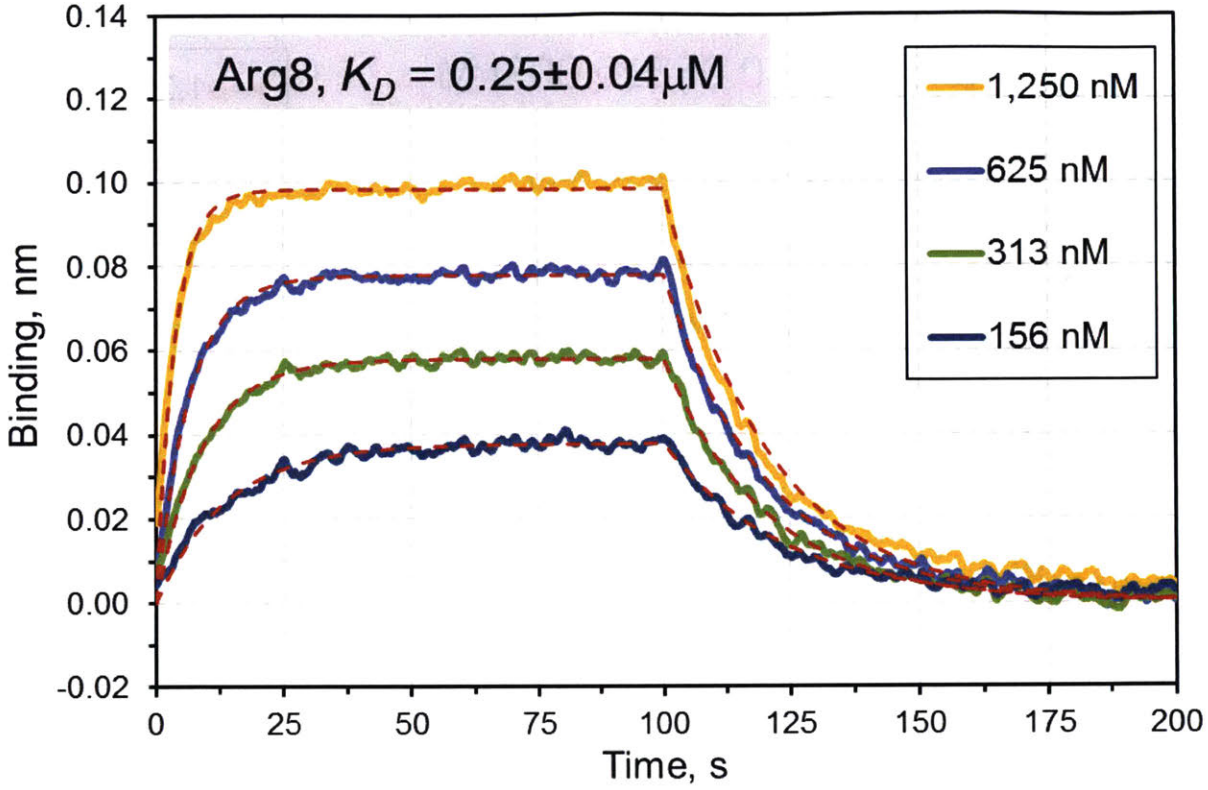
Unless otherwise noted, in each sensogram, the concentration of Z33 analog in each experiment was shown in the figure legend. The red dotted line was the fitted curve for each measurement. For experiments where no association was observed, K_D was reported as higher than the highest concentration tested.

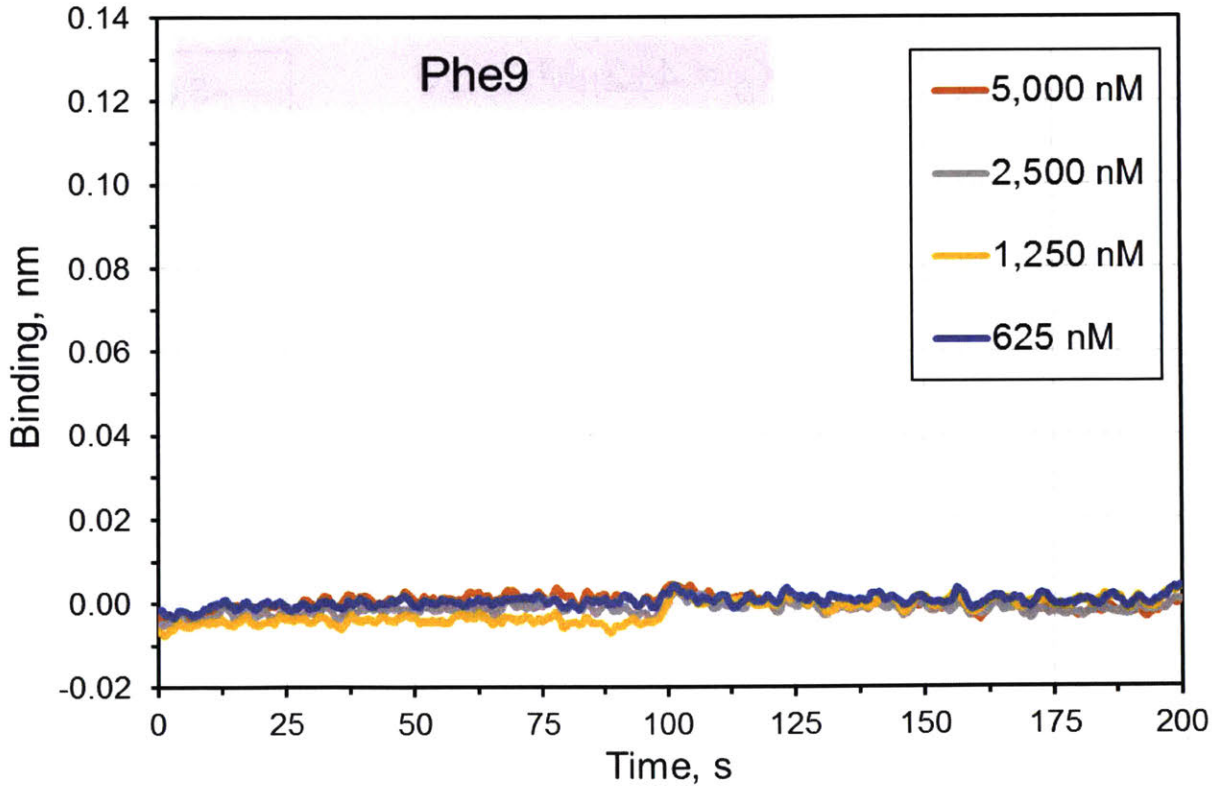


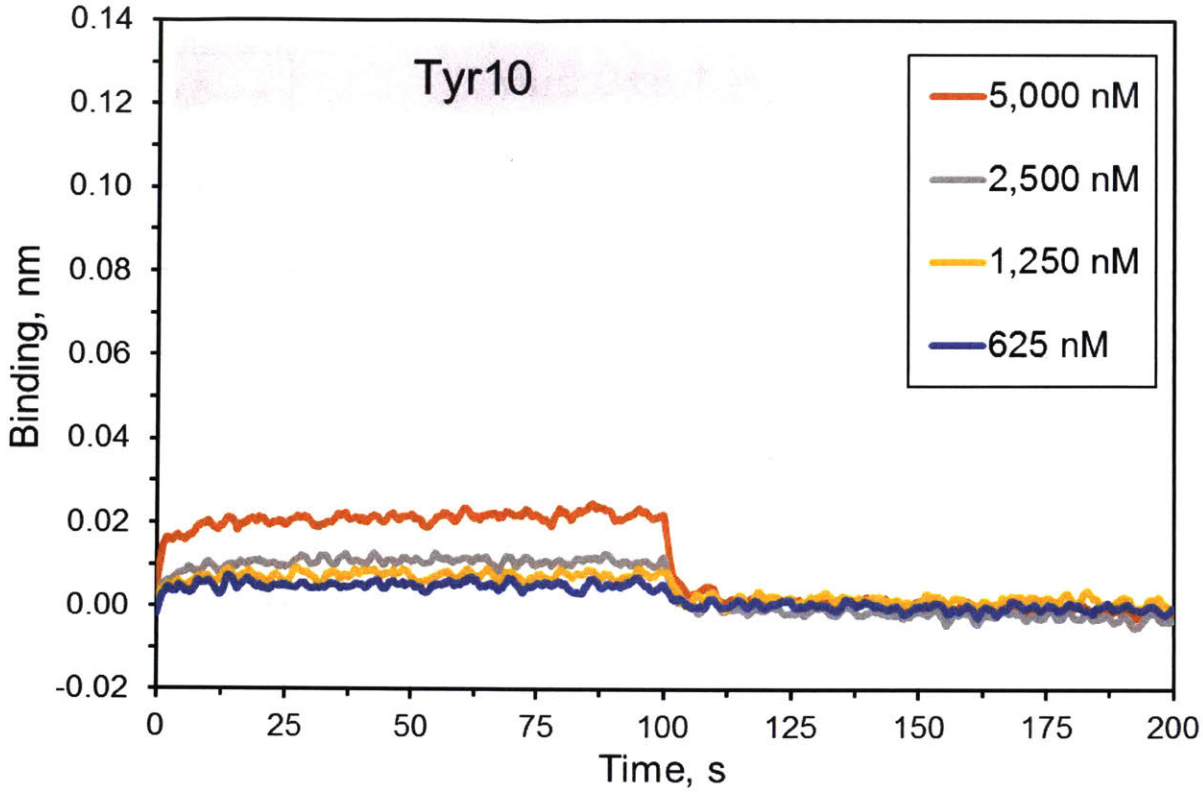


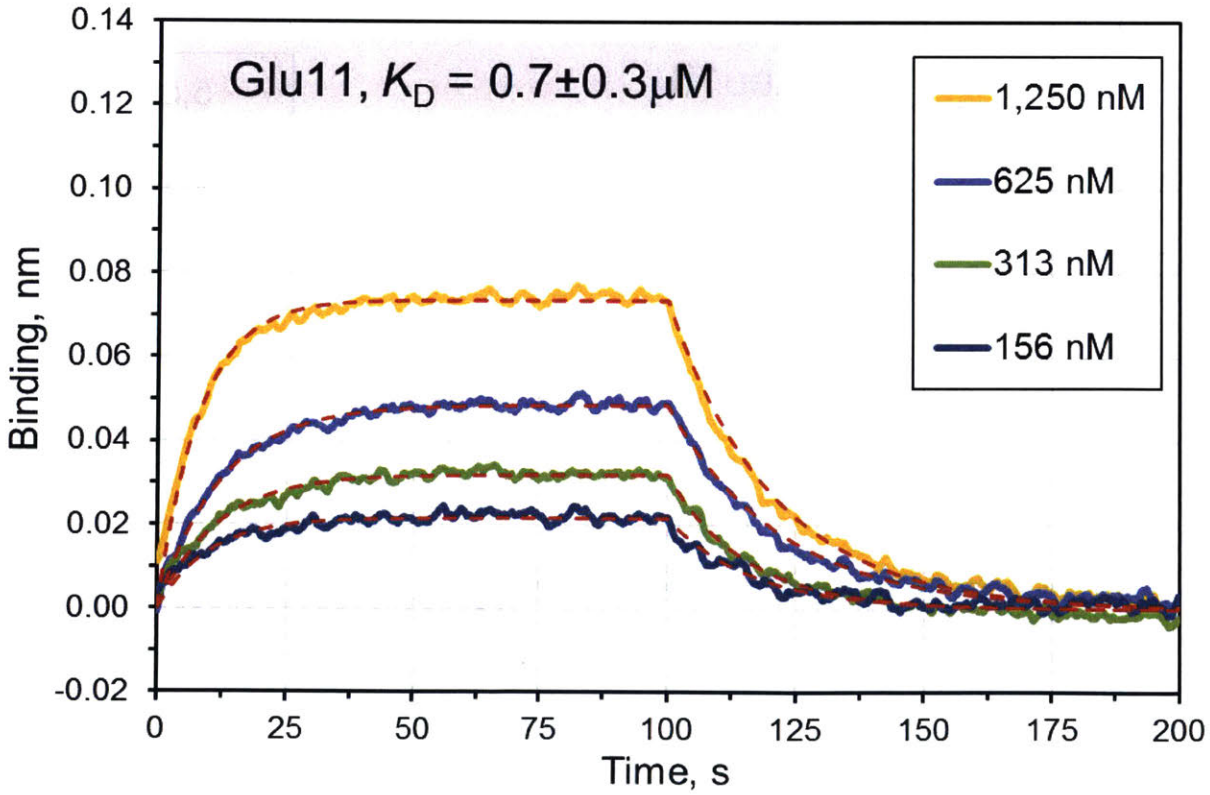

Met3, K_D = 0.088±0.003μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
25
50
75
100
125
150
175
200

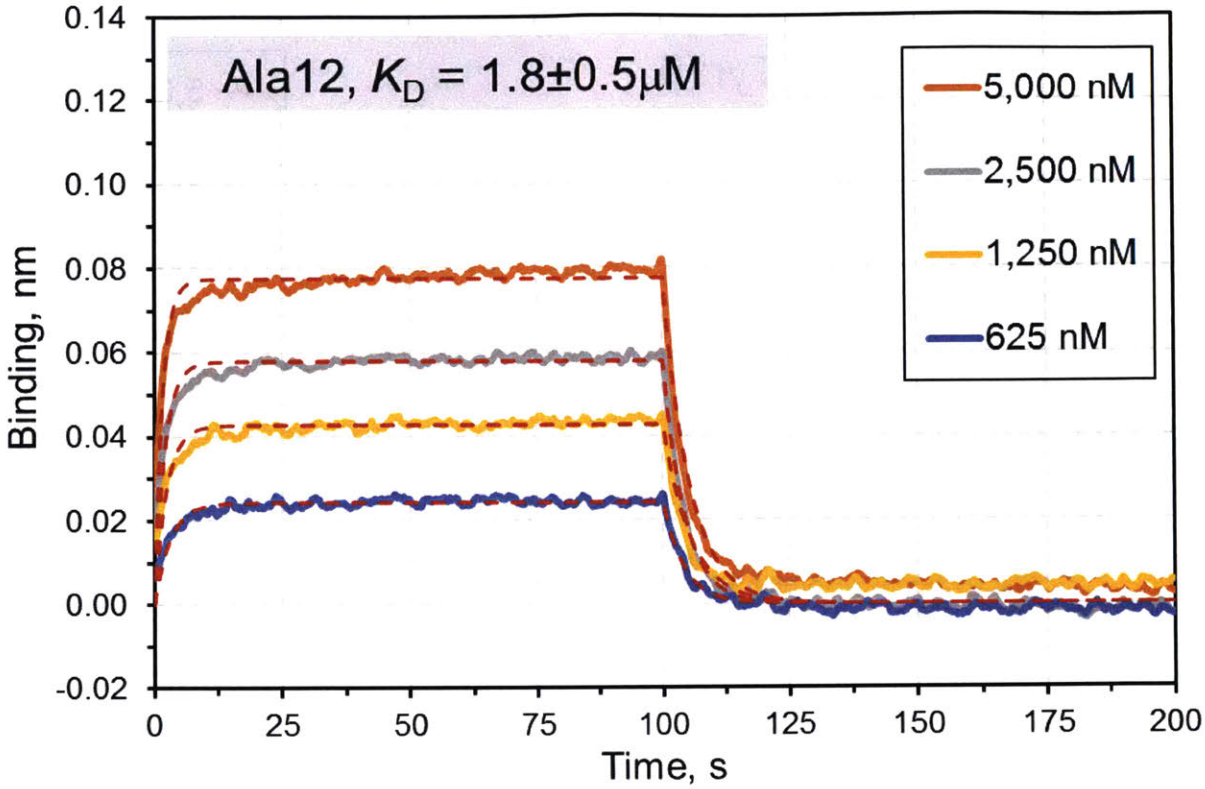

Gln6, K_D = 4±2μM
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

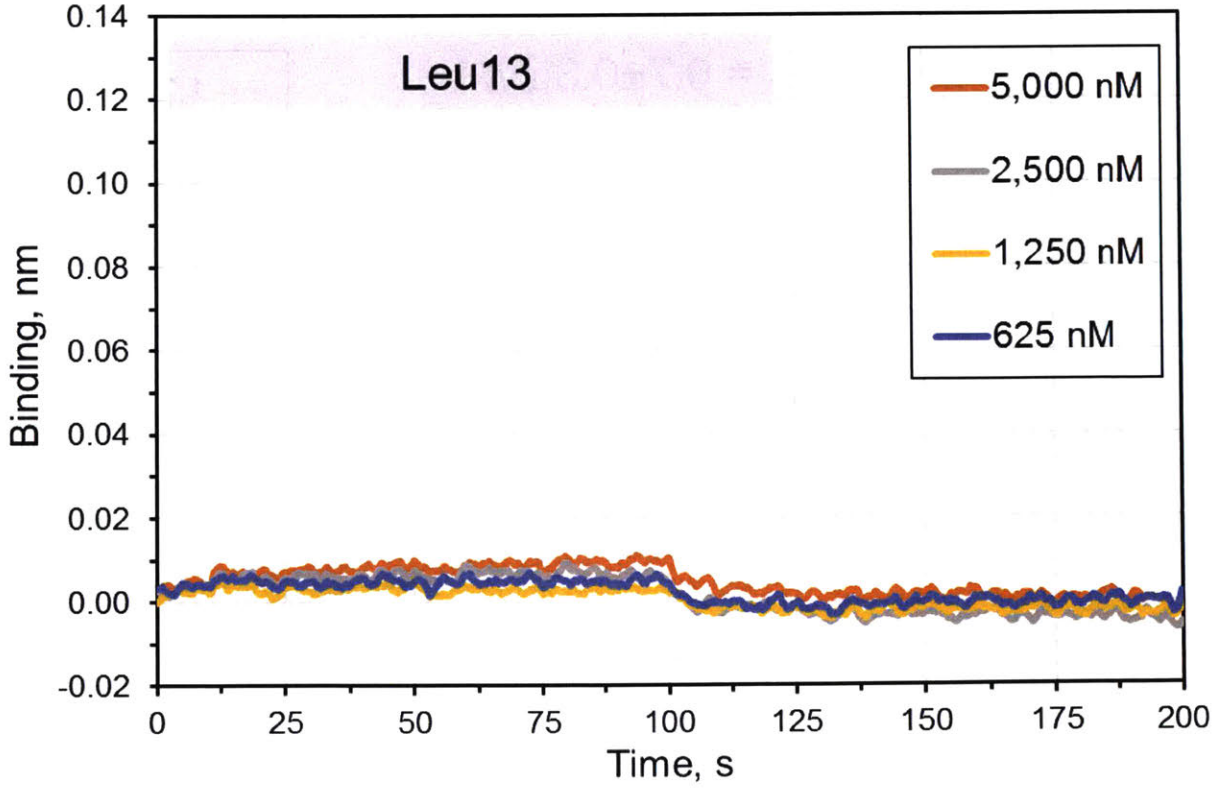

Arg8, K_D = 0.25±0.04μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

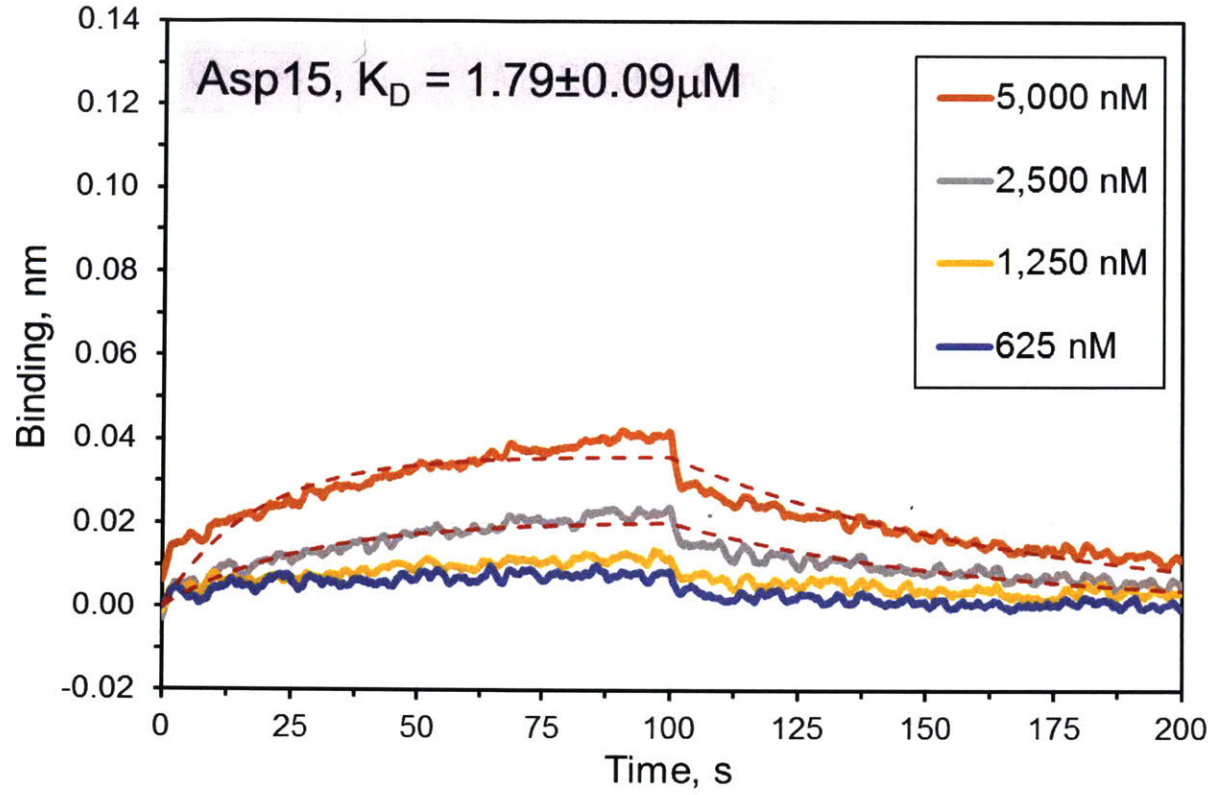

Phe9
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

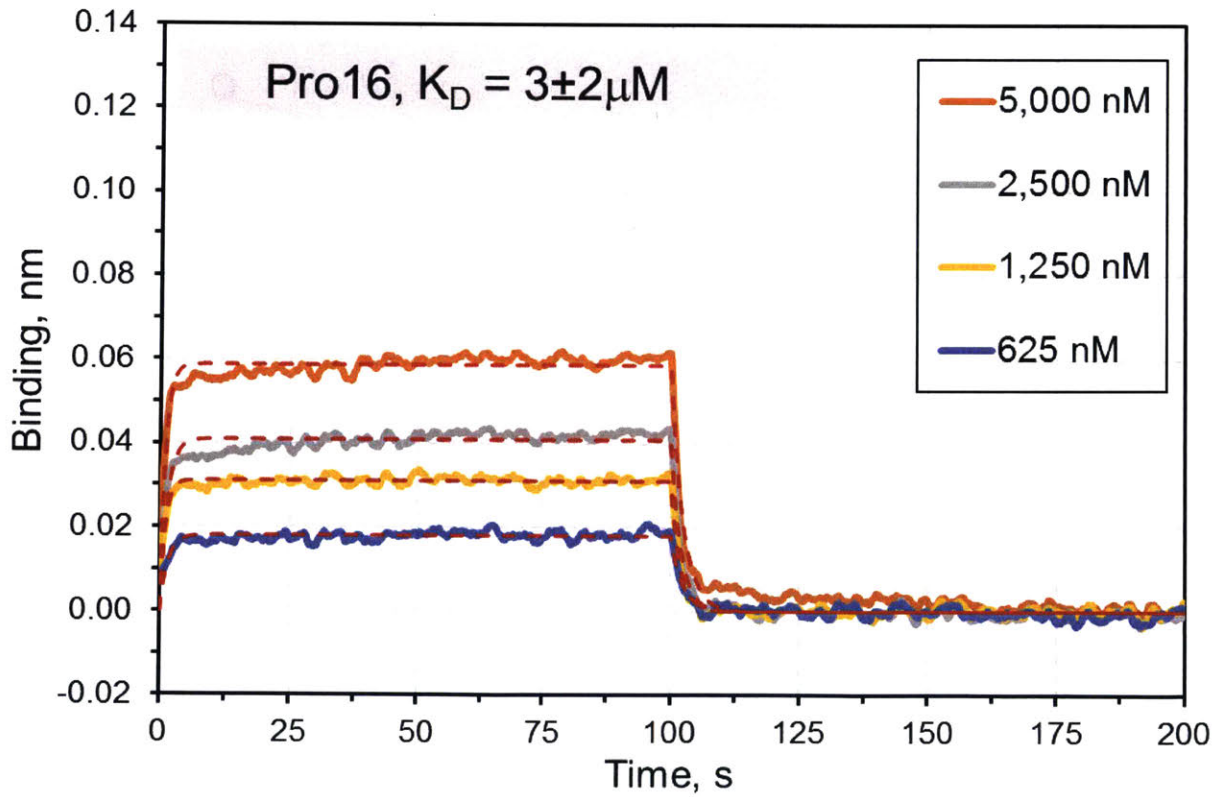

Tyr10
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200
Time, s

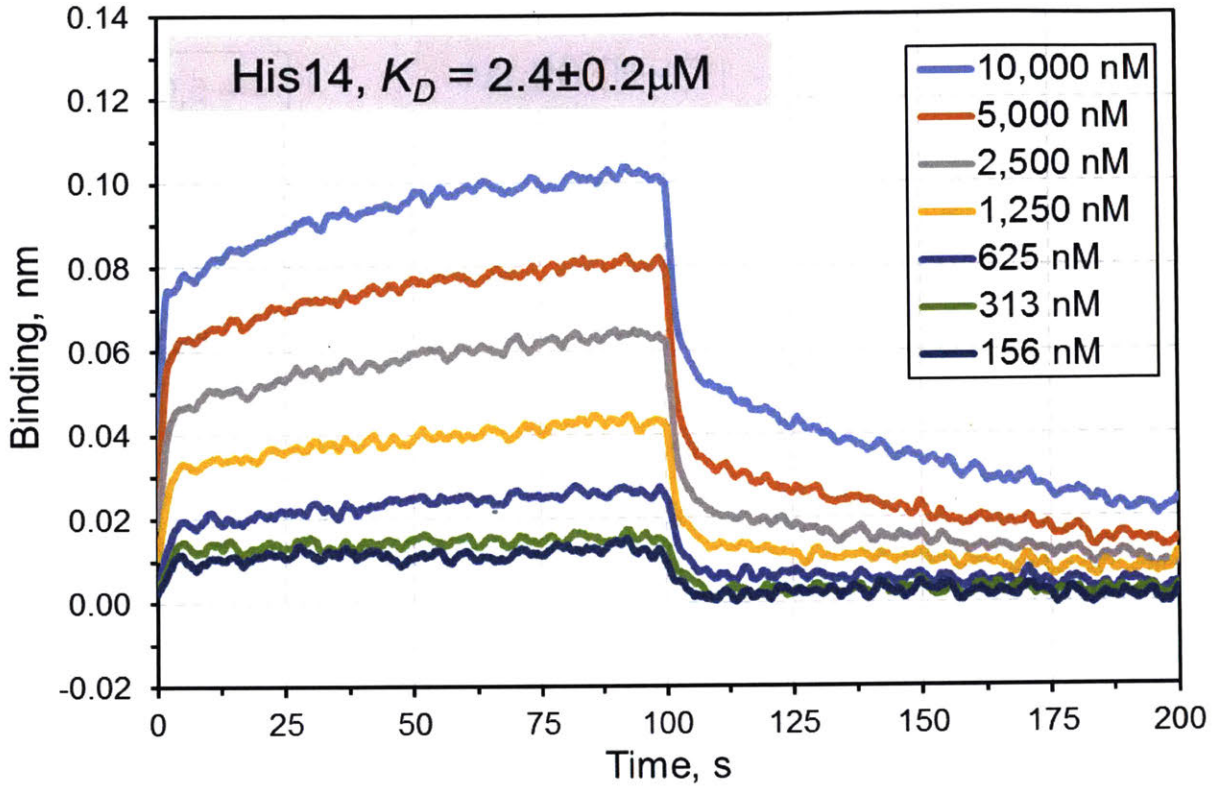

Glu11, K_D = 0.7±0.3μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200
Time, s

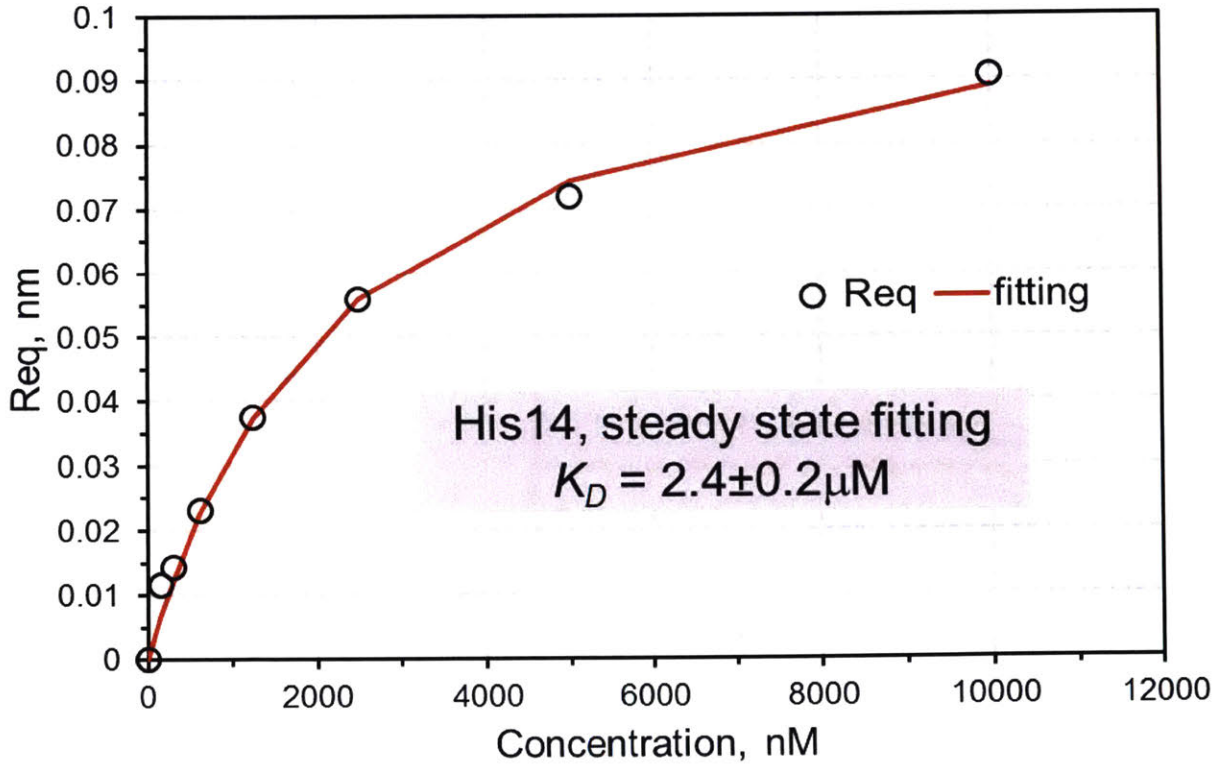

Ala12, K_D = 1.8±0.5μM
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

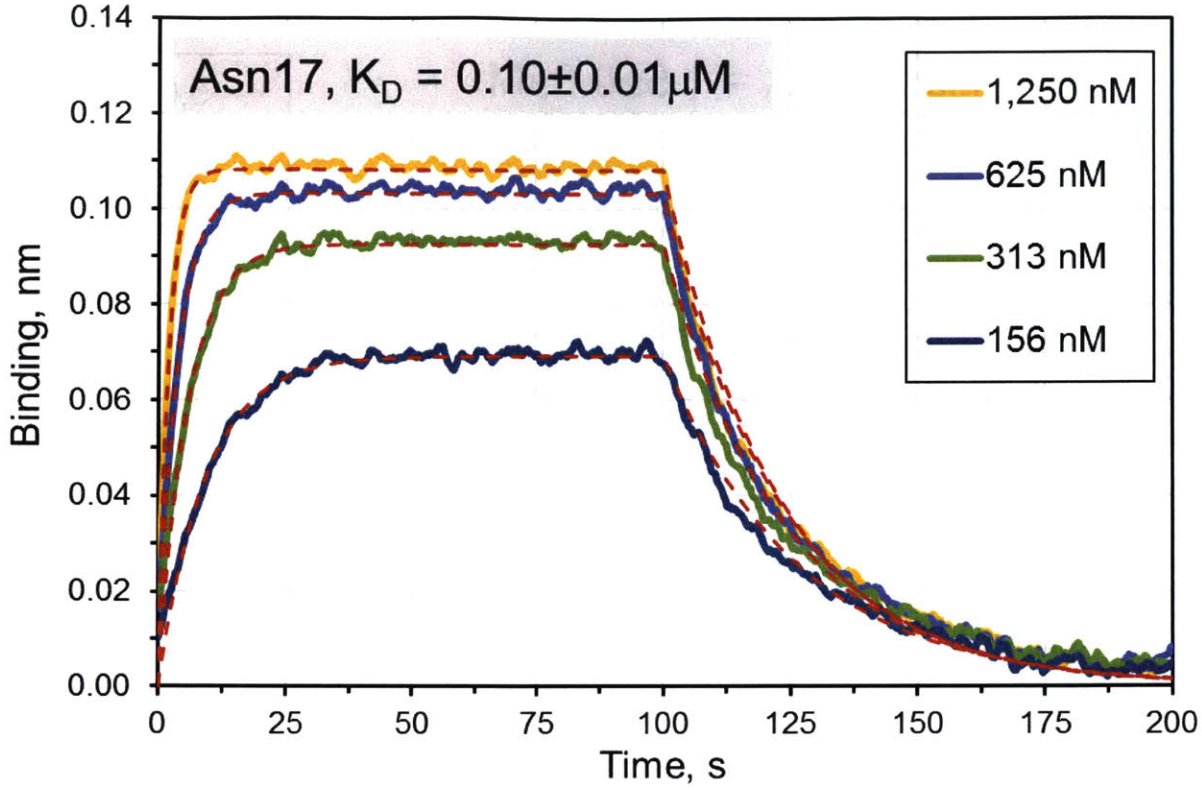

Leu13
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

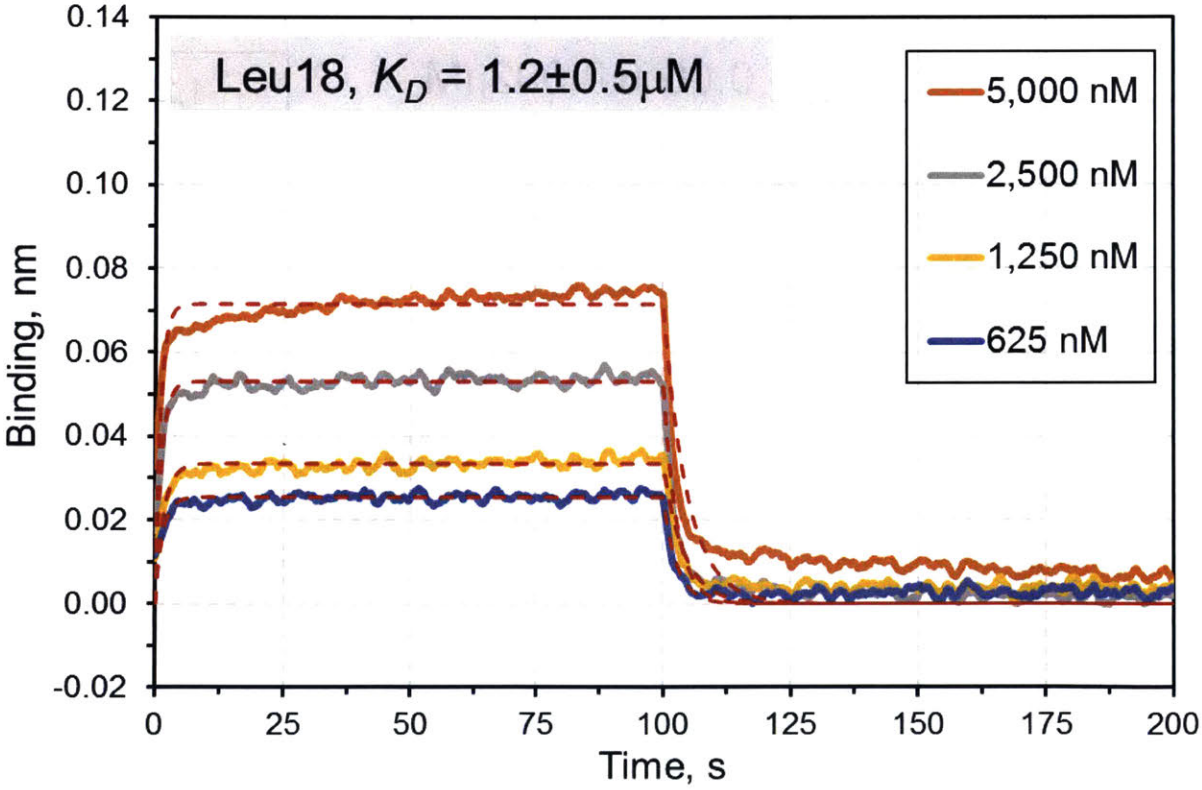

Asp15, K_D = 1.79±0.09μM
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200
Time, s

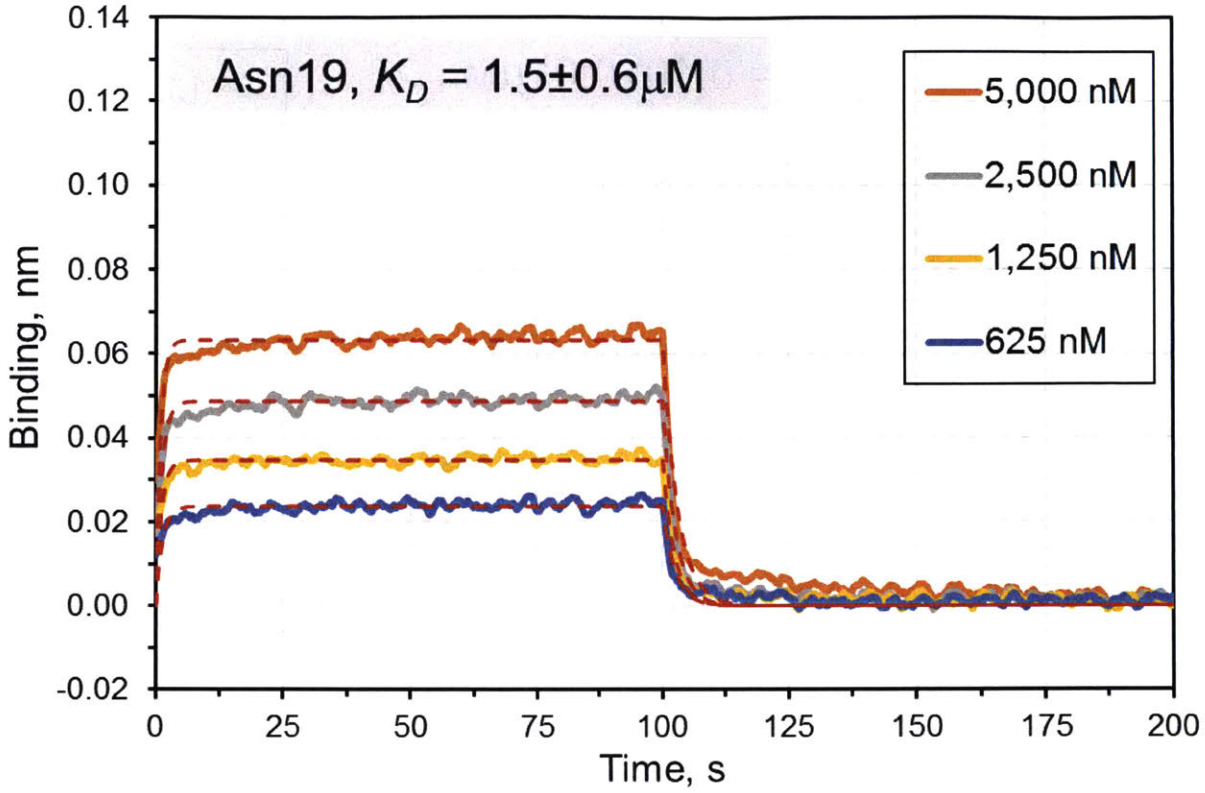

Pro16, K_D = 3±2μM
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200
Time, s

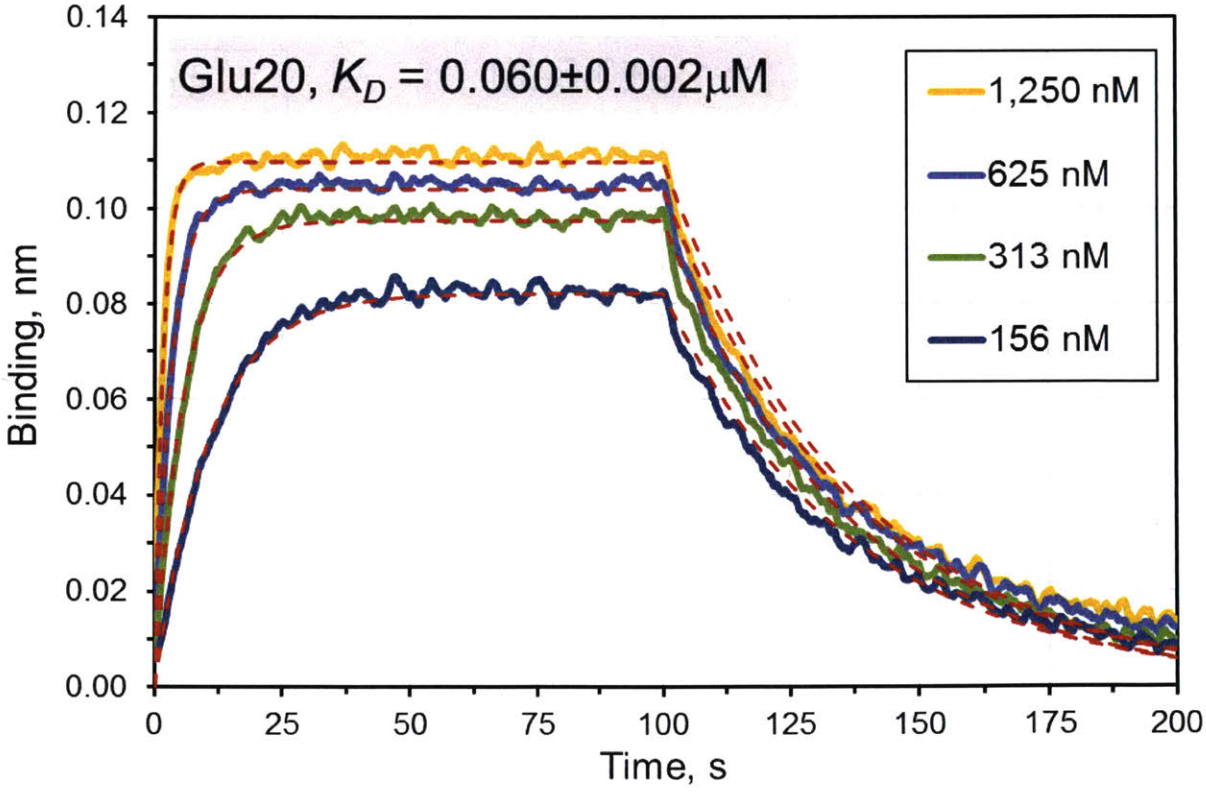

His14, K_D = 2.4±0.2μM
10,000 nM
5,000 nM
2,500 nM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

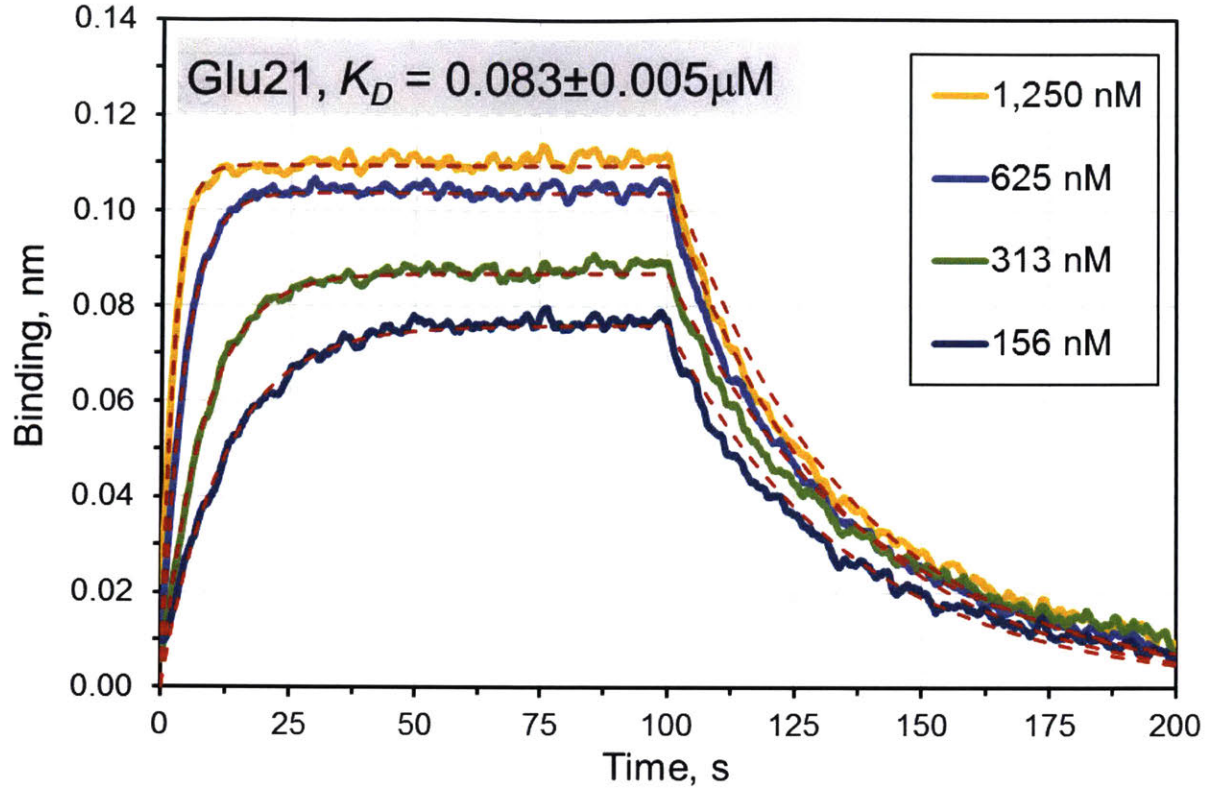

Req
fitting
His14, steady state fitting
K_D = 2.4±0.2μM
Req, nm
Concentration, nM
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
2000
4000
6000
8000
10000
12000

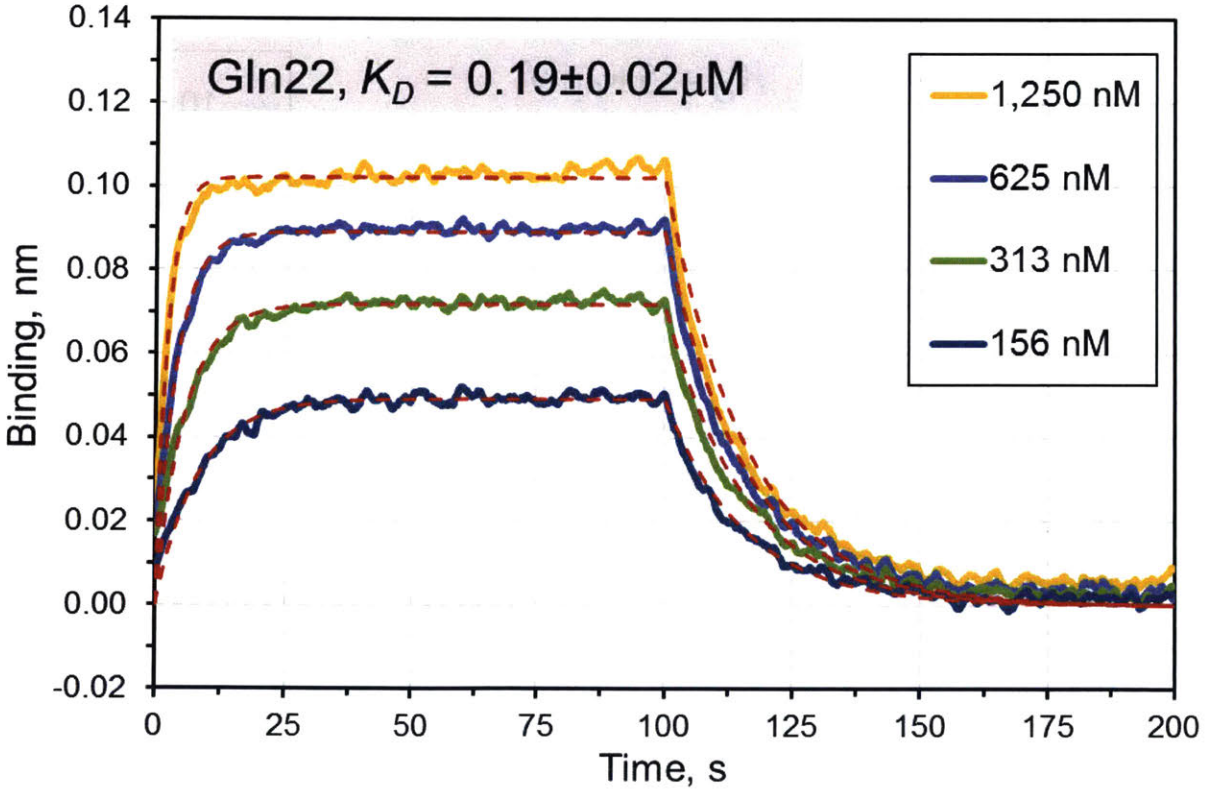

Asn17, K_D = 0.10±0.01μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
25
50
75
100
125
150
175
200

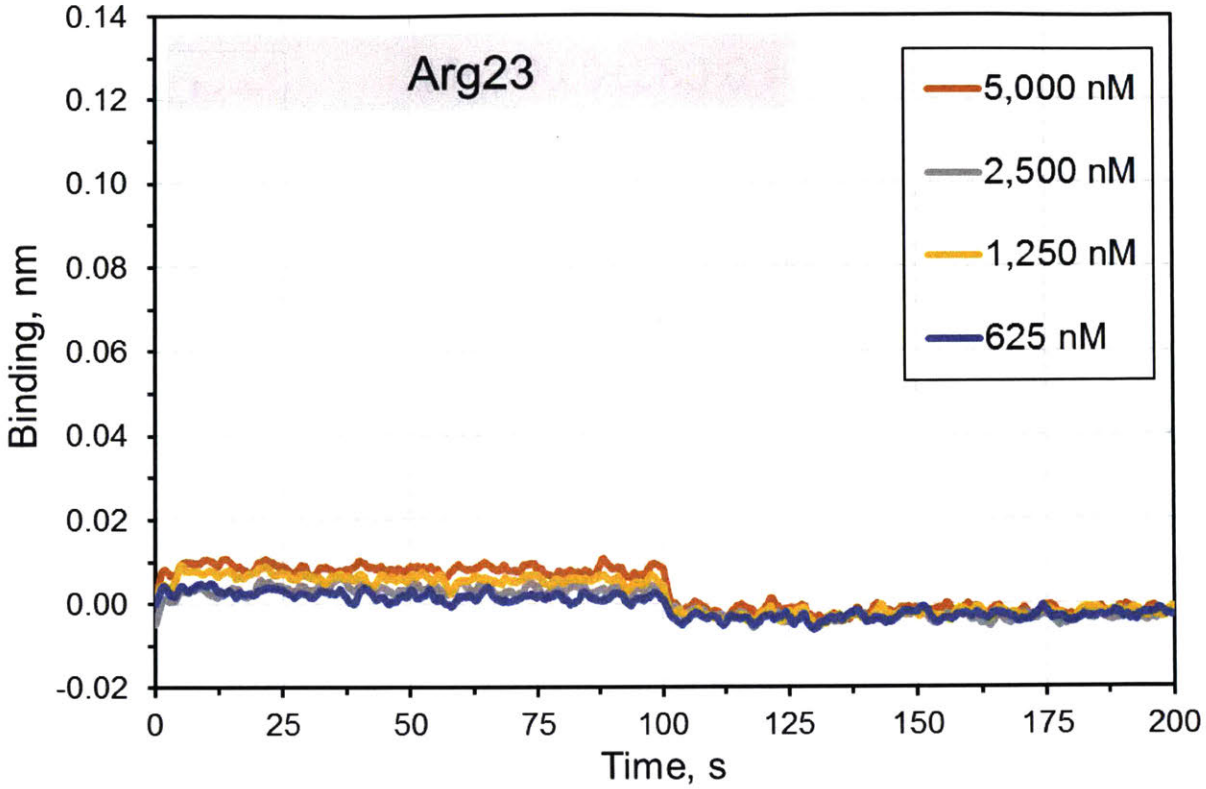

Leu18, K_D = 1.2±0.5μM
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

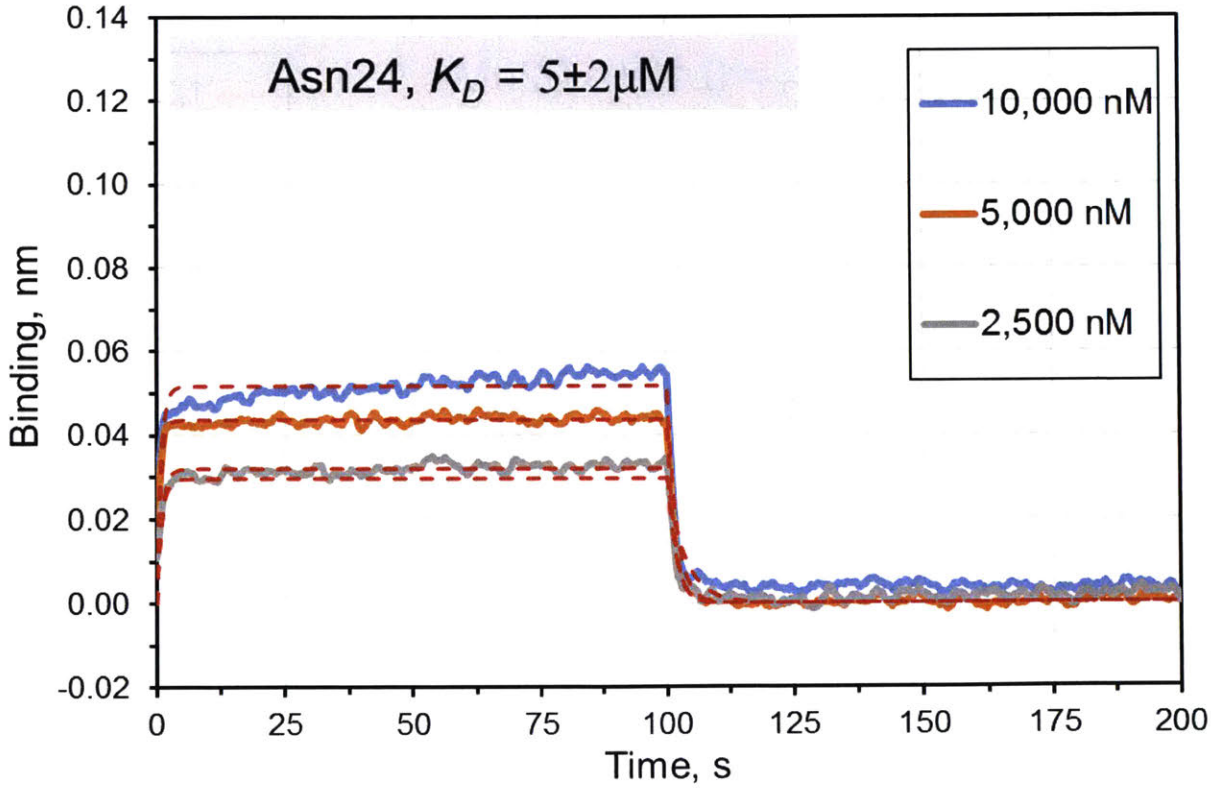

Asn19, K_D = 1.5±0.6μM
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

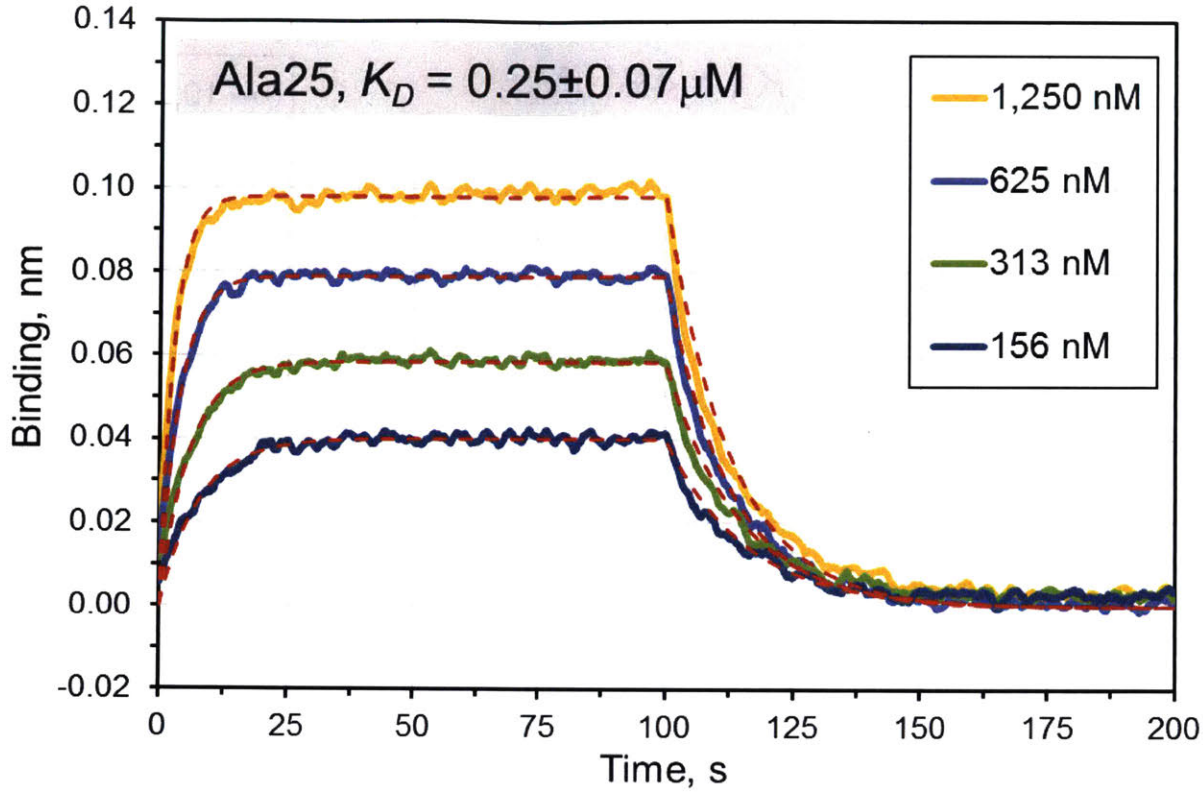

Glu20, K_D = 0.060±0.002μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
25
50
75
100
125
150
175
200

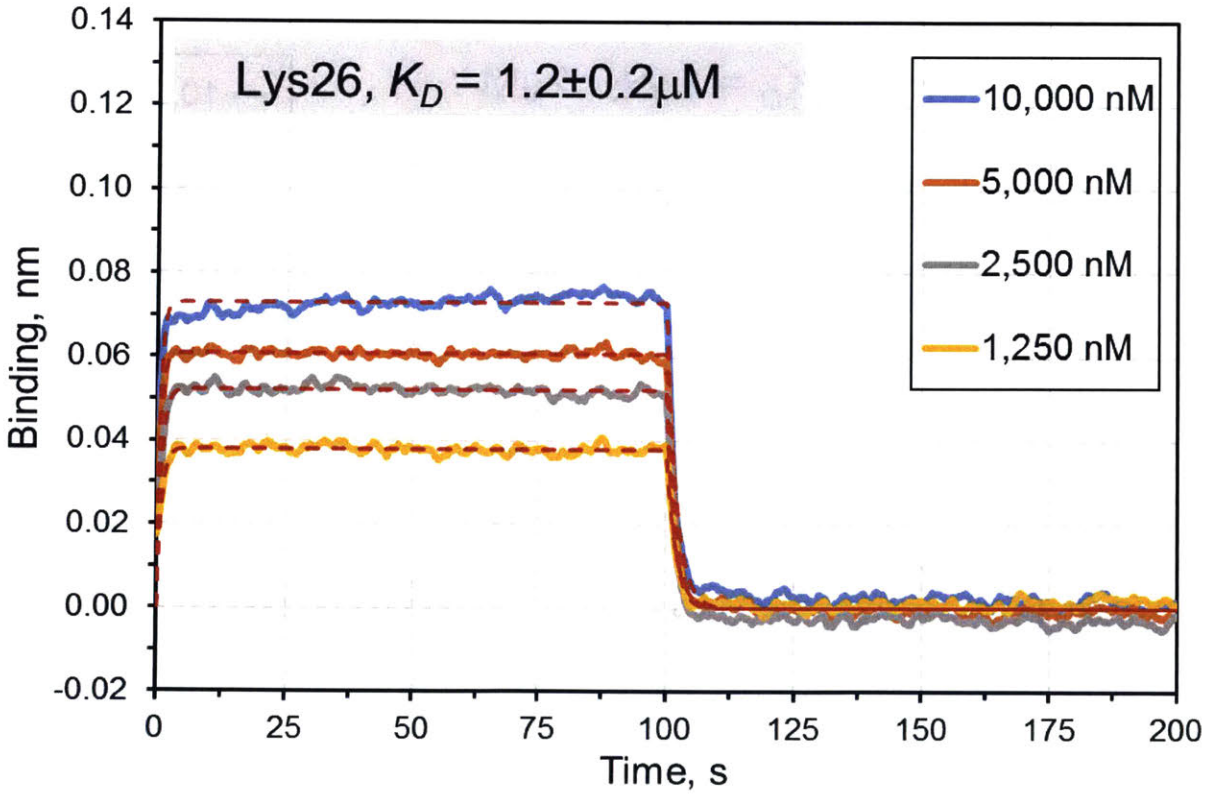

Glu21, K_D = 0.083±0.005μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
25
50
75
100
125
150
175
200

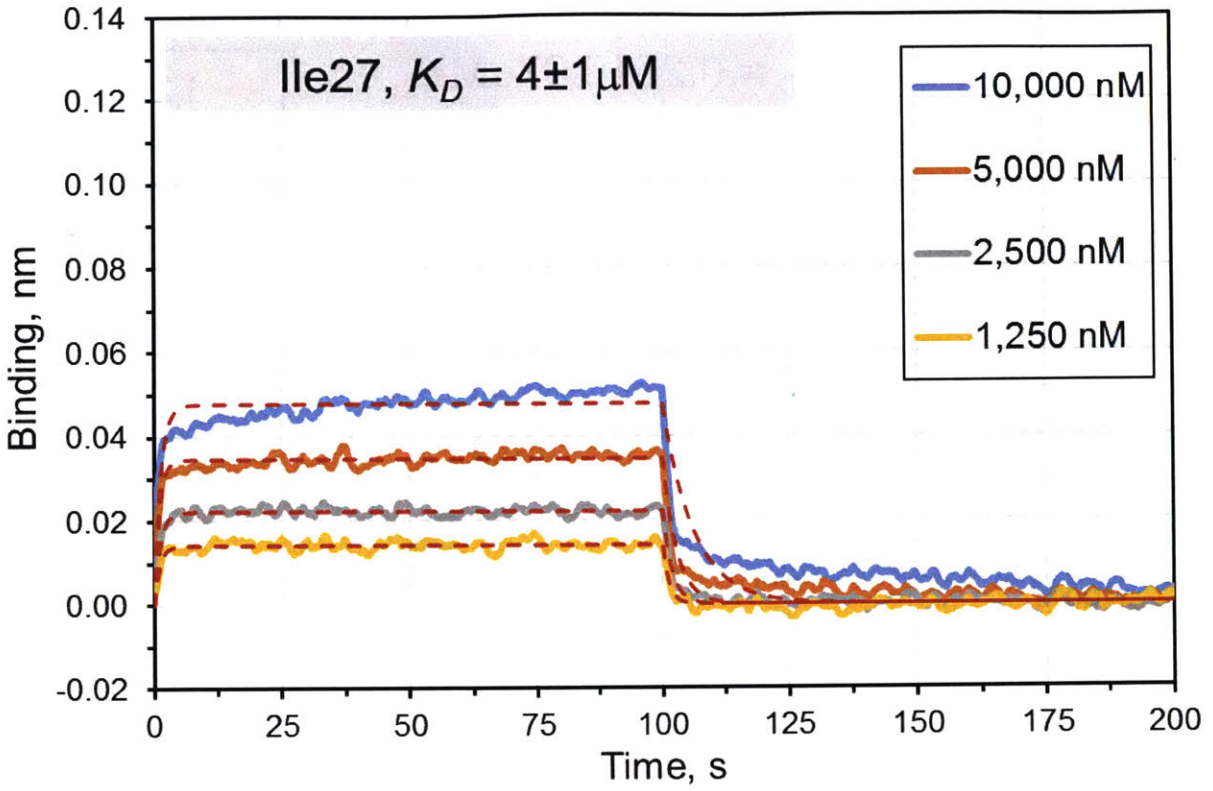

Gln22, K_D = 0.19±0.02μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

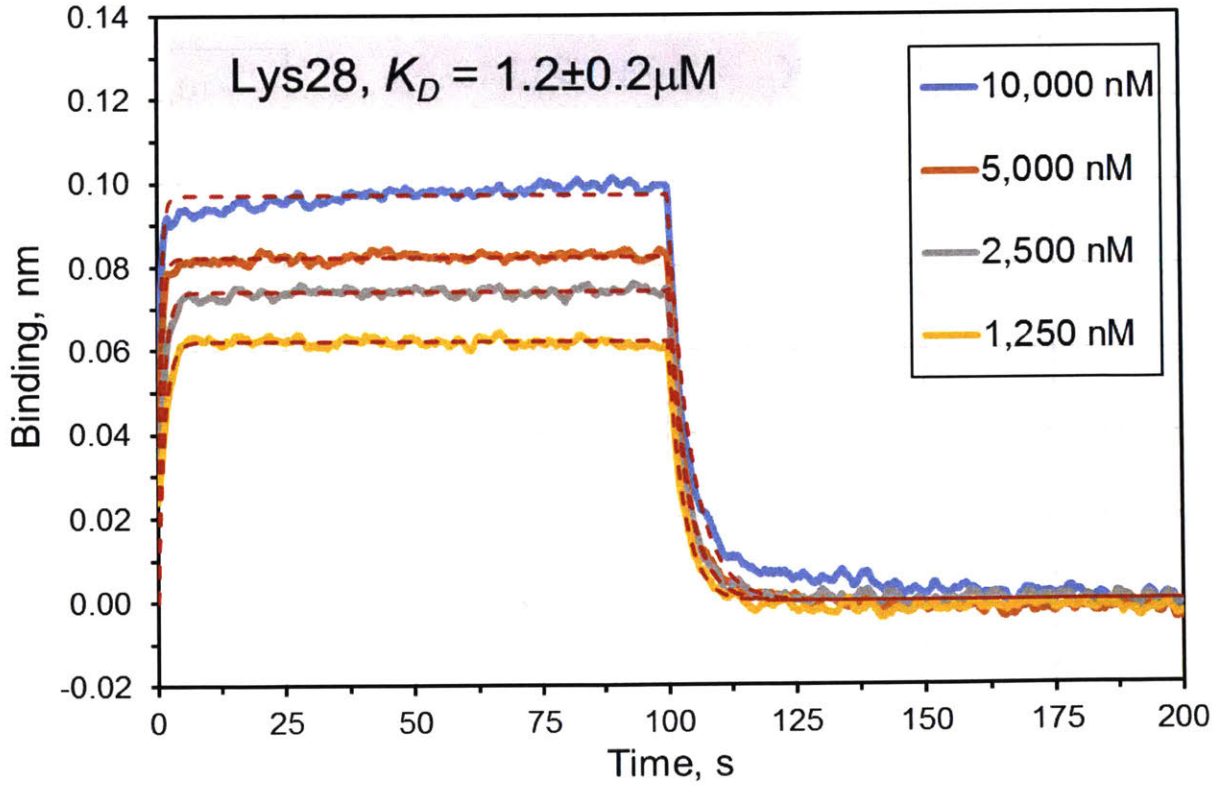

Arg23
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200
Time, s

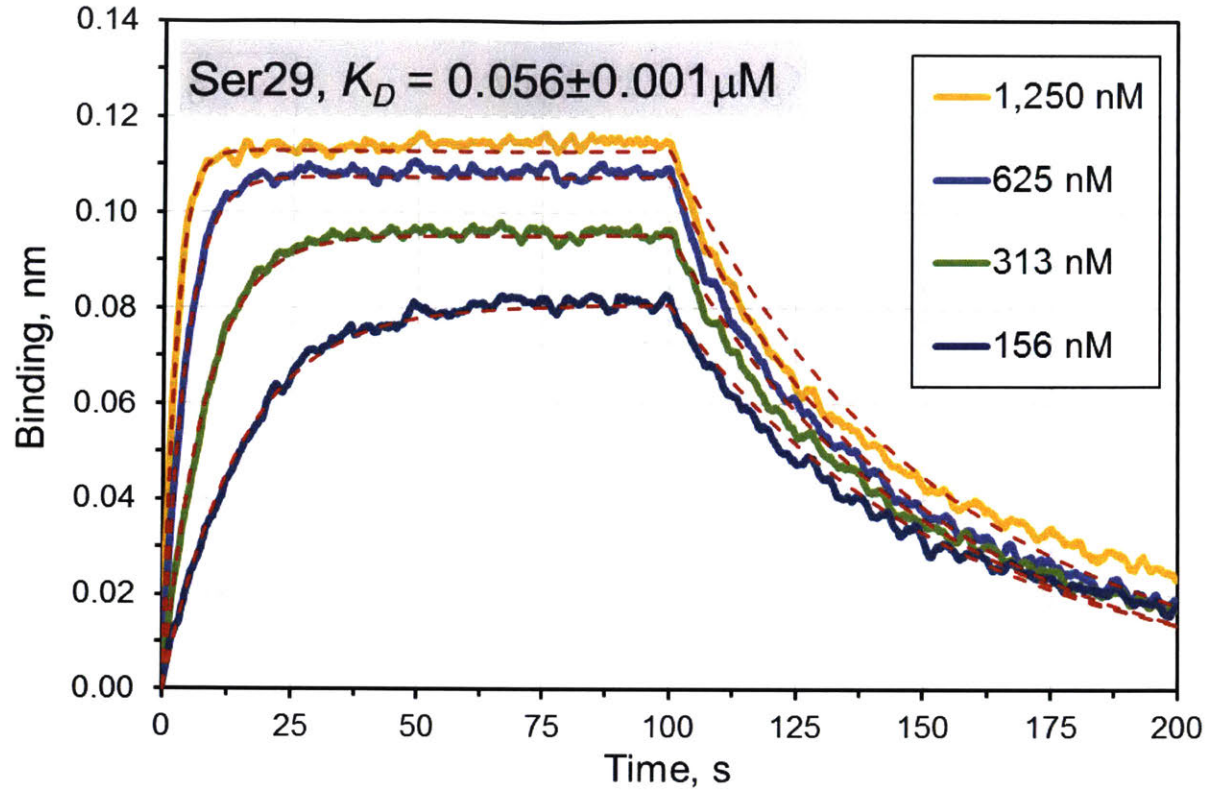

Asn24, K_D = 5±2μM
10,000 nM
5,000 nM
2,500 nM
Binding, nm
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200
Time, s

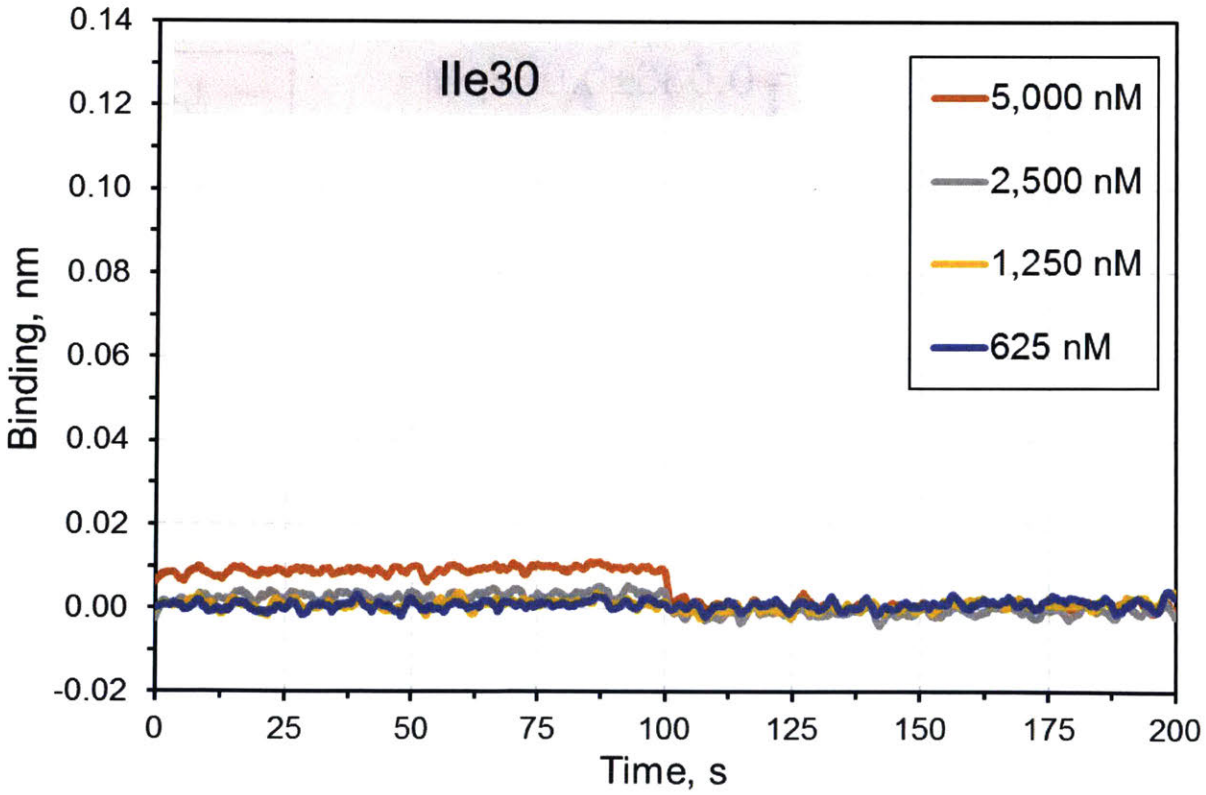

Ala25, K_D = 0.25±0.07μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

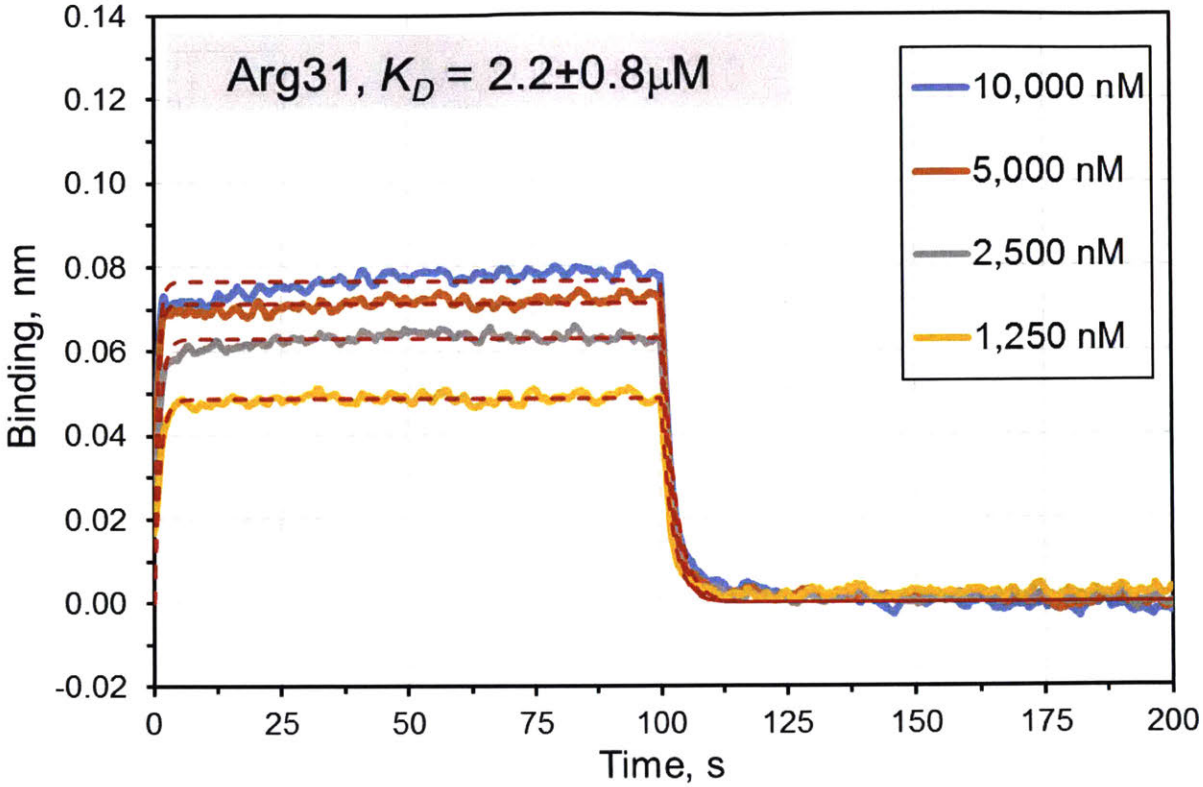

Lys26, K_D = 1.2±0.2μM
10,000 nM
5,000 nM
2,500 nM
1,250 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

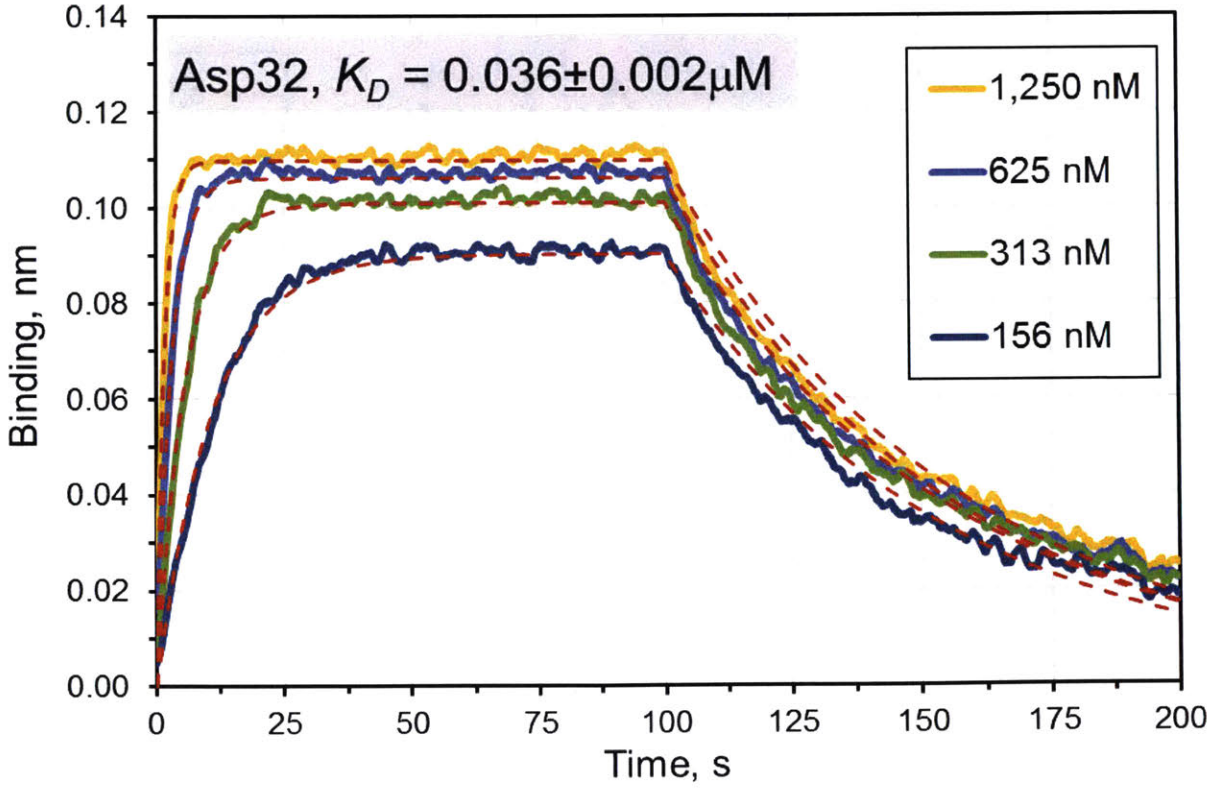

Ile27, K_D = 4±1μM
10,000 nM
5,000 nM
2,500 nM
1,250 nM
Binding, nm
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200
Time, s

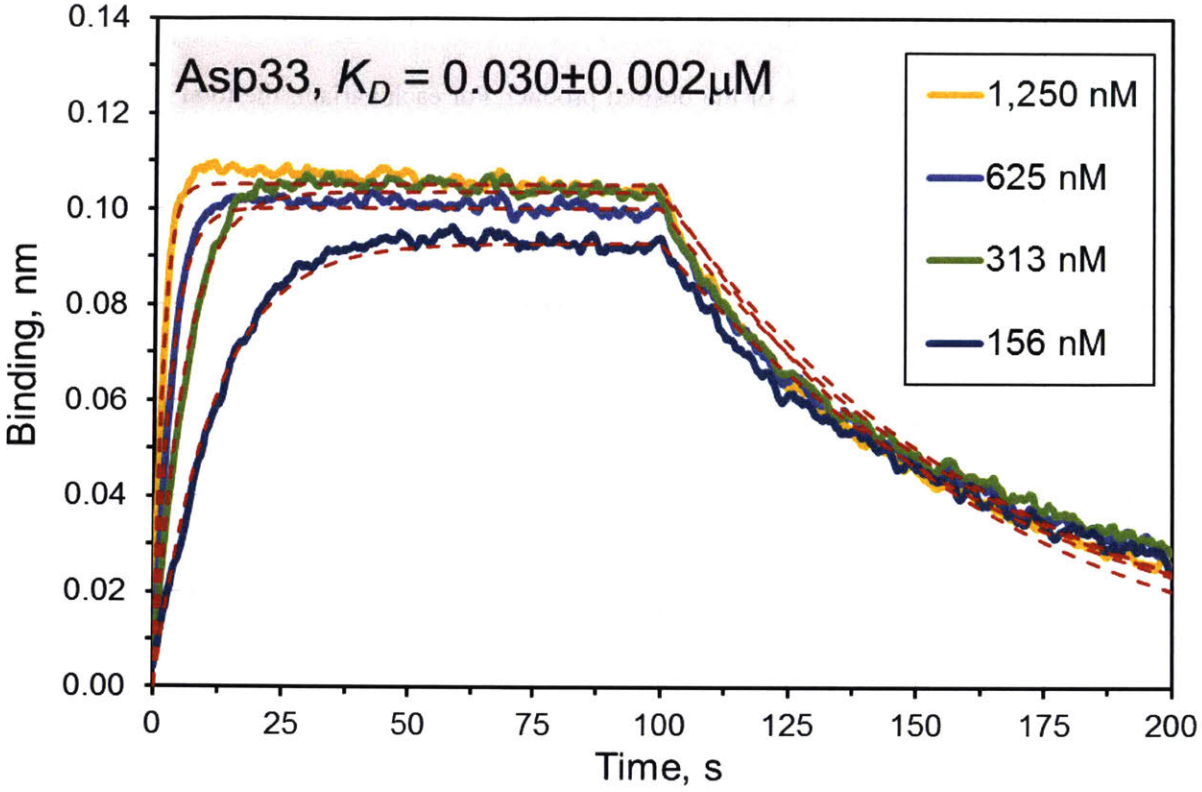

Lys28, K_D = 1.2±0.2μM
10,000 nM
5,000 nM
2,500 nM
1,250 nM
Binding, nm
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200
Time, s


Ser29, K_D = 0.056±0.001μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
25
50
75
100
125
150
175
200


Ile30
5,000 nM
2,500 nM
1,250 nM
625 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

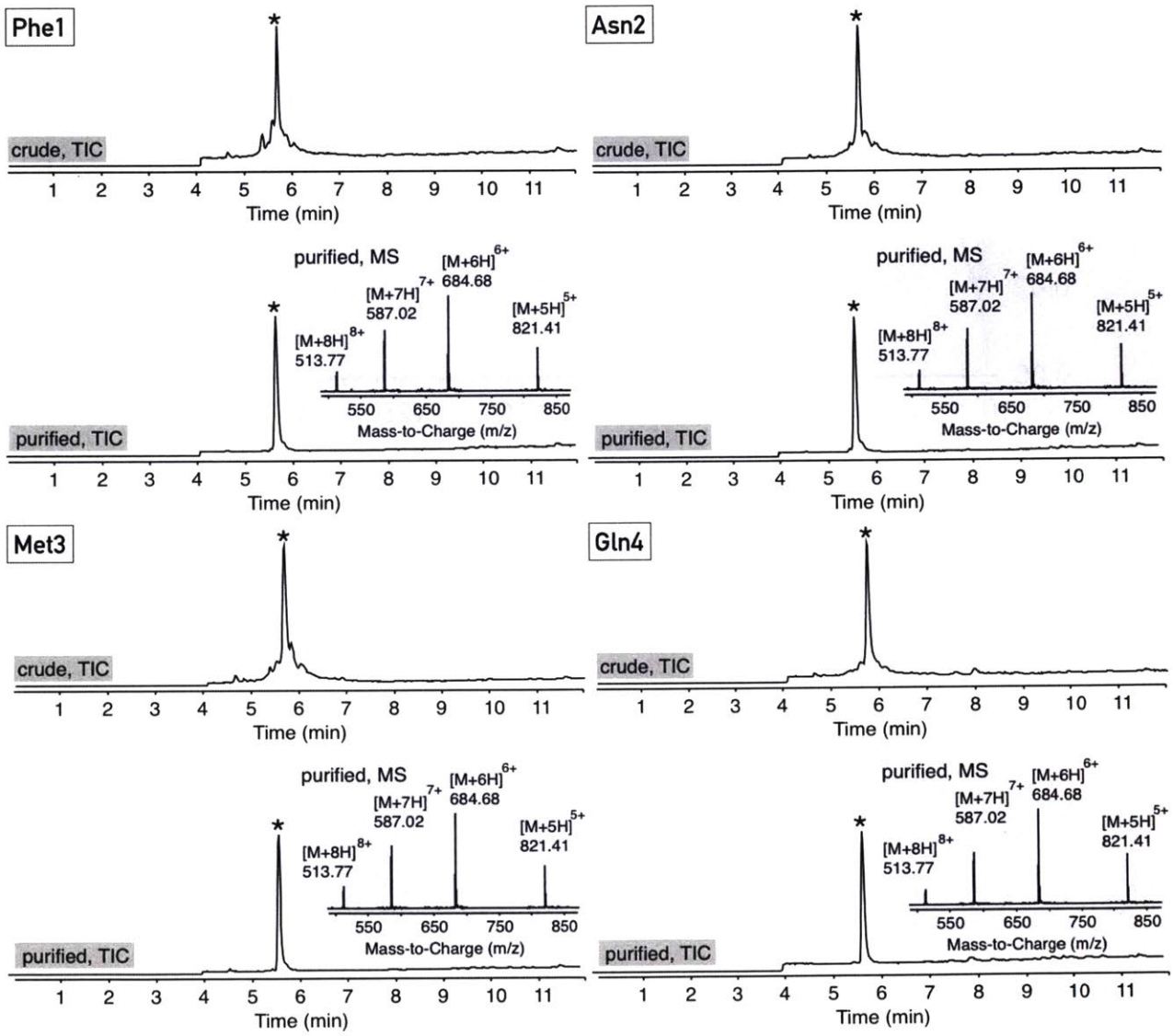

Arg31, K_D = 2.2±0.8μM
10,000 nM
5,000 nM
2,500 nM
1,250 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0
25
50
75
100
125
150
175
200

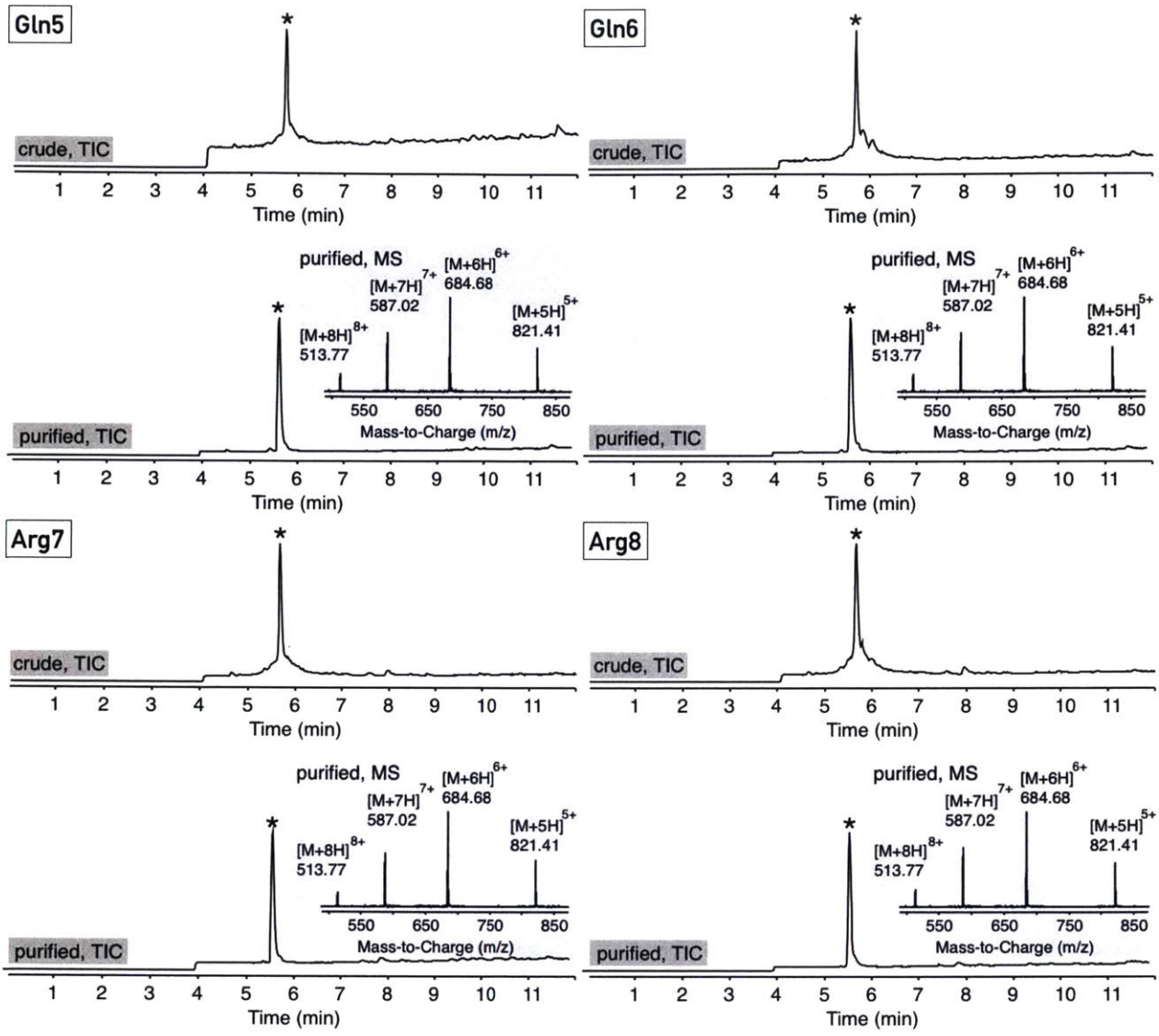

Asp32, K_D = 0.036±0.002μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
25
50
75
100
125
150
175
200

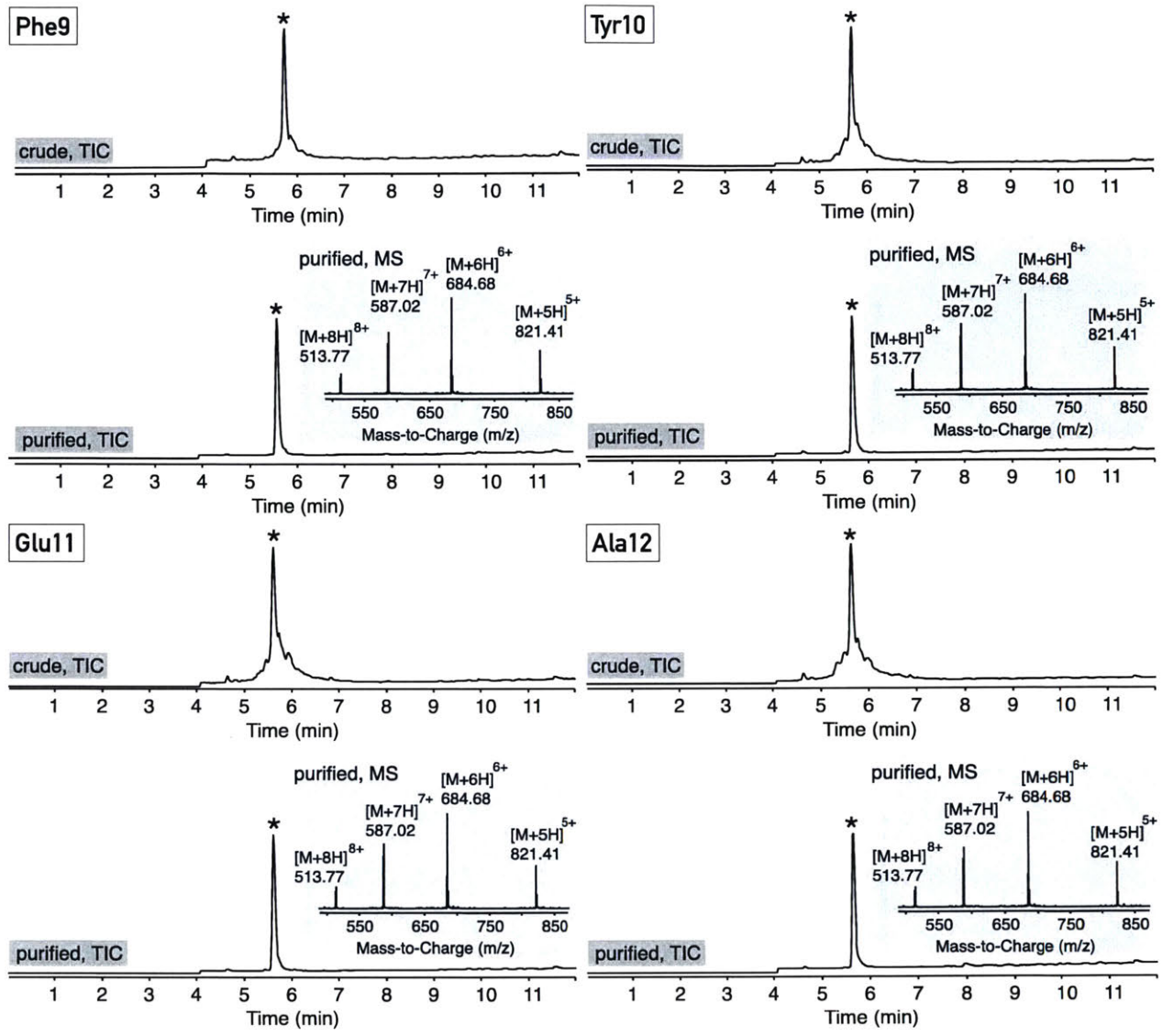

Asp33, K_D = 0.030±0.002μM
1,250 nM
625 nM
313 nM
156 nM
Binding, nm
Time, s
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
25
50
75
100
125
150
175
200

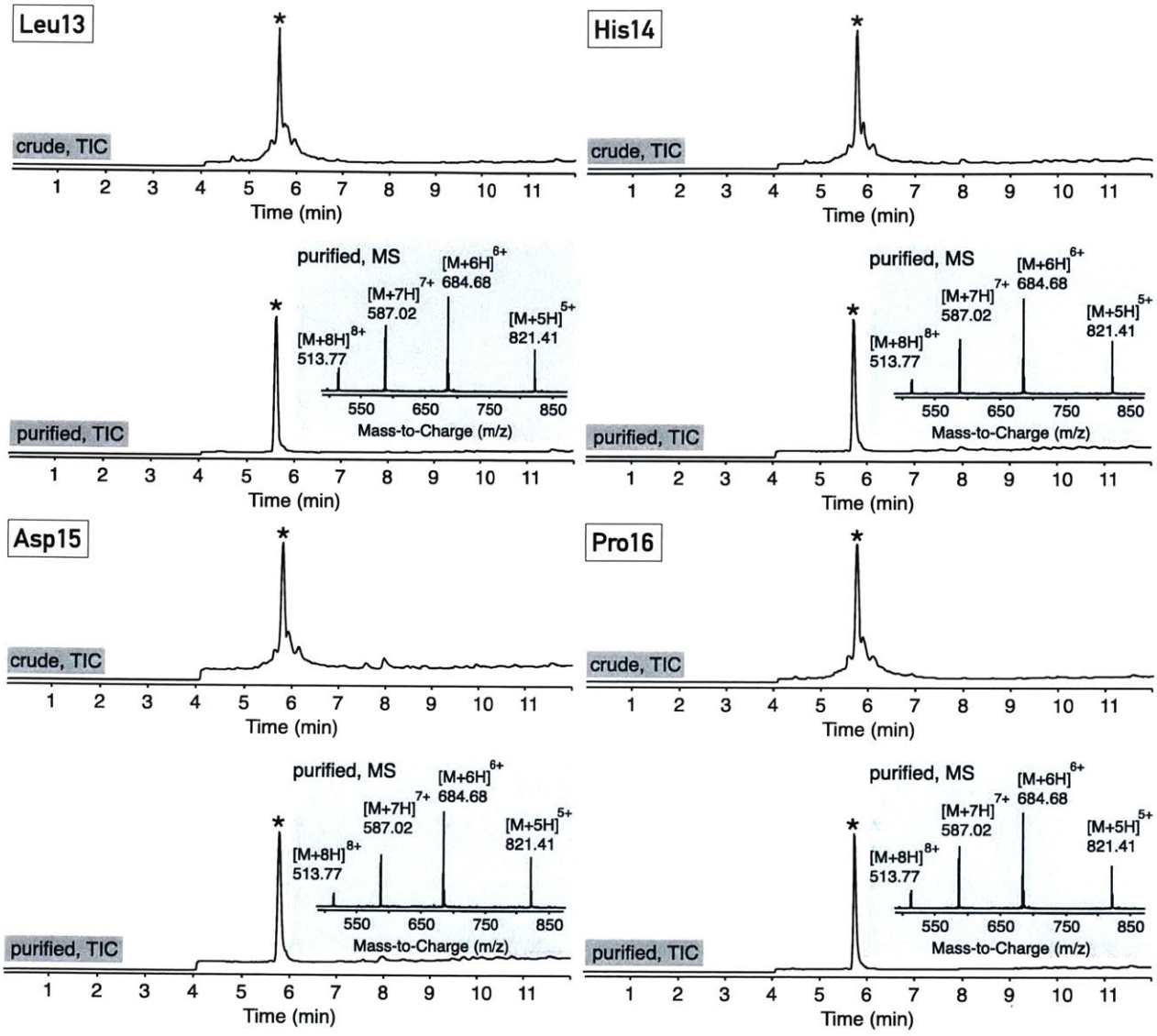
A4.17 LC-MS chromatograms of Crude and Purified Z33 variants

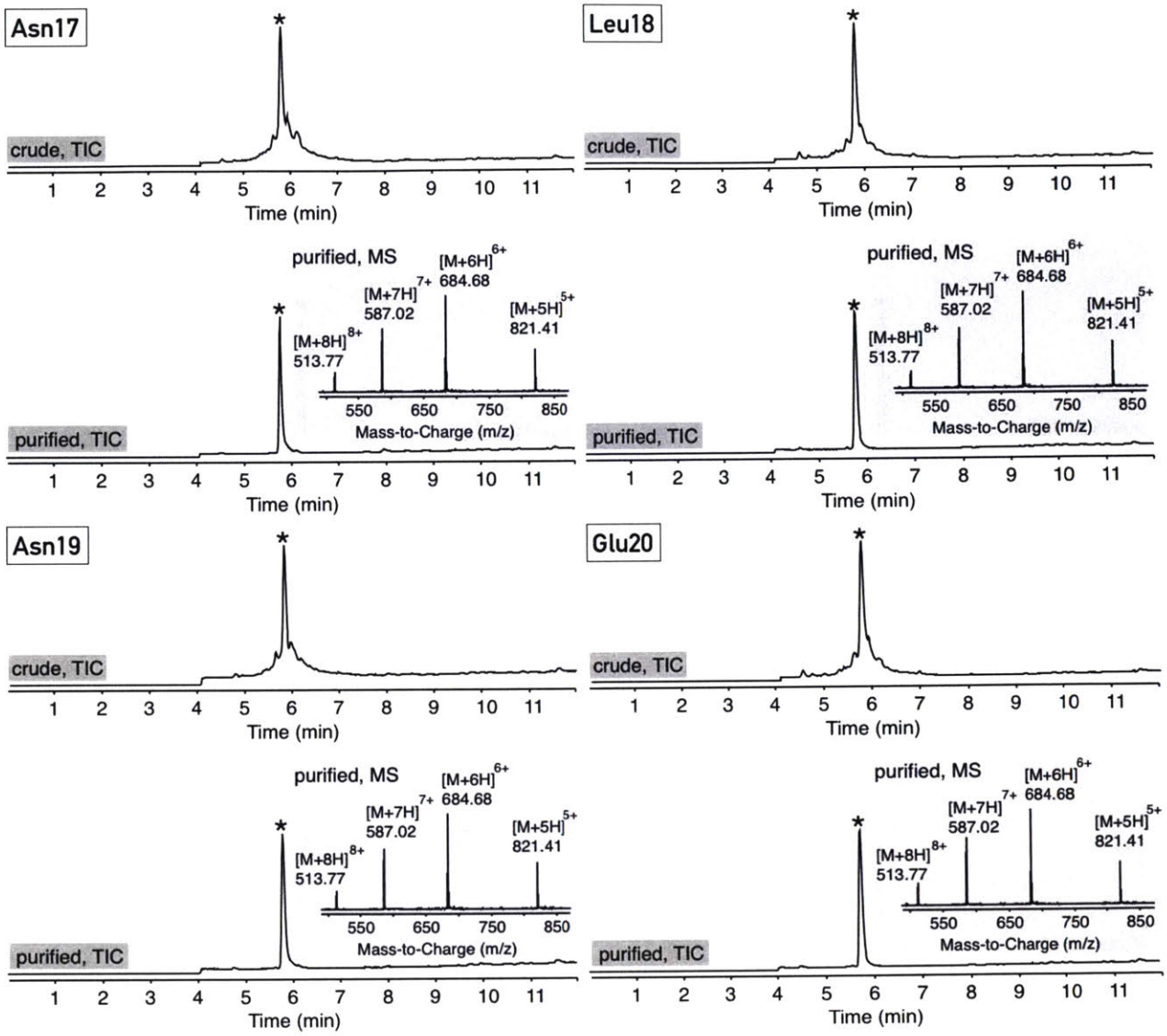
Note that "*" indicate the peak of the desired product. For each variant, the total ion current (TIC) chromatogram of the crude material is shown on the top and the TIC chromatogram of the purified material is shown at the bottom. The inset shows the integrated mass spectra of the whole TIC peak for the purified material.

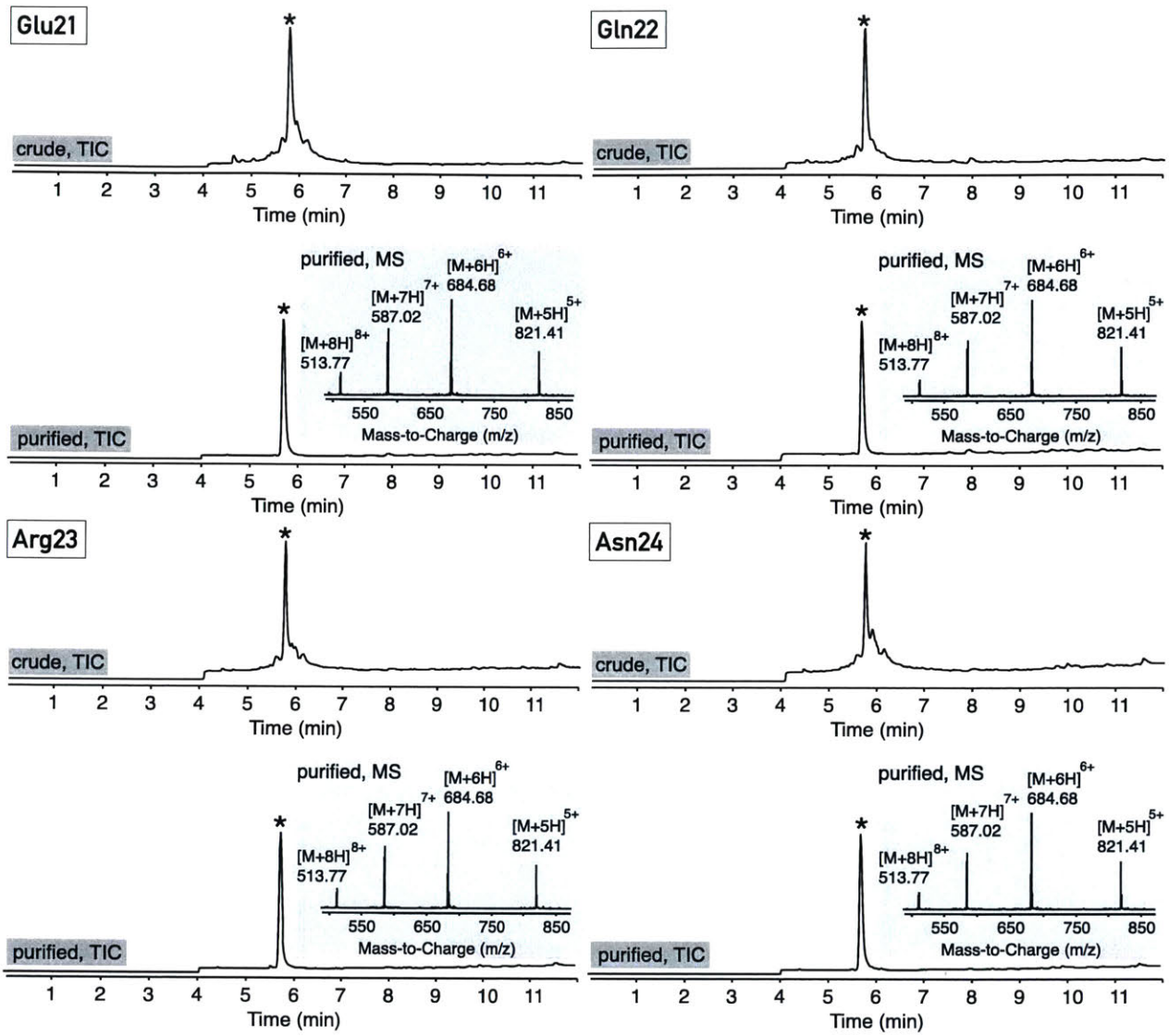


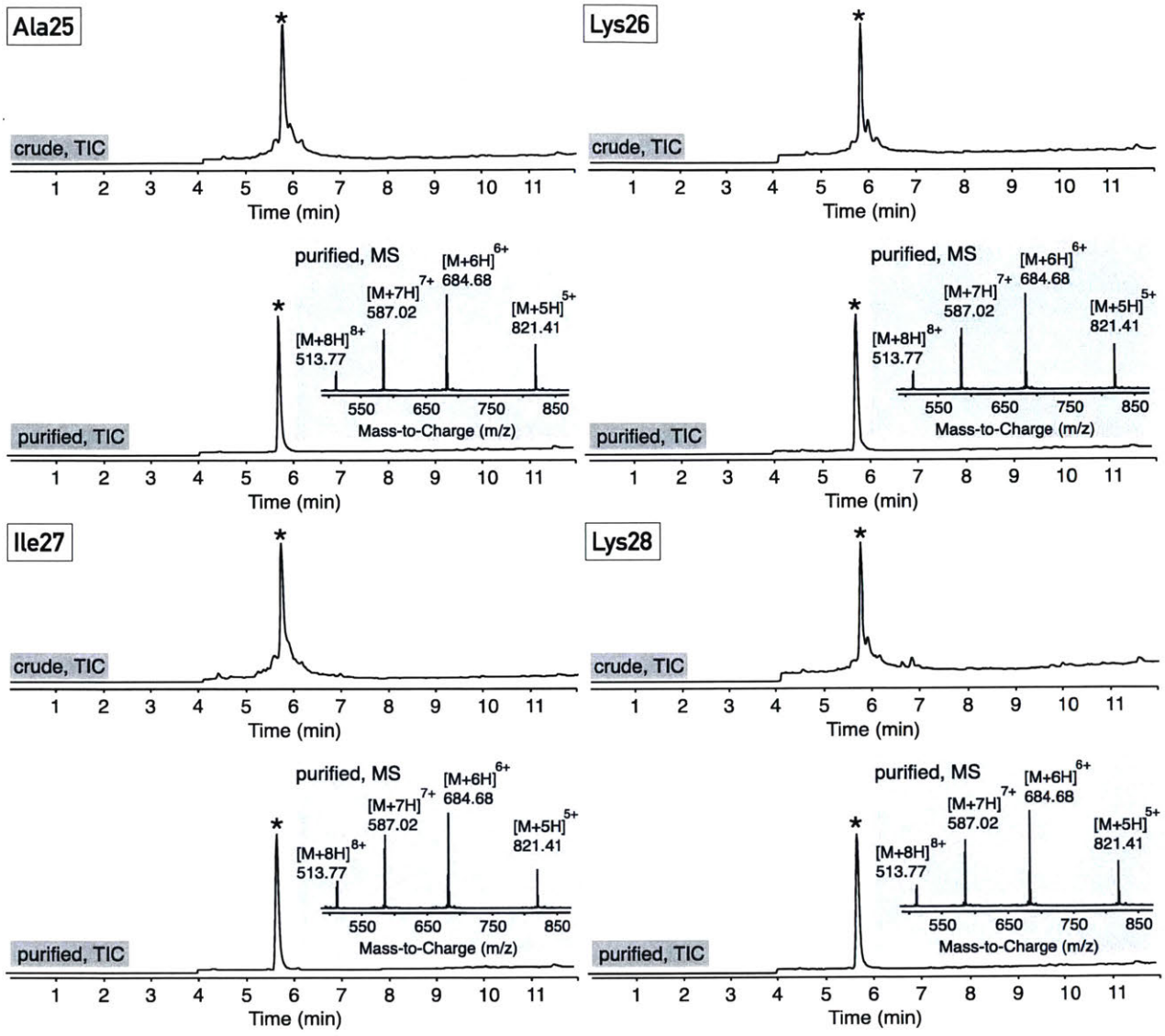

Gln5
crude, TIC
purified, MS
[M+6H]6+
684.68
[M+7H]7+
587.02
[M+5H]5+
821.41
[M+8H]8+
513.77
Mass-to-Charge (m/z)
purified, TIC
Time (min)
Gln6
Arg7
Arg8

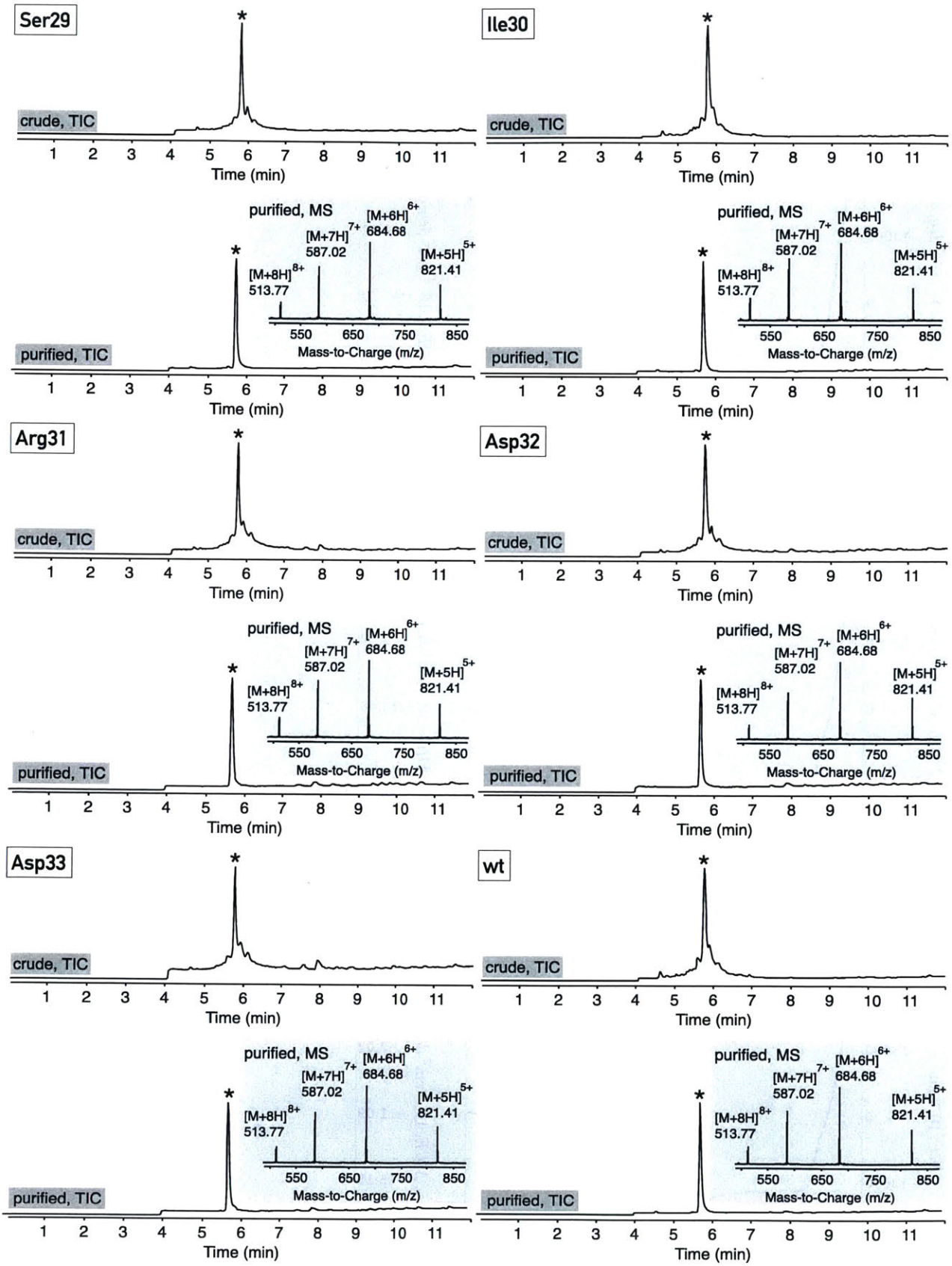

Phe9
crude, TIC
purified, MS
[M+6H]6+
684.68
[M+7H]7+
587.02
[M+5H]5+
821.41
[M+8H]8+
513.77
Mass-to-Charge (m/z)
purified, TIC
Time (min)
Tyr10
Glu11
Ala12


Leu13
crude, TIC
purified, MS
[M+6H]6+ 684.68
[M+7H]7+ 587.02
[M+5H]5+ 821.41
[M+8H]8+ 513.77
Mass-to-Charge (m/z)
purified, TIC
Time (min)
His14
Asp15
Pro16

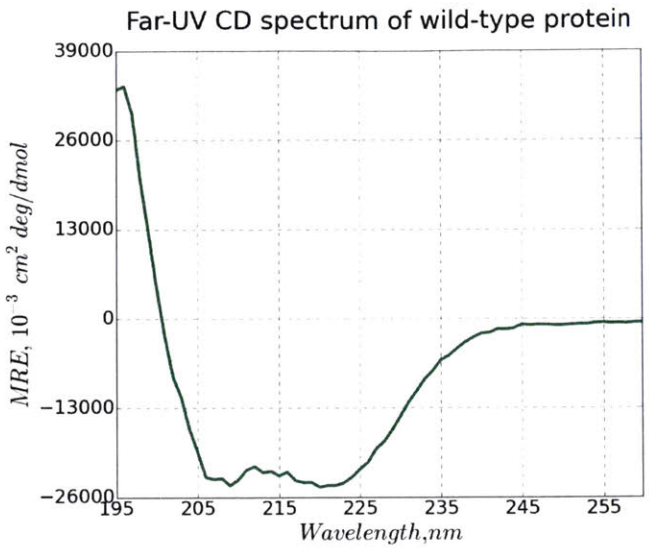

Asn17
crude, TIC
Time (min)
purified, MS
[M+6H]6+
684.68
[M+7H]7+
587.02
[M+5H]5+
821.41
[M+8H]8+
513.77
Mass-to-Charge (m/z)
purified, TIC
Leu18
Asn19
Glu20

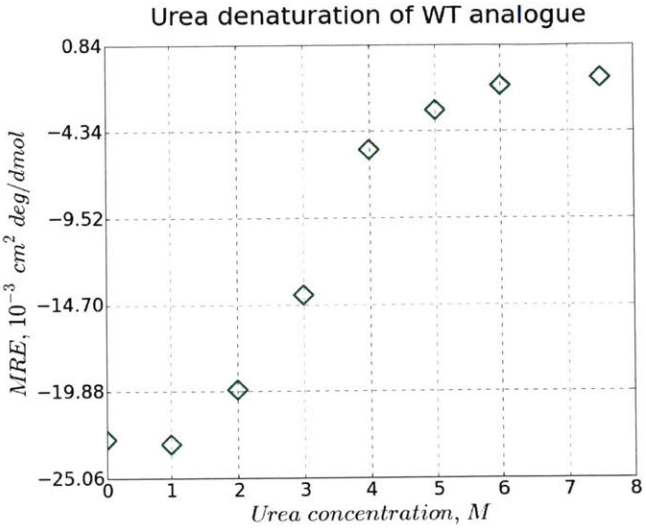

Glu21
crude, TIC
Time (min)
purified, MS
[M+6H]6+ 684.68
[M+7H]7+ 587.02
[M+5H]5+ 821.41
[M+8H]8+ 513.77
Mass-to-Charge (m/z)
purified, TIC
Gln22
Arg23
Asn24

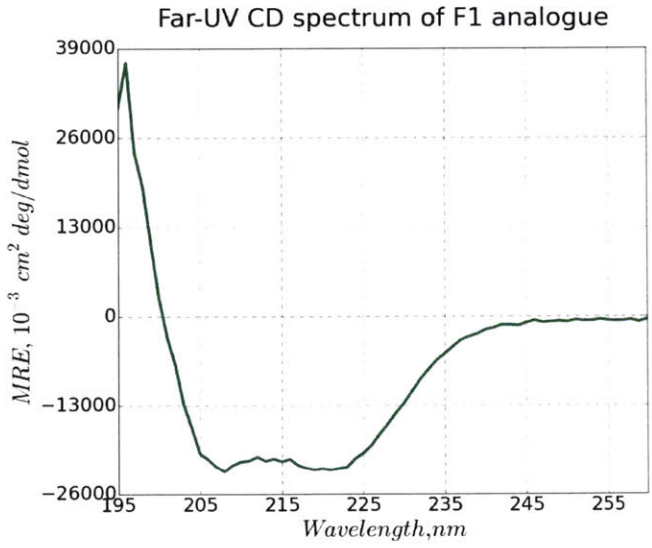

Ala25
crude, TIC
Time (min)
purified, MS
[M+6H]6+ 684.68
[M+7H]7+ 587.02
[M+5H]5+ 821.41
[M+8H]8+ 513.77
Mass-to-Charge (m/z)
purified, TIC
Lys26
Ile27
Lys28

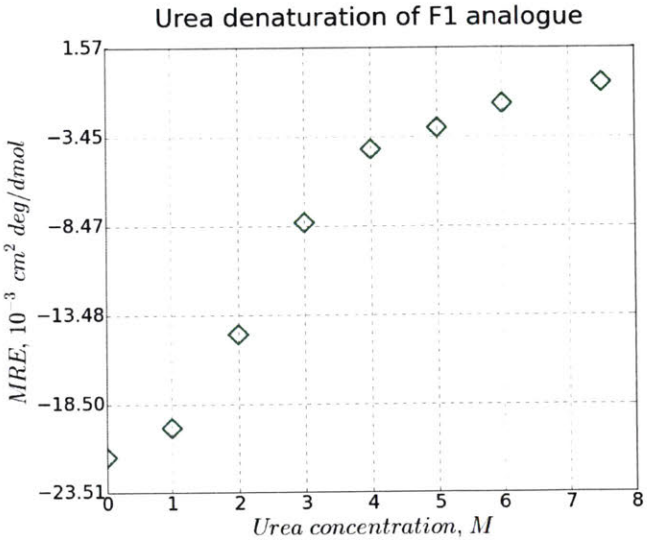

Ser29
crude, TIC
purified, MS
[M+6H]6+ 684.68
[M+7H]7+ 587.02
[M+5H]5+ 821.41
[M+8H]8+ 513.77
Mass-to-Charge (m/z)
purified, TIC
Time (min)
Ile30
Arg31
Asp32
Asp33
wt

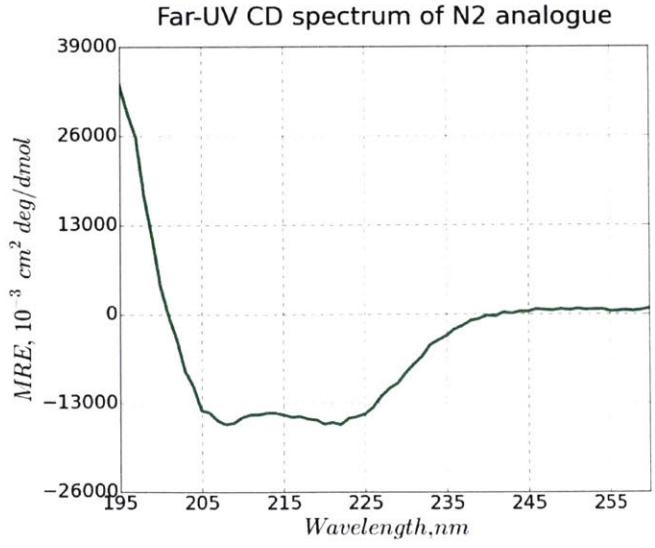
A4.18 Far-UV CD spectra and urea denaturation curves for Z33 analogs

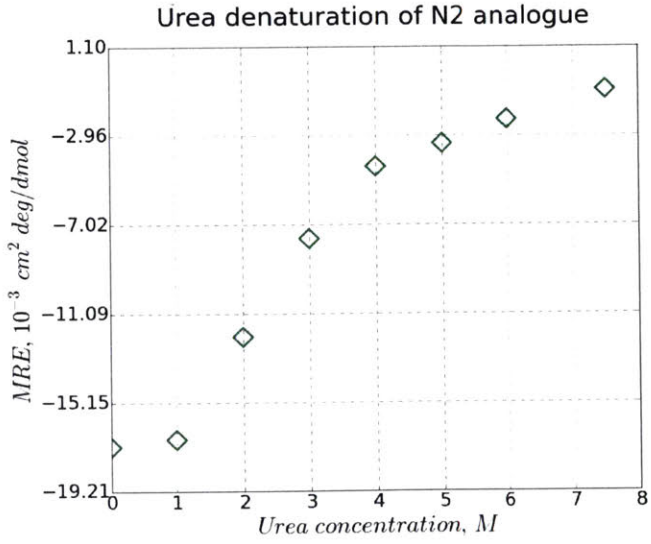


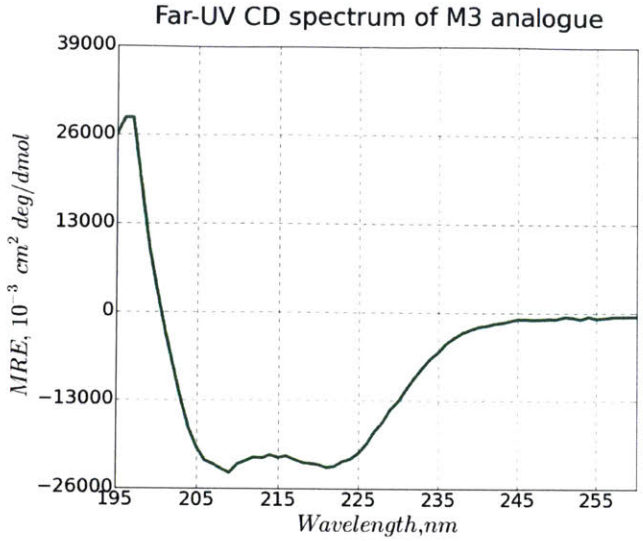


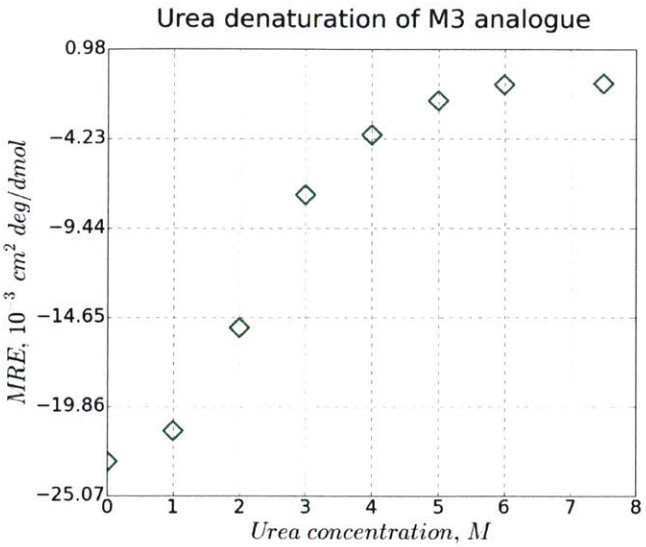


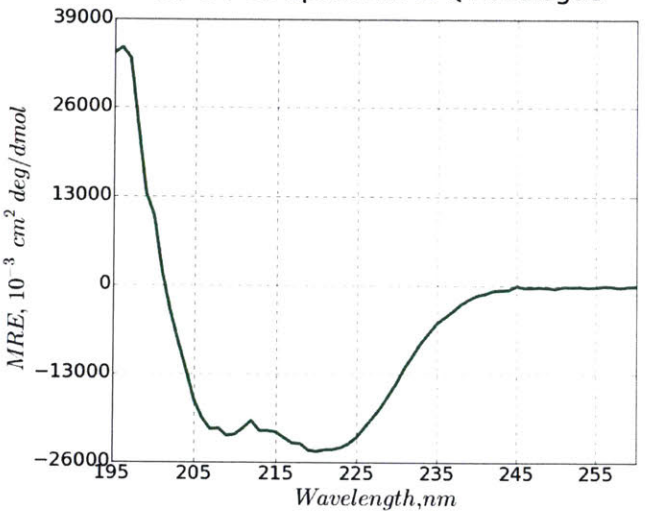


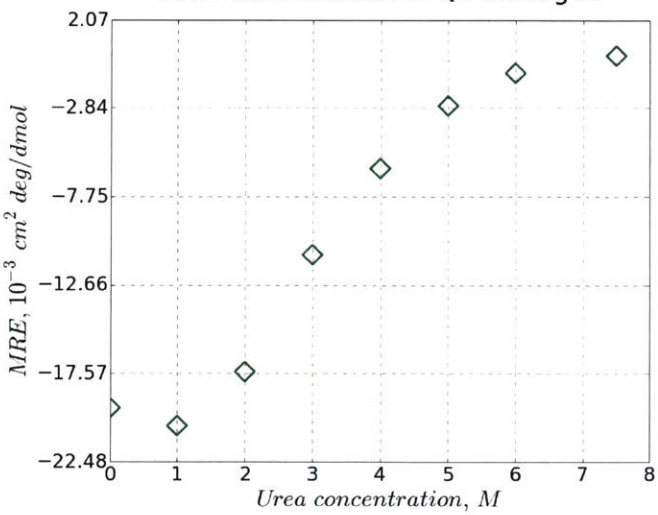


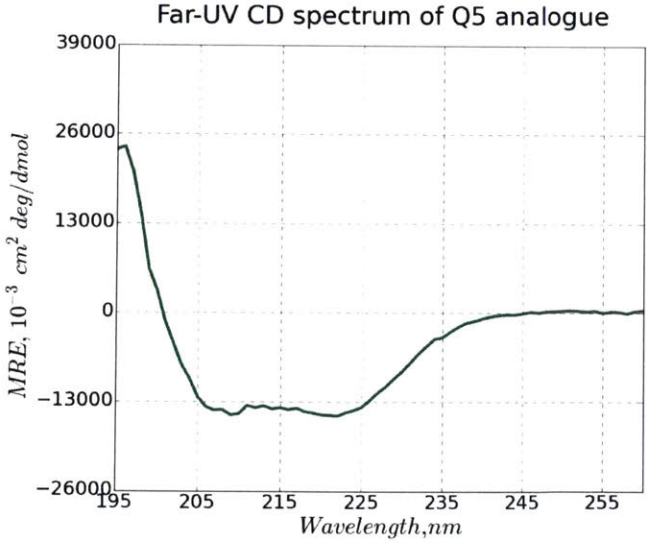


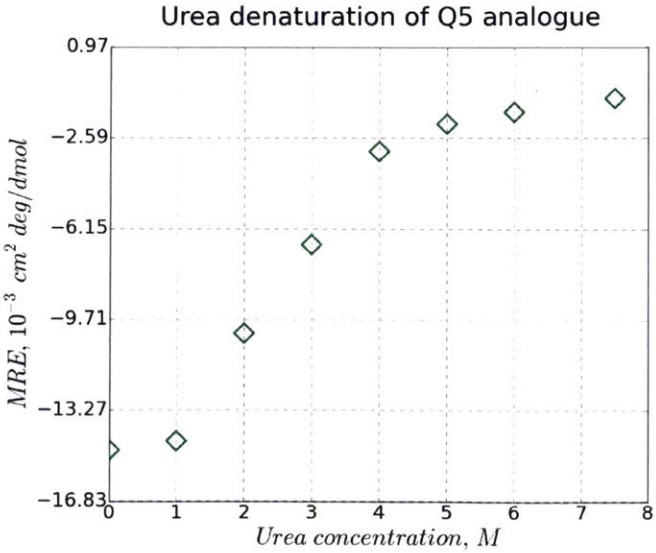

Far-UV CD spectrum of M3 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
MRE, 10 -3 cm2 deg/dmol
Wavelength,nm

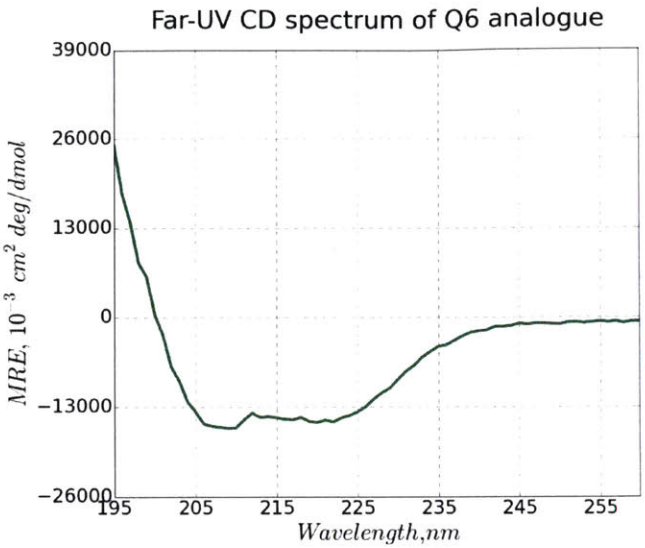

Urea denaturation of M3 analogue
0.98
−4.23
−9.44
−14.65
−19.86
−25.07
0
1
2
3
4
5
6
7
8
MRE, 10 -3 cm2 deg/dmol
Urea concentration, M

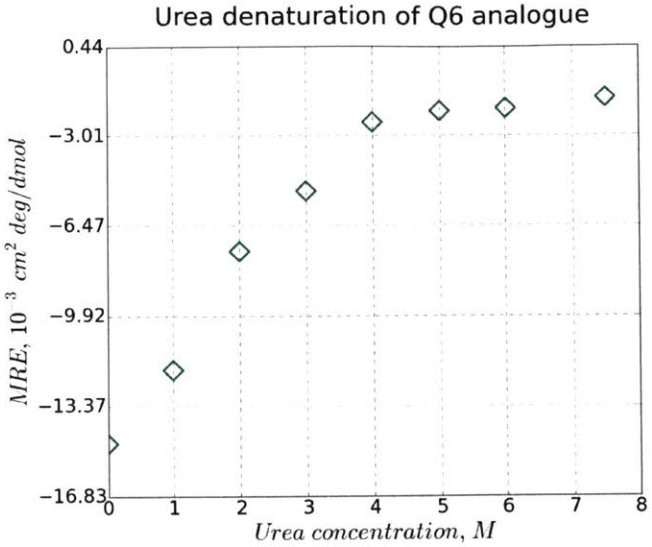

Far-UV CD spectrum of Q4 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
MRE, 10 -3 cm2 deg/dmol
Wavelength,nm

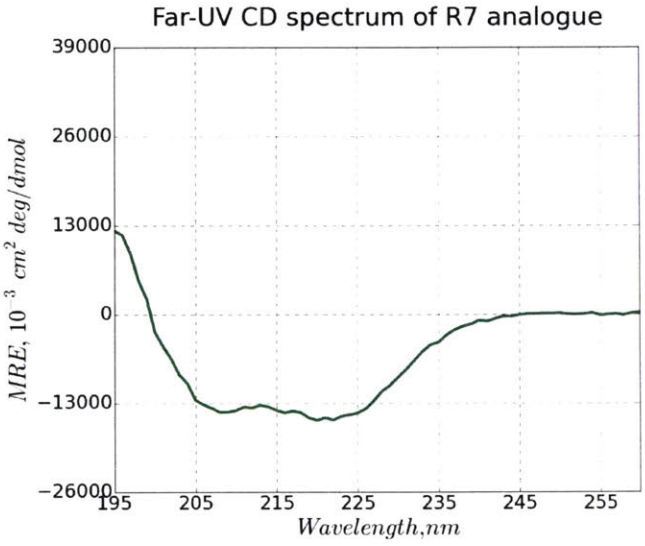

Urea denaturation of Q4 analogue
2.07
−2.84
−7.75
−12.66
−17.57
−22.48
0
1
2
3
4
5
6
7
8
MRE, 10−3 cm2 deg/dmol
Urea concentration, M

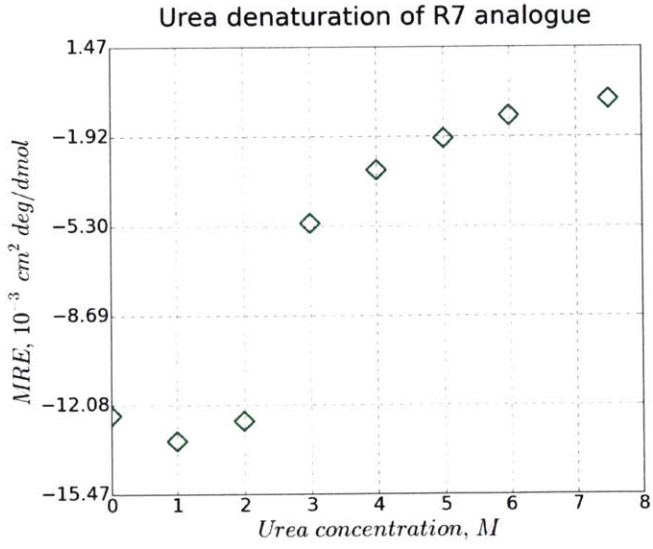

Far-UV CD spectrum of Q5 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
MRE, 10 -3 cm2 deg/dmol
Wavelength,nm

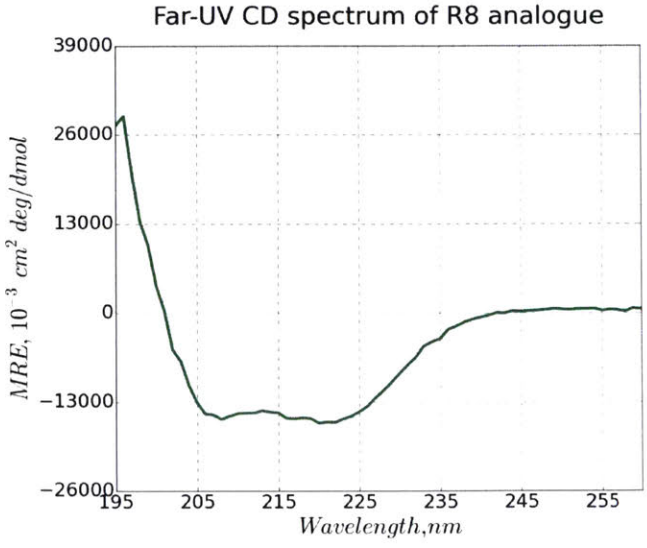

Urea denaturation of Q5 analogue
0.97
−2.59
−6.15
−9.71
−13.27
−16.83
0
1
2
3
4
5
6
7
8
MRE, 10 -3 cm2 deg/dmol
Urea concentration, M

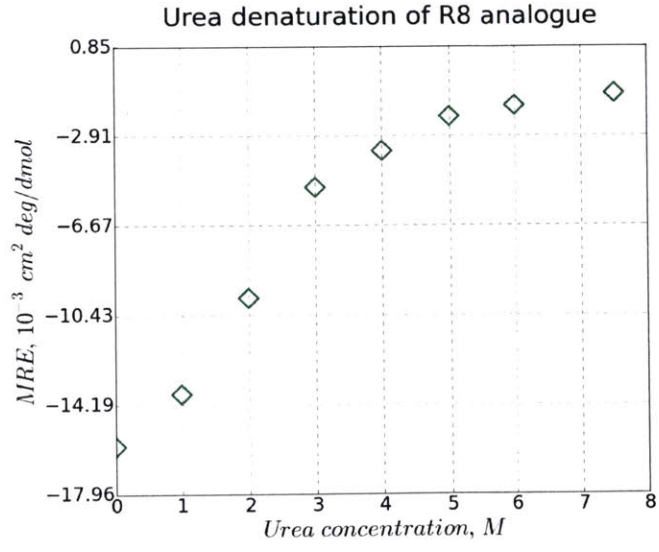

Far-UV CD spectrum of Q6 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol

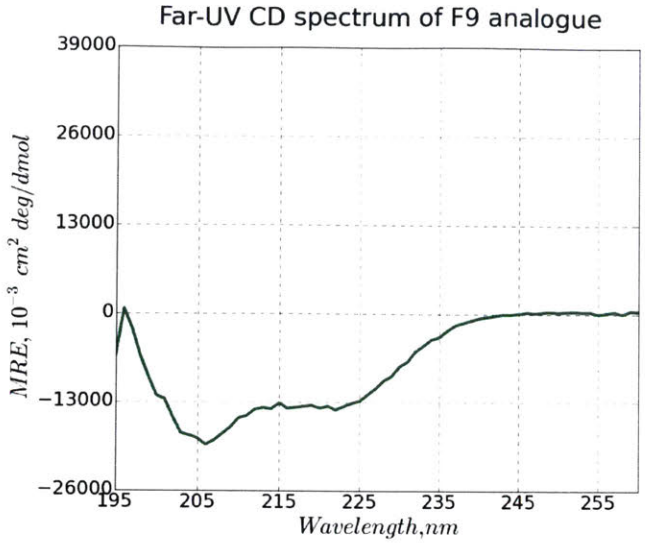

Urea denaturation of Q6 analogue
0.44
−3.01
−6.47
−9.92
−13.37
−16.83
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol

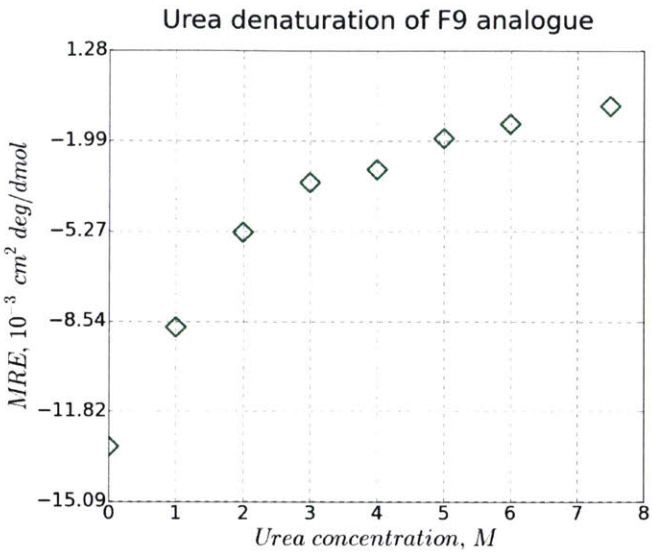

Far-UV CD spectrum of R7 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol

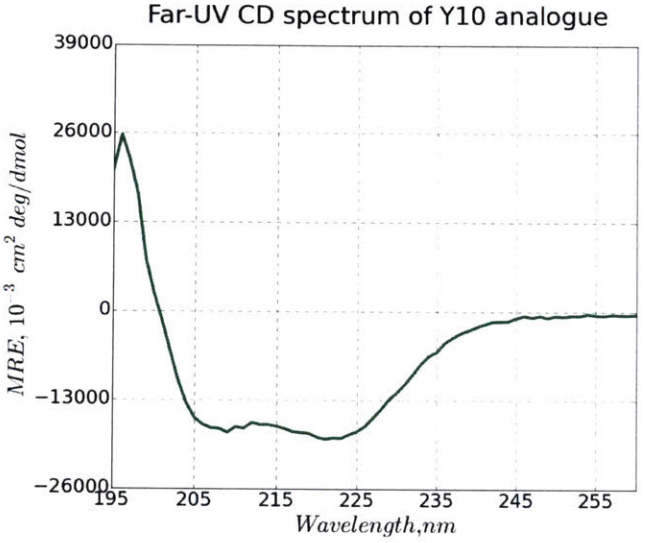

Urea denaturation of R7 analogue
1.47
−1.92
−5.30
−8.69
−12.08
−15.47
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol

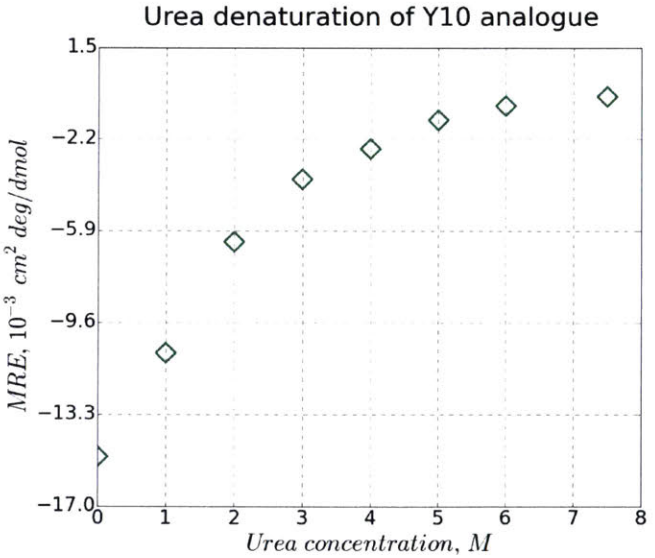

Far-UV CD spectrum of R8 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol

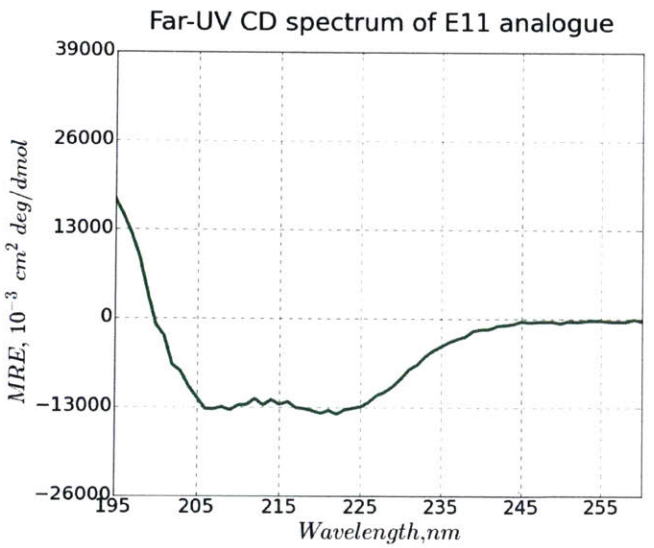

Urea denaturation of R8 analogue
0.85
−2.91
−6.67
−10.43
−14.19
−17.96
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol

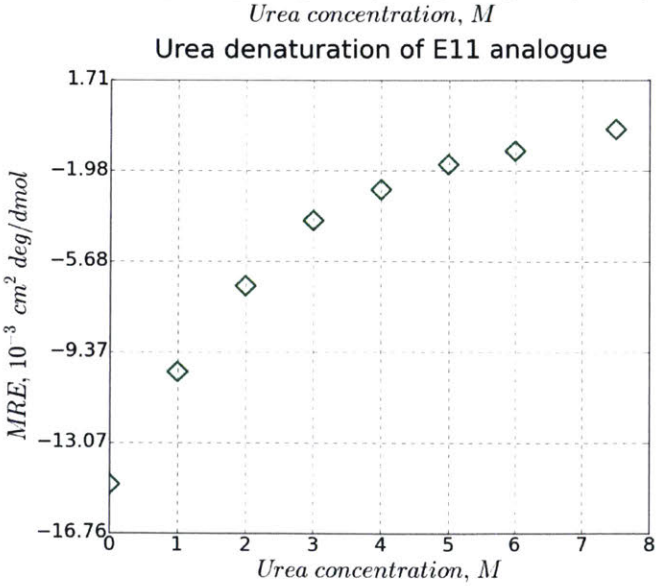

Far-UV CD spectrum of F9 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol

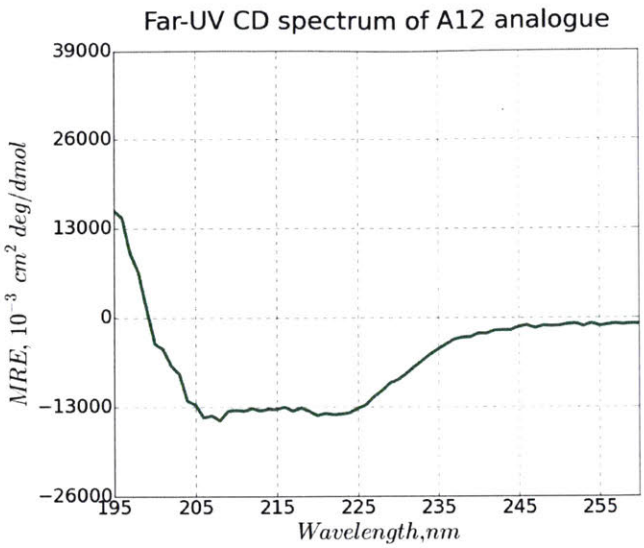

Urea denaturation of F9 analogue
1.28
−1.99
−5.27
−8.54
−11.82
−15.09
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol

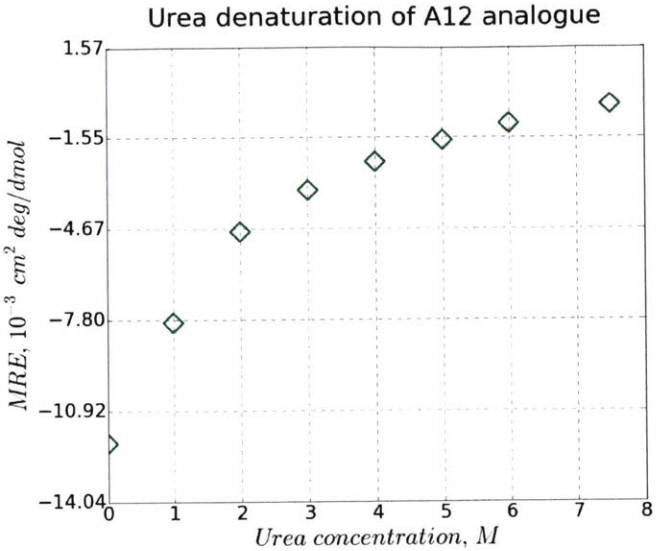

Far-UV CD spectrum of Y10 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol

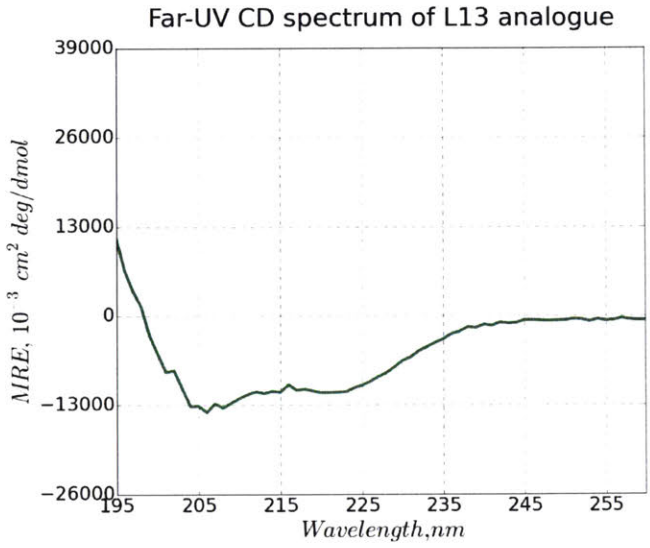

Urea denaturation of Y10 analogue
1.5
−2.2
−5.9
−9.6
−13.3
−17.0
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol

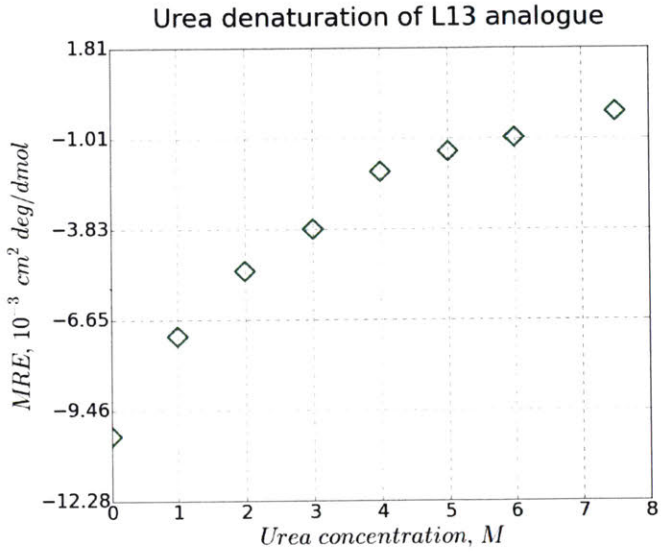

Far-UV CD spectrum of E11 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol

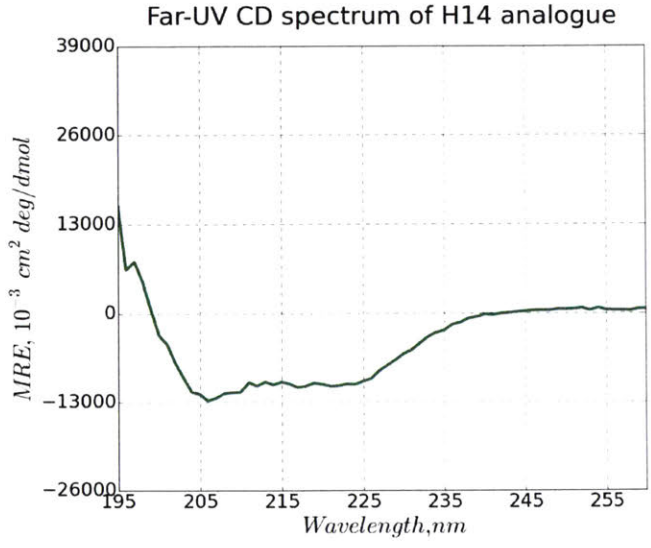

Urea denaturation of E11 analogue
1.71
−1.98
−5.68
−9.37
−13.07
−16.76
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol

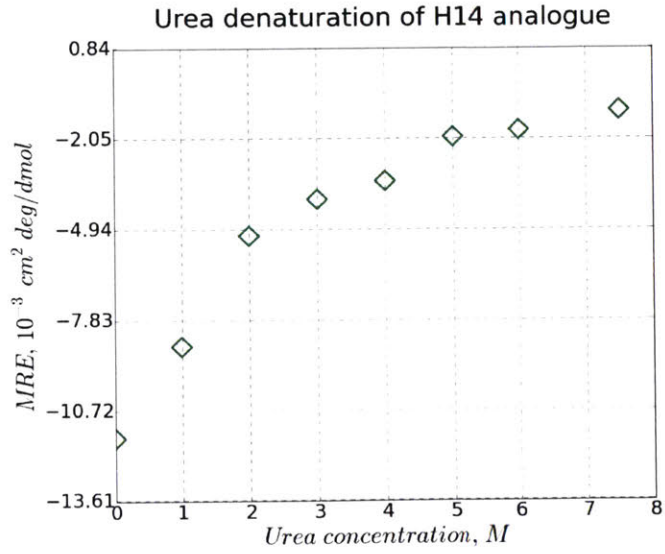

Far-UV CD spectrum of A12 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol

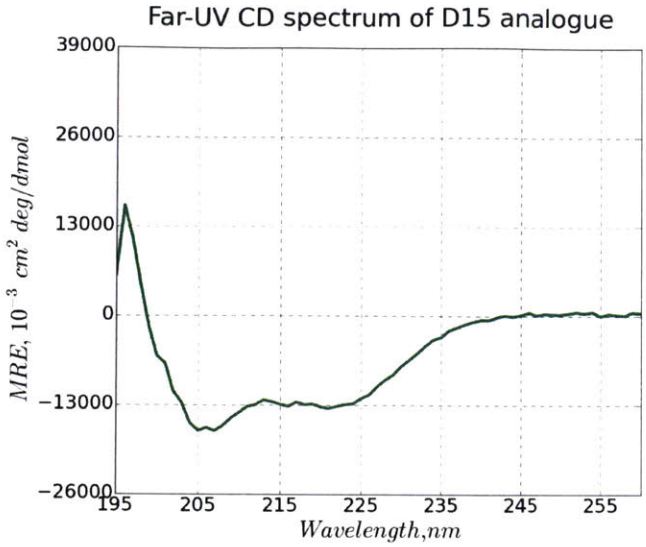

Urea denaturation of A12 analogue
1.57
−1.55
−4.67
−7.80
−10.92
−14.04
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol

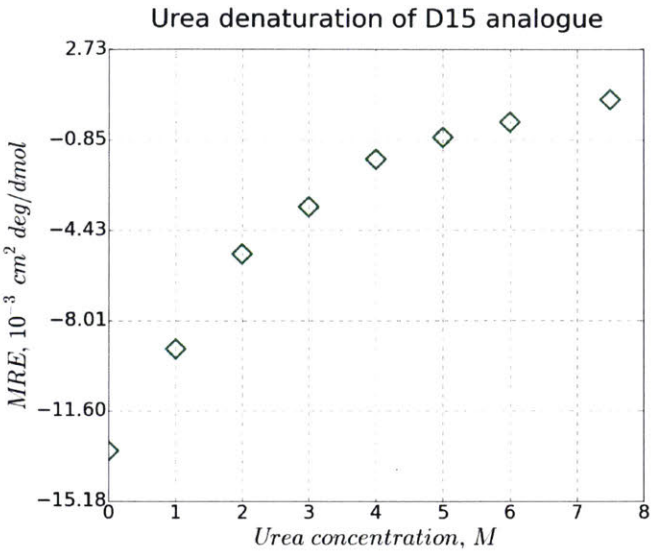

Far-UV CD spectrum of L13 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol

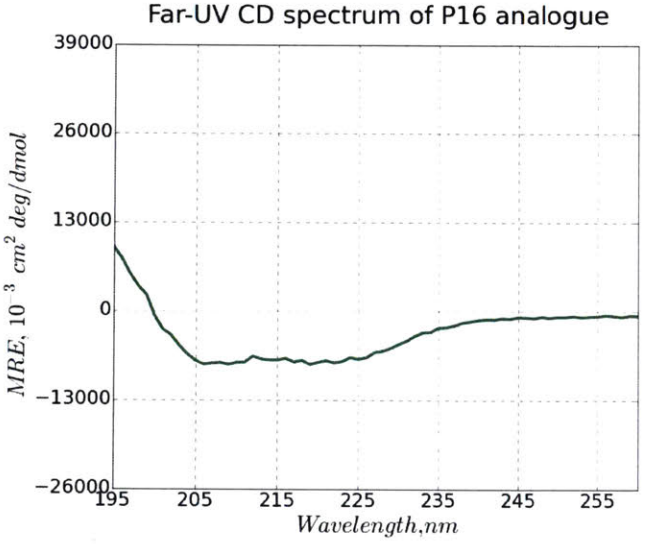

Urea denaturation of L13 analogue
1.81
−1.01
−3.83
−6.65
−9.46
−12.28
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol

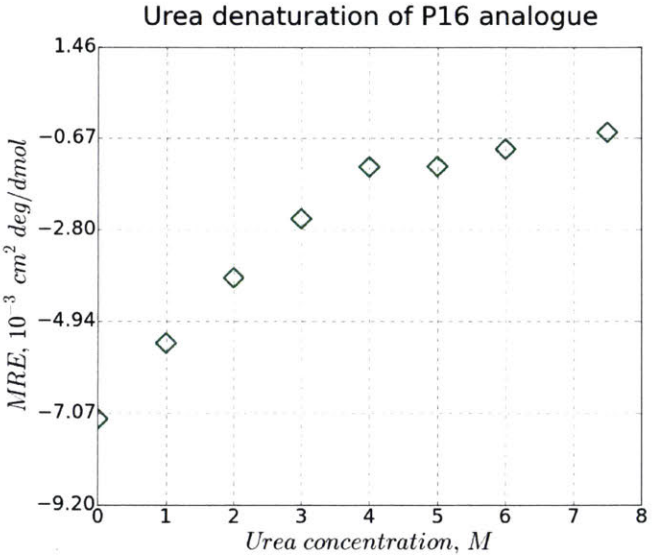

Far-UV CD spectrum of H14 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol

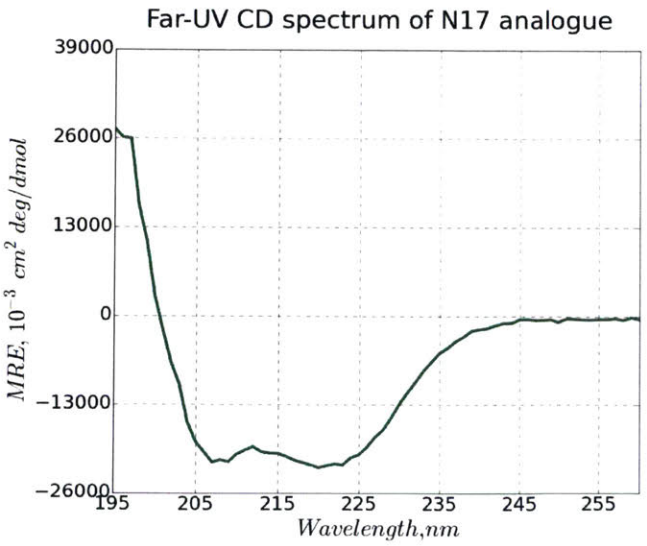

Urea denaturation of H14 analogue
0.84
−2.05
−4.94
−7.83
−10.72
−13.61
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol

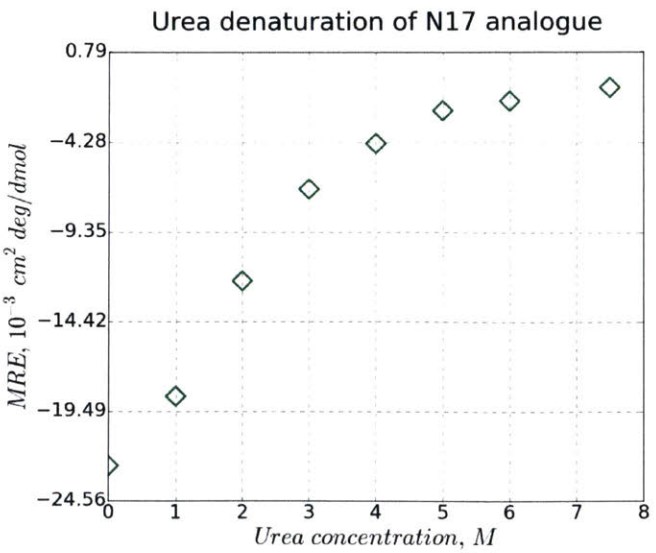

Far-UV CD spectrum of D15 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
MRE, 10 ^ 3 cm^2 deg/dmol
Wavelength,nm

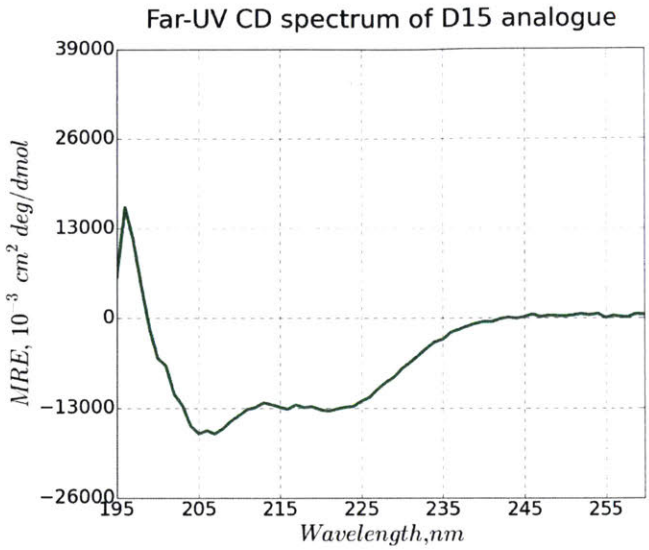

Urea denaturation of D15 analogue
2.73
−0.85
−4.43
−8.01
−11.60
−15.18
0
1
2
3
4
5
6
7
8
MRE, 10 ^ 3 cm^2 deg/dmol
Urea concentration, M

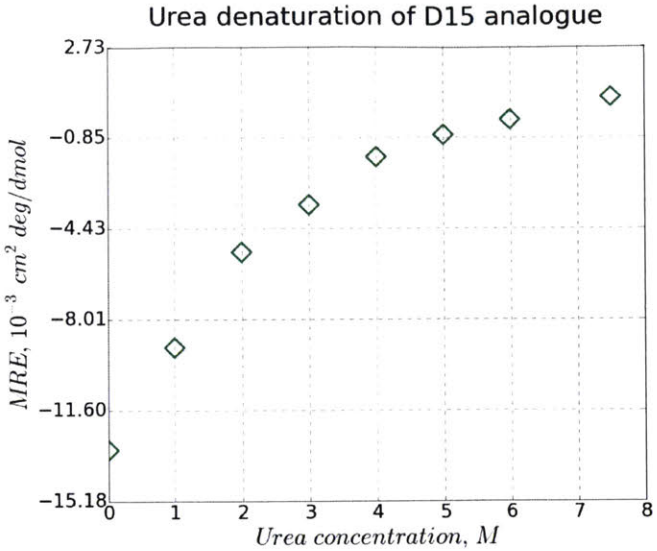

Far-UV CD spectrum of P16 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
MRE, 10 ^ 3 cm^2 deg/dmol
Wavelength,nm

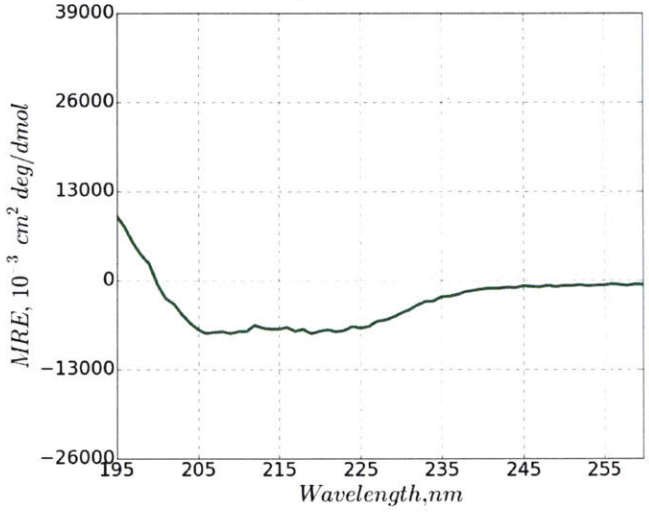

Urea denaturation of P16 analogue
1.46
−0.67
−2.80
−4.94
−7.07
−9.20
0
1
2
3
4
5
6
7
8
MRE, 10 ^ 3 cm^2 deg/dmol
Urea concentration, M

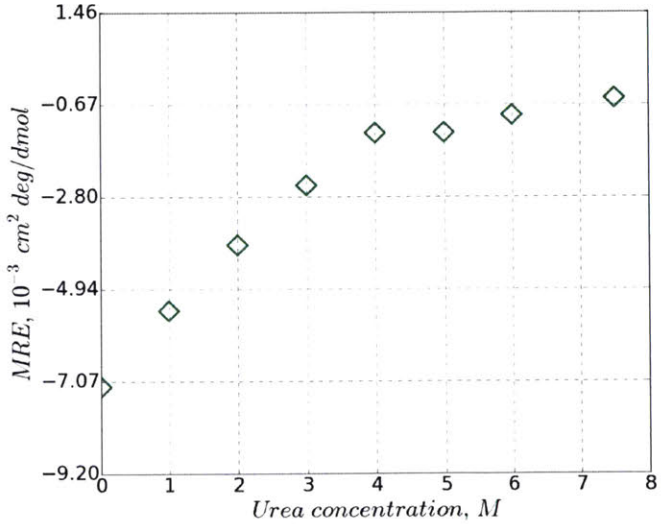

Far-UV CD spectrum of N17 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
MRE, 10 ^ 3 cm^2 deg/dmol
Wavelength,nm

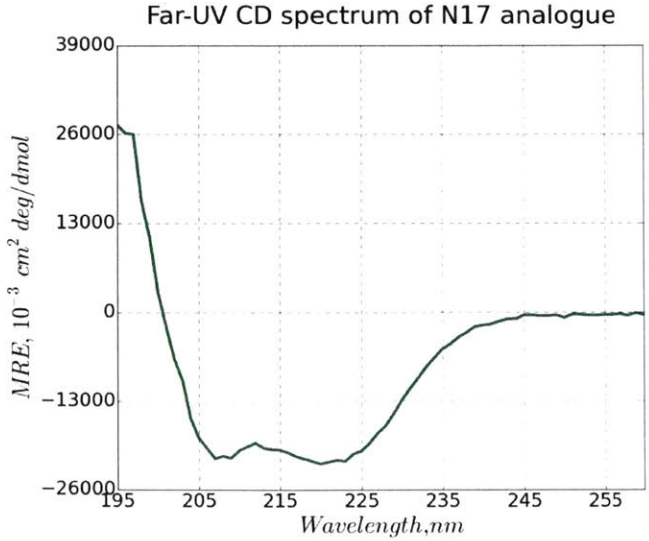

Urea denaturation of N17 analogue
0.79
−4.28
−9.35
−14.42
−19.49
−24.56
0
1
2
3
4
5
6
7
8
MRE, 10 ^ 3 cm^2 deg/dmol
Urea concentration, M

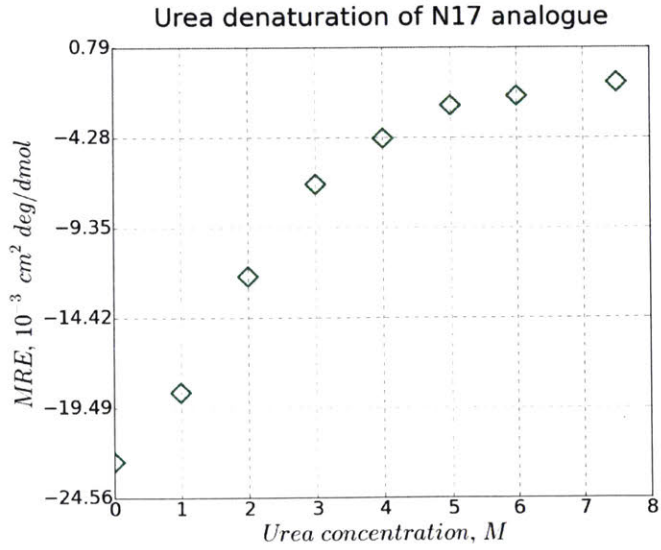

Far-UV CD spectrum of D15 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 ^ 3 cm^2 deg/dmol

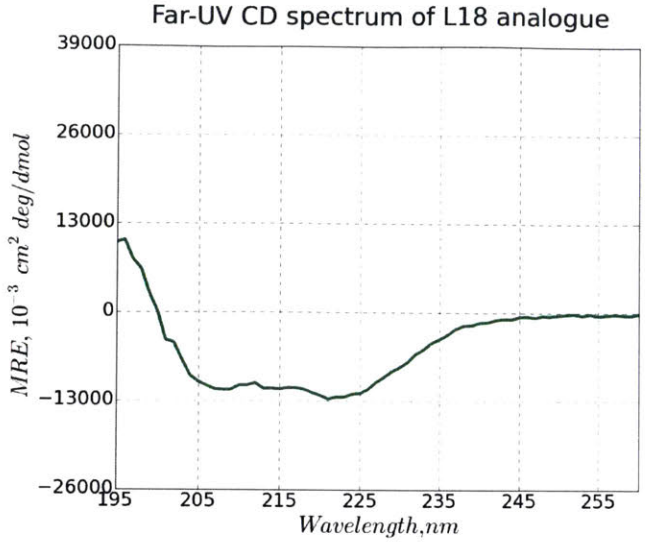

Urea denaturation of D15 analogue
2.73
−0.85
−4.43
−8.01
−11.60
−15.18
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 ^ 3 cm^2 deg/dmol

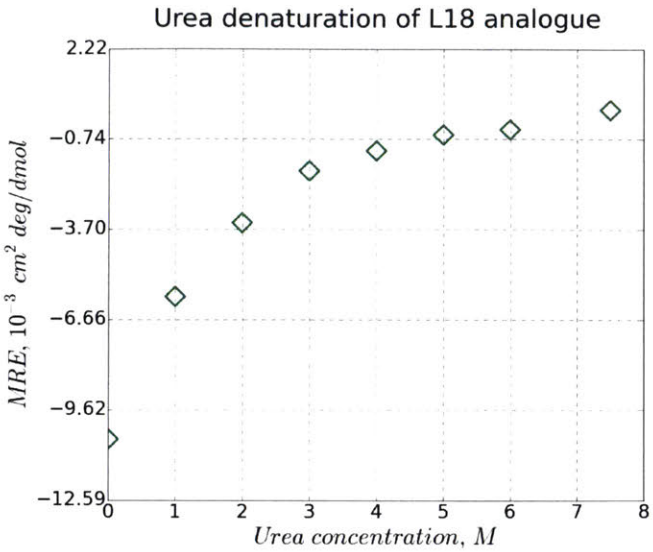

Far-UV CD spectrum of P16 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 ^ 3 cm^2 deg/dmol

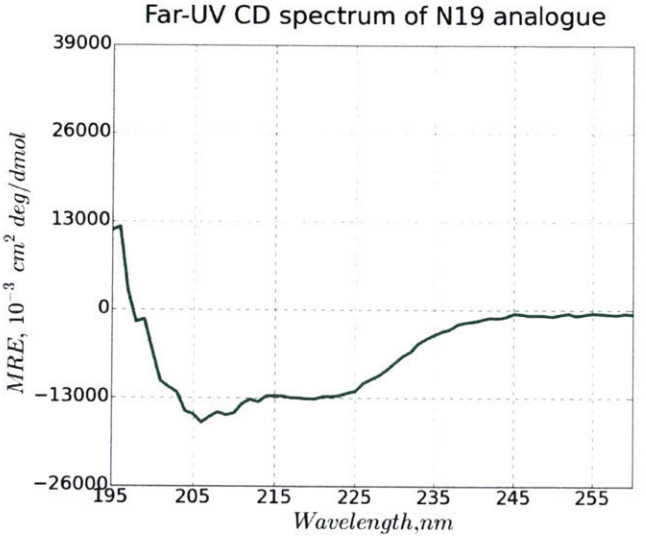

Urea denaturation of P16 analogue
1.46
−0.67
−2.80
−4.94
−7.07
−9.20
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 ^ 3 cm^2 deg/dmol

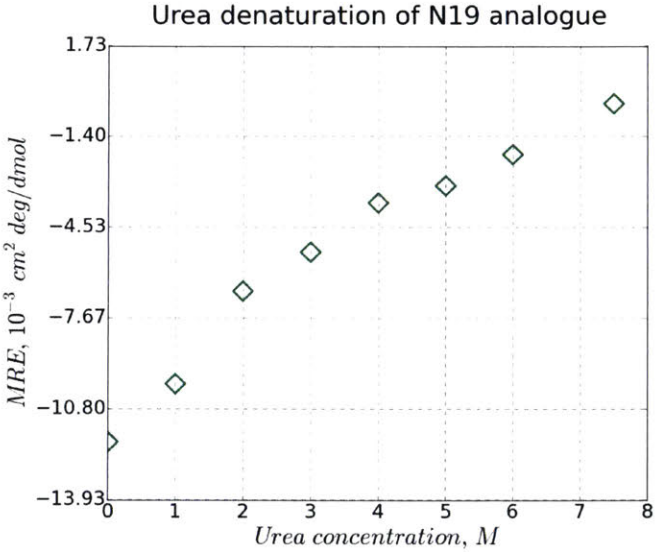

Far-UV CD spectrum of N17 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 ^ 3 cm^2 deg/dmol

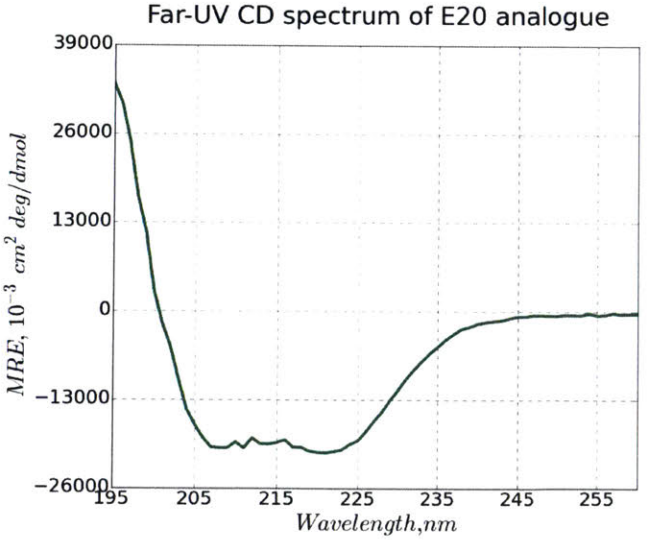

Urea denaturation of N17 analogue
0.79
−4.28
−9.35
−14.42
−19.49
−24.56
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 ^ 3 cm^2 deg/dmol

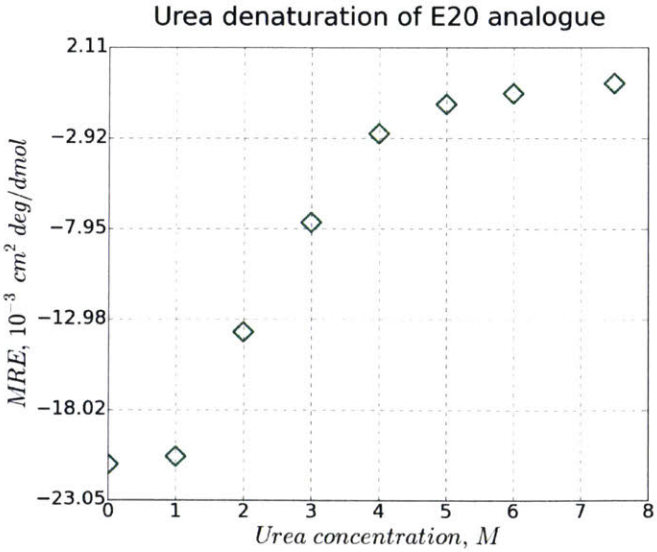

Far-UV CD spectrum of L18 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10^-3 cm^2 deg/dmol

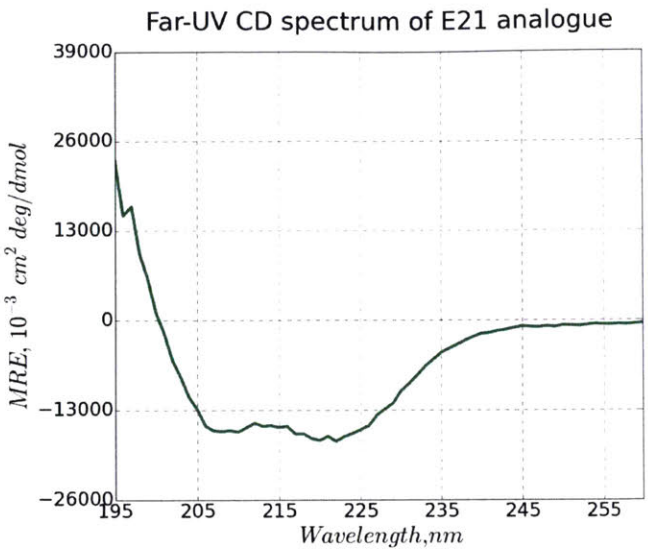

Urea denaturation of L18 analogue
2.22
−0.74
−3.70
−6.66
−9.62
−12.59
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10^-3 cm^2 deg/dmol

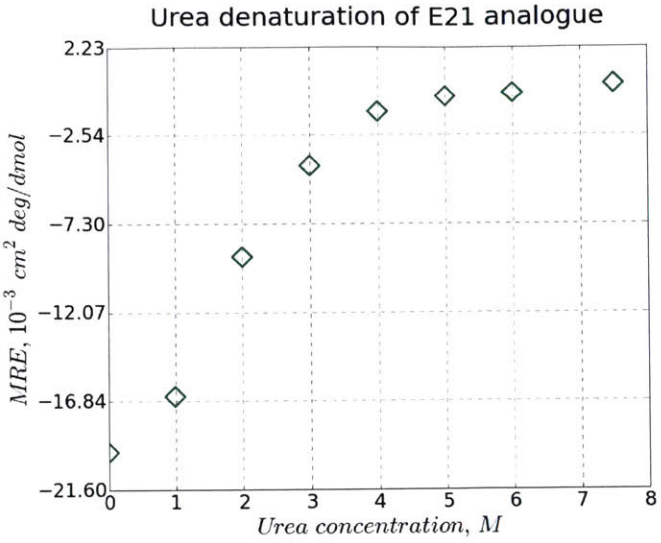

Far-UV CD spectrum of N19 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10^-3 cm^2 deg/dmol

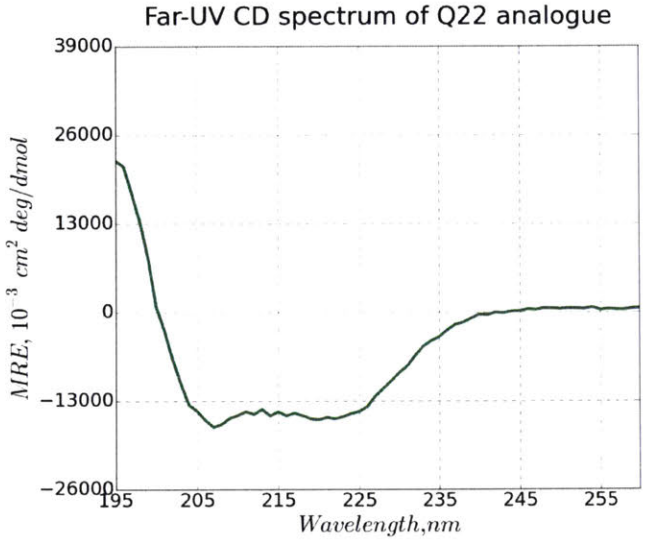

Urea denaturation of N19 analogue
1.73
−1.40
−4.53
−7.67
−10.80
−13.93
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10^-3 cm^2 deg/dmol

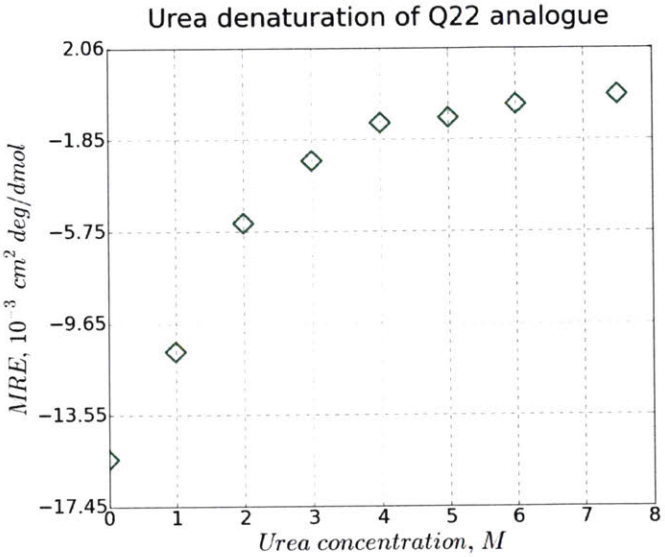

Far-UV CD spectrum of E20 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10^-3 cm^2 deg/dmol

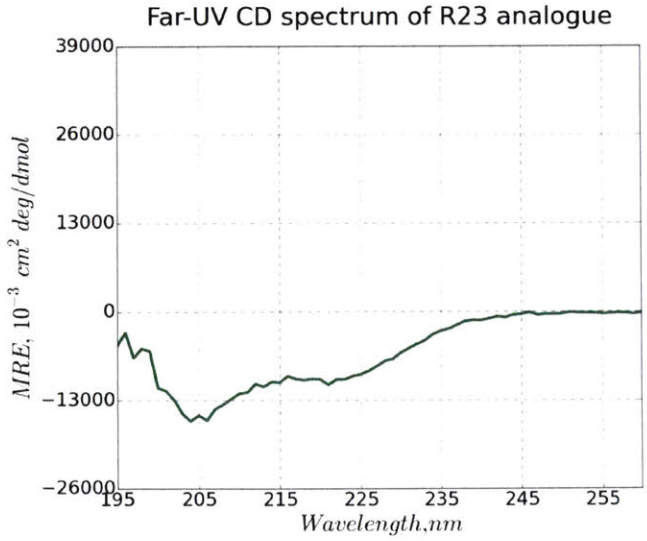

Urea denaturation of E20 analogue
2.11
−2.92
−7.95
−12.98
−18.02
−23.05
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10^-3 cm^2 deg/dmol

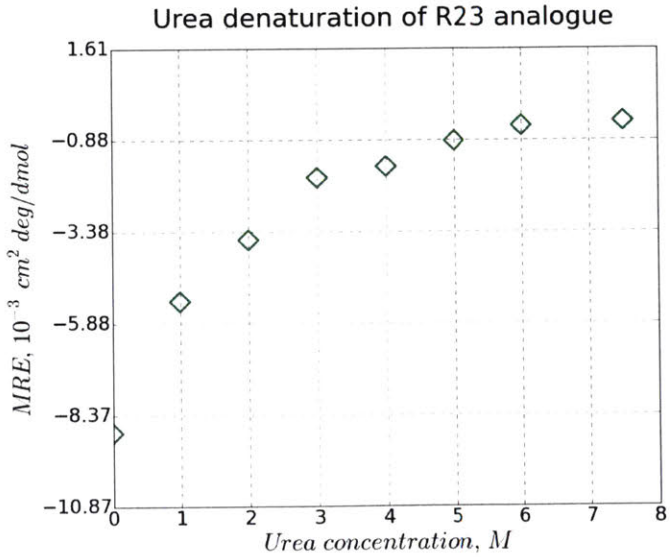

Far-UV CD spectrum of E21 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
$MRE, 10^{-3}\ cm^2\ deg/dmol$
$Wavelength, nm$

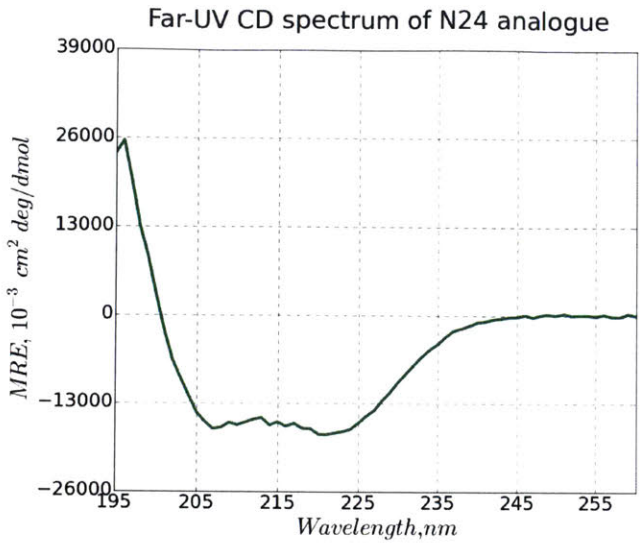

Urea denaturation of E21 analogue
2.23
−2.54
−7.30
−12.07
−16.84
−21.60
0
1
2
3
4
5
6
7
8
$MRE, 10^{-3}\ cm^2\ deg/dmol$
$Urea\ concentration, M$

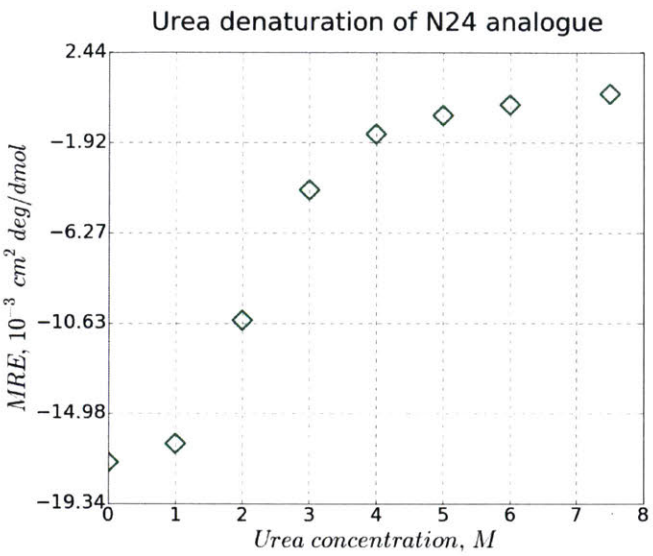

Far-UV CD spectrum of Q22 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
$MRE, 10^{-3}\ cm^2\ deg/dmol$
$Wavelength, nm$

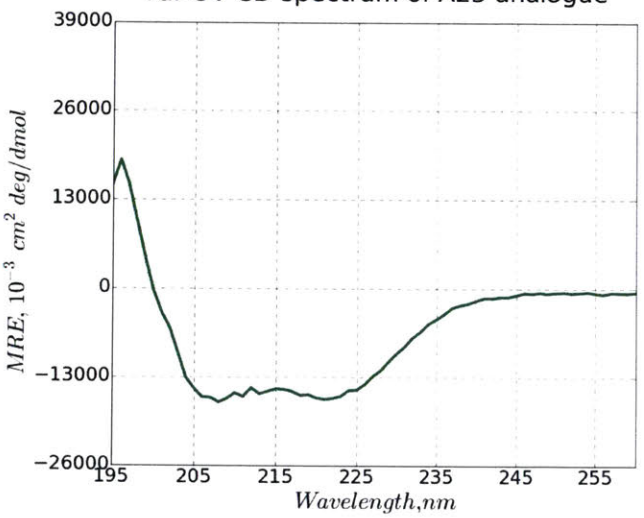

Urea denaturation of Q22 analogue
2.06
−1.85
−5.75
−9.65
−13.55
−17.45
0
1
2
3
4
5
6
7
8
$MRE, 10^{-3}\ cm^2\ deg/dmol$
$Urea\ concentration, M$

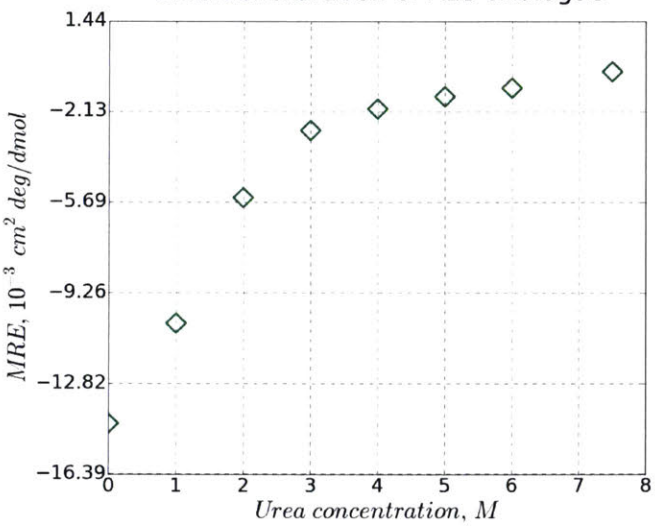

Far-UV CD spectrum of R23 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
$MRE, 10^{-3}\ cm^2\ deg/dmol$
$Wavelength, nm$

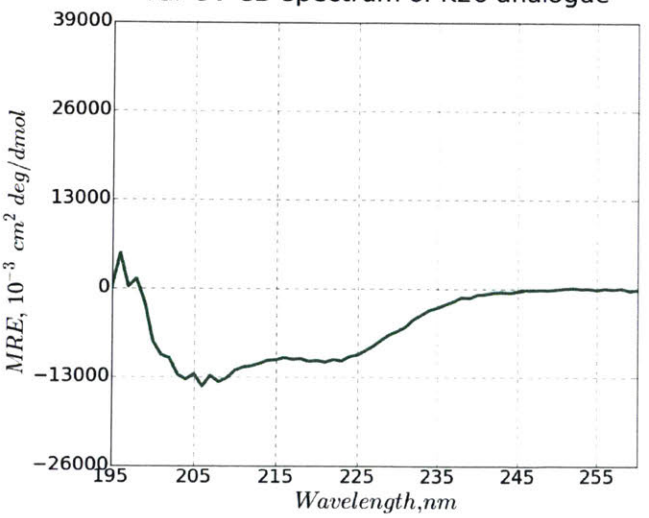

Urea denaturation of R23 analogue
1.61
−0.88
−3.38
−5.88
−8.37
−10.87
0
1
2
3
4
5
6
7
8
$MRE, 10^{-3}\ cm^2\ deg/dmol$
$Urea\ concentration, M$

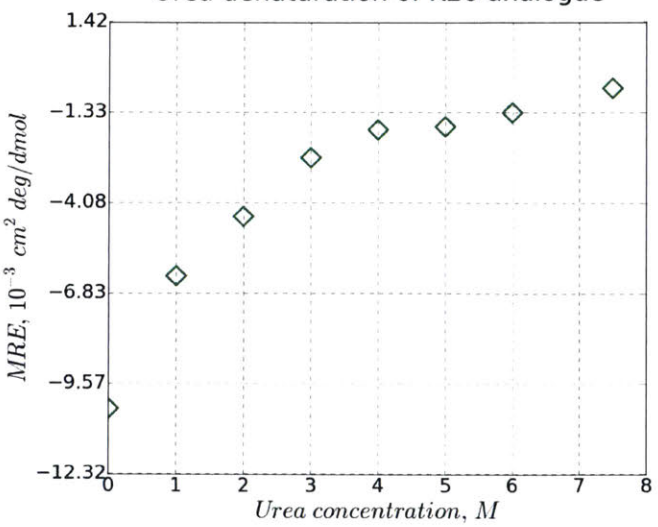

Far-UV CD spectrum of N24 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10^-3 cm^2 deg/dmol

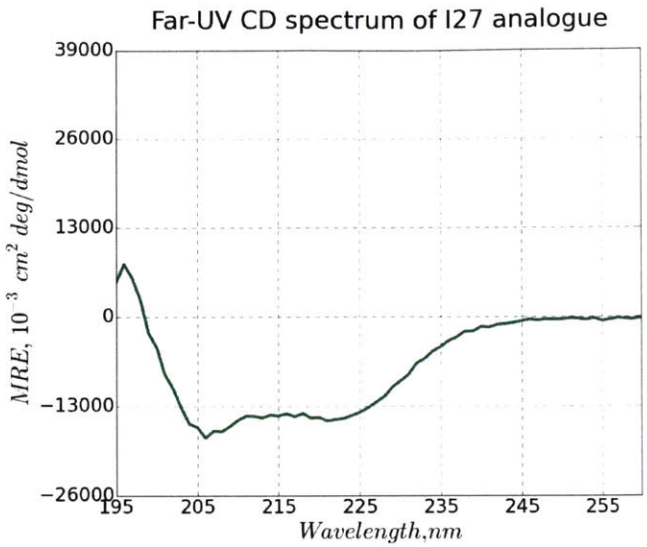

Urea denaturation of N24 analogue
2.44
−1.92
−6.27
−10.63
−14.98
−19.34
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10^-3 cm^2 deg/dmol

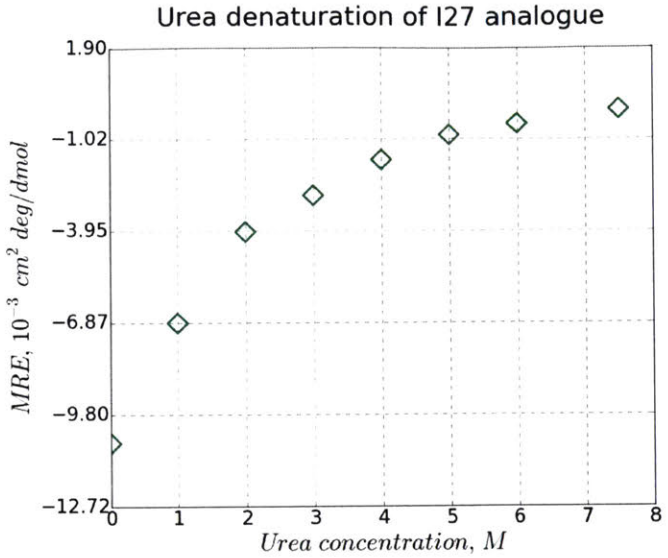

Far-UV CD spectrum of A25 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10^-3 cm^2 deg/dmol

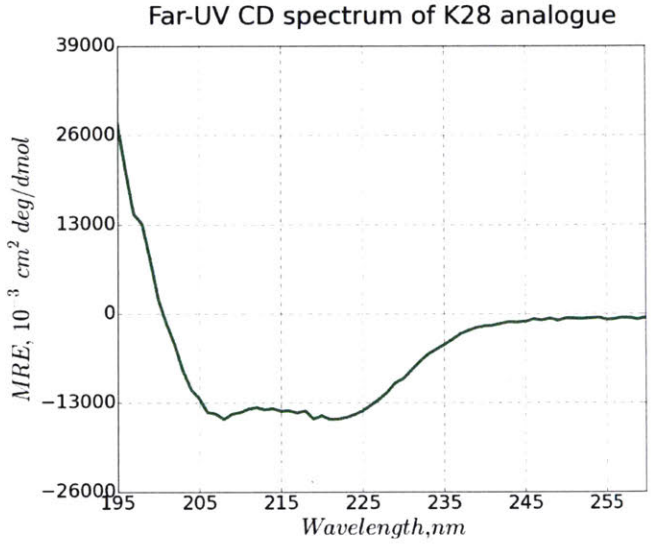

Urea denaturation of A25 analogue
1.44
−2.13
−5.69
−9.26
−12.82
−16.39
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10^-3 cm^2 deg/dmol

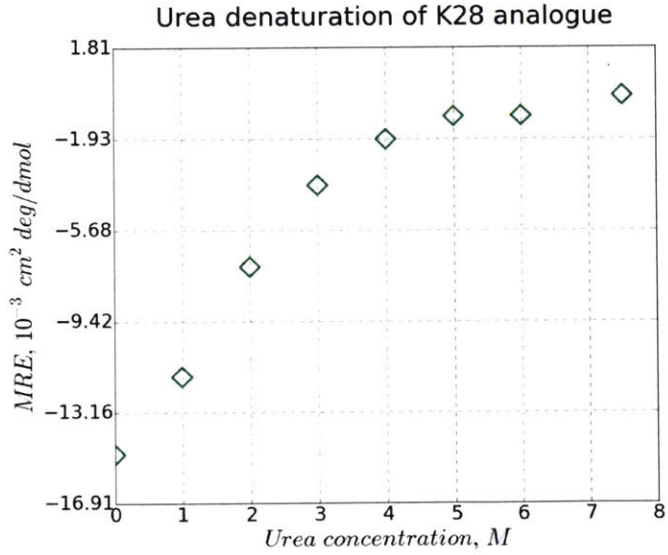

Far-UV CD spectrum of K26 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10^-3 cm^2 deg/dmol

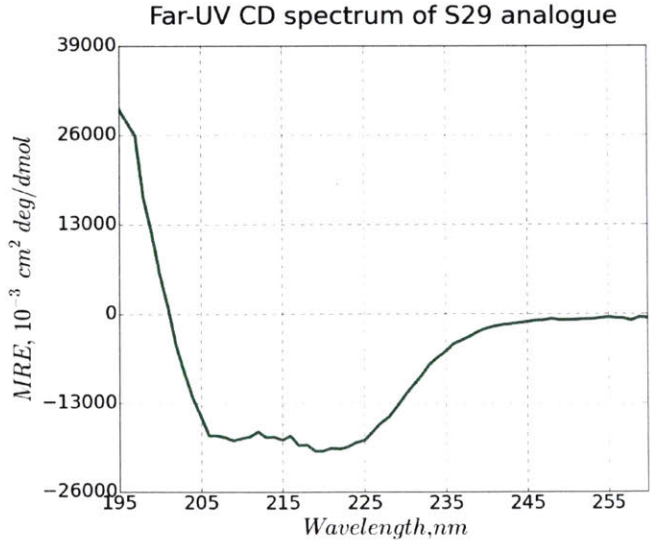

Urea denaturation of K26 analogue
1.42
−1.33
−4.08
−6.83
−9.57
−12.32
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10^-3 cm^2 deg/dmol

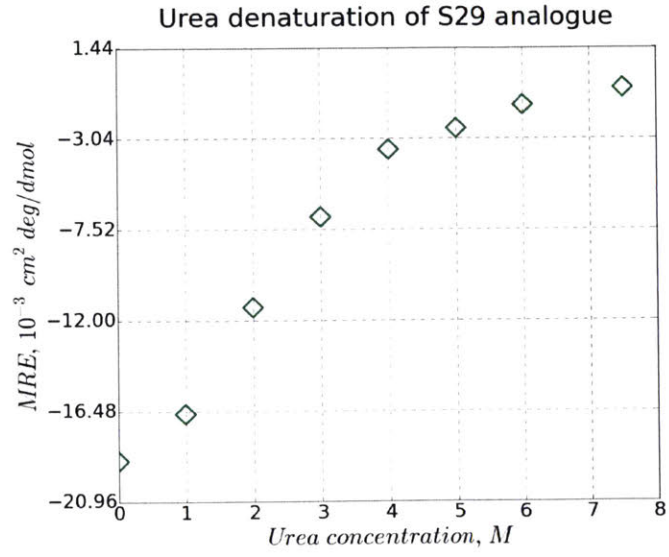

Far-UV CD spectrum of I27 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10^-3 cm^2 deg/dmol

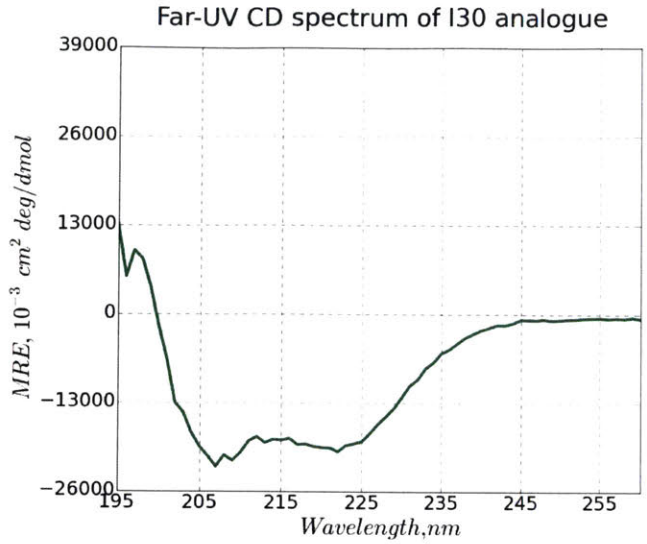

Urea denaturation of I27 analogue
1.90
−1.02
−3.95
−6.87
−9.80
−12.72
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10^-3 cm^2 deg/dmol

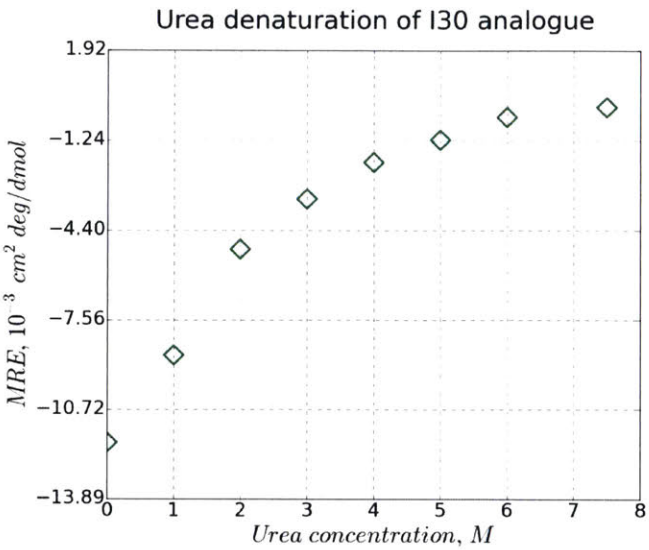

Far-UV CD spectrum of K28 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10^-3 cm^2 deg/dmol

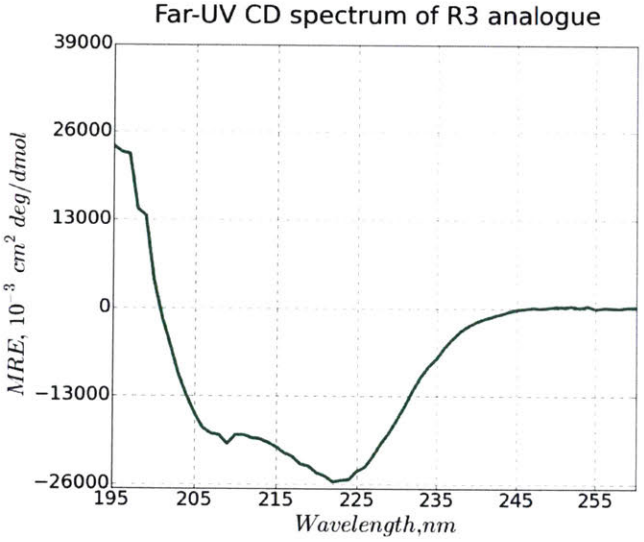

Urea denaturation of K28 analogue
1.81
−1.93
−5.68
−9.42
−13.16
−16.91
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10^-3 cm^2 deg/dmol

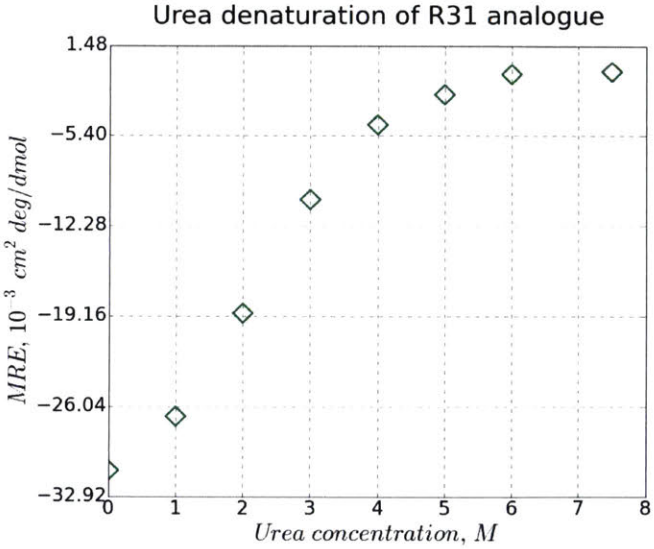

Far-UV CD spectrum of S29 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10^-3 cm^2 deg/dmol

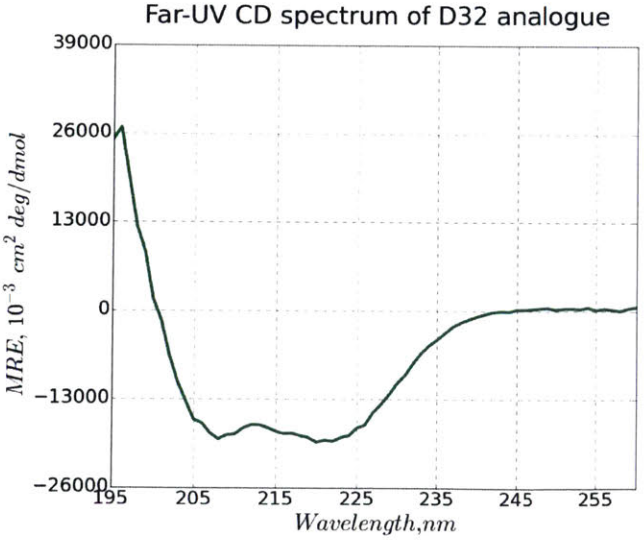

Urea denaturation of S29 analogue
1.44
−3.04
−7.52
−12.00
−16.48
−20.96
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10^-3 cm^2 deg/dmol

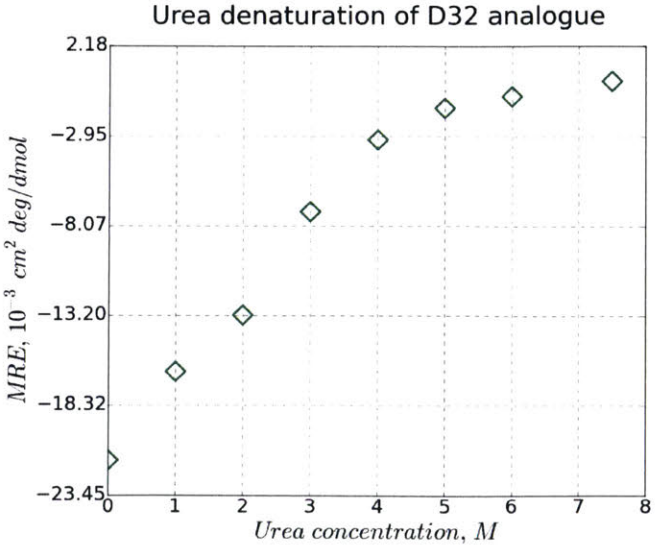

Far-UV CD spectrum of I30 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol

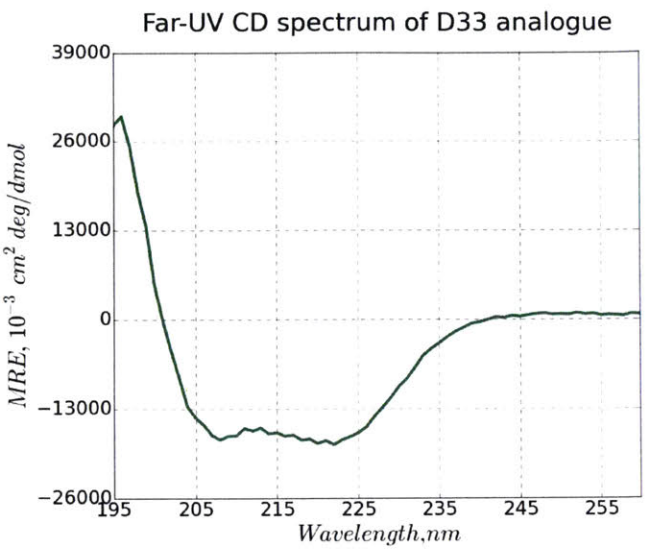

Urea denaturation of I30 analogue
1.92
−1.24
−4.40
−7.56
−10.72
−13.89
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol

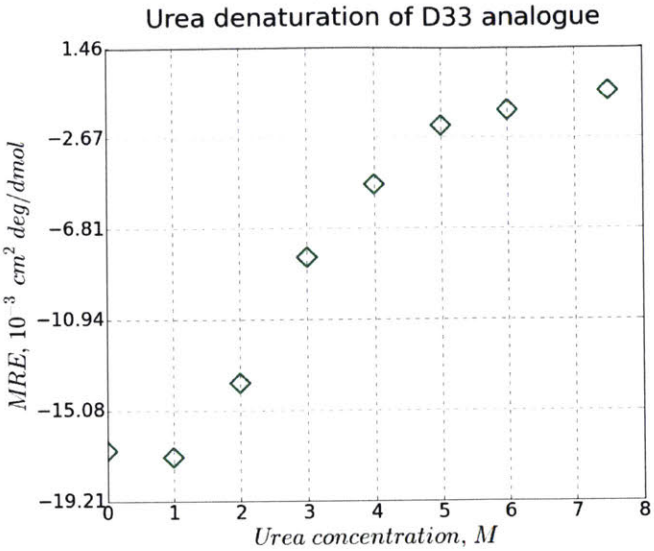

Far-UV CD spectrum of R3 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol


Urea denaturation of R31 analogue
1.48
−5.40
−12.28
−19.16
−26.04
−32.92
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol


Far-UV CD spectrum of D32 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
Wavelength,nm
MRE, 10 -3 cm2 deg/dmol


Urea denaturation of D32 analogue
2.18
−2.95
−8.07
−13.20
−18.32
−23.45
0
1
2
3
4
5
6
7
8
Urea concentration, M
MRE, 10 -3 cm2 deg/dmol


Far-UV CD spectrum of D33 analogue
39000
26000
13000
0
−13000
−26000
195
205
215
225
235
245
255
$MRE, 10^{-3}\ cm^2\ deg/dmol$
$Wavelength, nm$


Urea denaturation of D33 analogue
1.46
−2.67
−6.81
−10.94
−15.08
−19.21
0
1
2
3
4
5
6
7
8
$MRE, 10^{-3}\ cm^2\ deg/dmol$
$Urea\ concentration, M$

A4.19 Appendix 4 references

(1) Simon, M. D.; Heider, P. L.; Adamo, A.; Vinogradov, A. A.; Mong, S. K.; Li, X.; Berger, T.; Policarpo, R. L.; Zhang, C.; Zou, Y.; Liao, X.; Spokoyny, A. M.; Jensen, K. F.; Pentelute, B. L. Chembiochem 2014, 15 (5), 713.

(2) Durek, T.; Zhang, J.; He, C.; Kent, S. B. H. Org. Lett. 2007, 9 (26), 5497.

(3) Wentzel, A.; Christmann, A.; Krätzner, R.; Kolmar, H. J. Biol. Chem. 1999, 274 (30), 21037.

(4) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. J. Comput. Chem. 2004, 25 (13), 1605.

(5) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. J Chem Theory Coput. 2008, 4, 435.

(6) Jiang, F.; Wu, Y.-D. J. Am. Chem. Soc 2014, 136, 9536.

(7) Jiang, F.; Zhou, C.-Y.; Wu, Y.-D. J. Phys. Chem. B 2014, 118, 6983.

(8) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. J. Chem. Phys. 2004, 120 (20), 9665.

(9) Hoover, W. G. Phys. Rev. A 1985, 31 (3), 1695.

(10) Nosé, S. Mol. Phys. 1984, 52, 255

(11) Cheng, A.; Merz, K. M. J. Phys. Chem. 1996, 100, 1927

(12) Mor, A.; Ziv, G.; Levy, Y. J. Comput. Chem. 2008, 29, 1992.

(13) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. J. Chem. Phys. 1984, 81 (8), 3684.

(14) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. J. Comput. Chem. 1997, 18 (12), 1463.

(15) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. J. Chem. Phys. 1995, 103 (19), 8577.

(16) Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L. Nat. Chem. 2016, 8, 120-128.

(17) Louis-Jeune, C.; Andrade-Navarro, M. A.; Perez-Iratxeta, C. Proteins Struct. Funct. Bioinforma. 2012, 80, 374.

(18) Pace, C. N.; Shirley, B. A.; Thomson, J. A. Protein structure: A practical approach. 1989, pp 311–330.

(19) Allen, M. P.; Tildesley, D. J. Computer Simulations of Liquids; Oxford University Press, 1989.