11-11-2011

Synthesis of Oligopeptide Fragments for an Skp1 Mimetic

Jarret Curtis

Synthesis of Oligopeptide Fragments for an Skp1 Mimetic

Jarret Curtis

Department of Chemistry, Louisiana State University
Baton Rouge, LA 70803

Undergraduate Honors Thesis

Under Supervision of Dr. Carol Taylor

November 11, 2011

Abstract:

Gnt1 of *Dictyostelium discoideum* is a cytoplasmic enzyme responsible for transferring GlcNAc from UDP to Pro^{143} of Skp1, an important enzyme in the ubiquitin-proteasome system. This honors thesis describes the synthesis of oligopeptide fragments to be incorporated into an Skp1 mimetic. The mimetic has been designed to emulate the secondary structure of the full-length protein via a synthetic hydrogen bond surrogate (HBS) system that will covalently bind two of the amino acid residues together to induce the first turn of an alpha-helix. The design features for this system are expected to become a useful enzyme binding motif for the further study of enzyme activity and protein inhibition.

This project describes the synthesis of two major components of the mimetic: a glutamic acid trimer wherein the *C*-terminal residue has been *N*-allylated, and a Gln-Ile-Arg-Lys tetrapeptide. The tripeptide will be merged with another strategically placed alkene via a ring closing metathesis reaction in order to render the HBS system. The synthesis of the *N*-allyl glutamic acid residue is detailed, along with its incorporation into the tripeptide. A similar strategy was invoked to assemble the QIRK tetrapeptide.

Contents:

1. Introduction
    1.1 An Overview of the Ubiquitin-Proteasome Pathway
    1.2 SCF Complex and the Role of Skp1
    1.3 Post-Translational Modifications of Skp1 in *Dictyostelium* and Their Effect on Cell Function
    1.4 Implications of Gnt1-Mediated Glycosylation in Lower Eukaryotes
    1.5 Enzyme Kinetics of Gnt1 in *Dictyostelium discoideum*
    1.6 The α-Helix of Skp1 and the Hydrogen Bond Surrogate System
    1.7 Proposed Synthesis of Skp1 Mimetic

2. Discussion and Results
    2.1 Synthesis of the *N*-Allyl-Glutamic Acid Residue
        2.1.1 Esterification of the *C*-terminus
        2.1.2 Exchange of *N*-terminal Protecting Group
        2.1.3 Addition of the Allyl Group
        2.1.4 Removal of the Nosylate to Reveal the Target Residue
    2.2 Incorporation of *N*-Allyl-Glu into a Glutamic Acid Tripeptide
    2.3 Synthesis of the QIRK oligopeptide sequence
    2.4 Future Work: Ring Closing Metathesis and Elongation of the Polypeptide

3. Experimental
    3.1 General
    3.2 Fmoc-*L*-Glu(O^{t}Bu)-OMe
    3.3 *N*-Nosyl-*L*-Glu(O^{t}Bu)-OMe
    3.4 *N*-Nosyl-*N*-Allyl-Glu(O^{t}Bu)-OMe
    3.5 *N*-Allyl-Glu(O^{t}Bu)-OMe
    3.6 Fmoc-Glu(O^{t}Bu)-Glu(O^{t}Bu)-*N*-Allyl-Glu(O^{t}Bu)-OMe
    3.7 Fmoc-Gln(NHTrt)-Ile-Arg(NHPbf)-Lys(NHBoc)-ONH$_2$

4. References

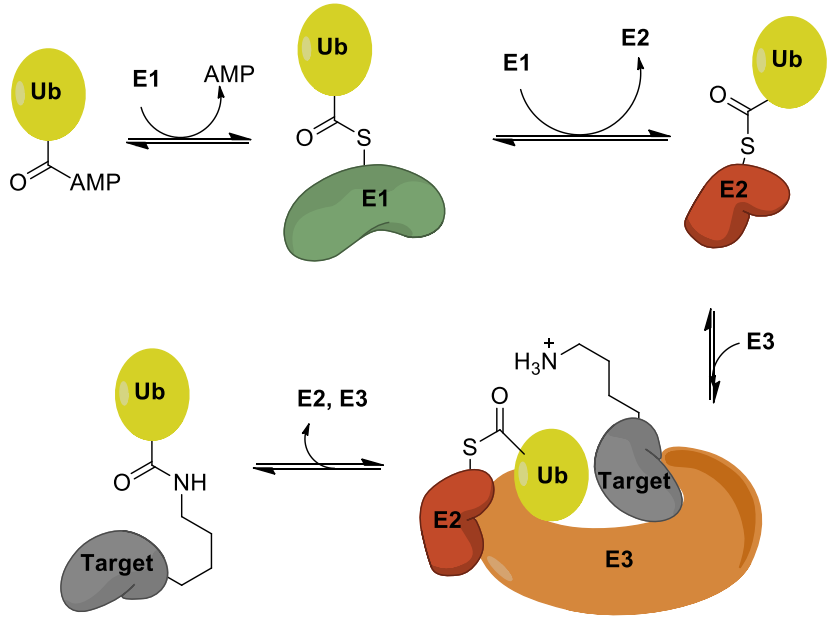
Appendix: Selected ^{1}H NMR spectra

1. Introduction

1.1 An Overview of the Ubiquitin-Proteasome Pathway

Proteins serve a multitude of functions within the cell. They can serve as structural support (collagen, actin, keratins), transporters (hemoglobin, serum albumin), messengers (insulin, glucagon, vasopressin), and the many enzymes that catalyze the chemical reactions necessary for survival and proliferation. Just as the cell must constantly synthesize new proteins as they become necessary, it degrades those that become old or damaged or are in excess. This degradation occurs in the cytoplasm of the cell through the metabolic pathway known as the ubiquitin-proteasome pathway.

The basic pathway of the ubiquitin system uses a messenger protein, ubiquitin, to signal a target protein for degradation in the proteasome. This is accomplished in two stages. The first stage involves covalent attachment of ubiquitin to the target via a series of transfers catalyzed by three 'helper enzymes', E1, E2, and E3. The second stage is the actual breakdown of the target in the proteasome.



Scheme 1.1: The Ubiquitination pathway, showing the transfer of ubiquitin (Ub) via the helper proteins and finally to the target molecule.

Formation of the covalent linkage between the target protein and ubiquitin is accomplished through an enzyme-mediated pathway (Scheme 1.1). After initial activation of ubiquitin via AMP binding, the next step of the pathway is binding of ubiquitin to E1, the ubiquitin-activating enzyme. The activated ubiquitin is then covalently transferred to E2, known as the ubiquitin-conjugating enzyme. E2, a carrier protein, delivers ubiquitin to the final enzyme, E3 – known as ubiquitin ligase. E3 binds both the E2-ubiquitin complex and the target protein, and catalyzes the transfer of ubiquitin to the target molecule. This process is repeated several times until a chain of four or more ubiquitin molecules is attached end to end. This poly-ubiquitination increases the effectiveness of the degradation signal for the target molecule.

The second stage involves the breakdown of the target protein by the proteasome. The proteasome is a large digestive enzyme made up of three main subunits: the catalytic core and two regulatory caps (Figure 1.1). As the target-ubiquitin complex enters the proteasome, ubiquitin molecules are cleaved and recycled back into the cytoplasm, but the target protein is sent through the catalytic center of the proteasome where it is broken down into its constituent amino acids. By this process, surplus and damaged proteins are broken down, and the resulting amino acids recycled to make new proteins.[1]

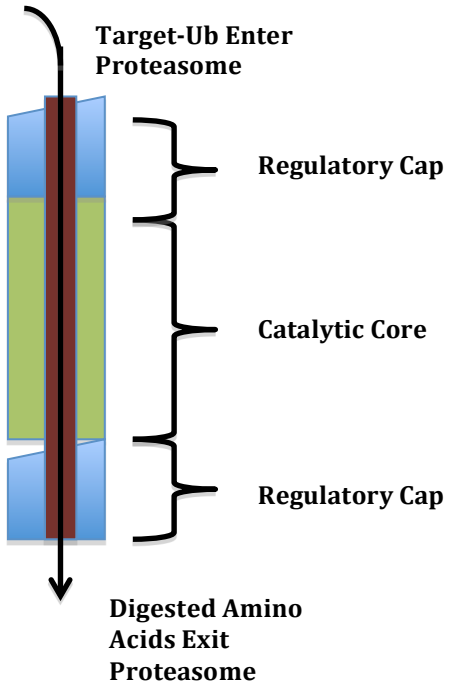
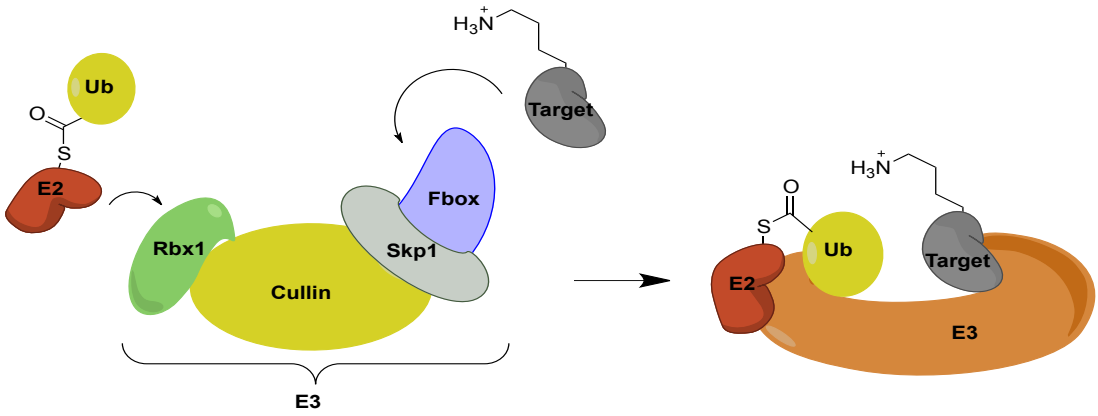


Figure 1.1: The proteosome with the catalytic core surrounded by the two regulatory caps

The focus of this research project is on one of the subunits of the E3 ubiquitin ligases. These are arguably the most important of the three types of 'helper' proteins. Throughout the entire process, they are the only proteins that are specific for the target proteins, *i.e.*, they are able to recognize and bind the target proteins, usually based on specific amino acid sequences at various loci on the target protein (most commonly, the *N*-terminal residue, cyclin destruction boxes, and PEST sequences). Owing to this specificity, in combination with the fact that E3 is the catalyst for formation of the target-ubiquitin isopeptide bond, E3 proteins are extremely important for functionality and proliferation of cells.[1]

1.2 SCF Complex and the Role of Skp1

In eukaryotes, E3 ubiquitin ligases contain a highly conserved structure known as the SCF complex. This complex constists of four major polypeptide subunits: Skp1, Cullin, Fbox protein, and Rbx1, as well as some other minor components (Figure 1.2). Each of the



Scheme 1.1: Schematic diagram of the structure of E3 SCF complex (left) linking with E2 and the target (right).

four core subunits is critical to the overall functionality of the protein, whether participating as a catalyst or regulator, or maintaining the structure of the protein. As described above, the function of this protein complex is to covalently link ubiquitin to the target molecule, and the horseshoe shape of the protein is essential to that functionality.

The proposed mechanism of action for the E3 proteins involves the Fbox subunit independently recognizing and binding the target molecule. The Fbox-target complex then associates itself with the rest of the E3 complex by binding the Skp1 subunit. At the same time, on the other side of the protein, the Rbx1 subunit of E3 binds the E2-ubiquitin complex. The structure of the protein places these two bound complexes in the appropriate proximity and alignment to catalyze the transfer of ubiquitin to the target.

The Skp1 subunit serves multiple purposes: it is a major structural component. Along with the Cullin subunit, Skp1 is also responsible for Fbox recognition and binding, and serves major regulatory functions within the overall protein (Figure 1.3).[2,3,15]

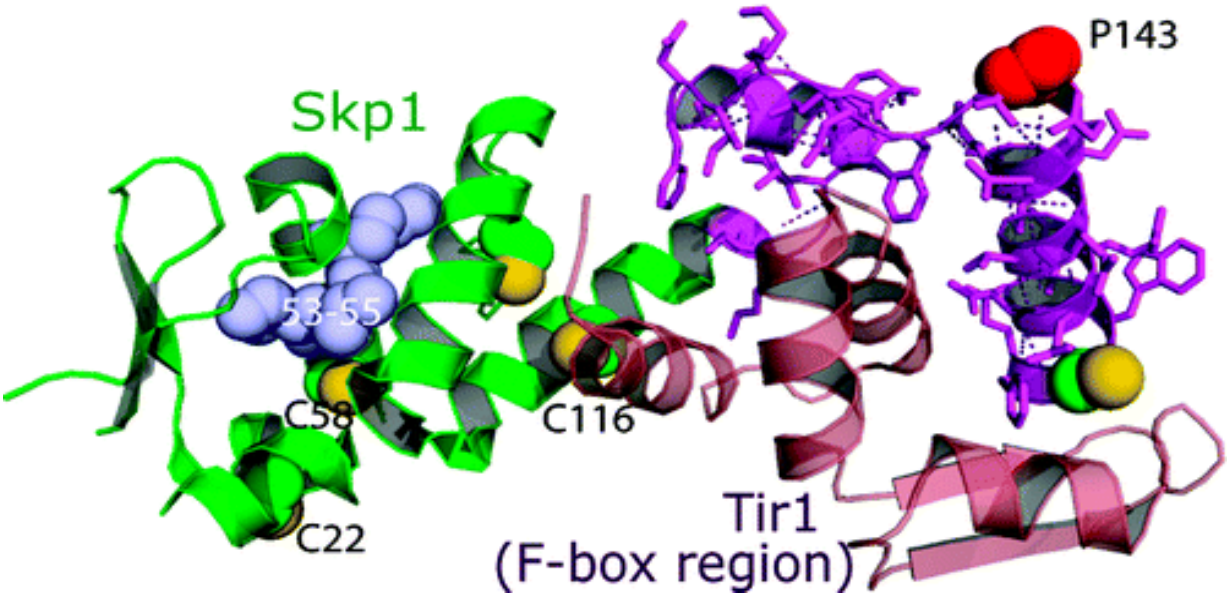


Figure 1.3: Ribbon Diagram of the Skp1-Fbox complex, highlighting Pro[143].[12]

1.3 Post-Translational Modifications of Skp1 in *Dictyostelium* and Their Effect on Cell Function

It has been shown by West *et al*. that the Skp1 subunit of SCF complexes is post-translationally modified in *Dictyostelium discoideum*. According to their studies, the Pro^{143} residue of Skp1 (Figure 1.4) is modified: initially it is hydroxylated, followed by the second modification of glycosylation (Figure 1.4, 1.5). The initial glycosylation is catalyzed by the enzyme Gnt1, adding *N*-acetylglucosamine as the initial sugar. This mono-glycosylated residue is then further glycosylated by other enzymes.

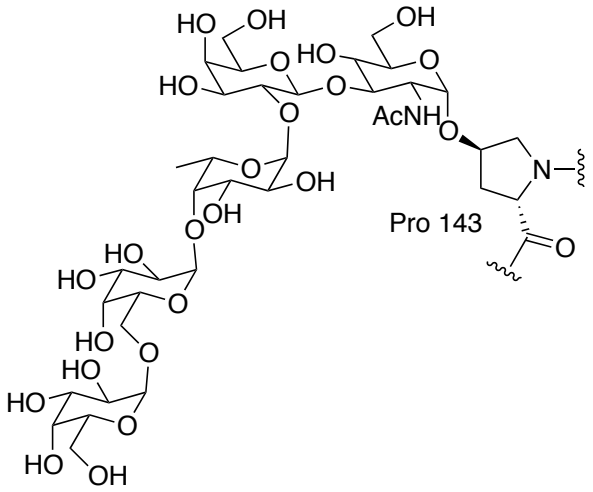


Figure 1.4: Chemical structure of Pro^{143} after post-translational modification.

The Gnt1-catalyzed glycosylation has been determined to occur under standard cellular temperature and pH conditions, using common cellular metabolites (*e.g.,* $MgCl_2$).

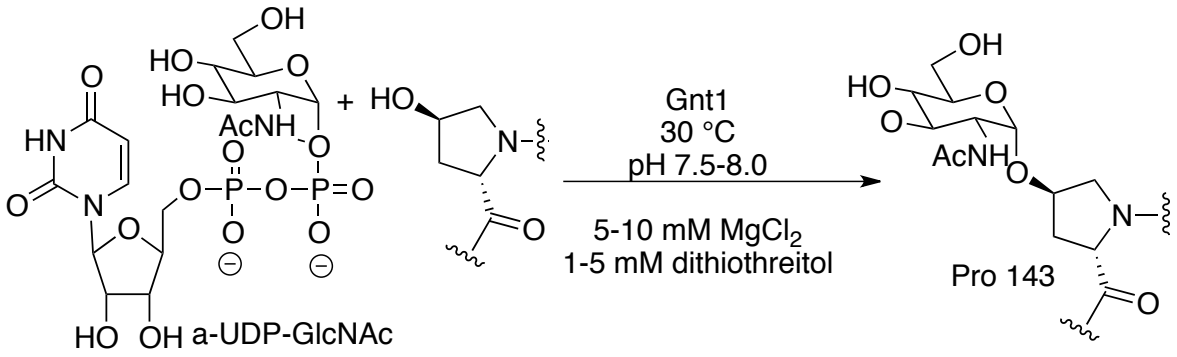


Figure 1.5: Gnt1-catalyzed glycosylation of Pro^{143}.

Gnt1 has also been shown to use activated uridine diphosphate (*i.e.*, sugar activated by covalent bond to UDP), another common metabolite within the cell. Under these conditions, the enzyme is able to transfer the *N*-acetylglucosamine from UDP to the hydroxylated proline residue of Skp1.

These post-translational modifications are essential to the regulatory function of Skp1. We hypothesize that they induce a conformational change in the rest of the protein that allows recognition and binding of the Fbox-target complex, thereby enabling successful ligation of the target molecule to ubiquitin. West *et al*. have shown that—regardless of mechanism—without these post-translational modifications, the organism is no longer able to culminate. This seems to indicate that without post-translational hydroxylation and glycosylation as described above, the ubiquitin pathway fails, leading to inadequate protein metabolism and an inability to reproduce.[11,12]

1.4 Implications of Gnt1-Mediated Glycosylation in Lower Eukaryotes

These enzyme and pathway analyses have led West and coworkers to identify Gnt1 as a key enzyme within the modification pathway of Skp1, and, moreover, of the ubiquitin pathway (due to the importance of the modifications on Skp1 to ubiquitin pathway functionality). Furthermore, sequencing analyses of the *skp1* gene have shown a potentially important trend in the amino acid sequences of various lower and upper eukaryotes (Table 1.1).

Organism	***Comment***	***Skp1 Sequence (133-155)***
Dictyostelium discoideum	Amoebazoan	KIFNIKNDFT**P**EEEEQIRKENEW
Saccharomyces cerevisiae	Yeast	RTFNIVNDFT**P**EEEAAIRRENEW
Cryptosporidium parvum	Diarrheal Pathogen	QIFNIENDFT**P**EEESAIREENKW
Toxoplasma gondii	Toxoplasmosis	RIFNIVNDFT**P**EEEAQVREENKW
Enterocella nidulans	Amoebic Dysentery	RTFNIVNDFT**P**EEEAAIRRENEW
Anopholes gambie	Malaria Vector	KTFNIKNDFT**P**AEEEQVRKENEW
Tetraodon nigroviridis	Puffer Fish	KRFNIKNDFT**E**EEEAQVRKENQW
Homo sapiens	Humans	KTFNIKNDFT**E**EEEAQVRKENQW

Table 1.1: Comparison of various eukaryotes and their respective Skp1 amino acid sequences (amino acid 133-155), with the key proline residue (or lack thereof) in bold

Many lower eukaryotes—most notably, various pathogenic and parasitic eukaryotes (Table 1.1)—have been shown to contain a highly conserved amino acid sequence (DFTPEEE) around the key proline residue (in *Dictyostelium*, residue 143). This is in contrast to the Skp1 amino acid sequence in higher eukaryotes that do not contain a proline residue at or around the 143-position. This leads to the conclusion that higher eukaryotes do not necessitate similar post-translational modifications of Skp1, and thus are not constrained by Gnt1-mediated glycosylation of Pro^{143} (*i.e.,* higher eukaryotes do not have the equivalent of Pro^{143}).

This realization has potentially important therapeutic implications. As noted above, many lower eukaryotes that are dependent upon Gnt1 are potentially pathogenic/parasitic/harmful to humans and other higher eukaryotic hosts. By developing a means of inhibiting the Gnt1-catalyzed glycosylation, it may be possible to inhibit lower

eukaryotic proliferation, without doing harm to the higher eukaryotic host (*e.g.,* humans). It may be possible to fight eukaryotic infections caused by organisms such as *Toxoplasma gondii* (responsible for toxoplasmosis) and *Saccharomyces cerevisiae* (yeast infections), along with a host of others (see Table 1.1 for other examples).[10,11,12]

1.5 Enzyme Kinetics of Gnt1 in *Dictyostelium discoideum*

Based on the determination that Gnt1-mediated glycosylation is essential to E3 functionality, West *et al*. conducted enzyme kinetics in an attempt determine if peptide fragments related to the substrate could be useful as probes to gain some insight into the mechanism of Gnt1. A study was performed using a synthetic 23-mer peptide with a sequence that matched that of Skp1 in *Dictyostelium discoideum* around the 143 residue and Skp1A-Myc, the hydroxylated full-length protein (Table 1.2)[11a,b].

Substrate	K_m **(mM)**	V_{max} **(nmol/h/mg)**
UDP-GlcNAc	**0.16**	**8.0**
Skp1A-Myc	**0.56**	**12.6**
23-mer Peptide	**1600**	**4.2**

Table 1.2: Enzyme kinetics comparison between the two biological substrates for Gnt1 and the synthetic 23mer.

The studies showed that small peptide fragments are not good substrates for Gnt1. While V_{max} values were within an order of magnitude compared to the full-length protein substrate, K_m values were significantly higher for the synthetic 23-mer. This indicates that the enzyme is technically able to bind the 23-mer and process it; however, the K_m difference implies that binding affinity is extremely low for the synthetic peptide.[10]

1.6 The α-Helix of Skp1 and the Hydrogen Bond Surrogate System

We hypothesize that the poor affinity of the synthetic 23-mer for Gnt1 is based on the lack of adequate secondary peptide structure. In general, short peptide sequences such as this are unable to form stable secondary structures (α-helices, β-sheets, turns, loops, etc.) as larger proteins are, and will interconvert between random arrangements.

Due to the secondary instability of short oligopeptides, we propose to induce α-helical conformation using a Hydrogen Bond Surrogate (HBS) system. This concept was advanced and developed by Arora *et al.*, and has been shown to effectively cause α-helical conformation in other peptides. However, their experiments have thus far been limited

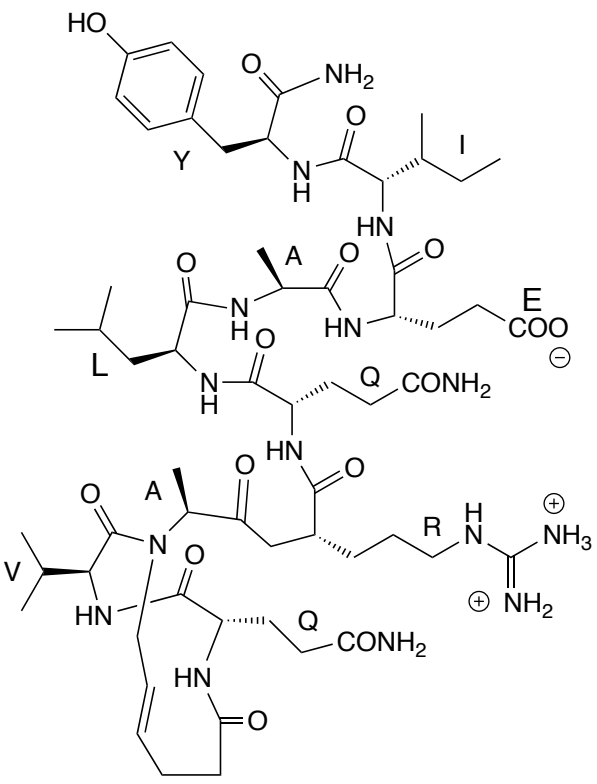


Figure 1.6: Synthetic oligopeptide in alpha-helical conformation due to HBS stabilization.[4]

to synthesizing an HBS system within oligopeptides that terminate with the end of the alpha helix (Figure 1.6). It has not yet been extended to longer peptide chains. We propose to use this system based on the hypothesis that it will stabilize α-helical conformation in small peptides related to Skp1.

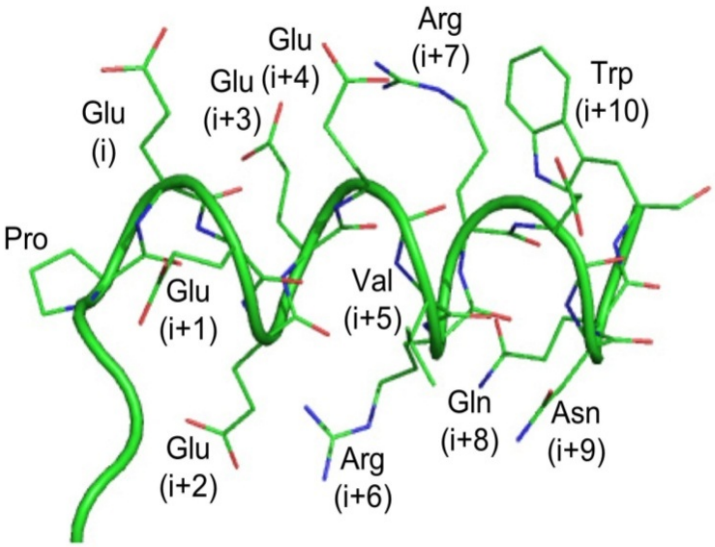
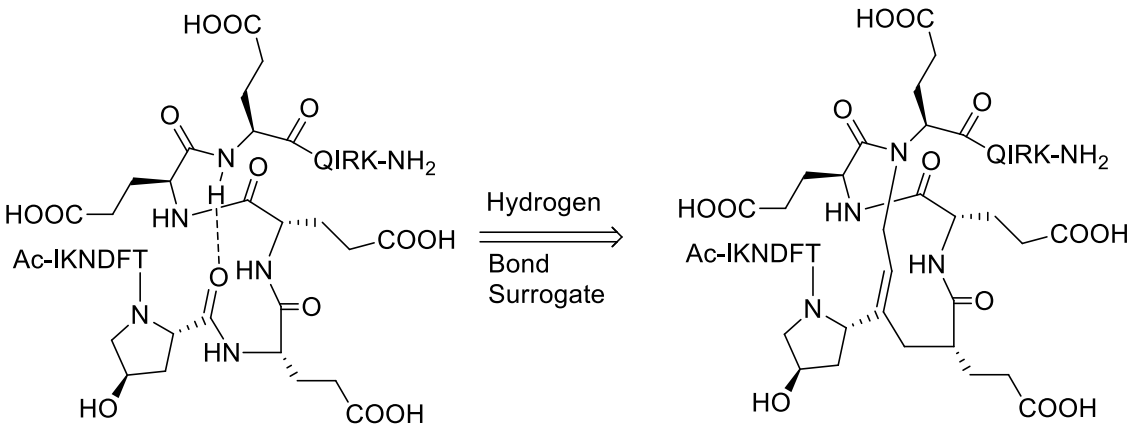


Figure 1.7: Crystal structure of the Skp1 alpha-helix with similar initiating amino acids of *Arabidopsis thaliana.*[11]

As the name suggests, the HBS system will replace the hydrogen bonds that normally stabilize secondary structure in peptides with covalent carbon-carbon bonds. The HBS system will provide the structural stability necessary to induce α-helical conformation by covalently linking the first two turns of the helix together in such a way that they are bound into an α-helix. We hypothesize that forcing the first segment of the sequence into this conformation will induce α-helical conformation in rest of the sequence.[3,4]

In the case of Skp1, it has been shown that the key proline residue appears at the *N*-terminus of an α-helix, and it is hypothesized that it breaks the helix (Figure 1.7). We believe this glutamic acid rich helix to be a crucial structural element for Gnt1 recognition. It is with this idea that this project focuses on the synthesis of an Skp1 mimetic that will share the secondary structural elements characteristic of the full-length protein. The key structural element that we are attempting to emulate is the α-helix comprised of the amino acid sequence PEEEEQIRK.



Scheme 1.2: Comparison of alpha helix stabilized by hydrogen bonds versus the HBS system

1.7 Proposed Synthesis of the Skp1 Mimetic

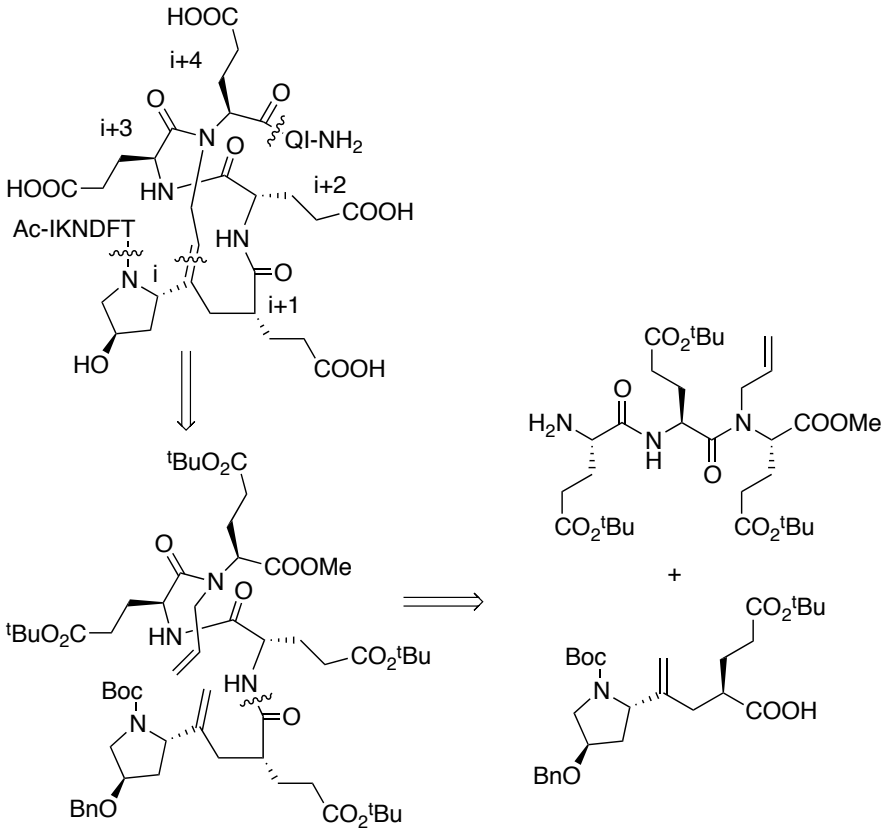
The completed Skp1 mimetic will be comprised of the equivalent of fifteen amino acids, with the HBS-stabilized α-helix mimetic built in. Our strategy involves using a fragment condensation approach to combine the central HBS fragment with *C*- and *N*-terminal peptide fragments prepared separately. As illustrated in the retrosynthetic analysis (Scheme 1.2, 1.3), a disconnect must be made within the HBS carbon chain, as well as on either end of the oligopeptide chain that makes up the initial turn of the helix. From there we are left with a linear pentapeptide that has been modified to be able to form an HBS system using a ring closing metathesis (RCM) reaction.[3,4]

For the purposes of this project, the pentapeptide can be further disconnected to yield a dipeptide (Pro-Glu) isostere and a tripeptide (Glu-Glu-Glu). My primary contribution to the project was synthesis of the *C*-terminal glutamic acid tripeptide.

Another member of the project group, Chamini Karunaratne, is currently working on synthesis of the *N*-terminal dipeptide isostere.

The reason for synthesizing these sections first lies in the HBS system. In order to successfully synthesize the system, very specific alterations to the amino acids must be made. For the HBS system to work, two amino acids, one turn of the helix apart, must be bound covalently as a replacement for the hydrogen bond that would normally link them. In the case of Skp1, those amino acids are the initial proline residue and the fourth glutamic acid residue (residue ***i*** and ***i+4***, **P**EEE**E**, Scheme 1.3).

For each of these two residues, critical changes must be made to the structure of the amino acid. The Pro-Glu dipeptide must be modified into an isostere where the peptide



Scheme 1.3: Retrosynthetic analysis of formation of a Hydrogen Bond Surrogate system using a ring closing metathesis reaction to close the system.

bond [–C(=O)NH–] is replaced with an all carbon group [–C(=CH_2)CH_2–]. This allows replacement of the carbonyl oxygen with an allyl group, and stabilizes the surrogate system that will later be created. The glutamic acid residue also requires modification with an additional *N*-allyl group.

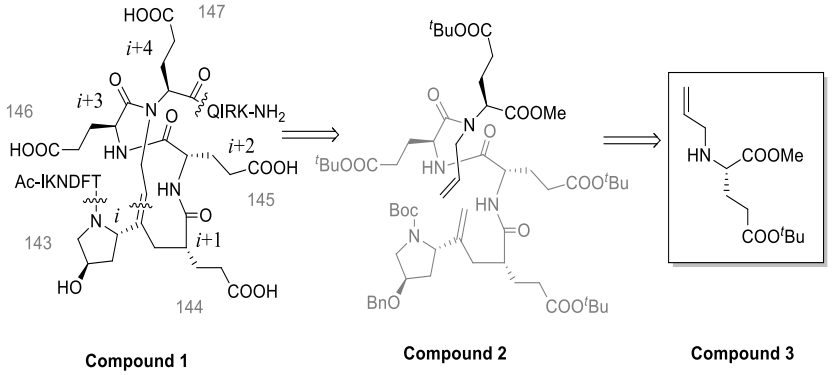
Once each of these two modified amino acid sequences has been made, they will be condensed to yield the modified linear pentapeptide, the precursor to the HBS system. In order to complete the system, the two allyl groups will be reacted via RCM reaction. This will close the ring and force the five amino acids to conform into the first turn of the α-helix (Scheme 1.3). Finally, the Gln-Ile-Arg-Lys peptide fragment will be added onto the *C*-terminus of the sequence, and we hypothesize that it will continue to conform to the secondary structure precedented by the HBS segment.

The focus of this individual undergraduate project was to synthesize the modified Glu-Glu-Glu tripeptide and the Gln-Ile-Arg-Lys tetrapeptide. The project was broken down into three primary phases: (a) synthesis of an *N*-allylated glutamic acid, (b) incorporation and synthesis of the *N*-allylated residue into the glutamic acid tripeptide, and (c) synthesis of the *C*-terminal tetrapeptide.

2. Discussion & Results

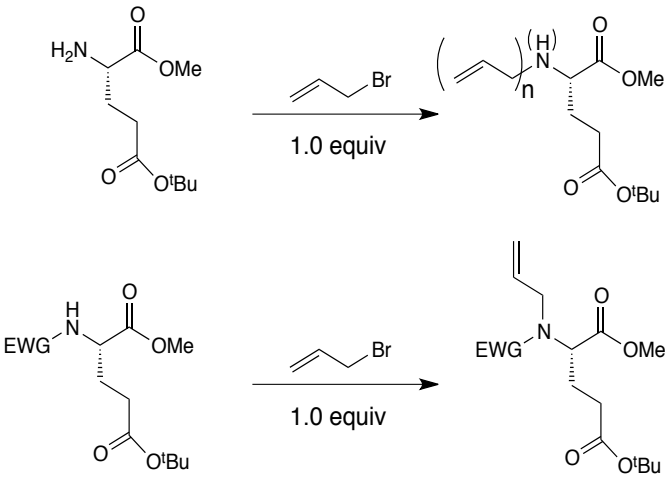
2.1 Synthesis of the *N*-Allyl Glutamic Acid Residue

Using retrosynthetic analysis (Scheme 2.1), we determined that the best course of action would be to start with synthesis of the *N*-allylated glutamic acid residue. From there we would be able to condense it with two other glutamic acid residues to attain the tripeptide.



Scheme 2.1: Retrosynthetic analysis highlighting the *N*-allyl glutamic acid residue

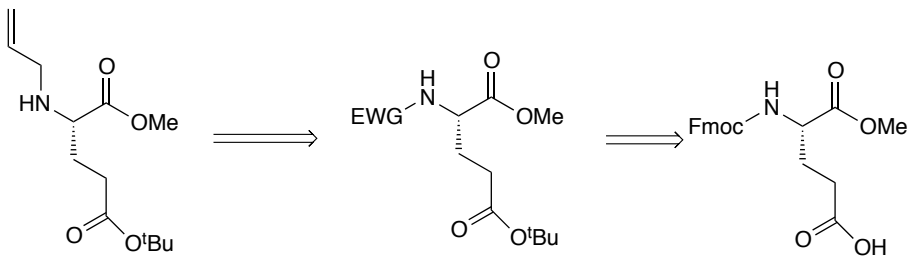
In order to allylate the amino group of the residue, we needed to protect it first. It is important to note that if allylation was attempted on a free amine (even using a 1:1 ratio), the result would be a random mixture of mono-, di-, and tri- allylated products, which would significantly reduce the usefulness of the product. In order to circumvent this issue, we proposed protecting the amine with an electron-withdrawing group (Scheme 2.2). This would reduce the nucleophilicity of the amine and permit high-yielding mono-allylation. By retrosynthetic analysis (Scheme 2.3), we



Scheme 2.2: Comparison of poly-allylation vs mono-allylation using an electron withdrawing group (EWG)

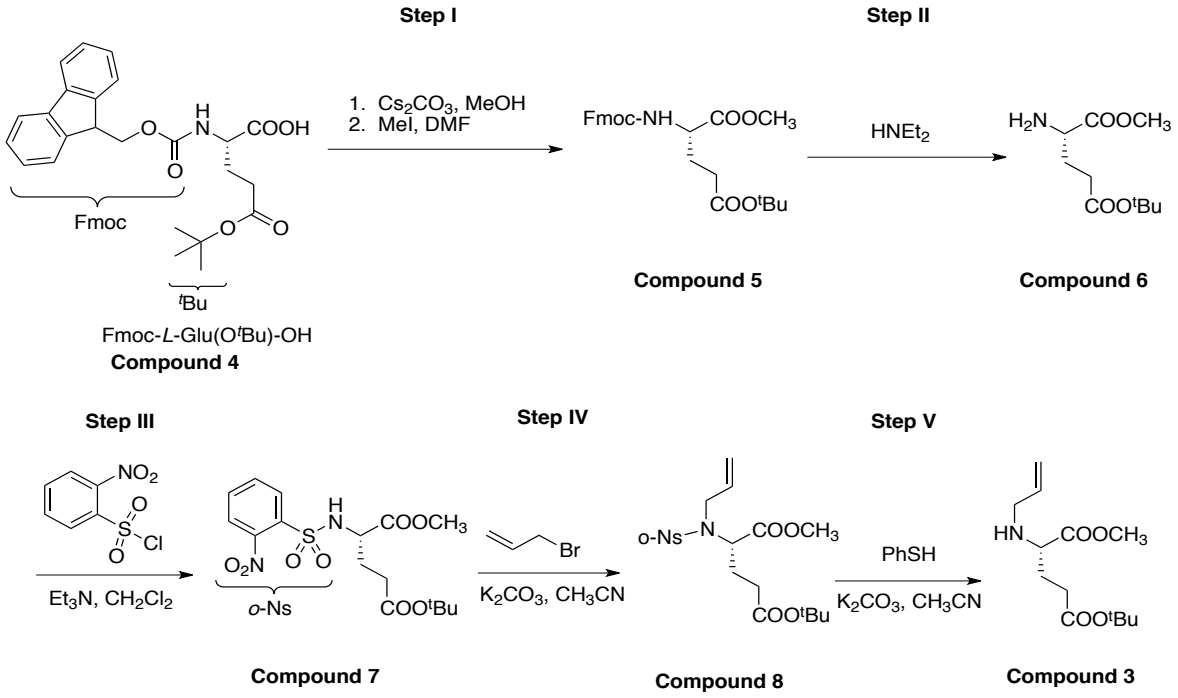
are able to go from the allylated residue, back to the non-allylated residue with an electron withdrawing group, and finally back to Fmoc-*L*-Glu(O^{t}Bu)-OH (commercially available at \$10.73/g).

Starting with Fmoc-*L*-Glu(O^{t}Bu)-OH, we proposed to first create the methyl ester via the intermediary of a cesium salt, thereby differentially protecting both carboxylate groups – one as a methyl ester, one as a *tert*-butyl ester. We would then be able to remove the fluorenylmethoxycarbonyl[5] (Fmoc) group and subsequently add an *ortho*-nitrophenyl sulfonate (nosyl) protecting group in its place.[8]



Compound 3 **Compound 4**

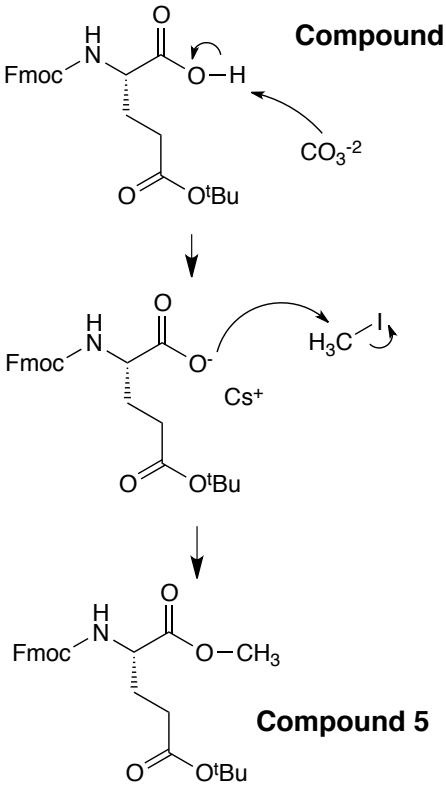
Scheme 2.3: Retrosynthetic analysis for the *N*-allylated glutamic acid residue



Scheme 2.4: Overall synthetic approach for the *N*-allylated glutamic acid residue

This protecting group is essential to the synthesis because attempting to alkylate the free amine would lead to a mixture of mono-, di-, and tri-alkylated products. The *N*-nosyl group would allow successful controlled alkylation of the amine, giving the mono-*N*-allylated residue as desired.[7] The final step was removal of the nosyl group in order to get our target molecule that would fit into the overall oligopeptide mimetic (Scheme 2.4).[9, 10]

2.1.1 Esterification of the *C*-terminus



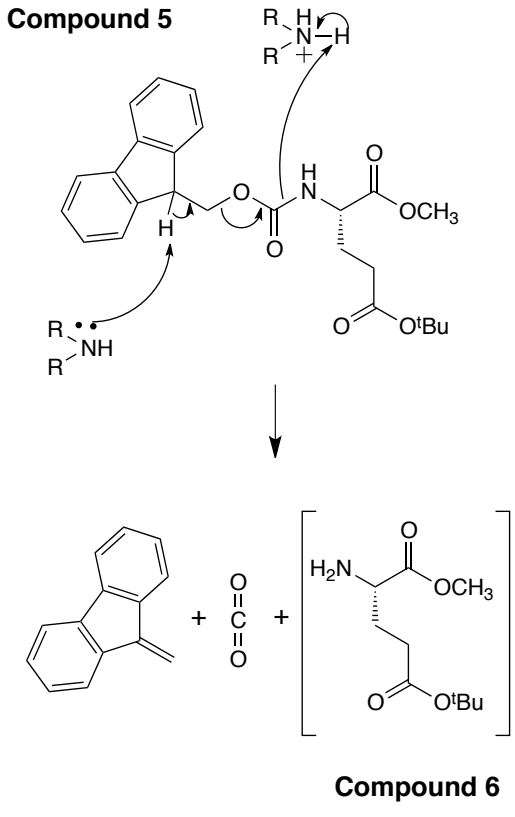
Scheme 2.5: Arrow pushing mechanism for creation of the methyl ester

This fundamental substitution reaction depends upon generation of a good nucleophile. We used cesium carbonate to deprotonate the carboxylic acid, and create a cesium salt in solution that would enhance the nucleophilicity of the carboxyl group. The negatively charged carboxylate group then became a good enough nucleophile to attack the methyl group of methyl iodide (Figure 2.3).

This compound was obtained fairly easily via aqueous work-up and purification by flash chromatography. The ^{1}H NMR showed a strong singlet at ∂ 3.76 ppm, which is diagnostic of the addition of the methyl ester (Figure A1).

This experiment was performed several times, each iteration having similar results. The products of the reaction were very clean each time with yields consistently between 80 and 90%. The significant difference in polarity between the starting material and product allowed for easy separation via chromatography, resulting in a very effective methylation reaction, which did not need to be further optimized. The reaction was performed multiple times in order to gain an appropriate amount of product for subsequent steps in the overall synthesis.

2.1.2 Exchange of the *N*-terminal Protecting Group



Scheme 2.6: Removal of the Fmoc group via E2 elimination reaction by diethylamine.

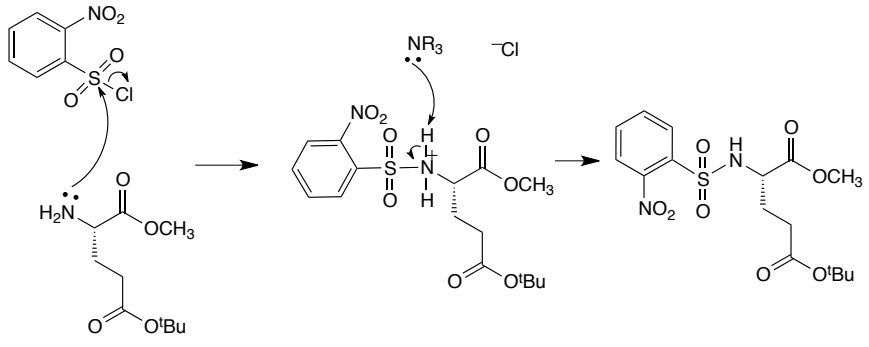
After methyl esterification, the next step (step II) was removal of the Fmoc protecting group as the beginning of a series of steps to mono-allylate the *N*-terminus. It is important to note that because the primary amine is very difficult to isolate, once we had a crude product, we immediately moved on to the next step of nosylation (step III). These two steps combined were much more troublesome than the first, so it took several attempts to isolate a good yield of *N*-nosylated product with acceptable purity from this series of reactions.

In Step II, the base cleaves the Fmoc group via an E2 reaction, affording tricyclic dibenzofulvene and carbon dioxide as by-products to the newly formed primary amine (Scheme 2.6). The amine acts as a weak base (due to a free electron pair) and abstracts the doubly benzylic proton of the Fmoc group with concomitant cleavage of the carbon-nitrogen bond, leaving the primary amine.

Step III of the synthesis was to add the nosyl protecting group to the *N*-terminus of Compound 6. As noted above, this was the second step in a two-part series (steps II and III) to exchange the protecting group from Fmoc to nosyl. We achieved this reaction by redissolving the crude product in dry dichloromethane and adding the appropriate amounts of triethylamine and *o*-nitrobenzenesulfonyl chloride.

This is a substitution reaction wherein the primary amine acts as a nucleophile and attacks the sulfur atom of *o*-nitrobenzenesulfonyl chloride. The chloride then departs, allowing successful addition of the protecting group (Scheme 2.5).

Originally attempted in the Fall 2010 semester, our early approaches to this series of reactions did not result in isolation of the desired product. A few compounds initially isolated by flash chromatography were shown, by NMR analysis, not to be Compound **7**. We eventually discovered that our desired product was eluting much earlier than anticipated, so we



Compound 6 **Compound 7**

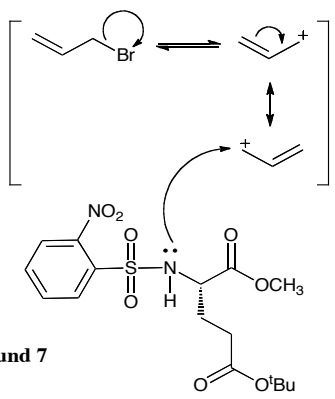
Scheme 2.7: Arrow-pushing mechanism for the addition of the *N*-nosyl group to the *N*-terminus of the amino acid residue

adjusted our mobile phase during chromatography to better isolate and purify the product. While our yield was unimpressive (48%), the product recovered was pure, as shown by the ^{1}H NMR. New peaks in the aromatic region also confirmed presence of the nosyl group in the produce (Figure A2). Mass spectrometry further confirmed the identity of the product.

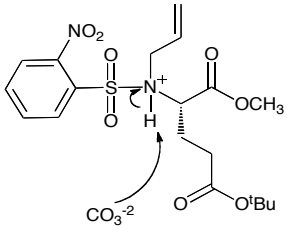
In the Spring 2011 semester, we tested some different methods for this series in order to improve the overall yield. Our original conditions consisted of *tris*-aminoethylamine (TAEA) in dry acetonitrile.[5] After repeated experiments with low yields, we changed the base to diethylamine, and isolated the amine by flash chromatography. We immediately saw significant increases in yield and purity of product. All iterations of this reaction since changing the procedure have had over 80% yields, with the most recent being around 96%.

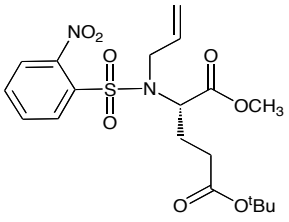
2.1.3 Addition of the Allyl Group

After successful nosylation of the residue, Step IV was to add the allyl group to the *N*-terminus of the residue. This reaction is conceptually related to the nosylation reaction (*i.e.*, it is a substitution reaction), with the notable difference that it is an S_N1 reaction. For this reaction mechanism, the carbon-bromine bond ionizes, leaving an allyl cation (stabilized by resonance). The lone pair of the sulfonamide nitrogen then attacks the stabilized carbocation, resulting in the *N*-allylated residue (Compound **8**, Scheme 2.8). With the nosyl group protecting and deactivating the amine, we prevent poly-allylation of the amine,[7] which would result in a product mixture of low utility.



Compound 7





Compound 8

Scheme 2.8: Allylation of the nosyl-protected amine to create a mono-*N*-allylated amino acid residue

For this reaction, we again began investigations on a small scale and were able to isolate some of our desired product on the first try. Initially, we worked up the reaction by washing the resulting mixture with water and concentrating the organic layer. While the product was impure (we believe there was residual dimethylformamide remaining even after purification and concentration), the ^{1}H NMR spectrum provided evidence for our desired molecule. There were several new peaks between 3 and 5 ppm that indicated presence of the new allyl group (Figure A3). However, we were only able to recover 35% yield for the reaction during the Fall 2010 semester.

There were two major changes made to this procedure during the Spring 2011 semester. Initially, to improve the yield, we altered the method of purification because we believed we were losing product to the aqueous layer during the work up. We circumvented exposure to water by evaporating the reaction mixture, and applying the residue directly to column chromatography.

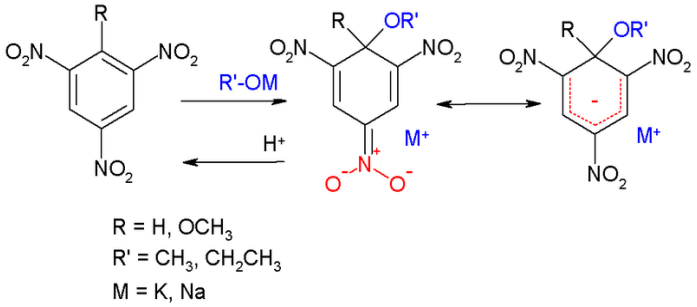
By using this method, we were able to repeatedly achieve greatly improved yields (95%) of very pure product (Compound **8**). When we scaled up this reaction during later iterations, we also changed the solvent from DMF to acetonitrile due to its similar characteristics (polar, aprotic). The higher volatility of acetonitrile also greatly facilitated

purification and concentration of the product (often very difficult with DMF) allowing us to easily obtain a very pure sample with yields ranging from 90-95%.

2.1.4 Removal of the Nosylate to Reveal the Target Residue

The final step in the synthesis of the *N*-allyl glutamic acid was to remove the nosyl protecting group from the *N*-terminus of Compound **8**, revealing the *N*-allylated amino acid residue (Compound **3**). We accomplished this by adding potassium carbonate—a weak base—and thiophenol. The mechanism for this reaction is described in more detail by Fukuyama *et al*.[7]

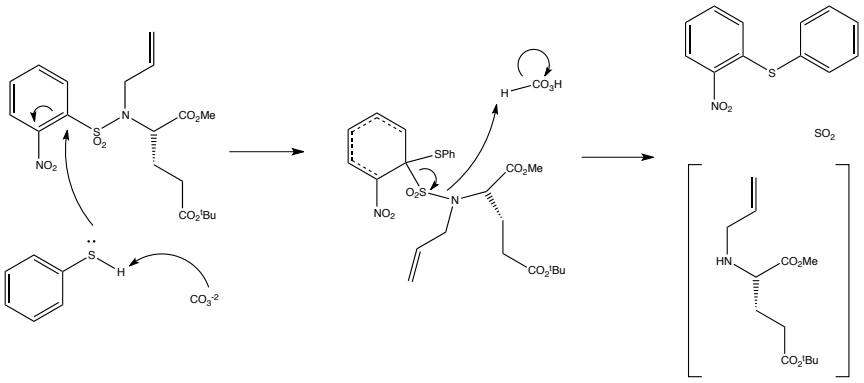
The original report of this reaction invoked formation of a Meisenheimer complex by reduction of a tri-nitro benzene derivative and an alkoxide anion (Scheme 2.9). While the complex in Scheme 2.9 is normally stabilized by three electron-withdrawing nitro groups at the *ortho* and *para* positions, the corresponding intermediate in our reaction is stabilized by only one nitro group at the *ortho* position. Our reaction employed a thiolate anion instead of an alkoxide anion. We hypothesize that due to the atypical size of the S-R groups used in our peptide, the addition of other nitro groups could sterically hinder the formation of the complex (Scheme 2.10).[13]



Scheme 2.9: Typical Meisenheimer complex[13]

We began investigations again on small scale (starting with 43 mg

Compound **8**). We began by following the procedure from Fukuyama *et al.*[6] Using DMF as the solvent and following their recommended reaction times, reagent ratios, and purification procedures, and were able to obtain a maximum yield of 41% very pure product. Due to the difficulty of removing DMF from the reaction mixture of larger-scale reactions, we began to investigate alternative procedures for this reaction.[16,17,18]



Scheme 2.10: Addition of thiophenol, forming a Meisenheimer complex, and eventual dissociation of *N*-allylated residue.

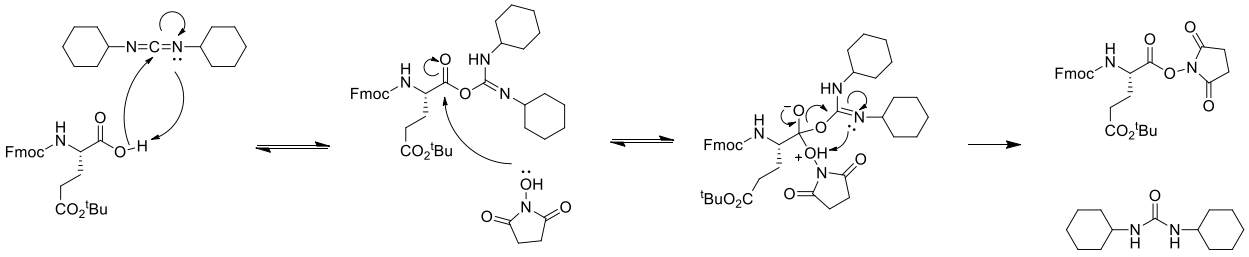
Because of the previous success we found using MeCN over DMF, we attempted to switch the solvent for our experiments. This had the added bonus of eliminating the need for aqueous work-up, which we believed to be a significant source of product loss. By using MeCN, and modifying the purification technique, as well as adding slightly more base to the reaction mixture, we were able to increase the yield to around 86% very consistently.

2.2 Incorporation of *N*-Allyl-Glu into a Glutamic Acid Tripeptide

Having successfully synthesized the *N*-allyl glutamic acid, we were able to progress to synthesis of the tripeptide. This was done using a standard two-part peptide coupling

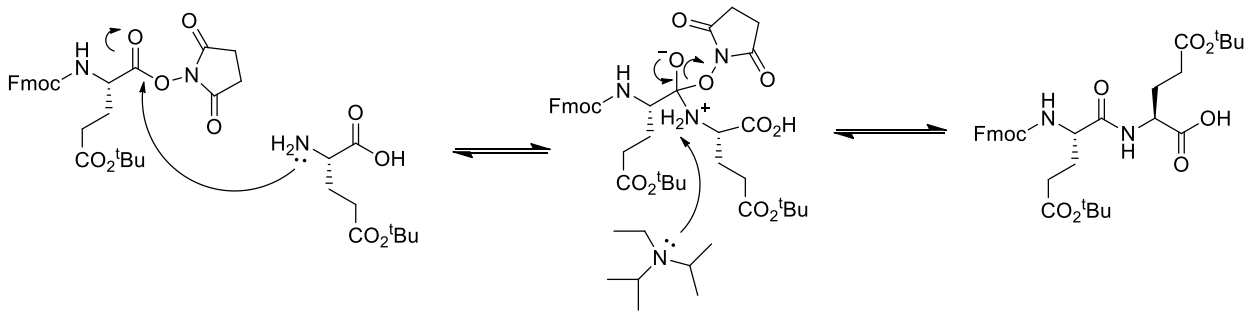
reaction. The first two glutamic acid residues were coupled using NHS-activation, and the third residue was added subsequently using HATU and collidine.[14] Through this reaction series, we were able to successfully couple each of the peptides, and synthesize the modified glutamic acid residue.

The first stage of our peptide coupling reaction involved activation of the carboxylic acid of the first residue by converting it to an *N*-hydroxysuccinimide (NHS) ester. The mechanism shows that the residue first is deprotonated, allowing nucleophilic attack on its carbonyl carbon. DCC and NHS sequentially attack the residue, causing formation of a tetrahedral complex. The equilibrium of this complex will lead to the loss of one of the substituents, and reform an ester with the remaining one. Due to the nature of the equilibrium, this process is repeated, until the NHS ester is formed, and the DCC is consumed via conversion to DCU (forms an insoluble solid). The reaction mechanism for the activation of the initial peptide in the first stage is shown in Scheme 2.11.



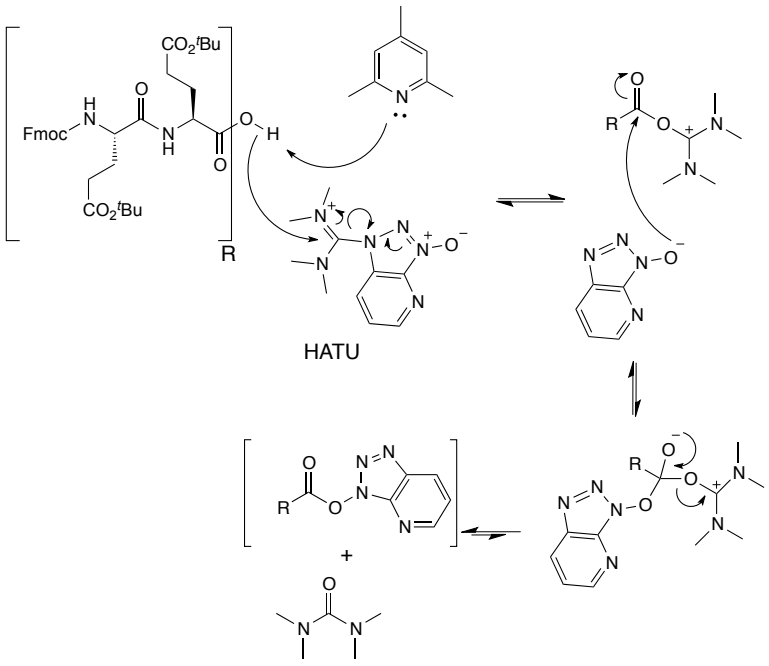
Scheme 2.11: Mechanism for the activation of the first amino acid

Once the NHS ester is formed, the second peptide residue is then added to the mixture, resulting in amide formation. Here, the amino group of the second residue nucleophilically attacks the activated acyl carbon, again causing formation of a tetrahedral complex. Again, the best leaving group departs as the –OAt ester is formed (Scheme 2.12).



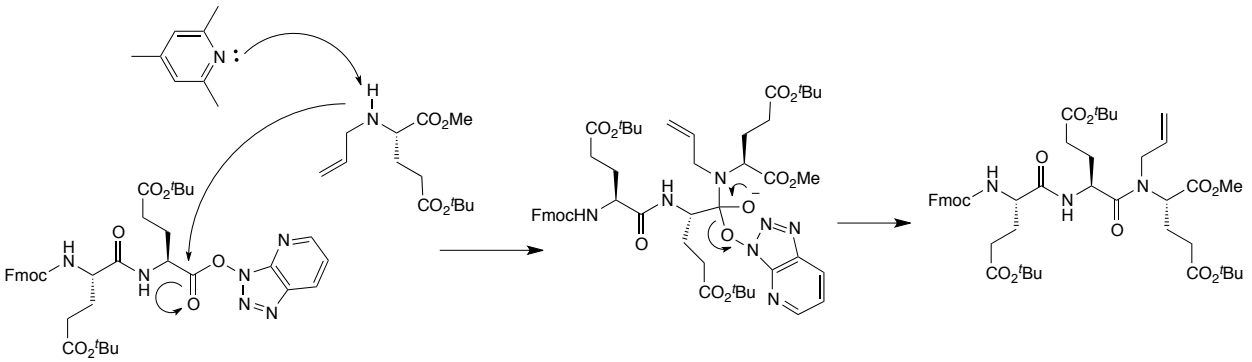
Scheme 2.12: Reaction in which the second peptide residue replaces NHS.

The second stage of this peptide coupling reaction involves combination of the dipeptide acid with the next amino acid. 2-(1H-7-Azabenzotriazol-1-yl)-1,1,3,3-tetramethyluroniumhexafluorophosphate Methanaminium (HATU) and collidine are incorporated into this reaction to allow successful coupling. In general, these are combined and stirred overnight and the peptide product purified via flash chromatography. Overall, the reaction is similar to the mechanism in the dipeptide synthesis (Scheme 2.13, 2.14).



Scheme 2.13: Mechanism for formation of the activated dipeptide ester using HATU and collidine

This series of reactions was performed several times with the dual purpose of optimizing the reaction and synthesizing the product in sufficient amount to be able to carry on subsequent experiments with it. Due to the length of the experiment and its



Scheme 2.14: Mechanism for transesterification to yield the tripeptide

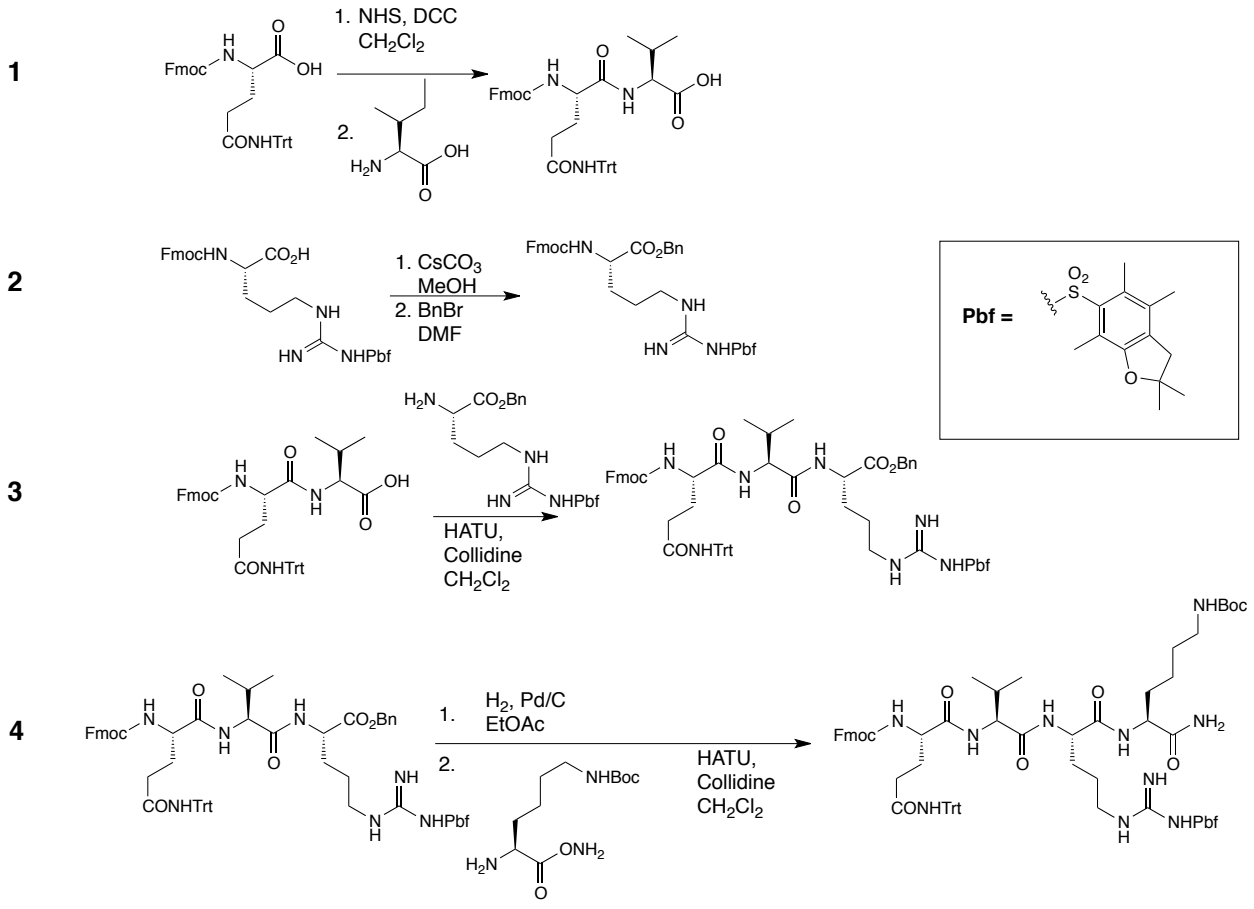
multiple steps, there is a great deal of room for various types of error in this reaction. While it is always important to have good lab technique, there are also many components to the chemistry of this experiment that could potentially reduce yields and purity. For the times we performed this experiment, the yields were generally above 30%, with the highest yield being 40%.

2.3 Synthesis of the QIRK Oligopeptide Sequence

In addition to synthesis of the *N*-allylated glutamic acid tripeptide, a second goal of my project was to synthesize the QIRK peptide sequence that will be directly *C*-terminal to the glutamic acid tripeptide. After successful synthesis of the HBS system, the first five peptides (PEEEE) would be covalently bound into the alpha-helical formation. These next four peptides (QIRK) would form the next turn of the helix. It will be important to have some stock of this tetrapeptide on hand to investigate the best methods of elongation of the

peptide sequence, as well as to test the conformation of the peptide residues past the HBS system.

The synthesis of the QIRK tetrapeptide was very similar to that of the glutamic acid tripeptide. In each case, the first two residues were coupled using the same NHS/DCC coupling reaction. The third amino acid was added using HATU and collidine, and in the case of the tetrapeptide, so was the fourth. Additionally, because of the need to add a fourth peptide, we protected the *C*-terminus of the arginine residue before the second coupling reaction (step 3), and deprotected it before the third coupling reaction (step 4). This also allowed characterization of the QIR tripeptide. For these reactions, the mechanisms are the same as above with synthesis of the glutamic acid tripeptide with the only difference being the amino acids used (Schemes 2.11-2.14, 2.15).



Scheme 2.15: Overall synthesis of QIRK tetrapeptide

The full synthesis of the QIRK tetrapeptide has only been performed once, and produced average yield. The protection of arginine was generally step 1 in the series (due to the relative stability of the product), and was similar in mechanism and procedure to the methylation of Fmoc-*L*-Glu(O^{t}Bu)-OH (Section 2.1.1, Scheme 2.5). The difference was the use of benzyl bromide instead of methyl iodide to esterify the *C*-terminus. Steps 2 and 3 were mediated coupling of glutamine and isoleucine (vis NHS ester), followed by HATU-mediated addition of benzylated arginine, just as with synthesis of the glutamic acid tripeptide. Step 4 was deprotection of the *C*-terminus (removal of benzyl group) via catalytic hydrogenation (for procedure see section 3.7). followed by addition of lysine (with the *C*-terminus already protected as an amide) using HATU and collidine to yield the tetrapeptide (Scheme 2.15).

Experimentally, step 1 in the synthesis was performed without any problems and with yields consistently in around 80%. The step 2-3 series went similarly to the glutamic acid tripeptide synthesis with the highest yield being around 40%. Step 4 went considerably better, giving a pure product of 72% yield. Overall, the tetrapeptide synthesis was successful, with high yields produced for all but step 2-3 series.

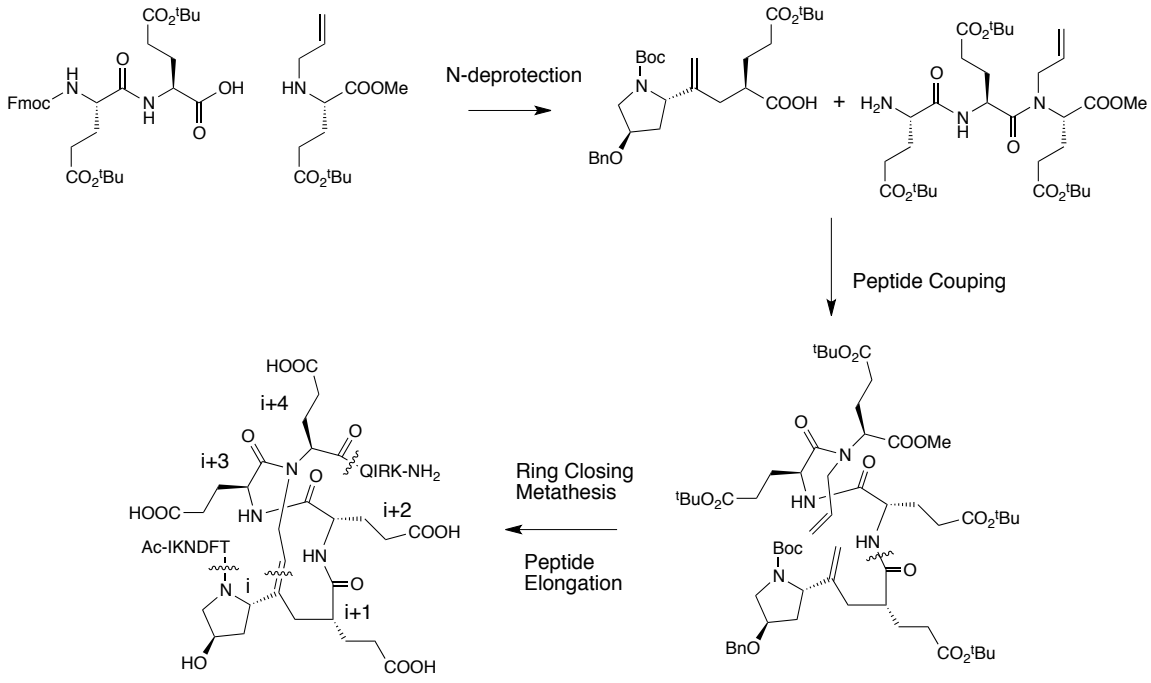
2.4 Future Work: Ring Closing Metathesis and Elongation of the Polypeptide

The next phase of this project will be combination of the glutamic acid tripeptide with the Hyp-Glu isostere that Chamini Karunaratne, is currently working on. These pieces together will form the hydrogen bond surrogate system[2,3] that will enable formation of the

critical alpha helix of Skp1 *in vitro* (Scheme 2.16). We will use the ring closing metathesis (RCM) reaction to create a second covalent bond between the bottom and top fragments of the HBS system. This type of reaction involves connection of the two allyl groups to close the ring. In the case of the HBS system, this additional linkage will be the substitute for the hydrogen bonds that normally stabilize alpha-helical conformation.

Once this peptidomimetic structure is formed, it will be elongated on either end. The QIRK tetrapeptide will be added onto the *C*-terminal end of the peptide, and the *N*-terminus will be elongated in solid-state. Based on Arora *et al.*'s work with the HBS system, the QIRK tetrapeptide will continue to conform to the alpha helix, and allow successful synthesis of an Skp1 mimetic with stable and accurate secondary structure.

Once this mimetic is complete, it will be assessed for its impact on recognition and processing by the enzyme Gnt1. This will give us a much better idea of the importance of



Scheme 2.16: Overall synthesis of the HBS Skp1 mimetic with incoporation of all fragments from this project, as well as the PE fragment from Chamini Karunaratne.

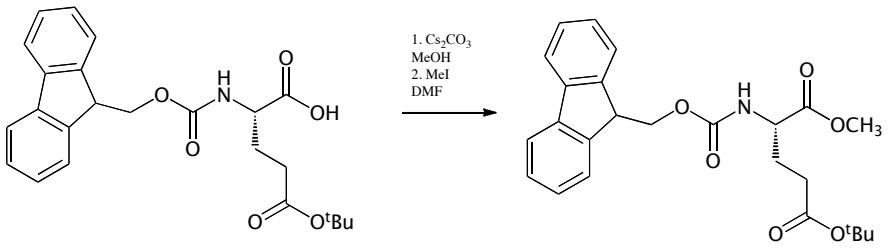
this Glu-rich alpha helix as a recognition element and possibly as an enzyme-binding domain. If it becomes evident that the helix is an important structure in Gnt1 binding, it will be developed as a permanent design feature in future investigative projects and protein inhibitors.

3. Experimental Procedures

3.1 General Procedures

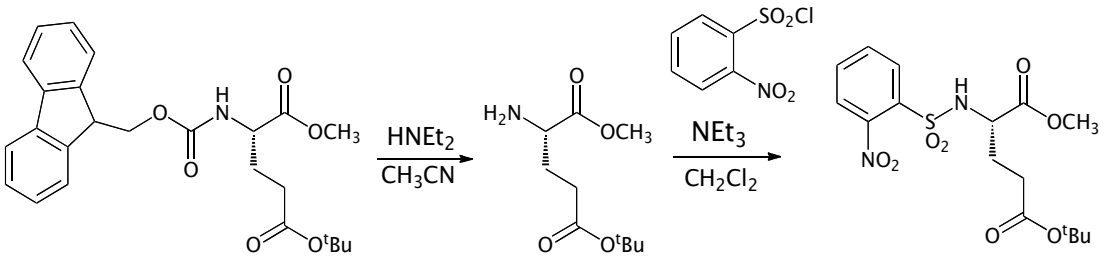
All reactions were performed under a dry nitrogen atmosphere unless otherwise noted. Reagents were obtained from commercial sources and used directly; exceptions are noted. Diethylamine, triethylamine, diisopropylethylamine, and collidine were dried and distilled from CaH_2 and stored over KOH pellets. Ethanol and methanol were distilled from Mg turnings and stored over 4 Å molecular sieves. Flash chromatography was performed using flash silica gel (32-63 μm) from Dynamic Adsorbents Inc. Reactions were followed by TLC on pre-coated silica plates (200 μm, F-254 from Dynamic Adsorbents Inc.). The compounds were visualized by UV fluorescence or by staining with anisaldehyde, ninhydrin, phosphomolybdic acid and potassium permanganate stains. NMR spectra were recorded on a Bruker AV-400-liquid spectrometer. Proton NMR data is reported in ppm downfield from TMS as an internal standard. High resolution mass spectra were recorded using either an Aglient 6210 time-of-flight MS, or a Hitachi MS-8000 3DQ LC-ion trap mass spectrometer with electrospray ionization.

3.2 Fmoc-*L*-Glu(O^{t}Bu)-OCH$_3$.



Fmoc-*L*-Glu(O^{t}Bu)-OH (1.00 g, 2.35 mmol, 1.00 equiv.) was dissolved in methanol (10 mL) and cooled in an ice bath. After the solution was cold, cesium carbonate (383 mg, 1.18 mmol, 0.50 equiv.) was added and stirred for 1 h. The solution was then evaporated and the residue dissolved in DMF (8 mL), combined with methyl iodide (148 μL, 340 mg, 2.35 mmol, 1.00 equiv.), and stirred overnight (~20 h). The mixture was partitioned between ethyl acetate (100 mL) and water (100 mL). The organic layer was washed with brine (100 mL), dried over $MgSO_4$, filtered and concentrated. The residue was purified via column chromatography eluting with 1:1 hexanes:ethyl acetate to give the methyl ester as a foam (820 mg, 1.89 mmol, 80%). R_f 0.65 (1:1 hexanes:ethyl acetate). $[\alpha]_D^{25}$ -0.12 (*c* = 0.5, MeOH). ^{1}H NMR ($CDCl_3$, 400 MHz) *∂* 1.45 (s, 9H), 1.92-2.01 (m, 1H), 2.12-2.20 (m, 1H), 2.25-2.40 (m, 2H), 3.76 (s, 3H), 4.22 (t, *J* = 7.0 Hz, 1H), 4.35-4.44 (m, 3H), 5.48 (d, *J* = 8.0, 1H), 7.31 (t, *J* = 7.4, 2H), 7.40 (t, *J* = 7.4, 2H), 7.58-7.61 (m, 2H), 7.76 (d, *J* = 7.5, 2H); ^{13}C NMR ($CDCl_3$, 100 MHz) *∂* 27.8, 28.3, 31.7, 47.4, 52.8, 53.7, 67.3, 81.0, 120.2, 125.3, 127.3, 127.9, 141.5, 143.9, 144.0, 156.2, 172.3, 172.8. HRMS (+ESI) calculated for $C_{25}H_{29}NO_6$ $(M+Na)^+$ 462.1887, observed 462.1900.

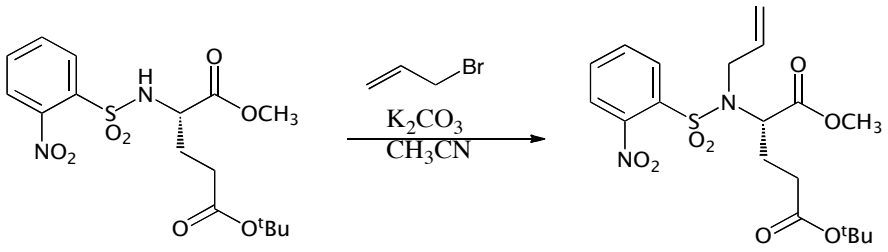
3.3 *N*-Nosyl-*L*-Glu(O^{t}Bu)-OCH$_3$.



Fmoc-*L*-Glu(O^{t}Bu)-OCH$_3$ (1.009 g, 2.290 mmol, 1 equiv.) was dissolved in dry acetonitrile (6 mL) and combined with diethylamine (3 mL, 2.12 g, 28.9 mmol, 13 equiv.) and stirred for 1.5 h. The product (H-*L*-Glu(O^{t}Bu)-OCH$_3$) is not easily purified, so we immediately moved on to the next stage of nosylation.

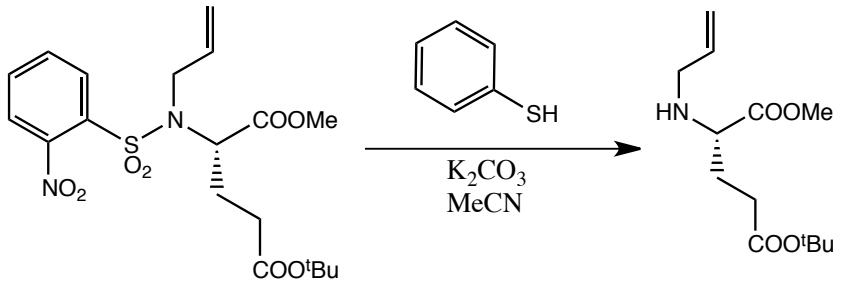
The crude H-*L*-Glu(O^{t}Bu)-OCH$_3$ (max: 498 mg, 2.29 mmol, 1 equiv.) was evaporated, dried, and dissolved in dry CH_2Cl_2 (6 mL), and cooled in an ice bath. *o*-nitrobenzenesulfonyl chloride (761 mg, 3.44 mmol, 1.5 equiv.) and triethylamine (3 mL, 2.184 g, 21.6 mmol, 9 equiv.) were added to the mixture, and stirred overnight (~24 h). The solution (dark brown) was evaporated and the residue purified via column chromatography, eluting progressively with 3:1, 2:1, and 1:1 hexanes:ethyl acetate to yield a clear, light yellow oil (918 mg, 2.28 mmol, 99% over 2 steps). R_f 0.47 (2:1 hexanes:ethyl acetate). $[\alpha]_D^{25}$ -0.22 (*c* 0.5, MeOH). ^{1}H NMR ($CDCl_3$, 400 MHz) ∂ 1.46 (s, 9H), 1.88-1.98 (m, 1H), 2.15-2.23 (m, 1H), 2.37-2.49 (m, 2H), 3.48 (s, 3H), 4.25 (dd, *J* = 9.0 Hz, 4.8 Hz, 1H), 6.21 (s, 1H), 7.70-7.77 (m, 2 H), 7.91-7.95 (m, 1H), 8.03-8.07 (m, 1H); ^{13}C NMR ($CDCl_3$, 100 MHz) ∂ 28.3, 31.0, 52.7, 56.2, 81.1, 125.8, 130.6, 133.1, 133.9, 134.0, 147.8, 171.4, 171.9. HRMS (+ESI) calculated for $C_{16}H_{22}N_2O_8S$ $(M+Na)^+$ 425.0989, observed 425.1000.

3.4 *N*-Allyl-*N*-Nosyl-*L*-Glu(O^{t}Bu)-OCH$_3$.



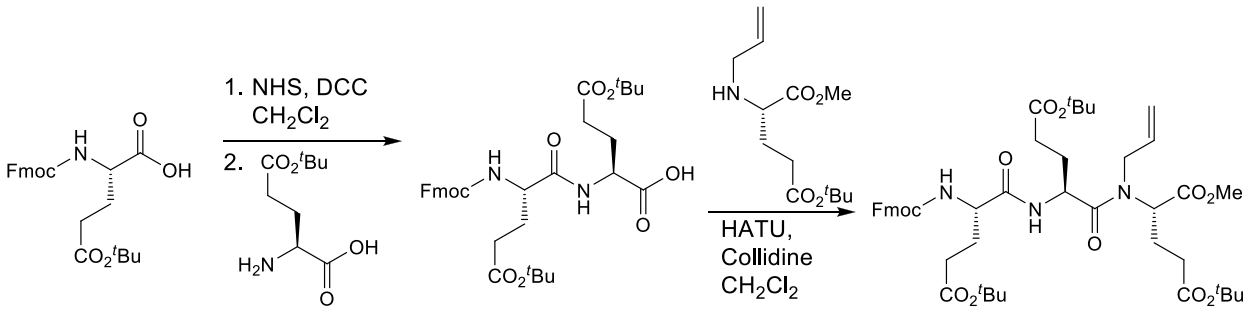
N-Nosyl-*L*-Glu(O^{t}Bu)-OCH$_3$ (139 mg, 0.345 mmol, 1 equiv.) was dissolved in DMF (2 mL), and cooled in an ice bath. Potassium carbonate (98.1 mg, 0.710 mmol, 2.02 equiv.), and allyl bromide (55.3 µL, 76.9 mg, 0.636 mmol, 1.84 equiv.) were added and the mixture was stirred overnight (~25 h). The mixture was evaporated and the product isolated by flash chromatography, eluting with 2:1 hexanes:ethyl acetate, and evaporated the appropriate fractions to give the product as a yellow oil (22 mg, 0.05 mmol, 14 %), R_f 0.46 (2:1 hexanes:ethyl acetate). $[\alpha]_D^{25}$ -0.05 (*c* 0.5, MeOH). ^{1}H NMR ($CDCl_3$, 400 MHz) ∂ 1.46 (s, 9H), 1.94-2.03 (m, 1H), 2.28-2.37 (m, 1H), 2.40-2.47 (m, 2H), 3.57 (s, 3H), 3.85 (dd, *J* = 16.5, 7.6 Hz, 1H), 4.13 (dd, *J* = 16.5, 5.5 Hz, 1H), 4.72 (dd, *J* = 10.3, 4.8 Hz, 1H), 5.13 (d, *J* = 10.2 Hz, 1H), 5.23 (d, *J* = 17.2 Hz, 1H), 5.88-5.97 (m, 1H), 7.59-7.63 (m, 1H), 7.66-7.72 (m, 2H), 8.05-8.07 (m, 1H); ^{13}C NMR ($CDCl_3$, 100 MHz) ∂ 24.9, 28.1, 31.5, 49.1, 52.3, 59.8, 80.8, 118.6, 124.0, 128.8, 131.3, 131.5, 133.2, 133.5, 134.7, 138.1, 148.0, 170.9, 171.7. HRMS (+ESI) calculated for $C_{19}H_{26}N_2O_8S$ $(M+Na)^+$ 465.1302, observed 465.1314.

3.5 *N*-Allyl-L-Glu(O^{t}Bu)-OCH$_3$.



N-Allyl-*N*-nosyl-L-Glu(O^{t}Bu)-OCH$_3$ (119 mg, 0.26 mmol, 1 equiv.) was dissolved in MeCN (2.5 mL) and combined with K_2CO_3 (73 mg, 0.528 mmol, 2.0 equiv.). Thiophenol (54.0 µL, 58 mg, 0.532 mmol, 1.98 equiv.) was then added, and the solution stirred in an ice bath and under nitrogen overnight (~22 h). The resulting solution was evaporated, and the product isolated via flash chromatography, eluting with 5:1 hexanes:ethyl acetate to yield the product (59 mg, 0.23 mmol, 86%). R_f 0.33 (2:1 hexanes:ethyl acetate). ^{1}H NMR ($CDCl_3$, 400 MHz) ∂ 1.44 (s, 9H), 1.80-1.87 (m, 1H), 1.90-1.97 (m, 1H), 2.35 (t, J = 7.5 Hz, 2H), 3.09 (dd, J = 13.9, 6.9 Hz, 1H), 3.28 (dd, J = 13.8, 6.9 Hz, 2H), 3.73 (s, 1H), 5.08 (dd, J = 10.2, 1.48 Hz, 1H), 5.17 (dq, J = 17.2, 1.6 Hz, 1H) 5.78-5.86 (m, 1H); ^{13}C NMR ($CDCl_3$, 100 MHz) ∂ 28.3, 28.8, 32.1, 50.9, 52.0, 59.9, 80.6, 116.6, 136.6, 172.6, 175.8. HRMS (+ESI) calculated for $C_{13}H_{24}NO_4$ $(M+Na)^+$ 257.1627, observed 257.1616.

3.6 Fmoc-Glu(O^{t}Bu)-Glu(O^{t}Bu)-*N*-Allyl-Glu(O^{t}Bu)-OMe.

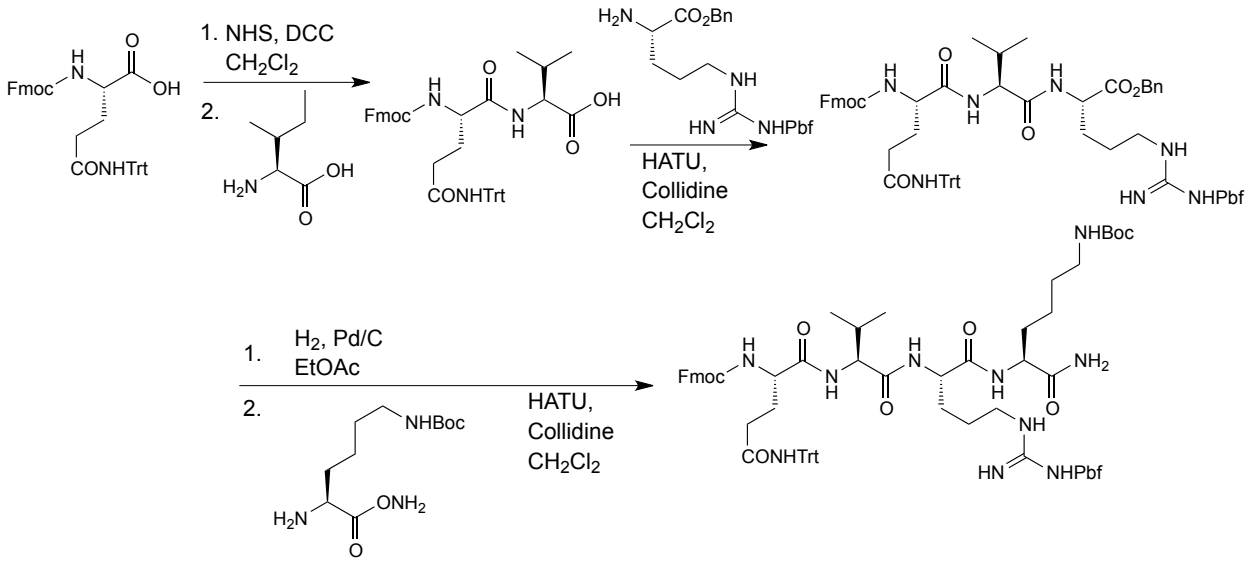


Fmoc-Glu(O^{t}Bu)-Glu(O^{t}Bu)-OH. Fmoc-Glu(O^{t}Bu)-OH (332 mg, 0.78 mmol, 1.0 equiv) was dissolved in dry CH_2Cl_2 (3 mL) in an ice bath. *N*-hydroxysuccinimide (NHS) (90 mg, 0.78 mmol, 1.0 equiv.) and DCC (162 mg, 0.78 mmol, 1.0 equiv.) were added and the reaction mixture stirred 4 h. The mixture was filtered to remove solid DCU and the filtrate evaporated down to approximately 1 mL and placed in the freezer for 2 h. The mixture was then refiltered, evaporated and redissolved in DMF (3 mL) in an ice bath. H-Glu(O^{t}Bu)-OH (158 mg, 0.78 mmol, 1.0 equiv.) was added to the mixture followed by addition of diisopropylethylamine (136 µL, 101 mg, 0.78 mmol, 1.0 equiv.) dropwise. The reaction mixture was stirred overnight (approx. 18 h). The mixture was diluted with ethyl acetate (50 mL) and washed with 2 M HCl (40 mL). The aqueous layer was back-extracted with ethyl acetate (50 mL) and each organic layer washed separately with water (40 mL each), filtered over $MgSO_4$, evaporated, and dried to yield a colorless foam. The dipeptide was used in the next step.

Fmoc-Glu(O^{t}Bu)-Glu(O^{t}Bu)-*N*-Allyl-Glu(O^{t}Bu)-OMe. Fmoc-Glu(O^{t}Bu)-Glu(O^{t}Bu)-OH (max. 490 mg, 0.78 mmol, 1.05 equiv.) and *N*-Allyl-Glu(O^{t}Bu)-OMe (190 mg, 0.74 mmol, 1.0 equiv.) were combined in dry CH_2Cl_2 (4 mL) in an ice bath, under a nitrogen

atmosphere. HATU (297 mg, 0.78 mmol, 1.05 equiv.) and collidine (195 μL, 179 mg, 1.48 mmol, 2.0 equiv.) were added sequentially and the mixture stirred overnight (approx. 17 h). The mixture was then concentrated and the residue purified via flash chromatography eluting with 2:1 hexanes-ethyl acetate, evaporated, and dried to yield a colorless foam (246 mg, 0.29 mmol, 40 %). R_f 0.19 (2:1 hexanes-ethyl acetate). ^{1}H NMR ($CDCl_3$, 400 MHz) ∂ 1.00 (t, J = 7.3 Hz, 2H), 1.05-1.16 (m, 1H), 1.24-1.30 (m, 1H), 1.42-1.45 (m, 27H), 1.59-1.77 (m, 3 H), 1.92-2.12 (m, 5H), 2.29-2.39 (m, 7H), 3.31 (app. t, J = 8.1 Hz, 1H), 3.48 (s, 1H), 3.66 (d, J = 6.0 Hz, 3H), 4.22-4.30 (m, 3H), 4.37 (d, J = 8.1 Hz, 2H), 4.58 (t, J = 6.6 Hz, 1H), 4.91 (s, 1H), 5.26 (d, J = 10.1 Hz, 1H), 5.34 (dd, J = 17.0, 8.5 Hz, 1H), 5.73 (d, J = 7.1 Hz, 1H), 5.82-5.93 (m, 1H), 6.98 (d, J = 8.3 Hz, 1H), 7.27 (d, J = 1.7 Hz, 2H), 7.32 (t, J = 7.4 Hz, 2H), 7.40 (t, J = 7.4 Hz, 2H), 7.60 (s, 2H), 7.76 (d, J = 7.4 Hz, 2H); ^{13}C NMR ($CDCl_3$, 100 MHz) ∂ 20.7, 24.7, 25.6, 26.9, 27.9, 27.9, 28.2, 31.4, 31.5, 33.4, 36.6, 47.0, 59.2, 51.8, 54.0, 67.0, 80.7, 80.8, 107.7, 119.6, 119.8, 120.9, 125.1, 126.9, 127.6, 128.6, 141.1, 143.6, 143.8, 156.1, 163.1, 171.6, 172.2, 172.8, 173.7, 174.9. HRMS (MALDI) calculated for $C_{46}H_{63}N_3O_{12}$ $(M+Na)^+$ 872.4309, observed 872.448.

3.7 Fmoc-Gln-Ile-Arg-Lys-OBz.



Fmoc-Arg(NHPbf)-OBn. Fmoc-Arg(NHPbf)-OH (200 mg, 0.31 mmol, 1.0 equiv.) was dissolved in dry MeOH (2 mL) and cooled in an ice bath under nitrogen. Cesium carbonate (50 mg, 0.15 mmol, 0.5 equiv.) was added and the reaction mixture stirred 1 h to form the salt. The solvent was evaporated and the residue redissolved in dry DMF (2 mL) under nitrogen. Benzyl bromide (37 µL, 53 mg, 0.31 mmol, 1.0 equiv.) was added to the mixture and stirred overnight (approx. 16 h). The mixture was then diluted with ethyl acetate (30 mL) and washed once with water (40 mL) and once with brine (30 mL) and filtered through $MgSO_4$. The filtrate was concentrated and purified via flash chromatography eluting with 2:1 ethyl acetate-hexanes to yield a colorless foam (188 mg, 0.25 mmol, 82 %). R_f 0.42 (2:1 ethyl acetate-hexanes). ^{1}H NMR ($CDCl_3$, 250 MHz) ∂ 1.25 (t, J = 7.1 Hz, 3H), 1.39 (s, 5H), 1.51 (s, 2H), 2.03 (s, 6H), 2.10 (s, 1H), 2.47 (s, 3H), 2.55 (s, 3H), 2.85 (s, 2H), 3.13 (s, 2H), 4.11 (q, J = 7.1 Hz, 3H), 4.28-4.33 (m, 3H), 4.79 (s, 1H), 5.08 (s, 2H), 5.84 (d, J = 7.9 Hz, 1H), 6.13 (s, 1H), 6.23 (s, 1H), 7.19-7.37 (m, 9H), 7.53 (d, J = 7.1 Hz, 2H), 7.71 (d, J = 7.3 Hz, 2H).

H_2N-Arg(NHPbf)-OBn. Diethylamine (0.75 mL, 0.53 g, 7.25 mmol, 72.5 equiv.) was added to a solution of Fmoc-Arg(NHPbf)-OBn (75 mg, 0.10 mmol, 1.0 equiv.) in dry MeCN (1.5 mL) at 0 ˚C. The mixture was stirred until TLC showed no remaining starting material (~3 h). The mixture was evaporated and the residue used directly in the Gln-Ile-Arg tripeptide synthesis (see below).

Fmoc-Gln(Trt)-Ile-OH. Fmoc-Gln(Trt)-OH (94 mg, 0.15 mmol, 1.0 equiv.) was dissolved in dry CH_2Cl_2 (2 mL). *N*-hydroxysuccinimide (17 mg, 0.15 mmol, 1.0 equiv.) and DCC (32 mg, 0.15 mmol, 1.0 equiv.) were added and the reaction mixture stirred 4 h. The mixture was filtered to remove insoluble DCU and the filtrate concentrated to approximately 1 mL and placed in the freezer for 2 h. The mixture was refiltered, and the filtrate evaporated and redissolved in DMF (2 mL) and the solution cooled in an ice bath under a nitrogen atmosphere. H-Glu(O^{t}Bu)-OH (20 mg, 0.15 mmol, 1.0 equiv.) was added followed via dropwise addition of diisopropylethylamine (27 μL, 20 mg, 0.15 mmol, 1.0 equiv.). The reaction mixture was stirred overnight (approx. 16 h). The mixture was then diluted with ethyl acetate (20 mL) and washed with 2 M HCl (15 mL). The aqueous layer was back-extracted with ethyl acetate (20 mL) and each organic layer washed separately with water (15 mL each), filtered over $MgSO_4$, evaporated, and dried to yield a colorless foam. The dipeptide acid was not isolated and was used directly in the next step

Fmoc-Gln(Trt)-Ile-Arg(NHPbf)-OBn. Fmoc-Gln(Trt)-Ile-OH (max. 107 mg, 0.154 mmol, 1.05 equiv.) and H_2N-Arg(NHPbf)-OBn (max. 52 mg, 0.10 mmol, 1.0 equiv.) were combined in dry CH_2Cl_2 (5 mL) in an ice bath, under a nitrogen atmosphere. HATU (58.0 mg, 0.15 mmol, 1.05 equiv.) and collidine (37 μL, 33.9 mg, 0.28 mmol, 2.0 equiv.) were added and the mixture stirred overnight (approx. 17 h). The mixture was then purified via flash chromatography eluting with 2:1 hexanes-ethyl acetate, evaporated, and dried to yield a colorless foam (110 mg, 0.090 mmol, 64 %). R_f 0.41 (3:1 ethyl acetate-hexanes). ^{1}H NMR ($CDCl_3$, 400 MHz) ∂ 0.81-0.86 (m, 6H), 1.04-1.20 (m, 6H), 1.24-1.37 (m, 4H), 1.44 (s, 5H), 1.58-1.59 (m, 1H), 1.64 (s, 8H), 1.69 (dt, *J* = 13.6, 3.8 Hz, 2H), 1.91-1.95 (m, 3H), 2.05 (s, 1H), 2.07 (s, 2H), 2.40 (s, 1H), 2.48 (s, 2H), 2.54 (s, 2H), 2.93 (s, 3H), 3.05-3.10 (m, 1H), 3.43-3.52 (m, 2H), 4.06-4.20 (m, 3H), 4.37 (d, *J* = 6.9 Hz, 2H), 4.56-4.61 (m, 1H), 5.10 (d, *J* = 12.0 Hz, 1H), 5.15 (d, *J* = 12.6 Hz, 1H), 5.30 (s, 1H), 5.56 (s, 3H), 5.86 (s, 1H), 6.61 (s, 1H), 7.17 (app. d, *J* = 7.5 Hz, 4H), 7.22-7.39 (m, 16H), 7.55 (d, *J* = 7.5 Hz, 2H), 7.75 (d, *J* = 7.3 Hz, 2H); ^{13}C NMR ($CDCl_3$, 100 MHz) ∂ 10.9, 12.4, 15.3, 17.9, 19.2, 28.5, 38.6, 43.2, 47.0, 53.4, 67.3, 70.7, 86.2, 117.3, 119.9, 125.1, 127.1, 127.2, 127.7, 128.9, 128.3, 128.3, 128.5, 128.6, 132.2, 135.0, 138.2, 144.1. HRMS (+ESI) calculated for $C_{13}H_{24}NO_4$ $(M+Na)^+$ 1222.5682, observed 1222.5696.

Fmoc-Gln(ONHTrt)-Ile-Arg(NHPbf)-OH. Fmoc-Gln(Trt)-Ile-Arg(NHPbf)-OBn (53 mg, 0.043 mmol, 1.0 equiv.) was dissolved in ethyl acetate (1.5 mL) and combined with palladium on carbon (10.6 mg, 0.10 mmol, 0.2 equiv w/w). The reaction flask was then evacuated using a vacuum pump, and subsequently opened to a balloon filled with

H_2 gas. The mixture was stirred overnight (~18 h). The catalyst was removed via filtration through a Celite cake and the filtrate concentrated and the crude acid used directly in the next reaction.

Fmoc-Gln(Trt)-Ile-Arg(NHPbf)-Lys(NHBoc)-NH_2. Fmoc-Gln(Trt)-Ile-Arg(NHPbf)-OH. (max. 49 mg, 0.043 mmol, 1.05 equiv.) and H_2N-Lys(NHBoc)-NH_2 (10.0 mg, 0.041 mmol, 1.0 equiv.) were combined in dry CH_2Cl_2 (3 mL) in an ice bath, under a nitrogen atmosphere. HATU (16.3 mg, 0.043 mmol, 1.05 equiv.) and collidine (11 μL, 9.9 mg, 0.082 mmol, 2.0 equiv.) were added and the mixture stirred overnight (approx. 17 h). The mixture was then purified via flash chromatography eluting with 19:1 dichloromethane-methanol, evaporated, and dried to yield a solid white foam (40 mg, 0.029 mmol, 72 %). R_f 0.43 (9:1 dichloromethane-methanol). ^{1}H NMR ($CDCl_3$, 400 MHz) ∂ 0.81-0.88 (m, 5H), 1.26-1.45 (m, 20H), 2.05-2.08 (m, 4H), 2.39 (s, 4H), 2.44 (d, J = 10.2 Hz, 3H), 2.52 (d, J = 9.9 Hz, 3H), 2.56 (s, 8H), 2.80 (s, 1H), 2.89-2.93 (m, 9H), 3.98-4.26 (m, 8H), 4.88 (s, 1H), 5.70 (s, 1H), 5.99 (s, 1H), 6.06 (s, 1H), 7.03 (s, 2H), 7.12-7.26 (m, 20H), 7.34 (t, J = 5.6 Hz, 2H), 7.55 (t, J = 5.7 Hz, 2H), 7.72 (t, J = 7.5 Hz, 2H); ^{13}C NMR ($CDCl_3$, 100 MHz) ∂ 12.5, 14.1, 14.2, 17.9, 19.3, 21.0, 21.4, 21.5, 22.6, 28.4, 31.5, 38.6, 53.4, 60.4, 86.4, 119.9, 123.4, 124.6, 127.1, 127.7, 127.9, 128.0, 128.5, 141.1, 144.3, 153.9, 154.7. HRMS (MALDI) calculated for $C_{75}H_{94}N_{10}O_{12}S$ $(M+Na)^+$ 1381.6671, observed 1381.691

4. References:

[1] Berg, Jeremy M. *et al.* Biochemistry. 6th ed. Freeman & Co New York. **2007**, 652.

[2] (a) West, C. M., "Evolutionary and functional implications of the complex glycosylation of Skp1, a cytoplasmic/nuclear glycoprotein associated with polyubiquitination," *Cell Mol. Life Sci.* **2003**, *60*, 229-240; (b) West, C. M.; Van der Wel, H.; Sassi, S.; Gaucher, E. A., "Cytoplasmic glycosylation of protein-hydroxyproline and its relationship to other glycosylation pathways," *Biochem. Biophys. Act.* **2004**, *1673*, 29-44.

[3] (a) Henchey, L. K., Jochim, A. L., Arora, P. S., "Contemporary strategies for the stabilization of peptides in the α-helical conformation," *Curr. Opin. Chem. Biol.* **2008**, *12*, 692-697; (b) Garner, J.; Harding, M. M., "Design and synthesis of α-helical peptides and mimetics," *Org. Biomol. Chem.* **2007**, *5*, 3577-3585.

[4] (a) Bao, J., Dong, X. Y., Zhang, J. Z. H., Arora, P. S., "Dynamical binding of hydrogen-bond surrogate derived Bak helices to antiapoptotic protein Bcl-X_L.," *J. Phys. Chem. B,* **2009**, *113*, 3565-3571; (b) Patgiri, A., Jochim, A. L., Arora, P. S., "A hydrogen bond surrogate approach for stabilization of short peptide sequinces in α-helical conformation," *Acc. Chem. Res.* **2008**, *41*, 1289-1300; (c) Wang, D., Chen, K., Kulp, J. L., Arora, P. S., "Evaluation of biologically relevant short α-helices stabilized by a main-chain hydrogen-bond surrogate," *J. Am. Chem. Soc.* **2006**, *128*, 9248-9256.

[5] Carpino, L. A.; Han, G. Y., "The 9-fluorenylmethoxycarbonyl amino-protecting group," *J. Org. Chem.* **1972**, *37*, 3404-3407.

[6] Carpino, L. A.; Sadat-Aalaee, D.; "Beyermann, M., "Tris(2-aminoethyl)amine as a substitute for 4-(aminomethyl)piperidine in the FMOC/polyamine approach to rapid peptide synthesis," *J. Org. Chem.* **1990**, *55*, 1673-1675.

[7] Fukuyama, T.; Jow, C. K.; Cheung, M., "2- and 4-Nitrobenzenesulfonamides: exceptionally versatile means for preparation of secondary amines and protection of amines," *Tetrahedron Lett.* **1995***, 36,* 6373-6374.

[8] Biron, E.; Kessler, H., "Convenient synthesis of *N*-methylamino acids compatible with Fmoc solid-phase peptide synthesis," *J. Org. Chem.* **2005**, *70*, 5183-5189.

[9] Chapman, R. N.,; Dimartino, G.; Arora, P. S., "A highly stable short α-helix constrained by a main-chain hydrogen-bond surrogate," *J. Am. Chem. Soc.* **2004**, *126*, 12252-12253.

[10] (a) Van der Wel, H.; Ercan, A.; West, C. M., "The Skp1 prolyl hydroxylase from *Dictyostelium* is related to the hypoxia-inducing factor-α class of animal prolyl 4-hydroxylases," *J. Biol. Chem.* **2005**, *280*, 14645-14655; (b) West, C. M.; Van der Wel, H.; Wang, Z. A., "Prolyl 4-hydroxylase-1 mediates O_2-signalling during development of *Dictyostelium*," *Develop*. **2007**, *134*, 3349-3358.

[11] (a) Teng-umnuay, P.; Van der Wel, H.; West, C. M., "Identification of a UDP-GlcNAc:Skp1-hydroxyproline GlcNAc-transferase in the cytoplasm of *Dictyostelium*," *J. Biol. Chem.* **1999**, *274*, 36392-36402. (b) Van der Wel, H.; Morris, H. R.; Panico, M.; Paxton, T.; Dell A.; Kaplan, L.; West, C. M., "Molecular cloning and expression of a UDP-*N*-acetylglucosamine(GlcNAc):hydroxyproline polypeptide GlcNAc-transferase that modifies Skp1 in the cytoplasm of *Dictyostelium*," *J. Biol. Chem.* **2002**, *277*, 46328-46337.

[12] Van der Wel, H.; Johnson, J.M.; Xu, Y.; Karunaratne, C.V.; Wilson, K.D.; Vohra, Y.; Boons, G.-J.; Taylor, C.M.; Bendiak, B.; West, C.M., "Requirements for Skp1 processing by cytosolic prolyl-4(*trans*)-hydroxylase and α-*N*-acetylglucosaminyltransferase enzymes involved in O_2-signaling in *Dictyostelium,*" *Biochemistry* **2011**, *50*, 1700-1713.

[13] Artamkina, G. A., Egorov, M. P., Beletskaya, I. P. "Some aspects of anionic σ-complexes". *Chem. Rev.* **1982,** *82,* 427–459.

[14] Carpino, L. A.; El-Faham, A., "Effect of tertiary bases on *O*-benzotriazolyluronium salt-induced peptide segment coupling," *J. Org. Chem.* **1994**, *59*, 695-698.

[15] West, C. M., *et al.* "Complex glycosylation of Skp1 in *Dictyostelium:* implications for the modification of other eukaryotic cytoplasmic and nuclear proteins". *Glycobiology,* **2002**. *12 (2),* 17R-27R.

[16] Fukuyama, T., Cheung, C., Chung-Kuang, J., Hidai, Y., Kan, T. "2,4-Dinitrobenzenesulfonamides: A Simple and Practical Method for the Preparation of a Variety of Secondary Amines and Diamines". *Tetrahedron Lett.* **1997***, 38(33),* 5831-5834.

[17] Guisado, C, *et al.* "The facile preparation of primary and secondary amines *via* an improved Fukuyama-Mitsunobu procedure. Application to the synthesis of a lung-targeted gene delivery agent". *Org. Biomol. Chem.* **2005**, *3*, 1049-1057.

[18] Prokhorov, D. I., *et al*. "Synthesis of Thymine-Containing Monomers of Negatively Charged Peptide Nucleic Acids". *Pharm. Chem. J.* **2005**, *39(6)*, 323-328.

Appendix: ^{1}H NMR Spectra

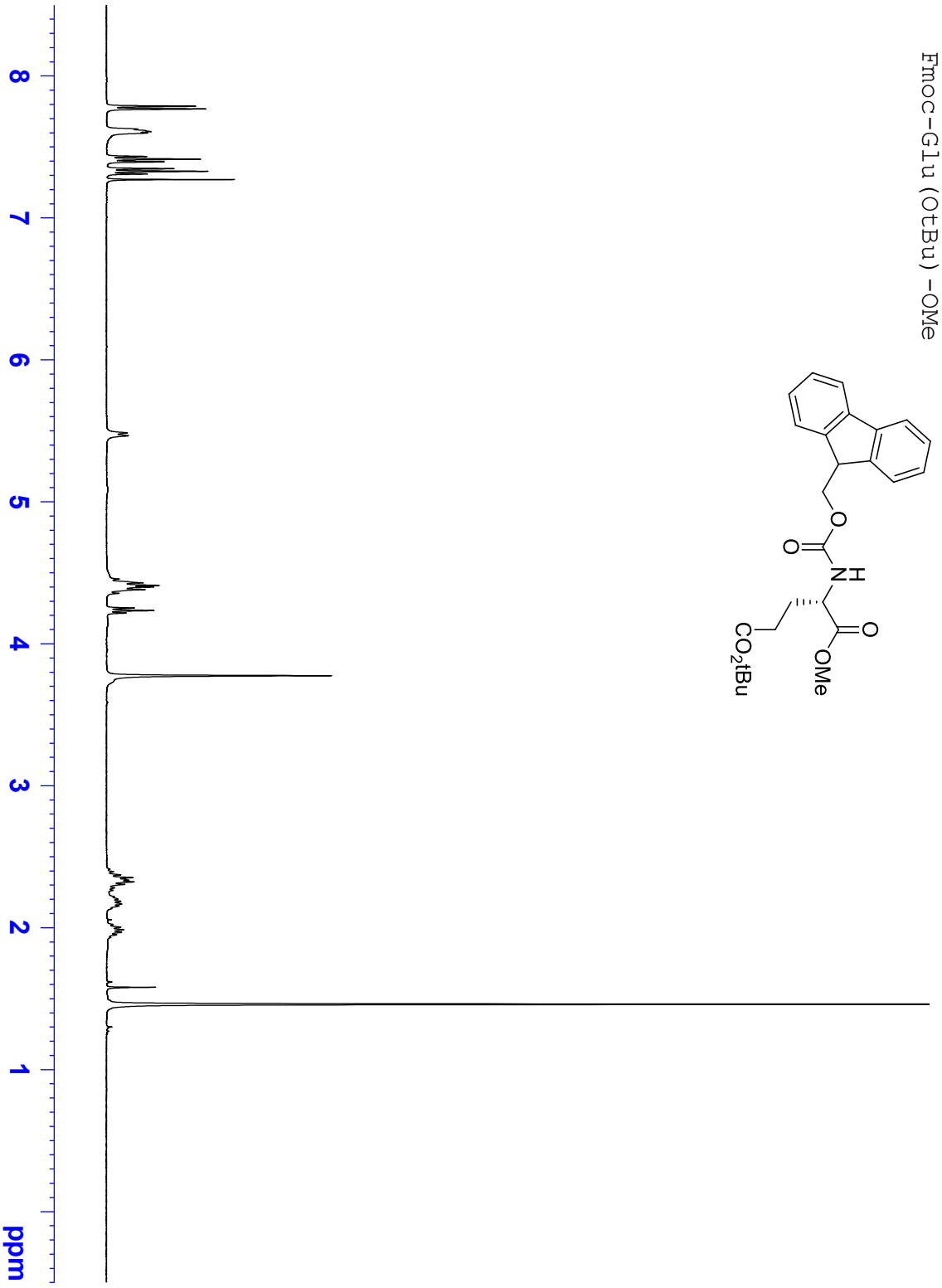


Figure A1

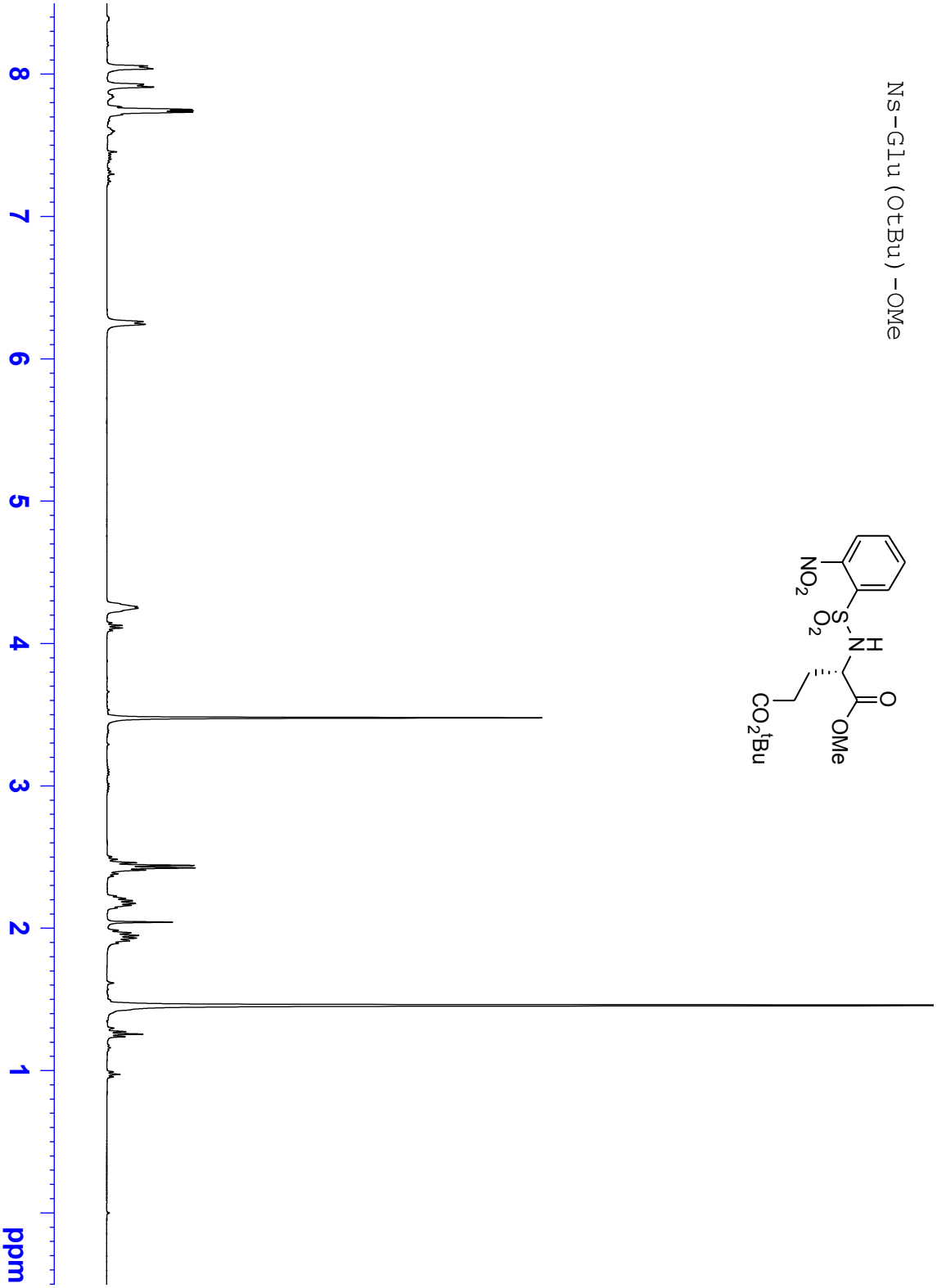


Figure A2

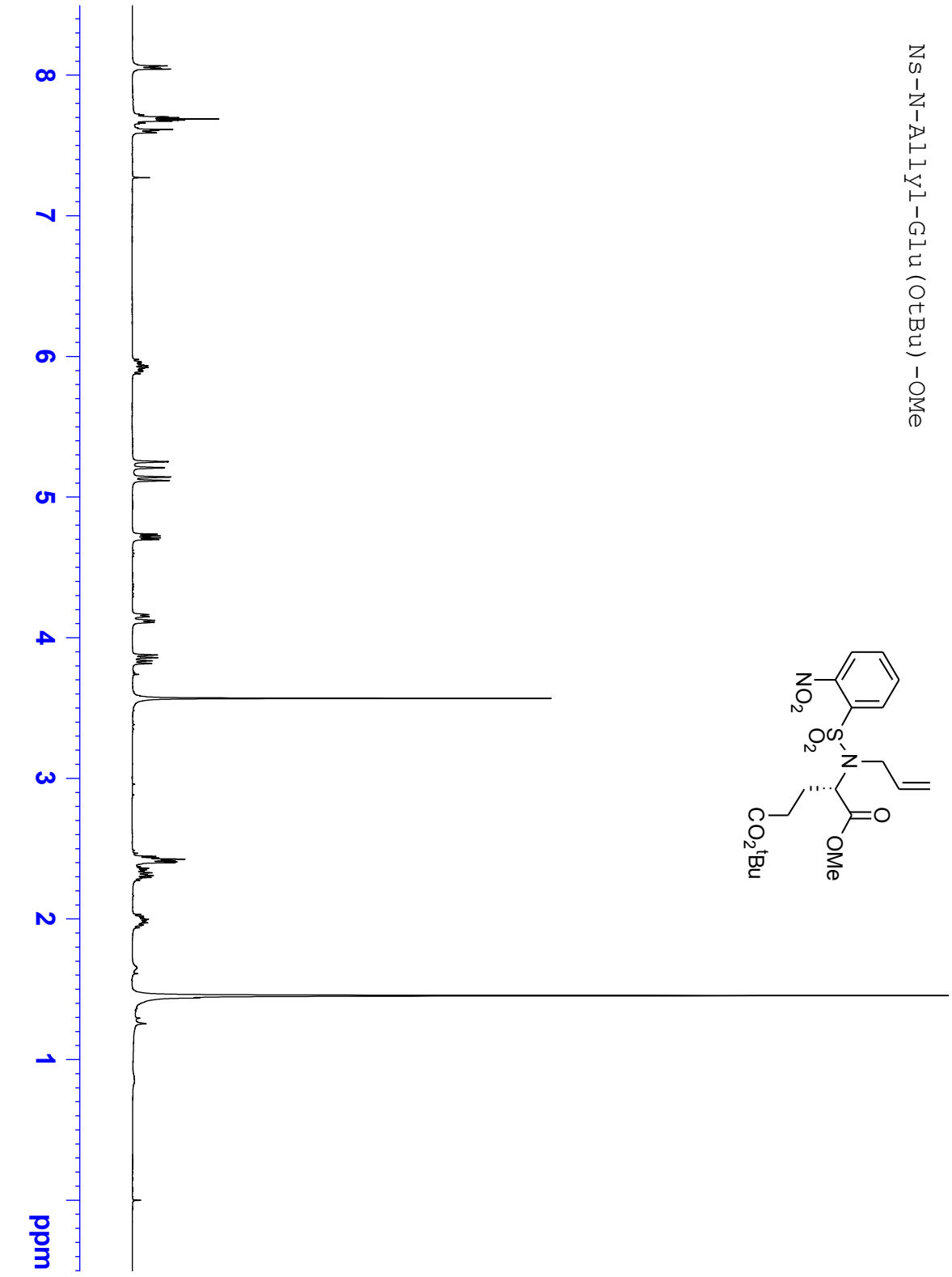


Figure A3

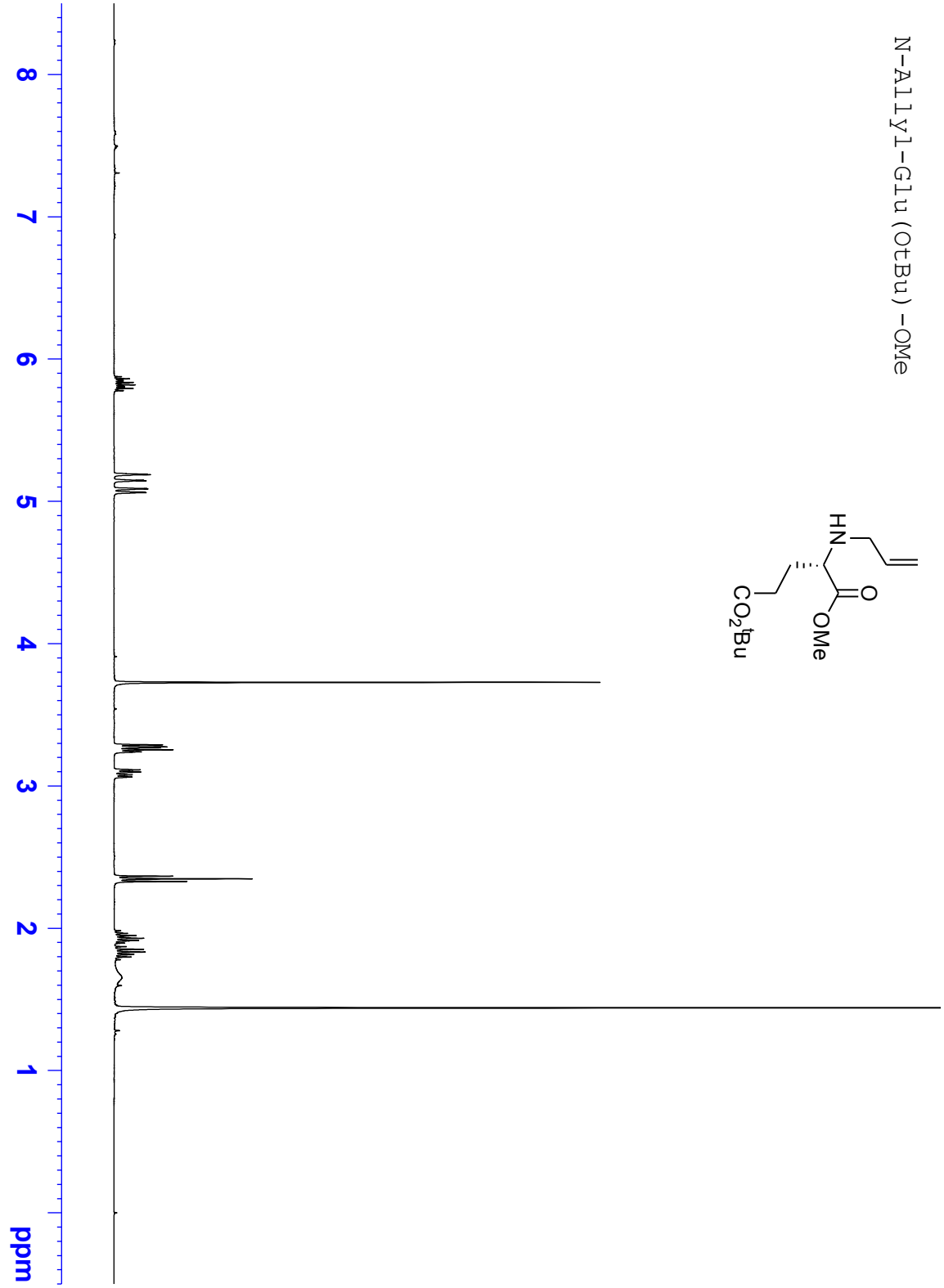


Figure A4

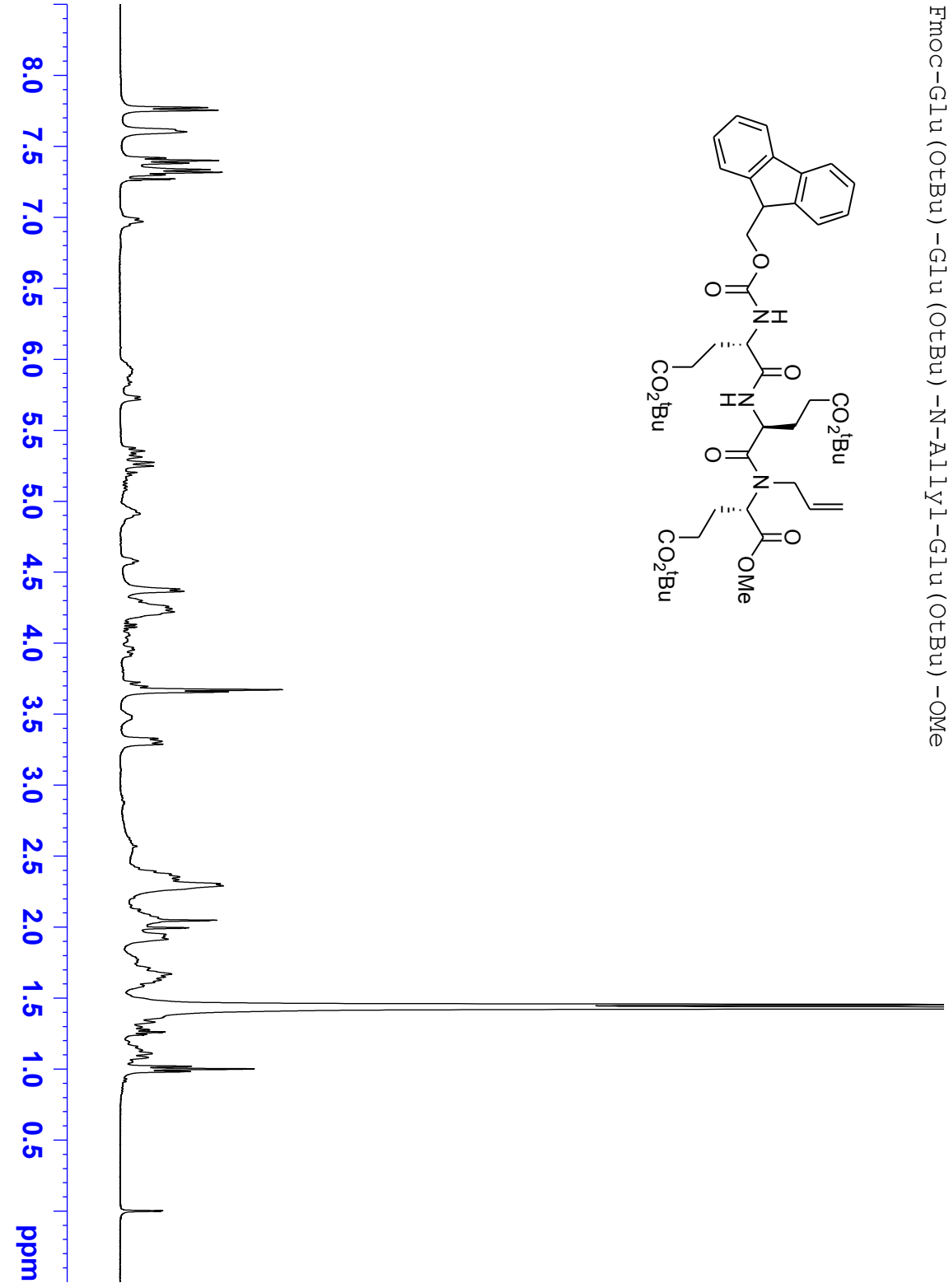

Fmoc-Glu(OtBu)-Glu(OtBu)-N-Allyl-Glu(OtBu)-OMe
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
ppm

Figure A5

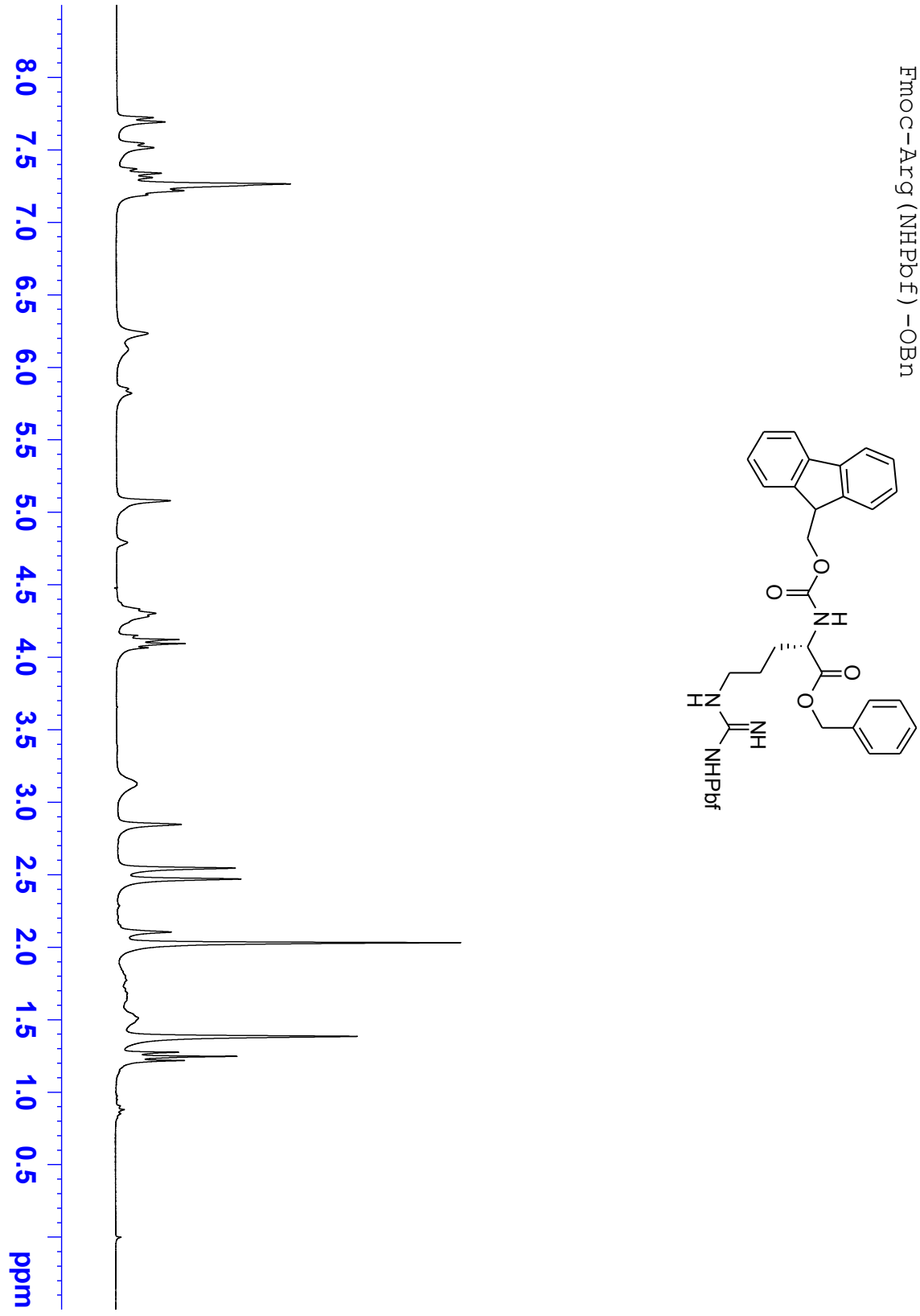

Fmoc-Arg(NHPbf)-OBn
NHPbf
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
ppm

Figure A6

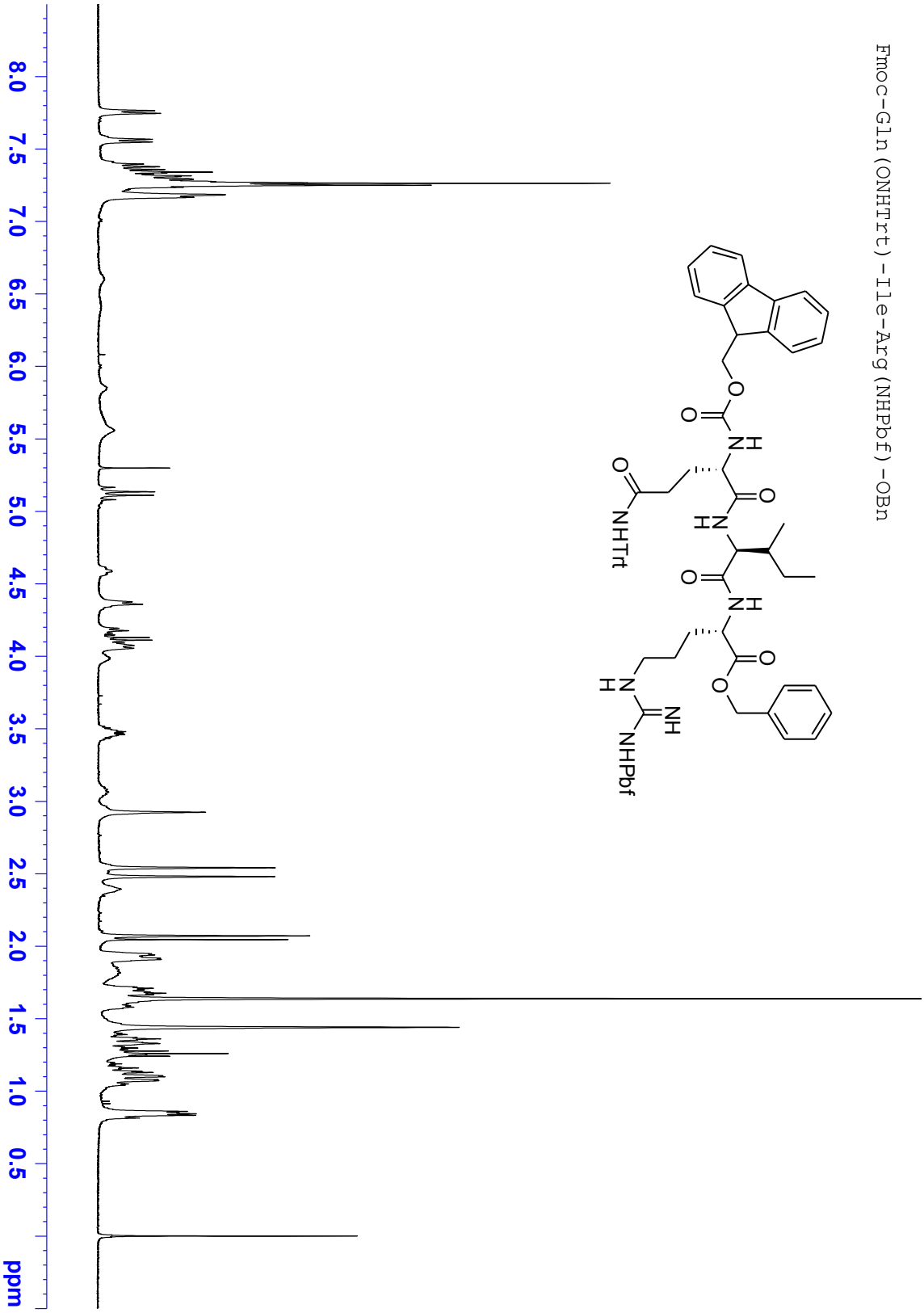

Fmoc-Gln(ONHTrt)-Ile-Arg(NHPbf)-OBn
NHTrt
NHPbf
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
ppm

Figure A7

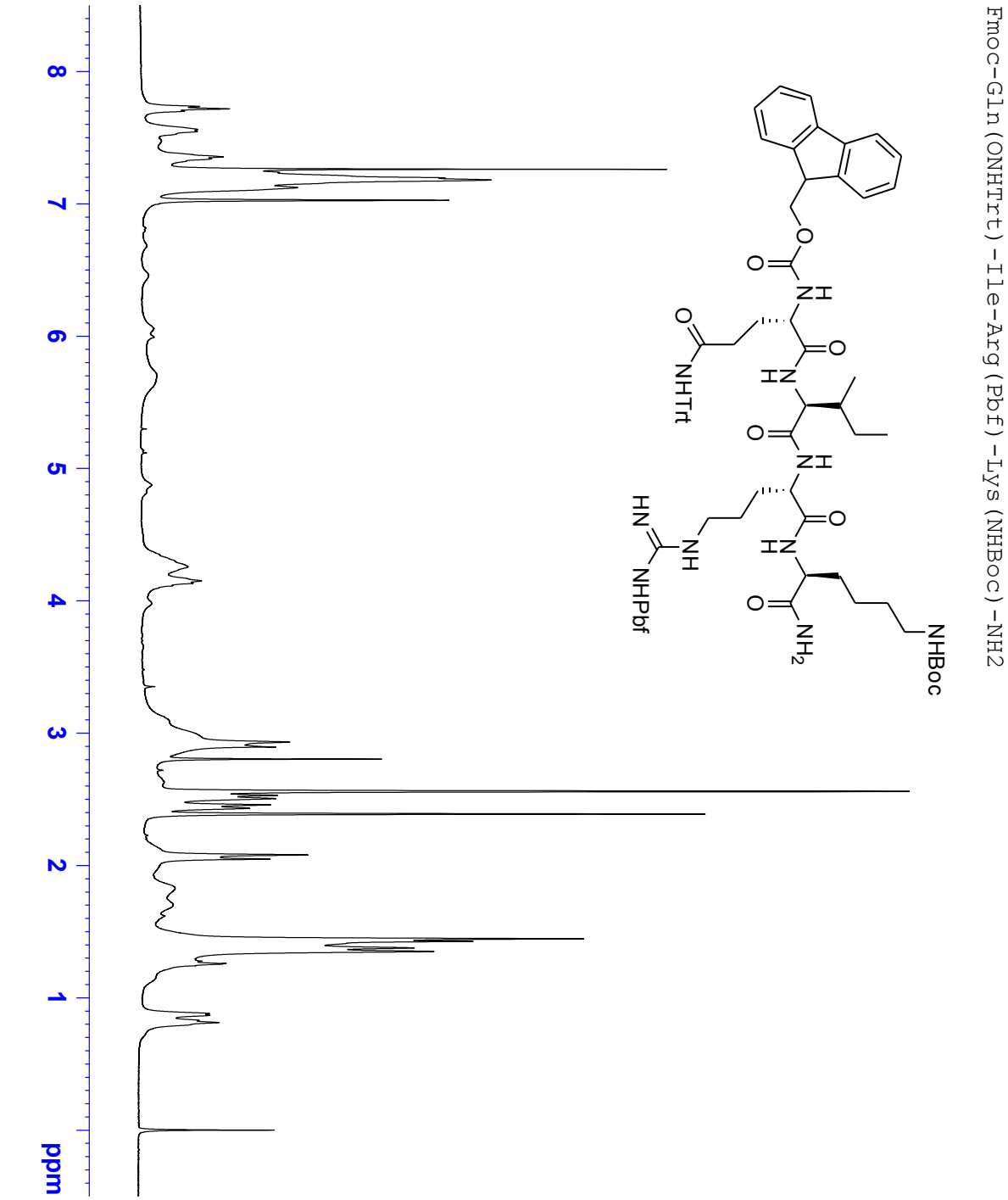

Fmoc-Gln(ONHTrt)-Ile-Arg(Pbf)-Lys(NHBoc)-NH2
NHTrt
NHPbf
NH2
NHBoc
8
7
6
5
4
3
2
1
ppm

Figure A8