DESIGN AND SYNTHESIS OF DNA MINOR GROOVE METHYLATING COMPOUNDS THAT TARGET PANCREATIC β-CELLS

Andrew McIver

A Thesis Submitted to the
University of North Carolina Wilmington in Partial Fulfillment
Of the Requirements for the Degree of
Master of Science

Department of Chemistry and Biochemistry

University of North Carolina Wilmington

2006

Approved by

Advisory Committee

_____Dr. Pamela Seaton_____ _____Dr. Paulo Almeida_____

____Dr. Sridhar Varadarajan___
Chair

Accepted by

_____
Dean, Graduate School

This thesis has been prepared in the style and format

Consistent with the journal

Journal of Organic Chemistry

TABLE OF CONTENTS

ABSTRACT

The design of compounds that form cytotoxic, non-mutagenic 3-methyladenine adducts in pancreatic ß-cells is being studied in this project for potential applications in the treatment of diseases such as diabetes and cancer. These compounds are composed of three components: 1) a cell-targeting moiety, glucosamine, which targets the insulin producing pancreatic β-cells by way of the GLUT-2 transporters present on these cells 2) a site-specific DNA methylating agent, Me-Lex, which has been shown to selectively produce cytotoxic, non-mutagenic N3-methyladenine adducts 3) a linker component that connects the two other components together. The linker is a critical component because it has to be such that the cell-targeting and DNA-methylating properties of the two functional components are maintained. A synthetic route was explored, which enables the easy introduction of various linkers into the molecules. Fluorescent compounds were also designed to bind weakly to DNA at the same positions as the DNA-methylating compounds. These fluorescent compounds will be used to calculate the binding constants of weakly binding compounds that bind to the minor groove of DNA at A/T rich regions. The design features and the synthesis of these compounds are described.

ACKNOWLEDGEMENTS

I would like to thank Dr. Sridhar Varadarajan for guiding me through this project as an honors student and as a graduate student. His ideas are so very interesting and the way he explains things makes it easy to learn and keeps me excited about chemistry. I would like to thank Dr. Pamela Seaton for being on my committee and being my first college chemistry teacher and captivating me in chemistry, especially organic chemistry, which I am excited about most. I would also like to thank Dr. Paulo Almeida for being on my committee and helping with the fluorescence experiments. The UNCW chemistry and biochemistry department has been wonderful, giving me such a great college chemistry experience. Last, I would like to thank my mother Lisa McIver for always encouraging me with school and being excited about what I tell her about my chemistry projects, even if she doesn't understand every thing I tell her.

LIST OF FIGURES

Figure Page

Figure Page

LIST OF SCHEMES

Scheme Page

CHAPTER 1. INTRODUCTION

Many Americans suffer from diabetes (about 18 million) of which about 5-10 % (about 5 million) have type-1 diabetes.[1] Diabetes is a disease in which the body does not properly use, or does not produce, insulin.[1] Insulin is a hormone secreted in the pancreas that enables cells in the body to take up glucose from the blood and create energy for everyday needs. The cells that produce insulin in the body are pancreatic β-cells. Type-1 diabetes results when insulin producing pancreatic β-cells are destroyed. Destruction of these cells is thought to occur because of an immune response caused by genetic factors, viral infections, or environmental factors, but the mechanism of destruction is still poorly understood.[2-3]

One of the commonly used animal models to study type-1 diabetes is one in which a DNA damaging drug, streptozotocin (STZ), is used to kill pancreatic β-cells. This compound, STZ, causes diabetes by two different mechanisms. In one mechanism a single large dose of STZ is used, that causes rapid and complete destruction of pancreatic β-cells resulting in diabetes.[4] This is a non-immune mechanism induced diabetes and is not reflective of human type-1 diabetes. In the second mechanism, multiple small doses of STZ is used, which results in an initial drop in insulin production, followed by an immune response and causes complete destruction of pancreatic β-cells.[4] This second mechanism, involving an immune reaction is more reflective of human type-1 diabetes.

STZ is a compound that damages (methylates) DNA at multiple sites. However, there is evidence in literature to show that the particular damage, caused by STZ, that is responsible for the immune response in pancreatic β-cells is the formation of the 3-methyladinine (3-MeA) DNA adduct.[4] It is difficult to study the direct correlation between the 3-MeA formation and the induction of the immune response resulting in diabetes in the animal models because STZ damages DNA at multiple sites. In fact, damage caused by STZ at some sites on DNA (other

than N3-adenine) is believed to cause mutations and lead to the formation of tumors. Thus, the damage at multiple sites and the consequent formation of tumors has complicated the study of type-1 diabetes using the STZ rodent model.

It would be possible to study the role of 3-MeA in triggering an immune response if one could generate exclusively 3-MeA in pancreatic β-cells. Understanding the factors that trigger the immune response in pancreatic β-cells would provide and invaluable tool for the study of type-1 diabetes. Furthermore, by understanding how to induce an immune response in a particular kind of cell by forming only 3-MeA DNA adducts in those cells can lead to the development of new drugs to target tumor cells, and destroy them by using the body's immune system.

The goal of this project was to design and synthesize compounds to produce 3-MeA DNA adducts in insulin producing pancreatic β-cells. The project involved combining an agent capable of causing exclusively one kind of damage on DNA (i.e. 3-MeA adduct) with a unit that can target this DNA-damaging agent preferentially to pancreatic β-cells. This thesis describes the design and progress towards the synthesis of such molecules.

CHAPTER 2. BACKGROUND AND SIGFIFICANCE

2.1. Background

The goal of this project is to make compounds that can cause a specific kind of damage on DNA (i.e. 3-MeA adducts), which is believed to be the cause of the immune response that is responsible for the destruction of cells, and that can target pancreatic β-cells. Such compounds would enable the investigation of the biological consequences of forming 3-MeA adducts in these insulin producing cells. In order to design these new compounds that can produce exclusively 3-MeA adducts on DNA, one must first have a good understanding of the structure of DNA.

2.1.1. DNA Structure and Damage

B-DNA (deoxyribose nucleic acid) is in the shape of a double helix (Figure 2.1) with each strand of the double helix being composed of a negatively charged phosphate sugar backbone. Connected to the sugars perpendicular to the double helix are the DNA bases adenine (A), thymine (T), guanine (G), and cytosine (C). Hydrogen bonding between base pairs draws together the two phosphate sugar strands to form the double helix. The base pairing is very specific with adenine always paired to thymine and guanine always paired to cytosine. As a result of this base pairing that brings the two strands together to form the double helix, two grooves are created. One groove is broad and is called the major groove. Sites within the major groove are easier to access and most proteins that interact with DNA do so in the major groove. The other groove is narrow and deep and is called the minor groove. Sites within this groove are more difficult to access.

As a result of the base pairing arrangement, certain sites on the DNA base pairs are exposed in the major groove, others are exposed in the minor groove, while

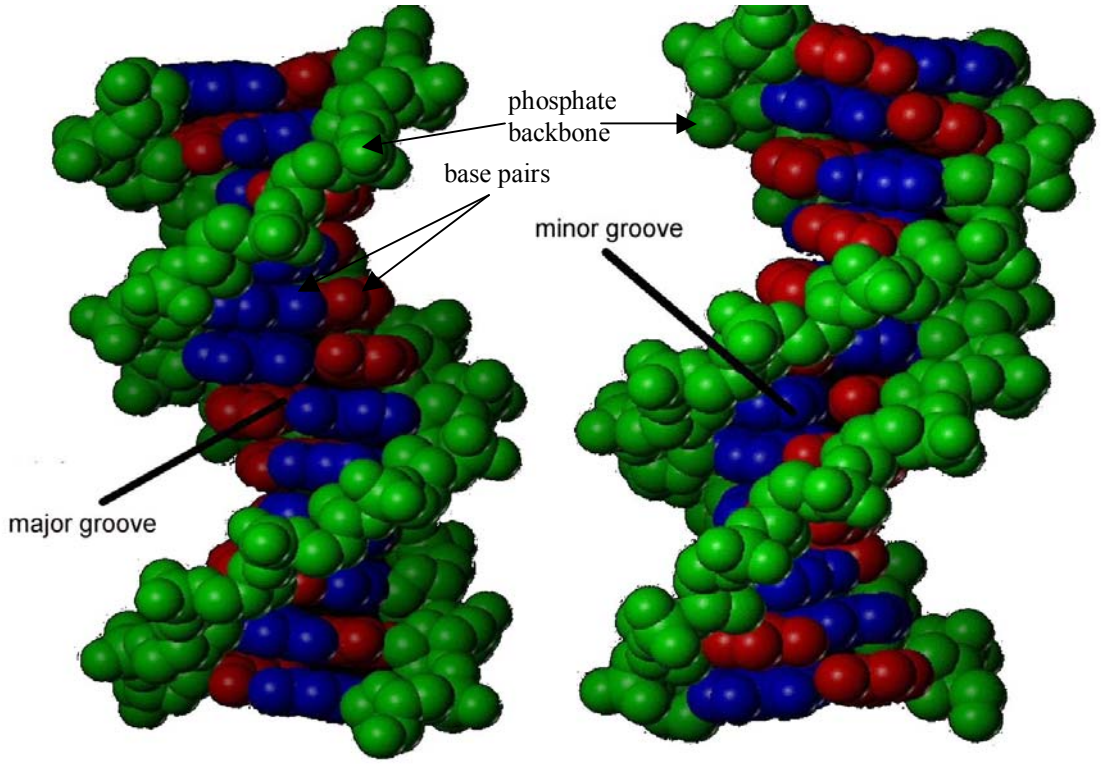


Figure 2.1. Structure of DNA showing the major and minor groove.

some sites are involved in the hydrogen bonding that draws the two DNA strands together, as can be seen from Figure 2.2. Many of these sites are targets for DNA methylation by DNA-methylating agents. For example some of the sites in the G/C base pairs which lie in the major groove and can be methylated are the N7 site of guanine, the O^6 site of guanine, and the exocyclic amine at the four position of cytosine. In fact the highly accessible and most nucleophilic N7 site of guanine is one of the most commonly methylated sites on DNA. The G/C base pair sites in the minor groove that can get methylated are N3 site of guanine and the exocyclic amine at the two position of guanine. Similarly, in the A/T base pairs, the sites in the major groove that can be methylated are the N7 site of adenine, the exocyclic amine at the six position of adenine, and O^4 site of thymine, and the sites which can be methylated in the minor groove are the N3 site of adenine and the O^2 site of thymine. In addition to the sites on the base pairs that can be methylated there are also sites on the phosphate backbone that can get methylated. The N3-adenine site, which is the target for methylation in this project, is indicated by the bold arrow in Figure 2.2.

Methylation on DNA can lead to different consequences depending on the site that is methylated. For example, the most common methylation seen on DNA, at the N7-guanine site is believed to be one which has little or no biological consequences.[5-8] On the other hand, methylation at the O^6 site of guanine is known to lead to both mutations and cell toxicity.[2-9] Similarly, there is evidence to show that methylation at N3-A results in cytotoxicity but does not lead to mutations.[5, 6, 9-12] Therefore, this 3-MeA adduct is a good choice when the goal is to kill a cell without causing other complications.

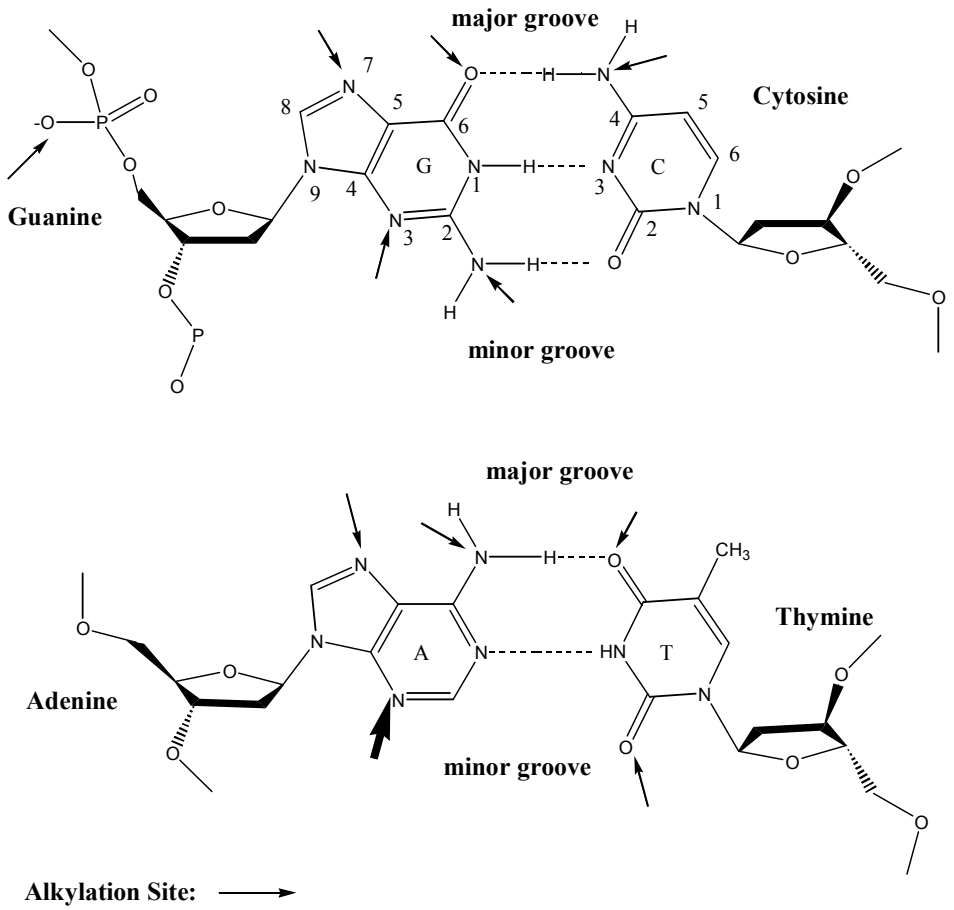


Figure 2.2. DNA base pairing, major and minor groove sites and sites that can be methylated on DNA. Each arrow indicates a site on DNA which can be alkylated.

2.1.2. STZ and its Properties

STZ, the compound that was used to induce diabetes by the destruction of pancreatic β-cells in animal models, is a methylating agent. It is an α-D-glucopyranose derivative of N-methyl-N-nitrosourea (Figure 2.3) that targets the pancreatic β-cells. STZ is believed to target pancreatic β-cells due to the selective uptake of the glucose moiety by the low affinity glucose transporter, GLUT-2, that is present on the surface of the pancreatic β-cells.[13-16]

The DNA methylation pattern by STZ is complex. It is known to methylate DNA at multiple sites such as N7-guanine, O^6-guanine, N7-adenine, and N3-adenine.[17] Of these adducts, the major adduct (70 %) which is formed by STZ is the N7-methylguanine.[17] N7-methylguanine, as discussed above is believed to be a benign adduct.[5-8] STZ also forms the O^6-methylguanine adduct that is known to lead to mutations and cause cell toxicity because it can miscode for thymine during DNA replication.[6-9] It is believed that this adduct is responsible for the formation of tumors in rodents treated with STZ, and this can lead to complications when studying diabetes using the STZ model. The small proportion of 3-MeA formed by STZ is believed to be responsible for the induction of the immune response caused by STZ in the STZ-rodent model. Since only a small percentage of the adducts formed by STZ cause the desired cell-toxicity, and since STZ treatment also induces tumor formation, it is not a very good choice for the study of type-1 diabetes.

The proof that the 3-MeA adduct was the adduct that was responsible for the immune response in the STZ-rodent model of type-1 diabetes came from studies with transgenic animals which were incapable of repairing 3-MeA adducts.[4] When these

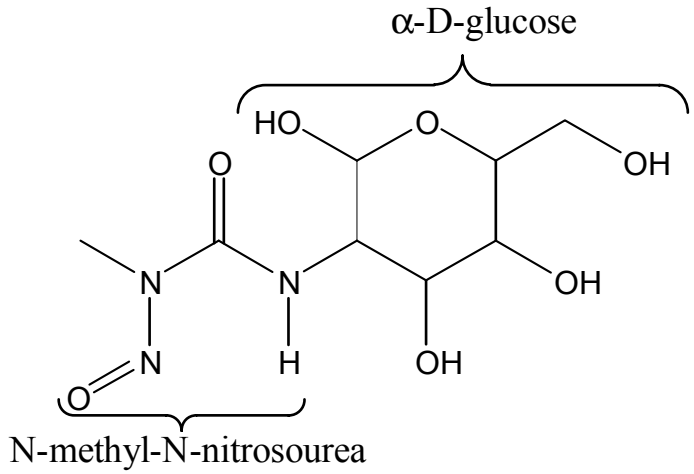


Figure 2.3. Structure of Streptozotocin (STZ).

animals were treated with low levels of STZ there was an initial drop in insulin levels, which soon recovered to normal levels. However, after a certain period of time these animals developed diabetes because of an immune response that destroyed the pancreatic β-cells.[4] When compared to wild type animals that did not develop diabetes when subjected to the same treatment, the only thing different between these transgenic animals and wild type animals was the inability of the transgenic animals to repair the 3-MeA adducts.[4] All other adducts formed by the low level treatment with STZ would have been repaired by the repair enzymes. This shows that, of the various DNA adducts formed by STZ, the 3-MeA adduct, when formed in low quantities and is unrepaired, is probably responsible for the immune response. This hypothesis that low levels of 3-MeA in pancreatic β-cells can trigger an immune response against those cells can be directly tested only if the 3-MeA adduct can be formed exclusively in those cells. This cannot be achieved by using reagents such as STZ, which form many different DNA adducts. Therefore, designing new compounds that can form exclusively 3-MeA adducts in pancreatic β-cells, would enable the investigation of the role of the 3-MeA adduct in causing the immune response and the factors that control the immune response.

2.1.3. 3-MeA Formation by Me-lex

Me-lex (Figure 2.4), a compound described in literature,[18] is known to selectively form 3-MeA adducts (> 95 %). Me-lex is a neutral DNA minor groove binding compound that is an N-methylpyrrolecarboxamide dipeptide (lex) with a methylating O-methyl sulfonate ester functionality attached at one end.[18] Me-lex targets specifically the minor groove at A/T rich regions on DNA and exclusively forms 3-MeA adducts in those A/T rich regions.[18] The reason it is exclusive to the minor groove in A/T

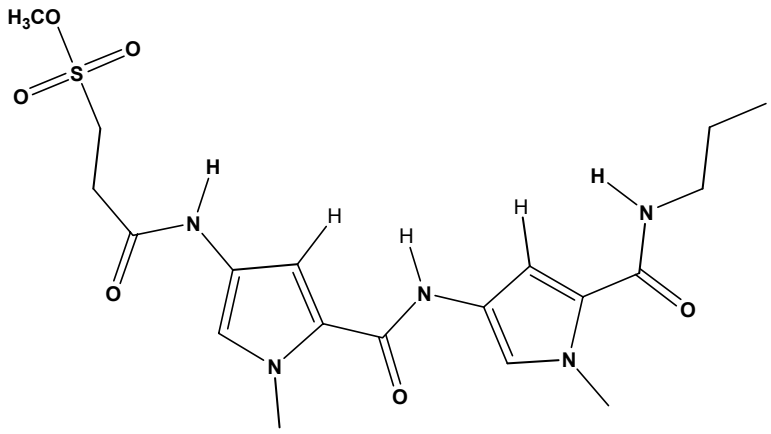


Figure 2.4. Structure of methylating lexitropsin (Me-Lex).

rich regions is because of its unique binding interactions with the DNA in those regions as illustrated in Figure 2.5. The interactions it has with DNA in A/T rich regions include H-bonds formed between the amide hydrogens of the pyrrole and the N3 of adenine and the exocyclic oxygen at the two position of thymine, and van der Waals interactions between the pyrrole hydrogens and the C2-position of adenine.[18] Once Me-lex binds in the minor groove, the methyl group is transferred to the most nucleophilic site in the minor groove at A/T rich regions, the N3-adenine by a concerted process to give the 3-MeA adduct. The compound then becomes a negatively charged sulfonate and is repelled away from the negatively charged DNA backbone and will have no other biological consequence of its own. Figure 2.6 shows a molecular model of Me-lex bound within the minor groove of DNA at A/T rich regions

The 3-MeA adducts that are formed in cells by Me-lex have been shown to be cytotoxic and non-mutagenic. These adducts are processed by the base excision repair pathway that triggers poly(ADP)-ribose polymerase (PARP-1) activation.[6] Over activation of PARP-1 due to high levels of 3-MeA depletes the cellular ATP levels and causes cell death by necrosis. If the PARP-1 is inhibited, or base excision repair is absent, then the cell dies by apoptosis.[6] It is believed that both these mechanisms of cell death, necrosis and apoptosis, are required for triggering an immune response.

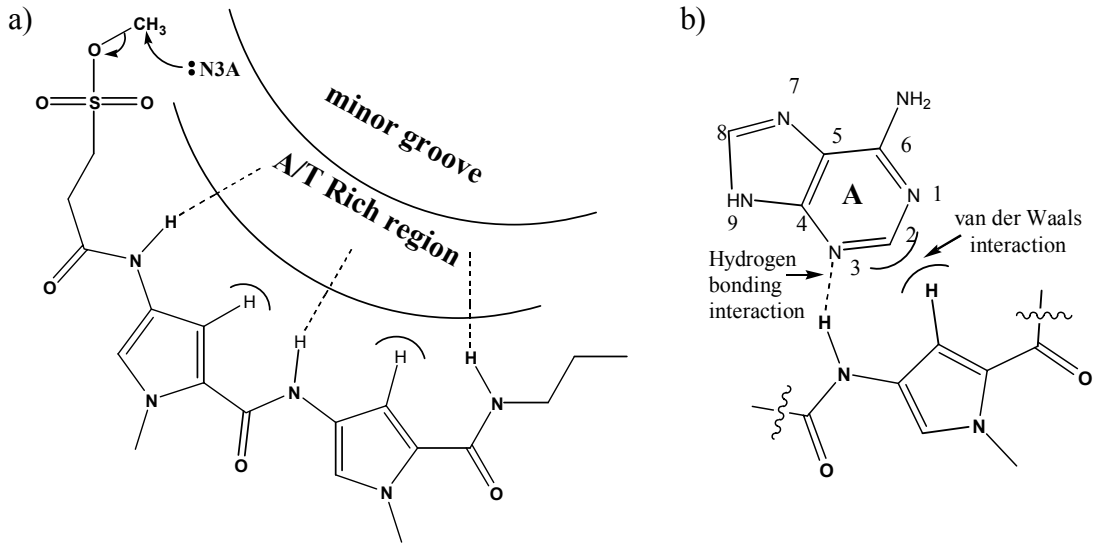


Figure 2.5. a) Interaction of Me-Lex at A-T rich regions of DNA, in the minor groove.
b) H-bonding and van der Waals interactions between adenines and the pyrrole amide units of Me-lex.

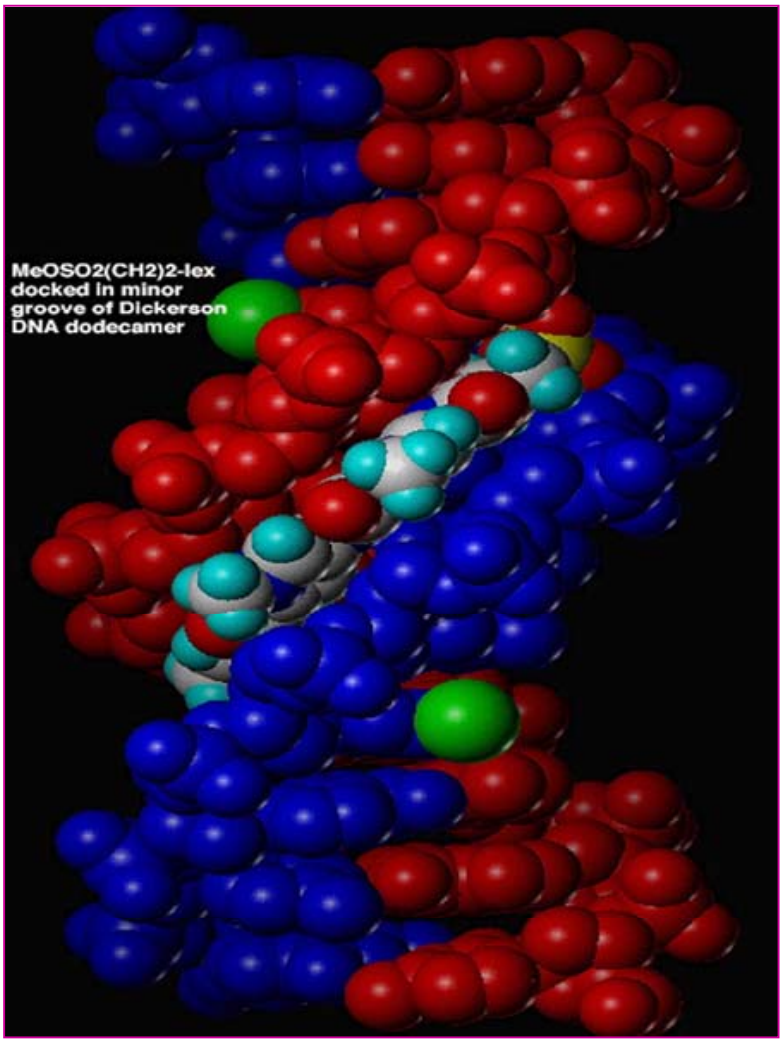


Figure 2.6. Molecular model showing Me-lex bound with the minor groove of DNA at A/T rich regions. This model was obtained by modification of a crystal structure of a similar compound bound to a DNA dodecamer.

2.2. Significance

This project describes the attempted design and preparation of new compounds that combine the pancreatic β-cell targeting ability of STZ with the selective N3-adenine methylating ability of Me-lex. The ability to successfully target and generate only 3-MeA adducts in pancreatic β-cells, and inducing diabetes in rodent models using these new molecules, would be very helpful in the research of type-1 diabetes. Also the low mutagenicity of 3-MeA, produced by these compounds would eliminate complications due to tumor formation that is reported for STZ rodent models of type-1 diabetes.[17]

If these compounds are successful in inducing an immune response against pancreatic β-cells on can investigate the factors needed to induce an immune response against particular cells. This could lead to the development of a strategy to use the immune system itself to target and destroy tumor cells.

CHAPTER 3. DESIGN AND SYNTHESIS OF DNA METHYLATING COMPOUNDS TARGETING PANCREATIC β-CELLS

3.1. Design of Molecules

The goal of this project is to make new molecules in order to deliver a sequence selective DNA methylating agent, namely one that can exclusively form 3-MeA DNA adducts, selectively to insulin producing pancreatic β-cells, by incorporating both cell targeting and DNA damaging properties in a single molecule. Pancreatic β-cells can be targeted by glucose units which are selectively taken up by the GLUT-2 glucose transporters that are known to be present on these cells. Exclusive DNA methylation at N3 of adenines can be achieved by using a compound described in literature, called Me-lex, which is capable of selectively producing 3-MeA adducts at A/T rich regions in the minor groove on DNA. Therefore, the strategy to make these new compounds capable of forming 3-MeA adducts in pancreatic β-cells is to incorporate, the selective DNA methylating ability of Me-lex with the pancreatic β-cell targeting ability of glucose. This strategy is schematically depicted in Figure 3.1.

The design of the new compounds has three important components:

1) a pancreatic β-cell targeting component which is the same as in STZ (see Figure 2.3), namely the glucose unit,

2) a component that can selectively produce 3-MeA, similar to Me-lex, which will replace the N-methylnitrosourea group of STZ, and

3) a variable linker component that will be used to attach the above two functional components together.

Molecules designed in this manner should be able to specifically target pancreatic β-cells, similar to STZ, and exclusively produce 3-MeA adducts in those cells, unlike STZ which produces multiple DNA adducts.

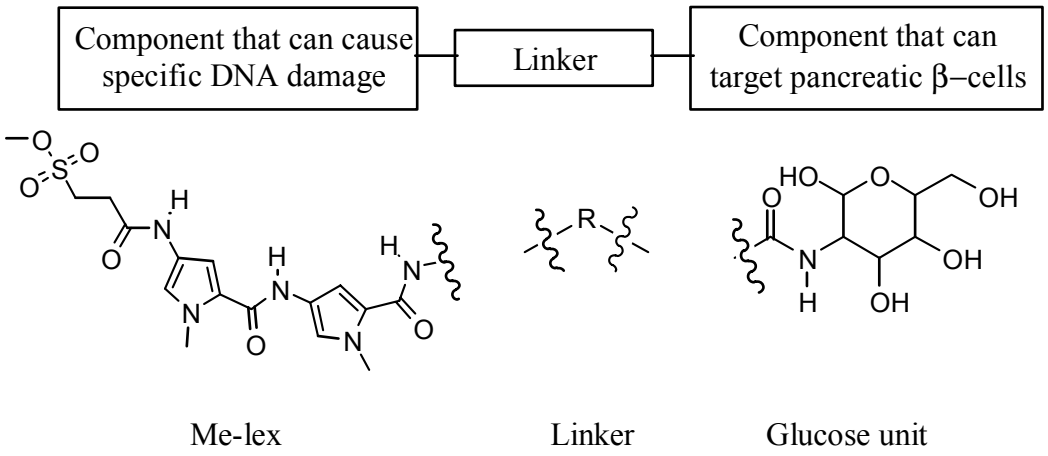


Figure 3.1. Design of compounds that can generate 3-MeA in pancreatic β-cells.

For compounds designed as shown in Figure 3.1 to achieve the desired goals successfully several factors need to be taken into account. One important consideration is whether the glucose unit, which is known to target the pancreatic β-cells when the N-methylnitrosourea or other small units are attached to it, will be able to target pancreatic β-cells when the methylating Me-lex unit is attached to it. The second important factor is whether the Me-lex unit (which is known to produce exclusive 3-MeA adducts) will be able to still selectively target sites on DNA and produce exclusively 3-MeA adducts efficiently, similar to the parent Me-lex molecule, even when the glucose unit is tethered to it. Therefore, the linker, which is used to connect the Me-lex unit to the glucose unit, assumes a critical role. This linker, which is the only variable component in the design, has to be modified in order to achieve the optimum cell-targeting and DNA-damaging properties that are desired. The composition of the linker can also be varied in order to improve the water solubility of the compounds which is required for biological applications.

The pancreatic β-cell targeting ability of the glucosamine unit with modifications at the nitrogen is well described in literature.[19-24] Several molecules such as those shown in Figure 3.2, with small units attached to the N of the glucosamine have been successfully targeted to cells with GLUT-2 glucose transporters.[19-24] However, there is not much evidence in literature of large units being attached to glucose units and being targeted to pancreatic β-cells. But, there is no evidence that indicates that such molecules, with large components attached to glucosamine cannot be targeted to pancreatic β-cells. A recent patent describes the use of a porphyrin-glucose conjugate (Figure 3.3) in selective photodynamic therapy for cancer.[25] Increased uptake of this

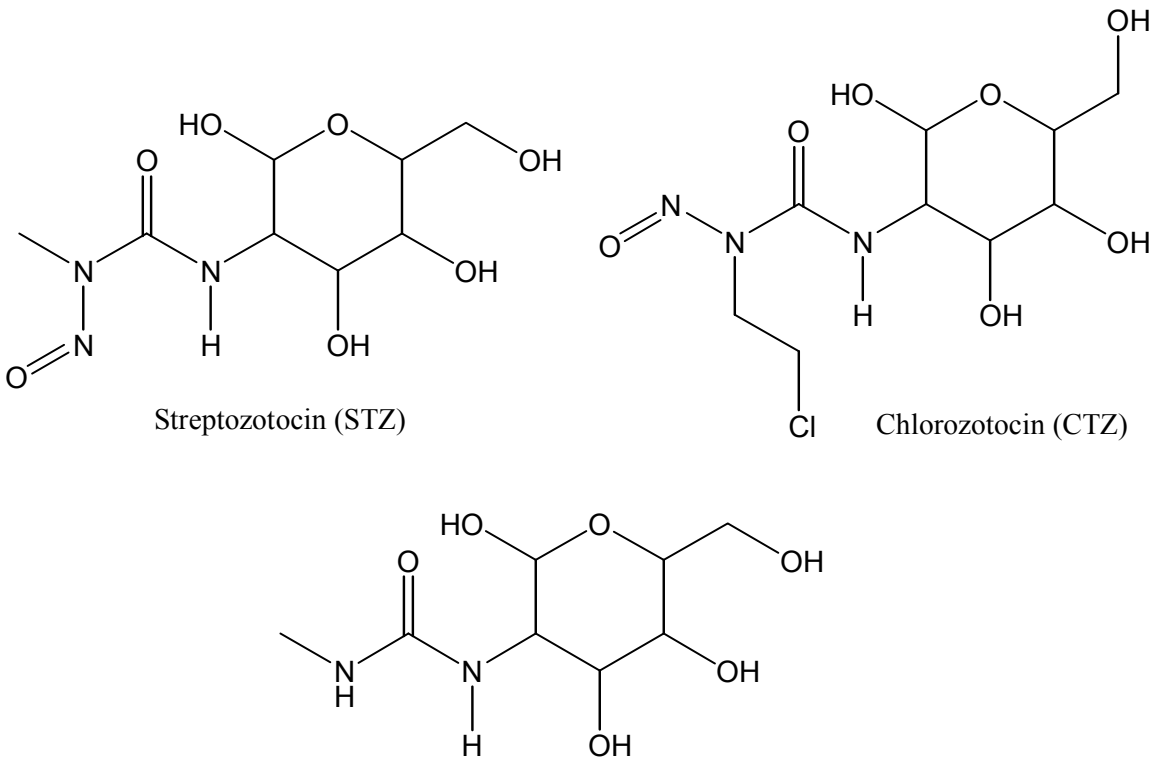


Figure 3.2. Small molecules attached to a glucose unit taken up by glucose transporters.

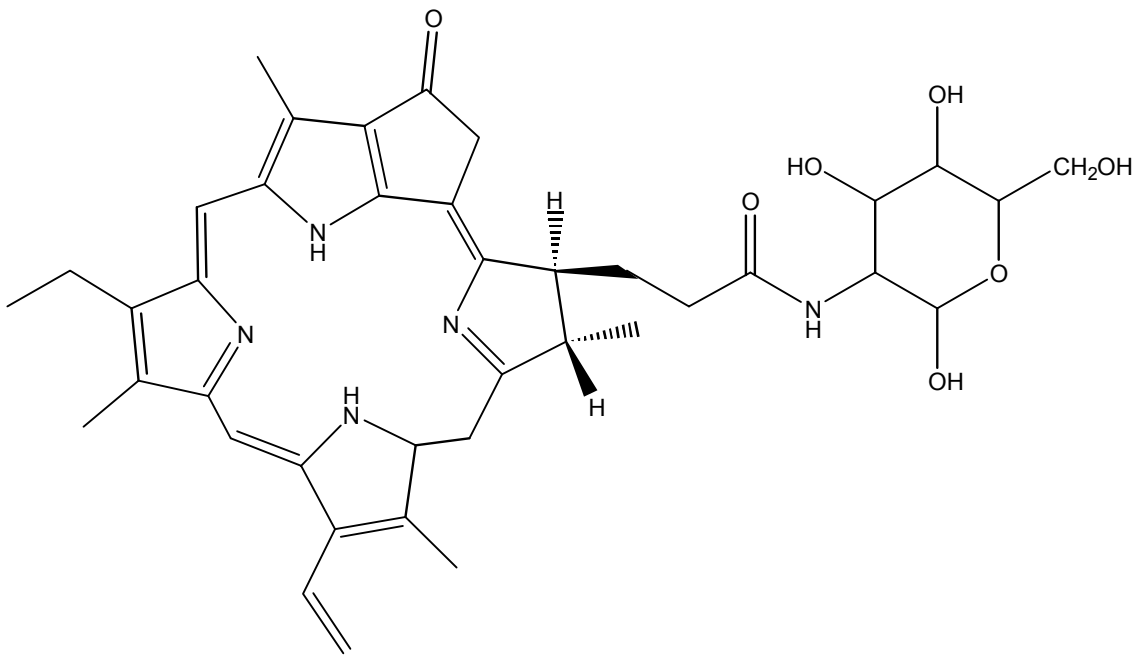


Figure 3.3. Structure of Pyro-2DG, where a large porphyrin ring is attached to a glucose unit.

molecule by certain cancer cell lines has been described. This increased uptake is attributed to the over expression of glucose transporters in these cancer cells,[25] though the actual glucose transporters that were used to transport this molecule into cells were not identified. If the large porphyrin unit is delivered within a cell by glucose transporters (possibly the GLUT-2 transporter), then it should be possible to deliver the Me-lex unit to pancreatic β-cells via the GLUT-2 transporter by tethering it to glucosamine.

The ability of the Me-lex unit to efficiently produce exclusively 3-MeA adducts on DNA, despite its attachment to the glucose unit via a tether, is crucial for the success of this project. The glucose unit is attached at the C-terminus of the Me-lex unit, and away from the reactive methylating group. It is not clear what effect the attached glucose unit will have on the methylating ability of the molecules. The selective binding of the new molecules to the desired A/T rich sites in the minor groove of DNA is critical. If the attached glucose unit also slides into the minor groove and forms favorable hydrogen bonding interactions within the groove, the binding affinity for these molecules will improve over the parent Me-lex molecule, and efficient methylation should be observed. If on the other hand, the glucose unit has steric conflicts with the DNA backbone and interferes with the binding of the Me-lex unit at the A/T rich regions of the minor groove, then the linker will have to be of sufficient length in order to suspend the glucose unit outside the DNA while allowing the Me-lex unit to still bind in the target site.

The methylation of DNA by these molecules is a bimolecular reaction. Therefore, strength of the binding of the molecules at the target site can be expected to directly affect the ability of these molecules to methylate the N3-A. These compounds are designed to be weakly binding so that once the methyl group is transferred to the DNA, the resulting negative charge on the molecule will cause it to be repelled from the negatively charged phosphate backbone, and be

eliminated from the DNA. However, a significant reduction in binding as compared to the parent Me-lex molecule, due to the attachment of the glucose unit, may compromise the methylating ability of the new molecules. If such is the case, a positively charged tetraalkyl ammonium group can be introduced into the linker in order to increase the binding affinity. On the other hand, the glucose unit may enhance the DNA binding due to favorable hydrogen bonding interactions within the groove, in which case one can expect efficient N3-A methylation. Computational methods are being employed in the laboratory to evaluate the binding of these new molecules at the desired target site. Furthermore, this thesis also describes the synthesis of new fluorescent probes designed to measure the binding of the new compounds and intermediates containing the tethered glucose unit to the desired target site.

Based on the design considerations and the factors described above, the molecules that were selected for synthesis in this particular project are shown in Figure 3.4. These two molecules, **1** and **2**, vary in the length of the tether by one CH_2 unit. These molecules will be used to investigate the effect of tethering a glucose unit to the DNA methylating Me-lex component on the ability of these molecules to methylate the DNA. Once the ability of these molecules to successfully methylate DNA is determined, these molecules can then be used in future studies to investigate their ability to target pancreatic β-cells via the GLUT-2 glucose transporter.

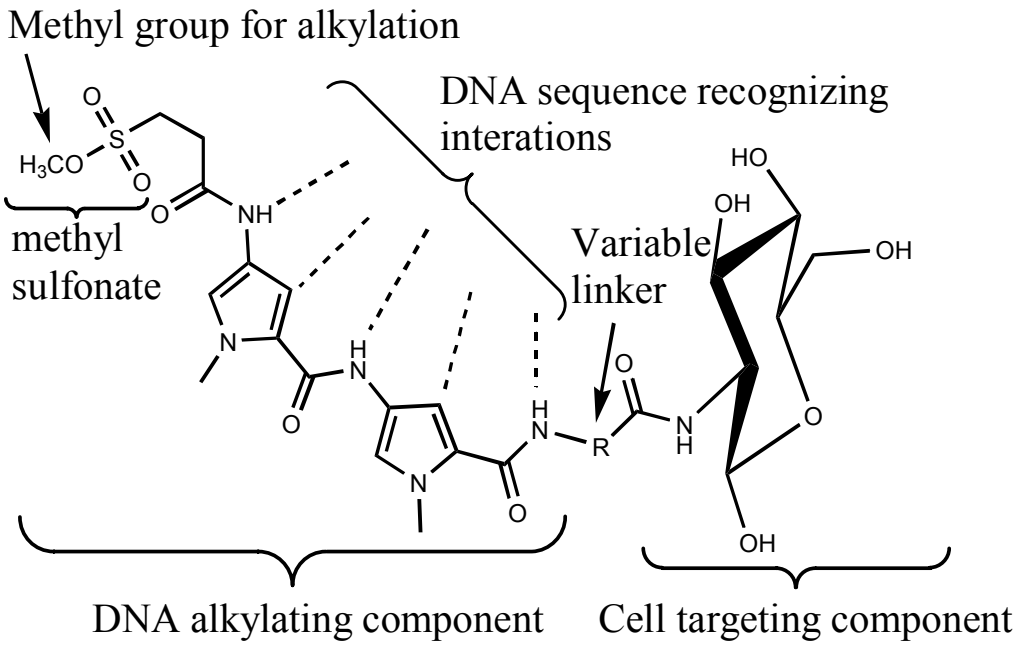


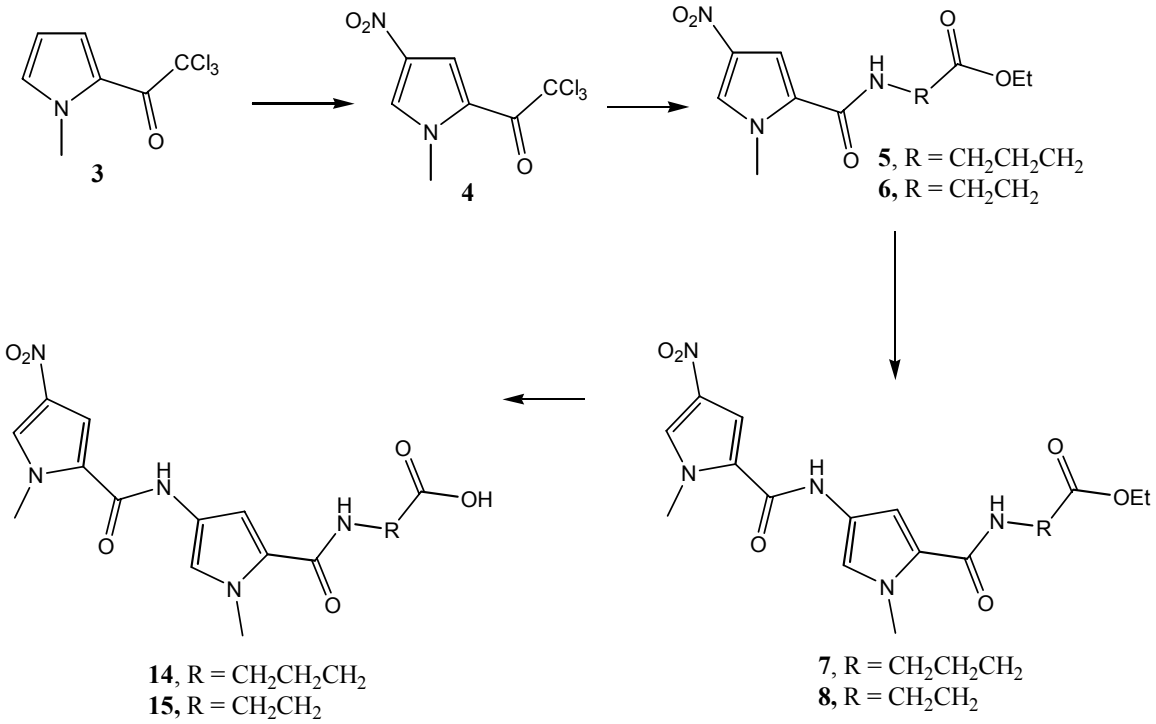
Figure 3.4. Target molecules for this project: **1**: R = $CH_2CH_2CH_2$ in, and **2**: R = CH_2CH_2.

3.2. Design of Synthetic Methodology

Compounds **1** and **2** were designed to test the effect of the variation in the length of the tether between the cell targeting glucose unit and the DNA methylating Me-lex unit by one CH_2 unit. It is expected that several variations of the linker would have to be tested before eventually identifying the optimum linker composition that provides favorable cell-targeting and DNA-methylating properties. Therefore, it is desirable to develop a synthetic methodology that allows for the easy introduction of modifications in the linker composition.

Two different approaches were taken for the synthesis of molecules **1** and **2**. The first approach was one that was similar to published procedures for similar compounds and would enable rapid completion of the target molecules. In this synthetic approach the linker is introduced very early in the overall synthesis. The second synthetic approach that was explored was one which would enable the introduction of variations in the linker at a late stage in the overall synthesis. This approach would facilitate the efficient preparation of several molecules bearing different linkers.

The first synthetic scheme that was adopted to make compounds **1** and **2** is outlined in Scheme 3.1. In this approach, the desired linker was added to the appropriate pyrrole unit, **4**, in the second step itself to yield compounds **5** and **6**. The second pyrrole unit was then added to compounds **5** and **6** followed by the hydrolysis of the ester on the linker to give **14** and **15**. These compounds can then be attached to the cell targeting glucose unit before the introduction of the reactive methyl sulfonate to form the final target molecules. The reactions outlined in Scheme 3.1 are similar to the ones followed for the synthesis of Me-lex as described in literature.[10] One disadvantage of this method



Scheme 3.1

is that the linker is added very early in the synthetic scheme. Therefore, every time a molecule with a new linker is required, the synthesis would have to be started all over right from beginning.

In order to be able to introduce different linkers into the molecules in a time efficient manor, an attempt to develop an alternative scheme in which the DNA sequence recognizing unit and the cell targeting unit would be made separately, and then assembled together at a late stage in the overall synthesis, with various linkers. After the assembly with the appropriate linker; the reactive methyl sulfonate group can be introduced in the final stages of the synthesis. A schematic diagram outlining this approach is shown in Figure 3.5. The DNA sequence recognizing unit would terminate as a carboxylic acid, linkers would be obtained as amino-esters, and the cell targeting unit would be obtained in the form of an amine. First, the DNA recognizing portion, as the carboxylic acid,

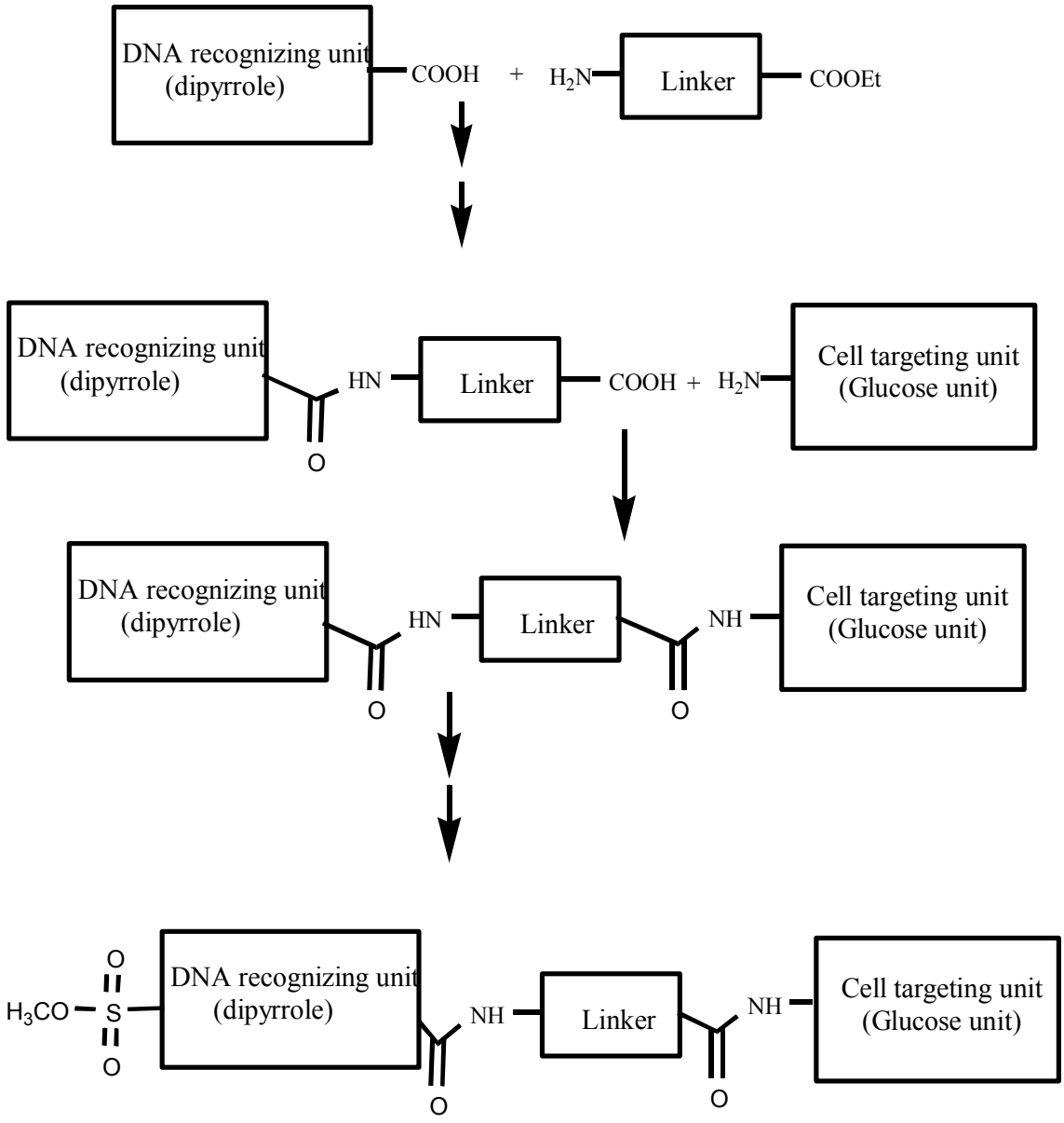
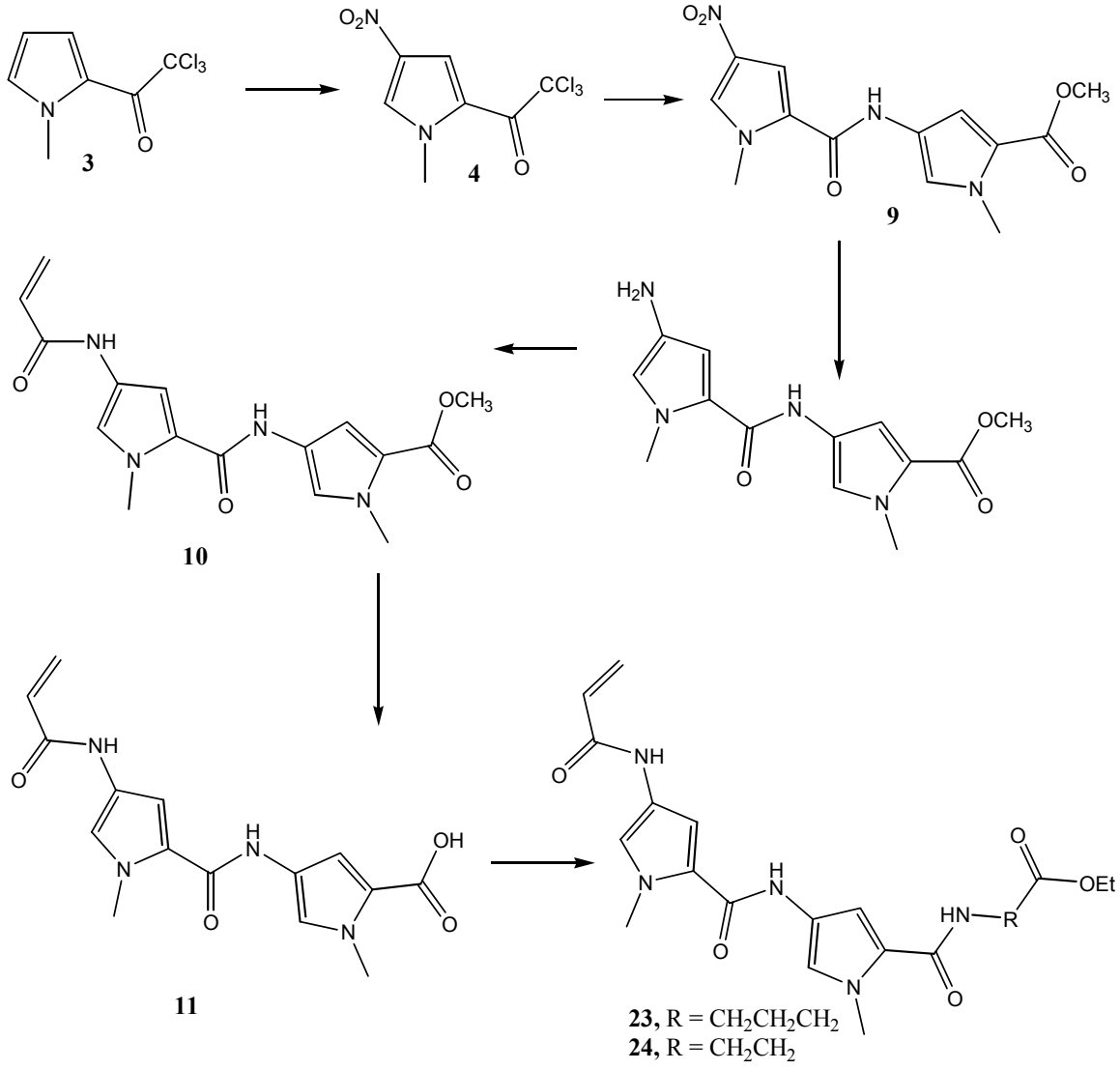


Figure 3.5. Outline of the design of the synthetic route to introduce linkers at a late stage in the synthesis.

would be condensed with the linker amine to form a new amide bond. At the other end of the linker, the ester would be hydrolyzed to the carboxylic acid and condensed with the amine on the cell targeting unit to form another amide bond. The final product containing the methyl sulfonate can then be obtained in a few steps.

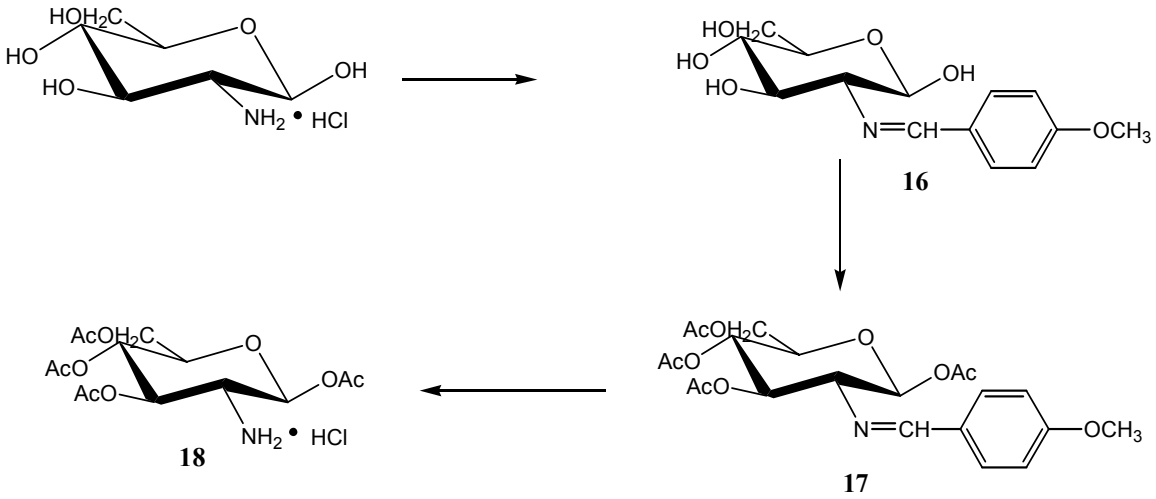
Based on this synthetic design shown in Figure 3.5, the new overall synthetic scheme that was attempted is outlined in Scheme 3.2. Starting from the same nitro



Scheme 3.2

compound, **4**, as in Scheme 3.1, the second pyrrole unit is attached to form the dipyrrole compound **9**. The reduction of the nitro to an amino, followed by condensation with acryloyl chloride gives the olefin, **10**. The ester at the other end can then be hydrolyzed to **11** and the different linkers can then be attached at this late stage in the overall synthesis to obtain compounds **23** and **24**.

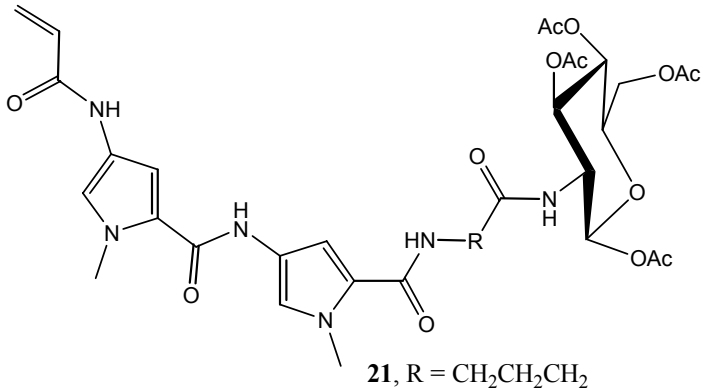
The cell-targeting glucose unit in the form of glucosamine can be attached to carboxylic acids which can be obtained by the hydrolysis of **23** and **24**. However, before this attachment reaction, the OH groups had to be protected. The protection of the OH groups on glucosamine, while maintaining a free amine was achieved by following procedures described in literature[26] as shown in Scheme 3.3. The amine of the glucosamine is first with masked with 4-methoxybenzaldehyde to give **16**. The OH groups of **16** are then protected with by reaction with acetic anhydride to give **17**. Finally the amine is and the amine is isolated and stored as the hydrochloride salt, **18**.



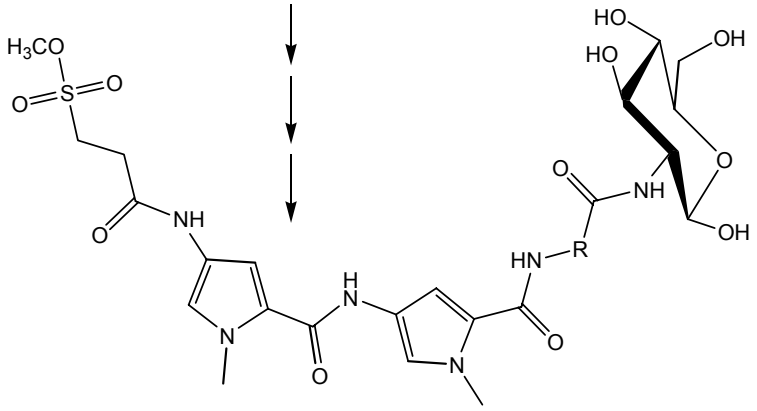
Scheme 3.3

When the protected glucosamine is attached to the linker, both the cell-targeting and the DNA sequence recognizing components are now present in the same molecule (**21** and **22**). The

OH groups on the glucose unit can then be deprotected and the DNA-methylating methyl sulfonate group can be introduced in a few steps to form the desired target molecules as shown in Scheme 3.4. The reactive methyl sulfonate has to be introduced in the least step after the deprotection of the glucose unit in order to minimize manipulations of the compound after introduction of the reactive methyl group. The glucose OH groups are not reactive to the methylating agent under the reaction conditions used to form the methyl sulfonate.



21, R = $CH_2CH_2CH_2$
22, R = CH_2CH_2

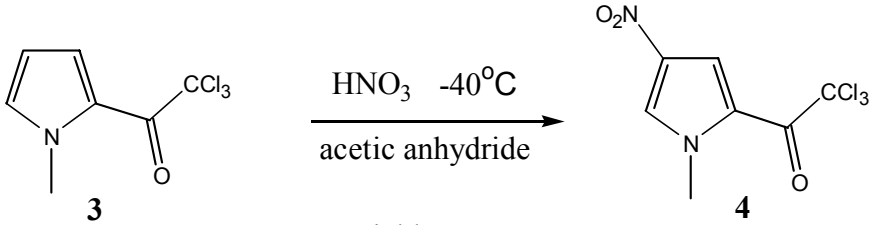


1, R = $CH_2CH_2CH_2$
2, R = CH_2CH_2

Scheme 3.4

3.3. Synthesis of the DNA-Recognizing Dipyrrole Component Following Scheme 3.1

The synthesis of the DNA recognizing dipyrrole unit, **7** or **8**, starts with the trichloroacetylation of N-methylpyrrole. This compound, **3**, has been made on a large scale in the laboratory by the reaction of trichloroacetyl chloride with N-methyl pyrrole following published procedures.[27] Compound **3** is first nitrated using fuming nitric acid with acetic anhydride as the solvent as shown in Scheme 3.5. This nitration results in



3 **4**

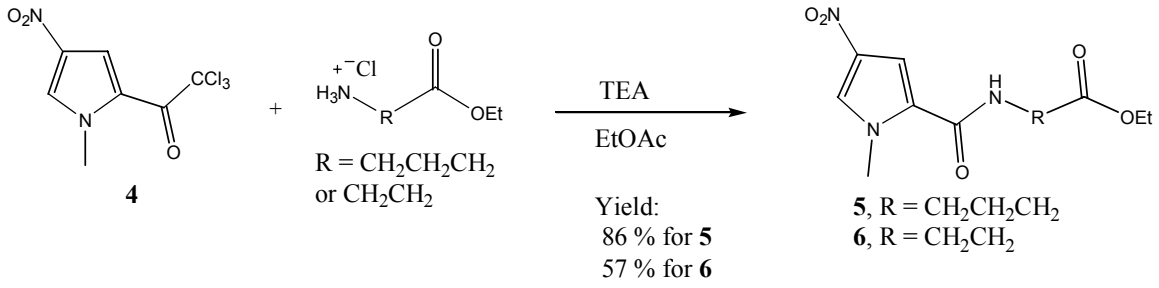
Yield: 84 %

Scheme 3.5

the formation of the desired 4-nitro product along with a small amount of the undesired 5-nitro isomer. Carrying out the reaction at low temperatures (- 40 °C), results in improving the yield of the desired 4-nitro isomer and minimizing the amount of the unwanted 5-nitro isomer formed. Upon completion of nitration, the reaction mixture is quenched with the addition of a specific amount of water and stirred overnight, which results in the hydrolysis of the acetic anhydride to acetic acid and precipitates the product. It turns out that the desired 4-nitro isomer is less soluble in water than the 5-nitro isomer and therefore, addition of an appropriate amount of water is essential for the isolation of pure product in this step. Addition of too much water results in the precipitation of both isomers. The trichloroacetyl group is stable to aqueous hydrolysis especially at room temperature, making this simple aqueous workup possible. Filtration of the precipitate through a büchner funnel then results in almost pure **4** in an 84 % yield. Occasionally small amounts (< 5 %) of the other isomer are also obtained but no attempt is made to further

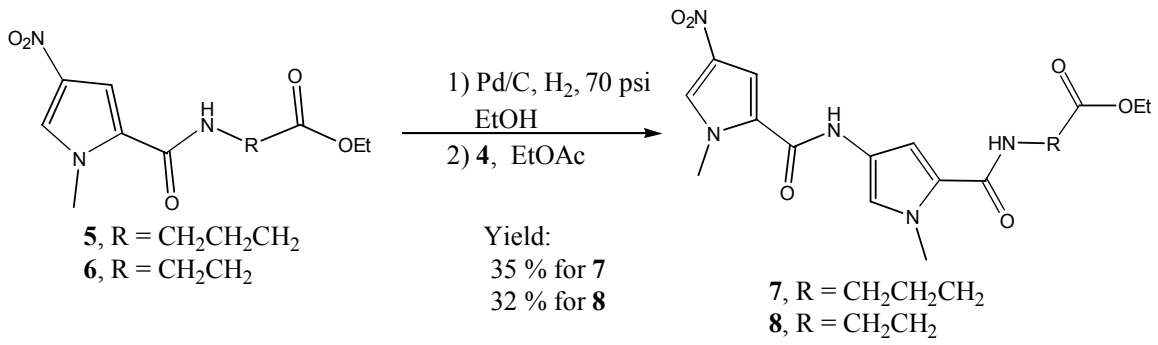
purify the product at this stage because the purification is easily accomplished in subsequent steps.

The next step of the synthesis was the addition of the linker as described above from Scheme 3.1. The nitropyrrole (**4**) was reacted with either ethyl 4-aminobutyrate hydrochloride (for the synthesis of **5**) or with ethyl β-alanine hydrochloride (for the synthesis of **6**) by stirring at room temperature in EtOAc using TEA as the HCl scavenger as shown in Scheme 3.6. These amines react readily with the trichloroacetyl groups



Scheme 3.6

eliminating chloroform, which can be easily removed by rotary evaporation. Upon completion of the reaction the, insoluble TEA hydrochloride salt can be filtered from the product. Yields were improved by bubbling Ar through the solution of EtOAc and TEA before the addition of the linker and **4** and carrying out the reaction under an Ar atmosphere. The product was purified from any unreacted starting material and any of the 5-nitro-N-methylpyrrole formed in the previous step, by flash column chromatography to give **5** in an 86 % yield and **6** in a 57 % yield. The addition of the second pyrrole unit is the next step of the synthesis and is accomplished as shown in Scheme 3.7 below. The nitro group in **5** or **6** was reduced with H_2 gas under high pressure in the presence of Pd/C catalyst. The reaction was followed by TLC and shortly after the disappearance of the nitro compound, the reaction was stopped, the catalyst

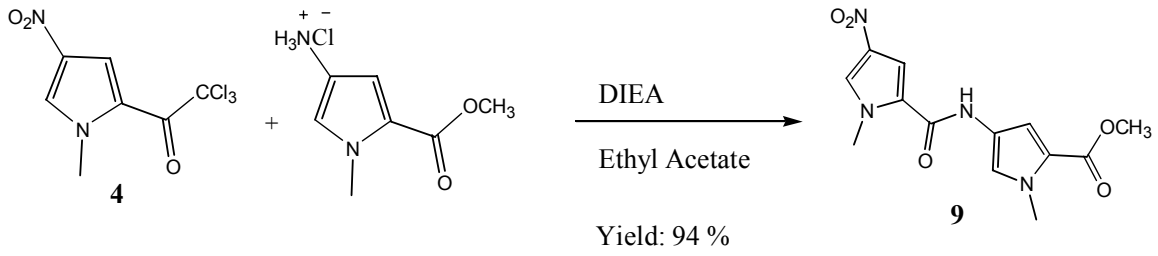


Scheme 3.7

was filtered off, the solvent was removed by rotary evaporation, and the product was evacuated under vacuum overnight. This intermediate product (presumably the amine) was reacted with another unit of **4** in EtOAc at room temperature. The desired product (**7** or **8**) fell out of solution as a yellow precipitate that was easily filtered off. The filtrate often contained some more of the product which could be precipitated and isolated upon concentration of the solution by rotary evaporation followed by cooling in the refrigerator. The filtered products, the dipeptide esters **7** and **8** were obtained in yields of 35 % and 32 % respectively.

3.4. Synthesis of the DNA-Recognizing Dipyrrole Component Following Scheme 3.2

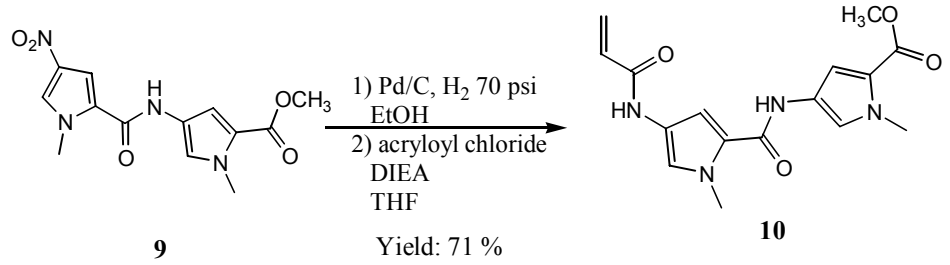
The alternate method of synthesis outlined in Scheme 3.2, which was developed in order to introduce the linker unit at a later stage, also starts with the same compound **3** which is nitrated to **4** as described above. Compound **4** is condensed with commercially available methyl-4-amino-1-methyl-1*H*-pyrrole-2-carboxylate, hydrochloride salt, to form the dipyrrole unit **9** as shown in Scheme 3.8. The reaction is carried out in EtOAc in the presence of DIEA which scavenges the HCl from 4-amino-1-methyl-1*H*-pyrrole-2-carboxylate, hydrochloride salt to form the free amine. The formation of **9** results in a



Scheme 3.8

yellow precipitate that can be filtered. The product has to be washed with water to remove the DIEA hydrochloride salt and dried to give **9** in a 94 % yield.

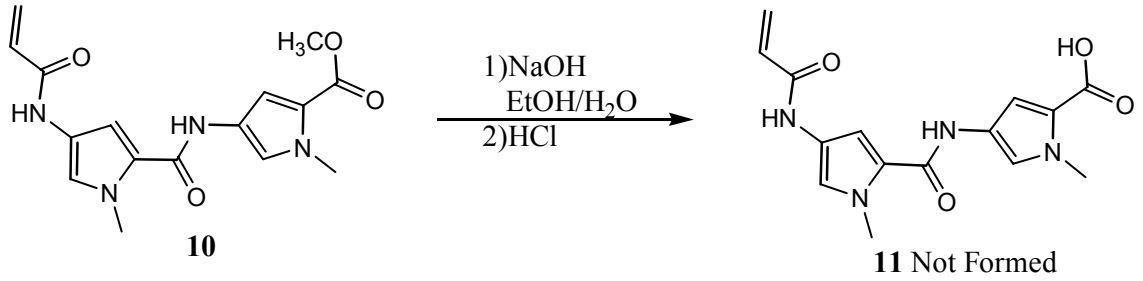
Nitro compound **9** was converted into the olefin by reducing the nitro group into the amine and then reacting the amine with acryloyl chloride in the presence of DIEA in THF as shown in Scheme 3.9. The DIEA scavenged the HCl that was formed



Scheme 3.9

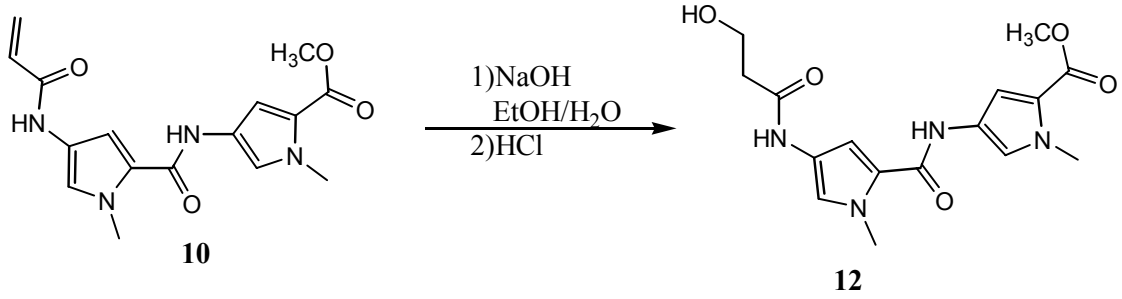
upon the reaction of the amine and the acryloyl chloride. Flash column chromatography was used to purify **10** which was obtained in a 71 % yield.

Once the methyl ester on **10** is hydrolyzed to the carboxylic acid, various linkers can be added. However, basic hydrolysis of **10** did not result in the desired olefin carboxylic acid **11** as shown in scheme 3.10.

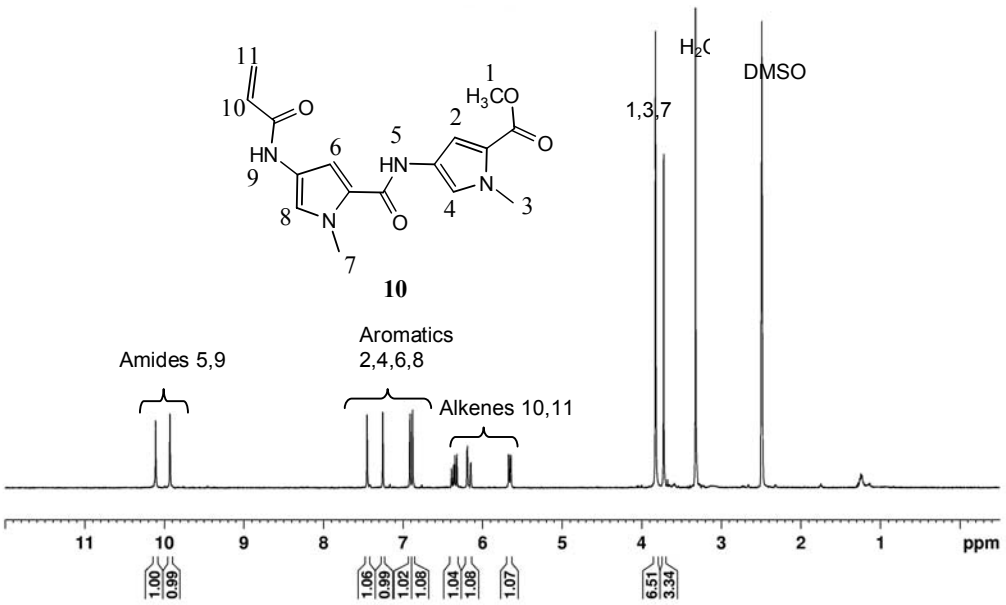


Scheme 3.10

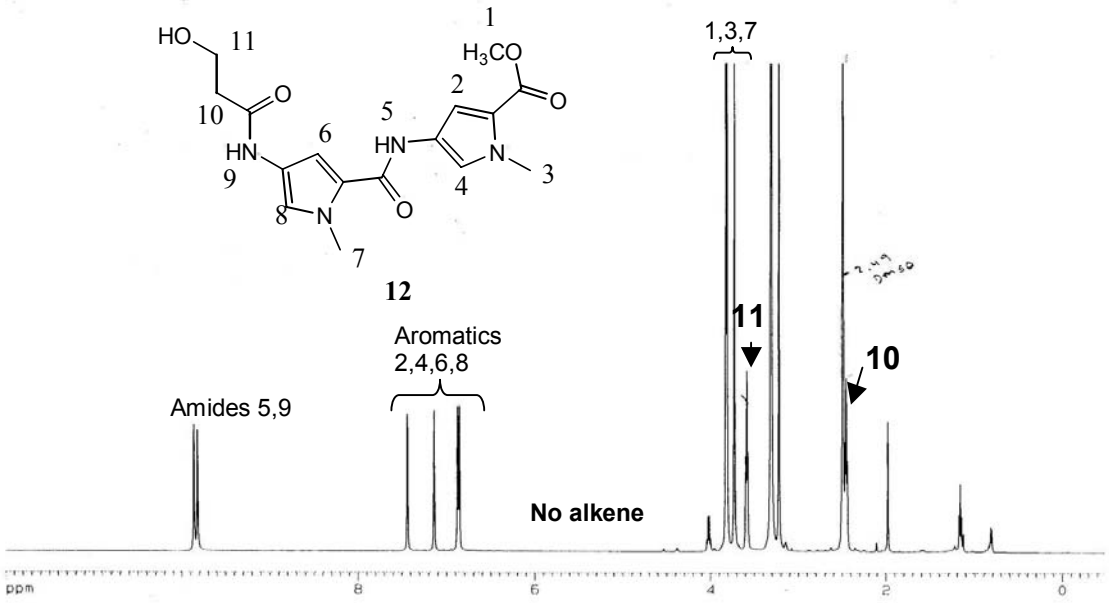
NMR analysis was used to determine the identity of the product obtained. The ^{1}H NMR spectra of the starting compound **10** and the product of the hydrolysis reaction are shown in Figures 3.6 and 3.7 respectively. Some of the key features in the ^{1}H NMR spectrum of **10**, as can be seen from Figure 3.6 a) are the two amide hydrogens around 10 ppm, the 3 olefinic hydrogens between 5.5-6.5 ppm, and the 3 methyl singlets between 3.5-4.0 ppm. In the ^{1}H NMR spectrum of the product obtained upon base hydrolysis of **10** (see Figure 3.6 b) the 3 methyl singlets are still present indicating that the ester has not been hydrolyzed. Instead, the olefinic hydrogens have disappeared. It is unlikely that the amide bond next to the olefin has been hydrolyzed off since both amide hydrogen peaks are still present at 10 ppm. It is therefore likely that a Michael addition of a hydroxide has taken place at the alkene resulting in the formation of compound **12** (Scheme 3.11).



Scheme 3.11



a)



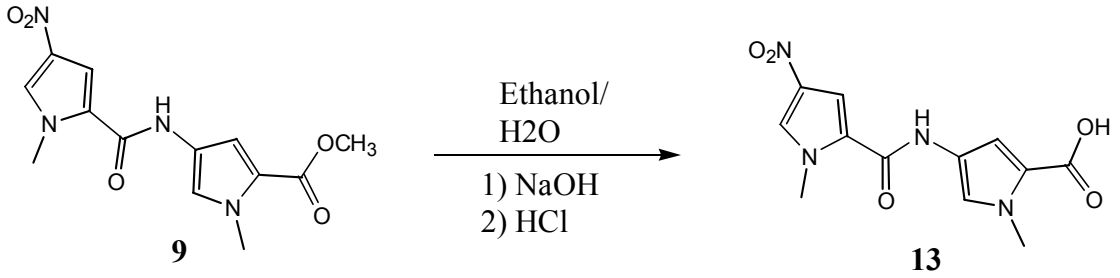
b)

Figure 3.6. ^{1}H NMR of a) **10** and b) products of reaction of **10** with NaOH.

This hypothesis is strengthened by the observation of two new peaks (labeled 10 and 11 in the spectrum) seen at 2.4 and 3.5 ppm which could represent the two new CH_2 hydrogens. No further attempt was made to identify compound **12**.

Further base hydrolysis of compound **12** resulted in the loss of one of the methyl groups indicating that the ester can be hydrolyzed under more rigorous conditions. However, in spite of attempting the hydrolysis of **10** under numerous different conditions with different reagents, the ester was unable to be hydrolyzed while leaving the alkene intact. Therefore, it was decided to add the linker and the targeting glucose unit before the introduction of the olefin in order to avoid all hydrolysis reactions after the addition of the olefin.

The methyl ester of **9** was then converted to the carboxylic acid with NaOH in EtOH and H_2O under reflux and then acidified and filtered to give **13** in yields close to 90 % as shown in Scheme 3.12. The carboxylic acid that is formed is yellow, but after drying on the filter paper the product often appears to be a dark brown solid.



Yield: 87 %

Scheme 3.12

The linkers can now be added to this nitro dipeptide carboxylic acid unit by standard peptide coupling methods. The coupling reagents that were employed here were EDCI and HOBT. EDCI is a water soluble carbodiimide that makes purification steps simple because it can be removed by liquid-liquid extraction. Figure 3.7 shows the

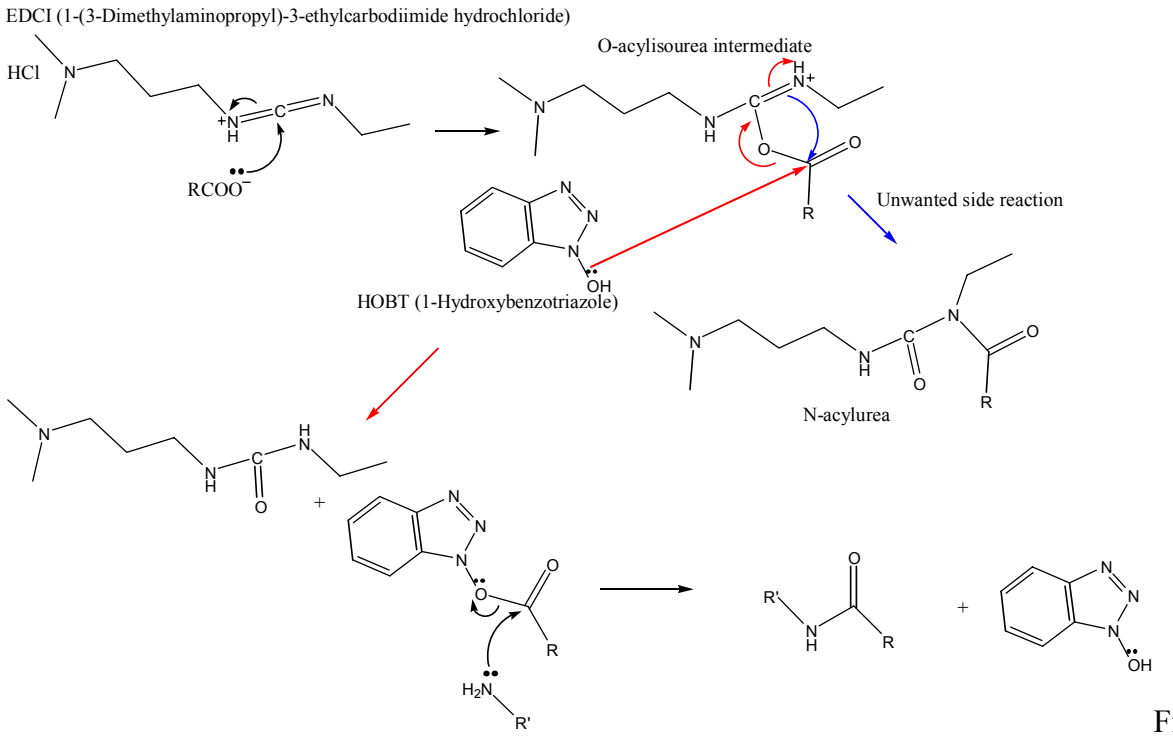
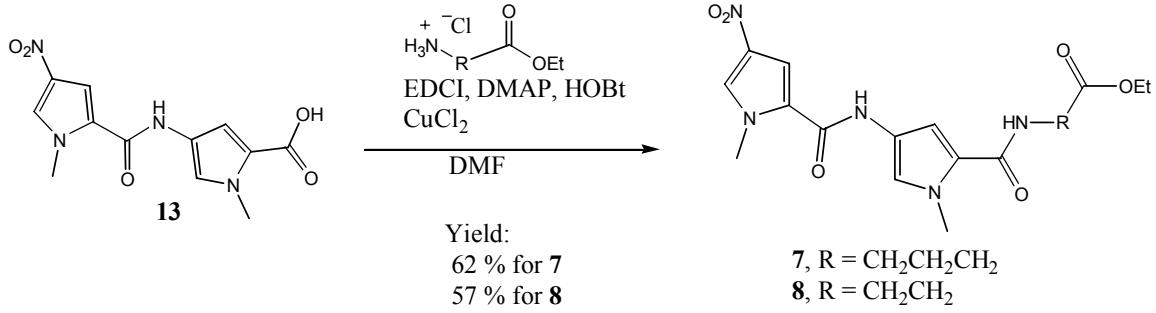


Figure 3.7. The mechanism by which EDCI and HOBT work to form an amide bond.

mechanism of action of EDCI and HOBT. The EDCI activates the carboxylic acid by forming the O-acylisourea intermediate.[28]

One complication that is typically seen which reduces yields in these reactions is the rearrangement of the O-acylisourea intermediate to give an N-acylurea that is then unreactive and cannot be used form the amide bond.[28] Use of HOBT prevents this from occurring. HOBT reacts with the O-acylisourea to form another intermediate, which does not undergo rearrangement and reacts with the amine to form an amide bond.[28-29] It has also been reported in literature that the use of $CuCl_2$ as a catalyst further reduces the occurrence of unwanted rearrangement reactions,[28-29] though the mechanism by which $CuCl_2$ achieves this is not clearly understood.

The two linkers were added to the dipeptide carboxylic acid **13** as shown in Scheme 3.13. Upon formation of the product as indicated by TLC the reaction mixture

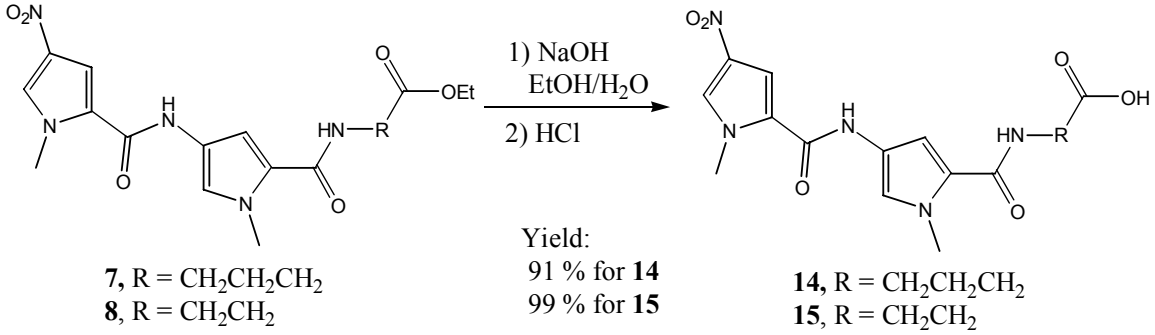


Scheme 3.13

was diluted with DCM, extracted with dilute base and dilute acid. Upon cooling the DCM in the freezer overnight the products crystallized out of solution to give **7** in a 62 % yield and **8** in a 57 % yield.

The carboxylate esters on the linkers in **7** and **8** had to be hydrolyzed to carboxylic acids before they could be coupled to the cell targeting unit. This hydrolysis is very efficiently

achieved using NaOH in EtOH/H_2O followed by acidification with yields than 90 % resulting in the formation of compounds **14** and **15** as shown in Scheme 3.14.



7, R = $CH_2CH_2CH_2$
8, R = CH_2CH_2

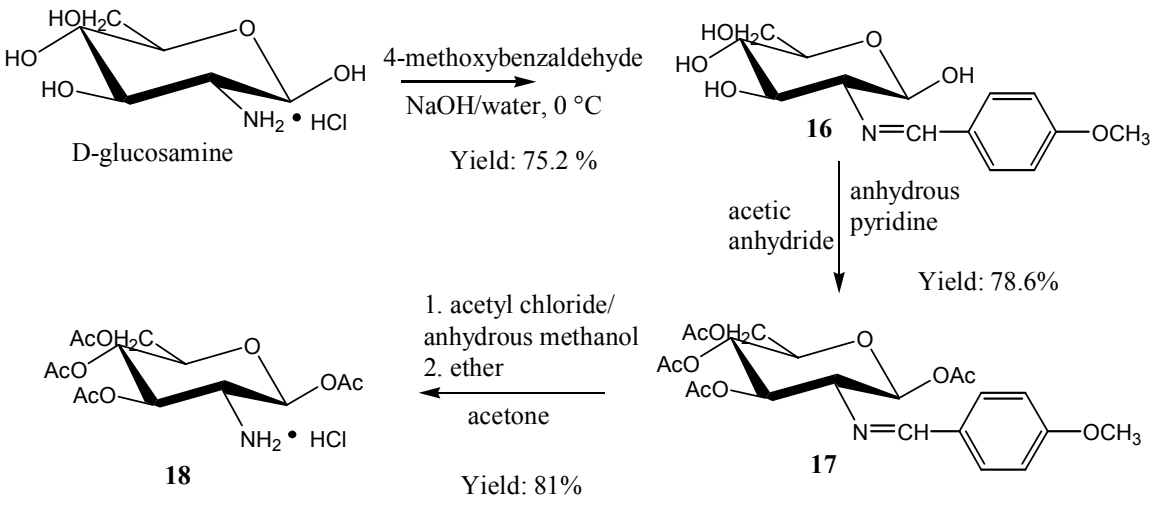
14, R = $CH_2CH_2CH_2$
15, R = CH_2CH_2

Scheme 3.14

Thus compounds **14** and **15** can be obtained following either Scheme 3.1 or 3.2. However, Scheme 3.2 has the advantage that the linker is added on at a later stage. Due to complications resulting from the Michael addition at the alkene in **10** during base hydrolysis, the linker had to be added at an earlier step than originally planned. Irrespective of the linker that is to be used, compound **13** can now be made on a large scale and then stored and used when needed with various linkers. Other approaches are being explored in the laboratory in which the linker can be added further down the line in the synthesis.

3.5. Synthesis of the Cell Targeting Glucose Unit

The targeting unit is introduced as D-glucosamine, which has both reactive hydroxyl groups and the reactive amine. The amine will be used in the formation of the amide bond that will link the glucose unit to compounds **14** and **15**. In order to prevent interference from the hydroxyl groups, they have to be protected while leaving the amine deprotected. This was accomplished as outlined below in Scheme 3.15 following



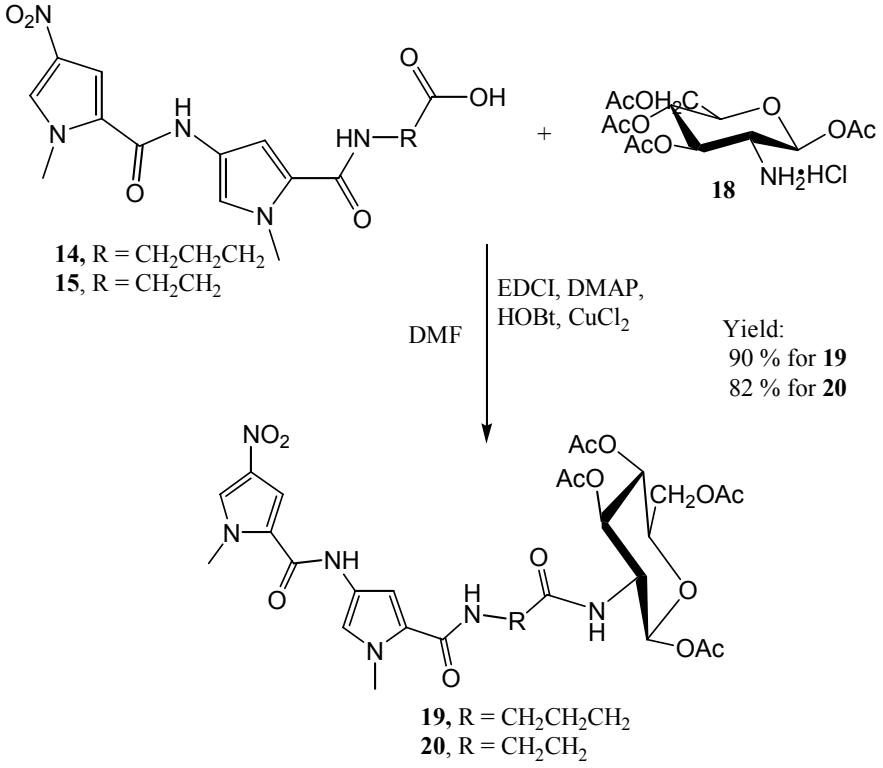
Scheme 3.15

procedures described in literature.[26] First, the amine of D-glucosamine was masked as 4-methoxybenzylidene using 4-methoxybenzaldehyde with NaOH/water at 0°C. The product falls out of the reactions mixture as a white precipitate upon cooling and can be easily filtered to give **16** in a 75 % yield. The OH groups were then protected with acetyl groups using acetic anhydride in anhydrous pyridine. Quenching the reaction with ice water causes the product to fall out as a white precipitate which can be filtered out to give **17** in a 79 % yield. Finally, the amine was unmasked by reacting **17** with acetyl chloride dissolved in anhydrous methanol using acetone as the solvent. Upon the addition of ether at the end of the reaction, a white precipitate was formed which was filtered out to give the protected glucosamine hydrochloride, **18**, in an 81 % yield. The overall yield of the three steps was 47.9%.

3.6. Assembly of the DNA Recognizing Unit with the Cell Targeting Glucose Unit

The protected glucosamine (**18**) from above can be coupled with carboxylic acid termini of the various linkers. The coupling reaction is carried out similar to the one described earlier for preparing **7** and **8** (Scheme 3.13). The protected glucosamine was reacted with the carboxylic

acid (**14** or **15**) in the presence of EDCI, HOBT, DMAP, and $CuCl_2$ in DMF as shown in Scheme 3.16. The products were isolated by crystallizing

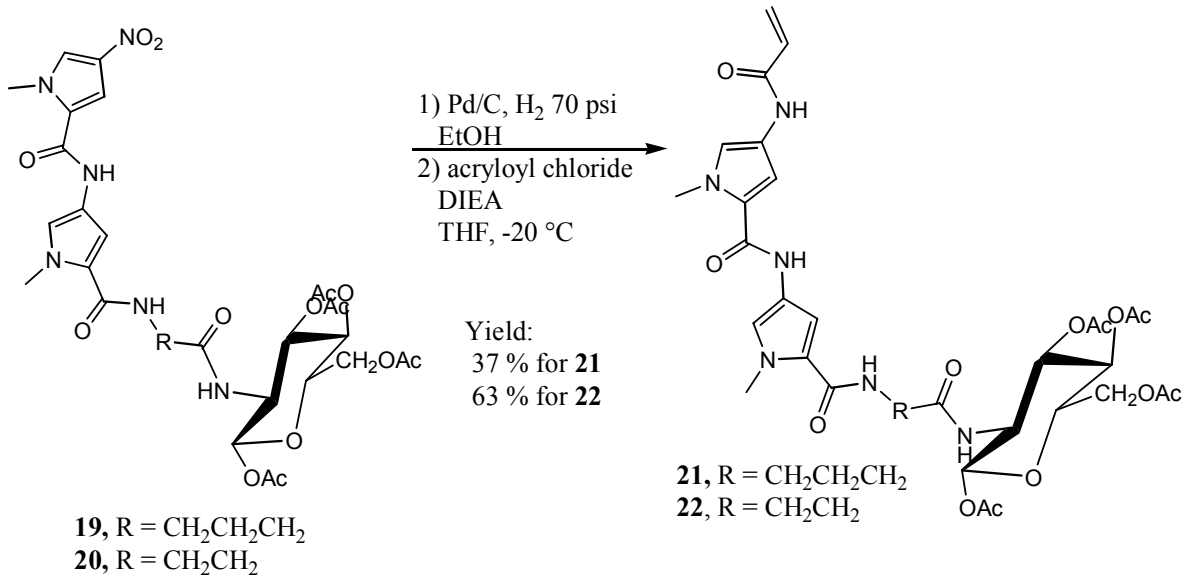


Scheme 3.16

from DCM as described earlier. The products **19** and **20** were obtained in very high yields (90 % and 82 % respectively.)

Compounds **19** and **20** incorporate both a DNA recognizing unit, which can recognize the A/T rich sequences on DNA, and a cell targeting unit (glucose unit), which upon deprotection of the hydroxyl groups should be able to target pancreatic β-cells through the GLUT-2 transporters. The next stage of the synthesis is to functionalize the N-terminus of the pyrrole dipeptide unit as the methyl sulfonate, which will serve as the DNA methylating agent (see Scheme 3.4).

The nitro group of **19** and **20** was reduced to the amine and reacted with acryloyl chloride in the presence of DIEA in THF as shown in Scheme 3.17 to form the olefin

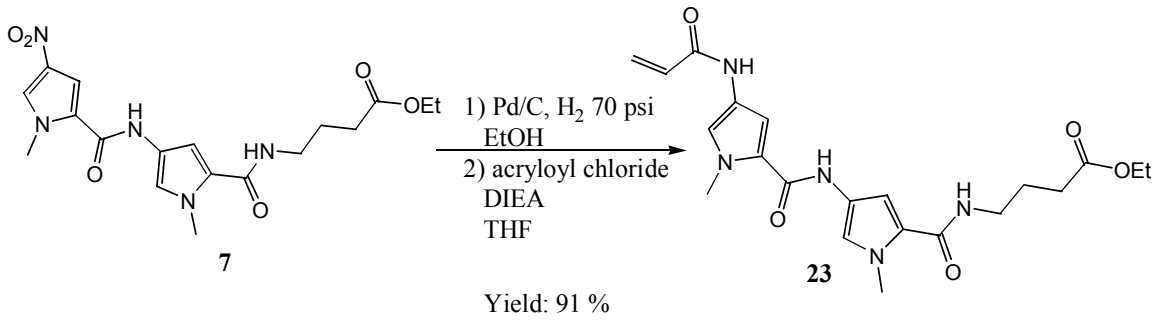


Scheme 3.17

products **21** and **22**. Thus the olefinic functionality, which underwent Michael addition if introduced earlier, has now been attached to the molecules. The products were then either crystallized out of the DCM or purified by flash column chromatography was to obtain **21** (37 %) and **22** (63 %).

These reactions of converting the nitro into the alkene typically give much higher yields in the absence of the sugar unit. For example, compound **7** when reacted under the same conditions as described above, gives the alkene compound **23** in a 91 % yield (Scheme 3.18). Similarly **9** can be converted into **10** in a 71 % yield (see Scheme 3.9).

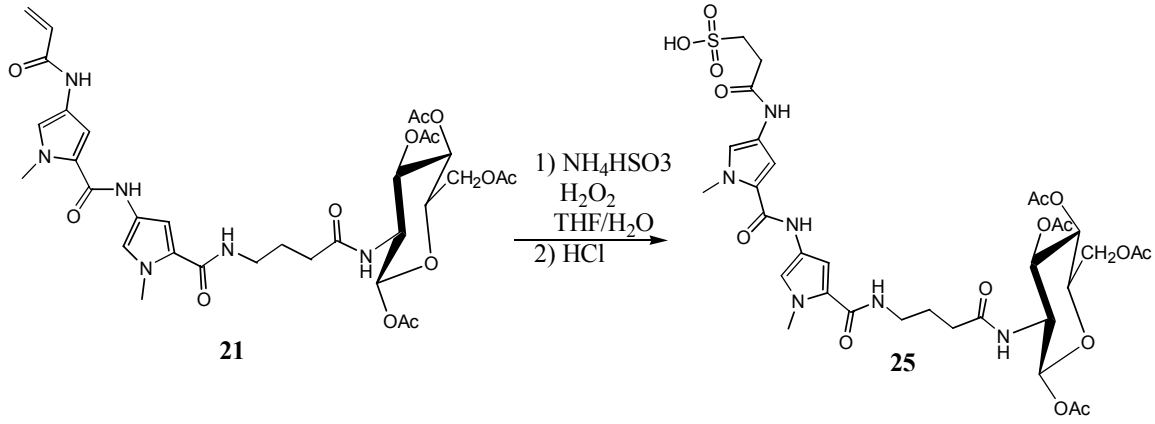
However, with the sugar attached, the yields for these reactions were much lower. These reactions are being further optimized in the laboratory.



Scheme 3.18

3.7. Introduction of the DNA Alkylating Methyl Sulfonate Group

The next step in the synthesis was to convert the alkene double bond into the sulfonic acid, and this is accomplished by the anti-Markovnikov addition of bisulfite (NH_4HSO_3) across the alkene double bond. This reaction is believed to proceed by a radical mechanism. Compound **21** was reacted with NH_4HSO_3 and H_2O_2 in 25 % THF/H_2O under reflux and then acidified as shown in Scheme 3.19. After flash column

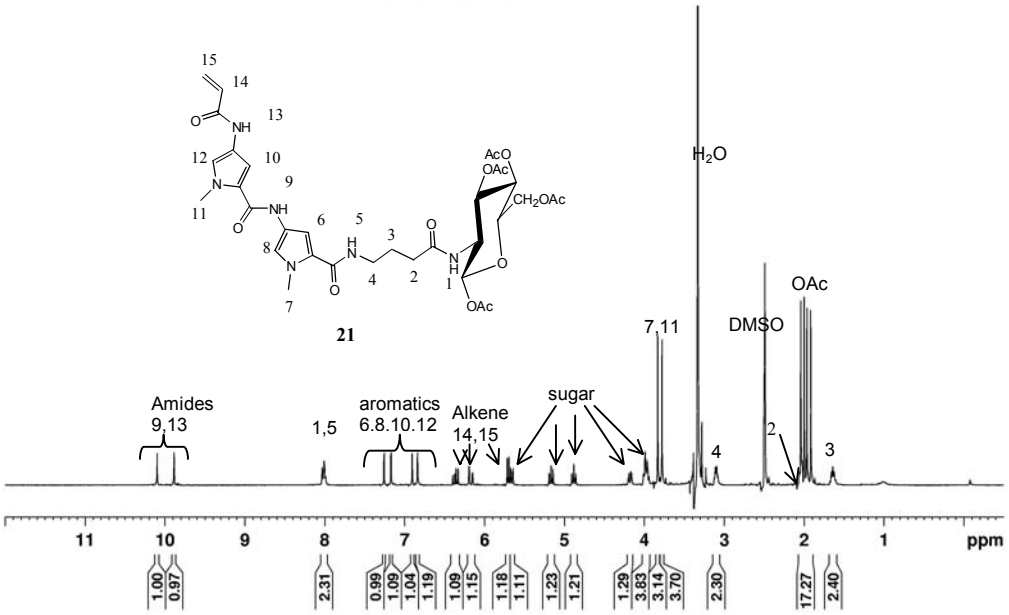


Scheme 3.19

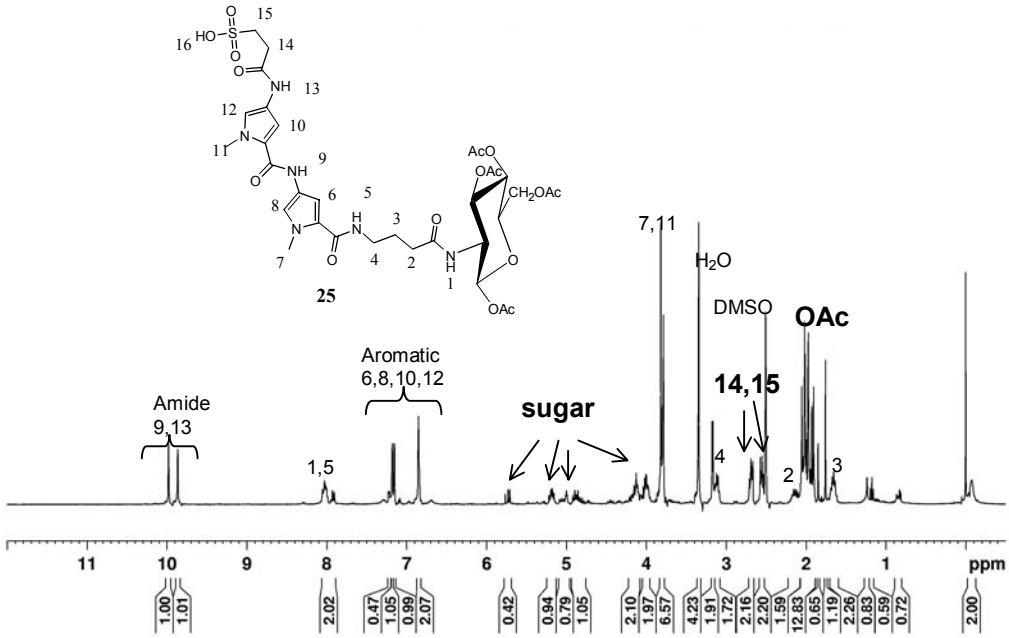
chromatography, a product was isolated as a yellow colored solid. Earlier procedures described in literature used DMF as a solvent.[10] However, the isolation of the product is more difficult when DMF is used and employing THF as solvent makes the purification and isolation of the product simpler.

The ^{1}H NMR spectra of **21** and the product of the reaction are shown in Figure 3.8 a) and b) respectively with all the key characteristics identified. The three olefinic hydrogens (5.7-6.3 ppm), the four different acetyl methyl hydrogens (1.9-2.0 ppm), the two aromatic amide hydrogens (9.8-10.0 ppm), and the two aliphatic amide hydrogens (8.0 ppm) can be clearly identified in the spectrum of **21** (Figure 3.8 a)) in which all the peaks are well resolved. The ^{1}H NMR spectrum of the product (Figure 3.8 b)) shows the disappearance of the alkene hydrogen peaks and the characteristic shifts of the CH_2CH_2 hydrogens next to the sulfonic acid between 2.6-2.8 ppm. However, in this spectrum there are more peaks than expected in the region where the acetyl methyl hydrogens of the protecting groups show up (1.8-2.2 ppm) and the sugar hydrogens are not well resolved. It is therefore possible that there is some level of deprotection of the acetyl groups, or epimerization of the sugar unit resulting in a mixture of compounds in the product that was isolated. This may have been caused by the treatment with concentrated HCl used in the acidification of the sulfonate salt. Therefore, the identity of **25** is yet to be confirmed and further purification of the product is also being attempted.

Since it seemed possible that some deprotection of the hydroxyl groups on the glucose unit was taking place possibly due to the treatment with strong acid, it was decided to attempt the deprotection reaction directly on the sulfonate salt before acidification. This deprotection can be easily accomplished by treatment with 7 N NH_3 in methanol, a procedure that was tested with compound **19** as shown in Scheme 3.20.

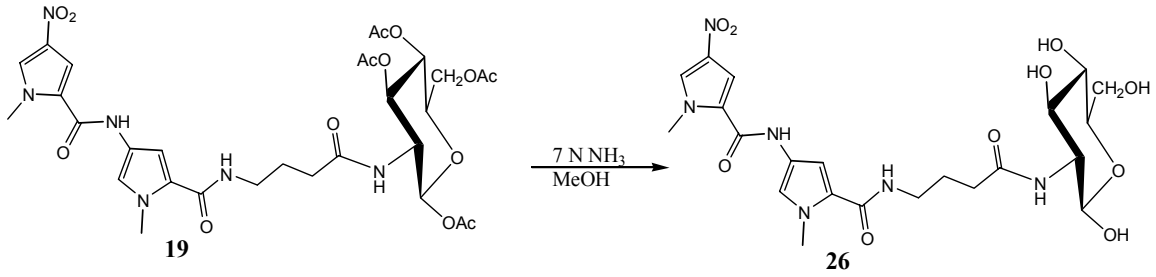


a)



b)

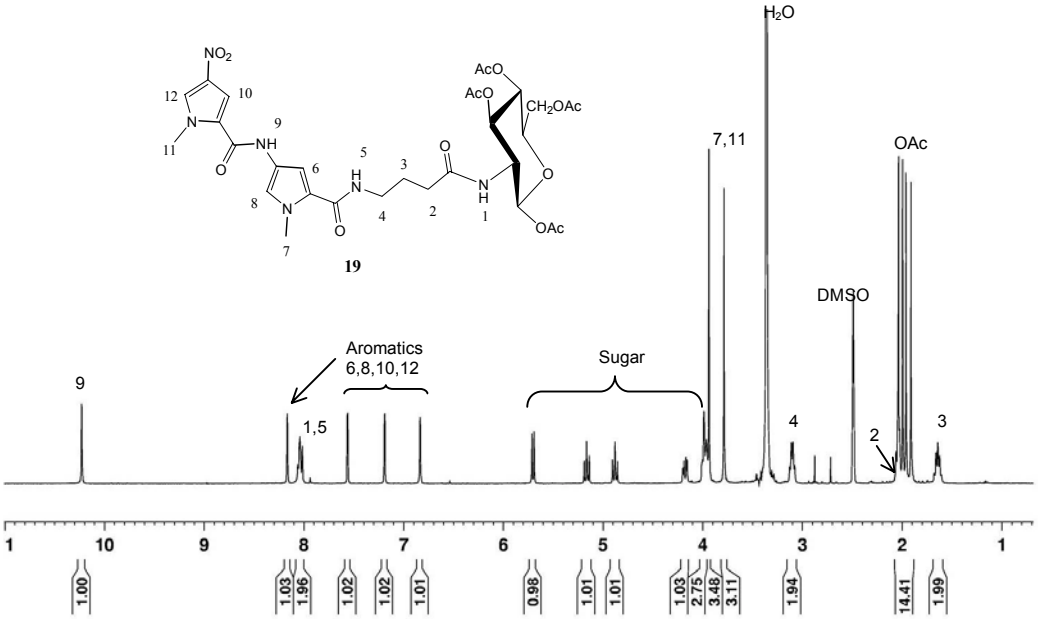
Figure 3.8. ^{1}H NMR of a) **21** and b) product of reaction of **21** with NH_4HSO_3 and H_2O_2.



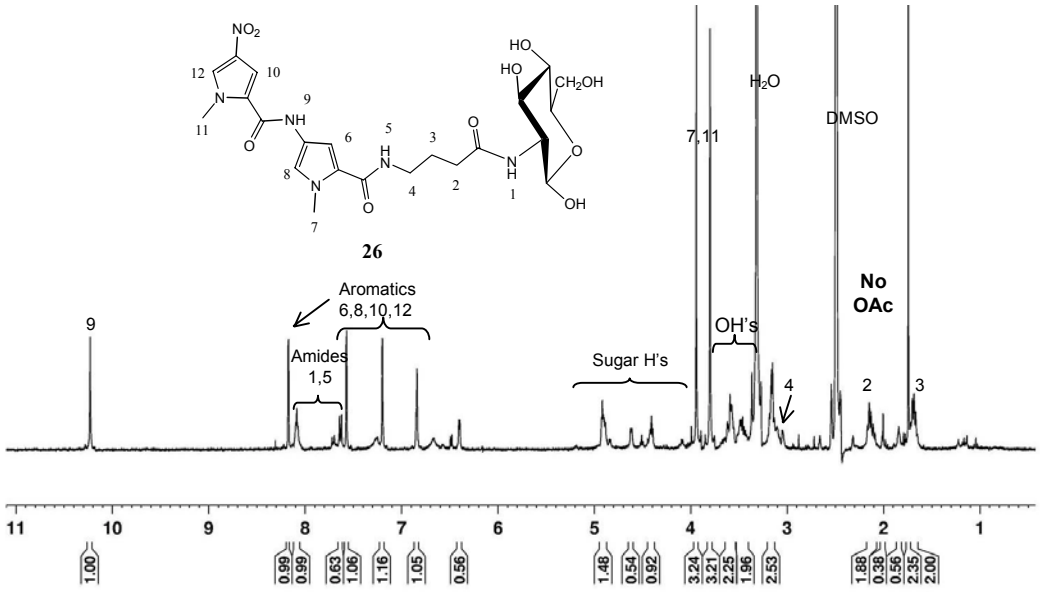
Scheme 3.20

The ammonia treatment of this compound results in the complete deprotection of the glucose unit and formation of **26**, as can be verified by a comparison of the ^{1}H NMR spectra of the two compounds **19** and **26** shown in Figures 3.9 a) and b) respectively. However, when this treatment was applied to the product of the reaction shown in Scheme 3.19, the expected well resolved ^{1}H NMR spectrum was not obtained. This procedure is still under investigation in the lab.

An alternate procedure for the conversion of an alkene to the sulfonic acid, using $NaHSO_3$ in $EtOH/H_2O$ at a basic pH, which is described in literature was tested.[30] Using this procedure, **23** can be converted to the sulfonic acid, **27**, as shown in Scheme 3.21 in a yield of 89 %. Surprisingly, under these basic reaction conditions, no Michael addition product is obtained. The use of these reaction conditions on compound **21** can, in principle, result in the conversion of the alkene into the sulfonic acid and the removal of

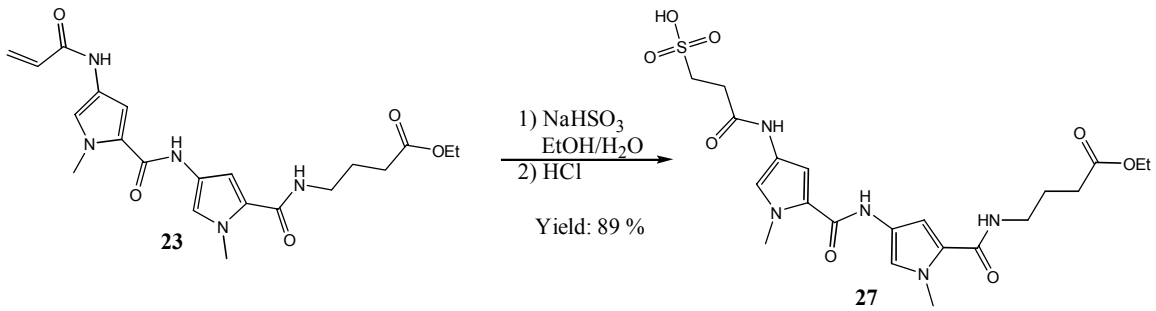


a)



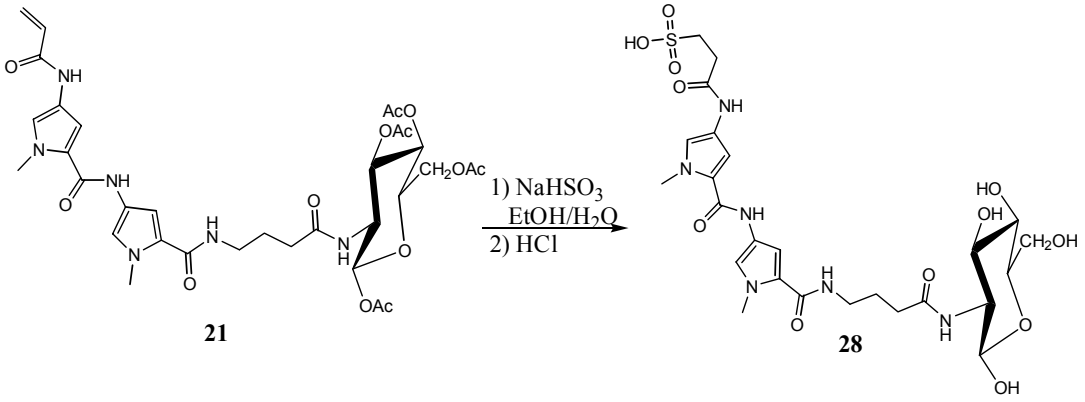
b)

Figure 3.9. ^{1}H NMR spectrum of a) **19** and b) **26**



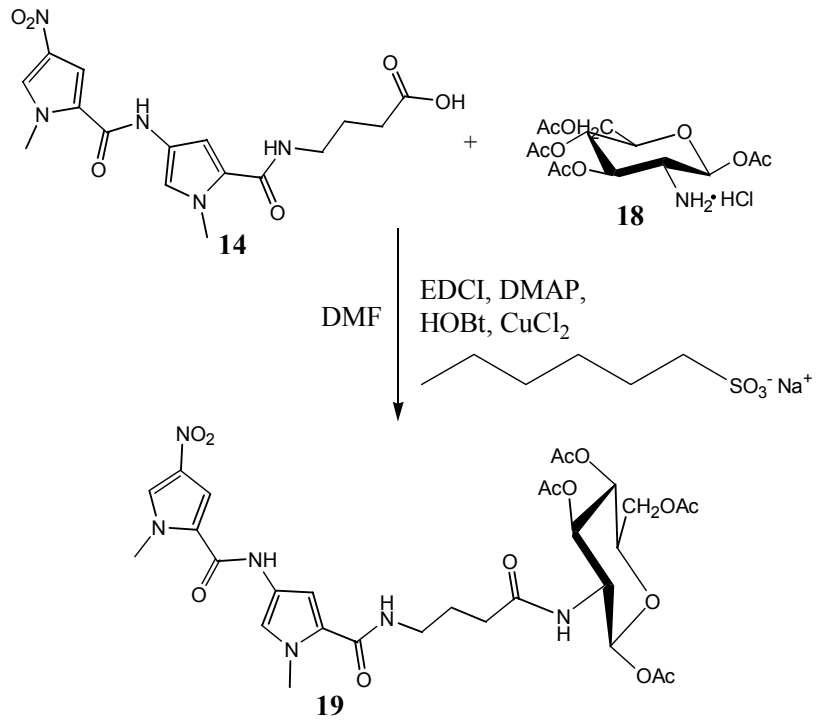
Scheme 3.21

the acetyl protecting groups on the glucose unit in a single step as shown in Scheme 3.22. This reaction is under investigation in the laboratory.



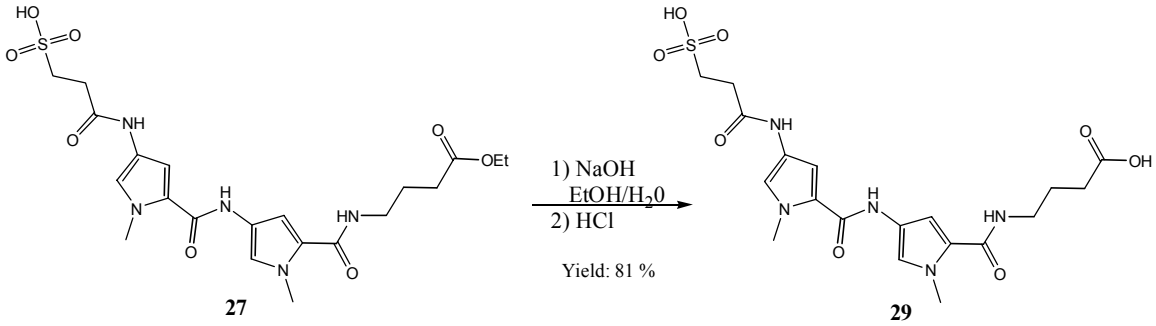
Scheme 3.22

There is evidence in literature that compounds containing both a sulfonic acid and a carboxylic acid can be reacted with amines under certain conditions which can result in the formation of an amide bond without complications due to the presence of the sulfonic acid.[31] This reaction was verified by coupling carboxylic acid **14** with **18** in the presence of equivalent amounts of the sodium salt of hexane sulfonic acid as shown in Scheme 3.23. This reaction resulted in the formation of **19** in excellent yields. Therefore, ester



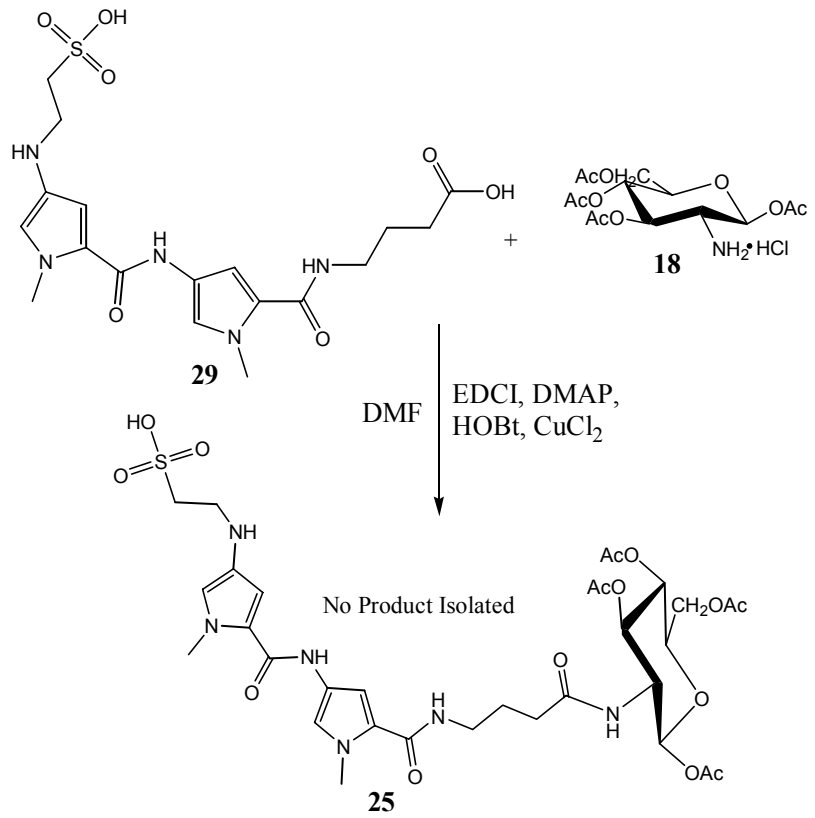
Scheme 3.23

27 was hydrolyzed to the carboxylic acid as shown in Scheme 3.24 in an 81 % yield.



Scheme 3.24

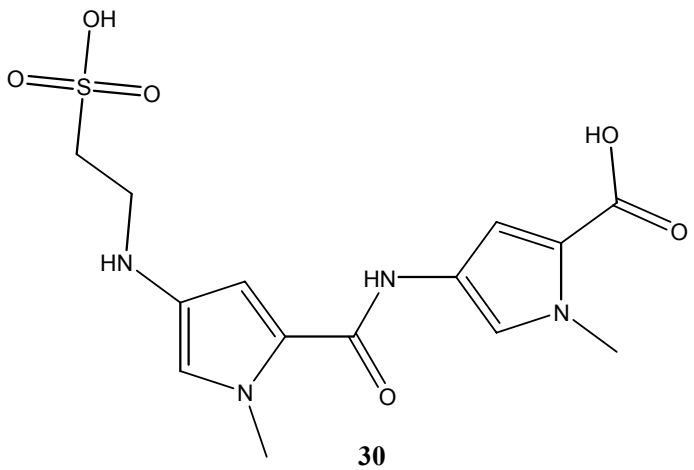
Subsequently, one equivalent of NaOH was added to the sulfonic acid carboxylic acid **29**, and the coupling reaction with the protected glucosamine **18** was attempted as outlined in Scheme 3.25. The initial attempt was inconclusive and this reaction is being further explored in the laboratory.



Scheme 3.25

If the coupling of **18** to a sulfonic acid carboxylic acid can be successfully accomplished, it would offer a tremendous advantage for the overall synthesis since compound **30** (Figure 3.10) could be prepared which could then be coupled with the appropriate linker when needed. The ester on the linker could then be hydrolyzed and coupled with glucosamine, this making the overall synthesis of molecules with several linkers very efficient.

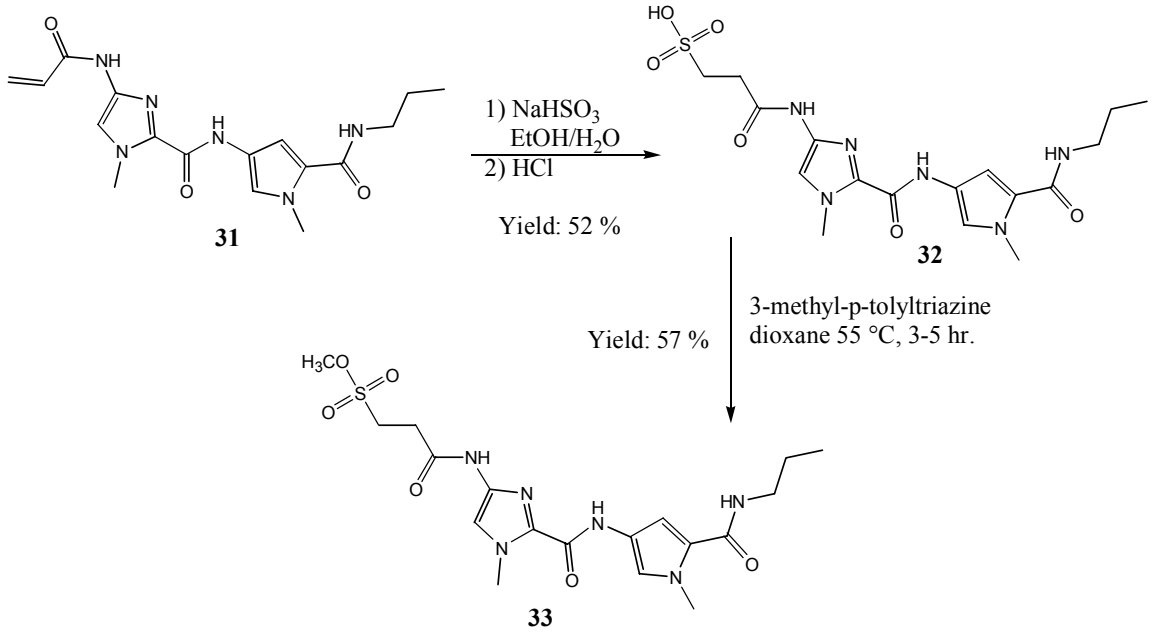
Once the desired sulfonic acid compound (**28**) with the deprotected glucose unit is prepared by one of the above methods, the reactive methyl group has to be introduced. The sulfonic acid is typically converted to the methyl sulfonate by treatment with 3-methyl-p-tolyltriazine in anhydrous dioxane. For example compound **31** (with one imidazole and one pyrrole ring) to compound **33** following procedures outlined in



30

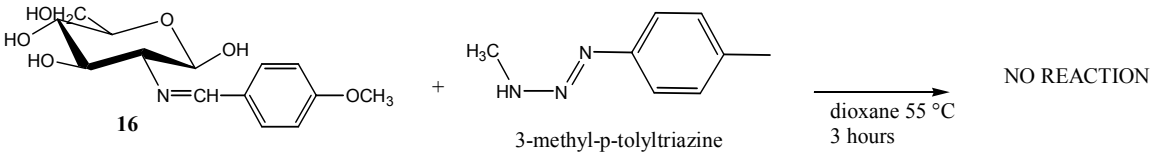
Figure 3.10. Structure of **30**.

Scheme 3.26. This reaction gives the product **33** in a 57 % yield after purification by flash column chromatography.



Scheme 3.26

. In order to verify that the hydroxyl groups on the glucose unit in the compounds are unaffected under these reaction conditions, the reaction was attempted on a test compound **16** as shown in Scheme 3.27. No reaction was obtained in this case and



Scheme 3.27

compound **16** was isolated intact from this reaction. This indicated that the OH groups on the glucose unit would not get methylated under the conditions used to convert the sulfonic acid to the methyl sulfonate.

The synthesis is two steps away from completion and will then be ready for testing with DNA to determine the methylation ability of the compounds on the DNA. Promising

possibilities for introducing the linker at a late stage in the overall synthesis is being further investigated to find the best way to construct these types of compounds with different linkers attached.

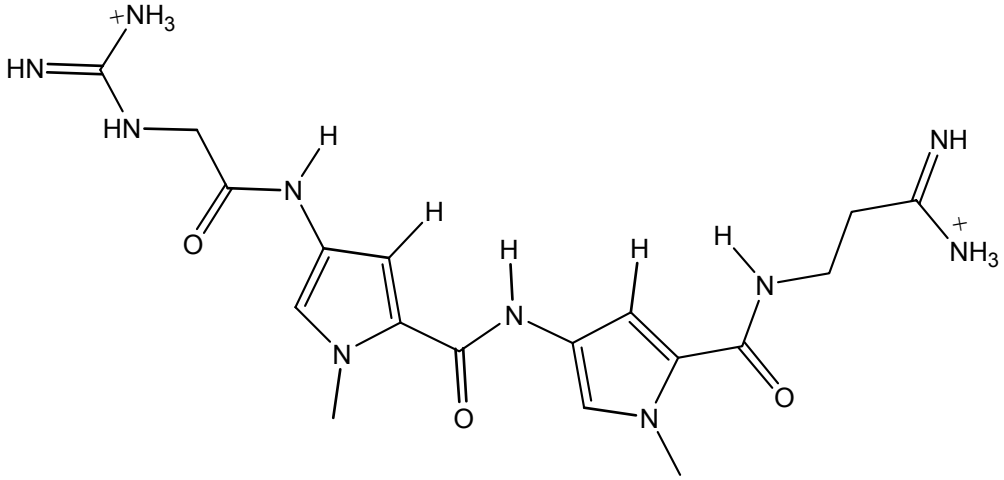
CHAPTER 4. DESIGN, SYNTHESIS, AND CHARACTERIZATION OF NEW DNA MINOR GROOVE BINDING FLUORESCENT PROBES

4.1. Design

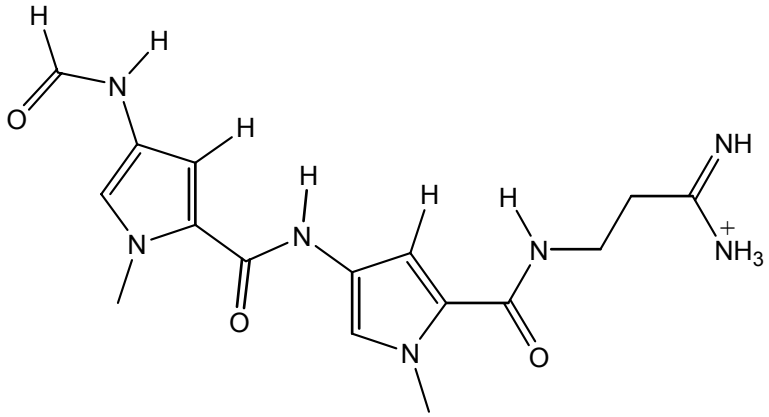
The binding of the Me-lex component of the new molecules to the DNA minor groove at A/T rich regions is expected to be critical for their ability to methylate DNA. The new compounds are, by design, weak DNA binders so that upon transferring the methyl group to DNA (which results in a negative charge on the molecule) they will depart without further interacting with the DNA. This would ensure that the biological outcomes observed are due to the methylation alone, and not due to other interactions between the molecules and DNA. However, the strength of binding will be directly correlated to the levels of methylation observed in these concerted alkylation reactions, and therefore sufficient strength of binding is necessary in order to achieve efficient methylation. Therefore, it was also decided to develop an assay to determine the DNA binding of these new compounds and understand how the binding is influenced by the attachment of the various linkers and the glucose unit to the Me-lex unit. This knowledge would enable the efficient design and synthesis of compounds which are likely to have a high chance of success in methylating DNA with the targeting ligand, glucose attached.

Fluorescence assays have been described in literature for measuring binding constants of compounds which bind to the DNA minor groove at A/T rich regions. For example, in the fluorescent compound Hoechst 33258 (Figure 4.1 a)) that binds to the A/T rich minor groove regions of DNA is used to measure the binding constants of compounds like netropsin and distamycin (Figure 4.1 b) and c)), which bind to the same sites on DNA.[32] The fluorescence properties of Hoechst 33258 when bound within the minor groove of DNA are different from those when it is free in solution.[32] This

a)



b)



c)

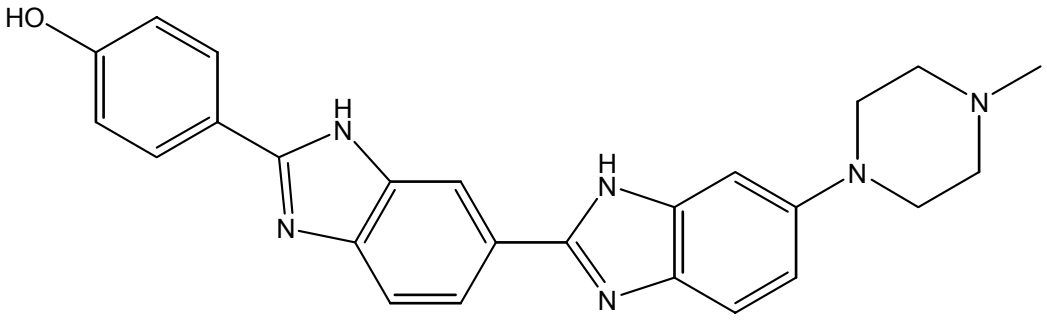


Figure 4.1. Structure of a) netropsin b) distamycin and c) Hoechst 33258.

difference in fluorescence has been used to determine the binding constants of the different compounds that compete for the same binding site.

The fluorescence assay described above is well suited for this study. However, the fluorescent probe used, Hoechst 33258, is a strong DNA binder ($K_b > 10^7$) while the compounds being investigated in this project are weak binders ($K_b \approx 10^5$) by design. Therefore, it is unlikely that the compounds being tested can displace the Hoechst 33258 from the binding site.

In order to use this assay to determine the binding of the new compounds, a fluorescent probe is needed whose strength of binding to the desired target site is comparable to the binding strength of the compound being tested. An extensive survey of literature did not lead to the identification of a suitable compound. Therefore, it was decided to synthesize new fluorescent probes suitable for the study.

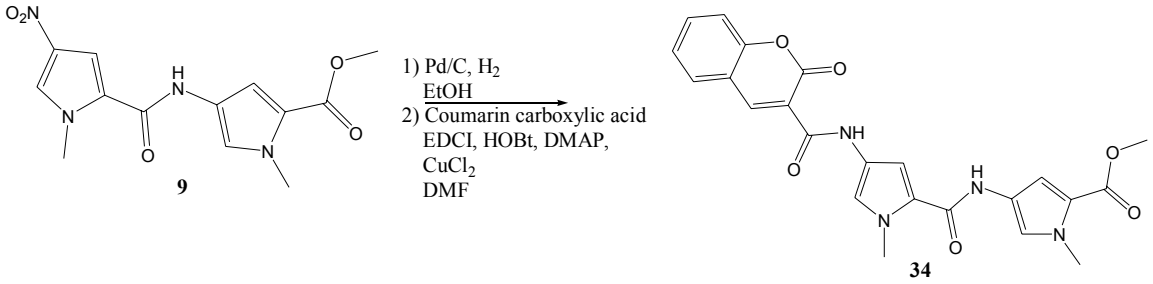
In order to design new fluorescent probes that would target the A/T rich sites in the minor groove of DNA it was decided to use the same DNA recognizing dipyrrole units as the compounds described in CHAPTER 3 and attach a fluorophore either at the C-terminus or the N-terminus of these compounds. Coumarin was selected as the fluorophore to be attached because of its favorable fluorescence properties, water solubility, and relatively planar structure which would enable it to slide within the DNA minor groove with minimal steric problems. The mode of attachment of the coumarin to the DNA recognizing dipyrrole units is important because the resultant compounds must retain their ability to bind to the minor groove at A/T rich regions of DNA, exhibit good fluorescence properties in solution, and exhibit a change in fluorescence when bound to DNA. With these characteristics in mind, four new compounds were designed and are shown in Figure 4.2 a)-d). In two of the compounds (**41**, **42**) the appropriate coumarin is attached at the C-terminus of the dipyrroles and the other two (**34**, **36**) the coumarin component

is attached to the N-terminus of the dipyrroles. If one of these compound exhibits sufficient binding to the target site and exhibits a change in fluorescence properties upon biding to the DNA, it can be used as a probe to investigate the DNA-binding ability of the various compounds and intermediates prepared in this project.

4.2. Synthesis

The synthesis of the fluorescent probes involved the attachment of the appropriate DNA recognizing dipyrrole component to the appropriate coumarin component. The two coumarin components that were used are shown in Figure 4.3 a) and b). When the coumarin was to be connected to the N-terminus, as in compounds **34** and **36**, coumarin-3-carboxylic acid (Figure 4.3 a)) was used. When the coumarin unit was to be connected to the C-terminus, as in compounds **39** and **43**, 7-amino-4-methylcoumarin (Figure 4.3 b)) was used.

Compound **34** was prepared earlier in the laboratory by condensation of **9** with coumarin -3-carboxylic acid as shown in Scheme 4.1. The nitro group on **9** was first



9 **34**

Scheme 4.1

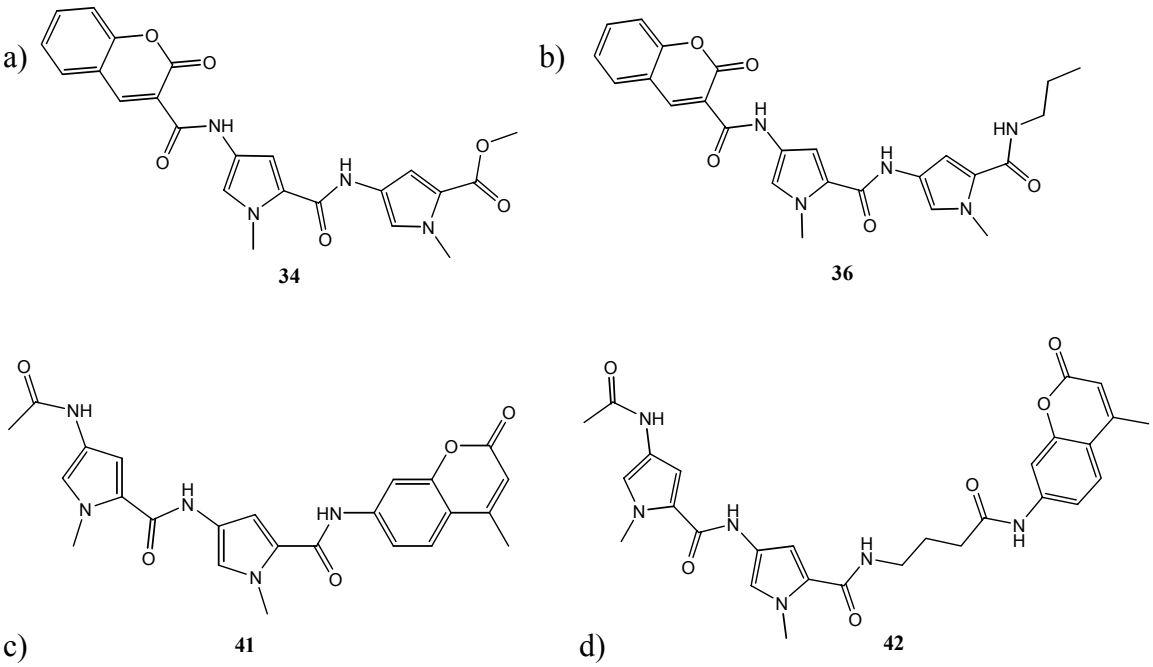
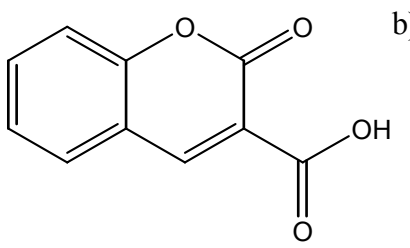


Figure 4.2. a) Coumarin attached at the N-terminus with a methyl ester at the C-terminus.
b) Coumarin attached at the N-terminus with a propylamide at the C-terminus.
c) Coumarin attached at the C-terminus with an acetamide at the N-terminus.
d) Coumarin attached at the C-terminus with a four carbon linker with an acetamide at the N-terminus.

a) 

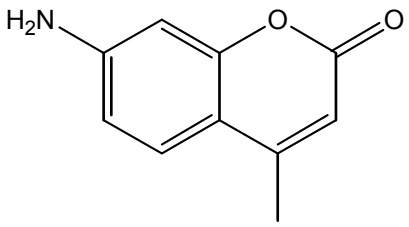
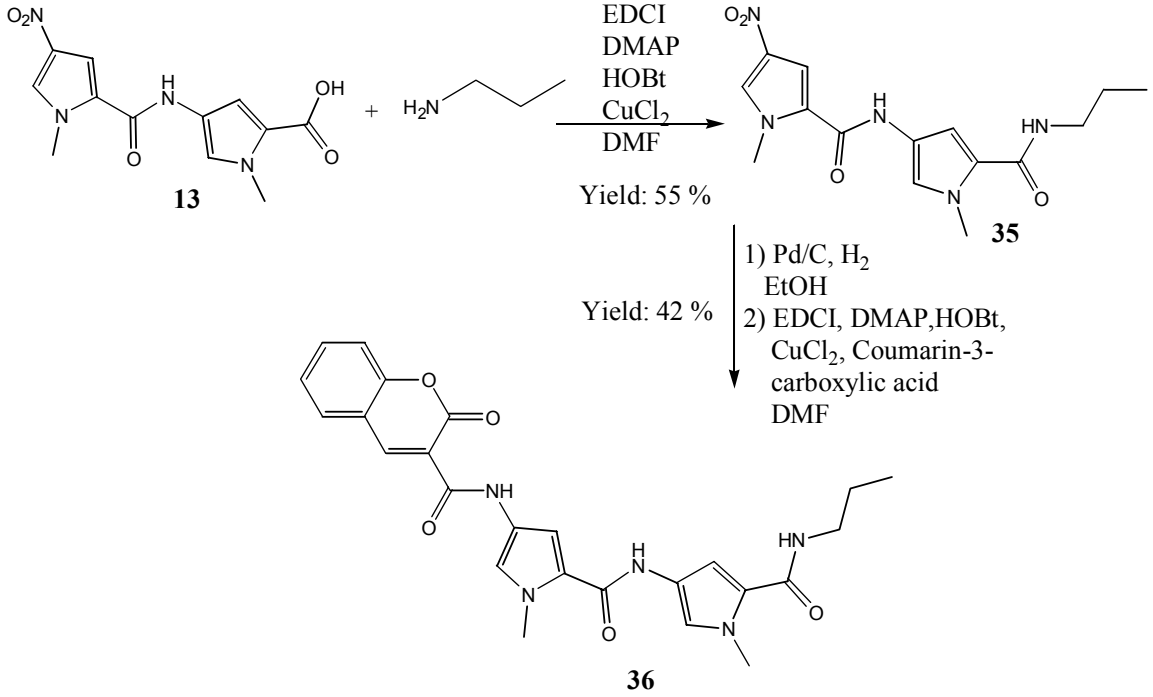
b) 

Figure 4.3. a) Coumarin-3-carboxylic acid and b) 7-amino-4-methylcoumarin

reduced to the amine and used for the next step without isolation. The amine was then reacted with the coumarin-3-carboxylic acid with the coupling reagents EDCI, HOBT, DMAP, and $CuCl_2$ in DMF to give **34**.

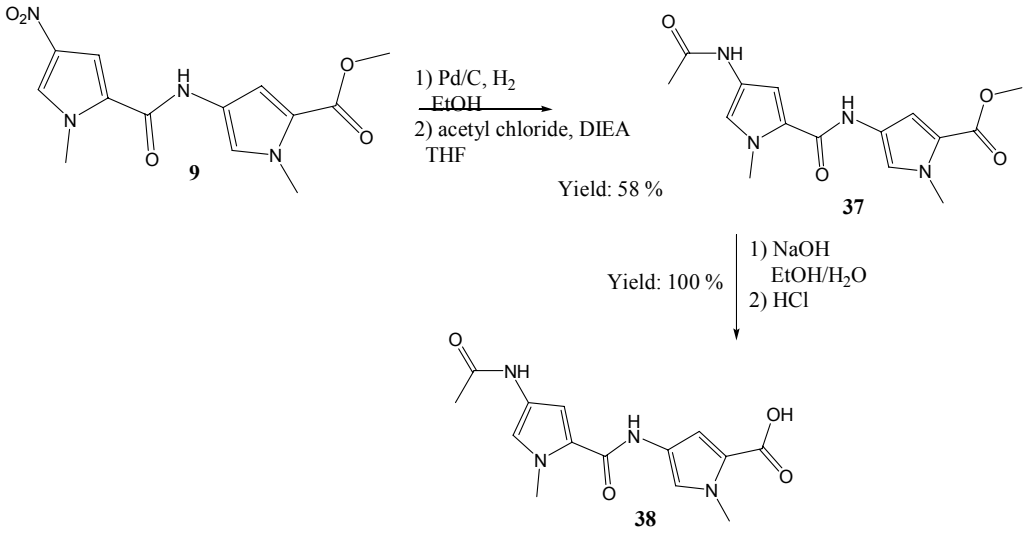
The synthesis of compound **36** was accomplished in three steps starting from compound **13** as shown in Scheme 4.2. The first step was a coupling reaction with **13** and N-propylamine using the coupling reagents EDCI, HOBT, DMAP, and $CuCl_2$ in DMF to give **35** in a 55 % yield. The next step involved a reduction of the nitro on **33** to



Scheme 4.2

the amine as in Scheme 4.2, and the last step was a coupling reaction with coumarin-3-carboxylic acid and the amine using the same coupling reagents as above to give **36** in a 42 % yield. This compound has a propylamide group on the C-terminus of the compound instead of a methylester as in **34**. This variation could make a difference in binding with DNA because of the amide NH on **36** that can form hydrogen binding in the minor groove of DNA.

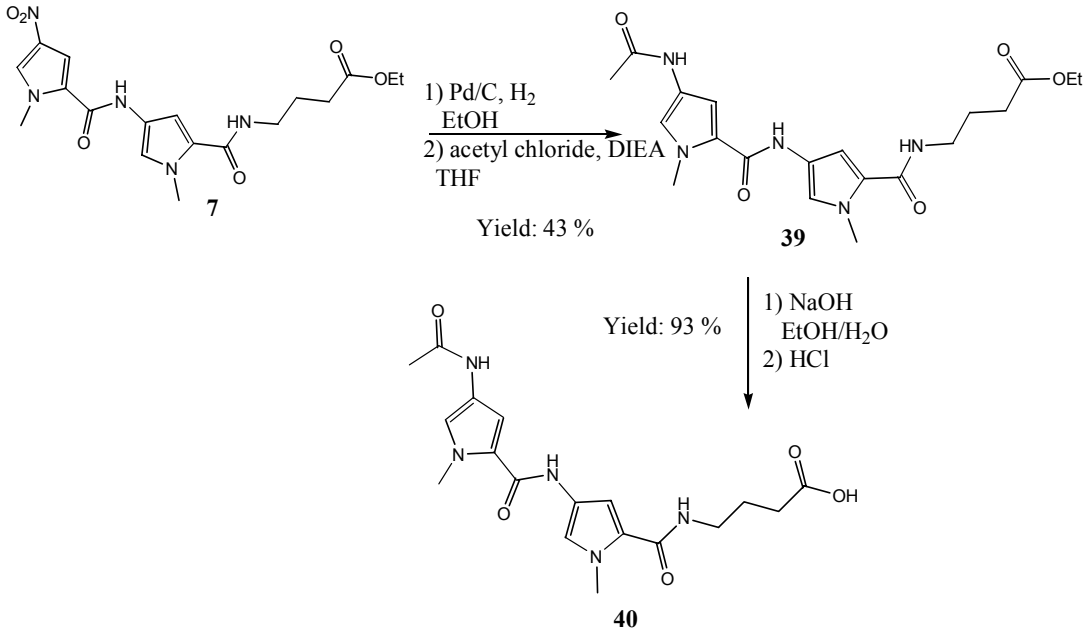
In order to synthesize compounds **41** and **42**, which have the coumarin units attached to the C-terminus, the N-terminus had to be first converted to an amide in order to contribute favorably to the DNA binding. In order to do this, the nitro group on **9** was first reduced to the amine and acetyl chloride was used to form an acetamide at the N-terminus to give **37** in a 58 % yield as shown in Scheme 4.3. The next step was a hydrolysis of the methylester to give the carboxylic acid **38** in a quantitative yield.



9 **37** **38**

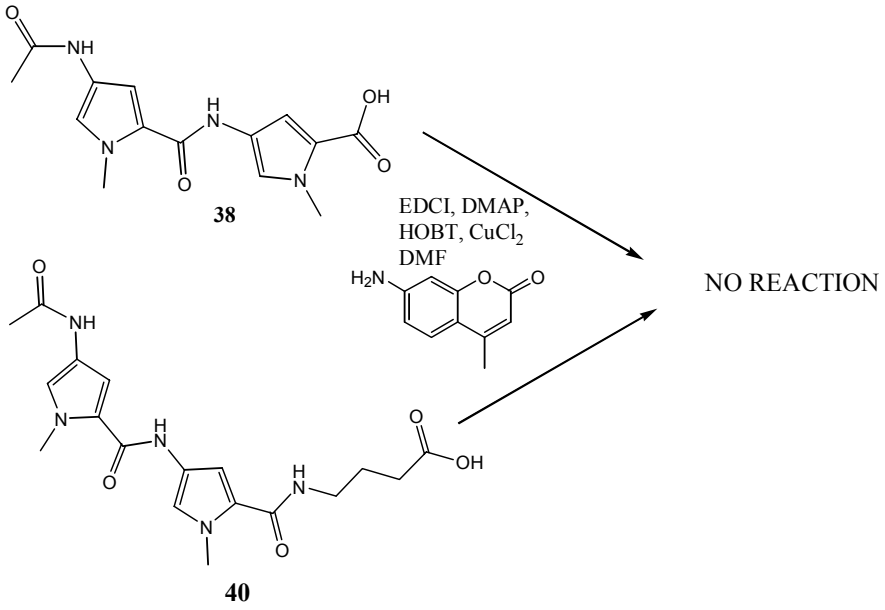
Scheme 4.3

In order to prepare compound **42** the nitro group on **7** was first reduced to the amine and acetyl chloride was used to form an acetamide at the N-terminus to give **39** in a 43 % yield as shown in Scheme 4.4. The next step was a hydrolysis of the ethylester to give



Scheme 4.4

the carboxylic acid **40** in a 93 % yield. However, attempts to couple 7-amino-4- methylcoumarin with carboxylic acids **38** and **40** (Scheme 4.5) have so far been

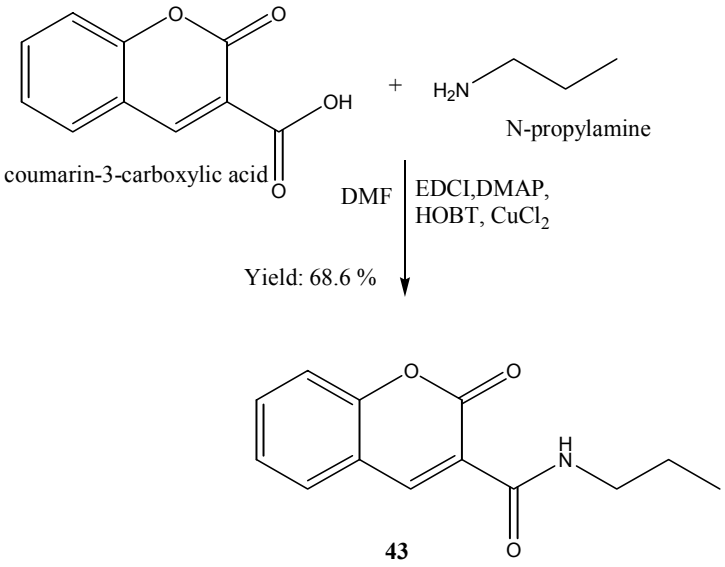


Scheme 4.5

unsuccessful, probably due to the unreactive nature of the 7-amino-4-methylcoumarin because of the lone pair of electrons on the amine being able to be delocalized into the rings of the

compound. Further efforts to accomplish this coupling reaction are currently in progress in the lab.

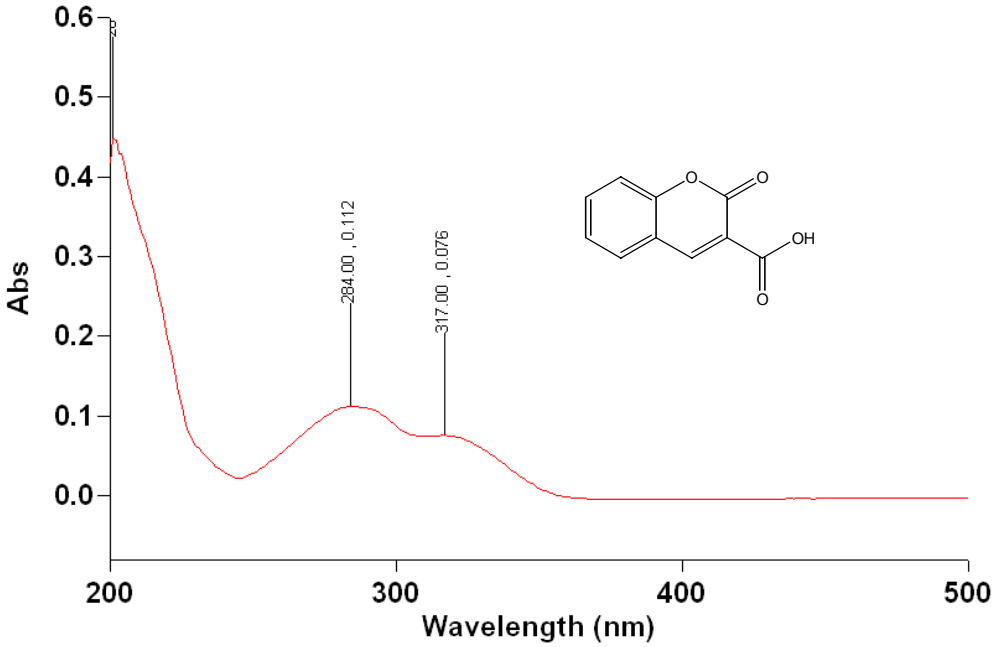
Coumarin-3-carboxylic acid itself is not very fluorescent, but as been reported to become fluorescent when the carboxylic acid is converted into an amide. In order to verify that this is indeed the case, compound **43** was made in a 69 % yield by condensing coumarin-3-carboxylic acid with N-propylamine using EDCI as the coupling reagent as shown in Scheme 4.6.



Scheme 4.6

4.3. Characterization of Spectral Properties

Compounds **34** and **36** in which the coumarin was attached to the N-terminus of the dipyrrole units were investigated further to see if they were suitable as fluorescent probes for the DNA binding assays. Coumarin-3-carboxylic acid, which was condensed with the amine on the dipyrrole units to form **34** and **36**, fluoresces very weakly. In order to verify that conversion of this compound into an amide results in a fluorescent compound, the fluorescence properties of **43**



a)

b)

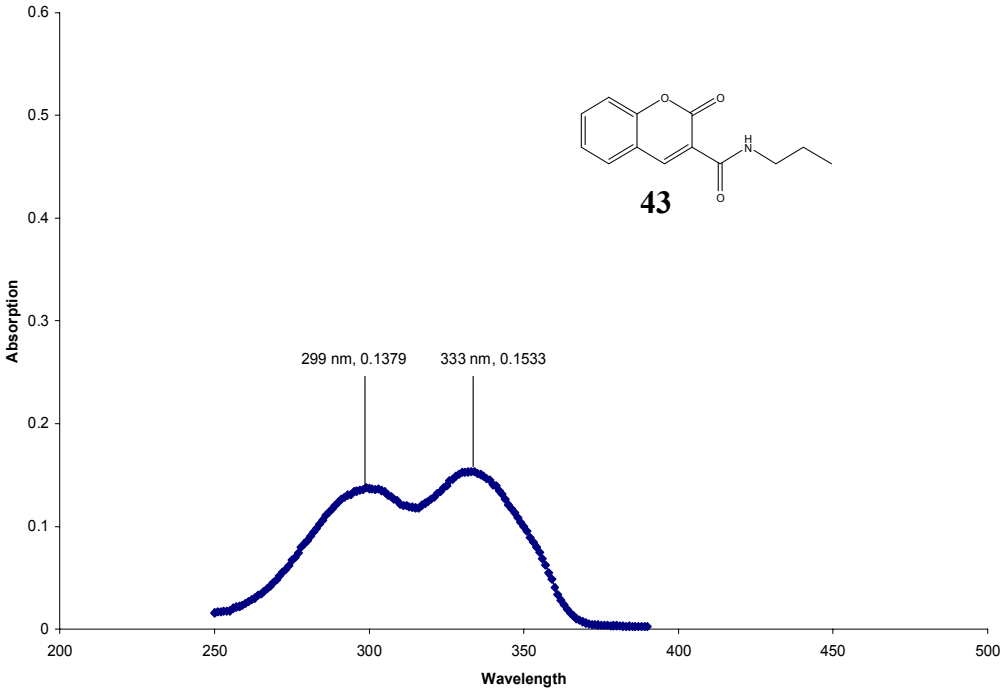


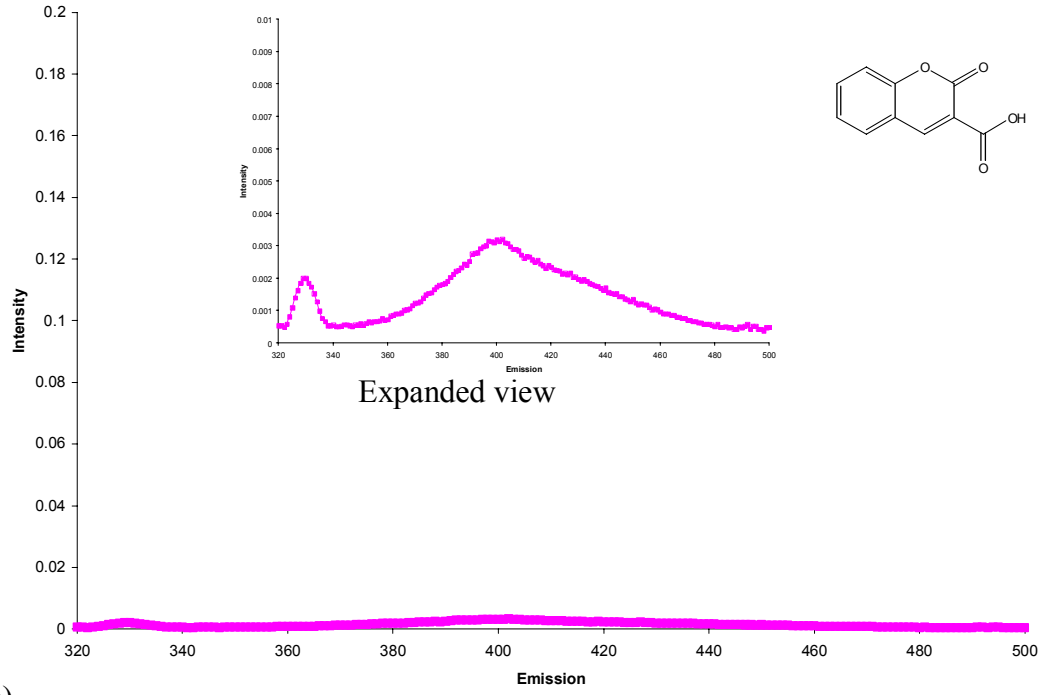
Figure 4.4. UV absorption spectra of a) coumarin-3-carboxylic acid and b) **43** at 10 µM in MeOH.

were tested. The UV spectra of coumarin-3-carboxylic acid and **43** are shown in Figure 4.4 a) and b) respectively. The carboxylic acid has a absorbance maxima at 284 and 317 nm, while the amide **43** has a absorbance maxima at 299 and 333 nm. These compounds were excited at 300 nm in a pH 7 buffer solution at 10 μM concentration, and the fluorescence spectra of the compounds are shown in Figure 4.5. As can be seen from the figure, conversion of the carboxylic acid to a simple amide results in a compound with strong fluorescence.

UV absorption spectra of compounds **34** and **36** are shown in Figure 4.6. The absorption profile of these two compounds was compared with the absorption profile of compound **38** (Figure 4.7, dipyrrole unit without attached coumarin) and that of the coumarin amide **43** (Figure 4.4 b)) in order to identify the absorption due to the coumarin unit in compounds **34** and **36**. This absorption maximum due to coumarin can then be used to determine the excitation wavelength for the fluorescent studies. However, neither compound (**34** or **36**) showed a strong absorbance attributable only to the coumarin unit. Both compounds had an absorption maximum around 300 nm and this wavelength was used as the excitation wavelength in the fluorescence experiments.

The fluorescence spectra of compounds **34** and **36** when excited at 300 nm in the pH 7 buffer solution at 10 μM concentration are shown in figure 4.8 a) and b) respectively. As can be seen from the figure, neither compound fluoresces strongly. The reason for this lack of fluorescence is not clear. Perhaps there is a lack of fluorescence due to the extended conjugation of the coumarin with the dipyrrole unit.

a)



b)

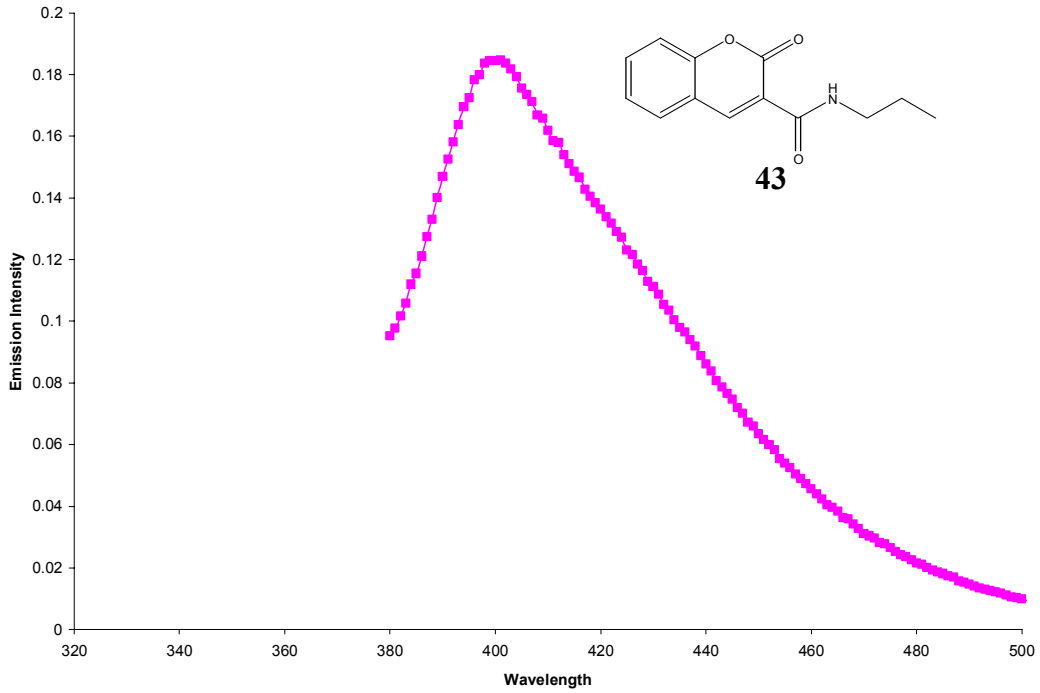
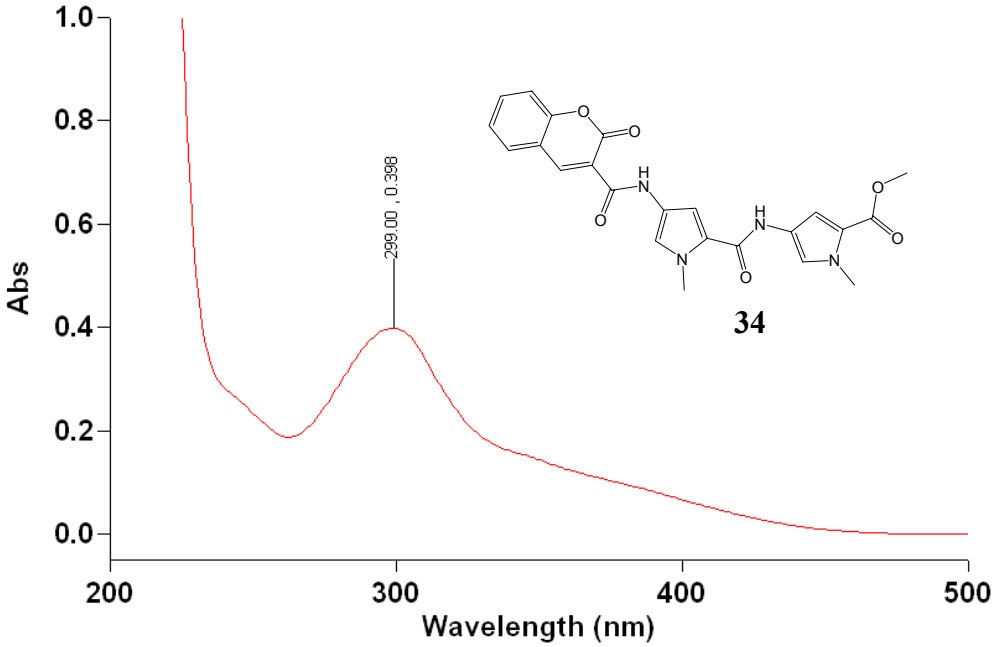


Figure 4.5. Fluorescence of a) coumarin-3-carboxylic acid and b) of **43** in MeOH at 10 μM concentration excited at 300 nm.



a)
b)

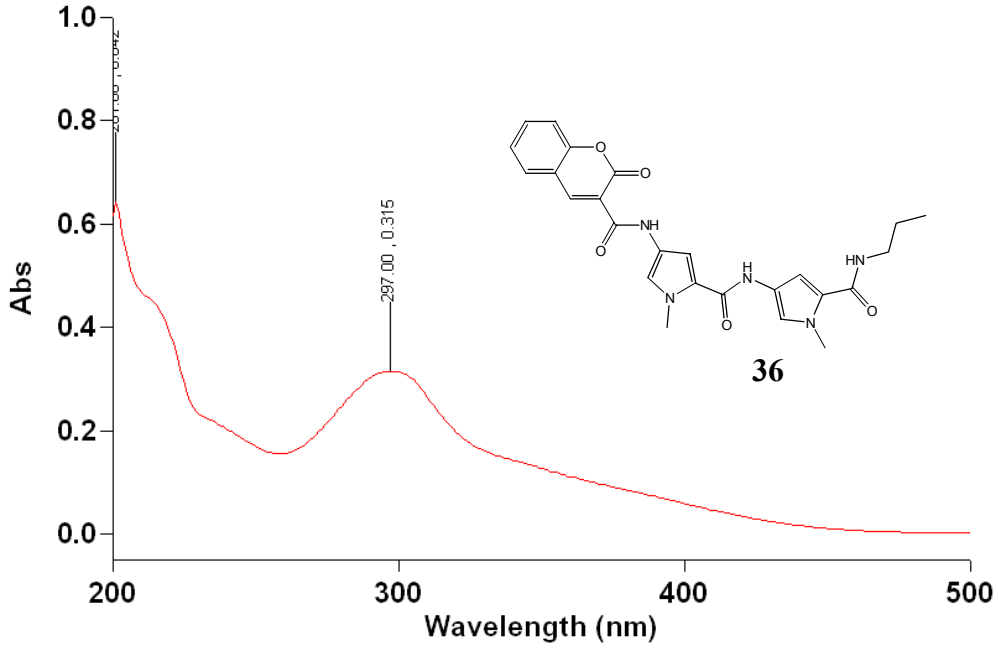


Figure 4.6. UV absorption spectra of a) **34** and b) **36** at a 10 μM concentration in MeOH.

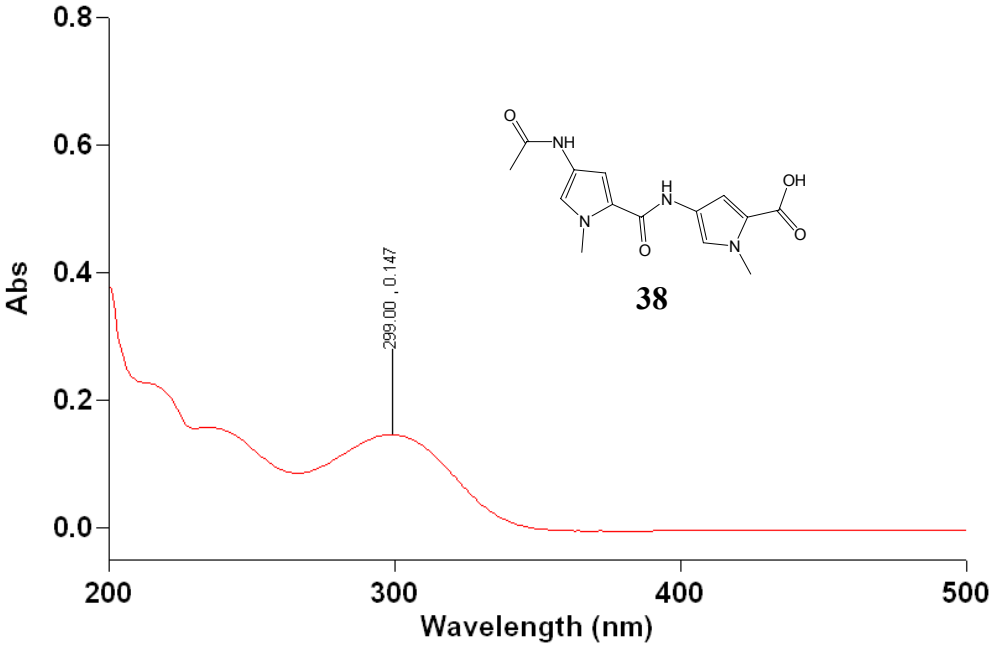


Figure 4.7. UV absorption spectrum of **38** at a 10 μM concentration in MeOH.

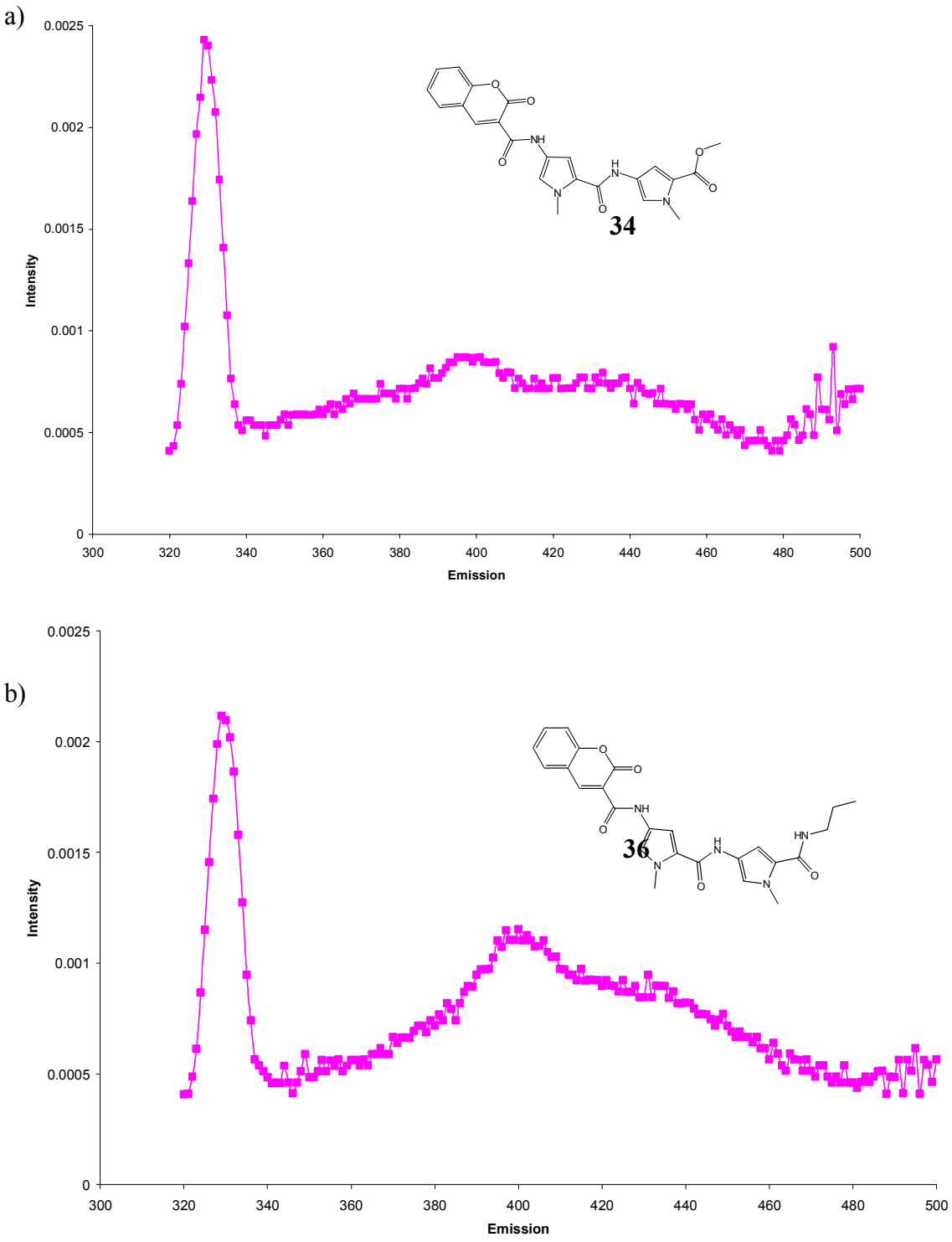


Figure 4.8. Fluorescence of a) **34** and b) **36** in buffer at 10 µM excited at 300 nm.

In order to test whether pyrrole compounds directly quench the fluorescence or whether the mode of connection of the coumarin unit to the pyrrole unit is responsible for the weak fluorescence of compounds **34** and **36**, 7-hydroxy-4-methylcoumarin (a model compound which fluoresces strongly) was titrated with compound **38**, and the change in fluorescence was recorded as shown in Figure 4.9. Aliquots of a 20 mM solution of **38** were added to a 10 μM MeOH solution of 7-hydroxy-4-methylcoumarin in MeOH. Each addition of **38** reduced the fluorescence intensity as shown in Figure 4.9. This decrease in fluoresce could be either due to quenching by compound **38**, or due to excitation wavelength intensity being filtered out due to absorption by increasing amounts of **38**. If fluorescence of 7-hydroxy-4-methylcoumarin is indeed being collisionally quenched by **38**, then a plot of F_o/F versus concentration of compound **38** (where F_o is the fluorescence observed in the absence of **38** and F is the fluorescence observed at a particular concentration of **38**) should be a straight line. If there is a complex formation, a curved line can result. As can be seen from Figure 4.10, this plot is not a straight line. If the reduction in fluorescence is due to excitation intensity being filtered out by increasing concentration of **38**, then a plot of the increase in UV absorbance at 330 nm (excitation wavelength) versus the decrease in fluorescence should be a straight line. The variation in the UV absorption profile upon addition of varying amounts of **38** is shown in Figure 4.11, and the plot of the change in absorbance versus the change in fluorescence is shown in Figure 4.12. Since this plot is a straight line, the decrease in fluorescence must be due to the filtering out of excitation wavelength intensity by increasing amounts of **38**, where **38** is absorbing all the incident light intensity.

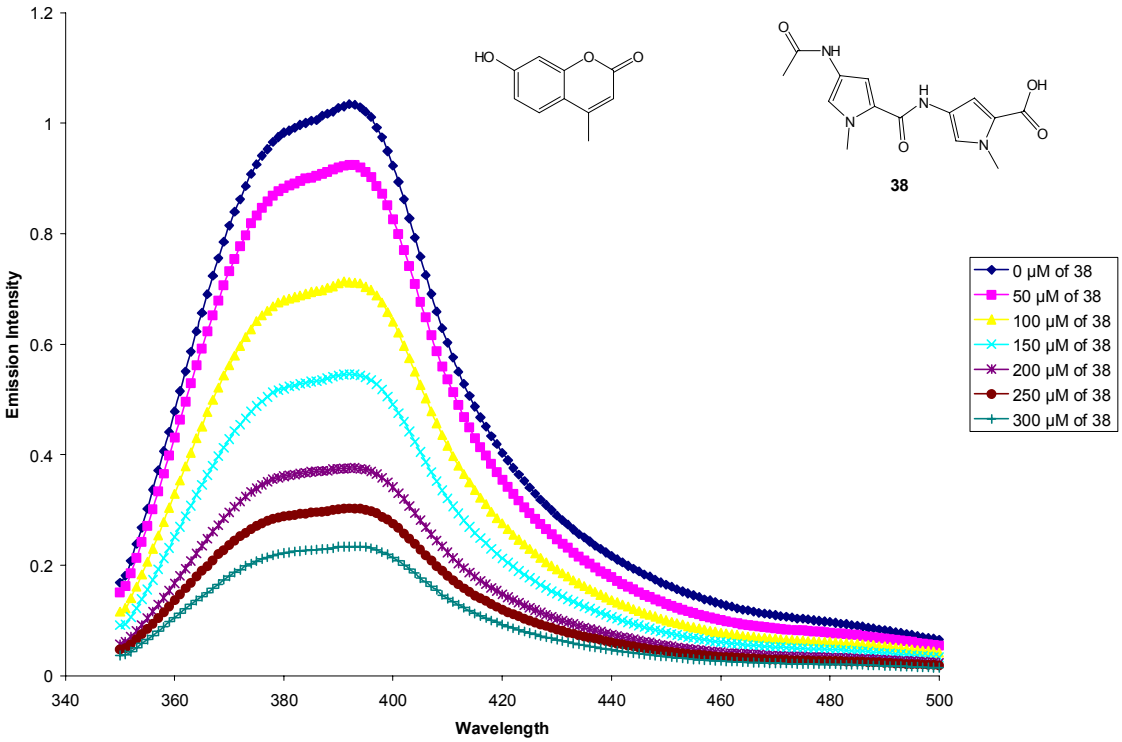


Figure 4.9. Fluorescence titration of compound **38** into a 10 µM solution of 7-hydroxy-4-methylcoumarin to show any fluorescent quenching excited at 330 nm.

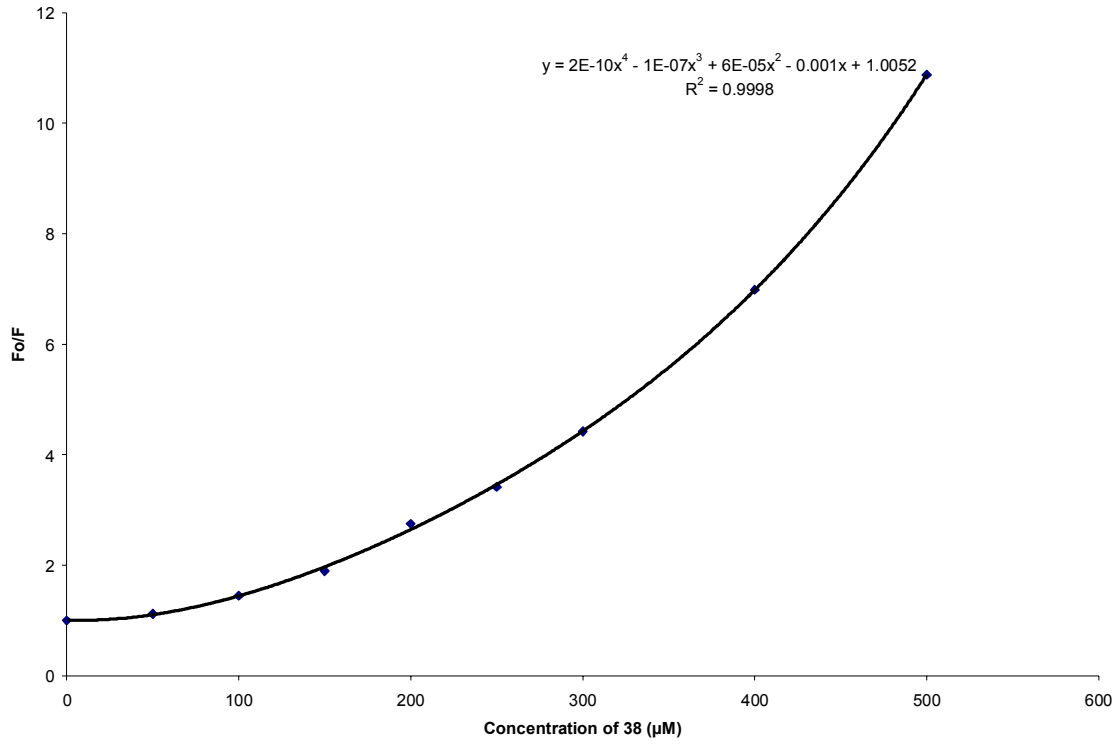


Figure 4.10. F_o/F vs. concentration of **38**, where F_o is the fluorescence observed in the absence of **38** and F is the fluorescence observed at a particular concentration of **38**.

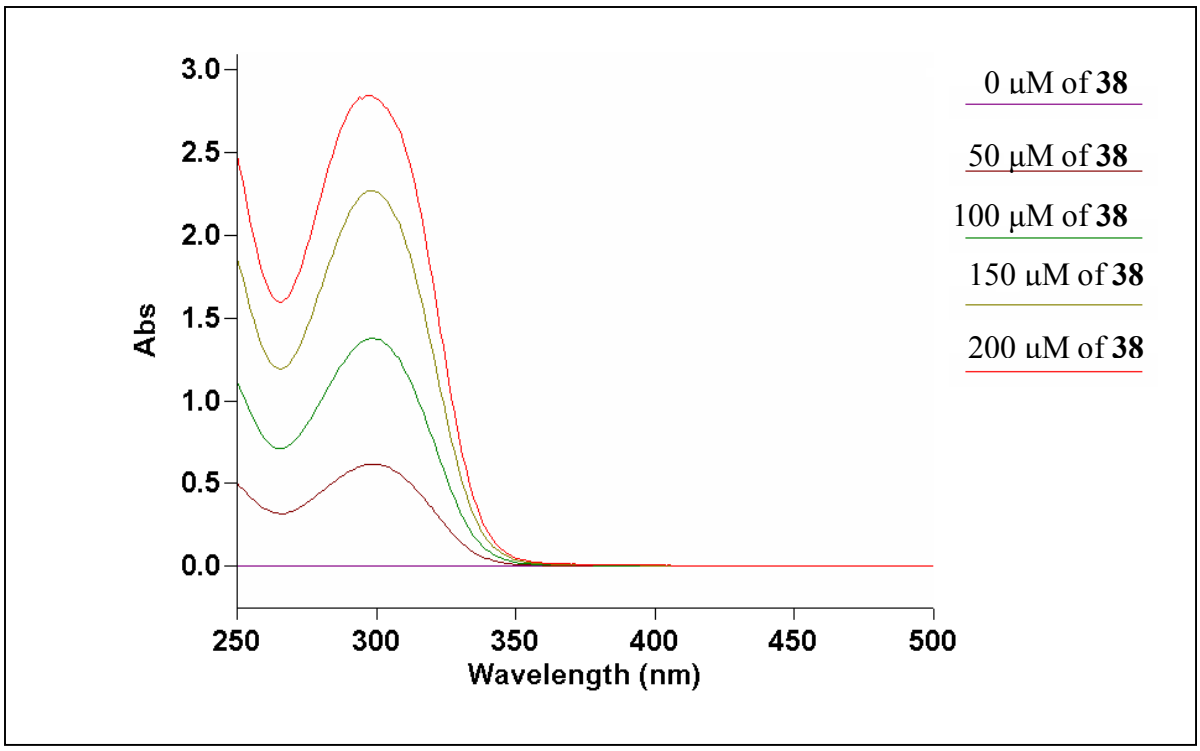


Figure 4.11. Variation in the UV absorption at varying concentrations of **38** in MeOH.

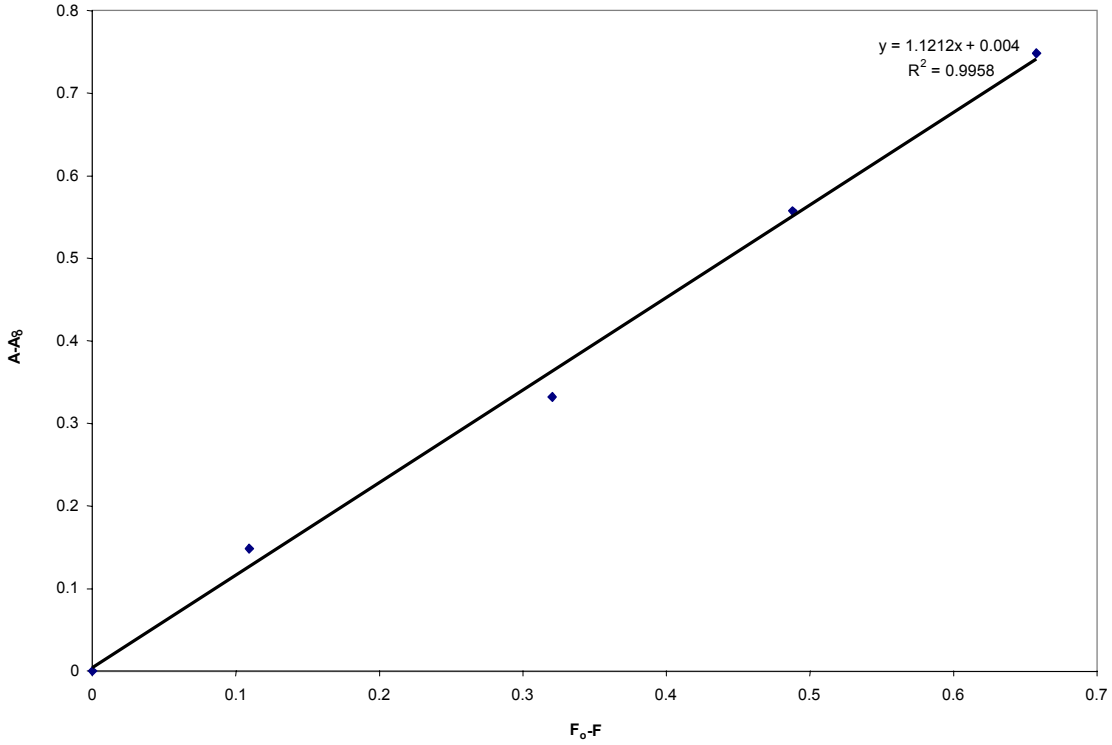


Figure 4.12. A-A_o of **38** vs. F_o-F of 7-hydroxy-4-methyl coumarin titrated with **38**, where A is the absorption of **38** at varying concentrations, A_o is the absorption with no **38**, F_o is the fluorescence observed in the absence of **38**, and F is the fluorescence observed at a particular concentration of **38**.

Since the mode of connection of coumarin to the dipyrrole unit appears to be responsible for the lack of fluorescence, it is possible that the introduction of methylene units between the coumarin and the dipyrroles, or the connection of the coumarin to the C-terminus (as in compounds **41** and **42**) may result in fluorescent compounds that can be investigated for use as probes for measuring DNA binding. The syntheses of such compounds are currently ongoing in the lab.

CHAPTER 5. EXPERIMENTAL

5.1. General

All solvents and reagents were purchased from VWR International (West Chester, Pennsylvania) or Sigma-Aldrich (Atlanta, Georgia) and were of the highest grade available unless otherwise noted. Flash chromatography was performed with silica gel (230/400 mesh, Life Force Inc.). TLC was performed on glass plates coated with silica gel (Whatman International, Maidstone, England, 150 A) that had a fluorescence indicator and were detected by UV visualization. All rotary evaporations were carried out using a Buchi R-3000 or a Buchi R-114 rotary evaporator equipped with a Brinkmann model B-16 vacuum aspirator. Hydrogenations were performed using a Parr Hydrogenation Apparatus in a 500 mL Parr jar. Melting points were determined using a Mel-Temp II.

All anhydrous reactions were carried out under positive pressure of argon. Glassware used for anhydrous reactions were dried overnight in an oven at 140 °C or dried over a flame, assembled while still hot, and cooled to room temperature under argon. Solvents and reagents (liquids) for anhydrous reactions were obtained in bottles with sure-seal caps and transferred by using oven-dried needles and glass syringes.

All ^{1}H NMR and ^{13}C NMR spectra were recorded with a Bruker Avance 400 MHz NMR spectrometer, using deuterated DMSO as the solvent. The deuterated DMSO was obtained in sealed ampoules from ACROS Organics or Sigma-Aldrich. The spectra are reported in ppm and referenced to deuterated DMSO (2.49 ppm for ^{1}H, 39.5 ppm for ^{13}C). The samples were contained in 5 mm pyrex glass tubes obtained from Wilmad- LabGlass Buena, New Jersey.

Ultraviolet absorbance analyses were performed on a Cary 1E UV-Visible spectrophotometer. Fluorescence emissions were recorded on a SLM 8100 spectrofluorometer.

5.2. Fluorescence Methods

5.2.1. Solutions of Compounds for UV and Fluorescence

All compounds were accurately weighed with 10 mg or more and dissolved in known volumes of appropriate solvents in order to make stock solutions of exact concentrations. All compounds were first dissolved in methanol except **34**, which was first dissolved in 1 mL of DMSO and then diluted into methanol. The final percentage of DMSO when used was always less than 0.3 % of the total volume. The stock solutions were diluted with a 0.01 M potassium phosphate buffer, pH 7.0, 0.01 M NaCl to a final concentration of 10 μM. These solutions of the compounds were used for UV and fluorescence experiments.

5.2.2. UV and Fluorescence Experiments

UV absorption wavelengths were scanned from 200 nm to 500 nm of each of the 10 μM solutions in the 0.01 M potassium phosphate buffer, pH 7.0, 0.01 M NaCl to find the absorption maxima. The excitation wavelength for all the compounds was set at 300 nm and the fluorescence was scanned from 330 nm to 500 nm.

For the titration of **38** with 7-hydroxy-4-methylcoumarin, the fluorescence was taken of a 10 μM solution of 7-hdyroxy-4-methylcoumarin. Aliquots of 2.5 μL of a solution containing **38** at 20 mM and 7-hydroxy-4-methylcoumarin at 10 μM in MeOH were titrated in until **38** was at 300 μM. Thereby each addition did not change the concentration of 7-hydroxy-4-methylcoumarin; however each addition changed the concentration of **38**. The excitation wavelength was 330 nm and the fluorescence was measured from 350 to 500 nm.

5.3. Synthesis

2,2,2-trichloro-1-(1-methyl-1*H*-pyrrol-2-yl)ethanone (3). This compound was prepared earlier in the laboratory using methods described in literature.[27]

2,2,2-trichloro-1-(1-methyl-4-nitro-1*H*-pyrrol-2-yl)ethanone (4). Pyrrole **3** (5.00 g, 0.0220 mol) was dissolved in acetic anhydride (50.0 mL) in a 250 mL round bottom flask and cooled to -40 ºC in a dry ice/acetone bath. Fuming nitric acid (2.5 mL) was added dropwise over 5 minutes with constant stirring while keeping the temperature of the reaction mixture lower than -30 °C by periodic addition of dry ice to the acetone bath. After the addition of HNO_3, the mixture was stirred for another 45 minutes at -40 °C. The solution was then allowed to warm to room temperature and was stirred for an additional hour. The solution was then cooled in an ice bath and then water (30 mL) was added slowly over one minute. A yellow solid precipitated and was filtered under vacuum and dried to give **4** (5.04 g, 84%): mp 112-120 ºC. TLC (1:1, EtOAc:Hexane) $R_f = 0.58$. ^{1}H NMR data: δ 8.56 (d, J = 1.73 Hz, 1H), 7.78 (d, J = 1.71 Hz, 1H), 3.98 (s, 3H). ^{13}C NMR data: δ 173.30, 134.73, 133.09, 121.09, 116.82, 95.02, 79.44.

Ethyl 4-{[(1-methyl-4-nitro-1*H*-pyrrol-2-yl)carbonyl]amino}butanoate (5). The nitro pyrrole **4** (8.00 g, 0.0352 mol) was dissolved in EtOAc (420 mL) and TEA (10 mL) was added while Ar was bubbled through the solution. Aminobutyrate hydrochloride (5.90 g, 0.0353 mol) was then added and the solution was allowed to stir overnight under Ar. The white precipitate that was formed was filtered off and the filtrate was concentrated by rotary evaporation to give a brown oil. The oil was kept under vacuum until a yellow solid formed to give **5** (8.58 g, 86%): mp 55-60 ºC. TLC (1:1, EtOAc:Hexane) $R_f = 0.27$. ^{1}H NMR data: δ 8.39 (t, J = 5.37, 5.41 Hz,1H), 8.11 (d, J = 1.67 Hz, 1H), 7.41 (d, J = 1.93 Hz, 1H), 4.02 (q, J = 7.12, 7.11, 7.11 Hz, 2H), 3.88 (s, 3H), 3.19 (q, J = 6.72, 6.04, 6.65 Hz, 2H), 2.33 (t, J = 7.4, 7.38 Hz,

2H), 1.73 (p, J = 7.17, 7.15, 7.15, 7.13 Hz, 2H), 1.16 (t, J = 7.10, 7.11 Hz, 3H). ^{13}C NMR data: δ 172.65, 159.85, 133.75, 127.85, 126.42, 107.30, 59.80, 37.94, 37.38, 30.97, 24.38, 14.11.

Ethyl *N*-[(1-methyl-4-nitro-1*H*-pyrrol-2-yl)carbonyl]-β-alaninate (6). Compound **6** was synthesized by a similar procedure similar to the one described above for **5** using 4.00 g (14.8 mmol) of **4** and 2.70 g (17.8 mmol) of β-alanine ethyl ester hydrochloride to obtain **6** (2.79 g, 57 %): mp 128-129 °C. TLC (2:1, EtOAc:Hexane) R_f = 0.45. ^{1}H NMR data: δ 8.48 (t, J = 5.37, 5.42 Hz, 1H), 8.12 (d, J = 1.88 Hz, 1H), 7.39 (d, J = 2.00 Hz, 1H), 4.05 (q, J = 7.11, 7.11, 7.13 Hz, 2H), 3.88 (s, 3H), 3.40 (q, J = 6.90, 5.61, 6.85, 2H), 2.53 (t, J = 6.90, 6.93 Hz, 2H), 1.16 (t, J = 7.12, 7.14 Hz, 3H). . ^{13}C NMR data: δ 171.20, 159.87, 133.74, 127.94, 126.23, 107.42, 59.99, 37.38, 34.95, 33.63, 14.09.

Ethyl 4-{[(1-methyl-4-{[(1-methyl-4-nitro-1*H*-pyrrol-2-yl)carbonyl]amino}-1*H*-pyrrol-2-yl)carbonyl]amino}butanoate (7). Following Scheme 3.1 the nitro pyrrole ester **5** (2.28 g, 8.06 mmol) was dissolved in ethanol (70 mL) in a 500 mL parr jar and Pd/C (2.00 g) was added to it. The mixture was shaken in a hydrogenator under pressurized hydrogen (60 psi) until the reaction was complete as indicated by TLC (EtOAc in 45 min.). The Pd/C was filtered out through celite, and the filtrate was concentrated by rotary evaporation and kept under vacuum overnight to yield a yellow solid. Ethyl acetate (20 mL) was added to this solid in a 250 mL round bottom flask and the nitro pyrrole **4** (2.20 g, 8.00 mmol) was added and the mixture was allowed to stir for 48 hours during which time an orange solid fell out of solution. After the reaction was complete as indicated by the disappearance of the starting material by TLC (4:1, EtOAc:MeOH), the orange product that fell out of solution was vacuum filtered to give pure **7** (1.15 g, 35.1%).

Alternate procedure following Scheme 3.2: In a flask flushed with Ar, DMAP (3.48 g, 2.5 eq.), EDCI (3.38 g, 1.5 eq.), HOBT (5.24 g, 3 eq.), $CuCl_2$ (0.153 g, 0.1 eq.), ethyl-4-aminobutyrate, HCl (1.91 g, 0.114 mol), and the carboxylic acid **10** (3.33 g, 0.0114 mol) were added. The solid mixture was dissolved in 75 mL anhydrous DMF and allowed to stir under Ar overnight at which time TLC (EtOAc) indicated the complete disappearance of **10**. The solution was diluted with 150 mL DCM and the organic solution was extracted with H_2O (2x 150 mL), saturated $NaHCO_3$ (2x 200 mL), and 1M HCl (2x 200 mL). The organic layers were dried over $MgSO_4$. The solution was concentrated on a rotary evaporator until the volume reduced to approximately 20 mL. This solution was then cooled in the freezer overnight when yellow crystals formed. The yellow crystals were vacuum filtered to give pure **7** (2.86 g, 62 %): mp 140-143ºC. TLC (EtOAc) $R_f = 0.67$. ^{1}H NMR data: δ 10.24 (s, 1H), 8.19 (s, 1H), 8.09 (t, J = 5.4, 5.59 Hz, 1H), 7.58 (t, J = 1.72, 1.45 Hz, 1H), 7.20 (s, 1H), 6.86 (t, J = 1.64, 1.48 Hz, 1H), 4.19 (q, J = 7.12, 7.10, 7.13 Hz, 2H), 4.04 (s, 3H), 3.80 (s, 3H), 3.18 (q, J = 6.47, 6.30, 6.22 Hz, 2H), 2.32 (t, J = 7.51, 7.45 Hz, 2H), 1.73 (p, J = 7.00, 7.09, 7.12, 7.04 Hz, 2H), 1.17 (t, J = 7.12, 7.12 Hz, 3H). ^{13}C NMR data: δ 173.09, 161.55, 157.21, 134.16, 128.60, 126.68, 123.55, 121.70, 118.35, 107.92, 104.39, 60.13, 51.64, 37.86, 36.41, 31.42, 25.07, 14.49.

Ethyl *N*-[(1-methyl-4-{[(1-methyl-4-nitro-1*H*-pyrrol-2-yl)carbonyl]amino}-1*H*-pyrrol-2-yl)carbonyl]-β-alaninate (8). Compound **8** was synthesized following Scheme 3.1 using a procedure similar to the one described above for **7** using 2.11 g (7.87 mmol) of **5** and 2.149 g (7.87 mmol) of **4** to give **8** (0.975 g, 32 %).

Following the alternate procedure in Scheme 3.2 compound **8** was made similar to the one described above for **7** using 0.5802 g (19.9 mmol) of **10** and 0.3216 g (19.9 mmol) of β-alanine ethyl ester hydrochloride to obtain **8** (0.443 g, 57 %): mp 200-201 °C. TLC (EtOAc) R_f =

0.62. ^{1}H NMR data: δ 10.22 (s, 1H), 8.16 (d, J = 1.20 Hz, 1H), 8.10 (t, J = 5.44, 5.58 Hz, 1H), 7.57 (d, J = 1.88 Hz, 1H), 7.20 (d, J = 1.54 Hz, 1H), 6.83 (d, J = 1.68 Hz, 1H), 4.05 (q, J = 7.11, 7.11, 7.12 Hz, 2H), 3.94 (s, 3H), 3.79 (s, 3H), 3.39 (q, J = 6.78, 5.89, 6.74 Hz, 2H), 2.52 (t, J = 6.92, 7.04 Hz, 2H), 1.70 (t, J = 7.10, 7.13 Hz, 3H). ^{13}C NMR data: δ 171.44, 161.24, 156.91, 133.85, 128.27, 126.34, 123.03, 121.44, 118.17, 107.62, 104.17, 59.96, 37.54, 36.10, 34.89, 34.07, 14.15.

Methyl 1-methyl-4-{[(1-methyl-4-nitro-1*H*-pyrrol-2-yl)carbonyl]amino}-1*H*-pyrrole-2-carboxylate (9). Argon was bubbled through 50 mL EtOAc containing 3.67 mL DIEA. Methyl 4-amino-1-methyl-1*H*-pyrrole-2-carboxylate, HCl (2.49 g, 0.0131 mol) was then added to the solution followed by the addition of the nitro compound **4** (3.58 g, 0.0131 mol) and the solution was allowed to stir under Ar for two days. The yellow solid that was formed, was filtered out, washed with H_2O, and dried under vacuum to afford the yellow solid **9** (4.05 g, 94 %): mp 231-235 °C. TLC (2:1, EtOAc:Hexane) R_f = 0.33 . ^{1}H NMR data: δ 10.26 (s, 1H), 8.18 (d, J = 1.85 Hz, 1H), 7.54 (d, 1.98 Hz, 1H), 7.45 (d, 1.92 Hz, 1H), 6.87 (d, 1.96 Hz,1H), 3.94 (s, 3H), 3.83 (s, 3H), 3.73 (s, 3H). ^{13}C NMR data: δ 160.73, 156.94, 133.81, 128.36, 122.15, 120.86, 118.85, 108.28, 107.66, 51.06, 37.47, 36.28.

Methyl-4-({[4-(acryloylamino)-1-methyl-1H-pyrrol-2-yl]carbonyl}amino)-1-methyl-1H-pyrrole-2-carboxylate (10). Compound **9** (0.454 g, 10.4 mmol) was dissolved in 50 mL EtOH and 100 mg Pd/C was added in a parr jar. The mixture was shaken under pressurized hydrogen (75 psi) until **9** had disappeared as indicated by TLC (EtOAc). The Pd/C was filtered out through celite, and the solution was rotary evaporated and the residue was kept under vacuum overnight. The flask with the residue was flushed with Ar and the residue was dissolved in anhydrous THF (5 mL) and anhydrous DIEA (0.418 mL) was added to the solution. The

solution was bubbled with Ar and cooled to -40 °C. Acryloyl chloride (0.156 mL) was added dropwise to the solution which as then allowed to stir below -20 °C, protected from light, until the reaction was complete as indicated by TLC (EtOAc). The solution was concentrated by rotary evaporation and the product was isolated by flash column chromatography (EtOAc) to afford pure **10** (0.387 g, 71 %): mp 199-203 °C. TLC (EtOAc) $R_f = 0.43$. ^{1}H NMR data: δ 10.10 (s, 1H), 9.92 (s, 1H), 7.45 (d, J = 2.00 Hz, 1H), 7.25 (d, J = 2.00 Hz, 1H), 6.91 (d, J = 2.00 Hz, 1H), 6.88 (d, J = 2.00 Hz, 1H), 6.35 (m, 1H), 6.17 (dd, J = 2.00, 15.2 Hz, 1H), 5.65 (dd, J = 2.4, 7.60 Hz, 1H), 3.82 (s, 6H), 3.72 (s, 3H). ^{13}C NMR data: δ 161.71, 160.86, 158.38, 131.55, 125.70, 122.92, 121.84, 120.84, 118.64, 108.40, 104.16, 79.54, 78.44, 51.04, 36.25, 36.23.

1-methyl-4-{[(1-methyl-4-nitro-1*H*-pyrrol-2-yl)carbonyl]amino}-1*H*-pyrrole-2-carboxylic acid (13). The nitro ester **9** (4.00 g, 0.0131 mol) was suspended in EtOH (200 mL), and a solution of NaOH (2.1 g, 4 eq.) in H_2O (150 mL) was added. This suspension was allowed to reflux until the disappearance of **9** was indicated by TLC (EtOAc). The solution was concentrated by rotary evaporation to remove the majority of the EtOH. The aqueous solution was cooled in an ice bath and then acidified with concentrated HCl until the final pH was approximately 1 when a precipitate fell out of solution. The mixture was allowed to cool in the freezer for a further two hours, and the yellow solid precipitate was filtered out and rinsed with ice cold water and air-dried to give pure **13** (3.33 g, 87 %): mp 200-202 °C. TLC (EtOAc) $R_f = 0.56$. ^{1}H NMR data: δ 12.24 (s, 1H), 10.27 (s, 1H), 8.18 (d, J = 1.84 Hz, 1H), 7.56 (d, J = 1.95 Hz, 1H), 7.41 (d, J = 1.86 Hz, 1H), 6.83 (d, J = 1.96 Hz, 1H), 3.94 (s, 3H), 3.82 (s, 3H). ^{13}C NMR data: δ 161.88, 156.88, 133.81, 128.29, 126.17, 121.89, 120.39, 119.84, 108.30, 107.65, 37.47, 36.24.

4-{[(1-methyl-4-{[(1-methyl-4-nitro-1*H*-pyrrol-2-yl)carbonyl]amino}-1*H*-pyrrol-2-yl)carbonyl]amino}butanoic acid (14). Compound **7** (1.00 g, 2.46 mmol) was dissolved in ethanol (28 mL) at room temperature in a 250 mL round bottom flask. To this solution, KOH (0.460 g, 8.61 mmol) in water (20 mL) was added and allowed to stir until all of the starting material had disappeared by TLC (3:1, CH_2Cl_2:MeOH). Excess concentrated HCl (approx. 1 mL) was added until the medium was acidic, when a yellow precipitate fell out of solution. This precipitate was filtered out, washed with ice cold water, and air-dried to give the product **14** (0.855 g, 91.9 %): mp 251-254 ºC. TLC (4:1, EtOAc: MeOH) $R_f = 0.35$. ^{1}H NMR data δ 12.05 (s, 1H), 10.21 (s, 1H), 8.17 (d, J = 1.87 Hz, 1H), 8.08 (t, J = 5.62, 5.72 Hz, 1H), 7.57 (d, J = 1.95 Hz, 1H), 7.19 (d, J = 1.73 Hz, 1H), 6.85 (d, J = 1.79 Hz, 1H), 3.94 (s, 3H), 3.79 (s, 3H), 3.18 (q, J = 6.62, 6.09, 6.66 Hz, 2H), 2.24 (t, J = 7.42, 7.36 Hz, 2H) 1.69 (p, J = 7.10, 7.10, 7.11, 7.21 Hz, 2H). ^{13}C NMR data: δ 174.31, 161.19, 156.85, 133.79, 128.22, 126.31, 123.23, 121.33, 117.97, 107.56, 104.01, 37.85, 37.50, 36.04, 31.15, 24.74.

***N*-({1-methyl-4-[2-(5-methyl-3-nitrocyclopenta-1,3-dien-1-yl)prop-2-en-1-yl]-1*H*-pyrrol-2-yl}carbonyl)-β-alanine (15).** Compound **15** was synthesized using a procedure similar to the one described above for **14** using 1.29 g (3.29 mmol) of **8** to give **15** (1.17 g, 99 %): mp 230-234 °C. TLC (6:1, EtOAc: MeOH) $R_f = 0.52$. ^{1}H NMR data: δ 12.19 (s, 1H), 10.23 (s, 1H), 8.17 (d, J = 1.83 Hz, 1H), 8.08 (t, J = 5.52, 5.60 Hz, 1H) 7.56 (d, J = 1.98, 1H) 7.20 (d, J = 1.78 Hz, 1H), 6.83 (d, J = 1.84, 1H), 3.94 (s, 3H), 3.80 (s, 3H), 3.35 (q, J = 7.04, 5.75, 7.13 Hz, 2H), 2.46 (t, J = 8.31, 7.09 Hz, 2H). ^{13}C NMR data: δ 173.01, 161.61, 156.85, 133.78, 128.24, 126.30, 123.02, 121.36, 118.08, 107.57, 104.10, 37.50, 36.07, 34.88, 34.02.

2-Deoxy-2-(4-methoxybenzylidene)amino-β-D-glucopyranose (16). Compounds **16** was synthesized from D-glucosamine hydrochloride as described in literature.[26] Sodium

hydroxide (11 g) was dissolved in 235 mL of water in a 500 mL round bottom flask that was cooled to 0 ºC and D-glucosamine hydrochloride (50.01 g, 0.232 mol) was added. The solution was stirred until clear (5 min.) and 4-methoxybenzaldehyde (31.0 mL, 0.264 mol) was added and stirred at 0 ºC for 30 min. and then left unstirred at 0 ºC for 48 hours. A white precipitate fell out of solution and was filtered under vacuum and dried to give **16** (59.76 g, 75.2 %): mp 154-155ºC. ^{1}H NMR data: δ 8.11 (s, 1H), 7.67 (d, J = 8.66 Hz, 2H) 6.97 (d, J = 8.68 Hz, 2H), 6.52 (d, J = 5.67 Hz, 1H), 4.93 (d, J = 5.18 Hz, 1H), 4.82 (d, J = 5.55 Hz, 1H), 4.69 (t, J = 6.50, 6.49 Hz, 1H), 4.55 (t, J = 5.79, 5.72 Hz, 1H), 3.78 (s, 3H), 3.72(dd, J = 5.40, 4.70, 5.39 Hz, 1H), 3.45 (m, 2H), 3.23 (m, 1H), 3.15 (m, 1H), 2.78 (t, J = 8.35, 8.60 Hz, 1H). ^{13}C NMR data: δ 161.25, 161.06, 129.64 (2C), 129.11, 113.91 (2C), 95.64, 78.19, 76.86, 74.60, 70.37, 61.27, 55.29.

1,3,4,6-Tetra-*O*-acetyl-2-deoxy-2-(4-methoxybenzylidene)amino- β-D-glucopyranose (17). Compound **17** was synthesized as described in literature.[26] The imine **16** (55.01 g, 0.161 mol) was dissolved in anhydrous pyridine (350 mL) at 0 ºC in a 1 L round bottom flask. Acetic anhydride (200 mL) was then added to the solution and stirred while allowing the solution to warm to room temperature overnight. This mixture was concentrated to half the original volume on a rotary evaporator maintaining the temperature below 30 °C and was then poured into ice water (~1 L) and stirred for 1 h. The white precipitate that was formed was filtered under vacuum to give **17** (62.52 g, 78.6%): mp 172-173ºC. ^{1}H NMR data: δ 8.27 (s, 1H), 7.64 (d, J = 8.77 Hz, 2H), 6.97 (d, J = 8.76 Hz, 2H), 6.06 (d, J = 8.24 Hz, 1H), 5.44 (t, J = 9.67, 9.68 Hz, 1H), 4.97 (t, J = 9.61, 9.65 Hz, 1H), 4.26 (m, 2H), 4.01 (d, J = 10.60 Hz, 1H), 3.78 (s, 3H), 3.44 (t, J = 8.48, 9.52 Hz, 1H), 2.00 (s, 3H), 1.97 (s, 6H), 1.81 (s, 3H). ^{13}C NMR data: δ 170.04, 169.44, 168.98, 168.59, 164.45, 161.83, 129.93 (2C), 128.27, 114.19 (2C), 92.54, 72.35, 72.26, 71.54, 67.82, 61.66, 55.53, 20.52, 20.43 (2C), 20.18.

1,3,4,6-Tetra-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranose hydrochloride (18). Compound **18** was synthesized as described in literature.[26] Acetyl chloride (0.156 mL) was added to anhydrous methanol (1.08 mL) with constant stirring at 0 ºC in a 100 mL round bottom flask. This mixture was then added to a stirred solution of the imine **17** (1.00 g, 0.002 mol) in acetone (30.4 mL) at room temperature. This mixture was stirred for 45 min. and then cooled to 0 ºC. Ether (10.8 mL) was then added to the cooled solution and stirred for an additional 45 min. at 0 °C. Vacuum filtration of the white precipitate gave **18** (0.665 g, 81%): mp: 120 °C-decomposition. ^{1}H NMR data: δ 8.83 (s, 3H), 5.90 (d, J = 8.64 Hz, 1H), 5.35 (t, J = 10.09, 9.48 Hz, 1H), 4.92 (t, J = 9.82, 9.36 Hz, 1H), 4.17 (dd, J = 4.33, 8.12, 4.29 Hz, 1H), 4.01 (m, 2H), 3.55 (t, J= 9.85, 9.15 Hz, 1H), 2.16 (s, 3H), 2.02 (s, 3H), 1.98 (s, 3H), 1.96 (s, 3H). ^{13}C NMR data: δ 169.98, 169.78, 169.32, 168.65, 90.10, 71.60, 70.33, 67.79, 61.26, 52.12, 20.96, 20.87, 20.51, 20.36.

1,3,4,6-tetra-*O*-acetyl-2-deoxy-2-[(4-{[(1-methyl-4-{[(1-methyl-4-nitro-1*H*-pyrrol-2-yl)carbonyl]amino}-1*H*-pyrrol-2-yl)carbonyl]amino}butanoyl)amino] hexopyranose (19). In a flask flushed with Ar, **18** (3.23 g, 8.4 mmol) was dissolved in 70 mL anhydrous DMF along with EDCI (2.49 g, 1.5 eq.), DMAP (2.57 g, 2.5 eq.), HOBT (3.86 g, 3 eq.), and $CuCl_2$ (0.113 g, 0.1 eq.). Once in solution, the carboxylic acid **14** (3.16 g, 8.4 mmol) was then added and the solution was allowed to stir at room temperature over two days until disappearance of starting material as indicated by TLC (3:1, EtOAc: MeOH). The solution was diluted with 100 mL DCM and the organic solutions was extracted with H_2O (2x, 100 mL), saturated $NaHCO_3$ (2x, 150 mL), and 1M HCl (2x, 150 mL). The organic layer was dried over $MgSO_4$. The resulting solution was rotary evaporated until a solid began to fall out of solution. The solution was then warmed up again until the solid redissolved and was then cooled slowly in the freezer when yellow solid

crystals fell out of solution, which were filtered and dried to give the product **19** (5.35 g, 90.1 %): mp 174-177 °C. TLC (6:1, $CHCl_3$: MeOH) $R_f = 0.51$. ^1H NMR data: δ 10.23 (s, 1H), 8.17 (s, 1H), 8.05 (m, 2H), 7.56 (d, J = 1.70 Hz,1H), 7.19 (s, 1H), 6.84 (s, 1H), 5.70 (d, J = 8.83 Hz, 1H), 5.16 (t, J = 9.90, 10.05 Hz, 1H), 4.88 (t, J = 9.83, 9.71 Hz, 1H), 4.18 (m, 1H), 3.99 (m, 3H), 3.94 (s, 3H), 3.78 (s, 3H), 3.09 (m, 2H), 2.06 (s, 3H), 2.03 (s, 3H), 1.98 (s, 3H), 1.91 (s, 3H), 1.64 (m, 2H). ^{13}C NMR data: δ 172.27, 170.06, 169.62, 169.29, 168.92, 161.13, 156.86, 133.78, 128.23, 126.30, 123.19, 121.32, 118.00, 107.56, 104.00, 91.73, 72.17, 71.52 , 68.06, 61.48, 51.80, 37.97, 37.49, 36.04, 33.26, 25.68, 20.51 (2C), 20.41, 20.32.

1,3,4,6-tetra-*O*-acetyl-2-deoxy-2-[(3-{[(1-methyl-4-{[(1-methyl-4-nitro-1*H*-pyrrol-2-yl)carbonyl]amino}-1*H*-pyrrol-2-yl)carbonyl]amino}propanoyl)amino] hexopyranose (20). Compound **20** was synthesized using a procedure similar to the one described above for **19** using 1.00 g (2.75 mmol) of **15** and 1.05 g (2.75 mmol) of **18** to give **20** (1.57 g, 82 %): mp 169-173 ° C. TLC (6:1, $CHCl_3$:MeOH) $R_f = 0.53$. ^1H NMR data: δ 10.27 (s, 1H), 8.19 (d, J = 1.78 Hz, 1H), 8.12 (d, J = 9.17 Hz, 1H), 8.05 (t, J = 5.52, 5.67 Hz, 1H), 7.59 (d, J = 1.87 Hz, 1H), 7.22 (d, J = 1.68 Hz, 1H), 6.82 (d, J = 1.75, 1H), 5.72 (d, J = 8.84 Hz, 1H), 5.18 (t, J = 9.65, 10.26 Hz, 1H), 4.89 (t, J = 9.77, 9.71 Hz, 1H), 4.19 (dd, J = 4.42, 8.04, 4.28 Hz, 1H), 3.98 (m, 3H), 3.95 (s, 3H), 3.80 (s, 3H), 3.30 (q, J = 6.40, 6.95 Hz, 2H), 2.31 (t, J = 7.45, 7.19 Hz, 2H), 2.01 (s, 3H), 2.00 (s, 3H), 1.97 (s, 3H), 1.88 (s, 3H). ^{13}C NMR data: δ 170.80, 170.05, 169.61, 169.28, 168.88, 161.14, 156.86, 133.77, 128.24, 126.29, 122.99, 121.38, 118.13, 107.63, 103.99, 91.68, 72.16, 71.51, 68.02, 61.49, 51.87, 37.50, 36.05, 35.55, 35.39, 20.52 (2C), 20.42, 20.25.

1,3,4,6-tetra-*O*-acetyl-2-{[4-({[4-({[4-(acryloylamino)-1-methyl-1*H*-pyrrol-2-yl]carbonyl}amino)-1-methyl-1*H*-pyrrol-2-yl]carbonyl}amino)butanoyl]amino}-2-deoxyhexopyranose (21). Compound **19** (1.00 g, 1.425 mmol) was dissolved in 100 mL EtOH

and 500 mg Pd/C was added in a parr jar. The mixture was shaken under pressurized hydrogen (75 psi) until **19** had disappeared as indicated by TLC (3:1, EtOAc:MeOH). The Pd/C was filtered out through celite, and the solution was concentrated by rotary evaporation and the residue was kept under vacuum overnight. The flask with the residue was flushed with Ar, and the residue was dissolved in anhydrous THF (30 mL) and anhydrous DIEA (0.869 mL) was added to the solution. The solution was bubbled with Ar and cooled to -40 °C. Acryloyl chloride (0.116 mL) was added dropwise to the solution which was then allowed to stir below -20 °C, protected from light, until the reaction was complete as indicated by TLC (6:1, $CHCl_3$:MeOH). The solution was concentrated by rotary evaporation and was dissolved in 30 mL DCM and extracted with 20 mL H_20. The DCM layer was crystallized to give pure **21** (0.3845 g, 37 %): mp 130-135 °C. TLC (6:1, $CHCl_3$:MeOH) $R_f = 0.48$. ^{1}H NMR data: δ 10.09 (s, 1H), 9.88 (s, 1H), 8.01 (m, 2H), 7.25 (d, J = 1.78 Hz, 1H), 7.17 (d, J = 1.77 Hz, 1H), 6.90 (d, J = 1.83 Hz, 1H), 6.84 (d, J = 1.80 Hz, 1H), 6.36 (m, 1H), 6.17 (m, 1H), 5.70 (d, J = 8.85 Hz, 1H), 5.64 (dd, J = 2.80, 7.94, 2.22 Hz, 1H), 5.16 (t, J = 10.35, 9.60 Hz, 1H), 4.88 (t, J = 9.52, 9.70 Hz, 1H), 4.17 (dd, J = 4.75, 8.08, 4.20 Hz, 1H), 4.00 (m, 4H), 3.83 (s, 3H), 3.78 (s, 3H), 3.10 (q, J = 6.65, 6.68, 6.32 Hz, 2H), 2.04 (s, 3H), 2.00 (s, 3H), 1.96 (s, 3H), 1.91 (s, 3H), 1.64 (p, J = 7.32, 7.20, 7.08, 7.37 Hz, 2H). ^{13}C NMR data: δ 172.23, 170.05, 169.61, 169.28, 168.91, 161.65, 161.24, 158.30, 131.50, 125.62, 123.03, 122.90, 121.99, 121.69, 118.44, 117.87, 104.16, 103.98, 91.74, 72.18, 71.52, 68.06, 61.48, 51.80, 37.94, 36.19, 35.96, 33.27, 25.72, 20.42 (3C), 20.33.

1,3,4,6-tetra-*O*-acetyl-2-{[3-({[4-({[4-(acryloylamino)-1-methyl-1*H*-pyrrol-2-yl]carbonyl}amino)-1-methyl-1*H*-pyrrol-2-yl] carbonyl}amino)propanoyl]amino}-2-deoxyhexopyranose (22). Compound **22** was synthesized using a procedure similar to the one described above for **21** using 0.500 g (0.723 mmol) of **20** and 59 μL (0.723 mmol) of acryloyl

chloride to give **22** (0.327 g, 63 %): mp 137-137 ° C. TLC (6:1, $CHCl_3$:MeOH) R_f = 0.30. ^{1}H NMR data: δ 10.11 (s, 1H), 9.90 (s, 1H), 8.09 (d, J = 9.16 Hz, 1H), 7.97 (t, J = 5.28, 5.69 Hz, 1H), 7.26 (d, J = 1.59 Hz, 1H), 7.19 (d, J = 1.61, 1H), 6.90 (d, J = 1.72 Hz, 1H), 6.80 (d, J = 1.67 Hz, 1H), 6.35 (m, 1H), 6.17 (dd, J = 1.99, 14.91, 2.15 Hz, 1H), 5.71 (d, J = 8.84 Hz, 1H) 5.66 (dd, J = 1.98, 8.02, 2.16 Hz, 1H), 5.17 (t, J = 9.73, 10.19 Hz, 1H), 4.88 (t, J = 9.80, 9.76 Hz, 1H), 4.18 (dd, J = 4.63, 8.07, 4.30 Hz, 1H), 3.97 (m, 3H), 3.83 (s, 3H), 3.78 (s, 3H), 3.29 (m, 2H), 2.29 (t, J = 7.47, 7.25 Hz, 2H), 1.99 (s, 3H), 1.98 (s, 3H), 1.96 (s, 3H), 1.87 (s, 3H). ^{13}C NMR data: δ 170.81, 170.05, 169.62, 169.29, 168.89, 161.64, 161.23, 158.29, 131.49, 125.62, 123.01, 122.70, 122.03, 121.68, 118.44, 118.00, 104.10, 103.99, 91.68, 72.16, 71.51, 68.01, 61.48, 51.85, 41.34, 36.19, 35.97, 35.58, 35.37, 20.52, 20.42 (2C), 20.25.

Ethyl 4-({[4-({[4-(acryloylamino)-1-methyl-1*H*-pyrrol-2-yl]carbonyl}amino)-1-methyl-1*H*-pyrrol-2-yl]carbonyl}amino)butanoate (23). The NO_2PyPyγOEt **7** (1.45 g, 3.58 mmol) was dissolved in ethanol (200 mL) and Pd/C (500 mg) was added in a parr jar. The mixture was shaken under pressurized hydrogen (75 psi) until **7** had disappeared as indicated by TLC (EtOAc). The Pd/C was filtered out through celite, and the solution was rotary evaporated down to afford the amine which was kept under vacuum overnight. The amine flushed with Ar was dissolved in anhydrous THF (44 mL) and anhydrous DIEA (2.07 mL) was added to the solution. The solution was bubbled with Ar and cooled to -40 °C. Acryloyl chloride (0.292 mL) was added dropwise to the solution and allowed to stir below -20 °C protected from light until the reaction was complete as indicated by TLC (6:1, $CHCl_3$:MeOH). The solution was concentrated by rotary evaporation dissolved in 30 mL EtOAc and extracted with 30 mL H_20. The EtOAc layer was dried over $MgSO_4$ and concentrated by rotary evaporation to give pure **23** (1.40 g, 91 %): mp 70-75 °C. TLC (EtOAc) R_f = 0.29. ^{1}H NMR data: δ 10.09 (s, 1H), 9.89 (s,

1H), 8.03 (t, J = 5.46, 5.59 Hz,1H), 7.26 (d, J = 1.41 Hz, 1H), 7.17 (d, J = 1.32 Hz, 1H), 6.90 (d, J = 1.52 Hz, 1H), 6.85 (d, 1.45 Hz, 1H), 6.36 (m, 1H), 6.19 (dd, J = 1.90, 15.04, 1.99 Hz, 1H), 5.65 (dd, J = 1.87, 2.02 Hz, 1H), 4.03 (q, J = 7.11, 7.11, 7.13 Hz, 2H), 3.83 (s, 3H), 3.78 (s, 3H), 3.17 (q, J = 6.49, 6.11, 6.47 Hz, 2H), 2.29 (t, J = 7.80, 7.49 Hz, 2H), 1.72 (p, J = 6.97, 7.08, 7.06, 7.04 Hz, 2H), 1.16 (t, J = 7.10, 7.10 Hz, 3H). ^{13}C NMR data: δ 172.73, 161.64, 161.30, 158.29, 131.51, 125.59, 123.04, 122.89, 122.00, 121.70, 118.44, 117.87, 104.18, 103.96, 59.76, 37.71, 36.20, 35.96, 31.06, 24.72, 14.13.

1,3,4,6-tetra-*O*-acetyl-2-deoxy-2-({4-[({1-methyl-4-[({1-methyl-4-[(3-sulfopropanoyl)amino]-1*H*-pyrrol-2-yl}carbonyl)amino]-1*H*-pyrrol-2-yl}carbonyl)amino]butanoyl}amino)hexopyranose (25). The alkene **21** (1.00 g, 1.37 mmol), was dissolved in THF (20 mL) and H_2O (80 mL) was added. To the solution, 45 % NH_4HSO_3 (6 mL) and 50 % H_2O_2 (0.5 mL) were added and the solution was allowed to reflux for 16 hours at which time TLC (6:1, $CHCl_3$, MeOH) indicated the complete disappearance of **21** and the formation of a baseline spot. The solution was concentrated by rotary evaporation and the solid was washed with MeOH, which gave a yellow solution leaving behind a white salt. The yellow solution was concentrated by rotary evaporation to give a yellow solid that was dissolved in a minimal amount of H_2O. The solution was acidified to a pH of approximately 1 with concentrated H_2SO_4 and the aqueous solution was evaporated off. Flash column chromatography (1:1:1, EtOAc:DCM:MeOH) was used to purify the resulting solid to give slightly impure **25** (294 mg, 26.5 %): mp – decomposition 170 °C. TLC (1:1:1, EtOAc:DCM:MeOH) R_f = 0.57. ^{1}H NMR data: δ 10.02 (s, 1H), 9.91 (s, 1H), 8.09 (m, 2H), 7.11 (s, 1H), 7.09 (s, 1H), 6.80 (s, 2H), 5.20 (m, 1H), 5.00 (m, 1H), 4.85 (m, 1H), 4.15 (m, 2H), 4.00

(m, 2H), 3.82 (s, 3H), 3.78 (s, 3H), 3.12 (m, 2H), 2.68 (m, 2H), 2.57 (m, 2H), 2.15 (m, 2H), 1.75 (m, 2H).

2-deoxy-2-[(4-{[(1-methyl-4- {[(1-methyl-4-nitro-1H-pyrrol-2-yl)carbonyl] amino}-1H-pyrrol-2-yl)carbonyl]amino}butanoyl)amino]hexopyranose (26). Compound **19** (10 mg) was dissolved in 7 N NH_3 in MeOH and allowed to stir at room temperature until **19** had disappeared as indicated by TLC (6:1, $CHCl_3$:MeOH). The solution was concentrated by rotary evaporation to give **26**. ^{1}H NMR data: δ 10.23 (s, 1H), 8.17 (d, J = 2.0 Hz, 1H), 8.09 (s, 1H), 7.63 (m, 1H), 7.57 (d, J = 2.0 Hz, 1H), 7.20 (d, J = 2.0 Hz, 1H), 6.85 (d, J = 2.0 Hz, 1H), 4.91 (m, 2H), 4.41 (m, 1H), 3.94 (s, 3H), 3.80 (s, 3H), 3.58 (m, 2H), 3.16 (m, 2H), 2.15 (m, 2H), 1.75 (m, 2H).

2-[(5-{[(5-{[(4-ethoxy-4-oxobutyl)amino]carbonyl}-1-methyl-1*H*-pyrrol-3-yl)amino]carbonyl}-1-methyl-1*H*-pyrrol-3-yl)amino]-2-oxoethanesulfonic acid (27). To a solution of alkene ester **23** (1.4 g, 3.26 mmol) in 65 mL of 4:1 EtOH: H_2O, $NaHSO_3$ (678 mg, 6.52 mmol) in 10 mL of H_2O was added. The pH was adjusted to about 8 with 5 % NaOH and the mixture was refluxed until **23** had disappeared and a spot appeared on the baseline by TLC (6:1 $CHCl_3$: MeOH). The solution was cooled in an ice bath and concentrated HCl was added until the pH was about 1. The solvents were removed by rotary evaporation and the residue was dried under vacuum. The residue was then stirred in EtOH and the insoluble white salt was filtered off. This yellow solution was then concentrated by rotary evaporatin and dried to afford the sulfonic acid ester **27** (1.48 g, 89 %): TLC (1:1:1 EtOAc:DCM:MeOH) R_f= 0.57. ^{1}H NMR data: δ 9.96 (s, 1H), 9.85 (s, 1H), 8.03 (t, 1H), 7.16 (d, 1H), 7.14 (d, 1H), 6.84 (s, 2H), 4.03 (q, 2H), 3.80 (s, 3H), 3.77 (s, 3H), 3.16 (m, 2H), 2.66 (m, 2H), 2.54 (m, 2H), 2.30 (t, 2H), 1.70 (q, 2H), 1.16 (t, 3H).

4-[({1-methyl-4-[({1-methyl-4-[(sulfoacetyl)amino]-1*H*-pyrrol-2l}carbonyl) amino]-1*H*-pyrrol-2-yl}carbonyl)amino]butanoic acid (29). The sulfonic acid ester **27** (1.48 g, 3.26 mmol) was dissolved in EtOH (60 mL) and NaOH (832 mg) dissolved in H_2O (50 mL) was added. This solution was allowed to stir until the disappearance of **27** indicated by TLC (1:1, EtOAc:MeOH). The solution was then cooled in an ice bath and concentrated HCl was added until acidic (pH approximately 1). The solution was concentrated by rotary evaporation to give a white and yellow solid, which was sitrred with EtOH which resulted in a yellow solution and an insoluble white solid. The solid was filtered off and the yellow solution was rotary evaporated down to give pure **29** (1.31 g, 81 %): mp 184-188 °C. TLC (1:1:1 EtOAc:DCM:MeOH) R_f= 0.55. ^{1}H NMR data: δ 9.97 (s, 1H), 9.85 (s, 1H), 8.03 (t, 1H), 7.16 (d, 1H), 7.14 (d, 1H), 6.84 (s, 2H), 3.80 (s, 3H), 3.77 (s, 3H), 3.18 (m, 2H), 2.67 (m, 2H), 2.54 (m, 2H), 2.23 (t, 2H), 1.69 (q, 2H). ^{13}C NMR data: δ 174.37, 168.32, 161.37, 158.43, 122.94, 122.72, 122.11, 122.09, 118.21, 117.88, 104.27, 103.98, 59.82, 47.55, 36.16, 36.00, 32.23, 31.23, 24.81.

[1-methyl-4-[1-methyl-4-(2-propenamido)-imidazole-2-carboxamido]pyrrole-2-carboxamido]propane (31). Compound **31** was previously made in the laboratory following procedures described in literature.[10]

[1-methyl-4-[1-methyl-4-(3-sulfopropanamido)imidazole-2-carboxamido]-pyrrole-2-carboxamido]propane (32). To a solution of alkene **31** (0.500 g, 1.4 mmol) in 50 mL of 4:1 EtOH: H_2O, $NaHSO_3$ (291 mg, 2.8 mmol) in 5 mL of H_2O was added. The pH was adjusted to about 8 with 5 % NaOH and the mixture was refluxed until **31** had disappeared and a spot appeared on the baseline on TLC (EtOAc). The solution was cooled in an ice bath and concentrated HCl was added until the pH was about 1 and the solution was cooled in the freezer for 2 h. A yellow precipitate fell out of solution which was filtered out to give **32** (0.364 g, 59

%): mp 264-268 °C. TLC (1:1:1 EtOAc:DCM:MeOH) R_f = 0.55. ^{1}H NMR data: δ 10.35 (s, 1H), 10.00 (s, 1H), 8.00 (s, 2H), 7.41 (d, J = 0.71 Hz, 1H), 7.20 (s, 1H), 6.93 (s, 1H), 3.92 (s, 3H), 3.78 (s, 3H), 3.10 (d, J = 5.03 Hz, 2H), 2.68 (m, 2H), 2.62 (m, 2H), 1.47 (sextet, J = 7.28, 7.06, 7.34, 7.23, 6.85 Hz, 2H), 0.85 (t, J = 7.37, 6.59 Hz, 3H). ^{13}C NMR data: δ 169.11, 161.15, 155.18, 135.37, 133.77, 123.45, 121.05, 118.01, 113.97, 104.20, 47.10, 36.08, 35.13, 31.77, 22.64, 11.51.

[1-methyl-4-[1-methyl-4-(3-(methoxysulfonyl)-propanamido)imidazole-2-carboxamido]pyrrole-2-carboxamido]propane (33). Compound **32** (0.136 g, 0.309 mmol) was suspended in anhydrous dioxane (22 mL) and warmed to 55 °C in an oil bath. To the suspension, 3-methyl-p-tolyltriazine (100 mg, 2.2 eq) was added and the mixture was allowed to stir for 5 h when a new spot appeared on TLC (EtOAc). The product was purified by flash column chromatography (EtOAc) to give **33** (79.2 mg, 56.6 %) which was stored under Ar in the freezer to prevent hydrolysis: ^{1}H NMR data: δ 10.48 (s, 1H), 9.94 (s, 1H), 8.03 (t, J = 5.70, 5.37 Hz, 1H), 7.41 (s, 1H), 7.26 (s, 1H), 6.89 (s, 1H), 3.92 (s, 3H), 3.85 (s, 3H), 3.78 (s, 3H), 3.66 (t, J = 12.41, 12.73 Hz, 2H), 3.12 (q, J = 6.77, 6.37, 6.16 Hz, 2H), 2.83 (t, J = 13.16, 14.38 Hz, 2H), 1.47 (sextet, J = 7.32, 7.36, 7.21, 7.24, 7.20 Hz, 2H), 0.85 (t, J = 7.42, 7.30 Hz, 3H).

Methyl 1-methyl-4-{[(1-methyl-4-{[(2-oxo-2*H*-chromen-3-yl)carbonyl] amino}-1*H*-pyrrol-2-yl)carbonyl]amino}-1*H*-pyrrole-2-carboxylate (34). This compound was made previously in lab using procedures similar to forming **36**.

1-methyl-4-{[(1-methyl-4-nitro-1*H*-pyrrol-2-yl)carbonyl]amino}-*N*-propyl-1*H*-pyrrole-2-carboxamide (35). To a solution of propylamine (0.281 mL, 3.42 mmol), EDCI (1.01 g, 1.5 eq), DMAP (1.04 g, 2.5 eq), HOBt (1.57 g, 3 eq), and $CuCl_2$ (46 mg, 0.1 eq) in dry DMF (30 mL), the carboxylic acid (**13**) (1.0 g, 3.42 mmol) was added. The solution was allowed

to stir under Ar until **13** had disappeared as indicated by TLC (EtOAc). The solution was diluted with 60 mL EtOAc and washed with H_2O (2x, 60 mL), saturated $NaHCO_3$ (2x, 115 mL), and 1M HCl (2x, 115 mL). The organic layer was dried over $MgSO_4$ and the resulting yellow solution was concentrated by rotary evaporation to give pure **35** (0.627 g, 55%): mp 222-226 °C. TLC (EtOAc) $R_f = 0.43$. ^{1}H NMR data: δ 10.22 (s, 1H), 8.17 (d, J = 1.85 Hz, 1H), 8.05 (t, J = 5.69, 5.68 Hz, 1H), 7.57 (d, J = 1.99 Hz, 1H), 7.19 (d, J = 1.78 Hz, 1H), 6.83 (d, J = 1.86 Hz, 1H), 3.94 (s, 3H), 3.79 (s, 3H), 3.11 (q, J = 6.48, 6.97, 6.43 Hz, 2H), 1.47 (sextet, J = 7.40, 7.34, 7.08, 7.22, 7.33 Hz, 2H), 0.85 (t, J = 7.36, 7.44 Hz, 3H). ^{13}C NMR data: δ 161.11, 156.85, 133.79, 128.24, 126.32, 123.37, 121.30, 117.87, 107.56, 103.92, 37.50, 36.03, 22.60, 11.46.

1-methyl-4-{[(1-methyl-4-{[1-(2-oxo-2*H*-chromen-3-yl)vinyl]amino}-1*H*-pyrrol-2-yl)carbonyl]amino}-*N*-propyl-1*H*-pyrrole-2-carboxamide (36). The NO_2PyPyNHPr **35** (0.200 g, 0.60 mmol) was dissolved in ethanol (20 mL) and Pd/C (160 mg) was added in a parr jar. The mixture was shaken under pressurized hydrogen (75 psi) until **35** had disappeared as indicated by TLC (EtOAc). The Pd/C was filtered out through celite, and the solution was concentrated by rotary evaporation to afford a residue which was kept under vacuum overnight. The residue in the flask was dissolved in dry DMF (10 mL) and EDCI (178 mg, 1.5 eq), DMAP (147 mg, 2 eq), HOBt (276 mg, 3 eq), and $CuCl_2$ (8 mg, 0.1 eq) were added. To this solution coumarin-3-carboxylic acid (114 mg, 0.6 mmol) was added and allowed to stir until the coumarin-3-carboxylic acid as indicated by TLC (EtOAc). The solution was diluted with 20 mL EtOAc and the organic solution was extracted with H_20 (2x, 20 mL), saturated $NaHCO_3$ (2x, 40 mL), and 10 % HCl (2x, 40 mL). The organic layer was dried over $MgSO_4$ and the solution was cooled in the freezer to crystallize the product and the crystals were filtered out to afford the pure yellow **36** (120 mg, 42 %): mp 200-202 °C. TLC (EtOAc) $R_f = 0.45$. ^{1}H NMR data: δ 10.52 (s,

1H), 9.86 (s, 1H), 8.89 (s, 1H), 8.01 (m, 2H), 7.76 (td, J = 1.57, 7.09, 1.16, 5.87, 1.6 Hz, 1H), 7.53 (d, J = 8.31 Hz, 1H), 7.48 (t, J = 6.89, 7.3 Hz, 1H), 7.41 (d, J = 1.78 Hz, 1H), 7.17 (d, J = 1.77 Hz, 1H), 7.09 (d, J = 1.84 Hz, 1H), 6.84 (d, J = 1.85 Hz, 1H), 3.87 (s, 3H), 3.79 (s, 3H), 3.11 (q, J = 6.48, 6.98, 5.98 Hz, 2H), 1.48 (sextet, J = 7.40, 7.33, 7.10, 7.21, 7.29 Hz, 2H), 0.86 (t, J = 7.34, 7.46 Hz, 3H). ^{13}C NMR data: δ 161.21, 160.75, 158.14, 158.09, 153.83, 147.00, 134.15, 130.27, 125.29, 123.29, 123.16, 121.93, 120.78, 119.45, 118.91, 118.60, 117.68, 116.25, 104.37, 104.01, 36.30, 35.95, 22.62, 11.47.

Methyl 4-{[(4-acetyl-1-methyl-1*H*-pyrrol-2-yl)carbonyl]amino}-1-methyl-1*H*-pyrrole-2-carboxylate (37). The nitro compound **9** (0.50 g, 1.63 mmol) was dissolved in 60 mL EtOH and Pd/C (0.40 g) was added in a parr jar. The mixture was shaken under pressurized hydrogen (75 psi) until **9** had disappeared as indicated by TLC (EtOAc). The Pd/C was filtered out through celite, and the solution was rotary evaporated down to afford a residue which was kept under vacuum overnight. The residue in the flask was dissolved in dry THF (15 mL) under Ar and DIEA was added to the solution (0.568 mL). The solution was cooled to -30 °C and Ar was bubbled through the solution for 10 minutes. Acetyl chloride (0.116 mL) was added drop wise to the solution and was allowed to stir as it warmed to room temperature until the residue spot had disappeared as indicated by TLC (EtOAc). The solution was concentrated by rotary evaporation and the solid was dissolved in 15 mL EtOAc. This was extracted with H_2O (2x, 15 mL) and 10 % HCl (2x, 15 mL) and the organic layer was dried over $MgSO_4$. The organic layer was concentrated by rotary evaporation and cooled in the freezer overnight to allow the product to fall out of solution, which was filtered to afford pure **37** (0.300 mg, 58 %): mp 182-183 °C. TLC (EtOAc) R_f = 0.22. ^{1}H NMR data: δ 9.88 (s, 1H), 9.82 (s, 1H), 7.44 (d, J = 1.78 Hz, 1H), 7.13 (d, J = 1.63 Hz, 1H), 6.88 (d, J = 1.87 Hz, 1H), 6.83 (d, J = 1.72 Hz, 1H), 3.82 (s, 3H), 3.80

(s, 3H), 3.72 (s, 3H), 1.96 (s, 3H). ^{13}C NMR data: δ 166.49, 160.81, 158.40, 122.90, 122.47, 122.19, 120.77, 118.51, 118.17, 108.36, 103.88, 50.97, 36.18, 36.09, 23.08.

4-{[(4-acetyl-1-methyl-1*H*-pyrrol-2-yl)carbonyl]amino}-1-methyl-1*H*-pyrrole-2-carboxylic acid (38). Compound **37** (0.185 g, 0.581 mmol) was dissolved in EtOH (10 mL) and NaOH (0.93 g) in H_2O (7.5 mL) was added to the solution. The solution was allowed to reflux until **37** had disappeared as indicated by TLC (EtOAc) in about 3 hr. The solution was concentrated by rotary evaporation and the aqueous solution was cooled in an ice bath. Cold concentrated HCl was added until the pH was about 1 when a yellow precipitate fell out. The precipitate was filtered to give pure **38** (0.177 g, 100 %): mp 135-138 °C. TLC (4:1, $CHCl_3$: MeOH) R_f = 0.31. ^{1}H NMR data: δ 12.14 (s, 1H), 9.85 (s, 1H), 9.82 (s, 1H), 7.40 (d, J = 1.9 Hz, 1H), 7.13 (d, J = 1.75 Hz, 1H), 6.82 (m, 2H), 3.80 (s, 6H), 1.96 (s, 3H). ^{13}C NMR data: δ 166.50, 161.97, 158.36, 122.60, 122.53, 122.16, 120.30, 119.49, 118.12, 108.37, 103.83, 36.15, 36.10, 23.08.

E**thyl 4-{[(4-{[(4-acetyl-1-methyl-1*H*-pyrrol-2-yl)carbonyl]amino}-1-methyl-1*H*-pyrrol-2-yl)carbonyl]amino}butanoate (39).** Compound **39** was synthesized using a procedure similar to the one described above for **37** using 0.500 g (1.23 mmol) of **7** and 88 μL (1.23 mmol) of acetyl chloride to give **39** (0.328 g, 64 %): mp 126-128 ° C. TLC (6:1, $CHCl_3$:MeOH) R_f = 0.52. ^{1}H NMR data: δ 9.85 (s, 1H), 9.83 (s, 1H), 8.02 (t, J = 5.45, 5.74 Hz), 7.16 (d, J = 1.69 Hz, 1H), 7.13 (d, J = 1.71 Hz, 1H), 6.84 (d, J = 1.78 Hz, 1H), 6.82 (d, J = 1.77 Hz, 1H), 4.03 (q, J = 7.15, 7.11, 7.11 Hz, 2H), 3.80 (s, 3H), 3.78 (s, 3H), 3.16 (q, J = 6.57, 5.79, 6.49 Hz, 2H), 2.31 (t, J = 7.44, 7.44 Hz, 2H), 1.96 (s, 3H), 1.72 (quintet, J = 7.16, 7.17, 7.14, 7.33 Hz, 2H), 1.16 (t, J = 7.10, 7.09 Hz, 3H). ^{13}C NMR data: δ 172.73, 166.47, 161.30, 158.36, 122.86, 122.69, 122.10, 122.05, 118.02, 117.84, 104.20, 103.76, 59.77, 37.70, 36.12, 35.95, 31.06, 24.72, 23.07, 14.14.

4-{[(4-{[(4-acetyl-1-methyl-1*H*-pyrrol-2-yl)carbonyl]amino}-1-methyl-1*H*-pyrrol-2-yl)carbonyl]amino}butanoic acid (40). Compound **40** was synthesized using a procedure similar to the one described above for **38** using 0.328 g (0.787 mmol) of **39** to give **40** (0.292 g, 93 %): mp 122-125 ° C. TLC (1:1:1, DCM:EtOAc:MeOH) R_f = 0.24. ^{1}H NMR data: δ 12.02 (s, 1H), 9.85 (s, 1H), 9.82 (s, 1H), 8.03 (t, J = 5.85 Hz, 1H), 7.16 (d, J = 1.73 Hz, 1H), 7.13 (d, J = 1.74 Hz, 1H), 6.84 (d, J = 1.80 Hz, 1H), 6.82 (d, J = 1.83 Hz, 1H), 3.80 (s, 3H), 3.78 (s, 3H), 3.16 (q, J = 6.60, 6.07, 6.62 Hz, 2H), 2.23 (t, J = 7.42, 7.38 Hz, 2H), 1.96 (s, 3H), 1.69 (quintet, J = 6.93, 7.28, 7.32, 7.29 Hz, 2H). ^{13}C NMR data: δ 174.33, 166.49, 161.31, 158.36, 122.91, 122.69, 122.10, 122.03, 118.02, 117.83, 104.81, 103.75, 37.83, 36.13, 35.96, 31.16, 24.76, 23.07.

2-oxo-*N*-propyl-2*H*-chromene-3-carboxamide (43). Compound **43** was synthesized using a procedure similar to the one described above for **35** using 100 mg (0.525 mmol) of coumarin-3-carboxylic acid and 31 mg (0.525 mmol) of N-propylamine to give **43** (83 mg, 68.6 %): mp 119-121 °C. TLC (EtOAc) R_f= 0.63. ^{1}H NMR data: δ 8.49 (s, 1H), 8.69 (m, 1H), 7.96 (d, J = 7.75 Hz, 1H), 7.73 (t, J = 7.35, 6.88 Hz, 1H), 7.50 (d, J = 8.32 Hz, 1H), 7.43 (t, J = 7.60, 7.46, 1H), 3.27 (q, J = 6.90, 6.60, 6.14 Hz, 2H), 1.53 (sextet, J = 7.39, 7.37, 7.27, 7.17, 7.07 Hz, 2H), 0.89 (t, J = 7.33, 7.46 Hz, 3H). ^{13}C NMR data: δ 161.11, 160.51, 153.91, 147.35, 134.08, 130.30, 125.20, 119.22, 118.57, 116.20, 40.88, 22.32, 11.44.

CHAPTER 6. CONCLUSION

The goal of this project was to design and synthesize compounds to produce 3-MeA adducts in insulin producing pancreatic β-cells. The project involved combining an agent capable of causing one kind of damage on DNA (i.e. 3-MeA adduct) with a unit that can target this damaging agent preferentially to pancreatic β-cells. Two different molecules were designed in which the cell targeting glucose unit and the DNA damaging Me-lex unit were to be linked by a three carbon linker or a four carbon linker shown in Figure 6.1. A synthetic route for these molecules was devised comprised of 12-15 steps. Out of the 12-15 steps of the synthesis, both of these molecules have been accomplished to within 2-3 steps of the final product. Compounds **21** and **22** shown below in Figure 6.2 have been synthesized.

A second goal of this project was to make these molecules in a way to introduce the linkers at a later stage. Therefore, the goal was to make compound **11** shown in Figure 6.3. However, this route had to be modified since the ester was unable to be hydrolyzed to the carboxylic acid without the Michael addition at the alkene taking place. Compound **24** was successfully made and there are reports that an amide can be made in the presence of a sulfonic acid this was verified with the test reaction described Scheme 3.23. This process is being further explored in the laboratory.

During the course of the project it was decided to develop a binding assay to measure the strength of binding of the new compounds with the glucose unit attached to the target site on DNA because this is expected to influence the efficiency of methylation of these compounds. A fluorescence assay was decided upon and two different potential fluorescent probes were synthesized in this project as shown in Figure 6.4. Initial experiments have indicated that neither of these compounds fluoresces very strongly and

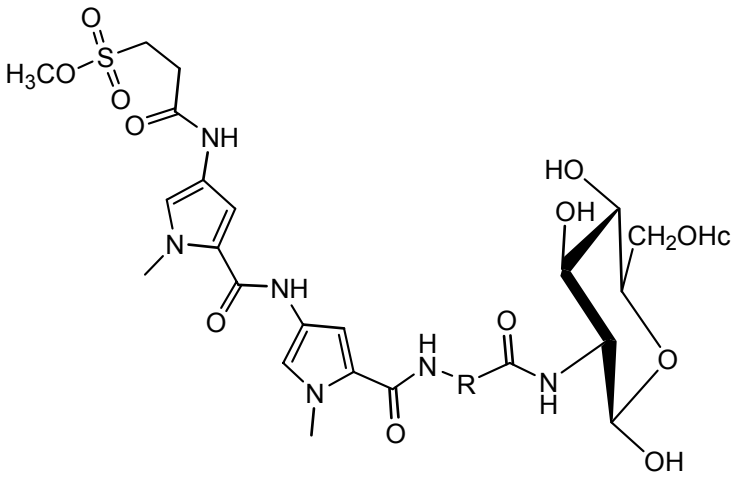
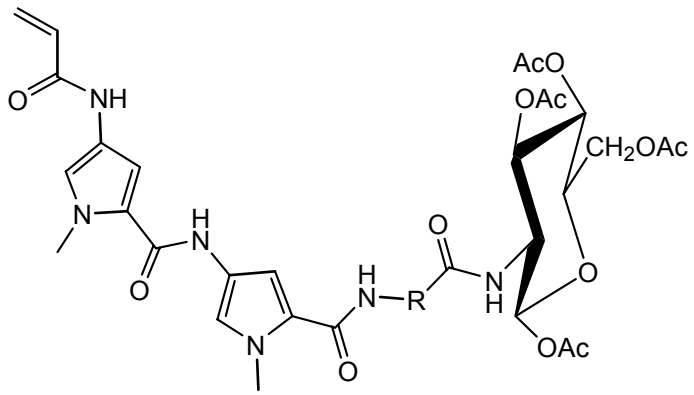
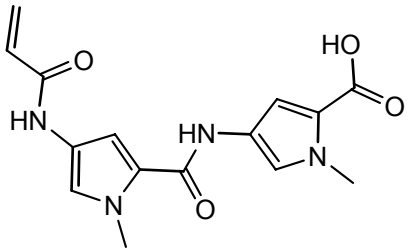


Figure 6.1. Structure of the final compounds to be made where R = $CH_2CH_2CH_2$ or CH_2CH_2.



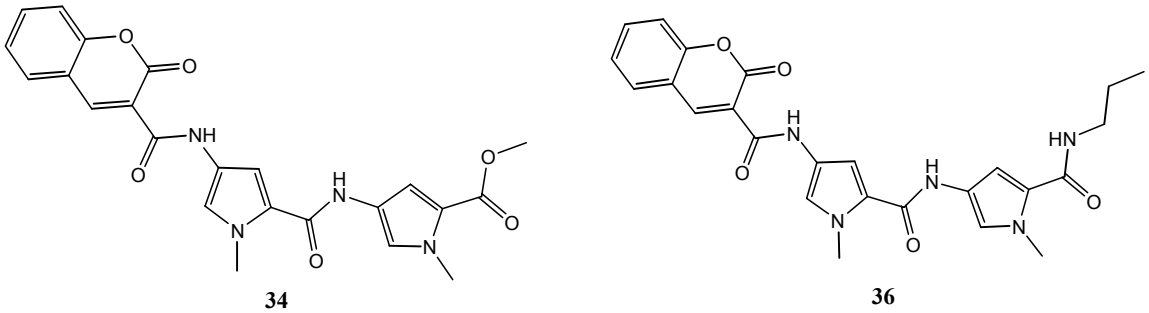
21, R = $CH_2CH_2CH_2$
22, R = CH_2CH_2

Figure 6.2. Structure of the compounds completed thus far.



11

Figure 6.3. Structure of **11**.



34 **36**

Figure 6.4. Structure of **34** and **36**.

therefore, they cannot be used directly for measurements of binding constants of the compounds made in this project. The experiments indicate that the mode of connection of the coumarin to the dipyrroles that quenches the fluorescence and new compounds are being designed shown in Figure 6.5 to identify the best probe to use for these studies.

Once the final compounds which have the DNA-methylating unit and the cell-targeting unit completed these molecules can be tested with DNA for their ability to methylate DNA. A measurement of these compounds obtained from the fluorescence assays can then be used to correlate the strength of the binding to the levels of methylation. When the best candidates are identified cell targeting ability of these compounds will be investigated using cell which have the GLUT-2 transporter and those which do not have the GLUT-2 transporter.

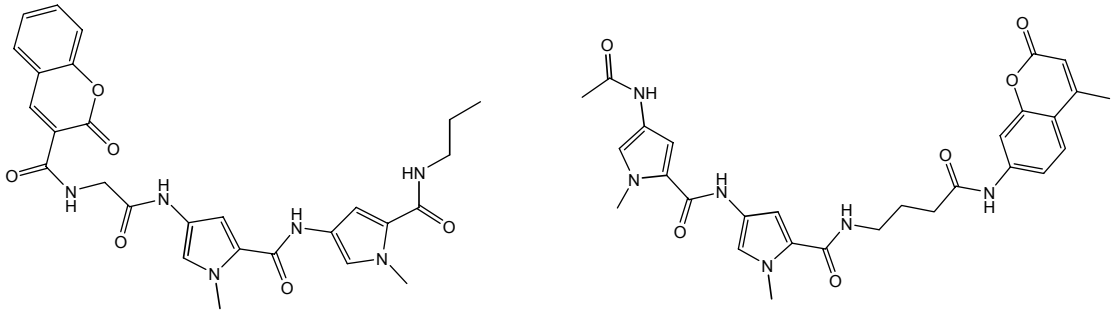


Figure 6.5. Compounds to be made to test for fluorescence.

REFERENCES

1. American Diabetes Association. "All About Diabetes." *Diabetes Information*. www.diabetes.org/about-diabetes.jsp. **2006**.
2. Simone, E.; Eisenbarth, G. S. "Chronic autoimmunity of type I diabetes." *Hormone and Metabolic Research*. **1996**, 28(7), 332-336.
3. Greiner, D.L.; Rossini, A.A.; Mordes, J.P. "Translating Data from Animal Models into Methods for Preventing Human Autoimmune Diabetes Mellitus: Caveat Emptor and Primum non Nocere." *Clinical Immunology*. **2001**, 100(2), 134-143.
4. Cardinal, J. W.; Margison, G. P.; Mynett, K. J.; Yates, A. P.; Cameron, D. P.; Elder R. H. "Increased susceptibility to streptozotocin-induced β-cell apoptosis and delayed autoimmune diabetes in alkylpurine-DNA-*N*-glycosylase-deficient mice." *Mol. Cell. Biol.* **2001**, 21, 5605-5613.
5. Fronza, Gilberto; Gold, Barry. "The Biological Effects of N3-Methyladenine." *Journal of Cellular Biochemistry*. **2004**, 91, 250-257.
6. Tentori, L.; Olindo F.; Fossile E.; Muzi A.;Vergati M.; Portarena I.; Amici C.; Gold, B.; Graziani G. "N3-Methyladenine Induces Early Poly(ADP-Ribosylation), Reduction of Nuclear Factor-κB DNA Binding Ability, and Nuclear Up-Regulation of Telomerase Activity." *Molecular Pharmacology*. **2004**, 67, 572-581.
7. Lawley, P. D.. "Carcinogenesis by Alkylating Agents." *Chemical Carcinogens*. **1984**, 1, 325-484, American Chemical Society, Washington, DC.
8. Beranek, D. T.; Weis, C. C.; Swenson, D. H.. "A Comprehensive Quantitative Analysis of Methylated and Ethylated DNA Using HPLC." *Carcinogenesis*. **1980**, 1, 595-605.

9. Shah, Dharini; Gold, Barry. "Evidence in *Escherichia coli* that N3-Methyladenine Lesions and Cytotoxicity Induced by a Minor Groove Binding Methyl Sulfonate Ester Can Be Modulated *in Vivo* by Netropsin." *Biochemistry.* **2003**, 42, 12610-12626.

10. Varadarajan, Sridhar; Shah, Dharini; Dande, Prasad; Settles, Samuel; Chen, Fa-Xian; Fronza, Gilberto; Gold, Barry. "DNA Damage and Cytotoxicity Induced by Minor Groove Binding Methyl Sulfonate Esters." *Biochemistry.* **2003**, 42, 14318-14327.

11. Monti, P; Iannone, R; Campomenosi, P; Ciribilli, Y; Varadarajan, S; Shah, D; Menichini, P; Gold, B; Fronza, G. "Nucleotide Excision Repair Defect Influences Lethality and Mutagenicity Induced by Me-Lex, a Sequence-Selective N3-Adenine Methylating Agent in the Absence of Base Excision Repair." *Biochemistry*. **2004**, 43, 5592-5599.

12. Engelward, B. P.; Allan, J. M.; Dreslin, A.J.; Kelly, J.D.; Gold, B.; Samson, L.D. "3-Methyladenine DNA Lesions Induce Chromosome Abberations, Cell Cycle Delay, and Apoptosis." *J. Biol. Chem.* **1998**, 273, 5412-5418.

13. Ito, Mikio. "Streptozotocin-induced diabetic animal models." *Sogo Rinsho*. **2004**, 53(4), 1479-1481.

14. Schnedl, W.J.; Ferber, S.; Johnson, J.H.; Newgard, C.B. "STZ transport and cytotoxicity. Specific enhancement in GLUT2-expressing cells." *Diabetes.* **1994**, 43(11), 1326-33.

15. Wang, Z.; Gleichmann, H. "GLUT2 in pancreatic islets: crucial target molecule in diabetes induced with multiple low doses of streptozotocin in mice." *Diabetes*. **1998**, 47(1), 50-6.

16. Elsner, M.; Guldbakke, B.; Tiedge, M.; Munday, R.; Lenzen, S. "Relative importance of transport and alkylation for pancreatic beta-cell toxicity of streptozotocin." Diabetologia (2000), 43(12), 1528-1533.

17. Bolzàn, Alejandro D.; Bianchi, Martha S. "Genotoxicity of Streptozotocin." *Mutation Research.* **2002**, 512, 121-134.

18. Zhang, Yi; Chen, Fa-Xian; Mehta, Pratibha; Gold, Barry. "Groove-and Sequence-Selective Alkylation of DNA by Sulfonate Esters Tethered to Lexitropsins." *Biochemistry.* **1993**, 32, 7954-7965.

19. Elsner, M.; Guldbakke, B.; Tiedge, M.; Munday, R.; Lenzen, S. "Relative importance of transport and alkylation for pancreatic beta-cell toxicity of streptozotocin." *Diabetologia.* **2000**, 43(12), 1528-1533.

20. Elsner, M.; Tiedge, M.; Guldbakke, B.; Munday, R.; Lenzen, S. "Importance of the GLUT2 glucose transporter for pancreatic beta cell toxicity of alloxan." *Diabetologia*. **2002**, 45(11), 1542-1549.

21. Mossman, Brooke T.; Wilson, Glenn L.; Craighead, John E. "Chlorozotocin: a diabetogenic analog of streptozocin with dissimilar mechanisms of action on pancreatic beta cells." *Diabetes*. **1985**, 34(6), 602-10.

22. Tutwiler, G. F.; Bridi, G. J.; Kirsch, T. J.; Burns, H. D.; Heindel, N. D. "Hyperglycemic activity of some non-nitrosated streptozotocin analogs." *Proceedings of the Society for Experimental Biology and Medicine*. **1976**, 152(2), 195-198.

23. Bhuyan, B. K.; Peterson, A. R.; Heidelberger, Charles. "Cytotoxicity, mutations and DNA damage produced in Chinese hamster cells treated with streptozotocin, its analogs, and N-methyl-N'-nitro-N-nitrosoguanidine." *Chemico-Biological Interactions*. **1976**, 13(2), 173-179.

24. Ficsor, Gyula; Zuberi, Riaz I.; Suami, Tetsuo; Machinami, Tomoya. "Differential mutagenicity of streptozotocin analogs of the carbohydrate moiety." *Chemico-Biological Interactions*. **1974**, 8(6), 395-402.

25. Zheng, Gang; Glickson, Jerry D.; Chance, Britton. "Preparation of 2-aminodeoxy-glucose derivatives as antineoplastic agents targeted via GLUT transporters." PCT Int. Appl. **2004**, 71 pp. CODEN: PIXXD2 WO 2004110255 A2 20041223 CAN 142:56618 AN 2004:1124548 CAPLUS

26. Medgyes, Adél; Farkas, Erzsébet; Liptá, András; Pozsgay, Vince. "Synthesis of the Monosaccharide Units of the O-Specific Polysaccharide." *Tetrahedron*. **1997**, 12, 4159-4178.

27. Xiao, Juhauna; Yuan, Gu; Huang, Weiqiang; Chan, Albert; Lee, Daniel. "A Convenient Method for the Synthesis of DNA-Recognizing Polyamides in Solution." *J. Org. Chem*. **2000**, 65, 5506-5513.

28. Grayson, Ian. "Water-soluble carbodiimide – an efficient agent for synthesis." Pharmaceutical Intermediates. **2000**, 86-88.

29. Miyazawa, Toshifumi; Otomatsu, Toshihiko; Fukui, Yoshimasa; Yamada, Takashi; Kuwata, Shigeru. "Racemization-free and Efficient Peptide Synthesis by the Carbodiimide Method using 1-Hydroxybenzotriazole and Copper (II) Chloride simultaneously as Additives." *J. Chem. Soc., Chem. Commun*. **1988**, 419-420.

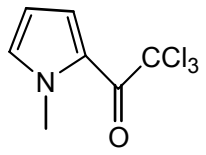
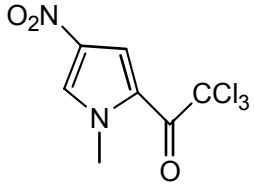
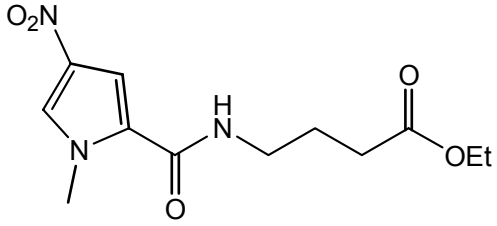
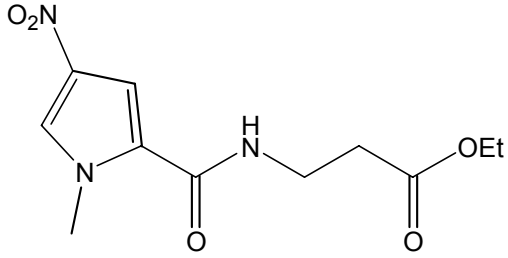
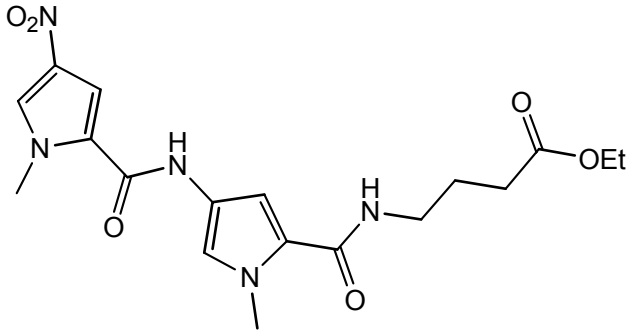
30. Fuhrhop, Jurgen-Hinrich; David, Hans-Hermann; Mathieu, Joachim; Liman, Ulrich; Winter, Hans-Jorg; Boekema, Egbert. "Bolaamphiphiles and Monolayer Lipid Membranes Made from 1,6,19,24-Tetraoxa-3,21-cyclohexatriacontadiene-2,5,20,23-tetrone." *J. Am. Chem. Soc*. **1986**, 108, 1785-1791.

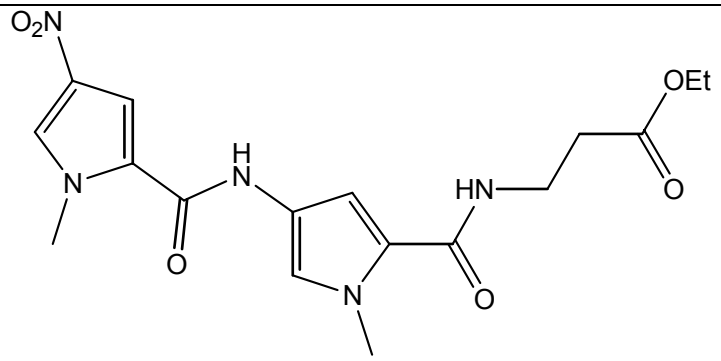
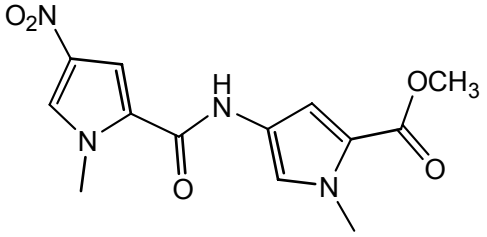
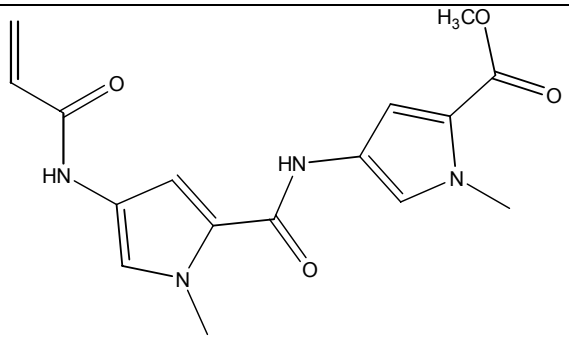
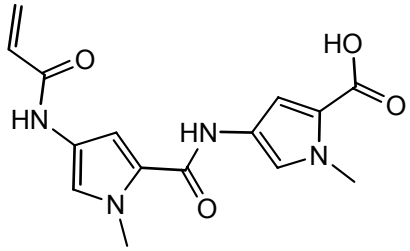
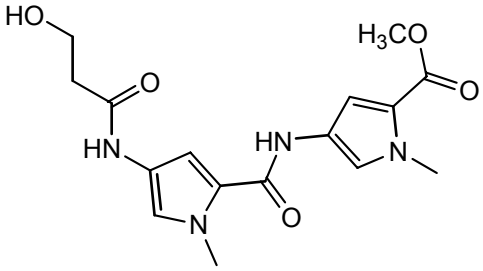
31. Tangallapally, Rajendra; Yendapally, Raghunandan; Lee, Robin; Hevender, Kirk; Jones, Victoria; Lenaerts, Anne; McNeil, Michael; Wang, Yuehong; Franzblau, Scott; Lee, Richard. "Synthesis and Evaluation of Nitrofuranylamides as Antituberculosis Agents." *J. Med. Chem*. **2004**, 47, 5276-5283.

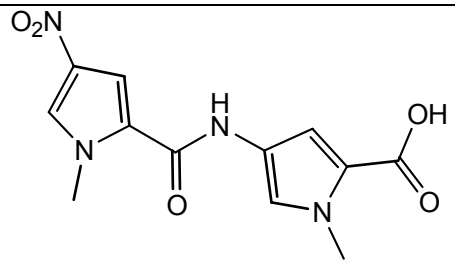
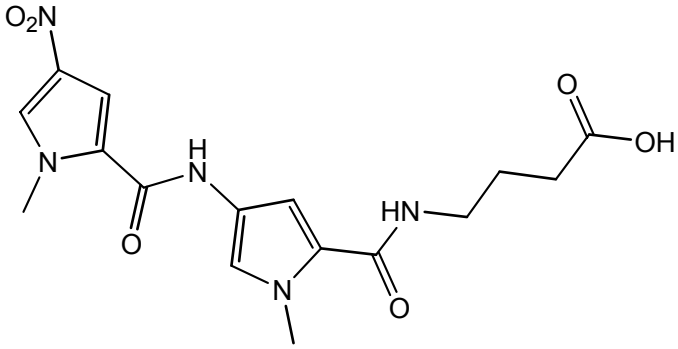
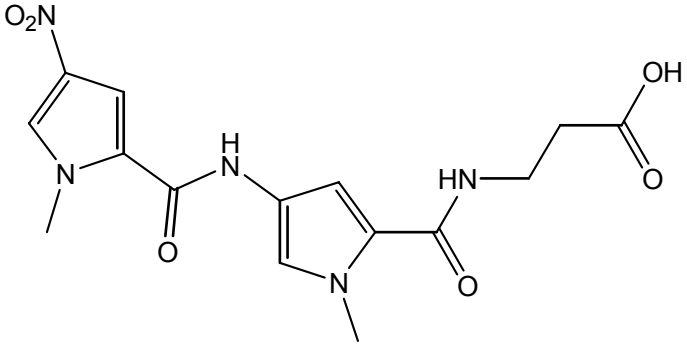
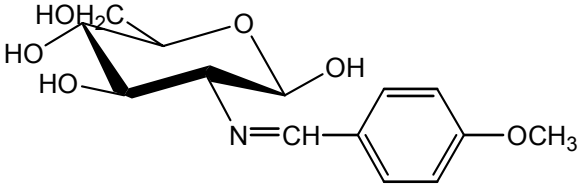
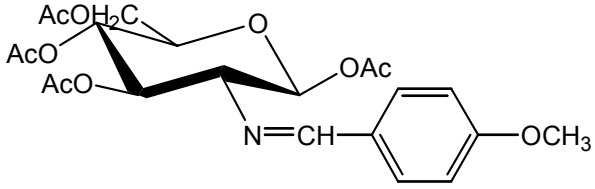
32. Brown, K. A.; He, G.; Bruice, T. C. "Microgonotropens and Their Interactions with DNA. 2. Quantittative Evaluation of Equilibrium Constants for 1:1 and 2:1 Binding of Dien-Microgonotropen-a, -b, and –c as well as Distamycin and Hoechst 33258 to d(GGCGCAAATTTGGCGG)/d(CCGCCAAATTTGCGCC)." *Journal of the American Chemical Society*. **1993**, 115, 7072-7079.

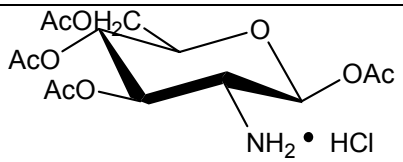
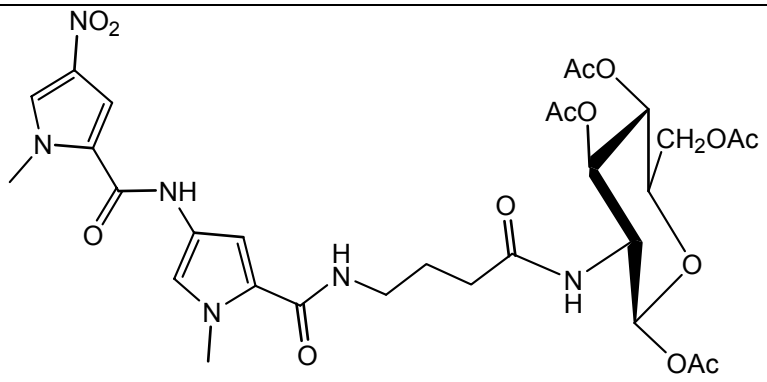
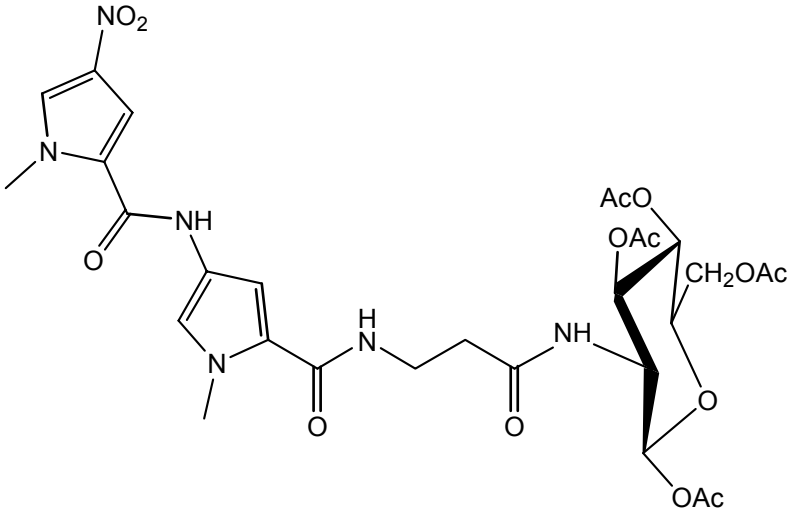
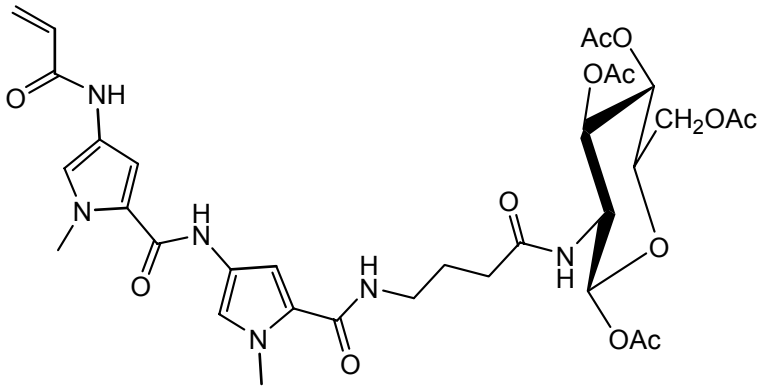
APPENDIX

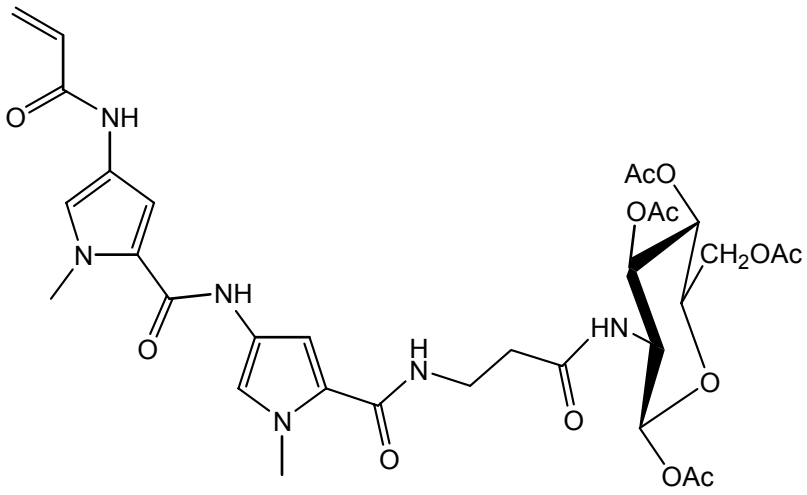
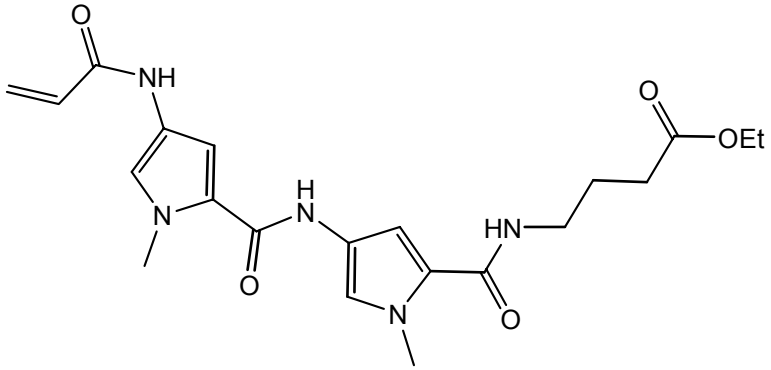
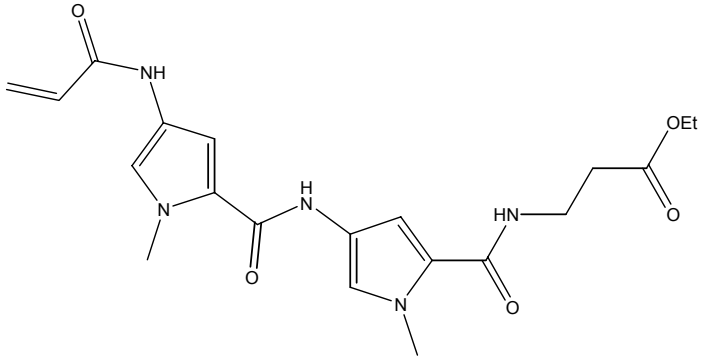
Appendix A. Structure, number, and percent yields of the compounds.

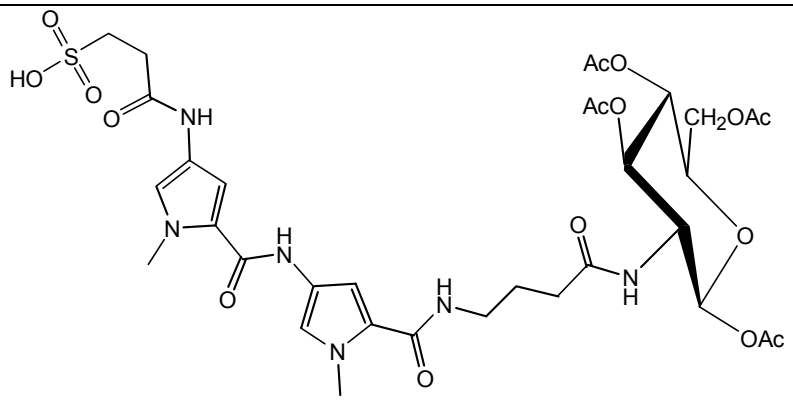
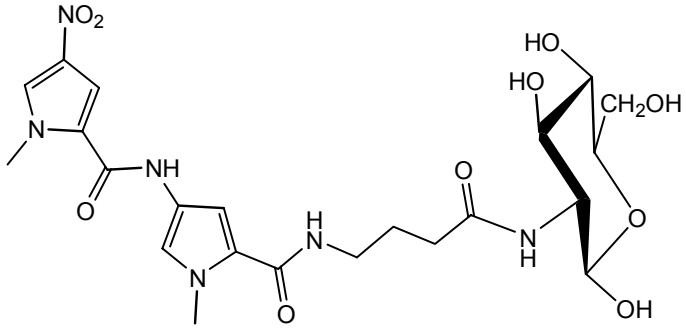
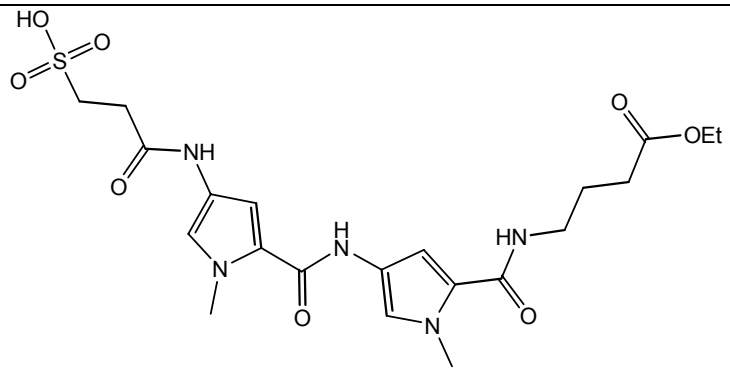
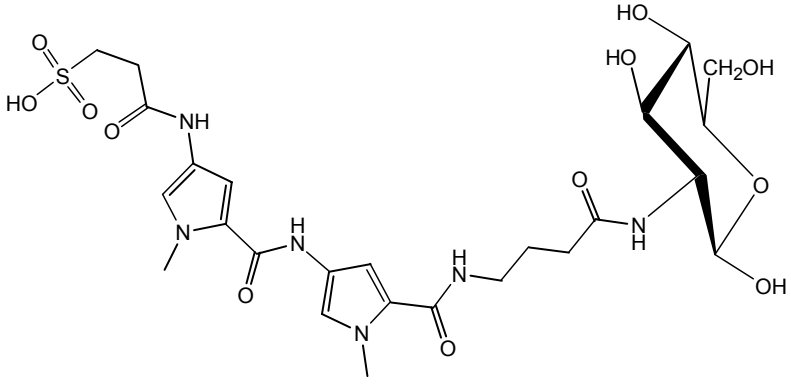
Structure	Number	Percent Yield
	**3**	N/A
	**4**	84 %
	**5**	86 %
	**6**	57 %
	**7**	35 % / 62 %

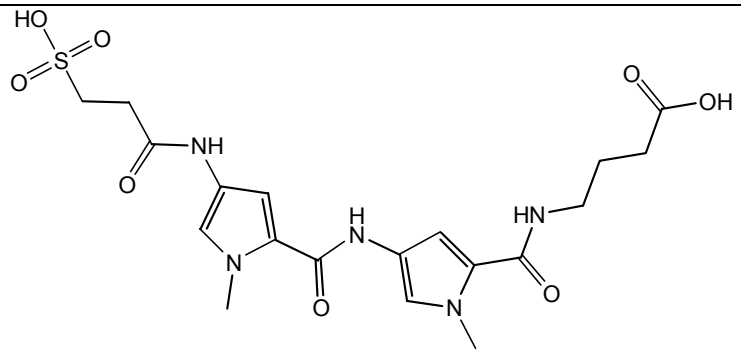
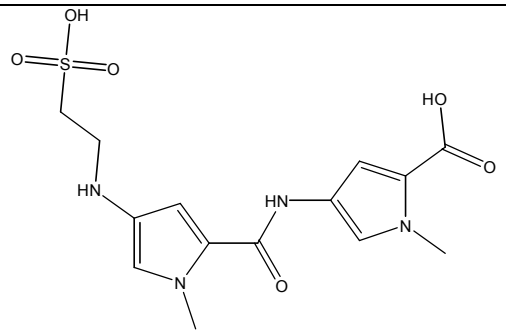
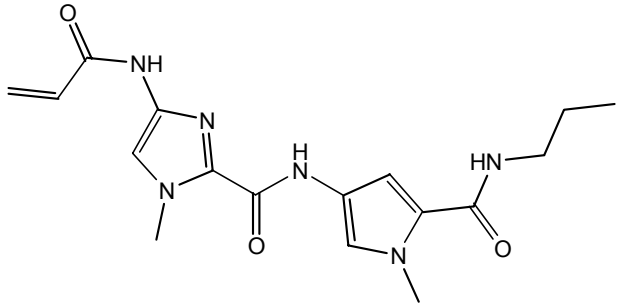
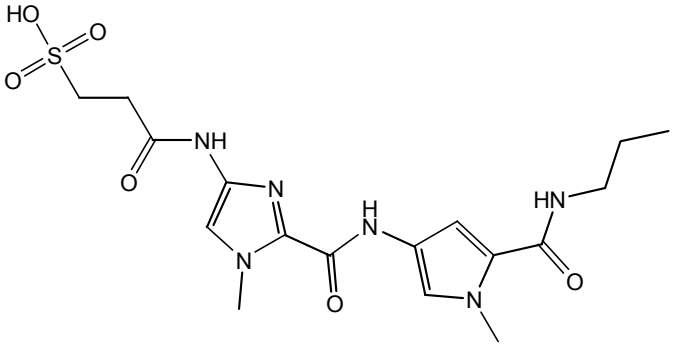
	**8**	32 % / 57 %
	**9**	94 %
	**10**	71 %
	**11**	N/A
	**12**	N/A

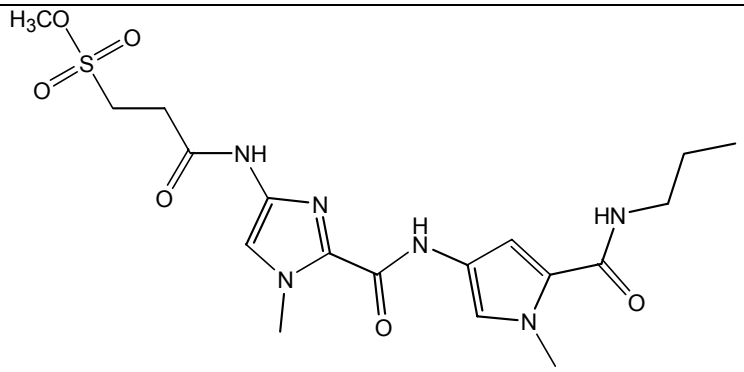
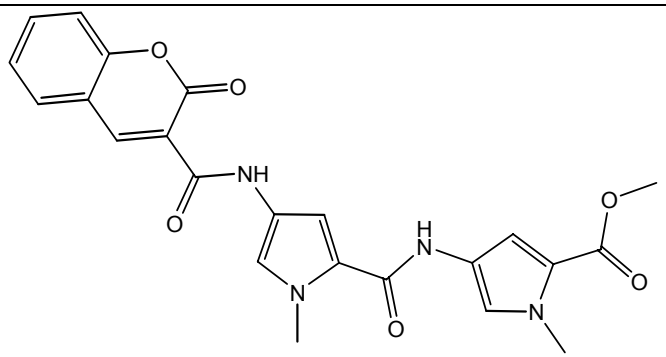
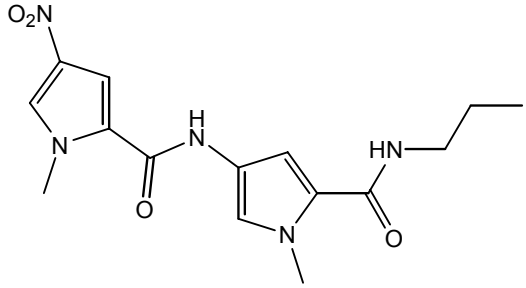
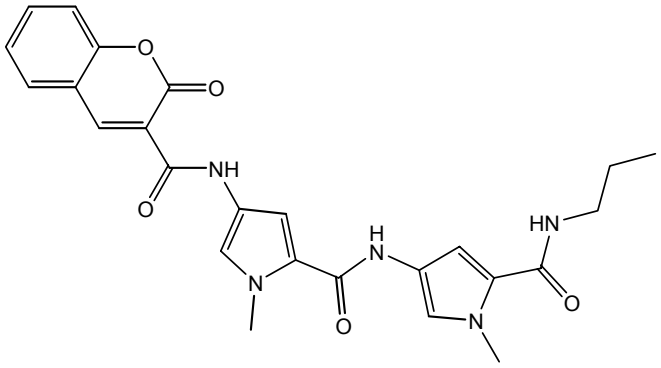
	**13**	87 %
	**14**	92 %
	**15**	99 %
	**16**	75 %
	**17**	79 %

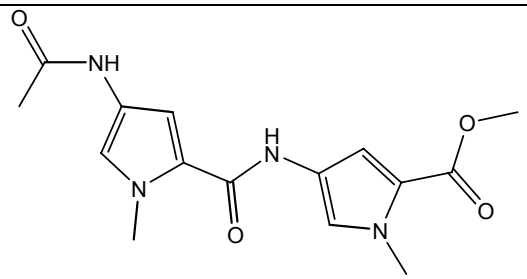
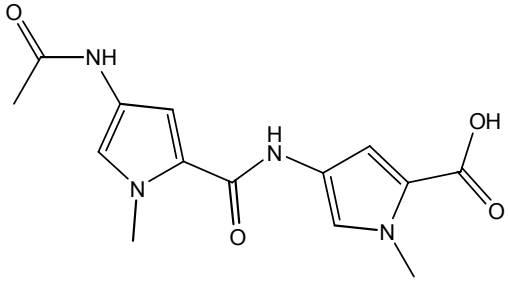
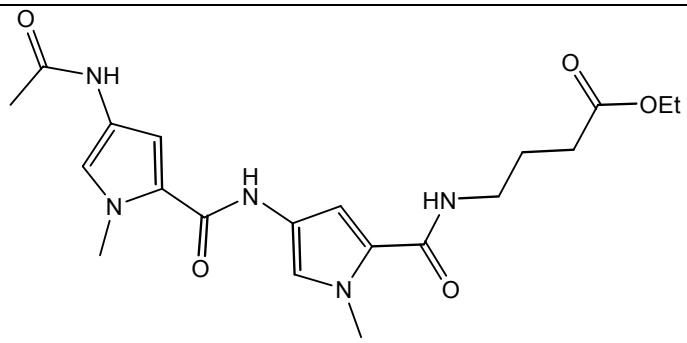
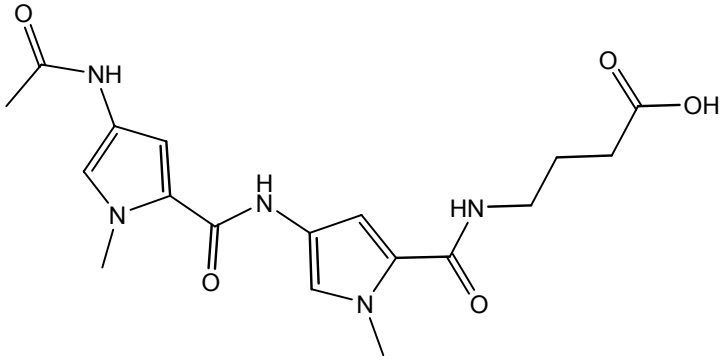
	**18**	81 %
	**19**	90 %
	**20**	82 %
	**21**	37 %

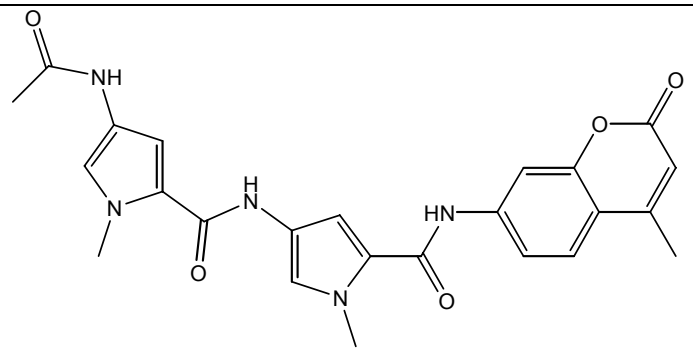
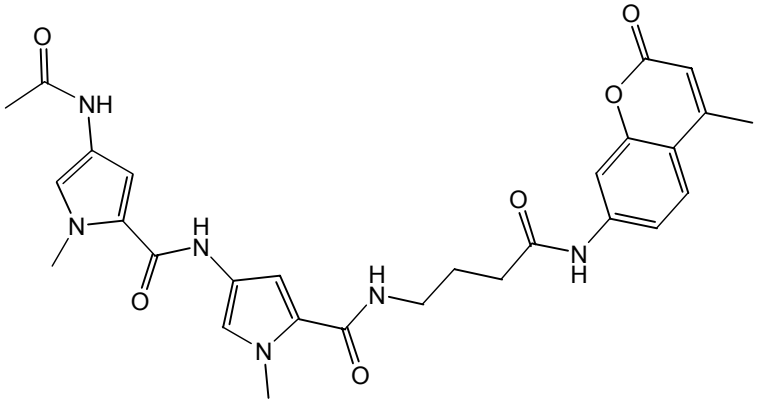
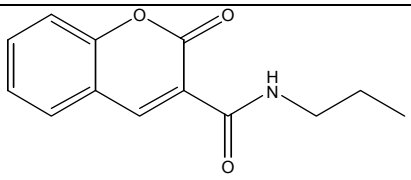
	**22**	63 %
	**23**	91 %
	**24**	N/A

	**25**	26 %
	**26**	N/A
	**27**	89 %
	**28**	N/A

	**29**	81 %
	**30**	N/A
	**31**	N/A
	**32**	59 %

	**33**	57 %
	**34**	N/A
	**35**	55 %
	**36**	42 %

	**37**	58 %
	**38**	100 %
	**39**	43 %
	**40**	93 %

	**41**	N/A
	**42**	N/A
	**43**	68.6 %

Appendix B. List of Abbreviations

3-MeA – 3-methyladenine

3-MeG – 3-methylguanine

6-MeG – 6-methylguanine

A - adenine

A/T – adenine-thymine

C - cytosine

DCM – dichloromethane

DIEA - diisopropylethylamine

DNA – deoxyribose nucleic acid

DMAP - 4-Dimethylaminopyridine

DMF – dimethylformamide

DMSO – methyl sulfoxide

EDCI - 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimidehydrochloride

EtOAc – ethyl acetate

EtOH – ethanol

G – guanine

HCl – hydrochloric acid

HOBT - Hydroxybenzotriazole

Me-lex – methyl lexitropsin

MeOH – methanol

NaCl – sodium chloride

NaOH – sodium hydroxide

NMR – nuclear magnetic resonance

Pd/C – palladium on carbon

STZ – streptozotocin

TEA – triethylamine

THF - tetrahydrofuran

TLC – thin layer chromatography

T - thymine

UV – ultra violet