5-2014

Porphyrin Cross-Linkers for Generating Soluble Molecularly Imprinted Polymers from Polyethyleneimine

Mohammed R. Elshaer
melshaer@fdu.edu

Porphyrin Cross-Linkers for Generating Soluble Molecularly Imprinted Polymers from Polyethyleneimine

Dissertation by
Mohammed Ragab Elshaer

In Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy
in
Chemistry

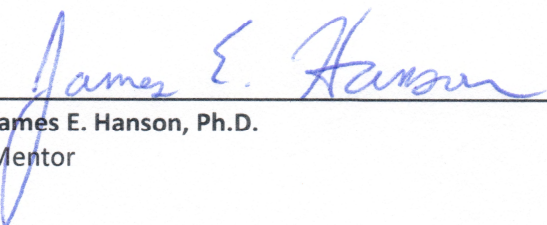Seton Hall University
South Orange, New Jersey

May 2014
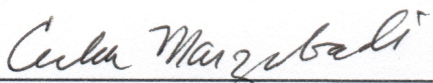
CERTIFICATION

We certify that we have read this thesis and that in our opinion it is adequate in scientific scope and quality as a dissertation for the degree of Doctor of Philosophy

APPROVED

_____
James E. Hanson, Ph.D.
Mentor

_____
Cecilia Marzabadi, Ph.D.
Reader

_____
Cosimo Antonacci, Ph.D.
Reader

_____
Nicholas H. Snow, Ph.D.
Chair, Department of Chemistry and Biochemistry

For my parents, Ragab and Nagwa

Thank you for all your love and support

Acknowledgements

Table of Contents

List of Figures

List of Tables

List of Schemes

List of Equations and Equilibria

Abbreviations and Definitions

DNA	Deoxyribonucleic Acids
RNA	Ribonucleic Acids
mRNA	Messenger Ribonucleic Acids
MIP	Molecularly Imprinted Polymer
A, G, T, C	Adenine, Guanine, Thymine, Cytosine
G4	G-Quadruplex
ssDNA	Single Stranded DNA
dsDNA	Double Stranded DNA
bp	Base pairs
kb	Kilobases
nt	Nucleotides
BMVC	3,6,-bis[2-(1-methypyridinium)vinyl] carbazole diiodide
FRET	Fluorescence Resonance Energy Transfer
SPR	Surface Plasmon Resonance
H_2TMPyP4	5,10,15,20-tetra(N-methyl-4-pyridyl)porphyrin
LCAO	Linear combination of atomic orbitals
HOMO	Highest occupied molecular orbital
LUMO	Lowest unoccupied molecular orbital
TPP	Tetraphenylporphyrin
DDQ	2,3-Dichloro-5,6-Dicyanobenzoquinone
H_2TPyP4	5,10,15,20-Tetra(4-pyridyl)-21H,23H-porphyrin
DCM	Dichloromethane
TFA	Trifluoroacetic acid
PEI	poly(ethyleneimine)
PVP	poly(vinylpyrolidine)
TLC	Thin Layer Chromatography
O.D.	Optical Density
$CHCl_3$	Chloroform
$CDCl_3$	Deuterated Chloroform
PTFE	Polytetrafluoroethylene
EDTA	Ethylenediaminetetraacetic acid
^{1}H NMR	Proton Nuclear Magnetic Resonance
ESI-MS	Electrospray Ionization-Mass Spectrometry
S_N2	Bimolecular nucleophilic substitution reaction
S_N1	Unimolecular nucleophilic substitution reaction
Et	Ethyl
DMF	Dimethylformamide
ddH_2O	Double-distilled water
*m*CPBA	*meta*-Chloroperbenzoic acid
DCM	Dichloromethane

DMSO	Dimethyl sulfoxide
BH$_3$-THF	Borane tetrahydrofuran
Ph$_3$P	Triphenylphosphine
CBr$_4$	Carbon tetrabromide
TEA	Triethylamine
DMT-MM	4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride
DCC	Dicyclohexylcarbodiimide
EDC	*N*-(3-Dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride
HPLC	High Performance Liquid Chromatography
RP	Reverse Phase
SEC	Size Exclusion Chromatography
CT-DNA	Calf Thymus DNA
TBE	Tris-Borate-EDTA buffer
DP	Degree of Polymerization
MW	Molecular Weight
Sulfo-NHS	*N*-Hydroxysulfosuccinimide
KPBS	Potassium Phosphate Buffered Saline
MWCO	Molecular Weight Cut Off
CD	Circular Dichroism

Abstract

Molecular recognition is vital to many biochemical processes and is at the heart of promising bio-medically related technologies. Molecular imprinting has a long-standing history as a successful method for mimicking the molecular recognition phenomena exhibited by nature, whereby artificial receptors are prepared for a given target molecule based on synthetic polymers. The molecularly imprinted polymer (MIP) contains a three dimensional network with a memorized cavity specific to the shape and functionality of the templated target molecule. The utility of traditional MIPs has been limited due to an inherent lack of solubility. We have worked toward developing a system that allows for the preparation of soluble MIPs targeting quadruplex DNA, specifically the human telomeric repeat $(TTAGGG)_4$. To do so we have synthesized a series of *meso*-substituted, water soluble, tetracationic pyridinium porphyrins which we have successfully coupled to polyethyleneimine (PEI), forming a condensation polyamide. We have demonstrated that one of these porphyrins can be used as an efficient polymer cross-linker, which provides a unique quadruplex DNA binding site in the polymer network. Unfortunately, the high cationic charge density found on PEI has been found to elicit potential dilemmas in the utility of this method. Attempts have been made to reduce this charge by increasing the cross-linking agent and partially acetylating the PEI. While the network structure of this soluble cross-linked polymer still requires optimization, it has shown promise and demonstrates the opportunities for new soluble molecularly imprinted polymer designs that include quadruplex binding sites.

Mohammed R. Elshaer
Doctor of Philosophy in Chemistry
Seton Hall University
Dr. James E. Hanson, Mentor

1. Introduction

1.1. Molecular Recognition

Molecular recognition events are ubiquitous and essential to many fundamental processes found in biology and chemistry.[1] The remarkable recognition processes that allows for the ability to selectively recognize complementary target molecules in a vast pool of similar molecules are driven by non-covalent forces such as hydrogen bonding, electrostatic interactions, hydrophobic interactions, van der Waals forces, π-π interactions and conformational energy. Examples of these processes include the binding of an enzyme to its substrate, a protein to a receptor, a drug to a biological target, antibody-antigen recognition, the hybridization of complementary nucleic acid oligomers, and the translation of DNA into mRNA.

Clearly the molecular recognition phenomenon is one of the driving forces that allows for life processes to occur.[2] The underlying mechanisms at the heart of molecular recognition are the chemical and geometrical complementarity between the interacting molecules.[3] Extensive research efforts have been put forth in an attempt to develop synthetic material with similar recognition abilities that mimic these naturally occurring processes.

1.2. Molecular Imprinting

The technique of molecular imprinting has been regarded as one of the most promising strategies for creating materials with molecular recognition capabilities comparable to those found in natural systems. This technique allows for the formation of a three dimensional network which includes a memorized cavity that is specific to the shape and functionality of the template molecule.

1.2.1. Historical perspective of Molecular Imprinting

The origins of molecular imprinting can be traced back to the "lock and key" theory, proposed by Emil Fischer in 1894, where he attributed the specificity of an enzymes to its substrate based on sterics.[4] He wrote "To use a picture, I would like to say that enzyme and glucoside have to fit to each

other like a lock and key in order to exert a chemical effect on each other." Through evolution, enzymes have developed binding sites that are complementary and specific to the shape, size and electron distribution of their particular substrate(s). The binding sites also contain functional groups that co-operatively interact non-covalently with the substrate.

While the molecular recognition demonstrated by enzymes may have taken millennia to develop, the recognition generated by antibodies to an antigen occurs on a time scale of days. In 1932 Mudd proposed his theories on how these antibodies were produced by the body in order to eliminate foreign molecules. He suggested that the formation of the antibody occurred on the surface of the antigen, where amino acids or peptides adapted physically and chemically to the antigen surface at the site of coupling. He stated that the specificity exhibited by the antibody was due to the stereochemcial correspondence with the antigen.[5] In the 1940's, Pauling further developed this theory by proposing a detailed mechanism for their formation, describing the induced folding of a polypeptide chain around an antigen.[6]

Pauling's theory inspired Frank Dickey to carry out experiments toward the development of synthetic antibodies with the aim of obtaining tailor made binding materials. He termed his devised technique "specific adsorption" which would later become known as molecular imprinting. Initially he was interested in the chiral separation of sugars using a silica polymer as the synthetic receptor. However, due to difficulties arising from the solubility of the sugars in the alcohol solvents used, he moved to the separation of dye molecules.[7] He was successful at preparing a material that adsorbed methyl orange 1.4 times greater than ethyl orange.[8]
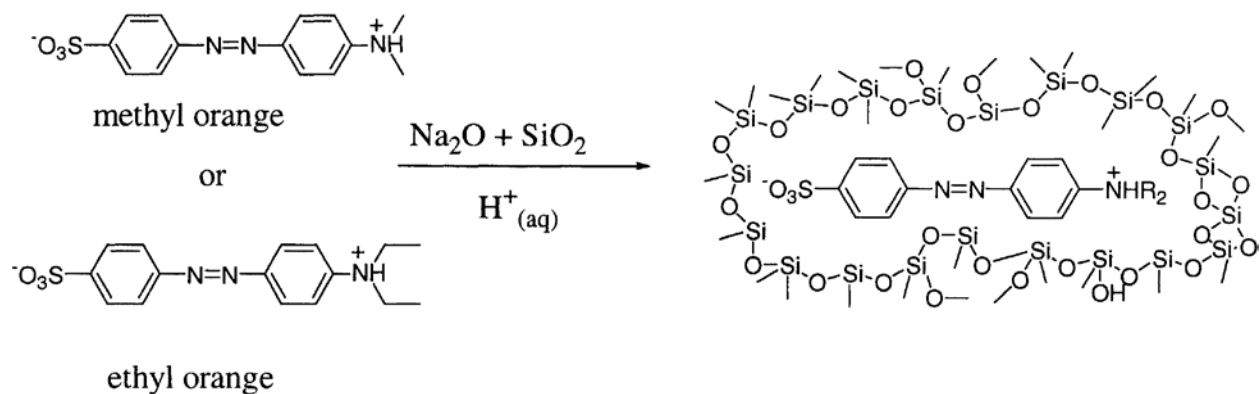
Figure 1. Specific Adsorption for selective binding of different dye molecules by silica polymers

1.2.2. Molecularly Imprinted Polymers

The modern strategy for molecular imprinting utilizes synthetic organic polymers as the imprinting material. This process allows for the preparation of organic polymers containing within them recognition sites for template small molecules.[9] The technique is conceptually simple; it involves the pre-organization of functional monomers around the template molecule through covalent or non-covalent interactions, which are subsequently polymerized in the presence of a cross-linking agent forming an inflexible polymer. After extraction of the template molecule, functional groups in the polymer matrix are arranged at defined positions that are in a spatial arrangement complementary to the template molecule.[10]

Most of the initial imprinting following this modern method were performed using the covalent binding approach developed by Wulff and Sarhan.[11] This approach uses a polymerizable derivative of the template molecule, which is obtained by forming covalent bonds between the template and suitable polymerizable monomers. Extraction of the template from the polymer is performed by chemically cleaving these covalent bonds, which are subsequently reformed during binding with the template molecule.

The alternative approach introduced by Mosbach and co-workers relies on the formation of a pre-polymerization complex that allows monomers with suitable functional groups to self-assemble,

through non-covalent interactions, with the template molecule.[12] After polymerization and cross-linking, the functional groups are held stationary in their position by the polymer matrix. The template can then be removed by simple extraction. The rebinding of the template to the molecularly imprinted polymer (MIP), as a result, will occur through non-covalent interactions.[13] Since most biomolecules interact through non-covalent interactions, this "self-assembly" approach is similar to processes found in nature. A depiction of both approaches can be found in Figure 2. It should be noted that both of these methods typically yield insoluble MIPs in powder or gel forms. Synthetic MIPs are capable of specific molecular recognition and have received much attention due to features such as excellent stability, ability to function in various media and their low cost of synthesis. They have found applications in chromatography, chiral separations, biomimetic sensors, analysis and catalysis.
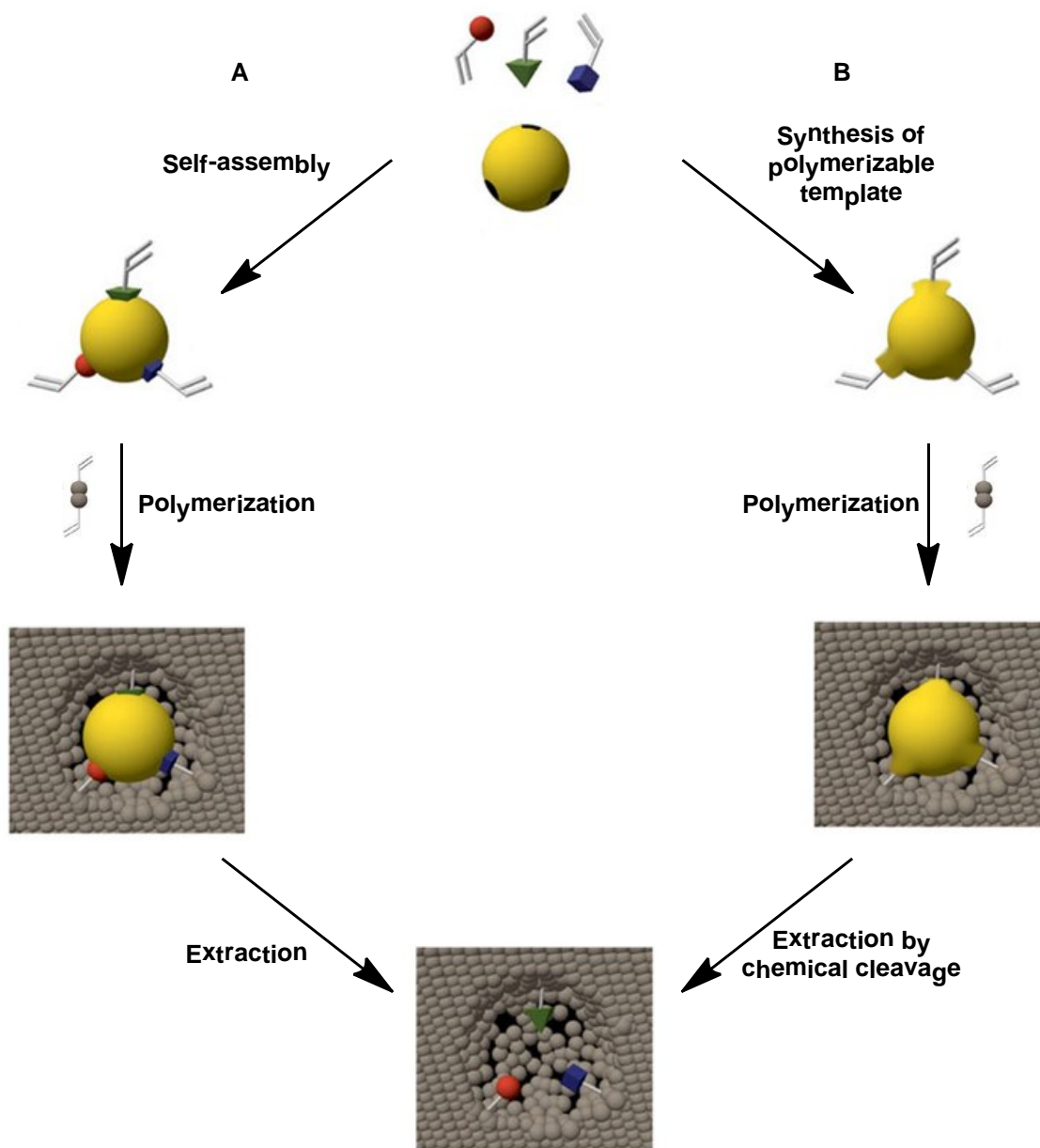
Figure 2. Schematic representation of molecular imprinting using the (A) non-covalent approach and (B) the covalent approach. In the non-covalent approach functional monomers are allowed to self-assemble around the template molecule via non-covalent interactions while in the covalent approach, a polymerizable derivative of the template is used.

1.3. Nucleic Acids

Nucleic acids are natural polyphosphoester biopolymers composed of phosphoric acid and sugars, to each of which a heterocyclic amine, which is referred to as a base, is attached. There are two principal types of nucleic acids, those containing D-ribose are known as ribonucleic acids (RNA), and those containing 2-deoxy-D-ribose are known as deoxyribonucleic acids (DNA).[14]

Each monomeric nucleotide unit contains a phosphate group coupled to the 5'-hydroxyl group on the ribose to yield a phosphate ester. The unit becomes distinct upon formation of a covalent bond, via hydrolysis, between C-1' of the ribose and the heterocyclic amine base, at position N-1 for pyrimidines (Cytosine, Thymine or Uracil) and at position N-9 for purines (Adenine or Guanine) resulting in the creation of an *N*-β-glycosyl bond.

The formation of DNA and RNA results from the polymerization of nucleotides through an esterification reaction between the 5'-phosphate group of one nucleotide and the 3'-hydroxyl group on the sugar of the adjacent nucleotide to form a 3',5'-phosophodiester linkage, yielding the primary sequence of DNA or RNA biopolymer. The adduct contains a backbone of alternating sugar and phosphate groups, and since nucleic acids are acidic they are completely ionized at the near neutral pH that occurs in living organisms, which results with a net charge of -1 per phosphate, making DNA and RNA polyanionic.
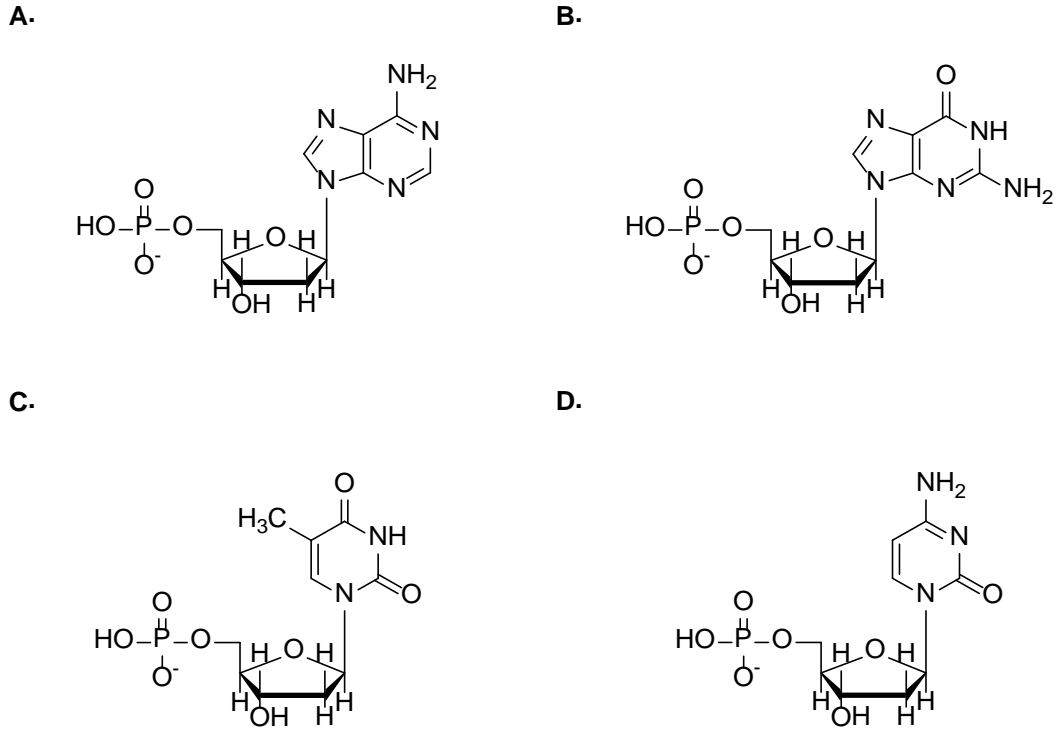
Figure 3 Structures of the monomeric nucleotide units of DNA (A) Adenosine 5'-monophosphate, (B) Guanosine 5'-monophosphate, (C) Thymidine 5'-monophosphate (D) Cytidine 5'-monophosphate

1.3.1. DNA Structure

The Watson and Crick model proposed in 1953 was a historic event in science.[15] Watson and Crick postulated that DNA was found as a helical duplex, two antiparallel polynucleotide chains complementary to each other held together by specific hydrogen bonding between base pairs of adenine (A) with thymine (T) and guanine (G) with cytosine (C). Two hydrogen bonds are formed between A and T and three between G and C, this is referred to as Watson-Crick base pairing. Figure 4 illustrates the original Watson-Crick Model for the structure of DNA. The hydrophobic bases are located along the axis of the double helix with the sugar-phosphate hydrophilic backbone winding along the periphery. The distance between the bases is 3.4 Å. Ten bases are required to make a complete rotation about the helical axis, meaning that one complete rotation about the axis would require 34 Å. They

arrived at their 3-dimensional model of DNA with the assistance of x-ray diffraction data collected on

DNA fibers by Franklin[16] and the A=T and G=C base equivalence identified by Chargaff.[17]

A.

B.
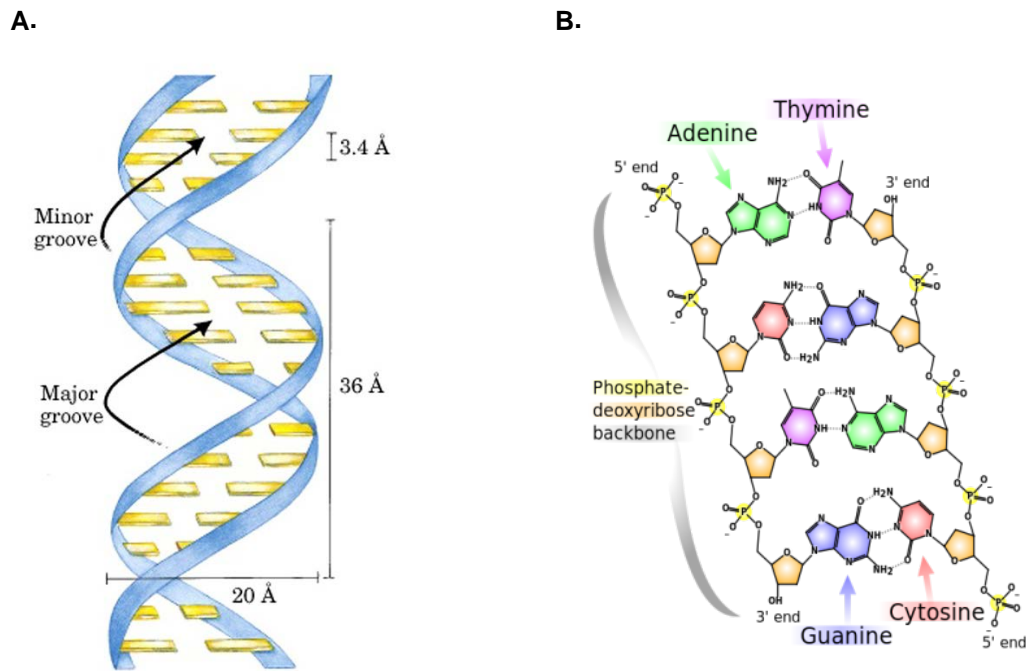
Figure 4. Watson-Crick model for the structure of DNA (A) Schematic representation showing helix dimensions (B) The complimentary antiparallel strands exhibiting the base equivalence of A=T and G≡C and the hydrogen bonding between the bases.[18]

1.3.2. Beyond Watson-Crick B-DNA

Due to the polymorphic nature of DNA, variations of the Watson-Crick DNA structure, referred to as B-DNA, do exist. These variations are possible due to the different conformations that can be taken by the deoxyribose via its rotation about the bonds that make up the phosphodeoxyribose backbone and the free rotation about the C-1'-*N*-glycosyl bond.

It is believed that most of the DNA in the human genome is present in the B-DNA form. However, DNA may form alternative non-B-DNA secondary structures within certain sequences, fueled by various dynamic molecular events.[19] This is typically a function of the nucleic acid sequence, topology, environmental conditions i.e. ionic conditions, protein binding, methylation, carcinogen

binding and other modifications on DNA.[20] Examples of these alternative secondary structures including Z-DNA, triplexes, cruciforms, and hairpins are shown in Figure 5.

A.　　　　　　　**B.**　　　　　　　**C.**　　　　　　　**D.**
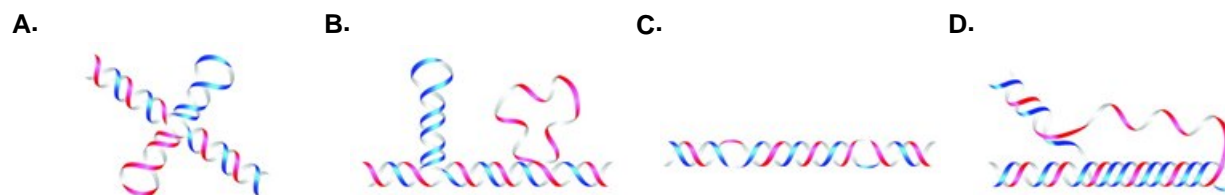
Figure 5. Schematic representations of some non-B DNA conformations in the genome:
(A) Cruiforms (B) Slipped Hairpin (C) Z-DNA (D) Triplex [21]

1.3.3.　DNA Absorption Spectra

DNA absorbs light in the UV region of the electromagnetic spectrum as a result of σ→σ*, π→π* and n→π* electronic transitions in the heterocyclic bases. The π→π* and n→π* lower energy transitions are more readily monitored because they are found at wavelengths that are removed from typical solvent absorptions. The electronic transitions of DNA bases have been well established and reviewed.[22,23] A typical absorbance spectra of DNA has a λ_{max} at approximately 260 nm, which can be used to calculate the concentration of DNA in solution or for monitoring chemical or physical perturbations to DNA structures that result with hypo- or hyperchromicity

1.4. G-Quadruplexes

G-quadruplexes, commonly referred to as G4 DNA, are yet another family of secondary nucleic acid structures that deviate from the Watson-Crick B-form dsDNA. They are formed within specific repetitive G-rich sequences that are widely distributed throughout the human genome. With the completion of the human genome project, scientists have been allowed to explore the occurrence and frequency of these unusual DNA secondary structures. One of the surprising conclusions from these studies is that only 1-2% of the human genome is believed to code for protein. This accounts for approximately 20,000-25,000 genes and raises questions to what the function of the remaining 98% of noncoding genomic DNA is.[24,25] Algorithmic predictions estimate the probability of DNA sequences that

possess the ability to form G-quadruplex structures at 375,000, if not more.[26,27,28] However, these sequences are not distributed evenly throughout the genome. Computational studies in various organisms have shown that G-quadruplexes tend to cluster in particular areas, specifically, certain functional regions such as telomeric regions and gene promoters, but apparently were not found to be widely present in the coding regions of DNA.[29] This data, which demonstrates the nonrandom distribution of G4 motifs, supports the suggestion that G4 structures have an evolutionary conserved function in living cells that may include the regulation of gene expression and the preservation of chromosomal integrity.[30,31]

The ability for guanosine and its derivatives to form self-assembled structures was originally report by Bang in 1910 when he discovered that concentrated solutions of guanylic acid formed gels.[32] The self-assembled tetrameric structure formed was resolved using X-ray diffraction by Gellert et al more than fifty years later in 1962.[33]

1.4.1. G-Quartets

G-quadruplexes are formed by the vertical stacking of G-quartets (also referred to as G-tetrads). Each G-quartet is comprised of 4 guanine bases that form a square planar array, considered as the basic building block of G4 structures. Hoogsteen hydrogen bonds between the N1, N7, O6 and N2 guanine bases associate each of the 4 guanines into a cyclic hydrogen bonding pattern that forms this planar structure.[33,34] Each of the guanine bases is both the donor and acceptor of two hydrogen bonds. The self-assembled G-quartet structure is depicted in Figure 6. The structure is further stabilized by monovalent cations that occupy the central pore between the stacked G-quartets. This coordination neutralizes the electrostatic repulsive forces exhibited by the oxygens found on the 4 guanines that point inwards towards the central pore.[35]
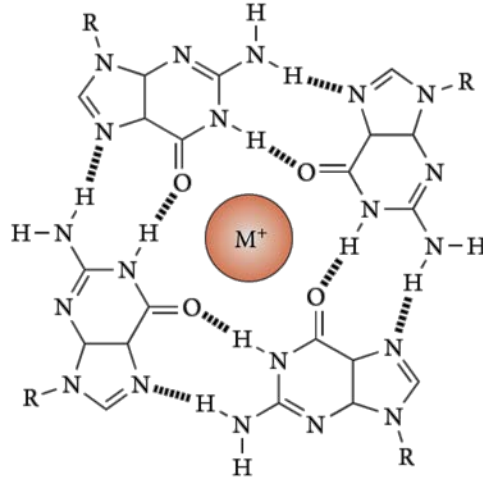
Figure 6. Illustration of the Hoogsteen hydrogen bonding pattern in a G-quartet

1.4.2. G4 Topology and Structure

G4 structure can adopt a number of various topologies, which is in part a consequence of variations in their molecularity, number of stacked G-quartets (at least 2 are required), the nature of the binding cations, strand orientation, sequence and the location and length of their loops.[36]

G4 DNA can be formed either by intramolecular folding of one DNA strand (unimolecular) or intermolecularly through the association of two (bimolecular) or four strands (tetramolecular). Technically, it is possible for a G4 DNA structure to form from three strands; however this occurrence is rare. Unimolecular G4 structures forming sequences can be described using the following[37]:

$$G_m X_n G_m X_o G_m X_p G_m \quad [1]$$

Where m is the number of guanine nucleotide residues in each short G-tract, usually containing 3-5 consecutive guanosines and X_n, X_o and X_p can be any combination of nucleotide residues (including guanines) that form the loops. It should be noted that the G-tracts themselves can vary in length. Guanines present in the G-tracts that do not participate in G-quartet interactions will be part of the loop region.

In theory, bimolecular and tetramolecular G4 structures can be formed from non-equivalent sequences. However, almost all bimolecular structures that have been reported in the literature are formed by the association of two identical nucleic acid sequences and can be described as:

$$X_n G_m X_o G_m X_p \; [2]$$
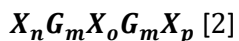
where both *n* and *p* can equal zero. Tetramolecular G4 structures are formed by the intermolecular association of four strands that can be described using either one of the following:

$$X_n G_m X_o \qquad [3]$$

$$G_m X_p G_m \qquad [4]$$
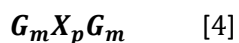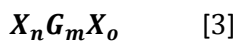
Figure 7 illustrates a variety of G4 topologies that can form from varying strand stoichiometries. A number of folding patterns are shown for a given molecularity, demonstrating the variations in both the loop arrangement and the strand orientations. Each linear or curved line represents a nucleic acid sequence. Arrows located at the end of each line indicate strand polarity in the 5'→3' direction.
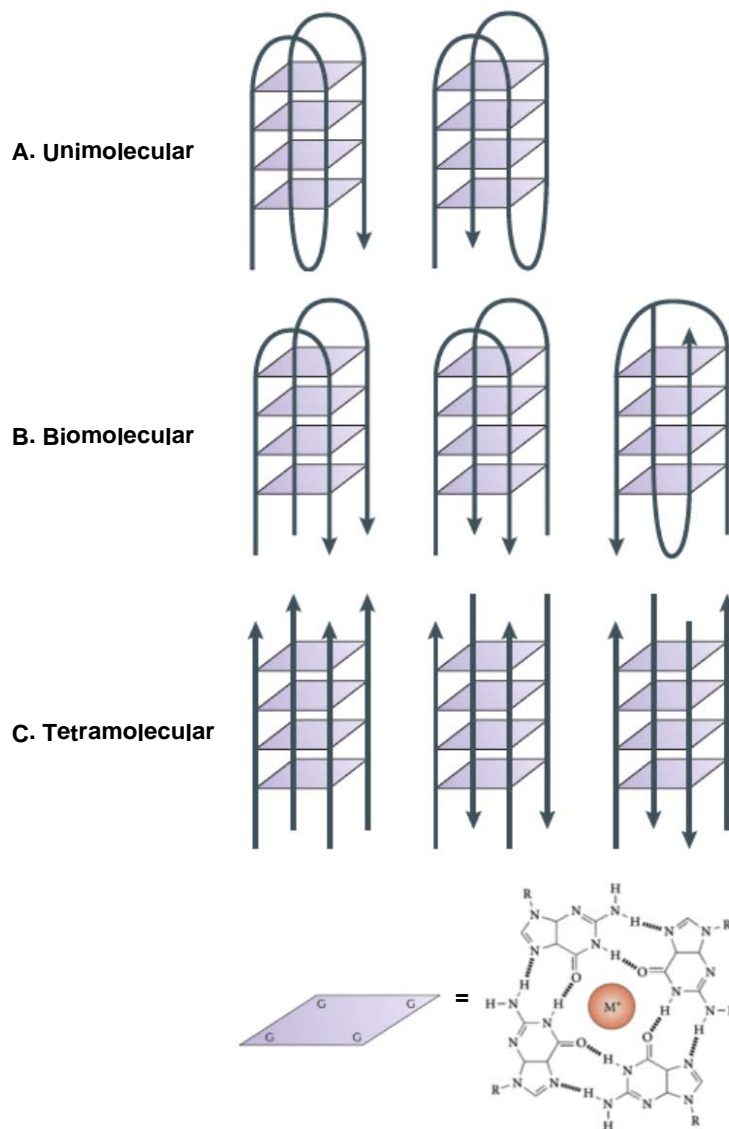
Figure 7. Schematic representation of G4 structures formed from varying strand stoichiometries

G4 structures, with the exception of the formation of tetramolecular quadruplexes, will have intervening DNA strands that form ssDNA loops with varying length and sequence composition. The loop length is typically between 1-7 nucleotides with shorter loops resulting in more stable structures.[38] There are three types of loop arrangements: (1) lateral loops which connect adjacent strands at one terminal G-quartet, (2) diagonal loops which span across the top of the G4 structure and (3) propeller loops which link the bottom G-quartet with the top G-quartet.[37,39] Figure 8 illustrates G4 structures that

demonstrate these various types of loop arrangements. It should be noted that certain G4 structures possess multiple loop types. The unimolecular quadruplex in Figure 8 used to present the diagonal loop also contains two lateral loops protruding from the opposing terminal G-quartet.
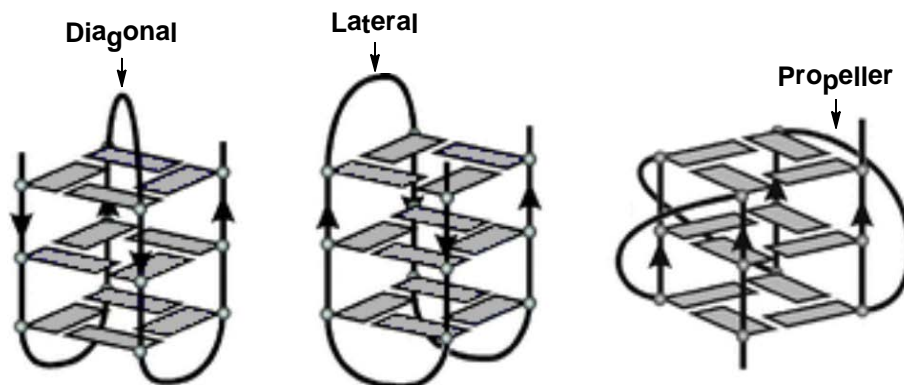
Figure 8. G4 structures illustrating the various types of loop arrangements[40]

Furthermore, the strand orientations (polarities) within the G4 can vary and as a result produce parallel, antiparallel or hybrid type structures. There are four fundamental strand orientations that are displayed in Figure 9.
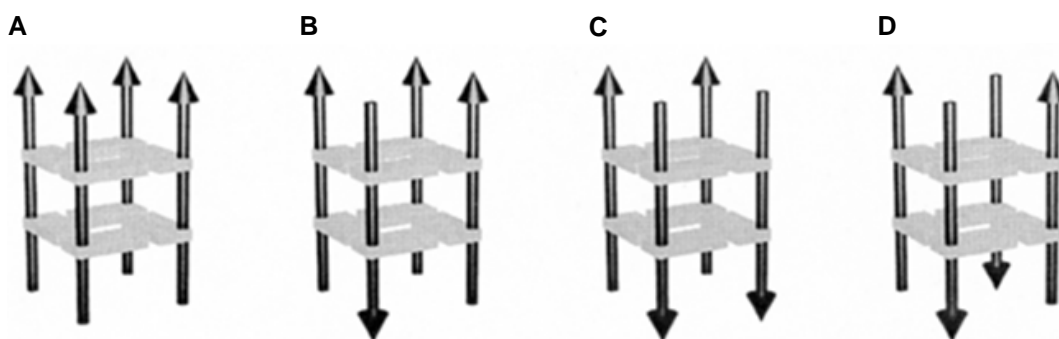
Figure 9. Strand orientations in G4 structures (A) all parallel (B) three parallel one anti-parallel (c) adjacent parallel (d) alternating antiparallel[41]

Investigations into the influence of loops on the folding and stability of G4 structures has provided some general trends (1) parallel structures are preferred by short loops (2) longer loops favor anti-parallel structures (3) loops formed from a single nucleotide residue have higher melting

14

temperatures i.e. are stable and (4) increasing the loop length decreases the melting temperatures i.e. decreases the structures stability.[42]

As mentioned previously during the discussion on the polymorphic nature of nucleic acids, free rotation about the C-1'-*N*-glycosyl bond is possible and does in fact occur in certain G4 structures. As a result both the *anti* and *syn* conformations can be present. This is not the case with B-DNA which is found held entirely in the *anti* conformation. Figure 10 portrays both of these conformations and, demonstrates the rotation about the glycosidic bond showing their interconversion. It should also be noted that the sugar pucker in most quadruplexes is in the C2'-endo conformation.[43]
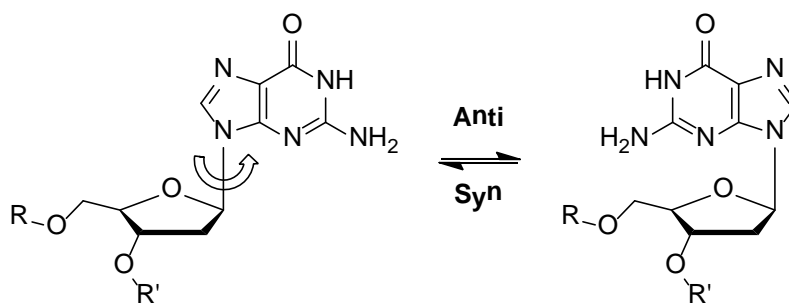
Figure 10. Rotation about the glycosidic bond allows guanine base to interconvert between the *syn* and *anti* conformations

The strand orientation directly influences the conformational state of the glycosidic bonds which in turn has an effect on the four grooves present in the G4 structure. These grooves describe the cavities bounded by the phosphodiester backbones. G4 having all parallel strand orientations will have their guanine glycosidic bonds in the *anti* conformation with grooves between the phosphodiester backbone(s) that are of all equal size. This makes this structure C_4 symmetric (see Figure 11A). When any one of the strands is antiparallel the guanine bases in that strand must assume the *syn* conformation in order to facilitate the formation of the Hoogsteen hydrogen bonds. As a result the orientation of the backbone in respect to the G4 is altered and the grooves become unequal in size. When consecutive guanines, starting with the one contributing N1 and N2, are both in same conformation (*anti* or *syn*) then the grove is of medium size. If the first guanine is *anti* and the second

15

guanine is *syn* then the groove is wide. If the first is *syn* and the second is *anti* then the grove is narrow.[44]

Therefore, if we consider a G4 structure that has alternating antiparallel strand orientation, then the corresponding glycosidic bond conformational states would be *anti-syn-anti-syn*. Following the guidelines previously stated this structure would have grooves that were wide, narrow, wide and narrow. In comparison, an adjacent antiparallel structure corresponds to an *anti-syn-syn-anti* conformation resulting with grooves that are wide, medium, narrow and medium. Both of these G4 structures are depicted in Figure 11.
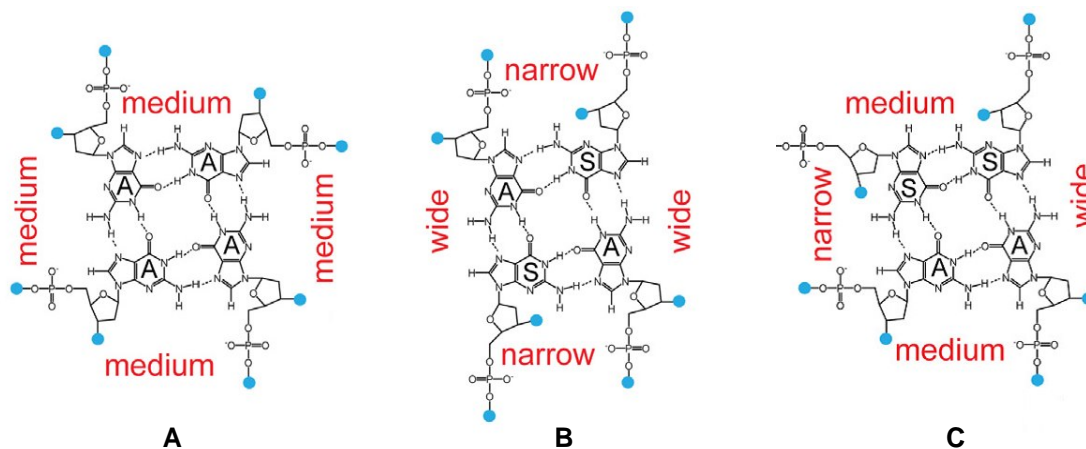
Figure 11. Relationship between conformational state of glycosidic bond, S=*syn* A=*anti*, and strand orientation (A) all parallel (B) alternating parallel (C) adjacent antiparallel[45]

Finally, G4 structures exhibit a significant dependence on cations for their formation and stability. K^+ and Na^+ both bind to and stabilize many G4 structures.[46] However, a higher ion binding preference is exhibited towards K^+ ions.[47] This has been attributed to the ability of K^+ to reside between two stacked G-quartets and coordinate with the eight carbonyl oxygens atoms that are present.[48] G4 structures depend on the cation identity and can form altered structures in the presence of K^+ or Na^+. Physiological conditions are favorable for the formation of G4 structures since the predominant ions

found in cells are K^+ and Na^+ and are present at intracellular concentration levels of 140 mM and 5-15 mM respectively.[49]

1.5. Biological Relevance of G4 DNA

1.5.1. Telomeres

Telomeres are nucleoprotein structures found at the end of linear eukaryotic chromosomes. They play an essential role in maintaining genomic stability by protecting chromosomes from degradation, recombination, end to end fusion, and from being recognized by cellular machinery as double stranded breaks.[50]

Other functions include preventing the loss of coding sequences at the end of chromosomes upon the completion of DNA replication. Dividing cells have been shown to lose base pairs (bp) following each round of cell division.[51] Instead, of losing vital coding sequences, non-coding telomeric DNA is lost. This loss of DNA bp is a direct consequence of the "end replication problem" inherent to DNA-polymerase (the inability of the enzyme to replicate the very end of linear DNA).[52] Telomeres as a result are progressively shorter after each round of cell division. After a certain number of cellular divisions the telomere will reach a critical length where it no longer can maintain its replicative capacity. This triggers the cell to halt the cell cycle and enter non-replicative senescence. This process is proposed to function as a biological clock, limiting the life span of somatic cells.[50]

Telomeric DNA contains a double stranded region and a single stranded 3' G-rich overhang. The dsDNA region is comprised of short tandem repeats of a DNA sequence. The distribution of guanine and cytosine is disproportionate between the two strands and are called the G and C strands accordingly. The G-strand runs in the 5'→3' direction and terminates with the 3' overhang. The tandem repeats appear to be highly conserved sequences because all vertebrates, including humans, have the same hexanucleotide sequence of (TTAGGG)$_n$. Table 1 shows some of the known telomeric DNA sequences.

Table 1. Some known telomeric DNA sequences of various species

Group	Organism	Telomeric Repeat
Vertebrates	*Human, mouse, Xenopus*	TTAGGG
Filamentous fungii	*Neurospora crassa*	TTAGGG
Slime moulds	*Physarum, Didymium,*	TTAGGG
	Dictyostelium	AG1-8
Kinetoplastid protozoa	*Trypanosoma, Crithidia*	TTAGGG
Ciliate protozoa	*Tetrahymena, Glaucoma*	TTGGGG
	Paramecium Oxytricha	TTGGG(T/G)
	Stylonychia, Euplotes	TTTTGGGG
Apicomplexan protozoa	*Plasmodium*	TTAGGG(T/C)
Higher plants	*Arabidopsis thaliana*	TTTAGGG
Green algae	*Chlamydomonas*	TTTTAGGG
Insects	*Bombyx mori*	TTAGG
Roundworms	*Ascaris lumbricoides*	TTAGGC
Fission yeasts	*Schizosaccharomyces pombe*	TTAC(A)(C)G_{1-8}
Budding yeasts	*Saccharomyces cerevisiae*	TG_{1-3}
	Candida glabrata	GGGGTCTGGGTGCTG
	Candida albicans	GGTGTACGGATGTCTAACTTCTT
	Candida tropicalis	GGTGTA[C/A]GGATGTCACGATCATT
	Candida maltosa	GGTGTACGGATGCAGACTCGCTT
	Candida guillermondii	GGTGTAC
	Candida pseudotropicalis	GGTGTACGGATTTGATTAGTTATGT
	Kluyveromyces lactis	GGTGTACGGATTTGATTAGGTATGT

The human telomeric DNA sequence repeat of 5'-TTAGGG-3' has a length, at birth, typically between 10-15 kilobases (kb) and ranges from 2-20 kb in adult cells. The 3' overhang is between 50-500 nucleotides (nt) and can fold back to pair with the complementary sequence on the C-strand forming what is known as the t-loop. The location where the G-strand overhang invades and displaces the G-strand in the duplex regime is known as the D-loop (displacement loop).[53] Figure 12 presents a schematic representation of the telomeric region of human chromosomes.
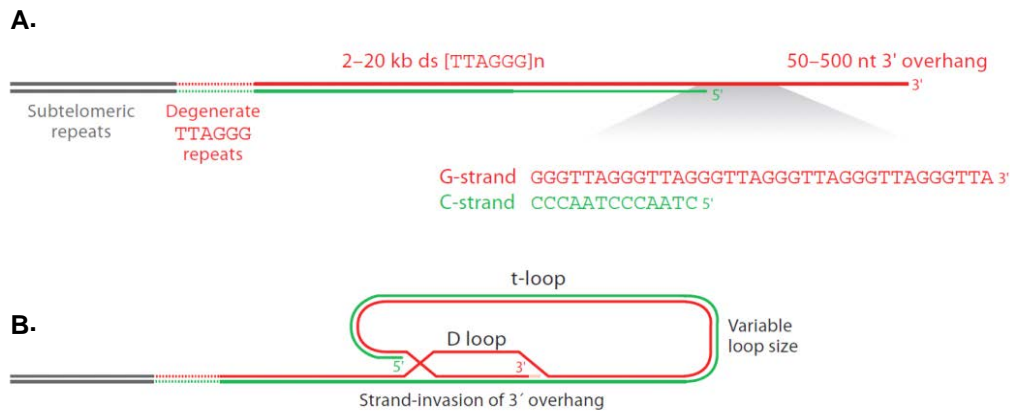
Figure 12. Structure of human telomeres (A) Chromosome end showing dsDNA region of TTAGGG repeats and the 3' overhang. (B) T-loop and D-Loop structures resulting from the interaction of the overhang with the dsDNA regime.[53]

There is a continuously growing number of proteins discovered that bind to telomeric DNA, including a six membered protein complex known as shelterin.[54] Its components include Telomeric Repeat binding Factor 1 and 2 (TRF1 and TRF2), Repressor/Activator 1 (Rap 1), TRF2- and TRF1-Interacting protein 2 (TIN2), Tripeptidyl peptase 1 (TPP1) and Protection of Telomeres (POT1). Additionally, a ribonucleoprotein composed of a reverse transcriptase protein and an RNA template, known as Telomerase, is responsible for telomere elongation. The activity of this enzyme in most cells is strongly regulated and typically will not interfere with the incremental shortening of telomeric DNA during each round of cell division.[55]

Telomeric sequences have been shown to form stable G4 structure *in vitro*. The human telomeric sequence, 5'-TTAGGG-3', folds spontaneously into an intramolecular G4 structure where the GGG nt participate in the formation of G-quartets while the TTA nt form the intervening loops.[56] This structure has been found to be stable under physiological condition and has a melting temperature of approximately 65 °C. Therefore, it has been frequently suggested that the 3'-overhang G-strand of telomeres form G4 structures *in vivo* and are physiological relevant.

1.5.2. Structure of Human Telomeric G4

Dai and coworkers reported the NMR solution structure of the intramolecular G4 formed by human telomeric DNA repeats, under physiological condition, in K[+] solution.[57] They investigated the folding of a 26 nt sequence, d[AGGG(TTAGGG)$_3$], composed of a 22-mer of four human telomeric repeats with flanking AA added to each end. Their studies reveal that the quadruplex, consisting of 3 stacked G-quartets, adopts a hybrid-type mixed parallel-antiparallel structure, where the third strand is antiparallel to the other three. The first two strands are linked via a TTA propeller-type loop and the other three strands are linked via TTA lateral loops. The lower 2 G-quartets are arranged in an *anti-anti-syn-anti* conformation and the top G-quartet is in a *syn-syn-anti-syn* conformation. A schematic representation of the folding topology is shown in Figure 13. A superimposition of the 10 lowest energy structures as determined by NOE-restrained structure calculation is presented in Figure 14 (PDB ID 2HY9). It is important to note that this hybrid-type G4 is the predominant form under physiologically relevant conditions; however, other conformations are possible, including the formation of higher order multimers.[58]
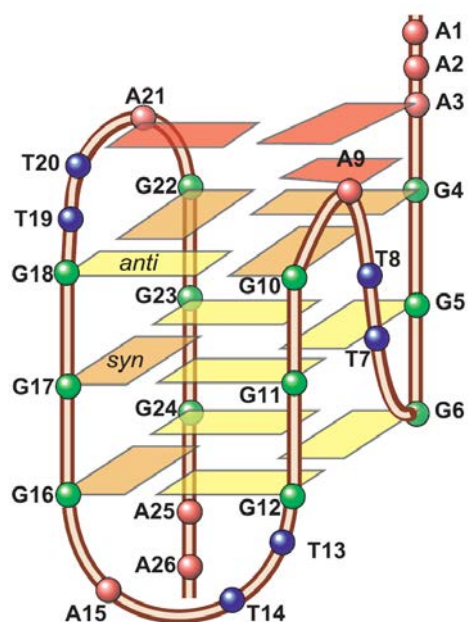
Figure 13. Folding topology of the unimolecular human Telomeric G4
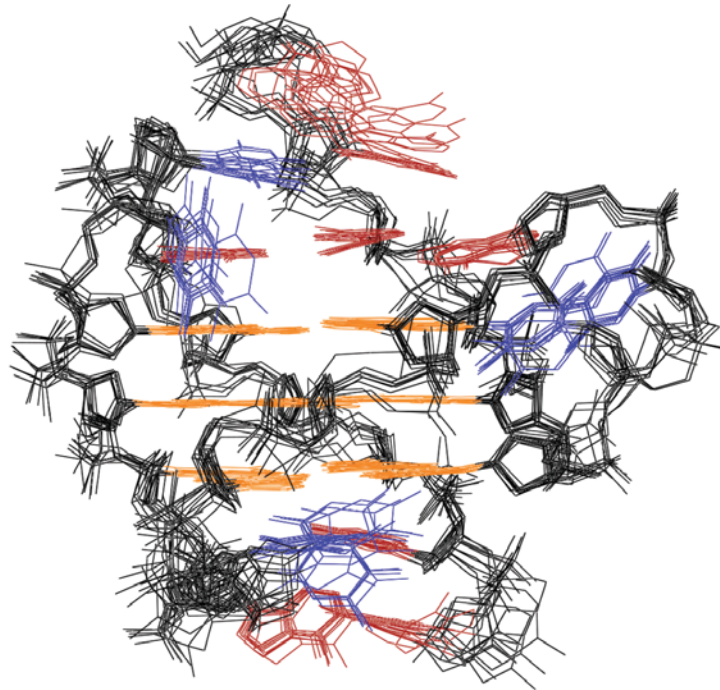
Figure 14. Superimposed 10 lowest energy structures of telomeric G4 in K$^+$ solution by NOE-restrained structure refinement (g = orange, a = red, t = blue)

1.5.3. Evidence of G4 Structures *in vivo* at Telomeres

It is widely believed that G4 structures do in fact form *in vivo*. Demonstrating this, however, has proven to be a difficult task for researchers. The strongest supporting and most direct evidence of their *in vivo* formation comes from studies on ciliated protozoa specifically, *Stylonychia lemnae*. These organisms contain high concentrations of telomeric DNA in their macronucleus, approximately 10^8 telomeres, possessing the repeat sequence d(TTTTGGGG)$_2$.[59] The studies performed on these organisms utilize developed antibodies, raised by ribosome, that display high binding specificity between parallel and antiparallel telomeric G$_4$T$_4$ structures.[60] The presence of *in vivo* G4 structures at telomeres was directly observed by *in situ* immunostaining. Only antibodies raised specific for antiparallel G4 structures were found to bind. This indicates that only the antiparallel and not the parallel G4 structures are present *in vivo*.[61] This data is consistent with previous observations that determined the G-strand

overhang would fold intermolecularly into an antiparallel structure. Furthermore, control experiments demonstrated that the antibodies themselves do not induce the formation of G4. Therefore, it was concluded that the staining of the *Stylonchia lemnae* telomeres with the antiparallel specific G4 antibodies was dependent on the natural expression of the *in vivo* G4 structure. One of the difficulties for applying this immunostaining technique used on *Stylonchia lemnae* to vertebrate cells is due to their low concentration of telomeres. Also, thus far, there have been no antibodies against human telomeric G4 DNA that have been isolated.

There does however, exist some less direct evidence for the formation of G4 structures at telomeres in human cells. 3,6,-bis[2-(1-methypyridinium)vinyl] carbazole diiodide (BMVC), a highly fluorescent biomarker that is known to bind and stabilize G4 structures *in vitro*, has been found to stain metaphase chromosomes in human lung adenocarcinoma cells in vivo.[62,63] This suggests that the BMVC is binding to the telomeres. However, it is unclear whether the BMVC is responsible for inducing the G4 formation. Research into finding more direct evidence for the presence of G4 structures at telomeres in organisms other than ciliates is ongoing.

1.5.4. Effects of G4 on DNA Replication

During DNA replication in eukaryotic cells the DNA double helix of the parent strand is unwound by a helicase creating two replication forks, one functions as a leading strand the other as a lagging strand. The leading strand is synthesized, by polymerases, in the 5'→3' direction continuously without interruption while the lagging strand is replicated discontinuously forming short single strands, known as Okazaki fragments, that are subsequently ligated together. The single stranded nature adopted by DNA during this process provides an opportunity for G4 structures to form. Some of these structures may have relevant functions in cellular processes while others may be simply a result of interactions between residues in the unwound strand sequence. Regardless of their origin, in order for replication to continue

these structures must be disrupted. The formation of these G4 structures influence DNA replication by slowing down or inhibiting the replication process.[64]

Helicases are most likely responsible for the unwinding of these G4 structures formed during the replication process. *In vitro* studies have shown that many helicases non-specifically bind or unwind G4 structures. WRN[65], BLM[66], FANCJ[67], PIF1[68] are all human helicases that have demonstrated this activity and have been associated with human diseases. (See section 1.5.6)

1.5.5. Effects of G4 on Transcription

Computational studies focusing on protein coding genes in humans suggest that more than 40% of these genes have at least one G4 motif within 1 kb upstream of their transcription start site.[69] Additionally, other studies reveal that the G4 motifs are proximal to or overlapping transcription factor recognition sites.[70] This data suggests that G4 structures may perhaps influence the transcriptional activity of the proximal gene in both a positive or negative manner. One theory postulates that the effect of the G4 is associated directly with the DNA strand on which it forms. If the G4 is found on the template strand, the structure will prevent the transcription machinery from proceeding and as a result inhibit the transcriptional process. Conversely, if the G4 is found on the non-template strand it can increase transcription level by maintaining the single stranded-ness of the template. Another theory suggests that G4 structures regulate transcription through the recruitment of proteins that either, one, encourage transcription by stimulating or initiating polymerase or, two, inhibit transcription by recruiting proteins that suppress the process. Figure 15 illustrates these presumed functions of G4 structures during transcription.

A. G4 on template strand, inhibiting transcription
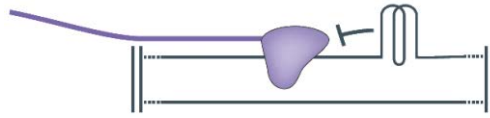
B. G4 on non-template strand, encouraging transcription
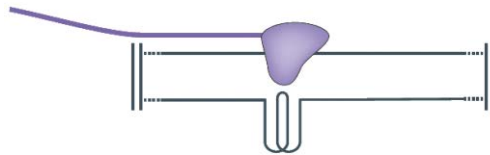
C. Protein binding to G4, stimulating transcription
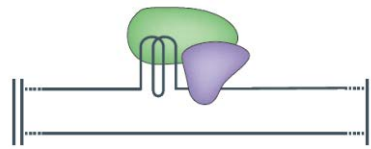
D. Protien binding to G4, suppressing transcription
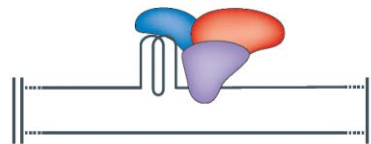
Figure 15. Presumed functions of G4 structures during transcription

1.5.6. G4 Structures and Disease

The instability of human telomere structures is associated with a number of human diseases. Most notably, G4 forming sequences have been implicated in cancer cell proliferation and as a result have attracted much interest over the years. Shorter telomere length is a risk factor for the development of cancer and telomerase is activated in approximately 90% of cancer cells. Furthermore different forms of anemia, hypertension, coronary heart diseases, chronic human immunodeficiency virus infection, ulcerative colitis and chronic liver disease have also been shown to relate to defects in telomerase or telomere stability.[71,72]

Less discussed G4 associated human diseases do exist. For example Werner syndrome, Fanconi anemia and Bloom's syndrome are diseases caused by genetic mutations that result in non-functional

24

WRN, FANCJ and RecQ helicases, respectively. The normal function of these helicases is to unwind G4 DNA.[73]

1.6. G4 Binding Ligands

Ligands that bind G4 structures efficiently frequently contain a large planar aromatic surface that is capable of participating in π-stacking interactions with the G-quartet motif. The larger this surface the greater the resulting aromatic-aromatic overlap becomes, resulting in improved selectivity of the ligand for G4 structures vs. duplex DNA. These ligands also usually carry positive charges that participate in electrostatic interaction with the polyanionic backbone(s) and grooves of the G4 structure. Known binding ligands have been classified according to their cationic state: (1) upon *in situ* protonation (2) via *N*-alkylation (3) due to the presence of a metal center or (4) non-cationic.[74]

G4 binding ligands usually exhibit both hydrophobic and hydrophilic character, as a result of their planar aromatic surface and their cationic charges, respectively. The hydrophilicity assists in increasing the water solubility of the ligand which is essential when targeting aqueous biological molecules such as these nucleic acids. The design of G4 binding ligands typically follows this assembly since it was first introduced by Neidle, Hurley and coworkers who demonstrated the binding and telomerase inhibition activity of a 2,6-diamidoanthraquinone.[75] The selectivity of this compound for G4 structures vs. duplex DNA was insufficient and the compound was found to be cytotoxic which deterred the researchers from further pursuing its biological application. However, the beneficial features possessed by this compound became mainstays in the future development of G4 binding ligands.
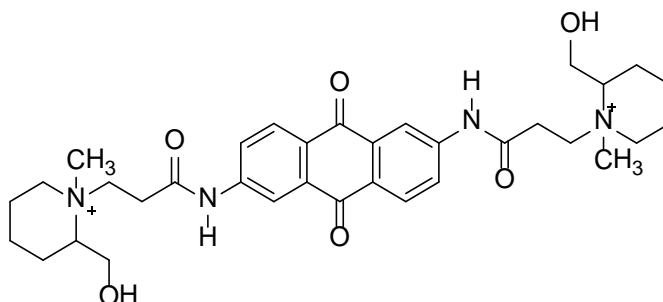
Figure 16. 2,6-diamidoanthaquinine prepared by Neidle and Hurley

1.6.1. *In Situ* protonated Ligands

Neidle continued his development of G4 binding ligands, initially, by modifying the planar aromatic core using fluorenone,[76] acridone[77] and acridine[78] in place of the anthraquinone. Of particular interest was a 3,6-disubstituted acridine, known as BSU6039 which contained two amine appendages that were protonable *in situ*. The acridine core was found to participate in π-stacking interactions with two of the guanine residues on the terminal G-quartet and the two appendages interacted electrostatically with two of the G4 grooves. Further optimization of this compound was achieved by increasing the number of protonable appendages to three which permitted electrostatic interactions to occur with three of the grooves. This structure, known as BRACO-19, was found to stabilize the G4 structure by FRET-melting assay ($\Delta T_{1/2}$ = 27°C)[79,80] and exhibited a 31-fold binding preference for G4 structures vs. duplex DNA as shown by SPR studies.[81] This compound also demonstrated a significant therapeutic efficacy as an inhibitor for cancer cell proliferation.[82] Figure 17 shows the structure of the BSU6039 and BRACO-19 acridine derivatives.
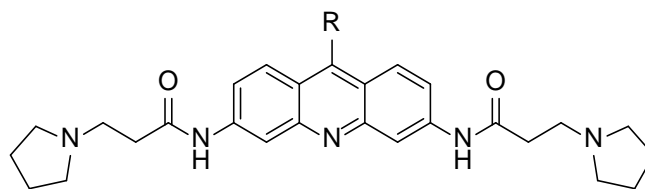
R = H, **BSU6039**
R = NH-p-C$_6$H$_4$-N(CH$_3$)$_2$, **BRACO-19**

Figure 17. Substituted acridine derivatives

Hurley and coworkers focused on the expansion of the aromatic core of the ligand, preparing perylenediimide compounds that possessed two protonable appendages. These included a compound known as PIPER which was found to be less effective as a telomerase inhibitor than BRACO-19, however, demonstrated an increase in selectivity, preferring G4 structures 42-fold more than duplex DNA.
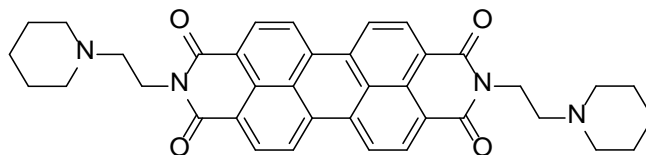
Figure 18. Structure of the perylene diimide derivative **PIPER**

Teulade-Fichou and coworkers developed ligands using pentacyclic quinacridines that are believed to maximize the π-π stacking interactions with the guanines of the G-quartet due to their crescent-like shape. Figure 19 shows their most promising compound, MMQ$_3$. This ligand has been shown to exhibit high G4 stabilization ($\Delta T_{1/2}$ = 20°C) and high telomerase inhibition activity.[83] NMR structural studies performed on MMQ$_1$, a dipropylamino derivative of this compound, indicates that the pentacylic quinacridine experiences aromatic overlap with three of the guanine residues in the G-quartet motif.[84] Other G4 binding ligands that similarly possess this crescent-like shape have also been studied and include indoloquinolines,[85] cryptolepine[86] and quindolines.[87]
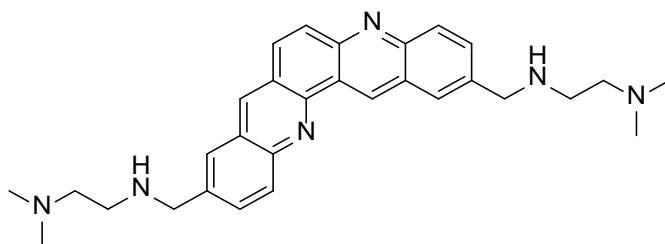
Figure 19. Structure of **MMQ₃**

A dimeric macrocylic quinacridine, BOQ$_1$, that was subsequently prepared by this group, showed an enhanced G4 stabilization ($\Delta T_{1/2}$ = 28°C) coupled with a 10-fold increase in G4 structure selectivity.[88] The expanded aromatic surface of this molecule, shown in Figure 20, demonstrates that with increased π-π stacking interactions, G4 stability is promoted and the large size of the aromatic surface sterically hinders duplex binding and as a result increases selectivity towards G4 structures. The binding efficiency of dimeric macrocylces, however, has been shown to be dependent on the aromatic unit which they contain. Dimers derived from quinacridine or acridines are efficient G4 binders while those containing phenanthroline of naphthalene demonstrate poor binding ability.[84]
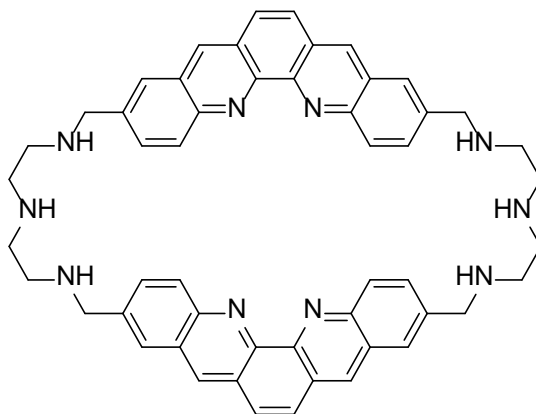
Figure 20. Structure of the dimeric macrocylcic quiancridine **BOQ₁**

In an attempt to further increase G4 selectivity Teulade-Fichou and coworkers focused on introducing new structural elements to known binding ligands that increase the recognition of G4 loops and grooves. NCQ is a neomycin capped quinacridine that has been designed following this principle and

does in fact show a preference towards loop-containing G4 structures vs. non-loop containing ones.[89] This finding along with its G4 stabilization ability ($\Delta T_{1/2}$ = 14°C) and telomerase inhibition activity validated and propelled the design of G4 binding ligands to contain moieties that possess recognition capability for both G-quartets and grooves. The structure of NCQ is shown in Figure 21.
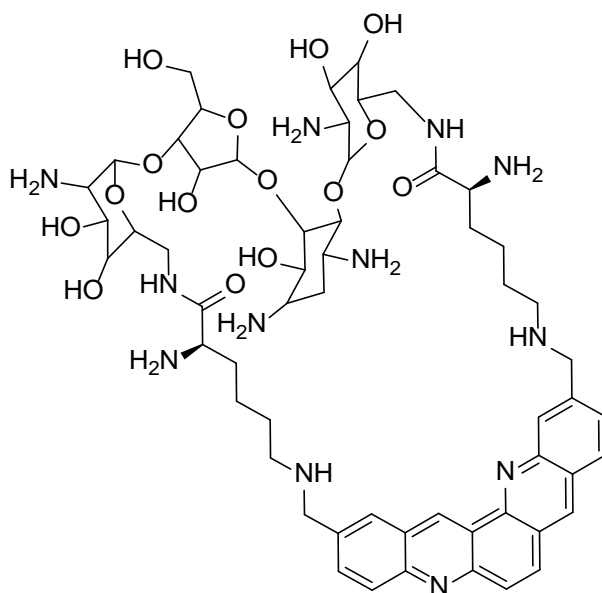
Figure 21. Structure of neomycin capped quinacridine, **NCQ**

Furthermore, groove interacting scaffolds can allow for the selectivity of particular G4 structures. This has been demonstrated by the efficiency in which tri-oxazole macrocycles containing three protruding amine appendages *cis* to one another exhibit a binding preference for G4 structures formed by the c-kit sequence vs. the human telomeric sequence.[90]
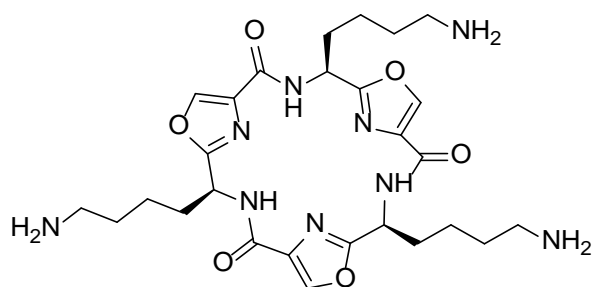
Figure 22. Tri-oxazole macrocycle scaffold containing three protruding *cis* amine appendages

1.6.2. *N*-alkylated Ligands

As an alternative to the protonation of amine side-chains *in situ*, the cationic charge found on the ligand can be introduced by the quaternization of aromatic nitrogens via alkylation. This approach affords two advantages; it produces a water soluble ligand with no need for cationic appendages and increases the π-stacking ability of the ligand by reducing the electron density of the aromatic moiety.[91]

Cationic porphyrins are found within this class of ligands and have been some of the most extensively studied, specifically, 5,10,15,20-tetra(*N*-methyl-4-pyridyl)porphyrin (TMPyP4), shown in Figure 23.[92]
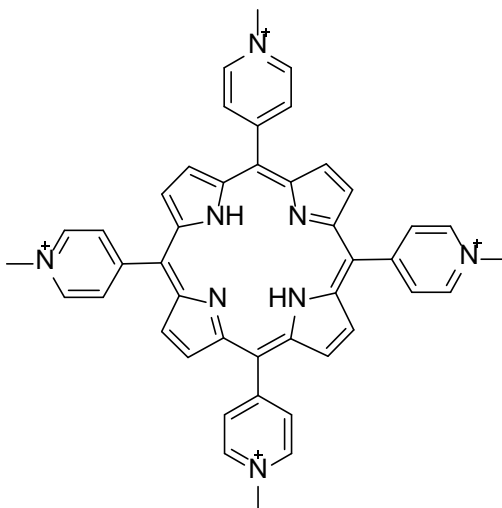
Figure 23. Structure of 5,10,15,20-tetra(*N*-methyl-4-pyridyl)porphyrin (H$_2$TMPyP4)

This tetracationic porphyrin has demonstrated a high affinity for G4 DNA (ΔT$_{1/2}$ = 17°C), inhibition of telomerase activity and the ability to downregulate the expression of oncogenes by inducing G4 formation.[93] However, the compounds selectivity for G4 structures is poor, in part due to the various G4 binding modes it can participate in, ranging from intercalation[94] to the more expected external stacking[95] and mixtures thereof. A more in-depth discussion of porphyrins can be found in section 1.7.

Structurally similar compounds to TMPyP4 have also been identified and include for example TQMP, a porphyrin in which the *N*-methylpyridinium substituents were replaced with (trimethylammonium)methylphenol arms as shown in Figure 24.[96]
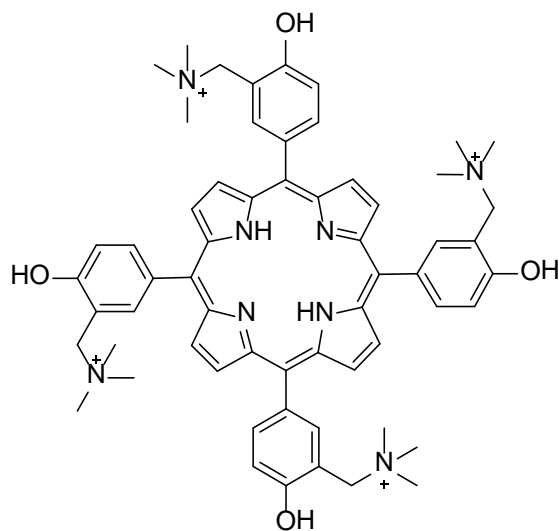
Figure 24. Structure of **TQMP**

Structural variations of the porphyrin macrocycle have also demonstrated efficient binding to G4 structures. These analogues differ by the nature and number of their quaternizing appendage as well as their ring-size. Figure 25 illustrates various polyheterocyclic macrocyclic cores used in the preparation of these analogues.
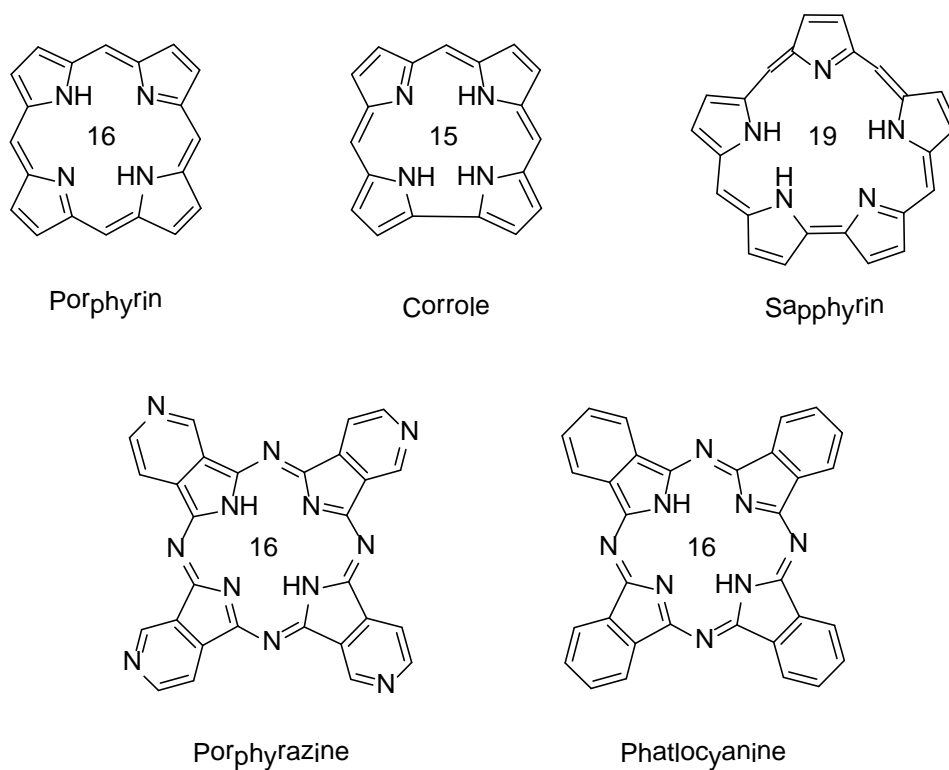
Figure 25. Polyheterocyclic cores used in the design of G4 ligands

Analogues that have been shown to bind efficiently with high selectivity for G4 structures include a diselenosapphyrin known as Se2SAP which has demonstrated a 50-fold selectivity for G4 DNA over duplex DNA. This compound has also been shown to have higher affinity for G4 structures found at promoters over those at telomeres. It should be noted that the synthesis of this molecule is challenging and low-yielding. The structure of Se2SAP is shown in Figure 26 along with the structure of 3,4-TMPyPZ, a porphyrazine derivative that demonstrated a 100-fold increase in affinity to G4 structures coupled with a 30 fold preference for G4 over duplex DNA.
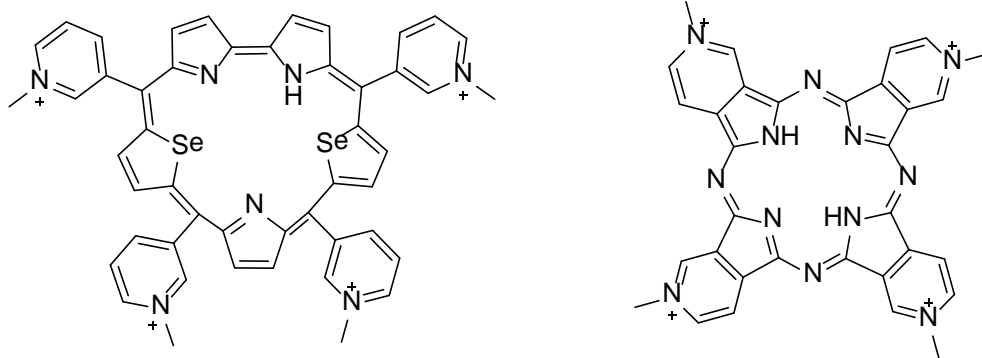
Figure 26. Structure of **Se2SAP** (left) and **3,4-TMPyPz** (right)

Aside from cationic macrocycles there are a number of *N*-alkylated small molecules that are known to bind G4 DNA. These include ethidium bromide as well as less toxic derivatives of this compound. Their selectivity for G4 structures is typically poor, mainly due to their preferred intercalative binding mode which does not substantially participate in distinguishing between G4 and duplex DNA.

Another notable series of ligands are bisquinolinium compounds that contain pyridodicarboxamide cores. One of the most favorable ligand identified from this series is known as 360A, shown in Figure 27. This compound has exhibited a high degree of G4 stabilization ($\Delta T_{1/2}$ = 21°C), a 150-fold selectivity of G4 DNA over duplex DNA and high inhibition of telomerase activity.[97] Interestingly, this compound has also demonstrated the ability to actively induce the formation of tetramolecular G4 structures, acting as a chaperone for the association of the four strands.[98]
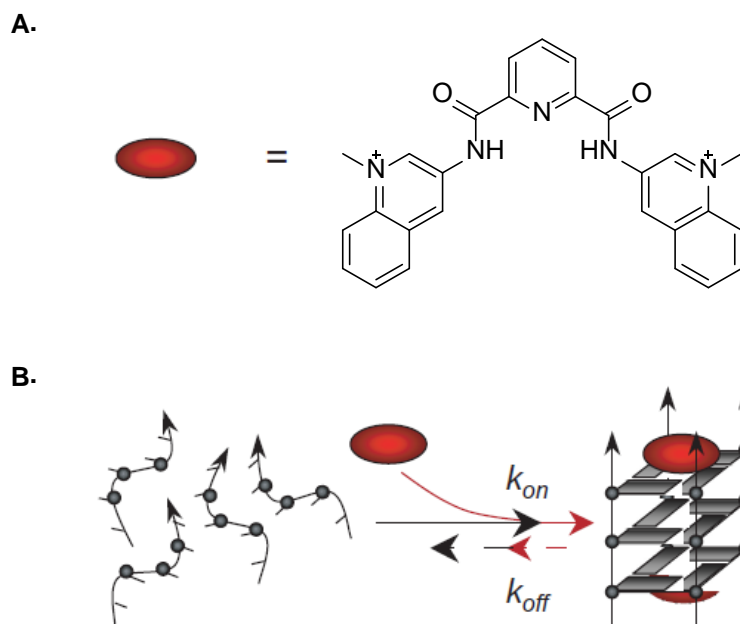
Figure 27. (A) Structure of 360A (B) Induced formation of tetramolecular quadruplex

Finally, expansion of the central 2,6-disubstituted pyridine to a 1,10-disubstituted phenanthroline results in a family of compounds known as Phen-DC, which have been shown to geometrically match G-quartets and as a result exhibit exceptional stabilization ($\Delta T_{1/2}$ = 29.7°C) along with a 150-fold binding preference for G4 DNA.[99]
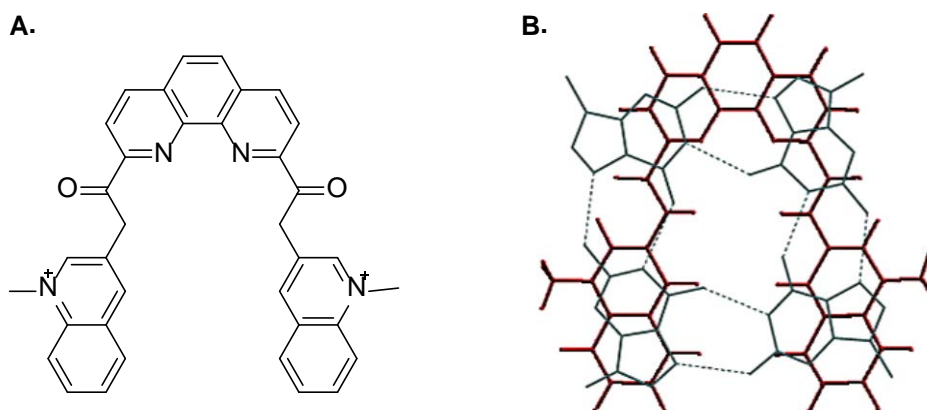
Figure 28. (A) Structure of a bisquinolinium pyridodicarboxamide compound with a central phenanthroline core, known as **Phen-DC₃** (B) Superposition of Phen-DC₃ and G-quartet depicting their geometrical match

1.6.3. Ligands with Metal Center

Organometallic complexes have also been shown to have a stabilizing effect on G4 structures. The underlying principle behind their affinity is the substitution of the cationic charge of the normally coordinated, centrally positioned cation (potassium or sodium) with partial charges found on the organometallic complex induced by the metal. According to molecular modeling predictions the complex is thought to coordinate its metal directly above the cation channel.[100] Furthermore, the cationic or polarized state of the complex promotes binding with anionic DNA.

The insertion of a metal in the central cavity of TMPyP4 increases its selectivity towards G4 DNA, specifically when coordinated with Cu(II),[101] Ni(II) and Mn(III).[102,103] In fact, Dixon and coworkers demonstrated that the pentacationic manganesen(III)-porphyrin complex, shown in Figure 29, had a remarkable 10,000-fold selectivity towards G4 structures over duplex DNA and is considered to be one of the most effective ligands for G4 binding.[104] This increase in G4 selectivity is also evident in a number of other organometallic ligands that contain Ru(II), Fe(III), Zn(II) and Pt(II), and Ni(II) transition metals.
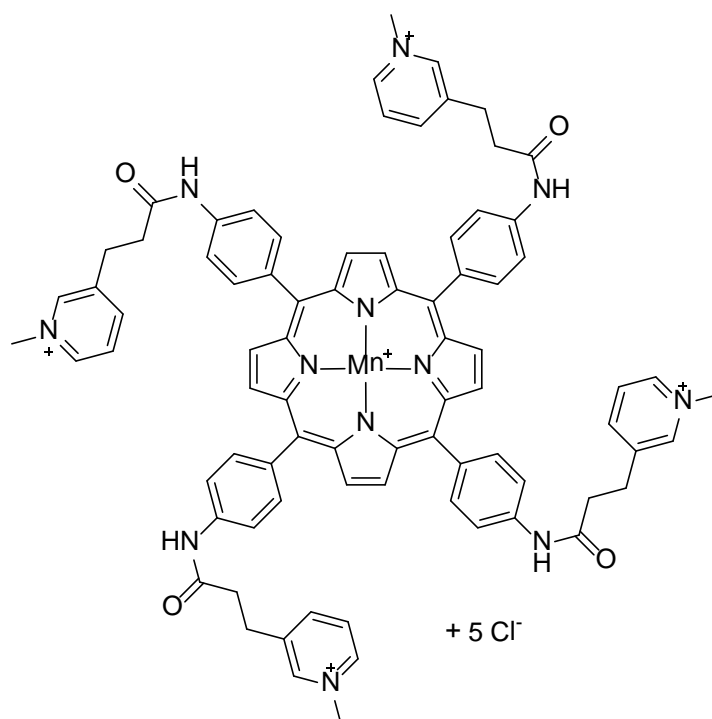
Figure 29. Structure of the pentacationic organometallic G4 ligand Mn(III) porphyrin

1.6.4. Non Cationic Ligands

Almost all compounds found within this classification of G4 ligands can be considered as telomestatin-like macrocycles. Telomestatin itself is a naturally occurring, 8-ring 24 membered macroheterocycle, isolated from *Streptomyces annulatus* 3533-SV4.[105] This polycyclic compound, shown in Figure 30, is comprised of five oxazole, two methyloxazole and one thiazoline rings. It has attracted much attention since it is considered to be one of the most selective G4 ligands, possessing high binding affinity for G4 DNA and none for duplex DNA. This was made evident by competitive FRET studies showing that the stabilization of G4 structures by telomestatin ($\Delta T_{1/2}$ = 24°C) was unchanged with the addition of 50 equivalents of competitive duplex DNA. This lack of affinity towards duplex DNA can be attributed to its neutral state and large shape. Telomestatin has also been shown to have the ability to inhibit telomerase activity and exhibits antiproliferative activity against a number of cancers.[93]

The drawback to telomestatin, is that its synthesis which has been reported by Doi and coworkers,[106] is complex and difficult to scale-up. However, there have been a few telomestatin-like synthetic analogues prepared by Rice and co-workers that exhibit G4 affinity and telomerase inhibition activity. [107]
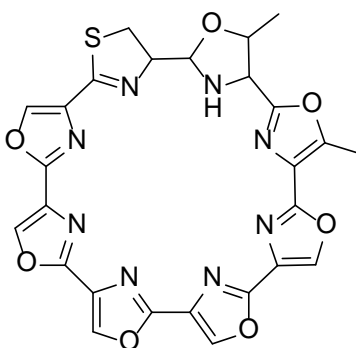
Figure 30. Structure of telomestatin

1.7. Porphyrins

1.7.1.   Basic Structure

Porphyrins are a class of naturally occurring macrocyclic compounds involved in a wide range of vital biological processes, and include the well-known heme and chlorophyll compounds. The common feature found in all porphyrins is their planar aromatic tetrapyrrolic macrocyclic core, comprised of four pyrrolic subunits connected *via* four methine bridges as shown in Figure 31.
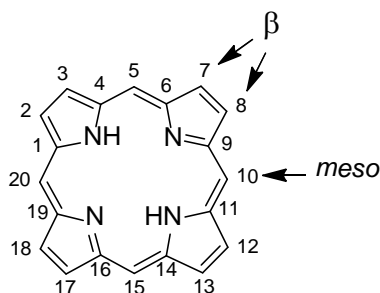
Figure 31. The structure of the Porphyrin macrocylic core

There are two distinct regions in these macrocycles. The β positions, which consist of the outermost pyrrolic carbons, eight in total, numbered 2, 3, 7, 8, 12, 13, 17 and 18; and the *meso-*positions, the central carbons in the methine bridges, four in total, numbered 5, 10, 15 and 20. The resulting aromatic macrocycle contains a total of 22 π electrons, 18 of which are delocalized within the system following Huckels 4n+2 rule for aromaticity. Four of these delocalized electrons are donated by two of the pyrroles found in the -3 oxidation state, with their lone pair electrons facing in towards the center of the macrocycle.

1.7.2.   UV-Visible Absorption Properties

Porphyrins are highly chromophoric compounds that are known for their intense colors. This is an expected consequence of their extensively conjugated macrocyclic system. The UV-Vis spectra of free base porphyrins typically consist of an intense absorbance band in the near UV-region (370-430 nm), commonly referred to as the Soret band, with a molar extinction coefficient usually on the order of 10^5

M^{-1} cm^{-1}; and four less intense absorption bands in the visible region (500-700 nm), known as the Q bands, with molar extinction coefficients usually on the order of 10^4 M^{-1} cm^{-1}. The resulting UV-Vis spectra of porphyrins have been explained using the Gouterman's 4-orbital LCAO model, which considers the possible excitation transitions between the two π-bonding HOMO orbitals (a_{1u} and a_{2u}) and the two degenerate π* antibonding LUMO orbitals (e_{gx} and e_{gy}).[108] The a_{1u} and a_{2u} orbitals are close enough in energy that they can be considered as equivalent, making their excitation energy to the e_g orbitals nearly identical. Excitation states based on the transition from $a_{1u} \rightarrow e_{gy}$ and $a_{2u} \rightarrow e_{gx}$ are X-axis polarized, while those between $a_{1u} \rightarrow e_{gx}$ and $a_{2u} \rightarrow e_{gy}$ are Y-axis polarized. Electron-electron interactions are mixed and split in energy by configuration interaction, resulting in constructive interference that forms the high intensity allowed Soret band (combination of B_x and B_y) and destructive interference that results in the low intensity forbidden Q bands. Free base porphyrins are D_{2h} symmetrical which results in further destructive splitting of the Q_x and Q_y bands leading to the four Q bands that are identified in order of increasing energy as $Q_x(0,0)$, $Q_x(1,0)$, $Q_y(0,0)$ and $Q_y(1,0)$.[109] The intensity and absorption wavelengths of a given porphyrin is concentration and solvent dependent. Also, the absorption properties vary according to the nature of the functional groups present at the β and *meso* positions.
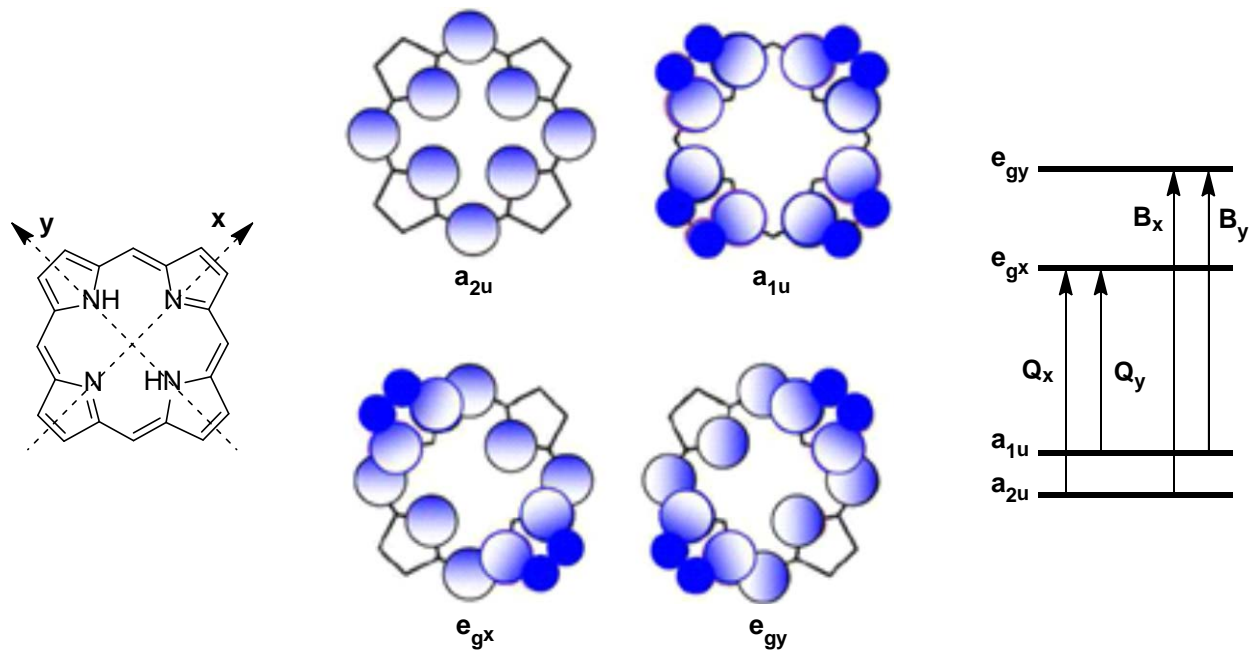
Figure 32. Left: Drawing representing the X and Y polarization of D_{2h} free base porphyrin Middle: depiction of the four Gouterman orbitals Right: illustration of the excitations from the HOMOs to the LUMOs and their ensuing bands
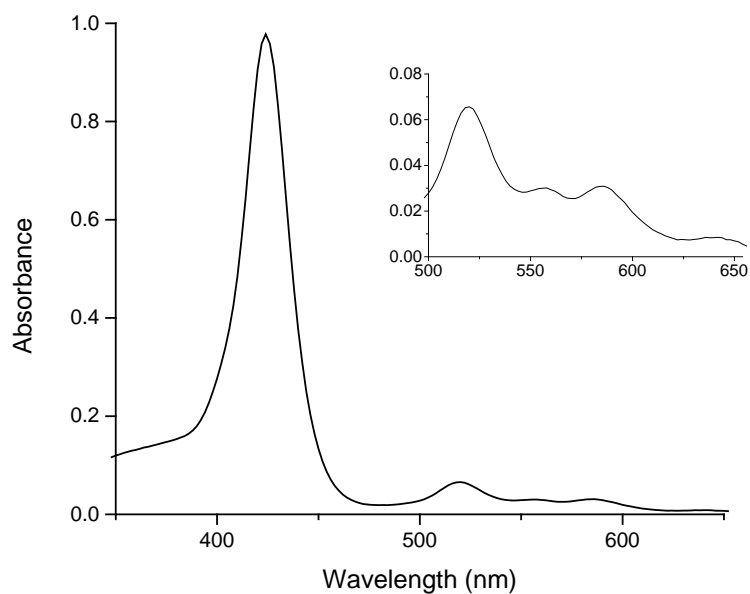
Figure 33. UV-Vis spectrum of a *meso*-substituted porphyrin (H$_2$mTPyP) displaying the intensity of the Soret band and Q bands. The four Q bands are expanded in the insert.

1.7.3. Synthesis of Symmetrical *meso*-substituted Porphyrins from Monopyrroles

The common objective in the synthesis and design of porphyrins usually centers on the preparation of porphyrin macrocycles possessing substituents specifically arranged around their periphery. The ability to control the substitution pattern allows for customizable porphyrins that can be prepared for specific applications. Figure 33 shows the retro-synthesis of porphyrins and illustrates how the macrocycle can be broken down into its primary building blocks of four pyrroles and four aldehyde bridging units. Substituents introduced at the β and *meso* position are bound to the pyrrole and aldehyde, respectively. Synthetic β substituted porphyrins resemble those that are naturally occurring; however, *meso*-substituted compounds are more readily synthesized and have found their way into a wide range of applications.
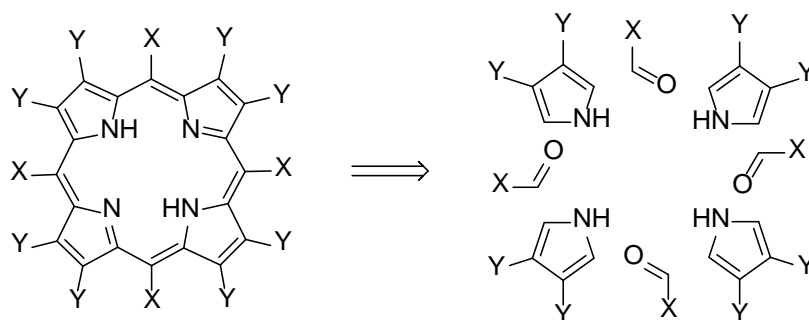
Figure 34. Retrosynthesis of porphyrin illustrating the 4 pyrrole and 4 aldehyde building blocks. X and Y represent substituents in the *meso* and β position, respectively, and are shown to originate from their aldehyde and pyrrole building blocks.

Rothemund reported the synthesis of symmetrical *meso*-substituted porphyrins *via* the condensation reaction between pyrrole and a number of aldehydes.[110] The reaction was carried out at high temperatures, under nitrogen, in a sealed tube with pyridine used as solvent. One of the products prepared was *meso*-substituted tetraphenylporphyrin (TPP) resulting from the reaction of pyrrole with benzaldehyde. These harsh reaction conditions typically resulted with very low yields (<10%).

1.7.4. Adler-Longo Synthesis

Adler and Longo later re-examined the preparation of porphyrins using the synthesis of TPP as their model. They successfully developed an optimized synthetic method by modifying the reaction conditions through adjustments to the acidity, solvent, temperature, initial concentrations of the starting reagents and the availability of atmospheric oxygen. [111]

Adler and Longo experimented by reacting pyrrole and benzaldehyde under various acidic conditions, and found propionic acid to be the most favorable solvent. It should be noted that while the use of acetic acid (bp 118 °C) typically resulted with higher yields (30-40%), propionic acid (bp 141 °C) allowed for easier purification since the resulting porphyrin crystallized out of the reaction mixture when cooled.[112] Furthermore, the reactions were performed under atmospheric oxygen, specifically using a Dean-Stark apparatus, which not only validated that the porphyrin was a result of condensation reactions, but also suggested that there was an apparent need for atmospheric oxygen, implying that oxidation reactions were also occurring. Other findings included: (1) the reaction preferred higher temperatures, (2) the use of equimolar initial concentrations of starting materials gave the highest yield, (3) reactions carried out under nitrogen gave a yield of 5% and (4) the reaction was faster in propionic acid than acetic acid however the yields were reduced by half.[113]

The mechanism proposed by Adler contained four chemical steps: (1) an addition step that results with a carbinol intermediate, (2) a condensation step between the carbinol and another pyrrole molecule, (3) the closure of the open-chain tetrapyrrole to form the cyclic porphyrinogen and (d) the oxidation of the porphyrinogen to the porphyrin.[114]

Dolphin further examined the role atmospheric oxygen plays in the preparation of porphyrins by performing the reactions under anaerobic conditions.[115] His findings proved that porphyrins were a product of the oxidation of porphyrinogen, which he determined to be an intermediate formed by the condensation of aldehyde with pyrrole. It was concluded that the presence of atmospheric oxygen was

necessary for this oxidation step to occur and in turn give the desired porphyrin compound. This finding assisted in determining the overall reaction stoichiometry for the synthesis of TPP following the Adler-Longo, as shown in Scheme 1.

$$4 \quad \underset{\text{pyrrole}}{\text{NH}} \quad + \quad 4 \quad \text{PhCHO} \quad + \quad 3/2 \, O_2 \quad \xrightarrow[\substack{\text{Reflux, 30min} \\ 20\%}]{\text{CH}_3\text{CH}_2\text{COOH}} \quad \text{TPP} \quad + \quad 7 \, H_2O$$
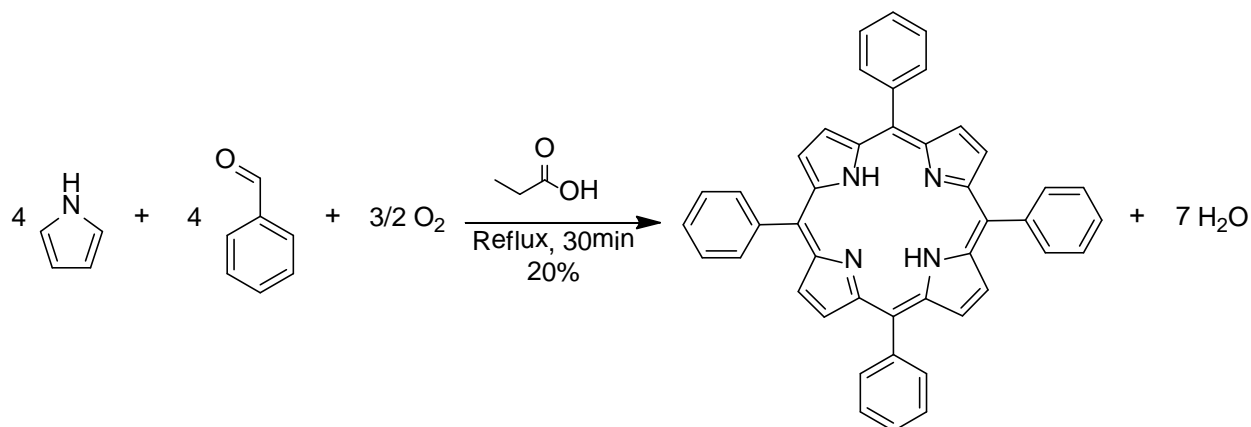
Scheme 1. Synthesis of TPP using the Adler-Longo Method

The primary disadvantage of the Adler-Longo method is the formation of chlorin byproducts. This impurity, however, can be easily oxidized to the porphyrin by treatment with 2,3-Dichloro-5,6-Dicyanobenzoquinone (DDQ) in refluxing toluene[112] or can be removed by column chromatography.[116]

A large number of symmetrical *meso*-substituted porphyrins have been synthesized following this method, including the 5,10,15,20-Tetra(4-pyridyl)-21*H*,23*H*-porphyrin (H$_2$TPyP4) which can be subsequently methylated to afford the *N*-quaternized G4 binding ligand H$_2$TMPyP4.

1.7.5.   Lindsey Method

Unlike the Adler-Longo method, the preparation of porphyrins using the Lindsey method is a two-step reaction between aldehydes and pyrroles, catalyzed by a Lewis or protic acid. This method is typically performed by reacting pyrrole and aldehyde, at dilute equimolar concentrations (10^{-2} M), in DCM at room temperature using TFA, BCl$_3$ or BF$_3$ as the acid catalyst. Originally, triethyl orthoacetate was added to the reaction mixture to function as a water scavenger, however, this is not required but

did slightly increase the percent yield. The first-step of the reaction results in the formation of a colorless porphyrinogen that is subsequently oxidized to the porphyrin in the following step by the stoichiometric addition of DDQ or *p*-chloranil. Lindsey and coworkers' strategy in developing their method centered around the following ideas: (1) that the porphyrinogen is the thermodynamically favored product, meaning that reaction conditions that favor reaching equilibrium prior to the oxidation step should afford porphyrins at a higher yield, (2) benzaldehyde and pyrrole are reactive starting reagents and do not require high temperatures in order for them to react and (3) the synthetic conditions should allow for the use of aldehydes that contain sensitive functional groups.[117] This synthetic method afforded TPP at yields of 45-50%, furthermore it allows for the preparation of symmetrical *meso*-substituted porphyrins using aldehydes that cannot withstand the harsh acidic and high temperature conditions of the Adler-Longo method.

1.7.6. Alternative Porphyrin Syntheses

The synthesis of *meso*-substituted porphyrins bearing up to four distinct substituents was recently reviewed by Lindsey.[118] One of the primary synthetic routes for their preparation is known as the [2+2] method, the condensation reaction of a dipyrromethane and a dipyrromethane-1,9-dicarbinol. The overall substitution pattern is governed by the functional groups introduced *via* acylation at the α-pyrrolic positions of a dipyrromethane. For a porphyrin bearing four different substituents a dipyrromethane containing one substituent, A, is acylated to give the 1-acyldipyrromethane now bearing substituents A-B. The subsequent acylation at positions 9 gives the 1-acyl-9-acyldipyrromethane, now containing the A-B-C substituents. This compound is then reduced to give the dipyrromethane-dicarbinol, followed by condensation with a dipyrromethane containing substituent D to give the porphyrinogen with an A-B-C-D substitution pattern. Subsequent oxidation of this compound affords the porphyrin.
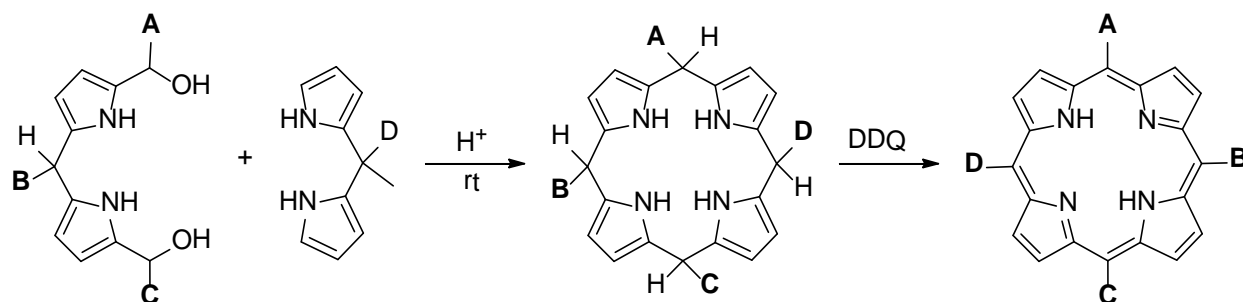
Scheme 2. ABCD-Porphrin prepared *via* the [2+2] method. The ABC dipyrromethan-1,9-carbinol, reacts with the D dipyrromethane to form the porphyrinogen. Oxidation affords the ABCD substituted porphyrin.

1.8. Research Purpose

In this era of genomics and nanomedicines, a huge demand exists for the development of systems that allow for the manipulation of biological organisms through modifications to their genetic material, accomplished by chemical or biochemical means. These techniques could prove to be powerful tools for molecular biology, biochemistry and nanomedicine. Ongoing efforts in our research group focus on employing polymers for this purpose, and studying their interactions with nucleic acids in hopes of gaining a better understanding on how they can participate in the development of such systems. Polymers represent a class of compounds that can be extensively modified to meet the needs for novel biotechnology methods including the manipulation, recognition and delivery of nucleic acids.[119] Polymer-based systems offer enhanced biosafety, biocompatibility, and high flexibility regarding the size of the delivered or targeted nucleic acids in comparison to other DNA interacting materials.[120] They currently serve as important vectors for a number of nucleic acid technologies, such as transfection, gene therapy, RNAi therapies, and we believe they have the potential to assist in the preparation of soluble molecularly imprinted nucleic acid recognition domains.

Spivak and Shea reported the synthesis of MIP receptors for DNA and RNA nucleotide bases in 1998, using ethylene glycol dimethacrylate/methacrylic acid copolymer as the imprinted material.[121] The

synthesis was performed in an organic medium using the covalent molecular imprinting approach. This afforded insoluble MIPs which exhibited binding affinity and selectivity to nucleotide bases. Unfortunately, most biomolecules are insoluble in organic solvents and because biological interactions are mainly based on non-covalent forces, there is a large focus on artificial recognition systems that use such interactions in aqueous solutions. Soluble MIP nanoparticles with affinities for specific nucleic acid sequences and their corresponding structures would be a much more useful biochemical reagent. These polymer nanoparticles may eventually act as a scaffold for artificial restriction enzymes for any arbitrary DNA sequence and as artificial transcription factors.

The interactions between DNA and polymers are of a rather complex nature involving numerous types of non-covalent interactions including van der Waals forces, hydrophobic interactions, hydrogen bonding, dipole-dipole interactions, charge transfer interactions, and electrostatic interactions.[122] In order to achieve the desired highly selective recognition by the MIP the design of the binding site to incorporate intermolecular interactions between template molecule and the functional groups on the imprinting polymer matrix is critical. Orientation in space of the functional groups participating in these non-covalent interactions facilitates a highly cooperative combination of the aforementioned non-covalent forces and improves selectivity by having a complementary fit between the two.

The use of the non-covalent molecular imprinting approach for creating MIPs is analogous to naturally occurring nucleic acid interactions and allows for the use of several different monomer units to simultaneously interact with the nucleic acid pre-polymerization by self-assembly. The greater the variety of non-covalent interactions between the DNA and polymer, the more efficient the artificial binding site becomes. This in turn should lead to a higher degree of selectivity by the imprinted recognition site. Therefore, a combination of monomers may be the best approach for designing artificial binding sites in MIPs. Also, it should be noted that in order to achieve sequence and structure specific nucleic acid recognition by a MIP, the non-covalent interactions should not be limited to the

sugar-phosphate backbone of the nucleic acid but should also encompass the nucleotide bases. Otherwise nucleic acids that form similar structures may exhibit the same binding properties to imprinted recognition sites. Finally, the polymer(s) must also have the potential for cross-linking. Some polymers are readily cross-linked while other may require synthetic alterations to allow for their cross-linking.

This research focuses primarily towards developing a method for generating soluble imprinted polymer nanoparticles possessing affinities for G4 structures formed by specific nucleic acid sequences. Work previously done by our group examined interactions between polymers known to bind DNA, such as poly(ethyleneimine) (PEI) and poly(vinylpyrrolidone) (PVP), with $(T_4G_4)_4$ tetramolecular G4 structures. Early findings suggested that strongly cationic polymers cause a disruption to the DNA structure while neutral polymers do not. We believed that this disruption is caused by strong electrostatic interactions between the polyanionic DNA and the polycationic polymer, resulting with the denaturation of the G4 structure. Later findings suggested that the cationic polymers may be inducing a conformational change to the G4 by altering strand orientations. The electrostatic interactions between cationic polymers and anionic nucleic acids are the primary interactions that results in the observed binding of the two species. Neutral polymers, however, lack the ability to partake in similar interactions. We originally assumed that the identification of polymers that interact non-covalently with template nucleic acids in aqueous media would be sufficient for the development of soluble MIPs that exhibit recognition capabilities under physiological conditions. However, early in our investigation, the screening of commercially available polymers, ranging from neutral to highly cationic, led us to conclude that constructing a MIP solely based on polymer-DNA interactions may prove to be problematic. As a result, the alternative approach we adopted modifies known binding ligands, specifically tetracationic porphyrins, in an attempt to incorporate these structures into the imprinting process. The logic behind this approach is based on our perception that binding ligands can introduce a number of beneficial characteristics to the MIP. (1) They

can act as stabilizing factors for the G4 structure, potentially overcoming the destabilization or conformational changes the polymer may induce on the G4 structure and (2) they can increase the types of non-covalent interactions between the imprinting material and the G4 structure by targeting the nucleotide bases, specifically the G-quartet through external stacking, and as a result significantly increase the material's selectivity.

2. Synthesis of cationic *meso*-substituted porphyrins

2.1. Materials and Methods

All starting reagents were purchased from Sigma-Aldrich or Alfa-Aesar and used without any further purification. Glassware was washed using dilute nitric acid and methanol prior to being rinsed with ddH$_2$O and oven dried. The nitric acid wash is essential for removing any remaining porphyrin that adheres to the glass during previous use. Thin-layer chromatography was performed on aluminum plates that had been pre-coated with silica gel F$_{254}$. Column chromatography was performed on Fisher 60 mesh alumina. Size exclusion chromatography was performed using Sephadex LH-20 purchased from GE Healthcare. Lyophilization was performed using a Labconco FreeZone Freeze Dry System by placing the water soluble compounds in a 20 mL scintillation vial and freeze drying overnight. Porphyrin salts were dried further in an oven to constant weight. UV-Vis spectra were collected on an HP 8452A Diode Array Spectrophotometer using 1 cm pathlength quartz cuvettes in chloroform or ddH$_2$O. ^{1}H NMR spectra were acquired on a Varian Inova AS500 (500 MHz) spectrometer in CDCl$_3$ or CD$_3$OD. High resolution ESI-MS was performed at the CUNY Mass Spectrometry Facility at Hunter College on an Agilent 6520 Q-TOF. MS samples were prepared in chloroform or methanol. All needles and syringes used were sterile. Samples were filtered using PTFE syringe filters or fritted glass filters with medium pore size. Reactions were performed under atmospheric conditions unless otherwise specified. Anhydrous DMF solvent packaged under argon was employed and transferred using an inert atmosphere syringe for syntheses performed under an inert atmosphere of nitrogen. Porphyrins were shielded from light during their synthesis and when stored thereafter.

2.2. Synthesis of 5,10,15,20-Tetra(4-pyridyl)-21*H*,23*H*-porphyrin (H₂mTPyP4)
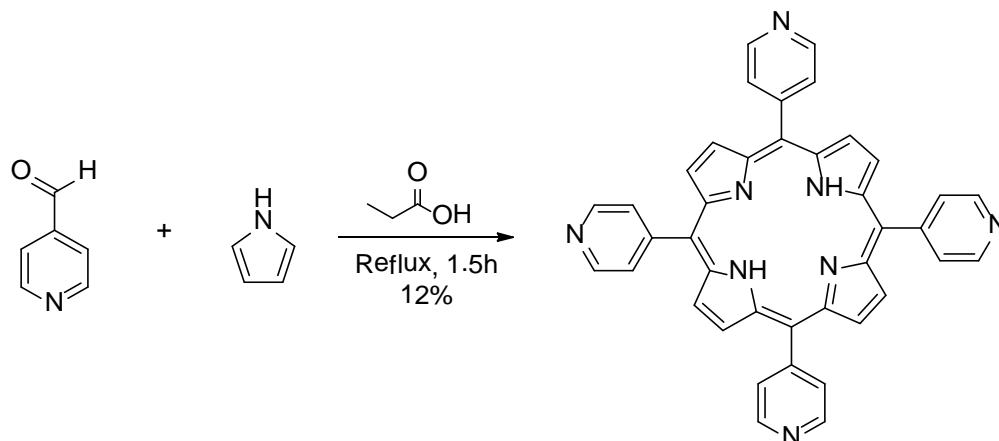
Scheme 3. Adler-Longo Synthesis of H₂mTPyP4

2.2.1. Experimental

The synthesis of the free base H₂mTPyP4 was carried out following the Adler-Longo method of condensation. Pyrrole (9.9 g, 0.15 mol) was added with 4-pyridinecarboxaldehyde (16.1 g, 0.15 mol) to 800 mL of propionic acid in a 2000 mL round bottom flask fitted with a condenser and refluxed for 1.5 h. The solution was then allowed to cool to room temperature and the reaction mixture was vacuum filtered. The filter cake was washed thoroughly with methanol and then hot water affording brilliant purple crystals that were collected and dried in a vacuum oven for 24 h. The crude product was purified on a 60 mesh alumina column using 5:95 methanol in chloroform as eluent. Collected fractions were analyzed by TLC using the same mobile phase. Fractions found to contain pure product, giving one spot on TLC, were combined and concentrated *in vacuo* affording (2.29 g, 3.7 mmol) a 12% yield.

2.2.2. Results and Discussion

The formation of the H₂TPyP4 was clearly demonstrated by the fluorescence it exhibits when dissolved in chloroform and exposed to long-wave UV light. Figure 35 displays the resulting red fluorescence of a prepared low concentration solution. This property allowed for the monitoring of the

compound as it traveled down the alumina column and was found to be a generally useful practice for

detecting the presence of porphyrins.
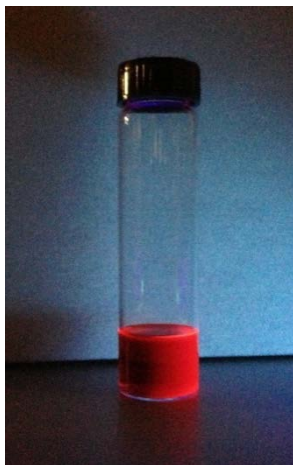
Figure 35. H$_2$mTPyP4 chloroform solution exhibiting fluorescence under long-wave UV

The appearance of the intense Soret band and less intense Q-bands was evident in UV-Vis

absorption data collected, and were found to be in agreement with the literature.[123,124] Figure 36

provides the spectrum of a 1 O.D. solution of H$_2$mTPyP4 prepared in chloroform. The absorption spectra

of all porphyrins were collected at this optical density and below to avoid Soret splitting, which was

commonly observed in solutions with higher porphyrin concentrations, which promote the formation of

aggregates.[125] It should also be noted that the product did not readily dissolve in CHCl$_3$, rather it

required a significant amount of agitation for complete dissolution.
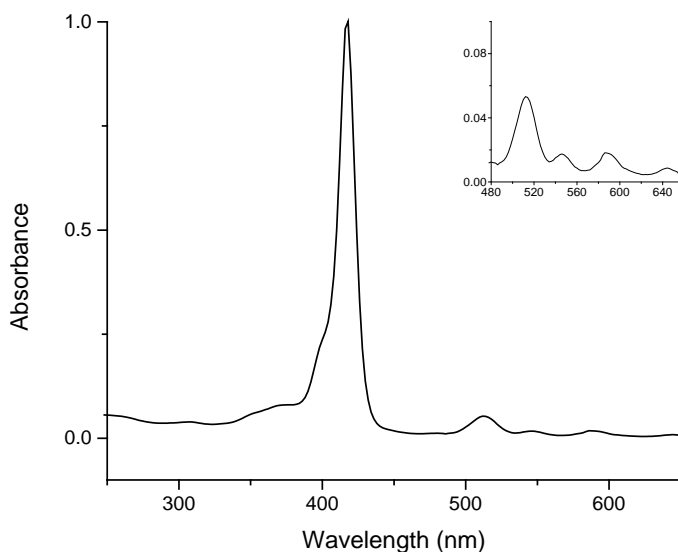
Figure 36. Free base H_2mTPyP4 UV-Visible spectrum collected in chloroform.
Q bands are shown expanded in the insert.

As is observed in Figure 36, the spectrum of H_2TPyP4 collected in $CHCl_3$ displayed a λ_{max} for the Soret band at 418 nm and for the Q bands at 512 nm for $Q_y(1,0)$, 546 nm for $Q_y(0,0)$, 588 nm for $Q_x(1,0)$ and 642 nm for $Q_x(0,0)$.

Further spectroscopic characterization was performed by ^{1}H NMR. The analyzed sample was prepared by dissolving 15 mg of the column-purified product in 2 mL of $CDCl_3$ and filtering the resulting solution through a 45 μm PTFE filter. The collected spectrum of H_2mTPyP4, displayed in Figure 37, contained the anticipated characteristic peaks. There are two sets of equivalent protons on each of the *meso*-substituent pyridine rings, labeled H_A and H_B, found *para* and *ortho* to the nitrogen, respectively. Each one of these sets contains a total of eight protons resulting in the observed low field signals at 9.08 ppm and 8.18 ppm. The 1H-pyrrole and 2H-pyrrole rings each contain two external protons, in respect to the macrocycle, at the β positions. In total they account for eight protons, labeled as H_c, with a corresponding signal found at 8.88 ppm. In principle these are nonequivalent protons, however, rapid transformations occur at room temperature and the signals for these eight β-protons is averaged.[126] It

should be noted that the two different signals that correspond to the pyrrole isomers can be resolved at low temperatures. The internal protons of the NH groups are shielded by the overall macrocyclic current and their resulting proton signals are generally weak and observed at very high field (δ from -1.4 to -4.4 ppm). As is the case with the collected spectrum for H_2mTPyP4, we rarely obtained signals for any of these internal protons.
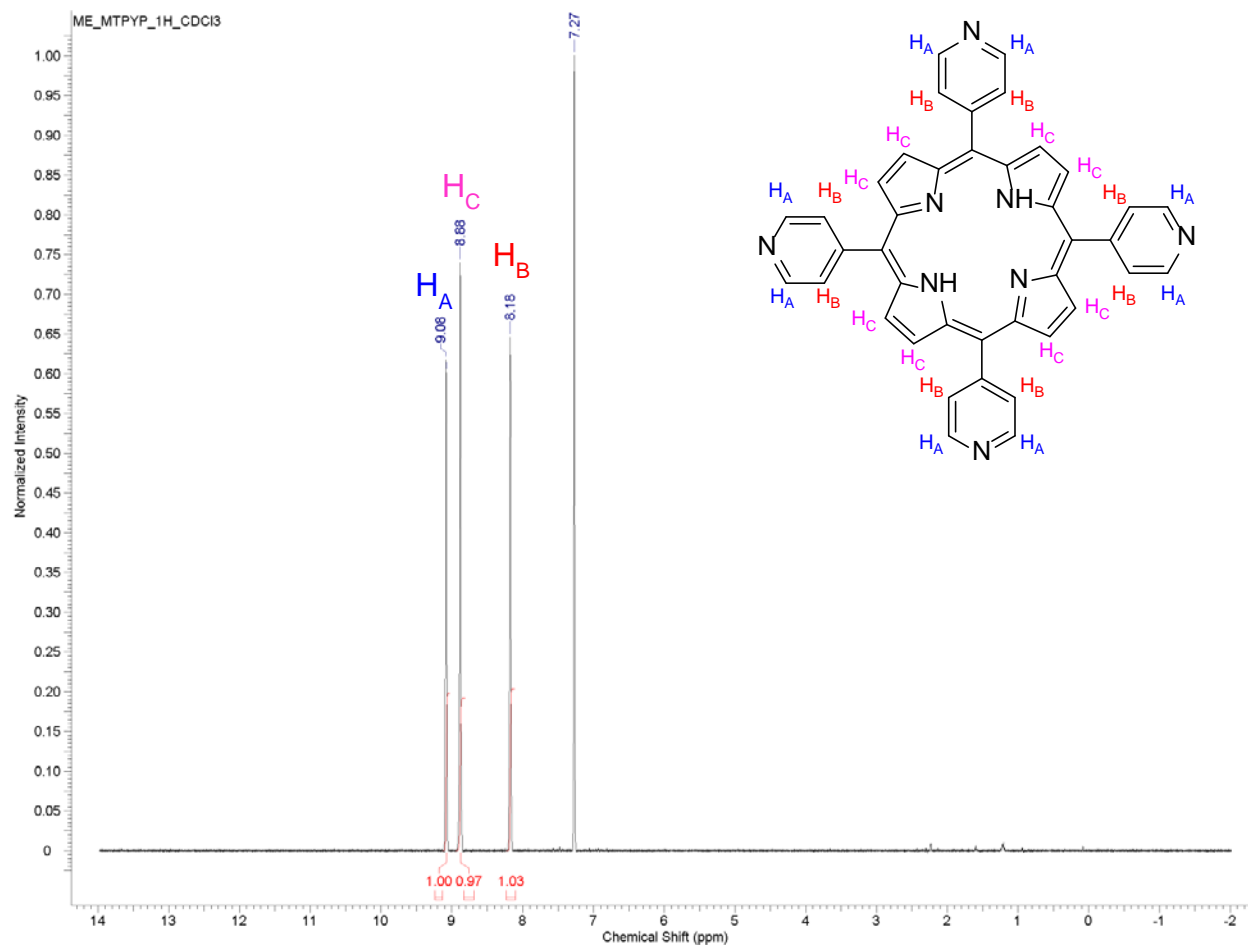
Figure 37. ^{1}H NMR spectrum of H_2mTPyP4. The structure is shown with the protons and their corresponding peaks labeled.

Mass spectrometry was also used to further confirm the successful synthesis of H_2TPyP4. Figure 38 below presents the ESI-MS full spectrum collected. Figure 39 is a zoomed spectrum focusing on the (M+H)+ and (M+2H)+2 peaks which were found to have masses of 619.2354 and 310.1220, respectively.

The target mass of the analyte was calculated to be 618.228 (exact mass calculated by ChemBioDraw

Ultra 12.0). A collection of zoomed spectra displaying the isotope patterns is presented in the appendix.
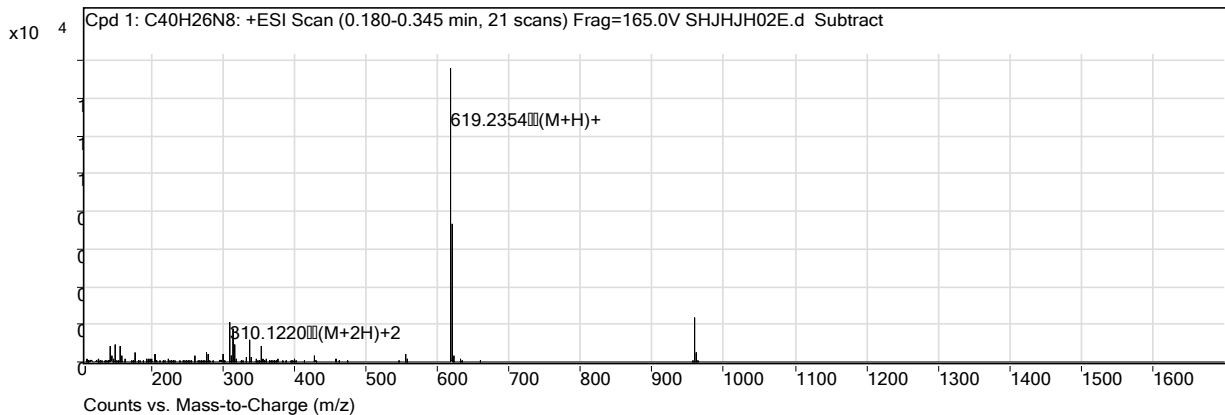
Figure 38. MS spectrum of H$_2$mTPyP4
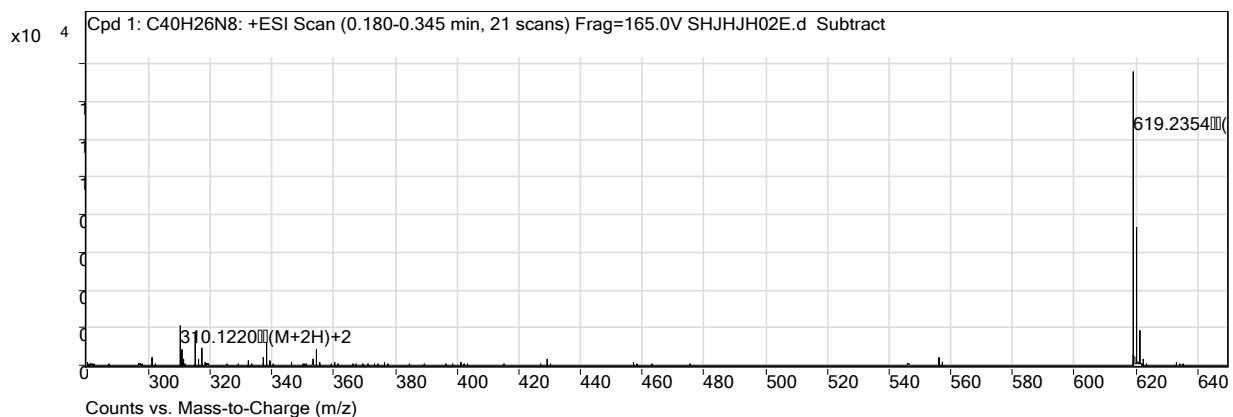
Figure 39. MS zoomed spectrum of H$_2$mTPyP4

The subsequent step towards the preparation of tetracationic porphyrins, which exhibit binding

affinity towards G4 DNA in a similar manner to H$_2$TMPyP4, result from introducing positive charges

through the *N*-alkylation of the pyridine *meso*-substitutents. This is typically accomplished *via* an S$_N$2

reaction, known as the Menschutkin reaction, carried out between tertiary amines and alkyl halides.[127]

The tertiary amines on the H$_2$TPyP4 pyridines are nucleophilic and react with the alkyl halide to yield the

desired quaternary pyridinium salt. This is the most common method for their preparation. The reaction

of substituted pyridines is shown in Figure 40.[128]
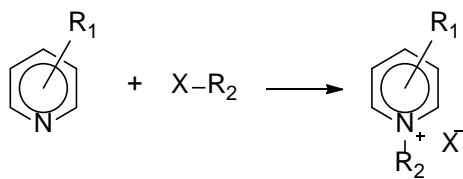
Figure 40. Menschutkin reaction of substituted pyridines

Interestingly, this reaction is considered to be an unusual S_N2 reaction where the reactants are uncharged and the product is charged, in contrast to the more conventional S_N2 reactions where both the reactants and products are charged. Furthermore, the solvent effect is the opposite of typical S_N2 reactions. Charge delocalization at the transition state, resulting from solvent interactions, slows down the reaction in usual S_N2, while in the Menschutkin reaction this solvent effect is favored due to the charge separation it provides.[129] It should be noted that some of the literature suggests an S_N1 and S_N2 mechanism duality.[130]

2.3. Synthesis of cationic tetra-aldehyde functionalized porphyrin

Branched PEI contains primary amines that are cross-linkable. Aldehydes are commonly used as cross-linking agents. Research conducted previously in our lab has employed isophthalaldehyde for this purpose. The cross-linking process is simple and straightforward, primarily due to the amine's high nucleophilicity, it is performed at room temperature by simply adding the cross-linking agent to the PEI. The reaction between amine and aldehyde results in the formation of an aldimine, as shown in Figure 41. Initially, nucleophilic addition results in a carbinolamine intermediate that is subsequently dehydrated to the product. Bifunctional aldehyde cross-linkers, such as isophthalaldehyde, react with two primary amines forming the cross-link. This can occur either inter or intra molecularly with respect to the PEI polymeric chain(s).
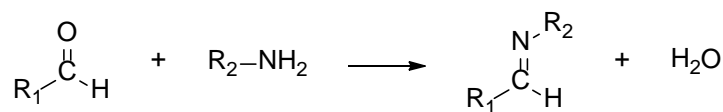
Figure 41. Preparation of aldimine from the reaction between aldehyde and primary amine

In an effort to prepare cationic porphyrin cross-linkers we attempted to synthesize a tetra-functionalized porphyrin bearing four aldehyde groups. We imagine that a structure functionalized in this manner can participate in several approaches in the design and development of soluble G4 MIPs. Theoretically, it can be employed as the cross-linking agent for primary amine containing polymers that have already self-assembled about the G4 structure, and thus holding the polymer chains in complementary spatial arrangement. More importantly, it will contribute to enhancement in selectivity by providing a G4 binding ligand. Furthermore, the adduct's dimensions should direct the cationic porphyrin to interact with the G4 structure *via* an end-stacking binding mode, providing specificity and preventing the porphyrins from intercalation into duplex DNA. Alternatively, the aldehyde functionalized cationic porphyrin can be used in conjunction with other cross-linking agents. Addition of the porphyrin to the G4 DNA should provide stabilization that can deter conformational changes induced by the subsequent addition of the polymer. Once the polymer has covalently reacted with the porphyrin and non-covalently oriented itself around the G4, the polymer itself can be cross-linked using another cross-linking agent, generating the imprinted material.

Our strategy in the preparation of this porphyrin involved the *N*-alkylation of the H$_2$mTPyP4 by the Menschutkin reaction using a protected haloaldehyde, bromoacetaldehyde diethyl acetal, followed by deprotection, *via* hydrolysis to yield the *meso*-substituted tetraaldehyde pyridinium porphyrin salt as presented in Scheme 4. To our displeasure, we were unable to successfully prepare this compound in this manner and to the best of our knowledge, there have been no *N*-aldehyde functionalized *meso*-substituted cationic porphyrins reported to date in the literature. Reactions between H$_2$TMPyP4 and bromoacetaldehyde were attempted under various conditions; changes in temperature (40 °C, 70 °C,

100 °C, reflux), time (3 h, 6 h, 24 h, 72 h), and changes in molar ratios of starting material all proved futile. Analysis of the reaction mixtures, by TLC, UV-Vis and MS, showed no reaction was transpiring and the starting material remained intact.
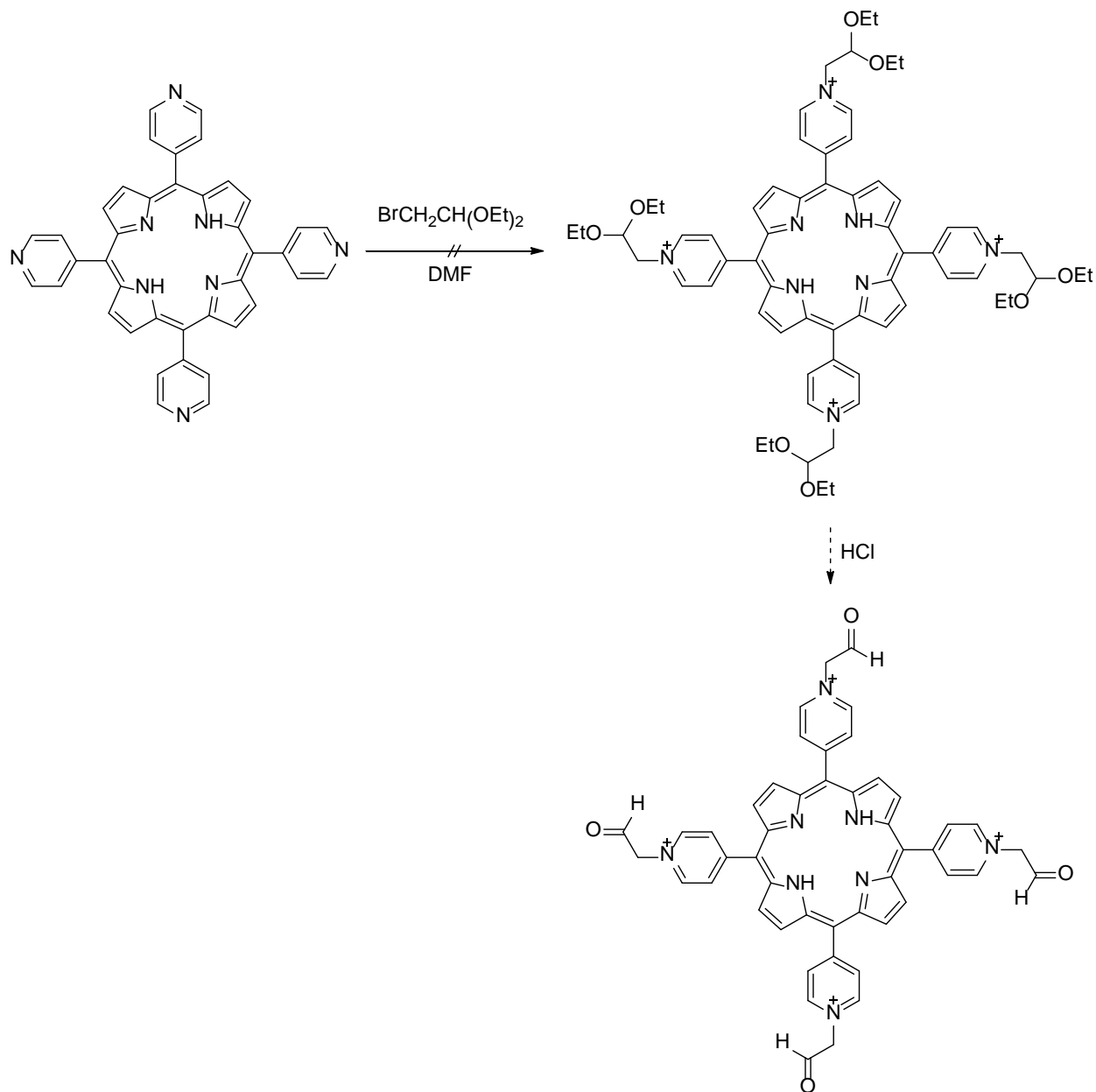
Scheme 4. Attempted synthesis of *meso*-tetra(4-pyridyl(*N*-acetaldehyde))-21*H*,23*H*-porphyrin

2.4. Synthesis of 5,10,15,20-Tetra(4-pyridyl(N-prop-2'-ene))-21H,23H-porphyrin (H2mTAlPyP4)

Since our initial attempts at preparing an aldehyde functionalized cross-linking porphyrin failed, we shifted our attention towards synthesizing a porphyrin bearing allyl functional groups. Commercially available allyl monomers are commonly used in the preparation of highly cross-linked allyl resins and as cross-linking agents for other polymers.[131] Furthermore the allyl groups have the potential to act as initiation sites for polymerization. Unlike the porphyrin cross-linking approach, which centered on the idea of the porphyrin reacting with an already formed polymer, here, the porphyrin derivative can act as a point of origin for polymerization for other DNA-interactive monomers, forming a tentacle porphyrin or a polymer network containing cationic porphyrin moieties. This of course, depends on the degree of polymerization. Initially, we synthesized the allyl functionalized porphyrin by alkylating the H_2mTPyP4 using allyl bromide (3-bromopropene) to yield the tetracationic pyridinium porphyrin H_2mTAlPyP4. Thereafter, we focused on further optimizing this structure for our purposes by attempting to react at the introduced allyl groups. Scheme 5 illustrates the synthesis of H_2mTAlPyP4.

2.4.1. Experimental

The synthesis of 5,10,15,20-Tetra(4-pyridyl(*N*-prop-2'-ene))-21*H*,23*H*-porphyrin (H_2mTAlPyP4) was carried out in a slightly modified fashion to previously reported methods found in the literature.[132] 100 mg of H_2mTPyP4 (161 µmol) and 140 µL of allyl bromide (10X molar excess), measured using a micropipette, was added to 50 mL of DMF in an oven dried 3-neck 100 mL round bottom flask, fixed with a condenser, thermometer and rubber septum. The reaction mixture was shielded from light and heated to 60 °C for 16 h while being magnetically stirred. The progress of the reaction was monitored by UV-Vis, with small aliquots being removed by syringe and analyzed, as well as by TLC using 6:1 $CHCl_3$-MeOH as eluent. Upon completion, the solution was allowed to cool to room temperature and precipitated by adding dropwise into 100 mL of stirred anhydrous diethyl ether. The precipitate was collected by vacuum filtration on a medium size (10-15 µm) glass frit and the filter cake was washed

thoroughly with diethyl ether and chloroform. Five mL of ddH$_2$O was added to the residue in 1 mL aliquots, agitated, and the water soluble salt product, taken up in water, was collected, transferred into an 8 dram vial and lyophilized. The resulting highly electrostatic, low density material was dried in the oven overnight to remove any remaining water affording an 84% yield.
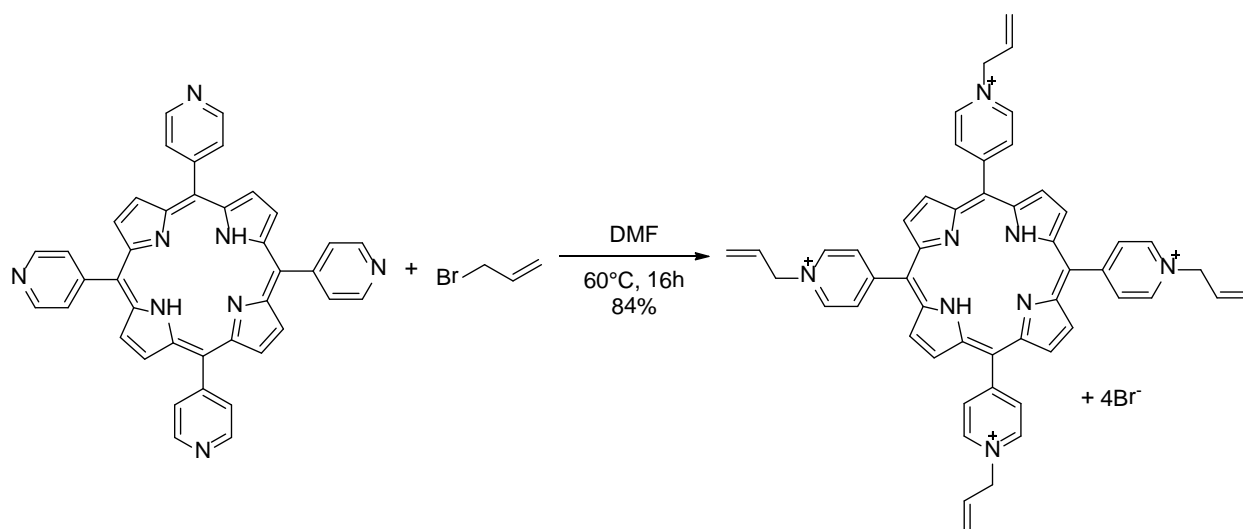
Scheme 5. Synthesis of H$_2$mTAlPyP4

2.4.2. Results and Discussion

One of the most facile techniques for testing the formation of the cationic porphyrins is by examining the solubility of the product in water. The starting H$_2$mTPyP4 exhibits no water solubility; however, the pyridinium salt product is highly soluble. Therefore, by taking an aliquot of the reaction mixture and adding it to a minimal amount of water, filtering the mixture through a 45 µM PTFE filter and checking the filtrate for fluorescent activity under long-wave UV light, the presence of the quaternized porphyrin species could be observed. An alternative technique, following the same premise, is to examine the partition of the reaction mixture between water and chloroform phases, looking for the aqueous phase to exhibit color. An indication that the reaction is complete is by observing a lack of color in the chloroform. Analysis by TLC also proved to be valuable. Using 6:1 CHCl$_3$-MeOH as eluent, the

58

polar $H_2mTAlPyP4$ was found to adhere strongly to silica, while the starting material demonstrated significant mobility. The gradual disappearance of the starting material spot was evident as the reaction proceeded.

Spectroscopically, the UV-Vis spectra collected on reaction mixture aliquots also allowed for the monitoring of reaction progress. As the pyridinium cations are progressively formed a gradual bathochromic shift to the λ_{max} of the Soret band is observed. Complete *N*-alkylation is confirmed when the λ_{max} becomes static. Figure 42 displays the UV-Vis spectrum of $H_2mTAlPyP4$ collected in ddH_2O; the λ_{max} absorption bands for the Soret band is found at 424 nm, 520 nm for $Q_y(1,0)$, 556 nm for $Q_y(0,0)$, 586 nm for $Q_x(1,0)$ and 642 nm for $Q_x(0,0)$. An overlay of the absorption spectra of $H_2mTAlpyP4$ and $H_2mTPyP4$ is presented in Figure 43, demonstrating the characteristic bathochromic shift observed.
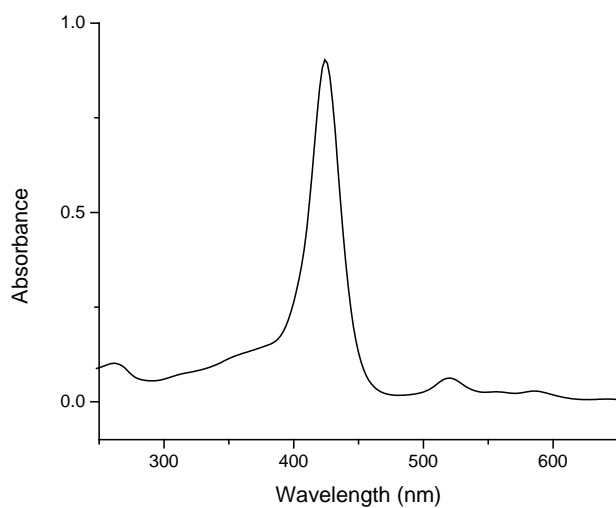
Figure 42. UV-Vis absorption spectrum of H$_2$mTAlPyP4
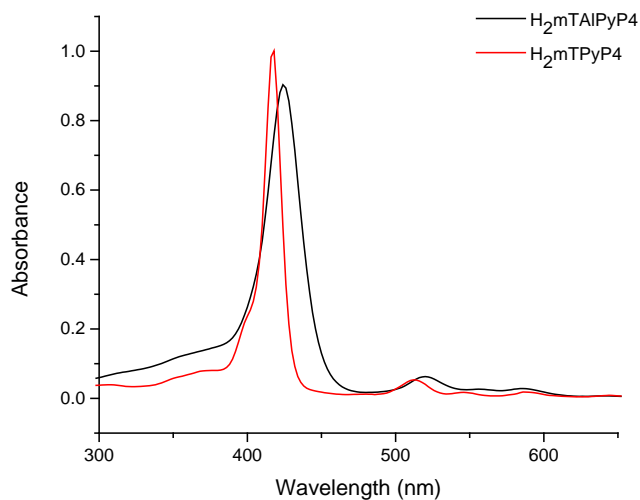
Figure 43. Overlaid UV-Vis spectrum demonstrating the observed bathochromic shift induced by the *N*-quaternization of pyridyl rings (i.e. formation of pyridiniums)

The collected MS spectrum of H$_2$mTAlPyP4 is shown in Figure 44. The target mass of the analyte is 782.358. The base peak is the triply charged species, (M+3H)+3, with a m/z ratio of 260.4592. The doubly charged species was also detected with a m/z ration of 390.1844. Furthermore, inspection of the

60

full spectrum presented in Figure 45 reveals that none of the starting porphyrin material was present in the sample.
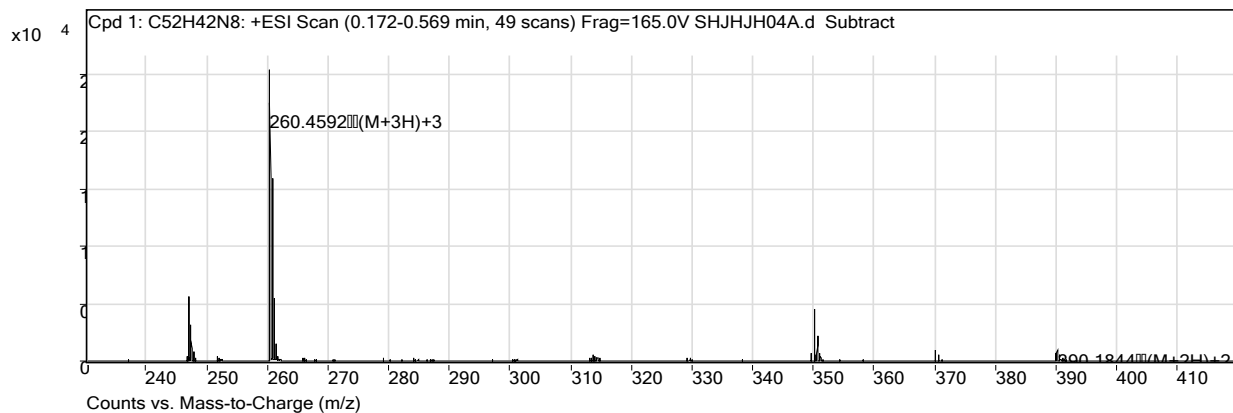
Figure 44. MS zoomed spectrum of $H_2mTAlPyP4$
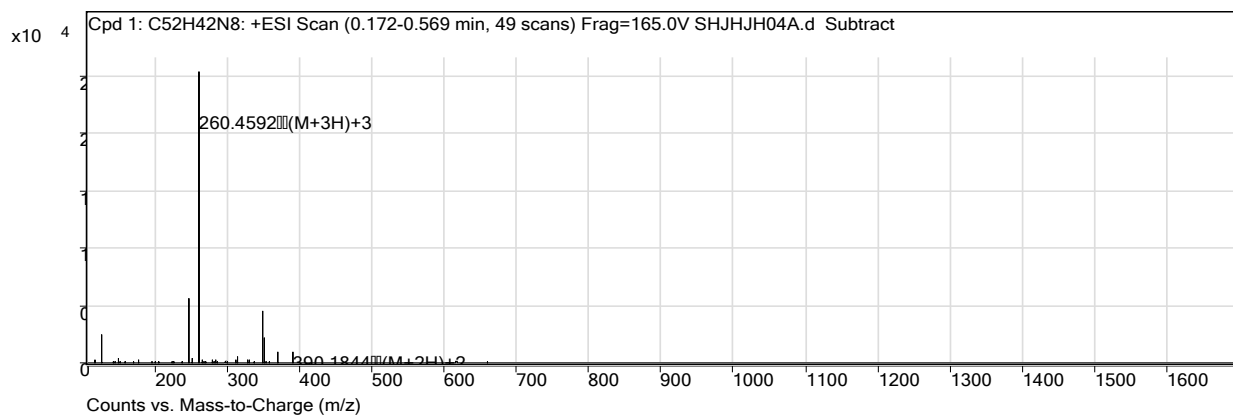
Figure 45. MS spectrum of $H_2mTAlPyP4$

The NMR sample was prepared by dissolving 10 mg of the compound in methanol-d_4 and filtering through a 45 µM PTFE filter prior to analysis. The resulting spectrum indicated the presence of twenty additional protons associated with the four quaternizing allyl groups. This includes a convincing multiplet resulting from the C2 proton found at 6.56 ppm. Further inspection of the spectrum reveals significant broadening of the peak correlating to the pyrrolic protons. Integration of this peak, however,

does in fact give the expected area under the curve corresponding to eight protons. Figure 46 shows a zoomed spectrum displaying the aforementioned peaks.
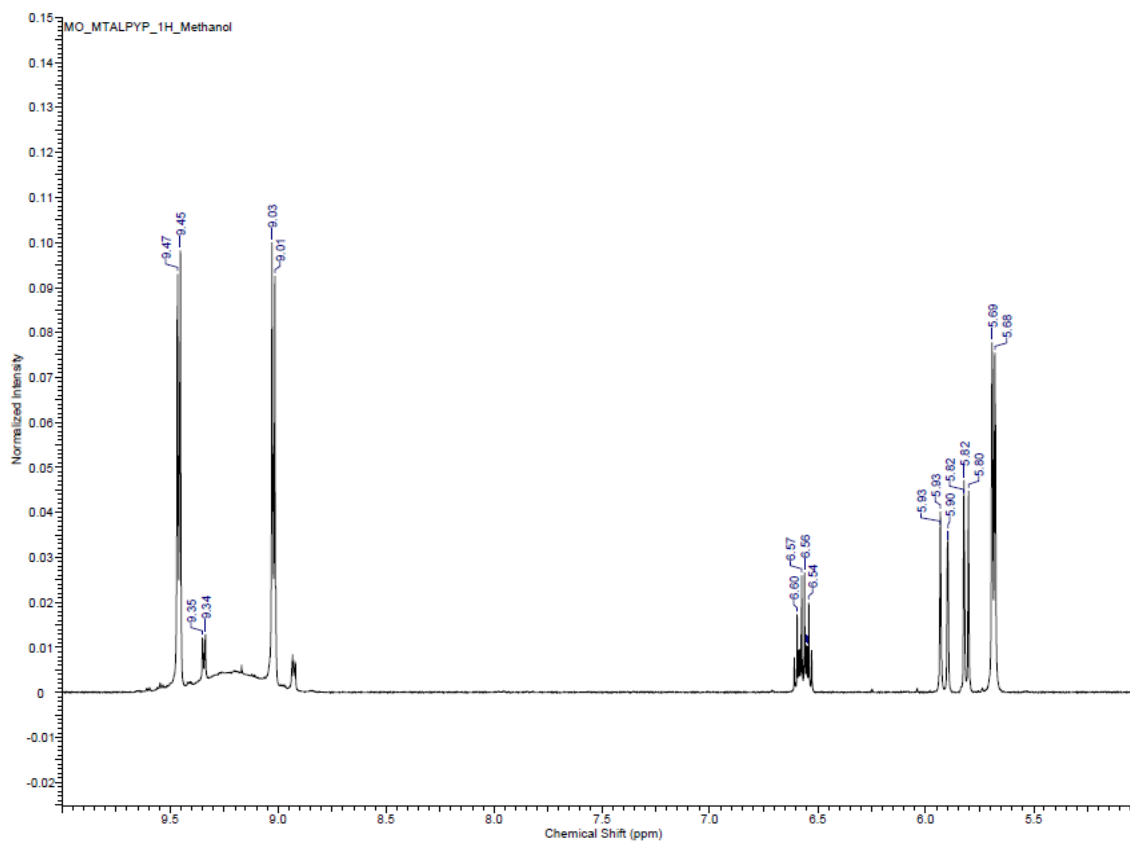
Figure 46. Zoomed NMR spectrum of H₂mTAlPyP4

2.5. Attempted epoxidation of H₂mTAlPyP4

While the direct addition of amines to alkenes is possible through metal catalyzed hydroamination reactions, and may have proven to be sufficient for binding the allyl bearing porphyrin to PEI, we were more inclined to prepare epoxides from the allyl groups due to their expanded scope of reactivity. For example, epoxidation could be followed by nucleophilic substitution reactions, with any polymer of our choosing containing nucleophilic moieties, not necessarily limited to just the amines of PEI. Furthermore, epoxides are more readily polymerizable than allyl groups and would enhance the

initiation sites for polymerization. We attempted the synthesis of a tetra-epoxide functionalized porphyrin by reacting H_2mTAlPyP4 with the peroxyacid mCPBA as shown in Scheme 6. This reaction is known as the Prilezhaez reaction.
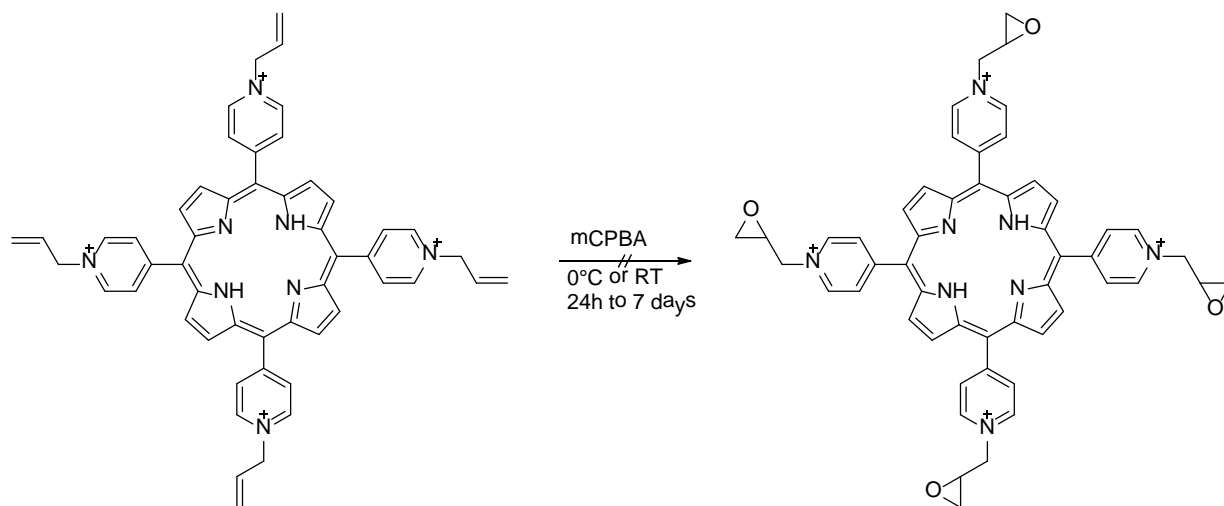
Scheme 6. Attempted epoxidation of H_2mTAlPyP4

We found that the main impediment to the epoxidation was the lack of solubility demonstrated by the cationic porphyrin in nonpolar organic solvents. Typically this reaction readily occurs in nonpolar solvents, where the formation of ions is inhibited. However, H_2mTAlPyP4 only exhibited solubility in methanol and water. We experimented with different solvent systems hoping to find a balance between the nonpolar solvent and solubility requirement of the reaction and the starting material, respectively. Reactions were performed in chloroform, DCM, benzene, acetone, DMSO and DMF following adaptations of procedures found in the literature.[133,134] Variation in the concentration of the starting materials, their stoichiometry, reaction temperature (0 °C and RT), as well as reaction time (24 h to 7 days), all proved to be ineffective as indicated by ^{1}H NMR and ESI-MS analysis.

2.6. Synthesis of *N*-hydoxy pyridinium porphyrins

As a result of the failed epoxidation of H_2mTAlPyP4, we were forced to adopt an alternative scheme for the functionalization of the cationic porphyrin's *meso*-substituents to allow for nucleophilic

substitution. We elected then to attempt the synthesis of a cationic tetra-brominated porphyrin. Since

bromines are good leaving groups, we assumed, they would be suitable for our purposes. Initially, we

attempted *anti*-Markovikov hydroboration-oxidation of H_2mTAlPyP4 using BH_3-THF and H_2O_2 to convert

the allyl groups into primary alcohols, subsequently reacting them *via* the Appel reaction, using Ph_3P and

CBr_4, to give the alkyl bromides. It was presumed that this would yield the tetra *N*-hydroxypropyl

pyridinium porphyrin, as shown in Scheme 7. However, all attempts at this synthesis were unsuccessful.

Consequently, we resorted to quaternizing H_2mTPyP4 with 2-bromoethanol to afford 5,10,15,20-

Tetra(4-pyridyl(*N*-hydoxyethyl))-21*H*,23*H*-porphyrin (H_2mTOEtPyP4) as shown in Scheme 8. Various

procedures found in the literature were performed and evaluated for their efficiency. We found the

most successful to be the method reported by Igarashi,[135] where the reaction is performed in DMF at

fairly dilute concentrations, in contrast to other reported procedures, and with a large excess of the

alkylating agent. This method proved successful and was employed in the preparation of *N*-hydroxy

pyridinium porphyrin with ethyl, propyl, and butyl alcohol *meso*-substituents.
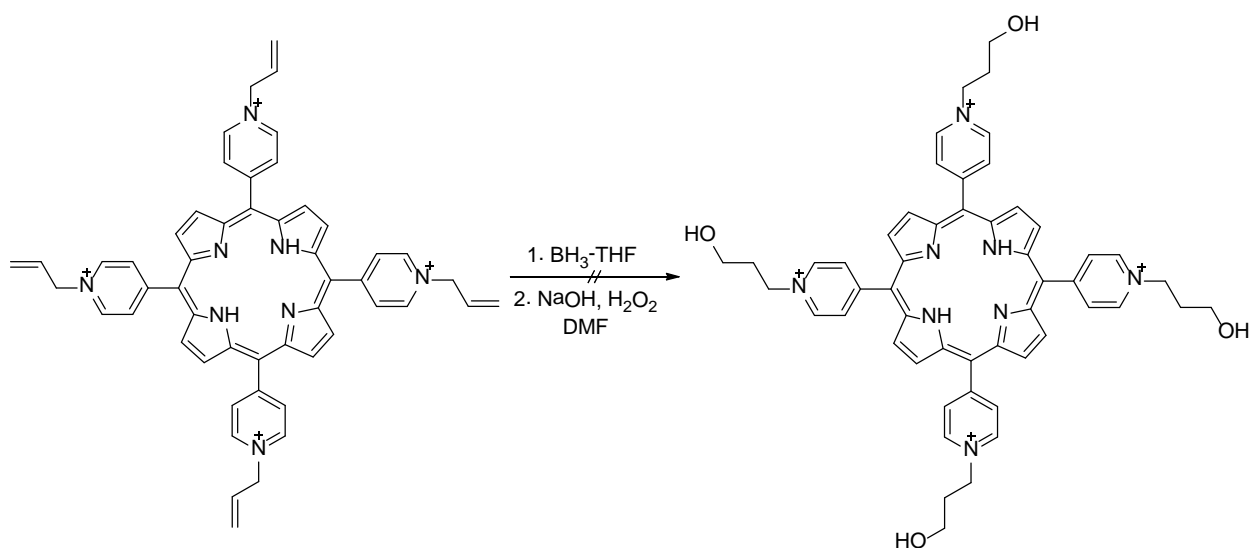
Scheme 7. Attempted hydroboration oxidation of H_2mTAlPyP4
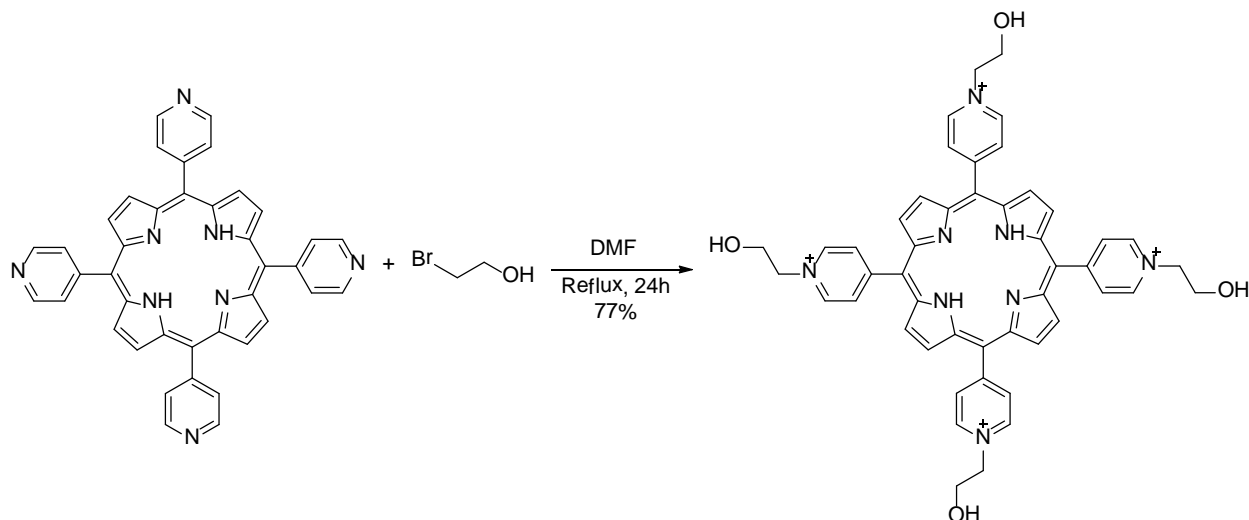
Scheme 8. Synthesis of H$_2$mTOEtPyP4

2.6.1. Experimental

5,10,15,20-Tetra(4-pyridyl(*N*-2-hydoxyethyl))-21*H*,23*H* porphyrin (H$_2$mTOEtPyP4), the hydroxypropyl pyridinium porphyrin (H$_2$mTOPrPyP4) and the butoxypropyl pyridinium porphyrin (H2mTOBuPyP4), were synthesized by refluxing 200 mg of H$_2$mTPyP4 (323 µmol) with 71 mmol of the alkylating agent (5 mL of 2-bromoethanol, 6.4 mL of 3-bromo-1-propanol, 6.5 mL 4-bromo-1-butanol, respectively) in 50 mL of DMF. The reactions were carried out in an oven-dried, 3-neck, 100 mL round bottom flasks each fixed with a condenser, thermometer and rubber septum. Reaction mixtures were shielded from light and refluxed for 24 h while being magnetically stirred. The progress of the reaction was monitored by UV-Vis, with small aliquots being removed by syringe and analyzed, as well as by TLC using 6:1 CHCl$_3$-MeOH as eluent. Upon completion, the solution was allowed to cool to room temperature and precipitated by adding dropwise into 100 mL of stirred anhydrous diethyl ether. The precipitate was collected by vacuum filtration on a medium size (10-15 µm) glass frit and the filter cake was washed thoroughly with diethyl ether and chloroform. Five mL of ddH$_2$O was added to the residue in 1 mL aliquots, it was agitated, and the water soluble salt product, taken up in water, was collected, transferred into an 8 dram vial and lyophilized. The products were then purified on an alumina column.

Fractions that gave one spot on TLC were combined, concentrated *in vacuo* and dried in the oven overnight.

2.6.2. Results and Discussion

UV-Vis spectra collected in ddH$_2$O of the prepared *N*-hydroxy pyridinium porphyrins are shown in Figure 47-49. The expected characteristic bathochromic shift associated with the quaternization of the pyridyl groups, i.e. the formation of the pyridinium salts, is clearly observed for the three compounds. All of the *N*-hydroxy pyridinum porphyrins were found to exhibit a Soret band λ$_{max}$ at 424 nm and to have similar absorption values for their Q bands as show in Table 2.

Table 2. Absorption spectra of *N*-hydroxy pyridinium porphyrins

Porphyrin	Soret	Qy(1,0)	Qy(0,0)	Qx(1,0)	Qx(0,0)
			λ$_{max}$ (nm)		
H$_2$mTOEtPyP4	424	520	554	584	648
H$_2$mTOPrPyP4	424	522	552	584	646
H$_2$mTOBuPyP4	424	520	558	586	646
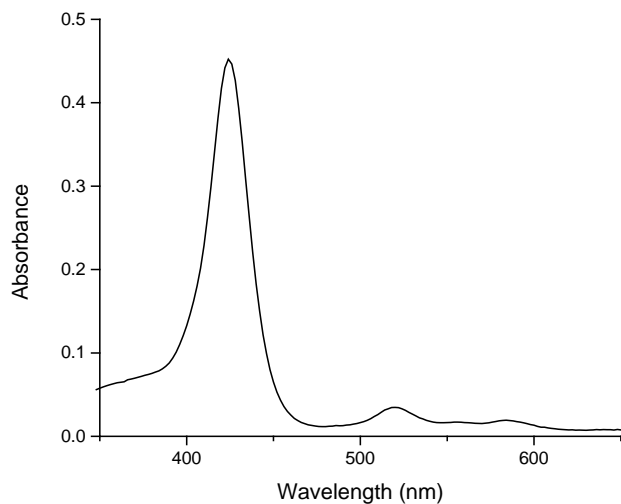
Figure 47. UV-Vis spectrum of H$_2$mTOEtPyP4 in ddH$_2$O
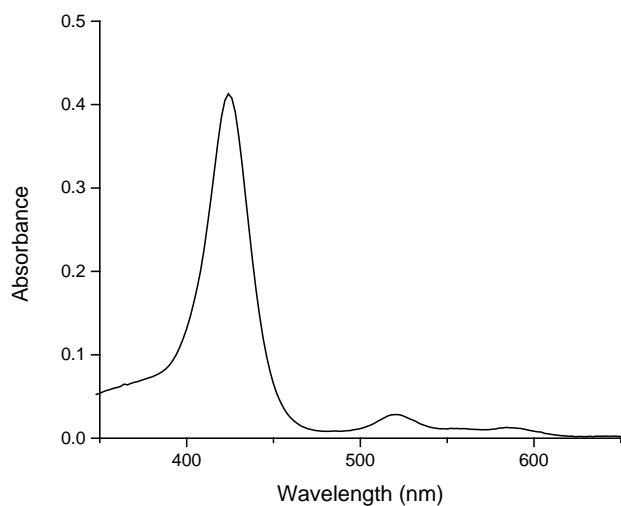
Figure 48. UV-Vis spectrum of $H_2mTOPrPyP4$ in ddH_2O
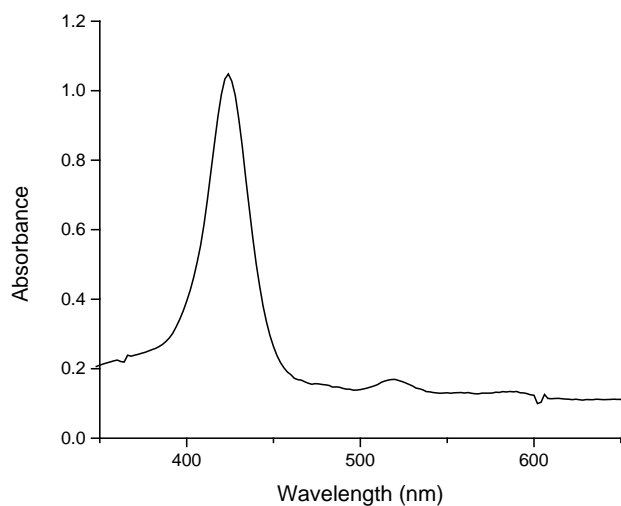
Figure 49. UV-Vis spectrum of $H_2mTOBuPyP4$ in ddH_2O

TLC analysis was found to be appropriate for monitoring the consumption of the porphyrin starting material. However, it proved insufficient for distinguishing between the alkylated species (i.e. mono to tetraalkylated) that are successively formed as the reaction proceeds. This is due to the lack of mobility equally exhibited by all of the alkylated species on polar silica plates. MS analysis played an essential

role for distinguishing between these species and in turn for evaluating the usefulness of the various methods. The first procedure employed in our synthesis, reported by Tovmasyan and coworkers[136], suggested a 98.6% yield, however, we were unable to reproduce these results. Their method used very minimal amounts of solvent, short reaction times and a large excess of the alkylating agent. MS analysis indicated that the crude product recovered, following their protocol, was mainly composed of the monoalkylated species with a small presence of the dialkylated species (approximately 8% dialkylated). Extending the reaction time did not further enable the degree of alkylation. Furthermore, a substantial amount of the starting $H_2mTPyP4$ remained. The MS spectrum collected on the crude product is shown in Figure 50. It clearly indicates the presence of starting material at 619.2353 m/z, the singly charged and doubly charged monoalkylated species at 663.2616 and 332.1356 m/z, respectively.
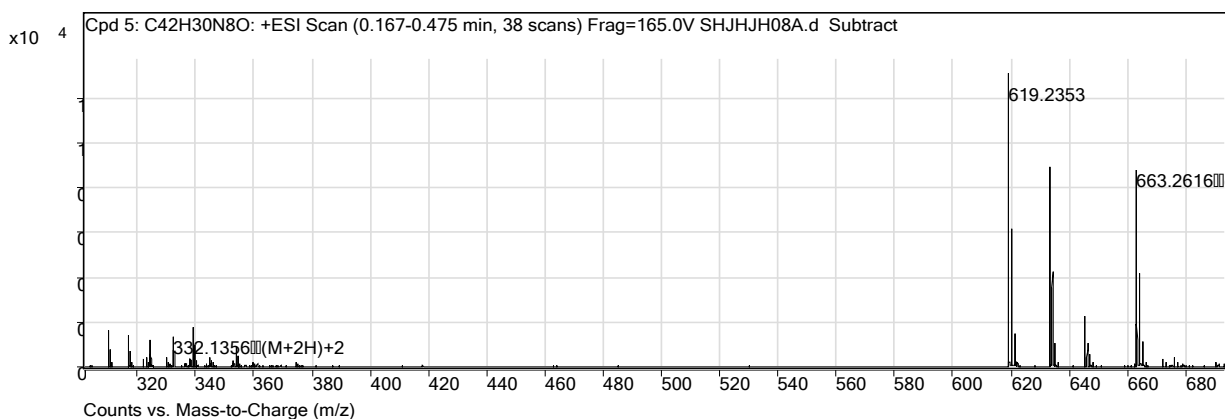
Figure 50. Partial *N*- hydroxyethyl alkylation of $H_2mTPyP4$

The alternative method we employed, reported by Igarashi, proved more successful at generating the tetraalkylated product. MS analysis revealed that the collected crude product consisted of 65% tetralkylated and 35% trialkylated pyridinium porphyrins. None of the monoalkylated or dialkylated species were detected. Figure 51 shows the spectrum collected with the peaks corresponding to the tetralkylated product labeled. Figure 52 shows the identical spectrum with the peaks associated with the trialkylated product labelled.
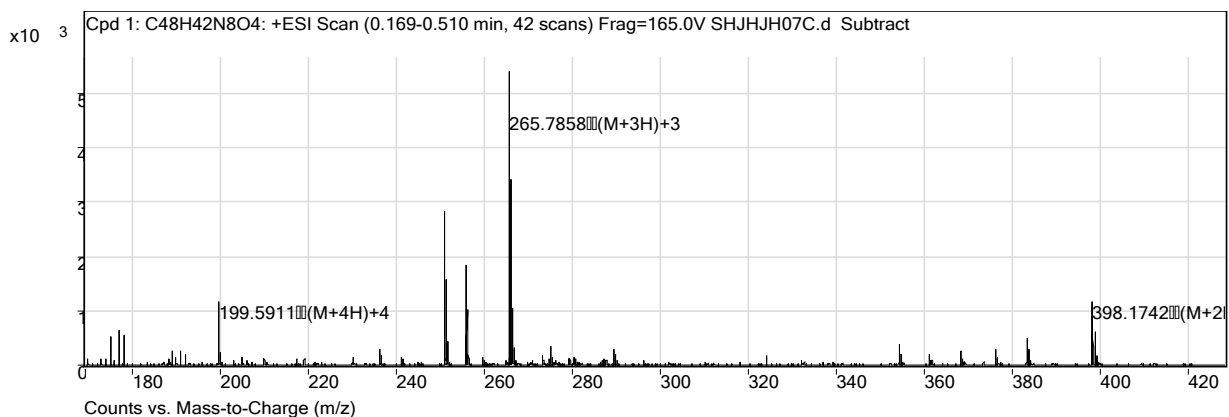
Figure 51. ESI-MS spectrum of H_2mTOEtPyP4
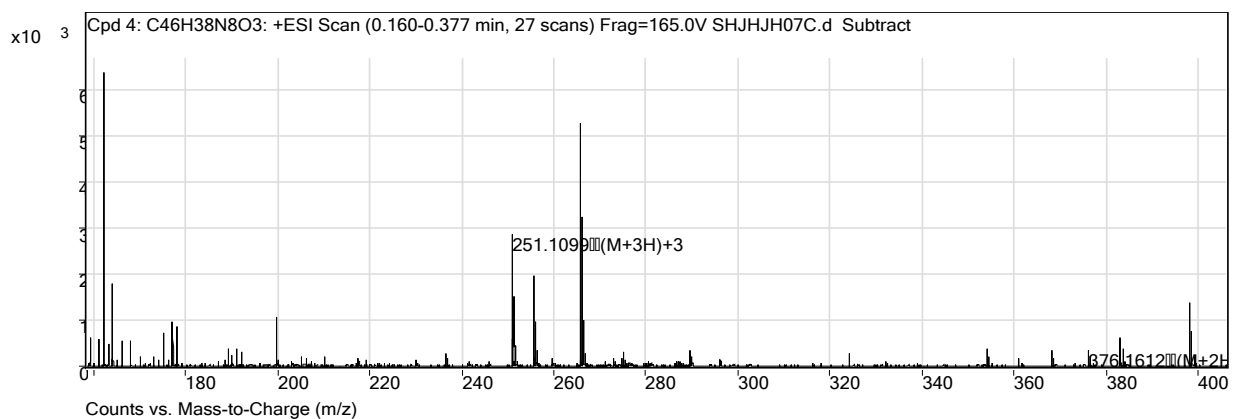
Figure 52. ESI-MS spectrum of the 5,10,15-trisubstituted *N*-hydroxyethyl pyridinium porphyrin

We identified a downside to this method, which is the presence of side product contaminants we believe result from reactions occurring between the 2-bromoethanol and DMF. This might be an explanation to Tovmasyan's decision to use minimal amounts of solvent. Purification of the crude product by column chromatography, using alumina, was sufficient for removing these contaminants. Their presence was not observed during the preparation of H_2mTAlPyP4, and may very well be attributed to the elevated temperatures at which the alkylation with the hydroxyethanols require in comparison to the allylbromide.

Characterization of H_2mTOEtPyP4 was further accomplished via ^{1}H NMR. The collected spectrum shown in Figure 53 is in agreement with the literature. Sixteen total protons corresponding to the -

OCH$_2$- and -NCH$_2$- exhibited chemical shifts of 4.39 (8H) and 5.15 ppm (8H), respectively. The unlabeled peaks are signals resulting from the Methanol-d_4 solvent and residual water. Hydroxyl protons were not detected due to rapid proton exchange with the solvent. The integrated spectra of the prepared *N*-hydroxy pyridinium porphyrins are provided in the appendix.
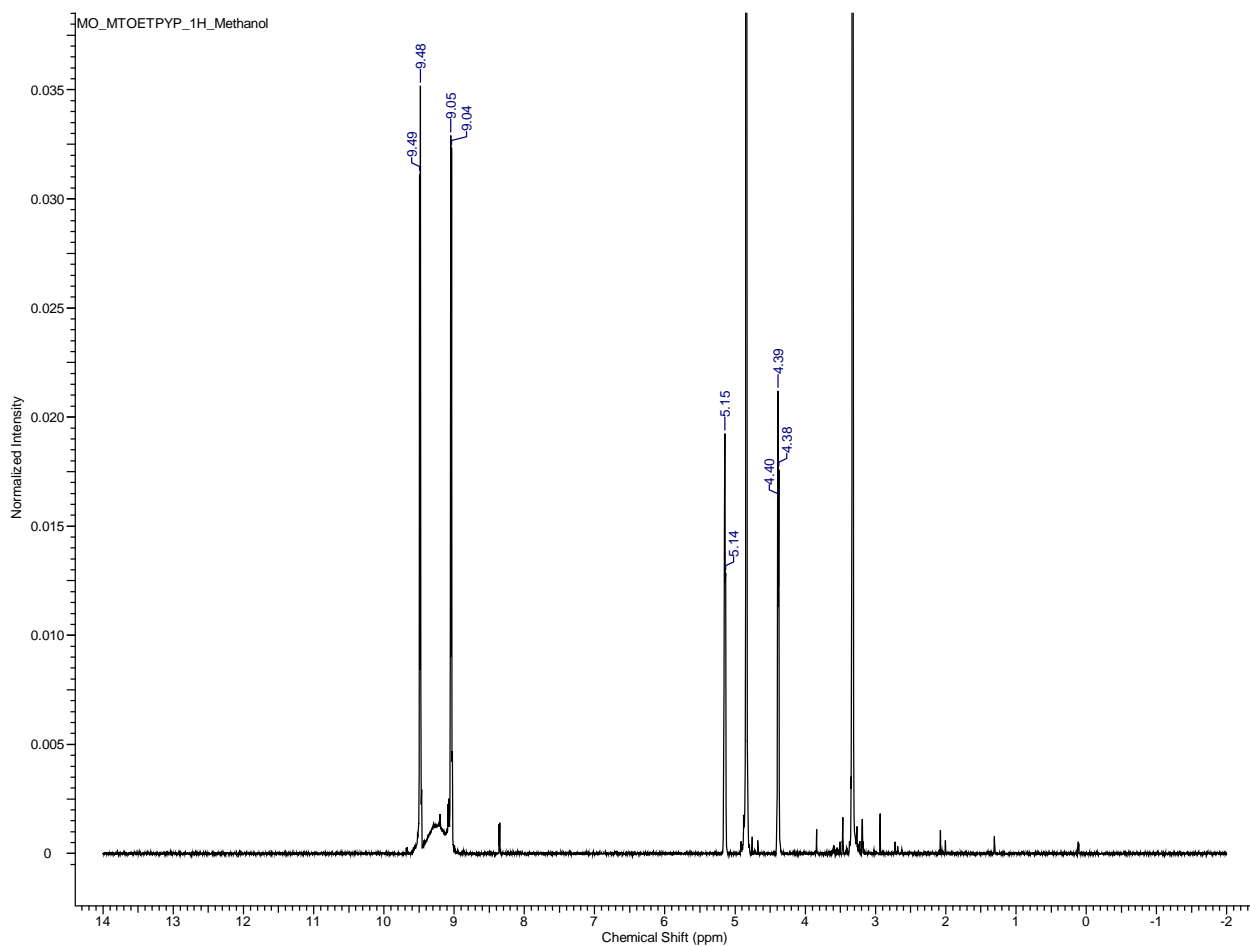
Figure 53. ^{1}H NMR spectrum of H$_2$mTOEtPyP4

2.7. Attempted Synthesis of *N*-alkyl bromide pyridinium porphyrins

Following the preparation of the *N*-hydroxy pyridinium porphyrins we attempted to convert the hydroxyl substituents to the corresponding alkyl bromides. Originally we anticipated that this would be

achieved by treating the alcohol with HBr, resulting in nucleophilic substitution reactions occurring at the hydroxyl groups, generating the alkyl bromides.

$$R\text{-}OH \ + \ H\text{-}X \ \rightleftharpoons \ R\text{-}X \ + \ H_2O \qquad\qquad [5]$$

While primary alcohols are the least reactive, they have been shown to undergo substitution in this manner *via* S_N2, especially with a reactive hydrogen halide such as HBr. Nevertheless, treatment of mTOEtPyP4 with either dilute or concentrated 48% $HBr_{(aq)}$ did not give the desired product. Reactions were attempted by stirring the porphyrin in aqueous solutions of HBr at room temperature and subsequently refluxing the mixture to no avail. Upon addition of HBr the reaction mixtures underwent halochromism, changing from dark purple to green, signifying the protonation of the internal pyrrolic nitrogens. An appropriate concentration of acid, however, was present which was sufficient to allow for the protonation of the hydroxyl groups as well, with the objective of forming the hydronium leaving groups that allows for nucleophilic attack by the bromines. Furthermore, attempts at treating the *N*-hydroxy porphyrins with NaBr and H_2SO_4 to generate the HBr *in situ* also proved futile.

Formation of alkyl halides from alcohols is also achievable using Appel reaction reagents, Ph_3P and CBr_4.[137] The reaction proceeds through an oxyphosphonium intermediate that converts the alcohol into a leaving group that subsequently undergoes S_N2 by a bromine ion generated *in situ*.

$$R\text{-}OH \ + \ Ph_3P \ + \ CBr_4 \ \longrightarrow \ R\text{-}Br \ + \ Ph_3PO \ + \ HCBr_3 \qquad\qquad [6]$$

Originally we were hesitant to attempt this reaction since it has been shown that the reaction between Ph_3P and quaternary salts of nitrogen heterocycles results in the dealkylation of the heterocycles.[138] Interestingly, neither halogenation nor dealkylation was observed in our attempts at this reaction.
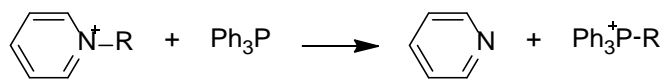
Figure 54. Dealkylation of quaternary nitrogen heterocycles by treatment with Ph_3P

These failed reactions, along with our previous attempts at the hydroboration-oxidation of mTAlPyP4, demonstrate the difficulties we encounter when reacting the porphyrins after their quaternization. Unexpectedly, well known and established methods seem to fail. At first, the solvent limitation was evident, as we encountered during epoxidation of mTAlPyP4. However, even reactions carried out in DMF and methanol were found to be ineffective. It is unclear whether this impediment is a direct result of the cationic charge on the nitrogens or inductive effects resulting from the aromatic pathway of the porphyrin macrocycle. Consequently, it was apparent that in order to prepare an *N*-alkyl bromide substituted pyridinium porphyrin we would be required to do so *via* alkylation and not by the conversion of the quaternizing functional groups to alkyl halides. Therefore, we investigated the quaternization of H_2mTPyP4 with 1,2-dibromoethane. Understandably, modifications to previously adopted procedures would be required in order to prevent polymerization from occurring, since the alkylating agent contains two equally reactive leaving groups. As a result we chose to add a hot solution of dilute H_2mTPyP4 dropwise into a concentrated solution of refluxing 1,2-dibromoethane, with both solutions prepared in DMF. The underlying reasoning behind this being that a dilute concentration of porphyrin in the reaction mixture should limit the alkylation of each porphyrin to react with one of the bromines on the dibromoethane and avoid the formation of dimers and so on. Unfortunately, we found this method to be inadequate and the results were identical to previous results obtained by simply refluxing the 1,2-dibromoethane with H_2mTPyP4, that is the formation of an insoluble polymer.
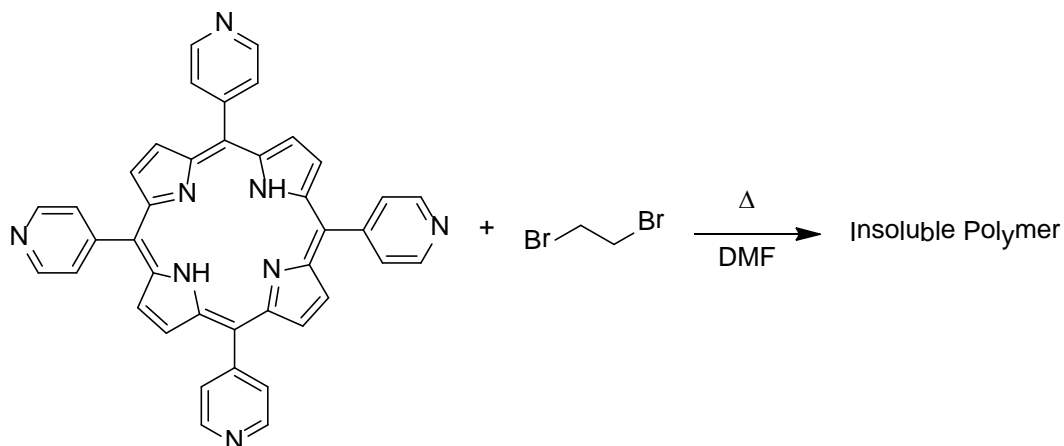
Scheme 9. Polymerization of H_2mTPyP4 with 1,2-dibromoethane

Mesylates are better leaving groups than halides. Therefore, the mesylation of 2-bromoethanol provides an alkylating agent that contains two leaving groups that are unequal in their reactivity. We prepared 2-bromoethyl methanesulfonate with the belief that the *N*-alkylation can be controlled to selectively react with the mesylate, resulting in the formation of the tetra-alkybromide. The synthesis of 2-bromoethyl methansulfonate was carried out following previously reported methods.[139,140] Equimolar amounts of bromoethanol (12.5 g, 0.1 mol) and methanane sulfonylchloride (11.5 g, 0.1 mol) were mixed in 100 mL of diethylether and stirred at 0 °C. To this solution, triethylamine (10.6 g, .105 mol) was added dropwise, at this temperature, and then stirred for an additional 2 hours at room temp. The resulting suspension was poured into 50 mL of ddH$_2$O and the organic phase was collected and subsequently washed with an additional 30 mL of ddH$_2$O. The separated organic phase was dried over Na$_2$SO$_4$, filtered and then concentrated to afford the 2-bromoethyl methansulfonate in the form of an oil that was isolated with an 80% yield.
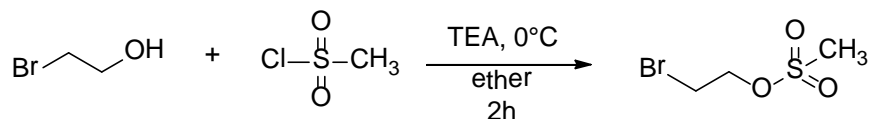
Scheme 10. Synthesis of 2-bromoethyl methanesulfonate

Characterization of the product was performed *via* GC-MS analysis conducted by the Separation Science Research Group.[*] The collected spectrum is shown in Figure 55. Referencing MS databases, the software provided a 98% probability of positive identification of the analyte. Additionally, the gas chromatogram consisted of one single peak, which was indicative of a highly pure product that did not require any further purification.
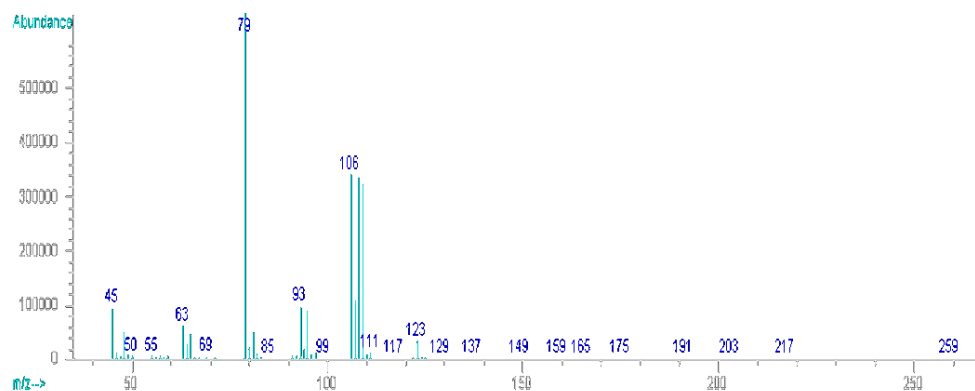
Figure 55. GC-MS spectrum of 2-bromoethyl methansulfonate

Initially, we conducted the alkylation of H_2mTPyP4 using our prepared mesylate ester, as we did previously with the allyl bromide, by adding the porphyrin and alkylating agent to DMF and heating at 60 °C. We found that this temperature was insufficient for allowing this particular quaternization reaction to occur, leaving us with the starting reagents unreacted after 24 hours. Progressively, we increased the reaction temperature until quaternization was detected by TLC and UV analysis, which was found to transpire at 110 °C. This temperature, however, did not produce the desired selectivity we sought. Instead it created an insoluble polymer. Consequently, we attempted the reaction by adding a dilute solution of the porphyrin to a concentrated solution of the mesylate ester. Unfortunately, this also proved unsuccessful and we ultimately concluded that we were unable to selectively alkylate the porphyrin in this manner and thus, diverted our efforts away from further pursuing the preparation of an *N*-alkyl bromide substituted pyridinium porphyrin.

2.8. Synthesis of 5,10,15,20-tetra(4-pyridyl(*N*-carboxypentylpyridinium))21*H*,23*H*-porphyrin

Our inability to prepare an appropriately quaternized H_2mTPyP4 with suitable functional groups, in order to generate a porphyrin cross-linker, forced us to explore alternative avenues. This led to the incorporation of carbodiimide coupling techniques into our strategy. This process is widely used in peptide coupling reactions where the carboxylic acid moiety of one amino acid is activated by a coupling reagent and then reacts with the amine moiety of a second amino acid to produce the desired peptide as illustrated in Scheme 11.[141]
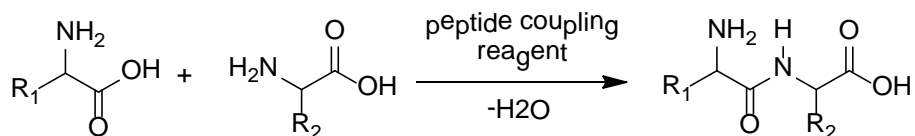
Scheme 11. Peptide coupling reaction

This technique, however, is not limited to the coupling of amino acid residues; rather, it has been commonly used in the formation of carboxamides and is considered as a method for the direct condensation between carboxylic acids and amines. Since PEI contains primary amines, preparing porphyrins functionalized with carboxylic acid groups would allow for the application of this coupling strategy for our intended purposes. Encouragingly, the conjugation of non-cationic porphyrins to peptides and proteins using peptide coupling regents has been previously reported.[142,143] Furthermore, a Fluorescein-PEI coated nanoparticle has been prepared using this technique, which demonstrates PEI's ability to be coupled.[144]

Taking into account the unanticipated lack of reactivity exhibited by previously prepared cationic porphyrins, we were inclined to alkylate H_2mTPyP4 with long carboxylic acid chains in order to minimize any of the inductive effects that may be caused by the porphyrin macrocycle and avoid any interactions with the cationic charge of the pyridiniums. In regards to the solvent limitations, resulting from the lack of solubility exhibited by the pyridinium salt in commonly used organic solvents, there are

a number of highly efficient, commercially available, carbodiimide coupling reagents that have been reported to prepare carboxamides in alcohol and water. These include 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM), dicyclohexylcarbodiimide (DCC), and *N*-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (EDC).[145,146] Accordingly, we decided that *N*-alkylation using 6-bromohexanoic acid, as shown in Scheme 12, would be appropriate for yielding a cationic pyridinium porphyrin that satisfies the criteria for our subsequent coupling reactions.
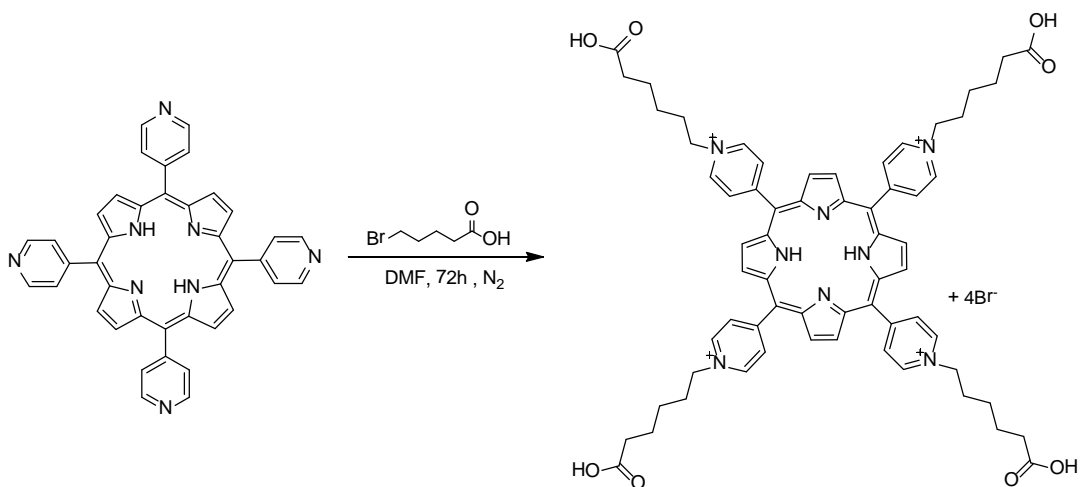
Scheme 12. Synthesis of H$_2$mTCAPyP4

The synthesis of 5,10,15,20-Tetra(4-pyridyl(*N*-carboxypentylpyridinium))-21*H*,23*H*-porphyrin (H$_2$mTCAPyP4) was carried out initially in a similar fashion to previous alkylations i.e. H$_2$mTPyP4 and excess alkylating agent are mixed together and heated in DMF. Experimentation with reaction conditions, assessed using TLC and UV monitoring of the ongoing reaction, revealed the reactions preference for more concentrated reaction mixtures. This was contrary to our inclination since we had originally believed that dilute mixtures would more readily facilitate the formation of the tetra-*N*-alkylated product. Our assumption was due to the gradual decrease in the solubility exhibited by the pyridinium porphyrin, in DMF, after each subsequent alkylation, leading us to postulate that more solvent is required to keep the charged alkylated species in solution. ESI-MS analysis on the crude

product, obtained from reactions using minimal solvent amounts, revealed the presence of a mixture of alkylated species. The di ,tri, and tetra *N*-alkyated porphyrins were detected with the trialkylated being the most prevalent. Only 10.3%, by abundance, of ~~the H₂m~~TPyP4 was found to have been fully quaternized as shown in Table 3. Figures 56 and 57 show the structures with their corresponding ESI-MS spectra and peak assignments, for the non-fully alkylated species.

Table 3. Abundance of alkylated species in the preparation of H_2mTCAPyP4

Compound	Formula	% Abundance
Tetraalkylated	$C_{64}H_{70}N_8O_8{}^{4+}$	10.3%
Trialkylated	$C_{58}H_{59}N_8O_6{}^{3+}$	70%
Dialkylated	$C_{52}H_{48}N_8O_4{}^{2+}$	22.7%
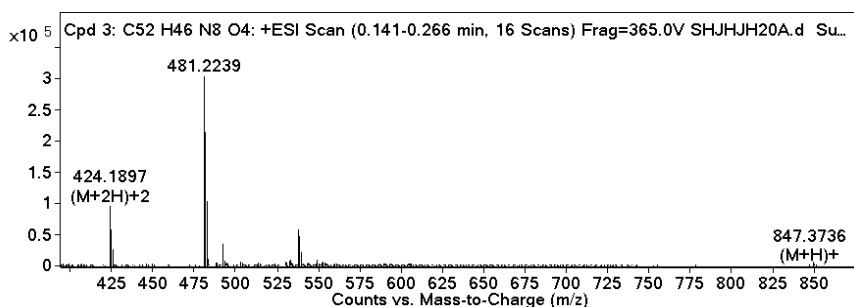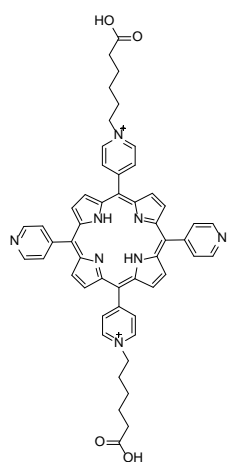
Figure 56. Structure and ESI-MS of 5,15-disubstituted *N*-carboxypentyl pyridinium porphyrin
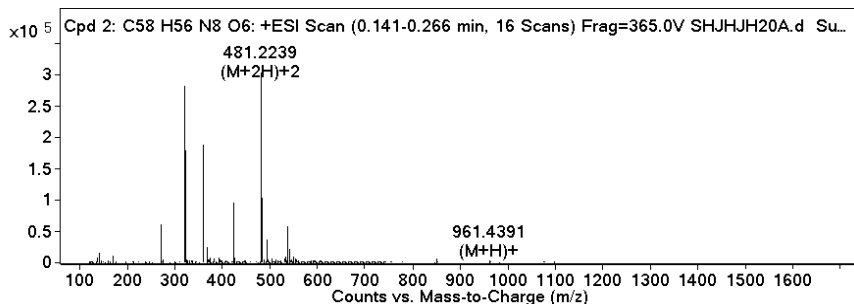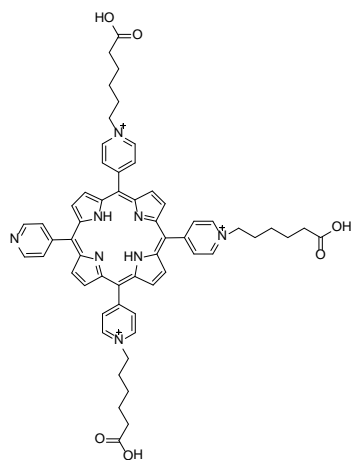
Figure 57. Structure and ESI-MS of 5,10,15-trisubstituted *N*-carboxypentyl pyridinium porphyrin

77

In regards to the fully alkylated porphyrin, the MS detected four peaks corresponding to various protonation states possible for this compound. Complete deprotonation of all four carboxylic acid moieties results in a total of four negatively charged carboxylates. These cancel out the four pyridinium charges, resulting in an overall net charge of zero, making it essentially a zwitterion. Similarly mono-deprotonated $H_2mTCAPyP4$ would have an overall charge of +3. The different protonation states are depicted in Figure 58 along with the ESI-MS spectrum of $H_2mTCAPyP4$, with their distinct peaks labeled.
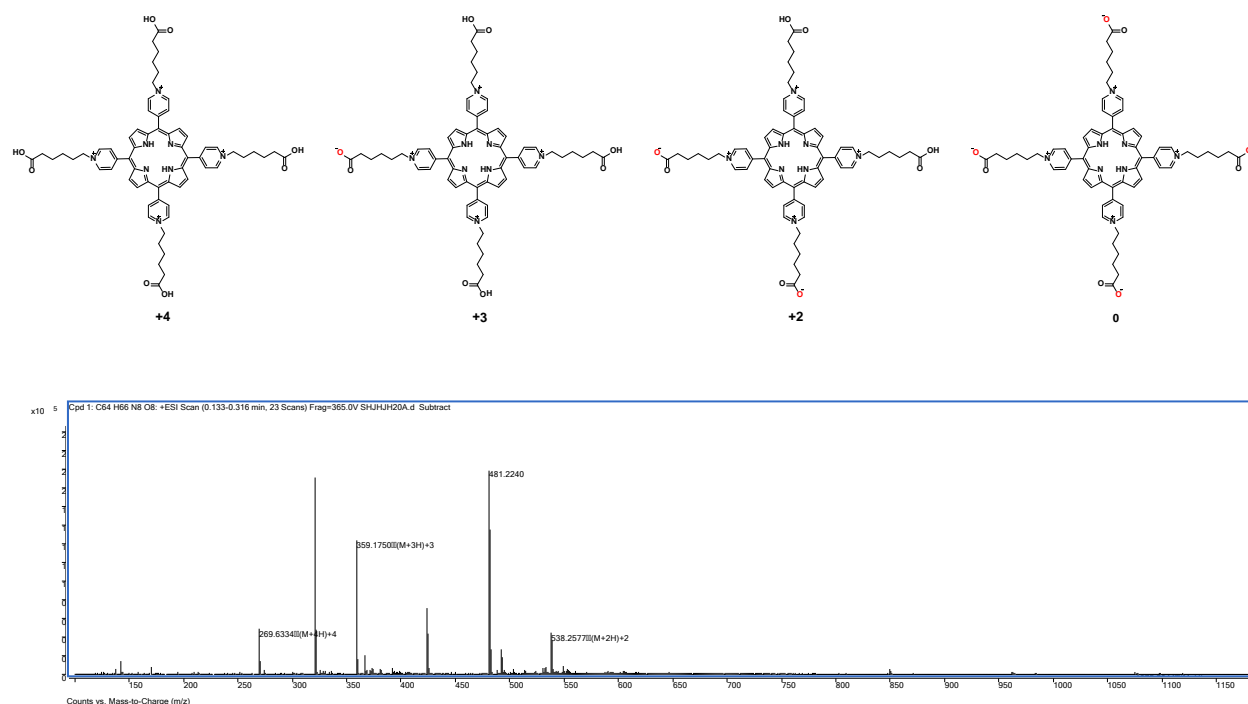
Figure 58. Structures showing the various protonation states of $H_2mTCAPyP4$ with the charges indicated below and ESI-MS displaying their associated peaks

Optimization of the reaction conditions was required to push the reaction forward and increase the yield of the fully quaternized product. We performed the synthesis from this point forward using fresh, individually packaged, small bottles of anhydrous DMF packed under argon and the reaction itself was carried out under an inert atmosphere (N_2). We found it advantageous to heat the $H_2mTPyP4$ in DMF to 100 °C, as to completely solubilize the porphyrin, prior to the addition of the alkylating agent. It

was also determined experimentally that 40X excess of 6-bromoheaxanoic acid (by mole ratio, 10:1 by functionality) and a reaction time of 72 hours, to be the most favorable conditions. Furthermore, upon completion, the reaction was cooled to 0 °C, instead of room temperature, as was done during previous alkylations. These modifications afforded a crude product that was found to be 54.5% tetraalkylated.

Isolation of the tetraalkylated species was a daunting task. Separation of the crude product components was possible by RP-HPLC, using previously reported conditions.[147] However, the obtained chromatograms were highly convoluted, consisting of multiple peaks that were difficult to assign compounds to without the advantage of an MS detector. The multiple peaks correlate with the degree of alkylation, the various protonation states that each one of these species can assume and possibly atropisomerization, resulting from the interaction between a carboxylate folding back to associate with the charge on the pyridinium. Attempts at isolating the components by preparative HPLC were unsuccessful. As a result we resorted to size exclusion chromatography (SEC), which we found to be a rather sufficient technique for isolating the tetraalkylated product. Exhaustive SEC was performed on Sephadex LH-20 columns, isocratically, with methanol as the eluent. Analysis by ion exchange chromatography on the column purified product, performed by Frontier Scientific (Salt Lake City, Ut), revealed that the isolated material was composed of 90.1% tetraalkylated pyridinium porphyrin.

The UV-Vis spectrum of H_2TCAPyP4, collected in ddH_2O, is shown in Figure 60. It exhibits the expected bathochromic shift associated with quaternization. The Soret band was found to have a λ_{max} of 424 nm, 520 nm for $Q_y(1,0)$, 556 nm for $Q_y(0,0)$, 584 nm for $Q_x(1,0)$ and 640 nm for $Q_x(0,0)$.
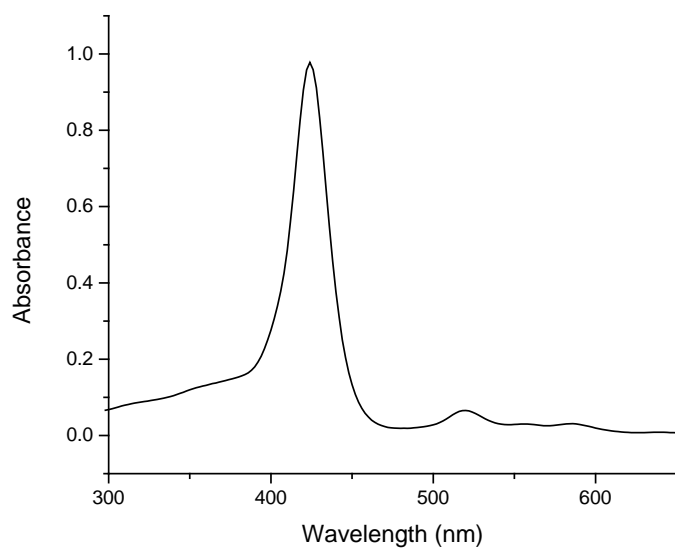
Figure 59. UV-Vis spectrum of H_2mTCAPyP4 in ddH_2O

Characterization by ^{1}H NMR was used to further confirm the successful synthesis of H_2mTCAPyP4. The sample was prepared by dissolving 5 mg of the SEC-purified product in methanol-d_4. The collected spectrum is shown in Figure 61. The compound contains 70 protons, 26 associated with the porphyrin macrocycle, 2 internal protons do not give signals and 44 associated with the four hexanoic acid substituents, with no signals detected for the 4 carboxylic acid protons. The integrated spectrum is provided in the appendix displaying the 66 detected protons.
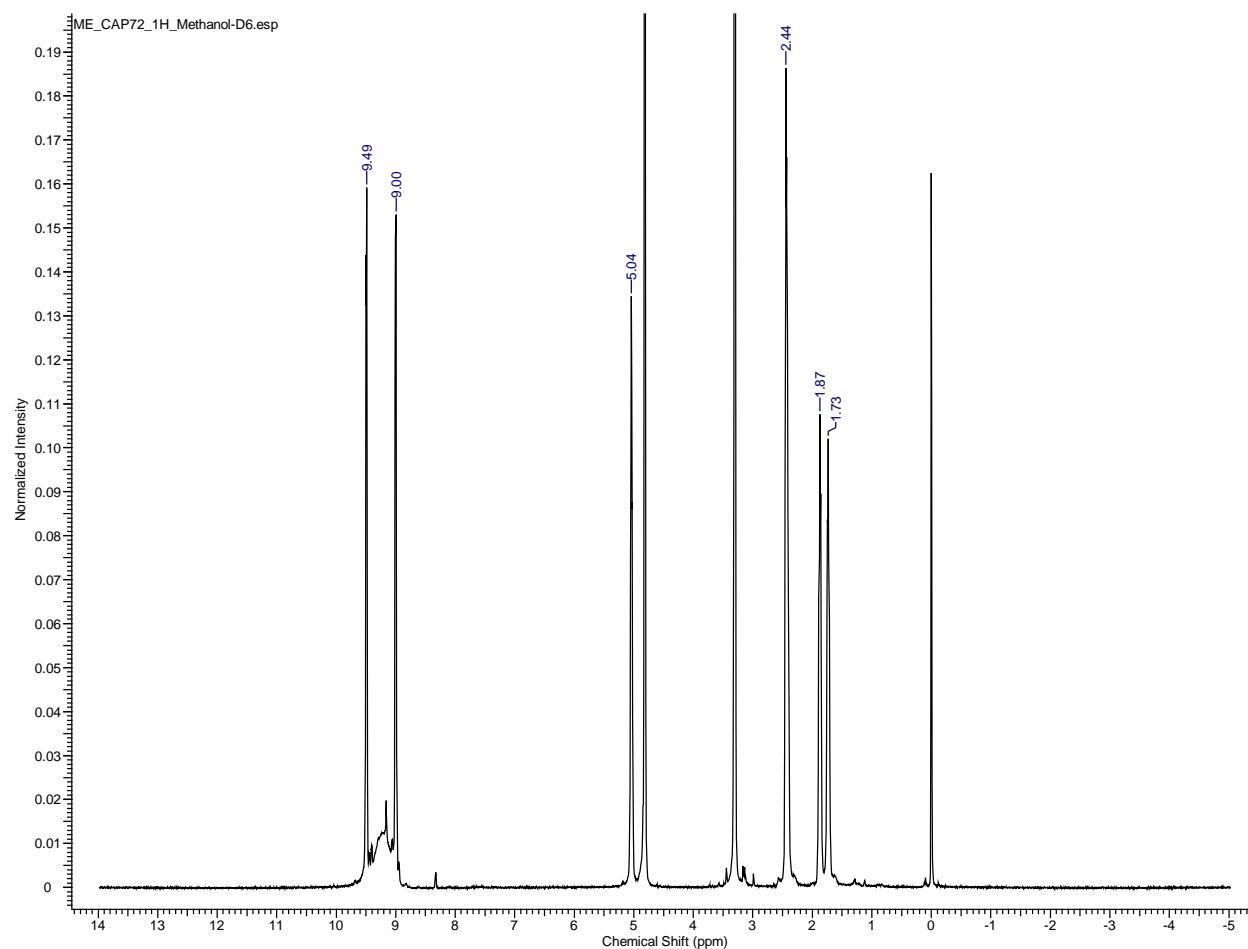
Figure 60. ^{1}H NMR spectrum of H$_2$mTCAPyP4

2.9. Summary

In our pursuit of synthesizing cationic *meso*-substituted porphyrin cross-linkers for the preparation of a molecularly imprinted polymer specific for G4 DNA, we have made and explored the reactivity of a number of *N*-alkylated pyridinium porphyrins. The experimental results suggest that the potential for subsequent reactions involving prepared tetracationic compounds is challenging. Therefore, alternative methods were adopted in order to functionalize cationic porphyrins with carboxylic acid moieties to allow for the use of coupling reagents to facilitate their ensuing reactions with amine containing polymers.

mTPyP4 was synthesized using the Adler-Longo method. *N*-alkylation was achieved using the Menschutkin reaction. The synthesis of mTAlPyP4 was performed as previously described in the literature.[132] Attempts at the hydroboration-oxidation of this compound and its epoxidation proved futile. Consequentially, we synthesized a series of *N*-hydroxy porphyrins, with oxyethyl (mTOEtPyP4), oxypropyl (mTOPrPyP4) and oxybutyl (mTOBuPyP4) substituents. Attempt to convert the alcohols to alkyl bromides by treatment with HBr or *via* the Appel reaction failed. As a result, we tried to alkylate using 1,2-dibromoethane. We found that we were unable to control this reaction in a manner that prevented polymerization from occurring. The preparation and use of 2-bromoethyl methanesulfonate as an alkylating agent did not exhibit the desired selectivity we sought and also resulted in an insoluble polymer. The only unsuccessful quaternization reaction we encountered was the alkylation of H_2mTPyP4 using bromoacetaldehyde diethyl acetal.

The alternative strategy employing amide coupling agents required us to prepare the carboxylic acid functionalized, mTCAPyP4. The initial reaction conditions were found to be insufficient and required optimization, which we were able to achieve, resulting in a significant increase in the formation of the tetralkylated product from 10.3% to 54.5%. Isolation of the product was accomplished by exhaustive SEC using Sephadex LH-20.

All reactions were adequately monitored by TLC and UV-Vis analysis. Characterization was accomplished by ^{1}H NMR and ESI-MS. The prepared *meso*-substituted cationic porphyrins were prepared for subsequent use in DNA binding studies and the mTCAPyP4 would be specifically examined for its cross-linking ability.

3. Polymer Cross-Linking

3.1. Screening of Polymers

At the onset of this research we recognized the importance of identifying polymers that interact favorably with nucleic acids, as to not induce conformational changes in the higher order structures and allow for the construction of an accurate imprint. Through our initial screening of commercially available polymers, we quickly accepted that polymers that bind DNA are cationic and interact primarily, non-covalently, through electrostatic interactions with anionic DNA. Consequentially, one of the perceived benefits to using cationic porphyrin cross-linkers is its ability to stabilize G4 structures and perhaps overcome the destabilizing effects of the polymer employed in the imprinting process. Table 5 below gives the list of polymers examined for their affinity towards DNA.

Table 4. Polymers screened for their affinity towards DNA

Polymer	MW
Polyethylenimine	50,000
Poly(4-vinylpyridine *N*-oxide)	200,000
Polyvinylpyrrolidone	40,000
Poly(2-vinylpyridine *N*-oxide)	300,000-400,000
Poly(*N*-Isopropyl acrylamide)	40,000
Poly[bis(2-chloroethyl) ether-alt-1,3-bis[3-(dimethylamino)propyl]urea]	N/A
Poly(diallydimethylammonium chloride)	100,000-200,000
Polyacrylamide	10,000
Poly(ethylene glycol) diglycidyl ether	526
Polyethylenimine	1,200
Polyethylenimine	25,000
Poly(acrylamide-co-diallyldimethylammonium chloride)	232.75
Poly(2-vinylpyridine)	50,000
Poly(4-vinylpyridine)	60,000
Chitosan	100,000-300,000

3.1.1. Materials and Methods

Agarose Gel Electrophoresis

0.5% agarose gel was prepared by adding 0.5 g SeaKem LE Agarose, purchased from Cambrex, to 100mL of 0.5X TBE buffer. Mixture was swirled and heated for 70 seconds in a microwave, allowed to slightly cool, and 10µL ethidium bromide was added. Gel was poured and solidified before samples were pipetted into the wells. Polyplex samples were prepared by adding 50 µL of 100 ng/µL CT DNA with 100 µL of 100 ng/µL of polymer. Samples were incubated for 30 minutes, mixed with loading buffer, and 10 µL were loaded onto the gel.

UV-Vis Spectroscopy

Hypo- and hyperchromicity studies were performed on a Cary Varian 3E UV-Vis spectrophotometer. 10 µL of 1 mg/mL polymer was added to 1 mL of 33 µg/mL DNA in a quart cuvette. A spectrum was obtained for each polymer-DNA solution ranging from 220-320 nm at 25 °C.

3.1.2. Results and Discussion

One of the most telling pieces of data collected during the screening of these polymers came from gel electorophoretic experiments that clearly demonstrated the hindered mobility of DNA resulting from polymer binding interactions. A picture of a 0.5% agarose gel, placed on a UV transilluminator, resulted in the gel mobility bands shown in Figure 61. Lane 1 contains 10 µL of 10 ng/µL CT-DNA (100 ng) in 0.5X TBE. Lane 2-15 contain 10 µL of 10 ng/µL CT-DNA with 1 µL of 100 ng/µL solution, prepared in the same buffer, of a given polymer added. Lanes 2, 9, 10, 13 and 15 demonstrate how polymers that bind DNA completely restricted its mobility. The DNA in these samples bound to the polymer so strongly they forced the DNA to remain within the well. The polymers corresponding to these lanes are PEIs of varying molecular weights in lanes 2, 9 and 10, poly(diallydimethylammonium chloride) in lane 10 and Poly[bis(2-chloroethyl) ether-alt-1,3-bis[3-(dimethylamino)propyl]urea] in lane 15. These three polymers are all cationic. Their structures are given in Figure 62. None of the neutral polymers exhibited this same

effect or showed any detectable disruption to the mobility of DNA. Their bands were found to migrate

similarly to the control DNA sample in lane 1.

Figure 61. 0.5% agarose gel used to detect DNA-polymer binding

Figure 62. Structure of (A) Poly[bis(2-chloroethyl) ether-*alt*-1,3-bis[3-(dimethylamino)propyl]urea] (B) PEI (C) poly(diallydimethylammonium chloride)

UV-Vis experiments revealed that these three cationic polymers have a hyperchromic effect on

CT-DNA, which is indicative of duplex denaturation. Even though this data suggests that the polymers

are participating in the destabilization of the structure, it does confirm the occurrence of DNA-polymer

interactions.

Figure 63. Hyperchromicity induced by cationic polymers on CT-DNA

Figure 64 provides the change in absorbance at 260 nm resulting from the addition of 10 µL of 1 mg/mL, of an investigated polymer, to 1 mL of 33 µg/mL CT-DNA. The data clearly reveals that neutral polymers have a hypochromic effect, which suggests contraction of the DNA helix. Whether this is due to repulsive forces exhibited on the duplex by the polymer or external stacking interactions is unclear. However, the prior scenario is more likely. Nevertheless, cationic PEI clearly has the most pronounced effect, indicating that it has the greatest amount of interaction with the DNA.
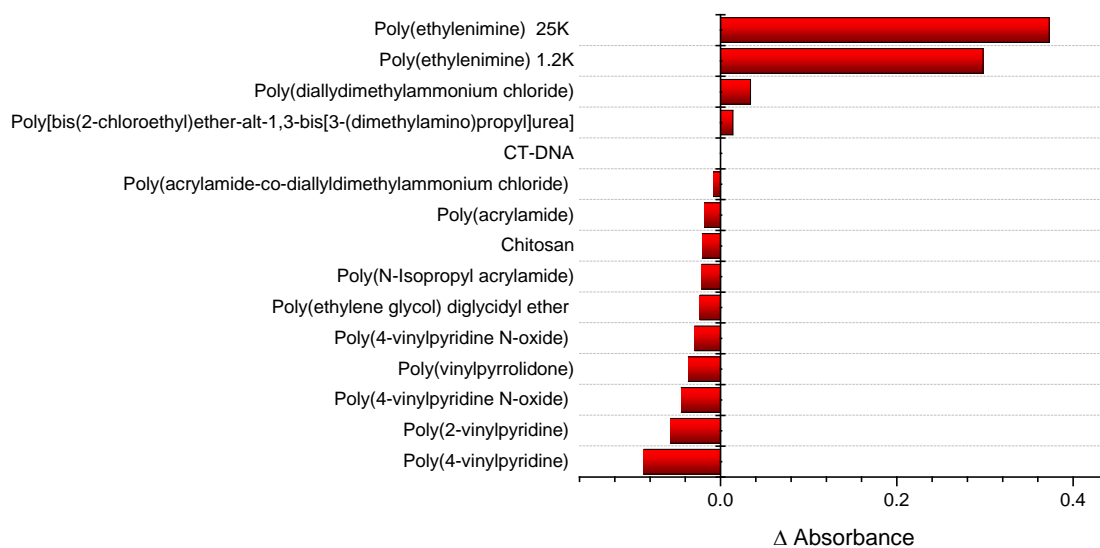
Figure 64. DNA and polymer difference spectra at 260 nm

From the cationic polymers that have been found to bind DNA, PEI is the only one that contains amines that are available to react. The cationic nature of PEI results from the protonation of its amines in aqueous media, while the cationic charge found on the other two polymers is due to their prior *N*-quaternization, preventing the ability to react these amines further.

3.1.3. Poly(ethylenimine)

PEI can be synthesized in different lengths, branched or linear, and undergo functionalized group substitution or addition.[148] Branched PEI is prepared either by the cationic polymerization of aziridine (ethylene imine) at high temperatures in aqueous or alcoholic solvents or at low temperatures by acid catalyzed polymerization, in bulk, as shown in Scheme 13.[149]
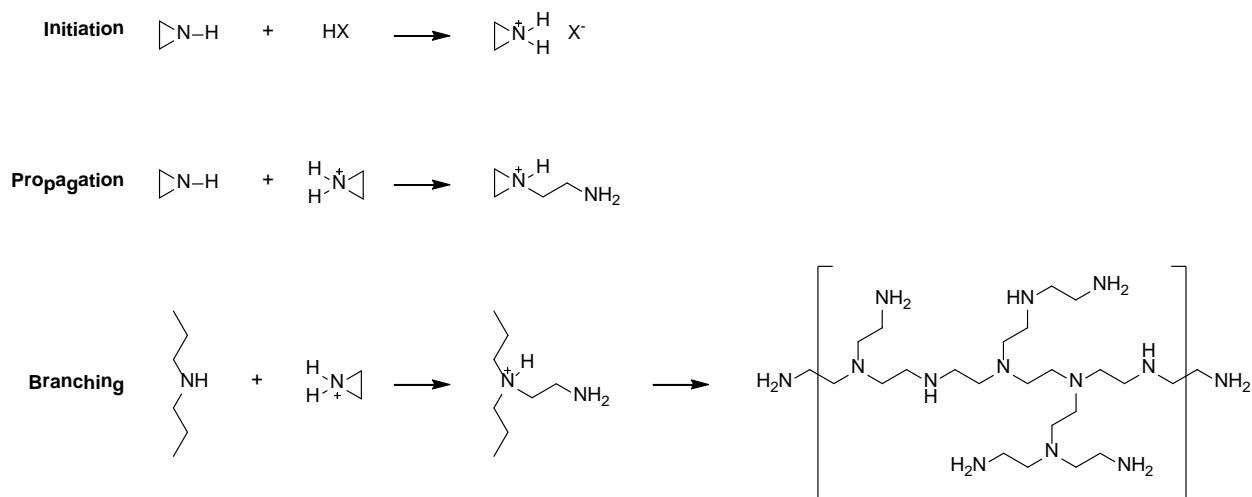
Scheme 13. Acid catalyzed bulk polymerization of ethylene imine in aqueous solutions

The resulting polymer contains primary, secondary and tertiary amines in approximately 1:2:1 ratio. Linear PEI is prepared by the hydrolysis of poly(2-ethyl-2-oxazoline) and as a result is composed predominantly of secondary amines, with at most two terminal primary amines.[150] Accordingly, we employed branched PEI for its primary amines and their potential for cross-linking *via* coupling reactions, with the carboxylic acid moieties found on H$_2$mTCAPyP4.

PEI has been widely used as a non-viral vector for gene delivery and is known to successfully complex DNA molecules. Boussif and coworkers demonstrated that PEI could be used to successfully deliver DNA *in vitro* and *in vivo*.[151] It was chosen for this use primarily due to its high cationic charge density, which condenses DNA to form a compact polyplex structure that shields the DNA from non-specific interactions with other molecules.[152,153] Furthermore, non-protonated amines on the PEI provide a buffering effect, a property that allows for its participation in the "proton sponge" mechanism, which has been postulated to facilitate and enhance transfection efficiency.[154] It should be noted that the basicity of amines in aqueous solutions follows the order: 2°>1°>3. The literature is inconsistent in regards to the reported protonation sates of PEI.[155] Suh and coworkers[156] suggested that 10-20% of branched PEI is protonated at the physiological pH of 7.4, while Nagaya and coworkers[157] later proposed

protonation of 30% of the amines. Choosakoonkriang[158] and von Harpe[159] report the chain length

dependent pKa's for a series of PEIs is in the range of 8.2 to 9.5, which suggests that most of the amines

(~90%) would be protonated at physiological pH, while other studies suggest that all forms of PEI have

approximately 50% of their amines protonated. This uncertainty in the actual PEI protonation state

causes difficulty in estimating the overall charge of PEI, which is an important factor to consider when

determining the optimal conditions for binding with anionic DNA without causing a disruption to its

native confirmation. Interestingly, Sun and coworkers[160] have demonstrated that PEI effectively induced

and stabilized G4 structures through the condensation effect. However, the topologies of the induced

G4 structures differed from the topology of the G4 structure formed without being under the influence

of PEI. This observation suggests that the G4 structure readjusts in a manner to adapt to the surrounding

conditions i.e. the condensation effect of PEI.

The degree of polymerization (DP) is defined as the total number of monomeric units in a

macromolecule. It can be calculated with respect to M_w or M_n as shown in Equations 6 and 7,

respectively. Where M_w is the weight average molecular weight, M_n is the number average molecular

weight and M_0 is the molecular weight of the monomeric unit. In polydispersed polymers the M_w is

always greater than the M_n. Accordingly, the calculated DP_w indicates a greater number of monomers

than DP_n..

$$DP_W \equiv X_w = \frac{M_w}{M_0} \qquad [7]$$

$$DP_n \equiv X_n = \frac{M_n}{M_0} \qquad [8]$$

The PEI used in our studies were classified according to their MWs, therefore, calculations are

based on DP_w. The molecular weight of aziridine is 43.07 g/mol. A 10,000 MW branched PEI polymer is

calculated to have a DP_w of approximately 232 monomers. Each monomeric unit contains one nitrogen

atom which corresponds to a total of 232 nitrogens per polymer chain, distributed appropriately

between classifications as shown in Table 5.

Table 5. Amine distribution in 10,000 MW PEI

Class of Amine	% of Total Amines	# of Amines
1°	25	58
2°	50	116
3°	25	58

3.2. Cross-linking

Cross-linking is a process by which polymer chains are linked together through covalent or ionic bonds. This mechanism decreases molecular freedom, forming a network polymer with an increased molecular weight. Cross-linking can be accomplished through numerous processes, all of which can be classified as either (1) cross-linking during the polymerization process by employing polyfunctional instead of bifunctional monomers or (2) cross-linking in a subsequent post-polymerization process using preformed polymers. Typically, the former process contains cross-links that possess the same structural features as the main chain, while the latter process may introduce entirely new structural elements. The cross-linking is usually accompanied by changes to the properties of the polymeric material, including altered solubility and flow properties. These characteristics correlate to the cross-link density, Γ, which is the number of cross-linked monomer units per polymer chain, as described in Equation 9, where $(\bar{M}_n)_0$ is the number average MW of uncross-linked polymer and $(\bar{M}_n)_c$ is the number average MW between cross-links.

$$\Gamma = \frac{(\bar{M}_n)_0}{(\bar{M}_n)_c} \qquad [9]$$

The higher the cross-link density the more rigid the material becomes, with very highly cross-linking densities leading to embrittlement. On the other hand, elastomers have a very low cross-link density which permits flexibility of the main polymer chains to allow for deformation.

3.2.1. Carbodiimide-mediated coupling using EDC and Sulfo-NHS

EDC is a water soluble carbodiimide, zero-length cross-linker, commonly used to couple carboxylates to primary amines. EDC reacts with carboxylic acids to form a highly reactive *O*-acylisourea

intermediate. This activated species can then react with a nucleophile, such as a primary amine to form an amide bond.[161] The *O*-acylisourea intermediate is susceptible to hydrolysis in aqueous solutions; therefore, activated groups that do not react rapidly with amines can be hydrolyzed back to carboxylate groups accompanied with the release of isourea. In order to prevent this competing reaction from occurring, *N*-hydroxysulfosuccinimide (sulfo-NHS) is often included in aqueous EDC coupling protocols.[162] Sulfo-NHS is a water soluble compound that improves the coupling efficiency by forming a more stable intermediate that is significantly less susceptible towards hydrolysis. This stable intermediate is formed from the reaction of the hydroxyl group on the sulfo-NHS with the *O*-acylisourea species resulting with a sulfo-NHS ester intermediate, which is yet another activated carboxylate species. This species reacts with primary amines to give the amide linkages and regenerates the sulfo-NHS. The sulfo-NHS ester intermediate has an extended half-life measured in hours instead of a hydrolysis rate constant measured in seconds for the EDC activated ester.[163] Furthermore, another advantage of adding sulfo-NHS is to maintain or increase the water solubility of the compound bearing the activated functional groups. EDC and EDC/sulfo-NHS cross-linking reaction schemes illustrating the coupling of carboxylates to primary amines are provided in Scheme 14. It should be noted that the isourea byproducts and the regenerated sulfo-NHS are not shown.
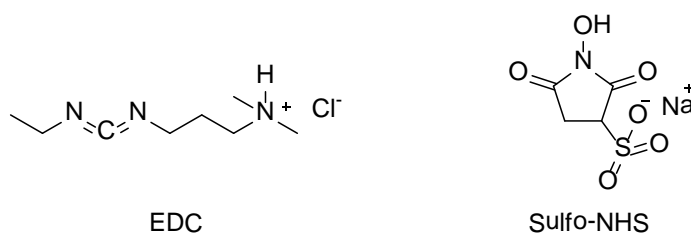
EDC   Sulfo-NHS

Figure 65. Structures of EDC (*N*-(3-Dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride) and Sulfo-NHS (*N*-hydroxysulfosuccinimide sodium salt)
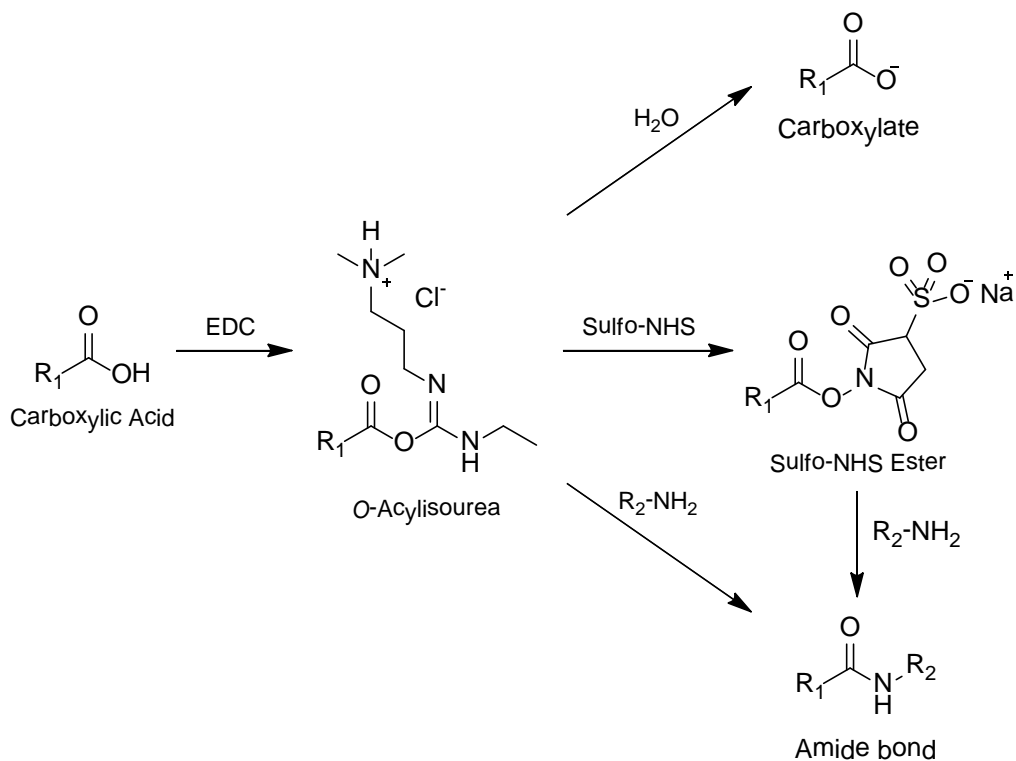
Scheme 14. EDC and EDC/Sulfo-NHS coupling of carboxylic acids to amines resulting with amide bond formation. The hydrolysis of the *O*-acylisourea reactive intermediate is a competing reaction when using EDC alone, as shown in the top-most pathway.

3.3. Cross-linking of PEI with H$_2$mTCAPyP4 using EDC/Sulfo-NHS

The human telomeric repeat (TTAGG)$_4$ has a MW of 7575. Each H$_2$mTCAPyP4 cross-linker contains four carboxylic acids moieties positioned directly above each of the grooves, assuming an end-stacking binding mode as shown in Figure 66. This provides the most π-stacking overlap between the porphyrin macrocycle and the terminal G-quartet. Taking these aspects into account, we rationalized that a 10,000 MW PEI would be the most appropriate polymer size for interaction with each of the four G4 grooves. Smaller PEI chains would not be sufficient for fully interacting with the entire groove, while larger PEI chains introduce an excess of cationic charges that result with a more profound condensation effect on the G4 structure.
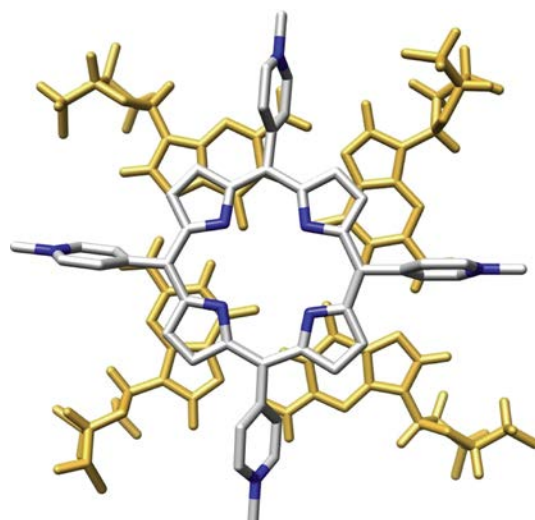
Figure 66. Interaction of cationic H_2mTMPyP4 and a G-quartet; the *meso*-substituents are located directly above the grooves.[93]

In theory, each H_2mTCAPyP4 cross-linker can intramolecularly cross-link a maximum of four PEI chains. The more likely scenario would be a combination of both inter- and intramolecular crosslinking due to the surplus of cross-linkable primary amines on each PEI chain. Initially, we performed the cross-linking reactions in 4:1 and 1:4 PEI to H_2mTCAPyP4 ratios. The former ratio representing all intermolecular cross-linking while the later would provide four porphyrin cross-linkers for each PEI chain, which would promote a greater possibility for intramolecular cross-linking. The cross-linking was accomplished using modified protocols from those found in the literature. It should be noted that while it has been shown that EDTA interferes with cross-linking using EDC alone and not with EDC/Sulfo-NHS, we chose to perform the reactions in EDTA free KPBS.[164] Furthermore, the concentration of coupling reagents used for the activation of the carboxylic acids was performed in respect to functionality i.e. the molar ratio is four times the functionality ratio.

The cross-linking was performed by dissolving 50 mg of 10,000 MW PEI in 1 mL of pH 7.2 EDTA free KPBS. To this solution 4X and .25X moles of H_2mTCAPyP were added to achieve the 4:1 and 1:4 PEI to cross-linker ratios, respectively. 4X sulfo-NHS with respect to carboxylic acid functionality, was added and the solution was shaken vigorously for 15 min. Then, a 10X EDC solution was added. The mixture

was vortexed intermittently for 2 hours and allowed to react further for 24 hours to ensure completion. The solution was then dialyzed using a 1000 MWCO dialysis membrane against two changes of 1 L ddH$_2$O in 48 hours. Dialyzed solutions were then collected and concentrated by lyophilization.

3.3.1. Results and Discussion

Dialysis of the reaction mixture was initially intended for removal of isourea byproducts and unreacted coupling reagents. Interestingly, it was visibly evident that the dialysate of the 1:4 reaction contained porphyrin cross-linker. Its presence indicates PEIs inability to encompass all of the cross-linking molecules at this ratio, resulting in the excess unincorporated cross-linker diffusing across the dialysis membrane. This implies that, on average, each polymer chain coupled to less than 4 cross-linkers, despite the fact that there are approximately 58 primary amines available for coupling per polymer chain. Furthermore, prior to lyophilization, mild centrifugation of the collected 1:4 dialyzed reaction mixture revealed the presence of a precipitate. This is undoubtedly a direct result of PEI cross-linking and an indication of the preparation of a highly cross-linked insoluble material. In regards to the 4:1 reaction, the presence of H$_2$mTCAPyP4 cross-linker was undetectable in the dialysate, visibly or by UV-Vis spectroscopy. Also, the dialyzed solution did not contain any precipitate. Figure 67 shows photos comparing the results of the dialysis and centrifugation for both reaction mixtures. It should be noted that while H$_2$mTCAPyP4 has an MW of 1079.29 (solvated state), which is greater than the 1000 MWCO of the dialysis tube, MWCO is determined as the solute size that is retained by at least 90%. Furthermore, the permeability of a solute is dependent on molecular shape, degree of hydration, polarity and ionic charge. It is typically recommended for the MWCO used to have half the size of the MW of the species to be retained.[165] Therefore, it is not surprising to find the unincorporated cross-linker diffusing out of the dialysis membrane.
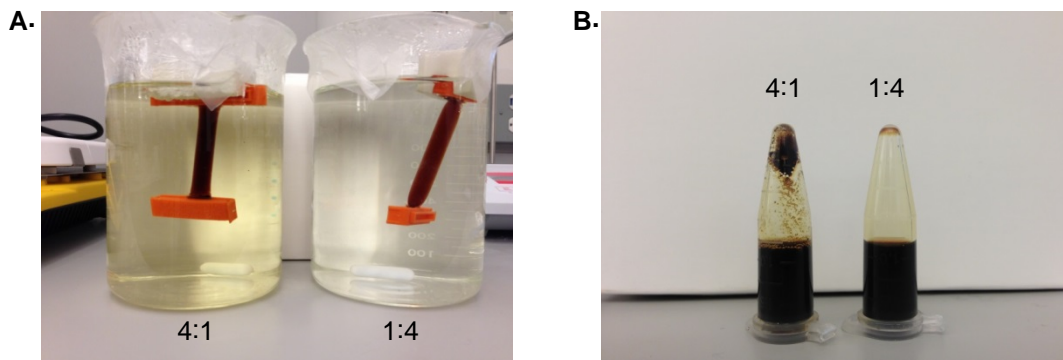
Figure 67. (A) Dialysis of the coupling reaction mixtures: The 4:1 reaction (excess porphyrin) demonstrates the presence of the porphyrin cross-linker in the dialysate (B) Centrifuged dialyzed solutions: The 4:1 reaction affords an insoluble highly cross-linked polymer precipitate

The lack of detectable porphyrin cross-linker in the 1:4 reaction dialysate suggests that all of the cross-linkers were successfully coupled to the PEI. However, unlike the 4:1 excess porphyrin reaction, the cross-linking was not so extensive as to cause the insolubility of the cross-linked polymer. ESI-MS analysis was required to confirm the successful preparation of this soluble cationic porphyrin cross-linked PEI. Small amounts of the lyophilized products were reconstituted in methanol, filtered through 22 μm PTFE filters and submitted for analysis. The resulting spectra obtained are shown in Figures 68.
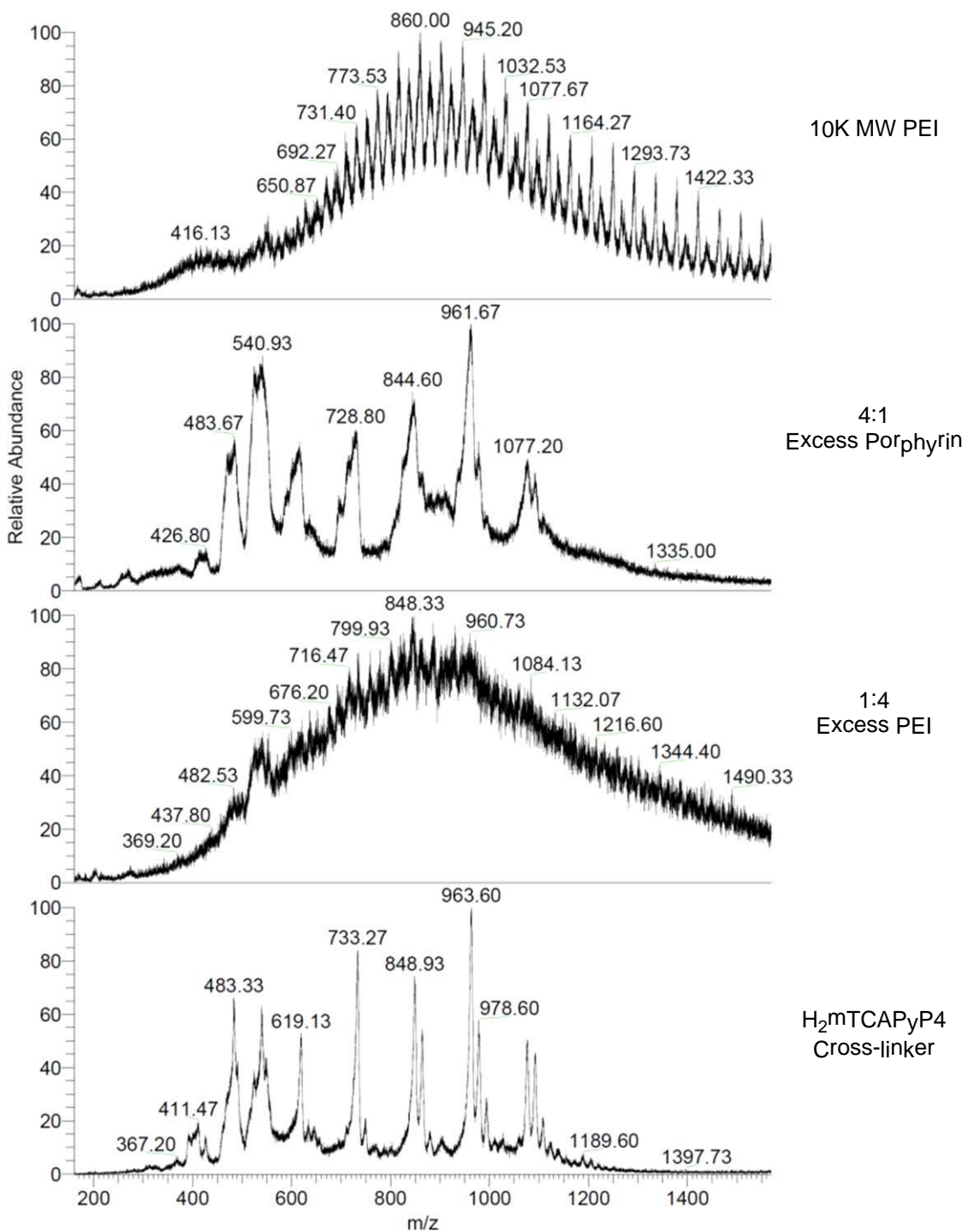
Figure 68. ESI-MS spectra of PEI, $H_2mTCAPyP4$ cross-linker, cross-linked 1:4 and 4:1 PEI to $H_2mTCAPyP4$ polymer solutions

The number of charges on a polymer in ESI depends on the MW and the potential charge sites, including those for protonation. While low MW polymers exhibit only a singly charged ion species, polymers with high MW can form multiply charged species. Therefore, the spectra of these high MW polymers contain superimposed peaks corresponding to the various species present, resulting with highly convoluted spectra. Nevertheless, ESI-MS analysis proved to be insightful. The spectrum of 10,000 MW branched PEI exhibited an average mass difference of 43 between individual polymer molecules, which corresponds to the MW of an ethylenimine monomer. The spectra of the soluble 1:4 porphyrin cross-linked PEI has mass differences between molecules that are significantly greater than 43 and the differences are inconsistent. This indicates that the polymer molecules differ by more than just an ethylenimine monomeric unit i.e. the cross-linker has been coupled to the polymer and incorporated into the mass. The inconsistency in mass differences from one molecule to the next is expected considering that each polymer chain is not necessarily cross-linked to the same extent, with some polymers being coupled to more cross-linkers than others in addition to the various combinations of inter- and intramolecular cross-linking possibilities. Furthermore, examination of the 4:1 PEI spectrum indicates that no PEI was present in the supernatant. The spectrum for this excess porphyrin reaction resembles the spectrum collected for the free H_2mTCAPyP4 cross-linker. This data suggests that the precipitate formed during this reaction, resulting with the insoluble highly cross-linking PEI, reacted completely with all of the PEI, dropping it out of solution and leaving behind only excess cross-linker.

NMR analysis was used to further confirm the successful cross-linking of PEI. The collected ^{1}H NMR spectra of both PEI and cross-linked PEI from the 1:4 reaction are shown in Figures 69 and 70, respectively. NMR bandwidths are sensitive to cross-linking. Analyses of cross-linked polymers usually yield a characteristic peak broadening effect as a direct result of the cross-linking polymer having restricted motion in solution at the segmental level.[166] Examination of the collected spectra clearly exhibited this peak broadening effect confirming success of our cross-linking reaction.
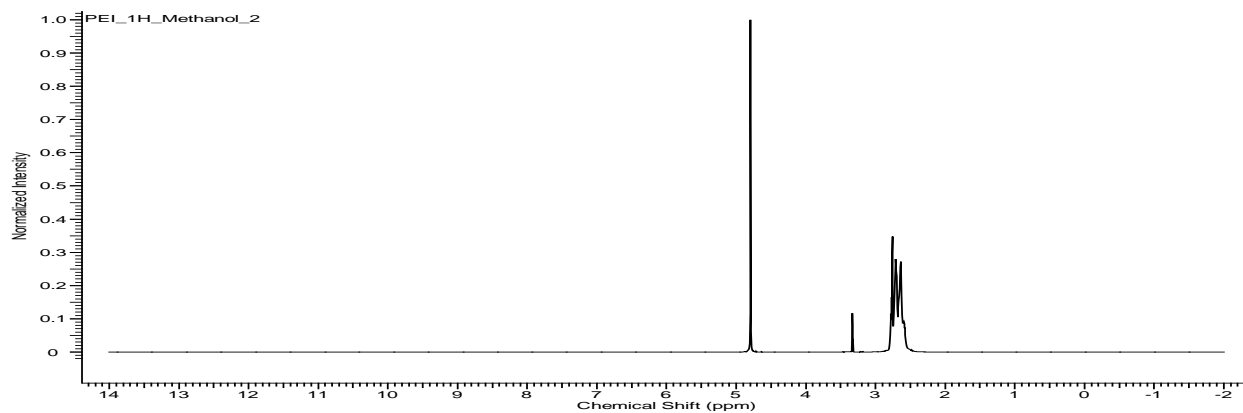
Figure 69. ^{1}H NMR of 10,000 MW branched PEI
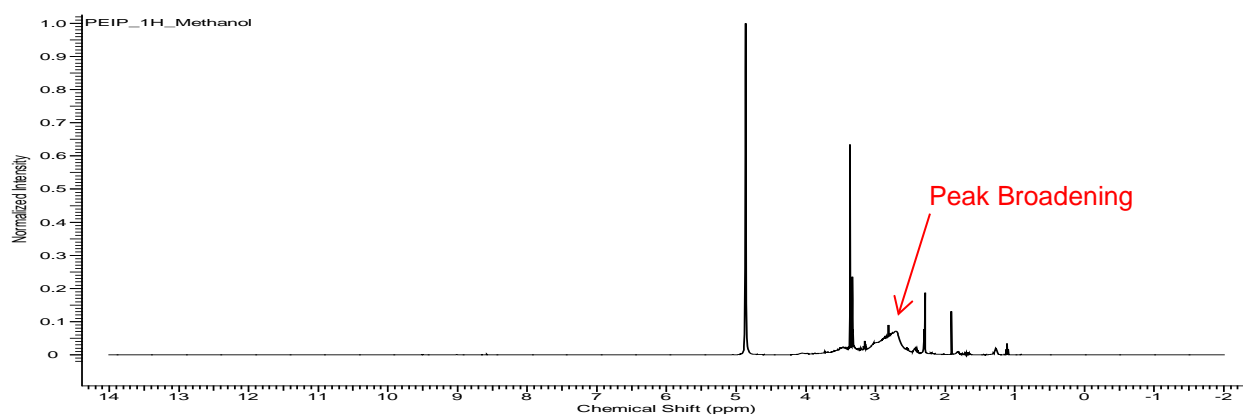
Figure 70. ^{1}H NMR of H$_2$mTPyP4 cross-linked PEI

Results from these initial coupling reactions were very pleasing. The EDC/sulfo-NHS coupling strategy allowed us to successfully react a cationic porphyrin at its quaternizing groups, a feat that was previously elusive. Furthermore, it allowed for the cross-linking of PEI with the tetracarboxylic acid functionalized cationic porphyrin, H$_2$mTCAPyP4. The cross-linking reactions also demonstrated that at lower cross-linking densities the cross-linked polymer maintained its water solubility, which is an essential property for the material to possess in hopes of generating a soluble MIP. A schematic representation of the cross-linked PEI polymer is shown in Figure 71.
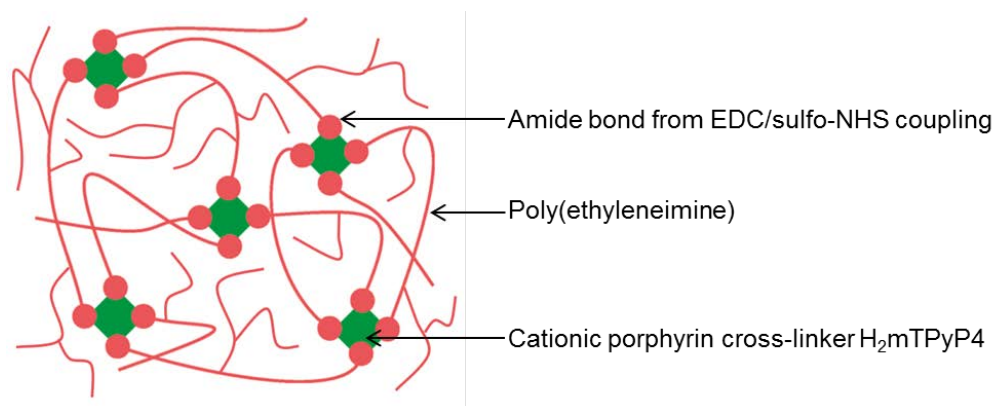
Figure 71. Schematic representation of H₂mTPyP4 cross-linked PEI. The H₂mTPyP4 cross-linker (green circles) forms amide bonds (red circles) via EDC/sulfo-NHS coupling with primary amines of the PEI (red lines).

We proceeded by experimenting with varying the degree of cross-linking using the 1:4 porphyrin H₂mTCAPyP4 to PEI as our benchmark. We gradually decreased the cross-linking density from 3.5% by mass to 0.35%. Table 6 provides information on the reagent stoichiometries used in the four cross-linking reactions we performed.

Table 6. Reaction stoichiometries varying the degree of cross-linking density

RXN	mTCAPyP: PEI (mol ratio)	PEI (μM)	mTCAPyP (μM)	Carboxcylic Acid (μM)	EDC 10X (μM)	NHS 4x (μM)	mTCAPyP (mg)	Cross-Linking by Mass (%)
1	1:40	5	0.125	0.50	5.0	2.0	0.17	0.350
2	1:10	5	0.500	2.00	20.0	8.0	0.70	1.40
3	1:7	5	0.714	2.86	28.6	11.4	1.00	2.00
4	1:4	5	1.25	5.00	50.0	20.0	1.75	3.50

In these cross-linking reactions the concentration of PEI was kept constant, while the concentration of the H₂mTPyP4 and coupling reagents were varied accordingly. As anticipated, the cross-linking performed at these densities all provided soluble products, forming no precipitates. Furthermore, no porphyrin was detectable by UV in the dialysate.

The mobilities of the resulting cross-linked PEI products were examined on a .5% agarose gel. Increasing the cross-linking density is expected to promote more intermolecular cross-linking and thus increasing the overall MW of the product. As a result higher cross-linking density results in decreased

99

mobility. Due to the overall cationic charge of the cross-linked PEI, electrophoresis was performed with the wells being placed at the positive electrode in order to facilitate the polymers migration towards the negative electrode. A constant voltage of 90 mA was applied to the gel, which was visualized on a UV transilluminator. This resulted in the following electrophoretic mobilities shown in Figure 72. Lane 1 contains PEI and lanes 2, 3, 4, and 5 contains .35% (1:40), 1.4% (1:10), 2.0% (1:7) and 3.5% (1:4) cross-linked PEI, respectively. The gel did not require any staining since the porphyrin cross-linker is UV active. The PEI alone in lane 1 produced no detectable bands. However, the lanes containing cross-linked PEIs were found to each produce a smear. This should be expected considering the polydispersity of the original polymer and the various combinations of inter- and intramolecular cross-links possible. The gel clearly demonstrated that the cross-linking density has a direct effect on the polymers mobility with the smears in each lane decreasing in length as a function of increasing cross-linking.
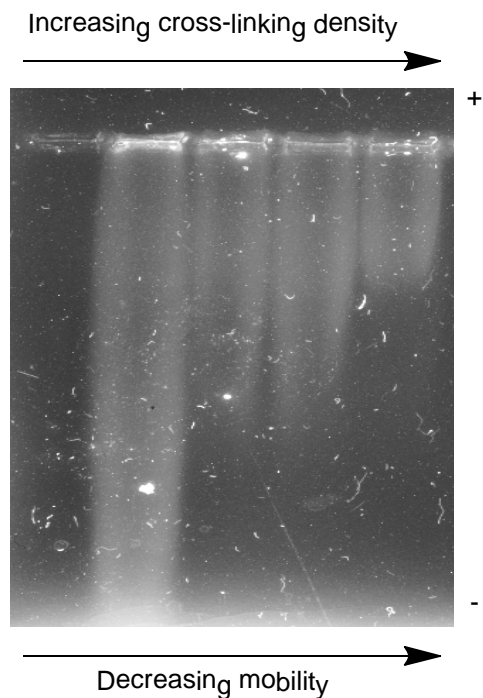
Figure 72. 0.5% agarose gel showing the different mobilities of cross-linked PEI resulting from varying the cross-linking density.

3.4. Summary

Screening of commercially available polymer for their affinity towards nucleic acids indicated that electrostatic interactions with cationic polymers are the most achievable way to elicit binding events. The three polymers that were found to induce DNA hyperchromicity were poly[bis(2-chloroethyl) ether-alt-1,3-bis[3-(dimethylamino)propyl]urea], poly(ethylenimine) and poly(diallydimethylammonium chloride), two of which carry cationic charges by previous *N*-quaternization. The cationic charge of PEI results from the protonation of its primary and secondary amines in aqueous solutions. The presence of primary amines in the PEI structure provides functionality that allows for coupling. Furthermore, PEI resulted in the greatest net change to DNA UV absorbance, indicating the highest degree of interaction. Carbodiimide chemistry using EDC/sulfo-NHS coupling reagents allowed for the coupling of our previously prepared cationic porphyrin H_2mTCAPyP4. The carboxylic acid moieties were activated using these coupling reagents to from activated NHS-esters that allowed for nucleophilic attack by the primary amines of PEI, forming amide bonds. Since each cationic porphyrin contains four carboxylic acids where coupling can occur, this reaction provides a method to achieve the cross-linking of the PEI polymer by forming four amide bond cross-links intra- or intermolecularly. Highly cross-linked PEI resulted with an insoluble polymer while PEI cross-linked at a 1:4 H_2mTCAPyP4 to PEI produced a soluble product. The cross-linking was confirmed by ESI-MS which exhibited changes to the mass difference between polymer molecules that did not correlate to the ethyleinimine monomeric molecular weight. Furthermore, ^{1}H NMR analysis of the cross-linked PEI exhibited characteristic peak broadening, signifying the restricted motion of the material in solution prompted by the cross-linking. Variations to the cross-linking density also produced soluble cross-linked PEIs that demonstrated different electrophoretic mobilities on agarose gels, correlating to the increase in effective size with increase in cross-link density. The complete pathway to achieving the H_2mTCAPyP4y cross-linked PEI is shown in Scheme 15. An example of intermolecularly cross-linked PEI is shown in the ball and stick model presented in Figure 73.
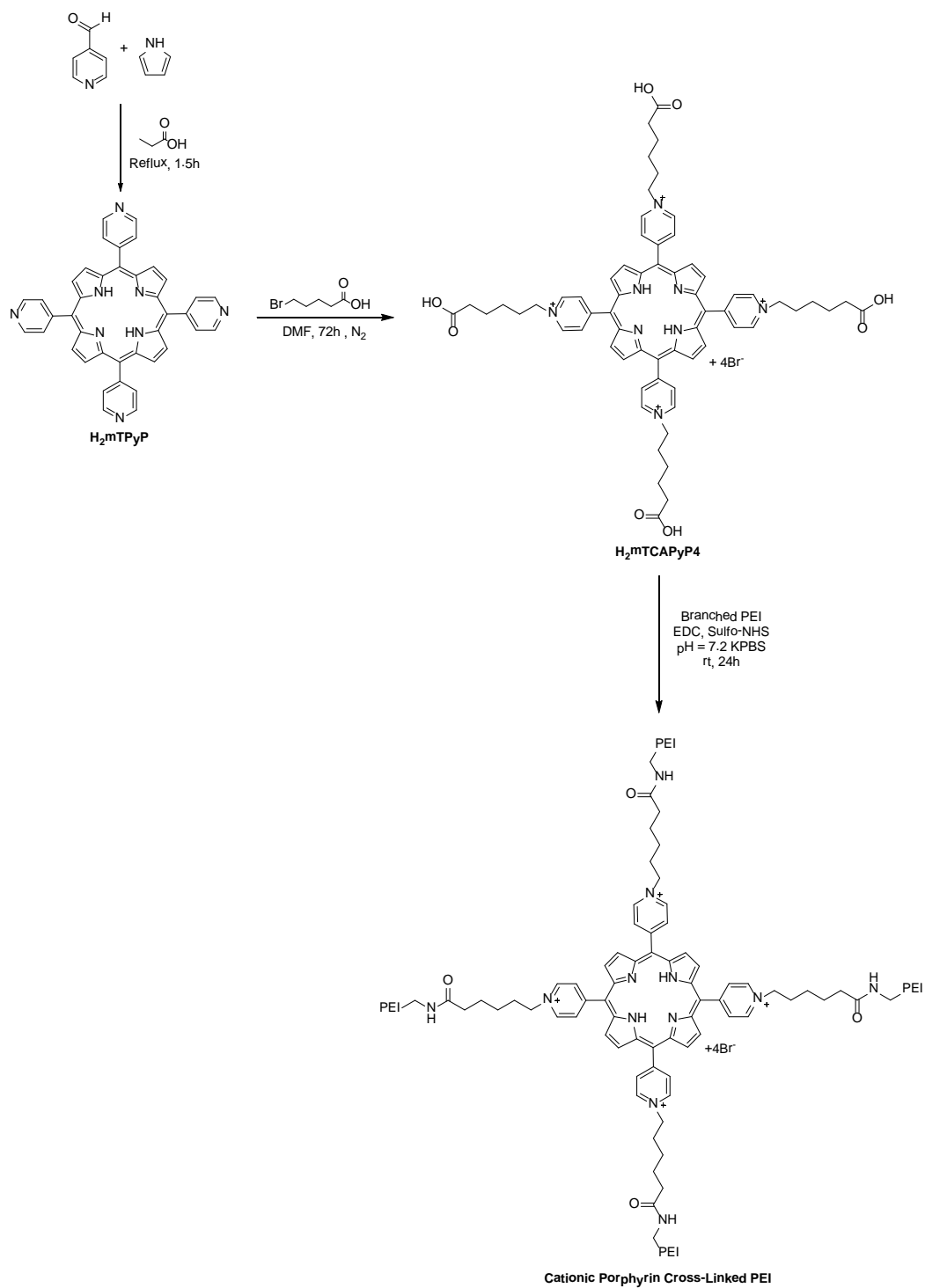
Scheme 15. Pathway to achieving cationic porphyrin cross-linked PEI
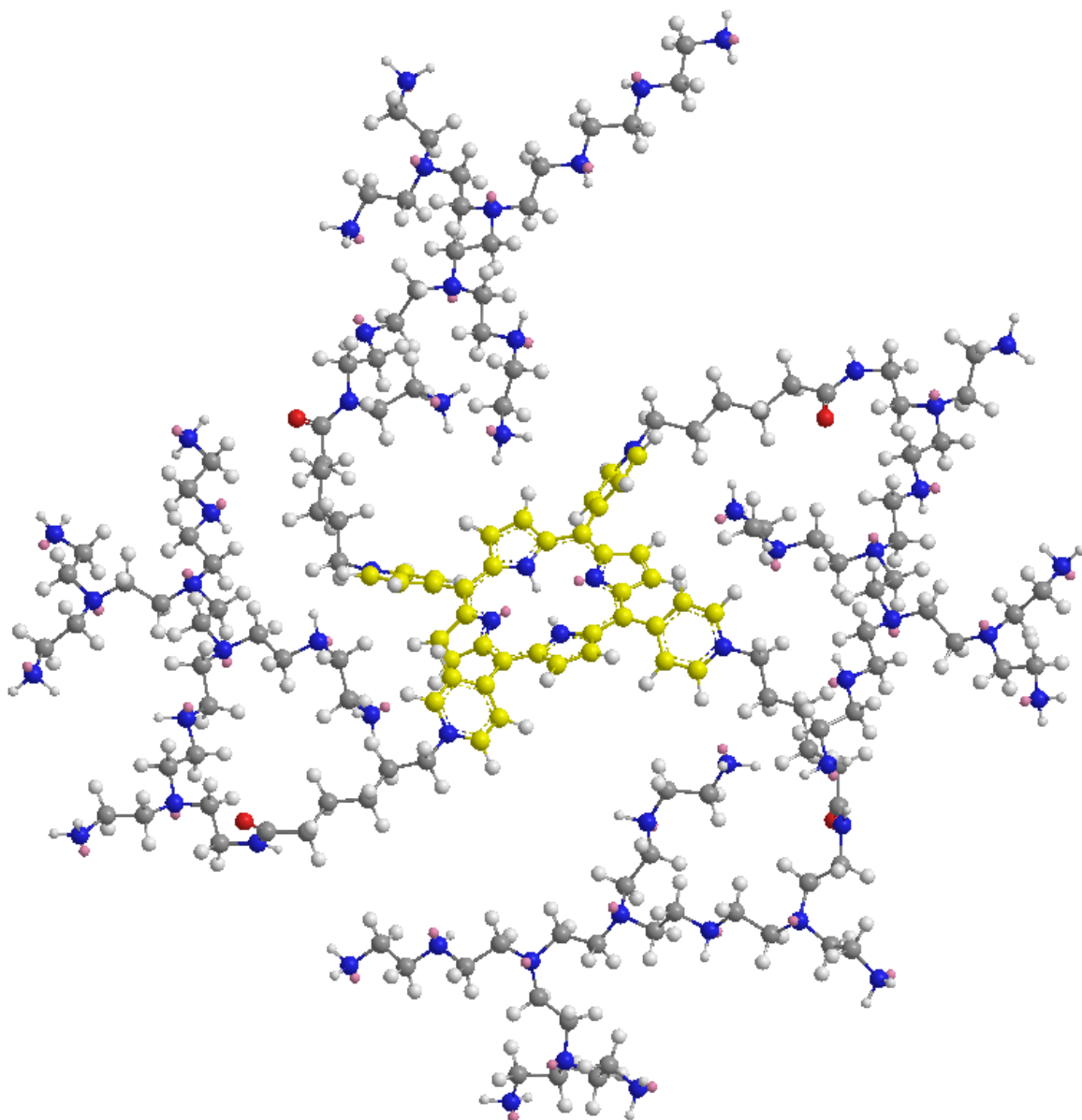
Figure 73. Ball and stick model illustrating H$_2$mTCAPyP4 intermolecularly cross-linked PEI. Porphyrin carbons are shown in yellow, the cross-linking amide bond oxygens in red, nitrogen atoms in blue and lone pair electrons in pink.

4. DNA Binding Studies

We investigated the DNA-binding properties of our prepared water-soluble cationic porphyrins and cross-linked PEI-porphyrin conjugates by employing UV-Vis and circular dichroism (CD) spectroscopy. These techniques are commonly used in thermodynamic studies and for determining the binding modes in porphyrin-nucleic acid interactions. We employed these methods to determine whether or not our prepared cationic porphyrins, primarily our porphyrin cross-linker, have affinities toward G4 structures. Specifically, we studied the interactions of $H_2mTAlPyP4$, $H_2mTOEtPyP4$, $H_2mTOPrPyP4$, $H_2mTCAPyP4$ and the $H_2mTCAPyP4$ cross-linked PEI with the human telomeric repeat G4 forming oligonucleotide $(TTAGGG)_4$.

4.1. Materials and Methods

<u>Preparation of Quadruplex $(TTAGGG)_4$</u>

The HPLC purified unquadruplexed 5'-$(TTAGGG)_4$ 24mer was purchased from Sigma and used without any further purification. DNA was quadruplexed by reconstituting the dry oligonucleotide in 150 mM KPBS buffer (1.001g K_2HPO_4, .5716g KH_2PO_4, .0405g K_2EDTA in 1L ddH_2O). The DNA was vortexed, to ensure complete dissolution, and was then heated to 95 °C where it was held at this temperature for 5 minutes. The heating block was then turned off and the sample was allowed to equilibrate to room temperature, a process that takes approximately 2.5 hours. The sample was then incubated at 5 °C for 48 hours. Verification of successful quadruplex formation was performed by CD spectroscopy.

DNA concentrations were determined by UV spectroscopy on a Cary 3E spectrophotometer. An extinction coefficient of 244,6000 M^{-1} cm^{-1} was used to quantify DNA concentrations using optical density measurements at 260 nm. This ε value was determined using Equation 10 which utilizes the extinction coefficients of the individual 5'-monophosphates at 260 nm and 25 °C. This formula equates the extinction coefficient of the strand to the sum of the products of the extinction coefficients of single

5'-monophosphate by the number of bases found in the DNA strand, where *N* is the number of bases of the specific base denoted in the subscript.

$$\varepsilon_{260}(\boldsymbol{strand}) = (\mathbf{15400} \times N_A) + (\mathbf{7400} \times N_C) + (\mathbf{11500} \times N_G) + (\mathbf{8700} \times N_T) \qquad [10]$$

The concentration of the stock quadruplex solution was determined at 95 °C using a 1:100 dilution. This elevated temperature causes the denaturation of the quadruplex structure resulting with a hyperchromic effect that provides an accurate optical density measurement suitable for DNA quantification. The stock solution was prepared at a concentration of 8.81 x 10^{-4} M and stored at 4 °C. The DNA was allowed to equilibrate to room temperature prior to use.

Porphyrin Solutions

The extinction coefficients used for porphyrin quantification were determined by producing Beer's Law plots. 1 mg of porphyrin was weighed out and dissolving into 100 mL of ddH$_2$O. A series of dilutions provided the linear relationship between absorbance and concentration. The value of the resulting slopes allowed for the determination of the molar absorptivity using the Beer-Lambert equation. The measurements were performed on an HP 8452A Diode Array spectrophotometer using 1 cm disposable cuvettes, as to prevent any residual porphyrin from the previous dilution interfering with the subsequent absorbance measurements. Data was collected via scan mode from 250 to 800 nm, averaging 10 repeated scans, using Olis Spectral Works version 4.4. After an absorbance measurement was collected on the original solution, 800 µL were pipetted with 200 µL of ddH$_2$O into a new cuvette and a subsequent scan was taken; this process was repeated until an absorbance reading of approximately .01 was obtained. Collected data was exported to Origin 9.0 for plotting and analysis. Accordingly, 100 mL of porphyrin stock solutions were prepared at 10 OD in KPBS, and were stored shielded from light at room temperature.

<u>UV-Vis Titrations</u>

Titrations were performed on an HP 8452A Diode Array spectrophotometer using 1 cm pathlength quartz cuvettes. Prior to each use, cuvettes were rinsed thoroughly using dilute nitric acid, methanol and then ddH$_2$O. All absorption titration experiments were carried out by the stepwise addition of aliquots of titrant solution to a cuvette containing 1 mL of a porphyrin solution with an OD of 1 (100 μL stock porphyrin solution in 900 μL KPBS) or .5 (50 μL stock porphyrin solution in 950 μL). For titrations with quadruplexed (TTAGGG)$_4$, a 200 μL titrant solution was prepared in 150 mM KPBS by adding 20 μL of porphyrin stock solution to 8X DNA concentration, determined accordingly from the initial absorbance measurement of the porphyrin solution to be titrated. Addition of 2.5 μL of titrant was followed by 30 seconds of vortexing and 5 minutes of incubation time prior to each measurement being recorded. Each spectrum is an average of 5 scans collected between 200-800 nm using Olis Spectral Works 4.4. Data was subsequently exported for analysis and plotting using Origin 9.0 software. Titration of H$_2$mTAlPyP4 with CT-DNA was performed by adding 20 μL aliquots of 33 mg/mL CT-DNA to a 5 μM solution of H$_2$mTAlPyP4.

<u>CD Spectroscopy</u>

CD spectra for evaluation of conformations of quadruplex structure were recorded at 25 °C using an Aviv Associates CD Spectropolarimeter. After 5 min equilibration at the sample temperature, the instrument collected spectral data in the 220-320 nm range every 0.5 nm, with an averaging time of 3 sec and bandwidth of 1. The data was exported to Origin 6.1 software for buffer blank subtraction and 3 point adjacent averaging filter smoothing. Induced CD spectra were performed at the Keck Biophysics Facility at Northwestern University on a Jasco J-815 Spectrometer. Spectra were collected in the range of 200-600 nm with 1 nm data intervals, integration time of 1 sec and a bandwidth of 2. The porphyrin-DNA titrations were performed by adding eight 12.5 μL aliquots (100 μL total) of the titrant solution (~2 x 10^{-6} M porphyrin and ~7.7 x 10^{-5} DNA) to 1 mL of a 0.5 OD porphyrin solution (~2 x 10^{-6} M) in a 5 mm

cuvette with 20 minute incubation times between additions. The experiment was prepared so that a DNA to porphyrin ratio of 4:1 was achieved after the addition of 100 µL of titrant solution. In a control experiment, a DNA only sample was scanned; this solution was prepared by diluting the stock DNA (8.81 x 10^{-4} M) such that the final DNA concentration in this sample matches the DNA concentration in the end of the titration point. All data was corrected for buffer absorbance and smoothed using the Savitzky-Golay filter, convolution width 13, using Origin 9.0 software.

4.2. Results and Discussion

The preparation of quadruplexed $(TTAGG)_4$ was confirmed by CD spectroscopy at 25°C in 150 mM KPBS buffer. The collected spectrum displayed in Figure 74 shows a peak at 287 nm, a shoulder at 273 nm and a trough at 236 nm. This CD spectrum is consistent with intramolecular quadruplex structures reported in the literature.[167]
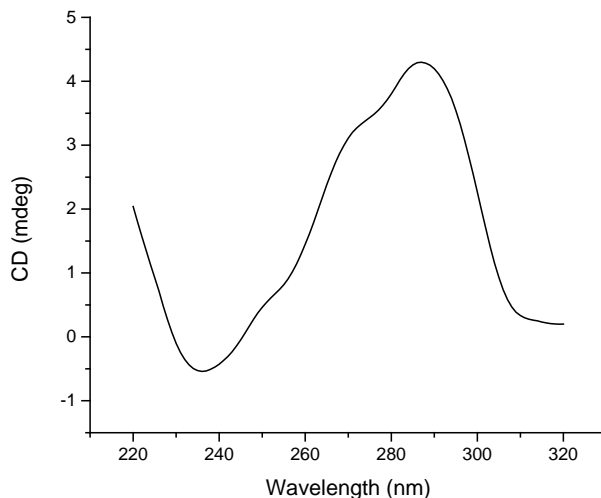
Figure 74. CD spectrum of $[(TTAGGG)_4]$ = 4.6 x 10^{-7} M in 150 mM KPBS at 25°C

Figure 75 provides the UV spectrum collected on the same $(TTAGG)_4$ sample. The quadruplex exhibits a λ_{max} at 257 nm. As with common practice, DNA concentrations were calculated using the measured absorbance at 260 nm.
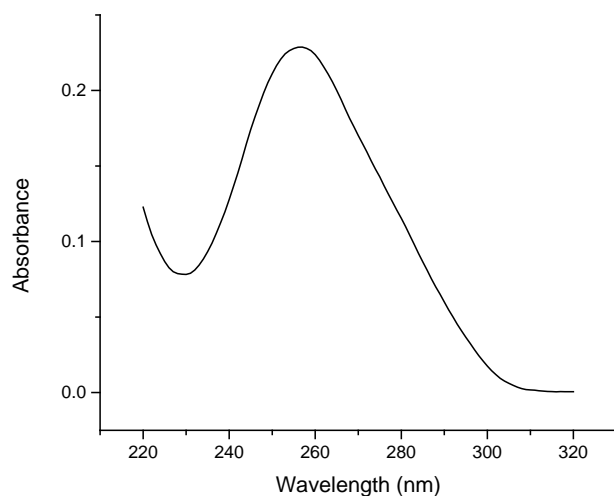
Figure 75. UV absorption spectrum of [(TTAGGG)$_4$] = 4.6 x 10^{-7} M in 150 mM KPBS at 25 °C

The absorption peaks of DNA and porphyrins are distinct, as demonstrated in Figure 76, which provides overlaid spectra of (TTAGGG)$_4$ and mTAlPyP4 from 220 nm to 475 nm.
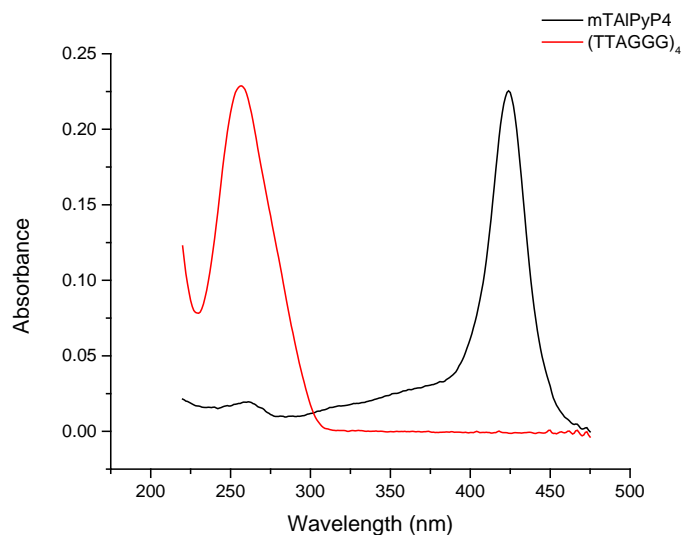
Figure 76. Overlaid absorbance spectra demonstrating the distinct absorption peaks of quadruplexed (TTAGGG)$_4$ and cationic porphyrin H$_2$mTAlPyP4, both at a concentration of 1.9 x 10^{-6} M.

Cationic porphyrin concentrations were calculated using molar extinction coefficients derived from the Beer's law plots shown in Figures 77-80. Table 7 provides the extinction coefficients calculated

using the Beer-Lambert law shown in Equation 11, where A is the measured absorbance, ε is the molar

extinction coefficient, b is the path length and c is the concentration.

$$A = \varepsilon bc \qquad [11]$$
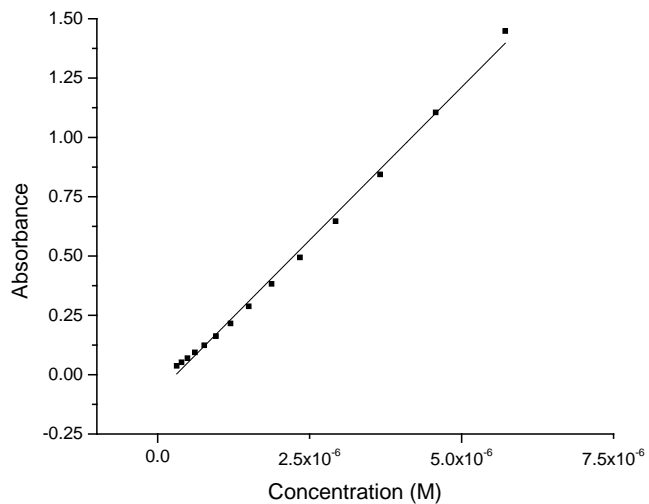
Figure 77. Beer's law plot of $H_2mTAlPyP4$ in ddH_2O
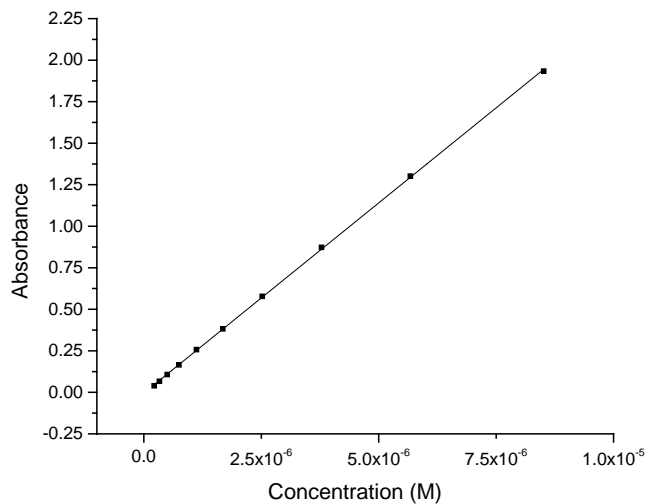
Figure 78. Beer's law plot of $H_2mTOEtPyP4$ in ddH_2O
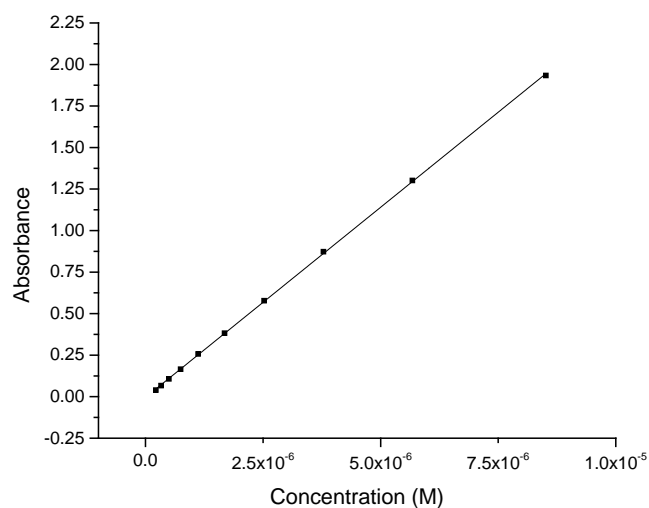
Figure 79. Beer's Law Plot of H2mTOPrPyP4 in ddH$_2$O
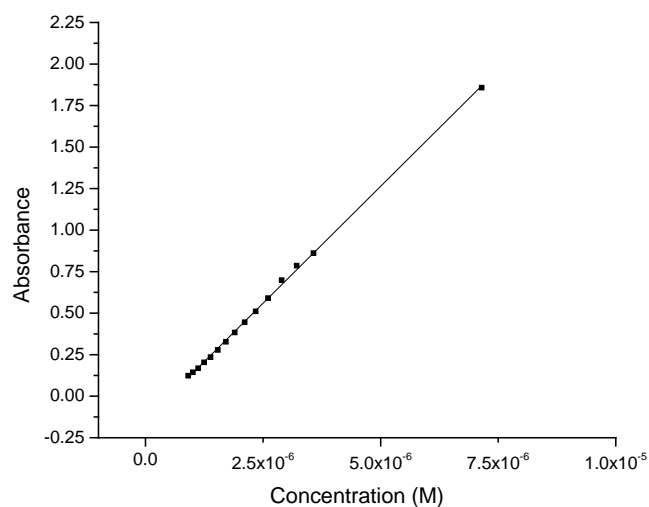
Figure 80. Beer's law plot of H$_2$mTCAPyP4 in ddH$_2$O

Table 7. Extinction coefficients of cationic porphyrins

Porphyrin	ε (M^{-1} cm^{-1})	R^2
H$_2$mTAlPyP4	257,852	.99613
H$_2$mTOEtPyP4	228,906	.99987
H$_2$mTOPrPyP4	228,933	.99987
H$_2$mTCAPyP4	264,430	.99918

The extinction coefficient calculated for mTOBuPyP4 deviated significantly from molar absorptivity

values expected for a porphyrin compound and was therefore excluded from any further studies.

The interactions of these porphyrins with quadruplex DNA were monitored from changes in the Soret region of the absorbance spectra. We recorded UV-Vis absorbance spectra for porphyrins in the presence and absence of $(TTAGGG)_4$ at room temperature by performing the titration of porphyrins with G4 DNA. The titrant solutions were prepared in a manner to provide a 4:1 molar ratio of $(TTAGGG)_4$ to porphyrin at the conclusion of the titration. Furthermore, in order to maintain a constant porphyrin concentration throughout the experiment, an identical concentration of porphyrin was added accordingly to the titrant solution to match the concentration of the solution to be titrated.

The 3D plot resulted from the titration of mTAlPyP4 with $(TTAGGG)_4$ is shown in Figure 77 and demonstrates the changes to the UV-Vis spectra at both the Soret band and DNA absorbance peak at 260 nm upon each aliquot addition. The plot clearly exhibits that with each addition there is an increase in DNA absorbance and a corresponding decrease in the intensity of the Soret maximum.
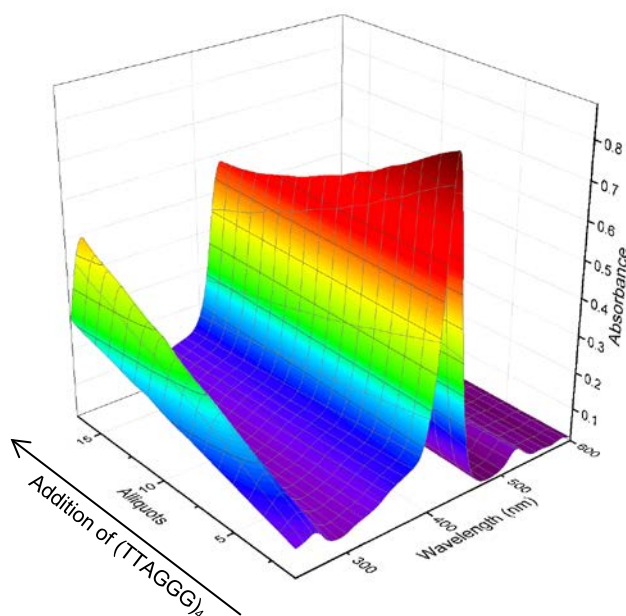
Figure 81. 3D plot of the titration of $H_2mTAlPyP4$ with $(TTAGGG)_4$ in 150 mM KPBS

Closer examination of the data focusing solely on the Soret band, as shown in Figure 82, reveals that with each G4-porphyrin binding event a substantial bathochromic shift is coupled to the

hypochromicity of the Soret λ_{max}. An arrow indicates the decrease in absorbance as a result of increasing G4 DNA concentration. The Soret λ_{max} of free H$_2$mTAlPyP4 found at 424 and the new red shifted λ_{max} of the fully bound H$_2$mTAlPyP4 found at 436 nm are labeled. Furthermore, the presence of an isosbestic point at 435 nm is indicative of a two-state system (free and fully bound). Figure 83 shows these two states and indicates the hypochromicity, ΔA, and change in the Soret λ_{max}, $\Delta\lambda$. The titration was deemed complete when the addition of titrant no longer resulted in bathochromic shift or hyperchromism.
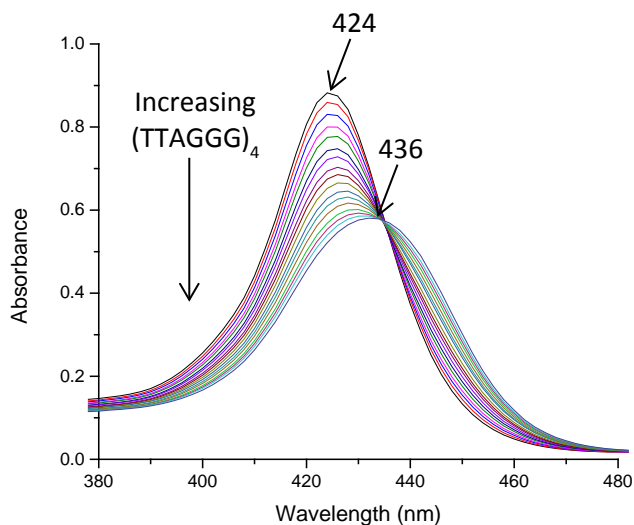
Figure 82. Titration of H$_2$mTAlPyP4 with (TTAGGG)$_4$ in 150 mM KPBS
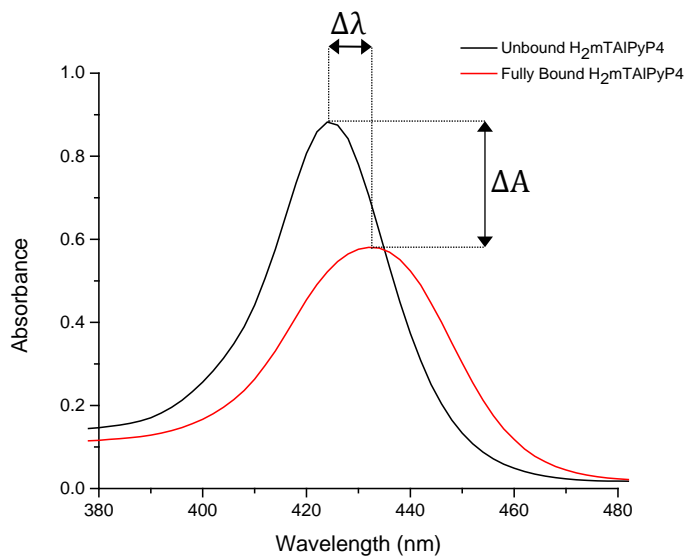
Figure 83. Free and fully bound H$_2$mTAlPyP4 in 150 mM KPBS

The percent hypochromicity is calculated using Equation 12 where ε_F and ε_B are the extinction coefficients of the free and fully bound states, respectively.

$$\%H = \left[\frac{\varepsilon_B - \varepsilon_F}{\varepsilon_F}\right] \times 100 \qquad [12]$$

The shift in the Soret band is calculated using Equation 13 where λ_B and λ_F are the λ_{max} values of the free and fully bound states, using the same subscript notation as the previous equation.

$$\Delta\lambda = \lambda_B - \lambda_F \qquad [13]$$

Results from the titration of mTAlPyP4 with (TTAGGG)$_4$ reveal a 12 nm red shift coupled with a 34% hypochromicity. These changes observed in the visible absorption spectra indicate that our prepared cationic porphyrin is binding to (TTAGGG)$_4$. The lack of specificity exhibited by mTAlPyP4 for G4 structures is demonstrated in Figure 84, which shows comparable changes to the Soret band when binding to non-quadruplex DNA through a titration carried out with CT-DNA. This result is as expected for a cationic porphyrin N-quaternized with a sterically non-hindering group. Similarly, the data indicates a 12 nm red shift, however, a more profound 45% hypochromicity is observed.
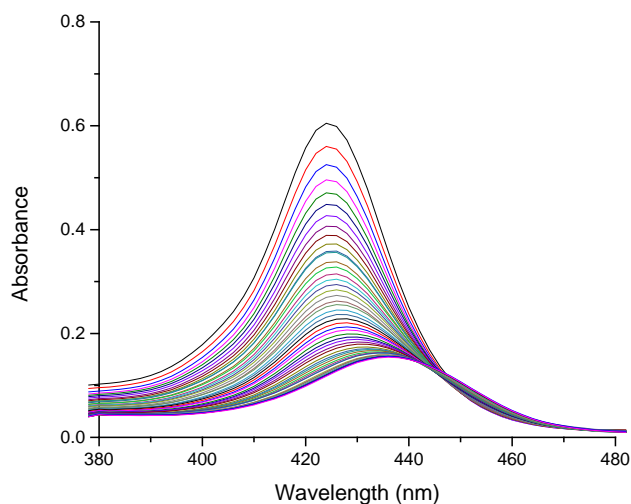
Figure 84. Titration of H$_2$mTAlPyP4 with CT-DNA in 150 mM KPBS

The spectra for the titrations performed using the remaining cationic porphyrins with (TTAGGG)$_4$ are presented in Figures 84-91. The data from these experiments are summarized in Table 8.
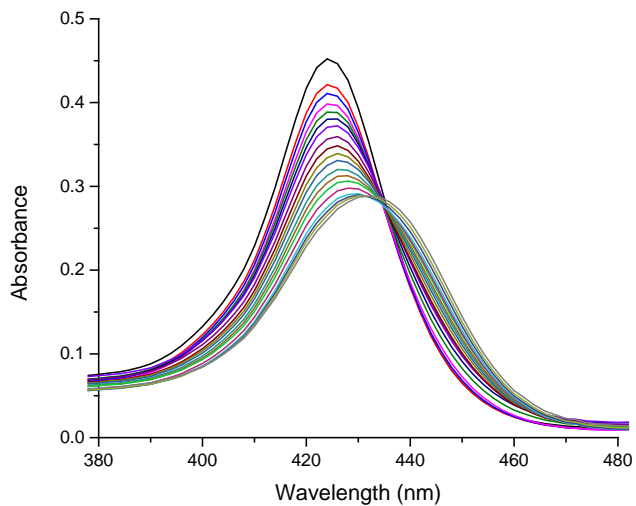
Figure 85. Titration of H$_2$mTOEtPyP4 with (TTAGGG)$_4$ in 150 mM KPBS
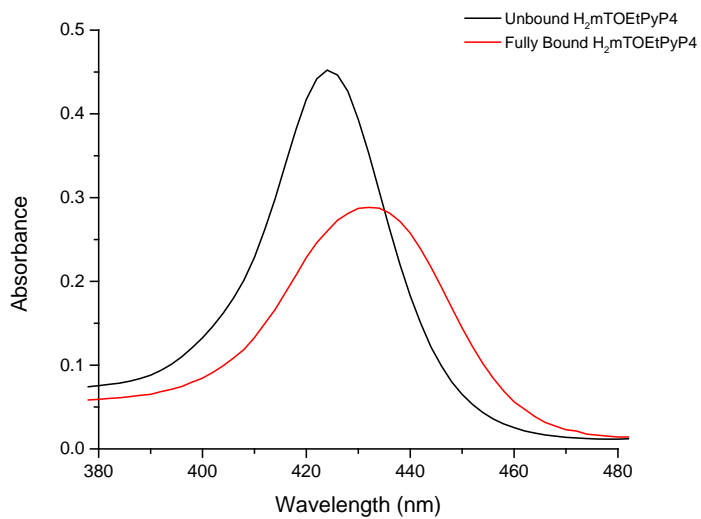
Figure 86. Free and fully bound H$_2$mTOEtPyP4 in 150 mM KPBS
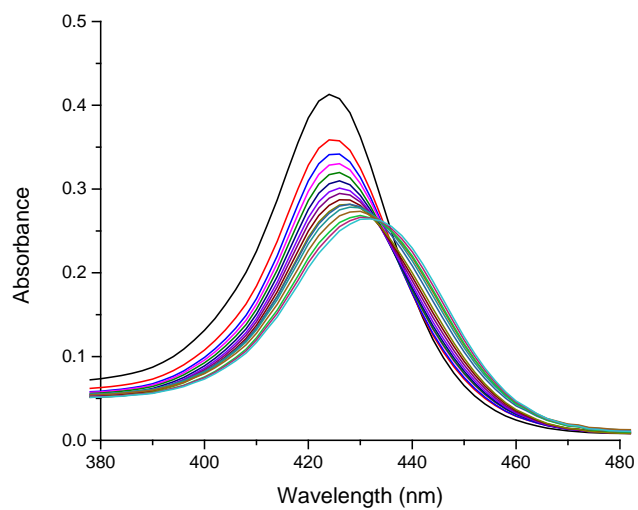
Figure 87. Titration of $H_2mTOPrPyP4$ with $(TTAGGG)_4$ in 150 mM KPBS
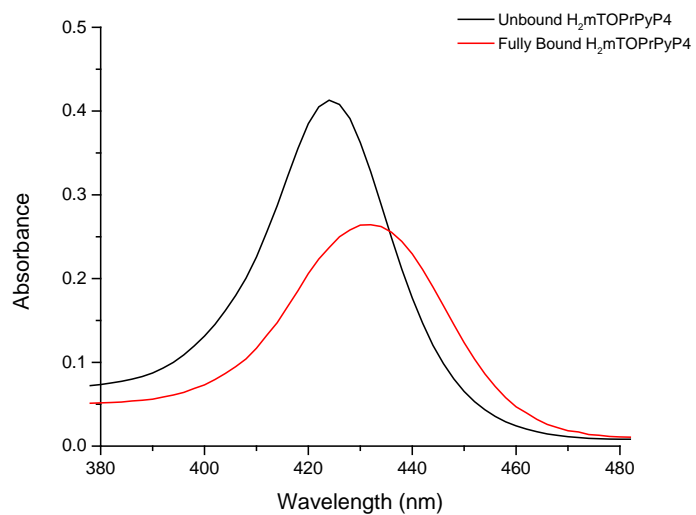
Figure 88. Free and fully bound $H_2mTOPrPyP4$ in 150 mM KPBS
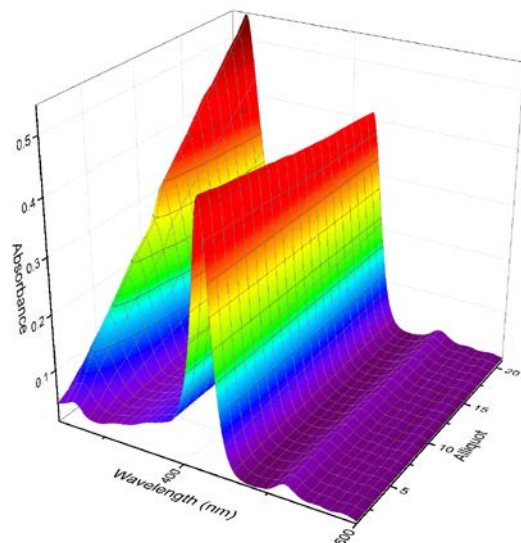
Figure 89. 3D plot of the titration of $H_2mTCAPyP4$ with $(TTAGGG)_4$ in 150 mM KPBS
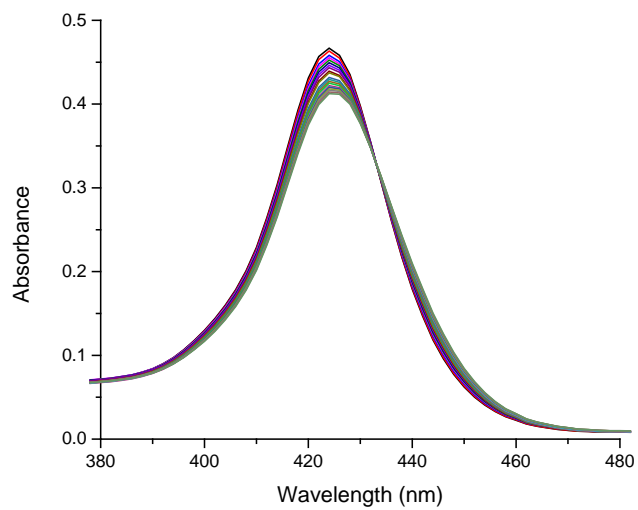
Figure 90. Titration of $H_2mTCAPyP4$ with $(TTAGGG)_4$ in 150 mM KPBS
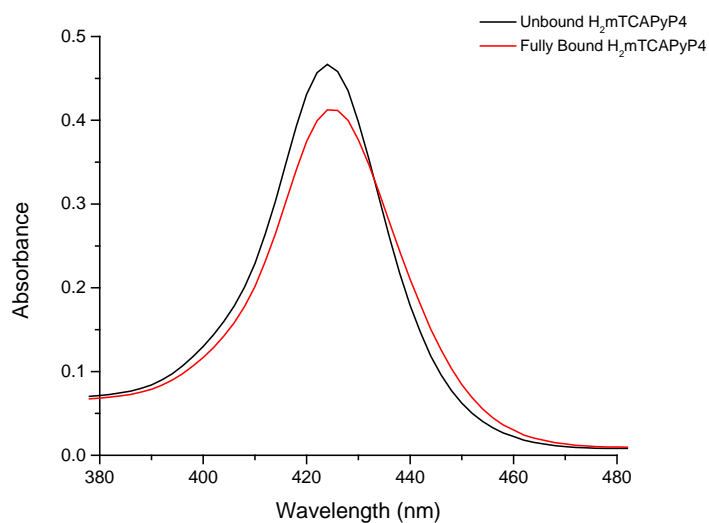
Figure 91. Free and fully bound $H_2mTCAPyP4$ in 150 mM KPBS

Table 8. Porphyrin hypochromicities and bathochromic shifts from titration experiments with G4 DNA

Porphyrin	λ_F	λ_B	$\Delta\lambda$	ε_F	ε_B	%H
$H_2mTAlPyP4$	424	436	12	257,852	161,188	34
$H_2mTOEtPyP4$	424	430	6	228,906	134,319	41
$H_2mTOPrPyP4$	424	432	8	228,933	146,494	36
$H_2mTCAPyP4$	424	424	0	264,430	237,898	10

Data from the UV titration experiments clearly demonstrated both bathochromic shifts and hypochromicities of Soret maximum resulting from the titration of $(TTAGGG)_4$ with the allyl and hydroxy *N*-quaternized porphyrins. $H_2mTCAPyP4$, on the other hand, only resulted with a 10% hypochromicity with no detectable shift to the λ_{max} of the Soret band. $H_2mTAlPyP4$ produced the largest bathochromic shift of 12 nm, followed by $H_2mTOPrPyP4$ with 8 nm and $H_2mTOEtPyP4$ with 6 nm. While $H_2mTOEtPyP4$ had the least red shifted Soret band, it experienced the most hypochromism, with an observed hypochromicity of 41%. $H_2mTOPrPyP4$ and $H_2mTAlPyP4$ followed with hypochromicities of 36% and 34%, respectively.

Over the past few decades numerous studies have investigated the interaction of porphyrins with nucleic acids and have documented that the primary modes of binding are intercalation, end

stacking and/or groove binding. The exact nature of the interaction is dependent on the G4 molecularity, DNA base content and the ionic strength and pH of the media.[168] It had previously been suggested that the modes of binding can be discriminated based on characteristic changes in the porphyrin Soret band. Intercalation is said to be indicated by large bathochromic shifts (≥ 15 nm) and substantial hypochromism (≥ 35%). Outside binding is characterized by smaller red shifts (≤ 8 nm) and small hypochromicity or even hyperchromicity. Interestingly, these values were determined for long pieces of duplex DNA,[169] where end-stacking interactions are not significant, yet these values have been regularly used as benchmarks and applied towards the characterization of G4-porphyrin binding interactions. The changes in the Soret absorption band alone are insufficient for determining the binding mode. However, they are adequate for indicating whether binding events are occurring, which satisfies our primary purpose for conducting these titration experiments.

The lack of a red shift during the titration of our cross-linking porphyrin, H_2mTCAPyP4, was concerning and forced us to further examine the interaction between the cross-linker and G4 DNA by performing CD experiments. Our prepared porphyrins are symmetrical achiral molecules that do not produce CD signals alone. However, a signal can be induced from interactions with chiral asymmetrical G4 structures. An induced signal in the Soret region is indicative of porphyrin-G4 binding.

Just as with the UV-Vis titrations, the titrant solutions were prepared in a manner to maintain a constant porphyrin concentration throughout the titration. CD scans were collected in 150 mM KPBS at 25 °C with 20 minute incubation time between each aliquot addition. The induced CD ellipticity spectra of the porphyrin-G4 DNA complexes are displayed in Figures 92-98. The first of these figure shows the resulting spectra collected from the titration of H_2mTAlPyP4 with $(TTAGGG)_4$. An arrow is used to indicate the incremental increase in DNA concentration with each aliquot addition and the resulting induced CD signals in the Soret region is identified.
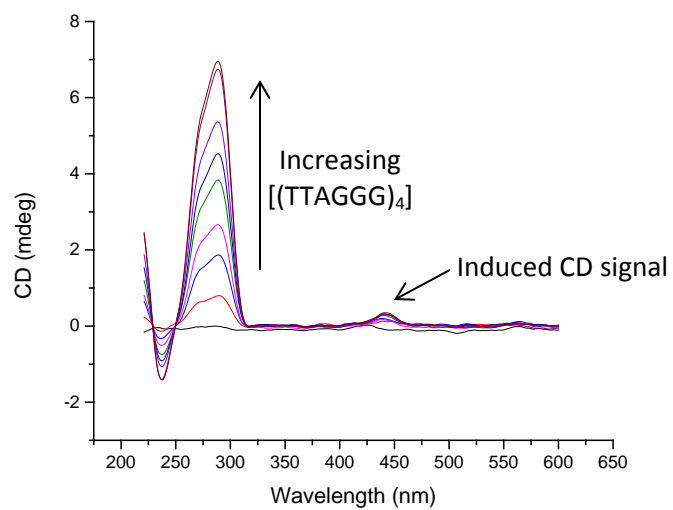
Figure 92. Induced CD spectra of H_2mTAlPyP4 at 25 °C in 150 mM KPBS
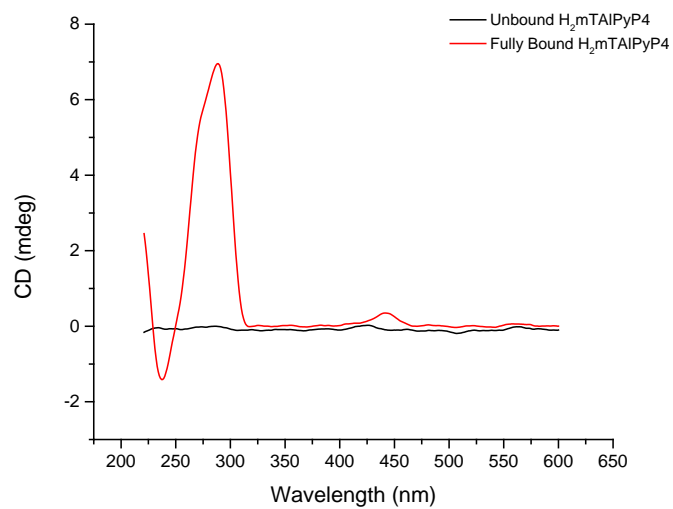
Figure 93. CD spectra of free and fully bound H_2mTAlPyP4 at 25 °C in 150 mM KPBS
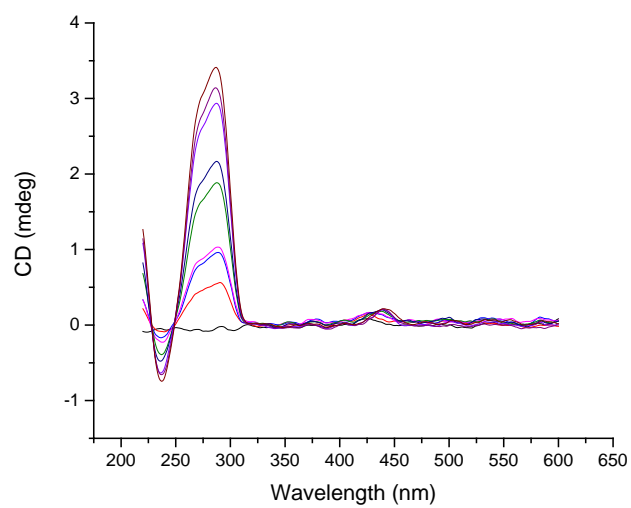
Figure 94. Induced CD spectra of $H_2mTOEtPyP4$ at 25 °C in 150 mM KPBS
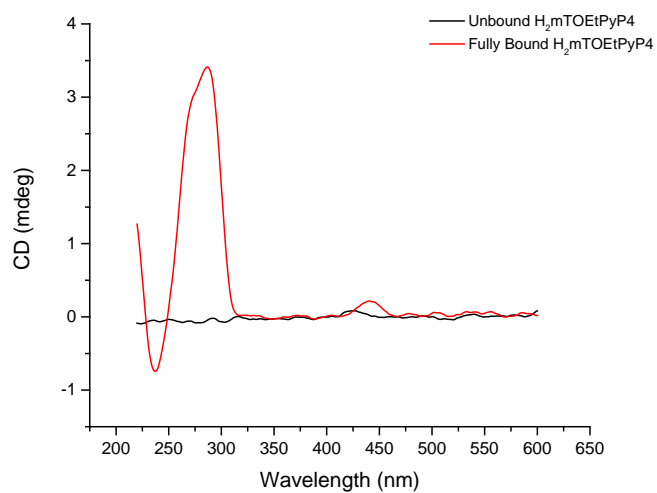
Figure 95. CD spectra of free and fully bound $H_2mTOEtPyP4$ at 25 °C in 150 mM KPBS
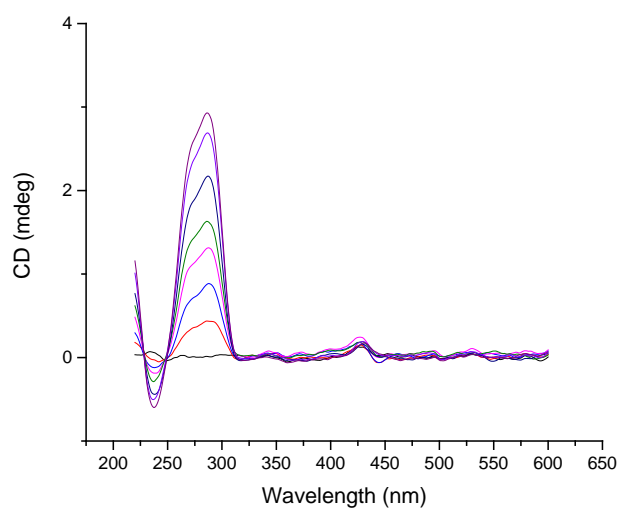
Figure 96. Induced CD spectra of H_2mTOPrPyP4 at 25 °C in 150 mM KPBS
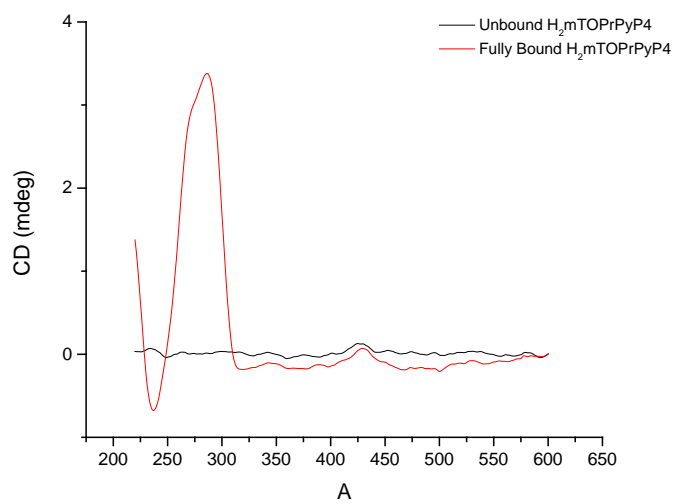
Figure 97. CD spectra of free and fully bound H_2mTOPrPyP4 at 25 °C in 150 mM KPBS
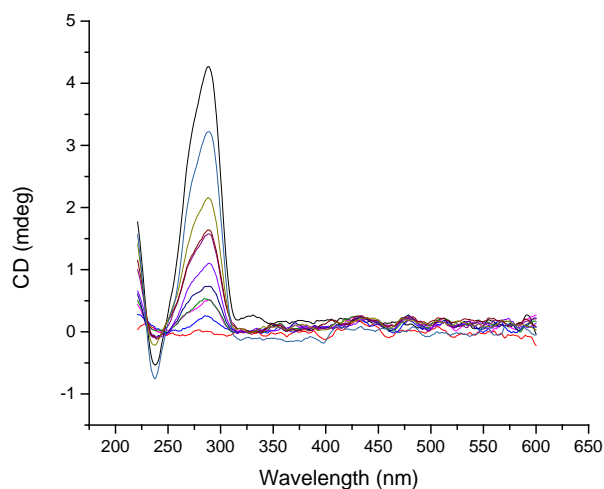
Figure 98. CD titration of $H_2mTCAPyP4$ at 25 °C in 150 mM KPBS

As anticipated, spectra collected on porphyrins prior to the addition of the first aliquot of titrant did not produce any CD signals. The incremental addition of $(TTAGG)_4$ was found to induce CD signals in the Soret region for the allyl and hydroxyl tetrasubstituted porphyrins at approximately 440 nm. The positive induced CD bands increase as the concentration of G4 DNA is increased. These results are consistent with findings published in the literature.[170] However, the CD spectra for $H_2mTCAPyP4$ remained unchanged irrespective of the concentration of G4 DNA added. The lack of any induced spectral changes as a function of increasing G4 concentration indicates that the porphyrin cross-linker does not bind to the quadruplex structure. This result taken in conjunction with the UV-Vis titration data that produced no detectable bathochromic shift further confirms that $H_2mTCAPyP4$ does not bind $(TTAGGG)_4$.

We speculate that this lack of binding is primarily due to electrostatic repulsion resulting from interactions between the deprotonated carboxylic acid moieties found on the $H_2mTCAPyP4$ and anionic DNA. Assuming that the carboxylic acids on the porphyrin cross-linker have similar pKa values to that of hexanoic acid (pKa = 4.88), then in pH 7.2 KPBS buffer these carboxylic acids would be found completely

deprotonated. Therefore, while we have been considering H$_2$mTCAPyP4 as being a tetracationic porphyrin, it in fact possesses an overall net charge of zero in our aqueous buffer solution, with the four carboxylates providing a -4 charge countering the +4 charge provided by the pyridiniums.

Interestingly, previous studies using competitive dialysis experiments demonstrated the structural selectivity of an anionic porphyrin, *N*-methyl mesoporphyrin IX, for quadruplex DNA.[171] This compound, shown in Figure 99, was found to selectively bind G4 DNA in the presence of ssDNA, dsDNA, DNA-RNA hybrids, Z-DNA and triplex DNA. While the level of absolute binding was low, it was not found to bind any form of nucleic acid except for quadruplexed G4 DNA. This suggests that binding to G4 structures by non-cationic porphyrins, even anionic porphyrins carrying charges that repel nucleic acids, is viable and can most likely be attributed to π-stacking interactions, which explains the selectivity exhibited by this porphyrin for quadruplex structures. It should be noted that the competitive dialysis experiments were performed in an aqueous 200 mM Na$^+$ phosphate buffer solution with a pH of 7.0.
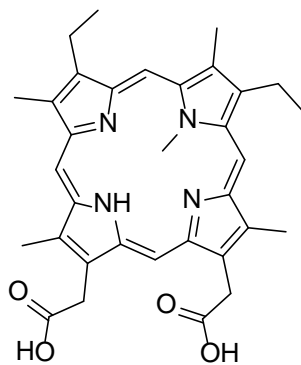
Figure 99. Structure of N-methyl mesoporphyrin IX

Unlike our porphyrin cross-linker, however, this compound contains two carboxylic acid β-substituents on adjacent pyrrole rings, leaving the opposite side of the porphyrin macrocycle free of any anionic charges. This provides a surface that can approach or be approached by the G4 structure without any ensuing repulsive forces. Since our porphyrin cross-linker is tetrasubstituted with an anionic charge emanating from a protruding arm on each side of the porphyrin macrocycle, we believe that

these deprotonated carboxylic acid appendages prevent the cross-linker from having any accessible surface. Thus, preventing the porphyrin and G4 structure from interacting as a result of the electrostatic repulsive forces encountered on all sides.

Alternatively, the length and free rotation of the appendages make it possible for the carboxylates to associate with the pyridinium charges, which would in turn hinder the porphyrin macrocyclic core from π-stacking. This, however, does not account for the 10% hyperchromic effect detected in the UV-Vis titrations. The titration of H_2mTCAPyP4 with (TTAGGG)$_4$ was repeated in order to eliminate the possibility that the hypochromicity observed was an experimental artifact. We found that this effect transpired consistently during the repeated trials. The hypochromisim can most likely be attributed to the formation of aggregates brought on by the electrostatic repulsion forces that promote the self-stacking of the porphyrin. This suggests that it is probably the electrostatic forces and not the association of the carboxylates to the pyridiniums that primarily impede the binding with G4 DNA.

While we were disappointed with the outcome of our binding studies and the realization that we would be unable to perform the EDC/Sulfo-NHS mediated crosslinking reaction in the presence of the G4 structure as we initially intended, we were optimistic that the cross-linked PEI would bind to the G4 structure. This would allow for us to perform a postcross-linking cross-linking, maintaining the spatial arrangement resulting from the interaction of the PEI and H_2mTCAPyP4 adduct with the G4 structure. Since the cross-linking of PEI using H_2mTCAPyP4 couples the carboxylic acids of the porphyrin cross-linker to the primary amines of the PEI by forming amide bonds, there are no longer negatively charged carboxylates present on the porphyrin to repel DNA. Instead the porphyrin moieties at this point each carry a +4 charge.

Upon examination of the binding of our cross-linked PEI with (TTAGGG)$_4$ we found that no bathochromic shift or hypochromicity of the Soret band transpired. Rather, what we observed was a hyperchromic effect on the DNA absorbance peak which is similar to spectra collected when examining

124

the binding interaction of (TTAGGG)$_4$ to uncross-linked branched PEI. Consequently, we decided to examine the effect the cross-linked PEI had on G4 DNA previously bound to cationic porphyrin. What we discovered was that the addition of the cross-linked PEI to G4 DNA bound to H$_2$mTOEtPyP4 caused a disruption in the binding between the two species and results in the displacement of the H$_2$mTOEtPyP4. This was evident with the hypsochromic shift of the Soret band back to its original absorbance maximum, from its fully bound state at 436 to it unbound state at a 424 nm. This effect is shown in Figure 100, where the spectrum for the cross-linked PEI addition is normalized.
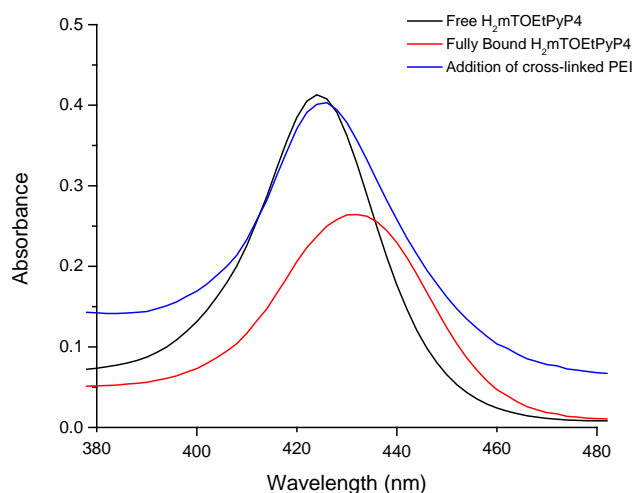
Figure 100. Disruption of porphyrin binding to G4 DNA upon the addition of cross-linked PEI

A similar result was observed when performing this assay using uncross-linked 10,000 MW PEI. Conversely, when adding neutral poly(2-ethyl-2-oxazoline), the unhydrolyzed precursor to linear PEI shown in Figure 101, which only contains tertiary amides, the bound porphryin was not displaced. Clearly, it can be deduced from these findings that the electrostatic interactions between PEI and DNA dominate the binding interaction and prevent the porphyrins from binding to the quadruplex. Ironically, PEI was chosen for its strong affinity towards nucleic acids. However, its high charge density does not provide an opportunity for the porphyrin moieties within the cross-linked PEI to bind with the G4

structure. Furthermore, the PEI interacts so strongly with G4 DNA that porphyrins previously associated with the quadruplex structure are displaced.
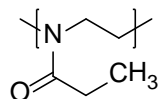
Figure 101. Structure of poly(2-ethyl-2-oxazoline)

In an attempt to resolve this dilemma we proceeded by partially acetylating PEI in order to reduce the overall cationic charge of the polymer. By forming amides from primary and secondary amines on the PEI, the reacted amines will no longer be available for protonation, and as a result will not contribute any cationic charges. This reaction is shown in Scheme 16 and was performed following methods found in the literature, which affords PEIs with varying degrees of acetylation.[172]
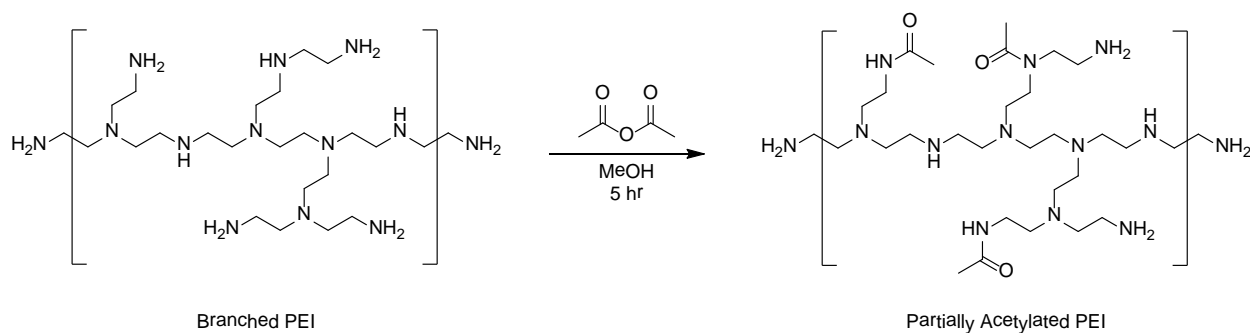
Scheme 16. Partial acetylation of PEI

Briefly, 500 mg of 10,000 MW PEI was added to 3 mL of anhydrous methanol in a scintillation vial. A sufficient amount of acetic anhydride required to achieve the desired degree of acetylation was added, assuming a 100% yield. The vial was sealed and the solution was stirred for 5 hours at 60 °C. The amount of PEI used for each reaction contained 8.7 mmol of primary and secondary amines. Table 9 provides the amounts of acetic anhydride added accordingly, to each reaction mixture, to obtain the indicated percent of acetylation.

Table 9. Conditions for the partial acetylation of 500 mg 10,000 MW PEI

Reaction	Acetic Anhydride (mmol)	% Acetylation
1	0.7	8.0
2	1.45	16.6
3	2.90	33.3
4	5.80	66.6

Upon the addition of these acetylated PEI's to fully porphyrin bound $(TTAGGG)_4$, the Soret band returned to the λ_{max} corresponding to the porphyrin in its unbound state. Clearly, these results suggest that the charge density of PEI remained high regardless of the partial acetylation of its primary and secondary amines. The electrostatic interactions between the polymer and the G4 structure prevent the porphyrin from maintaining its contact with the $(TTAGGG)_4$. This can be either the result of the PEI denaturing the G4 structure or the preference of the G4 for the PEI over the cationic porphyrin. Irrespective of the cause, this simple modification performed to the branched PEI, intending to reduce the overall cationic charge density of the polymer, was found to be insufficient in preventing the displacement of bound cationic porphyrins.

5. Conclusion

This research was originally intended for the development of a novel method for generating soluble molecularly imprinted polymers for G4 DNA. Our design focused primarily on employing a G4 binding ligand, specifically a cationic porphyrin, as a monomeric unit in the formation of the MIP. We successfully synthesized water soluble cationic porphyrins, which were tetra-functionalized in a manner to provide sites for polymerization or for binding to preformed polymers. The screening of commercially available polymers for use in generating the MIP was performed by UV-Vis and gel electrophoretic studies. Polymers identified as nucleic acid interacting polymers were all found to be cationic. All of these contained quaternized amines, with the exception of PEI, which is cationic when protonated in solution. The strong affinity exhibited by PEI for DNA and the availability of reactive amine groups in its structure, led to its use for the purposes of generating a MIP.

We successfully synthesized H_2mTPyP4 using the Adler-Longo method and were able to introduce the cationic charge to the porphyrin *via* alkylation of the pyridyl groups, which formed the charged pyridinium species. Alkylation with allyl bromide yielded H_2mTAlPyP4, while alkylation using the appropriate halohydrin allowed for the preparation of H_2mTOEtPyP4, H_2mTOPrPyP4 and H_2mTOBuPyP4. All attempts at further reacting these cationic porphyrins, post-quaternization, were unsuccessful. These attempts included the epoxidation of mTAlPyP4 with *m*CPBA and the preparation of alkyl bromides from H_2mTOEtPyP4 using PPh_3 and CBr_4.

Consequently, we synthesized H_2mTCAPyP4, a cationic porphyrin containing four carboxylic acid moieties that allows for carbodiimide-mediated coupling reactions. This porphyrin was prepared by quaternizing the H_2mTPyP4 with 6-bromohexanoic acid. In order to increase the yield of the tetraalkylated product, optimization of the reaction conditions employed in previous alkylating procedures was required and was achieved by performing the reaction using fresh anhydrous DMF, under an inert atmosphere, with 40X excess of 6-bromoheaxanoic acid with a reaction time of 72 hours.

Also, the H_2mTPyP4 was heated to 100 °C, in order to ensure complete dissolution of the porphyrin before the addition of the alkylating agent and cooled to 0 °C, in order to precipitate the product at completion. The prepared cationic porphyrins were characterized by ^{1}H NMR, high resolution ESI-MS, and UV-Vis.

The strategy behind the synthesis and use of H_2mTCAPyP4 was based on the ability of the carboxylic acid functional groups to couple with the primary amines of PEI using the peptide coupling agent EDC. One of the main challenges in the preparation of the H_2mTCAPyP4 was the isolation of the tetra-quaternized species, which we accomplished by performing exhaustive SEC using Sephadex LH-20, however, ion-exchange chromatography may have proven to be a superior method.

The primary amines of PEI were successfully coupled to the carboxylic acid moieties of the H_2mTCAPyP4 using EDC/sulfo-NHS mediated coupling reactions carried out in EDTA-free PBS at a pH of 7.2. The incorporation of the sulfo-NHS in the reaction increases the coupling efficiency by converting the *o*-acylisourea intermediate to a more stable amine-reactive sulfo-NHS ester. In essence these coupling reactions are cross-linking reactions; hence, we successfully cross-linked PEI using a soluble cationic porphyrin cross-linker.

Excess cross-linker produced insoluble products, while cross-linking with a 1:4 H_2mTCAPyP4 to PEI ratio gave soluble cross-linked PEI. The successful cross-linking was demonstrated by gel electrophoresis, which correlated the cross-linked PEI's electrophoretic mobility with the cross-linking density. Cross-linking was also confirmed by ESI-MS, which exhibited changes to the mass difference between polymer molecules that did not correlate to the ethyleinimine monomeric molecular weight. Furthermore, ^{1}H NMR revealed peak broadening due to the restricted motion of the polymer in solution, at the segmental level, as a direct result of the cross-linking. With this novel cross-linking strategy in hand, which allows for the cross-linking of primary amine containing polymers, such as PEI, with a

cationic porphyrin cross-linker, we were eager to perform the cross-linking reaction in the presence of G4 DNA.

DNA binding studies demonstrated that cationic porphyrins are not specific for quadruplex DNA and do bind duplex DNA. The most significant structural difference between these two nucleic acid structures is the G-quartet. Therefore, it is logical that the G-quartet be a key target for designing MIPs that can discriminate G4 structures from duplex DNA. Cationic porphyrins that contain sterically hindering bulky substituent groups would be more inclined to interact with G4 DNA *via* an external mode of binding rather than intercalation. We believe this to be the case for cationic porphyrin cross-linkers intra- or intermolecularly bound to large macromolecules. With the reduction in the ability of the cationic porphyrin to intercalate, the ability to bind duplex DNA is dramatically hindered. Furthermore, in order to discriminate a specific G4 structure from a diverse group of G4s, it is necessary for the MIP to possess groove recognition ability. This offers the potential for enhanced selectivity among various G4 structures. Therefore the MIP should simultaneously possess dual-site binding ligands, which can bind to both the terminal G-quartet and the grooves. This would most likely result with a highly specific material.

The preparation of a MIP using the EDC/sulfo-NHS coupling strategy was designed in a manner where the polymer that interacts with the grooves is subsequently coupled to our cross-linking cationic porphyrin that has been previously bound to the G4 structure, presumably by external binding with the terminal G-quartet. Unfortunately, DNA binding studies performed on our prepared cross-linker revealed that the cross-linking porphyrin, $H_2mTCAPyP4$, does not bind G4 DNA, specifically $(TTAGGG)_4$. The allyl and hydroxyl functionalized porphyrins produced hypochromic and bathochromic shifts when titrated with $(TTAGGG)_4$, which is indicative of porphyrin-G4 binding events. Conversely, $H_2mTCAPyP4$, exhibited only 10% hypochroism with no red shift. Furthermore, CD experiments did not produce induced CD signal when examining the interaction of the cross-linker with G4 DNA. All other cationic

porphyrins examined did exhibit the expected induced CD signal in the Soret region. We speculate that this lack of binding is due to electrostatic repulsion forces between the deprotonated carboxylic acid moieties found on the cross-linker and anionic DNA. The hypochromism observed in the UV-Vis titration experiments is most likely a direct result of the formation of porphyrin aggregates.

When examining the binding of the cross-linked PEI with G4 DNA, we found the interaction to be dominated by the PEI. UV-Vis data suggests that the porphyrin cross-linker is not provided the opportunity to associate with the G4 structure. The interaction of G4 DNA with cross-linked PEI appears to be similar to its interaction with uncross-linked PEI. Clearly, the high cationic charge density on PEI allows the polymer portion of the adduct to bind electrostatically with the G4 DNA, an interaction that is much stronger than the π-stacking interaction that the porphyrin is capable of partaking in. Attempts at reducing the cationic charge density of PEI by partial acetylation of its primary and secondary amines proved insufficient at resolving this dilemma. Furthermore, the cross-linked PEI was found to disrupt the binding interactions between cationic porphyrins previously bound to G4 DNA, specifically by displacing the ligand. This effect is also observed when examining uncross-linked PEI. However, the assessment of the neutral polymer, poly(2-ethyl-2-oxazoline), a linear PEI precursor, was found not to disrupt this binding interaction.

We were successful at synthesizing a tetracationic pyridinium porphyrin capable of cross-linking amine functionalized polymers, which allowed for the preparation of water soluble cross-linked PEI, using carbodiimide chemistry that incorporated cationic porphyrin structures into the polymer matrix. However, we did not realize our overall objective of preparing a water soluble MIP for G4 DNA. The need to explore alternative polymers for achieving this is warranted. Such polymers should bind favorably to nucleic acids as to not disrupt the G4 structure and not compete or interrupt the binding of the porphyrin moieties. This most likely will be a weakly charged cationic polymer that can be cross-linked with $H_2mTCAPyP4$. These include polymers such as poly(allylamine) and poly(1,3,5-triazine-2,4-

diamine). While the network structure of our prepared soluble cross-linked polymer still requires optimization, it has shown promise and demonstrates the opportunities for new soluble molecularly imprinted polymer designs that include quadruplex binding sites.

References

[1] Sreenivasan K. *React. Funct. Polym.* **2007**, *67*, 859-864.

[2] Chen B.; Piletsky S.; Turner A.P. *Comb. Chem. High Throughput Screening* **2002**, *5*, 409-427.

[3] Peppas N.A.; Bergmann N.M. *Prog. Polym. Sci.* **2008**, *33*, 271-288.

[4] Fischer E. *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985-2993.

[5] Mudd S. *J. Immunol* **1932**, *23*, 423-427.

[6] Pauling L. *J. Am. Chem. Soc.* **1940**, *62*, 2643-2657.

[7] Spivak D.A; Shea J.K. *Anal. Chim. Acta* **2001**, *435*, 65-74.

[8] Dickey F.H. *Proc. Natl. Acad. Sci. U.S.A.* **1949**, *35*, 227-229.

[9] Shea, K.J.; Hart, B.R. *J. Am. Chem. Soc.* **2001**, *123*, 2072-2073.

[10] Mathew, J.; Buchardt, O. *Bioconjugate Chem.* **1995**, *6*, 524-528.

[11] Wulff, G.; Sarhan, A. *Angew. Chem., Int. Ed. Engl.* **1971**, *11*, 341-344.

[12] Mosbach, K.; Ramstrom, O. *Nat. Biotechnol.* **1996**, *14*, 163.

[13] Mosbach, K.; Sellergen, B.; Lepisto, M. *J. Am. Chem. Soc.* **1988**, *110*, 5853.

[14] Stevens, M.P. *Polymers Chemistry an Intoduction*, 3rd Edition; Oxford University Press: New York, 1999; pp 502-509.

[15] Watson, J.D.; Cick, F.H. *Nature* **1953**, *171*, 737-738.

[16] Franklin, R.E.; Gosling, R.G. *Nature* **1953**, *171*, 964-967.

[17] Zamenhof, S.; Brawerman, G.; Chargaff, E. *Biochim. et Biophys. Acta* **1952**, *9*, 402.

[18] Nelson, D.L; Cox, M.M In *Lehninger Principles of Biochemistry*, 4th Edition; Katherine A., Morgan R., O'Neill J., Wong V., Ciprioni J., Mays M., Tymoczko N., Eds.; W.H Freeman and Company: New York, 2005; pp 282-291.

[19] Qin Y.; Hurley L.H. *Biochimie* **2008**, *90*, 1149-1171.

[20] Wells, R. D., Blakesley, R. W., Hardies, S. C., Horn, G. T., Larson, J. E., Selsing, E., Burd, J. F., Chan, H. W., Dodgson, J. B., Jensen, K. F., Nes, I. F., and Wartell, R. M. *Crit. Rev. Biochem.* **1977**, *4*, 305–340.

[21] Bacolla A.; Well R.D. *J. Biol. Chem.* **2004**, *279*, 47411-47414.

[22] Callis P.R. *Annu. Rev. Phys. Chem.* **1983**, *34*, 329-357.

[23] Shukla M.K.; Leszczynsko J. *J. Biomol. Struct. Dyn.* **2007**, *25*, 93-118.

[24] Lander E.S; Linton L.M.; Birren B.; Nusbaum C. *Nature* **2004,** *431*, 931-945.

[25] McConkey E.H; Varki A. *Science* **2005**, *309*, 1499-1501.

[26] Eddy J.; Maizels N. *Nucleic Acids Res.* **2006**, *34*, 3887-3896.

[27] Bourdoncle A.; Esteves Torres A.; Gosse C.; Lacroix L.; Vekhoff P.; Le Suax T.; Jullien L.; Mergeny J.L. *J. Am. Chem. Soc.* **2006**, *128*, 11094-11105.

[28] Huppert J.L.; Balasubramanian S. *Nucleic Acids Res.* **2005**, *33*, 2908-2916.

[29] Huppert, J.L. *Chem. Soc. Rev.* **2008**, *37*, 1375-1384.

[30] Nakken S.; Rognes T.; Hovig E. *Nucleic Acids Res.* **2009**, *37*, 5749-5756.

[31] Capra J.A; Paeschke K.; Singh M. Zakian V.A. G-Quadruplex DNA Sequences Are Evolutionarily Conserved and Associated with Distinct Genomic Features in Saccharomyces cerevisiae *PLoS Comput. Biol.* [Online] **2010**, *6*, e1000861. DOI: 10.1371/journal.pcbi.1000861. http://ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1000861 (accessed Apr 10, 2014).

[32] Bang. I. *Biochem. Z.* **1910**, *29*, 293-231.

[33] Gellert M.; Lipsett M.N.; Davies D.R. *Proc. Natl. Acad. Sci. U.S.A.* **1962**, *48*,2013-2018.

[34] Guschlbauer W.; Chantot J.F.; Thiele D. *J. Biomol. Struct. Dyn.* **1990**, *8*, 491-511.

[35] Williamson J.R.; Raghuraman M.K.; Cech T.R. *Cell* **1989**, *59*, 871.

[36] Bochman M.L.; Paeschke K.; Zakian V.A. *Nat.Rev. Genet.* **2012**, *1*, 770-780.

[37] Burge S.; Parkinson G.N.; Hazel P.; Todd A.K.; Neidle S. *Nucleic Acids Res.* **2006**, *34*, 5402-5414.

[38] Huppert J.L.; *FEBS J.* **2010**, *277*, 3452-3458.

[39] Cang X.; Sponer J. Cheatham T.E. *J. Am. Chem. Soc.* **2011**, *133*, 14270-14279.

[40] Balkwill G.D.; Garner T.P.; Searle M.S. *Mol. BioSyst.* **2009**, *5*, 542-547.

[41] Simonsson T. *Biol. Chem.* **2001**, *382*, 621-628.

[42] Cang X.; Sponer J.; Cheatham T.E. *J. Am. Chem. Soc.* **2011**, *133*, 14270-14279.

[43] Lane, A.N.; Jenkins T.C. *Curr. Org. Chem.* **2001**, *5*, 845-869.

[44] Huppert J.L. *Chem. Soc. Rev.* **2008**, *37*, 1375-1384.

[45] Karsisitotis A.I.; O'Kane C.; da Silva M.W. *Methods* **2013**, *64*, 28-35.

[46] Sen D.; Gilbert W. *Nature* **1990**, *344*, 410-414.

[47] Williamson, J.R. *Annu. Rev. Biophys. Biomol. Struct*. **1994**, *23*, 703-730.

[48] Hud N.V.; Smith F.W.; Anet F.A.; Feigon J. *Biochemistry* **1996**, *35*, 15383-15390.

[49] Qin Y.; Hurley L.H. *Biochimie* **2008**, *90*, 1149-1171.

[50] Hurley L.H.; Sun D. *Methods in Enzymol.* **2001**, *340*, 573-592.

[51] Harley C.B.; Futcher A.B.; Greider C.W. *Nature* **1990**, *345*, 458-469.

[52] Lingner J.; Cooper J.P.; Cech T.R. *Science* **1995**, *269*, 1533-1534.

[53] Palm W.; de Lange T. *Annu. Rev. Genet.* **2008**, 42, 301-334.

[54] De Lange T. *Genes Dev.* **2005**, *19*, 2100-2110.

[55] Palm W.; de Lange T. *Annu. Rev. Genet.* **2008**, *42*, 301-334.

[56] Neidle S.; Parkinson G.N. *Curr. Opin. Struct. Biol.* **2003**, *13*, 275-283.

[57] Dai I.; Punchihewa C.; Ambrus A.; Chen D.; Jones R.A.; Yang D. *Nucleic Acids Res.* **2007**, *35*, 2440-2450.

[58] Dai J.; Carver M.; Yang D. *Biochimie* **2008**, *90*, 1172-1183.

[59] Prescott D.M. *Microbiol. Rev.* **1994**, *58*, 233-237.

[60] Schaffitzel C.; Gerger I.; Postberg J.; Hanes J.; Lipps H.J.; Pluckthun A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8572-8577.

[61] Lipps H.J.; Rhodes D. *Trends Cell Biol.* **2009**, *19*, 414-422.

[62] Chang C.C.; Kuo I.C.; Lim J.J.; Chen C.T.; Back H.T.; Lou P.J.; Chang T.C. *Chem. Biodiversity* **2004**, *1*, 1377-1384.

[63] Yang Q.; Xiang J.; Yang S.; Zhou Q.; Li Q.; Tang Y.; Xu G. *Chem. Commun.* **2009**, *9*, 1103-1105.

[64] Bochman M.L.; Paeschke K.; Zakian V.A. *Nat. Rev. Genet.* **2012**, *13*, 770-780.

[65] Mohaghegh P; Karow J.K.; Brosch Jr. R.M.; Bohr V.A.; Hickson I.D. *Nucleic Acids Res.* **2001**, *29*, 2843-2849.

[66] Huber M.D.; Lee D.C.; Maizels N. *Nucleic Acids Res.* **2002**, *30*, 3954-3961.

[67] London T.B. *J. Biol. Chem.* **2008**, *283*, 36132-36139.

[68] Sanders C.M. *Biochem J.* **2010**, *430*, 119-128.

[69] Huppert J.L.; Balasubramanian S. *Nucleic Acids Res.* **2007**, *35*, 406-413.

[70] Eddy J.; Maizels N. *Nucleic Acids Res.* **2008**, *36*, 1321-1333.

[71] Lansdorp P.M. *EMBO J.* **2009**, *28*, 2532-2540.

[72] Armanios M. *Annu. Rev. Genomics Hum. Genet.* **2009**, *10*, 45-61.

[73] Wu Y.; Brosh R.M.; *FEBS J.* **2010**, *277*, 3470-3488.

[74] Monchaud D.; Teulade-Fichou M-.P. *Org. Biomol. Chem.* **2008**, *6*, 627-636.

[75] Sun D.; Thompson B.; Cathers B.E.; Salazar M.; Kerwin S.M; Trent J.O.; Jenkins T.C.; Neidle S.; Hurley L.H. *J. Med. Chem.* **1997**, *40*, 2113-2116.

[76] Perry P.J.; Read M.A.; Davies R.T.; Gowan S.M.; Reszka P.; Wood A.A.; Kelland L.R.; Neidle S. *J. Med. Chem*. **1999**, *42*, 2679-2684.

[77] Harrison R.J.; Reszka A.P.; Haider S.M.; Romangnoli B.; Morrell J.; Read M.A.; Gowan S.M.; Incles C.M.; Kelland L.R.; Neidle S. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5845-5849.

[78] Harrison R.J.; Gowan S.M.; Kelland L.R.; Neidle S. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2643-2646.

[79] Mergny J.-L.; Maurizot J.-C. *ChemBioChem* **2001**, *2*, 124-132.

[80] De Cian A.; Guittat L.; Kaise M.; Sacca B.; Amrane S.; Bourdoncle A.; Alberti P.; Teulade-Fichou M-.P.; Lacroix L.; Mergny J-.L. *Methods* **2007**, *42*, 183-195.

[81] White E.W.; Tanious F.; Ismail M.A.; Reszka A.P.; Neidle S.; Boykin D.W.; Wilson W.D. *Biophys. Chem.* **2007**, *126*, 140-153.

[82] Burher A.M.; Dai F.; Schultes C.M.; Raszka A.P.; Moore M.J.; Double J.A.; Neidle S. *Cancer Res.* **2005**, *65*, 1489-1496.

[83] Mergny J.L.; Lacroix L.; Teulade-Fichou M-.P.; Hounsou C.; Guittat L.; Hoarau M.; Arimondo P.B.; Vigneron J.P.; Lehn J.M.; Riou J.F.; Grastier T.; Helene C. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3062-3067.

[84] Hounsou C.; Guittat L.; Monchaud D.; Jourdan M.; Saettel N.; Mergny J-Fj.L.; Teulade-Fishou M-.P. *ChemMedChem* **2007**, *2*, 655-666.

[85] Alberti P.; Schmitt P.; Nguyen C-.H.; Rivalle C.; Hoarai M.; Grierson D.S.; Mergny J-.L. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1071-1074.

[86] Caprio V.; Guyen B.; Opoku-Boahen Y.; Mann J.; Gowan S.M.; Kelland L.M.; Read M.A. Neidle S. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2063-2066.

[87] Zhou J-.L.; Lu Y-.J.; Ou T-.M.; Zhou J-.M.; Huang Z-.S.; Zhu X-.F.; Du C-.J.; Bu X-.Z.; Ma L.; Gu L-.Q.; Li Y-.M.; Chan A. S-C. *J. Med. Chem.* **2005**, *48*, 7315-7321.

[88] Teulade-Fichou M-.P.; Carrasco C.; Guittat L.; Bailly C. Alberti P.; Mergny J-.L.; David A.; Lehn J-.M.; Wilson W.D. *J. Am. Chem. Soc.* **2003**, *125*, 4732-4740.

[89] Kaiser M.; De Cian A.; Sainlos M.; Renner C.; Mergny J-.L.; Teulade-Fichou M-.P. *Org. Biomol. Chem.* **2006**, *4*, 1049-1057.

[90] Jantos K.; Rodriguez R.; Ladame S.; Shirude P.S.; Balasubramanian S. *J. Am. Chem. Soc.* **2006**, *128*, 13662-13663.

[91] Monchaud D.; Teulade-Fichou M-.P. *Org. Biomol. Chem.* **2008**, *6*, 627-636.

[92] Shi D-.F.; Wheelhouse R.T.; Sun D.; Hurley L.H. *J. Med. Chem.* **2001**, *44*, 4509-4523.

[93] Monchaud D.; Granzhan A.; Saettel N.; Guedlin A.; Mergny J-.L.; Teulade-Fichou M-.P. "One Ring to Bind Them All" – Part I: The Efficiency of the Macrocyclic Scaffold for G-Quadruplex DNA Recognition *J. Nucleic Acids* [Online] **2010**, DOI: 10.4061/2010/525862. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2915875 (accessed Apr 10, 2014).

[94] Wei C.; Jia G.; Yuan J.; Feng Z.; Li C. *Biochemistry* **2006**, *45*, 6681-6691.

[95] Phan A.T.; Kuryavyi V.; Gaw H.Y.; Patel D.J. *Nat. Chem. Biol.* **2005**, *1*, 167-173.

[96] Wang P.; Ren L.; He H.; Ling F.; Zhou X.; Tan Z. *ChemBioChem* **2006**, *7*, 1155–1159.

[97] Pennarun G.; Granotier C.; Gauthier L.R.; Gomez D.; Hoffschir F.; Mandine E.; Riou J-.F.; Mergny J-.L.; Mailliet P.; Boussin F.D. *Oncogene* **2005**, *24*, 2917-2928.

[98] De Cain A.; Mergny J-.L.; *Nucleic Acids Res.* **2007**, *35*, 2483-2493.

[99] De Cian A.; DeLemos E.; Mergny J-.L.; Teulade-Fichou M-.P; Monchaud D. *J. Am. Chem. Soc.* **2007**, *129*, 1856-1857.

[100] Reed J.E.; Arnal A.A.; Neidle S.; Vilar R. *J. Am. Chem. Soc.* **2006**, *128*, 5992-5993.

[101] Evans S.E.; Mende M.A.; Turner K.B.; Keating L.R.; Grimes R.T.; Melchoir S.; Szalai V.A. *J. Biol. Inorg. Chem.* **2007**, *12*, 1235-1249.

[102] Maraval A.; Franco S.; Vialas C.; Pratviel G.; Blasco M.A.; Meunier B. *Org. Biomol. Chem*. **2003**, *1*, 921-927.

[103] Vialas C.; Pratviel G.; Meunier B. *Biochemistry* **2000**, *39*, 9514-9522.

[104] Dixon I.M.; Lopez F.; Tejare A.M.; Esteve J-.P; Blasco M.A.; Pratviel G.; Meunier B. *J. Am. Chem. Soc.* **2007**, *129*, 1502-1503.

[105] Shin-ya K.; Wierzba K.; Matsuo K.; Ohtani T.; Yamada Y.; Furihata K.; Hayakawa Y.; Seto H. *J. Am. Chem. Soc.* **2001**, *123*, 1262-1263.

[106] Doi T.; Yoshida M.; Shin-ya K.; Takahashi T. *Org. Lett.* **2006**, *8*, 4165-4167.

[107] Minhas G.S.; Pilch D.S.; Kerrigan J.E.; Lavoie E.J.; Rice J.E. *Bioorg. Med.Chem.Lett.* **2006**, *16*, 3891-3895

[108] Gouterman M. *J. Mol. Spectrosc.* **1962**, *6*, 138-163.

[109] Nemykin V.N.; Hadt R.G. *J. Phys. Chem. A* **2010**, *114*, 12062-12066.

[110] Rothemund P. *J. Am. Chem. Soc.* **1936**, *58*, 625-627.

[111] Adler A.D.; Longo F.R.; Shergalis W. *J. Am. Chem. Soc.* **1964**, *86*, 3145-3149.

[112] Adler A. D.; Longo F.R.; Finarelli J.D.; Goldmacher J.; Assour J.; Korakoff L. *J. Org. Chem.* **1967**, *32*, 476-477.

[113] Lindsey J.S. In *The Porphyrin Handbook: Synthesis and Organic Chemistry*, 1st Edition; Kadish K.M., Smith K.M., Guilard R., Eds.; Academic Press: San Diego, **2000**; Volume I, pp 49-55.

[114] Kim J.B.; Leonard J.J.; Longo F.R. *J. Am. Chem. Soc.* **1972**, *94*, 3986-3992.

[115] Dolphin D. *J. Heterocycl. Chem.* **1970**, *7*, 275.

[116] Rousseau K.; Dolphin D. *Tetrahedron Lett.* **1974**, *48*, 4251-4254.

[117] Lindsey J.S.; Schreiman I.C.; Hsu H.C.; Kearney P.C.; Marguerettaz A.M. *J. Org. Chem.* **1987**, *52*, 827-836.

[118] Lindsey J.S. *Acc. Chem. Res.* **2010**, *43*, 300-311.

[119] Gary D.J; Puri N.; Won Y. *J. Controlled Release* **2007**, *121*, 64.

[120] Kloeckner J.; Wagner E.; Ogris M. *Eur. J. Pharm. Sci.* **2006**, *29*, 414.

[121] Shea, K.J.; Spivak, D.A. *Macromolecules* **1998**, *31*, 2160.

[122] Wulff, G. In *Molecular Interactions in Bioseparations,* 1st Edition; Ngo T.T., Ed.; Plenum Press: New York, 1993; pp 363-381.

[123] Fleisher E.B. *Inorg. Chem.* **1962**, 1, 493-495.

[124] Sugata S.; Yamanouchi S.; Matsushima Y. *Chem. Pharm. Bull.* **1977**, *25*, 884-889.

[125] Zimmerman J.; Siggel U.; Fuhrhop J-.H.; Roder B. *J. Phys. Chem. B* **2003**, *107*, 6019-6021.

[126] Mamardashvili N.Zh.; Golubchikov O.A. *Russ. Chem. Rev.* **2001**, *70*, 577-606.

[127] Sola M.; Lledos A.; Duran M.; Bertran J.; Abboud J-.L.M *J. Am. Chem. Soc.* **1991**, *113*, 2873-2879.

[128] Dchug J.C.; Viers J.W. *J. Org. Chem.* **1983**, *48*, 4892-4899.

[129] Gonzalez-Lafont A.; Truhlar D.G. In *Chemical Reactions in Clusters*, 1st Edition; Bernstein E.R., Ed.; Oxford University Press: New York, 1996; pp 25-26.

[130] Kim S.H.; Yoh S.-D.; Lim C.; Mishima M.; Fujio M.; Tsuno Y. *J. Phys. Org. Chem.* **1998**, *11*, 254-260.

[131] Stevens, M.P. *Polymers Chemistry an Intoduction*, 3rd Edition; Oxford University Press: New York, 1999; pp 188-191.

[132] Picard N.; Ali H.; Lier J.E.; Klarskov K.; Paquette B. *Photochem. Photobiol. Sci.* **2009**, *8*, 224-232.

[133] Bellucci G.; Catelani G.; Chiappe C.; D'Andrea F.; Grigo G. *Tetrahedron: Asymmetry* **1997**, *8*, 765-773.

[134] Leemhuis M.; Akerody N.; Kruijtzer J.A.W.; van Nostrum C.F; Hennink W.E. *Eur. Polym. J.* **2008**, *44*, 308-317.

[135] Igarasgi S.; Yotsuyanagi Takao *Chem. Lett.* **1984**, *11*, 1871-1874.

[136] Tovmasyan A.G.; Babayan N.S.; Sahakyan L.A.; Shahkhatuni A.G.; Gasparyan G.H.; Aroutiounian R.M.; Ghazaryan R.K. *J. Porphyrins Phthalocyanines* **2008**, *12*, 1100-1110.

[137] Appel R.; Kleinstuck R.; Ziehn K.-D. *Chem. Ber.* **1971**, *104*, 1035.

[138] Deady L.W.; Finlayson W.L.; Korytsky O.L. *Aust. J. Chem*. **1972**, *32*, 1735-1742.

[139] Buchel K.H.; Homeyer B. Combating Nematodes with Alkanesulfonic Acid 2-bromoethyl esters. U.S. Pat. Appl. US3934026, Nov 21, 1973.

[140] Shealy Y.F.; Krauth C.A. Struck R.F.; Montgomery J.A*. J. Med. Chem.* **1983**, *26*, 1168-1173.

[141] Han S.Y; Kim Y.A *Tetrahedron* **2004**, *60*, 2447-2467.

[142] Giuntini F.; Alonso C.M.A; Boyle A. R. *Photochem. Photobiol. Sci.* **2011**, *10*, 759-791.

[143] Tome J.P.C; Neves M.G.P.M.S.; Tome A.C.; Cavalerio J.A.S.; Soncin M.; Magaraggis M.; Ferro S.; Jori G. *J. Med. Chem.* **2004**, *47*, 6649-6652.

[144] Xu W.; Park J.Y.; Kattel K.; Ahmad M.W.; Bony B.A.; Heo W.C.; Jin S.; Park J.W.; Chang Y.; Kim T.J.; Park J.A.; Do J.Y.; Chae K.S.; Lee G.H. *RSC Adv.* **2012**, *2*, 10907-10915.

[145] Kunishima M.; Kawachi C.; Hioko K.; Terao K.; Tani S. *Tetrahedron* **2001**, *57*, 1551-1558.

[146] Eiselt P.; Lee K.Y.; Mooney D.J. *Macromolecules* **1999**, *32*, 5561-5566.

[147] Datta A.; Quintavalla S.M.; Groves J.T. *J. Org. Chem.* **2007**, *72*, 1818-1821.

[148] Morille M.; Passirani C.; Vonarbourg A.; Clavreul A.; Benoit J.P. *Biomaterials* **2008**, *29*, 3477-3496.

[149] von Harpe A.; Petersen H.; Li Y.; Kissel T. *J. of Controlled Release* **2009**, *69*, 309-322.

[150] Lambermont-Thijs H.M.; van der Woerdt F.S.; Baumgaertel A.; Bonami L.; Du Prez F.E.; Schubert U.S.; Hoongenboom R. *Macromolecules* **2010**, *43*, 927-933.

[151] Boussif O.; Lezoualc'h F.; Zanta M.A.; Mergny M.D.; Scherman D.; Demeneix B.; Behr J.P. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *92*, 7297-7301.

[152] Wang Y.; Chen P.; Shen J. *Biomaterials* **2006**, *27*, 5292-5298.

[153] Lungwitz U.; Breunig M.; Blunk T.; Gopferich A. *Eur. J. Pharm. Biopharm.* **2005**, *60*, 247-266.

[154] Suh J.; Lee S.H.; Kim S.M.; Hah S.S. *Bioorg. Chem.* **1997**, *25*, 221–231.

[155] Ziebarth J.D.; Wang Y. *Biomacromolecules* **2010**, *11*, 29-38.

[156] Suh J.; Paik H.J.; Hwang B.K. *Bioorg. Chem.* **1994**, *22*, 318–327.

[157] Nagaya J.; Homma M.; Tanioka A.; Minakata A. *Biophys. Chem.* **1996**, *60*, 45-51.

[158] Choosakoonkriang S.; Lobo B.A.; Koe G.S, Koe J.G; Middaugh C.R. *J. Pharm. Sci.* **2003**, *92*, 1710-1722.

[159] von Harpe A.; Petersen H.; Li Y.; Kissel T. *J. Controlled Release* **2000**, *69*, 309-322.

[160] Sun H.; Zhou Q.; Xiang J.; Tang Y. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4669-4672.

[161] Williams A.; Ibrahim I.A. *J. Am. Chem. Soc.* **1981**, *103*, 7090-7095.

[162] Staros J.V. *Biochemistry* **1982**, *21*, 3950-3955.

[163] Hermanson G.T. In *Bioconjugate Techniques*, 3rd Edition; Audet J., Preap M., Eds.; Academic Press: London, 2013; pp 259-269.

[164] Grabarek Z.; Gergely J. *Anal. Biochem.* **1990**, *185*, 131-135.

[165] Fundamentals of Membrane Dialysis. http://www.spectrumlabs.com/dialysis/Fund.html (accessed Apr 10, 2014), part of SpectrumLabs.com http://www.spectrumlabs.com (accessed Apr 10, 2014).

[166] Lie K-. J.; Burlant W. *J. Polym. Sci. Part A-1: Polym. Chem.* **1967**, *5*, 1407-1413.

[167] Ambrus A.; Chen D.; Dai J.; Bialis T.; Jones R.A.; Yang D. *Nucleic Acids Res.* **2006**, *34*, 2723-2735.

[168] Sheardy R.D.; Anantha N.V.; Azam M *Biochemistry*, **1998**, *37*, 2709-2714.

[169] Pasternack R.F.; Gibbs E.J.; Villafranca J.J. *Biochemistry*, **1983**, *22*, 2406-2414.

[170] Ghazaryan A.A.; Dalyan Y.B.; Haroutiunian S.G.; Vardayan V.I.; Ghazaryan R.K.; Chalikian T.V. *J. Biomol. Struct. Dyn.* **2006**, *24*, 67-74.

[171] Chaires J.B.; Ren J. *Biochemistry* **1999**, *38*, 16067-16075.

[172] Forrest M.L.; Meister G.E.; Koerber J.T.; Pack D.W. *Pharm. Res.* **2004**, *21*, 365-371.

Appendix

MS and NMR Spectra

Cpd 1: C40H26N8: +ESI Scan (0.175-0.398 min, 28 scans) Frag=165.0V SHJHJH03A.d Subtract

x10 4

Counts vs. Mass-to-Charge (m/z)

223.0641

311.1292

338.3419

619.2356|(M+H)+

Cpd 1: C40H26N8: +ESI Scan (0.175-0.398 min, 28 scans) Frag=165.0V SHJHJH03A.d Subtract

x10 4

Counts vs. Mass-to-Charge (m/z)

619.2356|(M+H)+

619.2580

620.2395|(M+H)+

621.2477|(M+H)+

622.2530

623.2589

142

Cpd 1: C40H26N8: +ESI Scan (0.175-0.398 min, 28 scans) Frag=165.0V SHJHJH03A.d Subtract

x10 4

310.1215||(M+2H)+2

310.6237||(M+2H)+2

311.1292

311.6315

312.1342

312.6353

Counts vs. Mass-to-Charge (m/z)

143

Cpd 1: C52H42N8: +ESI Scan (0.172-0.569 min, 49 scans) Frag=165.0V SHJH-JH04A.d Subtract

x10 4

Counts vs. Mass-to-Charge (m/z)

260.4592⬜(M+3H)+3

390.1844⬜(M+2H)+2

Cpd 1: C52H42N8: +ESI Scan (0.172-0.569 min, 49 scans) Frag=165.0V SHJH-JH04A.d Subtract

x10 2

Counts vs. Mass-to-Charge (m/z)

390.1844⬜(M+2H)+2

390.6861⬜(M+2H)+2

391.1896⬜(M+2H)+2

391.6912

144

Cpd 1: C52H42N8: +ESI Scan (0.172-0.569 min, 49 scans) Frag=165.0V SHJHJH04A.d Subtract

×10 4

Counts vs. Mass-to-Charge (m/z)

260.4592▯(M+3H)+3

260.7935▯(M+3H)+3

261.1280▯(M+3H)+3

261.4634▯(M+3H)+3

Cpd 1: C48H42N8O4 +ESI Scan (0.169-0.510 min, 42 scans) Frag=165.0V SHJHJH07C.d Subtract

199.5911[(M+4H)+4

265.7858[(M+3H)+3

398.1742[(M+2H)+2

Counts vs. Mass-to-Charge (m/z)

Cpd 1: C48H42N8O4 +ESI Scan (0.169-0.510 min, 42 scans) Frag=165.0V SHJHJH07C.d Subtract

398.1742[(M+2H)+2

398.6751[(M+2H)+2

399.1775[(M+2H)+2

Counts vs. Mass-to-Charge (m/z)

146

x10 ³

Cpd 1: C48H42N8O4: +ESI Scan (0.169-0.510 min, 42 scans) Frag=165.0V SHJHJH07C.d Subtract

Counts vs. Mass-to-Charge (m/z)

265.7858‖(M+3H)+3

266.1199‖(M+3H)+3

266.4547‖‖(M+3H)+3

266.7893‖‖(M+3H)+3

x10 ³

Cpd 1: C48H42N8O4: +ESI Scan (0.169-0.510 min, 42 scans) Frag=165.0V SHJHJH07C.d Subtract

Counts vs. Mass-to-Charge (m/z)

199.5911‖(M+4H)+4

199.8418‖(M+4H)+4

200.0943‖(M+4H)+4

147

Cpd 1: C64 H66 N8 O8: +ESI Scan (0.133-0.316 min, 23 Scans) Frag=365.0V SH-JH-JH20A.d Subtract

x10 5

Counts vs. Mass-to-Charge (m/z)

269.6334[I](M+4H)+4
359.1750[II](M+3H)+3
481.2240
538.2577[III](M+2H)+2

Cpd 1: C64 H66 N8 O8: +ESI Scan (0.133-0.316 min, 23 Scans) Frag=365.0V SH-JH-JH20A.d Subtract

x10 5

Counts vs. Mass-to-Charge (m/z)

269.6334[I](M+4H)+4
359.1750[II](M+3H)+3
481.2240
538.257[III](M+2H)+2

Cpd 1: C64 H66 N8 O8: +ESI Scan (0.133-0.316 min, 23 Scans) Frag=365.0V SH-JH-JH20A.d Subtract

x10 4

Counts vs. Mass-to-Charge (m/z)

538.2570|(M+2H)+2
538.7596|(M+2H)+2
539.2615|(M+2H)+2
539.7639|(M+2H)+2
540.2658|(M+2H)+2

Cpd 1: C64 H66 N8 O8: +ESI Scan (0.133-0.316 min, 23 Scans) Frag=365.0V SH-JH-JH20A.d Subtract

x10 5

Counts vs. Mass-to-Charge (m/z)

359.1750|(M+3H)+3
359.5096|(M+3H)+3
359.8439|(M+3H)+3
360.1754|(M+3H)+3
360.5127|(M+3H)+3

149

MO_MTOETPYP_1H_Methanol

150

PEL_1H_Methanol_2

Normalized Intensity

Chemical Shift (ppm)

153

Normalized Intensity

Chemical Shift (ppm)

154