Appendix III

Imparting Sequence Specificity on Non-Specific DNA Enzymes with Hairpin Polyamide Conjugates

Abstract

Many of the enzymes responsible for maintaining and manipulating DNA do not possess inherent sequence specificity. Because of this, complexes of these enzymes with DNA substrates have been notoriously difficult to crystallize for structure determination. Because these structures would be of great use to the biochemical community, we report here initial efforts towards using DNA-binding polyamides to impart sequence specificity on these non-specific enzymes. We have synthesized a library of maleimide-conjugated polyamides for use in trapping cystine-labeled topoisomerase II. We also report the synthesis of chlorambucil-conjugated polyamides for use in trapping various members of the helicase family of enzymes. Biochemical and crystallographic studies are currently ongoing in the Berger lab at the University of California, Berkeley.

Introduction.

Many important enzymes and proteins contact DNA in a sequence-independent fashion. Chief among these are enzymes responsible for altering the topology of DNA during transcription. One example is topoisomerase II, which aids in solving the topological problems associated with DNA replication, transcription, and chromatin remodeling.[1] Topoisomerase II accomplishes this by binding duplex DNA, and creating
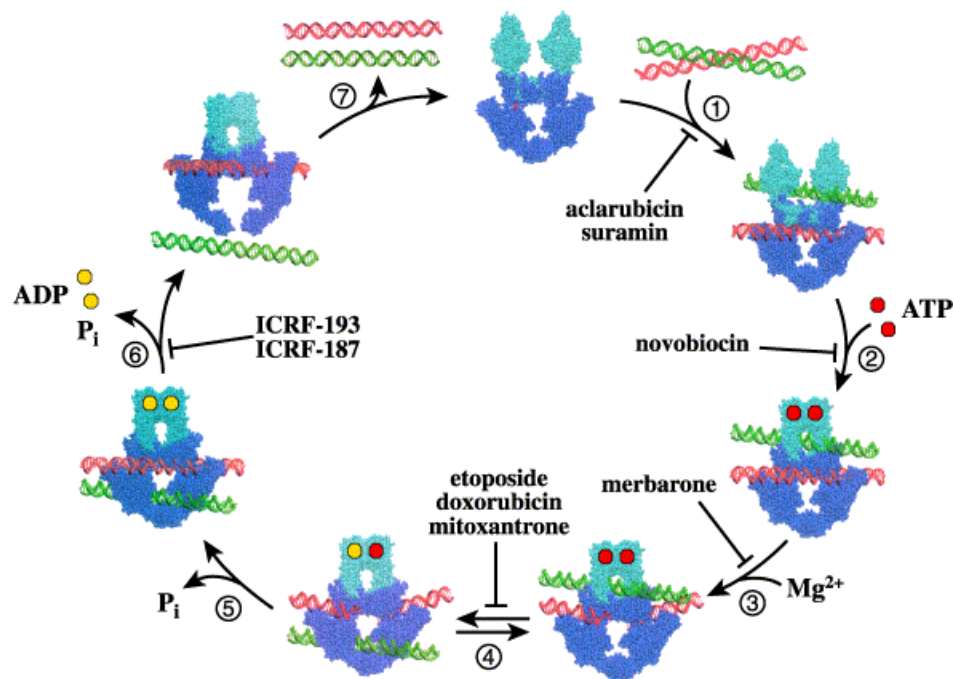
Figure III.1.[2] The catalytic cycle of topoisomerase II. 1 – The dimeric enzyme binds two strands of duplex DNA; 2, 3 – ATP is bound, and a double-strand break in one of the helices is formed; 4 – The intact strand is passed through the broken strand in an ATP-dependent fashion; 6, 7 – The cleaved strand is religated and the strands are released. Topoisomerase is an important cancer target, and many therapeutics have been developed to target various steps in this catalytic cycle (shown in the interior of the circle).

a double-strand break in the helix. The enzyme next passes a second double-helical region of DNA through this newly created break. Finally, the enzyme catalyzes the religation of the first helix (Figure III.1).[1, 3, 4] In this way, the enzyme helps relax DNA supercoils formed during transcription of circular DNA.

The helicase family of enzymes represents a second class of non-specific DNA-binding proteins. Helicases function as molecular motor proteins, moving unidirectionally along the double helix to unwind the energetically stable duplex DNA in an ATP-dependent fashion (Figure III.2).[5]
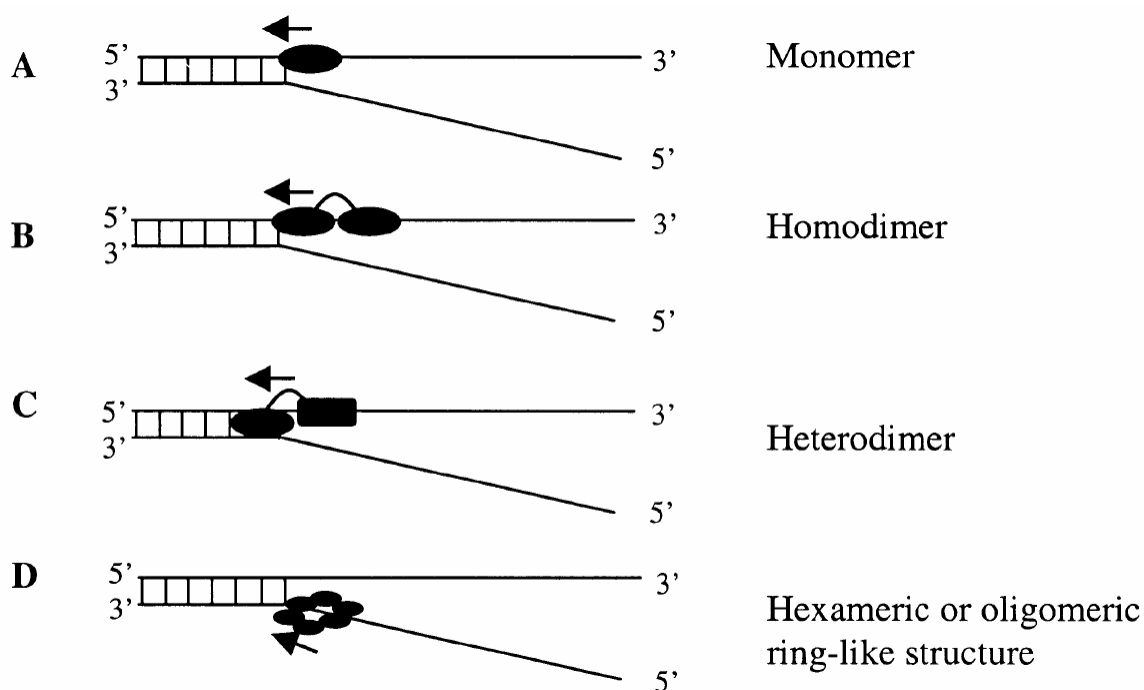
Figure III.2.[5] Schematic of the interaction of monomeric and oligomeric DNA helices with a forked DNA substrate. A – Monomeric helicase binds to both ssDNA and dsDNA. B – In homodimeric helicases, one subunit is always associated with the ssDNA along which it moves. C – Heterodimeric helicases contain a dsDNA-binding unit and an ssDNA-binding unit responsible for unwinding and translocation. D – Hexameric "ring-like" helicases encircle ssDNA, preventing local reannealing. One or more of the subunits binds the ssDNA/dsDNA junction.

Because topoisomerase II and helicase enzymes do not have any inherent sequence specificity, homogeneous crystals of the DNA:enzyme complexes have proven difficult to obtain. As a result, structural data on these complexes are currently unavailable.

Polyamides are small molecules capable of binding targeted sequences of DNA with high affinity and sequence specificity.[6] By conjugating protein-binding domains to

polyamides, researchers have been able to recruit proteins to targeted DNA sequences adjacent to binding sites for the polyamides.[7-10] Perhaps polyamides can be used to impart some sequence specificity on non-specific DNA enzymes, thereby leading to homogeneous populations of protein:DNA:polyamide complexes suitable for crystal growth and structure determination.
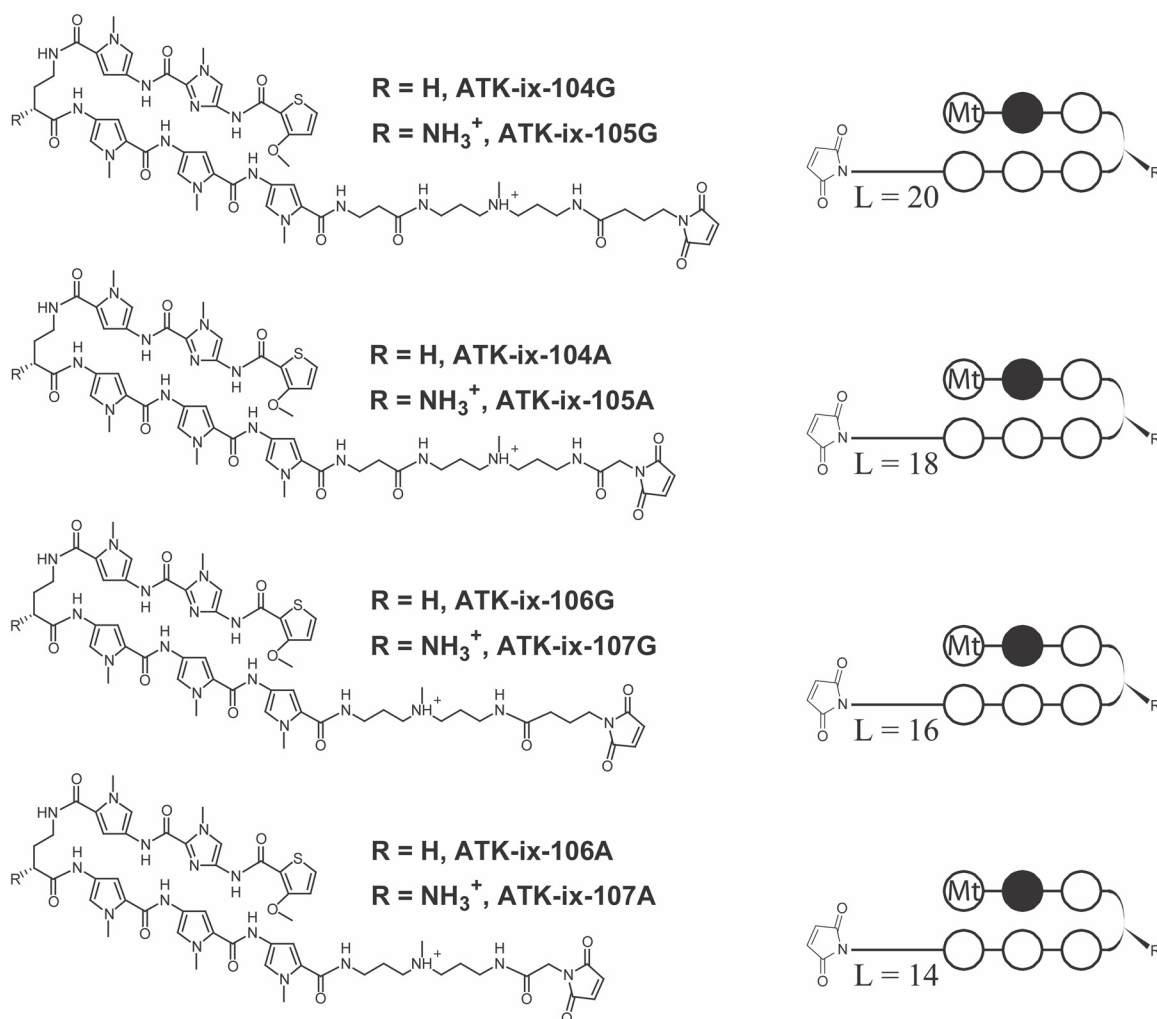

Topoisomerase II.

As shown in Figure III.1, topoisomerase II is thought to bind DNA as a dimer, with the two arms of the dimer fully encircling the double helix of DNA. Normally, this binding can occur at any sequence (although GC-rich regions are slightly preferred). In our efforts towards homogeneous topoisomerase II:DNA complexes for crystallization, we use sequence-specific polyamides to target topoisomerase II to a discrete site on a duplex template. Our experimental design is illustrated in Figure III.3.

Briefly, a series of topoisomerase II mutants will be engineered to contain a single reactive functionality at various sites flanking the DNA-binding domain. Polyamide conjugates containing a complementary reactive group will be targeted to sites flanking an optimal topoisomerase binding site. When both topoisomerase II and polyamide are bound, a covalent linkage will be created, transferring the sequence specificity of the polyamide to the enzyme, and effectively trapping the topoisomerase II enzyme at a site adjacent to the polyamide binding site. Because the newlyimparted sequence specificity should yield homogeneous populations of complexes, crystals will be grown, and structures determined.
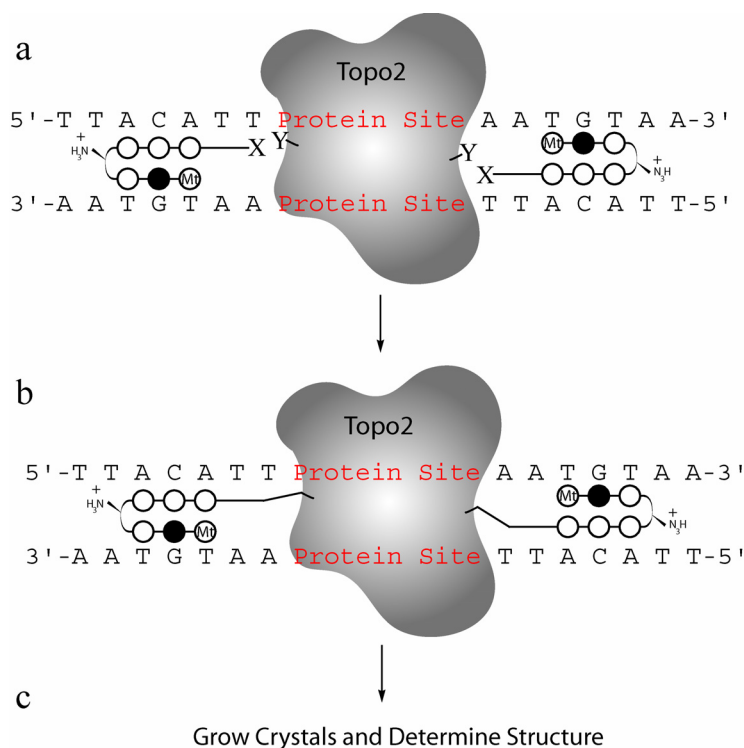
Figure III.3. Schematic representation of polyamide-mediated trapping of topoisomerase II. a – Polyamides functionalized with X and topoisomerase II functionalized with Y (where X and Y are complementary reactive groups) bind DNA at adjacent sites. b – Due to proximity, X and Y react, forming a covalent link that imparts the sequence specificity of the polyamide onto the topoisomerase II enzyme. c – Homogeneous complexes are then used for crystal growth and structure determination.

We have chosen a Michael addition between a cystine thiol residue on the protein and a maleimide functionality on the polyamide for initial attachment reactions. A series of mutant topoisomerase II enzymes will be created in which a single cystine mutation will be made at various residues flanking the DNA-binding domain (Figure III.5). Once these mutants have been tested to ensure the retention of their DNA-binding properties and function, they will be incubated with a DNA template and maleimide-functionalized polyamides. We chose to use the sequences 5'-CCG GT<u>T ACA T</u>T(G **GCC**)$_n$ **GAT CGG CCG ATC** (**GGC** C)$_n$A<u>A TGT A</u>AC CGG-3' for our studies. These sequences are palindromic, and contain a GC-rich domain for topoisomerase binding (bold), flanked by two polyamide binding sites (underlined). Because the precise distance between the

Figure III.4. Chemical and schematic structures of the polyamide-maleimide conjugates synthesized for topoisomerase II trapping. Distances from the DNA-binding portion to the reactive maleimide range from 14–20 atoms. Analogs of each compound were synthesized with and without the chiral turn (to modulate DNA-binding affinity). Shown below each schematic is the linker length in atoms (L = n).

polyamide and topoisomerase II binding sites optimal for conjugate addition is not known, several templates (n = 0, 1, 2, and 3) will be tested. The polyamides synthesized for thus study are illustrated in Figure III.4. We chose a polyamide directed towards the site 5'-WTGWW-3' (where W = A or T) because of the sequence's high AT content. This polyamide should thus not bind the GC-rich topoisomerase-binding region. The
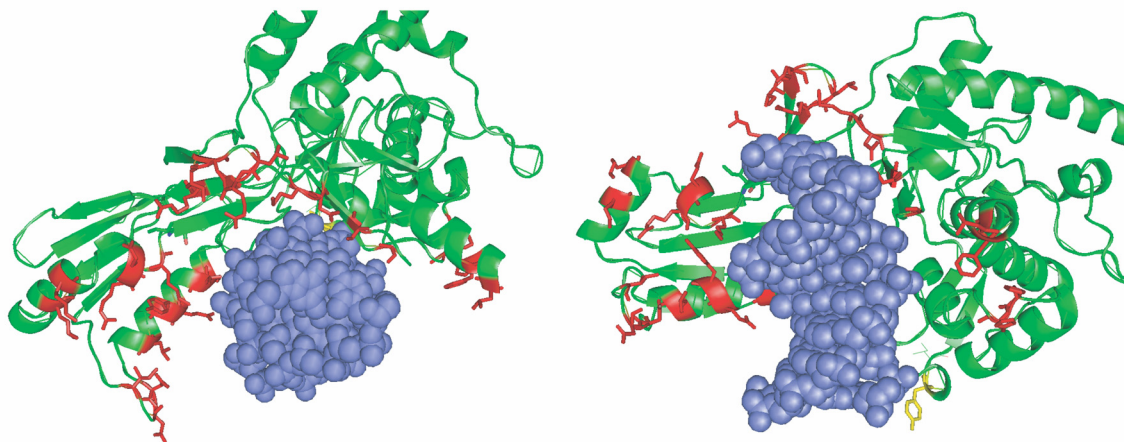
Figure III.5. Two views of a molecular model of the DNA-binding portion of topoisomerase II (green ribbons) complexed with duplex DNA (blue spheres). Highlighted in red are the residues chosen for cysteine mutations. Each mutant topoisomerase II will contain a single cysteine substitution at one of the residues highlighted in red. All residues chosen lay within 10 Å of the modeled DNA helix.

polyamide library consists of molecules with an identical polyamide core functionalized with several different linker lengths connecting the DNA-binding polyamide to the cystine-reactive maleimide. Analogs with and without the chiral turn were synthesized to give compounds with weak (no chiral turn) and strong (chiral turn) binding affinities. Compounds were sent to the Berger lab at the University of California, Berkeley, in November 2004.
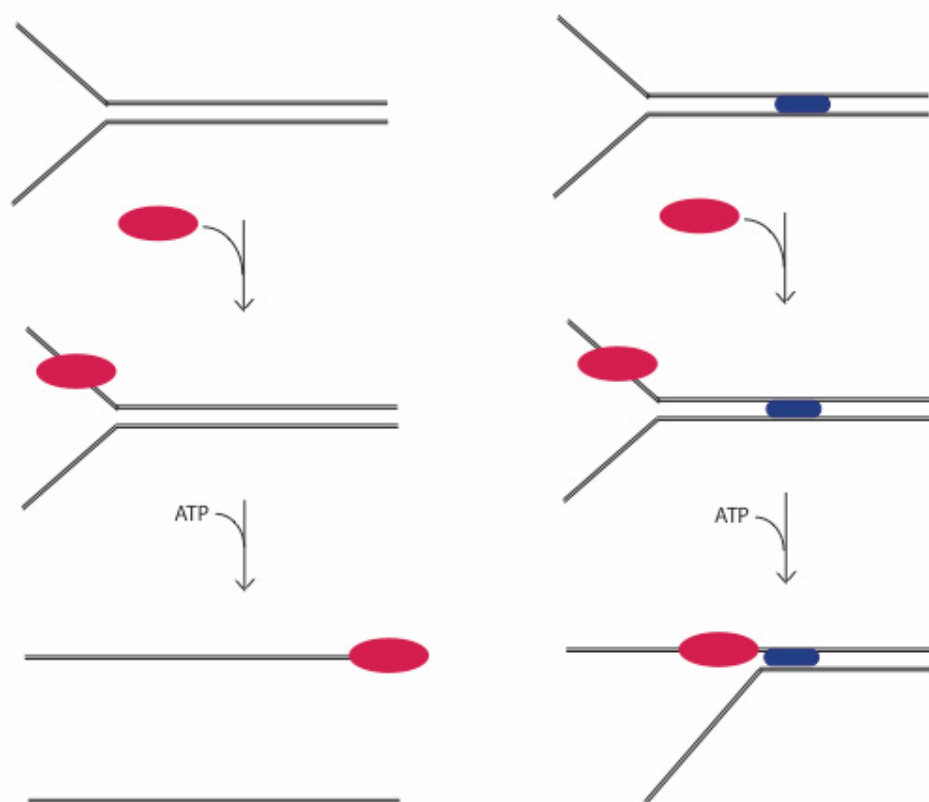
Helicase.

As shown in Figure III.2, helicases slide non-specifically along the double helix, melting the two strands as they progress. Polyamide-chlorambucil conjugates have been shown to create interstrand DNA crosslinks in a sequence-specific fashion (Figure III.6).[11, 12] In an effort to create homogeneous DNA:helicase complexes for crystallographic structure determination, we propose the use of polyamide-chlorambucil conjugates to act as molecular "chocks" to stop helicase progression at a predetermined site (Figure III.7). In such a scheme, polyamide-chlorambucil conjugates will be

a

5'-A T G G T C T A T-3'   5'-A T G G T C T A T-3'

3'-T A C C A G A T A-5'   3'-T A C C A G A T A-5'

b
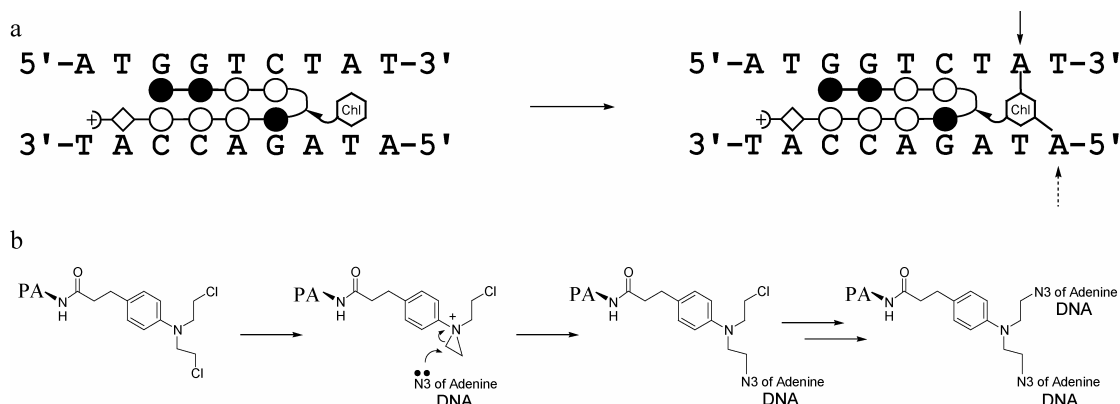
Figure III.6. a. Schematic of sequence-specific DNA crosslinking by a polyamide-chlorambucil conjugate.[12] At right, dark arrow indicates major alkylation site; dashed arrow indicates minor alkylation site. b. Mechanism of action of chlorambucil. The lone pair of electrons on N3 of adenine attacks the aziridine, forming a covalent bond. Because there are two chloroethyl functionalities, this reaction can occur twice to create an interstrand crosslink.

Figure III.7. Left column: Schematic of helicase activity on DNA. Helicase (red oval) binds to the ssDNA region of a fork template. The helicase then translocates in an ATP-dependent fashion to unwind the DNA. Right Column: Polyamide-chlorambucil conjugates will bind to a designed sequence in the dsDNA region of the fork template. The nitrogen mustard agent will then form an interstrand crosslink. Helicase will be added. The DNA will be unwound until the enzyme reaches the polyamide "chock," where it will be unable to continue. Homogeneous complexes will then by used for crystal growth and structure determination.

incubated with a DNA template for helicase unwinding. The two DNA strands will be crosslinked at a single site adjacent to the polyamide binding site. Helicase enzymes will then be allowed to progress towards the polyamide "stopper." Upon reaching the crosslink site, progression will be arrested. Because progression is arrested at a discrete site (dependent upon polyamide binding), homogeneous populations of complexes should result, which will then be used for structure determination.
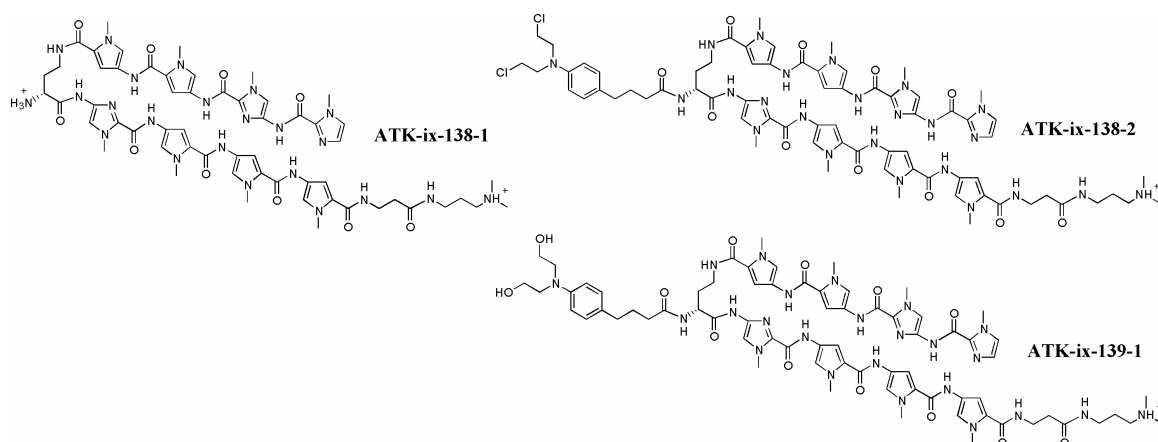
Figure III.8. Chemical structures of the polyamides synthesized for helicase trapping studies. Compound **ATK-iix-138-2** is the chlorambucil-containing polyamide. **ATK-iix-138-1** and **ATK-iix-139-1** are negative control compounds where the chlorambucil has been removed and hydrolyzed, respectively.

Because the DNA construct for this study is artificial, we chose to use the polyamide core ImImPyPy-γ-ImImPyPy-β-Dp for crosslinking studies. This polyamide has previously proven to bind DNA with high affinity and excellent specificity.[13] Because our main concern is homogeneity in the site of helicase arrest, this highly specific polyamide was chosen. Polyamide conjugates shown in Figure III.8 have been synthesized and sent to the Berger Lab, at the University of California, Berkeley, as of

April 2005. The polyamides will be tested against a library of helicases, including MCM, DnaB, and T7 phage helicase.

Conclusion.

Polyamide conjugates have been designed and synthesized to impart sequence specificity onto non-specific DNA enzymes in order to create homogeneous complexes that are able to be crystallized for structural determination. Topoisomerase II and helicase enzymes are the targets of the present research. Because no precise structures are currently available with these enzymes in complex with their DNA substrates, they would be of great use to the field.

Materials and Methods.

Polyamides were synthesized according to standard procedures on either PAM or oxime resin.[14, 15]   Polyamides **ATK-iix-104 – 107** were synthesized using the methoxythiophene cap as previously reported.[16]

Sample procedure for the synthesis of maleimide conjugates.

Base polyamides were cleaved from either oxime or PAM resin with methylamino dipropylamine and purified by preparative HPLC. 1 μmol polyamide amine was combined in 100 μL 10:1 DMF:DIEA with 5 μM of maleimide NHS esters (Pierce). The reaction was monitored by analytical HPLC, and when done, purified by HPLC. Typical yields after resin cleavage were ~ 70%.

ATK-iix-104A UV (H$_2$O) λ_{max} 310 nm (51540). MALDI-TOF-MS calcd. (M + H): 1191.3. Found 1191.3.

ATK-iix-104G UV (H$_2$O) λ_{max} 310 nm (51540). MALDI-TOF-MS calcd. (M + H): 1219.4 Found 1219.3.

ATK-iix-105A UV (H$_2$O) λ_{max} 310 nm (51540). MALDI-TOF-MS calcd. (M + H): 1207.3 Found 1207.8.

ATK-iix-105G UV (H$_2$O) λ_{max} 310 nm (51540). MALDI-TOF-MS calcd. (M + H): 1235.4. Found 1235.5.

ATK-iix-106A UV (H$_2$O) λ_{max} 310 nm (51540). MALDI-TOF-MS calcd. (M + H): 1120.2. Found 1120.1.

ATK-iix-106G UV (H$_2$O) λ_{max} 310 nm (51540). MALDI-TOF-MS calcd. (M + H): 1148.3. Found 1148.1.

ATK-iix-107A UV (H$_2$O) λ_{max} 310 nm (51540). MALDI-TOF-MS calcd. (M + H): 1136.2. Found 1136.2.

ATK-iix-107G UV (H$_2$O) λ_{max} 310 nm (51540). MALDI-TOF-MS calcd. (M + H): 1164.3. Found 1164.6.

ATK-iix-138-2

Polyamide **ATK-iix-138-1** was liberated from resin by aminolysis in neat dimethylamino propylamine (Dp) for 12 hours at 50 ºC and purified by preparative HPLC. **ATK-iix-138-1** (2 μmol) was dissolved in 300 μL DMF and 10 μL DIEA. Chlorambucil (4 μmol, 1.33 mg) was added, followed quickly by PyBOP (3.9 μmol, 2.04 mg). The reaction was mixed, and allowed to stand at room temperature for 45 minutes. The reaction was then diluted with 0.1% TFA in water and purified by reversed-phase HPLC. To avoid hydrolysis, as soon as the product eluted from the column, 500 μL aliquots were freeze-dried. Product was isolated in 50% yield. UV (H_2O) λ_{max} 310 nm (68720). MALDI-TOF-MS calcd. (M + H): 1525.5. Found 1525.7.

ATK-iix-139-1

Polyamide **ATK-iix-138-2** (500 nmol) was dissolved in 100 μL DMF. 1 mL of 25% NaOH in water was added, and the reaction mixed at room temperature for four hours. The reaction was acidified with TFA, diluted to 10 mL with water, and purified by reversed-phase HPLC (200 nmol yield).

References.

1. Champoux, J. J., *Annual Review of Biochemistry*, **2001,** *70*, 369–413.

2. Larsen, A. K.; Eseargueil, A. E.; Skladanowski, A., *Pharmacology & Therapeutics*, **2003,** *99*, 167–181.

3. Baird, C. L.; Gordon, M. S.; Andrenyak, D. M.; Marecek, J. F.; Lindsley, J. E., *Journal of Biological Chemistry*, **2001,** *276*, 27893–27898.

4. Fass, D.; Bogden, C. E.; Berger, J. M., *Nature Structural Biology*, **1999,** *6*, 322–326.

5. Tuteja, N.; Tuteja, R., *European Journal of Biochemistry*, **2004,** *271*, 1849–1863.

6. Dervan, P. B.; Edelson, B. S., *Current Opinion in Structural Biology*, **2003,** *13*, 284–299.

7. Ansari, A. Z.; Mapp, A. K.; Nguyen, D. H.; Dervan, P. B.; Ptashne, M., *Chemistry & Biology*, **2001,** *8*, 583–592.

8. Arndt, H. D.; Hauschild, K. E.; Sullivan, D. P.; Lake, K.; Dervan, P. B.; Ansari, A. Z., *Journal of the American Chemical Society*, **2003,** *125*, 13322–13323.

9. Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B., *Proceedings of the National Academy of Sciences of the United States of America*, **2000,** *97*, 3930–3935.

10. Wang, C. C. C.; Dervan, P. B., *Journal of the American Chemical Society*, **2001,** *123*, 8657–8661.

11. Wang, Y. D.; Dziegielewski, J.; Wurtz, N. R.; Dziegielewska, B.; Dervan, P. B.; Beerman, T. A., *Nucleic Acids Research*, **2003,** *31*, 1208–1215.

12. Wurtz, N. R.; Dervan, P. B., *Chemistry & Biology*, **2000,** *7*, 153–161.

13. Poulin-Kerstien, A. T. DNA-Templated Dimerizations of Minor Groove-Binding Polyamides. California Institute of Technology, Pasadena, **2005**.

14. Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society*, **1996,** *118*, 6141–6146.

15. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., *Bioorganic & Medicinal Chemistry*, **2002,** *10*, 2767–2774.

16. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B., *Bioorganic & Medicinal Chemistry*, **2003,** *11*, 4333–4340.