Chapter 1

Introduction

DNA Structure and Function

Deoxyribonucleic acid (DNA) is a biological macromolecule that contains genetic information encoding for the entire structure and function of all known organisms. Various biological characteristics are accounted for by individual genes made up of specific sequences of DNA that together provide a blueprint for life, the full set of which is called a genome. DNA's primary structure consists of a backbone of alternating 2-deoxyribose sugar units and phosphate residues, which are joined by phosphodiester bonds at the 3' and 5' positions of the sugar. Attached to each sugar at the 1' position is one of four nucleobases: adenine (A), cytosine (C), guanine (G), or thymine (T). As it typically exists in nature, double-stranded DNA (dsDNA) contains two antiparallel strands of DNA in a double helix, forming a series of specific hydrogen bonding interactions between complementary nucleobases: A and T are specific for each other, as are C and G.

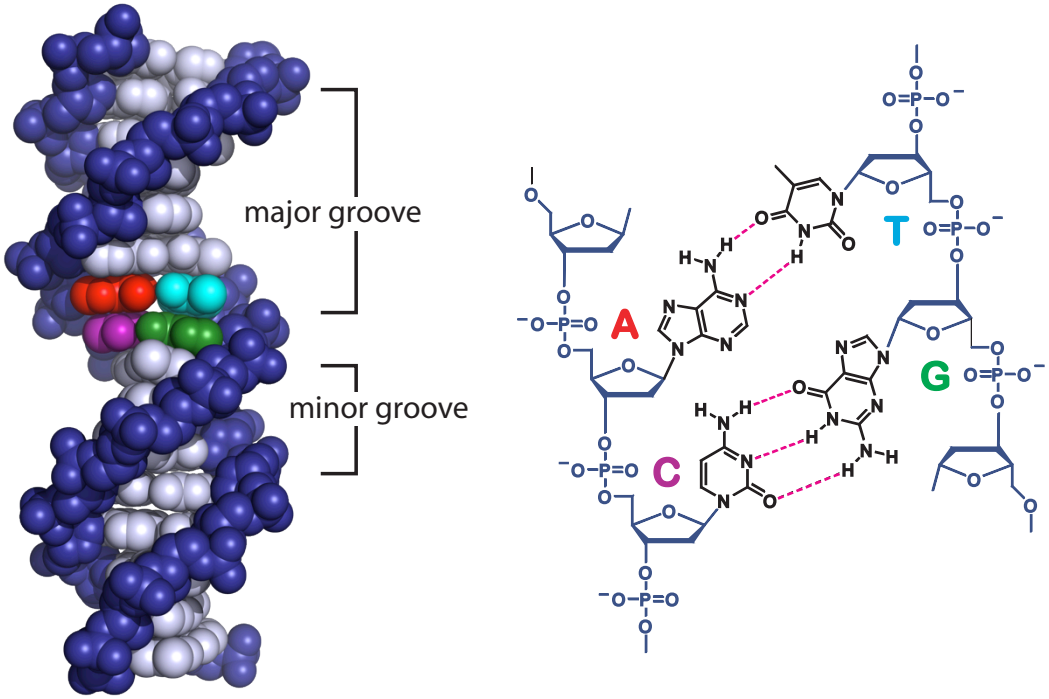


Figure 1.1: The 3-D structure of helical dsDNA[1] showing the backbone in blue and base pairs in gray, along with the minor and major grooves indicated and a set of base pairs highlighted (left), and the specific pairing of nucleobases attached to the sugar-phosphate DNA backbone (right). The two hydrogen bonding interactions between A and T and the three interactions between C and G are indicated in pink.

The Central Dogma of Molecular Biology

The term "central dogma" was first coined in 1958 by Walter Crick who later described it as dealing with "the detailed residue-by-residue transfer of sequential information [and] such information cannot be transferred from protein to either protein or nucleic acid."[2] Proteins are a type of biomacromolecule that consist of chains of linked amino acids that (depending on the sequence) can perform a variety of functions within organisms. Proven over many years, the basic concept behind the central dogma is that DNA serves as the genetic code, storing information in the form of genes that eventually lead to the production of specific proteins, but proteins can never transmit sequence information to lead directly back to the corresponding DNA. Encompassed in this idea are the processes of transcription and translation, which together provide the information transfer from DNA to protein. Transcription is the process that copies DNA sequence information into corresponding ribonucleic acid (RNA) chains. RNA is structurally similar to DNA, except for the presence of a 2' hydroxyl group on the sugar moiety. There are a variety of types of RNA that can perform different functions, but in particular messenger RNA (mRNA) can allow for the production of proteins. In this process, called translation, sequence information is transferred from mRNA into a corresponding protein with a specific sequence of amino acids. Because proteins serve so many functions and misregulated protein expression is a factor in a variety of diseases, strategies to disrupt the pathway from DNA to RNA to protein for specific genes or subsets of genes may be of significant therapeutic value.

Transcription Factors and Cellular Signaling

Transcription factors (TFs) are a type of protein that can control the levels of transcription of a given gene, and often sets including multiple genes. This process is typically accomplished by non-covalent binding to a specific DNA sequence adjacent to a target gene, and either blocking (repression) or promoting (activation) the recruitment of RNA polymerase, the main enzyme responsible for transcription. Natural transcription factors vary in structure but typically consist of proteins bearing multiple domains that each serve specific functions such as a DNA-binding domain (DBD), a transcriptional activation domain (TAD), or a signal sensing domain (SSD). Transcription factors can bind to either the wider major groove or narrower minor groove of DNA, but in most cases, specific DNA-binding occurs via interaction between the DBD and 4-8 DNA nucleotides in a sequence specific fashion. These binding events can involve hydrogen bonding, electrostatic interactions, or van der Waals interactions with the base pairs or phosphate backbone of DNA. Proteins that bind to DNA, such as leucine zippers and zinc fingers, can bind as monomers, dimers, or in a combinatorial fashion where multiple DNA-binding proteins are required to regulate a certain gene.

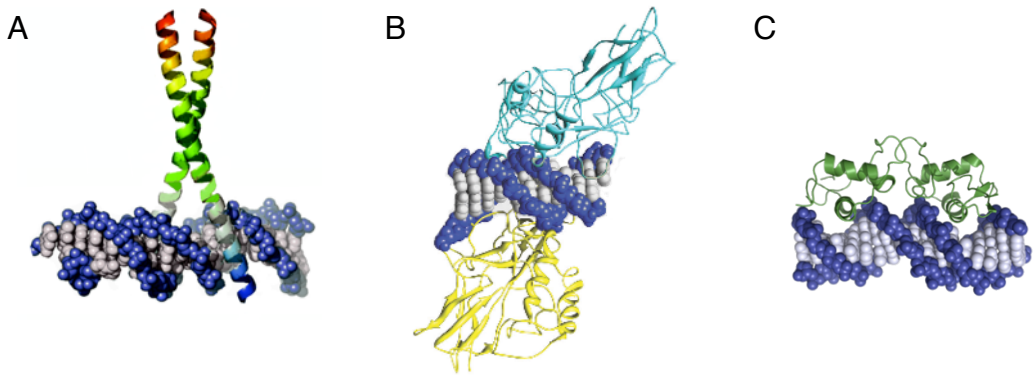


Figure 1.2: Structures of DNA-binding proteins; A) GCN4, a homodimeric leucine zipper; B) NF-κB, a heterodimer of p50/p65; c) Androgen Receptor, which includes a ligand binding domain.

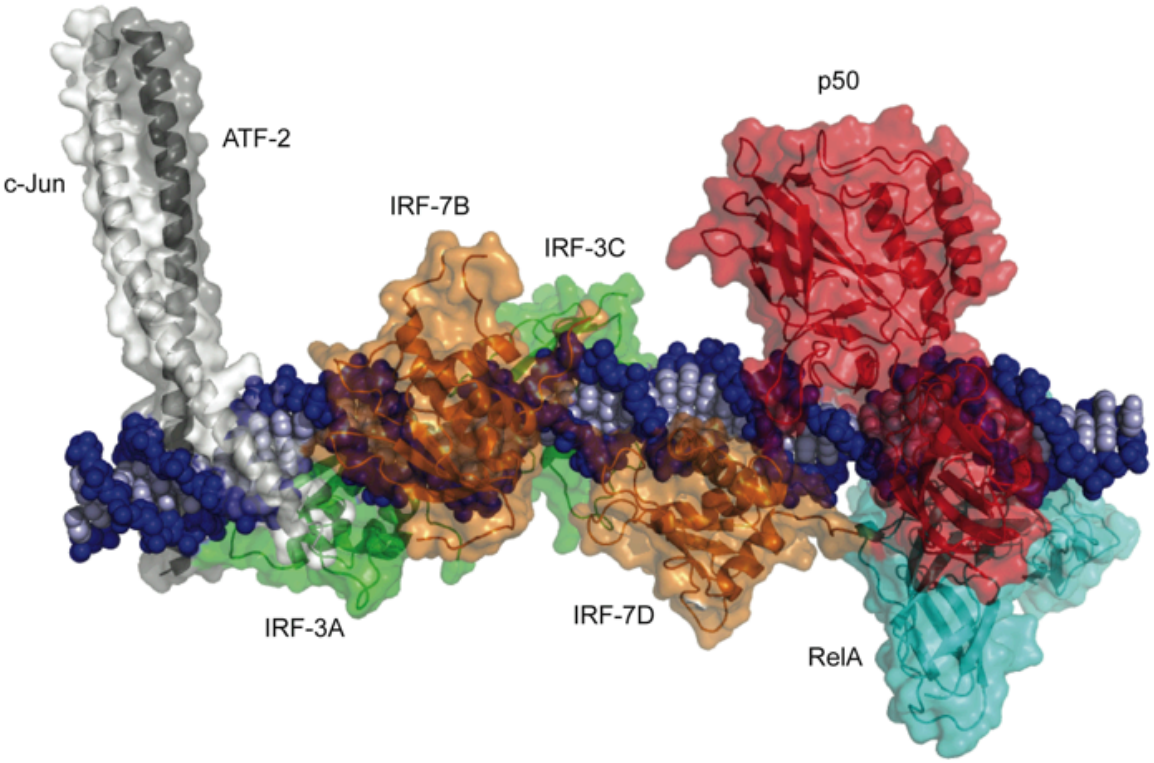
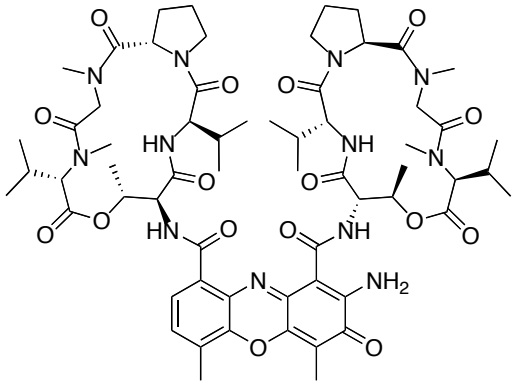


Figure 1.3: Structure of the interferon-ß enhanceosome[3], an example of combinatorial binding of a variety of proteins. This cooperative assembly occurs without any protein-protein contacts.

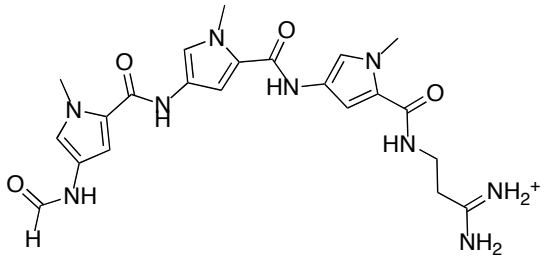
Nature has evolved complex networks of protein interactions called signaling pathways that can respond to external stimuli or allow for communication between cells. These pathways control processes crucial for the proper functioning of an organism, such as differentiation, stress response, cell growth, neurological function, metabolism, etc. Transcription factors are often major players in cellular signaling, so modulation of transcription can often cause major downstream (later in signaling pathway) effects or can counteract upstream (earlier in signaling pathway) effects. Therefore, disruption of transcription factor function is a potentially powerful strategy for developing therapeutics with many possible applications.

DNA-Binding Small Molecules

In addition to DNA-binding proteins, there are also many known small molecules capable of binding to DNA in a sequence specific fashion, most of which have been isolated as natural products. There are two major modes of non-covalent small molecule DNA-binding: intercalation and groove binding. Intercalation involves planar aromatic molecules which lodge between DNA base pairs, taking advantage of ring-stacking stabilization interactions with the adjacent nucleobases. In either major or minor groove binding, a variety of structures are possible, all of which take advantage of the unique electronic and steric environment that exists within the groove and vary with different sequences. An example of each binding mode are shown in Figure 1.4: actinomycin D, which preferentially binds to DNA by intercalation at 5'-GC-3' sequences;[4] and distamycin A, which targets AT tracts of DNA by binding to the minor groove.[5]



Actinomycin D
5'-GC-3'



Distamycin A
5'-$(A,T)_5$-3'

Figure 1.4: Structures of actinomycin D, an intercalator that targets 5'-GC-3', and distamycin A, a minor groove binder that targets 5'-$(A,T)_5$-3' sequences.

Recognition of the DNA Minor Groove

In the exposed portion of the minor groove, the plane of the base paired rings is perpendicular to the binding molecule, such that one side of the base pairing interaction is accessible (while the opposite side is accessible in the major groove). This imparts a unique electronic environment depending on the sequence of nucleotides, and a specific pattern can therefore be mapped to describe the electronic environment displayed by the minor groove of a specific sequence, as in Figure 1.5B.

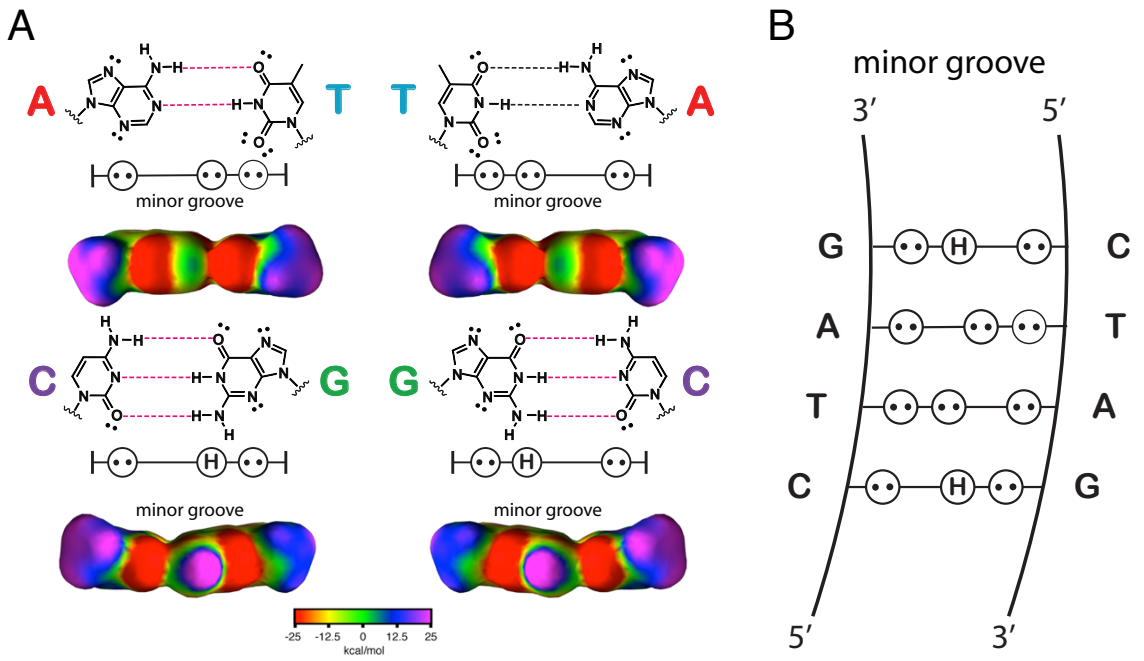


Figure 1.5: The electronic environment of the minor groove of DNA. A) For A•T base pairs, A displays one lone pair while T displays two lone pairs, whereas in C•G base pairs C displays one lone pair while G displays one lone pair and one N-H bond from the exocyclic amine. The unique electronic environments for each are clear from the corresponding electron density maps. B) The different patterns of lone pairs and N-H bonds displayed by a certain sequence can be mapped to provide the unique pattern of electron density for that sequence.

In the case of distamycin, the x-ray crystal structure of its interaction with DNA has been studied in detail.[6,7] However, it was found that distamycin binds DNA in two different stoichiometries: 1:1 binding with DNA or 2:1 distamycin:DNA binding. Interestingly, crystal structures revealed hydrogen bond contacts between the amide N-H bonds of

distamycin and electron lone pairs of exposed lone pairs from base pairs exposed in the minor groove. In the 2:1 binding mode, the distamycin molecules bind in an antiparallel fashion, and each of the amide N-H bonds forms a favorable hydrogen bond with a lone pair on the same side of the minor groove.

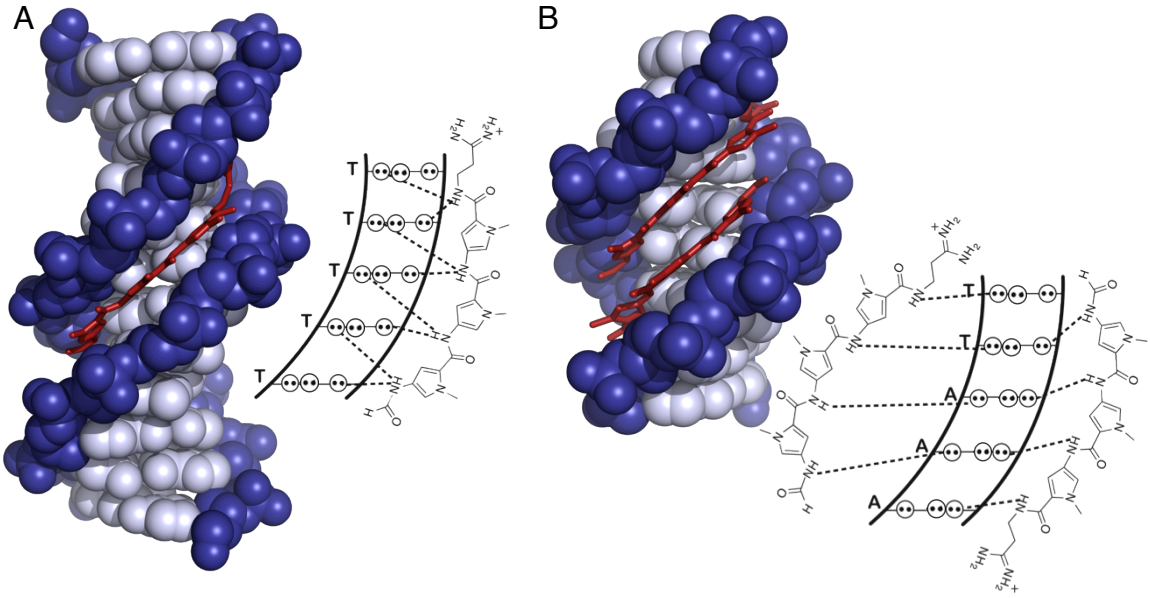


Figure 1.6: Recognition of the DNA minor groove by distamycin in A) a 1:1 complex and B) a 2:1 antiparallel complex.

This structural information from these crystal structures seem to indicate a favorable shape complementarity between distamycin and DNA, allowing for two molecules to bind the minor groove of DNA, each with monomeric pyrrole heterocyclic amino acids units oriented across from each other such that hydrogen bond interactions occur with a single base pair of DNA matched to a single pyrrole monomer unit on each molecule of distamycin.

Pyrrole-Imidazole Polyamides

Based on structure of distamycin, a new class of oligomeric heterocyclic DNA-binding small molecules called pyrrole-imidazole (Py-Im) polyamides have been developed that can be programmed to bind to specific sequences of DNA. In particular, by replacing a pyrrole subunit with an imidazole heterocycle, sequences including G•C or C•G base pairs can be targeted, and incorporation of a hydroxypyrrole heterocycle can distinguish between T•A and A•T base pairs.[8,9] A set of pairing rules has been developed that allows for programmable targeting specific DNA sequences: N-methylpyrrole/N-methylimidazole (Py/Im) targets C•G, Im/Py recognizes G•C, 3-hydroxy-N-methylpyrrole/N-methylpyrrole (Hp/Py) and Py/Hp target T•A and A•T respectively.[10] Finally, Py/Py pairs are degenerate for T•A and A•T.

In addition to the expanded range of sequences that can be targeted, a γ-aminobutyric acid (GABA) "turn" unit has been developed, to form a "hairpin" polyamide that provides a favorable entropic effect.[11] This structural motif has been shown to provide a 100-3600-fold increase in binding affinity for a target DNA sequence relative to the corresponding 2:1 motif.[12,13] Additionally, a GABA turn unit can allow for the targeting of non-palindromic DNA sequences by incorporating unsymmetrical ring pairings. Further development of a chiral amino substituted GABA turn allows for binding in a single orientation, where a polyamide N- to C- terminus aligns with DNA in a 5' to 3' direction, known as "forward" binding.[14,15]

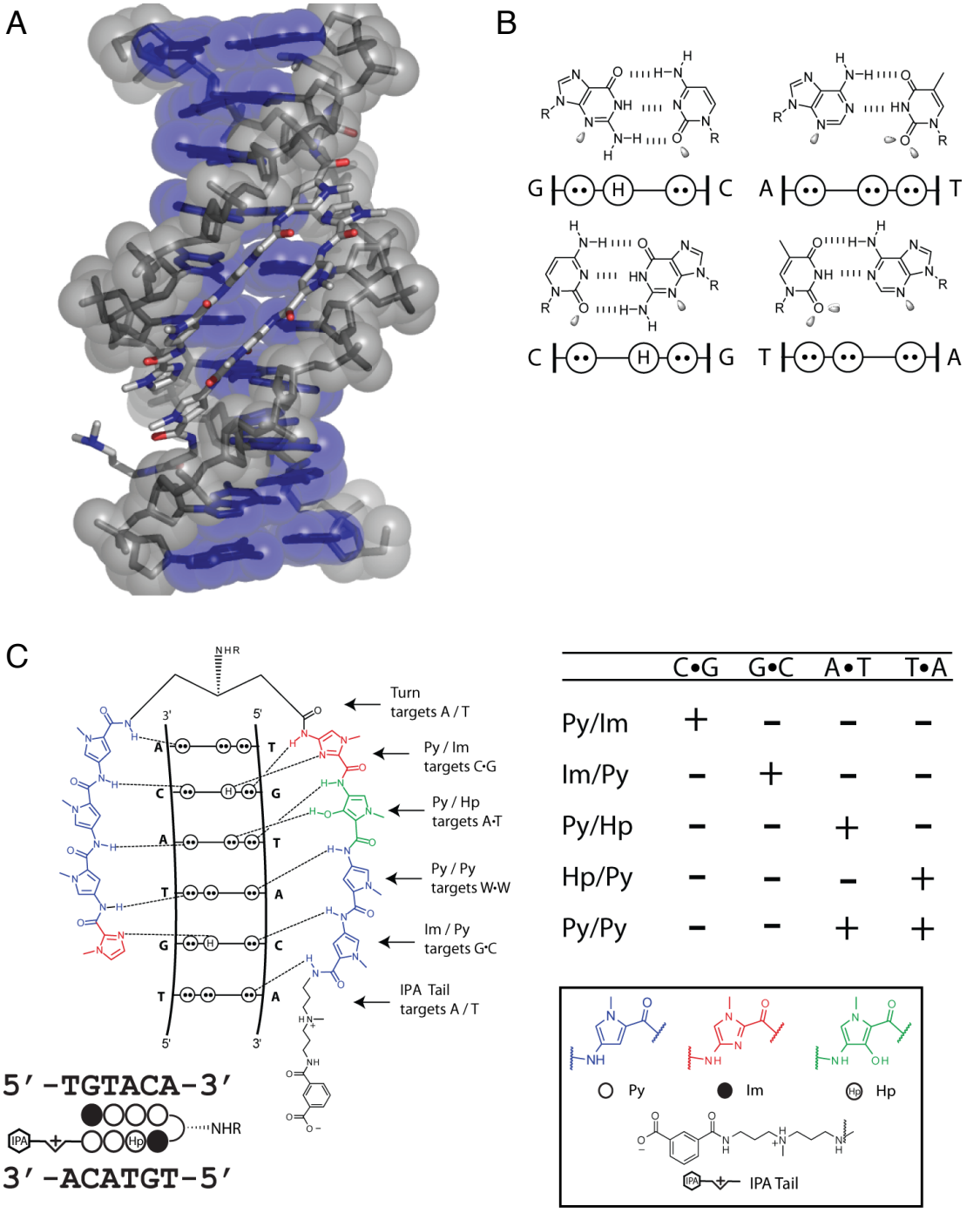


Figure 1.7: Recognition of the DNA minor groove by Py-Im polyamides. A) The structure of a Py-Im polyamide bound to DNA, showing clear ring stacking and shape complementarity. B) Electronic environment created by each base pair in the minor groove. C) Pairing rules for targeting all four base pairs with Py-Im polyamides. Ball-and-stick notation (alongside abbreviations) is also defined here, which is used as standard nomenclature for Py-Im polyamides.

As seen Figures 1.7 and 1.8, the Im subunit is specific for G nucleotides due to a hydrogen bonding interaction between the lone pair of the heterocyclic nitrogen of the Im and the N-H bond of the exocyclic amine of guanine. Hp monomers are specific for T nucleotides based on an interaction between the O-H bond of Hp (that is sterically accommodated by the cleft of the T•A base pair) and the extra lone pair displayed in the minor groove by thymine. Finally, the hairpin turn and the "tail" region at the C-terminus both target either the A•T or the T•A base pairs adjacent to the ends of the core of heterocycles, meaning that 8-heterocycle Py-Im polyamides targets six DNA base pairs.

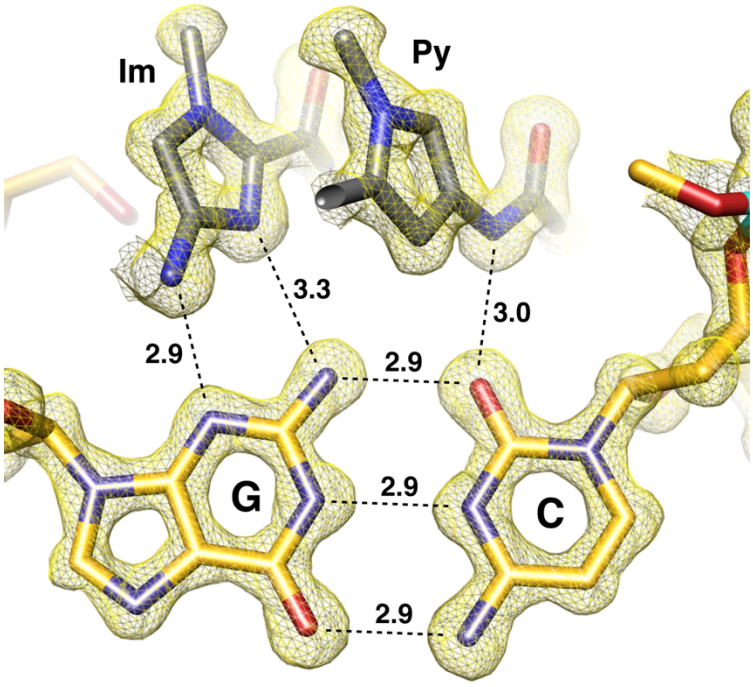


Figure 1.8: Crystal structure showing the orientation and hydrogen bonding interactions of an Im-Py heterocycle pairing adjacent to a G•C base pair.[16] The interactions between amide N-H bonds and DNA lone pairs are clearly visible, as well as between the lone pair of the heterocyclic nitrogen and the N-H bond of the exocyclic amine of guanine. Also indicated are the three hydrogen bonds formed by the G•C base pair.

Development of Py-Im Polyamides with Therapeutic Potential

For Py-Im polyamides to be viable in a biological context, it is critical that they are able to enter live cells, and bind to DNA inside the cell nucleus. Previous cell culture studies have indicated that many fluorophore-labeled Py-Im polyamides are able to enter a variety of live cell lines, and traffic to the nucleus.[17,18] In addition, crystal structures of Py-Im polyamides bound to nucleosome core particles demonstrated the ability of these compounds to bind to DNA in the context of chromatin, causing structural changes in nucleosomal DNA without significantly perturbing the structure of the histone octamer.[19]

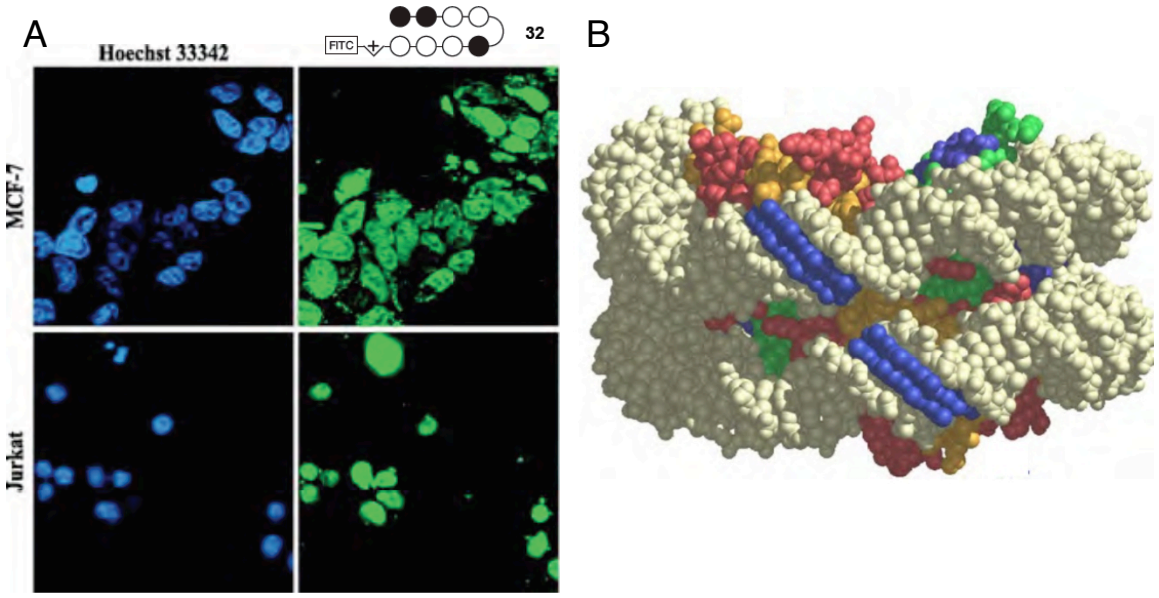


Figure 1.9: A) Py-Im polyamides can often enter live cells. In this experiment, a fluorescently labeled (FITC) polyamide is added to two live cell lines, each incubated with a nuclear stain (Hoechst 33342, left panels). In both cell lines, clear co-localization occurs.[20] B) Crystal structure of a nucleosome core particle, showing DNA-bound polyamide in blue.

Py-Im polyamides with isophthalic acid moieties attached at the tail region have also shown enhanced cell permeability properties, leading to a class of cell-permeable, sequence-specific, and high-affinity DNA-binding small molecules.[21]

Regulation of Gene Expression with Py-Im Polyamides

Due to the favorable biological properties of Py-Im polyamides and the power of modulating of gene expression for therapeutic purposes, the potential for using polyamides to repress or activate transcription is an attractive approach. In addition, the modular sequence-specificity of polyamides suggests that polyamides can be tailored to a variety of different biological targets for various therapeutic applications. Regulation of transcription could occur by two main routes: (1) repression of transcription by interfering with the function of existing transcription factors or (2) activation of transcription factor through conjugation of a DNA-binding polyamide to a TAD.

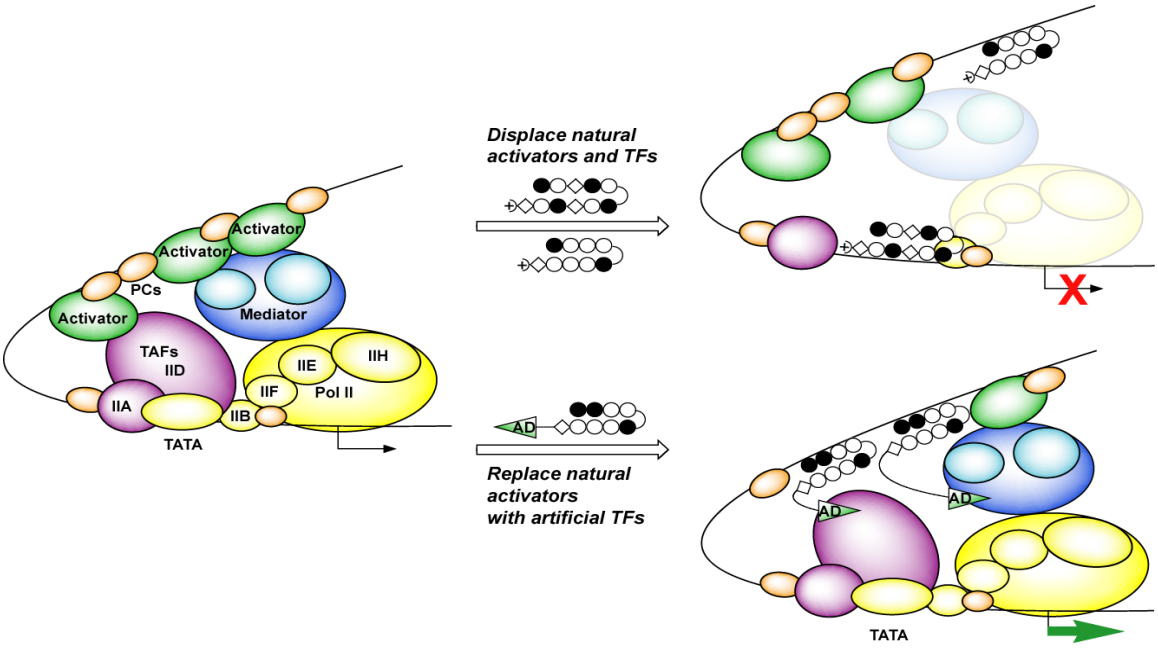


Figure 1.10: The two routes of modulating transcription. In the case of repression (top), a polyamide targeted to the binding site of a transcription factor may be able to interfere with the TF-DNA interaction, limiting the ability of the TF to recruit RNA polymerase and initiate transcription. For activation to occur (bottom), an activation domain could be attached to a polyamide targeted adjacent to a gene of interest, causing an artificial activation of transcription if the polyamide-TAD complex can properly recruit RNA polymerase.

While the strategy of interfering with the proper functioning of a transcription factor is a promising strategy for repression of transcription, there initially remained the question of exactly how a Py-Im polyamide could achieve this goal. Because many

transcription factors bind to the major groove of DNA, it is not immediately obvious how the process of a polyamide disrupting transcription factor function would work. A possible model for this to occur is evident when examining the crystal structure of polyamide-DNA complexes.

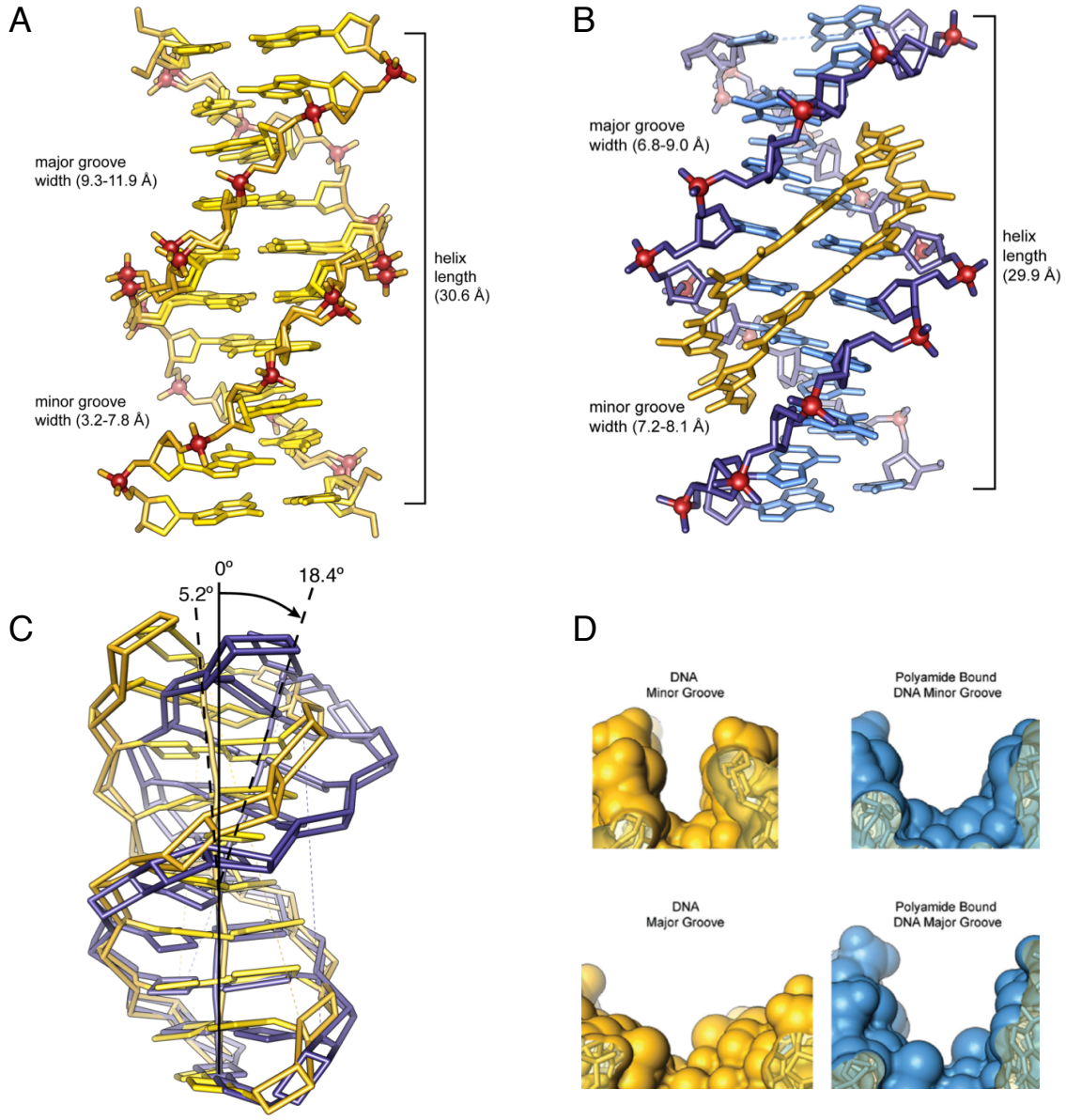


Figure 1.11: The allosteric effect of Py-Im polyamide-DNA binding.[16] A) The structure of DNA with a narrow minor groove and wide major groove. B) When DNA is in a complex with a polyamide, significant widening of the minor groove occurs in order to accommodate the polyamide, causing a concomitant narrowing of the major groove. C) In addition to the effect on groove shape and width, the DNA strand becomes bent along the helical axis upon polyamide binding. D) An alternate view of the size and shape of the major and minor groove before and after polyamide binding.

As seen in Figure 1.11, significant deformation of the structure of DNA in the vicinity of the polyamide occurs upon binding. It is believed that this allosteric effect is the molecular basis for how polyamides binding to the minor groove of DNA can cause a disruption of the transcription factor-DNA interface for major groove-binding transcription factors.

One of the initial transcription factors targeted for transcriptional downregulation by polyamides was hypoxia inducible factor 1α (HIF-1α), which activates a set of genes in response to low oxygen environments by binding to the Hypoxic Response Element (HRE) consensus sequence 5'-TACGTG-3'. One of these genes is vascular endothelial growth factor (VEGF), which when activated can lead to tumor vascularization. In an induced system, a polyamide targeted to the HRE (Figure 1.12A) was demonstrated to repress VEGF expression via qRT-PCR.[22,23] In contrast, treatment with a mismatch polyamide that does not bind the HRE did not significantly affect VEGF expression. On a global scale, microarray analysis indicated that the match compound downregulated a subset of genes activated by induction of HIF-1α mediated expression. Chromatin immunoprecipitation (ChIP) experiments further demonstrated a decrease in occupancy of HIF-1α at the VEGF HRE following treatment with the match compound, supporting a sequence-dependent mechanism.

Similar studies were conducted for transcription factors such as: (1) androgen receptor,[24] which drives the expression of a set of genes crucial in the development and progression of prostate cancer (Figure 1.12B), (2) glucocorticoid receptor,[25] which is a major regulator of inflammatory response (Figure 1.12C), and (3) NF-κB,[26] which plays a

role in imflammatory response, cell differentiation, and senescence, and its misregulation has been implicated in cancer (Figure 1.12D).

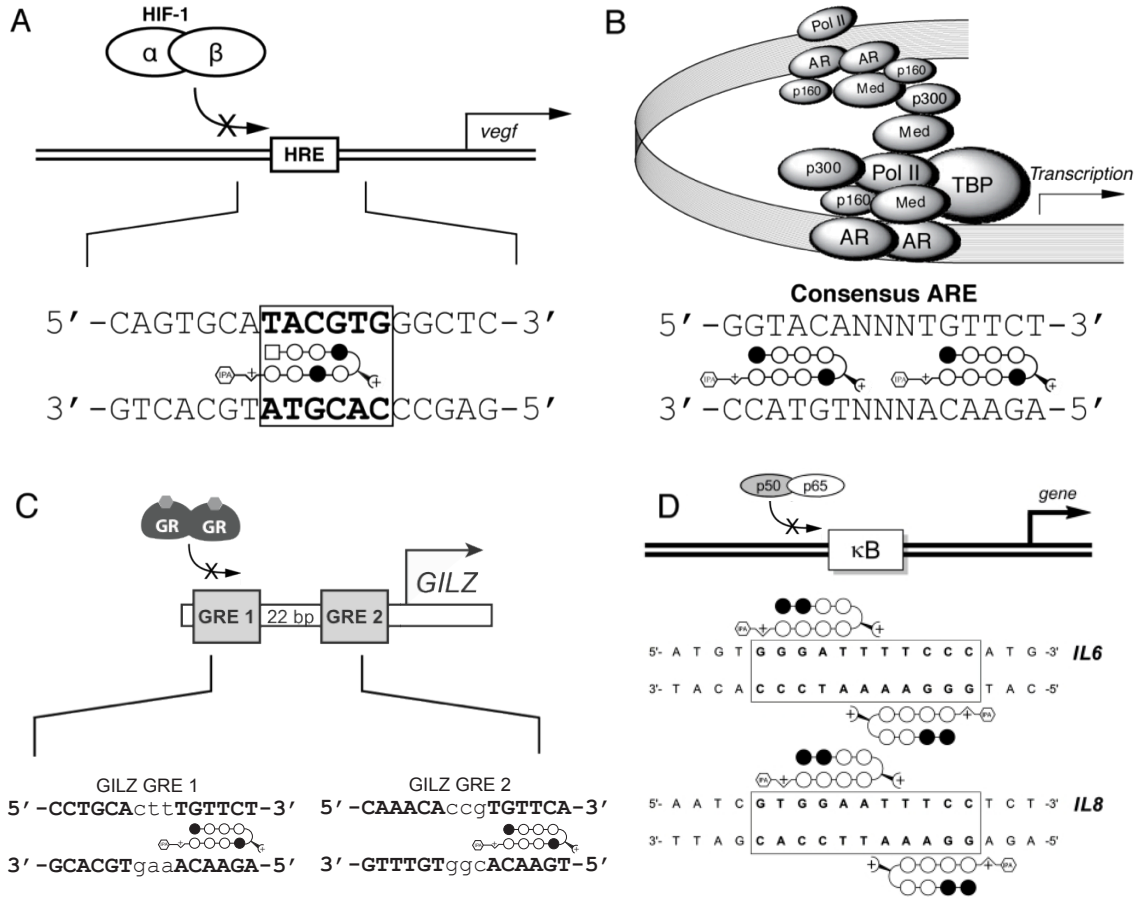


Figure 1.12: Various transcription factors that have been targeted with Py-Im polyamides, for which reduction of TF-mediated transcription has been demonstrated. In each case, the consensus sequence that was targeted is shown, along with the corresponding polyamide. A) Hypoxia inducible factor 1α, B) Androgen receptor, C) Glucocorticoid receptor, D) Nuclear factor κB.

Py-Im Polyamide-mediated Gene Activation

In contrast to repression of TF-mediated transcription, it may also be possible to activate specific genes using a Py-Im polyamide. Through the attachment of a transcriptional activation domain (TAD) capable of recruiting transcriptional machinery (Figure 1.13A), polyamides can be used to activate transcription at target sites. This construct is also referred to as an artificial transcription factor (ATF), as it functions like a transcription factor and contains similar domains, but can be entirely synthetic. These domains consist of a polyamide as the DNA-binding domain, and a synthetic molecule capable of recruiting activating transcription. Several polyamide-TAD conjugates have previously been demonstrated to upregulate transcription *in vitro*,[27-30] but a lack of efficient cellular uptake—likely due to the large molecular weight of these compounds—prevented *in vivo* applications.[10,31,32]

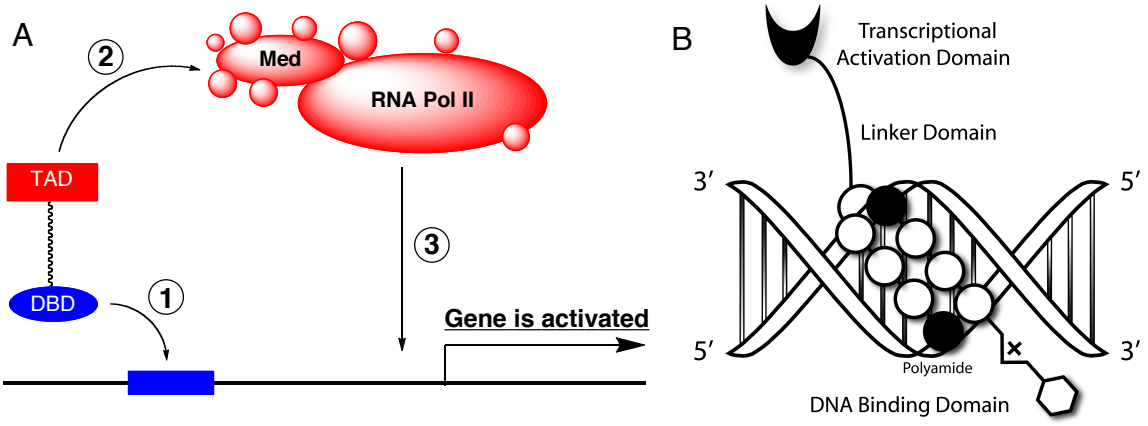


Figure 1.13: The concept of a polyamide based transcriptional activator. A) A small molecule containing DBD and TAD domains can bind to DNA (1), recruit transcriptional machinery (2), and allow for the assembly of the RNA Polymerase and initiation of transcription (3). B) General structure of a polyamide-TAD conjugate, indicating the DBD, linker domain, and TAD.

Scope of this Work

The work presented in this thesis involves various studies aimed at advancing Py-Im polyamides as a potential class of therapeutics. By making improvements that allow for increased biological activity in cancer cell lines, we have developed a more potent class of compounds with greatly increased cellular uptake. In Chapter 2, we present a new class of Py-Im polyamides bearing modifications to the γ-aminobutyric acid turn unit, with unprecedented levels of cytotoxicity in cell culture. We used confocal microscopy and flow cytometry to qualitatively and quantitatively assess levels of nuclear concentration, which strongly suggest that this effect was based on a significant increase in levels of nuclear uptake. We found that a simple modification to include an aryl-turn moiety can potentiate the biological effects of a Py-Im polyamide by up to two orders of magnitude. In Chapter 3, we develop an improved synthetic route that allows a practical method for unsymmetrical modifications to be made at the turn units of cyclic Py-Im polyamides. We then designed a panel of cyclic compounds to study the effects of a variety of substitutions that included aryl-turn technology, and found surprising results that further emphasize the dramatic effect that small structural changes can have on biological activity. In Chapter 4, we examine the aggregation propensity of a diverse library of Py-Im polyamides, finding that the solubility of polyamides does not seem to correlate with biological activity. We also investigate the effects of solubilizing agents, and find that this strategy can improve polyamide solubility and allow for improved injection conditions in mice. Finally, in Chapter 5 we develop a class of Py-Im polyamides conjugated to isoxazolidine moieties, and investigate their potential as artificial transcription factors.

References

1. Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Structure of a B-DNA dodecamer - conformation and dynamics. *Proc. Natl. Acad. Sci. USA* **1981,** 78, 2179-2183.

2. Crick, F., Central dogma of molecular biology. *Nature* **1970**, 227, 561.

3. Panne, D.; Maniatis, T.; Harrison, S. C., An atomic model of the interferon-beta enhanceosome. *Cell* **2007**, 129 (6), 1111-1123.

4. Kamitori, S.; Takusagawa, F., Multiple binding modes of anticancer drug actinomycin-D - X-ray, molecular modeling, and spectroscopic studies of d(GAAGCTTC)2-actinomycin D complexes and its host DNA. *J. Am. Chem. Soc.* **1994**, 116, 4154-4165.

5. Coll, M.; Frederick, C. A.; Wang, A. H. J.; Rich, A., A bifurcated hydrogen-bonded conformation in the d(A·T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc. Natl. Acad. Sci. USA* **1987**, 84, 8385-8389.

6. Mitra, S. N.; Wahl, M. C.; Sundaralingam, M., Structure of the side-by-side binding of distamycin to d(GTATATAC)(2). *Acta Crystallogr. Sect. D* **1999,** 55, 602-609.

7. Pelton, J. G.; Wemmer, D. E., Structural characterization of a 2-1 distamycin A·d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci. USA* **1989,** 86, 5723-5727.

8. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **1998,** 391, 468-471.

9. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., A structural basis for recognition of A·T and T·A base pairs in the minor groove of B-DNA. *Science* **1998,** 282, 111-115.

10. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003,** 13, 284-299.

11. Mrksich, M.; Parks, M. E.; Dervan, P. B., Hairpin peptide motif - a new class of oligopeptides for sequence-specific recognition in the minor-groove of double-helical DNA. *J. Am. Chem. Soc.* **1994,** 116, 7983-7988.

12. Trauger, J. W.; Baird, E. E.; Dervan, P. B., Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **1996,** 382, 559-561.

13. Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Cycle polyamide motif for recognition of the minor groove of DNA. *J. Am. Chem. Soc.* **1999,** 121, 1121-1129.

14. White, S.; Baird, E. E.; Dervan, P. B., Orientation preferences of pyrrole-imidazole polyamides in the minor groove of DNA. *J. Am. Chem. Soc.* **1997,** 119, 8756-8765.

15. Herman, D. M.; Baird, E. E.; Dervan, P. B., Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *J. Am. Chem. Soc.* **1998,** 120, 1382-1391.

16. Chenoweth, D. M.; Dervan, P. B., Allosteric modulation of DNA by small molecules. *Proc. Natl. Acad. Sci. USA* **2009,** 106, 13175-13179.

17. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B., Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res.* **2004,** 32, 2802-2818.

18. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg. Med. Chem.* **2002,** 10, 2767-2774.

19. Suto, R. K.; Edayathumangalam, R. S.; White, C. L.; Melander, C.; Gottesfeld, J. M.; Dervan, P. B.; Luger, K., Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. *J. Mol. Biol.* **2003,** 326, 371-380.

20. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B., Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci. USA* **2003,** 100, 12063-12068.

21. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Improved nuclear localization of DNA-binding polyamides. *Nucleic Acids Res.* **2007,** 35, 363-370.

22. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B., Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. USA* **2004,** 101, 16768-16773.

23. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem. Biol.* **2007,** 2, 561-571.

24. Nickols, N. G.; Dervan, P. B., Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc. Natl. Acad. Sci. USA* **2007,** 104, 10418-10423.

25. Muzikar, K. A.; Nickols, N. G.; Dervan, P. B., Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc. Natl. Acad. Sci. USA* **2009,** 106, 16598-16603.

26. Raskatov, J. A.; Meier, J. L.; Puckett, J. W.; Yang, F.; Ramakrishnan, P.; Dervan, P. B., Modulation of NF-kappaB-dependent gene transcription using programmable DNA minor groove binders. *Proc. Natl. Acad. Sci. USA* **2012,** 109, 1023-1028.

27. Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B., Activation of gene expression by small molecule transcription factors. *Proc. Natl. Acad. Sci. USA* **2000,** 97, 3930-3935.

28. Ansari, A. Z.; Mapp, A. K.; Nguyen, D. H.; Dervan, P. B.; Ptashne, M., Towards a minimal motif for artificial transcriptional activators. *Chem. Biol.* **2001,** 8, 583-592.

29. Arora, P. S.; Ansari, A. Z.; Best, T. P.; Ptashne, M.; Dervan, P. B., Design of artificial transcriptional activators with rigid poly-L-proline linkers. *J. Am. Chem. Soc.* **2002,** 124, 13067-13071.

30. Kwonj, Y.; Arndt, H. D.; Qian, M.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M., Small molecule transcription factor mimic. *J. Am. Chem. Soc.* **2004,** 126, 15940-15941.

31. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B., Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci. USA* **2003,** 100, 12063-12068.

32. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Improved nuclear localization of DNA-binding polyamides. *Nucleic Acids Res.* **2007,** 35, 363-370.