Chapter 1

Introduction

1.1 Deoxyribonucleic acid

In 1944, Oswald Avery, Colin MacLeod, and Maclyn McCarty discovered that deoxyribunocleic acid (DNA) is the chemical that makes up genes, the molecular unit of heredity (1). Then in 1953, the structure of DNA was solved, revealing it to be a double-stranded, anti-parallel, right-handed, double-helical polymer (Figure 1.1A) (2,3). Each DNA strand is composed of four monomers called nucleotides: deoxyadenosine (A), deoxythymidine (T), deoxycytidine (C), and deoxyguanosine (G), which are connected to one another by phosphodiester bonds with the deoxyribose sugar (Figure 1.1B). The two

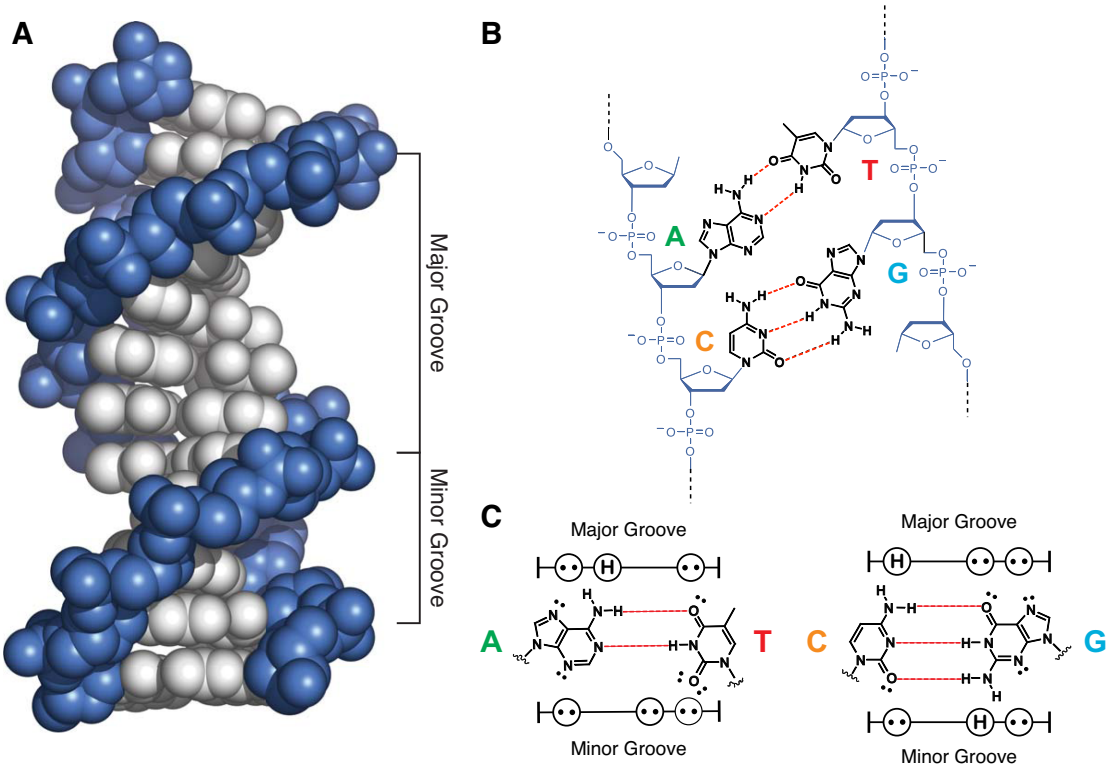


Figure 1.1 Structure of DNA and hydrogen bonding patterns of Watson-Crick base pairs. (**A**) Structure of B-form DNA (PDB 1BNA) (4). (**B**) Chemical structures duplex DNA showing all four bases. Adenine (A) is bonded to thymine (T) and cytosine (C) is bonded to guanine (G). Dashed lines indicate hydrogen bonds. (**C**) Major and minor groove hydrogen bonding patterns of the four Watson-Crick base pairs. Circles with dots represent lone pairs (hydrogen bond acceptor), and circles with an H represent hydrogen atoms from an exocyclic amine (hydrogen bond donor).

strands are held together by hydrogen bonds formed between the nitrogenous bases, which lie perpendicular to the helical axis. The structure of DNA also demonstrated the pairing rules between the bases of the two strands, namely A pairs with T and C pairs with G. These rules result in four possible base pairs, called Watson-Crick base pairs. Two hydrogen bonds can be formed between A and T, while three can form between C and G. The structure of DNA shown in Figure 1.1A is of B-form DNA which is believed to be the predominant form in cells. In this form of DNA, the asymmetric spacing of the strands caused by the nitrogenous bases creates major and minor grooves. The pattern of hydrogen bond donors and acceptors on the exocyclic portions of the bases allows for molecular recognition of all four base pairs in the two different grooves (Figure 1.1C), which is critical for all DNA functions.

DNA is vital to all life on Earth, and since the early discoveries of the 1940s and 1950s much effort has been put toward characterizing all of its functions. The central function of DNA was believed to be storing and coding the information for transcription of messenger RNA (mRNA) from genes that are then translated into the proteins needed for cellular processes. This is the so-called central dogma of molecular biology. Therefore, if one knew all of the genes within an organism's genome one could produce a list of all of the important factors. A milestone in this endeavor was the sequencing of the human genome in 2001 (5). Current estimates suggest that there are about 19,000 protein-coding genes encoded in the ~3 billion base pair human genome (6). Though, this number seems low relative to other less complex organisms, there is added information due to alternative splicing of exons. DNA also encodes for ~15,000 long non-coding RNAs,

which are believed to have regulatory functions (7). In addition, short peptides called short open reading frame encoded polypeptides (SEPs) are being discovered in higher order eukaryotes that are expressed from RNA transcripts previously thought to be non-coding (8). Clearly, after more than 70 years we still have much to learn about DNA. Continuing to learn about DNA's chemical properties and biological functions has been critical to understanding life and the origin of disease, as many diseases can be traced to aberrations in the genome. With enough knowledge, one may be able to develop chemical tools to combat a variety of diseases through targeting of the DNA.

1.2 Molecular recognition of DNA in nature

Transcription of DNA into RNA is one of the most critical processes in the cell. Transcription factors (TFs) are proteins that reversibly bind to DNA and regulate transcription by participating in the recruitment or blocking of RNA polymerase. TFs bind to specific sequences of DNA non-covalently through interactions with the nitrogenous bases in the minor groove, the major groove, or a combination of both.

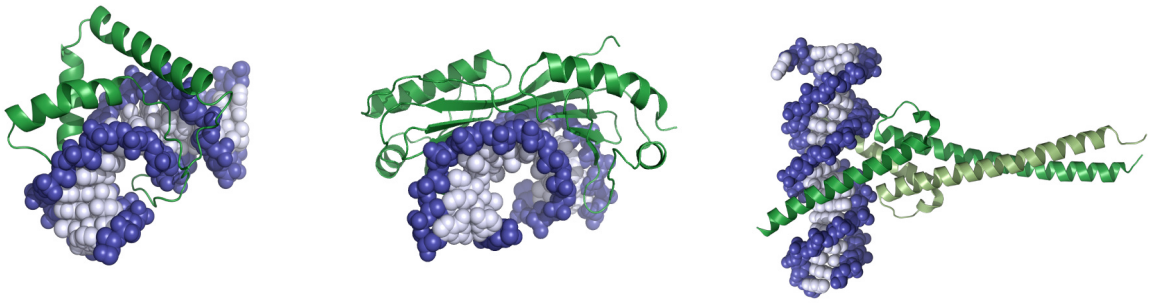


Figure 1.2 Structures of transcription factor:DNA complexes. Lef-1 (PDB 2LEF) (12), TBP (PDB 1TGH) (13), Myc/Max (PDB 1NKP) (14).

Through interactions with DNA as well as interactions with other proteins either directly or allosterically through DNA, TFs are able to achieve high binding affinities to DNA (9-11). TFs have a variety DNA-binding domain motifs, representing different strategies nature has evolved for molecular recognition of DNA. Such examples include the minor groove-binding high mobility group domain found in Lef-1, the minor groove-binding TATA binding protein, and the major groove-binding basic helix-loop-helix leucine zipper domain found in the Myc/MAX heterodimer (Figure 1.2) (12-14). In addition to their critical role in transcription, TFs offer insight into how sequence-selective recognition of DNA can be achieved by artificially engineered factors or synthetic molecules.

Another class of biomolecules that can reversibly bind to specific sequences of DNA is small molecule natural products. Like TFs, these small molecules are also able to

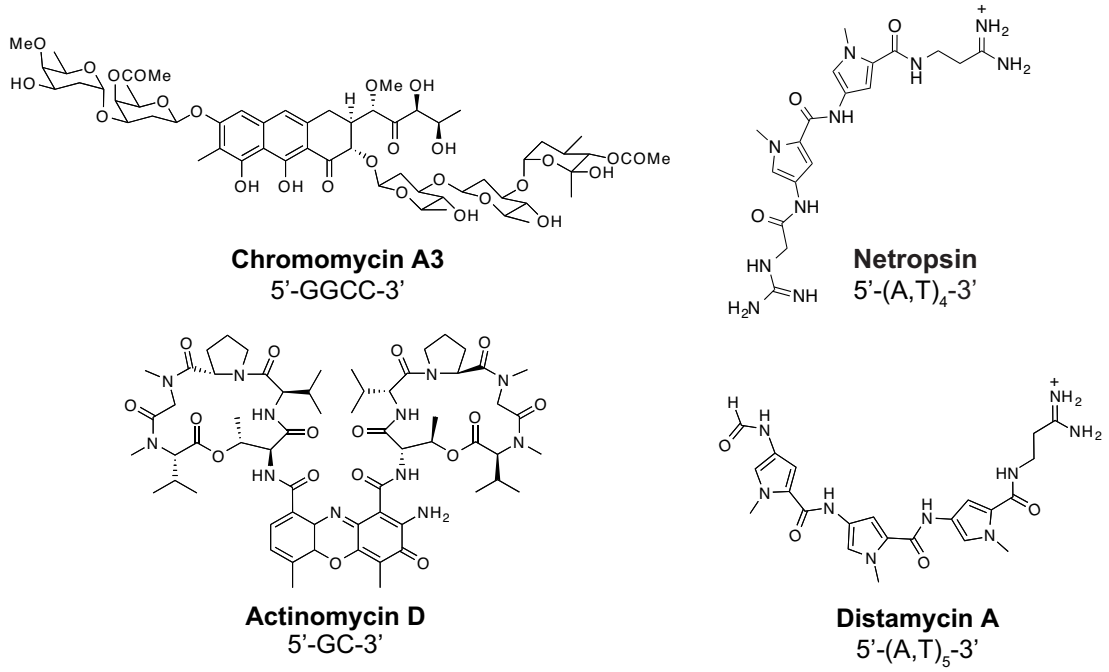


Figure 1.3 Chemical structures of DNA-binding small molecule natural products.

bind to DNA by various modes including intercalation, minor groove binding, major groove binding, and combinations thereof (15). Examples of DNA-binding natural products include the intercalator actinomycin D, and the minor groove binders chromomycin A3, netropsin, and distamycin A (Figure 1.3). Actinomycin D is selects for 5'-GC-3' sequences, while chromomycin selects for 5'-GGCC-3'. Both netropsin and distamycin bind specifically to A/T tracts.

DNA-binding natural products are often antibiotics synthesized by microorganisms in order to kill or inhibit the growth of competing organisms of a different species, and some have also been effective as anticancer agents (15,16). These DNA-binding molecules often inhibit critical DNA-dependent processes such as transcription and DNA replication. For example, actinomycin has been shown to inhibit both transcription and replication (17,18). Distamycin A has been shown to disrupt a variety of DNA-dependent processes as well through inhibition of RNA polymerase, DNA polymerase, topoisomerases I and II, and helicases (19-21).

DNA binding by distamycin is a particularly interesting example of molecular recognition, as the crescent-shaped molecule can bind the DNA minor groove in either a 1:1 or 2:1 stoichiometry (Figure 1.4). In the 2:1 complex, two molecules of distamycin are stacked in an anti-parallel fashion such that the N-mehtylpyrrole units are across from each other, which results in significant widening of the minor groove compared to the 1:1 complex (22-24). In both 1:1 and 2:1 configurations, distamycin binds preferentially to A•T base pairs; however, it was suggested that replacement of N-methylpyrrole with N-

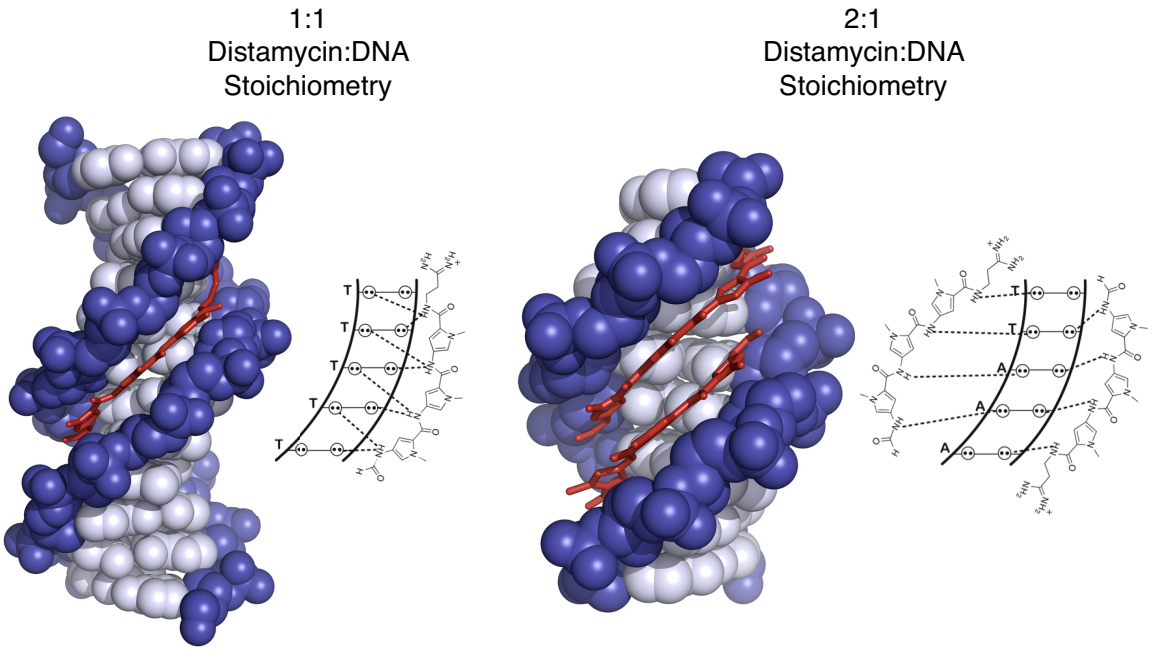


Figure 1.4 Structures of distamycin A bound to DNA. (**A**) Structure of 1:1 complex of distamycin with DNA (PDB 2DND) and a schematic of distamycin binding to the minor groove in the 1:1 mode. (**B**) 2:1 antiparallel complex of distamycin with DNA (PDB 378D) and a schematic of distamycin binding to the minor groove in the 2:1 mode.

methylimidazole would allow for recognition of G•C base pairs in the floor of the minor groove (25). Drawing inspiration from nature to synthesize small molecules capable of sequence-specific recognition of DNA could provide an effective means of combating disease states through the modulation of transcription.

1.3 Pyrrole-Imidazole polyamides

The initial observation of the 2:1 distamycin binding configuration to DNA led to the development of a new class of synthetic small molecules called pyrrole-imidazole (Py-Im) polyaimdes that are capable of recognizing all four Watson-Crick base pairs and bind to DNA with affinities comparable to DNA-binding proteins (26,27). Recognition of all four base pairs was achieved by modifying distamycin to incorporate N-

methylimidazole and 3-hydroxy-1-methylpyrrole (Hp) heterocycles. The side-by-side pairing of these monomers in the minor groove governs which base is preferentially recognized (Figure 1.5). An Im/Py pair preferentially recognizes a G•C base pair over C•G, A•T, and T•A, while Py/Im recognizes C•G. This preference exists due to the relief of the steric hindrance between the hydrogen at the C3 position of pyrrole and the guanine exocyclic amine provided by the substitution of imidazole. Hp/Py preferentially recognizes a T•A base pair over A•T, C•G, and G•C, while Py/Hp recognizes A•T. Specificity for T•A over A•T is likely due to steric accommodation of the exocyclic hydroxyl group on Hp. Another heterocycle that has been used to distinguish T•A over A•T is 3-chlorothiophene when paired across a Py (28). As with the 2:1 distamycin configuration, Py/Py pairs recognize A•T and T•A base pairs over G•C and C•G.

Many other modifications to the distamycin framework have also been made to improve properties such as affinity, specificity, solubility, and cellular/nuclear uptake. In the current generation of Py-Im polyamides, γ-diaminobutyric acid (γ-DABA) is used to link two chains of polyamides to help orient the molecule in a hairpin configuration when bound to DNA (Figure 1.5). This hairpin configuration mimics the 2:1 anti-parallel distamycin stacking and increases affinity to DNA significantly, in part due to reduced entropic penalty (29). In addition, linking two polyamide chains allows for two different sequences of Py and Im to be arranged on the top and bottom strands, as opposed to having the same strand stack against itself in a 2:1 fashion. The chiral alpha amine on γ-DABA also increases affinity to DNA and shows preferential binding to A•T and T•A base pairs (30). Polyamide chains have also been linked by (*R*)-3,4-diaminobutyric acid,

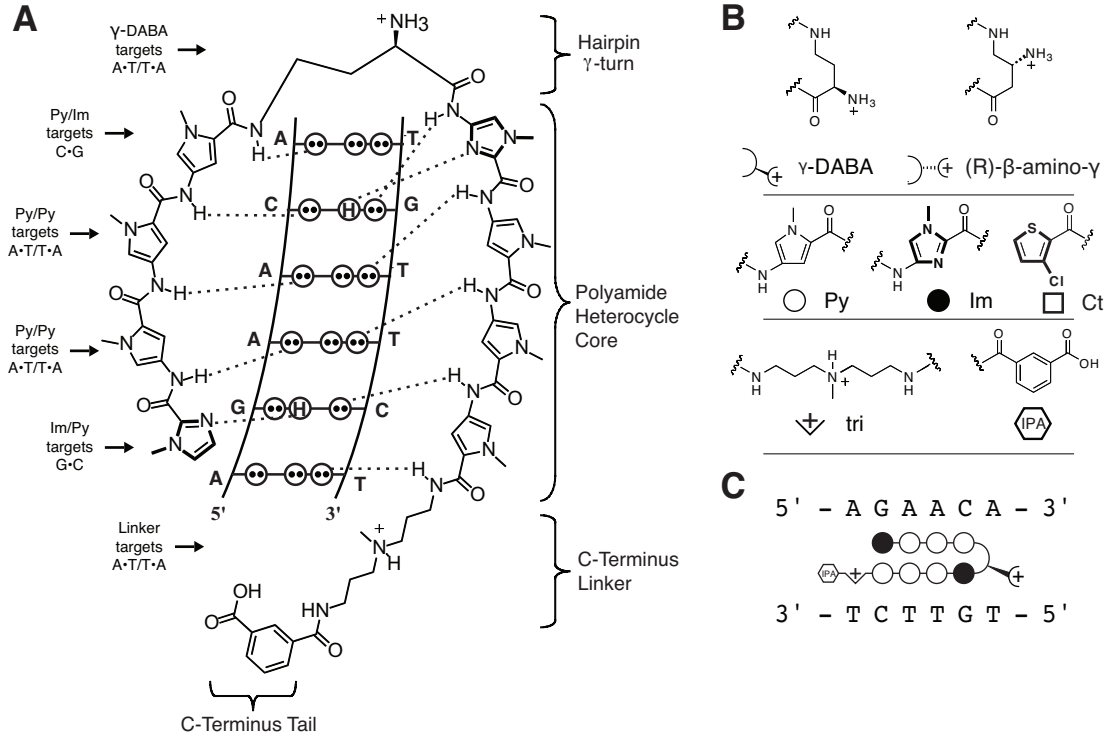


Figure 1.5 Molecular recognition of DNA by Py-Im polyamides. (**A**) Schematic of Py-Im polyamide targeted to 5'-WGWWCW-3' binding to the minor groove. (**B**) Chemical structures of monomers found in polyamides and their ball-and-stick representations. Abbreviation for each monomer is also listed. (**C**) Ball-and-stick representation of the Py-Im polyamide in (A) bound to its cognate DNA sequence.

which also show improvements to affinity (30,31). Another feature of current generation Py-Im polyamides is the conjugation of isophthalic acid to the C-terminus linked by 3,3'-diamino-N-methyldipropylamine. This C-terminus tail aids in the uptake of polyamides into live cells (32). Uptake into the nucleus of live cells was demonstrated by conjugating fluorescein to the C-terminus linker (33). The 3,3'-diamino-N-methyldipropylamine C-terminus linker also shows preferential binding for A•T and T•A base pairs. Therefore, currently used eight-ring hairpin Py-Im polyamides are capable of recognizing a specific six base pair sequence.

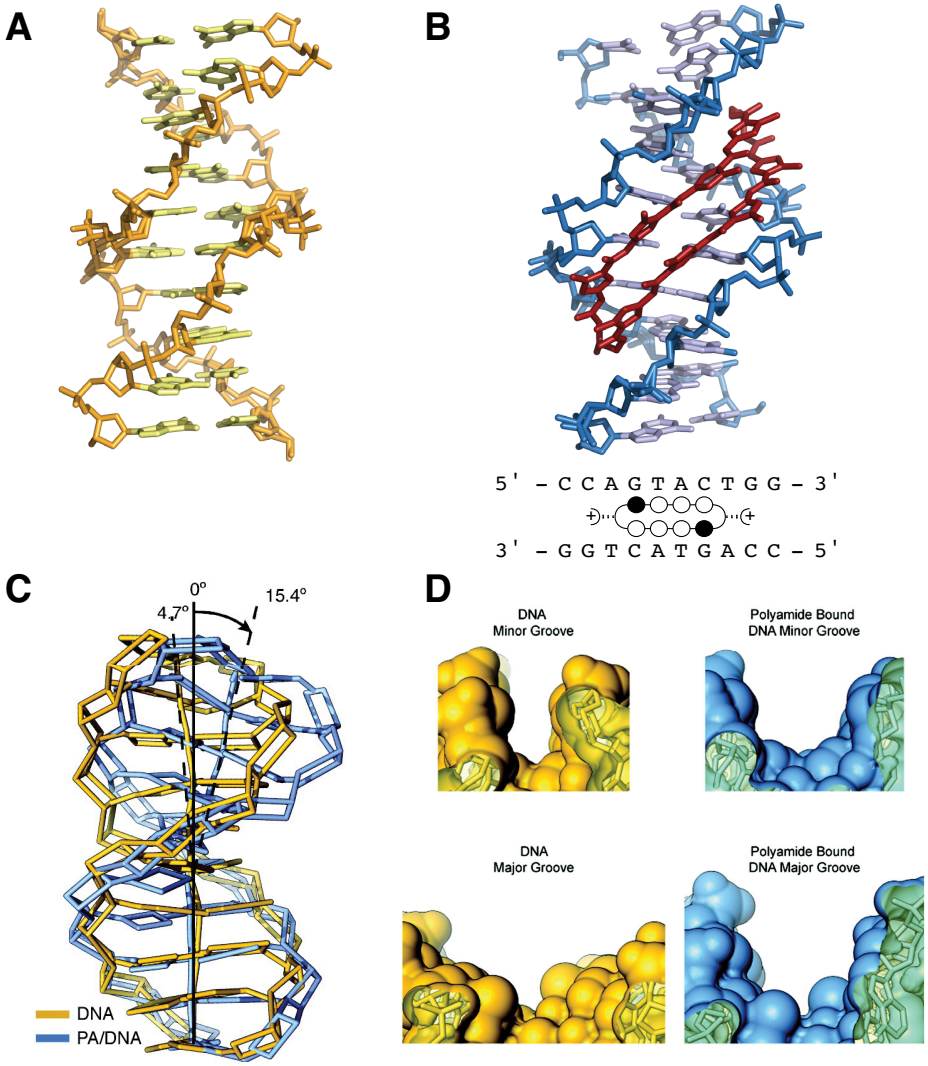


Figure 1.6 Py-Im polyamide binding alters the structure of DNA. (**A**) Native DNA crystal structure at 0.98 Å resolution (PDB 1D8G) (**B**) DNA/polyamide co-crystal structure at 0.95 Å resolution (PDB 3OMJ), ball-and-stick model of polyamide bound to DNA (**C**) Significant DNA bending is observed for polyamide-bound DNA (blue) versus unbound DNA (yellow). (**D**) Top: comparison of the minor-groove width for DNA in the absence of polyamide (yellow) and in the presence of bound polyamide (blue). Bottom: Comparison of the major-groove widths.

Features of the current generation Py-Im polyamide and its effects on DNA can be observed in the crystal structure of a macrocyclic eight-ring Py-Im polyamide bound to DNA (Figure 1.6) (34). Polyamide binding induces widening of the minor groove and narrowing of the major groove, as observed in the distamycin 2:1 structure (Figure 1.4),

as well as bending of the helix. This structure demonstrates how polyamides can act allosterically to inhibit binding of transcription factors and other DNA-binding proteins (34). Structural studies have also shown that polyamides can bind to DNA in a reconstituted nucleosomal core particle (35), suggesting that they are capable of binding to chromatin. The unique properties of these cell-permeable, sequence-specific, high affinity DNA-binding small molecules make them an attractive candidate for modulation of dysregualted gene expression in disease states.

1.4 Biological activity of pyrrole-imidazole polyamides

Py-Im polyamides have been used successfully to modulate gene expression in many cell culture models of disease. In human glioblastoma cells U251, for example, polyamides targeted to the HIF-1:DNA interface were shown to downregulate HIF-1-driven genes such as *vegf* (Figure 1.7A) (36). In total, 69 out of 297 induced transcripts were affected at least twofold by polyamide treatment in this cell line, and ~1500 total transcripts were affected. Polyamide treatment was also shown to reduce occupany of HIF-1 at the promoter binding site at some, but not all, genes, suggesting that polyamides may inhibit binding of HIF-1 to DNA as observed *in vitro* (37). Similar results were observed in human prostate cancer cells LNCaP using a polyamide targeted to the androgen response element (ARE) in order to disrupt androgen receptor (AR):DNA binding (Figure 1.7B) (38). More recently, next generation sequencing has been employed to analyze all of the genes affected by a polyamide targeted to the NF-κB:DNA interface in human non-small cell lung cancer A549 (39). By RNA-seq, polyamide

treatment was shown to downregulate 182 transcripts out of 650 induced transcipts, and in total the expression of ~1200 transcripts were affected.

Evidence that Py-Im polyamides are biologically active and potent in cell culture led to the investigations of polyamide effects in human tumor xenografts grown in mice. The first milestone was the demonstration that polyamides are able to circulate in a healthy mouse and have favorable pharmacokinetics (Figure 1.7C) (40-42). Py-Im polyamides were also shown to accumulate into tumors and modulate gene expression, as observed in cell culture (43-45). Interestingly, investigations of polyamide biodistribution

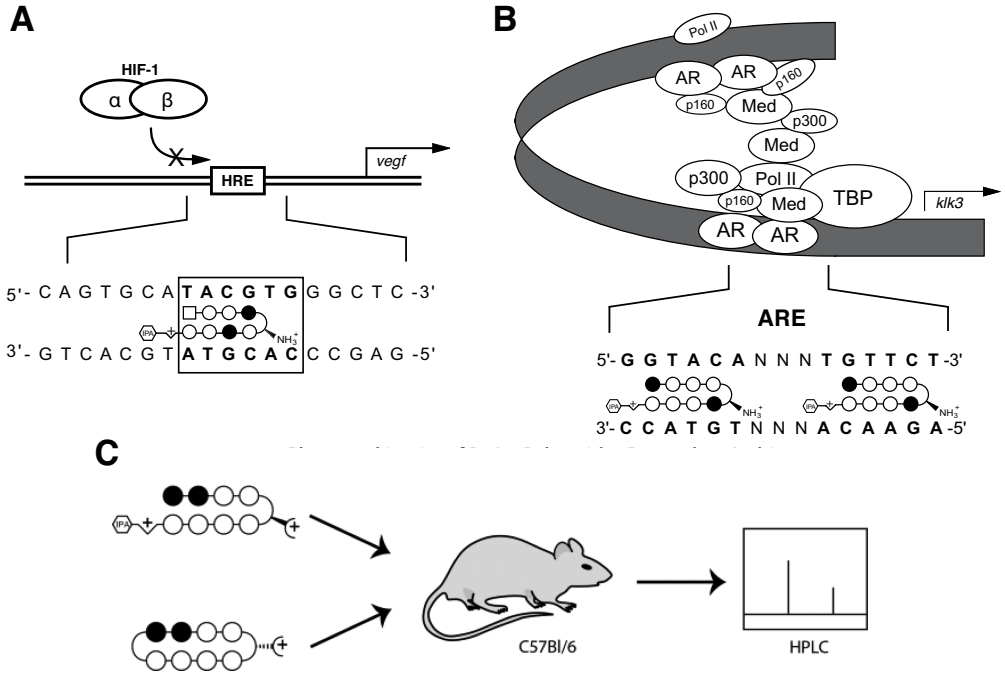


Figure 1.7 Examples of polyamide biological activities in tissue culture and animal studies. (**A**) A Py-Im polyamide targeted to the HIF response element (HRE) was able to downregulate *vegf* expression and reduce occupancy of HIf-1 at the *vegf* promoter (36). (**B**) A Py-Im polyamide targeted to the androgen response element (ARE) was able to downregulate *klk3* expression and reduce occupancy of AR at the *klk3* promoter and enhancer (38). (**C**) Py-Im polyamides injected into C57BL/6 mice were able to circulate for hours in the blood. Polyamide plasma levels varied with the architecture of the molecules (40).

in healthy mice showed showed measurable polyamide levels in all tissues except brain (46). Importantly, ARE-targeted polyamides were able to reduce the growth of human prostate cancer xenografts with favorable animal toxicity profiles (47,48).

The early success of animal experiments suggests that Py-Im polyamides could one day be an effective therapeutic for the treatment of human cancers. Before that point, however, it will be important to understand the mechanism of action of polyamides and any potential side effects. The abundance of transcripts affected by polyamide treatment that were not induced or directly driven by the TF of interest raised the question as to whether polyamide effects may be exerted by a non-specific mechanism rather than specific disruption of the TF:DNA interface to modulate gene expression. In addition, Py-Im polyamides targeted to different sequences have been found to induce cytotoxicity in a given cell line, despite affecting different genes and signaling pathways (39,44,49). Non-specific inhibition of DNA-dependent processes, such as transcription and replication, by other small molecule DNA-binders spoke to this concern as well (15,18). Non-specific inhibition of transcription was explored recently in LNCaP cells using ARE-targeted polyamides (47). In this study, polyamide treatment caused upregulation of p53 and PARP cleavage, suggesting induction of apoptosis. This was accompanied by degradation of the RNA polymerase (pol) II large subunit RPB1, which can occur when transcription is stalled. RNA pol II was also found to have reduced occupancy at transcription start sites. These results suggested that polyamide bound DNA may inhibit transcription by blocking RNA pol II, which leads to apoptosis. Inhibition of RNA pol II has also been shown with distamycin and actinomycin (50,51).

Another potential mechanism investigated to explain the non-specific effects of polyamides was inhibition of topoisomerases (52). Topoisomerases are enzymes that bind to DNA and introduce single- or double-strand breaks in DNA to relieve helical torsion that occurs during DNA replication and transcription (53). Many natural products exert their toxic effects through inhibition of topoisomerases (15). *In vitro* experiments showed that ARE-targeted polyamides inhibit Top2α-catalyzed relaxation of a super-coiled plasmid *in vitro*, and that the likely mechanism is through inhibition of Top2α binding to DNA. Knockdown of Top2α in hormone-insensitive prostate cancer cells also conferred resistance to polyamide treatment. Inhibition of RNA Pol II elongation and inhibition of topoisomerase DNA binding are two potential mechanisms of non-specific polyamide effects and cytotoxicity, but there are other critical DNA-dependent processes to explore as well.

1.5 Small Molecule Inhibitors of DNA replication

Previous studies suggested that while Py-Im polyamides are effective modulators of gene expression and potent inhibitors of cell growth, their mechanism of action might not occur by specific inhibition of TF:DNA binding. Another process potentially inhibited by polyamides that can explain non-specific effects is DNA replication. DNA replication is the process by which the genome is copied prior to cell division. Like transcription, it involves that action of numerous proteins, many of which interact with DNA to form a complex called the replisome (54). Some of the proteins in the replisome are the DNA polymerases, the replicative helicase, the sliding clamp processivity factor,

primase, and single-stranded binding protein. Inhibition of any of these factors will disrupt DNA replication. Even prior to replication initiation, inhibition of the any of the factors that are needed to begin DNA replication will prevent the process from moving forward. Damage to the DNA template itself by DNA base adducts can also prevent translocation of the replisome.

Numerous natural product and synthetic small molecules are capable of inhibiting DNA replication through a variety of mechanisms (Table 1.1). Aphidicolin binds directly to DNA polymerase to inhibit replication fork progression (55). Hydroxyurea and gemcitabine inhibit ribonucleotide reductase, which depletes the nucleotide pool needed for DNA synthesis (56,57). NSC 19630 inhibits werner helicase, which is important in DNA repair (59). T2AA prevents binding of PCNA, the DNA polymerase processivity factor, to DNA polymerase (60). ET-743 alkylates DNA, which then forms a double-stranded break when encountered during replication (61). Comparison of Py-Im polyamide effects on the cell cycle and DNA replication to those of these small molecules can provide insight into the non-specific effects of polyamides.

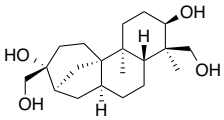
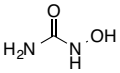
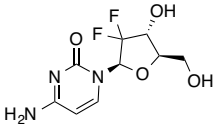
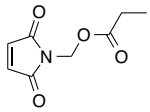
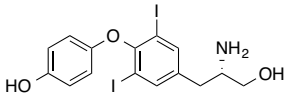
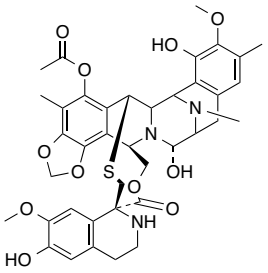
Inhibitor	Mechanism of Action
 Aphidicolin	Inhibits B-family DNA polymerases by binding near the nucleotide binding site
 Hydroxyurea	Inhibits ribonucleotide reductase by quenching the tyrosyl radical through a 1-electron transfer
 Gemcitabine	Inhibits ribonucleotide reductase through loss of tyrosyl radical or covalent reaction with alpha subunit; can also be incorporated into DNA
 NSC 19630	Inhibits WRN helicase
 T2AA	Inhibits of interaction of PCNA with PIP-box domain on DNA polymerase
 ET-743	Binds to minor groove of DNA and alkylates guanine at N2; adducts lead to replication-dependent double strand breaks

Table 1.1 Structures and putative mechanisms of a library of small molecule DNA replication inhibitors (55-61).

1.6 Scope of this work

The work presented in this thesis is primarily focused on elucidating the effects of Py-Im polyamides on DNA replication, and explores how mammalian cells respond to the unique stress exerted by these molecules. In Chapter 2, we present data demonstrating that androgen response element (ARE)-targeted polyamides inhibit DNA replication in

hormone insensitive prostate cancer cells. Cells respond to inhibited DNA replication by activating the ATR checkpoint signaling pathway. However, only part of the downstream signaling events observed in response to stress induced by other small molecules were activated in polyamide-treated cells. We conclude that the replication stress induced by polyamides in this cell line is low relative to other characterized inhibitors. Experiments testing polyamide effects *in vitro* suggest that polyamides may function by inhibiting the replicative helicase. In Chapter 3, we attempt to study the mechanism of polyamide-induced DNA replication stress in closer detail by utilizing cell-free extracts made from *X. laevis* oocytes. In this system, DNA replication is inhibited by treatment with a variety of Py-Im polyamides targeted to different six base pair sequences. However, ATR signaling is not activated. Imaging of the DNA in activated extracts revealed that chromatin failed to properly decondense in response to polyamide treatment, preventing replication initiation. Chapter 4 reports the effects of Py-Im polyamides in a cell line model of prostate cancer harboring the TMPRSS2-ERG gene fusion. We show that polyamides targeted to the ARE in the TMPRSS2 promoter decrease ERG expression, and that polyamides targeted to the ERG recognition sequence decrease expression of ERG-driven genes. A polyamide targeted to the ARE was found to reduce growth of tumors engrafted to SCID mice.

1.7 References

1. Avery, O.T., Macleod, C.M. and McCarty, M. (1944) Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *J Exp Med*, 79, 137-158.

2. Watson, J.D. and Crick, F.H. (1953) Genetical implications of the structure of deoxyribonucleic acid. *Nature*, 171, 964-967.

3. Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171, 737-738.

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

5. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, 409, 860-921.

6. Ezkurdia, I., Juan, D., Rodriguez, J.M., Frankish, A., Diekhans, M., Harrow, J., Vazquez, J., Valencia, A. and Tress, M.L. (2014) Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. *Hum Mol Genet*, 23, 5866-5878.

7. Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S. *et al.* (2012) GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*, 22, 1760-1774.

8. Chu, Q., Ma, J. and Saghatelian, A. (2015) Identification and characterization of sORF-encoded polypeptides. *Crit Rev Biochem Mol Biol*, 50, 134-141.

9. Nalefski, E.A., Nebelitsky, E., Lloyd, J.A. and Gullans, S.R. (2006) Single-molecule detection of transcription factor binding to DNA in real time: specificity, equilibrium, and kinetic parameters. *Biochemistry*, 45, 13794-13806.

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

11. Naar, A.M., Lemon, B.D. and Tjian, R. (2001) Transcriptional coactivator complexes. *Annu Rev Biochem*, 70, 475-501.

12. Love, J.J., Li, X., Case, D.A., Giese, K., Grosschedl, R. and Wright, P.E. (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature*, 376, 791-795.

13. Juo, Z.S., Chiu, T.K., Leiberman, P.M., Baikalov, I., Berk, A.J. and Dickerson, R.E. (1996) How proteins recognize the TATA box. *J Mol Biol*, 261, 239-254.

14. Nair, S.K. and Burley, S.K. (2003) X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. *Cell*, 112, 193-205.

15. Tse, W.C. and Boger, D.L. (2004) Sequence-selective DNA recognition: natural products and nature's lessons. *Chem Biol*, 11, 1607-1617.

16. Gurova, K. (2009) New hopes from old drugs: revisiting DNA-binding small molecules as anticancer agents. *Future Oncol*, 5, 1685-1704.

17. Bensaude, O. (2011) Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? *Transcription*, 2, 103-108.

18. Horikawa, S., Sakata, K., Uchiumi, F., Hatanaka, M. and Tsukada, K. (1987) Effect of actinomycin D on DNA replication and c-myc expression during rat liver regeneration. *Biochem Biophys Res Commun*, 144, 1049-1054.

19. Zimmer, C., Puschendorf, B., Grunicke, H., Chandra, P. and Venner, H. (1971) Influence of netropsin and distamycin A on the secondary structure and template activity of DNA. *Eur J Biochem*, 21, 269-278.

20. Beerman, T.A., McHugh, M.M., Sigmund, R., Lown, J.W., Rao, K.E. and Bathini, Y. (1992) Effects of analogs of the DNA minor groove binder Hoechst 33258 on topoisomerase II and I mediated activities. *Biochim Biophys Acta*, 1131, 53-61.

21. Brosh, R.M., Jr., Karow, J.K., White, E.J., Shaw, N.D., Hickson, I.D. and Bohr, V.A. (2000) Potent inhibition of werner and bloom helicases by DNA minor groove binding drugs. *Nucleic Acids Res*, 28, 2420-2430.

22. Coll, M., Frederick, C.A., Wang, A.H. and Rich, A. (1987) 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 U S A*, 84, 8385-8389.

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

24. Pelton, J.G. and Wemmer, D.E. (1989) Structural characterization of a 2:1 distamycin A.d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc Natl Acad Sci U S A*, 86, 5723-5727.

25. Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. and Dickerson, R.E. (1985) The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc Natl Acad Sci U S A*, 82, 1376-1380.

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

27. Hsu, C.F., Phillips, J.W., Trauger, J.W., Farkas, M.E., Belitsky, J.M., Heckel, A., Olenyuk, B.Z., Puckett, J.W., Wang, C.C. and Dervan, P.B. (2007) Completion of a Programmable DNA-Binding Small Molecule Library. *Tetrahedron*, 63, 6146-6151.

28. Foister, S., Marques, M.A., Doss, R.M. and Dervan, P.B. (2003) Shape selective recognition of T.A base pairs by hairpin polyamides containing N-terminal 3-methoxy (and 3-chloro) thiophene residues. *Bioorg Med Chem*, 11, 4333-4340.

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

30. Farkas, M.E., Li, B.C., Dose, C. and Dervan, P.B. (2009) DNA sequence selectivity of hairpin polyamide turn units. *Bioorg Med Chem Lett*, 19, 3919-3923.

31. Dose, C., Farkas, M.E., Chenoweth, D.M. and Dervan, P.B. (2008) Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit. *J Am Chem Soc*, 130, 6859-6866.

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

33. Best, T.P., Edelson, B.S., Nickols, N.G. and Dervan, P.B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc Natl Acad Sci U S A*, 100, 12063-12068.

34. Chenoweth, D.M. and Dervan, P.B. (2010) Structural basis for cyclic Py-Im polyamide allosteric inhibition of nuclear receptor binding. *J Am Chem Soc*, 132, 14521-14529.

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

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

37. Olenyuk, B.Z., Zhang, G.J., Klco, J.M., Nickols, N.G., Kaelin, W.G., Jr. and Dervan, P.B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A*, 101, 16768-16773.

38. Nickols, N.G. and Dervan, P.B. (2007) Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc Natl Acad Sci U S A*, 104, 10418-10423.

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

40. Raskatov, J.A., Hargrove, A.E., So, A.Y. and Dervan, P.B. (2012) Pharmacokinetics of Py-Im polyamides depend on architecture: cyclic versus linear. *J Am Chem Soc*, 134, 7995-7999.

41. Harki, D.A., Satyamurthy, N., Stout, D.B., Phelps, M.E. and Dervan, P.B. (2008) In vivo imaging of pyrrole-imidazole polyamides with positron emission tomography. *Proc Natl Acad Sci U S A*, 105, 13039-13044.

42. Synold, T.W., Xi, B., Wu, J., Yen, Y., Li, B.C., Yang, F., Phillips, J.W., Nickols, N.G. and Dervan, P.B. (2012) Single-dose pharmacokinetic and toxicity analysis of pyrrole-imidazole polyamides in mice. *Cancer Chemother Pharmacol*, 70, 617-625.

43. Raskatov, J.A., Nickols, N.G., Hargrove, A.E., Marinov, G.K., Wold, B. and Dervan, P.B. (2012) Gene expression changes in a tumor xenograft by a pyrrole-imidazole polyamide. *Proc Natl Acad Sci U S A*, 109, 16041-16045.

44. Nickols, N.G., Szablowski, J.O., Hargrove, A.E., Li, B.C., Raskatov, J.A. and Dervan, P.B. (2013) Activity of a Py-Im polyamide targeted to the estrogen response element. *Mol Cancer Ther*, 12, 675-684.

45. Raskatov, J.A., Puckett, J.W. and Dervan, P.B. (2014) A C-14 labeled Py-Im polyamide localizes to a subcutaneous prostate cancer tumor. *Bioorg Med Chem*, 22, 4371-4375.

46. Raskatov, J.A., Szablowski, J.O. and Dervan, P.B. (2014) Tumor xenograft uptake of a pyrrole-imidazole (Py-Im) polyamide varies as a function of cell line grafted. *J Med Chem*, 57, 8471-8476.

47. Yang, F., Nickols, N.G., Li, B.C., Marinov, G.K., Said, J.W. and Dervan, P.B. (2013) Antitumor activity of a pyrrole-imidazole polyamide. *Proc Natl Acad Sci U S A*, 110, 1863-1868.

48. Yang, F., Nickols, N.G., Li, B.C., Szablowski, J.O., Hamilton, S.R., Meier, J.L., Wang, C.M. and Dervan, P.B. (2013) Animal toxicity of hairpin pyrrole-imidazole polyamides varies with the turn unit. *J Med Chem*, 56, 7449-7457.

49. Muzikar, K.A., Nickols, N.G. and Dervan, P.B. (2009) Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc Natl Acad Sci U S A*, 106, 16598-16603.

50. Mote, J., Ghanouni, P. and Reines, D. (1994) A DNA Minor Groove-Binding Ligand Both Potentiates and Arrests Transcription by Rna-Polymerase-Ii - Elongation-Factor Sii Enables Readthrough at Arrest Sites. *Journal of Molecular Biology*, 236, 725-737.

51. Sobell, H.M. (1985) Actinomycin and DNA-Transcription. *P Natl Acad Sci USA*, 82, 5328-5331.

52. Phillips, J.W. (2011). PhD Thesis, California Institute of Technology.

53. Durand-Dubief, M., Svensson, J.P., Persson, J. and Ekwall, K. (2011) Topoisomerases, chromatin and transcription termination. *Transcription*, 2, 66-70.

54. Masai, H., Matsumoto, S., You, Z.Y., Yoshizawa-Sugata, N. and Oda, M. (2010) Eukaryotic Chromosome DNA Replication: Where, When and How? *Annual Review of Biochemistry, Vol 79*, 79, 89-130.

55. Baranovskiy, A.G., Babayeva, N.D., Suwa, Y., Gu, J., Pavlov, Y.I. and Tahirov, T.H. (2014) Structural basis for inhibition of DNA replication by aphidicolin. *Nucleic Acids Res*, 42, 14013-14021.

56. Lassmann, G., Thelander, L. and Graslund, A. (1992) EPR stopped-flow studies of the reaction of the tyrosyl radical of protein R2 from ribonucleotide reductase with hydroxyurea. *Biochem Biophys Res Commun*, 188, 879-887.

57. Artin, E., Wang, J., Lohman, G.J.S., Yokoyama, K., Yu, G.X., Griffin, R.G., Bar, G. and Stubbe, J. (2009) Insight into the Mechanism of Inactivation of Ribonucleotide Reductase by Gemcitabine 5 '-Diphosphate in the Presence or Absence of Reductant. *Biochemistry*, 48, 11622-11629.

58. Gowda, A.S.P., Polizzi, J.M., Eckert, K.A. and Spratt, T.E. (2010) Incorporation of Gemcitabine and Cytarabine into DNA by DNA Polymerase beta and Ligase III/XRCC1. *Biochemistry*, 49, 4833-4840.

59. Aggarwal, M., Sommers, J.A., Shoemaker, R.H. and Brosh, R.M. (2011) Inhibition of helicase activity by a small molecule impairs Werner syndrome helicase (WRN) function in the cellular response to DNA damage or replication stress. *P Natl Acad Sci USA*, 108, 1525-1530.

60. Punchihewa, C., Inoue, A., Hishiki, A., Fujikawa, Y., Connelly, M., Evison, B., Shao, Y.M., Heath, R., Kuraoka, I., Rodrigues, P. *et al.* (2012) Identification of Small Molecule Proliferating Cell Nuclear Antigen (PCNA) Inhibitor That Disrupts Interactions with PIP-box Proteins and Inhibits DNA Replication. *Journal of Biological Chemistry*, 287, 14289-14300.

61. Soares, D.G., Escargueil, A.E., Poindessous, V., Sarasin, A., de Gramont, A., Bonatto, D., Henriques, J.A.P. and Larsen, A.K. (2007) Replication and homologous recombination repair regulate DNA double-strand break formation by the antitumor alkylator ecteinascidin 743. *P Natl Acad Sci USA*, 104, 13062-13067.