ROLE OF A CHAIR-TYPE G-QUADRUPLEX IN C-*MYC* REGULATION

By

CORY LYLE GRAND

A Dissertation Submitted to the Faculty of the

GRADUATE INTERDISCIPLINARY PROGRAM IN CANCER BIOLOGY

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2003

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

THE UNIVERSITY OF ARIZONA ®
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Cory Lyle Grand

entitled Role of a Chair-Type G-Quadruplex in c-myc Regulation

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Laurence H. Hurley, Ph.D.	7/2/03 Date
Bernard W. Futscher, Ph.D.	7/2/03 Date
G. Timothy Bowden, Ph.D.	7/2/03 Date
Scot W. Ebbinghaus, M.D.	7/2/03 Date
Daniel D. Von Hoff, M.D.	7/2/03 Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director Laurence H. Hurley, Ph.D.	7/2/03 Date

STATEMENT BY THE AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Signed 

ACKNOWLEDGEMENTS

I would like to thank my graduate advisor and mentor Dr. Laurence H. Hurley for the invaluable support and guidance of this research project. I feel that Dr. Hurley has not only educated me on the scientific principles contained herein, but has also provided insight on improving my skills as a researcher. I would also like to thank the other members of my committee, Drs. Scot Ebbinghaus, Bernard Futscher, Daniel Von Hoff and G. Tim Bowden. In addition, I would like to thank the former members of my advisory committee, Drs. Scott Selleck and David Bearss. Dr. Bearss was especially valuable to my graduate career, both as an advisor and a friend. My appreciation also goes out to the members of the Cress and Bowden laboratories, for the use of their equipment and advice. The staff of the Hurley and Von Hoff laboratories also deserve recognition, as much of this research could not have been performed without their input and expertise. I would also like to acknowledge the input of Anne Cione and Dr. David Bishop, both of whom have made my graduate experience much easier with their assistance. Finally, I would like to thank my parents, Lyle Grand and Avadawn Wisminity, and my good friends, both old and new, for their unwavering support and kindness.

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

LIST OF FIGURES

LIST OF FIGURES – CONTINUED

LIST OF FIGURES – CONTINUED

LIST OF FIGURES – CONTINUED

LIST OF TABLES

ABSTRACT

As the importance of developing anticancer agents specific to tumor cells grows, a need has arisen for new and better targets. It has been determined that, in addition to tumorigenic proteins and enzymes, DNA itself can be used to direct antitumor therapeutic agents. In recent years, the G-quadruplex DNA secondary structure has become of interest in terms of its use as a promising target in tumors. G-quadruplex DNA structures have been proposed to form in the promoter regions of a number of genes involved in cellular proliferation and tumorigenesis, including the proto-oncogene c-*myc* and many others. In this dissertation, the roles of a chair-type G-quadruplex structure in c-*myc* regulation and dysregulation, and its use as a target for antitumor agents, are investigated. It was found that such a structure very likely plays the part of a repressor of c-*myc* transcription, as its absence results in significant upregulation of this gene. Similarly, stabilization of this G-quadruplex with small molecules results in a reduction in c-*myc* expression, an activity lost when this structure can no longer form. It appears that this structure acts as a barrier to transcription factor binding, and Sp1 or an Sp1-like protein(s) may be the critical factor that is not permitted to bind when the quadruplex is present. Disruption of the G-quadruplex in the c-*myc* promoter, and subsequent upregulation of the gene, may play a role in the development of human tumors, possibly in the later stages of tumorigenesis, as mutations that abrogate quadruplex formation are found in patient tumor specimens and cultured tumor cells. Finally, a means of stabilizing the c-*myc* quadruplex with small molecules is developed using rational structure-based

drug design, in order to take advantage of this repressor structure in a therapeutic context. Thus, there is elucidated a novel means of gene regulation heretofore unexplored, one that also represents a new frontier in anticancer drug design.

CHAPTER I

INTRODUCTION

I. The c-*myc* Proto-Oncogene

The c-*myc* oncogene was isolated in 1982 as the normal chicken homologue of a transforming gene (v-*myc*) found in the avian myelocytomatosis retrovirus MC29 (1, 2). The human homologue was later found in multiple copies in the DNA of a malignant promyelocyte cell line isolated from a patient with acute promyelocytic leukemia (3). The c-*myc* proto-oncogene has been mapped to chromosome 8q24.12- q24.13 (4), and was found to be translocated to chromosomes 2, 14 or 22 in Burkitt's Lymphoma cells (5, 6). In the following years, the c-*myc* gene has been found to be deregulated through amplification, translocation or overexpression in a wide variety of human malignancies, including colon and cervical cancers, myeloid leukemias, B-cell and T-cell lymphomas, and glioblastomas (7, 8). These data place c-*myc* in a positive light as a target for anti-tumor therapy.

a. c-MYC as a Transcription Factor

The c-*myc* gene is a member of a highly conserved family of proto-oncogenes, which also includes the neuroblastoma-related N-*myc* and small cell lung cancer-associated L-*myc* (9). Two other members, B-*myc* and S-*myc*, are less well-conserved (9), and are thought to represent pseudogenes, but very little has been published on their

contribution to tumorigenesis. c-*myc* contains a C-terminal basic-helix-loop-helix-zipper (bHLHZ) domain, and functions as a heterodimeric transcription factor with its partner, *myc*-associated protein X (MAX), binding to the E-box DNA sequence (CACGTG) (9-11). The mechanisms through which *myc* activates transcription are many; it can participate in histone acetylation, by recruiting the histone acetyltransferase complex, in chromatin remodeling through interactions with the SWI/SNF complex, and in promoter clearance of RNA polymerase II in a complex with PTEF-b, an RNA polymerase II C-terminal-domain kinase (10, 11). The *Myc*/MAX complex can also repress transcription, by interacting with transcription factors at Inr (initiator elements) or enhancer elements and blocking gene activation (10), though the mechanisms of this activity are less well understood. Originally, c-*myc* was thought to act as an oncogene, promoting cellular growth and division through its role as a transcription factor, activating proliferation-associated genes like *cad*, *cdc25a*, *cdk4*, *elF-4E* and *ISGF3*, and repressing other anti-proliferation genes such as *C/EBPα* and *gadd153/CHOP* (12-15). The c-MYC protein as a transcription factor alone regulates a variety of cellular functions, including vesicular trafficking, DNA replication, metabolism, controlled protein degradation, response to oxidative stress and many more. (For a review, see (10, 16)).

b. Protein-Protein Interactions Involving c-MYC

The activities of this very complex protein are not limited to its role as a transcription factor. c-MYC also displays interactions with other proteins, including the

DNA repair proteins MLH1 and MSH2 (17), retinoblastoma (18) and p53 (9), all tumor suppressor genes of a sort. It is thought that perhaps these interactions sequester the c-MYC protein and thus act as 'anti-oncogenes' in this fashion. The c-MYC protein also complexes with ZNF151 (also called MIZ-1) and negates its growth-inhibitory function, perhaps another means by which c-MYC causes cellular proliferation (19). A similar effect results from c-MYC interactions with SMAD2 and SMAD3, involved in signal transduction through the TGF-β pathway. Interactions of MYC with other proteins can lead to MYC activation or repression, and can affect a myriad of cellular functions, including DNA replication (with MSSP1, a replication origin-binding protein) regulation of cellular structure and transport (with α-tubulin) and signaling by cytokines (with Nmi, an interferon-inducible protein) (10).

c. c-MYC and Apoptosis

Perhaps most counterintuitive is the fact that c-*myc* is also capable of sensitizing cells to apoptotic stimuli, such that deregulated c-*myc* can have opposite effects, based on the internal and external environments of the cell at the time. For example, ectopic c-*myc* expression in keratinocytes can occur with no sign of apoptosis at all, while overexpression of *myc* in murine pancreatic β cells *in vivo* leads to extensive apoptosis and eventually a diabetic phenotype (20). This may act as a 'fail-safe' mechanism to ensure that the upregulation of c-*myc*, without the concomitant presence of the appropriate growth factors and conditions of genomic stability, cannot have an

uncontrolled proliferative effect on the cell (21). For example, c-*myc* overexpression sensitizes cells to apoptosis in the presence of DNA-damaging agents or under serum starvation conditions (22).

The mechanisms through which c-*myc* induces sensitivity to apoptotic stimuli are not yet fully understood at this time. However, in recent years, research has revealed a number of possible candidate pathways for this effect. In the past few years, studies revealed that potentiation of apoptosis by c-*myc* is accompanied by a loss in mitochondrial membrane potential (23) and the release of cytochrome c, an integral player in the apoptotic response (24). This sensitivity is repressed by Bcl-2, which can inhibit cytochrome c release from the mitochondria (25). The nuclear respiratory factor-1 (NRF-1) transcription factor gene bears a non-canonical E-box sequence in its promoter, and NRF-1 regulates a number of mitochondrial genes, including cytochrome c (26, 27). A group at the Fred Hutchison Cancer Research Center has shown that NRF-1 is indeed a direct target of c-MYC, and that inhibition of NRF-1 function by introduction of a dominant-negative NRF-1 construct abrogates the ability of c-*myc* to induce apoptosis (28). Therefore, NRF-1 expression is the most likely agent through which c-*myc* exerts its apoptosis-stimulating effects. However, c-MYC is also known to interact with Bin-1, pRb, p53 and p73, and activates p53 expression; each of these proteins is known to play a role in apoptosis as well (9, 29-31).

d. c-*myc* Control and Regulation

A transcription factor with genome-wide effects must necessarily be under tight control from a number of sources, and c-*myc* is no exception. There are a number of signals that are known to regulate the expression of c-*myc*, culminating in a large number of transcription factors that activate the expression of this gene. Among these are TCF4 (activated by the APC/β-catenin signaling pathway) (32), YY1 (also known as CF1) (33), FBP (far upstream element binding protein) (34), NFκB (35), and CDR-BP (coding region determinant-binding protein) (36). Other factors are known to inhibit *c-myc* transcription, and include MBP-1 (*myc*-binding protein 1) (9, 37), MIF-1 (*myc* intron factor 1) (9, 38, 39), TNF-α signaling (40), TGF-β (through Smad3 and p300 protein (41, 42)) and c-MYC itself (9, 43). The production of c-*myc* is also regulated at the level of transcriptional elongation (attenuation) (7, 44, 45), mRNA stability, partially dependent on the 5' untranslated leader of c-*myc* mRNA (7, 9, 46), translation, including internal ribosome entry sites (47), and protein stability (7). Sparse literature also suggests that a c-*myc* antisense oligonucleotide is transcribed from the first exon of c-*myc*, providing another means through which c-*myc* can regulate its own expression (48, 49).

Aside from the vast number of transcription factors that affect expression of c-*myc*, an additional level of control exists on transcription initiation, in the form of four unique promoters (Figure 1.1). The two major promoters are the P1 and P2 promoters, the latter being responsible for 75-90% of all c-*myc* transcripts in untransformed cells and

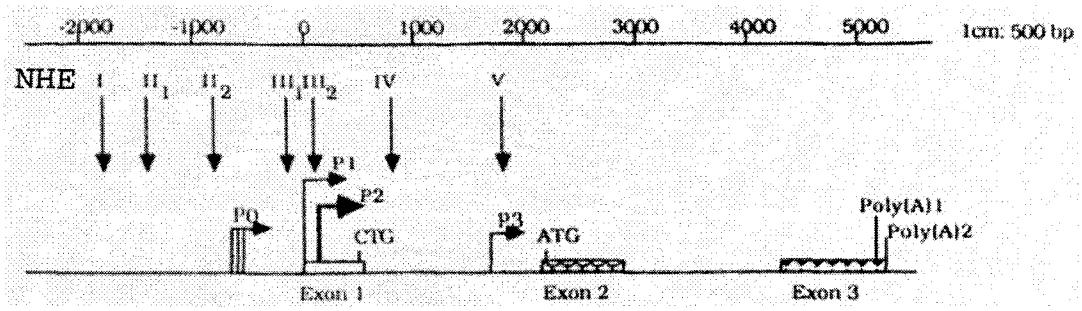


Figure 1.1: Map of the *c-myc* proto-oncogene (adapted from (9))

lying 161 base pairs downstream of P1 (7). In certain transformed cells (7, 50, 51), in response to specific physiological conditions (52-54) and as a result of tumor suppressor loss of function (55) however, a promoter switch is observed, such that P1 becomes the dominant promoter. The best-studied example of this promoter switch occurs in Burkitt's Lymphoma (BL), which is characterized by translocation of the c-*myc* gene to an immunoglobulin gene on chromosome 2, 14 or 22 (6). Whether this is due solely to the juxtaposition of the immunoglobulin regulatory elements or not is still not clear (7, 56). It appears that a region between the P1 and P2 start sites is responsible for controlling the differential promoter usage seen (55).

e. The Importance of the NHE III_1 in c-*myc* Regulation

Chromatin structure surrounding the c-*myc* gene has also been implicated in playing a role in regulation of c-*myc*, based on the emergence of certain nuclease hypersensitivity elements (NHEs) as the c-*myc* gene is activated or repressed (7, 57, 58) (see Figure 1.1). One of these elements, the NHE III_1, has been shown by deletion analysis to be required for anywhere from 75% to 95% of c-*myc* transcription, regardless of whether the P1 or P2 promoter is used (59-61). This sequence is conserved 100% in the chimpanzee gene, and the core guanines (Figure 1.2) are conserved in the mouse gene, as determined by a

Figure 1.2: Sequence of the NHE III_1 (60)
Runs of guanines are underlined.

BLAST search using the NHE III_1 sequence. The fact that this sequence is so integral to c-*myc* expression makes the NHE III_1 an appealing target for anti-cancer agents. This particular sequence is especially unique, in that it contains an inordinate number of guanine residues on a single strand, as shown in Figure 1.2. This G-rich sequence imparts a special property to the DNA in this region: the ability to form a G-quadruplex structure (60, 61). In an *in vitro* system, members of the Hurley laboratory have shown that the double-stranded NHE III_1 sequence will spontaneously convert to the G-quadruplex on the purine-rich strand, and an as yet unidentified structure on the pyrimidine-rich strand.

II. G-Quadruplexes

a. Introduction

In vitro, guanine-rich DNA sequences are able to form G-quadruplexes, four-stranded secondary structures in which guanines from each strand interact with one another through Hoogsteen or reverse Hoogsteen base-pairing to form stacked guanine tetrads (Figure 1.3). Such structures can be formed by a single strand folding back on itself (intramolecular), by two hairpin strands, or by four separate strands (intermolecular) (62). Even within a single type of quadruplex, there is a great deal of polymorphism; for example, intramolecular quadruplexes can exist in chair, basket, propeller or heptad-tetrad forms (63-65). Two of these, the basket and chair forms, are shown in Figure 1.3. Each form has a unique topology, both in cross-section (i.e. the sizes

of each of the grooves between each of the four strands) and in terms of overall shape (i.e. three-dimensional space-filling) (64).

b. Biological Relevance of G-Quadruplexes

Quadruplex DNA sequences with the propensity to form G-quadruplexes have been found in telomeric DNA (66-68), immunoglobulin heavy chain switch regions (67, 69), the control region of the retinoblastoma susceptibility gene (70), and in the promoters of many oncogenes, including c-*abl*, c-*fos*, c-*jun*, c-*src*, c-*myb*, c-*sis* and, as mentioned earlier, c-*myc* (60). More recently, a French research group has postulated that the formation of a G-quadruplex is an integral step in the pre-integration step of HIV-1 (71). This hypothesis is strengthened by the previous discovery of an effective anti-HIV-1 oligonucleotide that forms a G-quadruplex under physiological conditions (72, 73). Roles for G-quadruplex structures have been proposed in a number of diverse cellular processes, including recombination (the immunoglobulin type switch and HIV-1 integration), transcription (the insulin gene) and telomere maintenance (64). In 2001, antibodies with specificity for G-quadruplex DNA were found to recognize the macronuclei of the lower eukaryote *Stylonychia lemnae*, leading researchers to believe that the telomeres of this organism adopt a G-quadruplex structure during its non-replicative phase (74). This evidence suggests that G-quadruplexes are not simply chance structures that may form in the DNA, but are instead biologically significant regulatory structures.

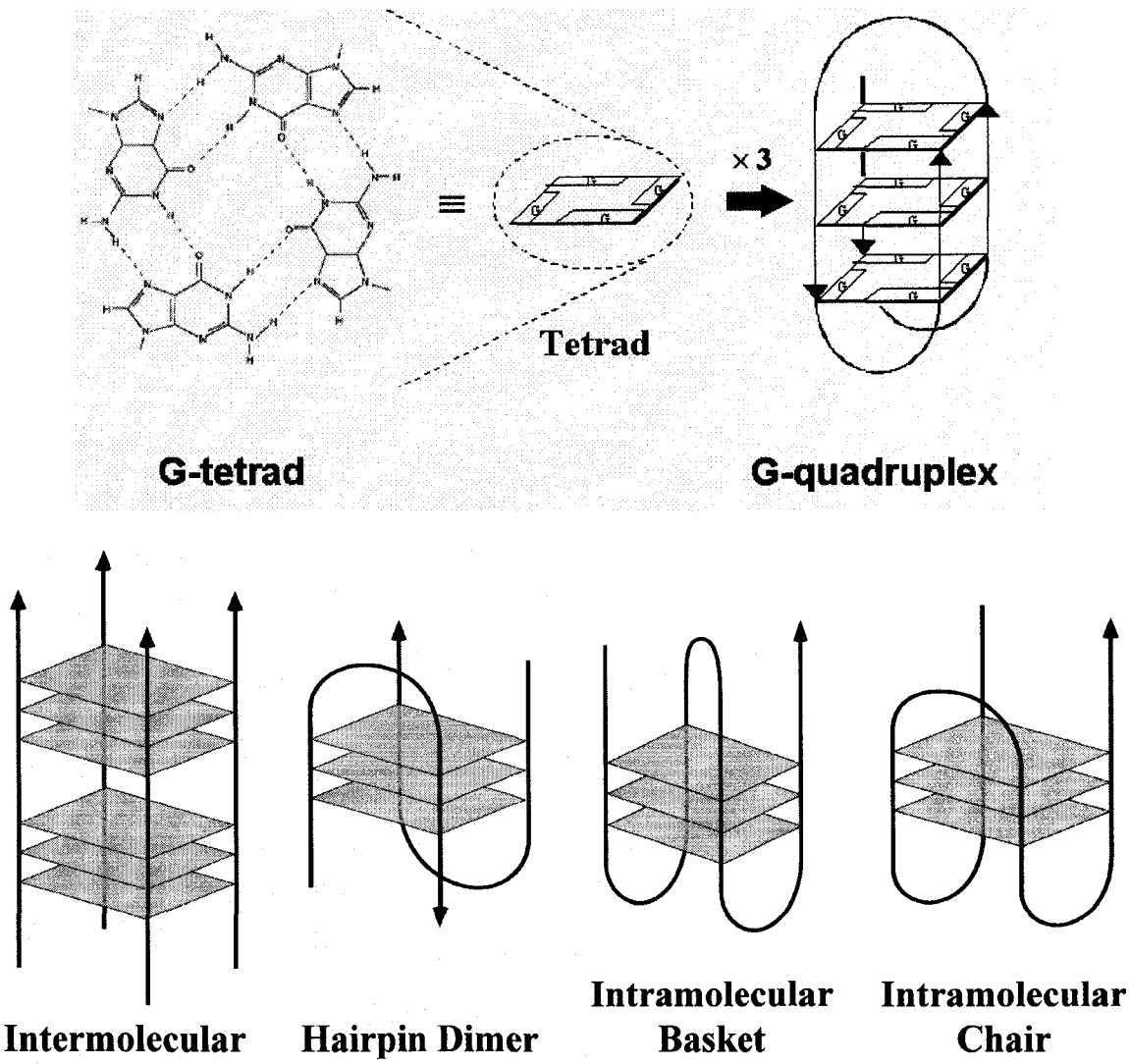


Figure 1.3: The formation of a G-quadruplex and polymorphism among quadruplexes.

G-quadruplex DNA may also play a role in human disease. Both Bloom's and Werner's syndromes are associated with a loss of function of a helicase that is believed to unwind inappropriate quadruplex DNA structures (75-78). Triplet repeat diseases, including Unverricht-Lundborg type progressive myoclonous epilepsy (79) and Fragile X syndrome (80), a mental retardation disease which affects approximately one of every 2000 people (81) are both believed to be caused by quadruplex formation in their respective genes. Also, formation of quadruplex DNA may disrupt transcription of the insulin gene, which may be a mechanism for the development of insulin-dependent diabetes mellitus (IDDM) (82, 83). If such structures can indeed cause human disease, there must exist some means by which they can be resolved, in order to prohibit these conditions from developing.

c. Proteins That Bind to and Act upon G-Quadruplex DNA

The existence of sequences with the propensity to form G-quadruplexes suggests that these structures are important biological signaling molecules, but in order for them to act as such, there must be mechanisms in place to allow their formation from and dissolution back to unstructured single-stranded DNA. A number of helicases exist that preferentially unwind G-quadruplex DNA, including the SV40 large T antigen helicase (84), the human Bloom's syndrome helicase (BLM) (77) and Werner's syndrome

helicase (WRN) (78) and the yeast Sgs1 helicase (85). As mentioned above, if either the Bloom's or Werner's syndrome helicase function is lost, disease is the result. As well, G-quadruplex-interactive small molecules can inhibit the helicase activity of both of these proteins (86).

Similarly, other proteins are known to bind to and cause the formation and stabilization of G-quadruplexes. These include the β subunit of the *Oxytricha nova* telomere-binding protein (87-89), the *Tetrahymena thermophila* quadruplex binding protein TGP (90), RAP1, STM1 and Hop1 in yeast (91-93), and the mammalian proteins LR1 (involved in recombination) (94, 95), MyoD (96), QUAD (97, 98), two heterogeneous nuclear ribonucleoproteins, qTBP42 and uqTBP25 (99) and possibly human topoisomerase I (100). G-quadruplex specific nucleases have also been characterized, and included among them are the murine XRN1p (101), yeast KEM1 (102, 103) and human GQN1 proteins (104). Again, that these proteins are present in nature suggests that their substrate, the G-quadruplex, has an important role as a biological regulatory molecule. This then implies that G-quadruplexes may be viable targets for pharmacological agents, given the regions of the genome in which G-quadruplex-amenable sequences have been found.

d. Small Molecules That Bind to G-Quadruplexes: TMPyP4

In recent years, small molecules have been developed that stabilize G-quadruplexes in vitro. Among these are the perylene derivative PIPER, which can also drive G-quadruplex formation (105) and the cationic porphyrin 5,10,15,20-tetra-(N-

methyl)-4-pyridyl porphine (TMPyP4, see Figure 1.4) (106). Of interest is the finding that a positional isomer of TMPyP4, 5,10,15,20-tetra-(N-methyl)-2-pyridyl porphine (TMPyP2, see Figure 1.4) is much less able to interact with and stabilize G-quadruplex structures (106). Of note is the fact that TMPyP4 is also able to bind I-tetraplex DNA, with TMPyP2 much less able to do so (107). These compounds were originally developed as telomerase inhibitors, due to the propensity of telomeric sequences to form G-quadruplexes. It was believed that sequestration of one DNA strand of the telomere by TMPyP4 would inhibit telomerase activity, and this was later found to be true. Since TMPyP4 has been shown to be a G-quadruplex-interactive agent, and since the NHE III_1, which has an important role in allowing c-*myc* transcription, also has the ability to form a G-quadruplex *in vitro*, the effect of TMPyP4 on c-MYC production has come under investigation. As I shall show later, TMPyP4 is able to reduce c-*myc* mRNA and protein in cell culture in a time-dependent manner. It is this effect of TMPyP4 that I intended to characterize and address.

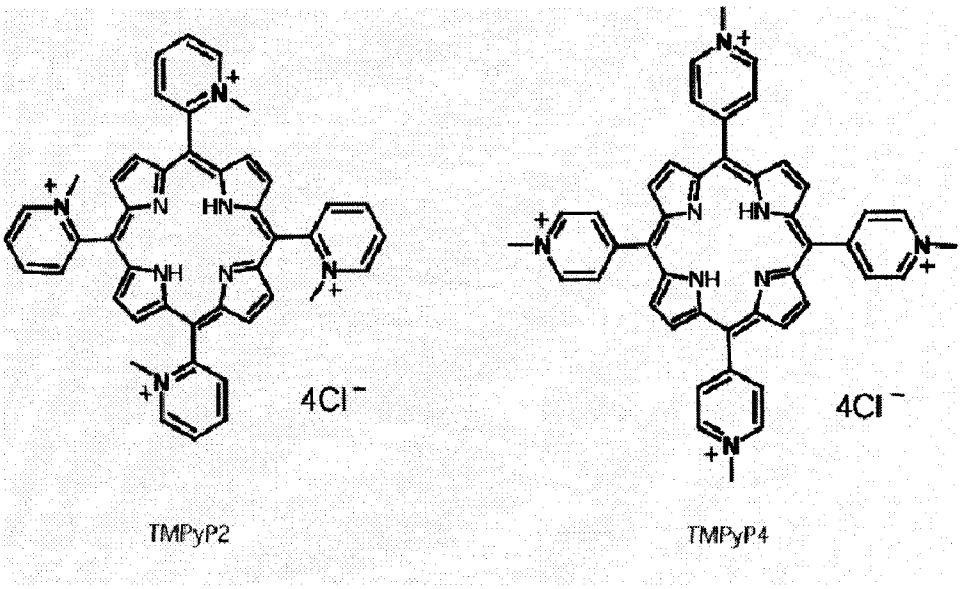


Figure 1.4: Structures of TMPyP2 and TMPyP4

III. Protein Factors That Bind to the NHE III_1 of the c-*myc* Promoter

As it has been shown that the NHE III_1 of the c-*myc* promoter is required for nearly all transcription of this gene, the protein factors that interact with it *in vivo* or *in vitro* have been studied and, in many cases, characterized in-depth. These proteins include hnRNP K, hnRNP A1, hnRNP A2, hnRNP B1, NM23-H2, CNBP, NSEP-1, Sp1, Sp3, CTCF, MAZi, THZif-1 and c-MYB. (Reviewed in (61)). It is of interest that many of these factors do not bind with highest affinity to the double-stranded form of the NHE III_1, but instead to one or the other single strand (or to both). Of the list above, hnRNP K (108, 109), MAZi (110) and THZif-1 (111) recognize the pyrimidine-rich strand, hnRNP A1 (60, 108) and CNBP (112) recognize the purine-rich strand, and NM23-H2 has the ability to bind to both the purine-rich (113) and pyrimidine-rich (114) strands, as well as the duplex NHE III_1 DNA (114). This suggests that a melting of the duplex DNA in this region may be important in the regulation of c-*myc*. Three of these proteins, hnRNP K, NM23-H2 and CNBP, are known to have other activities aside from DNA binding, and have been characterized in detail.

a. hnRNP K Binds the C-Rich Strand and May Recruit RNA Polymerase II

In the late 1980s, a factor with the characteristics of a ribonucleoprotein (RNP) was found to interact with the pyrimidine-rich strand of the NHE III_1 sequence and was proposed to be involved in c-*myc* regulation (115). This protein was later determined to be heterogeneous nuclear RNP K (hnRNP K), which can activate transcription from the NHE III_1 region (116-118). The mechanism for this activation may involve two different actions, working independently and concurrently. First, it is thought that hnRNP K binding to the pyrimidine-rich strand of the NHE III_1 may facilitate bending of this region, allowing it to act as a DNA "hinge", and bringing activator proteins bound further upstream to the TATA box (118). Interestingly, hnRNP K has been shown to interact with TATA Box-Binding Protein (TBP) of the eukaryotic RNA Polymerase II complex, implying that it may recruit the basal transcription machinery as a means of carrying out this transcriptional activation (117). It has not yet been shown that this occurs at the c-*myc* promoter, however. hnRNP K also has molecular interactions with a number of transcriptional activators (like YY1), repressors (such as ZIK-1), signal transducers (including c-SRC) and many others (117, 119-125).

hnRNP K has been suggested to be a "docking platform" for proteins on DNA and RNA, regulating both transcription and translation at the sites to which it binds (117). That hnRNP K binds to the non-duplex form of the NHE and has RNA Polymerase II-recruiting ability again suggests that this region of the c-*myc* promoter may not be in B-DNA form during part of the activation process of this gene; this region of DNA must at

least be single-stranded to allow recruitment of hnRNP K. It is known *in vitro* that the C-rich strand of the NHE III_1 may also have secondary structure, forming an i-tetraplex (Figure 1.5) under certain conditions (61). The i-tetraplex has as its building block the hemiprotonated cytosine$^+$-cytosine base pair and like the G-quadruplex, can be formed within a single nucleic acid molecule or among two or more molecules.

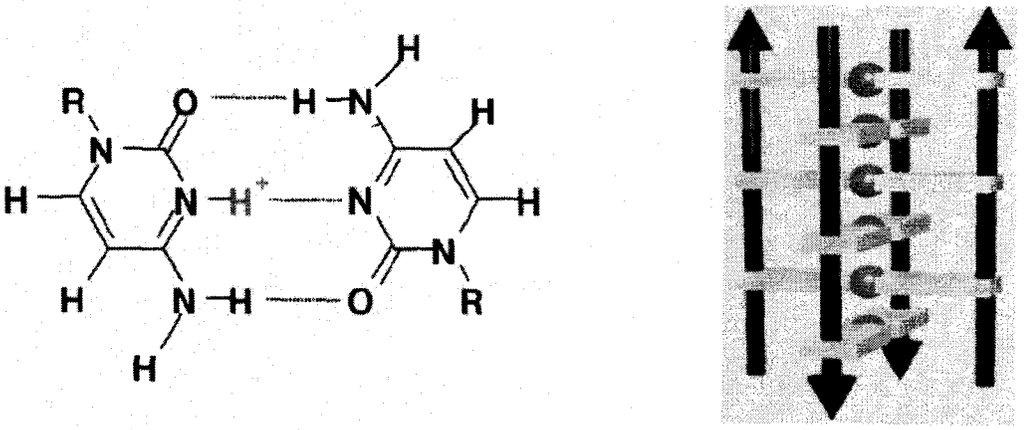


Figure 1.5: Hemiprotonated C^+-C base pair, and i-tetraplex (from (61))

b. The NM23-H2 Factor: a Multifunctional Transition Protein

NM23-H2 (non-metastatic 23-human isoform 2) (59) is one of several members of a large family of nucleoside diphosphate kinase (NDPK) proteins, of which at least five and as many as eight have been characterized and identified, and others are still under study (126-128). The NM23 proteins are found in many types of organisms, from mice and humans to amphibians, insects and even bacteria (126, 129-133). These proteins were first documented as metastasis suppressors in certain tumor types (126, 128, 131, 133-139), but are likely involved in normal cellular development and in the differentiation of cells (126). NM23-H2 itself is known to act as a transcriptional regulator of a number of genes, including the platelet-derived growth factor alpha subunit (PDGF-A) (133, 140) and c-*myc* (59, 126, 133, 137, 139-144). In this latter respect, NM23-H2 was initially given the name purine-binding factor, or PuF (142).

In recent years, NM23-H2/NDPK-B/PuF (hereafter known as NM23-H2) has been found to have a number of distinct and separate activities. Like its homologues, it can catalyze the transfer of phosphate groups from nucleoside triphosphates to nucleoside diphosphates, through a "ping-pong" mechanism involving phosphorylation of histidine-118 as an intermediate (59, 113, 114, 126, 127, 130, 132, 133, 138, 144-149). As its role as a transcription factor suggests, this hexameric protein also bears a sequence-specific DNA recognition surface along its equator (113, 138, 144, 147, 149). NM23-H2 can recognize DNA in many forms, including duplex, single-strand and possibly G-

quadruplex, with approximately equal affinities (113, 133, 144, 145, 147). In addition, it bears a second, covalent, DNA-binding domain, which overlaps the nucleotide-binding site involved in its NDPK activity (138). There is some evidence, though debated, that suggests that NM23 proteins may have a protein kinase activity, at serine, histidine, aspartate or glutamate residues, and that this activity correlates with motility suppression in the case of NM23-H1 (126, 138).

Finally, most applicable to this dissertation, NM23-H2 has the ability to catalyze a DNA cleavage/religation reaction, through a coordinate and reversible glycosylase/lyase function, possibly involved in a DNA repair function of this enzyme (127, 133, 134, 138). We believe that this activity may be involved in a transition within the NHE III_1 region, between the duplex/partial single-stranded forms and the G-quadruplex form. Other researchers have also postulated that the ability of NM23-H2 to activate c-*myc* transcription is linked to its ability to change DNA promoter structure (126, 133, 134, 138, 144). One group has reported that NM23-H2 can catalyze the conversion of the duplex NHE III_1 to complexes of single-stranded DNA and protein (114), so a conversion from quadruplex to unstructured single-strand DNA may also be possible. Since it has been shown in this laboratory that the NHE III_1 DNA sequence can spontaneously convert to the G-quadruplex form under certain conditions (elevated temperature or possibly negative supercoiling strain, data not shown) we posit that NM23-H2 is involved in the back-transition, from G-quadruplex to duplex/partial single strand, perhaps to allow other transcription factors access to their respective binding sites within the c-*myc* promoter.

c. CNBP Activates c-*myc* from the Purine-Rich Strand of the NHE III_1

Cellular Nucleic Acid Binding protein (CNBP) is a 19 kDa RNA- and DNA-binding protein, consisting of seven consecutive zinc finger repeats (150). Two isoforms of this protein exist, both of which are produced from the same gene by differential splicing (150). CNBP was first identified through its ability to bind to and activate transcription from sterol response elements (SRE) in certain promoters (151). Of relevance to this project, a protein factor is known to activate transcription of the c-*myc* gene from the NHE III_1 purine-rich strand, and this activation requires the presence of divalent metal cations, specifically zinc, as EGTA abolishes this effect (112). It was later found that this protein factor was indeed CNBP, which could stimulate transcription from the NHE III_1 in a chloramphenicol reporter assay (112). The ability of CNBP to bind to quadruplex DNA has not yet been studied, but the purine-rich strand of the NHE III_1 readily adopts a quadruplex structure under physiological conditions of salt and temperature and in the absence of other influences, in experiments performed in the Hurley laboratory. It is yet unclear if CNBP would be prohibited from binding to this secondary structure, or if the secondary structure is the form of the NHE III_1 that CNBP recognizes. As the consensus SRE sequence (AAAGTGGGGAAAA) (151) does not appear to have intramolecular quadruplex-forming amenability, the former may be true.

d. hnRNP A1 May Be a Quadruplex-Specific Protein

The hnRNP A1 protein, as with other hnRNPs, has a variety of activities, involving itself in RNA splicing, pre-mRNA packaging and transport, renaturation of single-stranded DNA and RNA, and may be hijacked by the mouse hepatitis virus for its own RNA-dependent RNA transcription, to name a few (152, 153). Recently, evidence has surfaced that suggests hnRNP A1 may have functions in binding to and acting upon G-quadruplex DNA (and perhaps RNA as well). In late 2002, the mouse protein UP1, which is a proteolytic product representing the first 195 amino acids of hnRNP A1, was found to specifically unfold G-quadruplex structures in the mouse minisatellite Pc-1, and in the human telomeric repeat, allowing DNA replication to progress normally (154). As mentioned earlier, the telomeric sequence is known to form such structures *in vitro*, and hnRNP A1 is required in part for maintenance of telomere length (118), suggesting that it may be involved in unwinding such secondary structures to allow access of telomerase. Also, hnRNP A1 bears close homology to the aforementioned uqTBP25, another proposed quadruplex-binding protein (155). Finally, in a study designed to determine the binding sequence for hnRNP A1, the G-rich sequence with the greatest affinity was found not only to aggregate into higher-order structures in solution, but this higher-order structure appeared to contribute to the ability of hnRNP A1 to bind with high affinity (152). These "higher order structures" may well represent guanine quadruplexes, and the authors of this study consider this possibility. All of these data at least imply that hnRNP

A1 may be involved in the recognition and perhaps the formation or dissolution of G-quadruplex DNA.

IV. Preliminary Studies

These data have been included in this dissertation in order to lead the reader through the experiments that led us to study the G-quadruplex-forming region of the c-*myc* promoter. Because these results do not relate to any other chapter of this dissertation, they are included in this introductory chapter as means of introducing the reader to the project discussed in the rest of the text.

a. TMPyP4 Reduces Telomerase Activity in Cell Culture (from (156))

It has already been established that TMPyP4 can bind to and stabilize DNA G-quadruplex structures in human telomeric sequences. Knowing this, we were interested in the effect that this interaction might have on telomerase activity. MiaPaCa-2 cells were treated with 100 μM TMPyP4 for 12, 24, 36 or 48 hours, and total protein was extracted for use in a TRAP assay. As seen, it was found that TMPyP4 could indeed inhibit telomerase activity in a time-dependent manner (Figure 1.6). A similar pattern was seen when HeLa S_3 cells were treated with TMPyP4 (data not shown). TMPyP2, an analogue of TMPyP4 that is less able to interact with G-quadruplex structures, had a much less pronounced effect on telomerase. Although this inhibition cannot be directly linked to an interaction with or stabilization of G-quadruplex DNA, the fact that TMPyP2, with

essentially the same characteristics as TMPyP4 save for its ability to bind G-quadruplexes, was less able to have a hindering effect on telomerase than TMPyP4 suggests that G-quadruplex DNA is involved.

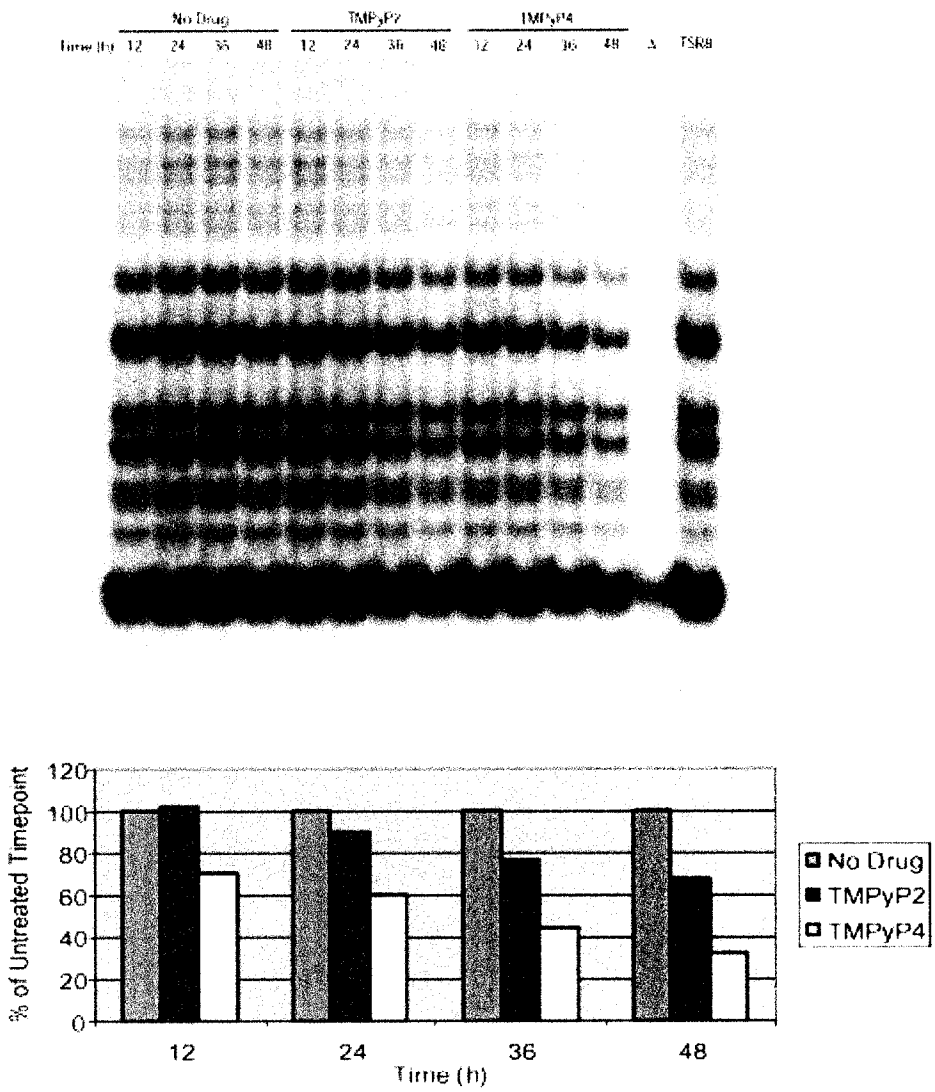


Figure 1.6: Effect of TMPyP2 and TMPyP4 on telomerase activity in MiaPaCa-2 cells.

b. cDNA Microarray To Determine Whole-Genome Effects of TMPyP4 (from (156))

In order to determine the effects aside from telomerase inhibition that TMPyP4 might have, cDNA microarray analysis was performed on a 5000-gene array, as well as on a 500 cancer-associated gene array, using HeLa S_3 cells treated with 100 μM TMPyP2/TMPyP4 or H_2O for 48 hours. The results are shown in Figure 1.7. A number of genes were affected similarly by both TMPyP2 and TMPyP4, and are listed in Table 1.1A. This list indicates the induction of certain genes involved in responses to reactive oxygen species, which comes as no surprise, as both TMPyP2 and TMPyP4 are capable of producing these species when exposed to light. Both porphyrins also caused a decrease in the expression of metallothioneins, which is also expected, as TMPyP2 and TMPyP4 can coordinate certain divalent cations in the center of the porphine core, including Mg^{2+}, Mn^{2+}, and Sr^{2+}. The titration of these ions from solution likely signals a decrease in the need for the metallothioneins, hence the downregulation seen. An abridged list of genes affected by TMPyP4, but not by TMPyP2, is given in Table 1.1B. It was expected that any changes resulting from treatment with TMPyP4 but not TMPyP2 may be due to the ability of TMPyP4 to bind to and stabilize quadruplex DNA. Among the genes downregulated by TMPyP4 is c-*myc*, which is known, like the human telomere, to contain DNA with the capability of forming a G-quadruplex structure. It was also noted that TMPyP4 treatment similarly affected c-*fos* and c-*myb*, two other genes that bear G-

quadruplex-amenable sequences in their promoters (60). This warranted further investigation into the effect of TMPyP4 on c-*myc*.

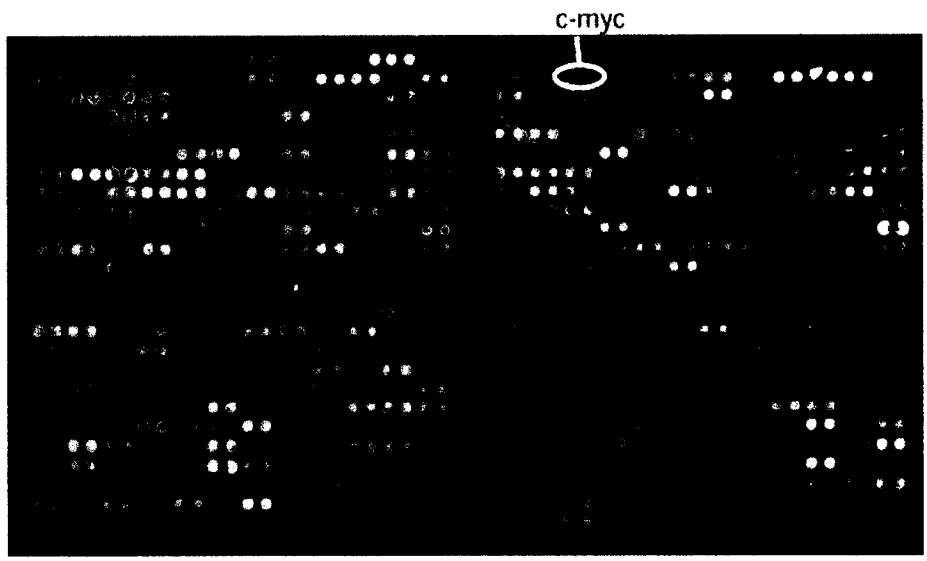


Figure 1.7: cDNA microarray for TMPyP4 treatment of HeLa S_3 cells.

Induced Genes	Downregulated Genes
Oxidation Reduction Genes	**Metallothionein Genes**
Cystathionase	Metallothionein 2H
Lactate dehydrogenase	Metallothionein 1L
Cytochrome p450	Metallothionein 1H
Thioredoxin	Metallothionein 2A
Superoxide dismutase-1	
Superoxide dismutase-2	
Glutathione S-transferase M4	
Proteasome Genes	
Macropain alpha	
Macropain 26S	

Table 1.1A. Genes Affected by Both TMPyP4 and TMPyP2 Treatment

Induced Genes	Downregulated Genes
Apoptosis Genes	**c-*MYC*-Associated Genes**
Caspase 1	c-*MYC*
	Ornithine decarboxylase
	CDC25A
Cell Signaling Genes	**Cell Cycle Genes**
TGF-β	*CDK-4*
CD47	*CDK-6*
CD9	Cyclin B1
CO-29	
RAB-1A	
RAB9	
Proliferation-associated gene A	
DNA Repair Genes	**Cell Signaling and Oncogenes**
MLH1	c-*FOS*
ERCC5	c-*MYC*
	JUN-B
	c-*MYB*
	STAT-1
Heat Shock Genes	**Oxidation Reduction Genes**
HSP27	Cytochrome C1
HSP10	

Table 1.1B. Genes Affected by Only TMPyP4

c. TMPyP4 Causes a Decrease in c-*myc* Expression; Northern Blot Analysis (from (156))

In order to verify the effect on c-*myc* expression by TMPyP4, northern blot analysis was performed for the c-*myc* transcript. HeLa S_3 cells were treated with 100 μM TMPyP4 for 12, 24, 36 or 50 hours, and the total RNA extracted, electrophoresed, and used as a target for northern blot analysis, with labeled c-*myc* cDNA as a probe. The results are shown in Figure 1.8. As is evident, there is an overall decrease in the abundance of c-*myc* mRNA with continued exposure to TMPyP4. Due to the time and resource investment involved in northern blot analysis, we chose to use RT-PCR for further assays of the transcriptional effects of TMPyP4.

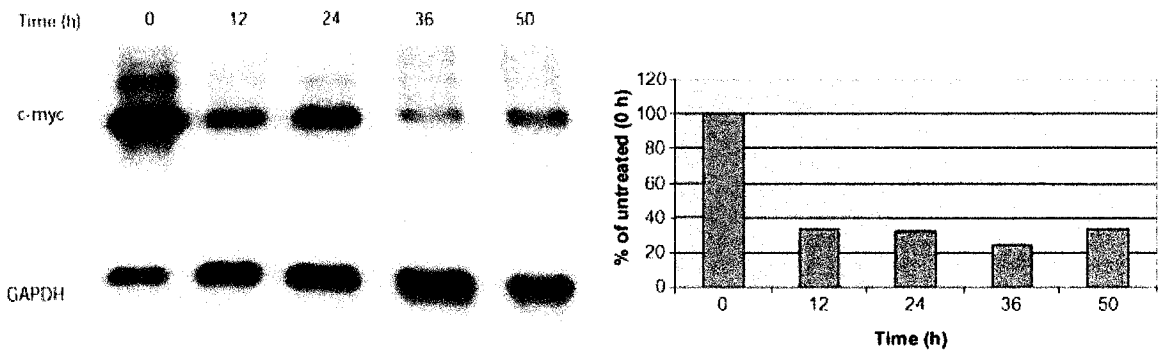


Figure 1.8: Effect of TMPyP4 on c-*myc* mRNA in HeLa S_3 cells – northern Blot

V. Statement of the Problem

We had determined that TMPyP4, a quadruplex-interactive agent, could reduce c-*myc* expression in cell culture, while its positional isomer, TMPyP2, could not. The purpose of this research was to determine why this occurs. **Hypothesis: There is a G-quadruplex in the NHE III_1 of the c-*myc* promoter, and stabilization of this structure leads to a reduction of c-*myc* transcription due to blockade of transcription factor binding, while loss of this structure leads to c-*myc* overexpression and tumorigenesis.** To test this hypothesis, our studies focused on the following aims: **Specific Aim 1:** Verify the effect of TMPyP4 on c-*myc* expression, and the existence and effect of the G-quadruplex secondary structure, in the c-*myc* promoter. **Specific Aim 2:** Characterize the interaction between TMPyP4 and this secondary structure, and establish the requirement of this structure and the NHE III_1 for the effect of TMPyP4. **Specific Aim 3:** Determine the role of the chair-type G-quadruplex in tumorigenesis. **Specific Aim 4:** Begin to investigate the mechanism through which the G-quadruplex exerts its effects. **Specific Aim 5:** Develop a G-quadruplex-interactive compound superior to TMPyP4 in its ability to reduce c-*myc* expression.

CHAPTER II

MATERIALS AND METHODS (modified from (156, 157))

Cell Culture

a. Human Cells

HeLa S_3 (human cervical carcinoma metastasis), MiaPaCa-2 (human pancreatic tumor), Ramos and CA46 (both human Burkitt's Lymphoma) were obtained from ATCC; ForF cells (human foreskin fibroblast cells) were prepared fresh at the Arizona Cancer Center from newborn foreskins. ForF, MiaPaCa-2, Ramos and CA46 cells were cultured at 37 °C in RPMI medium (Cellgro) with 10% fetal bovine serum (FBS), 50 U/ml penicillin G sodium, and 50 U/ml streptomycin sulfate (Gibco/BRL). HeLa S_3 cells were cultured in Dulbecco's modification of Eagle's Medium (DMEM, Cellgro) with 10% FBS, 50 U/ml penicillin G sodium, 50 U/ml streptomycin sulfate. Adherent cells were grown to 80% confluency and passaged at 1:10 in the following fashion. The medium was aspirated, and the cells were washed with 1× PBS (Cellgro). Sufficient trypsin (Gibco/BRL) was added to cover the cells, and cells were incubated at room temperature for approximately 3 min, or until cells detached from the flask with firm tapping. The trypsin was neutralized with an equal volume of culture medium, and the cells were counted using a haemocytometer. $2 \times 10^5 - 1 \times 10^6$ cells were removed for cell cycle analysis, if needed. Suspension cells were grown to approximately 5×10^6 cells/ml, and passaged at 1:10 into fresh medium. The remaining cells were then pelleted by

centrifugation at 500× gravity (g), the supernatant was aspirated, and the pellet was washed in PBS, recentrifuged, and frozen at –80 °C.

b. Insect Cells

S2 Schneider's *Drosophila* cells were purchased from Invitrogen, and were cultured in Schneider's *Drosophila* medium (Invitrogen) supplemented with 10% FBS, 50 U/ml penicillin G sodium, and 50 U/ml streptomycin sulfate (Gibco/BRL). Cells were passaged at 1:5 when they reached approximately 10^7 cells/ml. When cells were to be frozen for stock maintenance, 1.1 x 10^7 cells/ml were resuspended in freezing medium (45% conditioned complete Schneider's *Drosophila* medium, 45% fresh complete Schneider's *Drosophila* medium, 10% DMSO) and frozen to -80°C in a control rate-freezing container (1°C per minute).

Treatment with Drugs

Cell cultures were permitted to reach ~50% confluency before drug was added. Drug concentrations were as listed below in Table 2.1. Drugs were diluted in media in which the cells were normally cultured. Cells were washed once with PBS, and new medium containing drug was added directly to the flask. Cells were harvested as noted above 12, 24, 36, 48, or 50 h after initial treatment, concurrently with untreated cells. Timepoints were collected in duplicate for each treatment.

Chemical Name	Abbreviated Name	Concentrations Used
5,10,15,20-tetra-(N-methyl)-4-pyridyl porphine	TMPyP4	50 μM, 100 μM
5,10,15,20-tetra-(N-methyl)-2-pyridyl porphine	TMPyP2	50 μM, 100 μM
Doxorubicin	Doxorubicin	0.2 mg/ml
Taxotere/Docetaxol	Taxotere	100 nM
Gemcitabine/Gemzar	Gemcitabine	10 nM
5,10,15,20-tetra-(1,2-dimethylpyrozolium-4-yl)-21H, 23H-porphine tetrafluorosulfonate	OMPzP4	100 μM
5,10,15,20-tetra-(N-methyl)-3-pyridyl –26-28-diselena sapphyrin	Se_2SAP	5 μM, 10 μM, 25 μM, 50 μM, 100 μM
5,10,15,20-tetra-(N-methyl)-3-pyridyl –21-23-diselena porphyrin	Se_2Py3	5 μM, 10 μM, 25 μM, 50 μM, 100 μM
5,10,15,20-tetra-(N-methyl)-3-pyridyl –21-monoselena porphyrin	SeHPy3	100 μM
5,10,15,20-tetra-(N-methyl)-3-pyridyl –21-23-dithia porphyrin	S_2Py3	100 μM
5,10,15,20-tetra-(N-methyl)-3-pyridyl –21-monothia porphyrin	SHPy3	100 μM
5,10,15-tri-(N-methyl-3-pyridyl)-corrole	Corrole P3	100 μM
5,10,15-tri-(N-methyl-4-pyridyl)-corrole	Corrole P4	100 μM

Table 2.1 Drugs Tested and Concentrations Used

RNA Extraction

Cell pellets were lysed in Buffer RLT from the RNeasy RNA Mini Extraction Kit (QIAGen) and homogenized using a QIAShredder (QIAGen). RNA was extracted according to the protocol included with the RNA Extraction Kit and eluted in distilled, deionized water with 0.1% diethyl pyrocarbonate (DEPC-H_2O, Sigma) to a final volume of 30 µl. RNA was quantitated by ultraviolet spectrophotometry, resuspended to a convenient concentration for downstream applications, and stored at –80 °C. Later in the study, we began using the Clontech/Machery-Nagel Nucleospin RNA II Kit. Cell pellets were lysed in Buffer RA I supplemented with 1% β-mercaptoethanol. RNA was extracted according to the protocol included with the kit, and eluted in 50 µl of nuclease-free water included in the kit. RNA was quantified, resuspended and stored as above.

Protein Extraction and Quantitation

Cell pellets were lysed in 100–150 µl Nonidet-P40 lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.02% sodium azide, 100 µg/ml PMSF, 1% Nonidet-P40), the lysates were centrifuged at >14,000× g, and the supernatants were collected and stored frozen at –80 °C. Lysates were quantified in 96-well plates by BCA Protein Assay (Pierce).

TRAP Assays

1.5 μL of protein was used to determine the amount of active telomerase enzyme in each total protein extract, using the TRAPeze radioisotopic detection kit (Intergen), a telomeric repeat amplification protocol (TRAP). The protocol included with the kit was followed.

Reverse Transcription

Total RNA was used as a template for reverse transcription, using the following protocol: each 20 μl reaction contained 1× Omniscript RT Buffer (QIAGen), 500 μM each of dCTP, dATP, dGTP, and dTTP (QIAGen), 1 μM Oligo dT primer (Ambion), 1 μM random decamer primers (Ambion), 1 U Omniscript reverse transcriptase (QIAGen), 0.1% diethyl pyrocarbonate in water (DEPC-H_2O), and 2 μg of total RNA. Mixtures were incubated at 37 °C for 60 min for reverse transcription, then at 92 °C for 10 min to inactivate the enzyme. Later in the study, for financial reasons, a different company supplied the reagents for these studies. For these later studies, the following protocol was used: Each 20 μl reaction contained 1× Reaction buffer (Fermentas), 500 μM each of dCTP, dATP, dGTP, and dTTP (Fermentas), 0.5 μg Oligo$(dT)_{18}$ primer (Fermentas), 1 U M-MuLV Reverse Transcriptase (Fermentas), 20 U Ribonuclease Inhibitor (Fermentas), DEPC-H_2O, and 2 μg of total RNA. RNA, oligo$(dT)_{18}$, and DEPC-H_2O were preheated to 70°C for 5 minutes. Reaction buffer, ribonuclease inhibitor and deoxynucleotides were added, and the mixture was incubated at 37°C for 5 minutes. Finally, reverse transcriptase was added, and the reaction was allowed to proceed for 1 hour at 37°C. The

reaction was stopped by a 10 minute incubation at 72°C. All incubations were carried out in a DNA Engine Peltier Thermal Cycler (MJ Research). Reaction products were held at 4 °C and stored at -20°C.

Polymerase Chain Reaction (PCR)

PCR was performed according to the following protocol: each 50-µl reaction contained 1× PCR Buffer (Promega), 50 µM each of dCTP, dATP, dGTP, and dTTP (Promega), 0.5 µl β-actin primer pair (Ambion), 2.5 U Taq Polymerase (Promega), 0.5 µM of each appropriate primer (Table 2.2), 0.1% DEPC-H_2O, and 2 µl of the reverse transcriptase reaction detailed above. The reactions were incubated in a DNA Engine Peltier Thermal Cycler as follows: 95 °C, 5 min; (95 °C, 1 min; 59 °C, 1 min, 10 sec; 72 °C, 1 min, 30 sec) 30× for c-*myc* and hTERT, 40× for sequencing; 72 °C, 5 min. PCR products were held at 4 °C and stored at -20°C. To visualize the reaction products, DNA loading dye (Fermentas) was added to the samples to a final concentration of 1×, and 10-15 µl of this mixture was electrophoresed through a 1% agarose/1× TAE gel. DNA in the gel was visualized with ethidium bromide under ultraviolet light, and photographed using the Eagle Eye II system (Stratagene).

Real-Time Polymerase Chain Reaction

Real-time PCR was performed using the QuantiTect SYBR Green PCR Kit (Qiagen). Briefly, each 25 µl reaction contained 12.5 µl of 2× reaction buffer, 0.5 µl of each mycRT primer (see table 2.2) or 1 µl of Actin Primer Pair (Ambion), 1 µl of cDNA,

and 10.5 µl of distilled, deionized water. The reactions were incubated in a DNA Engine Opticon Continuous Fluorescence Detector (MJ Research) as follows: 95 °C, 14 min; (95 °C, 1 min; 59 °C, 1 min, 10 sec; 72 °C, 1 min, 30 sec; plate read for fluorescence) 40×; 72 °C, 10 min. A melting curve was then performed over the temperature range of 55°C to 95°C, with a plate read at each 1°C increment, to ensure that only a single PCR product had been produced in each well. Quantitation was performed using the integral Opticon software, and data are presented as -fold expression of c-*myc* (normalized to β-actin) versus an untreated control.

Primer Name	Sequence (5' to 3')	Product Size (bp)
c-*myc*-1	AGAGAAGCTGGCCTCCTACC	2166
c-*myc*-2	AGCTTTTGCTCCTCTGCTTG	
hTERT-1	GCCTCTTCGACGTCTTCCTA	1493
hTERT-2	CCCCAATTTGACCCACAG	
*myc*RT-1	CTTGGCGGGAAAAAGAACGC	189 (for real-time PCR)
*myc*RT-2	TATTCGCTCCGGATCTCCCT	
NHEseqfw	GACAAGGATGCGGTTTGTCA	275
NHEseqrv	GAGATTAGCGAGAGAGGATC	

Table 2.2 Primers for Polymerase Chain Reaction and Expected Product Sizes

Northern Blot Analysis

Selected results obtained by cDNA microarray analysis were confirmed experimentally by northern blot analysis by Dr. Haiyong Han. Briefly, equal amounts of total RNA from HeLa S_3 were size-separated on a 1% denaturing formaldehyde agarose gel, transferred to a nylon membrane, and hybridized with radio-labeled probes specific to the gene to be analyzed. Probe templates were obtained by PCR amplification of the cDNA insert from the respective IMAGE Consortium clone used on the cDNA microarrays.

Western Blot Analysis of c-MYC Protein (from (156))

Depending on protein concentration, 50 μg of protein from each sample and untreated control was brought up to 40 μl or 80 μl total volume with NP-40 lysis buffer and 10 μl or 20 μl of 5× SDS-PAGE sample treatment buffer (30% v/v glycerol, 0.3M Tris pH 6.8, 10% w/v SDS, 25% v/v 2-mercaptoethanol, 0.1% w/v bromophenol blue), respectively. These preparations were heated at 100 °C for 10 min to denature the proteins and then placed on ice to condense until ready for loading onto the gel. Solutions for 10% Laemmli SDS-polyacrylamide gels were prepared, and gels were poured according to published procedure (158). Depending on the amount of protein loaded, 0.75-mm- or 1.5-mm-thick × 15-cm-long gels were poured, with a 3-cm stacking gel and a 12-cm running gel. Denatured protein solutions were loaded onto the gel, along with 15 μl of Kaleidoscope Protein Standards (Bio-Rad), resuspended in NP-40 lysis

buffer to equal the volume of protein loaded. Empty wells were loaded with an equal volume of 5× SDS-PAGE sample treatment buffer diluted in NP-40 lysis buffer. Gels were run on an adjustable vertical slab gel apparatus (CBS Scientific Co.) at 100–120V until the dye front ran off the bottom of the gel. After electrophoresis, gels were cropped of their stacking component and immersed in cold (4 °C) Towbin transfer buffer (158). Protran nitrocellulose membrane material (Schleicher and Schuell Inc.) and 3MM Whatman filter paper (Bio-Rad) were cut to match the size of the gels and immersed in cold transfer buffer as well. The transfer stack was assembled according to published procedure (158), and proteins were transferred for 2.5 h in a model EBU-102 transfer tank (CBS Scientific Co.) at 300 mA using an EC-3000P power source (EC Apparatus Corp.). Membranes were blocked overnight with 5% nonfat dry milk in TBS-T and blotted with a 1:1000 dilution of c-MYC monoclonal antibody-2 (Neomarkers) in blocking buffer for 1 h. Membranes were then washed 3 times in blocking buffer, and bound antibodies were localized with a 1:5000 solution of HRP-labeled goat–anti-mouse immunoglobulin. Protein bands were then detected using the Phototype-HRP Detection Kit (Cell Signaling).

cDNA Microarray

All cDNA microarray studies were performed by the Microarray Core Facility at the Arizona Cancer Center. Probe preparation, microarray fabrication and target preparation are as in (159) with the following modifications: Poly A^+ RNA from TMPyP2- or TMPyP4-treated and untreated HeLa S_3 cells (see RNA extraction and

Treatment with drugs) was used as a template for labeled first-strand cDNA production. Poly A^+ RNA from treated cells was labeled with Cy5-dCTP, and that from untreated cells with Cy3-dCTP. Hybridizations were performed in duplicate for each timepoint (12, 24, 36 and 48 hours).

Plasmids

The Del-4 plasmid was a gift from Dr. Bert Vogelstein at Johns Hopkins University. Dr. Robert Tjian, a Howard Hughes Medical Institute investigator at the University of California at Berkeley, provided the pPacSp1 plasmid. The pRL-TK plasmid was purchased from Promega.

Transfections

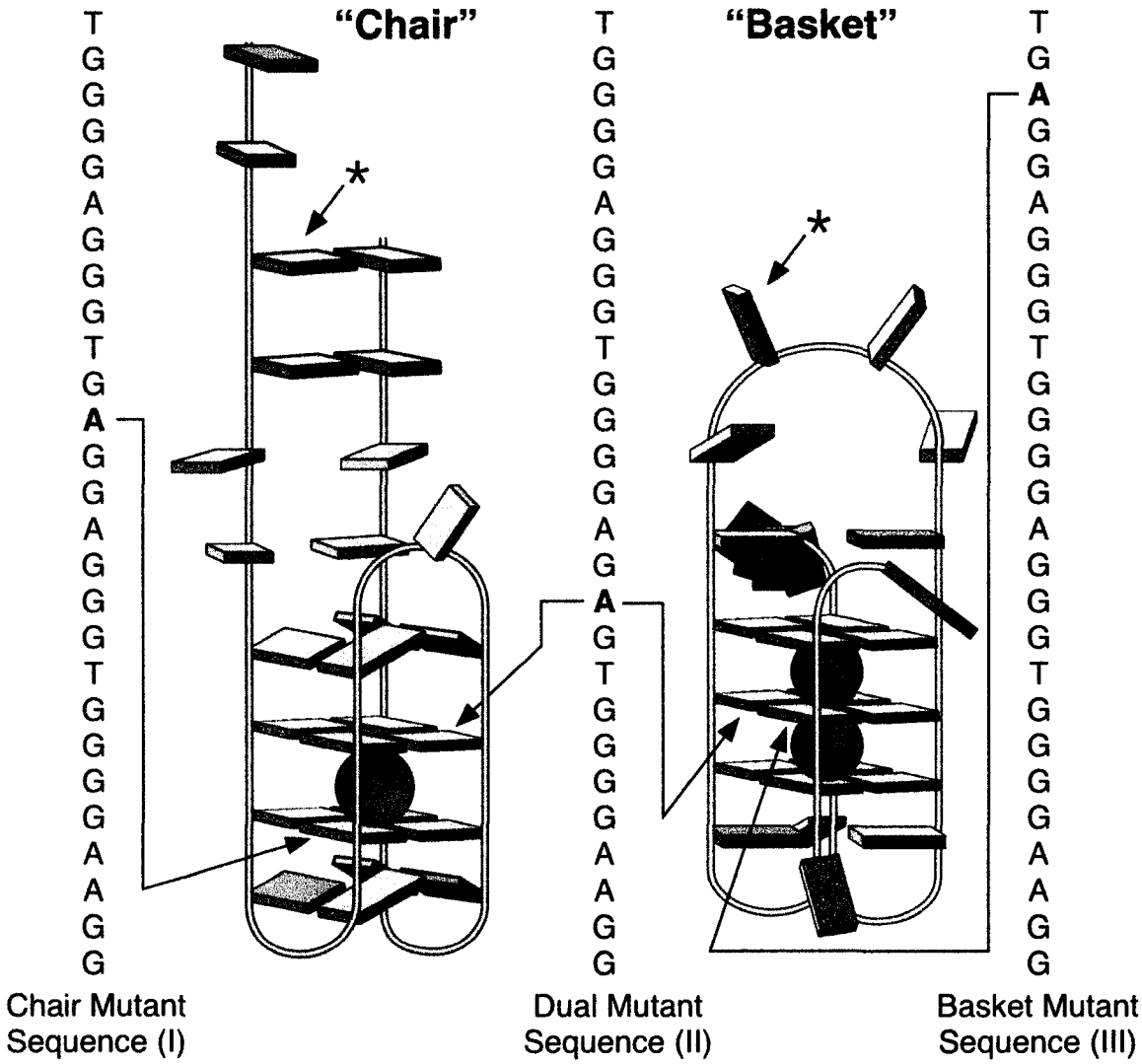
Cells were transfected using the Effectene lipid-based system (QIAgen) according to the manufacturer's protocol. Briefly, 0.1 μg of pRL-TK (Renilla luciferase reporter plasmid) and 0.9 μg of the Del-4 (wild-type) or mutated plasmids were used to transfect HeLa S_3 cells at approximately 50% confluence. For S2 Schneider cells, the same lipid-based system and protocol was used to transfect 5 x 10^6 cells in a 60mm cell culture dish. When the pPacSp1 plasmid was co-transfected with the Del-4 plasmids, 0.1 μg of the pPacSp1 plasmid was included in the mixture, and only 0.8 μg of the appropriate Del-4 plasmid was used. Transfection complexes were left on the cells for 18 hours, after which the medium was replaced. Transfectants were allowed to grow to 80% confluency and lysed with Passive Lysis Buffer (Promega).

Luciferase Assays

Firefly and Renilla luciferase activities were assayed using the Dual Luciferase Reporter Assay System (Promega) according to the manufacturer's protocol. Briefly, 20 µl of total lysate was placed in a 1.6 ml microcentrifuge tube. Firefly luciferase activity was assayed by adding 100 µl of Luciferase Assay Reagent II to the lysate and reading luminescence on a TD 20/20 luminometer (Turner Designs) or an FB12 luminometer (Berthold Detection Systems) with a 3 second delay and 10 second read time. Renilla activity was assayed by adding 100 µl of Stop-N-Glo reagent and reading as before. Relative firefly luciferase activity was determined by taking the ratio of firefly luciferase relative light units (RLU) to Renilla luciferase RLU.

Site-Directed Mutagenesis

Single guanine mutations were made to the NHE III_1 of this plasmid using the QuikChange XL Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's protocol. Briefly, the plasmid was propagated in subcloning efficiency *E. coli* DH5α cells and extracted by alkaline lysis. Plasmids were mutated using Pfu Turbo and mutagenic primers containing the desired single base change. Parental plasmid DNA was digested with Dpn I and the remaining mutant plasmids were used to transform DH5α cells for propagation. The resulting sequences of the entire mutant c-myc inserts were verified by the automated DNA sequencing facility at the University of Arizona, and are shown in Figure 2.1.



Wild-Type (Pu27): TGGGGAGGGTGGGGAGGGTGGGGAAGG
Null Mutant Sequence (IV): TGGGGAGGGTGGGGAGGGTGGGGAA**TT**

Figure 2.1 Site-directed mutations made to the Del-4 plasmid, and the specific quadruplex(es) they disrupt.

Laser Capture Microdissection (LCM) and Sequencing

LCM was used to collect cells from patient colorectal cancer specimens stored at the Arizona Cancer Center. Paraffin blocks containing both tumor as well as untransformed surrounding tissue were examined by Dr. Ray Nagle at the Arizona Cancer Center for their viability as candidates for LCM. These paraffin blocks were then cut and mounted on glass slides by members of Dr. Nagle's laboratory, and stained with hematoxylin and eosin, in order to make visualization easier. Dr. Nagle again reviewed these slides, in order to determine the best sections for subsequent LCM. Laser capture was performed by both Tiffanie Powell, an undergraduate student in the Chemistry Department at the University of Arizona and myself. A PixCell II Laser Capture system (Arcturus) was utilized, with the following settings: 30 μm spot size, 50 mW power, and 6.2 ms duration. Approximately 1500 pulses were taken, and adhered to a CapSure HS LCM cap (Arcturus).

To extract genomic DNA from the laser captured cells, the Pico Pure DNA Extraction Kit was used (Arcturus). Briefly, cells were incubated in 10 μl of Proteinase K solution for at least 16 hours at 65°C. This lysate was used in a subsequent PCR, with the NHEseqfw and NHEseqrv primers (Table 2.2): Each 50-μl reaction contained 1× High Fidelity PCR Buffer (Invitrogen), 50 μM each of dCTP, dATP, dGTP, and dTTP (Fermentas), 2 mM $MgSO_4$ (Invitrogen), 2.5 U Platinum Taq High-Fidelity Polymerase (Invitrogen), 0.5 μM of each primer, distilled/deionized water, and 2-4 ml of the lysate from above. The reactions were incubated in a DNA Engine Peltier Thermal Cycler as

follows: 95 °C, 5 min; (95 °C, 1 min; 59 °C, 1 min, 10 sec; 72 °C, 1 min, 30 sec) 45×; 72 °C, 5 min. PCR products were held at 4 °C and stored at -20°C. PCR products were resuspended in 100 ml of nuclease-free water and submitted for cleanup and sequencing with the NHEseqrv primer to the University of Arizona DNA Sequencing Core Facility.

Electrophoretic Mobility Shift Assays (EMSA)

Oligonucleotide probes were synthesized by the custom oligonucleotide synthesis service provided by Sigma-Genosys. Sequences of the Pu27 oligomers are identical to the sequences presented in Figure 2.1 above. Py27 oligomers were the exact antiparallel complement to these sequences. Pu45 oligomers had the following sequence:
5' GCTTA [appropriate Pu27 sequence] TGGGGAGGAGACT 3'
Py45 oligomers were the exact antiparallel complement to these sequences. The oligonucleotides were labeled according to the protocol from Promaga's Gel Shift Assay System.

Unincorporated label was removed by centrifugation through a Sephadex G-25 Quick Spin Column (Roche) into TE buffer according to the manufacturer's protocol. 1 µl of labeled and purified oligonucleotide was removed for determination of total activity, using a Top Count NXT microplate scintillation and luminescence counter (Packard). For double-stranded oligonucleotides, 33 µl of each labeled single strand were mixed and heated to 95 °C for 5 minutes to denature any secondary structures, then the mixture was allowed to cool to room temperature to anneal the two strands. Each DNA binding reaction included: 3 µl HeLa Nuclear Extract, Gel-Shift Assay Grade (Promega),

3 µl of Gel Shift Binding 5 x Buffer (20% glycerol, 5 mM $MgCl_2$, 2.5 mM EDTA, 2.5 mM dithiothreitol, 250 mM NaCl, 50 mM Tris-Hcl pH 7.5, 0.25 mg/ml poly [dI-dC]), 50,000 to 100,000 counts per minute (cpm) of oligonucleotide, water to a final volume of 15 µl. As well, when required, one or more of the following reagents was included: 1 µl anti-Sp1 antibody (Biocarta), 1 µl consensus Sp1 oligonucleotide (Promega), 1 µl SRE-A oligonucleotide (consensus binding site for CNBP protein, sequence 5' AAAGTGGGGAAAA 3'), 1 µM TMPyP2 or 1 µM TMPyP4. All of the reagents, save for the labeled oligonucleotide, were mixed together and incubated at room temperature for 20 minutes. After this, the appropriate oligonucleotide was added, and the binding reaction was allowed to progress at room temperature for an additional 30 minutes.

While the binding reaction was proceeding, a 1.5mm thick polyacrylamide gel was poured. The gel was pre-run at 150 V for 30 minutes, and the wells cleaned using a syringe and 20-gauge needle. The binding reactions, as well as 15 µl of 1 x DNA loading dye, were then loaded onto the gel, and electrophoresed at 150 V until the bromophenol blue in the loading dye had traveled about 80%-90% of the way down the gel. Gels were dried at 80 °C for 2 hours under vacuum and exposed to radiographic film overnight to visualize the radioactive label.

DMS Footprinting (from (157))

Dr. Adam Siddiqui-Jain, a post-doctoral researcher in this laboratory, performed all DMS footprinting. Each band of interest was excised and soaked in 100 mM KCl solution (300 µl) for 6 h at 4°C. The solutions were filtered (microcentrifuged) and

30,000 cpm (per reaction) of DNA solution was diluted further with 100 mM KCl in 0.1× TE (10 mM Tris/1 mM EDTA, pH 7.5) to a total volume of 70 μl (per reaction). After the addition of 1 μl of salmon sperm DNA (0.1 μg/μl), the reaction mixture was subjected to 1 μl of dimethyl sulfate solution (DMS:ethanol; 4:1, vol/vol) for the times shown. Each reaction was quenched with 18 μl of stop buffer (3 M - mercaptoethanol:water:NaOAc; 1:6:7, vol/vol). After ethanol precipitation (twice) and piperidine cleavage, the reactions were separated on a preparative gel (16%) and visualized on a PhosphorImager. DMS analysis of unstructured Pu27 was performed essentially in the same way, by using heat-denatured 3'-end-labeled Pu27 in 0.1× TE buffer.

Taq Polymerase Stop Assay (from (157))

Adam Siddiqui-Jain performed all Taq polymerase stop assays discussed here. The procedure is a modification of that used in (160). Briefly, a reaction mixture of template DNA (77-mer with a Pu27 or mutant Pu27 insert) (50 nM), Tris·HCl (50 mM), MgCl (10 mM), DTT (0.5 mM), EDTA (0.1 mM), BSA (60 ng), and 5′-end-labeled 18-mer template (~18 nM) was heated to 90°C for 5 min and allowed to cool to ambient temperature over 30 min. Taq polymerase in storage buffer containing 100 mM KCl (1 μl) was added to the reaction mixture (giving a final KCl concentration of 10 mM), and the reaction was maintained at a constant temperature (45°C or 60°C) for 30 min. (For the porphyrin comparison, after cooling to ambient temperature, the requisite amount of porphyrin solution was added to the reaction mixture and left at ambient temperature for

30 min prior to the addition of polymerase.) After the addition of 10 μl stop buffer [formamide (20 ml), 1 M NaOH (200 μl), 0.5 M EDTA (400 μl), 10 mg bromophenol blue], the reactions were separated on a preparative gel (12%) and visualized on a phosphorimager. Adenine sequencing (indicated by "A" at the top of the gel) was performed using double-stranded DNA Cycle Sequencing System from Life Technologies. The general sequence for 77-mer template strands was 5' TCCAACTATGTATAC *–INSERT-* TTAGCGACACGCAATTGCTATAGTGAGTCGTATTA 3', and the 18-mer primer had the sequence 5' TAATACGACTCACTATAG 3'.

Photocleavage

Photocleavage with TMPyP4 and TMPyP2 was conducted by Adam Siddiqui-Jain essentially as outlined in (161), except that salmon sperm DNA was used instead of calf thymus DNA in the stop buffer.

CHAPTER III

A CHAIR-TYPE QUADRUPLEX EXISTS IN THE C-*MYC* PROMOTER

I. Introduction

In our previous work, we have shown that the quadruplex-interactive agent TMPyP4 can reduce c-*myc* expression in cell culture, while its positional isomer, TMPyP2, is much less able to accomplish this. This was first realized using two cDNA microarrays, and was verified with a northern blot. In order to determine if this downregulation was biologically relevant, an examination of genes activated by c-*myc* was required. The catalytic protein subunit of human telomerase reverse transcriptase, hTERT, is one such gene, and is thought to play a very important role in tumorigenesis itself (162, 163). Also, as c-*myc* acts as a protein transcription factor, it was necessary to determine that the decrease in c-*myc* mRNA seen in the microarray and northern blot experiments was accompanied by a concomitant decrease in c-*myc* protein.

The c-*myc* gene is responsive to a variety of growth-inducing stimuli, acting as an "immediate early response" gene, bridging the gap between stimulation by serum growth factors and transition from a quiescent (G_0 phase) cell state to an actively dividing state (9). When c-*myc* is inappropriately activated, it can confer growth factor independence to cells, one of the hallmarks of cancer (163, 164). We were concerned that a blockade of the cell cycle might then reduce c-*myc* expression, as a means of reverse regulation; if TMPyP4 significantly changes the cell cycle profile, this may provide a non-quadruplex-mediated mechanism for this small molecule.

Since the NHE III_1 in the c-*myc* promoter is very rich in guanine residues, and preliminary studies have shown that TMPyP4, a known quadruplex-interactive agent, can reduce expression of this gene, a thorough characterization of the secondary structures this region can form was mandatory, along with their biological relevance and function. Using a variety of tools from both biology and chemistry, we set out with that task in mind. First, the NHE III_1 was examined *in vitro,* in an attempt to determine what, if any, secondary structures are possible in this region. Plasmids were constructed, bearing the NHE III_1 sequence in control of a reporter gene, and mutations to this sequence were characterized in terms of their effect on quadruplex formation and transcription activation. With these experiments, we hoped not only to characterize the NHE III_1 structure, but also to gain some insight as to the role quadruplex DNA might play in biological systems.

II. Results

a. TMPyP4 Causes a Decrease in the Expression of c-*myc* and hTERT: RT-PCR (from (156))

Having shown an inhibitory effect of TMPyP4 on c-*myc* expression in HeLa S_3 cells, we wished to determine if this effect would be carried over to other cell lines. As well, since hTERT, the catalytic subunit of telomerase, is partially under the control of the c-MYC transcription factor, we were curious as to the effect of TMPyP4 on the expression of this gene as well. MiaPaCa-2 pancreatic cancer cells were treated with 100

μM TMPyP2 or TMPyP4 for 12, 24, 36 or 48 hours, the total RNA extracted and subjected to poly-A specific reverse transcription to generate cDNA. This cDNA was then used as a template for specific PCR amplification of the c-*myc* and hTERT sequences. The results can be seen in Figure 3.1. The effect of TMPyP4 on the expression of each of these genes is quite evident; as the time of exposure to this small molecule increases, the expression of both hTERT and c-*myc* decrease to nearly nothing, as assayed by RT-PCR. TMPyP2, on the other hand, has a much less pronounced effect on these two genes. A similar pattern of gene downregulation for both c-*myc* and hTERT was seen in HeLa cells, and for c-*myc* in normal human foreskin fibroblast (ForF) cells (hTERT was not expressed to a detectable level in untreated or treated ForF cells) when these cell lines were treated with these two porphyrins (data not shown). Since hTERT is under transcriptional control in part by the c-MYC transcription factor, the effect on hTERT does not come as a surprise, and likely explains at least some of the inhibition of telomerase we saw earlier by the TRAP assay. The promoter sequence of the hTERT gene was surveyed for "quadruplex amenability" (guanine richness) and was found to contain no polyguanine tracts sufficient for quadruplex formation.

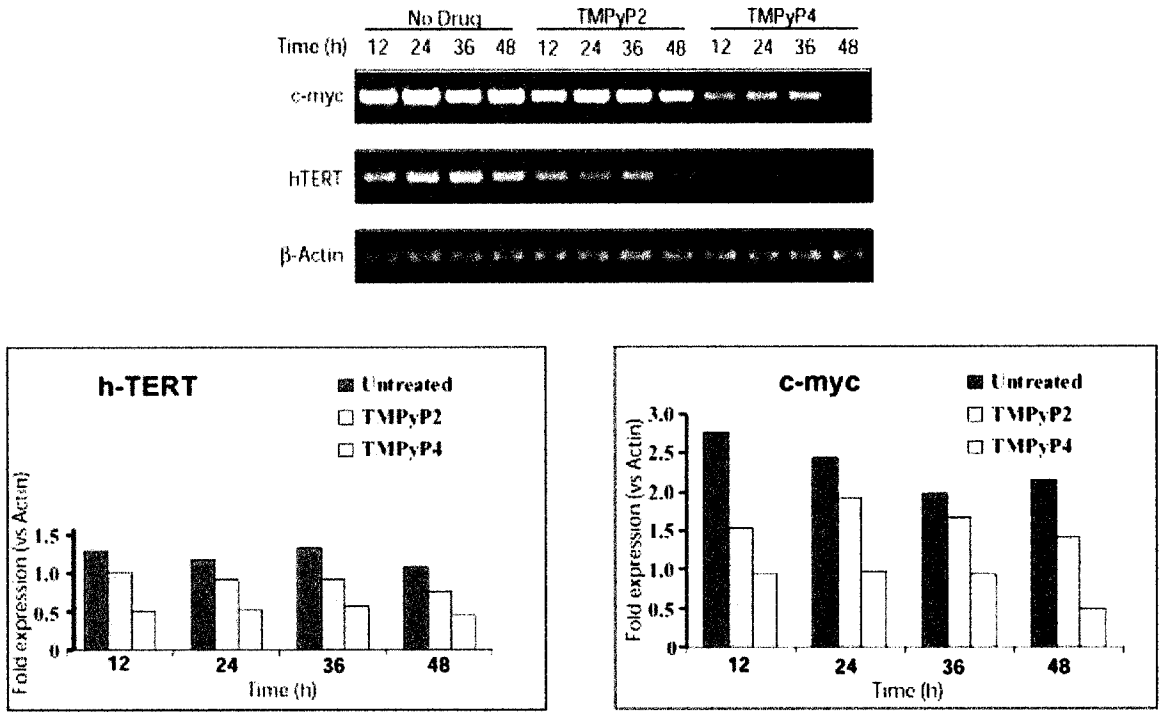


Figure 3.1: Effect of TMPyP2 and TMPyP4 on c-*myc* and hTERT mRNA in MiaPaCa-2 cells – RT-PCR.

MiaPaCa-2 pancreatic cancer cells were treated with 100 μM TMPyP2 or TMPyP4 for the times indicated. Total mRNA was collected and subjected to reverse transcription followed by PCR for c-*myc*, hTERT and β-actin. TOP: Agarose gel electrophoresis of PCR products. BOTTOM: Quantification of PCR product intensity. Expression of c-*myc* and hTERT are expressed as –fold expression versus β-actin.

b. TMPyP4 Causes a Time-Dependent Decrease in c-MYC Protein in Cell Culture (from (156))

Although we have been able to show a decrease in the amount of c-*myc* message after treatment with TMPyP4, it does not automatically follow that there is a corresponding decrease in the c-MYC protein. We wanted to ensure that TMPyP4 was having an effect on the presence of the c-MYC transcription factor, and so western blot analysis was performed. MiaPaCa-2 pancreatic cancer cells were treated with 100 μM TMPyP2 or TMPyP4 for 12, 24, 36 or 48 hours, the total protein extracted and electrophoresed by SDS-PAGE, and the electrophoresed protein used as a target for western blot analysis. A monoclonal antibody to human c-MYC was used as a probe to specifically detect this transcription factor. The results are shown in Figure 3.2. Parallel with the RT-PCR and northern results presented earlier, TMPyP4 caused a decrease in the amount of c-MYC protein present, which became more pronounced as time went on, reducing c-MYC levels to near undetectable levels by 48 hours. TMPyP2 had a similar, but much less prominent effect, consistent with what we have seen thus far. We performed a similar experiment using ForF cells (Figure 3.3), and found that TMPyP4 could produce the same effect on c-MYC protein levels in this cell line, although not as dramatically. It must be noted, however, that basal c-MYC levels in ForF cells are already quite low, so we would not expect nearly as impressive an effect in these cells.

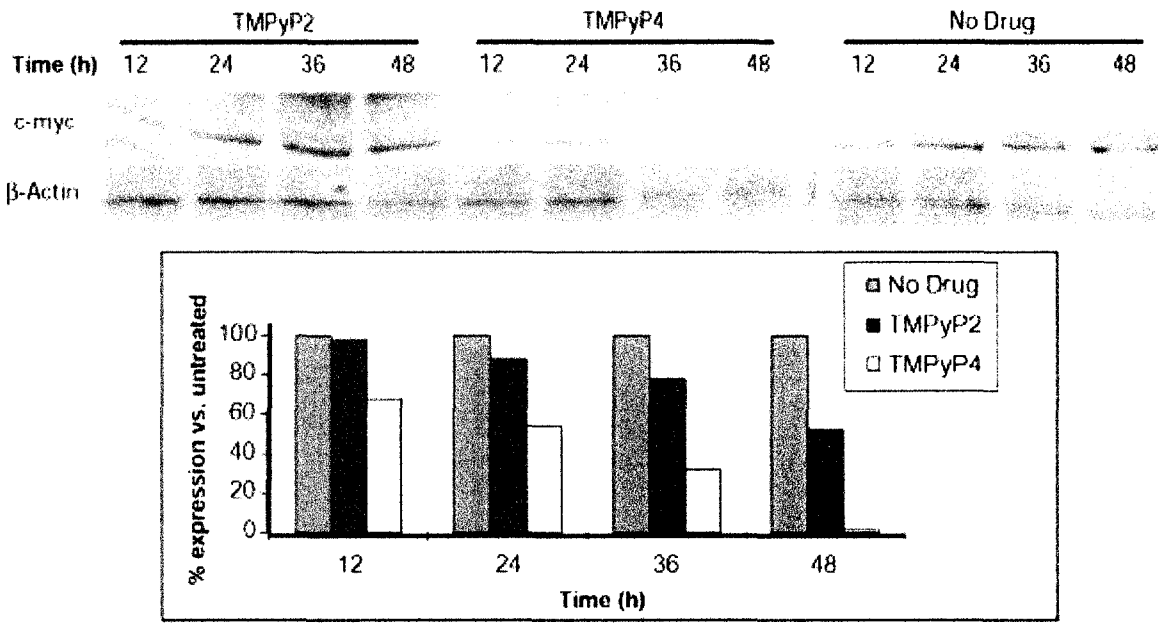


Figure 3.2: Effect of TMPyP2 and TMPyP4 on c-MYC protein in MiaPaca-2 cells – western Blot.

MiaPaCa-2 pancreatic cancer cells were treated with 100 μM TMPyP2 or TMPyP4 for the times indicated. Total protein was extracted and subjected to western blot analysis for c-MYC and β-actin. TOP: western blot for c-MYC and β-actin. BOTTOM: Quantification of western blot. Results are presented as c-MYC protein, normalized to actin, as a proportion of each untreated timepoint.

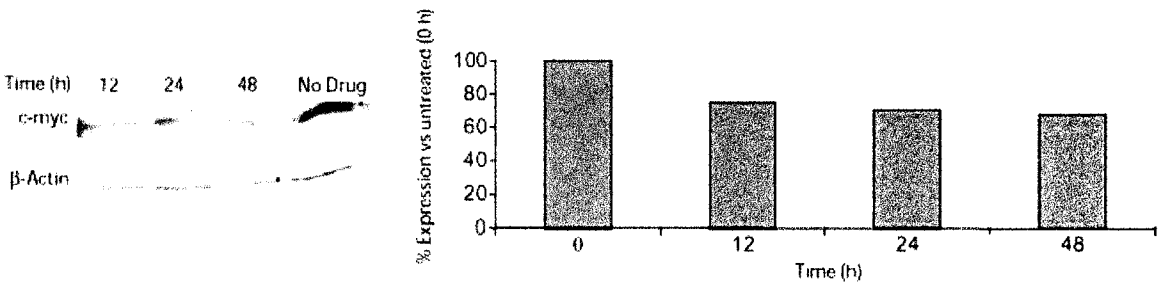


Figure 3.3: Effect of TMPyP4 on c-MYC protein in ForF cells – western Blot.

Foreskin fibroblast cells were treated with 100 μM TMPyP2 or TMPyP4 for the times indicated. Total protein was extracted and subjected to western blot analysis for c-MYC and β-actin. LEFT: western blot for c-MYC and β-actin. RIGHT: Quantification of western blot. Results are presented as c-MYC protein, normalized to actin, as a proportion of each untreated timepoint.

c. Cell Cycle Effects of TMPyP4

While the above data suggested that TMPyP4 is having some specific effects on c-*myc* and, via this effect, on telomerase as well, the possibility existed that this porphyrin could be mediating its effects through interruption of the cell cycle. For this reason, HeLa S_3 cells were treated with 100 μM TMPyP2 or TMPyP4 for 2, 6, 12 and 24 hours and subjected to fluorescence-activated cell sorting (FACS) analysis. The results of this experiment can be seen in Figure 3.4. Neither TMPyP4 nor its analogue TMPyP2 caused a significant change in the cell cycle profile, even after 24 hours. This experiment was repeated in the normal colon epithelial cell line CCD-18co, with similar results (not shown). MiaPaCa-2 cells were also treated with 100 μM TMPyP4, out to a time of 48 hours. Still no effect on the cell cycle could be seen (not shown). Therefore, we believe that the effects of TMPyP4 on c-*myc* and the other genes mentioned earlier are not due to the porphyrin acting as a general cell cycle inhibitor, but instead due to some specific activity on the genes or proteins themselves.

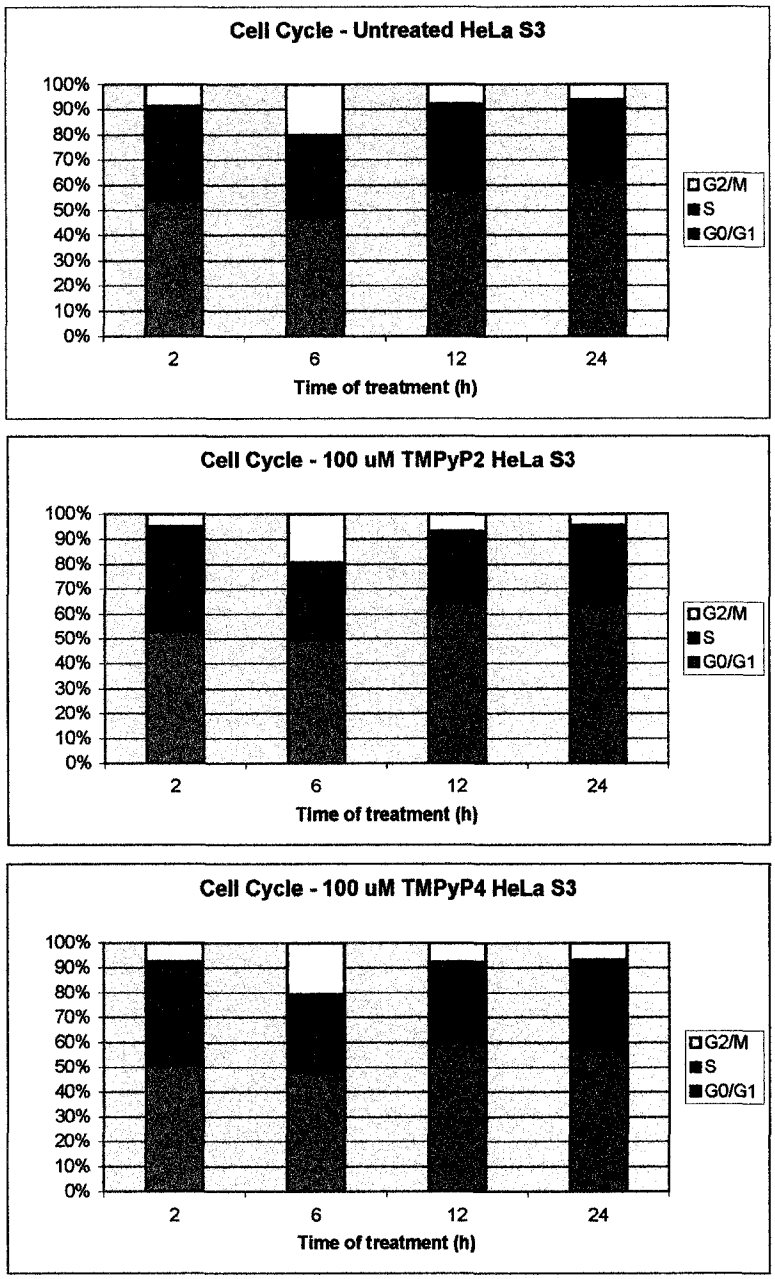


Figure 3.4: Effect of TMPyP4 on cell cycle in MiaPaca-2 cells.
MiaPaca-2 pancreatic cancer cells were treated with 100 μM TMPyP2 or TMPyP4 for the indicated times, then harvested and subjected to propidium iodide staining followed by fluorescence-activated cell sorting analysis. Graphs show % of cells in each cell cycle phase.

d. Effect of TMPyP4 on a Reporter Gene under Control of the c-*myc* Promoter

In order to separate the effects of TMPyP4 on c-*myc* transcription from the c-MYC protein itself, we have obtained two plasmids which contain ~3.75 kb and ~0.85 kb of the c-*myc* promoter upstream of a luciferase reporter gene. These plasmids were obtained from Bert Vogelstein at the Johns Hopkins Oncology Center, and have been named Del-1 and Del-4 respectively (See Figure 3.5). HeLa S_3 cells were transfected with these two plasmids, as well as a control vector, pGL3- Control (Promega), which bears the luciferase gene under the control of an SV40 promoter. The cells were treated 24 hours after transfection with 100 μM TMPyP2, TMPyP4 or an equal volume of H_2O. The cells were harvested and lysed after 24 hours of treatment, and equal amounts of total protein were assayed for luciferase activity. The results can be seen in Figure 3.6. We show that TMPyP4, and to a lesser extent TMPyP2, reduces luciferase expression from both of the Del-1 and Del-4 vectors. The effect on the control plasmid, pGL3-Control, was the same for both TMPyP2 and TMPyP4, suggesting that there is a non-quadruplex-related effect of both porphyrins on luciferase activity. Moreover, we show that the reporter vectors Del-1 and Del-4 behave similarly to the endogenous c-*myc* gene in response to TMPyP2 and TMPyP4, making them useful models for future studies.

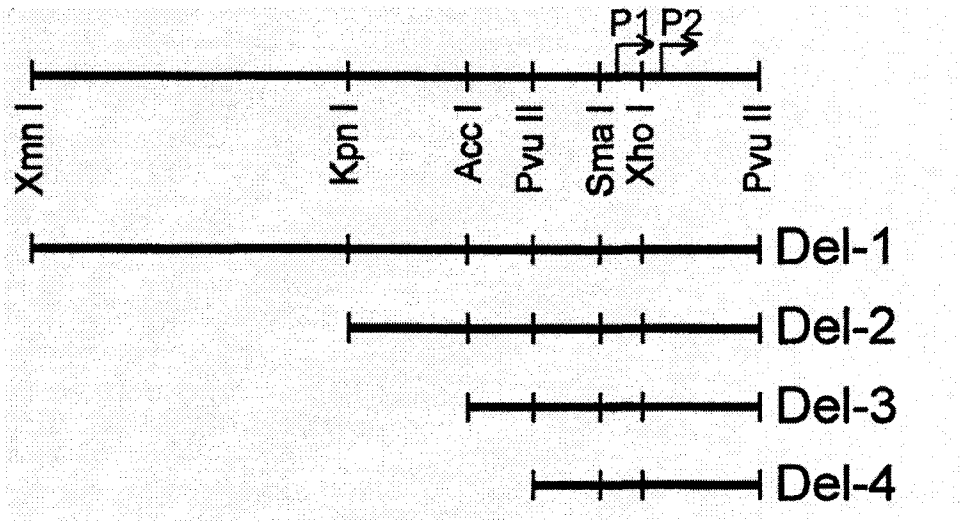


Figure 3.5: Maps of the c-*myc* reporter constructs (From (32)).

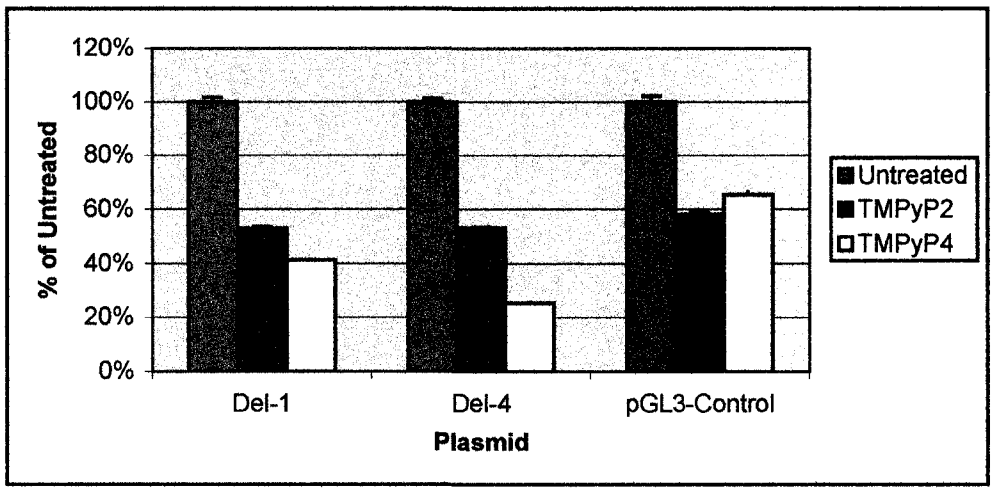


Figure 3.6: Effect of TMPyP2 and TMPyP4 on luciferase activity under c-*myc* promoter control.

HeLa S_3 cells were transfected with the Del-1 and Del-4 plasmids, along with a Renilla luciferase vector as a control for transfection efficiency. Transfectants were then treated with 100 μM TMPyP2 or TMPyP4 for 24 hours. Results are presented as firefly luciferase expression, normalized to Renilla luciferase expression, as a percent of the expression in untreated transfectants.

e. Characterization of the G-Quadruplex in the NHE III_1

In 1998, the laboratory of Dr. Tomas Simonsson demonstrated that the 27 base pair sequence of the NHE III_1 had the ability, *in vitro,* to fold into an intramolecular G-quadruplex structure (60). Briefly, the purine-rich strand of the NHE III_1 (hereafter known as Pu27) was shown to be specifically accelerated during electrophoresis, only when the oligonucleotide was incubated for 48 hours at 37 °C in the presence of 100 mM KCl. Based on the sensitivity of this high-mobility DNA species to dimethylsulphate (DMS), Dr. Simonsson's group concluded that it represented a DNA quadruplex, specifically a basket-type quadruplex.

In our laboratory, we sought to reproduce the results from the Simonsson laboratory. The same Pu27 sequence was electrophoresed under three different conditions (Figure 3.7). The same high-mobility species found in Simonsson's laboratory was again recreated under the same conditions, and named "Band 1" (Figure 3.7, Lane 3). When the oligonucleotide was incubated without potassium ions, or in potassium for a much shorter period of time, this high-mobility Band 1 was not found. A second intense band, presumed to be unstructured single-stranded Pu27, was seen under all three conditions, and termed "Band 2".

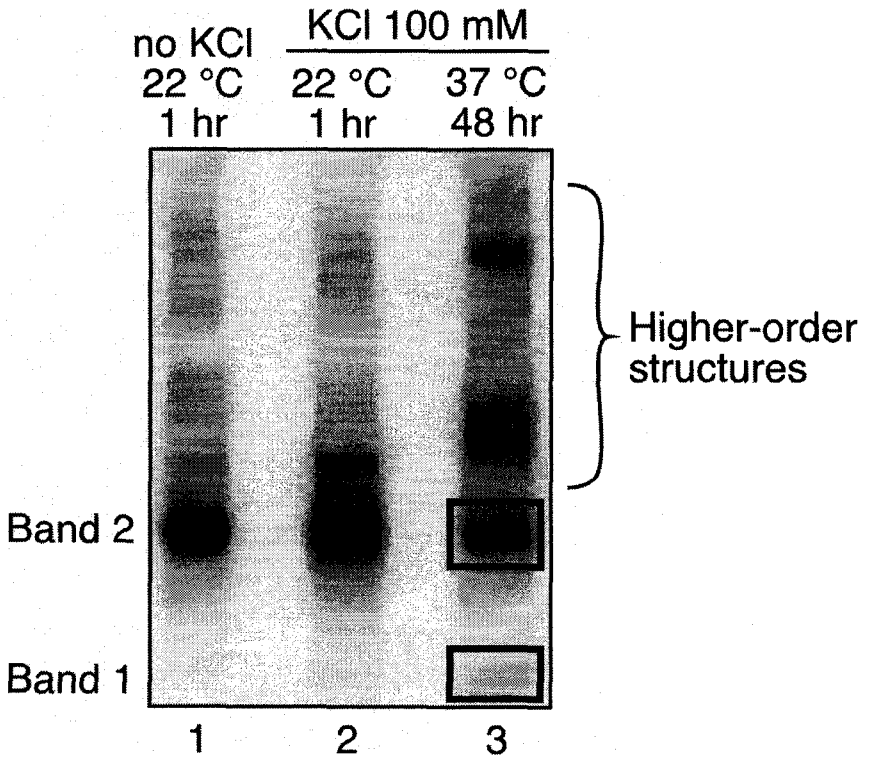


Figure 3.7 (from (157)): Electrophoresis of Pu27.
The Pu27 oligonucleotide was incubated under the conditions indicated above each lane and electrophoresed on a polyacrylamide gel. Two bands of interest (Bands 1 and 2) were chosen for further study.

Both of these bands were excised and subjected to DMS footprinting to verify their structures. DMS specifically methylates the N_7 position of guanine residues in a DNA molecule. When this treated DNA is exposed to piperidine and heat, the DNA sequence is cleaved specifically at these guanine residues, giving a distinctive "footprint" of the location of guanine residues in the DNA molecule. The formation of a G-tetrad involves occupation of the N^7 of guanine, which protects it from methylation by DMS. Hence, we can determine which guanines in a DNA molecule are involved in the G-quadruplex by observing its DMS footprint. Footprinting of Band 1 revealed what had already been shown by Dr. Simonsson; the pattern of protected guanine residues indicated that this high-mobility species represented an intramolecular basket-type quadruplex (Figure 3.8). Because of the relative overabundance of Band 2 under each of the three conditions tested, it too was subjected to DMS footprinting to determine if it was indeed merely single-stranded Pu27, or another species. DMS footprinting of this band is shown in Figure 3.9. Interestingly, this band was found not to be single-stranded at all, but instead a previously undescribed chair-type quadruplex structure. Based on the ability of this structure to form under a variety of conditions, and over a very short time, it was believed that formation of the chair-type quadruplex was kinetically very favorable. Dr. Adam Siddiqui later verified this, showing that this structure could form in a matter of seconds under physiological ionic conditions (data not shown).

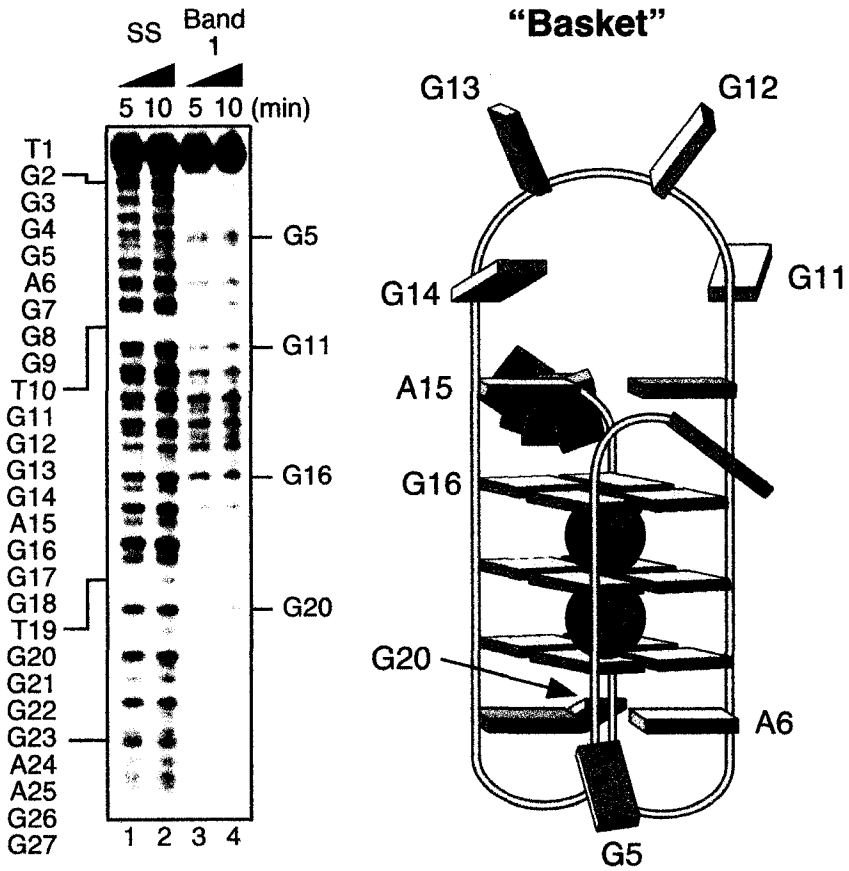


Figure 3.8 (from (157)): DMS footprinting of Band 1 from Figure 3.7.
Based on the pattern of protection of certain guanine residues in the DNA molecule from DMS-mediated nucleolytic cleavage, Band 1 was determined to be an intramolecular basket-type quadruplex, shown at the right.

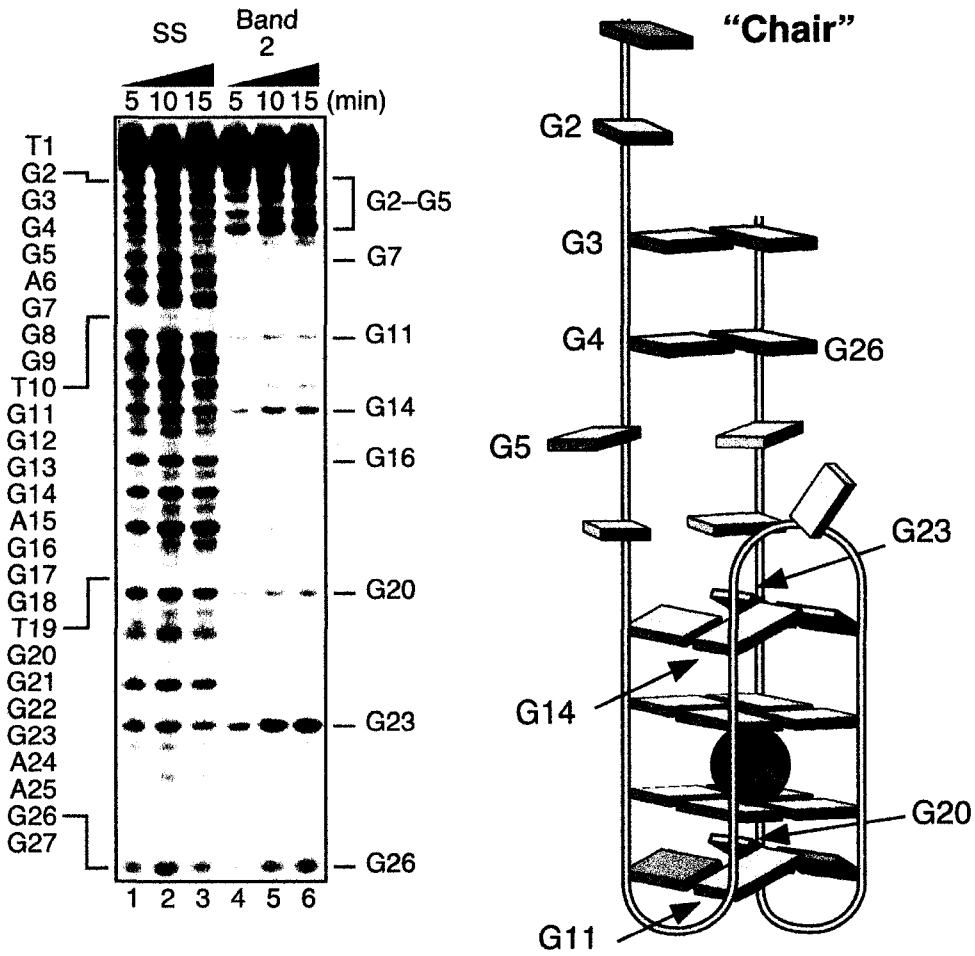


Figure 3.9 (from (157)): DMS Footprinting of Band 2 from Figure 3.7.
Based on the pattern of protection of certain guanine residues in the DNA molecule from DMS-mediated nucleolytic cleavage, Band 2 was determined to be an intramolecular chair-type quadruplex, shown at the right.

f. Importance of the Chair-Type Quadruplex in c-*myc* Regulation

Up to this point, the characterization of the G-quadruplex in the NHE III_1 had only been done in *in vitro* systems. The role (and existence) of the chair-type quadruplex still needed to be studied in a living system. To accomplish this, we turned again to the c-*myc* reporter vectors used earlier (see Figure 3.5). Because the Del-4 vector contains the shortest portion of c-*myc* promoter in control of luciferase, while still retaining the NHE III_1, it was used for the following experiments. In order to characterize what effect, if any, a quadruplex plays in the normal regulation of *c-myc*, the Del-4 plasmid was mutated in such a way that the quadruplexes would not be stable. Four different site-directed mutations were made in the Del-4 plasmid (see Figure 2.1): The Chair Mutant contained a G to A transition mutation in the third run of guanines in the NHE III_1, that would destabilize only the chair-type quadruplex, while leaving the basket-type quadruplex, in which this run of guanines is not involved in quadruplex formation, unaffected. Similarly, the Basket Mutant contained a G to A transition mutation in the first run of guanines that would abrogate formation of only the basket-type quadruplex. Since the chair-type quadruplex does not involve this run of guanines, it should not be affected by the basket mutation. A third mutation, termed the Dual Mutant, involved a G to A transition mutation in the fourth run of guanine residues. This guanine plays a role in the formation of both the chair and basket quadruplexes, and thus this mutation should destabilize both. The final mutation, resulting in the Null Mutant, entailed G to A transition mutations in the last two guanines of the NHE III_1, which are involved in

neither quadruplex, and so should have no effect at all in the formation of the basket or chair quadruplexes.

These mutant plasmids were transfected into HeLa S_3 cells, and the luciferase activity from each reporter was observed. The results of this experiment can be found in Figure 3.10. Mutating the Del-4 vector in a such a way that neither quadruplex was affected seemed to have little to no consequence in the basal regulation of c-*myc,* since the luciferase expression from the Null Mutant did not differ appreciably from that of Del-4. Interestingly, the Basket Mutant also showed no difference in activity from Del-4, suggesting that, at least in this context, the basket-type quadruplex does not play a role in the regulation of c-*myc*. When mutations were made that disrupted the chair-type quadruplex, however, a quite different pattern was seen; in both the Dual Mutant and the Chair Mutant, which both have lost the ability in the NHE III_1 to form a chair-type G-quadruplex, a 3-fold increase in basal c-*myc* promoter activity is seen. The fact that this upregulation occurs as a result of two different single-base mutations, which are separated in the DNA sequence of the NHE III_1, suggested to us that the chair-type quadruplex might play an inhibitory role in the normal regulation of c-*myc*. When this structure can no longer form, a constitutive increase in expression from the c-*myc* promoter results.

These same experiments were repeated in the context of four other cell lines, to ensure that this effect is not cell-type specific, but is a general phenomenon. Figure 3.11 shows the results of these experiments. MCF-10 is an immortalized but nontransformed mammary ductal epithelium cell line. 293 cells are also nontransformed and are derived

from human kidney. The HT29 and Panc-1 cell lines are both transformed cell lines, obtained from human colorectal and pancreatic tumors, respectively. In all four cell lines, we find the same pattern as with the HeLa S_3 cells: The Dual and Chair Mutant plasmids, both of which abrogate the formation of the chair-type G-quadruplex in the NHE III_1 of c-*myc,* have greater luciferase activity than the wild-type Del-4 plasmid, while the Null and Basket Mutant plasmids have much less effect. The induction of expression caused by the chair quadruplex mutations varies, from approximately 175% to 350% of wild-type activity. These data suggest that a chair-type quadruplex in the NHE III_1 of c-*myc* plays an important role in the regulation of this proto-oncogene, regardless of cell type or transformation status.

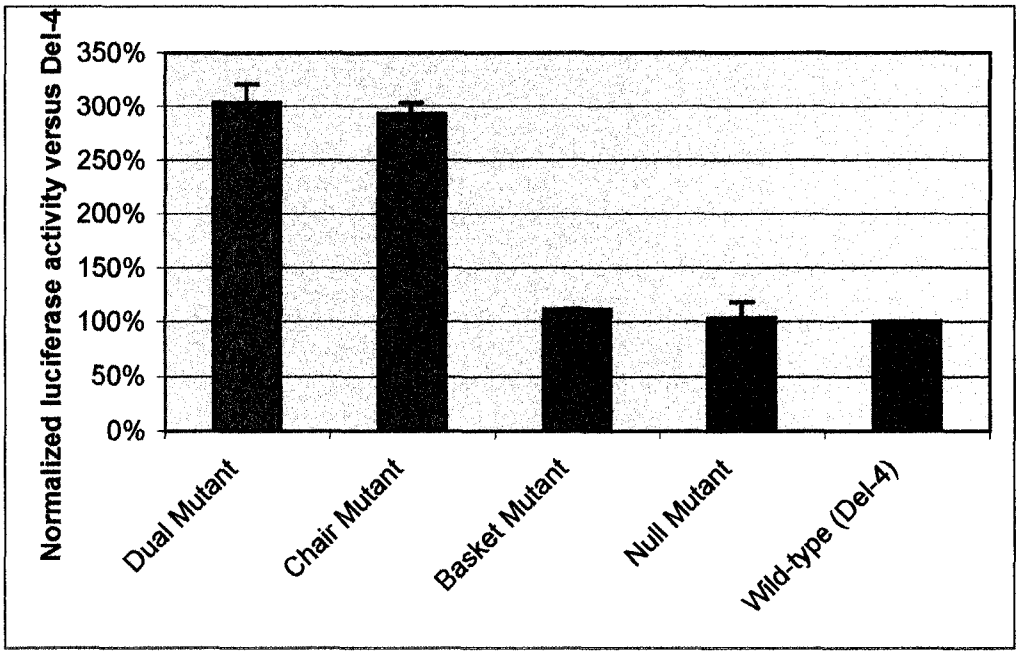


Figure 3.10. Effect of single-guanine quadruplex mutations on c-*myc* promoter activity – HeLa S_3.

The various mutant versions of the Del-4 plasmid were co-transfected into HeLa S_3 cells with a Renilla luciferase vector (as a control for transfection efficiency). Transfectants were allowed to grow for 24 hours post-transfection, and luciferase activity measured. Activity is reported relative to the wild-type (Del-4) vector.

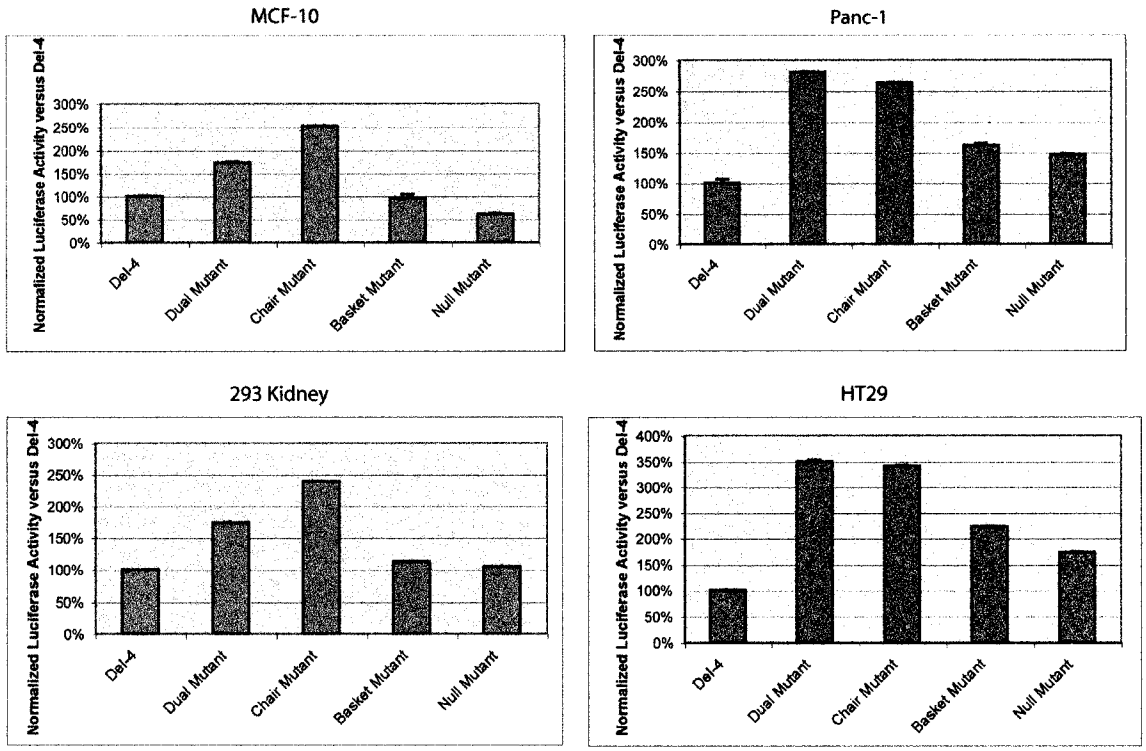


Figure 3.11. Effect of single-guanine quadruplex mutations on c-*myc* promoter activity in a variety of cell lines.

The various mutant versions of the Del-4 plasmid were co-transfected into four different cell lines with a Renilla luciferase vector as a control for transfection efficiency. After 24 hours, luciferase activity was determined from each transfectant. Data are presented as luciferase activity relative to the Del-4 (wild-type) vector.

Having established that the Chair and Dual mutations cause an increase in the expression of luciferase from a c-*myc* promoter, it was necessary to ensure that these mutations did indeed inhibit the formation of the chair-type quadruplex. In order to determine this, the Taq polymerase stop assay was used. The Pu27 sequences corresponding to each of the four mutations made were used as inserts in the template strand of this assay, to see if the stability of the secondary structure it forms (which we have now determined is a G-quadruplex) would be affected. These data can be found in Figure 3.12. Concordant with the transfection data, we find that the Chair and Dual mutations disrupt the stability of the secondary structure that was found in the earlier Taq polymerase stop assays. The Basket and Null mutations, however, have no effect on the stability of this structure. This indicates that this secondary structure is indeed a chair-type G-quadruplex, and the mutations that increase c-*myc* promoter activity in the Del-4 plasmid also abrogate the formation of this structure. Therefore, we have established that disruption of a chair-type quadruplex that forms in the NHE III_1 of the c-*myc* promoter leads to an increase in the activity of this promoter, and may provide a means by which a cell might suffer upregulation of this proto-oncogene by a single-base mutation. Such upregulation may be an important event in the process of cellular transformation.

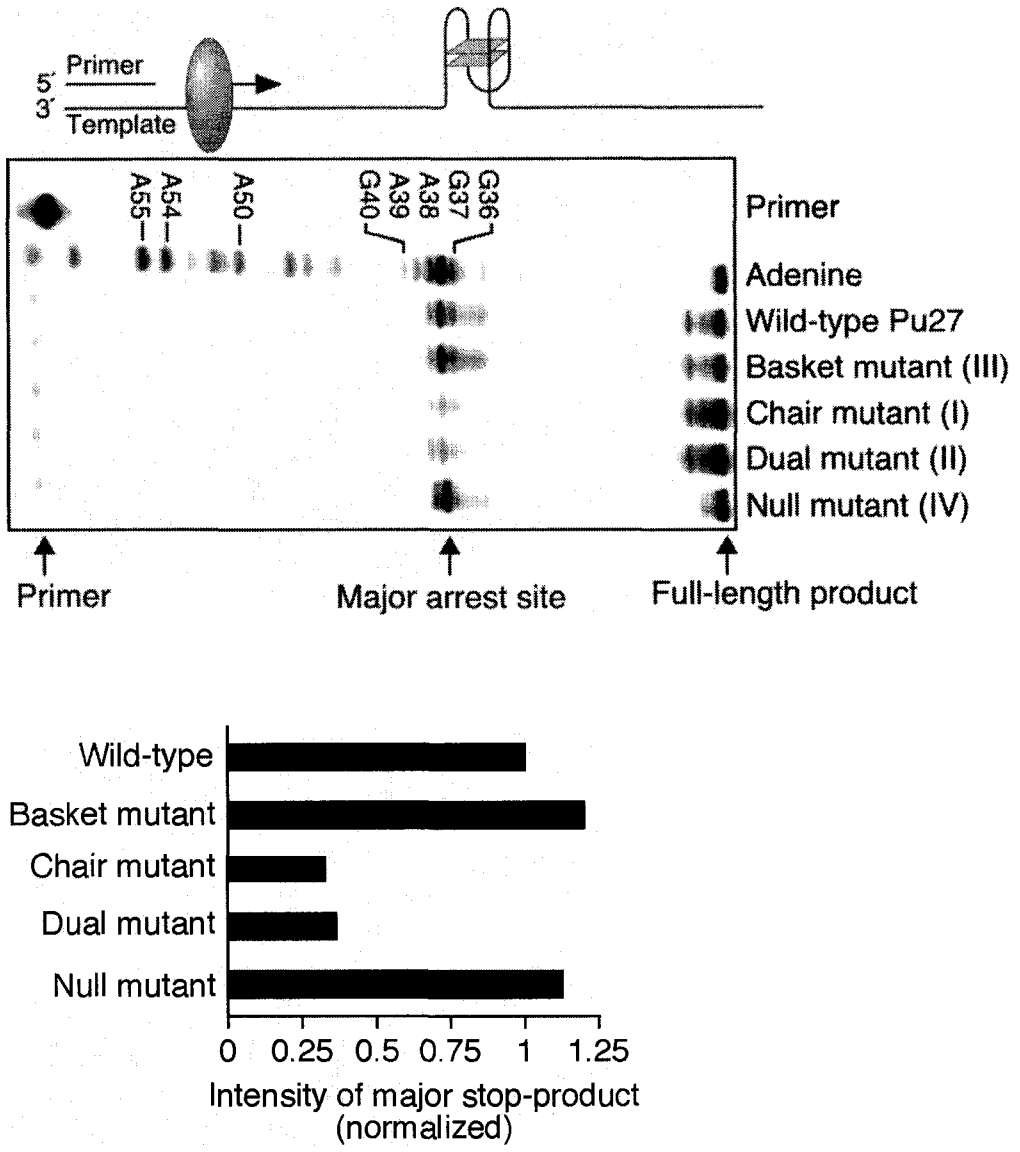


Figure 3.12 (from (157)). Taq polymerase stop assay to determine the stability of the G-quadruplex in the NHE III_1 in response to single-guanine mutations.

The concept of the polymerase stop assay is shown graphically above the gel. The G-quadruplex in this illustration is shown in approximately the position it would be found on the gel. The graph below the gel indicates the relative intensity of the quadruplex-induced stop product as a result of each mutation.

III. Discussion

We have shown here evidence that a G-quadruplex structure can form in the NHE III_1 of the c-*myc* promoter, and that this structure may be biologically relevant. A G-quadruplex-interactive agent, TMPyP4, but not its positional isomer TMPyP2, causes a decrease in both c-*myc* mRNA and protein, and also reduces the expression of a downstream target of c-*myc,* the telomerase catalytic subunit hTERT. We have determined that this effect is very likely due to the formation of a G-quadruplex in the c-*myc* promoter, and not due to a non-specific cell cycle-related toxicity. This same effect on expression is seen with a reporter vector under the control of the c-*myc* promoter, narrowing the region responsible for this effect to an approximately 850 base pair long region of the c-*myc* promoter, including the guanine-rich NHE III_1. We have characterized secondary structure formation in the NHE III_1, and have found that two distinct G-quadruplex structures are possible; a chair-type and a basket-type. The chair-type, while thermodynamically less stable, forms readily in solution, while the basket-type is much slower to form, but is more stable over time.

In an effort to determine which, if either, of these two quadruplexes is biologically relevant, we have constructed single-base mutations of the Del-4 reporter vector that have lost much of their ability to form one or both of these quadruplexes. It was found that, by abrogating formation of the chair-type quadruplex, but not the basket, c-*myc* expression could be increased to as much as 350% of wild-type levels. This effect was found in cell lines from different sources, both transformed and non-transformed,

suggesting that whatever effect the quadruplex has on c-*myc* expression is common to all cell types, and is not tumor specific. These results suggest that the chair-type G-quadruplex is a repressor structure in the c-*myc* promoter, concordant with the fact that a G-quadruplex-interactive agent is able to reduce c-*myc* levels in cell culture. Further studies, discussed in the next chapter, will confirm the role of the chair-type quadruplex in the effect of TMPyP4 on c-*myc*.

CHAPTER IV

THE CHAIR-TYPE G-QUADRUPLEX IS A REPRESSOR ELEMENT FOR c-*MYC* TRANSCRIPTION, AND IS REQUIRED FOR THE ACTIVITY OF TMPyP4

I. Introduction

It has now been established that a G-quadruplex-amenable sequence exists in the c-*myc* promoter, and this G-quadruplex is of the chair type. Further, we have shown that abrogation of the formation of this quadruplex structure in the context of the c-*myc* promoter sequence leads to a 3-fold increase in expression. This is consistent with a repressive role of this structure on c-*myc* gene activation. Thus, the effect of TMPyP4 on c-*myc* expression can be explained simply through stabilization of the repressor. However, more data are needed to prove this hypothesis. In other words, proof that the chair-type G-quadruplex structure is required for the activity of TMPyP4 must be found. To this end, we have obtained a pair of Burkitt's Lymphoma cell lines, one of which lacks the NHE III_1 region of the c-*myc* promoter, and these cell lines will be treated with TMPyP4. If our hypothesis is correct, TMPyP4 should only be able to repress c-*myc* expression from NHE III_1-driven c-*myc*. This will narrow our scope of study to the NHE III_1, and provide further evidence that the chair-type G-quadruplex is responsible for the inhibitory effect of TMPyP4.

In order to show conclusively that an interaction of TMPyP4 with the chair-type G-quadruplex is required for the effect of this porphyrin, such an interaction must be

found. In other words, it must be shown that TMPyP4 interacts very specifically with the chair-type quadruplex, to provide proof of principle. Accomplishing this, we can then make further use of the specific mutations discussed in the previous chapter to bolster our case; if the chair-type quadruplex plays an integral role in downregulation of c-*myc* by TMPyP4, then abrogation of this structure should reduce, if not totally abolish, this effect. Similarly, these mutations should result in a loss of stabilization of the chair-type quadruplex by TMPyP4. If all of these predictions hold true, then we can finally say with confidence that TMPyP4 interacts with the chair-type quadruplex in the NHE III_1 of c-*myc* and effects downregulation of c-*myc* **as a result** of this interaction.

II. Results

a. The Effect of TMPyP4 on c-*myc* Expression Requires the NHE III_1

The previous experiments with the Del-1 and Del-4 expression vectors had narrowed the region of DNA required for the effect of TMPyP4 on c-*myc* expression to ~850 bp of the c-*myc* promoter, which included the NHE III_1. To provide more evidence for the role of this element in downregulation of c-*myc* by TMPyP4, two Burkitt's Lymphoma cell lines were chosen for further study. Burkitt's Lymphoma is a non-solid tumor characterized by a set of specific genetic rearrangements. In this disease, a translocation occurs between chromosome 8, which bears the c-*myc* gene, and one of chromosome 2, 14 or 22, which carry the genes encoding immunoglobulins (the kappa light chain, the heavy chain and the lambda light chain, respectively) (7, 9, 19). It is

thought that these translocations may result from aberrant immunoglobulin gene rearrangements (165), the t(8;14) recombination event occurring most often (7, 9). The normal allele of c-*myc* is not expressed in this disease; only the translocated c-*myc* gene is active (7, 9, 46). The breakpoint regions are variable in this disease, but result in abnormally high expression of c-*myc* in antibody-producing B cells, hence leading to lymphoma. We exploited these different translocations in an effort to better ascertain the region of c-*myc* promoter that is required for TMPyP4 to exert its repressive effect.

Two different Burkitt's Lymphoma cell lines were used: Ramos and CA46. The translocation in the Ramos cell line spares the NHE III_1 of c-*myc*, such that it remains associated with the c-*myc* gene, now on chromosome 14. In CA46 cells, the translocation does not include the NHE III_1, and hence the translocated c-*myc* allele is no longer under the influence of this regulatory element (see Figure 4.1A). Both of these cell lines were exposed to 100 μM TMPyP2 or TMPyP4 for 24 hours, and RT-PCR was performed on their total RNA for c-*myc*. In the Ramos cell line, which bears the NHE III_1 upstream of the translocated c-*myc* allele, TMPyP2 and TMPyP4 can be seen to have approximately the same effect as was shown earlier; the poorly quadruplex-interactive TMPyP2 has very little effect, while TMPyP4 decreases c-*myc* substantially (Figure 4.1B). Contrast this to the effect of TMPyP2 and TMPyP4 on c-*myc* in CA46, in which the NHE III_1 is no longer present: Neither TMPyP2 nor TMPyP4 appear to have an effect on c-*myc* in this cell line, suggesting again that the NHE III_1, and possibly the G-quadruplex structure therein, is required for TMPyP4 to downregulate the expression of this gene.

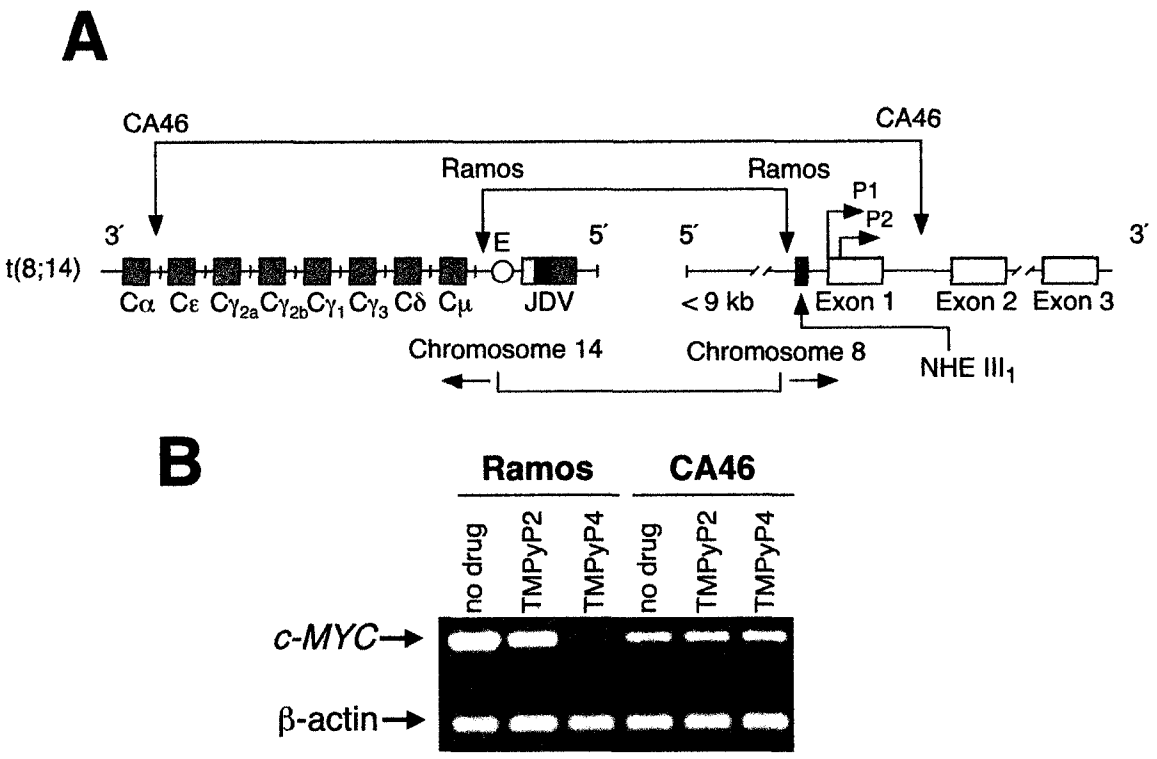


Figure 4.1 (taken from (157)). Effect of TMPyP2 and TMPyP4 on c-*myc* expression in two Burkitt's Lymphoma cell lines.
A. Genetic map of translocations in Ramos and CA46 cells (modified from (7)). B. RT-PCR for c-*myc* in these cell lines after 24-hour treatment with 100 μM TMPyP2 or TMPyP4.

b. A Secondary Structure in the c-*myc* Promoter Is Stabilized by TMPyP4

The fact that TMPyP4 is able to inhibit c-*myc* expression, while TMPyP2 is much less able to do so correlates quite well with the relative abilities of these molecules to interact with G-quadruplex DNA structures (105). At this point, evidence was needed that a G-quadruplex can indeed form in the NHE III_1 of c-*myc*, as we suspected. The Taq polymerase stop assay is a powerful tool for identifying secondary structures in DNA, though it cannot give any information on the actual identity (i.e. type) of secondary structure. Briefly, a template strand containing the region of interest is incubated with a radioactively labeled primer which is complementary to its the 3' end. Taq polymerase is added to this partial duplex, and begins adding nucleotides to the primer strand in the 5' to 3' direction. If a secondary structure exists in the template strand, the polymerase will be unable to read through this point, and will stall and eventually disengage from the DNA. When the products of this reaction are electrophoresed, this stop site will be visualized as a labeled "band" with a mobility less than the primer oligonucleotide, but greater than the expected full-length product, had the polymerization reaction gone to completion. This stop product indicates that a secondary structure forms in the sequence of interest.

Dr. Adam Siddiqui, using the NHE III_1 as an insert on the template strand, performed the Taq polymerase stop assay. TMPyP2 and TMPyP4 were added to the reaction mixture, in order to determine whether they could cause stabilization of any secondary structure that might form. The results can be seen in Figure 4.2. In each lane, even without the addition of TMPyP2 or TMPyP4, a mid-length stop product is apparent.

The size of this stop product is consistent with a polymerization block at the NHE III_1 insert. When TMPyP2 is added prior to polymerization, we see an increase in stabilization of this secondary structure with increased dose. TMPyP4 has a similar effect, but much stronger, such that by 5 μM, the full-length polymerization product is no longer seen; all of the template strand is in the secondary structured form, and polymerization cannot continue past this point. This is consistent with the effects of TMPyP2 and TMPyP4 on c-*myc* expression; TMPyP4 has a much stronger effect on c-*myc* expression than TMPyP2. Also, the affinity of TMPyP4 for quadruplex DNA is much greater than that of TMPyP2. For these reasons, we are led to believe that the secondary structure formed in the Pu27 sequence of the c-*myc* NHE III_1 is indeed a quadruplex, and that it is this quadruplex that mediates the effect of TMPyP4 on c-*myc*.

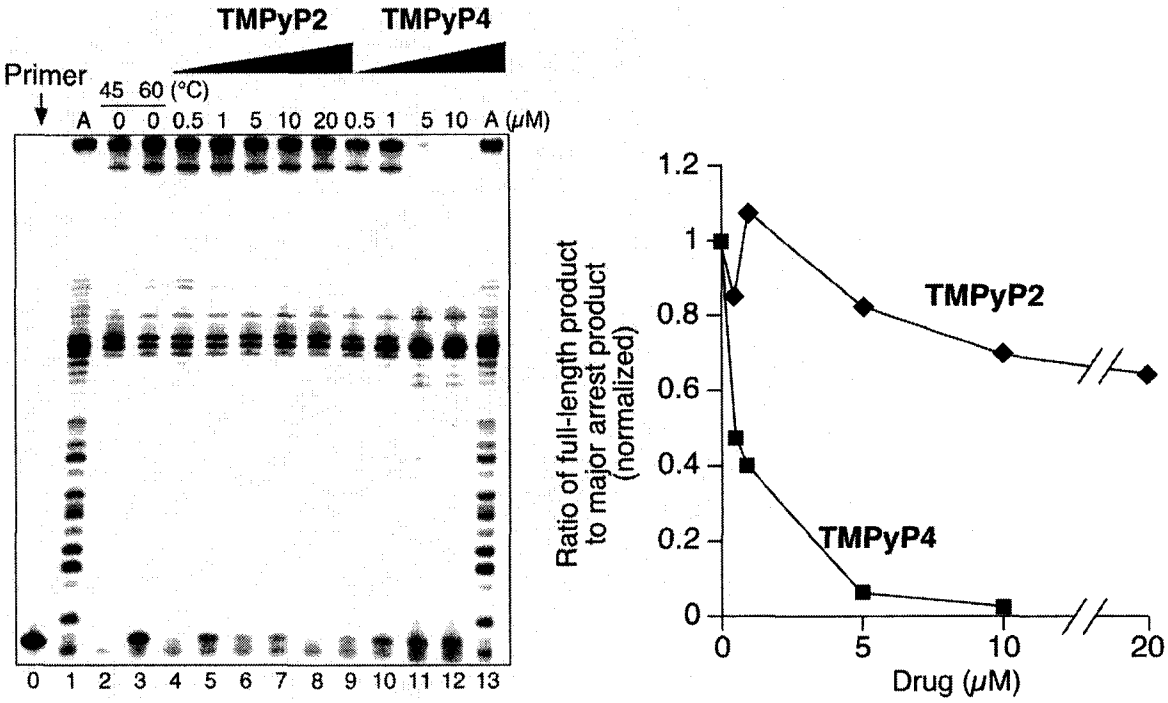


Figure 4.2 (from (157)). Taq polymerase stop assay to determine stabilization of the chair-type quadruplex by TMPyP2 and TMPyP4.
The Pu27 DNA sequence was inserted into the template strand, in order to determine if this sequence has the potential to form a secondary structure and thus block polymerization. Several doses of TMPyP2 and TMPyP4 were used in an attempt to stabilize the structure. The electrophoresis products are shown to the left, and a representation of how well each porphyrin was able to stabilize the secondary structure is shown to the right. The letter 'A' in the phosphorimage to the left indicates a Maxam-Gilbert sequencing reaction for adenine residues, which we used to determine size of the stop product.

c. The Chair-Type Quadruplex and TMPyP2/TMPyP4

Having found two different possible quadruplex structures in the NHE III_1, we sought next to determine which of these quadruplexes, if any, was responsible for the effect of TMPyP4 on c-*myc* expression. First, the interactions between each quadruplex and the two porphyrins, TMPyP2 and TMPyP4, were studied. Both of these porphyrins have the ability to induce cleavage of DNA molecules to which they are in close proximity when energy in the form of the proper wavelength of light is introduced. This activity is known as photocleavage. When the treated DNA molecule is electrophoresed, these site-specific cleavage events manifest as a specific "footprint" of the porphyrin on the DNA. Using this property of TMPyP2 and TMPyP4, we were able to determine where on the quadruplex each porphyrin was binding, by observing the pattern of photocleavage produced when the porphyrin/DNA complex was exposed to light. Dr. Adam Siddiqui performed these photocleavage experiments. First, the basket quadruplex was examined (Figure 4.3). The pattern of photocleavage with either porphyrin was the same, and indicated that they both bind in the loop region of the quadruplex, as evidenced by the fact that guanines 11 through 16 are strong sites of cleavage. Therefore, both porphyrins have a very specific mode of binding to the basket quadruplex, but there is no difference in binding between the two porphyrins. Recall that these two porphyrins have very different activities in terms of their relative abilities to inhibit c-*myc* expression. This suggests that the effect of TMPyP4 on c-*myc* is not mediated through the basket-type quadruplex.

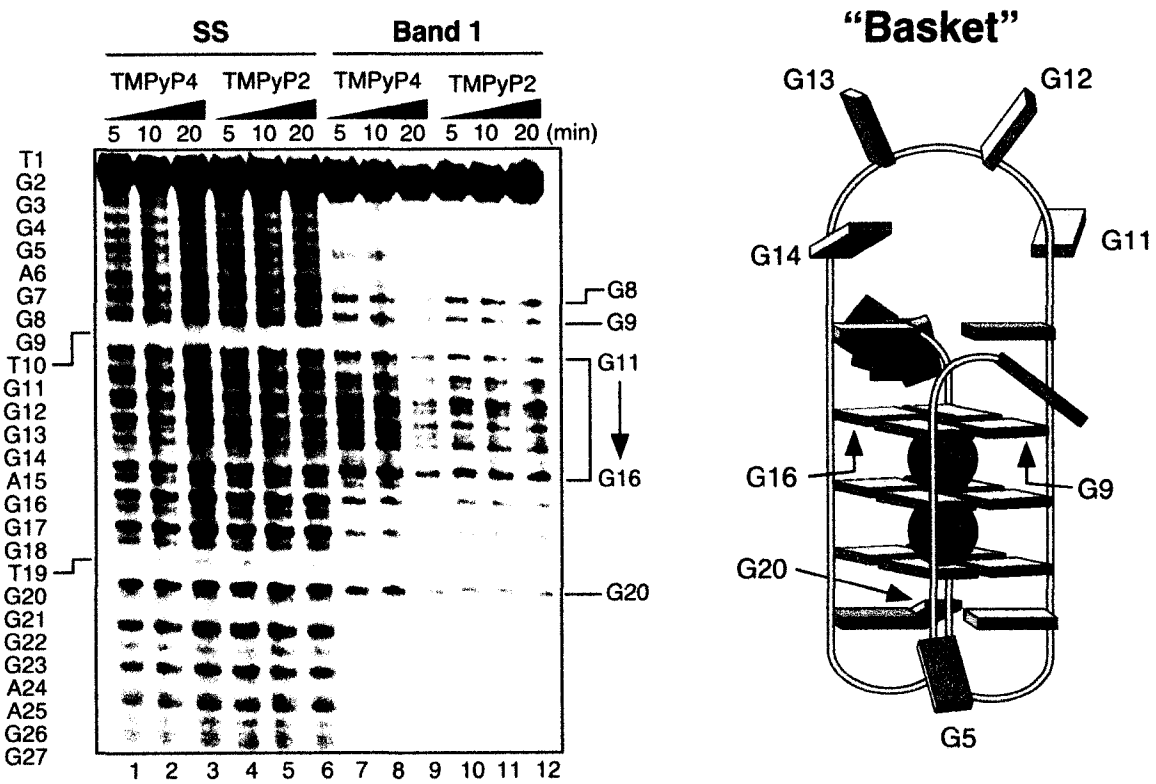


Figure 4.3 (from (157)). Photocleavage patterns of TMPyP2 and TMPyP4 on the basket-type quadruplex.

SS refers to single-stranded (unstructured) Pu27 DNA. Band 1 refers to the band shown in Figure 3.7, which was determined to be a basket-type quadruplex.

Next, the photocleavage patterns of TMPyP2 and TMPyP4 on the chair-type quadruplex were determined in the same way, and are shown in Figure 4.4. In the case of this quadruplex, we found that TMPyP2 had little to no specificity in of binding; each guanine residue was photocleaved approximately equally by TMPyP2, indicating that cleavage was probably mediated by chance encounters between the porphyrin and the quadruplex, and not through any stable specific interaction between the two. TMPyP4, on the other hand, shows a very specific pattern of photocleavage with the chair-type quadruplex. TMPyP4 cleaves guanines at the top and the bottom of the quadruplex, while those in the middle of the structure are protected from cleavage. This indicates that TMPyP4 binds directly above and below the quadruplex, as shown by the large arrows in Figure 4.4. This differential capacity of TMPyP2 and TMPyP4 to bind specifically to the chair-type quadruplex is in accord with the relative abilities of these two molecules to inhibit c-*myc* expression, implying that the chair-type quadruplex may be mediating this biological effect.

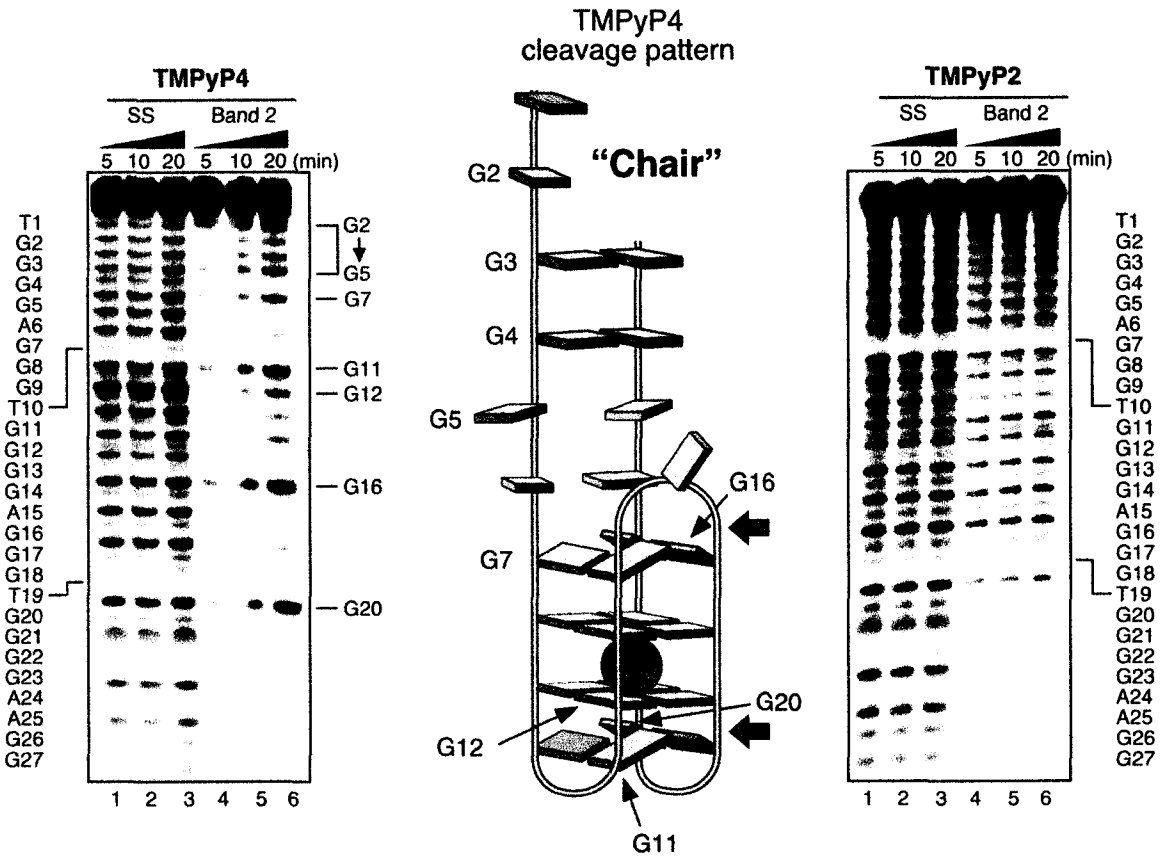


Figure 4.4 (from (157)). Photocleavage patterns of TMPyP2 and TMPyP4 on the chair-type quadruplex.

SS refers to single-stranded (unstructured) Pu27 DNA. Band 2 refers to the band shown in Figure 3.7, which was determined to be a chair-type quadruplex. The photocleavage footprint for TMPyP2 is shown to the right of the quadruplex diagram, and that for TMPyP4 is shown to the left.

d. The Effect of TMPyP4 on c-*myc* Expression Is Dependent upon the Formation of a Chair-Type G-Quadruplex.

It has been shown thus far that a chair-type G-quadruplex can form in the NHE III_1 of the c-*myc* promoter, and that inhibition of the formation of this structure leads to an increase in the basal expression of c-*myc*. However, it has yet to be determined that this structure mediates the effect of TMPyP4 on this gene. To do this, we again turned to the luciferase reporter vector, Del-4, and the quadruplex-specific mutations therein. These plasmids were transfected into HeLa S_3 cells, and the resultant transfectants were treated with TMPyP2 and TMPyP4 for 24 hours, in order to determine if abrogation of either quadruplex would have an effect on the downregulation of c-*myc* by these porphyrins. These data are presented in Figure 4.5.

Note first the effect of these porphyrins on a control luciferase reporter vector, pGL3-Control. This plasmid bears neither G-quadruplex-amenable sequences in its promoter, nor any homology to the c-*myc* promoter, yet both TMPyP2 and TMPyP4 reduce luciferase activity from this vector, to approximately 75% to 80% of untreated levels. This suggests that there is a non-specific, porphyrin-related effect of these molecules on luciferase activity, unrelated to G-quadruplex interactivity. This same level of downregulation also appears with the Del-4 and mutant plasmids, with TMPyP2, suggesting, in this system, that TMPyP2 has little effect on c-*myc*–related expression over and above the non-specific porphyrin-related effect.

When we turn our attention to the Del-4 plasmid, we see that TMPyP4 has an additional effect, bringing luciferase activity down to approximately 40% of the untreated control level, which is in common with what was presented in the last chapter. The same pattern of downregulation is seen with the Null Mutant and Basket Mutant plasmids, with TMPyP4 reducing luciferase activity to approximately 35% to 40% of untreated controls, suggesting that these two mutations have no effect on the ability of TMPyP4 to reduce expression from the c-*myc* promoter. Treatment of the Dual Mutant and Chair Mutant transfectants with TMPyP4, however, yields a very different result. In the context of these two mutations, TMPyP4 is much less able to reduce luciferase expression, bringing luciferase levels down to only approximately 60% of untreated controls. This implies that the single guanine mutations in these two plasmids cause some sort of change in the promoter that partially ablates the effect of TMPyP4. Both of these mutations also disrupt the chair-type quadruplex in the c-*myc* promoter, and we have already shown that TMPyP4 binds very well and very specifically to this structure, stabilizing it *in vitro*. Therefore, we suspect that these mutations reduce the ability of TMPyP4 to bind to a chair-type quadruplex in the promoter, and that this is the cause of this porphyrin's reduced effect.

The question remains, however, of why these mutations do not completely abrogate the effect of TMPyP4 on luciferase expression. If, as we have stated, the Chair and Dual mutations disrupt the formation of the chair-type quadruplex, we should see no effect of TMPyP4 on luciferase expression from these plasmids. This assumes, of course, that the mutations that were made totally overcome the ability of the chair-type

quadruplex to form. Looking back at Figure 13.12, however, we see that these single-guanine mutations only strongly inhibit, but do not completely abrogate, the formation of the quadruplex. Therefore, it is logical to expect that TMPyP4 could bind to the pseudostable mutated chair quadruplex and still inhibit expression from the c-*myc* promoter, but to a lesser extent than the inhibition seen with the wild-type c-*myc* promoter.

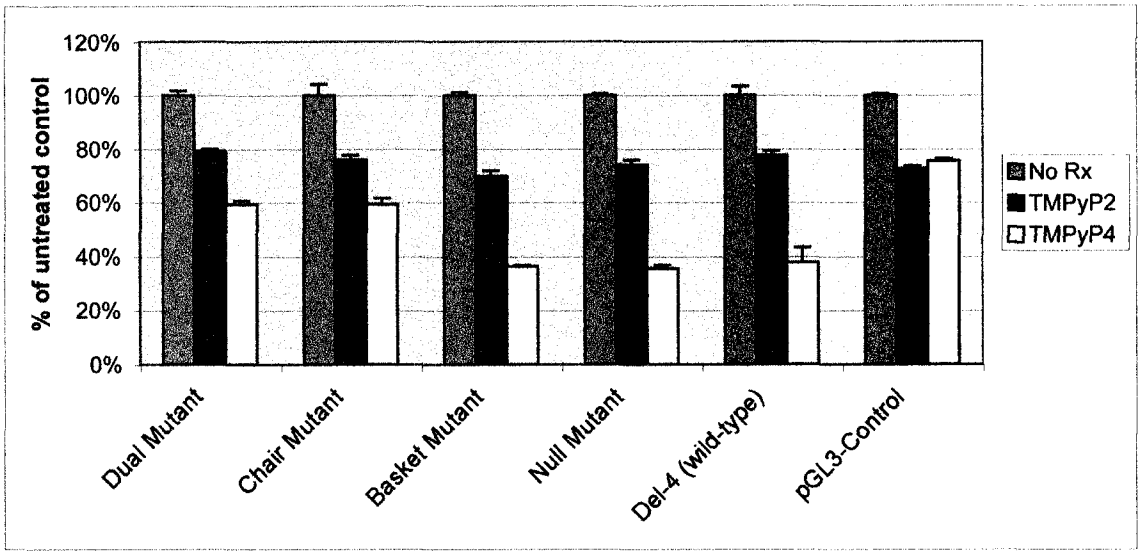


Figure 4.5. Effect of TMPyP2 and TMPyP4 on luciferase expression from mutated c-*myc* promoters.

HeLa S_3 cells were transfected with the Del-4 and mutant vectors, along with a Renilla luciferase vector (as a control for transfection efficiency), and treated 24 hours later with 100 μM TMPyP2 or TMPyP4. Luciferase activity was assayed after 24 hours of porphyrin treatment. Results are normalized for transfection efficiency, and presented as percent of untreated control for each plasmid.

To test the hypothesis that TMPyP4 is able to partially stabilize the mutated chair-type quadruplex, we tested these mutant sequences in the Taq polymerase stop assay, adding TMPyP4 to the enzyme buffer. The resultant gel is shown in Figure 4.6. It was found that, even in the presence of the Chair and Dual mutations, TMPyP4 was able to stabilize the chair-type quadruplex structure, albeit to a lesser extent than shown with the Basket and Null mutations, or with the wild-type sequence. This is in concordance with the results shown above with the Del-4 mutant plasmids in the presence of TMPyP4. Therefore, our predictions appear to be valid; though the chair quadruplex is sufficiently destabilized to remove this level of regulation from the c-*myc* promoter, it can still be "assisted" by binding of a quadruplex-stabilizing ligand molecule like TMPyP4. Hence, we do not see total abrogation of the activity of TMPyP4 when the quadruplex-forming sequence is mutated.

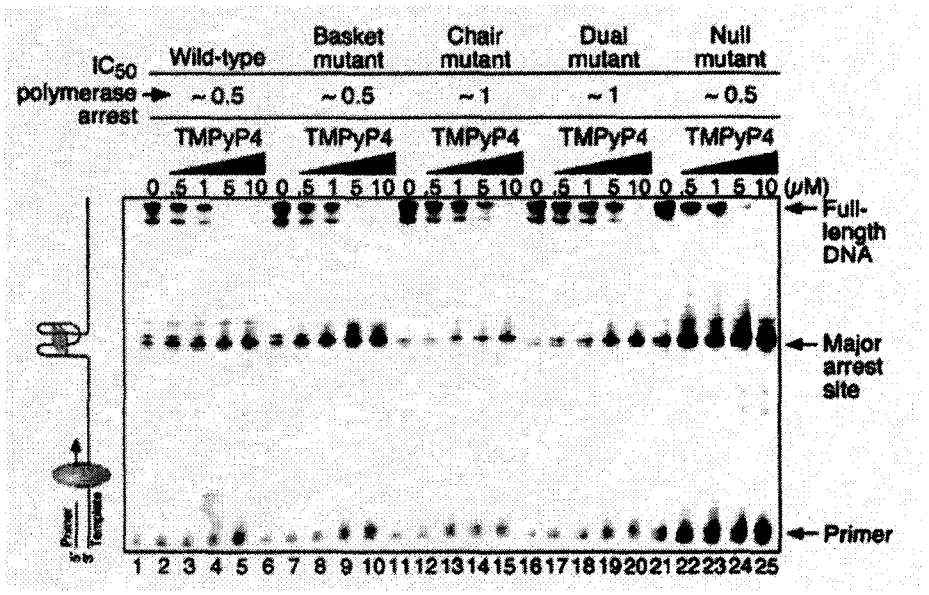


Figure 4.6 (from (157)). Effect of TMPyP4 on the stability of the mutated chair-type quadruplex in the c-*myc* NHE III_1.

The Taq polymerase stop assay was performed, using the purine-rich strand of the NHE III_1 as template, in the presence and absence of TMPyP4. The diagram to the left of the gel shows the principle of the assay, and the approximate position of the quadruplex stop product on the gel.

III. Discussion

Here we have provided evidence that the effect of TMPyP4 on c-*myc* expression is dependent on the formation of a chair-type G-quadruplex in the NHE III_1 of this gene. First, we have shown that, without the presence of the NHE III_1, TMPyP4 no longer has an effect on c-*myc,* using paired Burkitt's Lymphoma cell lines. Following this, we presented data showing that TMPyP4, but not TMPyP2, significantly stabilizes the chair-type quadruplex in the polymerase stop assay. Further, the chair-type quadruplex is the only structure to which TMPyP4 binds specifically; that is, TMPyP4 shows a very specific photocleavage pattern with this quadruplex, while TMPyP2 does not, consistent with our data on c-*myc* downregulation by these molecules. Finally, we find that mutation of the chair-type quadruplex in a c-*myc* promoter-driven plasmid significantly reduces the effect of TMPyP4, while mutation of the basket-type quadruplex does not affect the repressive ability of this drug. Corresponding to this abrogation of effect, these same chair-specific mutations also greatly reduce the stabilization of the quadruplex by TMPyP4. Thus, we can state with confidence that the effect of TMPyP4 on c-*myc* involves stabilization of a chair-type G-quadruplex in the c-*myc* promoter.

These findings are the first to indicate that stabilization of a G-quadruplex secondary structure is a viable means of regulating expression of a gene. While TMPyP4 is not the optimal drug to use for this purpose, as will be discussed in a later chapter, this provides us with a new foundation from which to build new strategies for anti-cancer drug development. As discussed earlier, there are a number of genes, many of which are

involved in cellular proliferation, that bear G-quadruplexes in their promoter sequences. The findings presented in this chapter suggest that these genes might be susceptible to repression by a new class of chemotherapeutic agent: the G-quadruplex-interactive ligands. For this to hold true, however, we must show that the G-quadruplex is a biologically relevant structure in human tumors. Following this, more effective G-quadruplex-interactive compounds, in terms of specificity and therapeutic window, must be developed in order to take advantage of this novel anti-tumor target. These two obstacles will be dealt with, along with the mechanism by which the chair-type G-quadruplex exerts its effects, in the following chapters of this dissertation.

CHAPTER V

THE CHAIR-TYPE G-QUADRUPLEX IS MUTATED IN HUMAN TUMORS

I. Introduction

Having established that a chair-type quadruplex exists in the c-*myc* promoter, and that disruption of this structure leads to increase expression, we posited a new model: Since abrogation of quadruplex formation *in vitro* causes upregulation of c-*myc* expression, perhaps such mutations *in vivo* have similar consequences, providing a means by which c-*myc* may be upregulated in a developing tumor cell. In other words, such mutations may be tumorigenic in nature. To examine this possibility, we began to characterize the NHE III_1 sequence in a number of colorectal cancer specimens taken from patients, as well as corresponding normal tissue. Colon adenomas (polyps) were also examined in this way, in order to determine if mutations to the quadruplex-forming region of the c-*myc* promoter might be early events in the genesis of cancer.

We chose colorectal cancer for a number of reasons. First, and foremost, c-*myc* expression is known to be elevated in colorectal cancer, but this elevation is not due to amplification or translocation of the genomic c-*myc* sequence (166). Therefore, we suspected that, if G-quadruplex mutation is indeed a means by which tumor cells upregulate this gene, colorectal cancer specimens would be the most likely tissue type to find evidence of this. Second, the progression of human colorectal cancer has been characterized extensively by the Vogelstein group and is one of the best-characterized human tumor systems (167). Thus, early precancer forms of this tumor, adenomatous

polyps, are known, and can be included in our study (167). Finally, both tumor and polyp specimens are readily available from researchers in the Arizona Cancer Center, either as unused portions of tumor tissue originally collected for diagnostic purposes, or as part of an earlier study, abrogating the need to obtain new tissue, and accelerating our study significantly.

In addition to characterizing the NHE III_1 sequence from these tumors, we also intended to quantify the amount of c-*myc* expressed for each, to determine whether a correlation exists between c-*myc* NHE III_1 sequence and actual protein expression. We also examined the localization of the β-catenin protein, a downstream effector of the APC pathway, which is known to activate c-*myc* expression (32, 168). Since the APC pathway is frequently mutated in colon cancer, it was necessary to examine its status in our specimens. Finally, in order to correlate our previous results with *in vivo*-relevant data, we have characterized a cell line that bears a mutation in the c-*myc* NHE III_1 sequence that should abrogate G-quadruplex formation. With these data we should be able to definitively state that the chair-type G-quadruplex is a biologically relevant structure *in vivo*, and represents a valuable tumor suppressor element worth studying in other contexts.

II. Results

a. Mutations That Disrupt the Chair-Type G-Quadruplex Exist in Human Tumors

Since mutations to the chair-type quadruplex in the c-*myc* promoter lead to increased transcriptional activation *in vitro*, we hypothesized that similar mutations *in vivo* might have the same effect, resulting in increased expression of c-*myc,* a consequence that may bring about a higher risk of cellular transformation. To examine this possibility, we obtained 21 patient colon tumor specimens, with surrounding normal tissue, and performed laser capture microdissection (LCM) to specifically dissect out either tumor or non-transformed cells. The genomic DNA was harvested from these cells, and PCR was performed to amplify the NHE III_1 region of the c-*myc* promoter. DNA sequencing analysis was then carried out on the PCR products. A flowchart of this procedure is shown in Figure 5.1. Of the 21 tumor samples, 6 were found to bear mutations that should disrupt the chair-type quadruplex structure, which do not appear in the normal tissue. These sequences are shown in Table 5.1. In none of the 21 cases did we find a mutation that would disrupt the chair-type quadruplex in non-transformed cells, indicating that these mutations are selected for in tumor cells. This is expected if such mutations cause c-*myc* expression to be increased, as we have found in our plasmid studies with Del-4. Tumor stage at resection was also examined, to determine if these mutations correlate with the degree of invasiveness or aggression of the tumors. While

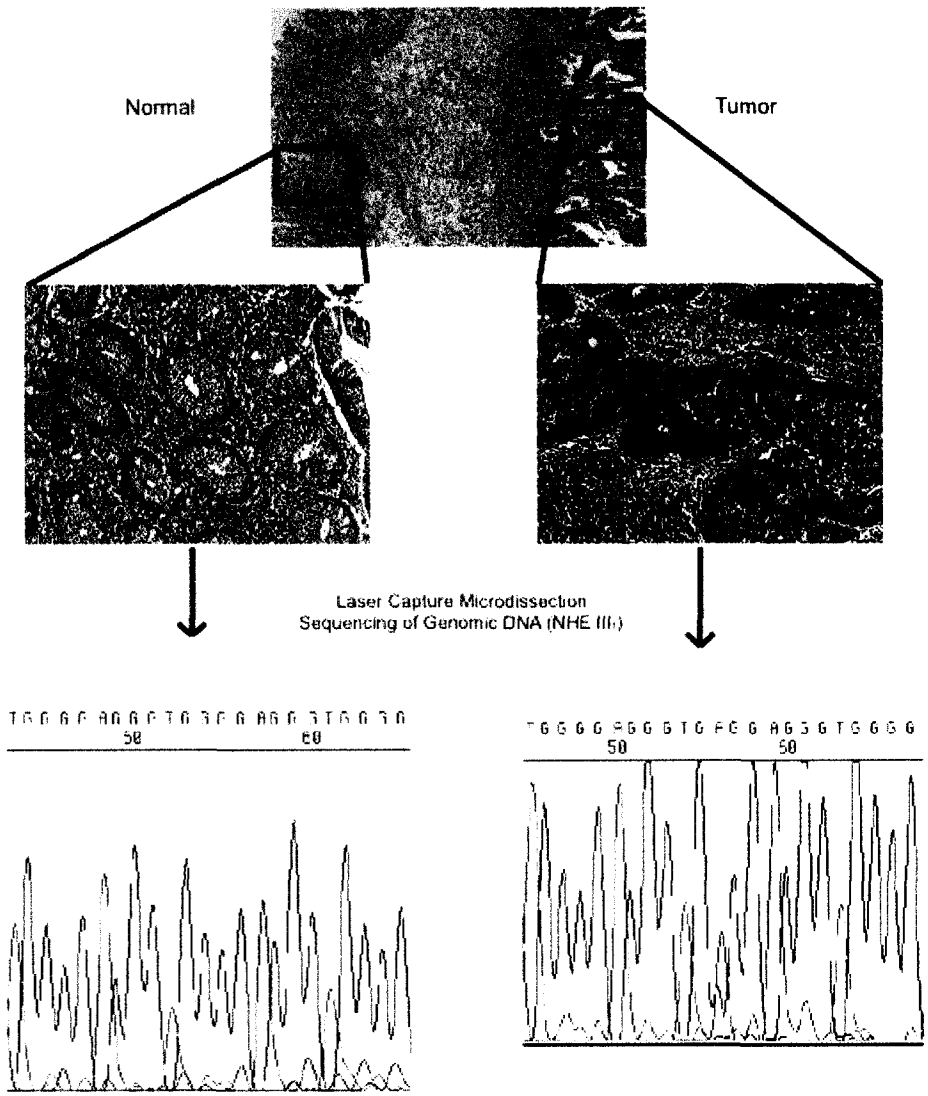


Figure 5.1. Using Laser Capture Microdissection to determine the sequence of the NHE III_1 in human colorectal tumor specimens.

First, tumor specimens were mounted on slides and stained with hematoxylin and eosin for visualization (TOP). Areas of these slides were chosen as candidates for LCM and designated Normal (left side of diagram) or Tumor (right side of diagram). These regions were microdissected, genomic DNA was harvested from these dissected cells, and PCR for the NHE III_1 was performed. The PCR products were sent to the University of Arizona Sequencing facility. The NHE III_1 sequences from normal and tumor cells of each patient were compared to one another (BOTTOM), in an effort to determine if mutations that would disrupt the chair-type quadruplex existed specifically in tumor cells.

the tumors with chair-type quadruplex mutations did show appreciable involvement of the colon tissue (T2 and T3) and local lymph nodes (N2 in two cases), the same was found for tumors without chair-type mutations, indicating that there is no correlation between the mutations and degree of tumor severity.

As c-*myc* is thought to act as an oncogene, we might expect that these mutations that are suspected to upregulate it need only be found in one of the two alleles. In order to determine if the mutations are heterozygous, the PCR products used for sequencing were subcloned into a shuttle vector, pCR2.1 TOPO, so that individual sequences could be examined. Bacteria were transformed with the plasmid plus insert constructs, and 10 individual colonies, each representing a single PCR product, were selected. These colonies were grown overnight and the plasmid DNA was extracted and again subjected to sequencing. The proportion of mutant sequences out of the total 10 subclones is presented in Table 5.1. Note that the only mutations found in 250 base pairs of sequence are those shown in Table 5.1, indicating that these changes are indeed significant and not due to chance single-base alterations.

In order to determine if a correlation exists between the G-quadruplex-destabilizing mutations and c-*myc* expression, immunohistochemistry was performed, using an antibody to the c-MYC protein (Figure 5.2, bottom). In each tumor studied, at least punctate staining for c-MYC was found; while all samples exhibited expression of c-*myc*, no correlation could be found between the intensity of c-MYC staining and the mutations discussed above. This does not come as a surprise, as c-*myc* is very commonly overexpressed in colorectal cancer; the G-quadruplex-specific mutations likely act as

only one of many possible mechanisms through which this gene may be dysregulated. An alternate pathway through which c-*myc* may also be upregulated in colorectal tumors involves abrogation of the APC pathway (32). To investigate if disruption of normal APC function correlates with the quadruplex mutations found in this study, immunohistochemistry for the APC-regulated protein responsible for c-*myc* expression, β-catenin, was also performed (Figure 5.2, top). In 5 of the 21 tumors, β-catenin was not found in the nucleus, where it is expected to localize if it is playing a role in regulation of the c-*myc* promoter. Of these, three tumor samples (3, 4 and 6 in Table 5.1) bore quadruplex-destabilizing mutations. This may suggest that disruption of the repressive chair quadruplex structure in the NHE III_1 can act as a surrogate or complement for APC mutations in upregulating c-*myc* expression.

Sample #	Sequence	Tumor Grade	Heterozygosity
1	TGGGGAGGGTGAGGAGGGTGGGG	T2N0	8/10 mutant
2	TGGGGAGGGTGGGGAGAGTGGGG	T3N0	4/10 mutant
3	TGGGGAGGGTGGGGAGAGTGGGG	T3N0	5/10 mutant
4	TGGGGAGGGTGGGGAGAGTGGGG	T3N0	9/10 mutant
5	TGGGGAGGGTGGGGAGAGTGGGG	T3N2	7/10 mutant
6	TGGGGAGGGTGAGGAGGGTGGGG	T3N2	6/10 mutant

Table 5.1. c-*myc* NHE III_1 Mutations in Human Colorectal Tumors

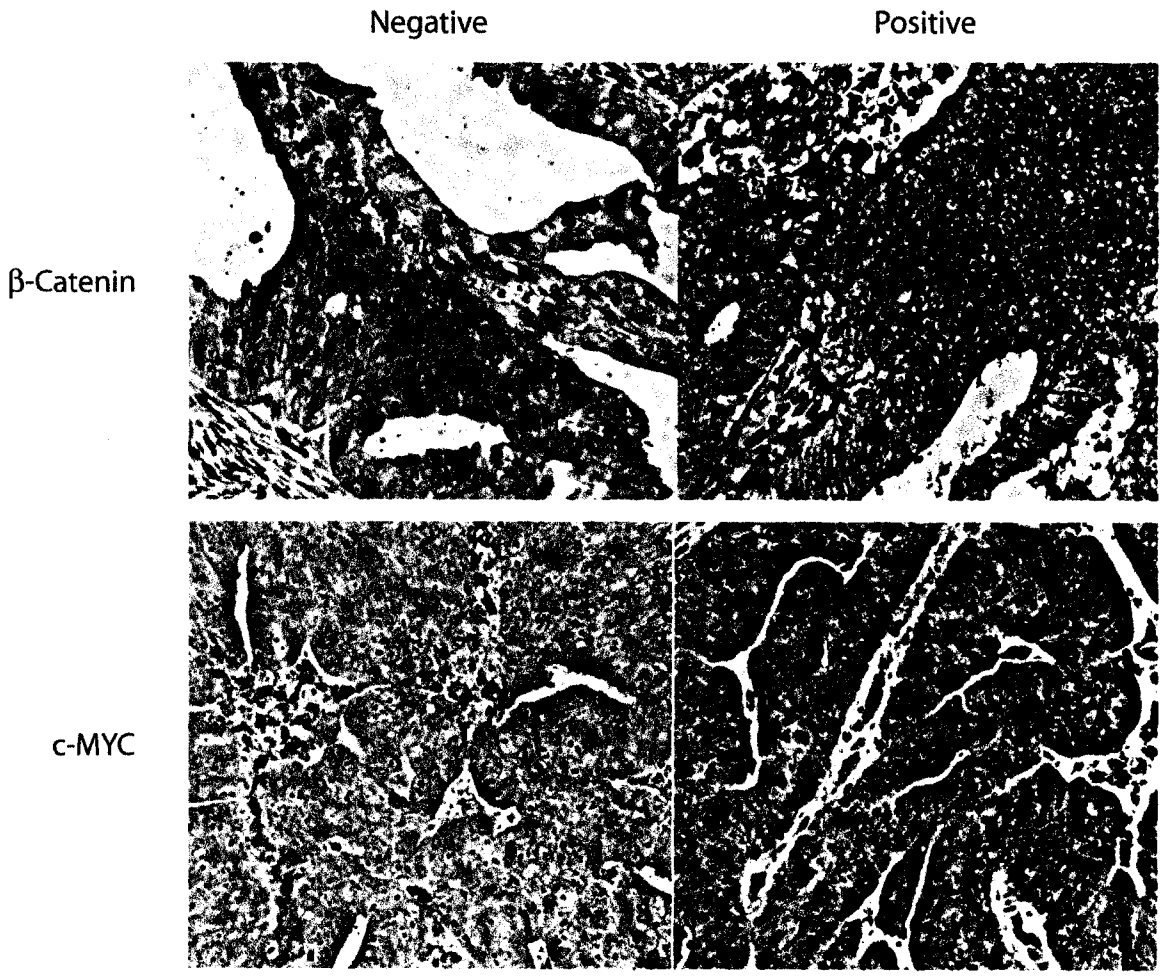


Figure 5.2. Immunohistochemistry for c-MYC and β-catenin in laser capture microdissected specimens.

Immunohistochemistry was performed on the 21 patient specimens laser capture microdissected for sequencing, using antibodies to c-MYC and β-catenin. Representative positive and negative staining tissues are shown.

b. The NHE III_1 Chair Quadruplex Mutation Is a Late Event in Colorectal Tumorigenesis

We had found that a proportion of human colorectal tumors bear mutations which disrupt the chair-type quadruplex and which have been shown *in vitro* to upregulate c-*myc* promoter activity. However, the point in a tumor's development at which these mutations occur was not yet known. That is, we did not know if such mutations were an early or late occurrence in colorectal tumorigenesis. To determine this, twenty human colon adenoma (polyp) samples were procured and tested in the same manner as the colorectal tumors discussed above. Cells were collected by laser capture microdissection, lysed, and the DNA analyzed for NHE III_1 sequence abnormalities. In this case, however, no mutations to the NHE III_1 were found in any of the twenty samples. This indicates that quadruplex-disruptive mutations in the c-*myc* promoter are a later event in tumorigenesis. While finding no mutations of any sort in these polyp samples was surprising, this is not inexplicable. It is widely accepted that c-*myc* has effects beyond induction of proliferation; it also activates apoptotic pathways, possibly to safeguard against upregulation of c-*myc* in the absence of appropriate growth signals, perhaps through interactions with Bin-1, pRb, p53 and p73, and through activation of p53 expression (9, 29-31). In colorectal cancer, p53 mutation or inactivation is a late event in tumorigenesis (167), therefore, overexpression of c-*myc* may be expected to occur later in a tumor's development, after such apoptotic checkpoints have been compromised through other mutations, when tumor cells begin growing out of control. Thus, our results suggest that

the chair-type quadruplex-specific mutations found in colorectal malignancies are not simply chance single nucleotide polymorphisms, but biologically relevant alterations that most likely have important consequences later in the development of this tumor type. To make certain of this, a number of single nucleotide polymorphism databases were searched in regards to the NHE III_1 region of c-*myc*. In no case was any polymorphism discovered, indicating that the mutations we have found do not occur normally in healthy cells. This suggests that mutations in this region are not normally tolerated, and thus are likely dealt with through apoptosis if they occur in the context of a non-dysplastic cell.

c. The Mutated Chair-Type G-Quadruplex Can Be Stabilized in Tumor Cells To Achieve Limited Therapeutic Benefit

We have already shown that mutant NHE III_1 sequences that cannot form a chair type G-quadruplex can still be partially stabilized by quadruplex-interactive agents like TMPyP4 (see Chapter IV). These experiments, however, were performed using plasmids expressing a luciferase reporter gene. Since we have demonstrated that similar mutations exist *in vivo*, we were curious to see if mutant promoter sequences in the context of the c-*myc* gene would behave the same way. A survey of cancer cell lines was taken, examining the sequence of the NHE III_1 in each, to find a line that bears a mutant promoter. One such line was found, with an NHE III_1 sequence as follows:

5' TGGGGA**T**GGTGGGGAGGGTGGGG 3'

This guanine to thymine transversion mutation was found in the polymerase stop assay to disrupt chair-type G-quadruplex formation as well as the mutations we created in the Del-4 plasmid (data not shown) although it occurs in the split tetrad of the chair quadruplex. This cell line was treated with TMPyP2 and TMPyP4, to determine the ability of these two molecules to stabilize the pseudostable G-quadruplex structure and reduce c-*myc* expression (Figure 5.2). We found approximately the same effect as was seen in our site-directed mutagenesis experiments. TMPyP4 was able to reduce c-*myc* levels in this cell line, but not nearly to the extent shown earlier in cells with a wild-type c-*myc* promoter sequence. Similar experiments were performed using doses of TMPyP2 and TMPyP4 of up to 200 μM, for times of up to 72 hours, with comparable results (maximum inhibition of approximately 24% at 200 μM, data not shown). This provides further proof that the effect of TMPyP4 on c-*myc* expression depends on the formation of a chair-type G-quadruplex structure in the NHE III_1. This also shows that even when the c-*myc* promoter is mutated, such that the chair-type quadruplex is unstable (and c-*myc* expression is hence increased) a G-quadruplex-interactive compound like TMPyP4 can be used to reduce expression.

Since this reduction is much less pronounced than that in cells with wild-type sequence, the use of TMPyP4 to counteract c-*myc* overexpression in these cases may not be prudent. However, if other quadruplex-stabilizing agents can be found that stabilize the mutated quadruplex well enough to significantly reduce c-*myc* expression to therapeutically useful levels, such agents might be useful in a clinical setting, regardless of NHE III_1 sequence. With this end in mind, AsPc-1 cells were treated with Se_2SAP, an

analogue of TMPyP4 that has much greater quadruplex-stabilizing properties than TMPyP4 (for a more thorough discussion of this molecule, see Chapter VII). Given that this analogue enters cells rather slowly, a long-term time-course study was performed. These data are presented in Figure 5.3. It was determined that this molecule was also able to reduce c-*myc* expression, to approximately the same levels as TMPyP4, but at a much lower dose. This effect, however, was not realized until 120 hours of drug treatment, similar to what we have found using cell lines with wild-type NHE III_1 sequence (see Chapter VII).

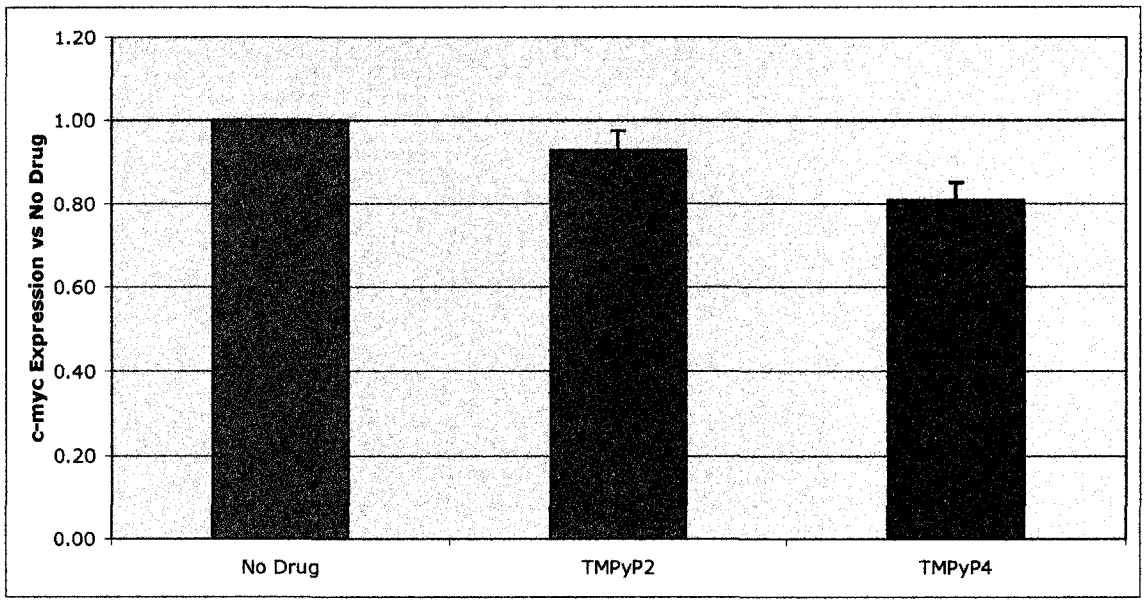


Figure 5.3. Effect of TMPyP2 and TMPyP4 on c-*myc* expression in AsPc-1 cells.

AsPc-1 cells, bearing a G-T transversion mutation in the NHE III_1, were treated with 100 μM TMPyP2 or TMPyP4 for 24 hours. Expression of c-*myc* was assayed using real-time PCR from cDNA, normalized to β-actin expression and is presented as the proportion of normalized expression in untreated cells.

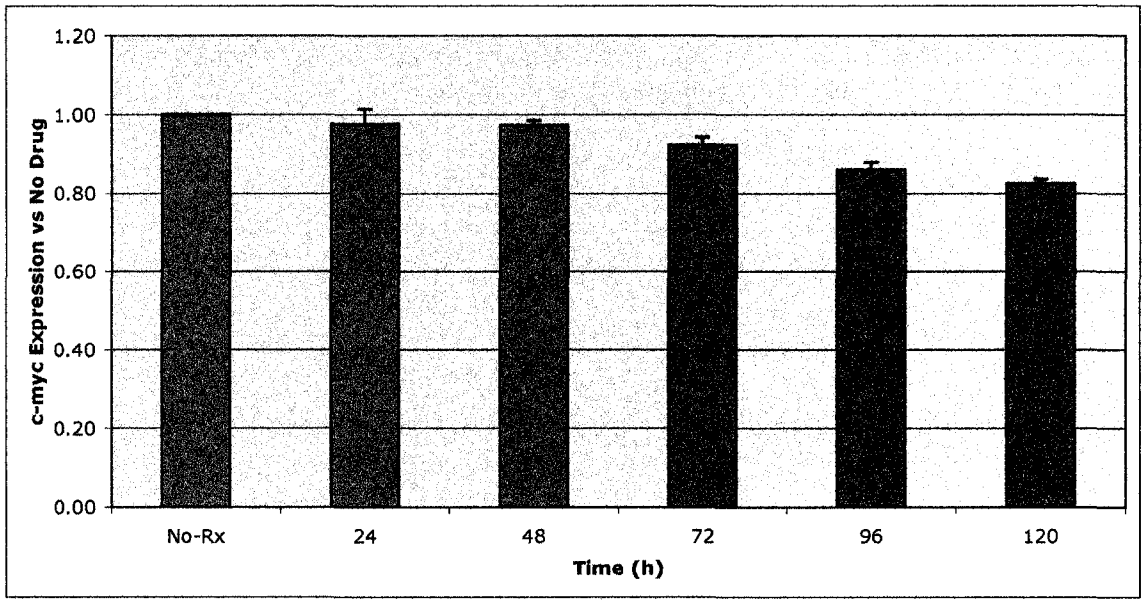


Figure 5.4. Effect of Se_2SAP on c-*myc* expression in AsPc-1 cells.
AsPc-1 cells, bearing a G-T transversion mutation in the NHE III_1, were treated with 10 μM Se_2SAP for 24, 48, 72, 96 or 120 hours. Expression of c-*myc* was determined using real-time PCR from cDNA, normalized to β-actin expression and is presented as the proportion of normalized expression in untreated cells.

III. DISCUSSION

With the data presented here, we have provided strong evidence that a chair-type quadruplex exists in the c-*myc* promoter in living human tissues, and that disruption of such a structure may provide a means through which c-*myc* is upregulated in the process of tumorigenesis. A significant number of colorectal tumor samples studied (nearly 30%) bear a mutation in the NHE III_1 region that destabilizes the quadruplex, a characteristic that has already been shown to increase c-*myc* promoter activity. Corresponding mutations were not found in non-transformed tissue, suggesting that such events are important in tumor development. These mutations were also not found in colon adenoma (polyp) tissue, which suggests that destabilization of the G-quadruplex is a late event in tumorigenesis, perhaps occurring after the apoptotic pathways that c-*myc* would normally stimulate have been compromised.

We have also encountered a pancreatic cancer cell line, AsPc-1, which bears a quadruplex-disrupting mutation in the NHE III_1 region. This line was used to determine if the unstable quadruplex, in the context of the c-*myc* gene, would behave as it had in our luciferase reporter system, Del-4, in response to quadruplex-stabilizing agents. It was found that both TMPyP4 and the core-expanded Se_2SAP could reduce c-*myc* expression in this cell line, but to a much lesser extent than has been seen with cells bearing a wild-type NHE III_1. This is exactly in accordance with what we have seen in Del-4, and suggests that quadruplex-stabilizing agents can be used to reduce c-*myc* levels, even in

cells that have an unstable G-quadruplex, though more effective agents need to be developed.

What remains to be done is a more thorough survey of tumor tissues from many different sources, in order to determine if these quadruplex-specific mutations are found globally in tumors, or only in specific types. Colorectal cancer is known to have disproportionate c-*myc* expression, making it the most likely tumor type in which to find G-quadruplex-disrupting mutations, which is why it was initially chosen for this study. However, this characteristic also makes it very difficult to correlate c-*myc* levels with destabilization of the quadruplex, since all tumors in our study exhibited staining for c-*myc*. Therefore, now that we know such mutations exist *in vivo*, we must continue to look in other tumor types, which only express c-*myc* in a fraction of cases, so that we might correlate these two characteristics.

CHAPTER VI

THE MECHANISM THROUGH WHICH THE CHAIR-TYPE G-QUADRUPLEX EXERTS ITS EFFECTS

I. Introduction

At this point, we had established that the NHE III_1 sequence in the c-*myc* promoter has the potential to form a chair-type G-quadruplex structure, and that this structure acts as an inhibitor of transcription; stabilization of this structure with G-quadruplex-interactive ligands leads to decreased expression, while abrogation causes upregulation and lowered sensitivity to such compounds. We became curious, then, as to the mechanism of action of this 3-dimensional DNA structure. A number of possible models were hypothesized: The chair-type G-quadruplex may recruit protein factors that repress transcription from the c-*myc* promoter. If this is the case, disruption of the quadruplex would remove this binding site, and allow transcription to proceed unchecked. Such a model would predict that the chair-type G-quadruplex could be used as a probe to pull out this factor. If, on the other hand, G-quadruplex formation in the NHE III_1 prohibits activator protein binding, disruption of the quadruplex would restore the binding site for such a factor, allowing transcription to carry on. For this model to hold true, mutant NHE III_1 sequences that cannot form the chair-type quadruplex could be used to selectively bind to and recover such a protein. The possibility has also been suggested that the quadruplex acts simply as a physical block to polymerization by RNA

polymerase II, inhibiting transcription by simply being "in the way." While such a possibility exists for other genes with G-quadruplex amenable sequences, it is doubtfully the case for c-*myc,* since the NHE III_1 region lies upstream of the two major start sites for transcription.

In order to determine whether the quadruplex structure in the NHE III_1 is acting as a repressor through binding of an inhibitory factor or displacement of an activator, we proposed to use electrophoretic mobility shift assays (EMSAs) to determine 1) if there is a G-quadruplex-specific protein that binds to this sequence, and 2) if a protein exists that binds only when the chair quadruplex cannot form, with the understanding that both or neither might be true. Having already synthesized sequences that could not form one or both of the chair-and basket-type quadruplexes, we simply needed to devise an assay that would allow us to select for specific protein binding to these sequences from a pool of likely candidates. Once this had been accomplished, we intended to use mass spectroscopy to sequence the proteins identified in the EMSA. As we later found, the amount of protein binding to these oligonucleotides was insufficient for sequencing, leaving us with the option of using a "guess and check" method of identifying specific protein shifts. Finally, we determined that even the EMSA was too non-discriminating a means for identifying proteins using the protocol we designed, leaving us with the decision to test a single likely candidate protein in a cell culture model. Fortunately, this gamble paid off in the end, resulting in a protein that likely is at least partially responsible for the effects of the chair-type G-quadruplex we had observed thus far.

II. Results

a. Correlation of Protein Binding to the NHE III_1 with Activity of the c-*myc* Promoter

We have now determined that there is a chair-type G-quadruplex in the c-*myc* regulatory region that appears to act as an inhibitory element. It is now necessary to determine how this structure might be effecting such a function. Since there has been no evidence suggesting that RNA polymerase II must traverse the NHE III_1 during c-*myc* transcription, it is unlikely that the quadruplex structure is simply acting as a physical block to polymerization, as it does in the Taq polymerase stop assay. Therefore, we hypothesized that the formation of this structure may be affecting the ability of one or more proteins to bind to this region. Perhaps the formation of the chair quadruplex prevents binding of an activator protein to one or both of the strands of the NHE III_1, or conceivably the G-quadruplex or the free C-rich strand (possibly involved in a secondary structure itself) is instead a recognition site for an inhibitor protein. To determine which, if either of these hypotheses is true, electrophoretic mobility shift analyses (EMSAs) were performed, using the purine-rich strand (Pu27 or Pu45) the pyrimidine-rich strand (Py27 or Py45) or the double stranded (ds27 or ds45) versions of the NHE III_1.

We first studied the purine-rich strand of the NHE III_1, as this contains the G-quadruplex-forming region. Oligonucleotides corresponding to 27 bases of the c-*myc* promoter, including the G-rich strand of the NHE III_1, were synthesized, including mutated sequences corresponding to each of the Dual, Chair, Basket and Null Mutants of

Del-4. These oligonucleotides were used in EMSA studies with HeLa nuclear extracts and 1 μM TMPyP2 or TMPyP4. The gel is presented in Figure 6.1. A similar pattern was obtained when a 45 base oligonucleotide from the c-*myc* promoter, including the NHE III_1, was used in this assay (data not shown). It is first apparent that there is no shift that is specific to the formation of a quadruplex structure, either the basket or the chair-type. This indicates that the effects of the quadruplex mutations in Del-4 are not the result of binding of a quadruplex-binding protein. We do find, however, a mobility shift that is common only to the Chair and Dual mutant oligonucleotide sequences, indicated by an arrow in Figure 6.1. This shift is reduced in intensity, but not entirely lost, with the addition of TMPyP4, which correlates well with the ability of TMPyP4 to stabilize the chair-type quadruplex in both of these mutant sequences.

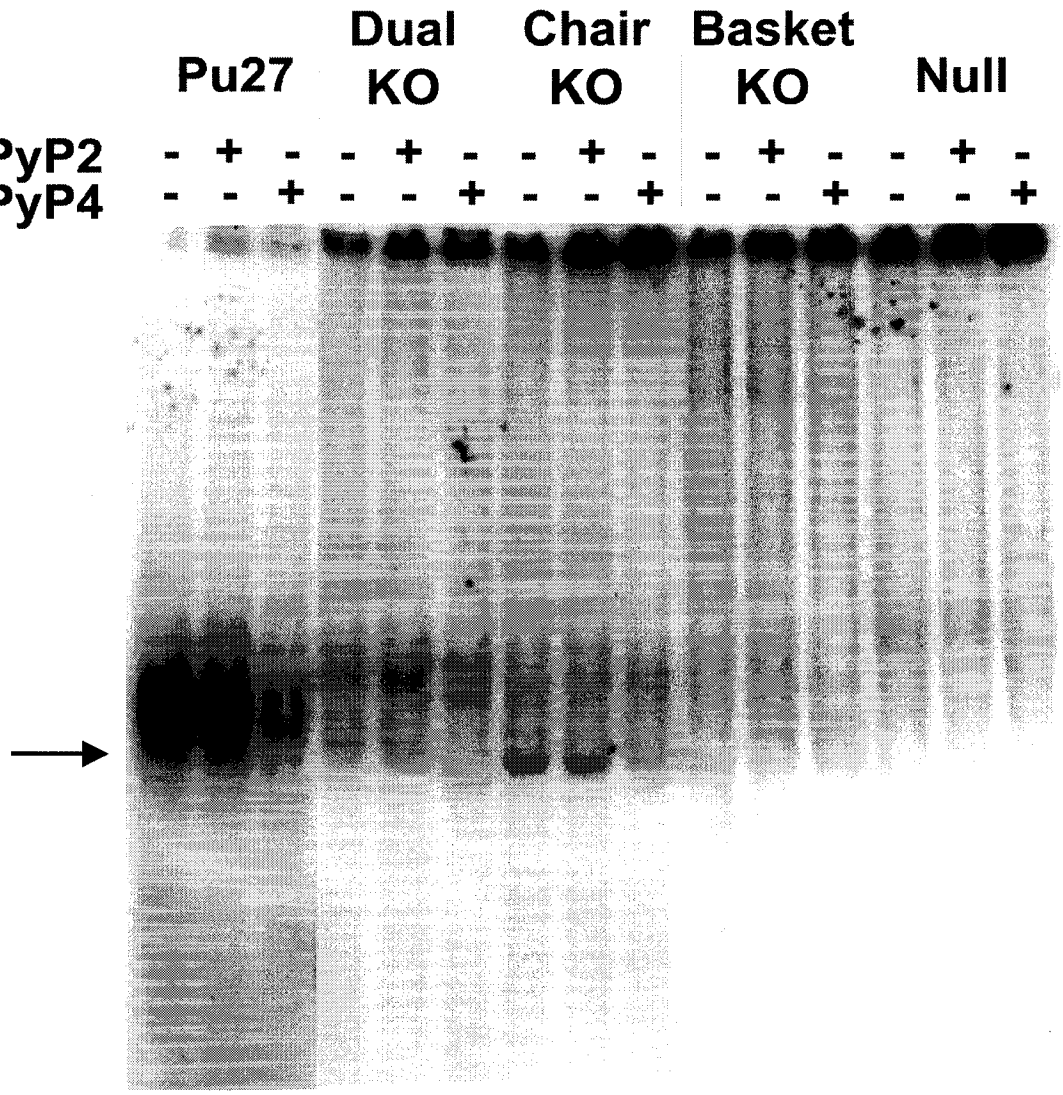


Figure 6.1. Electrophoretic mobility shift assay, using the Pu27 oligonucleotide and single-guanine mutants.

HeLa nuclear extracts were used to determine what, if any, proteins bound to each of these sequences. TMPyP2 and TMPyP4 were also added to evaluate the effect of stabilization of the G-quadruplex on protein binding. KO refers to knockout mutations of the quadruplex indicated. The arrow indicates the position of a shift present only for the Chair and Dual mutant oligonucleotides.

In order to determine the effect of quadruplex formation on the ability of this unknown protein to bind the NHE III_1, a new oligonucleotide was synthesized. This oligonucleotide, called Pu27-PPG, contains the nucleoside analog PPG (Figure 6.2) in place of the three guanine residues mutated in the Dual, Chair and Basket mutants. PPG is identical to guanine, save for the nitrogen atom at the 7 position. This nitrogen atom is instead found at the 8 position. This small change disallows G-quadruplex formation, since the N_7 of guanine is required for both Hoogsteen and reverse Hoogsteen base pairing. Canonical Watson-Crick base pairing is unaffected. Therefore, Pu27-PPG will form neither a chair- or basket-type G-quadruplex. Pu27 and Pu27-PPG were compared by EMSA in terms of their response to TMPyP2 and TMPyP4. Unfortunately, this gel and others following, despite repeated attempts, was very unclear, especially when converted to an image file, and exhibited different patterns of shifts with different attempts. The gel is shown in Figure 6.3 for the purpose of demonstration. A mobility shift corresponding to that seen with the Dual and Chair mutant oligomers is seen with Pu27-PPG. This shift also appears with Pu27, but is much weaker. Interestingly, this shift is insensitive to TMPyP2 and TMPyP4 treatment when Pu27-PPG is used as a probe, but disappears with TMPyP4 treatment when Pu27 is used, consistent again with the relative abilities of these two oligonucleotides to form a G-quadruplex structure. Therefore, the binding of this particular protein is dependent upon the inability of a chair-type quadruplex to form in the NHE III_1 sequence.

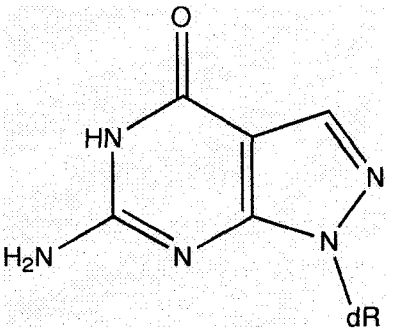


Figure 6.2. Structure of PPG (7-deaza-8-aza-deoxyguanosine).
Note that a nitrogen atom has been moved from its normal position 7 in deoxyguanosine to position 8 in this molecule.

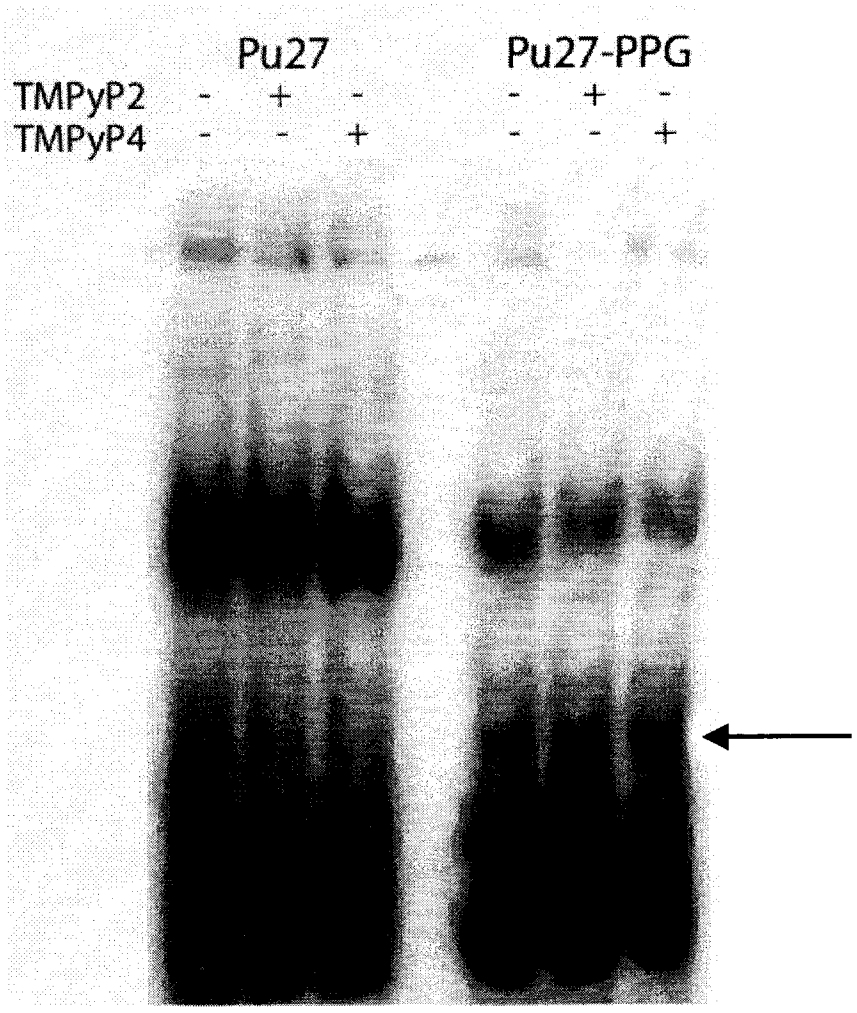


Figure 6.3. Effect of quadruplex formation on binding of proteins to the NHE III_1 of c-*myc*.

Pu27 and Pu27-PPG were used in an EMSA with HeLa nuclear extract and TMPyP2 or TMPyP4. The arrow represents a specific shift that is very weak and sensitive to TMPyP4 with Pu27, but which is quite strong and insensitive to porphyrin treatment with Pu27-PPG.

This may mean that a particular protein or protein complex is able to bind to the NHE III_1 purine-rich strand only when the chair-type quadruplex cannot form. This implies that perhaps the increase in c-*myc* promoter activity seen with these mutations in Del-4 is a result of the increased binding of an activator, which is unable to bind when the quadruplex is present.

For completeness, EMSA assays were also performed on the pyrimidine-rich strand of the NHE III_1 and on the double-stranded NHE III_1. These results were also repeated, using oligonucleotides of length 27 from the c-*myc* promoter (data not shown, due to poor image quality). In each case, though a number of shifts were seen using HeLa nuclear extract, none were specific to the formation or disruption of either quadruplex, and none correlated with the data already presented from the Del-4 mutant plasmid transfections. In the case of the double-stranded NHE III_1, we attempted to determine if NM23-H2 differed in its ability to bind and cleave each of the mutant sequences, but found that it did not cause a shift with any of the 45 base oligonucleotides. This is likely due to poor protein quality. Dr. Edith Postel of Princeton University has specialized in the study of this multifunctional protein, and we have forged a collaboration with her; her laboratory has been contacted and will hopefully be able to succeed in characterizing the interaction of NM23-H2 with each of our mutant sequences.

b. Sp1 Binding Correlates to Increased Activity of the c-*myc* Promoter

A single protein, then, was found whose ability to bind to the purine-rich strand of the NHE III_1 is increased when the chair-type quadruplex cannot form. The identity of this protein, however, still remained unknown. To establish this, more electrophoretic mobility shift assays were performed, using specific competitor oligonucleotides that represented consensus binding sites for various proteins known to bind to this region. Two such proteins, cellular nucleic acid binding protein (CNBP) and Sp1, have been studied with some rigor, and so represented the most logical proteins to study. Again, the 45 nucleotide-long purine rich oligonucleotides containing the NHE III_1 were labeled and incubated with HeLa nuclear extracts. To these binding reactions, a commercially available Sp1 consensus oligonucleotide or a consensus serum response element (SRE), to which CNBP is known to bind, was added, unlabeled, to act as a specific competimer for either of these two proteins. This gel was again very difficult to read and to reproduce, and so the image is not included in this dissertation, as the image on paper is impossible to evaluate. However, both of these proteins were identified as specific shifts in this assay. CNBP was not identified as the unknown protein, as it binds to all of the mutant oligonucleotides. It was found, however, that the shift which appears when the chair-type quadruplex cannot form (the Dual and Chair Mutants) is competed away with the addition of unlabeled Sp1 consensus oligonucleotide. A similar EMSA was performed, using an Sp1 blocking antibody, and the unknown shift was again competed away (data not shown). This suggested that the unknown protein is indeed Sp1.

c. Sp1 Binding Is Required for Increased c-*myc* Promoter Activity When the Chair-Type G-Quadruplex is Mutated

To validate that Sp1 is the protein responsible for binding to the purine-rich strand of the NHE III_1 when the chair-type quadruplex is mutated and consequently upregulating c-*myc*, a way to remove Sp1 from our experiments was needed. The Schneider's S2 cell line is a *Drosophila* cell line that lacks expression of this protein. These cells were transfected with the same Del-4 and mutant plasmids as were used earlier, to determine if the absence of Sp1 would have an effect on luciferase expression from these vectors. If Sp1 was responsible for the increase in c-*myc* activity when the chair-type quadruplex is prevented from forming, we would expect no increase in this cell line that lacks Sp1. The results of these transfection experiments are presented in Figure 6.4. It was found that, without the influence of the Sp1 transcription factor, there was no appreciable difference in activity between any of the mutant plasmids. This contrasts with what was in earlier experiments, in which the Chair and Dual Mutant plasmids had increased c-*myc* promoter activity in comparison to the wild-type plasmid and the other mutants, and implies that Sp1 is important in creating this increased activity when the chair-type quadruplex cannot form.

If this were indeed the case, we would expect that TMPyP4 should have much less effect on the activity of these plasmids as well, if our model is correct. Schneider's S2 cells were again transfected with the Del-4 and mutant series of plasmids, and these

transfectants were treated with 100 μM TMPyP2 and TMPyP4. The effects on luciferase expression can be seen in Figure 6.5. In the context of this cell line, TMPyP2 and TMPyP4 have very little effect on c-*myc* promoter activity, as measured by this assay, suggesting that the effect of these two porphyrins on c-*myc* expression is dependent on the presence of Sp1.

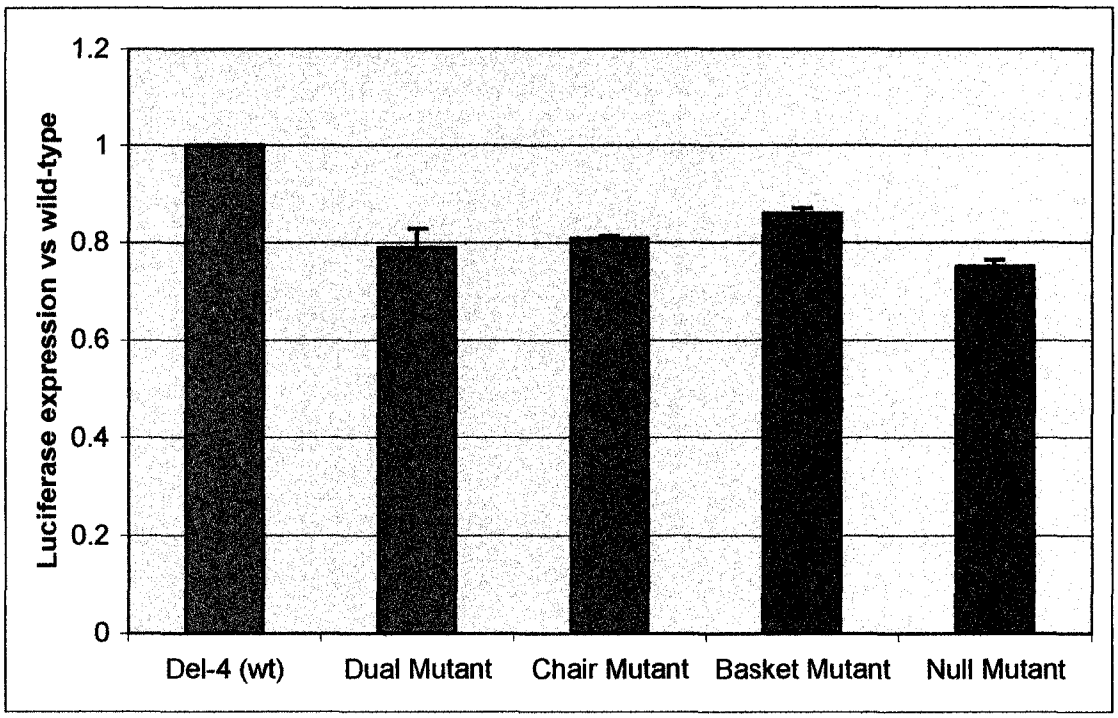


Figure 6.4. Effect of single-guanine mutations on c-*myc* promoter activity in Schneider's S2 cells.

Cells were transfected with the Del-4 and mutant plasmids and assayed for luciferase activity. Results are corrected for transfection efficiency and displayed versus the luciferase activity of the parent (Del-4) vector. Error bars represent one standard error above the mean.

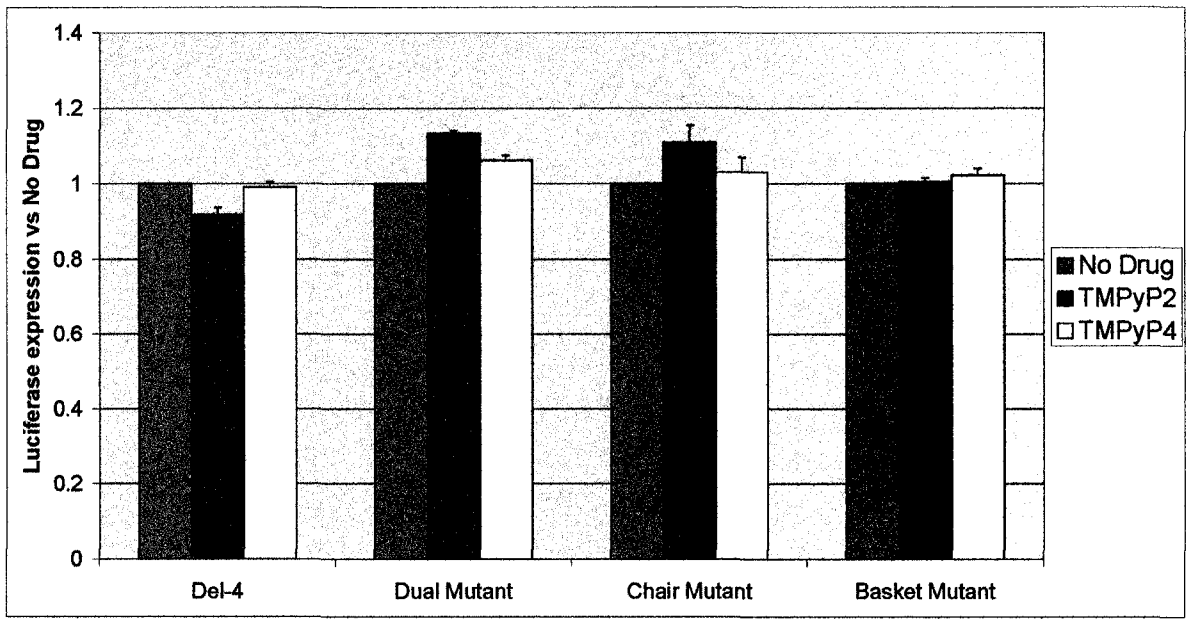


Figure 6.5. Effect of TMPyP2 and TMPyP4 on c-*myc* promoter activity in Del-4 and mutant vectors in Schneider's S2 cells.

Cells were transfected with the Del-4 and mutant vectors, then treated with 100 μM TMPyP2 or TMPyP4, and cell extracts were assayed for luciferase activity. Results are corrected for transfection efficiency and displayed versus the No Drug control for each plasmid. Error bars represent one standard error above the mean.

If Sp1 is responsible for the increase in c-*myc* expression when formation of the chair-type quadruplex is hampered, then adding Sp1 expression to the Schneider's S2 cell line should reestablish this effect. The pPacSp1 plasmid was provided by Dr. Robert Tjian of the University of California at Berkeley, and expresses the Sp1 transcription factor from an actin promoter. This plasmid was co-transfected into Schneider's S2 cells, along with the Del-4 and mutant plasmids, to determine if this prediction is true. The results of this experiment are found in Figure 6.6. It was discovered that, when Sp1 is reintroduced to the system, the effect of mutating the chair-type G-quadruplex is also reestablished; luciferase activity from the Chair and Dual Mutant plasmids is approximately 140% of that from the parent vector, while the Basket and Null Mutant plasmids show no change in activity compared to Del-4. These data, taken together, confirm that the Sp1 transcription factor is responsible for the effect of mutating the chair-type quadruplex in the c-*myc* NHE III_1; Sp1 is likely inhibited from binding to the NHE III_1 when the quadruplex is present, but upon abolition of this structure, Sp1-mediated stimulation of c-*myc* expression can proceed. When small molecules like TMPyP4 stabilize the quadruplex, the ability of Sp1 to bind is reduced even further. Thus, a relatively ubiquitous transcription factor can provide a means to regulate the expression of a gene under tight control, through modulation of a secondary structure at that factor's binding site.

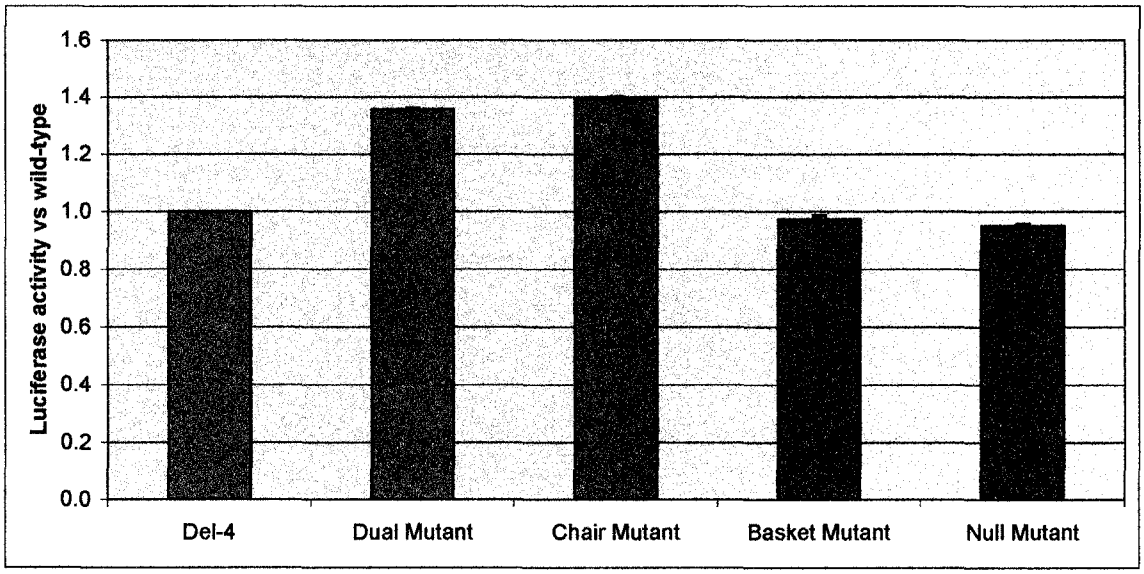


Figure 6.6. Effect of Sp1 on luciferase expression from Del-4 and mutant plasmids in Schneider's S2 cells.

The pPacSp1 plasmid, a gift from Dr. Robert Tjian was used to co-transfect Schneider's S2 cells, along with Del-4 and the mutant series of plasmids. Data are presented as luciferase activity corrected for transfection efficiency, and as a proportion of the activity seen in the Del-4 transfectant. Error bars represent one standard error above the mean.

III. Discussion

Here we have established a possible mechanism through which the chair-type G-quadruplex in the c-*myc* promoter mediates at least some of its effect. The binding ability of the Sp1 transcription factor to the NHE III_1 sequence is enhanced when this sequence is mutated, such that it can no longer form the G-quadruplex. This effect was found with two separate mutations, corresponding to the mutations made in our earlier reporter assay studies, and so is likely not due to a change in the primary structure of the Sp1 binding site. The biological effect of this enhanced binding was evaluated using a cell line that lacks expression of Sp1, and to which exogenous Sp1 was added. It was found that, in the absence of Sp1, reporter constructs with chair quadruplex mutations in the NHE III_1 behaved identically to constructs lacking such mutations. When Sp1 was used to reconstitute this system, the reporter vectors with chair-type G-quadruplex mutations showed increased activity, as we have seen in our earlier studies. This suggests that the repressive effect of the chair-type quadruplex is mediated through inhibition of Sp1 binding; Sp1 auxotrophy can substitute for quadruplex formation in repression of c-*myc* activity.

Therefore, the second model put forth in the introduction of this chapter appears to be the accurate one: The chair-type G-quadruplex acts as a physical block to transcription factor binding in the c-*myc* promoter, and only resolution or disruption of this structure will lead to high levels of transcriptional activity. It is of significance to note that Sp1 binding sites have also been identified in the promoter regions of the RET

proto-oncogene (169, 170) and HMGA2, a gene implicated in hyperproliferative disorders and hyperobesity (171-173). These same promoter regions have been shown in this laboratory to form G-quadruplex structures *in vitro*, suggesting that this mechanism of quadruplex-mediated inhibition of gene expression is not limited to c-*myc*. What remains to be seen is exactly how quadruplex formation affects the binding of Sp1; the EMSA results suggested that the unknown protein, identified as Sp1, binds best **to the purine-rich strand** of the NHE III_1 when quadruplex formation is inhibited. Sp1, however, is a double-strand DNA-binding protein. It may be that Sp1 bears some ability to bind to hairpin structures, essentially double-stranded regions of DNA formed by a single molecule, though such an activity has not been reported to date. This hairpin binding activity might play a role in stopping the quadruplex from forming, allowing proteins such as CNBP and hnRNP K to access their binding sites and activate transcription. Alternatively, Sp1 binding in itself may be sufficient to activate transcription. Further study of the binding modes of Sp1, as well as the other proteins known to bind in this region, is underway in this laboratory, to elucidate the role of Sp1 and the chair-type G-quadruplex in c-*myc* regulation.

In addition to the role of Sp1 in quadruplex-mediated repression of gene expression, we have also learned a great deal about methodology from these studies. It appears that the electrophoretic mobility shift assay as performed here is not the optimal means of searching for quadruplex-specific (or quadruplex non-binding) proteins. In the future, I believe that the use of specific recombinant proteins, instead of crude nuclear extracts, will yield much cleaner and more reproducible shifts. For such studies, however,

we will need to have a specific subset of proteins in mind for possible candidates. In the case of the c-*myc* promoter, a number of factors are already known to bind to this region and have important roles in transcription, so the Sp1, CNBP, hnRNP K and NM23-H2 factors can be directly tested for their abilities to bind to the mutated and non-mutated NHE III_1 sequences. Other genes with quadruplex-amenable sequences, including RET and HMGA-2, and those listed in the Introduction to this dissertation, have also been studied to some extent, and protein factors that bind to these regions have been identified, giving us at least an idea of where to start looking. However, a means to select for and identify unknown quadruplex-specific proteins is still required. Using the single stranded Pu27 oligonucleotide as a probe for protein precipitation appeared to be such a means, but too little protein was obtained by this method to be of use for protein sequencing. Perhaps attaching concatemers of this DNA sequence to a bead would yield enough protein for sequencing, but the possibility then arises of formation of other irrelevant G-quadruplex structures in such a sequence. In short, to identify the proteins that may play a role in the action of G-quadruplexes in gene promoters, there are two choices: 1) We must engage in "fishing expeditions" using purified proteins, in the hope that we know enough about these promoters to infer which proteins are likely candidates, or 2) a better protocol for precipitating proteins from a crude mixture using G-quadruplex-forming oligonucleotides must be designed.

CHAPTER VII

DEVELOPMENT OF G-QUADRUPLEX-INTERACTIVE COMPOUNDS TO INHIBIT c-*MYC* EXPRESSION

I. Introduction

Although we have used TMPyP4 as a demonstration molecule throughout the studies that have been discussed so far, this molecule is far from ideal in terms of clinical use. As mentioned earlier, TMPyP4 is a photoactive compound; because of the planarity and conjugation (entirely double bonded character) of its porphine core, it can absorb energy from light of particular wavelengths, with a peak at 411nm. This absorption can have one of many effects, three of which are of interest to us: The energy can be re-released as a particle, or photon, of light, it can be transferred to an oxygen molecule, creating a singlet oxygen free radical, or it can transfer this energy directly into a guanine residue of a nucleic acid, eventually creating a guanine cation. The latter of these two consequences can cause cleavage of DNA at guanine moieties, resulting in the photocleavage discussed in Chapter IV. However, though this photocleavage is a useful tool for determining the binding mode of TMPyP4 to DNA, it is also a hindrance to the use of this molecule as a therapeutic agent, acting as a source of toxicity. In fact, the ability of porphyrins to cause light-mediated cleavage of DNA has been used in a quite different way to treat disease, termed photodynamic therapy, in which this phototoxicity is exploited for killing undesired cells (174). Nevertheless, we intend to use small

molecules in a more specific way, to bind to and stabilize G-quadruplex DNA structures, making phototoxicity an obstacle rather than a boon.

In addition to ridding our small molecules of phototoxicity, we desire to find ways to tailor them, increasing their specificity and affinity for the chair-type G-quadruplex in the NHE III_1 of the c-*myc* promoter. One of the reasons for the high stability of G-quadruplex DNA, in addition to the Hoogsteen base pairs between guanines in each G-tetrad, lies in the planarity and conjugation of the guanine nucleotides themselves. Because of these two characteristics, there is essentially a "cloud" of electron density above and below each of the tetrads, called a π electron cloud. Interactions of these π clouds with one another, termed π-π stacking, leads to increased stability of a G-quadruplex structure. TMPyP4 stably binds to the ends of a G-quadruplex in a similar way, taking advantage of this π-π stacking, resulting in a durable interaction with the DNA. It is this characteristic of TMPyP4 that we wish to maintain in any new molecules that are synthesized.

With these considerations in mind, two chemists in the Hurley laboratory, Drs. Jeyaprakashnarayanan Seenisamy and Sridevi Bashyam, set out to create new porphyrin analogues that would take advantage of, and improve upon, the beneficial features of TMPyP4, while abrogating phototoxicity. Modifications to the porphine core atoms were made, such that phototoxicity would be reduced or removed altogether, in the hopes that this would not hinder π-π stacking between the molecule and the G-tetrads of the c-*myc* G-quadruplex. These changes are represented by the thia- and selena-porphyrins SHPy3, S_2Py3, SeHPy3 and Se_2Py3, as shown in Figure 7.1, Group III. Photoactivity of these

porphyrins is nearly undetectable; however, the porphine core is no longer planar in the molecules, but is buckled up or down by the presence of the selenium or sulfur atoms. The porphine core was also contracted or expanded, in order to better maximize the π-π interactions between the porphyrin analogues and the G-tetrads. These changes are modeled in the corroles (TMPyCor3 and TMPyCor4, Group I), which bear a contracted core ring, and in the diselenasapphyrin Se_2SAP, Group II, which has an expanded ring and also includes the core atom modifications described earlier to reduce photoactivity. Finally, modifications to the meso groups of the porphyrin were made, in order to better fit the grooves of the G-quadruplex and better interact with the negatively charged phosphate-sugar DNA backbone. Such a change is reflected in the pyrazole porphyrin OMPzP4, Group IV. Note that other members of these four groups have been or are in the process of being synthesized, but the study of such a large number of compounds is beyond the scope of this dissertation, so the compounds listed above have been chosen as representatives for this structure-activity relationship study.

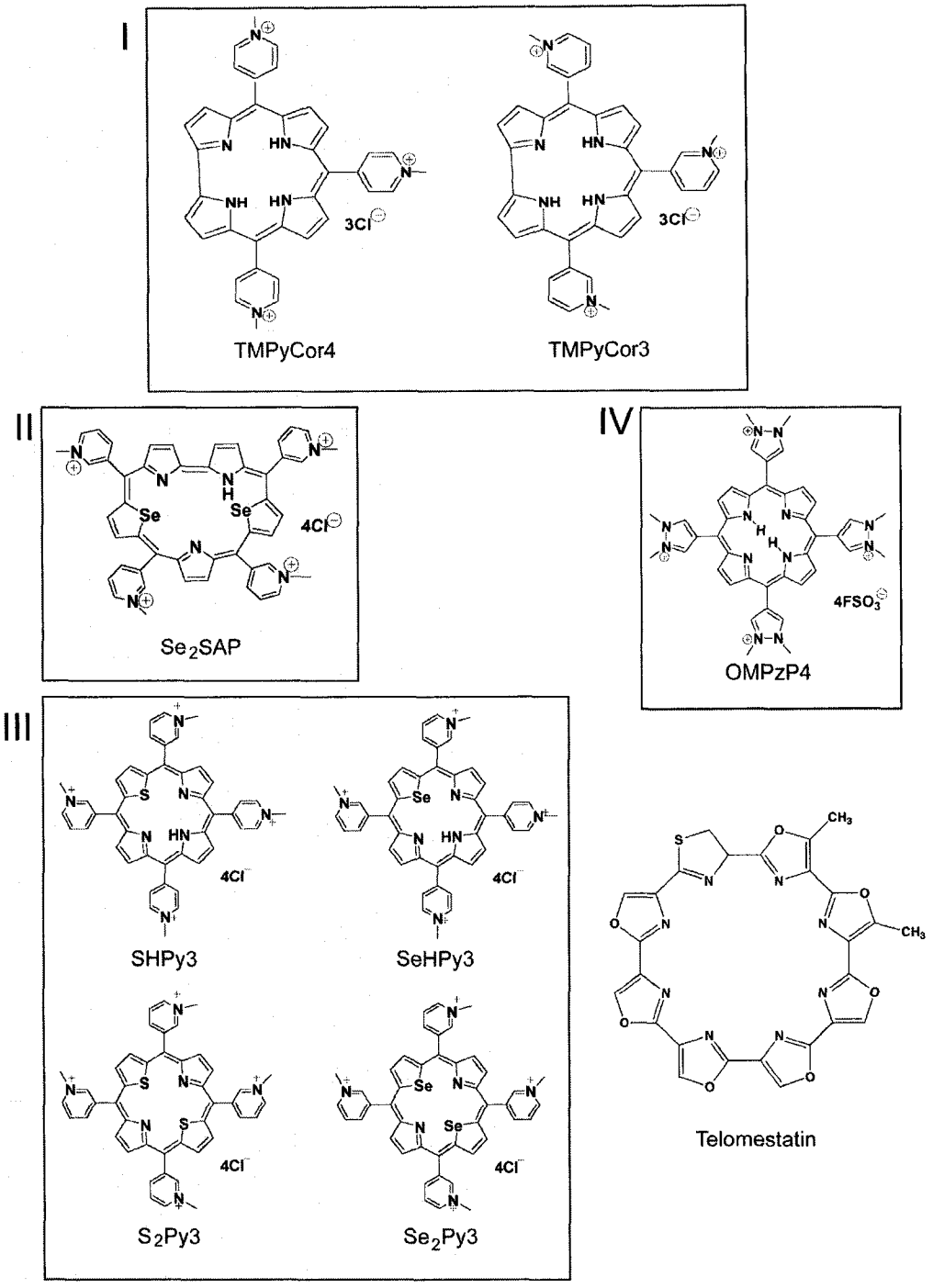


Figure 7.1. Structures of the porphyrin analogues used in this study.
Group I represents the core-contracted corrole compounds, Group II the core-expanded sapphyrin compounds, Group III the core-substituted thia- and selena-porphyrin series, and Group IV the meso substituted porphyrin OMPzP4. The structure of Telomestatin, a drug used as a control in the polymerase stop assay studies, is also included.

II. Results

a. Determination of the G-Quadruplex-Binding Ability of Nine Porphyrin Analogues

The purpose of this study was twofold: First, we wished to add credence to the hypothesis that the quadruplex-binding ability of TMPyP4 is responsible for the downregulation of c-*myc* mentioned in the Introduction to this manuscript. In order to do this, the ability of each new synthesized compound to bind to the NHE III_1 quadruplex needed to be determined, and correlated to c-*myc* downregulation. Second, we wished to synthesize compounds superior to TMPyP4 in terms of activity (reduction of c-*myc* expression) and toxicity (lower to no photoactivity). For this dissertation, focus will be placed on the former of these two purposes, the correlation of quadruplex binding to c-*myc* reduction.

Drs. Seenisamy and Bashyam performed Taq polymerase stop assays, using a number of porphyrin concentrations and the NHE III_1 sequence, in order to determine the ability of each analogue to bind to and stabilize the quadruplex structure and the duplex DNA. Representative charts, showing the relative abilities of the Group III porphyrin Se_2Py3 and the Group II sapphyrin Se_2SAP to stabilize the quadruplex and duplex are shown in Figure 7.2.

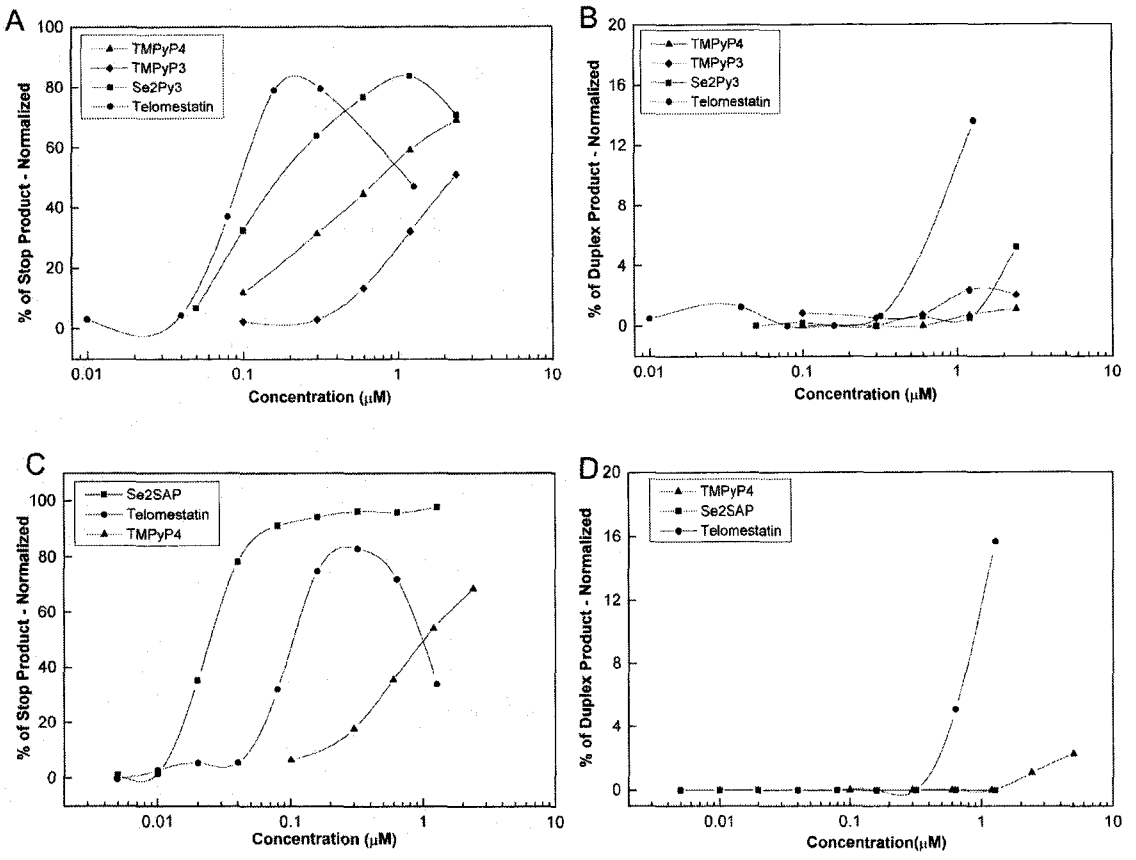


Figure 7.2. Taq polymerase stop assays for NHE III_1 quadruplex stabilization by Se2Py3 and Se_2SAP.

Polymerase stop assays were performed with various concentrations of porphyrin analogs, in order to determine the abilities of these molecules to stabilize the duplex and quadruplex forms of the NHE III_1. Other known compounds (TMPyP4, TMPyP3 and Telomestatin) were also used as positive controls. A. Quadruplex stabilization by Se_2Py3. B. Duplex binding by Se_2Py3. C. Quadruplex stabilization by Se_2SAP. D. Duplex binding by Se_2SAP. Each data point represents the percent of the indicated product compared to the full-length polymerization product.

b. Effect of Core-Contracted (Group I) and Meso-Substituted (Group IV) Analogues on c-*myc* Expression

The first set of compounds tested in this structure-activity relationship study included the core-contracted porphyrin analogues (TMPyCor3 and TMPyCor4) and the meso group-substituted porphyrin (OMPzP4), as these were the first to be synthesized and made available. These compounds were tested at a single concentration, 100 μM, in order to compare their activities to that of TMPyP4 and TMPyP2. The results are shown in Figure 7.3. Neither the core-contracted corroles nor OMPzP4 appeared to be as potent as TMPyP4 at this concentration, though all three did have a repressive effect, albeit small, on c-*myc* expression, with OMPzP4 having the greatest activity among the three (approximately 66% of the untreated control).

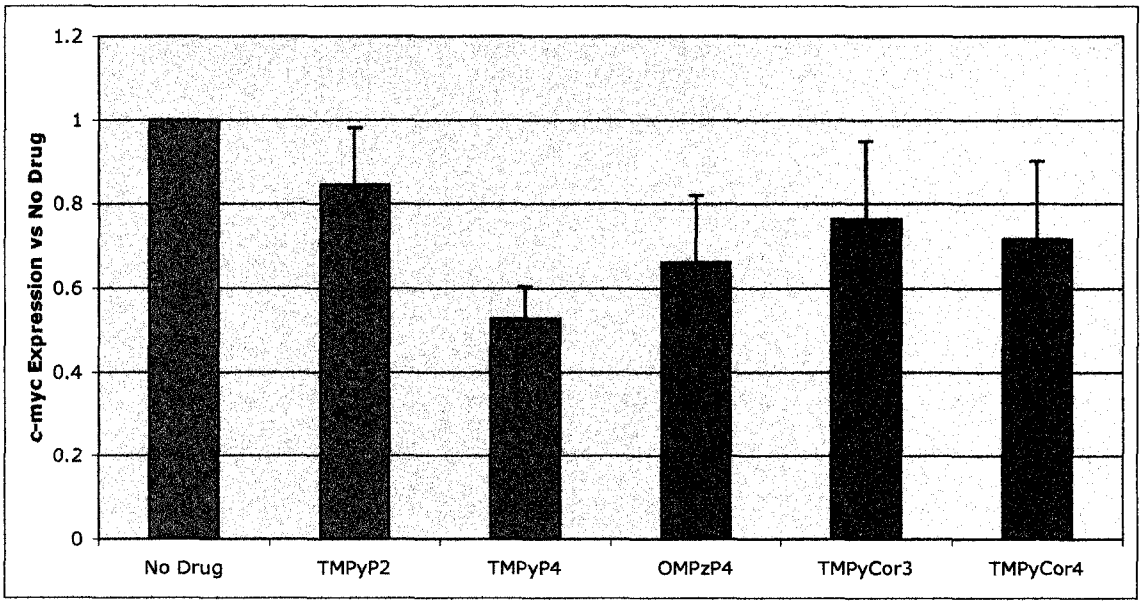


Figure 7.3. Effect of Core-Contracted and Meso-Substituted Porphyrins on c-*myc* Expression in HeLa S_3 cells.

The core-contracted corrole compounds (Group I in Figure 7.1) and the meso-substituted porphyrin OMPzP4 (Group IV in Figure 7.1) were added to HeLa S_3 cells at a concentration of 100 μM for 24 hours. Results are expressed as proportion of c-*myc* expression compared to the No Drug control, measured by real-time RT-PCR.

c. Effect of Core-Modified (Group III) Porphyrins on c-*myc* Expression

The core-modified di-thia and di-selenaporphyrins (SHPy3, S_2Py3, SeHPy3 and Se_2Py3) were examined for their relative abilities to inhibit c-*myc* gene expression. In Figure 7.2, we showed that these four molecules exhibit different activities at different concentrations. At low concentrations, the quadruplex is stabilized, and represents the major stop product in the polymerase stop assay. At higher concentrations, however, this arrest site disappears as the duplex DNA stop site increases in intensity, suggesting that the compounds have the ability to stabilize duplex DNA. We were curious as to the effect this might have on c-*myc* transcription, so HeLa S_3 cells were treated with a range of concentrations, from 5 μM to 100 μM, of these porphyrins. The outcome is shown in Figure 7.4. The two selenaporphyrins appear to downregulate c-*myc* expression at the lowest concentration used (5 μM), with an effect approximately identical to that of TMPyP4 at 100 μM, implying that these two molecules are superior to TMPyP4. However, all four porphyrin analogues cause an increase in c-*myc* expression as concentration increases, up to an apparent maximum of approximately 300% in all four cases. This correlates with the increased tendency of these molecules to stabilize the duplex DNA in the polymerase stop assay in Figure 7.2, suggesting that stabilization of the NHE III_1 in a duplex form causes an increase in c-*myc* transcription. This also implies that these four molecules may not be superior to TMPyP4 in terms of c-*myc* inhibition after all, if the effective concentration at the target cannot be controlled; if the

concentration of these molecules is too high, c-*myc* expression may be activated, obviously counterproductive to our cause.

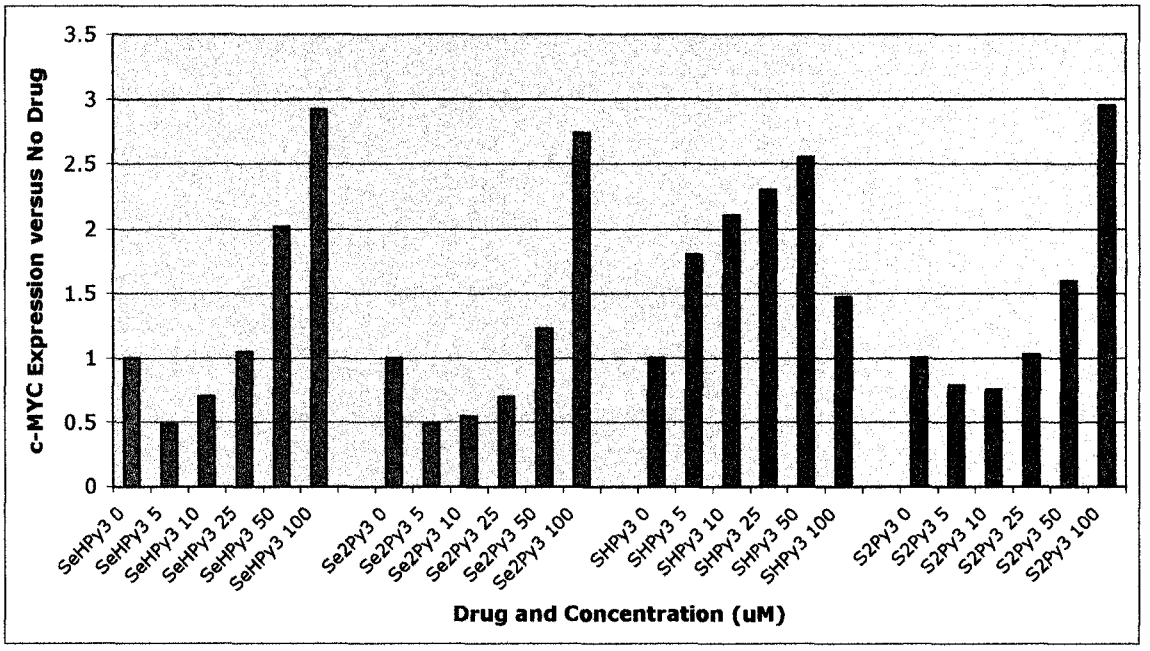


Figure 7.4 Effect of core-substituted porphyrins on c-*myc* expression in HeLa S_3 cells.

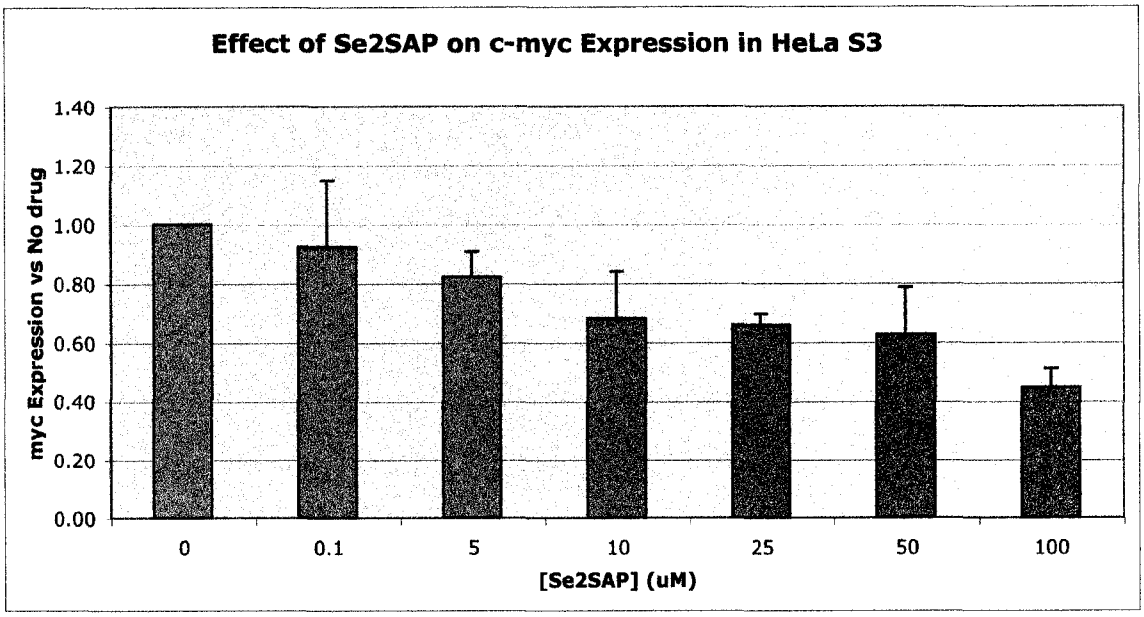
HeLa S_3 cells were treated with 5, 10, 25, 50 or 100 μM of each of the core-substituted porphyrins for 24 hours, and real-time RT-PCR was performed. Results are shown as proportion of c-*myc* expression relative to the No Drug (0 μM) control.

d. Effect of Se_2SAP (Group II) on c-*myc* Expression

In the hopes of finding a compound that lacked the photoactivity of TMPyP4, and the duplex-binding activity of the Group III porphyrins, Se_2SAP was tested for its ability to downregulate c-*myc* production. Notice from Figure 7.2 that this molecule has a very strong quadruplex-stabilizing property that only increases with increased dose, unlike the Group III core-modified porphyrins discussed above. Thus, we expected that Se_2SAP would only reduce c-*myc* expression, no matter the effective dose at the target. Also, as it appears that Se_2SAP has a much stronger quadruplex-stabilizing potential than the Group III porphyrins at equal doses, we expect that Se_2SAP will be much more potent than these other molecules in reducing c-*myc* activity in cell culture. HeLa S_3 cells were treated with a range of doses of Se_2SAP for 24 hours, and c-*myc* expression was measured. These data are found in Figure 7.5. As expected, as the dose of Se_2SAP was increased, the inhibitory effect of this molecule on c-*myc* increased as well. What was surprising, however, was the magnitude of this effect. Even at high (100 μM) concentration, Se_2SAP was scarcely as effective as TMPyP4, despite being more potent in the polymerase stop assay. This suggests that Se_2SAP may not be getting to the DNA target as well as TMPyP4 does.

This does not come as a surprise, as sapphyrins are not biologically relevant molecules in human cells, while porphyrins, such as heme, are quite common, and may have a specific transport mechanism for uptake. Therefore, we might expect that Se_2SAP is simply not being allowed to enter cells as readily as TMPyP4. In order to verify this, a time course was performed with Se_2SAP, and c-*myc* expression quantified. The results of

this experiment are presented in Figure 7.6. Over time, we see that the inhibitory effect of Se_2SAP increases, from nearly negligible after 24 hours to an apparent maximum of 55% after 120 hours. This effect appears to plateau out at this point, suggesting that this is the greatest effect we can expect from this molecule. This also implies that Se_2SAP is not taken up readily by cells, as we hypothesized, but is able to diffuse in slowly over time. However, its maximum effect is greater than that of TMPyP4, which, when considered with its lack of phototoxicity, makes Se_2SAP a valuable therapeutic lead compound. To increase the potency of this sapphyrin, however, a means must be found to increase its uptake into human cells.



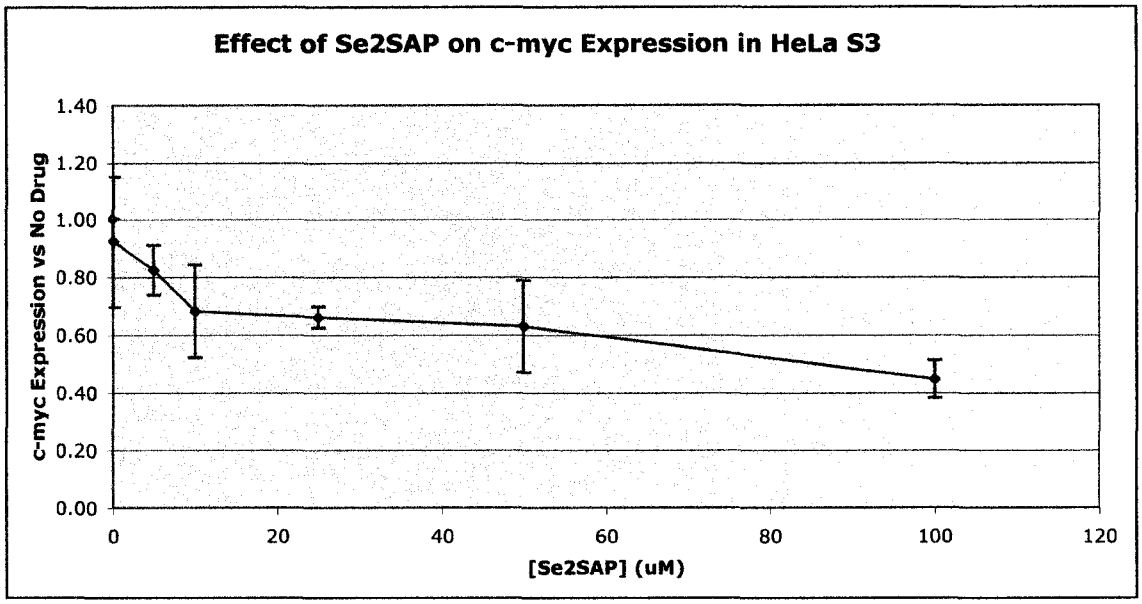
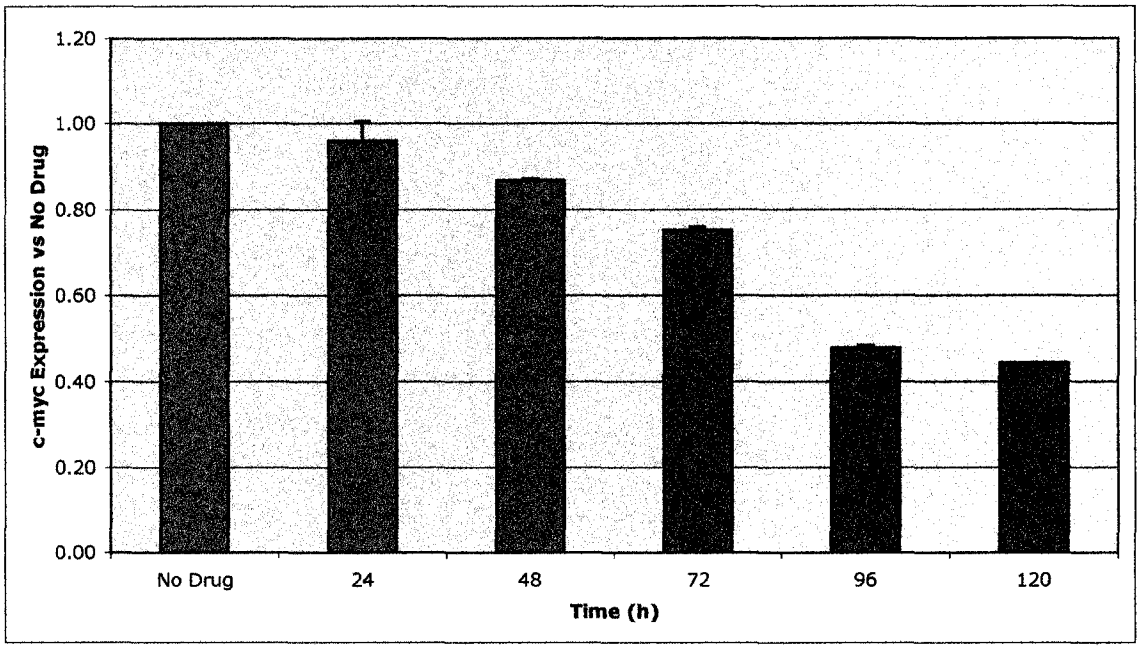


Figure 7.5. Dose-responsive effect of Se_2SAP on c-*myc* expression in HeLa S_3 cells.

Cells were treated with the listed concentrations of Se_2SAP for 24 hours, and c-*myc* expression measured by real-time RT-PCR. Results are presented as proportion of c-*myc* activity in the untreated (0 μM) sample.



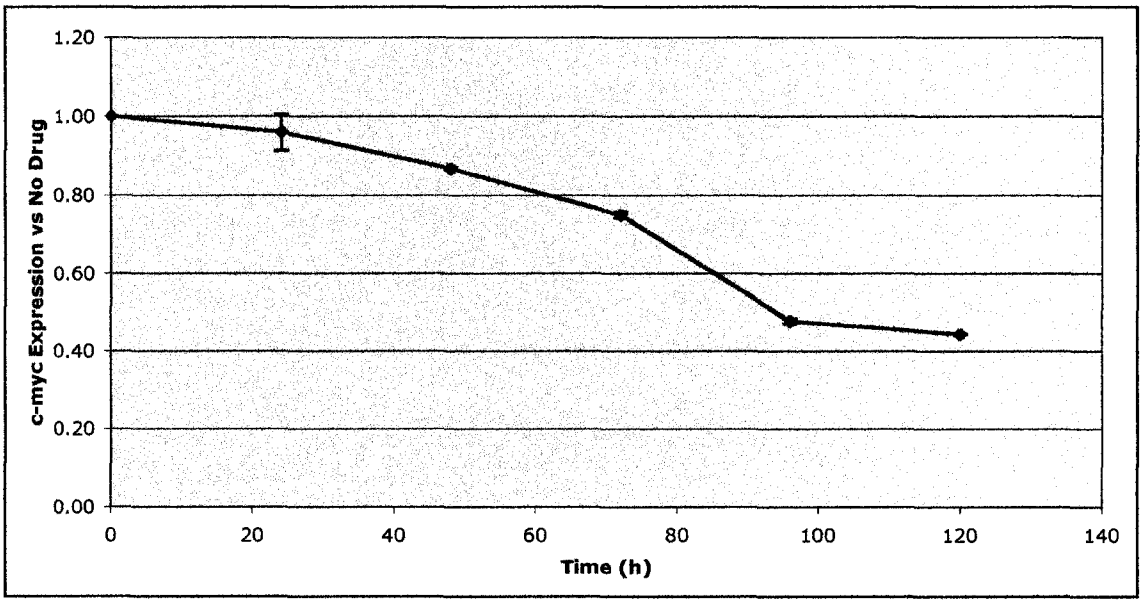


Figure 7.6. Time-dependent effect of Se_2SAP on c-*myc* expression in HeLa S_3 cells.

Cells were treated with 10 μM Se_2SAP for the times indicated, and c-*myc* expression was measured by real-time RT-PCR. Results are presented as proportion of c-*myc* activity in the untreated (No Drug) sample.

III. Discussion

We began this structure activity relationship study with two goals in mind: First, to develop a compound that can reliably reduce c-*myc* levels in tumor cells without high toxicity, and second, to ascertain that the effect of the G-quadruplex-interactive compound TMPyP4 on c-*myc* expression is indeed due to G-quadruplex interaction. In order to accomplish this, derivatives of TMPyP4 were synthesized and examined for their ability to bind to and stabilize the G-quadruplex in the c-*myc* NHE III_1, their phototoxicity, and their potency in the reduction of c-*myc*. In the process, a number of interesting observations were made. First, we found that derivatives of TMPyP4, which all interact with the G-quadruplex, are indeed able to reduce c-*myc* in cell culture, suggesting that the effects of TMPyP4 are not confined to this molecule. Second, compounds that have the ability to stabilize duplex DNA (or destabilize quadruplex DNA) in the c-*myc* promoter cause an increase in c-*myc* expression, consistent with what we have seen in the past with quadruplex-specific mutations. Third, a compound Se_2SAP, which is superior to TMPyP4 in its affinity for the specific chair-type G-quadruplex in the c-*myc* promoter is also superior to TMPyP4 in c-*myc* downregulation. However, this activity depends largely on the ability of the compound to enter cells; the maximum effect of Se_2SAP is seen only after a long incubation period. As mentioned earlier, this is likely due to the biological relevance of our G-quadruplex-interactive drugs to cells: The porphyrin core is used in biomolecules such as heme, and so it is no surprise that uptake of a porphyrin analogue occurs with high efficiency. This is seen in pellets of TMPyP2-

and TMPyP4-treated cells, both of which exhibit the color of the porphyrin molecule. The molecule synthesized in this study, Se_2SAP, however, has a sapphyrin core, which is not biologically relevant, a likely reason for its slow uptake into cells. Quite probably, TMPyP4 is entering cells through a specific transporter process, while Se_2SAP is forced to enter cells by passive diffusion.

The task that stands before us now in the development of Se_2SAP, then, is not to increase its affinity for the chair-type G-quadruplex, but instead to find a way to increase cellular uptake. To this end, a variety of means are available. Modifications of Se_2SAP itself that make it more likely to enter cells (such as increased lipid-solubility, for example) may increase its efficacy in terms of c-*myc* reduction. Novel means of drug delivery, e.g. cationic lipid encapsulation, might also be worthy of investigation.

Perhaps the most important piece of information that can be taken from this structure-activity relationship is the fact that we have developed a novel means of controlling gene activity: modulation of a hitherto unknown and thus unexploited secondary DNA structure, specific to that gene's promoter. The sapphyrin analogue Se_2SAP has been shown to bind preferentially to the chair-type quadruplex formed in the c-*myc* promoter over other proposed quadruplexes, such as the human telomeric basket-type quadruplex, a preference not exhibited by our original lead, TMPyP4. In this way, we believe that we have come upon a new frontier in therapeutic drug design.

CHAPTER VIII

CONCLUSIONS

With the research described herein, we have found compelling evidence for the existence of a regulatory structure heretofore unexplored: the G-quadruplex. We have shown that a G-quadruplex-interactive compound, TMPyP4, can reduce the levels of the proto-oncogene c-*myc,* the promoter region of which bears a sequence amenable to the formation of a chair-type G-quadruplex *in vitro.* Further study has shown that TMPyP4, but not its positional isomer TMPyP2, interacts very specifically with the chair-type quadruplex, which is in accordance with the effects of each of these molecules on c-*myc* expression. When as few as 850 base pairs of the c-*myc* promoter, including the NHE III_1, the region proposed to form the G-quadruplex, are taken out of the context of the rest of the c-*myc* gene and placed in control of a foreign gene, the effect of these porphyrin molecules is maintained. Similarly, this effect remains when the c-*myc* gene is translocated to a different chromosome, under control of different regulatory elements, as in Burkitt's Lymphoma, but only if the NHE III_1 is included in the translocation, again cementing the importance of this region for the effect of the quadruplex-interactive agent TMPyP4. This initially provides very strong evidence that a G-quadruplex exists in the c-*myc* promoter.

In order to make this evidence yet more convincing, the effect of disrupting this secondary structure was evaluated. When mutations are made to the c-*myc* promoter, such that the chair-type quadruplex is destabilized, a significant increase in

transcriptional activity, on the order of 1.75- to 3.5-fold, is witnessed, along with a marked reduction, though not complete abrogation, of the ability of TMPyP4 to reduce it. This suggests that the quadruplex acts as a control measure for c-*myc*, keeping expression down until needed. Once the quadruplex is resolved, by mutation or through a purposeful cellular mechanism, c-*myc* levels can rise. It is proposed that this resolution may be afforded by the glycosylase/lyase/nucleoside diphosphate kinase NM23-H2. Thus, the chair-type G-quadruplex in the NHE III_1 of the c-*myc* promoter is hypothesized to act as a transcriptional repressor structure.

If this hypothesis holds true, then one might expect that mutations *in vivo* that disrupt the chair-type G-quadruplex would be selected for in developing tumor cells, given that c-*myc* overexpression is found to play a role in tumorigenesis. Such is indeed the case; 6 of 21 colorectal tumors examined bear mutations in the c-*myc* promoter that destabilize the quadruplex. This is expected to be a late event in tumorigenesis, as the c-MYC protein can also induce apoptosis if expressed in the absence of the appropriate growth stimuli; only after these apoptotic pathways have been subjugated can c-*myc* be overexpressed with impunity. Evidence for this is found: no quadruplex-destabilizing mutations were observed in adenomatous polyps, an early lesion that precedes the development of colorectal cancer.

The mechanism through which the quadruplex exerts its effects is as yet poorly understood. It appears that it acts as a barrier to the binding of one or more protein transcription factors, and that one of these factors is Sp1 or an Sp1-like protein. When Sp1 is removed from the system, mutating the quadruplex no longer has the effect of

increasing c-*myc* promoter activity. This is remedied by supplementation with exogenous Sp1. These data also suggest, however, that Sp1 is able to bind to the purine-rich strand of the NHE III_1 alone, a characteristic unexpected for a known double-strand DNA-binding factor. Thus, more study of this system, and of Sp1, is needed.

Finally, the requirement for a G-quadruplex for the effect of TMPyP4 is further realized upon attempting to synthesize superior compounds. Other quadruplex-interactive compounds are also able to reduce c-*myc* expression, with differing efficacy, dependent both on quadruplex stabilization and cellular uptake. It is also determined that drugs which stabilize duplex DNA, or inhibit quadruplex formation, can cause c-*myc* expression to rise. These data again suggest that the chair-type quadruplex acts as a governor of the c-*myc* gene; when stabilized, this repression is enhanced, and when disrupted or destabilized, it is lifted. Thus, we show that a chair-type G-quadruplex exists in the c-*myc* NHE III_1, it is biologically relevant, and represents a novel and viable target for anticancer drug design.

From the data presented in this dissertation, we have posited a model for the role of a chair-type G-quadruplex in the NHE III_1 of the c-*myc* promoter, which is shown in Figure 8.1A. Under normal conditions in non-transformed cells, the c-*myc* promoter is normally in an "off" state. In which the G-quadruplex is present on the purine-rich strand. The state of the pyrimidine-rich strand was not studied, and is thus yet unknown, though under examination by other members of the laboratory. In this state, transcription factors are not bound to the NHE III_1 and transcription initiation cannot proceed. Under the action of NM23-H2, acting as a DNA remodeling factor, this secondary structure can be

resolved to the separate single strands, to which transcription factors, such as CNBP on the purine-rich strand and hnRNP K on the pyrimidine-rich strand can bind. This places the promoter in an "on" state, and transcription proceeds. The transition from quadruplex to separate single strands may be mediated through a hairpin or other double-stranded intermediate, to which Sp1 can bind. This intermediate might also explain the ability of Se_2Py3 and its Group III analogues (see Chapter VII) to increase c-*myc* expression, by stabilizing this intermediate and inhibiting the transition forward to the quadruplex state. The forward reaction, from single strands to quadruplex, is believed to occur without the involvement of proteins, as the quadruplex is the kinetically favored state of this sequence and likely forms readily under the influence of the negative superhelical stress present normally throughout the genome. It is only the reverse reaction, back to single strands, that requires energy and the assistance of protein factors. When a quadruplex-stabilizing agent such as TMPyP4 is introduced, the equilibrium is driven even further to the right, holding the promoter in an "off" state, and making the back reaction more energetically difficult. Also, as the binding mode of TMPyP4 places it in direct contact with the major cleavage site of NM23-H2, it may inhibit the activity of this enzyme, another means by which the promoter is kept inactive.

In the process of tumorigenesis, one of the mechanisms that leads to upregulation of c-*myc* may involve mutation of the c-*myc* NHE III_1 (Figure 8.1B). In this scenario, the chair-type quadruplex is destabilized, such that the equilibrium is driven instead to the left, with the quasi-stable G-quadruplex present only a small fraction of the time. This results in a higher level of basal transcription from the c-*myc* promoter, increased c-MYC

protein, and the increased level of proliferation associated with this upregulation, contributing to the transformed phenotype. This assumes that the apoptosis-inducing effects of c-*myc* upregulation have been abrogated somehow, through loss or disruption of the apoptotic effector mechanisms, e.g. p53 mutation. This predicts, then, that G-quadruplex destabilizing mutations in c-*myc* would be a late event in tumorigenesis, which our studies of adenomatous polyps suggest. When a G-quadruplex-stabilizing agent such as TMPyP4 is introduced, the quasi-stable quadruplex is further stabilized, and thus a downregulation of c-*myc* can be achieved, even in the face of these mutations. This effect is not as pronounced as that of TMPyP4 on wild-type c-*myc*, likely because the single strand-quadruplex equilibrium is shifted to the left, resulting in less substrate for TMPyP4 to bind to overall. Therefore, if G-quadruplex-interactive compounds are to be used in a therapeutic setting to downregulate c-*myc*, a knowledge of the NHE III_1 sequence in the patient would be useful in determining if such a treatment is worthwhile.

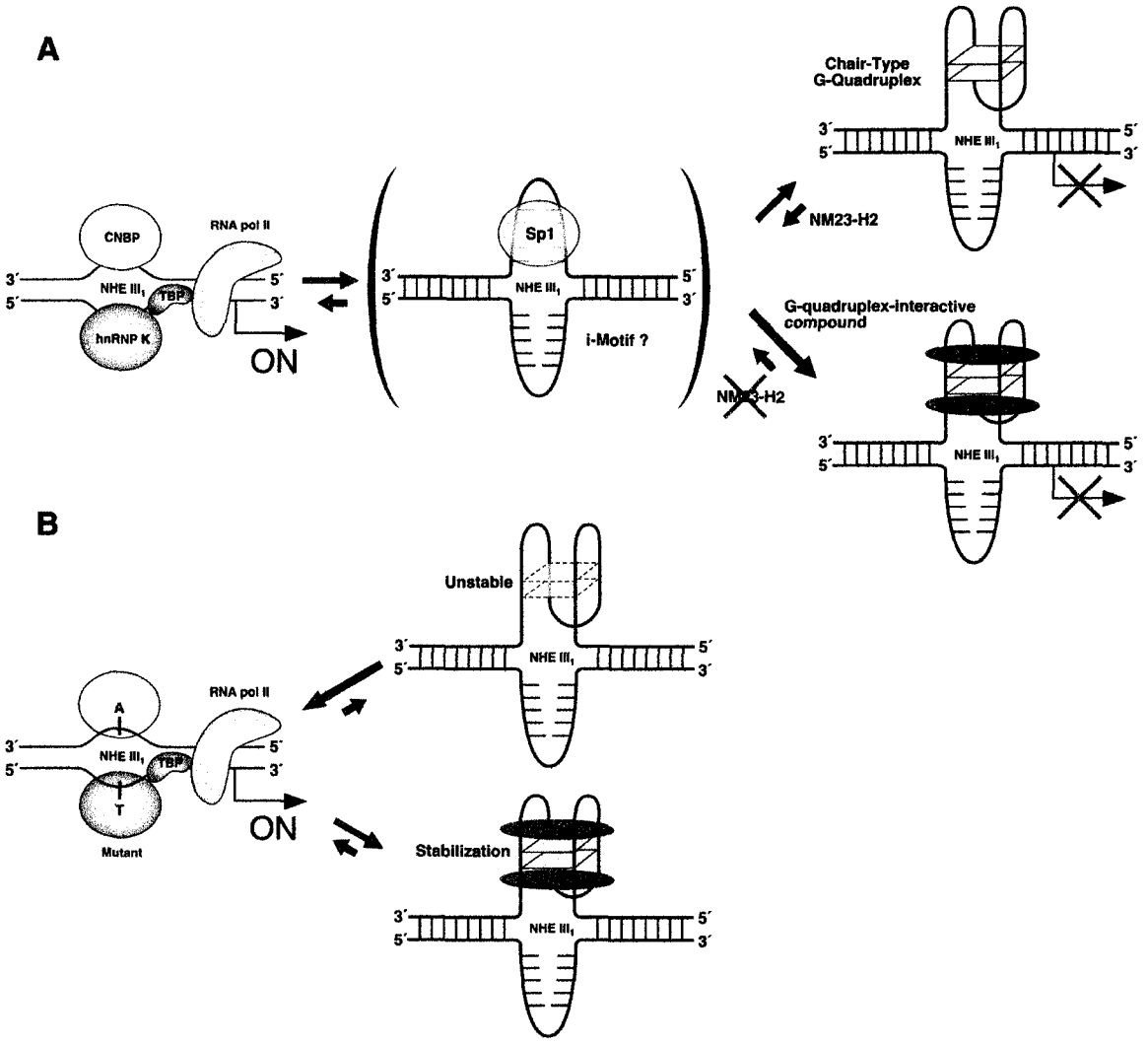


Figure 8.1. Model for the role of the c-*myc* chair-type G-quadruplex structure (see text).

A. The normal state of the c-*myc* promoter, with and without quadruplex stabilization. B. The effect of quadruplex destabilization by mutation. Arrow size (length and thickness) indicate intensity of the reaction and direction of equilibrium.

APPENDIX A

ABSTRACTS AND MANUSCRIPTS

PUBLICATIONS

Cory L. Grand, Haiyong Han, Rubén M. Muñoz, Steven Weitman, Daniel D. Von Hoff, Laurence H. Hurley, David J. Bearss. The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth *in vivo*. Molecular Cancer Therapeutics 1(8):565-73. Jun 2002.

Adam Siddiqui-Jain, **Cory L. Grand**, David J. Bearss, Laurence H. Hurley. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. Proc Natl Acad Sci U S A. 3;99(18):11593-8 Sep 2002.

Cory L. Grand, Tiffanie J. Powell, Raymond B. Nagle, David J. Bearss, Laurence H. Hurley. Identification of a Novel c-*MYC* Mutational Mechanism for Colorectal Tumorigenesis Involving Destabilization of the G-Quadruplex Silencer Element. Nature (submitted).

Jeyaprakeshnarayanan Seenisamy, Sridevi Bashyam, Hariprasad Vankayalapati, Daekyu Sun, **Cory L. Grand**, Laurence H. Hurley. Design and Synthesis of an Expanded Porphyrin that Has Selectivity for the c-*MYC* Chair-Type over the Human Telomeric Basket-Type Intramolecular G-Quadruplex Structure. J Amer Chem Soc (submitted).

Specific inhibition of c-MYC expression by the cationic porphyrin TMPyP4 results in downregulattion of hTERT expression and reduced telomerase activity.

By **Cory L. Grand**, David J. Bearss Ph.D., Rubén M. Muñoz M.S., Haiyong Han Ph.D., Daniel D. Von Hoff M.D., Laurence H. Hurley Ph.D. Proceedings of the American Association for Cancer Research (Abstract #1123) Vol. 42, March 2001. American Association for Cancer Research, New Orleans, LA.

Telomerase activity is associated with escape from cellular senescence and can confer the immortality implicated in the process of malignant transformation of cells. We have shown previously that the cationic porphyrin, TMPyP4, stabilizes G-quadruplexes in telomeric DNA and hampers the telomerase enzyme from interacting with its substrate. In addition, in vitro treatment of cells with this porphyrin results in lower levels of telomerase activity. To better understand the mechanisms by which TMPyP4 affects telomerase activity, changes in gene expression following treatment with TMPyP4 were evaluated using cDNA microarray analysis, and it was found that treatment with this small molecule resulted in the downregulation of the c-MYC gene and several c-MYC target genes, including the telomerase catalytic subunit hTERT. This effect was not seen with a related porphyrin, TMPyP2, which is unable to interact with G-quadruplex DNA as well as TMPyP4, due to steric hindrances. Interestingly, the c-MYC gene promoter region contains a possible G-quadruplex forming region, which may mediate the effects of TMPyP4 on c-MYC gene expression. To verify this effect, the effect of TMPyP4 on c-MYC and hTERT expression, telomerase activity, and cell cycle dynamics, was determined; HCT-116 and MiaPaCa cells were grown in 25cm^2 culture flasks to 60% confluency, and treated with 0.2μg/mL doxorubicin, 100μM TMPyP4, 100nM taxotere, 10nM gemcitabine, or 100μL H_2O (control) for a time course of 12, 24 and 48 hours. Cell pellets were collected and total RNA and total protein were extracted. RT-PCR for c-MYC and hTERT was performed using 2μL of total RNA. TRAP assays were run with 1.5μg of total protein, using the TRAPeze assay kit (Intergen). 2 X 10^5 cells were also removed, suspended in 0.5mg/mL propidium iodide in Krishan's buffer, and sent for cell cycle analysis by FACS. We found that c-MYC and hTERT expression were both reduced in a time-dependent fashion by TMPyP4, as was telomerase enzymatic activity. TMPyP4 did not, however, have an appreciable effect on the cell cycle, indicating that TMPyP4 does not mediate its effects as a general cell poison. (Supported by NCDDG grant CA67760 from the NCI DHHS)

The Cationic Porphyrin TMPyP4 Inhibits *c-Myc* Transcription, Probably Through a G-Quadruplex-Dependent Mechanism

By **Cory L. Grand**, Adam Siddiqui Ph.D., David J. Bearss Ph.D., Rubén M. Muñoz M.S., Haiyong Han Ph.D., Daniel D. Von Hoff M.D., Laurence H. Hurley Ph.D.
Proceedings of the 2001 AACR-NCI-EORTC International Conference; Molecular Targets and Cancer Therapeutics: Discovery, Biology, and Clinical Applications (Abstract #274, Oral Presentation). AACR-NCI-EORTC, Miami Beach, FL.

As the emphasis on developing drugs that target specific tumor-related molecules grows, a push has come to discover more viable marks to aim for. It is now known that, in addition to tumor-associated protein products, structures in DNA itself can be useful targets for novel chemotherapeutic agents. In the past few years, one such structure, the G-quadruplex, has emerged as a potential target in tumor cells. There is an increasing body of evidence that suggests that G-quadruplexes may have distinct and important roles in regulation of genes and processes involved in cellular proliferation and growth. These structures have been found in telomeric DNA sequences, and in the promoters of many proto-oncogenes, including *c-abl, c-jun, c-fos, c-src, c-myb, c-sis* and *c-myc*. It is our hypothesis that, by creating small molecules that can interact with or stabilize G-quadruplexes, we can alter the expression of these oncogenes, most significantly *c-myc*. The *c-myc* proto-oncogene has been found to be inordinately upregulated in a variety of human malignancies, making it an ideal target for chemotherapeutic agents. A G-quadruplex-interactive compound, TMPyP4, has been shown in this laboratory to inhibit the expression of *c-myc* in a time- and dose-dependent manner, and so we have pursued the possibility that targeting the G-quadruplex in the *c-myc* promoter will provide a meaningful way to reverse the oncogenic properties of this gene. We have reporter constructs that bear a luciferase gene under the control of a portion of the *c-myc* promoter that contains the putative G-quadruplex-forming region, and we have shown that these too show downregulation in response to TMPyP4. Interestingly, when these constructs are mutated, such that they can no longer form a G-quadruplex structure, the effect of TMPyP4 is no longer seen. Also, destruction of the G-quadruplex sequence abrogates expression from our reporter constructs 50-fold, even without drug treatment, suggesting that the formation of a G-quadruplex is important in *c*-myc gene control. Regulation of *c-myc* by this region of the promoter is discussed, and implications of G-quadruplexes as targets are brought forth.

Transcriptional Inhibition of c-MYC by a G-Quadruplex and a Compound, TMPyP4, That Stabilizes Its Structure

By **Cory L. Grand**, Adam Siddiqui-Jain Ph.D., David J. Bearss Ph.D., Laurence H. Hurley Ph.D. Proceedings of the American Association for Cancer Research (Abstract #LB-110) Vol. 43, April 2002. American Association for Cancer Research, San Francisco, CA.

The NHE III upstream of the P1 promoter of c-MYC controls up to 85-95% of the transcriptional activation of this gene. We have demonstrated that the DNA in this region can form two different intramolecular G-quadruplex structures, only one of which appears to be biologically relevant. The biologically relevant structure is the kinetically favored chair form G-quadruplex, which when mutated with a single G to A transition is destabilized, resulting in a 3-fold increase in basal transcriptional activity of the c-MYC promoter. A cationic porphyrin, TMPyP4, that is shown to stabilize this G-quadruplex structure is able to further suppress c-MYC transcriptional activation. These results provide evidence that a specific G-quadruplex structure formed in the c-MYC promoter region functions as a transcriptional repressor element. Furthermore, we establish the principle that c-MYC transcription can be controlled by ligand-mediated G-quadruplex stabilization.

Quadruplex Formation in the c-*MYC* Promoter Inhibits Protein Binding and Correlates with *in vivo* Promoter Activity

By **Cory L. Grand**, David J. Bearss Ph.D., Laurence H. Hurley Ph.D. European Journal of Cancer, Vol. 38, Supplement 7 (Abstract #358) November 2002. AACR-NCI-EORTC, Frankfurt, Germany.

Previously, we have determined that a G-quadruplex interactive compound, the cationic porphyrin TMPyP4, can cause down-regulation of the c-*MYC* proto-oncogene in tumor cell lines. Subsequently, we have found that a region of the c-*MYC* promoter, termed the NHE III_1, is able to form two different intramolecular G-quadruplex structures *in vitro*; a chair- and basket-type. Through site-directed mutagenesis of the c-*MYC* promoter, we have provided evidence that the chair-type quadruplex is a biologically relevant structure *in vivo*. Here, we show that a specific protein, identity to be determined, is able to bind to a 27-base long oligomer corresponding to the NHE III_1 only if the oligomer cannot form a G-quadruplex; when the oligomer is mutated such that the chair-type quadruplex can no longer form, the ability of this protein to bind is increased several fold. However, mutations to the basket-type quadruplex, or of bases not involved in formation of either quadruplex, have no effect on binding. This exactly correlates with c-*MYC* promoter activity from constructs bearing these same mutations. We hypothesize that the ability of TMPyP4 to stabilize the chair-type quadruplex in the NHE III_1 prohibits normal binding of this protein, and results in the decrease in c-*MYC* expression we have seen previously. A model is proposed, which explains how quadruplex formation occurs normally in the c-MYC promoter and regulates expression through the relative ability/inability of this protein to bind to the DNA.

Identification of a G to A mutation in the c-*MYC* repressor element that results in inactivation of the G-quadruplex repressor element and overexpression of c-*MYC* in colorectal cancer

By **Cory L. Grand**, Tiffanie J. Powell, Raymond B. Nagle M.D., David J. Bearss Ph.D., Laurence H. Hurley Ph.D. Proceedings of the American Association for Cancer Research (Abstract #LB-224) Vol. 44, March 2003, 2nd Edition. American Association for Cancer Research, Toronto, ON. (Later rescheduled to July 2003, Washington, DC.)

In a recent publication (Siddiqui-Jain et al., *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 11593–11598, 2002), we have demonstrated that a specific G-quadruplex structure formed in the c-*MYC* promoter region functions as a transcriptional repressor element. Furthermore, a specific G to A mutation in the element results in destabilization of this G-quadruplex repressor element and an increase in basal transcriptional activity. In order to validate this model in an *in vivo* context, we have examined the sequence of this promoter of c-*MYC* in human colon tumor samples and in the surrounding normal tissue. Tumor and nontransformed cells were collected by laser capture microdissection, and the genomic DNA was extracted for sequencing. We have found that X of 20 tumors contain mutations, not present in the corresponding surrounding normal tissue, that disrupt the chair-type quadruplex, which would give rise to the abnormally high expression of c-*MYC* in these cells. In addition to the significance of this finding for the etiology of colorectal cancer, these mutations have two further important implications. First, this further establishes the role of G-quadruplex structures as a novel mechanism for silencing of oncogenes, and second, this provides an appealing approach, i.e., restabilization of the G-quadruplex repressor element by drugs, for treating cancers with genetic defects.

APPENDIX B

PERMISSION TO USE COPYRIGHTED MATERIAL

Reproduced with permission, from:

Cory L. Grand, Haiyong Han, Rubén M. Muñoz, Steven Weitman, Daniel D. Von Hoff, Laurence H. Hurley, David J. Bearss, 2002, Molecular Cancer Therapeutics, **1**(8):565-73.

AND

Adam Siddiqui-Jain, Cory L. Grand, David J. Bearss, Laurence H. Hurley, 2002, Proceedings of the National Academy of Sciences of the United States of America, **99**(18):11593-8.

REFERENCES

1. Vennstrom, B. and Bishop, J. M. Isolation and characterization of chicken DNA homologous to the two putative oncogenes of avian erythroblastosis virus. Cell, *28:* 135-143, 1982.

2. Vennstrom, B., Sheiness, D., Zabielski, J., and Bishop, J. M. Isolation and characterization of c-myc, a cellular homolog of the oncogene (v-myc) of avian myelocytomatosis virus strain 29. J Virol, *42:* 773-779, 1982.

3. Collins, S. and Groudine, M. Amplification of endogenous myc-related DNA sequences in a human myeloid leukaemia cell line. Nature, *298:* 679-681, 1982.

4. Takahashi, E., Hori, T., O'Connell, P., Leppert, M., and White, R. Mapping of the MYC gene to band 8q24.12----q24.13 by R-banding and distal to fra(8)(q24.11), FRA8E, by fluorescence in situ hybridization. Cytogenet Cell Genet, *57:* 109-111, 1991.

5. Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S., and Leder, P. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proceedings of the National Academy of Sciences of the United States of America, *79:* 7837-7841, 1982.

6. Dalla-Favera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R. C., and Croce, C. M. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proceedings of the National Academy of Sciences of the United States of America, *79:* 7824-7827, 1982.

7. Spencer, C. A. and Groudine, M. Control of c-myc regulation in normal and neoplastic cells. Advances in Cancer Research, *56:* 1-48, 1991.

8. Magrath, I. The pathogenesis of Burkitt's lymphoma. Advances in Cancer Research, *55:* 133-270, 1990.

9. Marcu, K. B., Bossone, S. A., and Patel, A. J. myc function and regulation. Annual Review of Biochemistry, *61:* 809-860, 1992.

10. Oster, S. K., Ho, C. S., Soucie, E. L., and Penn, L. Z. The myc oncogene: MarvelouslY Complex. Adv Cancer Res, *84:* 81-154, 2002.

11. Levens, D. Disentangling the MYC web. Proc Natl Acad Sci U S A, *99:* 5757-5759, 2002.

12. Schuhmacher, M., Kohlhuber, F., Holzel, M., Kaiser, C., Burtscher, H., Jarsch, M., Bornkamm, G. W., Laux, G., Polack, A., Weidle, U. H., and Eick, D. The transcriptional program of a human B cell line in response to Myc. Nucleic Acids Res, *29:* 397-406, 2001.

13. Boyd, K. E. and Farnham, P. J. Identification of target genes of oncogenic transcription factors. Proc Soc Exp Biol Med, *222:* 9-28, 1999.

14. Coller, H. A., Grandori, C., Tamayo, P., Colbert, T., Lander, E. S., Eisenman, R. N., and Golub, T. R. Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. Proceedings of the National Academy of Sciences of the United States of America, *97:* 3260-3265, 2000.

15. Hermeking, H., Rago, C., Schuhmacher, M., Li, Q., Barrett, J. F., Obaya, A. J., O'Connell, B. C., Mateyak, M. K., Tam, W., Kohlhuber, F., Dang, C. V., Sedivy, J. M., Eick, D., Vogelstein, B., and Kinzler, K. W. Identification of CDK4 as a target of c-MYC. Proc Natl Acad Sci U S A, *97:* 2229-2234, 2000.

16. O'Connell, B. C., Cheung, A. F., Simkevich, C. P., Tam, W., Ren, X., Mateyak, M. K., and Sedivy, J. M. A large scale genetic analysis of c-Myc-regulated gene expression patterns. J Biol Chem, 2003.

17. Partlin, M. M., Homer, E., Robinson, H., McCormick, C. J., Crouch, D. H., Durant, S. T., Matheson, E. C., Hall, A. G., Gillespie, D. A., and Brown, R. Interactions of the DNA mismatch repair proteins MLH1 and MSH2 with c-MYC and MAX. Oncogene, *22:* 819-825, 2003.

18. Auvinen, M., Jarvinen, K., Hotti, A., Okkeri, J., Laitinen, J., Janne, O. A., Coffino, P., Bergman, M., Andersson, L. C., Alitalo, K., and Holtta, E. Transcriptional regulation of the ornithine decarboxylase gene by c-Myc/Max/Mad network and retinoblastoma protein interacting with c-Myc. Int J Biochem Cell Biol, *35:* 496-521, 2003.

19. OMIM v-myc avian myelocytomatosis viral oncogene homolog; myc. Vol. 2003: OMIM, 2000.

20. Pelengaris, S., Rudolph, B., and Littlewood, T. Action of Myc in vivo - proliferation and apoptosis. Current Opinion in Genetics & Development, *10:* 100-105, 2000.

21. Evan, G. and Littlewood, T. A matter of life and cell death. Science, *281:* 1317-1322, 1998.

22. Evan, G. I., Wyllie, A. H., Gilbert, C. S., Littlewood, T. D., Land, H., Brooks, M., Waters, C. M., Penn, L. Z., and Hancock, D. C. Induction of apoptosis in fibroblasts by c-myc protein. Cell, *69:* 119-128, 1992.

23. Fulda, S., Lutz, W., Schwab, M., and Debatin, K. M. MycN sensitizes neuroblastoma cells for drug-induced apoptosis. Oncogene, *18:* 1479-1486, 1999.

24. Juin, P., Hueber, A., Littlewood, T. D., and Evan, G. c-Myc-induced sensitization to apoptosis is mediated through cytochrome c release. Genes Dev, *13:* 1367-1381, 2000.

25. Bissonnette, R. P., Echeverri, F., Mahboubi, A., and Green, D. R. Apoptotic cell death induced by c-myc is inhibited by bcl-2. Nature, *359:* 552-554, 1992.

26. Virbasius, C. A., Virbasius, J. V., and Scarpulla, R. C. NRF-1, an activator involved in nuclear-mitochondrial interactions, utilizes a new DNA-binding domain conserved in a family of developmental regulators. Genes Dev, *7:* 2431-2445, 1993.

27. Scarpulla, R. C. Transcriptional activators and coactivators in the nuclear control of mitochondrial function in mammalian cells. Gene, *286:* 81-89, 2002.

28. Morrish, F., Giedt, C., and Hockenbery, D. c-MYC apoptotic function is mediated by NRF-1 target genes. Genes Dev, *17:* 240-255, 2003.

29. Lee, S. G. and Rho, H. M. Transcriptional repression of the human p53 gene by hepatitis B viral X protein. Oncogene, *19:* 468-471, 2000.

30. Elliott, K., Ge, K., Du, W., and Prendergast, G. C. The c-Myc-interacting adaptor protein Bin1 activates a caspase-independent cell death program. Oncogene, *19:* 4669-4684, 2000.

31. Watanabe, K., Ozaki, T., Nakagawa, T., Miyazaki, K., Takahashi, M., Hosoda, M., Hayashi, S., Todo, S., and Nakagawara, A. Physical interaction of p73 with c-Myc and MM1, a c-Myc-binding protein, and modulation of the p73 function. J Biol Chem, *277:* 15113-15123, 2002.

32. He, T. C., Sparks, A. B., Rago, C., Hermeking, H., Zawel, L., da Costa, L. T., Morin, P. J., Vogelstein, B., and Kinzler, K. W. Identification of c-MYC as a target of the APC pathway. [see comments]. Science, *281:* 1509-1512, 1998.

33. Riggs, K. J., Saleque, S., Wong, K. K., Merrell, K. T., Lee, J. S., Shi, Y., and Calame, K. Yin-yang 1 activates the c-myc promoter. Molecular & Cellular Biology, *13:* 7487-7495, 1993.

34. He, L., Liu, J., Collins, I., Sanford, S., O'Connell, B., Benham, C. J., and Levens, D. Loss of FBP function arrests cellular proliferation and extinguishes c-myc expression. Embo J, *19:* 1034-1044, 2000.

35. Bourgarel-Rey, V., Vallee, S., Rimet, O., Champion, S., Braguer, D., Desobry, A., Briand, C., and Barra, Y. Involvement of Nuclear Factor kappaB in c-Myc induction by Tubulin polymerization inhibitors. Mol Pharmacol, *59:* 1165-1170, 2001.

36. Coulis, C. M., Lee, C., Nardone, V., and Prokipcak, R. D. Inhibition of c-myc expression in cells by targeting an RNA-protein interaction using antisense oligonucleotides. Molecular Pharmacology, *57:* 485-494, 2000.

37. Ray, R. and Miller, D. M. Cloning and characterization of a human c-myc promoter-binding protein. Molecular & Cellular Biology, *11:* 2154-2161, 1991.

38. Zajac-Kaye, M. and Levens, D. Phosphorylation-dependent binding of a 138-kDa myc intron factor to a regulatory element in the first intron of the c-myc gene. Journal of Biological Chemistry, *265:* 4547-4551, 1990.

39. Zajac-Kaye, M., Gelmann, E. P., and Levens, D. A point mutation in the c-myc locus of a Burkitt lymphoma abolishes binding of a nuclear protein. Science, *240:* 1776-1780, 1988.

40. Kronke, M., Schluter, C., and Pfizenmaier, K. Tumor necrosis factor inhibits MYC expression in HL-60 cells at the level of mRNA transcription. Proceedings of the National Academy of Sciences of the United States of America, *84:* 469-473, 1987.

41. Kolli, S., Buchmann, A. M., Williams, J., Weitzman, S., and Thimmapaya, B. Antisense-mediated depletion of p300 in human cells leads to premature G1 exit and up-regulation of c-MYC. Proceedings of the National Academy of Sciences of the United States of America, *98:* 4646-4651, 2001.

42. Pietenpol, J. A., Holt, J. T., Stein, R. W., and Moses, H. L. Transforming growth factor beta 1 suppression of c-myc gene transcription: role in inhibition of keratinocyte proliferation. Proceedings of the National Academy of Sciences of the United States of America, *87:* 3758-3762, 1990.

43. Luscher, B. and Eisenman, R. N. New light on Myc and Myb. Part I. Myc. Genes & Development, *4:* 2025-2035, 1990.

44. Eick, D. and Bornkamm, G. W. Transcriptional arrest within the first exon is a fast control mechanism in c-myc gene expression. Nucleic Acids Research, *14:* 8331-8346, 1986.

45. Bentley, D. L. and Groudine, M. Novel promoter upstream of the human c-myc gene and regulation of c-myc expression in B-cell lymphomas. Mol Cell Biol, *6:* 3481-3489, 1986.

46. Cole, M. D. The myc oncogene: its role in transformation and differentiation. Annual Review of Genetics, *20:* 361-384, 1986.

47. Creancier, L., Mercier, P., Prats, A. C., and Morello, D. c-myc Internal ribosome entry site activity is developmentally controlled and subjected to a strong translational repression in adult transgenic mice. Molecular & Cellular Biology, *21:* 1833-1840, 2001.

48. Shliakhova, L. N., Itkes, A. V., Chernov, B. K., and Kiselev, L. L. [Transcription of antisense RNA for the human c-myc gene]. Mol Biol (Mosk), *28:* 909-917, 1994.

49. Shliakhova, L. N., Itkes, A. V., and Kiselev, L. L. [Mapping starts of antisense transcription of the human c-myc gene]. Dokl Akad Nauk, *337:* 263-265, 1994.

50. Nepveu, A., Fahrlander, P. D., Yang, J. Q., and Marcu, K. B. Amplification and altered expression of the c-myc oncogene in A-MuLV-transformed fibroblasts. Nature, *317:* 440-443, 1985.

51. Yang, J. Q., Bauer, S. R., Mushinski, J. F., and Marcu, K. B. Chromosome translocations clustered 5' of the murine c-myc gene qualitatively affect promoter usage: implications for the site of normal c-myc regulation. EMBO Journal, *4:* 1441-1447, 1985.

52. Chang, Y., Spicer, D. B., and Sonenshein, G. E. Effects of IL-3 on promoter usage, attenuation and antisense transcription of the c-myc oncogene in the IL-3-dependent Ba/F3 early pre-B cell line. Oncogene, *6:* 1979-1982, 1991.

53. DesJardins, E. and Hay, N. Repeated CT elements bound by zinc finger proteins control the absolute and relative activities of the two principal human c-myc promoters. Molecular & Cellular Biology, *13:* 5710-5724, 1993.

54. Broome, H. E., Reed, J. C., Godillot, E. P., and Hoover, R. G. Differential promoter utilization by the c-myc gene in mitogen- and interleukin-2-stimulated human lymphocytes. Molecular & Cellular Biology, *7:* 2988-2993, 1987.

55. Marcu, K. B., Patel, A. J., and Yang, Y. Differential regulation of the c-MYC P1 and P2 promoters in the absence of functional tumor suppressors: implications for mechanisms of deregulated MYC transcription. Current Topics in Microbiology & Immunology, *224:* 47-56, 1997.

56. Spencer, C. A. and Groudine, M. Transcription elongation and eukaryotic gene regulation. Oncogene, *5:* 777-785, 1990.

57. Michelotti, G. A., Michelotti, E. F., Pullner, A., Duncan, R. C., Eick, D., and Levens, D. Multiple single-stranded cis elements are associated with activated chromatin of the human c-myc gene in vivo. Molecular & Cellular Biology, *16:* 2656-2669, 1996.

58. Pullner, A., Mautner, J., Albert, T., and Eick, D. Nucleosomal structure of active and inactive c-myc genes. Journal of Biological Chemistry, *271:* 31452-31457, 1996.

59. Berberich, S. J. and Postel, E. H. PuF/NM23-H2/NDPK-B transactivates a human c-myc promoter-CAT gene via a functional nuclease hypersensitive element. Oncogene, *10:* 2343-2347, 1995.

60. Simonsson, T. DNA tetraplex formation in the control region of c-myc. Nucleic Acid Res., *26:* 1167-1172, 1998.

61. Simonsson, T., Pribylova, M., and Vorlickova, M. A nuclease hypersensitive element in the human c-myc promoter adopts several distinct i-tetraplex structures. Biochem. Biophys. Res. Comm. 158-166, 2000.

62. Kipling, D. The Telomere, p. 208. Oxford: Oxford University Press, 1995.

63. Parkinson, G. N., Lee, M. P., and Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. Nature, *417:* 876-880, 2002.

64. Arthanari, H. and Bolton, P. H. Functional and dysfunctional roles of quadruplex DNA in cells. Chem Biol, *8:* 221-230, 2001.

65. Matsugami, A., Ouhashi, K., Kanagawa, M., Liu, H., Kanagawa, S., Uesugi, S., and Katahira, M. An intramolecular quadruplex of (GGA)(4) triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction. J Mol Biol, *313:* 255-269, 2001.

66. Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., Jr., and Blackburn, E. H. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. Cell, *51:* 899-908, 1987.

67. Sundquist, W. I. and Klug, A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops [see comments]. Nature, *342:* 825-829, 1989.

68. Blackburn, E. Structure and function of telomeres. Nature, *350:* 569-573, 1991.

69. Sen, D. and Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature, *334:* 364-366, 1988.

70. Murchie, A. I. and Lilley, D. M. Retinoblastoma susceptibility genes contain 5' sequences with a high propensity to form guanine-tetrad structures. Nucleic Acids Research, *20:* 49-53, 1992.

71. Lyonnais, S., Hounsou, C., Teulade-Fichou, M. P., Jeusset, J., Le Cam, E., and Mirambeau, G. G-quartets assembly within a G-rich DNA flap. A possible event at the center of the HIV-1 genome. Nucleic Acids Res, *30:* 5276-5283, 2002.

72. Rando, R. F., Ojwang, J., Elbaggari, A., Reyes, G. R., Tinder, R., McGrath, M. S., and Hogan, M. E. Suppression of human immunodeficiency virus type 1 activity in vitro by oligonucleotides which form intramolecular tetrads. J Biol Chem, *270:* 1754-1760, 1995.

73. Bishop, J. S., Guy-Caffey, J. K., Ojwang, J. O., Smith, S. R., Hogan, M. E., Cossum, P. A., Rando, R. F., and Chaudhary, N. Intramolecular G-quartet motifs confer nuclease resistance to a potent anti-HIV oligonucleotide. J Biol Chem, *271:* 5698-5703, 1996.

74. Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H. J., and Pluckthun, A. In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. Proc Natl Acad Sci U S A, *98:* 8572-8577, 2001.

75. Auerbach, A. D. and Verlander, P. C. Disorders of DNA replication and repair. Curr Opin Pediatr, *9:* 600-616, 1997.

76. Chakraverty, R. K. and Hickson, I. D. Defending genome integrity during DNA replication: a proposed role for RecQ family helicases. Bioessays, *21:* 286-294, 1999.

77. Sun, H., Karow, J. K., Hickson, I. D., and Maizels, N. The Bloom's syndrome helicase unwinds G4 DNA. Journal of Biological Chemistry, *273:* 27587-27592, 1998.

78. Fry, M. and Loeb, L. A. Human werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d(CGG)n. Journal of Biological Chemistry, *274:* 12797-12802, 1999.

79. Saha, T. and Usdin, K. Tetraplex formation by the progressive myoclonus epilepsy type-1 repeat: implications for instability in the repeat expansion diseases. FEBS Lett, *491:* 184-187, 2001.

80. Fry, M. and Loeb, L. A. The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure. Proc Natl Acad Sci U S A, *91:* 4950-4954, 1994.

81. Brown, B. A., 2nd, Li, Y., Brown, J. C., Hardin, C. C., Roberts, J. F., Pelsue, S. C., and Shultz, L. D. Isolation and characterization of a monoclonal anti-quadruplex DNA antibody from autoimmune "viable motheaten" mice. Biochemistry, *37:* 16325-16337, 1998.

82. Kennedy, G. C., German, M. S., and Rutter, W. J. The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. Nat Genet, *9:* 293-298, 1995.

83. Lew, A. Unusual DNA structure of the diabetes susceptibility locus IDDM2 and its effect on transcription by the insulin promoter factor Pur-1/MAZ. Journal of Biological Chemistry, *273:* 34992-34999, 1998.

84. Baran, N., Pucshansky, L., Marco, Y., Benjamin, S., and Manor, H. The SV40 large T-antigen helicase can unwind four stranded DNA structures linked by G-quartets. Nucleic Acids Research, *25:* 297-303, 1997.

85. Sun, H., Bennett, R. J., and Maizels, N. The Saccharomyces cerevisiae Sgs1 helicase efficiently unwinds G-G paired DNAs. Nucleic Acids Research, *27:* 1978-1984, 1999.

86. Li, J. L., Harrison, R. J., Reszka, A. P., Brosh, R. M., Jr., Bohr, V. A., Neidle, S., and Hickson, I. D. Inhibition of the Bloom's and Werner's syndrome helicases by G-quadruplex interacting ligands. Biochemistry, *40:* 15194-15202, 2001.

87. Fang, G. and Cech, T. R. Characterization of a G-quartet formation reaction promoted by the beta-subunit of the Oxytricha telomere-binding protein. Biochemistry, *32:* 11646-11657, 1993.

88. Fang, G. and Cech, T. R. The beta subunit of Oxytricha telomere-binding protein promotes G-quartet formation by telomeric DNA. Cell, *74:* 875-885, 1993.

89. Laporte, L. and Thomas, G. J., Jr. Structural basis of DNA recognition and mechanism of quadruplex formation by the beta subunit of the Oxytricha telomere binding protein. Biochemistry, *37:* 1327-1335, 1998.

90. Schierer, T. and Henderson, E. A protein from Tetrahymena thermophila that specifically binds parallel-stranded G4-DNA. Biochemistry, *33:* 2240-2246, 1994.

91. Giraldo, R. and Rhodes, D. The yeast telomere-binding protein RAP1 binds to and promotes the formation of DNA quadruplexes in telomeric DNA. EMBO Journal, *13:* 2411-2420, 1994.

92. Hayashi, N. and Murakami, S. STM1, a gene which encodes a guanine quadruplex binding protein, interacts with CDC13 in Saccharomyces cerevisiae. Mol Genet Genomics, *267:* 806-813, 2002.

93. Muniyappa, K., Anuradha, S., and Byers, B. Yeast meiosis-specific protein Hop1 binds to G4 DNA and promotes its formation. Mol Cell Biol, *20:* 1361-1369, 2000.

94. Dempsey, L. A., Hanakahi, L. A., and Maizels, N. A specific isoform of hnRNP D interacts with DNA in the LR1 heterodimer: canonical RNA binding motifs in a sequence-specific duplex DNA binding protein. J Biol Chem, *273:* 29224-29229, 1998.

95. Dempsey, L. A., Sun, H., Hanakahi, L. A., and Maizels, N. G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D, A role for G-G pairing in immunoglobulin switch recombination. J Biol Chem, *274:* 1066-1071, 1999.

96. Walsh, K. and Gualberto, A. MyoD binds to the guanine tetrad nucleic acid structure. J Biol Chem, *267:* 13714-13718, 1992.

97. Weisman-Shomer, P. and Fry, M. Stabilization of tetrahelical DNA by the quadruplex DNA binding protein QUAD. Biochem Biophys Res Commun, *205:* 305-311, 1994.

98. Weisman-Shomer, P. and Fry, M. QUAD, a protein from hepatocyte chromatin that binds selectively to guanine-rich quadruplex DNA. J Biol Chem, *268:* 3306-3312, 1993.

99. Weisman-Shomer, P., Naot, Y., and Fry, M. Tetrahelical forms of the fragile X syndrome expanded sequence d(CGG)(n) are destabilized by two heterogeneous nuclear ribonucleoprotein-related telomeric DNA-binding proteins. J Biol Chem, *275:* 2231-2238, 2000.

100. Arimondo, P. B., Riou, J. F., Mergny, J. L., Tazi, J., Sun, J. S., Garestier, T., and Helene, C. Interaction of human DNA topoisomerase I with G-quartet structures. Nucleic Acids Research, *28:* 4832-4838, 2000.

101. Bashkirov, V. I., Scherthan, H., Solinger, J. A., Buerstedde, J. M., and Heyer, W. D. A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. J Cell Biol, *136:* 761-773, 1997.

102. Liu, Z. and Gilbert, W. The yeast KEM1 gene encodes a nuclease specific for G4 tetraplex DNA: implication of in vivo functions for this novel DNA structure. Cell, *77:* 1083-1092, 1994.

103. Liu, Z., Lee, A., and Gilbert, W. Gene disruption of a G4-DNA-dependent nuclease in yeast leads to cellular senescence and telomere shortening. Proc Natl Acad Sci U S A, *92:* 6002-6006, 1995.

104. Sun, H., Yabuki, A., and Maizels, N. A human nuclease specific for G4 DNA. Proc Natl Acad Sci U S A, *98:* 12444-12449, 2001.

105. Han, H., Cliff, C. L., and Hurley, L. H. Accelerated assembly of G-quadruplex structures by a small molecule. Biochemistry, *38:* 6981-6986, 1999.

106. Han, F., Wheelhouse, R., and Hurley, L. Interactions of TMPyP4 and TMPyP2 with quadruplex DNA. Structural basis for the differential effects on telomerase inhibition. J Am. Chem. Soc., *121:* 3561-3570, 1999.

107. Fedoroff, O. Y., Rangan, A., Chemeris, V. V., and Hurley, L. H. Cationic porphyrins promote the formation of i-motif DNA and bind peripherally by a nonintercalative mechanism. Biochemistry, *39:* 15083-15090, 2000.

108. Takimoto, M., Tomonaga, T., Matunis, M., Avigan, M., Krutzsch, H., Dreyfuss, G., and Levens, D. Specific binding of heterogeneous ribonucleoprotein particle protein K to the human c-myc promoter, in vitro. Journal of Biological Chemistry, *268:* 18249-18258, 1993.

109. Tomonaga, T. and Levens, D. Activating transcription from single stranded DNA. Proceedings of the National Academy of Sciences of the United States of America, *93:* 5830-5835, 1996.

110. Tsutsui, H., Sakatsume, O., Itakura, K., and Yokoyama, K. K. Members of the MAZ family: a novel cDNA clone for MAZ from human pancreatic islet cells. Biochemical & Biophysical Research Communications, *226:* 801-809, 1996.

111. Sakatsume, O., Tsutsui, H., Wang, Y., Gao, H., Tang, X., Yamauchi, T., Murata, T., Itakura, K., and Yokoyama, K. K. Binding of THZif-1, a MAZ-like zinc finger protein to the nuclease-hypersensitive element in the promoter region of the c-MYC protooncogene. J Biol Chem, *271:* 31322-31333, 1996.

112. Michelotti, E. F., Tomonaga, T., Krutzsch, H., and Levens, D. Cellular nucleic acid binding protein regulates the CT element of the human c-myc protooncogene. Journal of Biological Chemistry, *270:* 9494-9499, 1995.

113. Agou, F., Raveh, S., Mesnildrey, S., and Veron, M. Single strand DNA specificity analysis of human nucleoside diphosphate kinase B. Journal of Biological Chemistry, *274:* 19630-19638, 1999.

114. Hildebrandt, M., Lacombe, M. L., Mesnildrey, S., and Veron, M. A human NDP-kinase B specifically binds single-stranded poly-pyrimidine sequences. Nucleic Acids Research, *23:* 3858-3864, 1995.

115. Davis, T. L. Ribonucleoprotein and protein factors bind to an H-DNA-forming c-myc DNA element: possible regulators of the c-myc gene. Proceedings of the National Academy of Sciences of the United States of America, *86:* 9682-9686, 1989.

116. Shnyreva, M., Schullery, D. S., Suzuki, H., Higaki, Y., and Bomsztyk, K. Interaction of two multifunctional proteins. Heterogeneous nuclear ribonucleoprotein K and Y-box-binding protein. Journal of Biological Chemistry, *275:* 15498-15503, 2000.

117. Bomsztyk, K. Diverse molecular interactions of the hnRNP K protein. Journal of Biological Chemistry, *277:* 1560-1567, 2002.

118. Krecic, A. M. and Swanson, M. S. hnRNP complexes: composition, structure, and function. Curr Opin Cell Biol, *11:* 363-371, 1999.

119. Ritchie, S. A., Pasha, M. K., Batten, D. J., Sharma, R. K., Olson, D. J., Ross, A. R., and Bonham, K. Identification of the SRC pyrimidine-binding protein (SPy) as hnRNP K: implications in the regulation of SRC1A transcription. Nucleic Acids Res, *31:* 1502-1513, 2003.

120. Da Silva, N., Bharti, A., and Shelley, C. S. hnRNP-K and Pur(alpha) act together to repress the transcriptional activity of the CD43 gene promoter. Blood, *100:* 3536-3544, 2002.

121. Braddock, D. T., Baber, J. L., Levens, D., and Clore, G. M. Molecular basis of sequence-specific single-stranded DNA recognition by KH domains: solution structure of a complex between hnRNP K KH3 and single-stranded DNA. Embo J, *21:* 3476-3485, 2002.

122. Ostareck-Lederer, A., Ostareck, D. H., Cans, C., Neubauer, G., Bomsztyk, K., Superti-Furga, G., and Hentze, M. W. c-Src-mediated phosphorylation of hnRNP K drives translational activation of specifically silenced mRNAs. Mol Cell Biol, *22:* 4535-4543, 2002.

123. Expert-Bezancon, A., Le Caer, J. P., and Marie, J. Heterogeneous nuclear ribonucleoprotein (hnRNP) K is a component of an intronic splicing enhancer complex that activates the splicing of the alternative exon 6A from chicken beta-tropomyosin pre-mRNA. J Biol Chem, *277:* 16614-16623, 2002.

124. Chang, C. J., Luh, H. W., Wang, S. H., Lin, H. J., Lee, S. C., and Hu, S. T. The heterogeneous nuclear ribonucleoprotein K (hnRNP K) interacts with dengue virus core protein. DNA Cell Biol, *20:* 569-577, 2001.

125. Miau, L. H., Chang, C. J., Shen, B. J., Tsai, W. H., and Lee, S. C. Identification of heterogeneous nuclear ribonucleoprotein K (hnRNP K) as a repressor of C/EBPbeta-mediated gene activation. J Biol Chem, *273:* 10784-10791, 1998.

126. Lombardi, D., Lacombe, M. L., and Paggi, M. G. nm23: unraveling its biological function in cell differentiation. J Cell Physiol, *182:* 144-149, 2000.

127. Postel, E., Abramczyk, B., Levit, M., and Kyin, S. Catalysis of DNA cleavage and nucleoside triphosphate synthesis by NM23-H2/NDP kinase share an active site that implies a DNA repair function. Proc. Natl. Acad. Sci., *97:* 14194-14199, 2000.

128. Roymans, D., Willems, R., Van Blockstaele, D. R., and Slegers, H. Nucleoside diphosphate kinase (NDPK/NM23) and the waltz with multiple partners: possible consequences in tumor metastasis. Clin Exp Metastasis, *19:* 465-476, 2002.

129. Joosten, M., Vankan-Berkhoudt, Y., van den Broek, M., Beijen, A., Lowenbe, B., von Lindern, M., and Delwel, R. The gene NM23-M2 is frequently mutated in leukemia. Exp Hematol, *28:* 1491, 2000.

130. Lacombe, M. L., Milon, L., Munier, A., Mehus, J. G., and Lambeth, D. O. The human Nm23/nucleoside diphosphate kinases. J Bioenerg Biomembr, *32:* 247-258, 2000.

131. de la Rosa, A., Williams, R. L., and Steeg, P. S. Nm23/nucleoside diphosphate kinase: toward a structural and biochemical understanding of its biological functions. Bioessays, *17:* 53-62, 1995.

132. Levit, M. N., Abramczyk, B. M., Stock, J. B., and Postel, E. H. Interactions between Escherichia coli nucleoside-diphosphate kinase and DNA. J Biol Chem, *277:* 5163-5167, 2002.

133. Postel, E. H. NM23-NDP kinase. International Journal of Biochemistry & Cell Biology, *30:* 1291-1295, 1998.

134. Postel, E. H. Cleavage of DNA by human NM23-H2/nucleoside diphosphate kinase involves formation of a covalent protein-DNA complex. Journal of Biological Chemistry, *274:* 22821-22829, 1999.

135. Hartsough, M. T. and Steeg, P. S. Nm23/nucleoside diphosphate kinase in human cancers. J Bioenerg Biomembr, *32:* 301-308, 2000.

136. MacDonald, N. J., de la Rosa, A., and Steeg, P. S. The potential roles of nm23 in cancer metastasis and cellular differentiation. Eur J Cancer, *31A:* 1096-1100, 1995.

137. Postel, E. H., Berberich, S. J., Flint, S. J., and Ferrone, C. A. Human c-myc transcription factor PuF identified as nm23-H2 nucleoside diphosphate kinase, a candidate suppressor of tumor metastasis [see comments]. Science, *261:* 478-480, 1993.

138. Postel, E. H., Abramczyk, B. A., Gursky, S. K., and Xu, Y. Structure-based mutational and functional analysis identify human NM23-H2 as a multifunctional enzyme. Biochemistry, *41:* 6330-6337, 2002.

139. Caligo, M. A., Cipollini, G., Fiore, L., Calvo, S., Basolo, F., Collecchi, P., Ciardiello, F., Pepe, S., Petrini, M., and Bevilacqua, G. NM23 gene expression correlates with cell growth rate and S-phase. Int J Cancer, *60:* 837-842, 1995.

140. Ma, D. NM23-H1 and NM23-H2 repress transcriptional activities of nuclease-hypersensitive elements in the platelet-derived growth factor-A promoter. Proceedings of the National Academy of Sciences of the United States of America, *97:* 12508-12512, 2000.

141. Ji, L., Arcinas, M., and Boxer, L. M. The transcription factor, Nm23H2, binds to and activates the translocated c-myc allele in Burkitt's lymphoma. Journal of Biological Chemistry, *270:* 13392-13398, 1995.

142. Postel, E. H., Mango, S. E., and Flint, S. J. A nuclease-hypersensitive element of the human c-myc promoter interacts with a transcription initiation factor. Molecular & Cellular Biology, *9:* 5123-5133, 1989.

143. Postel, E. H. NM23/Nucleoside diphosphate kinase as a transcriptional activator of c-myc. Curr Top Microbiol Immunol, *213 (Pt 2):* 233-252, 1996.

144. Postel, E. H., Berberich, S. J., Rooney, J. W., and Kaetzel, D. M. Human NM23/Nucleoside diphosphate kinase regulates gene expression through DNA binding to nuclease-hypersensitive transcriptional elements. J Bioenerg Biomemb, *32:* 277-284, 2000.

145. Agou, F., Raveh, S., and Veron, M. The binding mode of human nucleoside diphosphate kinase B to single-strand DNA. J. Bioenerg. Biomemb., *32:* 285-292, 2000.

146. Postel, E. H. and Ferrone, C. A. Nucleoside diphosphate kinase enzyme activity of NM23-H2/PuF is not required for its DNA binding and in vitro transcriptional functions. J Biol Chem, *269:* 8627-8630, 1994.

147. Raveh, S., Vinh, J., Rossier, J., Agou, F., and Veron, M. Peptidic determinants and structural model of human NDP kinase B (Nm23-H2) bound to single-stranded DNA. Biochemistry, *40:* 5882-5893, 2001.

148. Ouatas, T., Abdallah, B., Gasmi, L., Bourdais, J., Postel, E., and Mazabraud, A. Three different genes encode NM23/nucleoside diphosphate kinases in Xenopus laevis. Gene, *194:* 215-225, 1997.

149. Postel, E. H., Weiss, V. H., Beneken, J., and Kirtane, A. Mutational analysis of NM23-H2/NDP kinase identifies the structural domains critical to recognition of a c-myc regulatory element. Proc Natl Acad Sci U S A, *93:* 6892-6897, 1996.

150. Flink, I. L. and Morkin, E. Alternatively processed isoforms of cellular nucleic acid-binding protein interact with a suppressor region of the human beta-myosin heavy chain gene. Journal of Biological Chemistry, *270:* 6959-6965, 1995.

151. Rajavashisth, T. B., Taylor, A. K., Andalibi, A., Svenson, K. L., and Lusis, A. J. Identification of a zinc finger protein that binds to the sterol regulatory element. Science, *245:* 640-643, 1989.

152. Abdul-Manan, N., O'Malley, S. M., and Williams, K. R. Origins of binding specificity of the A1 heterogeneous nuclear ribonucleoprotein. Biochemistry, *35:* 3545-3554, 1996.

153. Li, H. P., Zhang, X., Duncan, R., Comai, L., and Lai, M. M. Heterogeneous nuclear ribonucleoprotein A1 binds to the transcription-regulatory region of mouse hepatitis virus RNA. Proceedings of the National Academy of Sciences of the United States of America, *94:* 9544-9549, 1997.

154. Fukuda, H., Katahira, M., Tsuchiya, N., Enokizono, Y., Sugimura, T., Nagao, M., and Nakagama, H. Unfolding of quadruplex structure in the G-rich strand of the minisatellite repeat by the binding protein UP1. Proc Natl Acad Sci U S A, *99:* 12685-12690, 2002.

155. Erlitzki, R. and Fry, M. Sequence-specific binding protein of single-stranded and unimolecular quadruplex telomeric DNA from rat hepatocytes. Journal of Biological Chemistry, *272:* 15881-15890, 1997.

156. Grand, C. L., Han, H., Munoz, R. M., Weitman, S., Von Hoff, D. D., Hurley, L. H., and Bearss, D. J. The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth in vivo. Mol Cancer Ther, *1:* 565-573, 2002.

157. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. Proc Natl Acad Sci U S A, *99:* 11593-11598, 2002.

158. Instruments, H. S. Protein Electrophoresis Applications Guide, p. 106. San Francisco, CA, USA: Hoefer Scientific Instruments, 1994.

159. Watts, G. S., Futscher, B. W., Isett, R., Gleason-Guzman, M., Kunkel, M. W., and Salmon, S. E. cDNA microarray analysis of multidrug resistance: doxorubicin selection produces multiple defects in apoptosis signaling pathways. J Pharmacol Exp Ther, *299:* 434-441, 2001.

160. Weitzmann, M. N. and Howell, R. M. The development and use of a DNA polymerase arrest assay for the evaluation of parameters affecting intrastrand tetraplex formation. The chicken beta-globin gene promoter forms a novel "cinched" tetrahelical structure. Journal of Biological Chemistry, *271:* 20958-20964, 1996.

161. Sun, D., Thompson, B., Cathers, B. E., Salazar, M., Kerwin, S. M., Trent, J. O., Jenkins, T. C., Neidle, S., and Hurley, L. H. Inhibition of human telomerase by a G-quadruplex-interactive compound. Journal of Medicinal Chemistry, *40:* 2113-2116, 1997.

162. Greenberg, R. A., O'Hagan, R. C., Deng, H., Xiao, Q., Hann, S. R., Adams, R. R., Lichtsteiner, S., Chin, L., Morin, G. B., and DePinho, R. A. Telomerase reverse transcriptase gene is a direct target of c-Myc but is not functionally equivalent in cellular transformation. Oncogene, *18:* 1219-1226, 1999.

163. Lutz, W., Leon, J., and Eilers, M. Contributions of Myc to tumorigenesis. Biochim Biophys Acta, *1602:* 61-71, 2002.

164. Hanahan, D. and Weinberg, R. A. The hallmarks of cancer. Cell, *100:* 57-70, 2000.

165. Cory, S. Activation of cellular oncogenes in hemopoietic cells by chromosome translocation. Adv Cancer Res, *47:* 189-234, 1986.

166. Erisman, M. D., Rothberg, P. G., Diehl, R. E., Morse, C. C., Spandorfer, J. M., and Astrin, S. M. Deregulation of c-myc gene expression in human colon carcinoma is not accompanied by amplification or rearrangement of the gene. Mol Cell Biol, *5:* 1969-1976, 1985.

167. Kinzler, K. W. and Vogelstein, B. Lessons from hereditary colorectal cancer. Cell, *87:* 159-170, 1996.

168. Pignatelli, M. The adenomatous polyposis coli tumour suppressor gene regulates c-MYC transcription in colon cancer cells. Gut, *44:* 596, 1999.

169. Andrew, S. D., Capes-Davis, A., Delhanty, P. J., Marsh, D. J., Mulligan, L. M., and Robinson, B. G. Transcriptional repression of the RET proto-oncogene by a mitogen activated protein kinase-dependent signalling pathway. Gene, *298:* 9-19, 2002.

170. Andrew, S. D., Delhanty, P. J., Mulligan, L. M., and Robinson, B. G. Sp1 and Sp3 transactivate the RET proto-oncogene promoter. Gene, *256:* 283-291, 2000.

171. Rustighi, A., Tessari, M. A., Vascotto, F., Sgarra, R., Giancotti, V., and Manfioletti, G. A polypyrimidine/polypurine tract within the Hmga2 minimal promoter: a common feature of many growth-related genes. Biochemistry, *41:* 1229-1240, 2002.

172. Rustighi, A., Mantovani, F., Fusco, A., Giancotti, V., and Manfioletti, G. Sp1 and CTF/NF-1 transcription factors are involved in the basal expression of the Hmgi-c proximal promoter. Biochem Biophys Res Commun, *265:* 439-447, 1999.

173. Chau, K., Arlotta, P., Patel, U. A., Crane-Robinson, C., Manfioletti, G., and Ono, S. J. A novel downstream positive regulatory element mediating transcription of the human high mobility group (HMG) I-C gene. FEBS Lett, *457:* 429-436, 1999.

174. Dima, V. F., Vasiliu, V., and Dima, S. V. Photodynamic therapy: an update. Roum Arch Microbiol Immunol, *57:* 207-230, 1998.