J|A|C|S
ARTICLES
Published on Web 06/22/2004

The Dynamic Character of the G-Quadruplex Element in the c-MYC Promoter and Modification by TMPyP4

Jeyaprakashnarayanan Seenisamy,[†] Evonne M. Rezler,[†] Tiffanie J. Powell,[†] Denise Tye,[†] Vijay Gokhale,[†] Chandana Sharma Joshi,[‡] Adam Siddiqui-Jain,*[,§] and Laurence H. Hurley*[,†,‡,||]

Contribution from the College of Pharmacy, The University of Arizona, 1703 East Mabel, Tucson, Arizona 85721, Department of Chemistry, The University of Arizona, Tucson, Arizona 85721, Cylene Pharmaceuticals, 11045 Roselle Street, Suite C, San Diego, California 92121, and Arizona Cancer Center, 1515 North Campbell Avenue, Tucson, Arizona 85724

Received January 20, 2004; E-mail: hurley@pharmacy.arizona.edu; adamsj@cylenepharma.com

Abstract: The nuclease hypersensitivity element III_1 (NHE III_1) upstream of the P1 and P2 promoters of c-MYC controls 80−90% of the transcriptional activity of this gene. The purine-rich strand in this region can form a G-quadruplex structure that is a critical part of the silencer element for this promoter. We have demonstrated that this G-quadruplex structure can form a mixture of four biologically relevant parallel-loop isomers, which upon interaction with the cationic porphyrin TMPyP4 are converted to mixed parallel/antiparallel G-quadruplex structures.

Introduction

Overexpression of the c-MYC oncogene is linked with cellular proliferation and inhibition of differentiation, leading to its association with a wide range of human cancers, including colon, breast, small-cell lung, osteosarcomas, glioblastomas, and myeloid leukemia.[1−4] In transfection experiments, c-MYC has been demonstrated to immortalize normal fibroblasts, and in concert with other genes, such as v-ABL and v-RAS, it causes malignant transformation.[4] Because of this, c-MYC transcriptional control has emerged as an attractive target for anticancer therapeutic strategies.

The NHE III_1 of the c-MYC promoter controls 80−90% of c-MYC transcription and has been the subject of considerable research over the past two decades.[5−15] The NHE III_1 is a 27-base-pair sequence, located −142 to −115 base pairs upstream of the P1 promoter (Figure 1). This duplex element can equilibrate between transcriptionally active forms (duplex and single-stranded DNA) and a silenced form.[16] We have previously shown that single G-to-A mutations, which destabilize the G-quadruplex-forming unit, result in a 3-fold increase in basal transcriptional activity.[17] In contrast, agents that stabilize the specific G-quadruplex structure are able to suppress c-MYC transcriptional activity.[17] Furthermore, on the basis of DMS

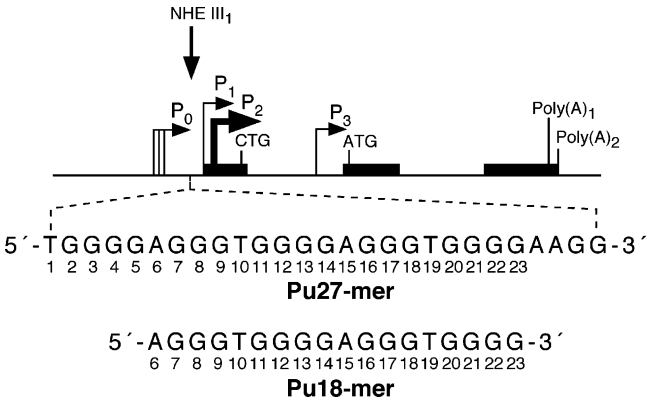

Figure 1. Promoter structure of the c-MYC gene and location of the NHE III_1. Shown in the inset are the NHE III_1 and numbering of the Pu27-mer sequence on the purine-rich strand. The numbering sequence of the truncated Pu18-mer is shown below.

[†] College of Pharmacy, The University of Arizona.
[‡] Department of Chemistry, The University of Arizona.
[§] Cylene Pharmaceuticals.
[||] Arizona Cancer Center.

(1) Pelengaris, S.; Rudolph, B.; Littlewood, T. *Curr. Opin. Genet. Dev.* **2000**, *10*, 100−105.
(2) Spencer, C. A.; Groudine, M. *Adv. Cancer Res.* **1991**, *56*, 1−48.
(3) Marcu, K. B.; Bossone, S. A.; Patel, A. J. *Annu. Rev. Biochem.* **1992**, *61*, 809−860.
(4) Facchini, L. M.; Penn, L. Z. *FASEB J.* **1998**, *12*, 633−651.
(5) Sakatsume, O.; Tsutsui, H.; Wang, Y.; Gao, H.; Tang, X.; Yamauchi, T.; Murata, T.; Itakura, K.; Yokoyama, K. K. *J. Biol. Chem.* **1996**, *271*, 31322−31333.
(6) Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *214*, 456−459.
(7) Siebenlist, U.; Henninghausen, L.; Battey, J.; Leder, P. *Cell* **1984**, *37*, 381−391.
(8) Boles, T. C.; Hogan, M. E. *Biochemistry* **1987**, *26*, 367−376.
(9) Simonsson, T.; Pribylova, M.; Vorlickova, M. *Biochem. Biophys. Res. Commun.* **2000**, *278*, 158−166.
(10) Simonsson, T.; Pecinka, P.; Kubista, M. *Nucleic Acids Res.* **1998**, *26*, 1167−1172.
(11) Ji, L.; Arcinas, M.; Boxer, L. M. *J. Biol. Chem.* **1995**, *270*, 13392−13398.
(12) Tomonaga, T.; Levens, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5830−5835.
(13) Bossone, S. A.; Asselin, C.; Patel, A. J.; Marcu, K. B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7452−7456.
(14) Michelotti, E. F.; Tomonaga, T.; Krutzsch, H.; Levens, D. *J. Biol. Chem.* **1995**, *270*, 9494−9499.
(15) Postel, E. H.; Berberich, S. J.; Rooney, J. W.; Kaetzel, D. M. *J. Bioenerg. Biomembr.* **2000**, *32*, 277−284.
(16) Collins, I.; Weber, A.; Levens, D. *Mol. Cell. Biol.* **2001**, *21*, 8437−8451.
(17) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593−11598.

10.1021/ja040022b CCC: $27.50

footprinting, we had proposed a model in which the biologically relevant structure was a chair G-quadruplex,[17] in accordance with the proposal of similar structures for the HIV aptamer T30695 and the thrombin binding aptamer (TBA).[18−21] On the basis of circular dichroism (CD) and further mutational studies, we now demonstrate that this G-quadruplex structure is a highly dynamic mixture of parallel-loop isomers, all of which appear to be biologically relevant. Upon incubation with TMPyP4, the parallel-loop isomers convert to a mixed parallel/antiparallel structure, which probably is the basis for the inhibition of c-MYC expression by this compound.

Results

CD Spectroscopy Shows That the Pu27-Mer NHE III$_1$ Is a Parallel Intramolecular G-Quadruplex Structure. T30695 and TBA, as well as the heptad-tetrad-forming $(GGA)_4$ oligonucleotide, were previously shown by NMR and X-ray crystallography to form monomeric or intramolecular G-quadruplexes.[18−24] Most importantly, in the previous NMR solution studies, buffering and high salt conditions similar to those that we use in our current study were employed, with the exception that much higher oligonucleotide concentrations of the millimolar order were used in the published studies. Therefore, by analogy, an oligonucleotide such as the Pu27-mer, which is longer than any of the aforementioned oligonucleotides and which has a G-content similar to that of those sequences, is also likely to form a monomeric or intramolecular G-quadruplex. In fact, the conditions used in our current study are particularly geared to the formation of intramolecular G-quadruplexes by the Pu27-mer because strand concentrations did not exceed 10 μM. Thus, it is reasonable to assume that the G-quadruplex structures studied here are in an intramolecular, rather than an intermolecular, form. Further evidence of this was obtained by comparing the EMSA mobility of the suspected intramolecular band with an equivalent 27-mer that forms an intermolecular hairpin dimer structure (unpublished results).

Our previous assignment of a chair G-quadruplex model for the Pu27-mer guanine-rich strand of the NHE III$_1$ was based on DMS footprinting and similar chair structures proposed for T30695 and TBA. The chair structure of TBA was proposed on the basis of NMR[18] and X-ray crystallography,[22] while the chair structure of T30695 was proposed only on the basis of NMR studies.[21] However, upon examination by CD, it was found that while the spectra of T30695 and the Pu27-mer were coincident, the TBA gave a very different CD signature, i.e., the T30695 and Pu27-mer showed absorption maxima at 262 nm, while the TBA had a maximum at 295 nm. (Figure 2A). Only the CD of the TBA is in accordance with an antiparallel structure,[25] while the CD signatures of the T30695 and the c-MYC Pu27-mer are consistent with a parallel structure,[25,26]

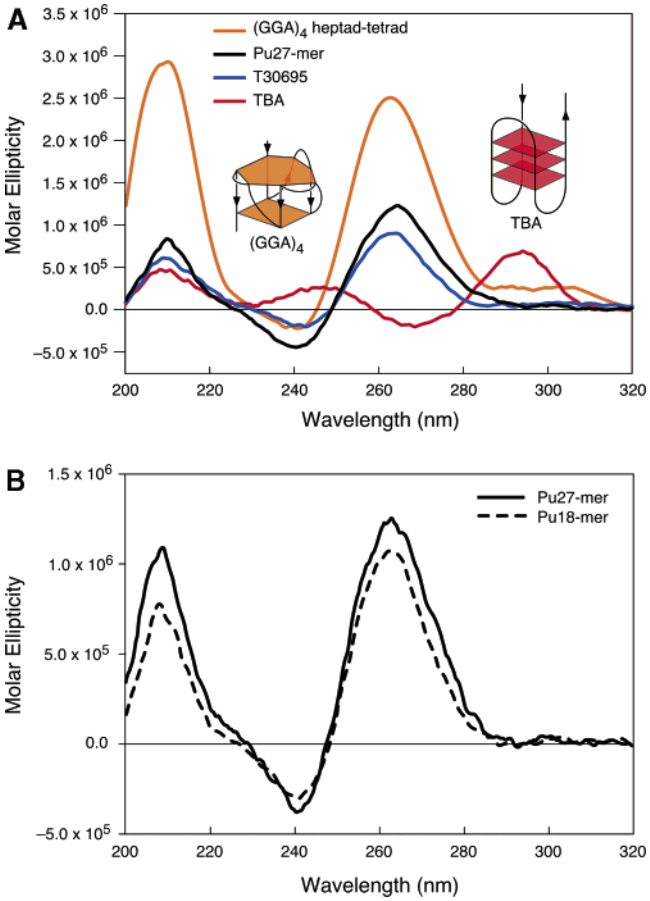

Figure 2. (A) CD spectra of the d$(GGA)_4$ oligonucleotide (orange graph and the orange G-quadruplex represent the parallel-stranded heptad-tetrad structure determined from NMR studies), the Pu27-mer (black graph), the T30695 oligonucleotide (blue graph), and the thrombin binding aptamer (TBA) (red graph and the red G-quadruplex represent the antiparallel chair structure determined from X-ray and NMR studies). All CD data were obtained with a 10 μM strand concentration in the presence of 100 μM KCl at 25 °C. (B) The CD spectra of the Pu27-mer (solid line) and Pu18-mer (dashed line) from the NHE III$_1$ of the c-MYC promoter. The CD data were obtained with a 10 μM strand concentration in the presence of 100 mM KCl at 25 °C.

which is typified by the heptad-tetrad structure determined previously by NMR[23,24] and shown in Figure 2A.

DMS cleavage of the Pu27-mer shows that G2−G5 are not involved in the G-quadruplex structure,[17] so a Pu18-mer containing the four consecutive 3′ runs of guanines (A6 through G23) (Figure 1) was made and used along with the Pu27-mer in subsequent studies. Oligomers containing just these four 3′ runs of guanines show a DMS cleavage pattern identical to that of the 27-mer (unpublished results). A comparison of the CD of the 18- and 27-mer, shown in Figure 2B, reveals that the 18-mer retains the parallel structure of the 27-mer.

Last, a comparison of melting temperatures (T_m) of the Pu27-mer with the those of T30695 and TBA was made by CD under uniform conditions of 100 mM KCl and a strand concentration of 10 μM. The Pu27-mer and T30695 showed similar T_m values of 85 and 90 °C, respectively, perhaps characteristic of parallel structures, such as the heptad-tetrad G-quadruplexes $(GGA)_4$ and $(GGA)_8$, for which the T_m values determined under similar conditions were 88 and 86 °C, respectively.[23,24] In contrast, TBA

(18) Schultze, P.; Macaya, R. F.; Feigon, J. *J. Mol. Biol.* **1994**, *235*, 1532−1547.
(19) Kelly, J. A.; Feigon, J.; Yeates, T. O. *J. Mol. Biol.* **1996**, *256*, 417−422.
(20) Jing, N. J.; Gao, X. L.; Rando, R. F.; Hogan, M. E. *J. Biomol. Struct. Dyn.* **1997**, *15*, 573−585.
(21) Jing, N. J.; Hogan, M. E. *J. Biol. Chem.* **1998**, *273*, 34992−34999.
(22) Padmanabhan, K.; Padmanabhan, K. P.; Ferraras, J. D.; Sadler, J. E.; Tulinsky, A. *J. Biol. Chem.* **1993**, *268*, 17651−17654.
(23) Matsugami, A.; Okuizumi, T.; Uesugi, S.; Katahira, M. *J. Biol. Chem.* **2003**, *278*, 28147−28153.
(24) Matsugami, A.; Ouhashi, K.; Kanagawa, M.; Liu, H.; Kanagawa, S.; Uesugi, S.; Katahira, M. *J. Mol. Biol.* **2001**, *313*, 255−269.
(25) Dapic, V.; Abdomerovic, V.; Marrington, R.; Peberdy, J.; Rodger, A.; Trent, J. O.; Bates, P. J. *Nucleic Acids Res.* **2003**, *31*, 2097−2107.
(26) Kettani, A.; Gorin, A.; Majumdar, A.; Hermann, T.; Skripkin, E.; Zhao, H.; Jones, R.; Patel, D. J. *J. Mol. Biol.* **2000**, *297*, 627−644.

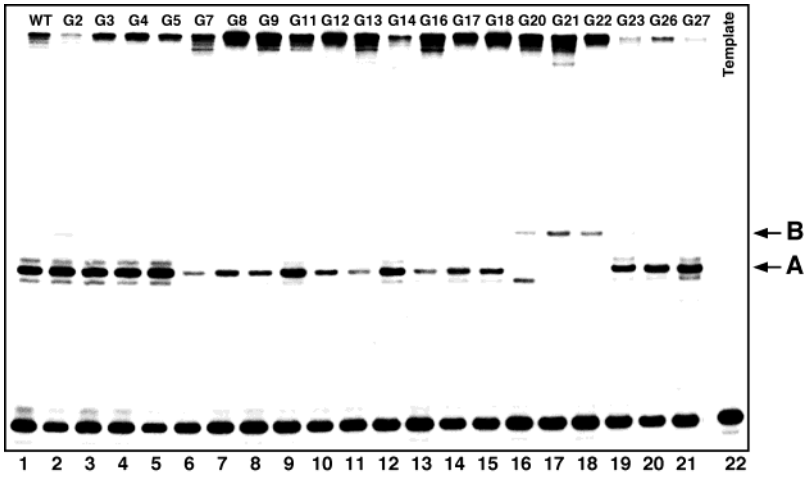

Figure 3. Effect of single G-to-A mutations at positions G2−G27 (lanes 2−21) on the stability of the Pu27-mer sequence in the polymerase stop assay.[17] Each of these lanes contained the mutant oligomer incubated with Taq polymerase at 37 °C (10 mM KCl and NaCl). Lane 22 lacked Taq polymerase. Arrows A and B refer to the polymerase stop products corresponding to either the four 3′ runs of guanines (A) or the four 5′ runs of guanines (B).

showed a much lower T_m value of 47 °C under the same conditions. Such a low T_m value may be a characteristic property of antiparallel intramolecular chair structures because it is similar to the T_m value of 56 °C found previously for another antiparallel intramolecular, albeit basket, G-quadruplex that was formed by the human telomeric sequence $dAG_3(T_2AG_3)_3$ and under similar salt conditions of 100 mM NaCl.[27] Thus, it would seem likely that the structures of the Pu27-mer and T30695 are similar, but in the absence of comparative NMR data, this conclusion must be still considered tentative.

Within the Four Consecutive 3′ Runs of Guanines Involved in G-Tetrad Formation, G11, G14, G20, and G23 Are Less Critical for G-Quadruplex Stability. The availability of N7 of guanine for methylation by DMS in duplex and single-stranded DNA, but not in G-tetrads, is a diagnostic indicator of G-quadruplex structures. In the previously reported DMS-induced cleavage experiment,[17] within the four consecutive 3′ runs of guanines, only G11, G14, G20, and G23 showed cleavage, suggesting that they were not involved in stable G-tetrad formation. If these guanines are not involved in the G-tetrads, then mutation to adenines should not significantly affect G-quadruplex stabilization. A polymerase stop assay was used to evaluate the effect of single guanine mutations (G-to-A) on G-quadruplex stability. The results (Figure 3) show that the 5′ run of guanines (G2−G5) is not required for formation of a stable G-quadruplex structure, as predicted from the DMS cleavage results.[17] However, for the four consecutive 3′ runs of guanines, the pattern is more complex. G7−G9, G12−G13, G16−G18, and G21−G22 are clearly critical for G-quadruplex stability, while G11, G14, G23, and perhaps G20 appear less critical. The upper bands (**B** in Figure 3) for the G-to-A mutations at G20, G21, and G22 correspond to a new polymerase stop site that presumably results from an alternative structure that uses the four consecutive 5′ runs of guanines. The G20 mutation is unique in that it shows that both types of G-quadruplex structures can be formed using either the four consecutive 5′ or 3′ runs of guanines. The lack of dependency on G11, G14, and G23—and perhaps G20—for G-quadruplex stability can be explained if we assume a degeneracy in use of the two runs of four guanines (i.e., G11−G14 and G20−G23), such that only three of the four guanines are required in each case. It should be noted that the other two consecutive 3′ runs of guanines (G7−G9 and G16−G18) have only three guanines, not allowing for degeneracy in these cases.

A Parallel G-Quadruplex Structure for the Pu18-Mer Is Compatible with Four Different Loop Isomers, Each of Which Has a Lower Melting Temperature Than the Wild-Type Sequence. The four possible loop isomers of the Pu18-mer parallel structure are shown in the upper row of structures in Figure 4A. These three-tetrad-containing loop isomers result from the degeneracy of the two runs of four guanines (G11−G14 and G20−G23), such that guanine slippage can occur so that either G11 or G14 (or G20 or G23) is involved in tetrad formation, while the other guanine is located within a loop. To eliminate this degeneracy, four different dual G-to-T mutations in the 18-mer were designed and synthesized (G→T-14,23; -14,-20; -11,23; and -11,20) so that each of the two runs of four guanines is reduced to three, and the resulting mutant thymines are forced to assume loop positions in order to preserve the three tetrad structures (lower row of structures in Figure 4A). The wild-type Pu18-mer sequence and the four dual mutants were subjected to an electromobility shift assay (EMSA), and the results are shown in Figure 4B. The wild-type sequence (lane 1) shows two bands, while each of the four different dual mutants shows only one band (lanes 2−5). On the basis of a comparison of the mobility of the bands corresponding to the wild-type and the four mutant 18-mers, band B in the wild-type sequence corresponds to loop isomers where G23 is intratetrad, while for band A, G23 is external. DMS footprinting of these four different dual mutants failed to show any cleavage, which indicates that, for each 18-mer, all of the guanines are involved in formation of G-tetrads (unpublished results).

Finally, each of the dual-mutant loop isomers has a significantly lower melting temperature than the wild-type sequence. The T_m value of each of the four loop isomers (A−D in Figure 4A) was determined by CD under the same conditions described for the Pu27-mer, and all the values were found to be between 68 and 72 °C.

Molecular Modeling of the 1:2:1 Loop Isomer. The c-MYC G-quadruplex parallel-type structure with a 1:2:1 arrangement of loops (two single base loops and one loop with two bases;

(27) Mergny, J.-L.; Phan, A.-T.; Lacroix, L. *FEBS Lett.* **1998**, *435*, 74−78.

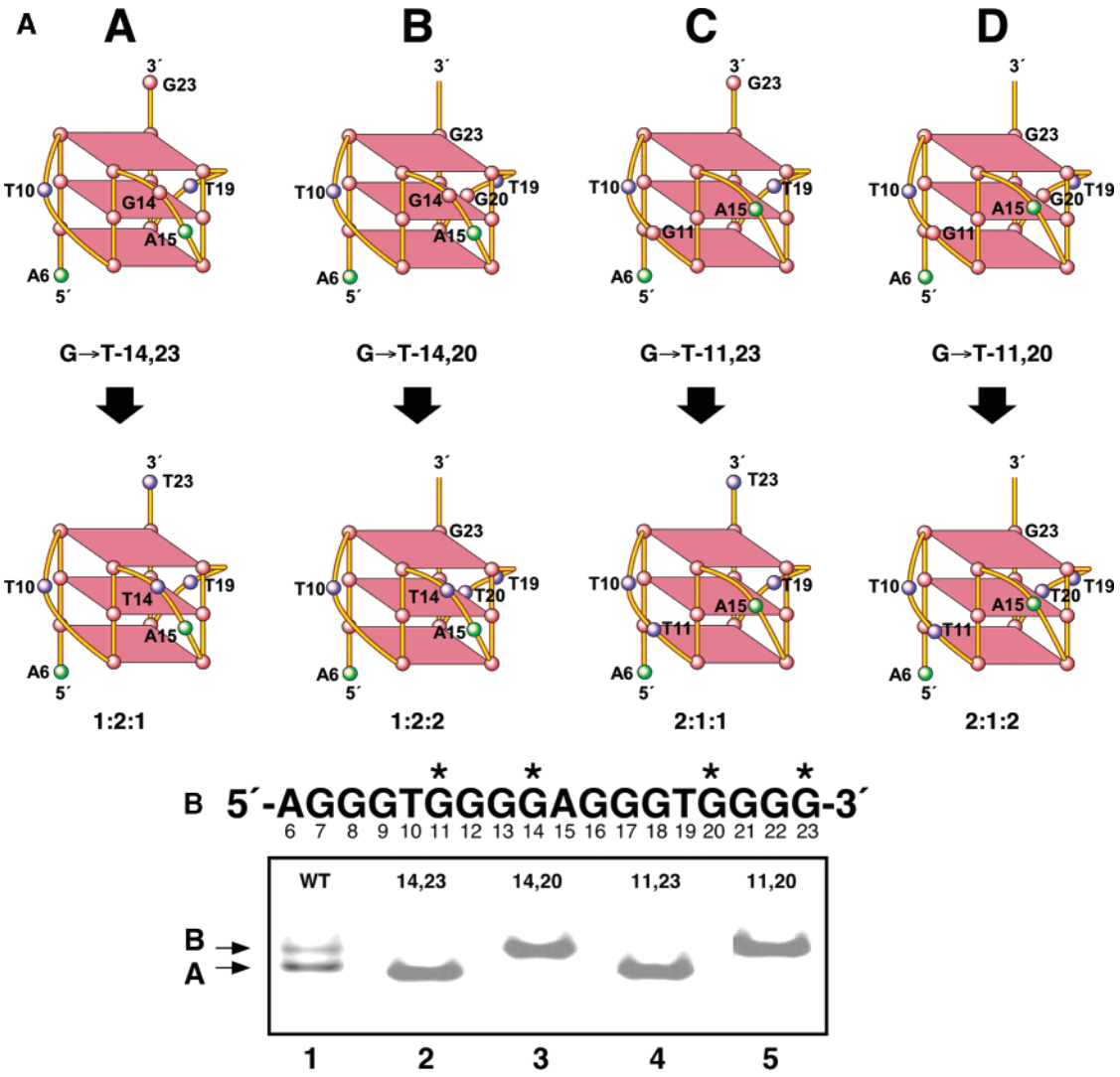

Figure 4. (A) Proposed structures of the four different loop isomers, in which dual G-to-T mutations at positions 11, 14, 20, and 23 result in defined loop isomers. The upper row shows the four proposed isomers and the lower row shows the result of dual G-to-T mutations. Base colors: red = guanine, blue = thymine, green = adenine, and yellow = cytosine. (B) EMSA of the wild-type (lane 1) and the four Pu18-mer loop isomers (lanes 2−5) containing the dual mutants G→T-14,23, -14,20, -11,23, and -11,20 respectively The Pu18-mer sequence and the positions of the four G-to-T mutations are indicated with asterisks above the EMSA.

see Figure 4A) was found to be stable during molecular minimization and dynamics calculations (Figure 5). All the guanine bases of the tetrads exist in an anti conformation. Single base loops connecting parallel guanine strands, as well as all guanine bases of the strands connected by single base loops in the tetrad arrangement, were found to be stable.

Upon Addition of TMPyP4 to the Parallel Pu27-Mer Sequence, There Is a Conversion to a Mixed Parallel/Antiparallel Structure with Which TMPyP4 Stacks Externally within the Lateral Loops. Titration of TMPyP4 into a solution of the Pu27-mer in the absence of added salt shows a significant change in the CD spectra, such that after addition of 4 mol equiv of drug, an absorption at 285 nm emerges as a shoulder (Figure 6). There are incremental changes in the CD up to the addition of 4 mol equiv. The signature is very similar to a mixed parallel/antiparallel structure (Figure 6), where one loop is internal and the two adjacent loops are external.[28] We therefore conclude that the product after addition of TMPyP4 corresponds most likely to a mixed parallel/antiparallel G-quadruplex structure. The induced CD absorption at 360 and 450 nm is due to TMPyP4 in an asymmetric chemical environment.

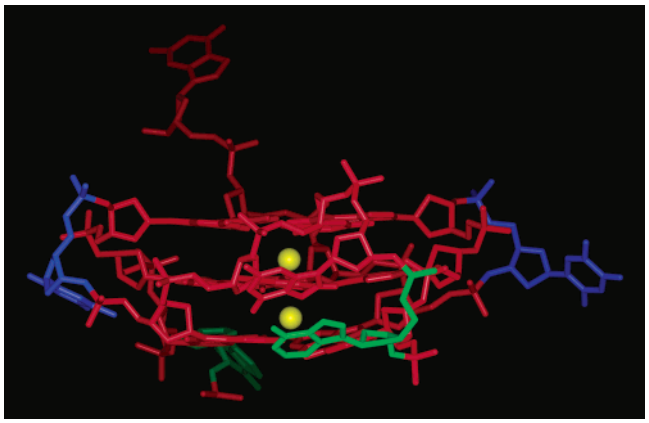

Figure 5. Most stable low-energy structure of the c-MYC Pu18-mer parallel quadruplex (color coding: adenine = green, cytosine = yellow, guanine = red, and thymine = blue) with potassium atoms (as CPK model). For clarity, hydrogen atoms have not been shown.

(28) Wang, Y.; Patel, D. J. *Structure* **1994**, *2*, 1141−1156.

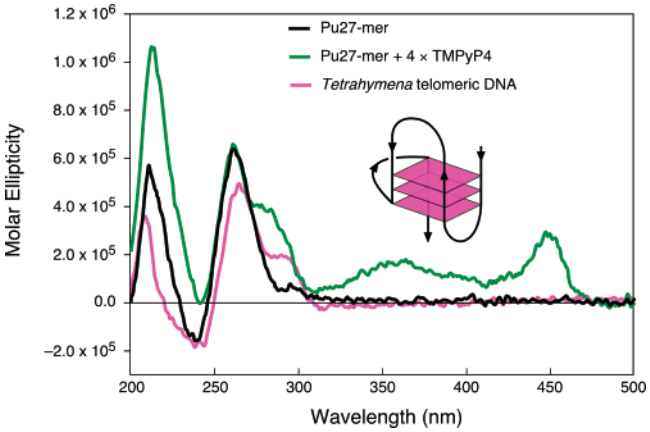

Figure 6. The CD spectra of the Pu27-mer (black line), the Pu27-mer with a 4 mol equivalence of TMPyP4 (green line), and the *Tetrahymena* telomeric DNA sequence (magenta line) in NaCl (100 mM), which, from NMR studies, was previously determined to form the mixed parallel/antiparallel G-quadruplex with two lateral loops and one propeller loop (inset).

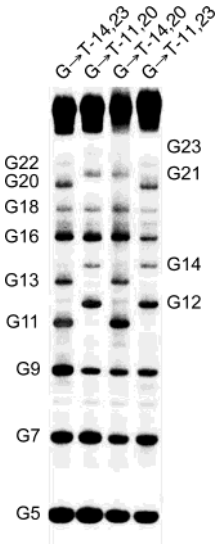

Figure 7. TMPyP4 photoinduced cleavage of the four dual-mutant loop isomers (G→T-14,23, -11,20, -14,20, and -11,23). The numbering of key guanine residues is shown on each side of the gel. A 27-mer was used, and G-to-T mutations at 3 and 4 were also made in each oligomer to obtain a single isomer. The band for G23 is present at higher sensitivity for G→T-11,20 and -14,20 but absent in the other two mutants. In contrast, G22 is present for G→T-14,23 and -11,23 only.

To determine the sites of binding of TMPyP4 to the mixed parallel/antiparallel structure, photocleavage of the DNA by TMPyP4 was carried out.[17] TMPyP4 catalyzes the oxidation of DNA upon exposure to light, which results in DNA strand breakage in proximity to the binding sites. The results of photocleavage of the four different dual-mutant Pu27-mers are shown in Figure 7. Each of these mutants shows a distinct TMPyP4 photocleavage pattern in which 8 of the 12 guanines are cleaved. The eight cleavage sites correspond to the external tetrads, with the interior tetrad protected from cleavage (Figure 7). The different positions of photocleavage for each dual-mutant Pu27-mer is dependent upon the site of mutation (e.g., 11 or 12 for G-to-T mutants at G14 or G11 and 20 or 21 for G-to-T mutants at G23 or G20). These photocleavage patterns are exactly those predicted from the known positions of mutations in these dual mutants. In the absence of additional data, it would appear that these cleavage patterns are indicative of binding of

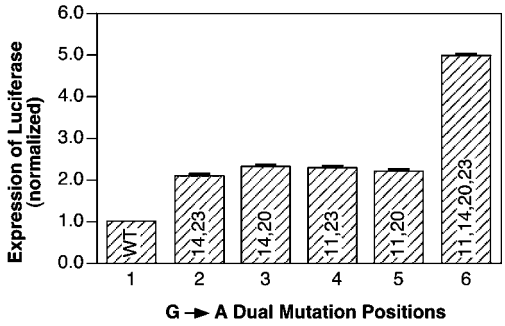

Figure 8. Effect of the four dual-loop mutants G→A-14,23, -14,20, -11,-23, and −11,20 (lanes 2−5, respectively) and tetra mutant (lane 6) on luciferase expression relative to the wild-type sequence (lane 1). The values are the average of three experiments. The wild-type sequence was assigned a nominal value of 1.0.

TMPyP4 to the two exterior faces of the parallel G-quadruplex structure. However, the pronounced changes in the CD upon addition of TMPyP4 to the Pu27-mer (Figure 6) are inconsistent with a parallel structure but are consistent with a mixed parallel/antiparallel structure. Since cleavage at both ends of the G-quadruplex structure cannot differentiate between an unchanged parallel structure and a loop inversion at both ends of the G-quadruplex, a similar experiment was carried out with an expanded porphyrin containing two selenium atoms, Se2SAP (5,10,15,20-[tetra-(*N*-methyl-3-pyridyl)]-26-28-diselena sapphyrin chloride), which binds only to one face of the G-quadruplex structure. In this case, cleavage occurred at G9 instead of G7, confirming the inversion of a one-base thymine loop from internal to external (unpublished results).

All Four Dual-Mutant Loop Isomers Are Biologically Relevant to c-MYC Transcriptional Control. We have previously reported that G-to-A mutations at either G12 or G17 resulted in a transcriptional activation of about 3-fold using a luciferase reporter assay.[17] To evaluate the transcriptional repressor role of each of the four loop isomers (see Figure 4A), the same dual mutants that dictate the individual loop isomers were evaluated for their effect on transcriptional activation using the same luciferase reporter system. In each case, an increase in basal gene expression of about 2.2−2.4 was seen (see Figure 8), implicating a role for a mixture of the loop isomers in stabilization of the parallel G-quadruplex structure. Mutation of all four guanines involved in the isomeric transitions (i.e., G11, G14, G20, and G23) produced a 5-fold increase in basal gene expression, further establishing the importance of these guanines in stabilization of the parallel G-quadruplex structure.

Discussion

We have previously demonstrated that the purine-rich strand of the NHE III_1 in the c-MYC promoter in the form of the G-quadruplex is a silencer element and that stabilization by TMPyP4 serves to inhibit c-MYC gene expression.[17,29] Conversely, specific G-to-A mutations at G12 and G17, which destabilize the G-quadruplex silencer element, result in a 3-fold increase in basal gene expression.[17] More recently, we have demonstrated that these specific mutations are found in 30% of colorectal tumors, which results in overexpression of c-MYC in these tumor cells.[30] For future work in understanding the

(29) Grand, C. L.; Han, H.; Muñoz, R. M.; Weitman, S.; Von Hoff, D. D.; Hurley, L. H.; Bearss, D. J. *Mol. Cancer Ther.* **2002**, *1*, 565−573.

structural basis for the drug-induced lowering of c-MYC and the effects of mutations on gene expression, it will be important to know the precise structures of both the naturally occurring G-quadruplex and its complex with drugs.

In a previous publication, we suggested a model for the G-quadruplex in the NHE III_1 such that a chair form of a G-quadruplex is the naturally occurring structure with which TMPyP4 interacts.[17] This was based on the DMS cleavage pattern and a comparison to proposed structures for TBA and T30695,[18−21] which have similar sequences. In this same paper, on the basis of mutational analysis and in combination with a promoter assay, we eliminated the possibility of a previously proposed basket structure[20] as being biologically relevant. In the present contribution, we have reexamined the DMS cleavage data and performed CD and additional mutational experiments, and the results compelled us to revise the proposed models of both the naturally occurring G-quadruplex in the NHE III_1 and its TMPyP4−G-quadruplex complex. Surprisingly, the naturally occurring structure is a dynamic mixture of four parallel G-quadruplex loop isomers. Upon incubation with TMPyP4, these structures appear to convert to a corresponding mixture of parallel/antiparallel G-quadruplex structures in which two molecules of TMPyP4 are most likely bound externally within the two external loops. There are six possible two-loop isomers; however, until firm structural data are obtained, which of these structures occurs can only be suggested, although modeling results suggest that a loop isomer with two adjacent external loops is most likely.

While the DMS cleavage pattern was largely consistent with our proposed chair structure, the partial, or unprotected, nature of some of the guanines (e.g., G11, G14, G20, and G23) proposed to be involved in tetrad formation was perplexing. Furthermore, subsequent literature[25] on the HIV aptamer T30695 and TBA, in which the CD showed different characteristics (i.e., parallel for T30695 and mixed parallel/antiparallel for TBA), was also perplexing and prompted us to reexamine our proposed models. A comparison of the Pu27-mer with these two aptamers showed that the Pu27-mer was similar to T30695 but not to TBA. Also, a comparison of the CD spectra of the Pu27-mer and T30695 with that of a known intramolecular parallel G-quadruplex in the heptad-tetrad structure was highly suggestive of a parallel, rather than an antiparallel, chair structure for both Pu27-mer and T30695 oligonucleotides. Furthermore, upon addition of TMPyP4, the CD of the Pu27-mer changed considerably to produce a signature of a mixed parallel/antiparallel structure, which clearly showed that our previous model, in which a chair structure was proposed for the substrate *and* product of the binding with TMPyP4, could not possibly be correct.

A full mutational analysis of the Pu27-mer provided further insight into the complexity of the problem. In accordance with our previous conclusions,[17] only the four consecutive 3′ runs of guanines were involved in the kinetically favored G-quadruplex structure. However, the lack of a G-quadruplex destabilizing effect of specific G-to-A mutations in runs of four guanines suggested inherent degeneracy within the system and the possibility of a dynamic equilibrium between different loop isomers. In these loop isomers, the 3′ and 5′ guanines from both these runs could assume either loop or tetrad positions. To test this postulate, four specific sets of dual-guanine mutants from the two runs of four guanines were designed and constructed. These Pu18-mers were shown by a combination of EMSA and one-dimensional ^{1}H NMR (unpublished results) to each give rise to single-loop isomers. For the four dual G-to-T mutant isomers (G→T-14,23; -14,20; -11,23; and -11,20), the numbers of bases in each loop are 1:2:1, 1:2:2, 2:1:1, and 2:1:2, respectively (Figure 4A). These loop patterns could be predicted on the basis of the individual dual mutants selected. Indeed, upon incubation with TMPyP4, the photoinduced cleavage patterns (Figure 7) in each of the dual mutants confirmed these proposed loop isomer types.

With the loop isomers having been defined through the dual-mutant studies, it was now possible to reconstruct in the wild-type sequence the dynamic state of individual guanines that can interchange between loop and tetrad positions. Thus, G11, G14, G20, and G23 can be either within a loop (cleavage by DMS) or part of a G-tetrad structure (no DMS cleavage). This explains the only partial protection of certain guanines in the Pu27-mer sequence following DMS cleavage[17] and, conversely, the ability of the four dual mutants to form separate, defined G-quadruplex structures. It is unlikely that all four isomers occur in equal proportions. The pronounced DMS cleavage of G14 and G23 suggests that these guanines are either within the loops (G14) or external to the quadruplex (G23), and therefore the A loop isomer shown in Figure 4A is probably predominant within the mixture. Indeed, the predominance of band A over band B in lane 1 of Figure 4B also supports this conclusion. Furthermore, the increased complexity of the TMPyP4 photoinduced cleavage can be explained if we assume that different loop isomers are trapped out before photocleavage takes place. Indeed, the DMS and TMPyP4 photoinduced cleavage products in the wild-type sequence[17] are an approximate composite of the corresponding cleavage products from each of the four dual-mutant sequences.

Taken alone, the TMPyP4 photoinduced cleavage data cannot differentiate between binding to the parallel G-quadruplex structure or an alternative structure in which one or two of the internal loops has undergone a strand inversion to form an external loop. However, the CD data are *only* consistent with the occurrence of such inversions. In order for both external faces of the G-quadruplex structure to be cleaved by TMPyP4, we have considered the possibility that two consecutive loop inversions at the ends of the G-quadruplex may occur such that the external loops are adjacent. Why does TMPyP4 interact with the looped-out structure rather than the parallel G-quadruplex? On the basis of comparative molecular modeling experiments with TMPyP4 and the parallel versus mixed parallel/antiparallel structure, the following observations were made. Without the external loop, stable binding is not achieved. The major interaction between the porphyrins and G-quadruplexes is through the positive charges on the pyridyl rings in porphyrins with the negative charges in the phosphate backbone of DNA (i.e., the porphyrins do not show any binding without the positive-charged meso-substituted rings).[31] In the parallel G-quadruplex structure, the internal loops are not available for interaction with the porphyrin meso cationic group. Therefore, to provide these interactions, TMPyP4 traps out a looped-out structure where the anionic backbone stabilizes the TMPyP4.

(30) Grand, C. L.; Powell, T. J.; Nagle, R. B.; Bearss, D. J.; Tye, D.; Gleason-Guzman, M.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6140−6145.

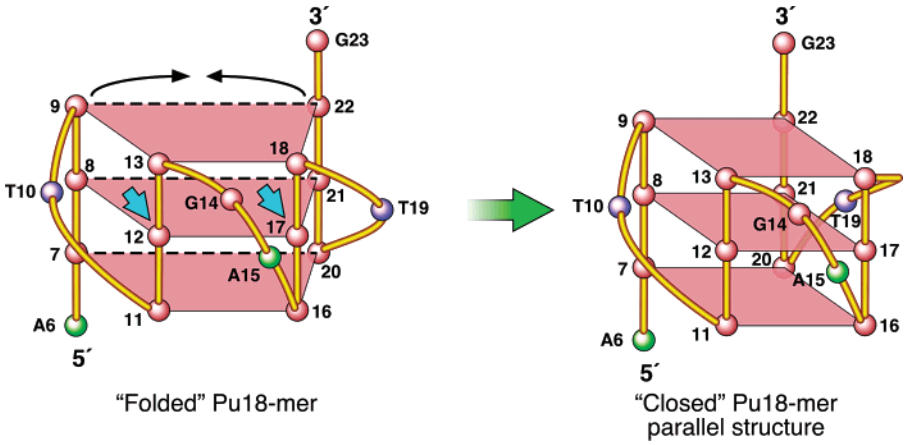

Figure 9. Folding and closure of the Pu18-mer to form the parallel-stranded structure. The blue arrows in the "folded" structure point to the G-to-A mutation sites that are selected in colorectal tumors.[30] Only the 1:2:1 loop isomer is shown as an example. Base colors: red = guanine, blue = thymine, green = adenine, and yellow = cytosine.

This new insight is important from at least three biological perspectives. First, the parallel structure of the G-quadruplex in the NHE III_1 now places the two classes of G-quadruplex mutants found in colorectal tumors in critical, equivalent side-by-side positions in the central tetrad, such that they are in pivotal positions in the initial folding scaffold prior to closure to form the G-quadruplex structure (Figure 9).[30] Second, all four parallel-loop isomers contribute to the stability of the silencer element in the NHE III_1 of the c-MYC promoter. Each dual-mutation isomer was found to contribute to a >2-fold increase in basal transcriptional activity. We propose it is this dynamic equilibrium between the loop isomers that provides the entropy and, thus, the stability (T_m = 85 °C) of this G-quadruplex structure. This proposal is further supported by the observation that each of the four dual isomers has a significantly lower (about 15 °C) melting temperature than the wild-type sequence. Third, the change in structure of the naturally occurring G-quadruplex from a parallel to a mixed parallel/antiparallel G-quadruplex in the presence of TMPyP4 provides an attractive hypothesis for how TMPyP4 might inhibit c-MYC transcription. Simply, the silencer element may no longer be recognized as a substrate for NM23-H2, which is required for transformation of this element to the transcriptionally active form.[32] While this manuscript was under review, an independent NMR study by Dinshaw Patel's group confirmed the assignment of the parallel G-quadruplex structure proposed here.[33]

Experimental Procedures

Materials for Biochemistry. Compound solutions were prepared as 1 mM stock solutions in distilled water and stored at −20 °C. These stock solutions were diluted to working concentrations immediately before use. Electrophoretic reagents (acrylamide/bisacrylamide solution and ammonium persulfate) were purchased from BioRad, and *N,N,N′,N′*-tetramethylethylenediamine (TEMED) was purchased from Fisher. *Taq* DNA polymerase, terminal transferase, and T4 polynucleotide kinase were purchased from Promega. [γ-^{32}P]ATP and [α-^{32}P]ATP were purchased from Amersham Biosciences and Perkin-Elmer.

Circular Dichroism Spectroscopy. The oligonucleotides TBA d[GGTTGGTGTGGTTGG], T30695 d[GGGTGGGTGGGTGGGT], and Pu27-mer d[TGGGGAGGGTGGGGAGGGTGGGGAAGG] were obtained PAGE purified from Sigma Genosys, dissolved in autoclaved double-distilled water (100 μM), and kept at −20 °C for no longer than 6 weeks. CD spectra were recorded on a Jasco-810 spectropolarimeter (Jasco, Easton, MD), using a quartz cell of 1 mm optical path length, an instrument scanning speed of 100 nm/min, with a response time of 1 s, and over a wavelength range of 200−330 or 200−600 nm for the titration experiments with drugs. All DNA samples were dissolved in Tris-HCl buffer (50 mM, pH 7.6) to a strand concentration of 10 μM, and where appropriate, the samples also contained 100 mM KCl. The compound was dissolved from previously prepared 10 mM stock solutions to a concentration of 0.5 mM with double-distilled water and titrated into the DNA sample at half mole equivalents up to a 4 mol excess. The CD spectra herein are representations of four averaged scans taken at 25 °C and are baseline corrected for signal contributions due to the buffer.

Preparation and End-Labeling of Oligonucleotides. Prior to the experiment, all oligonucleotides were treated in 10 mM NaOH for 30 min at 37 °C, followed by neutralization with 10 mM HCl and ethanol precipitation, in order to disrupt the self-associated structures. The 3′- or 5′-end-labeled single-strand oligonucleotides were obtained by incubating the oligomer with terminal transferase and [α-^{32}P]ATP or T4 polynucleotide kinase and [γ-^{32}P]ATP for 1 h at 37 °C. Labeled DNA was purified with a Bio-Spin 6 chromatography column (BioRad) after inactivation of the kinase by heating at 70 °C for 8 min.

Polymerase Stop Assay. The DNA primer d[TAATACGACTCACTATAGCAATTGCGTG], the Pu27 c-MYC template sequence (WT) d[TCCAACTATGTATAC(TGGGGAGGGTGGGGAGGGTGGGGAAGG)TTAGCGGCACGCAATTGCTATAGTGAGTCGTATTA], and the mutant sequences (G2−G27, where each of the guanines of the 27-mer insert in the template sequence was mutated with adenine) (Figure 3) were synthesized and purified as mentioned above. 5′-Labeled primer (100 μM) and template DNA (100 μM) were annealed in an annealing buffer (50 mM Tris-HCl, pH 7.5, 10 mM NaCl) by heating to 95 °C and then slowly cooled to room temperature. DNA formed by annealing the primer to the template sequence was purified using gel electrophoresis on a 12% native polyacrylamide gel. The purified DNA was then diluted to a concentration of 2 nM and mixed with the reaction buffer (10 mM $MgCl_2$, 0.5 mM DTT, 0.1 mM EDTA, and 1.5 μg/μL BSA) and 0.1 mM dNTP. KCl and NaCl (10 mM each) were added to the reaction. *Taq* DNA polymerase was added, and the mixture was incubated at 37 °C for 20 min. The polymerase extension was stopped by adding 2× stop buffer (10 mM EDTA, 10 mM NaOH, 0.1% xylene cyanole, and 0.1% bromophenol blue in formamide solution) and loaded onto a 16% denaturing gel.

(31) Han, H.; Langley, D. R.; Rangan, A.; Hurley, L. H. *J. Am. Chem. Soc.* **2001**, *123*, 8902−8913.
(32) Postel, E. H. *J. Bioenerg. Biomembr.* **2003**, *35*, 31−40.
(33) Phan, A. T.; Modi, Y. S.; Patel, D. J. *J. Am. Chem. Soc.* **2004**, *126*, 8710−8716.

Electromobility Shift Assay and DMS Footprinting. The 18-mer sequences d[AGGGTGGGGAGGGTGGGG] (MYC18), d[AGGGTGGGTAGGGTGGGT] (G→T-14,23), d[AGGGTGGGTAGGGTTGGG] (G→T-14,20), d[AGGGTTGGGAGGGTGGGT] (G→T-11,23), and d[AGGGTTGGGAGGGTTGGG] (G→T-11,20) were 3′-end-labeled and diluted to a final concentration of 2 μM. Each of the 3′-end-labeled sequences was heated to 95 °C and slowly cooled to room temperature in a buffer (100 mM Tris-HCl/1 mM EDTA, pH 7.5, 100 mM KCl). They were further incubated for 2 h at room temperature and then subjected to preparative gel electrophoresis (15% polyacrylamide, 25 mM KCl, 20 h, 4 °C). Each band of interest was excised and soaked in 300 μL of buffer (10 mM Tris-HCl/1 mM EDTA, pH 7.5, 100 mM KCl) for 6 h at room temperature. The solutions were filtered (microcentrifuged), and 50 000 cpm (per reaction) of DNA solution was then used to perform the DMS footprinting as outlined in ref 17.

Photomediated Strand-Cleavage Reaction. The 27-mer sequences d[TGTTGAGGGTGGGTAGGGTGGGTAAGG] (G→T-14,23), d[TGTTGAGGGTGGGTAGGGTTGGGAAGG] (G→T-14,20), d[TGTTGAGGGTTGGGAGGGTGGGTAAGG] (G→T-11,23), and d[TGTTGAGGGTTGGGAGGGTTGGGAAGG] (G→T-11,20) were 5′-end-labeled and diluted to a final concentration of 2 μM. Each of the 5′-end-labeled sequences was heated to 95 °C and slowly cooled to room temperature in a buffer (100 mM Tris-HCl/1 mM EDTA, pH 7.5, 100 mM KCl). For each reaction, 50 000 cpm of DNA solution was diluted further with the buffer (10 mM Tris/1 mM EDTA, pH 7.5, 100 mM KCl). A 1 μM final concentration of TMPyP4 was added, the solution was incubated at room temperature for 2 h in the absence of light, and then it was transferred to a 24-well Titertek microtiter plate (ICN). This plate was placed on top of a Pyrex glass shield and irradiated for 2 h with an 85 W xenon lamp placed under the Pyrex glass. Pyrex glass was used to filter the UV light under 300 nm, thereby eliminating DNA damage caused directly by UV irradiation. During the irradiation, the Titertek plate was rotated three times to eliminate light heterogeneity. Reactions were terminated by the addition of 10 μg of calf thymus DNA, followed by ethanol precipitation. The resulting samples were subjected to treatment with 0.1 M piperidine. The samples were then loaded onto a 19% sequencing gel.

Imaging and Quantification. The dried gels were exposed on a phosphor screen. Imaging and quantification were performed using a PhosphorImager (Storm 820) and ImageQuant 5.1 software from Amersham Biosciences.

Luciferase Reporter Assays. Site-directed mutagenesis was performed to create site-specific mutations in the c-MYC promoter using the protocol described previously.[17] These reporter vectors were transfected into HeLa S3 cells, and luciferase activity was detected using the assay described in ref 17.

Molecular Modeling. The X-ray crystal structure coordinates for the human telomeric sequence, d[$AG_3(T_2AG_3)_3$], from the PDB data entry (1KF1)[34] were used as a starting model for the construction of Pu27 c-MYC parallel G-quadruplex. Necessary replacements and deletions of bases were carried out. Remaining bases were added using standard B-DNA geometry using the Biopolymer module of Insight II 2000.1 (Accelrys Inc., San Diego, CA).[35] The modeled structure was then subjected to a series of minimization and molecular dynamics protocols.

Modeled Pu27 c-MYC quadruplex structures were then subjected to energy refinement. All polar hydrogen atoms were energy minimized using 2500 steps of conjugate gradient minimizer using Discover 3.0 using the Amber force field.[36] This was followed by subjecting the entire structural model to 2 × 2500 steps of conjugate gradient minimization.

Fully refined quadruplex structure was then neutralized by the addition of an appropriate number of sodium ions, and this structure was immersed in a box of TIP3P water molecules.[37] Two-stage molecular dynamics simulations were performed at 300 K. This involved restrained MD simulations with 20 ps equilibration and 100 ps simulations. Distances and angles for hydrogen bonds involving G-quadruplex tetrad bases were restrained by means of the upper-bound harmonic restraining function with a force constant of 10 kcal mol^{-1} $Å^{-2}$ for distances and 30 kcal mol^{-1} rad^{-2} for angles. The second run of MD simulations involved unrestrained simulations with 20 ps equilibration and 100 ps simulations at 300 K. The most stable low-energy structural model was then refined using 2500 steps of conjugate gradient minimization.

Acknowledgment. This research was supported by grants from the National Institutes of Health (CA88310, CA94166, and IS10RR16659). We thank Dinshaw Patel for communicating his results prior to publication, and we are grateful to David Bishop for preparing, proofreading, and editing the final version of the manuscript and figures. This manuscript is dedicated to Professor Heinz G. Floss on the occasion of his 70th birthday.

JA040022B

(34) Parkinson, G. N.; Lee, Michael, P. H.; Neidle; S. *Nature* **2002**, *417*, 876−880.
(35) Insight II 2000 Molecular Modeling Software, Accelrys Inc., 9685 Scranton Rd., San Diego, CA, 92121.
(36) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(37) Jorgensen, W. L.; Chandresekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.