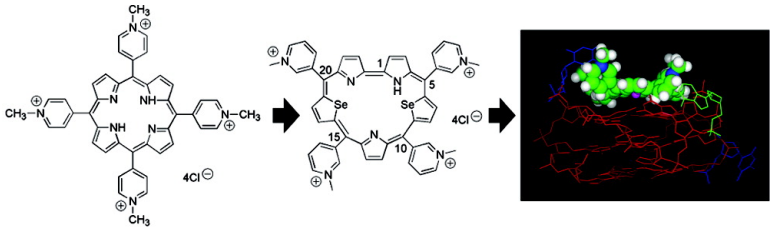
Article

Design and Synthesis of an Expanded Porphyrin That Has Selectivity for the c-MYC G-Quadruplex Structure

Jeyaprakashnarayanan Seenisamy, Sridevi Bashyam, Vijay Gokhale, Hariprasad Vankayalapati, Daekyu Sun, Adam Siddiqui-Jain, Nicole Streiner, Kazuo Shin-ya, Elizabeth White, W. David Wilson, and Laurence H. Hurley

J. Am. Chem. Soc., **2005**, 127 (9), 2944-2959• DOI: 10.1021/ja0444482 • Publication Date (Web): 05 February 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 23 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 02/05/2005

Design and Synthesis of an Expanded Porphyrin That Has Selectivity for the c-MYC G-Quadruplex Structure

Jeyaprakashnarayanan Seenisamy,[†] Sridevi Bashyam,[†] Vijay Gokhale,[†] Hariprasad Vankayalapati,[†] Daekyu Sun,[†] Adam Siddiqui-Jain,[‡] Nicole Streiner,[‡] Kazuo Shin-ya,[§] Elizabeth White,[||] W. David Wilson,[||] and Laurence H. Hurley*[,†,⊥,▽]

Contribution from the College of Pharmacy, The University of Arizona, 1703 East Mabel, Tucson, Arizona 85721, Arizona Cancer Center, 1515 North Campbell Avenue, Tucson, Arizona 85724, Department of Chemistry, The University of Arizona, Tucson, Arizona 85721, Cylene Pharmaceuticals, 11045 Roselle Street, Suite C, San Diego, California 92121, Institute of Molecular and Cellular Biosciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan, and Department of Chemistry, Georgia State University, Atlanta, Georgia 30303

Received September 13, 2004; E-mail: hurley@pharmacy.arizona.edu

Abstract: Cationic porphyrins are known to bind to and stabilize different types of G-quadruplexes. Recent studies have shown the biological relevance of the intramolecular parallel G-quadruplex as a transcriptional silencer in the c-MYC promoter. TMPyP4 also binds to this G-quadruplex and most likely converts it to a mixed parallel/antiparallel G-quadruplex with two external lateral loops and one internal propeller loop, suppressing c-MYC transcriptional activation. To achieve therapeutic selectivity by targeting G-quadruplexes, it is necessary to synthesize drugs that can differentiate among the different types of G-quadruplexes. We have designed and synthesized a core-modified expanded porphyrin analogue, 5,10,15,20-[tetra(*N*-methyl-3-pyridyl)]-26,28-diselenasapphyrin chloride (Se2SAP). Se2SAP converts the parallel c-MYC G-quadruplex into a mixed parallel/antiparallel G-quadruplex with one external lateral loop and two internal propeller loops, resulting in strong and selective binding to this G-quadruplex. A *Taq* polymerase stop assay was used to evaluate the binding of TMPyP4 and Se2SAP to G-quadruplex DNA. Compared to TMPyP4, Se2SAP shows a greater selectivity for and a 40-fold increase in stabilization of the single lateral-loop hybrid. Surface plasmon resonance and competition experiments with duplex DNA and other G-quadruplexes further confirmed the selectivity of Se2SAP for the c-MYC G-quadruplex. Significantly, Se2SAP was found to be less photoactive and noncytotoxic in comparison to TMPyP4. From this study, we have identified an expanded porphyrin that selectively binds with the c-MYC G-quadruplex in the presence of duplex DNA and other G-quadruplexes.

Introduction

Emerging evidence for the involvement of G-quadruplex structures in cellular processes such as transcriptional control of c-MYC[1] has stimulated the development of drugs that have selectivity for different G-quadruplex structures. The promoter regions of some important genes, such as the insulin gene,[2a] c-MYC,[1,2b] PDGF,[2c] HER-2/neu,[2d] c-MYB,[2e] the human and chicken β-globin genes,[2f,g] the rat preproinsulin II gene,[2f] adenovirus serotype 2,[2h] and retinoblastoma susceptibility genes,[2i] have been found to contain sequences that have the ability to form G-quadruplex structures under physiological conditions. Direct evidence for the involvement of G-quadruplexes in transcriptional control has been obtained only for c-MYC and the insulin gene.[1,2a] Due to the structural polymorphism in G-quadruplex structures, different G-quadruplex topologies may be associated with different signaling pathways.[3,4] The intramolecular G-quadruplex structures (Figure 1A,B) observed in the human telomeric sequence d[TTAGGG]$_4$

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

(1) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593−11598.
(2) (a) Castasi, P.; Chen, X.; Moyzis, R.; Bradbury, E.; Gupta, G. *J. Mol. Biol.* **1996**, *264*, 534−545. (b) Simonsson, T.; Pecinka, P.; Kubista, M. *Nucleic Acids Res.* **1998**, *26*, 1167−1172. (c) Rezler, E.; Hurley, L. H. *Abstracts of the 94th Annual Meeting*; American Association for Cancer Research: Philadelphia, 2003; Vol. 44, pp 1505−1506. (d) Basye, J.; Memmott, R.; Hurley, L. H.; Ebbinghaus, S. W. *Abstracts of the 94th Annual Meeting*; American Association for Cancer Research: Philadelphia, 2003; Vol. 44, p 1283. (e) Memmott, R.; Basye, J.; Hurley, L. H.; Ebbinghaus, S. W. *Abstracts of the 94th Annual Meeting*; American Association for Cancer Research: Philadelphia, 2003; Vol. 44, p 1282. (f) Evans, T.; Schon, E.; Gora-Maslak, G.; Patterson, J.; Efstratiadis, A. *Nucleic Acids Res.* **1984**, *12*, 8043−8058. (g) Howell, R. M.; Woodford, R. M.; Weitzmann, R. M.; Usdin, K. *J. Biol. Chem.* **1996**, *271*, 5208−5214. (h) Kilpatrick, M. W.; Torri, A.; Kang, D. S.; Engler, J. A.; Wells, R. D. *J. Biol. Chem.* **1986**, *261*, 11350−11354. (i) Murchie, A. I.; Lilley, D. M. *Nucleic Acids Res.* **1992**, *20*, 49−53.
(3) Arthanari, H.; Bolton, P. H. *Chem. Biol.* **2001**, *8*, 221−230.
(4) (a) Han, H.; Hurley, L. H. *Trends Pharm. Sci.* **2000**, *21*, 136−142. (b) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature (London)* **2002**, *417*, 876−880. (c) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263−282.

10.1021/ja0444482 CCC: $30.25

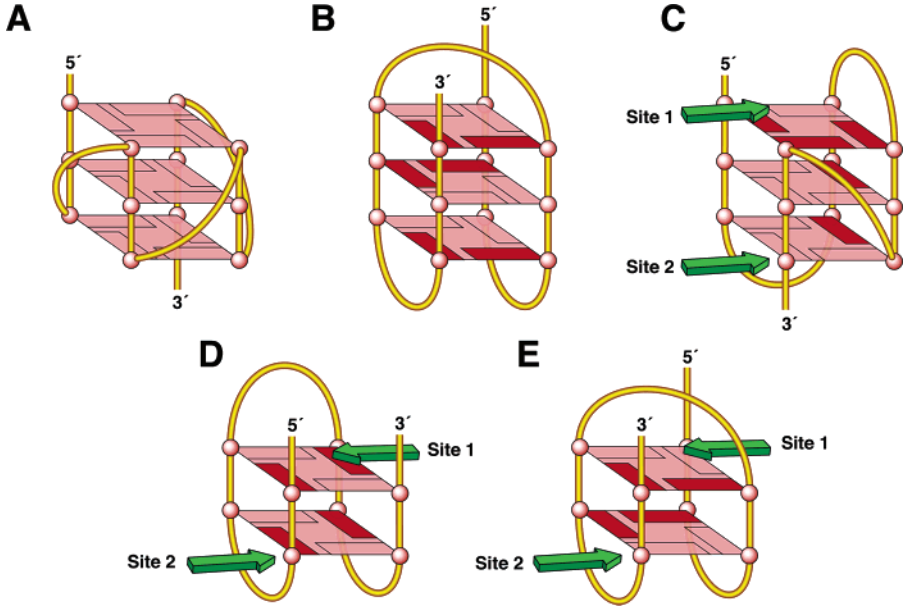


Figure 1. Structural polymorphism in G-quadruplex structures: (A) parallel propeller-type G-quadruplex, (B) basket-type G-quadruplex (HumTel), (C) double-loop hybrid G-quadruplex, (D) chair-type G-quadruplex, (E) basket-type G-quadruplex (G_2T_4). Sites 1 and 2 are alternative ligand-binding sites (see the text).

(HumTel) are proposed to inhibit telomerase activity by sequestration of the substrate required for enzyme activity.[4,5] Telomerase is an important enzyme involved in telomere maintenance in tumor cells, and telomerase inhibition by targeting G-quadruplex leads to telomere shortening and senescence in tumor cells.[5,6] The solution structure of the *Tetrahymena* telomeric DNA sequence d[TTGGGG]$_4$ (TetTel) forms a mixed parallel/antiparallel G-quadruplex with three tetrads and two external lateral loops[7a] (here termed a "double-loop hybrid") (Figure 1C). A thrombin-binding aptamer (TBA), having the sequence d[GGTTGGTGTGGTTGG], has been characterized as a chair-type G-quadruplex in both solution and the solid state (Figure 1D).[7b−f] Also, an independent sequence, d[(GGTTTT)$_3$GG] (G_2T_4), forms a basket-type G-quadruplex (Figure 1E).[7g,h]

Overexpression of c-MYC results in increased cellular proliferation in a variety of different malignancies, including breast, colon, cervix, small-cell lung, osteosarcomas, glioblastomas, and myeloid leukemias.[8] The nuclease hypersensitivity element (NHE) III$_1$ upstream of the P1 promoter of c-MYC controls up to 90% of the transcriptional activation of this gene.[1,2,9] Although the purine-rich strand of the DNA in this region can form different types of intramolecular parallel G-quadruplex structures, only the G-quadruplexes formed with four consecutive 3′ runs of guanines are biologically relevant, functioning as a transcriptional repressor element.[1] Also, the biologically relevant G-quadruplexes exist as a mixture of four different loop isomers, among which the 1:2:1 isomer exists as a major form[10a] (Figure 2A). The folding pattern of this parallel G-quadruplex structure has been defined by NMR.[10b] Upon interaction with TMPyP4 (Figure 3), the parallel G-quadruplex is converted into a mixed parallel/antiparallel G-quadruplex structure.[1,10a] We have demonstrated that c-MYC transcription can be controlled by stabilizing the G-quadruplex structure using TMPyP4.[1,10c] Cationic porphyrins are known to bind to and stabilize different types of G-quadruplexes and, in some cases, to facilitate G-quadruplex formation.[1,11,12] The interaction of

(5) (a) Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S.; Hurley, L. H. *J. Med. Chem.* **1997**, *40*, 2113-2116. (b) Perry, P. J.; Read, M. A.; Davies, R. T.; Gowan, S. M.; Reszka, A. P.; Wood, A. A.; Kelland, L. R.; Neidle, S. *J. Med. Chem.* **1999**, *42*, 2679−2684. (c) Gowan, S. M.; Heald, R.; Stevens, M. F. G.; Kelland, L. R. *Mol. Pharmacol.* **2001**, *60*, 981−988. (d) Mergny, J.-L.; Lacroix, L.; Teulade-Fichou, M.-P.; Hounsou, C.; Guittat, L.; Hoarau, M.; Arimondo, P. B.; Vigneron, J.-P.; Lehn, J.-M.; Riou, J.-F.; Garestier, T.; Hélène, C. *Proc. Natl. Acad Sci. U.S.A.* **2001**, *98*, 3062−3067. (e) Riou, J. F.; Guittat, L.; Mailliet, P.; Laoui, A.; Renou, E.; Petitgenet, O.; Megnin-Chanet, F.; Hélène, C.; Mergny, J.-L. *Proc. Natl. Acad Sci. U.S.A.* **2002**, *99*, 2672−2677. For recent reviews see (f) Rezler, E. M.; Bearss, D. J.; Hurley, L. H. *Curr. Opin. Pharmacol.* **2002**, *2*, 415−423. (g) Rezler, E. M.; Bearss, D. J.; Hurley, L. H. *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 359−379. (h) Hurley, L. H. *Nat. Rev. Cancer* **2002**, *2*, 188−200. (i) Neidle, S.; Parkinson, G. *Nat. Rev. Drug Discovery* **2002**, *1*, 383−393. (j) Mergny, J.-L.; Riou, J.-F.; Mailliet, P.; Teulade-Fichou, M.-P.; Gilson, E. *Nucleic Acids Res.* **2002**, *30*, 839−865.

(6) (a) Bryan, T. M.; Cech., T. R. *Curr. Opin. Cell Biol.* **1999**, *11*, 318−324. (b) Hahn, W. C.; Stewart, S. A.; Brooks, M. W.; York, S. G.; Eaton, E.; Kurachi, A.; Beijersbergen, R. L.; Knoll, J. H. M.; Meyerson, M.; Weinberg, R. A. *Nat. Med.* **1999**, *5*, 1164−1170. (c) de Lange, T.; Jacks, T. *Cell* **1999**, *96*, 273−275.

(7) (a) Wang, Y.; Patel, D. J. *Structure* **1994**, *2*, 1141−1156. (b) Schultze, P.; Macaya, R. F.; Feigon, J. *J. Mol. Biol.* **1994**, *235*, 1532−1547. (c) Kelly, J. A.; Feigon, J.; Yeates, T. O. *J. Mol. Biol.* **1996**, *256*, 417−422. (d) Jing, N. J.; Gao, X. L.; Rando, R. F.; Hogan, M. E. *J. Biomol. Struct. Dyn.* **1997**, *15*, 573−585. (e) Jing, N. J.; Hogan, M. E. *J. Biol. Chem.* **1998**, *273*, 34992−34999. (f) Marathias, V. M.; Bolton, P. H. *Nucleic Acids Res.* **2000**, *28*, 1969−1977. (g) Dapic, V.; Abdomerovic, V.; Marrington, R.; Peberdy, J.; Rodger, A.; Trent, J. O.; Bates, P. J. *Nucleic Acids Res.* **2003**, *31*, 2097−2107. (h) Marathias, V. M.; Bolton, P. H. *Biochemistry* **1999**, *38*, 4355−4364.

(8) (a) Pelengaris, S.; Rudolph, B.; Littlewood, T. *Curr. Opin. Genet. Dev.* **2000**, *10*, 100−105. (b) Spencer, C. A.; Groudine, M. *Adv. Cancer Res.* **1991**, *56*, 1−48. (c) Marcu, K. B.; Bossone, S. A.; Patel, A. J. *Annu. Rev. Biochem.* **1992**, *61*, 809−860. (d) Facchini, L. M.; Penn, L. Z. *FASEB J.* **1998**, *12*, 633−651.

(9) (a) 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. (b) Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *214*, 456−459. (c) Siebenlist, U.; Henninghausen, L.; Battey, J.; Leder, P. *Cell* **1984**, *37*, 381−391. (d) Tomonaga, T.; Levens, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5830−5835. (e) Bossone, S. A.; Asselin, C.; Patel, A. J.; Marcu, K. B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7452−7456. (f) Postel, E. H.; Berberich, S. J.; Rooney, J. W.; Kaetzel, D. M. *J. Bioenerg. Biomembr.* **2000**, *32*, 277−284.

(10) (a) Seenisamy, J.; Rezler, E. M.; Powell, T. J.; Tye, D.; Gokhale, V.; Joshi, C. S.; Siddiqui-Jain, A.; Hurley, L. H. *J. Am. Chem. Soc.* **2004**, *126*, 8702−8709. (b) Phan, A. T.; Modi, Y. S.; Patel, D. J. *J. Am. Chem. Soc.* **2004**, *126*, 8710−8716. (c) 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.

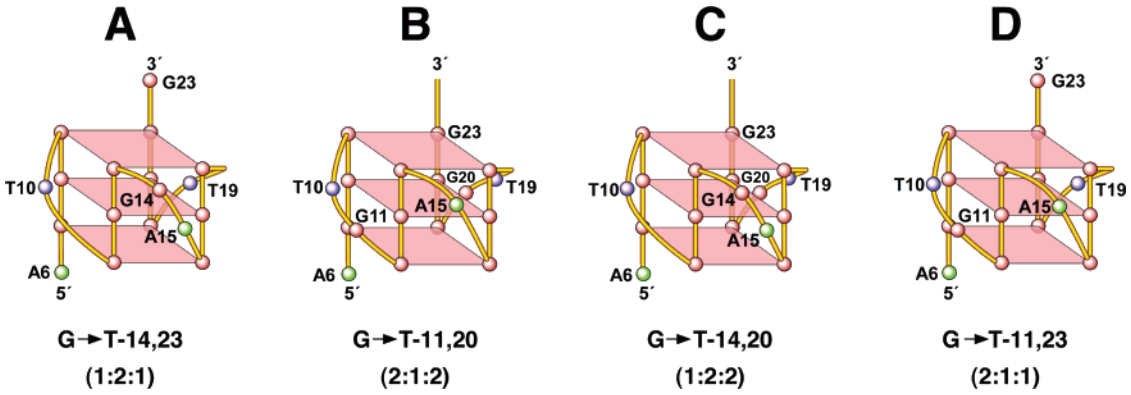


Figure 2. Proposed structures of the four different loop isomers, in which dual G-to-T mutations at positions 11, 14, 20, and 23 result in defined loop isomers. Below each isomer are the appropriate mutations (e.g., G → T-14,23) that result in the defined isomer and the loop size arrangement (e.g., 1:2:1).

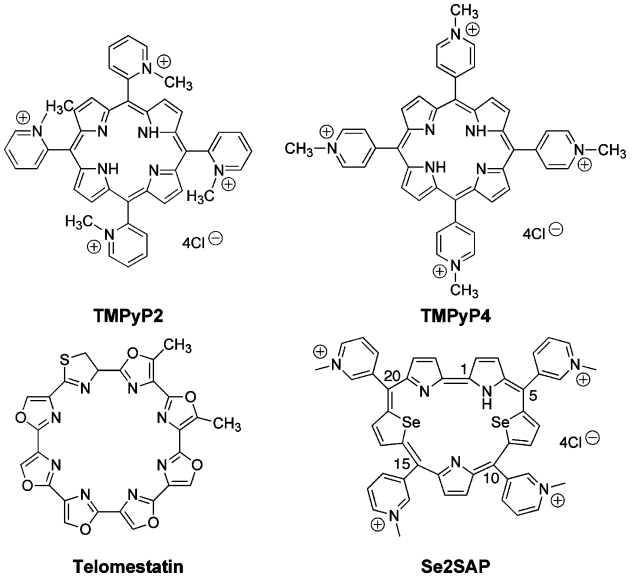


Figure 3. Structures of the G-quadruplex-interactive compounds TMPyP2, TMPyP4, telomestatin, and Se2SAP.

TMPyP4 with the intramolecular basket-type G-quadruplex formed in the human telomeric sequence results in telomerase inhibition.[12a,b,f] In a previous study, we explored the recognition of the telomeric basket-type structure by analogues of TMPyP4 and subsequent inhibition of telomerase activity.[12a] It is critical to design drugs that can differentiate among the different types of G-quadruplex structures to achieve therapeutic selectivity by targeting specific G-quadruplexes.

Telomestatin (Figure 3), a natural product isolated from *Streptomyces anulatus* 3533-SV4,[13a] has a larger ring system than TMPyP4 and is known to bind strongly to and stabilize intramolecular basket-type G-quadruplex structures.[13b,c] Porphyrins are known to produce photoinduced cytotoxicity.[14] The substitution of selenium for nitrogen in the porphyrin core can eliminate this photocytotoxicity.[15] On the basis of modeling and comparative analysis of the binding of TMPyP4 and telomestatin to G-quadruplex, a selenium-substituted expanded porphyrin, Se2SAP (Figure 3), was designed and synthesized. It has been subsequently demonstrated to have dramatically improved selectivity for an induced parallel/antiparallel G-quadruplex structure with one external lateral loop (here termed a "single-loop hybrid") in comparison to both duplex DNA and other G-quadruplex structures.

Results

Design of a Core-Modified Expanded Porphyrin. A previous study from our laboratory showed the differential binding of positional isomers of the cationic porphyrins TMPyP4 and TMPyP2 [5,10,15,20-tetra(*N*-methyl-2-pyridyl)porphine] with parallel and antiparallel intermolecular G-quadruplex structures.[11] TMPyP4 showed selectivity for intermolecular rather than intramolecular G-quadruplexes.[11] Recent studies with telomestatin showed preferential interaction with the intramolecular over the intermolecular G-quadruplexes.[13c] The strong and selective binding of telomestatin with intramolecular G-quadruplexes may be attributed to its larger ring system, which overlaps completely with the four guanines in the G-tetrad and allows the molecule to occupy the whole area created between the loop and the top tetrad region.[13b] Because of the smaller ring size of TMPyP4 (10.1 Å), and the positioning of TMPyP4 above the G-tetrad (12.06 Å), it produced an overlap with only two of the four guanines. Also, the center of mass of TMPyP4 is displaced from the G-quadruplex helical axis by ~2.5 Å.[11] On the basis of these observations, the design and synthesis of a larger porphyrin ring system could be anticipated

(11) Han, H.; Langley, D. R.; Rangan, A.; Hurley, L. H. *J. Am. Chem. Soc.* **2001**, *123*, 8902−8932.
(12) (a) Shi, D. F.; Wheelhouse, R. T.; Sun, D.; Hurley, L. H. *J. Med. Chem.* **2001**, *44*, 4509−4523. (b) Wheelhouse, R. T.; Sun, D.; Han, H.; Han, F. X.; Hurley, L. H. *J. Am. Chem. Soc.* **1998**, *120*, 3261−3262. (c) Han, H.; Cliff, C. L.; Hurley, L. H. *Biochemistry* **1999**, *38*, 6981−6986. (d) Han, F. X.; Wheelhouse, R. T.; Hurley, L. H. *J. Am. Chem. Soc.* **1999**, *121*, 3561−3570. (e) Anantha, N. V.; Azam, M.; Sheardy, R. D. *Biochemistry* **1998**, *37*, 2709−2714. (f) Izbicka, E.; Wheelhouse, R. T.; Raymond, E.; Davidson, K. K.; Lawrence, R. A.; Sun, D.; Windle, B. E.; Hurley, L. H.; Von Hoff, D. D. *Cancer Res.* **1999**, *59*, 639−644.
(13) (a) Shin-ya, K.; Wierzba, K.; Matsuo, K.; Ohtani, T.; Yamada, Y.; Furihata, K.; Hayakawa, Y.; Seto, H. *J. Am. Chem. Soc.* **2001**, *123*, 1262−1263. (b) Kim, M.; Vankayalapati, H.; Shin-ya, K.; Wierzba, K.; Hurley, L. H. *J. Am. Chem. Soc.* **2002**, *124*, 2098−2099. (c) Kim, M.; Guzman, M. G.; Izbicka, E.; Nishioka, D.; Hurley, L. H. *Cancer Res.* **2003**, *63*, 3247−3256.
(14) (a) Sessler, J. L.; Weghorn, S. J. *Expanded, Contracted & Isomeric Porphyrins*; Tetrahedron Organic Chemistry Series, Vol. 15; Pergamon: New York, 1997. (b) Croke, D. T.; Perrouault, L.; Sari, M. A.; Battioni, J. P.; Mansuy, D.; Hélène, C.; Le Doan, T. *J. Photochem. Photobiol., B* **1993**, *18*, 41−50. (c) Nussbaum, J. M.; Newport, M. E.; Mackie, M.; Leontis, N. B. *Photochem. Photobiol.* **1994**, *59*, 515−528. (d) Armitage, B. *Chem. Rev.* **1998**, *98*, 1171−1200. (e) Pushpan, S. K.; Venkatraman, S.; Anand, V. G.; Sankar, J.; Parmeswaran, D.; Ganesan, S.; Chandrashekar, T. K. *Curr. Med. Chem. Anti-Cancer Agents* **2002**, *2*, 187−207. (f) Patito, I. A.; Rothmann, C.; Malik, Z. *Biol. Cell* **2001**, *93*, 285−291.
(15) (a) Stilts, C. E.; Nelen, M. I.; Hilmey, D. G.; Davies, S. R.; Gollnick, S. O.; Oseroff, A. R.; Gibson, S. L.; Hilf, R.; Detty, M. R. *J. Med. Chem.* **2000**, *43*, 2403−2410. (b) Hilmey, D. G.; Abe, M.; Nelen, M. I.; Stilts, C. E.; Baker, G. A.; Baker, S. N.; Bright, F. V.; Davies, S. R.; Gollnick, S. O.; Oseroff, A. R.; Gibson, S. L.; Hilf, R.; Detty, M. R. *J. Med. Chem.* **2002**, *45*, 449−461.

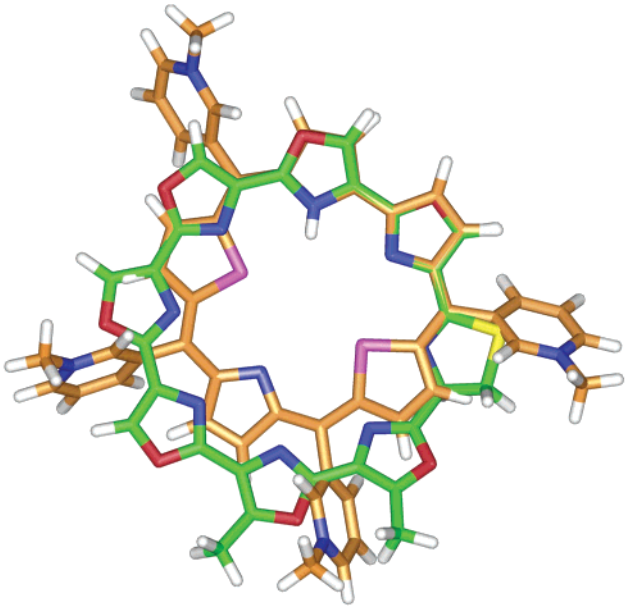


Figure 4. View of the field-fit alignment model of telomestatin and Se2SAP (color-by-atom). Telomestatin is represented in green and Se2SAP in orange.

to produce a compound with selectivity for intramolecular over intermolecular G-quadruplex structures. Although we have demonstrated that TMPyP4 has in vivo antitumor activities,[10c] it is potentially problematic in that it induces long-lasting skin toxicity, due to the retention of porphyrin molecules in cutaneous tissues.[14,15] This skin toxicity is related to the photoactivity of porphyrins and in principle can be eliminated by designing and synthesizing nonphotoactive porphyrins. Literature reports suggest that the introduction of heavy atoms such as selenium or sulfur into the porphyrin core greatly reduces the quantum yield for fluorescence.[15] Also, selenium analogues have not shown photosensitization in animal studies.[15] Consequently, Se2SAP was designed and synthesized.

Structural Comparison of Se2SAP and Telomestatin. The structural similarity between telomestatin and Se2SAP was confirmed by atom-by-atom superimposition and the electrostatic field-fit alignment method. The results from these studies show a root-mean-square deviation of 0.13 Å, indicating a considerable degree of similarity at the oxazole ring of telomestatin and the bipyrrole ring of Se2SAP (Figure 4). The measured distance between carbon atoms of the two opposite pyridyl rings attached at the meso position of Se2SAP is 11.08 Å, which is comparable to the distances (10.63−12.06 Å) between the two hydrogen atoms of the non-hydrogen-bonding G-NH2 functional group of the opposite guanine bases in the G-tetrad. Consequently, it was anticipated that Se2SAP would overlay very well with the entire G-tetrad in a manner similar to that of telomestatin,[4b] which should lead to strong binding of Se2SAP to intramolecular G-quadruplex structures.

Synthesis and Characterization of Se2SAP. In general, selenophene was lithiated in the 2 and 5 positions and converted to the corresponding diols by reaction with pyridyl aldehydes (Scheme 1).[15,16] The yields of each diol were dependent upon the aldehyde used (35% for 3-pyridyl-substituted diol and 20% for 4-pyridyl-substituted diol). The diol **1** was then condensed with pyrrole in the presence of acid to obtain the porphyrin **2** and demethylsapphyrin **3** in the yields 3% and 1%.[17] The presence of pyridyl rings restricted us to use only propionic acid. The use of other acids, such as TFA, $BF_3{\cdot}OEt_2$, and acetic acid, did not yield any porphyrins. The difficult task resided in purification of this minor product **3** from the more abundant oligomeric polypyrroles and porphyrin **2**. Repeated purification with preparative TLC was required since the R_f values for both **2** and **3** were very close. Subsequently purified **3** was subjected to further methylation and ion exchange with chloride ions to obtain Se2SAP. Initially, we attempted to synthesize 5,10,15,-20-[tetra(*N*-methyl-4-pyridyl)]-26,28-diselenasapphyrin chloride. Condensation of 2,5-bis[(4-pyridyl)hydroxymethyl]selenophene with pyrrole in propionic acid did not yield the expected diselenasapphyrin. Consequently, 3-pyridyl-substituted sapphyrin Se2SAP was synthesized.

^{1}H NMR showed a 2-fold symmetry at room temperature due to the NH tautomerism in the bipyrrole ring.[16] The Soret and Q bands in the absorption spectra are shifted to red compared to those of TMPyP4, which is due to the extended conjugation in the 22 π ring[16] of Se2SAP.

G-Quadruplex Binding of Se2SAP. A polymerase stop assay can be used to evaluate the relative stabilization of intramolecular G-quadruplex structures in the presence of ligands.[18] (The principle of this assay is shown to the left of the gel in Figure 5A.) DNA extension by *Taq* polymerase was paused at the G-quadruplex-forming site, the extent of which can be used as a measure of the stabilization of G-quadruplex structures by various ligands. The relative binding affinity of porphyrins to the intramolecular G-quadruplex structure was determined by incubating increasing concentrations of each porphyrin or telomestatin with a DNA template containing the Pu27 c-MYC sequence d[TGGGGAGGGTGGGGAGGGTGGGGAAGG] at 55 °C in the presence of *Taq* polymerase. Since there was a considerable amount of pausing (attributable to the presence of a K^+ ion in the buffer that facilitates G-quadruplex formation) in the absence of drug, we used an elevated temperature (55 °C) to partially destabilize the arresting G-quadruplex structure and thereby permit a larger window for drug stabilization of the G-quadruplex structure. The G-quadruplex-binding ability of Se2SAP was compared with that of TMPyP4 and telomestatin.

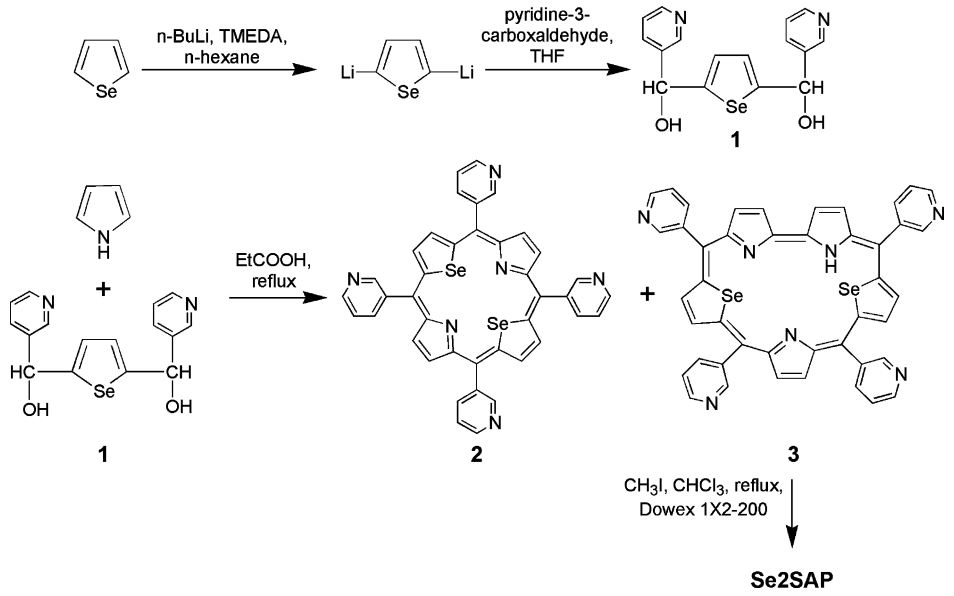
Figure 5A shows the concentration-dependent inhibition of *Taq* polymerase DNA synthesis by ligand stabilization of the G-quadruplex structure formed in the DNA templates containing the Pu27 c-MYC sequence. For each ligand there is significantly greater pausing at the G-quadruplex site in the presence of increasing concentrations of these compounds. For the Pu27 c-MYC sequence, there is pausing at the primer position in the absence of drug, which is due to the presence of nonpolymerase extendable free primer, even after purification. These results are quantified in Figure 5B,C. The pausing at the G-quadruplex-forming site is quantified as the percentage of normalized stop product with respect to the total intensity per lane. The enhanced pausing of polymerase processing at the primer position in the presence of higher concentrations of compound is most likely

(16) (a) Seenisamy, J.; Sridevi, B.; Chandrashekar, T. K.; Vij, A.; Roy, R. *J. Am. Chem. Soc.* **1999**, *121*, 9053−9068. (b) Seenisamy, J.; Bashyam, S.; Chandrashekar, T. K.; Vij, A.; Roy, R. *Angew. Chem., Int. Ed.* **1998**, *37*, 3394−3397.

(17) Alder, A. D.; Longo, F. R.; Finarelli, J. D.; Goldmacher, J.; Assour, J.; Korsakoff, L. A. *J. Org. Chem.* **1967**, *32*, 476.

(18) Han, H.; Salazar, M.; Hurley, L. H. *Nucleic Acids Res.* **1999**, *27*, 537−542.

Scheme 1



due to ligand binding with single- or double-stranded DNA. This is quantified as the percentage of normalized duplex product with respect to the total intensity per lane. The effective doses (ED_{50}) to produce 50% inhibition values were calculated from the graphs in Figure 5B,C, and Se2SAP showed the most potent inhibition ($ED_{50} = 0.03\ \mu M$) in comparison to telomestatin ($ED_{50} = 0.11\ \mu M$) and TMPyP4 ($ED_{50} = 0.99\ \mu M$). As expected, compared to TMPyP4, Se2SAP showed an increase in potency (40-fold) toward the c-MYC G-quadruplex (Figure 5B). While there is considerable apparent duplex interaction for telomestatin, the absence of duplex binding for Se2SAP in the c-MYC sequence shows selective recognition of the G-quadruplex structure by this molecule (Figure 5C).

Selectivity of Se2SAP for the c-MYC Sequence over Duplex DNA by Surface Plasmon Resonance Experiments. Since the pausing in the primer site (Figure 5A) in the *Taq* polymerase assay cannot distinguish between the binding of the molecule with the primer or the duplex, we conducted biosensor surface plasmon resonance (SPR) experiments to independently determine the selectivity of binding of Se2SAP with the c-MYC sequence relative to the duplex. SPR is a powerful technique to monitor intermolecular interactions in real time;[19,20] hence, it was used to study the interaction of Se2SAP with a 19-mer c-MYC sequence and a duplex sequence (CGAATTCG, as a hairpin duplex). These two oligonucleotides were immobilized in different flow cells on the same sensor chip, and a range of Se2SAP concentrations were injected to monitor the interactions with DNA. Suitable blank control injections with running buffer (HEPES buffer containing 200 mM KCl at pH 7.4) were also performed, and the resulting sensorgrams were subtracted from the compound sensorgrams to obtain the final concentration-dependent graphs (Figure 6). Sensorgrams [resonance units (RU) versus time] for the concentration-dependent binding of Se2SAP on the c-MYC sequence and the duplex DNA are shown in parts A and B, respectively, of Figure 6. The binding stoichiometry arises naturally from the RU as saturation of binding sites is approached in SPR experiments. During SPR titration, the increase of RU values is directly proportional to the amount of drug bound to DNA molecules immobilized on the sensor chip. The plot of RU versus the unbound concentration of Se2SAP was fitted to the binding models with one or two binding sites (Experimental Section). The best fit to the interaction results for the c-MYC sequence is for a single, strong binding site, and the expected RU_{max} was obtained for the single binding site (on the basis of the amount of DNA on the chip). At higher concentrations, additional Se2SAP binds to the DNA through a significantly weaker second site. The binding constant for the second site is over a factor of 10 less than for the strong site, and the dissociation rate is much faster, a characteristic of weak, nonspecific binding. The measured steady-state equilibrium binding constant, $K = 6.2 \times 10^{-7}\ M^{-1}$, indicates the high affinity of Se2SAP for the c-MYC sequence. The dissociation constant $k_d = 0.0029\ s^{-1}$ obtained from the kinetics analysis of the sensorgrams also further confirmed the single, strong binding site of Se2SAP with the c-MYC G-quadruplex.

The binding of Se2SAP with the duplex DNA sequence showed much faster kinetics ($k_d > 0.2\ s^{-1}$) than binding with the c-MYC sequence. The very fast off kinetics and the multisite binding at higher concentrations clearly show the nonspecific, weaker binding of Se2SAP to the duplex DNA rather than the c-MYC G-quadruplex. The steady-state K value calculated for the binding of Se2SAP with duplex is on the order of 10^6. SPR experiments further confirm the selectivity (~50-fold) of Se2SAP for the c-MYC G-quadruplex over the duplex DNA.

Se2SAP Shows a Preferred Binding to the 1:2:1 Loop Isomer of the c-MYC G-Quadruplex. Our recent studies indicate that biologically relevant parallel c-MYC G-quadruplexes can exist as a mixture of four loop isomers (Figure 2),

(19) (a) Wang, L.; Bailly, C.; Kumar, A.; Ding, D.; Bajic, M.; Boykin, D. W.; Wilson, W. D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12−16. (b) Wang, L.; Carrasco, C.; Kumar, A.; Stephens, C. E.; Bailly, C.; Boykin, D. W.; Wilson, W. D. *Biochemistry* **2001**, *40*, 2511−2521. (c) Mazur, S.; Tanious, F. A.; Ding, D.; Kumar, A.; Boykin, D. W.; Simpson, I. J.; Neidle, S.; Wilson, W. D. *J. Mol. Biol.* **2000**, *300*, 321−337. (d) Nguyen, B.; Tardy, C.; Bailly, C.; Colson, P.; Houssier, C.; Kumar, A.; Boykin, D. W.; Wilson, W. D. *Biopolymers* **2002**, 63, 281−297. (e) Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, W. D. *J. Am. Chem. Soc.* **2002**, *124*, 2153-2163.

(20) (a) Carrasco, C.; Facompré, M.; Chisholm, J. D.; Van Vranken, D. L.; Wilson, W. D.; Bailly, C. *Nucleic Acids Res.* **2002**, *30*, 1774−1781. (b) Teulade-Fichou, M.-P.; Carrasco, C.; Guittat, L.; Bailly, C.; Alberti, P.; Mergny, J.-L.; David, A.; Lehn, J.-M.; Wilson, D. W. *J. Am. Chem. Soc.* **2003**, *125*, 4732−4740. (c) Davis, T. M.; Wilson, D. W. *Anal. Biochem.* **2000**, *284*, 348−353.

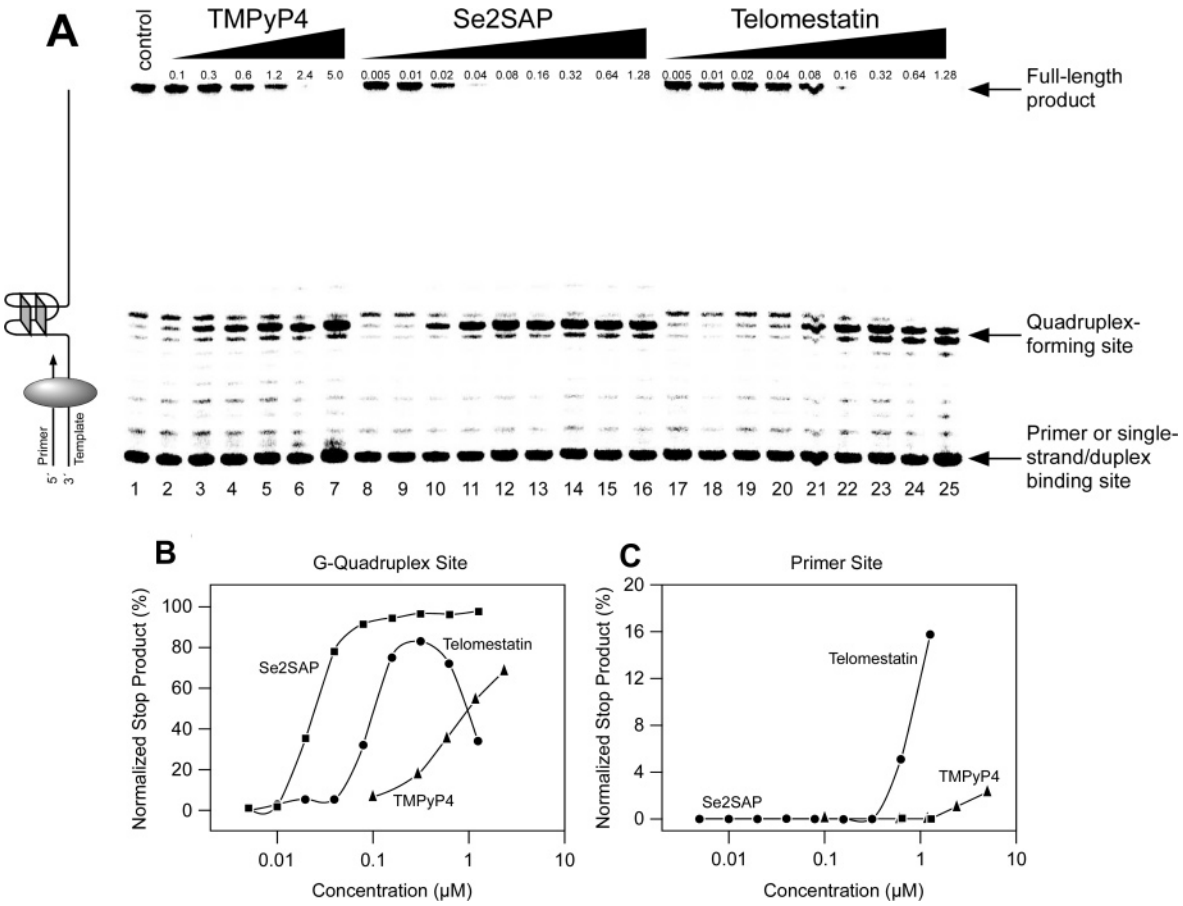


Figure 5. (A) Concentration-dependent inhibition of *Taq* polymerase DNA synthesis by stabilization of the c-MYC G-quadruplex structure with TMPyP4 (0.1−5.0 μM, lanes 2−7), Se2SAP (0.005−1.28 μM, lanes 8−16), and telomestatin (0.005−1.28 μM, lanes 17−25) using a DNA template containing the Pu27 c-MYC sequence at 55 °C (10 mM KCl/NaCl). The cartoon to the left of the gel illustrates the principle of the polymerase stop assay. The increase in intensity of all three bands relative to the control, but particularly the lower band, associated with the *Taq* polymerase stop at the quadruplex-forming site is expected as the G-quadruplex is stabilized by drug binding. In previous experiments, we have demonstrated that polymerase pausing does not occur with oligomers that cannot form a G-quadruplex structure.[1] (B, C) Graphical representations of the quantification of the autoradiogram shown in (A), showing the normalized percentage of stop product at the G-quadruplex site (B) and the primer site (C) as a percentage of the total intensity per lane.

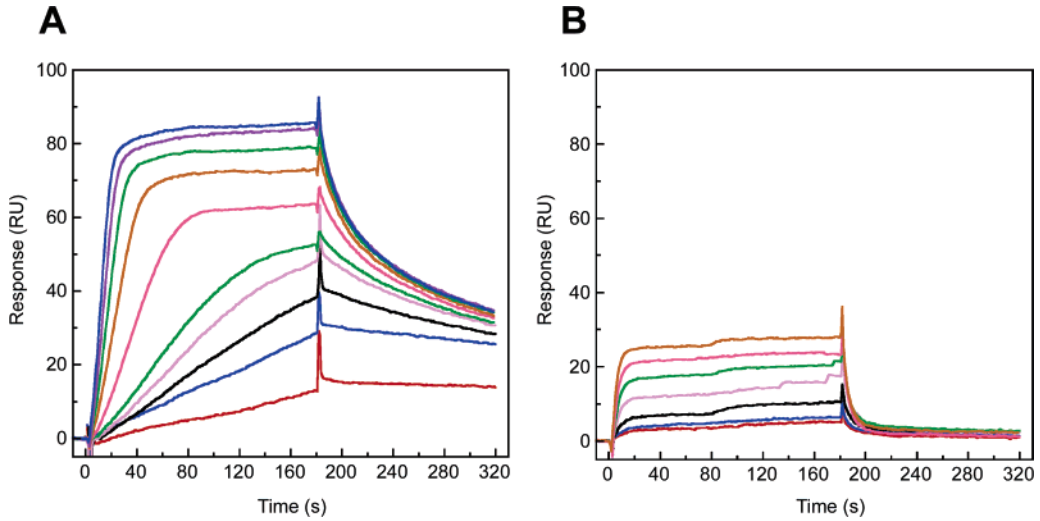


Figure 6. SPR sensorgrams for binding of Se2SAP to the immobilized G-quadruplex formed by a 19-mer c-MYC sequence (A) and CGAATTCG, as a hairpin duplex (B), in HEPES−KCl buffer at 25 °C. The c-MYC curves range in Se2SAP concentration from 2 nM for the lowest curve to 100 nM for the top curve, and for the duplex, they range in Se2SAP concentration from 8 to 100 nM, from bottom to top.

with the predominance of 1:2:1 and 2:1:1 isomers, where G23 is not part of the G-tetrad.[10a] Independent and parallel work conducted by the Patel lab using high-field NMR also showed the predominance of the 1:2:1 loop isomer in the c-MYC parallel G-quadruplex.[10b] The individual loop isomers can be trapped out by specifically mutating the guanines in the loops to thymines (e.g., mutating guanines 14 and 23 traps out the G-quadruplex with the 1:2:1 loop isomer).[10a] Accordingly, the

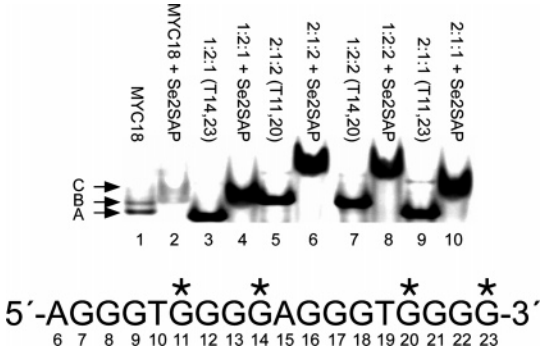


Figure 7. EMSA of the wild-type MYC18 without (lane 1) and with (lane 2) Se2SAP and the four Pu18-mer loop isomers (lanes 3, 5, 7, and 9) containing the dual mutants G → T-14,23, G → T-11,20, G → T-14,20, and G → T-11,23, respectively, and their corresponding Se2SAP incubations (lanes 4, 6, 8, and 10). Arrows at A, B, and C to the left of the gel point to the predominant and minor mixed isomer bands and the predominant Se2SAP-modified band, respectively. The Pu18-mer sequence with the positions of the four G-to-T mutations (indicated with asterisks) is shown below the EMSA.

wild-type 18-mer sequence (MYC18) and the mutated 18-mer sequences were incubated with Se2SAP in the absence of light for 4 h and then subjected to electromobility shift assay (EMSA) (Figure 7).

In the absence of Se2SAP, MYC18 (lane 1) showed two bands, with the bottom major band (band A) corresponding to the 1:2:1 and 2:1:1 isomers (lanes 3 and 9, respectively) and the minor band (band B) corresponding to the 2:1:2 and 1:2:2 isomers (lanes 5 and 7, respectively). Incubation of MYC18 with Se2SAP produced only one new band (band C in lane 2) that corresponds to the bands produced by the incubation of Se2SAP with the 1:2:1 (lane 4) and 2:1:1 (lane 10) loop isomers. The binding of Se2SAP with the 2:1:2 and 1:2:2 isomers produced by mutating guanines 11/20 and 14/20 gave low-mobility bands in lanes 6 and 8, respectively. Since the wild-type sequence incubated with Se2SAP did not show a band corresponding to that found in lanes 6 and 8, this infers that Se2SAP binds preferentially to the 1:2:1 and 2:1:1 isomers present in the wild-type sequence.

To determine the selectivity of Se2SAP within the loop isomers, a competition experiment was conducted. Initially, 5′ fluorescent (FAM) labeled primer was annealed to the template containing the 18-mer c-MYC sequence and allowed to form a G-quadruplex in the presence of 10 mM KCl. *Taq* DNA polymerase was used to extend the primer. A G-quadruplex-binding ligand (Se2SAP or TMPyP4) was added at the concentration previously established (from polymerase stop assay) to produce the 1:1 ratio of stop product to full-length product. Various concentrations of competitor nucleic acids (18-mer dual mutant sequences) were added, and the primer extension was conducted. The stop product and the full-length product were analyzed using capillary electrophoresis. The stronger the binding of the drug to the competitor, the less G-quadruplex pausing is produced. A CC_{50} for each nucleic acid competitor was determined. The CC_{50} is defined as the concentration of competitor required to change the ratio of stop product to full-length product from 1:1 to 1:2. Therefore, an increase in $1/CC_{50}$ values shows an increase in binding of the drug to the dual mutant nucleic acid competitor compared to the wild-type sequence.

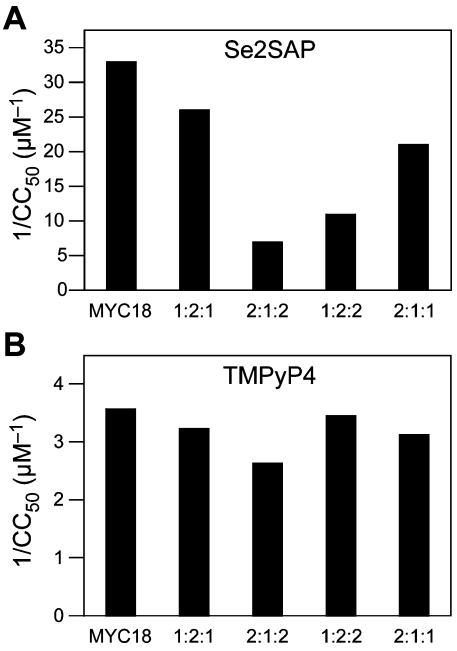


Figure 8. Graphical representation for the competitive binding of (A) Se2SAP and (B) TMPyP4 to the wild-type MYC18 and the four dual mutants G → T-14,23 (1:2:1), G → T-11,20 (2:1:2), G → T-14,20 (1:2:2), and G → T-11,23 (2:1:1).

In accordance with the results from the EMSA, Se2SAP showed preferential binding with the 1:2:1 and 2:1:1 loop isomers (Figure 8A), with a slightly larger selectivity for the 1:2:1 loop isomer. Of the remaining two loop isomers, Se2SAP showed the least binding to the 2:1:2 loop isomer. Since the major isomer in the MYC18 sequence is the 1:2:1 isomer, the majority of binding in the wild-type MYC18 stems from the binding of Se2SAP with this loop isomer. Presumably, Se2SAP would shift the equilibrium of loop isomers to those favored by interaction with this ligand. In contrast, TMPyP4 did not show any preferential binding to a particular loop isomer (Figure 8B).[21] TMPyP4 showed almost comparable $1/CC_{50}$ values with the 1:2:1, 1:2:2, and 2:1:1 loop isomers ($\sim$3.3 μM^{-1}) and slightly less binding with the 2:1:2 loop isomer (2.8 μM^{-1}).

Se2SAP Induces a Structural Change from Parallel to Single-Loop Hybrid G-Quadruplex. Generally, photoactive porphyrins catalyze the oxidation of DNA upon exposure to light, which results in DNA strand breakage in proximity to the binding sites of the porphyrins.[11,14] Se2SAP is much less photoactive than TMPyP4; however, it showed sufficient cleavage for analytical purposes.[10a] The specific cleavage pattern of the G-quadruplex structures can be used to identify both the binding sites and binding modes of porphyrins. Photocleavage experiments on the c-MYC sequence with TMPyP4 showed cleavage of guanines in both the bottom and top G-tetrads that corresponds to the binding of two molecules to the double-loop hybrid G-quadruplex.[10a] The DMS footprinting and photocleavage experiments were carried out with the major 1:2:1 loop isomer in the presence of Se2SAP, and the resulting cleavage patterns are shown in Figure 9A. The protection of guanines from G7 to G22 in DMS footprinting indicates the formation of a G-quadruplex using these guanines. In the 1:2:1 loop isomer, bases G9, G11, G16, and G20 are predominantly cleaved by Se2SAP. Since, in the parallel G-quadruplex structure, G7,

(21) Ren, J.; Chaires, J. B. *Biochemistry* **1999**, *38*, 16067−16075.

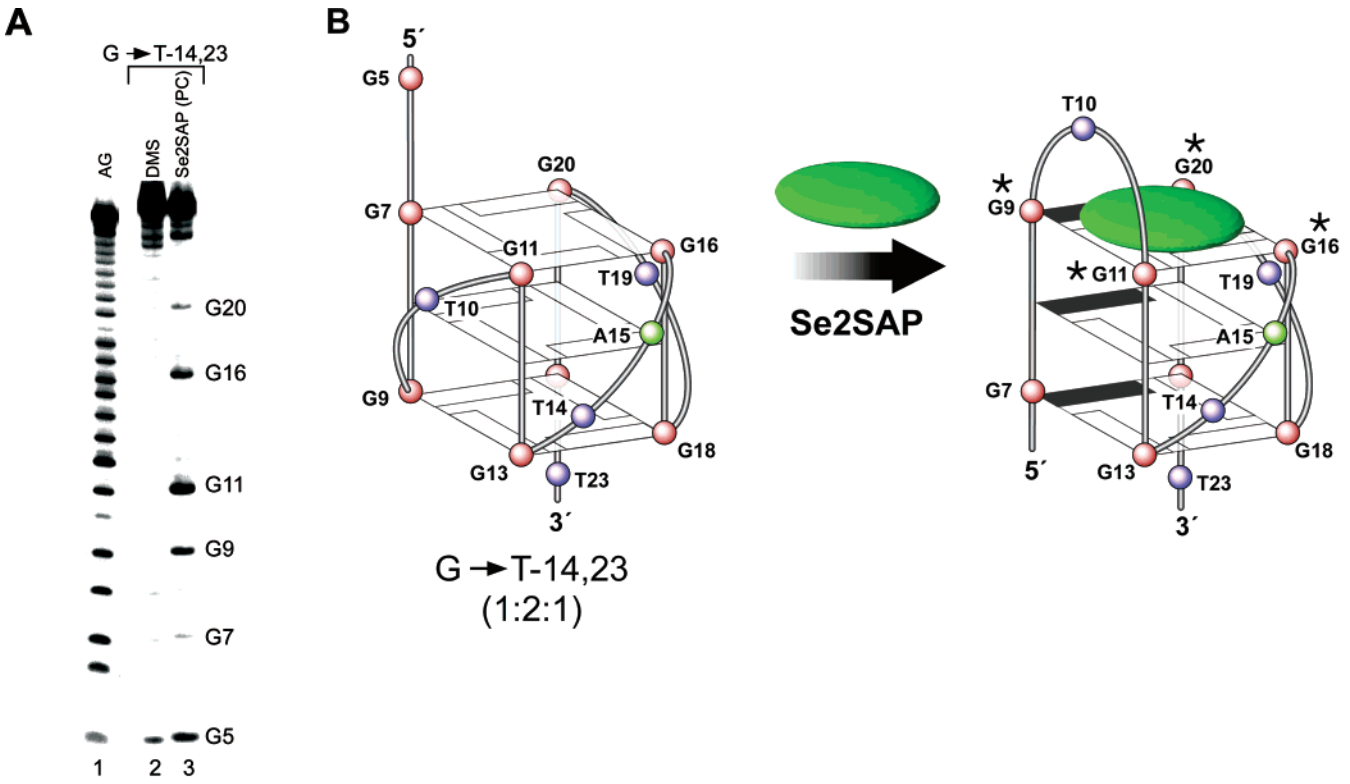


Figure 9. (A) Comparison of DMS footprinting and Se2SAP photoinduced cleavage of the c-MYC (27-mer) dual mutant loop isomer G → T-14,23. Lane 1 shows the A and G sequencing reaction for the Pu27-mer sequence without mutations. Lane 2 shows DMS footprinting for the dual mutant G → T-14,23 in the absence of Se2SAP. Lane 3 indicates the photocleavage of the dual mutant G → T-14,23 in the presence of Se2SAP. The numbering of key guanine residues is shown to the right of the gel. The 27-mer was used, and G-to-T mutations at 3 and 4 as well as 14 and 23 were made to obtain a single isomer. (B) Cartoon representing the conversion of a parallel G-quadruplex to the single-loop hybrid G-quadruplex in the dual mutant G → T-14,23. Shaded rectangles in tetrads indicate *syn*-guanine residues, and unshaded rectangles indicate *anti*-guanine residues.

G11, G16, and G20 are arranged in one G-tetrad, cleavage at G9 instead of G7 indicates that the sapphyrin has trapped a loop inversion at the 5′ end. This cleavage pattern is only compatible with a G-quadruplex structure where guanines G9, G11, G16, and G20 are placed in one G-tetrad to which one molecule of the sapphyrin binds (Figure 9B). The Se2SAP-trapped structure has one external lateral loop running in an antiparallel configuration, while the other two are in a parallel configuration (Figure 9B). This is in accord with the 1:1 stoichiometry for the binding of Se2SAP to the c-MYC sequence determined by SPR experiments.

Molecular Modeling of the Se2SAP Bound to the Single-Loop G-Quadruplex. On the basis of the competition and photocleavage experiments, the 1:2:1 loop isomer of the single-loop hybrid G-quadruplex was selected for modeling studies (Figure 9B). Since neither NMR nor crystallographic information for this novel G-quadruplex structure is available, the model was built from the known, closely related NMR structure of the *Tetrahymena* telomeric G-quadruplex[7a] and the crystal structure of the human telomeric G-quadruplex.[4b] Initial modeling was carried out by modifying the crystal structure of the human telomeric repeat sequence. Guanines at the 5′ end (G7, G8, and G9) were modeled using a similar strand from the *Tetrahymena* telomeric repeat structure. Bases G7, G8, and G9 exhibit a *syn* orientation, while other guanine bases of the quadruplex show an *anti* orientation, resulting in one *syn* and three *anti* orientations of guanines in each of the three tetrads of the quadruplex. This model structure was then refined using energy minimization and molecular dynamics and used for docking studies.

On the basis of the pattern of photocleavage of the 18-mer c-MYC sequence by Se2SAP, the molecule was docked with the single-loop hybrid G-quadruplex near the loop formed by T10 (Figure 9B). Initially, different orientations of Se2SAP were used as a starting point for energy minimization. The most stable orientation, showing the best interaction energy and optimum interactions between Se2SAP and the c-MYC G-quadruplex, was selected and subjected to molecular dynamics (MD) simulations. Trajectories collected during MD simulations were analyzed on the basis of potential energy. Ten low potential energy frames were then subjected to minimization, and interaction energy values were calculated. The best docking orientation of the Se2SAP molecule with the single-loop hybrid G-quadruplex shows an interaction energy value of −296.4 kcal/mol. The model obtained is shown in the Supporting Information. The positively charged pyridyl rings attached at C10 and C15 (for the numbering of Se2SAP, see Figure 3) interact with the negatively charged phosphate backbone of the T10 loop. One of these pyridyl rings (C15) also shows a stacking interaction with T10 and its methyl group. The third positively charged pyridyl ring at C5 shows an interaction with the 3′ phosphate of the guanine from the flanking G-tetrad. Finally, the expanded porphyrin core of Se2SAP exhibits good overlap with the same G-tetrad of the quadruplex.

SPR Demonstrates That Se2SAP Binds More Selectively to the c-MYC and Similar G-Quadruplexes Than to Other G-Quadruplex Structures. The selective trapping out of the single-loop hybrid G-quadruplex by Se2SAP and the strong binding of this molecule, as implied from the modeling studies,

suggested that this molecule may be selective for this G-quadruplex structure, which has the tetrad arrangement *syn*−*anti*−*anti*−*anti* associated with the one-base lateral loop. To compare the selectivity of Se2SAP with the c-MYC G-quadruplex versus other G-quadruplex structures, an SPR experiment was performed using three established and different G-quadruplex-forming sequences. Since the HumTel sequence forms at least two G-quadruplex structures,[4b,c,22] we have used the G_2T_4 sequence[7g,h] to determine the binding of Se2SAP to the basket structure. G_2T_4 is known to form the basket-type (Figure 1E) G-quadruplex structure in both NaCl and KCl solutions.[7g,h] For comparison with the chair-type G-quadruplex (Figure 1D), TBA was used.[7b−f] Finally, the TetTel sequence, which forms a G-quadruplex structure with one internal and two external loops (Figure 1C),[7a] was used as a mimic of the G-quadruplex structure formed upon addition of Se2SAP to the parallel c-MYC G-quadruplex. One end of this G-quadruplex has a structure with a tetrad arrangement similar to that proposed to be trapped out by Se2SAP. The three G-quadruplex-forming oligonucleotides, TetTel, TBA, and G_2T_4, were immobilized in different flow cells on the same sensor chip, and the SPR experiments were conducted as described above with a range of Se2SAP concentrations. HEPES buffer containing 200 mM KCl or 200 mM NaCl at pH 7.4 was used for all the sequences. Since the NMR structure of the mixed parallel/antiparallel G-quadruplex for the TetTel sequence was obtained in NaCl,[7a] the SPR experiment was conducted in HEPES buffer containing 200 mM NaCl for the TetTel sequence. The SPR experiment showed similar results for the binding of Se2SAP with the TetTel sequence in the presence of buffer containing KCl or NaCl. The final concentration-dependent sensorgrams (analogous to those in Figure 6) for the binding of Se2SAP to these three sequences are given in the Supporting Information.

In all cases, the G-quadruplexes exhibited two binding sites: a single strong binding site and a site that is at least 10-fold weaker. Since the second binding site in all cases exhibited much faster dissociation kinetics, which is characteristic of weak, nonspecific binding, only the results from the primary strong binding site are considered for further discussion. The dissociation rate constant (k_d) obtained from the kinetics method and the equilibrium binding constant (K) derived from the steady-state method for the different sequences are tabulated in Table 1 (for comparison, the values for c-MYC and hairpin duplex DNA are also included). The differences in the dissociation of Se2SAP from that of the TetTel, TBA, and G_2T_4 sequences can be easily visualized from their corresponding sensorgrams (Supporting Information). Among the G-quadruplexes, the c-MYC G-quadruplex showed the highest equilibrium constant ($K = 6.2 \times 10^{-7}$ M^{-1}) and the slowest dissociation constant ($k_d = 0.0029$ s^{-1}) indicative of the strongest binding of Se2SAP with this sequence. The kinetics ($k_d = 0.0041$ s^{-1}) and the steady-state ($K = 4.2 \times 10^{-7}$ M^{-1}) values for the TetTel sequence indicate that Se2SAP also binds very strongly to a single site in the double-loop hybrid G-quadruplex with a slightly lower binding constant than with the c-MYC G-quadruplex. With the G_2T_4 sequence, Se2SAP showed a much faster dissociation rate (k_d value is ~12 times larger than that for the c-MYC sequence), and its K value is similar to that of the duplex DNA (Figure 6B, Table 1), indicating the very weak binding of Se2SAP to the basket-type G-quadruplex structure. In contrast, the SPR experiments indicated that Se2SAP showed moderate binding to the chair-type G-quadruplex formed in the TBA sequence. In comparison to the c-MYC sequence, TBA showed an approximate 3-fold increase in the dissociation constant value for Se2SAP and a 4-fold decrease in the equilibrium constant value. The SPR results for the binding of Se2SAP to the four different G-quadruplex-forming sequences were converted to steady-state binding isotherm plots and are shown in Figure 10. The steady-state isotherms were constructed by averaging the observed SPR responses in the steady-state region, and the responses were then converted to r, moles of Se2SAP bound per mole of G-quadruplex DNA ($r = RU_{obs}/RU_{max}$), and plotted against the free Se2SAP concentration in the flow solution. The fitting results shown in Figure 10 are for the single strong binding site. As seen in Table 1, the plots clearly indicate the order of binding of Se2SAP to the G-quadruplex as c-MYC ≥ TetTel ≫ TBA ≫ G_2T_4, i.e., single-loop hybrid ≥ double-loop hybrid ≫ chair-type ≫ basket-type G-quadruplexes.

Table 1. Equilibrium and Dissociation Kinetics Constants for Se2SAP Binding to G-Quadruplex and Hairpin Duplex DNA[a]

DNA	$K \times 10^{-7}$ (M^{-1})	k_d (s^{-1})	$k_a \times 10^{-3}$ (M^{-1} s^{-1})
c-MYC	6.2	0.0029	1.8
TetTel	4.2	0.0041	1.7
TBA	1.4	0.0094	1.3
G_2T_4	0.31	0.034	1.1
hairpin duplex	0.16	>0.2	>3

[a] Binding experiments were conducted by biosensor SPR methods at 25 °C in HEPES−K^+ buffer. All experiments were repeated at least twice with fresh samples. Fitting errors due to random point scatter in any experiment are less than ±5%. Experimental errors estimated from reproducibility of results are ±10% for K values between 1×10^6 and 2×10^7 M^{-1} and for k_d values between 0.1 and 0.001 s^{-1}. Errors increase to ±20% for K values between 3×10^7 and 1×10^8 M^{-1}. It is difficult to accurately determine k_d values that are significantly greater than 0.1 s^{-1} by biosensor SPR methods. Methods for determining K and k_d are given in the Experimental Section. Because of some adsorption of Se2SAP in the biosensor flow system on initial injection, the k_a values are difficult to determine directly and the values in the table are estimated from K and k_d.

Competition Binding of Se2SAP to the Single-Loop Hybrid G-Quadruplex in Comparison to the Chair, Parallel, and Double-Loop Hybrid G-Quadruplex-Forming Sequences, the Human Telomeric Sequence, and Single-Stranded and Duplex DNA. To compare the competitive selectivity of Se2SAP with the c-MYC G-quadruplex versus other G-quadruplex structures, a competition experiment was performed using the different G-quadruplex-forming sequences. The results of the competition assay showed that Se2SAP has selectivity for the single-loop hybrid G-quadruplex formed in the c-MYC sequence (Figure 11A) and also a comparative selectivity for the TetTel sequence that has a similar external loop at one end, confirming the results from the SPR experiments. GGA_4 forms exclusively the parallel heptad−tetrad G-quadruplex structure, which cannot produce an induced structure (due to the unavailability of loop bases to form loops),[23a,b] and because of this, Se2SAP and TMPyP4 showed very weak binding with this sequence. In contrast to GGA_4, GGA_8 forms the two heptad−tetrad G-quadruplex structures that

(22) (a) Redon, S.; Bombard, S.; Elizondo-Riojas, M.-A.; Chottard, J.-C. *Nucleic Acids Res.* **2003**, *31*, 1605−1613. (b) Ying, L.; Green, J. J.; Li, H.; Klenerman. D.; Balasubramanian, S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14629−14634. (c) Phan, A. T.; Patel, D. J. *J. Am. Chem. Soc.* **2003**, *125*, 15021−15027.

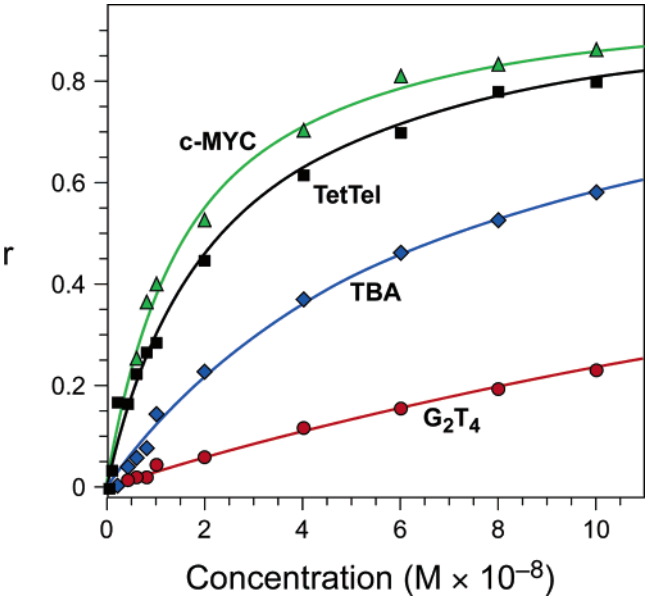


Figure 10. SPR steady-state binding plots. The concentration axis is for Se2SAP in the flow solution and is the unbound compound concentration; *r* represents the moles of compound bound per mole of G-quadruplex DNA. The points are for experimental data, and the lines are the best-fit results for nonlinear least-squares application of a single-site model to the data. Fitting errors due to random point scatter in any experiment are less than ±5%. The equilibrium constants determined from fitting data sets for each DNA are collected in Table 1. The buffer for all DNAs except TetTel was HEPES with 200 mM KCl. TetTel results are for the same buffer but with the KCl replaced by 200 mM NaCl. Experimental and fitting details are given in the Experimental Section.

are bridged by a single-base loop,[23a] which provides a binding pocket into which the porphyrin can be inserted. Se2SAP and TMPyP4 showed moderate binding with the GGA_8 sequence ($CC_{50} \approx 0.2\ \mu M$). In comparison to the human telomeric sequence ($CC_{50} = 1.3\ \mu M$), Se2SAP showed a 26-fold selectivity for the c-MYC sequence ($CC_{50} = 0.05\ \mu M$). Since the competition assay involves the polymerase extension at 65−70 °C in the final step, the G_2T_4 sequence, which forms the low-melting (40 °C) basket-type G-quadruplex, could not be used, and for the chair-type G-quadruplex structure, we have used the sequence d[GGGGTTGGGGTGTGGGGTTGGGG] (G4TBA, melting point 86 °C).[23c,d] Se2SAP showed a 2-fold selectivity for the c-MYC sequence in comparison to the G4TBA chair-type structure. Confirming the results from the SPR experiments, Se2SAP showed the least selectivity for or no binding to the duplex or the single-stranded DNA sequence. For comparison, the same experiment was carried out with TMPyP4 (Figure 11B) and telomestatin (unpublished results). TMPyP4 and telomestatin did not show any appreciable selectivity among the c-MYC, TetTel, G4TBA, and GGA_8 sequences (Figure 11B and unpublished results). In contrast to Se2SAP, TMPyP4 showed only an 8-fold selectivity for the c-MYC sequence in comparison to the HumTel sequence.

Comparative Molecular Modeling of the Binding of Se2SAP to Chair, Basket, and Double-Loop Hybrid G-Quadruplex Structures. The NMR solution structures of the TetTel[7a] and TBA[7f] were used for the modeling studies. These structures were obtained from PDB and refined using MD simulations. The G_2T_4 sequence forms the basket-type G-quadruplex,[7g,h] and a model of this sequence was constructed using the *Oxytricha* telomeric repeat basket-type G-quadruplex.[25] To construct this model, two guanine tetrads of the structure were deleted and the loops were rejoined. This model was refined using an MD simulation protocol similar to that described above.

Docking of Se2SAP with the different G-quadruplex structures was performed using manual rotation and translation. Initially, various orientations of Se2SAP (differing in orientation of the core with respect to the loops of the quadruplex) were obtained. These structures were then minimized, and the most stable structure showing the best interaction energy and optimum interactions between Se2SAP and G-tetrad was selected. This structure was subjected to MD simulations. For G-quadruplexes with two binding sites, docking of Se2SAP was performed separately in each loop. Table 1S (Supporting Information) shows the interaction energy values for the binding of Se2SAP with different G-quadruplex structures.

In the TetTel double-loop hybrid G-quadruplex, docking of Se2SAP was performed at sites 1 and 2 of the structure (Figure 1C). The orientations of guanines in the G-tetrad of site 2 are similar to that of the single-loop hybrid G-quadruplex structure (*syn*−*anti*−*anti*−*anti*). Se2SAP shows the best interaction at site 2, with an interaction energy of −322.3 kcal/mol. The two positively charged pyridyl rings attached at C5 and C20 (connecting with the bipyrrole ring system) interacted with the loop bases (four-base loop), while the pyridyl ring attached to C15 interacted with the phosphate group of the guanine at the 3′ end. This guanine also showed stacking interactions with the sapphyrin core of Se2SAP. Thus, Se2SAP is stacked between the G-tetrad and the loop bases along with the guanine at the 3′ end. This preferential interaction of Se2SAP with site 2 is also supported by the fact that Se2SAP interacts at site 1 with an interaction energy of only −231.4 kcal/mol (Table 1).

In the case of TBA, the Se2SAP molecule showed intermediate interactions at site 1 (−203 kcal/mol) and less favorable interactions at site 2 (−139 kcal/mol) of the chair-type G-quadruplex structure (Figure 1D) compared to the c-MYC and TetTel G-quadruplexes. The two pyridyl rings attached to C5 and C20 showed interactions with the phosphate backbone of the loop at site 1.

Docking of Se2SAP with the G_2T_4 basket-type G-quadruplex structure was attempted at the diagonal loop (site 1) and across the two lateral loops (site 2) of the structure (Figure 1E). The interaction of Se2SAP with both lateral loops was less favorable (−120 kcal/mol at site 1 and −59.5 kcal/mol at site 2) than with the loop in the c-MYC G-quadruplex structure. This indicates the very weak binding of Se2SAP to the basket-type G-quadruplex structure compared to other G-quadruplex structures. Thus, the molecular modeling results are in accordance with the experimental results obtained through the SPR and competition experiments.

Se2SAP Is Much Less Cytotoxic Than TMPyP4. The cytotoxicity assay was performed using a HeLa cell line. A comparison of the cytotoxicities of TMPyP4 and Se2SAP is

(23) (a) Matsugami, A.; Okuizumi, T.; Uesugi, S.; Katahira, M. *J. Biol. Chem.* **2003**, *278*, 28147−28153. (b) Matsugami, A.; Ouhashi, K.; Kanagawa, M.; Liu, H.; Kanagawa, S.; Uesugi, S.; Katahira, M. *J. Biol. Chem.* **2001**, *313*, 255−269. (c) Petraccone, L.; Erra, E.; Esposito, V.; Randazzo, A.; Mayol, L.; Nasti, L.; Barone, G.; Giancola, C. *Biochemistry* **2004**, *43*, 4877−4884. (d) Randazzo, A.; Galeone, A.; Esposito, V.; Varra, M.; Mayol, L. *Nucleosides Nucleotides* **2002**, *21*, 535−545.

(24) Insight II 2000, Molecular Modeling Software, Accelrys Inc., 9685 Scranton Rd., San Diego, CA 92121.

(25) Wang, Y., Patel, D. J. *J. Mol. Biol.* **1995**, *251*, 76−94.

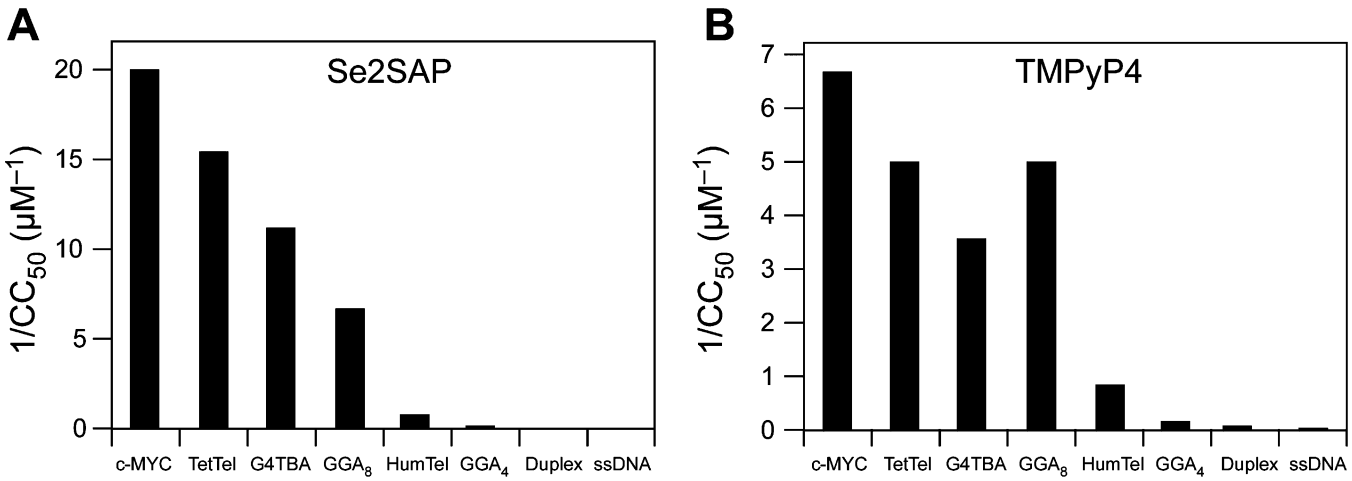


Figure 11. Graphical representation of the competitive binding of (A) Se2SAP and (B) TMPyP4 with the 27-mer c-MYC sequence and other sequences that form different G-quadruplex structures and duplex and single-stranded DNA. For the type of DNA, see the text.

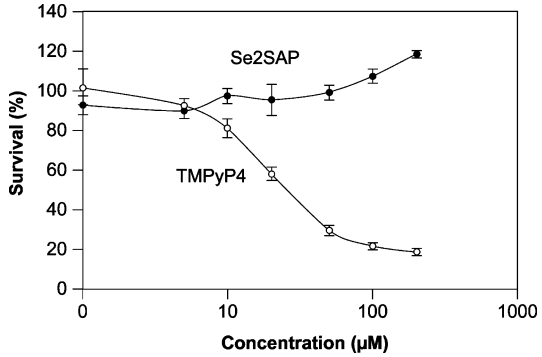


Figure 12. Cytotoxicity of porphyrins on HeLa S_3 cells. Cells were incubated with the indicated concentrations of Se2SAP and TMPyP4 for 96 h. Cytotoxicity was then measured using an MTS assay (see the Experimental Section).

shown in Figure 12. While TMPyP4 shows cytotoxicity, with an IC_{50} value of 43 μM, Se2SAP is not cytotoxic, even at 200 μM.

Discussion

In this study we have identified a ligand, Se2SAP, that has selectivity for the single-loop hybrid G-quadruplex over other G-quadruplex structures and duplex or single-stranded DNA. Molecular modeling has been used to rationalize the selectivity of Se2SAP for this induced c-MYC G-quadruplex over the other G-quadruplex structures. Our initial lead compound, TMPyP4, is known to bind to both the c-MYC and human telomeric G-quadruplex structures, leading to repression of c-MYC transcription[1,10c] and telomerase inhibition,[12] respectively. To achieve therapeutic selectivity by targeting G-quadruplexes, it is necessary to synthesize drugs that can differentiate among the different types of G-quadruplex structures formed in different sequences. TMPyP4 has several liabilities as a potential drug candidate. It shows moderate selectivity for the c-MYC G-quadruplex versus duplex DNA (Figure 11B), is photoactive, and, as a consequence of these properties, shows moderate cytotoxicity (Figure 12). The larger ring structure of telomestatin in comparison to TMPyP4 (Figure 3) and the observation that substitution of selenium in the porphyrin core decreases photoactivity[15] encouraged us to design a new generation of cationic porphyrins with improved selectivity and photoactive properties.

Comparison of G-Quadruplex Binding and Biological Effects. Se2SAP shows less photocytotoxicity and correspondingly less cytotoxicity than TMPyP4 (Figure 12). Among the loop isomers formed in the c-MYC sequence, Se2SAP showed a preference for the 1:2:1 and 2:1:1 loop isomers where G23 is external to the G-tetrads, which are the major isomers present in the c-MYC sequence. The interaction of the positive charge on the *meso*-pyridyl ring in porphyrins and the negative charge of the phosphate backbone in DNA is primarily responsible for the binding of *meso*-substituted cationic porphyrins to G-quadruplex structures (*meso*-substituted porphyrins without positive charges did not show any binding, unpublished results). Due to the lateral orientation of parallel loops in the c-MYC parallel G-quadruplex structure (Figure 1A), the phosphate anions are not available for interaction with the pyridyl positive charge; therefore, the porphyrin must trap out a lateral loop structure to successfully bind. This is probably the main reason that Se2SAP traps out the single-loop hybrid G-quadruplex in the c-MYC sequence. SPR technology not only confirmed the strong binding ($K = 6.2 \times 10^{-7}$ M^{-1}) of Se2SAP to this novel c-MYC G-quadruplex structure but also provided solid quantitative information regarding the selectivity of Se2SAP for the single-loop hybrid G-quadruplex over the chair- and basket-type G-quadruplex structures and duplex DNA. The analysis of the steady-state binding isotherms from SPR further established the order of selectivity for binding of Se2SAP to G-quadruplexes, as single-loop hybrid (c-MYC) $\geq$ double-loop hybrid (TetTel) $\gg$ chair-type (TBA) $\gg$ basket-type (G_2T_4) and duplex DNA (Table 1 and Figure 10). Se2SAP showed more than 600-fold selectivity for the single-loop hybrid G-quadruplex over the duplex or single-stranded DNA in the competition assay. TMPyP4 showed an 8-fold and Se2SAP a 26-fold preference for the G-quadruplex formed in the c-MYC sequence over the human telomeric sequence (Figure 11). Thus, the results from SPR and competition assay experiments confirmed the selective and strong binding of Se2SAP to the c-MYC G-quadruplex structure in comparison to other G-quadruplex structures (parallel, basket, and chair) or more common forms of DNA (single-stranded and duplex).

Although telomestatin showed a strong stabilization of the c-MYC chair-type G-quadruplex at low concentrations (about 0.1 μM), it also appeared to exhibit duplex or single-stranded

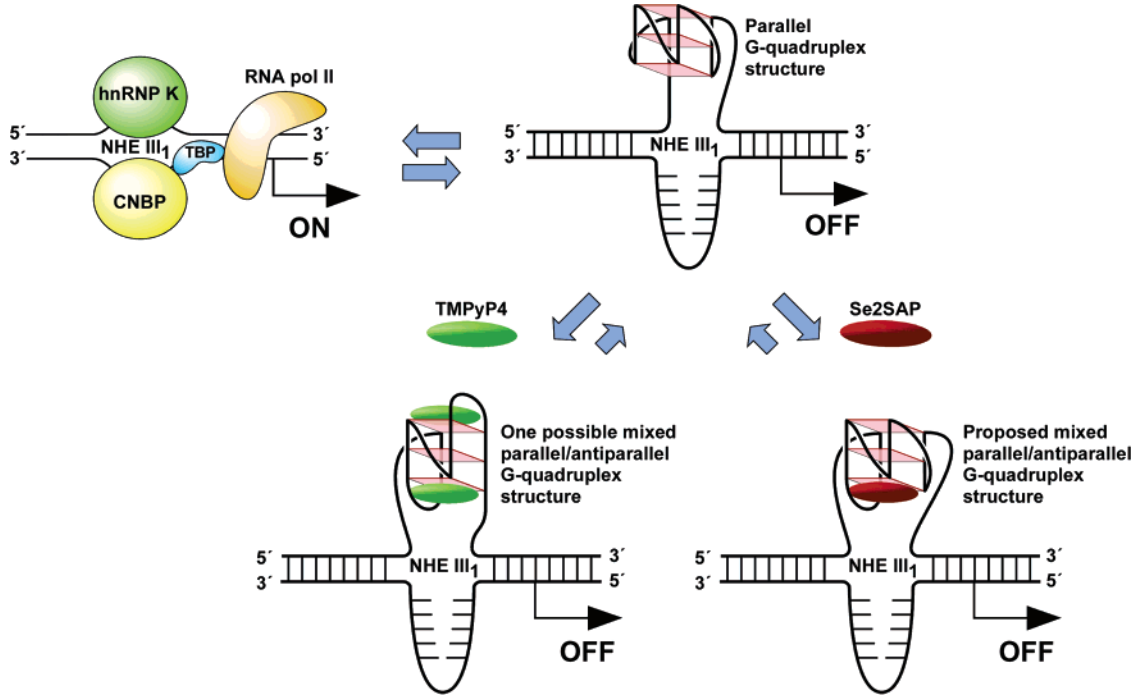


Figure 13. Model for the activation and repression of c-MYC gene transcription involving the conversion of the paranemic secondary DNA structures (gene off) to purine and pyrimidine single-stranded DNA forms for transcriptional activation. hnRNP K and CNBP are single-stranded-DNA-binding proteins involved in transcriptional activation. Interaction of the parallel G-quadruplex structure with TMPyP4 and Se2SAP stabilizes the gene-off form by conversion to proposed respective double-loop and single-loop hybrid G-quadruplex structures that stabilize the silencer element and results in transcriptional repression.

DNA binding at somewhat higher concentrations (1 μM) in the polymerase stop assay (Figure 5C). However, Se2SAP showed a very specific affinity for the c-MYC sequence in the polymerase stop assay, with little to no duplex binding activity (Figure 5C). Although Se2SAP showed a much more potent stabilization of the c-MYC G-quadruplex than TMPyP4, it showed only slightly better c-MYC inhibition at 100 μM (unpublished results). Most likely, Se2SAP is not being taken up into the cells as effectively as TMPyP4. This is perhaps not surprising, as the expanded nucleus of sapphyrin is not as biologically relevant to human cells as the tetrapyrrole porphin nucleus. Another reason for the lower than expected c-MYC lowering by Se2SAP may be the trapping out of different G-quadruplex structures by TMPyP4[10a] and Se2SAP. The strong and selective interaction of one Se2SAP molecule with the single-loop hybrid G-quadruplex stabilizes this G-quadruplex structure to a greater extent ($ED_{50} = 0.03$ μM) than TMPyP4 and hence did not allow for a second looping out required for binding of a second molecule of Se2SAP. In contrast, since TMPyP4 is not a very strong binding molecule ($ED_{50} = 0.99$ μM), the presence of more than 1 equiv of TMPyP4 can lead to the trapping out of a second lateral loop, thus binding one molecule in each of the loops.[10a] The double-loop hybrid structure induced by TMPyP4 completely changes the originally formed parallel G-quadruplex structure in the c-MYC NHE III_1 region compared to the single-loop hybrid structure induced by Se2SAP and hence might produce a more dramatic effect on the recognition and transformation of the transcriptionally inactive form (the parallel G-quadruplex structure) into an active form (linear or duplex DNA) by proteins involved in c-MYC transcriptional activation (Figure 13).

Molecular Basis for the Selectivity of Se2SAP for the c-MYC G-Quadruplex over the Chair- and Basket-Type G-Quadruplex Structures. There are two types of interactions involved in the binding of Se2SAP to G-quadruplex: stacking-type interactions between the porphyrin core and the G-tetrad guanines and interactions between the positively charged pyridyl rings and the negatively charged phosphate backbone. Since Se2SAP does not bind to the $d(GGA)_4$ sequence, which forms a parallel structure where loop phosphate anions are not available for interaction, it can be deduced that Se2SAP binding requires the presence of ionic interactions for recognition of G-quadruplex. The orientations of the phosphate anions in the loops with respect to the pyridyl positive charges on the Se2SAP are governed by the G-tetrad arrangements. For example, in the case of the single-loop hybrid G-quadruplex (Figure S1, Supporting Information), the *syn*−*anti*−*anti*−*anti* arrangement of guanines in the G-tetrad near the lateral loop positions the phosphate anions (5′ and 3′ phosphate groups of T10) favorably for strong interactions with the positive charges on the pyridyl groups at C10 and C15. The distance between the 5′ and 3′ phosphate groups of the T10 loop is 6.9 Å, which is similar to the distance between the positively charged nitrogens of the pyridyl rings attached to C10 and C15 (8.6 Å). Hence, the selectivity of Se2SAP may be correlated with the G-tetrad arrangement and the number of bases in the loop. The external loop containing T10 is also involved in the stacking-type interaction with one of the pyridyl rings. Finally, the bipyrrole and selenophene parts of Se2SAP extend over the G-tetrad of the quadruplex, allowing for better stacking interactions and resulting in enhanced binding (−296.4 kcal/mol). The absence of a second lateral loop similar in orientation to the T10 lateral loop avoids the steric interaction of the other two pyridyl rings (C5 and C20) with this loop and establishes the orientation of the pyridyl rings at C10 and C15 with respect to the phosphate anions of T10.

In the case of the double-loop hybrid G-quadruplex formed by the TetTel sequence (Figure 1C), the arrangement of guanines in the G-tetrad at site 2 is similar to the arrangement of guanines in the single-loop hybrid G-quadruplex of c-MYC (Figure 9B). Se2SAP also binds strongly to this TetTel G-quadruplex (Table 1S). Modeling of Se2SAP with the TetTel G-quadruplex showed that the two pyridyl rings attached at C5 and C20 bind well at site 2 (with an energy difference of 90 kcal/mol in comparison to site 1) of the G-quadruplex. The selective binding at site 2 is further confirmed by the observation of a single, strong binding site in the SPR experiments for the TetTel sequence. Site 2 consists of a four-base loop having a 15.1 Å distance between phosphate groups of the neighboring loop bases, which is similar to the distance between the positively charged pyridyl rings attached to C5 and C20 (14.6 Å). This interaction is further enhanced by the third pyridyl ring, which interacts with the phosphate of the guanine at the 3′ end. This guanine is also involved in the stacking interaction with Se2SAP, and these additional interactions increase the binding energy for the TetTel G-quadruplex compared to the c-MYC G-quadruplex.

Unlike the c-MYC and TetTel G-quadruplex structures, the TBA chair- and G_2T_4 basket-type G-quadruplexes have two lateral loops at site 2 (parts D and E, respectively, of Figure 1). The presence of a second loop at site 2 produces a steric hindrance for the pyridyl rings in Se2SAP. During MD simulations, repulsion between the loop bases and the 3-pyridyl rings of the Se2SAP was observed, which was responsible for the weak binding of Se2SAP at site 2 in both the TBA chair-type (−139 kcal/ mol) and G_2T_4 basket-type (−59.5 kcal/mol) G-quadruplex structures. Also, the SPR experiments showed a very weak or nonspecific second binding site in addition to the strong primary binding site for these two sequences.

The TBA G-quadruplex is a chair-type structure with a different backbone configuration of guanines (*syn*−*anti*−*syn*−*anti*) in site 1 (Figure 1D). Modeling of this structure with Se2SAP at site 1 indicates that only the two positively charged pyridyl rings (at C10 and C15) interact with the phosphate backbone. The distance between the terminal phosphate groups of the loop bases at site 1 is 11.5 Å, which is significantly greater than the distance between the pyridine rings attached to the C10 and C15 atoms (8.6 Å). This results in weaker binding of Se2SAP to the TBA chair-type G-quadruplex at site 1 (−203 kcal/mol) as compared to the c-MYC (−296.4 kcal/mol) or TetTel (−322.3 kcal/mol) G-quadruplex structure.

Since the distance between the terminal phosphate groups of the four-base diagonal loop in site 1 (12.1 Å) of the basket-type G-quadruplex formed in the G_2T_4 sequence is much less than the distance between the interacting, positively charged pyridyl groups at C10 and C20 (15.3 Å), Se2SAP shows very weak binding compared to that with the other G-quadruplex structures evaluated. Also, the tetrad arrangement in the basket-type G-quadruplex is *syn*−*syn*−*anti*−*anti*.

In summary, the selective binding and interactions of Se2SAP with the various G-quadruplexes is largely governed by the strength of the electrostatic interactions between the positively charged pyridyl rings in Se2SAP and the negatively charged phosphate backbones in the loops of the G-quadruplexes, which in turn are controlled by the relative orientations of guanines in the G-tetrads and the availability of loops. Hence, Se2SAP is selective to a G-quadruplex with a single lateral loop and the *syn*−*anti*−*anti*−*anti* arrangement of guanines in the G-tetrads.

Conclusion

We have identified an expanded porphyrin molecule that can selectively bind to the single-loop hybrid G-quadruplex. This G-quadruplex binding ability of Se2SAP presumably results in the inhibition of transcription of c-MYC by this molecule. Also, Se2SAP is not photoactive and shows no cytotoxicity. Analysis of these results shows that proper positioning of the positive charges in the molecule in relation to the negatively charged phosphate atoms in the loops can lead to selectivity for a particular type of G-quadruplex structure. This establishes the important principle that it is possible to design molecules that are selective for different G-quadruplex structures.

Experimental Section

Chemistry. ^{1}H NMR spectra were run on a Varian Unity 300 or 600 MHz NMR spectrophotometer. The chemical shifts are relative to the trace proton signals of the deuterated solvent. Coupling constants (*J*) are reported in hertz and refer to apparent peak multiplicity rather than coupling constants. UV−vis spectra were recorded on a Varian Cary 3E UV−vis spectrophotometer. Fourier transform ion cyclotron resonance experiments were carried out on an Ionspec 4.7 T instrument. Singly and multiply protonated porphyrin and sapphyrin molecules were generated by electrospray ionization (ESI). Fast atom bombardment (FAB) measurements were carried out on a JEOL HX-110 sector instrument equipped with a conventional Xe gun. A mixture matrix of glycerol/thioglycerol/*m*-nitrobenzyl alcohol (50:25:25) containing 0.1% TFA was used as the FAB matrix. For accurate mass measurements, poly(ethylene glycol) was used as the internal standard. Flash column chromatography was performed on silica gel 60, 230−400 mesh, purchased from Spectrum. Basic alumina was used for purifying porphyrins in column chromatography. All starting materials were obtained from commercial sources unless otherwise specified.

Synthesis. a. Preparation of 2,5-Bis(3-pyridylhydroxymethyl)-selenophene (1). A 150 mL sample of anhydrous hexane, 14.4 mL (0.09 mol) of TMEDA, 60 mL (1.6 M in hexane) (0.09 mol) of *n*-butyllithium, and 3.5 mL (0.04 mol) of selenophene were added to a three-necked, round-bottomed flask flushed with argon at room temperature, and the mixture was refluxed for 1 h. After the mixture was cooled to room temperature, the suspension formed was slowly transferred dropwise via a cannula to a degassed solution of 9.04 mL (0.09 mol) of 3-pyridinecarboxaldehyde in 250 mL of anhydrous THF and stirred overnight. Ice cold NH_4Cl solution was added to quench the reaction. The phases were separated, and the water layer was extracted with a 1:6 methanol−chloroform mixture. The organic layers were combined, washed with water, and dried over Na_2SO_4. After removal of solvent, the residue was purified by chromatography on silica gel using chloroform−methanol (8:1) as an eluent to give 4.6 g (35%) of **1**: solid; mp 185 °C; ^{1}H NMR (DMSO-d_6) δ 8.59 (s, 2H), 8.45 (d, J = 4.7 Hz, 2H), 7.76 (d, J = 4.7 Hz, 2H), 7.35−7.32 (m, 2H), 6.88 (s, 1H), 6.85(s, 1H), 6.43 (s, 2H), 5.90 (s, 2H); HRMS (FAB) *m*/*z* calcd for $C_{16}H_{15}O_2N_2Se$ 347.0299, found 347.0297.

b. Preparation of 2,5-Bis(4-pyridylhydroxymethyl)selenophene. 2,5-Bis(4-pyridylhydroxymethyl)selenophene was prepared by a reaction of 9.04 mL (0.09 mol) of 4-pyridinecarboxaldehyde with 3.5 mL (0.04 mol) of selenophene using a procedure similar to that described above. The 2,5-bis(4-pyridylhydroxymethyl)selenophene was obtained in 20% yield: solid; mp 183 °C; ^{1}H NMR (DMSO-d_6) δ 8.48 (d, J = 2.5 Hz, 4H), 7.36 (d, J = 2.5 Hz, 2H), 6.93 (s, 2H), 6.44 (br s, 2H), 5.82 (s, 2H); HRMS (FAB) (M^+) *m*/*z* calcd for $C_{16}H_{15}O_2N_2Se$ 347.0299, found 347.0301.

c. Preparation of 5,10,15,20-Tetra(3-pyridyl)-21,23-diselenaporphyrin (2) and 5,10,15,20-tetra(3-pyridyl)-26,28-diselenasapphyrin

(3). A mixture of 1.2 g (3.49 mmol) of **1** and 0.247 mL (3.49 mmol) of pyrrole was dissolved in 200 mL of propionic acid, and the resulting solution was refluxed for 48 h. After the solution was cooled to room temperature, the solvent was evaporated to dryness under high vacuum. The residue obtained was chromatographed on basic alumina using 5% methanol and dichloromethane to give a mixture of diselenaporphyrin **2** and diselenasapphyrin **3**. This mixture was purified by repeated column chromatography and separated by preparative TLC (silica gel). Yields of 3% and 1% were obtained for diselenaporphyrin **2** and diselenasapphyrin **3**, respectively. Data for diselenaporphyrin **2**: solid; mp > 250 °C; ^{1}H NMR ($CDCl_3$) δ 9.95 (s, 4H), 9.56 (s, 4H), 9.13 (d, J = 4.7 Hz, 4H), 8.88 (s, 4H), 8.65−8.63 (m, 4H), 7.90−7.86 (m, 4H); UV−vis (CH_2Cl_2) λ_{max} (nm) (ϵ (mol^{-1} cm^{-1})) 443 (28900), 524 (4900), 626 (1000), 688 (800), 723 (500); HRMS (FAB) (M^+) m/z calcd for $C_{40}H_{25}N_6Se_2$ 749.0480, found 749.0491. Data for diselenasapphyrin **3**: solid; mp > 250 °C; ^{1}H NMR ($CDCl_3$) δ 10.50 (d, J = 5.1 Hz, 2H), 10.38 (d, J = 5.4 Hz, 2H), 10.12 (d, J = 3.8 Hz, 2H), 9.77 (s, 2H), 9.67 (s, 2H), 9.38 (d, J = 4.5 Hz, 2H), 9.27 (d, J = 3.9 Hz, 2H), 9.22 (d, J = 3.8 Hz, 2H), 9.00 (s, 2H), 8.82 (d, J = 4.7 Hz 2H), 8.74−8.71 (m, 2H), 8.05−7.97 (m, 2H), 7.94−7.90 (m, 2H); UV−vis (CH_2Cl_2) λ_{max} (nm) (ϵ (mol^{-1} cm^{-1})) 476 (203500), 593 (15900), 632 (8700), 733 (4000), 821 (14100); HRMS (FAB) (M^+) m/z calcd for $C_{44}H_{28}N_7Se_2$ 814.0747, found 814.0745.

d. Preparation of 5,10,15,20-Tetra(*N*-methyl-3-pyridyl)-26,28-diselenasapphyrin Chloride (Se2SAP). A 15.0 mg (0.018 mmol) sample of diselenasapphyrin **3** was dissolved in 5.0 mL of chloroform and the resulting solution diluted with 3.0 mL of nitromethane. A 3.0 mL (20 mmol) sample of iodomethane was added, and the mixture was heated to reflux under argon for 24 h. After removal of solvent to dryness, the residue was washed thoroughly with chloroform. A 20 mL sample of water was added to the residue, which was then treated with 2.0 g of Dowex 1×2−200 anion-exchange resin in chloride form, shaking slowly for 2 h. The resin was filtered off and washed with water, and the filtrate was lyophilized to give the chloride salt Se2SAP (70%). The Se2SAP obtained was further purified by preparative HPLC: solid; mp > 250 °C; ^{1}H NMR (D_2O) δ 10.84 (d, J = 4.2 Hz, 2H), 10.66 (d, J = 4.2 Hz, 2H), 10.59 (br s, 2H), 10.01 (br d, 2H), 9.91 (br d, 2H), 9.62 (br s, 2H), 9.52 (m, 4H), 9.36 (m, 4H), 9.33 (s, 2H), 8.64 (m, 2H), 8.59 (m, 2H), 4.76 (s, 6H) 4.71 (s, 6H); UV−vis (DMSO/TFA) λ_{max} (nm) (ϵ (mol^{-1} cm^{-1})) 484 (431000), 599 (35200), 634 (18700), 729 (11400), 830 (24700); HRMS (ESI) [(M − 2H)/2] m/z calcd for $C_{48}H_{37}N_7Se_2/2$ 435.5726, found 435.5721.

Molecular Modeling Studies of c-MYC G-Quadruplex-Interactive Agents. a. Structural Model of the Pu27 c-MYC G-Quadruplex and G_2T_4 and MD Simulations. Since no structural information such as NMR or X-ray data is available for the c-MYC mixed parallel/antiparallel quadruplex, the c-MYC G-quadruplex structure was modeled using X-ray crystal structure coordinates of the human telomeric sequence d[$AG_3(T_2AG_3)_3$] (PDB entry 1KF1)[4b] and the solution structure of the *Tetrahymena* telomeric G-quadruplex d(T_2G_4)$_4$ (PDB entry 186D).[7a] Necessary deletions and replacement of bases were performed on the human telomeric parallel G-quadruplex to generate the parallel strands of the c-MYC G-quadruplex. Orientations and conformations of guanines G7, G8, and G9 were taken from the *Tetrahymena* telomeric G-quadruplex.[7a] The single-base loop of T10 was modeled using standard B-DNA geometries. All modeling was carried out using the Biopolymer module of the Insight II modeling software (Accelrys, Inc.).[24]

The sequence d[$(G_2T_4)_3GG$] is known to form a basket-type G-quadruplex in the presence of sodium or potassium ions. The solution structure of the *Oxytricha* telomeric basket-type G-quadruplex (PDB entry 201D) was used as a starting structure.[25] Necessary replacements and deletions of bases were carried out using the Biopolymer module of Insight II.

The structures of the *Tetrahymena* telomeric G-quadruplex (PDB entry 186D)[7a] and TBA (PDB entry 1C34)[7f] were used from the Protein Data Bank. Necessary sodium or potassium ions were added as part of the quadruplex structure.

Modeled G-quadruplex structures were subjected to energy refinement protocols. All polar hydrogen atoms were energy minimized with the Amber force field of Discover_3.0 using 2500 steps of conjugate gradient minimizer.[26] This was followed by subjecting the entire structural model to 2 × 2500 steps of conjugate gradient minimization.

The fully refined quadruplex structure was neutralized by adding the appropriate number of sodium ions, and this structure was immersed in a 10 Å layer of TIP3P water molecules.[27] Two-stage MD simulations were performed at 300 K. This involved restrained MD simulations with 40 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 40 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.

b. Structure of Se2SAP. The Se2SAP molecule was built using known crystal structures with similar ring systems.[16b] The porphyrin analogues were optimized using the extensible systematic force field (ESFF) implemented in Discover_3.[28] Positive charges for the 3-pyridyl-*N*-methyl groups were added using the charge option defined within Insight II.

c. Docking of Se2SAP with the Different G-Quadruplex Structures. Docking of Se2SAP with the single-loop hybrid G-quadruplex structure was performed using Insight II. Noncovalent complexes of the Pu18-mer with Se2SAP (1:1 complex) were modeled on the basis of photocleavage data for c-MYC. Se2SAP was manually placed between the G-tetrads and the lateral T10 loop. Various orientations of Se2SAP were obtained by rotation and translation. Orientations showing the maximum interactions between Se2SAP and the G-quadruplex were selected. These orientations were then subjected to 2500 steps of minimization using Discover_3. The orientation that gave the better interaction energy value and more favorable docking was selected for further MD simulations.

The complex structure obtained from docking was hydrated with a 10 Å layer of TIP3P water molecules. This structure was then subjected to a molecular dynamics protocol with 40 ps equilibration and 100 ps simulations at 300 K using the Verlet method and NVT ensemble.[28] The ESFF parameters were used during the simulations. Trajectories were collected after every 100 fs, and the lowest energy trajectories were collected from the entire set. The 10 lowest potential energy structures were selected and energy minimized using 2500 steps of conjugate gradient minimization using ESFF within Insight II. These energy-minimized structures were then used to observe the docking orientation of Se2SAP and calculate the intermolecular interaction energy values. The best structure was selected on the basis of the better interaction energy value and optimum interactions between Se2SAP and the G-quadruplex. The intermolecular interaction energy was calculated as the difference between the energy of the complex in an aqueous environment and the individual energies of the quadruplex and Se2SAP.

Procedures similar to the above were used for docking Se2SAP with the TBA, TetTel, and G_2T_4 sequences. For G-quadruplex structures where more than one binding site is available, docking was performed separately for each site using the protocol described earlier. A similar protocol was used for MD simulations.

(26) 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*, 5179−519.
(27) Jorgensen, W. L.; Chandresekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926−935.
(28) Discover_3 ESFF (extensible systematic force field), Molecular mechanics force fields, Insight II, 2000.3L, Molecular Modeling Software, Accelrys Inc., 9685 Scranton Rd., San Diego, CA 92121.

Immobilization of DNA and Biosensor SPR Experiments. Biosensor SPR experiments were performed with a four-channel BIAcore 3000 optical biosensor system (BIAcore, Inc.) and streptavidin-coated sensor chips (BIAcore SA or CM5 with linked streptavidin). All DNA samples, for either duplex- or quadruplex-binding experiments, were used as single strands (fold-back structures) to prevent dissociation in the SPR flow system. The concentration in all cases refers to the strand concentration, which is also the duplex or quadruplex concentration. The chips were prepared for use by conditioning with three to five consecutive 1 min injections of 1 M NaCl in 50 mM NaOH followed by extensive washing with buffer. 5′-Biotinylated DNA samples (25 nM) in HBS buffer were immobilized on the flow cell surface by noncovalent capture as previously described.[19] Three flow cells were used to immobilize DNA oligomer samples, and a fourth cell was left blank as a control. Interaction analysis was performed by using both kinetics and steady-state methods with multiple injections of different compound concentrations over the immobilized DNA surface at 25 °C. DNA-binding experiments were performed in two sterile filtered and degassed HBS buffers: 0.01 M HEPES, (pH 7.4), 3 mM EDTA, and 0.005% surfactant P20 with either 0.2 M KCl or 0.2 M NaCl. Compound solutions were prepared in the desired buffer by serial dilutions from stock solution and injected from 7 mm plastic vials with pierceable plastic crimp caps (BIAcore, Inc.). To remove any remaining bound compound after the dissociation phase of the sensorgram, a low-pH glycine regeneration buffer was used (10 mM glycine at pH 2). The baseline was then reestablished, and the next compound concentration sample was injected.

The instrument response (RU) in the steady-state region is proportional to the amount of bound drug and was typically determined by linear averaging over a 10−20 s or longer time span, depending on the length of the steady-state plateau. The predicted maximum response per bound compound in the steady-state region (RU_{max}) was determined from the DNA molecular weight, the amount of DNA on the flow cell, the compound molecular weight, and the refractive index gradient ratio of the compound and DNA, as previously described.[20c] In most of the cases, the observed RU values at high concentrations were greater than RU_{max}, pointing to more than one binding site in these DNA sequences. The number of binding sites was estimated from Scatchard plots of RU/C_{free} versus RU or from fitting plots of RU versus C_{free}. These methods can also be used to determine an empirical RU_{max} value. The RU_{max} value is required to convert the observed response (RU) to the standard binding parameter r (moles of drug bound per moles of DNA hairpin)

$$r = RU/RU_{max}$$

which is useful for comparison of a compound binding to different DNAs. To obtain the affinity constants, the data were fitted to the following interaction model using Kaleidagraph for nonlinear least-squares optimization of the binding parameters:

$$r = (K_1C_{free} + 2K_1K_2C_{free}^2)/(1 + K_1C_{free} + K_1K_2C_{free}^2)$$

where K_1 and K_2 are equilibrium constants for two types of binding sites and C_{free} is the concentration of the compound in equilibrium with the complex and is fixed by the concentration in the flow solution. For a single dominant binding site model, K_2 is equal to zero.

After injection of the compound and prior to reinjection of buffer, the association kinetics of compound binding to immobilized DNA can be monitored, and after reinjection of buffer the dissociation kinetics can be determined. To minimize possible mass transport effects, SPR kinetics experiments were conducted at flow rates of 50−100 μL/min with low surface densities of immobilized DNA. Global fitting (BIA Evaluation Software, BIAcore, Inc.) of the association and dissociation curves was done in a concentration range where the compound binds significantly only to the strongest binding site: up to 400 nM for G_2T_4 and 100 nM for TetTel and TBA. Both the association rate constant, k_a, and the dissociation rate constant, k_d, could be obtained for these complexes, and the calculated equilibrium constant ($K_{kin} = k_a/k_d$) was within a factor of 2 of the steady-state value. For binding of Se2SAP to duplex DNAs, the kinetics of association and dissociation are too fast for accurate analysis by SPR methods. For binding of the compound to the c-MYC sequence, we were unable to obtain satisfactory global fits, on the basis of the analysis of residuals and quality of fit values, for the lowest concentrations required in the 1:1 binding region with this strong interaction. In the low-concentration c-MYC association region, the curves were initially slightly concave, which may be due to a fraction of the compound adsorbing to the flow system at injection. This problem becomes much less significant at the higher concentrations used for the weaker binding DNAs. Since the compound is discarded on dissociation, the adsorption problem is less significant, and a k_d value can be accurately determined for the c-MYC complex. With the steady-state K and the k_d value, a k_a value can be calculated, but with less confidence than for the experimentally determined values.

Materials for Biochemistry. Compound solutions were prepared as 1 mM stock solutions in water for porphyrins and in dimethyl sulfoxide for telomestatin and stored at −20 °C. These stock solutions were diluted to working concentrations in distilled water immediately before use. Electrophoretic reagents (acrylamide/bisacrylamide solution and ammonium persulfate) were purchased from BioRad, and *N*,*N*,*N*′,*N*′-tetramethylethylenediamine 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.

Preparation and End-Labeling of Oligonucleotides. The oligonucleotides were obtained PAGE purified from Sigma Genosys, dissolved in autoclaved doubly distilled water (100 μM), and kept at −20 °C for no longer than 6 weeks. The 5′-end-labeled single-strand oligonucleotide was obtained by incubating the oligomer with T4 polynucleotide kinase and [γ-^{32}P]ATP for 1 h at 37 °C. Terminal deoxynucleotidyl transferase was used for the 3′-end-labeling of the single-strand oligonucleotide. Labeled DNA was purified with a Bio-Spin 6 chromatography column (BioRad) after inactivation of the kinase by heating for 8 min at 90 °C.

Polymerase Stop Assay. The DNA primer d[TAATACGACTCACTATAGCAATTGCGTG] and the Pu27 c-MYC template sequence d[TCCAACTATGTATAC(TGGGGAGGGTGGGGAGGGTGGGGAAGG)TTAGCGGCACGCAATTGCTATAGTGAGTCGTATTA] were purified as mentioned above. 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 cooling 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 a reaction buffer (10 mM $MgCl_2$, 0.5 mM DTT, 0.1 mM EDTA, 1.5 μg/μL BSA) and 0.1 mM deoxynucleotide triphosphate. KCl and NaCl (10 mM each) were added to the reaction. The drugs were added, and the reaction mixture was incubated for 20 min at room temperature. *Taq* DNA polymerase was added, and the mixture was incubated at 55 °C for 30 min. The polymerase extension was stopped by adding 2× stop buffer (10 mM EDTA, 10 mM NaOH, 0.1% xylene cyanole, 0.1% bromophenol blue in formamide solution) and loaded onto a 16% denaturing gel.

Electromobility Shift Assay. 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 4 h with Se2SAP in the absence of light at room

temperature and then subjected to preparative gel electrophoresis (15% polyacrylamide, 25 mM KCl, 20 h, 4 °C).

Photomediated Strand Cleavage Reaction and DMS Footprinting. The 27-mer sequence d[TGTTGAGGGTGGGTAGGGTGGGTAAGG] (G → T-14,23) was 5′-end-labeled and diluted to a final concentration of 2 μM. The 5′-end-labeled sequence 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]. A DNA solution [50000 cpm (per reaction)] was diluted further with a buffer [10 mM Tris/1 mM EDTA (pH 7.5), 100 mM KCl]. One portion of this solution was used for the DMS footprinting assay, as explained in ref 1. A 10 μM final concentration of Se2SAP was added to the other portion of the DNA solution, and this was incubated at room temperature for 2 h in the absence of light and then transferred to a 24-well Titertek microtiter plate (ICN). This plate was placed on top of a Pyrex glass shield and irradiated for 1 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 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 Storm 820 PhosphorImager and ImageQuant 5.1 software (Molecular Dynamics).

Competition Assay. The 5′-fluorescent (FAM) labeled primer (15 nM) was mixed with template DNA containing the c-MYC G-quadruplex sequence (15 nM) in a Tris−HCl buffer (15 mM Tris, pH 7.5) containing 10 mM $MgCl_2$, 0.1 mM EDTA, and 0.1 mM mixed deoxynucleotide triphosphates. The mixture was denatured at 95 °C for 5 min and, after being cooled to room temperature, was incubated at 37 °C for 15 min. After the mixture was again cooled to room temperature, 1 mM KCl and various concentrations of G-quadruplex-binding ligand were added. The primer extension was performed by adding a further 10 mM KCl and *Taq* DNA polymerase (2.5 U/reaction) and incubating at 65 °C for 30 min. The reaction was stopped by adding 1 μL of the reaction mixture to 10 μL of Hi-Di Formamide and 0.25 μL of LIZ120 (size standard). The reactions were conducted in a 96-well plate, and the products were separated and analyzed using capillary electrophoresis. An IC_{50} value for the G-quadruplex-binding ligand was determined as the concentration of the ligand required to induce a 1:1 ratio of G-quadruplex arrest to full-length product.

For the competition experiments, the method was essentially the same, with the addition of various concentrations of competitor nucleic acids at the first step, along with the primer and template sequences. The G-quadruplex binding ligand was added at the concentration previously established to produce a 1:1 ratio of stop product to full-length product (IC_{50}). A CC_{50} for each nucleic acid competitor is defined as the concentration of competitor required to change the ratio of arrest product to full-length product from 1:1 to 1:2. Column charts are generated by inverting the IC_{50} values for each competitor.

Cytotoxicity Assay. In vitro cytotoxicity assays were performed using the CellTiter 96 nonradioactive cell proliferation assay (Promega Corp., Madison, WI). Cells were plated in 0.1 mL of medium on day 0 in 96-well microtiter plates (Falcon, no. 3072). On day 1, 10 μL of serial dilutions of the investigational compound was added in replicates of four to the plates. After incubation for 4 days at 37 °C in a humidified incubator, 20 μL of a 20:1 mixture of 2 mg/mL MTS [3-(4,5-dimethyl-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-tetrazolium, inner salt] and an electron coupling reagent, phenazine methosulfate (0.92 mg/mL in DPBS), was added to each well followed by incubation for 4 h at 37 °C. Absorbance was measured using a Wallac Victor multilabel counter (Perkin-Elmer) at 490 nm. Data were expressed as the percentage of survival of the control calculated from the absorbance corrected for background absorbance. The surviving fraction of cells was determined by dividing the mean absorbance values of the test agents by the mean absorbance values of the untreated control.

Cell Culture, Human Cells. HeLa S_3 (human cervical carcinoma metastasis) was obtained from ATCC and cultured in Dulbecco's modified Eagle's medium (DMEM; Cellgro) with 10% FBS, 50 U/mL penicillin G sodium, and 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 by vacuum, and the cells were washed with 1× PBS (Cellgro). Sufficient trypsin (Gibco/BRL) was added to cover the cells, and the cells were incubated at room temperature for approximately 3 min, or until the 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 and passaged at 1:10 into fresh medium. The remaining cells were then pelleted by centrifugation at 500*g*, the supernatant was aspirated, and the pellets were washed in PBS, recentrifuged, and frozen at −80 °C for storage.

Treatment with Drugs. Cell cultures were permitted to reach ∼50% confluency before drug was added for 24 h treatments. This was adjusted as necessary, depending on the treatment times. The drugs were diluted in media in which the cells were normally cultured. The cells were washed once with PBS, and new medium containing drug was added directly to the flask. The cells were harvested as noted, concurrently with untreated cells. Time points were collected in duplicate for each treatment.

Acknowledgment. This research was supported by the National Institutes of Health (Grants CA94166, CA95060, and CA88310) and the Arizona Disease Control Research Commission (Grant 5015). We thank Mary Gleason-Guzman for carrying out the cytotoxicity experiments. We are grateful to David Bishop for preparing, proofreading, and editing the final version of the manuscript and figures.

Supporting Information Available: ^{1}H NMR for Se2SAP and other precursors, HPLC for Se2SAP, SPR sensograms for the binding of Se2SAP with TetTel, TBA, and G_2T_4 sequences (PDF), one modeling figure, and one table. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0444482