2010

Identification of natural product Fonsecin B as a stabilizing ligand of c-myc G-quadruplex DNA by high-throughput virtual screening

Ho-Man Lee

Daniel Shiu-Hin Chan

Fang Yang

Ho-Yu Lam

Siu-Cheong Yan

See next page for additional authors

Identification of Natural Product Fonsecin B as a Stabilizing Ligand of c-*myc* G-quadruplex DNA by High-Throughput Virtual Screening†

Ho-Man Lee,‡[a] Daniel Shiu-Hin Chan,‡[a] Fang Yang,[a] Siu-Cheong Yan,[a] Dik-Lung Ma*[a,b] and Chung-Hang Leung*[a]

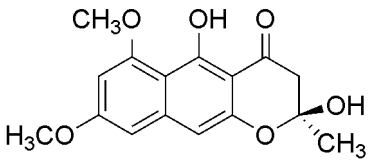
Received (in XXX, XXX) Xth XXXXXXXXX 200X, Accepted Xth XXXXXXXXX 200X`
First published on the web Xth XXXXXXXXX 200X
DOI: 10.1039/b000000x

Fonsecin B has been identified as a stabilizing ligand of c-*myc* G-quadruplex DNA using high-throughput virtual screening of a natural product database, and was found to inhibit *Taq* polymerase-mediated DNA extension *in vitro* through stabilization of the G-quadruplex secondary structure.

G-quadruplexes are DNA secondary structures formed from guanine-rich sequences, consisting of a planar arrangement of four guanines stabilized by Hoogsteen hydrogen bonding.[1] G-quadruplex DNA can be found throughout the human genome, particularly in the telomeres and in the promoter regions of growth control genes.[2] Notably, G-quadruplex promoter sequences have been identified for several proto-oncogenes, including c-*myc*, *bcl*-2, *VEGF*, *KRAS*, and c-*kit*.[3] Thus, there has been an interest in the development of small molecule ligands that can stabilize specific promoter G-quadruplex DNA sequences. Examples include c-*myc* G-quadruplex DNA stabilization by cationic porphyrins,[3a,4] quindoline derivatives[5] and platinum complexes,[6] while isoalloxazines[7] and triarylpyridine compounds[8] have been demonstrated to bind c-*kit* G-quadruplex DNA.

Nature provides a wonderfully rich source of chemically diverse scaffolds for the medicinal chemist.[9] It is estimated that approximately 50% of all approved drugs in the years 2000–2006 were related to the natural products field in some way.[9a] While the natural product telomestatin is the most potent G-quadruplex ligand known to date,[10] the majority of G-quadruplex binding ligands reported in the literature have been synthetic in origin.[11] Encouraged by the high molecular diversity of natural products, and their potentially lower toxicity profiles compared to synthetic compounds, we sought to develop high-throughput, ligand docking-based virtual screening methods to identify small molecule leads from natural product databases. Using this computer-aided approach, potent and specific ligands can be rapidly identified, potentially reducing the number of compounds that need to be evaluated *in vitro*. We have previously performed high-throughput virtual screening on a telomeric G-quadruplex DNA model to identify drug-like G-quadruplex ligands.[12]

The nuclear hypersensitivity element III_1 (NHE III_1) is a guanine-rich 27 base-pair sequence located upstream of the c-*myc* P1 promoter, and controls 80–90% of c-*myc* transcription.[13] In order to develop a high-throughput screening platform for G-quadruplex binding ligands, a model of the intramolecular G-quadruplex loop isomer of NHE III_1 was constructed using the X-ray crystal structure of the intramolecular human telomeric G-quadruplex DNA (PDB code: 1KF1). This NHE III_1 G-quadruplex isomer model has been previously employed to design quindoline compounds that stabilize c-*myc* G-quadruplex DNA.[5] As part of our effort, the aforementioned model has been utilized in conjunction with *in silico* hit-to-lead optimization to develop new platinum(II) Schiff-base complexes with superior c-*myc* regulatory properties, presumably through stabilization of the NHE III_1 G-quadruplex structure.[6]



1 Fonsecin B

Fig. 1 The chemical structure of fonsecin B (**1**).

Using this approach, over 20,000 compounds in a natural product database[14] were screened *in silico*. The continuously flexible ligands were docked to a grid representation of the receptor and assigned a score reflecting the quality of the complex according to the ICM method [ICM-Pro 3.6-1d molecular docking software (Molsoft)].[15] The 5 highest-scoring compounds (Fig. S1†) were tested in a preliminary polymerase stop assay (*vida infra*) to assess their ability to stabilize the c-*myc* G-quadruplex, and fonsecin B (**1**) emerged as the top candidate (Fig. 1). Fonsecin B **1** is a naphthopyrone pigment originally isolated from the fungus *Aspergillus fonsecaeus* in 1974.[16a] The related naphthopyrone fonsecin[16b,c] (8-*O*-desmethyl **1**) moderately inhibits interleukin-4 (IL-4) signal transduction, displaying IC_{50} values of 30 μM and 34 μM against IL4-induced STAT6-driven gene transcription and germline Cε expression, respectively, whereas **1** itself showed no inhibition under the same conditions.[17] To the best of our knowledge, no G-quadruplex DNA-binding activity nor any other biological activity of fonsecin B **1** has been reported in the literature.

The ability of **1** to interact with c-*myc* G-quadruplex DNA was first studied by an absorption titration experiment. An intramolecular c-*myc* G-quadruplex structure was prepared by incubating oligonucleotide Pu27† in Tris/KCl buffer, which was heated to 95 °C for 10 min and cooled to room temperature overnight. The expected G-quadruplex secondary DNA structure was confirmed by a positive CD peak at 262 nm and a negative CD peak at *ca*. 240 nm. Isosbestic spectral

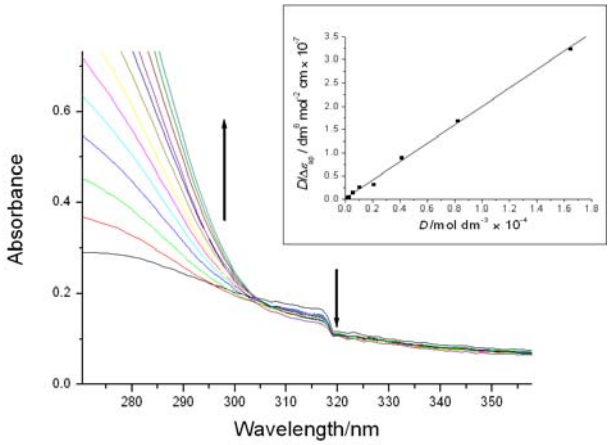


Fig. 2 UV-Vis absorption titration of **1** (50 μM) in Tris/KCl buffer (100 mM KCl, 10 mM Tris-HCl, pH 7.5) with increasing amounts of c-*myc* oligomer Pu27 (0–164 μM). Inset: plot of $D/\Delta\varepsilon_{ap}$ versus D. Absorbance was monitored at 316 nm.

changes and hypochromic effects were observed in the UV-visible absorption spectral changes as depicted in Figure 2. This hypochromic phenomenon is attributed to the strong interaction between fonsecin B **1** and G-quadruplex DNA. Using the Scatchard equation,[18] the plot of $D/\Delta\varepsilon_{ap}$ vs D revealed the binding constant K of **1** with G-quadruplex to be 1.1×10^6 dm^3 mol^{-1} at 20.0 °C.[19] To investigate the selectivity of **1** for G-quadruplex DNA over duplex DNA, a parallel UV-visible absorption titration experiment with calf thymus (ct) DNA was performed (Fig. S2†). The K value for ct DNA was calculated to be 2.0×10^5 dm^3 mol^{-1}, which is approximately 5.5-fold lower than the binding constant of **1** for the c-*myc* G-quadruplex. The binding constant of **1** for a random single-stranded 22-bp sequence SS22† was similarly determined to be 6.9×10^4 dm^3 mol^{-1} (Fig. S3†). Taken together, these data reveal that fonsecin B **1** selectively binds c-*myc* G-quadruplex DNA over duplex and single-stranded DNA, with a 5.5- and 16.5-fold higher binding affinity, respectively, as determined from UV-visible absorption titration.

We performed molecular modeling of fonsecin B **1** with the NHE III$_1$ intramolecular G-quadruplex loop isomer model in order to further investigate the mode of binding. Hurley[20a] and Patel[20b] showed the predominance of the 1:2:1 loop isomer in the c-*myc* parallel G-quadruplex structure. Since neither NMR nor X-ray crystallographic information for the NHE III$_1$ 1:2:1 loop isomer is available, a model was built from the known, closely related X-ray crystal structure of the human intramolecular telomeric G-quadruplex DNA. A truncated 18 base-pair sequence [5′-AGGGTGGGGAGGGTGGGG-3′] was used for this work since nucleotides G2–G5 in the c-*myc* sequence [5′-TGGGGAGGGTGGGGAGGGTGGGGAAGG-3′] are not involved in the G-quartet structure. The molecular docking results show that **1** binds strongly to c-*myc* G-quadruplex DNA with a binding energy of –47.88 kcal mol^{-1}, and that the relatively flat scaffold of **1** is stacked on the ends of the G-quadruplex at the 3′-terminus (Figure 3). Interestingly, in this model the phenolic and carbonyl oxygen atoms of **1** are situated close (4.8–5.1 Å) to a potassium ion in the central channel of the G-quadruplex, possible contributing favourable electrostatic interactions to the binding score. The binding energies for intercalation (*ca*. 25 kcal mol^{-1}) are markedly higher, suggesting that fonsecin B **1** is unlikely to bind c-*myc* G-quadruplex DNA through an intercalative binding mode. Stacking near the 5′-terminal (–36.49 kcal mol^{-1}) is also predicted to be disfavoured relative to 3′-end stacking (Table S1†).

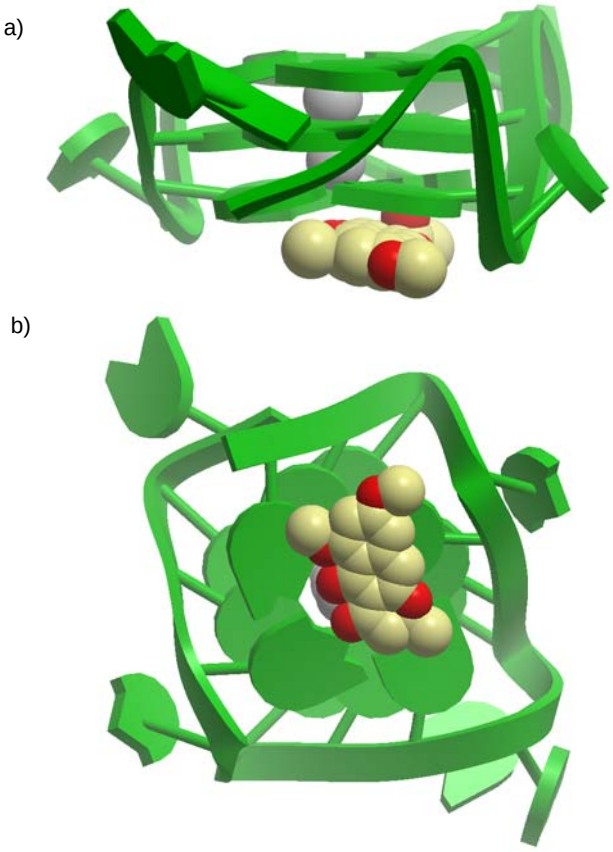


Fig. 3 Hypothetical molecular models showing the a) Side view; b) Top view of the interactions of fonsecin B **1** with the c-*myc* G-quadruplex structure. The G-quadruplex is depicted as a ribbon representation (green), while **1** is depicted as a space-filling representation showing carbon (beige) and oxygen (red) atoms. Potassium ions (grey) are located in the central ionic channel of the G-quadruplex.

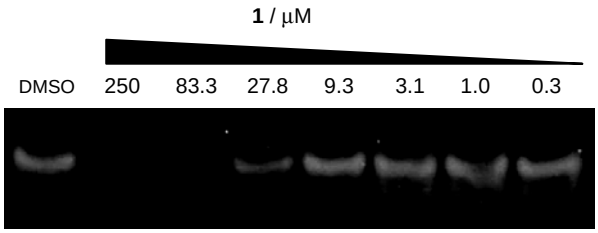


Fig. 4 Incubation of c-*myc* G-quadruplex oligomer with fonsecin B **1** (0.3–250 μM) caused a dose-dependent decrease of the PCR amplification product at 43 bp. Representative gel photograph image of duplicate independent experiments shown.

To validate our molecular modeling results, a dose response experiment with **1** was performed using the polymerase stop assay. In this assay, fonsecin B **1** was incubated with oligomer Pu27 in the presence of *Taq* polymerase. Formation of the intramolecular G-quadruplex prevents DNA annealing and subsequent *Taq* polymerase-mediated DNA extension. This is manifested as a reduction in the 43 bp PCR product observed after agarose gel electrophoresis. The results show that the addition of fonsecin B **1** led to a dose-dependent decrease in the 43 bp PCR product, with complete inhibition observed at

83 μM of **1** (Figure 4). The IC_{50} value of **1** against *Taq* polymerase-mediated DNA amplification due to stabilization of the c-*myc* G-quadruplex is estimated to be approximately 20 μM. To confirm the impact of fonsecin B **1** on c-*myc* G-quadruplex stabilization, a parallel experiment using oligomer Pu27m† was performed. This oligonucleotide is unable to form the G-quadruplex structure. Under the same reaction conditions, no inhibition was observed (data not shown). Importantly, fonsecin B **1** displayed comparable potency to the well-known G-quadruplex binder TMPyP4[3a] under the same experimental conditions (Fig. S4†). We envisage that this natural product scaffold may serve as a useful template for future lead optimization studies.

In conclusion, this intramolecular G-quadruplex NHE III_1 loop isomer model, constructed using the X-ray crystal structure of the intramolecular human telomeric G-quadruplex DNA, is a unique model developed by our group that has been successfully validated in previous reports. We have utilized this model in the present study to perform high-throughput virtual screening on a natural product library of over 20,000 compounds. Fonsecin B **1** emerged as the top candidate and was demonstrated to inhibit *Taq*-mediated extension through stabilization of the G-quadruplex secondary structure (IC_{50} = *ca.* 20 μM), with potency comparable to TMPyP4.[21] To our knowledge, this is the first large-scale application of high-throughput virtual screening of a natural product database for c-*myc* G-quadruplex stabilizing ligands. Computer-based hit-to-lead optimization is currently being carried out in order to generate further analogues for *in vitro* testing.

This work is supported by the Area of Excellence Scheme established under the University Grants Committee of the Hong Kong Special Administrative Region, China (AoE/P-10/01), the University of Hong Kong (University Development Fund), the University of Hong Kong Seed Funding Programme for Applied Research, and the University of Hong Kong Seed Funding Programme for Basic Research.

Notes and references

[a] *Department of Chemistry and Open Laboratory of Chemical Biology of the Institute of Molecular Technology for Drug Discovery and Synthesis, The University of Hong Kong, Pokfulam Road, Hong Kong. E-mail: duncanl@hku.hk*

[b] *Department of Chemistry, Hong Kong Baptist University, Kowloon Tong, Hong Kong (China). E-mail: edmondma@hkbu.edu.hk*

† Electronic Supplementary Information (ESI) available: Experimental details, chemical structures, UV-visible absorption titration spectra, and polymerase stop assay with TMPyP4. See DOI: 10.1039/b000000x/

‡ These authors contributed equally to this work.

* D.-L. Ma and C.-H. Leung contributed equally to this work.

1 (*a*) G. N. Parkinson, M. P. H. Lee and S. Neidle, *Nature*, **2002**, *417*, 876–880; (*b*) T. Simonsson, *Biol. Chem.*, 2001, **382**, 621–628.

2 (*a*) A. T. Phan and J.-L. Mergny, *Nucleic Acids Res.*, 2002, **30**, 4618–4625; (*b*) J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.,* 2007, **35**, 406–413; (*c*) J.-L. Mergny, C. Hélène, *Nat. Med.*, 1998, **4**, 1366–1367.

3 (*a*) A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci.*, 2002, **99**, 11593–11598; (*b*) J. Dai, D. Chen, R. A. Jones, L. H. Hurley and D. Yang, *Nucleic Acids Res.*, 2006, **34**, 5133–5144; (*c*) D. Sun, K. Guo, J. J. Rusche and L. H. Hurley, *Nucleic Acids Res.*, 2005, **33**, 6070–6080; (*d*) S. Cogoi and L. E. Xodo, *Nucleic Acids Res.*, 2006, **34**, 2536–2549; (*e*) S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian and S. Neidle, *J. Am. Chem. Soc.*, 2005, **127**, 10584–10589.

4 D. P. N. Gonçalves, R. Rodriguez, S. Balasubramanian and J. K. M. Sanders, *Chem. Commun.*, 2006, 4685–4687.

5 T.-M. Ou, Y.-J. Lu, C. Zhang, Z.-S. Huang, X.-D. Wang, J.-H. Tan, Y. Chen, D.-L. Ma, K.-Y. Wong, J. C.-O. Tang, A. S.-C. Chan and L.-Q. Gu, *J. Med. Chem.*, 2007, **50**, 1465–1474.

6 P. Wu, D.-L. Ma, C.-H. Leung, S.-C. Yan, N. Zhu, R. Abagyan and C.-M. Che, *Chem.–Eur. J.*, 2009, **15**, 13008–13021.

7 M. Bejugam, S. Sewitz, P. S. Shirude, R. Rodiguez, R. Shadid and S. Balasubramanian, *J. Am. Chem. Soci.*, 2007, **129**, 12926–12927.

8 Z. A. E. Waller, S. A. Sewitz, S.-T. D. Hsu and S. Balasubramanian, *J. Am. Chem. Soc.*, 2009, **131**, 12628–12633.

9 (*a*) D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2007, **70**, 461–477; (*b*) F. E. Koehn and G. T. Carter, *Nat. Rev. Drug. Discov.*, 2005, **4**, 206–220; (*c*) J. W.-H. Li and J. C. Vederas, *Science*, 2009, **325**, 161–165; (*d*) M. Butler, *Nat. Prod. Rep.*, 2008, **25**, 475–516.

10 (*a*) K. Shin-ya, K. Wierzba, K. Matsuo, T. Ohtani, Y. Yamada, K. Furihata, Y. Hayakawa and H. Seto, *J. Am. Chem. Soc.*, 2001, **123**, 1262–1263; (*b*) M.-Y. Kim, M. Gleason-Guzman, E. Izbicka, D. Nishioka and L. H. Hurley, *Cancer Res.*, 2003, **63**, 3247–3256.

11 (*a*) T.-M. Ou, Y.-J. Lu, J-.H. Tan, Z.-S. Huang, K.-Y. Wong and L.-Q. Gu, *Chem. Med. Chem.*, 2008, **3**, 690–713; (*b*) S. Balasubramanian and S. Neidle, *Curr. Opin. Chem. Biol.*, 2009, **13**, 345–353; (*c*) J. Dash, P. S. Shirude and S. Balasubramanian, *Chem. Commun.*, 2008, 3055–3057; (*d*) I. M. Dixon, F. Lopez, A. M. Tejera, J.-P. Estève, M. A. Blasco, G. Pratviel and B. Meunier, *J. Am. Chem. Soc.*, 2007, **129**, 1502–1503; (*e*) Q. Yang, J. Xiang, S. Yang, Q. Zhou, Q. Li, Y. Tang and G. Xu, *Chem. Commun.*, 2009, 1103–1105; (*f*) D. Monchaud and M.-P. Teulade-Fichou, *Org. Biomol. Chem.*, 2008, **6**, 627–636; (*g*) S Müller, G. D. Panto, R. Rodriguez and Shankar Balasubramanian, *Chem. Commun.*, 2009, 80–82; (*h*) I. M. Dixon, F. Lopez, J.-P. Estève, A. M. Tejera, M. A. Blasco, G. Pratviel and B. Meunier, *ChemBioChem*, 2005, **6**, 123–132; (*i*) P. Yang, A. De Cian, M.-P. Teulade-Fichou, J.-L. Mergny and D. Monchaud, *Angew. Chem. Int. Ed.*, 2009, **48**, 2188–2191; (*j*) D.-L. Ma, C.-M. Che and S.-C. Yan, *J. Am. Chem. Soc.*, 2009, **131**, 1835–1846.

12 D.-L. Ma, T.-S. Lai, F.-Y. Chan, W.-H. Chung , R. Abagyan, Y.-C. Leung and K.-Y. Wong, *ChemMedChem*, 2008, **3**, 881–884.

13 (*a*) E. H. Postel, S. E. Mango and S. J. Flint, *Mol. Cell. Biol.*, 1989, **9**, 5123–5133; (*b*) T. L. Davis, A. B. Firulli and A. J. Kinniburgh, *Proc. Natl. Acad. Sci .USA*, 1989, **86**, 9682–9686.

14 This chemical library was obtained from AnalytiCon Discovery GmbH. See the ESI for details.

15 M. Totrov and R. Abagyan, *Proteins Suppl*., 1997, **29**, 215–220.

16 (*a*) O. L. Galmarini, I. O. Mastronardi and H. A. Priestap, *Cell. Mol. Life. Sci.*, 1974, **30**, 568; (*b*) O. L. Galmarini, F. H. Stodola, K. B. Raper and D. I. Fennel, *Nature*, 1962, **195**, 503–504; (*c*) O. L. Galmarini and F. H. Stodola, *J. Org. Chem.*, 1965, **30**, 112–115.

17 M. Sakurai, J. Kohno, K. Yamamoto, T. Okuda, M. Nishio, K. Kawano and T. Ohnuki, *J. Antibiot*., 2002, **55**, 685–692.

18 C. V. Kumar and E. H. Asuncion, *J. Am. Chem. Soc*., 1993, **115**, 8547–8553.

19 Due to the weak absorbance of **1** (ε_{316} = *ca.* 2000 cm^{-1} M^{-1}) partially overlapping with the absorbance of G-quadruplex DNA at 316 nm, these binding constants are estimated values.

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

21 An RT-PCR experiment was also performed to assess if **1** could inhibit c-*myc* expression in a cellular system. However, no significant inhibition was observed under our experimental conditions.

Electronic Supporting Information

Identification of Natural Product Fonsecin B as a Stabilizing Ligand of c-*myc* G-quadruplex DNA by High-Throughput Virtual Screening

Ho-Man Lee,‡[a] Daniel Shiu-Hin Chan,‡[a] Fang Yang,[a] Siu-Cheong Yan,[a] Dik-Lung Ma*[a,b] and Chung-Hang Leung*[a]

[a] *Department of Chemistry and Open Laboratory of Chemical Biology of the Institute of Molecular Technology for Drug Discovery and Synthesis, The University of Hong Kong, Pokfulam Road, Hong Kong. Email: duncanl@hku.hk.*
[b] *Department of Chemistry, Hong Kong Baptist University, Kowloon Tong, Hong Kong (China). E-mail: edmondma@hkbu.edu.hk*
‡These authors contributed equally to this work.
** D.-L. Ma and C.-H. Leung contributed equally to this work.*

Experimental section

Materials. Calf thymus DNA (ct DNA) was purchased from Sigma Chemical Co. Ltd. and purified by the literature method.[1a] The DNA per base pair concentration was determined by UV/Vis absorption spectroscopy using the following molar extinction coefficients at the indicated wavelengths: calf thymus DNA, ε_{260} = 13200 bp cm^{-1} M^{-1} (base pair).[1b] DNA oligomers were obtained from Tech Dragon Limited (Carlsbad, CA). The sequences for oligomers Pu27, Pu27m and SS22 are:

Pu27 = [5′-TGGGGAGGGTGGGGAGGGTGGGGAAGG-3′]
Pu27m = [5′-TGGGGAGGGTGGAAAGGGTGGGGAAGG-3′]
SS22 = [5′-TAAGCCGCCACATCTTGCGAAT-3′]

Fonsecin B **1** and the other tested compounds was obtained from Analyticon Discovery GmbH (Postdam, Germany). **1** was obtained in 96% purity by HPLC-ELSD (Fig. S5†). This database, containing over 20,000 natural product/natural product-like structures, is publicly available and can be accessed free of charge. Unless otherwise stated, spectroscopic titration experiments were performed in 10 mM Tris/HCl (pH 7.5) containing 100 mM KCl. *Taq* DNA polymerase was purchased from QIAGEN (Valencia, CA). Stock solution of **1** (10 mM) was made in dimethyl sulfoxide (DMSO). Further dilutions to working concentrations were made with double-distilled water.

Physical measurement. Absorption spectra were recorded on a Perkin-Elmer Lambda 19 UV/Visible spectrophotometer.

Absorption titration. A solution of fonsecin B **1** (50 μM) was prepared in Tris/HCl buffer (10 mM, pH 7.4) containing 100 mM KCl, and aliquots of a millimolar stock solution of Pu27 in Tris/KCl buffer (0–164 μM) were added. Absorption spectra were recorded in the spectral range λ = 200–600 nm after equilibration at 20.0 °C for 10 min. The intrinsic binding constant, *K*, was determined from a plot of $D/\Delta\varepsilon_{ap}$ vs *D* according to equation (1):[2]

$$D/\Delta\varepsilon_{ap} = D/\Delta\varepsilon + 1/(\Delta\varepsilon \times K) \qquad (1)$$

where D is the concentration of DNA, $\Delta\varepsilon_{ap} = |\varepsilon_A - \varepsilon_F|$, $\varepsilon_A = A_{obs}/[\text{ligand}]$, and $\Delta\varepsilon = |\varepsilon_B - \varepsilon_F|$; ε_B and ε_F correspond to the extinction coefficients of DNA–ligand adduct and unbound ligand, respectively.

A similar absorption titration experiment was performed using ct DNA (0–24 μM) and SS22 (0–99 μM).

Polymerase stop assay. The polymerase stop assay was performed by using a modified protocol of the previously reported method.[3] The reactions were performed in 1× PCR buffer, containing each pair of oligomers (10 mmol), deoxynucleotide triphosphate (0.16 mM), Taq polymerase (2.5 U), and increasing concentrations of the fonsecin B **1** (from 0.3 to 250 μM). The reaction mixtures were incubated in a thermocycler under the following cycling conditions: 94 °C for 3 min followed by 30 cycles of 94 °C for 30 s, 58 °C for 30 s, and 72 °C for 30 s. The amplified products were resolved on 1.3% agarose gel and visualized by ethidium bromide staining.

Molecular modeling. Molecular docking was performed by using the ICM-Pro 3.6-1d program (Molsoft).[4] According to the ICM method, the molecular system was described by using internal coordinates as variables. Energy calculations were based on the ECEPP/3 force field with a distance-dependent dielectric constant. The biased probability Monte Carlo (BPMC) minimization procedure was used for globalenergy optimization. The BPMC global-energy-optimization method consists of 1) a random conformation change of the free variables according to a predefined continuous probability distribution; 2) local-energy minimization of analytical differentiable terms; 3) calculation of the complete energy including nondifferentiable terms such as entropy and solvation energy; 4) acceptance or rejection of the total energy based on the Metropolis criterion and return to step (1). The binding between **1** and DNA was evaluated by binding energy, including grid energy, continuum electrostatic, and entropy terms. The initial model of loop isomer was built from the X-ray crystal structures of human intramolecular telomeric G quadruplex (PDB code: 1KF1),[5] according to a previously reported procedure.[3,6] Briefly, the structure of human intramolecular telomeric G quadruplex was imported into Insight II package (Accelrys Inc., San Diego, CA), and necessary modifications were carried out including replacements and deletions of bases. Missing loop nucleotides were added using single-strand B-DNA geometry using the Biopolymer module. Potassium ions were placed between the G-tetrad planes to stabilize the tetrad structure. The initial models were then immersed in a box of TIP3P water molecules, and an appropriate number of sodium ions was added to neutralize the negative charge of the phosphate backbone. The molecular dynamics simulations were carried out in NAMD with VMD monitoring the process. The CHARMM force field parameter was assigned to every atom, and the Particle Mesh Ewald electrostatics was used to compute long-range electrostatic interactions. Hydrogen atoms were added and minimized by 3000 steps of conjugate gradient minimization. After 4000 steps of conjugate gradient minimization, two stages of molecular dynamics simulations were carried out at 300 K. In the first stage, only the loop area atoms were allowed to move, and this process involved a 20 ps equilibration and 100 ps simulations. The second stage involved unrestrained molecular dynamics simulations with 20 ps equilibration and 100 ps simulations at 300 K. Trajectories were recorded every 0.1 ps, and the most stable structure was extracted and further refined by 2500 steps of conjugate gradient minimization. In the docking analysis, the binding site was assigned across the entire structure of the DNA molecule. The ICM docking was performed to find the most favorable orientation. The resulting trajectories of the complex between **1** and G-quadruplex DNA were energy minimized, and the interaction energies were computed.

References

1 (a) J. Sambrook, E. F. Fritsch, T. E. Maniatis, *Molecular Cloning, A Laboratory Manual, 2nd Ed.,* **1989**, E.3 and E.10; (b) G. Felsenfeld, S. Z. Hirschman, *J. Mol. Biol.,* **1965**, *13*, 407.

2 C. V. Kumar, E. H. Asuncion, *J. Am. Chem. Soc.,* **1993**, *115*, 8547.

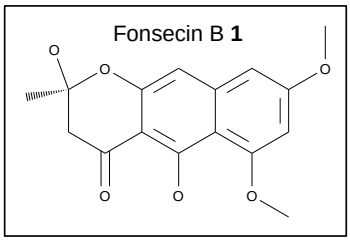
3 T.-M. Ou, Y.-J. Lu, C. Zhang, Z.-S. Huang, X.-D. Wang, J.-H. Tan, Y. Chen, D.-L. Ma, K.-Y. Wong, J. C.-O. Tang, A. S.-C. Chan, L.-Q. Gu, *J. Med. Chem.* **2007**, *50*, 1465.

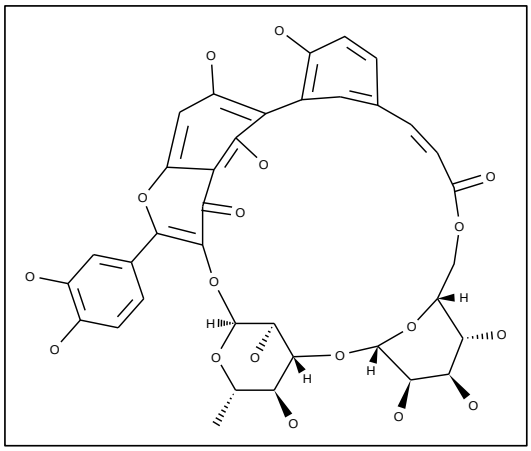
4 a) M. Totrov, R. Abagyan, Proteins **1997**, *29*, 215; b) D.-L. Ma, T.-S. Lai, F.-Y. Chan, W.-H. Chung, R. Abagyan, Y.-C. Leung, K.-Y. Wong, *ChemMedChem* **2008**, *3*, 881.

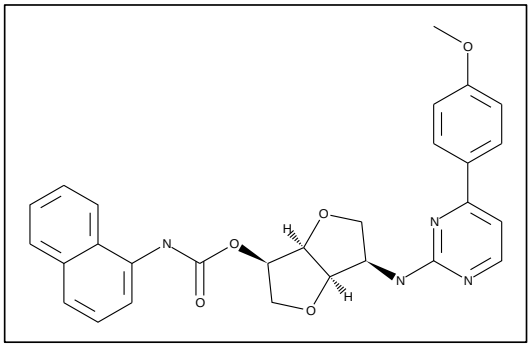
5 G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876.

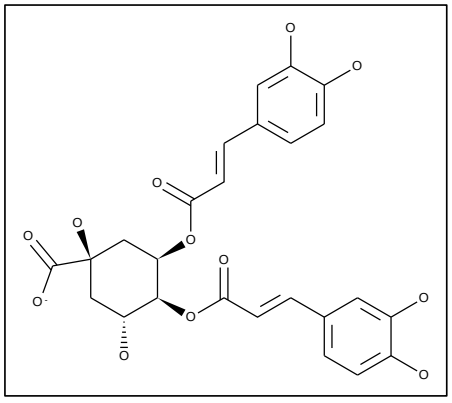
6 P. Wu, D.-L. Ma, C.-H. Leung, S.-C. Yan, N. Zhu, R. Abagyan, C.-M. Che, *Chem.–Eur. J.*, 2009, **15**, 13008.

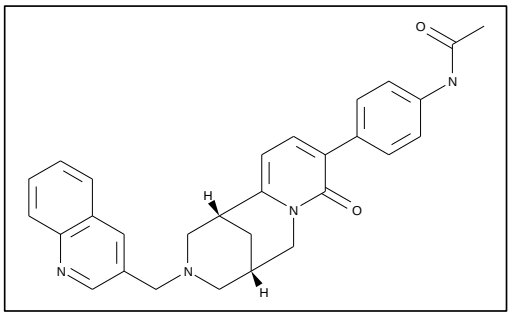
Fig S1. Chemical structures of the top-five scoring compounds (highlighted) identified using high-throughput virtual screening of a natural product database against the c-*myc* G-quadruplex. The structures ranked 6th-10th are also included for comparison.

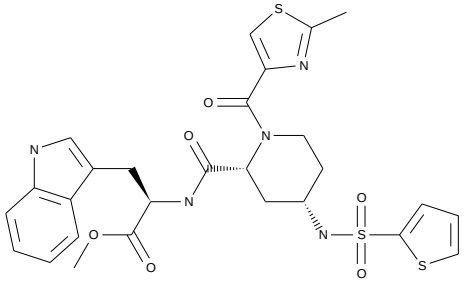


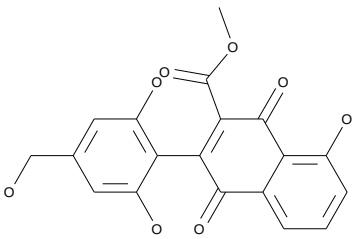


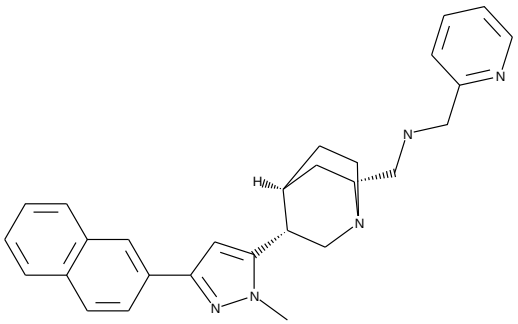


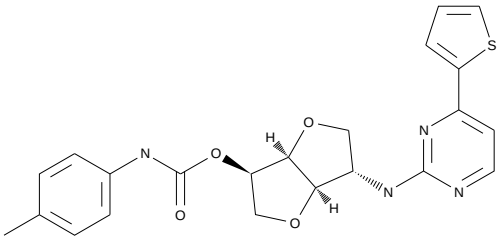












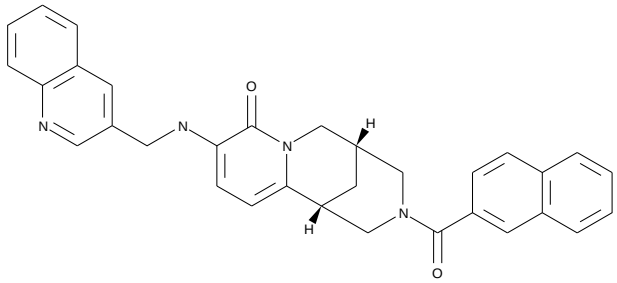


Fig S2. UV-Vis absorption titration of **1** (50 μM) in Tris/KCl buffer (100 mM KCl, 10 mM Tris-HCl, pH 7.5) with increasing amounts of ct DNA (0–24 μM). Inset: plot of $D/\Delta\varepsilon_{ap}$ versus D. Absorbance was monitored at 316 nm.

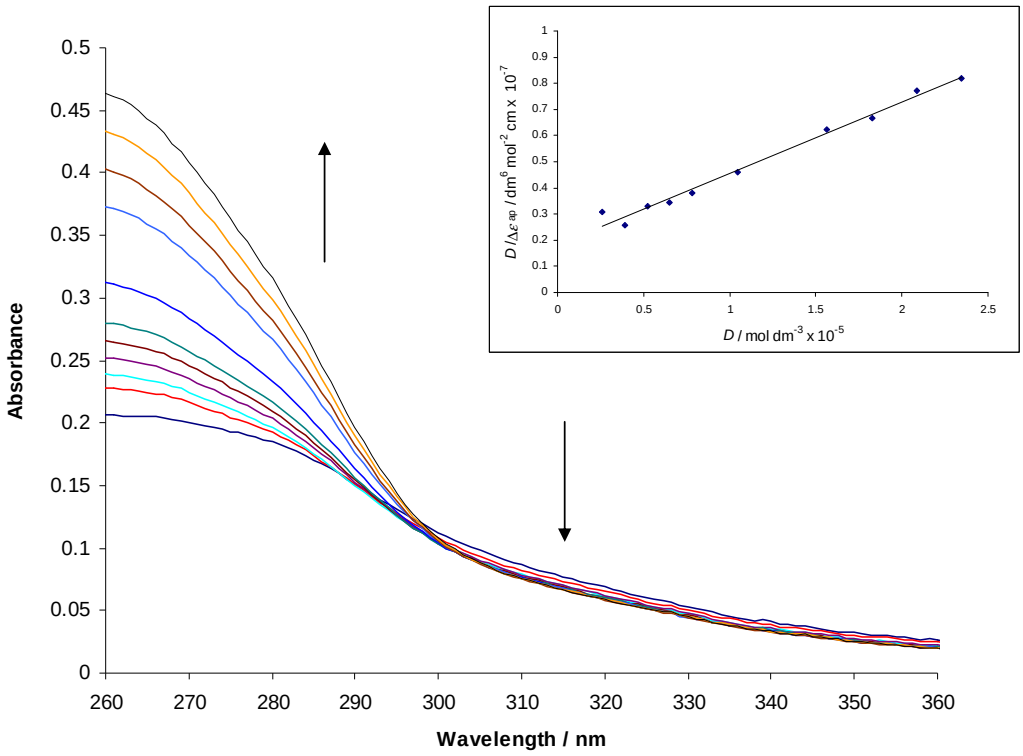


Fig S3. UV-Vis absorption titration of **1** (50 μM) in Tris/KCl buffer (100 mM KCl, 10 mM Tris-HCl, pH 7.5) with increasing amounts of SS22 (0–99 μM). Inset: plot of $D/\Delta\varepsilon_{ap}$ versus D. Absorbance was monitored at 316 nm.

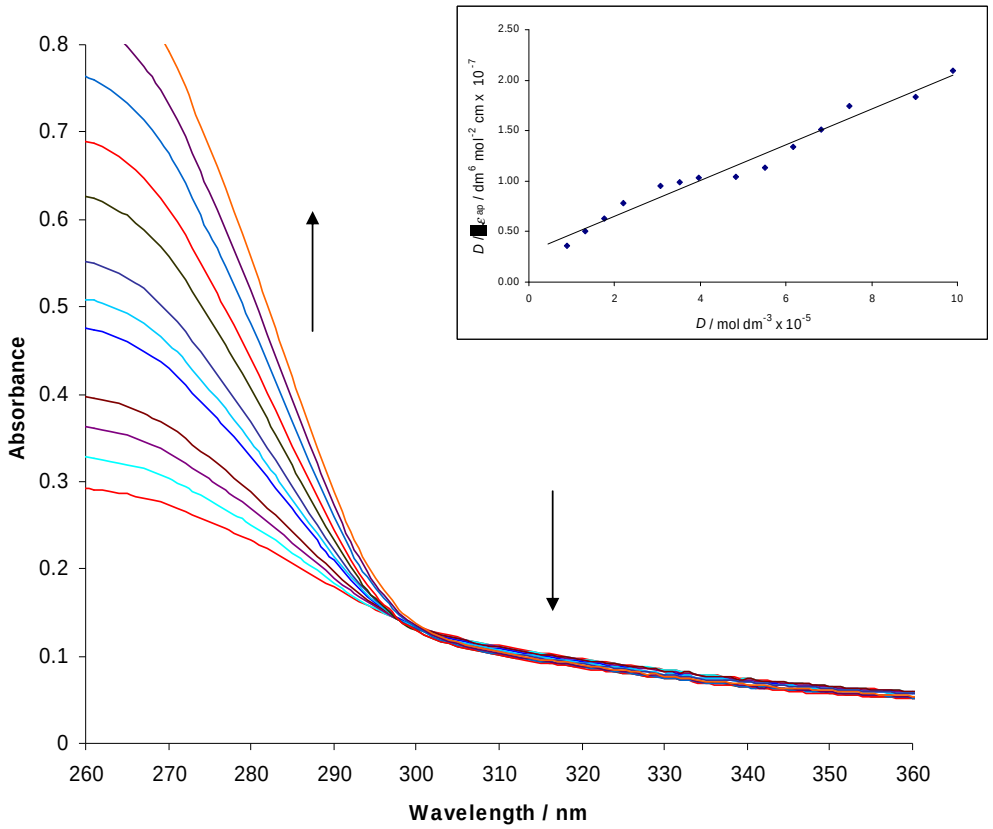


Fig S4. Polymerase stop assay with TMPyP4 (positive control) compared to fonsecin B **1**.

Deleted: chain reaction

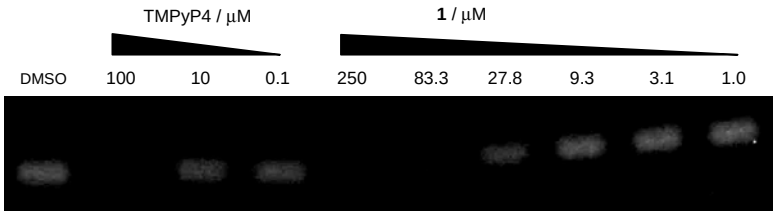
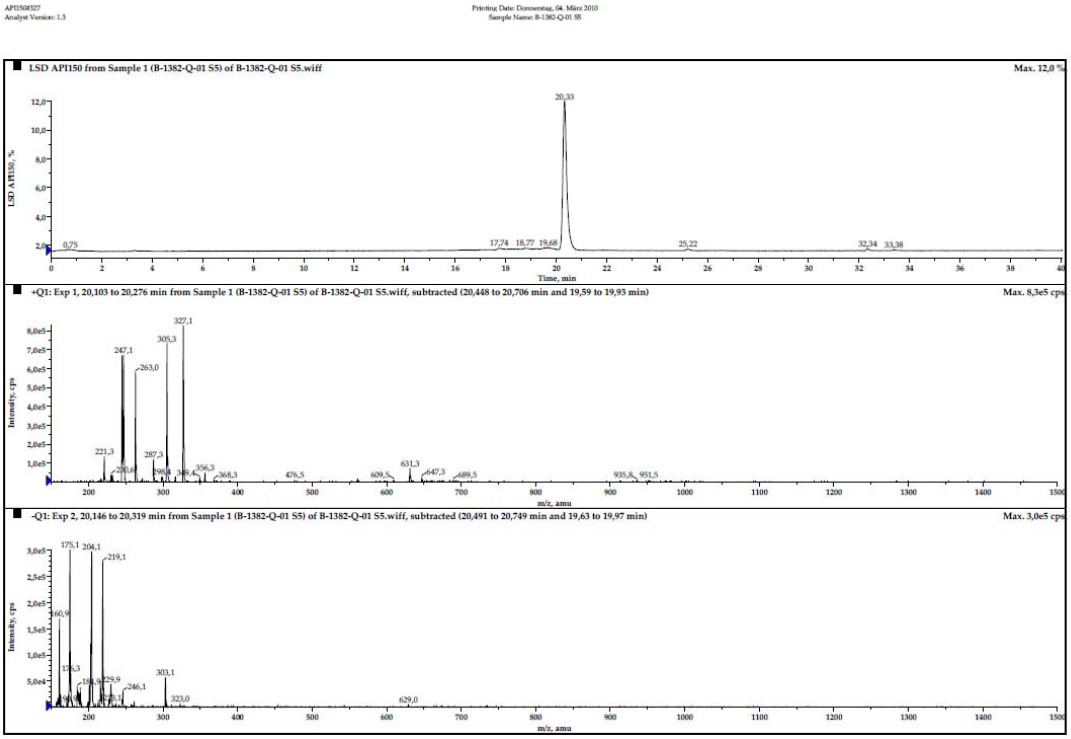
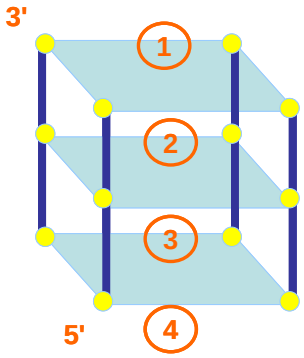


Figure S5.



Track 1: The HPLC-ELSD Chromatogram for purity detection.
Track 2: Detection of the positive ions generated during the ionisation of the compound in the MS: [MW+1H]+ and [MW+Na]+ can be detected.
Track 3: Detection of the negative ions generated during the ionisation of the compound in the MS: [MW-1H]-.

Table S1. Calculated binding energies (in kcal/mol) for **1** bound to different sites of the intramolecular c-*myc* G-quadruplex.



G-quadruplex-**1**	Binding energy (kcal/mol)			
	End-stacking at 3′ (1)[a]	intercalation near 3′ (2)[a]	intercalation near 5′ (3)[a]	End-stacking at 5′ (4)[a]
G-quadruplex–**1**	–47.88	25.15	24.96	–36.49

[a] Binding sites.