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).

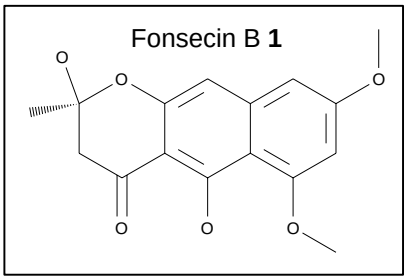
Polymerase stop assay. The polymerase stop assay was performed by using a modified protocol of the previously reported method.[3] The reactions were performed in 1× PCR buffer, containing each pair of oligomers (10 mmol), deoxynucleotide triphosphate (0.16 mM), Taq polymerase (2.5 U), and increasing 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.

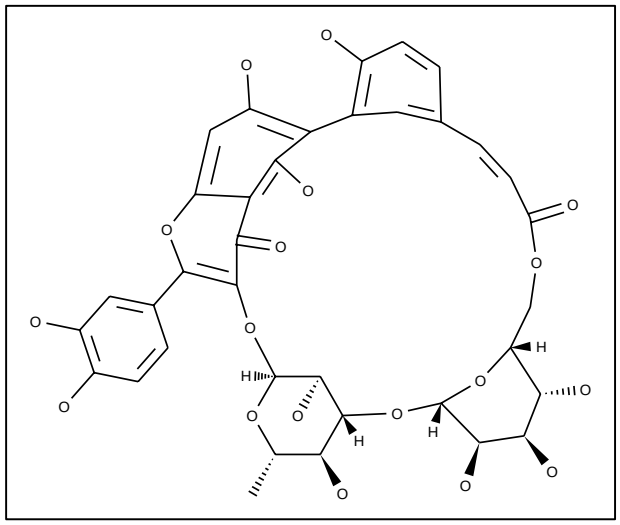
Molecular modeling. Molecular docking was performed by using the ICM-Pro 3.6-1d program (Molsoft).[4] According to the ICM method, the molecular system was described by using internal coordinates as variables. Energy calculations were based on the ECEPP/3 force field with a distance-dependent dielectric 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.

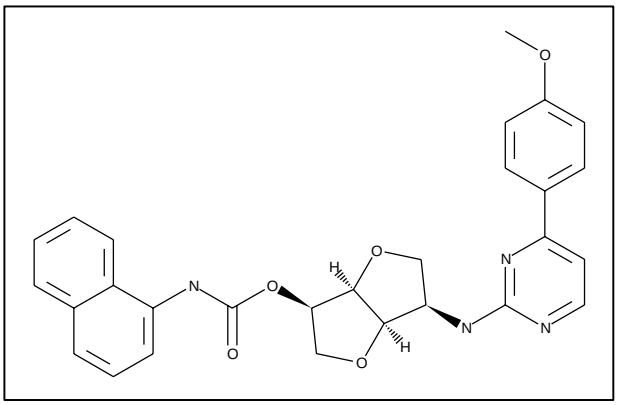
References

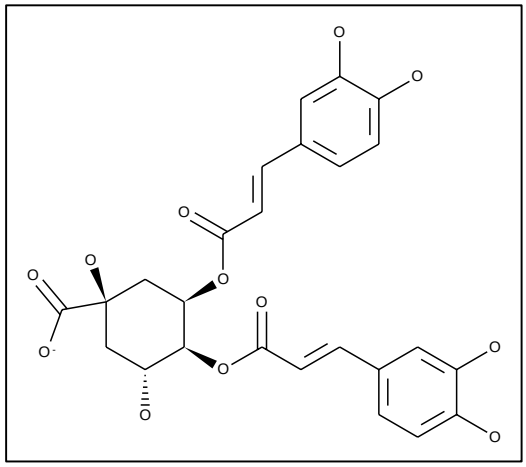
1 (a) J. Sambrook, E. F. Fritsch, T. E. Maniatis, *Molecular Cloning, A Laboratory Manual, 2nd Ed.,* **1989**, E.3 and E.10; (b) G. Felsenfeld, S. Z. Hirschman, *J. Mol. Biol.,* **1965**, *13*, 407.
2 C. V. Kumar, E. H. Asuncion, *J. Am. Chem. Soc.,* **1993**, *115*, 8547.
3 T.-M. Ou, Y.-J. Lu, C. Zhang, 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.

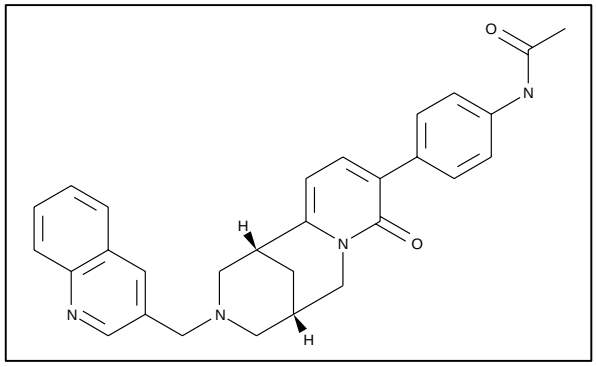
Fig S1. Chemical structures of the top-five scoring compounds (highlighted) identified using high-throughput virtual screening of a natural product database against the c-*myc* G-quadruplex. The structures ranked 6th-10th are also included for comparison.

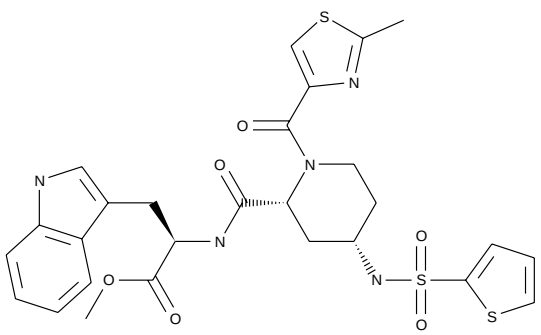


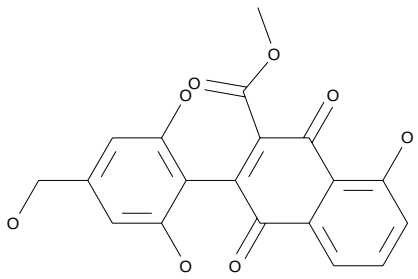


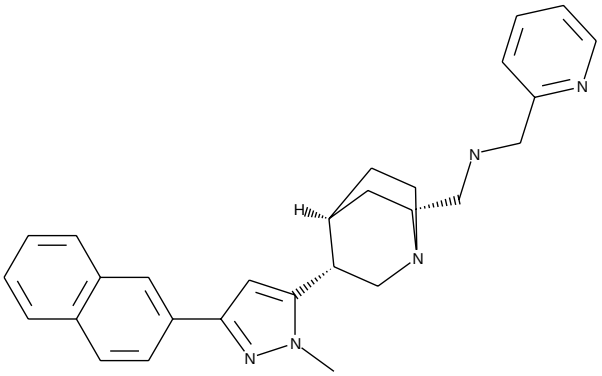


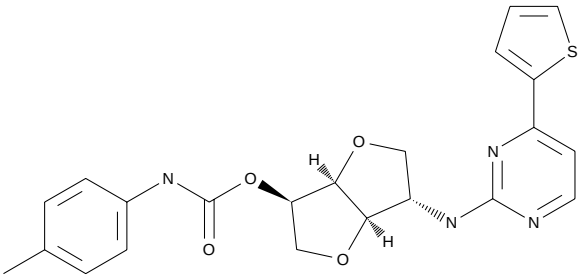












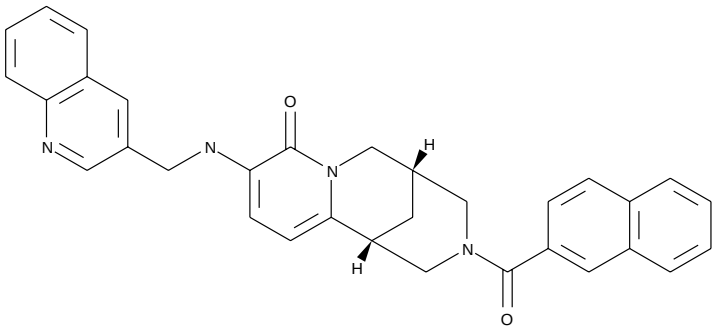


Fig S2. UV-Vis absorption titration of **1** (50 μM) in Tris/KCl buffer (100 mM KCl, 10 mM Tris-HCl, pH 7.5) with increasing amounts of ct DNA (0–24 μM). Inset: plot of $D/\Delta\varepsilon_{ap}$ versus D. Absorbance was monitored at 316 nm.

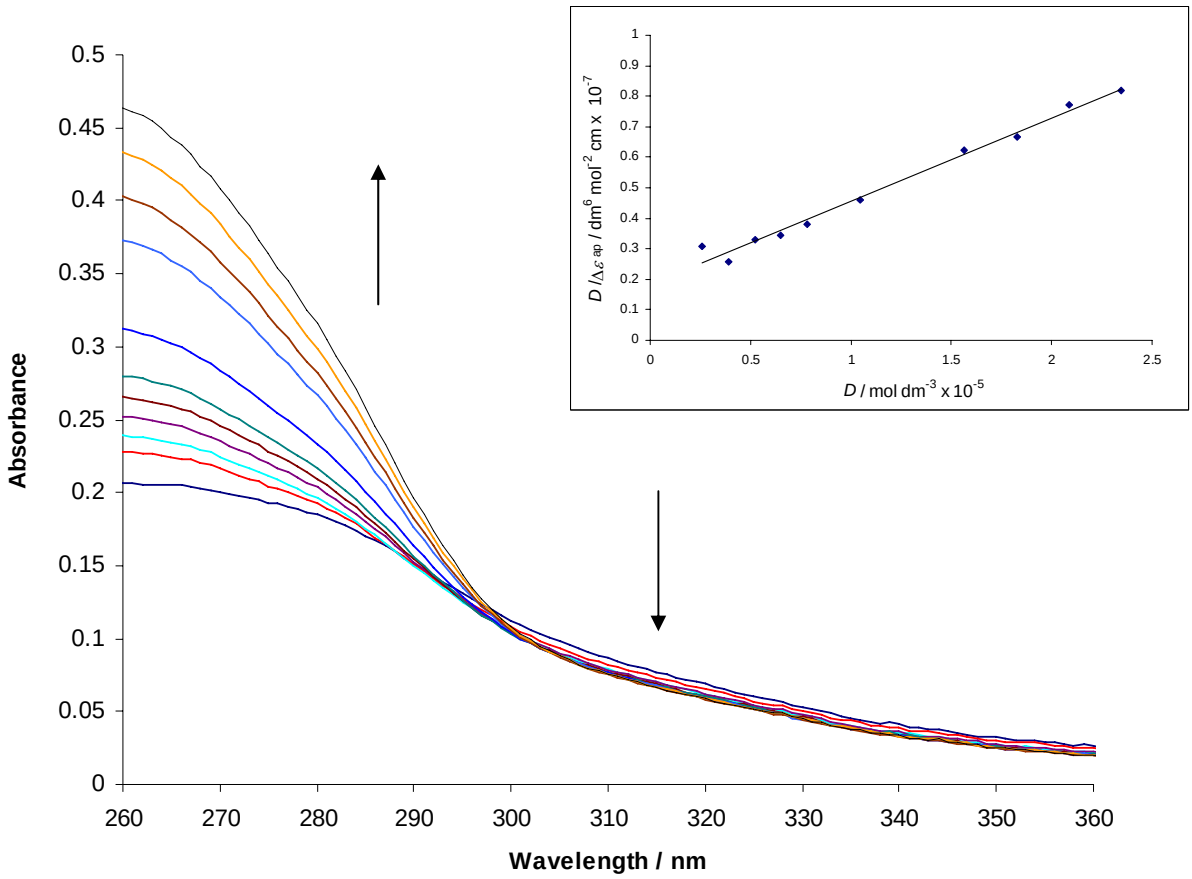


Fig S3. UV-Vis absorption titration of **1** (50 μM) in Tris/KCl buffer (100 mM KCl, 10 mM Tris-HCl, pH 7.5) with increasing amounts of SS22 (0–99 μM). Inset: plot of $D/\Delta\varepsilon_{ap}$ versus D. Absorbance was monitored at 316 nm.

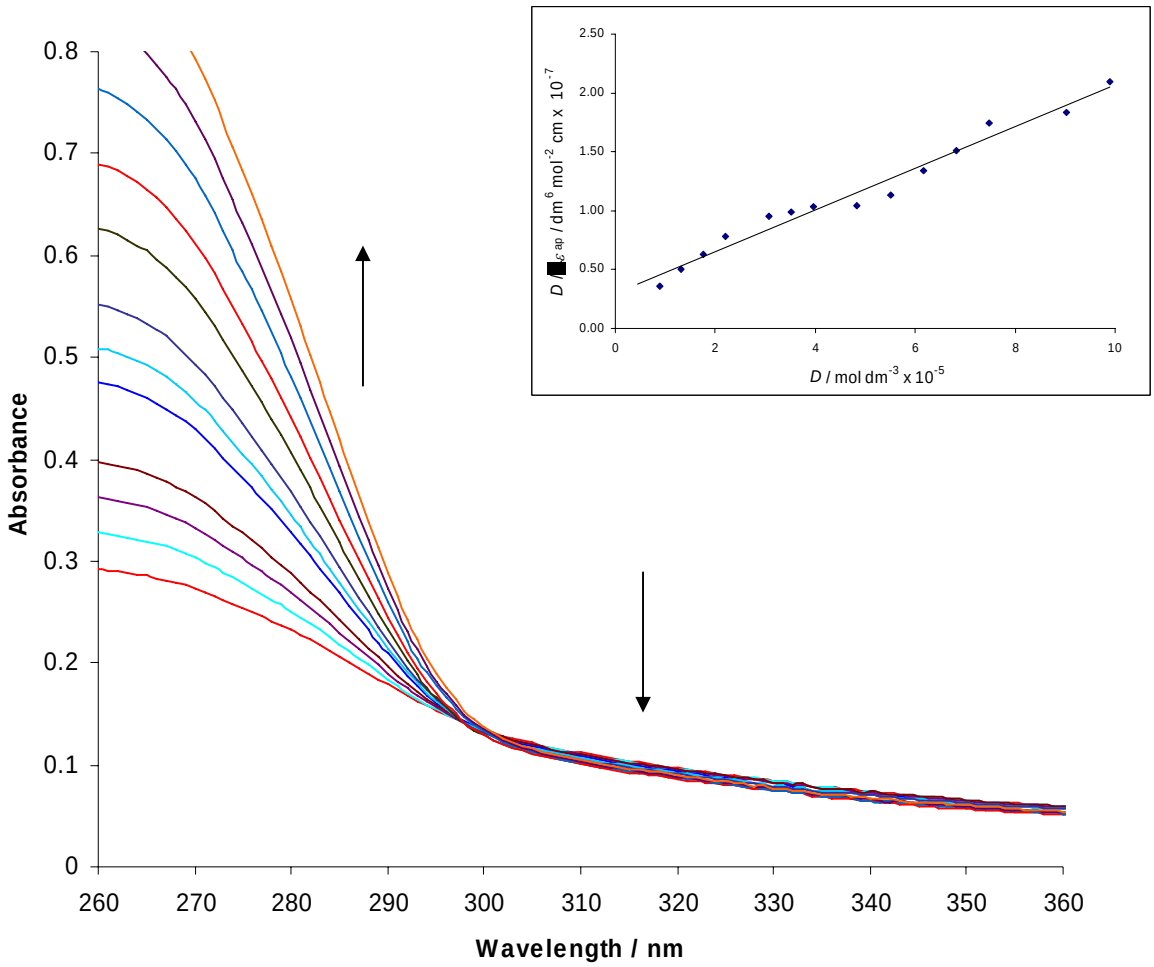


Fig S4. Polymerase stop assay with TMPyP4 (positive control) compared to fonsecin B **1**.

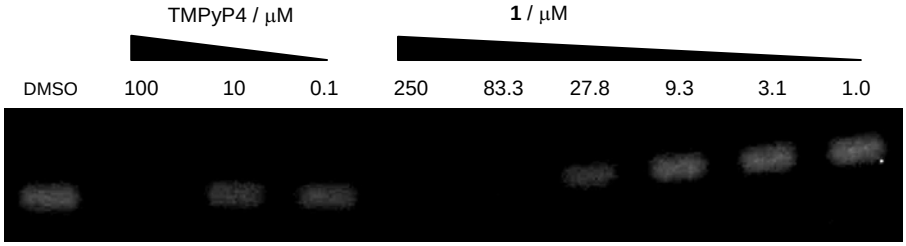
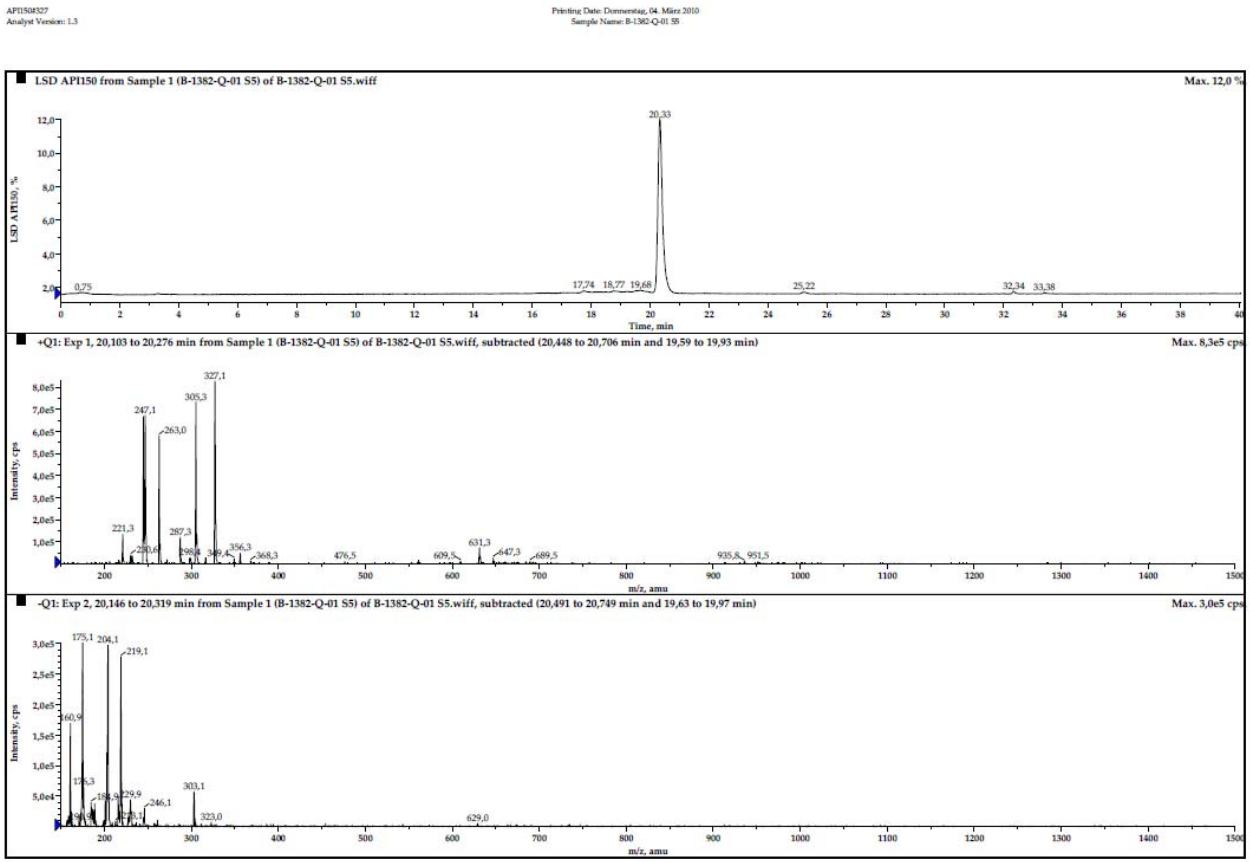
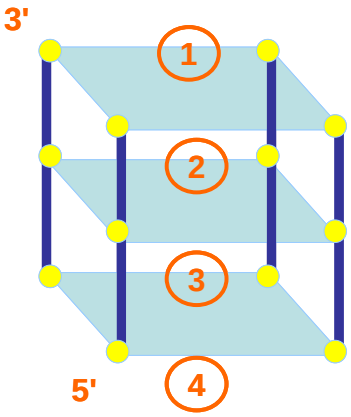


Figure S5.



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.