

Molecular Design of Synthetic Benzimidazoles for the Switchover of the Duplex to G-quadruplex DNA Recognition

Basudeb Maji^a and Santanu Bhattacharya^{*abc}

Abstract: Benzimidazole derivatives are well known for their antibacterial, antiviral, anticonvulsant, antihistaminic, anthelmintic and antidepressant activities. Benzimidazole's unique base-selective DNA recognition property has been studied widely. However, most of the early benzimidazole systems have been targeted towards the binding of duplex DNA. Here we have shown the evolution and progress of the design and synthesis of new benzimidazole systems towards selective recognition of the double-stranded DNA first. Then in order to achieve selective recognition of the G-quadruplex DNA and utilize their potential as future anti-cancer drug candidates, we have demonstrated their selective cytotoxicity towards the cancer cells and potent telomerase inhibition ability.

Keywords: Anti-cancer · Benzimidazole · Cytotoxicity · DNA · G-quadruplex · Telomerase · TRAP-LIG assay



Santanu Bhattacharya obtained his PhD degree from Rutgers University, New Brunswick, NJ under the supervision of Prof. Robert A. Moss. He had a three-year stint as a NIH postdoctoral fellow at the Massachusetts Institute of Technology under the tutelage of Nobel laureate Prof. Har Gobind Khorana. He is currently a full Professor and Chair of the Department of Organic Chemistry at the Indian Institute of Science, Bangalore. He is a recipient of TWAS prize in Chemistry and an elected fellow of the INSA, New Delhi and the IASc, Bangalore. His research interests include synthetic molecular design, chemical biology, molecular medicine and supramolecular chemistry.

*Correspondence: Prof. S. Bhattacharya^{abc}

^aDepartment of Organic Chemistry
Indian Institute of Science
Bangalore 560012, India
Tel.: +91 80 2293 2664
E-mail: sb@orgchem.iisc.ernet.in

^bChemical Biology Unit
Jawaharlal Nehru Centre for Advanced Scientific
Research

Bangalore 560012, India

^cJ. C. Bose Fellow, Department of Science and
Technology
New Delhi, India

Basudeb Maji obtained his BSc (Hons.) in Chemistry from the Asutosh College, University of Calcutta in 2006 and MSc in Organic Chemistry from the Scottish Church College, University of Calcutta in 2008. After that he joined Indian Institute of Science, Bangalore, India for his PhD degree and currently he is working as a senior research fellow under Prof. Santanu Bhattacharya since August, 2008. His research interest lies in developing human telomeric DNA targeted anti-cancer drugs and their prototypes.

1. Introduction

One of the most valuable properties that a living human being possesses is the genetic information which is hidden in the nucleic acid databank in the genome. Nucleic acids employ five nucleobases

to preserve the huge genetic information database. These nucleobases are based on two types of heterocyclic molecules, pyrimidine and purine. The preservation and transfer of the genetic information is crucial for a living being to sustain and survive. Unfortunately, it is not only important to survive but there is a necessity to maintain good health as well. This is exactly where we sometimes need drugs.

Scientists have been engaged in the discovery of drugs which can regulate and tinker with the function of nucleic acid ever since the discovery of the DNA double-helix by Watson and Crick.^[1] As far as DNA is concerned, the discovery of Hoechst compounds as blue fluorescent dyes for DNA staining, was a milestone for nucleic acids research.^[2] The sequence-specific DNA binding ability of Hoechst 33342 (**1a**, Fig. 1) and Hoechst 33258 (**1b**, Fig. 1) are well documented in literature.^[3]

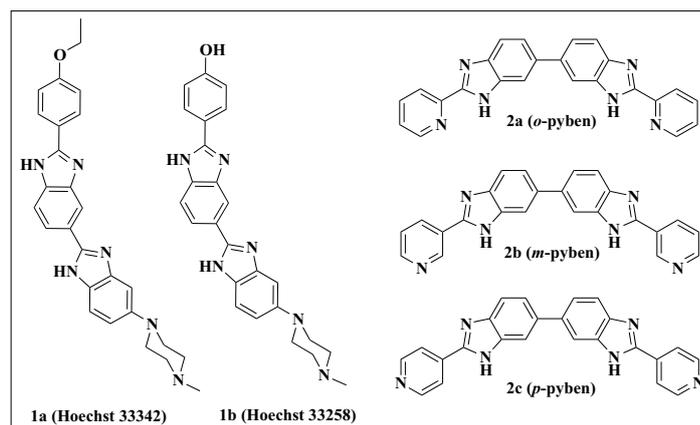


Fig. 1. Chemical structures of **1a** (Hoechst 33342), **1b** (Hoechst 33258) and the positional isomers, **2a** (o-pyben), **2b** (m-pyben) and **2c** (p-pyben).

The benzimidazole unit, which is quite similar to one of the DNA building blocks, purine base, has some exceptional features which make it a potent DNA binding moiety. Keeping its DNA recognition ability in mind unlike other synthetic templates that we developed earlier,^[4] we have recently synthesized new molecular entities based on benzimidazoles that not only recognize DNA sequences specifically but the basis of such molecular recognition could also be extended in their selective affinity towards other forms of DNA such as the G-quadruplex DNA.

2. Symmetrical Benzimidazole Ligands for DNA Binding

There are several reports on the modification of the original Hoechst 33258 molecule keeping the skeleton intact.^[5–7] Scrutiny of its molecular structural features reveals that a bis-benzimidazole unit is necessary for the double-helical DNA recognition. Further Hoechst 33258 is distinctly unsymmetrical with one end having a phenolic moiety while the other end possesses a piperazine residue. We began our investigation to understand the DNA binding property with a symmetrical dimer of benzimidazole. Towards this end we considered biphenyl-based positional isomers of the symmetric bis-2-(pyridyl)-1H-benzimidazole derivatives [*p*-pyben (**2a**), *m*-pyben (**2b**) and *o*-pyben (**2c**)].^[8a] Keeping the central biphenyl bis-benzimidazoles intact, we varied the *N*-termini in the form of various pyridine isomers to give an overall crescent shape with little deviation from planarity in the center (Fig. 1). Importantly, even after symmetrization, all the molecules preserved their A-T sequence specificity in the duplex DNA. The other aspect was the effect of positional isomers on their DNA binding capability. The spectroscopic titrations of each ligand with DNA (CD, fluorescence and DNA melting studies) have established that the *meta* and the *para* isomers act as efficient DNA binders whereas the *ortho* isomer possesses a poorer affinity towards DNA (Table. 1). It should be noted that each ligand molecule can adopt two canonical structures with respect to its benzimidazole protons. When we looked at the energy-minimized structure, we observed that both the *meta* and *para* isomers had insignificant energy differences in the two forms. For the *ortho* isomer, however, it was found to be significantly high due to the intramolecular H-bonding between the imidazole proton and pyridine nitrogen making one form more stable and inaccessible towards H-bonding interaction with the duplex DNA. This observation demonstrated the importance of the benzimidazole N-H residues towards the hydrogen bonding with duplex DNA during their binding. We have subsequently demonstrated the utility of related benzimidazoles towards topoisomerase inhibition.^[8b]

Table 1. Summary of DNA melting temperature in presence of **2a** (*o*-pyben), **2b** (*m*-pyben), **2c** (*p*-pyben) and (**1b**) Hoechst 33258.

	[Poly (dA-dT)] ₂		[Poly(dA).poly(dT)]	
	T _m [°C]	ΔT _m [°C]	T _m [°C]	ΔT _m [°C]
DNA (in buffer)	67.1	–	74.7	–
DNA + <i>p</i> -pyben	75.5	8.1	91.1	16.4
DNA + <i>m</i> -pyben	73.0	5.9	87.4	12.7
DNA + <i>o</i> -pyben	67.5	0.4	75.4	0.7
DNA + Hoechst	80.0	12.9	90.4	15.7

azole N-H residues towards the hydrogen bonding with duplex DNA during their binding. We have subsequently demonstrated the utility of related benzimidazoles towards topoisomerase inhibition.^[8b]

3. Metal Ion-promoted ‘Switch-off’ of the DNA Binding

We explored the duplex DNA binding of the *ortho*-pyridine-based bis-benzimidazole compound (**2a**) in the presence of various bivalent metal ions in solution.^[9] Unlike the *meta*- and the *para*-pyridine-based isomeric counterparts the *ortho* ligand showed complete ‘switch-off’ and ‘switch-on’ properties of DNA recognition in the presence of a few bivalent metal ions and EDTA respectively. The bivalent metal ions chelate with the imidazole N and the *ortho* N of the pyridine moiety in the inner concave surface of the ligand forcing the imidazole N–H to be on the outer convex surface (away from the surface of the minor groove wall of the duplex DNA) making the ligand inefficient towards the DNA binding (Fig. 2). Thus we concluded that not only the accessibility of the imidazole N–H, but the position of the imidazole N–H are both important for optimal DNA binding.

4. How Can One Design Ligands for a Specific Secondary Conformation of DNA?

The discovery of the double-helical structure of DNA by Watson and Crick has provided the first idea about the mode of heredity.^[1] This model suggests that double-stranded DNA exists in *B*-form, where the two complementary strands associate with each other *via* H-bonded nucleotide base pairing and other non-covalent interactions. This allows the synthesis and reconstruction of new DNA duplexes on the basis of distinct base sequence and strand complementarity while it is generally believed that DNA mostly adopts the *B*-form within cellular environments, DNA can also exist in other conforma-

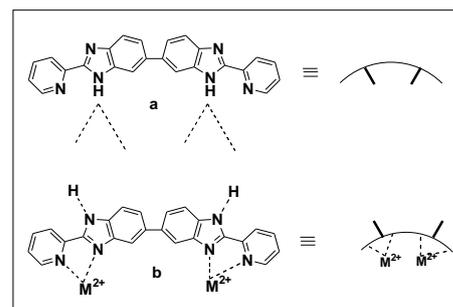


Fig. 2. ‘Switching-on’ and ‘switching-off’ mechanism of DNA recognition by **2a** in presence of bivalent metal ion. Adapted with permission from ref. [9], copyright 2007, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

tions.^[10] The dynamic topology of DNA allows formation of both non-*B*-form and non-helical secondary structures. For instance alternative helical conformers such as the A- and Z-forms of DNA have been demonstrated to exist under specific experimental conditions. The formation of non-conventional DNA secondary structures, *e.g.* G-quadruplexes,^[11] triple helix,^[12] i-motifs^[13] or A-motifs^[14] *etc.* have indeed been shown and specific ligands can be designed to stabilize such energetically less favorable DNA structures.

Towards this end, we first explored the telomeric *Tetrahymena thermophila* sequence d(T₂G₄) which is known to adopt polymorphic DNA structure through Hoogsteen mode of H-bonding between the guanine bases in presence of physiologically available monovalent cations like Na⁺ and K⁺ known as the G-quadruplex.^[15] The G-quadruplex morphology is totally different from the normal *B*-DNA duplex structure. We chose an aromatic ring in the center to make the molecule planar and attached a benzimidazole moiety at the benzene ring in different positions (*p*-phen-bbim, *m*-phen-bbim).^[16] The 1,3-isomer gave a ‘V’-shaped molecule whereas the 1,4-isomer afforded the corresponding linear form (Fig. 3).

The absorption titration showed that both ligands have very low binding affinity towards the duplex DNA irrespective of their base sequences. However, both ligands showed strong hypochromicity when these were titrated against the above

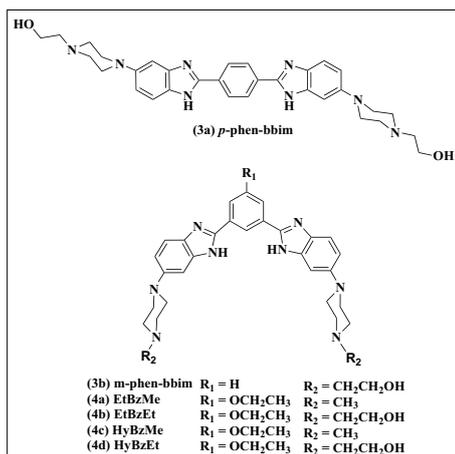


Fig. 3. Molecular structure of **3a** (*p*-phen-bbim), **3b** (*m*-phen-bbim), **4a** (EtBzMe), **4b** (EtBzEt), **4c** (HyBzMe) and **4d** (HyBzEt).

G-quadruplex DNA giving an estimation of >100 times preference towards the G-quadruplex over the duplex DNA. The poor binding towards duplex DNA was confirmed further through viscometric titrations with calf thymus DNA which eliminated the possibility of an intercalation mode of binding. Thermal melting studies showed hardly any elevation of the melting temperature by these ligands compared to the standard groove binder, Hoechst 33258. The circular dichroism spectral data showed the excellent molecular property of the *meta* isomeric ligand which induced a change in the G-quadruplex topology from the hybrid to parallel structure upon incubation with the preformed G-quadruplex. This was further substantiated by the electrophoretic mobility shift assay (Fig. 4). Importantly, the *meta* ligand **3b** could induce the formation of parallel G-quadruplex structure even in absence of any monovalent metal ion (Fig. 5). The CD melting profile showed a high degree of thermal stabilization upon incubation with the *meta* ligand even in LiCl buffer. The molecular dynamics (MD) simulation of the G-quadruplex DNA and the ligand **3b** revealed that the *meta* compound was able to stack efficiently over three guanine residues in the G-tetrad giving rise to higher

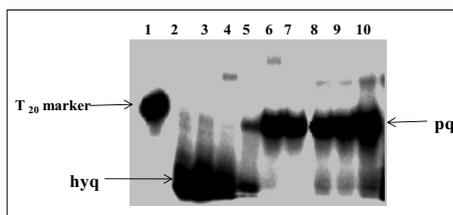


Fig. 4. Electrophoretic mobility shift assay showing G-quadruplex structure alteration from hybrid quadruplex (hyq) to parallel quadruplex (pq) upon interaction with **3b** (*m*-phen-bbim). Adapted with permission from ref. [15], copyright 2009, American Chemical Society.

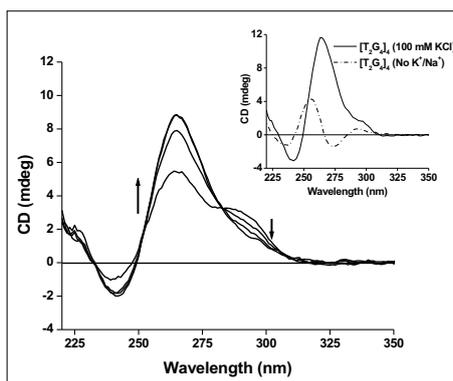


Fig. 5. Formation of G-quadruplex from $d(T_2G_4)_2$ upon incubation with ligand **3b** (*m*-phen-bbim) in absence of any monovalent ion. Adapted with permission from ref. [16], copyright 2010, American Chemical Society.

stabilization of the quadruplex-DNA. In contrast the *para* compound **3a** could cover only two guanine residues of the G-tetrad making it less efficient as a G-quadruplex DNA stabilizer. Both compounds showed some inhibition of the human telomere DNA telomerase indicating their potential towards drug design. Because the angular shaped **3b** was a stronger inhibitor than the linear isomer **3a**, we focused our attention on 1,3-substituted 'V'-shaped molecules with appropriate modifications in the central aromatic nucleus and incorporated an end piperazine moiety to get the optimum out of such molecular entities.^[17] We introduced four ligands (**4a-d**) and each of them (Fig. 3) showed better binding

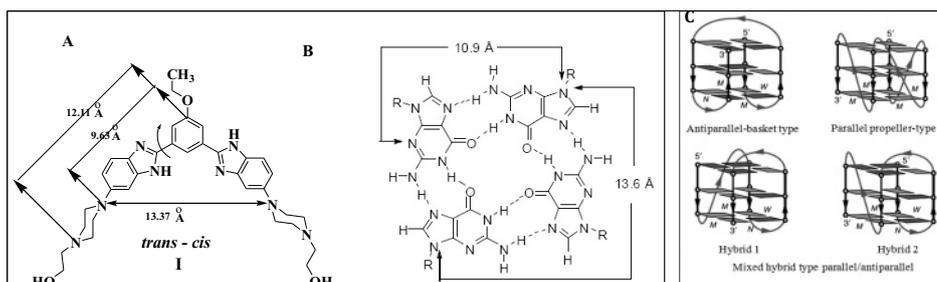


Fig. 6. (A) Energy-minimized structure of **4b** (EtBzEt) at B3LYP/6-31G* level of theory and (B) structure of the G-tetrad with the interatomic distances obtained from crystal structure.^[17] (C) Different polymorphs of the intramolecular human telomeric G-quadruplex DNA as obtained from NMR or X-ray crystal structural analysis. Adapted with permission from ref. [28], copyright 2006, American Chemical Society.

as well as distinct selectivity towards the G-quadruplex. The ligands also induced alteration in the G-quadruplex DNA topology.

5. Human Telomeric G-quadruplex DNA Targeting and Anti-cancer Drug Design

We then addressed the problem of targeting human telomeric G-quadruplex DNA $d(T_2AG_3)$ which is a potential anti-cancer target.^[18] It can form different types of structures in different salt conditions (Fig. 6c).^[19] Based on the G-quadruplex DNA crystal structure, the G-tetrad is $10.9 \times 13.6 \text{ \AA}$ in dimension with the three tetrads of 8.1 \AA height (Fig. 6).^[20] Based on this, we synthesized six dimeric ligands where two ligand units were connected by oxyethylene spacers of different lengths (Scheme 1). These ligands were compared with their monomeric counterpart towards the G-quadruplex DNA binding.^[21,22] We found a dramatic change in their binding capability compared to their monomers. Each dimeric ligand (**5a-f**) (Scheme 1) showed a high degree of thermal stabilization of the G-quadruplex DNA along with the structural alteration property.

The above compounds also showed efficient telomerase inhibition capability. Among these the dimer with a hexaethylene glycol spacer (**5f**) showed the best telomerase inhibition activity with a IC_{50} value as low as $5.9 \mu\text{M}$. Further such ligand molecules demonstrated selective cytotoxicity towards cancer cells (Fig. 7).

MD simulation reveals the mode of interaction of each ligand with the G-quadruplex DNA. This explains why an increase in the length of the intervening oxyethylene spacer resulted in increased binding efficiency. This was also supported by all the spectroscopic studies. MD simulation of the dimeric ligand with the dimeric G-quadruplex DNA further showed that one arm of the ligand prefers to interact with the G-tetrad whereas the other one has affinity for the groove (Fig. 8).

6. Non-planar Chiral Drugs for Targeting the G-quadruplex DNA Groove

One thing that must be noted is that the G-quadruplex binding compounds prefer a 'V'-shape and planar geometry and such a shape is applicable only if the target is a planar G-tetrad. The other potential binding sites for the G-quadruplex comprises the groove and the loop regions. There are few reports^[9] on the G-quadruplex DNA groove or loop binder ligands including oli-

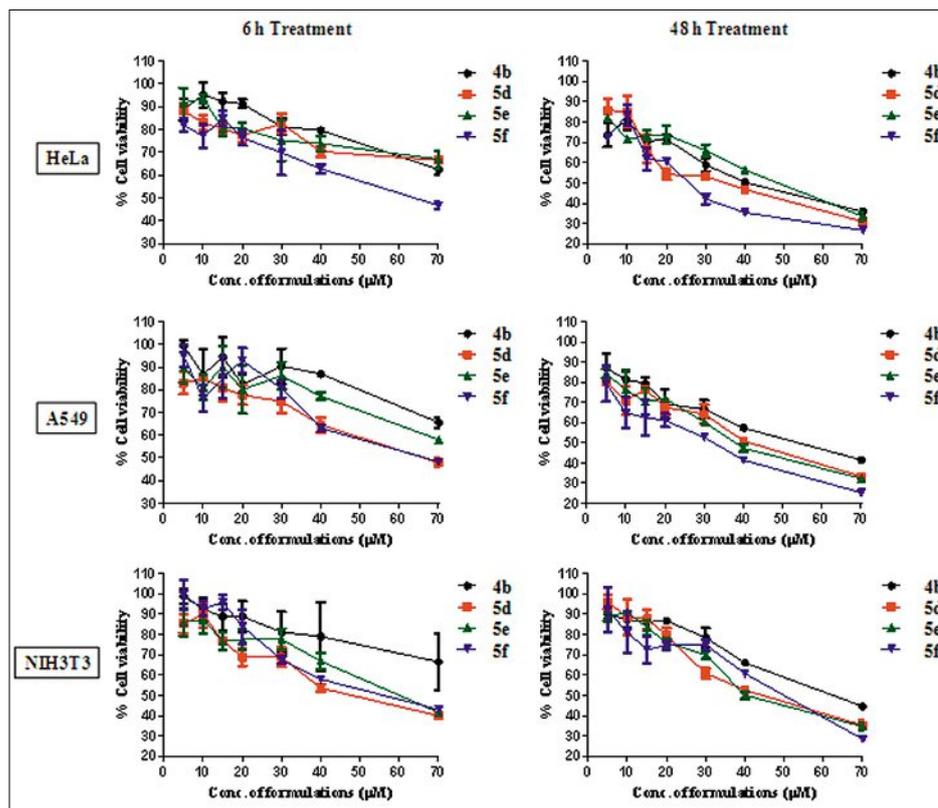


Fig. 7. Effect of the ligands on the cell viability after 6 h and 48 h exposure of HeLa, A549 and NIH3T3 cells with indicated ligands at various concentrations as measured by MTT assay. Adapted with permission from ref. [22], copyright 2012, Public Library of Science.

gopeptides (netropsin, distamycin-A), carbocyanine dyes *etc.*^[23–26] Towards this end we have introduced a novel Tröger's base chiral scaffold having the benzimidazole moiety (Scheme 2) to investigate their efficiency towards G-quadruplex DNA recognition.^[27] As it was not possible to separate the two enantiomers, we proceeded with the racemic mixtures. Both ligands (TBBZ1 **6** and TBBZ2 **7**) showed significant selectivity and strong affinity towards

the human telomeric G-quadruplex over the corresponding duplex DNA sequence as evidenced from UV-Vis, CD spectral, thermal denaturing and the membrane dialysis studies. Interestingly, the MD simulation showed that the (*S,S*) isomer had a

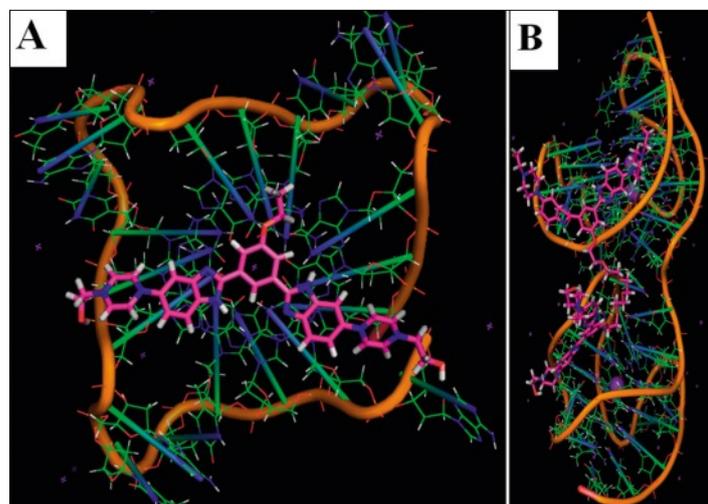


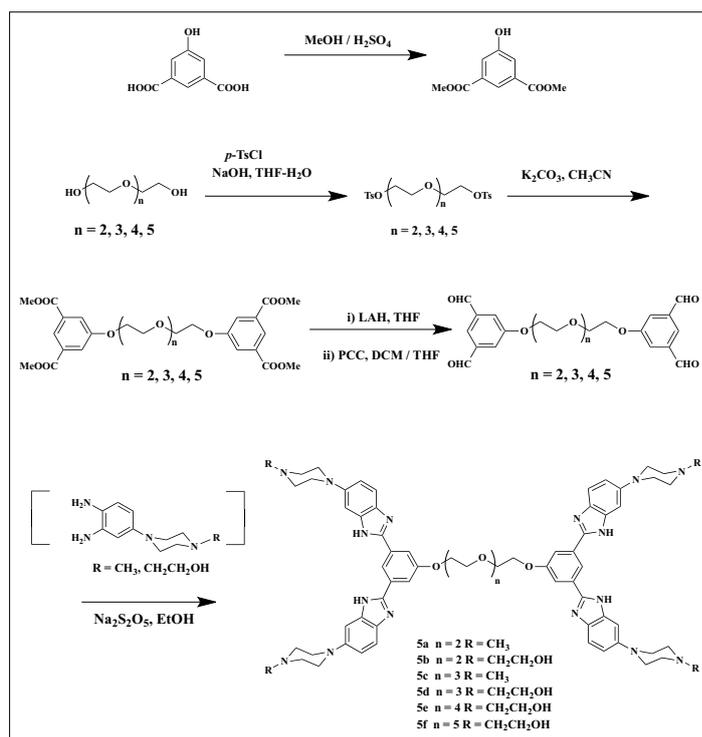
Fig. 8. Ligand **3b** (EtBzEt)-G-quadruplex DNA complex after 8 ns MD simulation (A) and **4e**-dimeric-G-quadruplex DNA complex after 8 ns MD simulation (B). Adapted with permission from refs [27, 22], copyright 2012, American Chemical Society and copyright 2012, copyright 2012, Public Library of Science.

preference of interaction with the G-tetrad whereas the (*R,R*) isomer preferably went to the groove (Fig. 9). The spectroscopic results were also in good agreement with the biological activities measured with the TRAP-LIG assay and cell viability assay (Fig. 10).

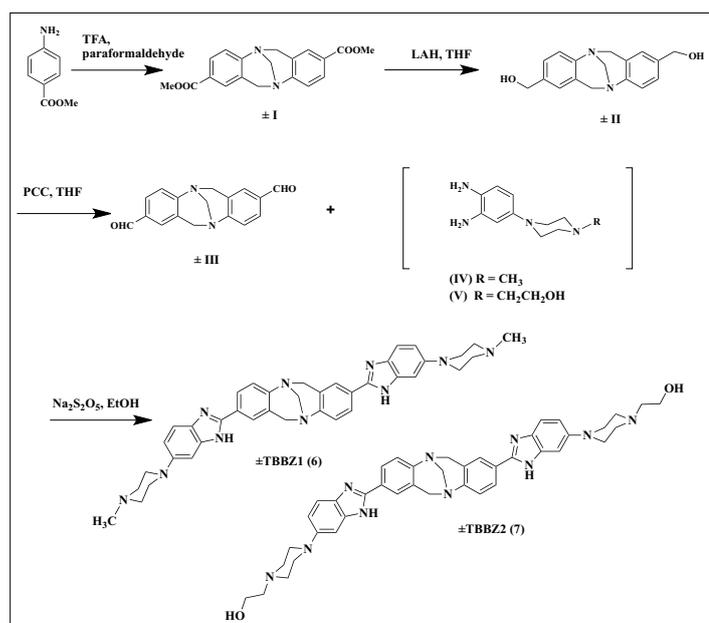
7. Conclusions and Outlook

The benzimidazole moiety, originally discovered by Hoechst AG, has now been used extensively towards the targeting of specific DNA sequences and also as DNA specific fluorophores. Here we have shown how the potency of such molecular designs could be extended, honed and tailored to achieve selective human telomeric G-quadruplex DNA in preference to duplex DNA, which is a fascinating strategy to selectively target the cancer cell. Currently we are working on carbazole-based new benzimidazole systems to increase their efficiency in the telomerase inhibition. Simultaneously, we are also looking for flexible amino acid-based systems towards selective G-quadruplex DNA recognition.

Received: October 7, 2012



Scheme 1. The synthetic outline of dimeric ligands **5a-5f**.



Scheme 2. Synthetic outline of the Tröger's base compounds, **6** (TBBZ1) and **7** (TBBZ2).

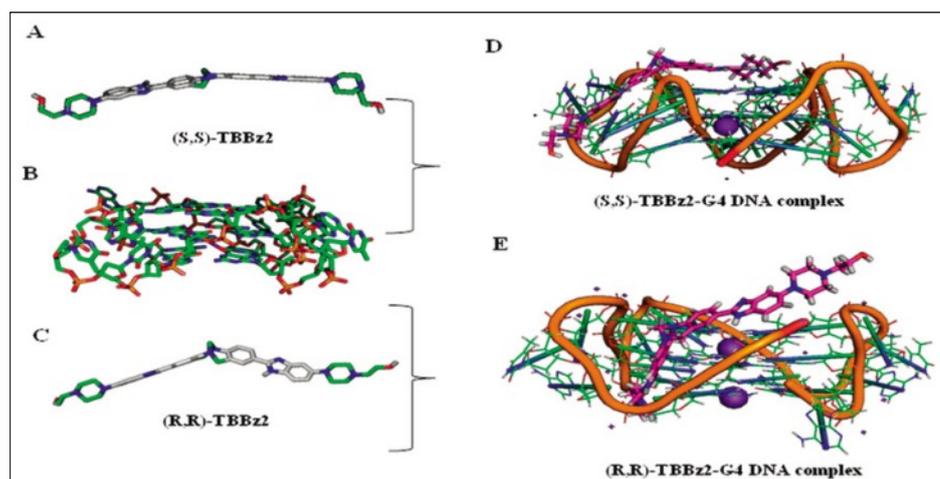


Fig. 9. Energy optimized structures at the B3LYP/6-31G* level of theory of (S, S)-TBBZ2, and (R, R)-TBBZ2 respectively (A and C); Structures of the 1KF1 (B); MD simulated ligand-DNA complex of (S,S)-TBBZ2, and (R,R)-TBBZ2 with 1KF1 respectively (D, E). Adapted with permission from ref. [27], copyright 2012, American Chemical Society.

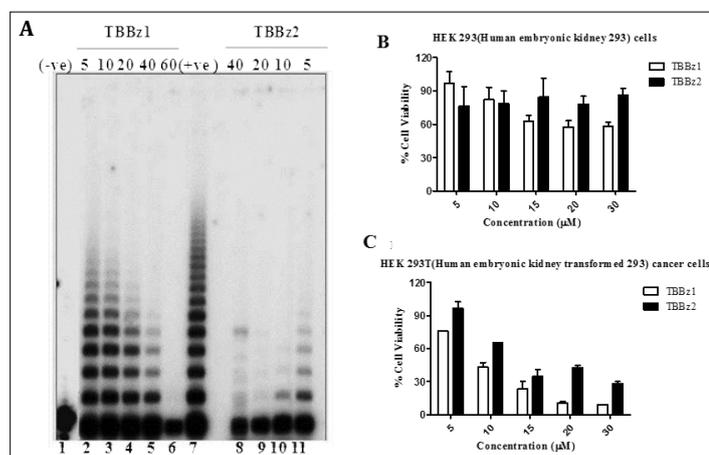


Fig. 10. Telomerase inhibition by **6** (TBBZ1) and **7** (TBBZ2) as determined from a TRAP-LIG assay (A) and Selective cancer cell cytotoxicity by **6** (TBBZ1) and **7** (TBBZ2) (B, C). Adapted with permission from ref. [27], copyright 2012, American Chemical Society.

- [1] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737.
- [2] (a) S. A. Latt, G. Setten, *J. Histochem. Cytochem.* **1976**, *24*, 24; (b) S. Bhattacharya, P. Chaudhuri, *Curr. Med. Chem.* **2008**, *15*, 1762.
- [3] N. Spink, D. G. Brown, J. V. Skelly, S. Neidle, *Nucleic Acids Res.* **1994**, *22*, 1607.
- [4] (a) S. Bhattacharya, M. Thomas, *Tetrahedron Lett.* **2000**, *41*, 5571; (b) S. Bhattacharya, M. Thomas, *Tetrahedron Lett.* **2001**, *42*, 5525; (c) S. Bhattacharya, M. Thomas, *Chem. Commun.* **2001**, 1446.
- [5] A. L. Satz, T. C. Bruice, *J. Am. Chem. Soc.* **2001**, *123*, 2469.
- [6] P. M. Reddy, P. T. Jindra, A. L. Satz, T. C. Bruice, *J. Am. Chem. Soc.* **2003**, *125*, 7843.
- [7] M. Tanada, S. Tsujita, S. Sasaki, *J. Org. Chem.* **2006**, *71*, 125.
- [8] (a) P. Chaudhuri, B. Ganguly, S. Bhattacharya, *J. Org. Chem.* **2007**, *72*, 1912; (b) P. Chaudhuri, H. K. Majumder, S. Bhattacharya, **2007**, *50*, 2536.
- [9] S. Bhattacharya, P. Chaudhuri, *Chem. Asian J.* **2007**, *2*, 648.
- [10] (a) S. Kendrick, L. H. Hurley, *Pure Appl. Chem.* **2010**, *82*, 1609; (b) A. K. Jain, S. Bhattacharya, *Bioconjugate Chem.* **2010**, *21*, 1389.
- [11] R. K. Ralph, W. J. Connors, H. G. Khorana, *J. Am. Chem. Soc.* **1962**, *84*, 2265.
- [12] H. Htun, J. E. Dahlberg, *Science* **1989**, *243*, 1571.
- [13] M. Gueron, J. L. Leroy, *Curr. Opin. Struct. Biol.* **2000**, *10*, 326.
- [14] J. Luo, M. H. Sarma, R-D. Yuan and R. H. Sarma, *FEBS Lett.* **1992**, *306*, 223.
- [15] C. C. Hardin, E. Henderson, T. Watson, J. K. Prosser, *Biochemistry* **1991**, *30*, 4460.
- [16] S. Bhattacharya, P. Chaudhuri, A. K. Jain, A. Paul, *Bioconjugate Chem.* **2010**, *21*, 1148.
- [17] A. K. Jain, V. V. Reddy, A. Paul, K. Muniyappa, S. Bhattacharya, *Biochemistry* **2009**, *48*, 10693.
- [18] E. H. Blackburn, *Nature* **1991**, *350*, 569.
- [19] N. K. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *J. Am. Chem. Soc.* **2006**, *128*, 9963.
- [20] G. N. Parkinson, M. P. Lee, S. Neidle, *Nature* **2002**, *417*, 876.
- [21] A. K. Jain, A. Paul, B. Maji, K. Muniyappa, S. Bhattacharya, *J. Med. Chem.* **2012**, *55*, 2981.
- [22] A. Paul, A. K. Jain, S. K. Misra, B. Maji, K. Muniyappa, S. Bhattacharya, *PLoS ONE* **2012**, *7*, e39467.
- [23] A. K. Jain, S. Bhattacharya, *Bioconjugate Chem.* **2011**, *22*, 2355.
- [24] A. K. Jain, S. Bhattacharya, *Curr. Pharm. Des.* **2012**, *18*, 1917.
- [25] A. Randazzo, A. Galeone, V. Esposito, M. Varra, L. Mayol, *Nucleosides, Nucleotides & Nucleic Acids* **2002**, *21*, 535.
- [26] A. Randazzo, A. Galeone, L. Mayol, *Chem. Commun.* **2001**, *11*, 1030.
- [27] A. Paul, B. Maji, S. K. Misra, A. K. Jain, K. Muniyappa, S. Bhattacharya, *J. Med. Chem.* **2012**, *55*, 7460.
- [28] N. K. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *J. Am. Chem. Soc.* **2006**, *128*, 9963.