J|A|C|S COMMUNICATIONS

Published on Web 10/26/2009

Tandem Application of Virtual Screening and NMR Experiments in the Discovery of Brand New DNA Quadruplex Groove Binders

Sandro Cosconati,[†] Luciana Marinelli,[†] Roberta Trotta,[‡] Ada Virno,[‡] Luciano Mayol,[‡] Ettore Novellino,[†] Arthur J. Olson,[§] and Antonio Randazzo*[,‡]

Dipartimento di Chimica Farmaceutica e Tossicologica, Università degli Studi di Napoli "Federico II", via D. Montesano 49, I-80131 Napoli, Italy, Dipartimento di Chimica delle Sostanze Naturali, Università degli Studi di Napoli "Federico II", via D. Montesano 49, I-80131 Napoli, Italy, and Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received July 29, 2009; E-mail: antonio.randazzo@unina.it

G-quadruplexes are four-stranded helical DNA or RNA structures, comprising stacks of G-tetrads, which are the planar association of four guanines in a cyclic Hoogsteen hydrogen-bonding arrangement.[1] These structures are formed by the folding of one DNA or RNA strand or by the association of two or more strands. This results in different combinations of relative strand orientations, with consequent formation of grooves of different widths, and a number of loop arrangements. From a biological point of view, G-quadruplexes are widespread in genome and they seem to play a role in a number of processes, such as replication, recombination transcription, and translation.[2] Furthermore, quadruplexes are also found in telomeric DNA. Telomers consist of an ensemble of proteins and specialized noncoding DNA sequences which protect the ends of the chromosome from damage and recombination, and their shortening is implicated in cellular senescence. It has been demonstrated that the elongation of telomeric DNA operated by the enzyme telomerase leads cancer cells to an infinite lifetime. Hence, the inhibition of telomerase, which is overactive in ~85% of tumors, is expected to move into the forefront of research for new effective anticancer drugs. Since this enzyme requires a single-stranded telomeric primer, the formation of G-quadruplex complexes by telomeric DNA inhibits the telomerase activity. In this respect, it has been found that small molecules that stabilize G-quadruplex structures are effective telomerase inhibitors.[3] Quadruplex structures offer several recognition sites. Most of the interacting molecules discovered so far have been found to interact with the wide π-stacking surface of the G-tetrads at the 5′ and/or 3′ edges of the quadruplex.[4] Only recently, our research group has unambiguously demonstrated a groove binding mode between distamycin A and the quadruplex $[d(TGGGGT)]_4$.[5] So far, this compound remains the only comprehensively documented groove binder for the DNA quadruplex. Classical intercalation, as observed in the duplex DNA structures, has not been demonstrated to date, most probably due to the presence of cations like K^+ or Na^+ in the very center cavity of the quadruplex structures, that prevents such a binding mode.

Generally, stacking interaction and groove-binding modes are characterized by very different specificities. Groove-binding recognition generally offers a higher extent of selectivity, since it can more easily recognize different DNA sequences and, in the case of quadruplex structures, can also discriminate among several quadruplex topologies, taking advantage of their different groove widths. However, to find new quadruplex groove binding agents, the chemical nature of the quadruplex grooves must be considered. Interestingly, all known quadruplex structures are characterized by grooves that are chemically and conformationally very different from the minor groove of the duplex DNA. This means that searching for new quadruplex groove binding agents among duplex minor groove binders may not be the most successful strategy. For this reason, a number of research groups have recently focused their attention on finding alternative molecular scaffolds able to recognize the groove of the quadruplex.[6,7] Here, we attempted to search for brand new molecular scaffolds able to interact with the groove of quadruplex structures by means of a structure-based virtual screening (VS) approach.[8] Actually, whereas there are numerous studies using these screening methods for targeting proteins, only few VS campaigns have been undertaken targeting nucleic acids, and, to the best of our knowledge, none targeting the groove of the quadruplex. In a recent review, Trent and co-workers[9] have outlined that the software Autodock optimally balances docking accuracy and ranking. Thus, as a starting point, this program was used in VS experiments aimed at targeting a very simple quadruplex, namely $[d(TGGGGT)]_4$ (PDB code 1S45).[10] This quadruplex possesses a 4-fold symmetry with all strands parallel to each other, which afford four grooves of identical medium width, and all nucleosides in an *anti* glycosidic conformation. Thus, the docking software Autodock4 (AD4) was used to dock a diversity set from the commercially available Life Chemicals database (6000 compounds). To avoid finding redundant information, a search area large enough to enclose only one of the four identical grooves was used (see Figure 1S in Supporting Information). The VS results were sorted on the basis of their predicted binding free energies (ΔG_{AD4}) which ranged from −0.95 to −9.55 kcal/mol. Solutions with a predicted binding free energy greater than −6.0 kcal/mol and a cluster size lower than 10 out of 100 individuals were discarded. Based on these criteria only 137 individuals were retained for further consideration. The binding poses calculated for these compounds were then visually inspected to discard all the individuals which were not predicted to establish tight interactions with the groove of the quadruplex structure. More precisely, compounds that were not able to form H-bonds with any of the guanine bases and/or to establish an electrostatic interaction with the backbone phosphate groups were not considered. After this final step, 30 compounds corresponding to 0.5% of the original Life Chemicals database were selected and purchased for further analysis.

The experimental test of the top computational "hits" for binding has been performed by NMR, which has the significant advantages

[†] Dipartimento di Chimica Farmaceutica e Tossicologica, Università degli Studi di Napoli "Federico II".
[‡] Dipartimento di Chimica delle Sostanze Naturali, Università degli Studi di Napoli "Federico II".
[§] Department of Molecular Biology, The Scripps Research Institute.

10.1021/ja9063662 CCC: $40.75

that it can detect weak binders and readily identify the ligand binding sites, consequently verifying the binding specificity.[11]

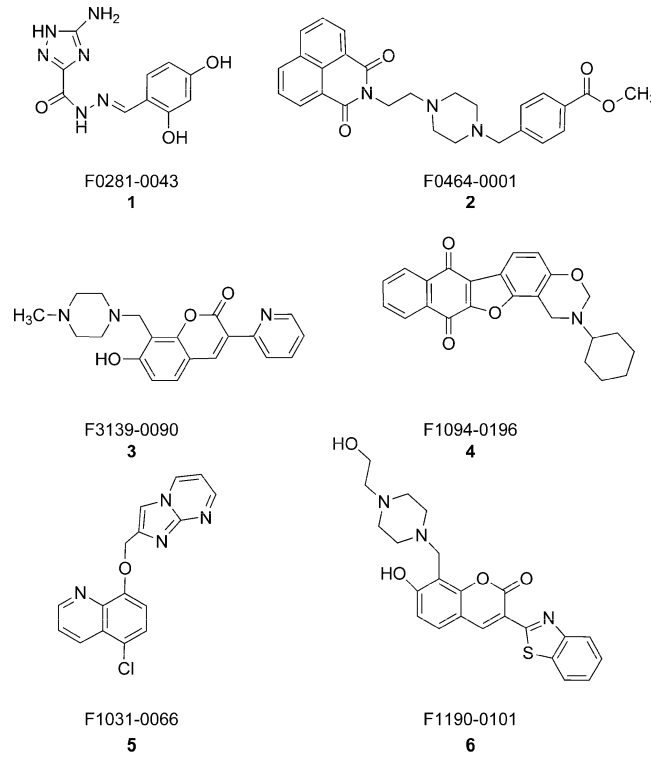
Thus, we prepared 30 identical DNA samples containing 6.8 mM of d(TGGGGT) (meaning 1.7 mM concentration of quadruplex), in 0.2 mL (H_2O/D_2O 9:1) of buffer solution having 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA, at pH 7.0. The first problem that we encountered in testing the selected compounds was solubility (most of them were not soluble in water). Nevertheless, all samples turned out to be soluble in DMSO. Therefore, we tested the stability of the quadruplex $[d(TGGGGT)]_4$ in buffers containing different percentages of DMSO. The quadruplex turned out to be perfectly structured even in buffer solutions containing 30% of DMSO. Therefore, we dissolved each equivalent of the selected compounds in 5 μL of DMSO, in such a way to gain an overall percentage of DMSO at the end of the titration not higher than 15%. The NMR titrations (700 MHz, $T = 25$ °C) were carried out monitoring resonance chemical shift changes of DNA, which were used to estimate whether a given compound is able to interact with the quadruplex and to determine the binding site.

In line with the expected results of a VS campaign, a number of false positives were found: (i) 8 molecules do not significantly interact with the quadruplex, (ii) 11 molecules caused a shift of the signal of Ts (T1 and T6), (iii) 3 molecules caused a shift of the resonances belonging to external bases (T1 and G2, T6 and G5), most probably due to an end-stacking interaction.

Outstandingly, six molecules (**1**−**6**, Chart 1) were found to cause an appreciable shift, among the others, of the G3 and G4 resonances signals, thus suggesting, as expected, a groove binding interaction. Particularly, compounds **1**−**5** provided NMR titration profiles very similar to each other, causing mainly drifting of the signals of G3, G4, G5, and T6, and so indicating that the recognition process involves mostly the 3′ side of the grooves. This is a very interesting result, since it is surprisingly consistent with the mode of binding calculated by the VS (see Supporting Information). On the other hand, **6** seems to entirely span the grooves, perturbing more uniformly all the residues of the quadruplex. In addition, **6** causes a major change of the resonances of the quadruplex, suggesting a higher affinity. Interestingly, while **1**−**4** and **6** are positively charged as are most of the already known groove binders, **5** does not possess any charge, so that the binding might be more driven by the presence of a number of H-bond acceptor heteroatoms. The titration of all 6 molecules turned out to be virtually completed at 4 equiv.

In summary, the application of VS calculations together with NMR experiments proved to be a successful strategy in the identification of new molecular chemotypes able to bind the grooves of DNA quadruplex structures. The structural diversity of these inhibitors has provided valuable alternative series for ongoing lead optimization aimed at the identification of brand new pharmacological tools; endowed with better affinity and a pharmacokinetic profile; and useful in the clarification of the mechanism, targeting, and therapeutic potential of G-quadruplexes. In the future we plan to apply this lead discovery approach targeting different and more complex G-quadruplex structures.

Chart 1. Structures of the Newly Identified Groove Binders[a]



[a] The Life Chemicals codes are reported in plain text. Numerals used in this paper are reported in bold.

Acknowledgment. This work is supported by Italian M.U.R.S.T. (P.R.I.N. 2007 and 2008) and Regione Campania (L.41, L.5).

Supporting Information Available: Experimental procedures, NMR titration of compounds **1**−**6**; 3D representation of the G-quadruplex searched area, and of the calculated (VS) complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402–5415.
(2) Johnson, J. E.; Smith, J. S.; Kozak, M. L.; Johnson, F. B. *Biochimie* **2008**, *90*, 1250–1263.
(3) Kelland, L. R. *Eur. J. Cancer* **2005**, *41*, 971–979.
(4) Ou, T.; Lu, Y.; Tan, J.; Huang, Z.; Wong, K.; Gu, L. *ChemMedChem* **2008**, *3*, 690–713.
(5) Martino, L.; Virno, A.; Pagano, B.; Virgilio, A.; Di Micco, S.; Galeone, A.; Giancola, C.; Bifulco, G.; Mayol, L.; Randazzo, A. *J. Am. Chem. Soc.* **2007**, *129*, 15950–15956.
(6) Dash, J.; Shirude, P. S.; Hsu, S. D.; Balasubramanian, S. *J. Am. Chem. Soc.* **2008**, *130*, 16048–16056.
(7) Li, Q.; Xiang, J.; Li, X.; Chen, L.; Xu, X.; Tang, Y.; Zhou, Q.; Li, L.; Zhang, H.; Sun, H.; Guan, A.; Yang, Q.; Yang, S.; Xu, G. *Biochimie* **2009**, *91*, 811–819.
(8) Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. *J. Comput. Chem.* **2007**, *28*, 1145–1152.
(9) Dailey, M. M.; Hait, C.; Holt, P. A.; Maguire, J. M.; Meier, J. B.; Miller, M. C.; Petraccone, L.; Trent, J. O. *Exp. Mol. Pathol.* **2009**, *86*, 141–150.
(10) Caceres, C.; Wright, G.; Gouyette, C.; Parkinson, G.; Subirana, J. A. *Nucleic Acids Res.* **2004**, *32*, 1097–1102.
(11) Pellecchia, M.; Bertini, I.; Cowburn, D.; Dalvit, C.; Giralt, E.; Jahnke, W.; James, T. L.; Homans, S. W.; Kesler, H.; Luchinat, C.; Meyer, B.; Oschkinat, H.; Peng, J.; Schwalbe, H.; Siegal, G. *Nat. Rev. Drug Discovery* **2008**, *7*, 738–745.

JA9063662