Proc. Natl. Acad. Sci. USA
Vol. 90, p. 3124, April 1993

Commentary

G-quartets in biology: Reprise

James R. Williamson

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139

The G-quartet is a cyclic hydrogen-bonded array of four guanine nucleotides that can be formed by G-rich sequences in both DNA and RNA. Several G-quartets can stack upon each other to form quadruple helical structures. Although many biological roles for G-quartets have been proposed (1–4), there is no proven role for this structure in biological function. However, recent work by Sundquist and Heaphy (5) leads one step closer to demonstrating a biological role for G-quartets. They have shown that human immunodeficiency virus (HIV) genomic RNA dimerizes via G-quartet formation *in vitro* and proposed that this dimerization occurs as well *in vivo*.

Although the G-quartet structure was first proposed almost 30 years ago (6), there has been a recent resurgence of interest because of its possible role in telomere function. Telomeric DNA sequences from most organisms are G-rich tandem repeats, such as $d(TTTTGGG-G)_n$, $d(TTGGGG)_n$, or $d(TTAGGG)_n$. Oligonucleotides from a variety of telomeric DNA sequences readily form G-quartets *in vitro* under physiological salt concentrations, and much of our knowledge about G-quartet structures is the result of biophysical studies on telomeric oligonucleotides. The most interesting characteristic of the G-quartet structure is their selective interaction with certain monovalent ions. G-quartet structures greatly prefer potassium ions to sodium ions, and the potassium-specific stabilization of G-quartets has become their hallmark.

A variety of proteins have been found that interact with telomeric DNAs, but none appear to interact directly with the G-quartet structure. A telomere-binding protein has been found in a variety of protozoans that binds to a 3′-terminal single-stranded overhang. However, formation of G-quartet structures inhibits binding of this protein (7). Telomere terminal transferase also recognizes single-stranded telomeric DNAs as substrates, and in this case, G-quartet formation inhibits the enzyme (8). In addition, healing of telomeres in yeast by addition of telomeric DNA sequences to plasmid substrates does not seem to require G-quartet formation (9). Thus, although telomeric DNAs exhibit a strong tendency to form G-quartet structures *in vitro*, evidence for a functional role for these structures remains elusive.

G-quartet structures can form by the association of G-rich segments from one, two, or four different strands. Unimolecular G-quartet formation requires four nearby G-rich segments that fold back on themselves to form the quadruple helical structure (3). Such structures have been proposed for telomeric DNAs during replication, but it is uncertain what functional role this type of structure might play. Formation of G-quartets by four different DNA molecules is also possible, and it has been suggested that such an interaction might align chromosomes during meiosis for recombination (1).

Dimerization of two sets of G-rich sequences is the type of G-quartet structure most likely to be observed biologically. The G-rich regions in each strand first fold back on themselves to form G·G base-paired hairpins, and these two hairpins can then dimerize by forming G-quartets. It has been proposed that the 3′ terminal G-rich single-stranded regions of telomeres can dimerize, resulting in the association of two chromosomes (2, 3). Two different modes of G-quartet-mediated dimerization have been observed in recent structures of telomeric oligonucleotides determined by x-ray and NMR (10, 11). The dimerization of telomeric DNA sequences might be useful in organizing chromatin in the nucleus or for alignment of chromosomes during replication or recombination. The association of chromosomes via telomeres may be particularly important in the macronuclei of ciliated protozoans that contain millions of chromosomes (12, 13).

The recent report by Sundquist and Heaphy reveals that dimerization of HIV genomic RNA *in vitro* is dependent on the presence of G-rich sequences (5, 14). Since dimerization is facilitated by the presence of potassium ions, Sundquist and Heaphy conclude that HIV genomic RNAs can dimerize *in vitro* via G-quartet formation. Packaging of retroviral RNA involves dimerization via RNA·RNA contacts as well as association with the gag protein. The RNA·RNA contacts and the RNA structure appear to be important for dimerization, since the dimer remains associated after removal of the proteins. In general, RNAs tend to exhibit a more complex tertiary structure than do DNAs, and it may be likely to find G-quartets in the context of other RNA structures as well. Although there is no firm evidence that G-quartets are involved in retroviral dimerization *in vivo*, this work is a significant step toward understanding dimerization process and suggests a new biological role for G-quartets.

G-quartet-mediated dimerization of two similar or identical strands is an appealing biological role for G-quartet structures. The requirements for G-quartet formation are simply single-stranded G-rich sequences and the presence of potassium ions. In the case of telomeres, two like chromosomal ends can be associated by the G-rich telomeric sequences. In the case of the retroviral genome, two identical strands must be packaged into a virion. The presence of highly conserved G-rich sequences in telomeres is at present the only hint at a biological role for G-quartets in telomere function. We can now include dimerization of retroviral genomes as a contender for finally providing a biological role for this most interesting nucleic acid structure.

1. Sen, D. & Gilbert, W. (1988) *Nature (London)* **334**, 364–366.
2. Sundquist, W. I. & Klug, A. (1989) *Nature (London)* **342**, 825–829.
3. Williamson, J. R., Raghuraman, M. K. & Cech, T. R. (1989) *Cell* **59**, 871–880.
4. Sen, D. & Gilbert, W. (1991) *Nature (London)* **344**, 410–414.
5. Sundquist, W. I. & Heaphy, S. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 3393–3397.
6. Gellert M., Lipsett, M. N. & Davies, D. R. (1962) *Proc. Natl. Acad. Sci. USA* **48**, 2013–2018.
7. Raghuraman, M. K. & Cech, T. R. (1990) *Nucleic Acids Res.* **18**, 4543–4552.
8. Zahler, A. M., Williamson, J. R., Cech, T. R. & Prescott, D. M. (1991) *Nature (London)* **350**, 718–720.
9. Lustig, A. J. (1992) *Nucleic Acids Res.* **20**, 3021–3028.
10. Kang, C. H., Zhang, X., Ratliff, R., Moyzis, R. & Rich, A. (1992) *Nature (London)* **356**, 126–131.
11. Smith, F. W. & Feigon, J. (1992) *Nature (London)* **356**, 164–168.
12. Oka, Y. & Thomas, C. A., Jr. (1987) *Nucleic Acids Res.* **15**, 8877–8898.
13. Acevedo, O. L., Dickinson, L. A., Macke, T. J. & Thomas, C. A., Jr. (1991) *Nucleic Acids Res.* **19**, 3409–3419.
14. Marquet, R., Baudin, F., Gabus, C., Darlix, J.-L., Mougel, M., Ehresmann, C. & Ehresmann, B. (1991) *Nucleic Acids Res.* **19**, 2349–2357.