ELSEVIER

Biophysical Chemistry 126 (2007) 140–153

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

Structure-specific recognition of quadruplex DNA by organic cations: Influence of shape, substituents and charge

Elizabeth W. White [a,1], Farial Tanious [a,1], Mohamed A. Ismail [a], Anthony P. Reszka [b], Stephen Neidle [b], David W. Boykin [a], W. David Wilson [a,*]

[a] *Department of Chemistry, Georgia State University, Atlanta, GA 30303, United States*
[b] *Cancer Research UK Biomolecular Structure Group and The School of Pharmacy, University of London, London WC1N 1AX, UK*

Received 7 March 2006; received in revised form 13 June 2006; accepted 13 June 2006
Available online 20 June 2006

Abstract

Combining structure-specific recognition of nucleic acids with limited sequence reading is a promising method to reduce the size of the recognition unit required to achieve the necessary selectivity and binding affinity to control function. It has been demonstrated recently that G-quadruplex DNA structures can be targeted by organic cations in a structure-specific manner. Structural targets of quadruplexes include the planar end surfaces of the G-tetrad stacked columns and four grooves. These provide different geometries and functional groups relative to duplex DNA. We have used surface plasmon resonance and isothermal titration calorimetry to show that binding affinity and selectivity of a series of quadruplex end-stacking molecules to human telomeric DNA are sensitive to compound shape as well as substituent type and position. ITC results indicate that binding is largely enthalpy driven. Circular dichroism was also used to identify a group of structurally related compounds that selectively target quadruplex grooves.

Keywords: Synthetic cations; Telomere; DNA-quadruplex binding; G-tetrad; BIAcore; ITC

1. Introduction

Specific recognition of double helical sequences of base pairs in DNA, for example, by transcription control proteins [1] and designed synthetic molecules [2,3], has been thoroughly studied and many sequence-specific recognition complexes are now characterized in molecular detail. Therapeutic intervention or control of cellular function through specific recognition of a cellular duplex DNA sequence requires interaction with a large number of base pairs. This is routinely accomplished by cellular proteins but is more difficult to achieve with relatively small synthetic compounds that must have properties that allow them to pass through barriers, including cell membranes, in order to bind to DNA. Incorporating nucleic acid structure into the recognition motif, however, is a promising method to reduce the size of the recognition sequence required to gain the necessary specificity. Specific recognition of RNA structures by small molecules that interact with only a few bases or base pairs, for example, is well established for aminoglycoside antibiotics that target ribosomal RNA [4–6]. The recent discovery of small metabolites and analogs that bind specifically to RNA riboswitch structures is now a demonstrated method for specific control of translation [7]. Selection methods that yield nucleic acids that are capable of highly specific binding to small molecules provide an additional example of recognition by structure-specific motifs [8]. Examples of structure-specific targeting are also emerging for selective recognition of DNA. The mitochondrial kinetoplast DNA of kinetoplastid eukaryotic parasites, which cause serious diseases that affect millions of people, for example, has a complex structure that is composed of a few large circular DNAs in an interlocked, catenated array with thousands of minicircular DNAs [9–11]. The minicircular

* Corresponding author. Tel.: +1 404 651 3903; fax: +1 404 651 2751.
E-mail address: wdw@gsu.edu (W.D. Wilson).
[1] These two authors contributed equally to this paper.

0301-4622/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2006.06.006

DNAs have phased AT sequence tracts that bend the DNA duplex [12] and provide an optimum sequence-structure target for drug design [13].

Eukaryotic cells of the parasites as well as of higher organisms have a 3′ terminal G-rich, single-stranded telomere DNA sequence that performs several essential functions in chromosome maintenance and protection. The G-rich strand can fold into a four-stranded quadruplex structure that is an attractive potential structure-specific target in rapidly dividing cells, such as eukaryotic parasites and cancer [14]. The key structural feature of the quadruplex is a series of stacked guanine tetrads held together in a coplanar cyclic array by Hoogsteen and Watson–Crick hydrogen bonds (Fig. 1). The quadruplex is also stabilized through π–π stacking interactions of the stacked tetrads as well as by coordination with cations located between or within the tetrads. Guanine-rich sequences, which are capable of forming quadruplex structures, are present in biologically significant regions of the genome including immunoglobulin switch regions [15], the transcriptional regulatory regions of a number of genes such as the insulin gene [16], and also the promoter regions of certain oncogenes, such as c-MYC [17,18].

Because of the critically essential roles of telomere DNAs in both cancer and parasitic cells, the telomere is a particularly attractive target for drug design [19–26]. The telomeric sequence and structure varies depending on the organism. In humans and other vertebrates and the eukaryotic parasites, telomeres consist of tandem T_2AG_3 repeats that can adopt a G-quadruplex conformation *in vitro* under physiological conditions [27,28]. The discovery of proteins such as transcription factors, nucleases and helicases that can bind to and even promote the formation of telomeric quadruplexes suggests that these structures may exist *in vivo* under certain conditions [29–31]. A very exciting recent finding is that a radiolabeled G-quadruplex binding ligand accumulated in nuclei of cultured cells and preferentially bound to the terminal regions of the chromosomes, indicating that G-quadruplexes do exist in vivo and are accessible to drugs [32]. Each time a cell replicates, DNA polymerase is unable to copy the extreme end of the 3′, lagging strand. This "end replication problem" results in shortening of the telomere by about 30–200 bases per cell doubling. After 60–70 rounds of cell replication, the telomeres reach a critical length and the cells enter a non-dividing state called senescence, which leads to apoptosis and eventually cell death [33]. However, in 85–90% of cancer cells and in parasites, telomerase is activated [34]. This enzyme is inactive in most normal somatic cells, providing a potentially specific therapeutic target. In order to extend telomere DNA the telomerase enzyme requires the telomere primer to be single stranded. The formation of higher ordered structures such as G-quadruplexes prevents hybridization of the telomerase RNA template onto the primer and thus inhibits telomerase activity through an indirect topological mechanism [35].

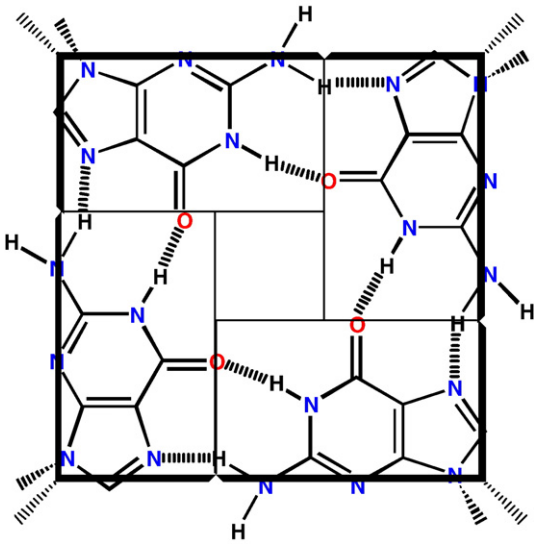

Fig. 1. A molecular representation of an H-bonded G-tetrad is shown overlaid with a schematic of the tetrad that we will use for a size reference for synthetic molecules in figures that follow. As can be seen, the G-tetrad has eight hydrogen bonds with a mix of Watson–Crick and Hoogsteen-type interactions. The four sugar-phosphate chains project away from the tetrad and create four grooves that can have different dimensions depending on the quadruplex geometry.

Stabilization of the quadruplex conformation of telomeres, such as by binding small molecules, has been shown to be an effective method to inhibit telomerase activity [19–26,36]. The development of small molecules that can selectively bind to and stabilize the G-quadruplex conformation of the telomere is therefore a current area of interest in anticancer as well as antiparasitic drug design. Compounds that have been shown to bind to quadruplex DNA have traditionally been planar, aromatic compounds that bind via external end-stacking to the G-quartet on either one end or both ends of the quadruplex [19–26]. These compounds, which include anthraquinones, cationic porphyrins, acridines, macrocyclic compounds and analogs, have planar aromatic surface areas that mimic the large planar surface of the G-tetrads in quadruplex DNA (Fig. 1) [37–40]. Since essentially all known quadruplex DNA binders are based on, or derived from duplex intercalators, many exhibit little selectivity for quadruplex over duplex structures and this can result in nonspecific cytotoxicity. Increasing the selectivity of telomerase inhibitors for their quadruplex targets is an important focus of research.

Groove binding has been a useful way to selectively recognize duplex DNA with relatively low nonspecific toxicity [2,3,13,41–43], but groove binding is an under-exploited design mode with DNA quadruplexes. The structural differences between duplex and quadruplex DNA grooves offer an attractive strategy for development of compounds to differentiate between these two structures. Since groove dimensions vary according to the type of quadruplex [44], groove binding also offers the opportunity for obtaining increased selectivity for a particular quadruplex structure. Unfortunately, no compounds to date have been found which bind to the grooves of the intramolecular human or parasite telomeres or to oncogene control quadruplexes. Shafer et al. obtained spectroscopic data suggesting that the dye 3,3′-diethyloxadicarbocyanine (DODC) binds in the quadruplex grooves of a dimeric hairpin G-quadruplex [45]. Satellite hole spectroscopy studies support the groove-binding model for this particular compound/DNA pair