Abstract

A largely increased number of d(G_4C_2) repeats located in the non-coding region of *C9orf72* gene has been identified as the leading cause of two related neurological disorders, familial amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). Non-canonical structures including G-quadruplexes adopted by expanded repeats are hypothesized to be crucial in pathogenesis. We examined the G-quadruplex forming ability of d(G_4C_2)-repeat containing oligonucleotides with four guanine tracts, chosen as the smallest possible model that can form a unimolecular G-quadruplex. Oligonucleotides formed a mixture of G-quadruplex structures in the presence of K^+ ions, making detailed structural characterization difficult. Oligonucleotide wt22, d[$(G_4C_2)_3G_4$], displayed the most favorable ^{1}H NMR spectra and showed formation of two predominant G-quadruplex structures with antiparallel orientation of strands and *syn* glycosidic conformation of some guanine residues. Since 8Br-dG residue adopts a *syn* conformation, we designed the sequence sl21, d[$(G_4C_2)_3GG^{Br}GG$], with dG to 8Br-dG substitution at position 21, that was expected to adopt a *syn* conformation in only one of predicted topologies. Surprisingly, while dG to 8Br-dG substitution did reduce formation of low populated species thus significantly improving NMR spectral characteristics of sl21 compared to wt22, it did not drive the equilibrium to one of the two structures. Instead, relative populations of two structures depend on folding conditions. Two distinct folding condition were established that favor formation of mostly a single structure, which facilitated their structural characterization with NMR. Neutral pH and slow cooling (annealing) of sl21 lead to formation of a major G-quadruplex structure denoted NAN, and 30% of minor structure, denoted AQU. Acidic pH and fast cooling (quenching) lead to formation of 80% AQU and 20% NAN. Structure of NAN determined with NMR to high-resolution showed a compact antiparallel fold with four G-quartets, where one of the cytosine residues in each lateral C-C loop is stacked over the nearby G-quartet. Novel AQU fold is antiparallel and comprises four G-quartets and three edgewise loops. Although topologically similar to NAN, high-resolution structure of AQU revealed specific structural features. Lateral C11-C12 loop is stacked on the outer G-quartet on the 5' and 3' side of AQU structure, while the outer G-quartet on the opposite side of the structure features cytosine residues arranged in two C•C base pairs. Interestingly, protection of imino protons assessed by hydrogen-deuterium exchange showed distinct dynamics of base pairs within G-quartets in AQU compared to NAN. Oligonucleotides with sequences extended from four to eight d(G_4C_2) repeats fold into unimolecular antiparallel G-quadruplexes. The distribution of imino proton signals in the ^{1}H NMR spectra of longer oligonucleotides is similar to distribution of signals corresponding to AQU and NAN in wt22, which suggests that the two structures could form in the presence of additional d(G_4C_2) repeats.