Proc Indian Natn Sci Acad, **70, A,** No. 4, July 2004, pp. 605–613
© Printed in India.

NOVEL STRUCTURAL MOTIFS IN TELOMERE DNA SEQUENCES

NEEL S BHAVESH[†], PRASANTA K PATEL[‡] AND RAMAKRISHNA V HOSUR[†*]

[†]*Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005 (India)*

(Received 13 March 2003 ; Accepted 28 March 2003)

Maintenance of linear DNA in chromosomes during replication and other biological functions of telomeric DNA and its associated proteins have been widely investigated in recent years. Advances in NMR methodologies have led to characterization of many telomeric repeat sequences, which have exhibited a great degree of structural polymorphism. Many new structural motifs and folding topologies have been discovered in telomeric DNA repeats. A number of proteins have also been identified which have specific binding for higher order structures formed by telomeric DNA repeat sequences. In this review, we have discussed the various new motifs experimentally observed in telomeric repeat sequences.

Key Words : Telomere; DNA; Quadruplex; Tetrad; *i*-Motif; NMR

Introduction

The ultimate mystery of life is how a single cell, dividing and subdividing again and again, in the well-ordered drama of conception, birth, and development, becomes in time a complex, sentiment organism such as man. Within a living cell there go on a large number and variety of tightly regulated biochemical processes, almost all of which involve, or are controlled by, large helical molecules called deoxyribose nucleic acid (DNA). Four bases, which make nucleic acids, are, with minor reservations, same throughout *Nature*. They carry the genetic information in any organism, although certain small viruses use RNA as their genetic material. Much of this information is used to determine the amino acid sequence of the proteins of that organism.

DNA exists as a compact mass, occupying a limited volume; and its various activities, such as replication and transcription, are confined within these confines. The organization of this type accommodates transitions between inactive and active states. So, in contrast with the customary picture of DNA as an extended double helix, structural deformation of DNA to bend or fold into a more compact form is a rule rather than an exception. In eukaryotes, DNA is organized with the help of certain basic proteins like histones into nucleosome and further into chromosome. An essential feature of all chromosomes is the 'telomere', which 'seals' the end and confers stability to the *linear* molecule. It has a special structure as it is sticky and tends to react with other chromosomes, whereas natural ends are stable.

Telomeres are nucleo-protein structures containing a long stretch of duplex DNA (hundreds to thousands of base pair) with evolutionarily conserved tandem repeats of short G-rich sequences (6-10 nucleotide in each repeat) in one strand (in 5' to 3' direction) and their complimentary sequences in the other strand[1-10]. The G-rich strand contains about 2-3 repeats of a single strand overhang at 3' end except in the case of telomere in human somatic cells, which have 20-30 repeats[11-13]. A unique feature of this 3' overhang is its exceptional resistance to nuclease degradation[14-15]. Although the telomeric DNA contains evolutionarily conserved tandem repeats of short G-rich sequences, the exact telomeric sequence repeat varies among species[4,5] (Table I). Detailed information on telomere and telomere proteins can be found at http://www.genlink.wustl.edu/teldb/index.html.

Telomeres have two important duties. First, they are caps that protect chromosome ends from vandalistic nucleases and other damaging activities such as unwanted end-joining events. Second, they enable the ends of chromosomes to be completely replicated. It has also been shown that telomeres may function as regulators for gene expression[16]. The length of telomeric DNA is related to aging and diseases like cancer[1-5,17-20]. Thus, detailed structural and dynamic

*Author for correspondence : *E-mail :* hosur@tifr.res.in
Phone : 091-22-2215-2971 Ext. 2488
Fax : 091-22-2215-2110

[‡]Present Address : University of Massachusetts, Program in Molecular Medicine, Biotech-II, Suite 309, 373 Plantation Street, Worcester, MA 01605 (USA)

Table I

Telomeric Repeat Sequences

Group/organism	Telomeric repeats
Ciliated protozoa (*Euplotes, Oxytricha, Stylo*nichia)	...T_4G_4...
Ciliated protozoa (*Tetrahymena, Glaucoma*)	...T_2G_4...
Ciliated protozoa (*Paramecium*)	...T_2G_3(G/T)...
Coccidial protozoa (*Plasmodium*)	...T_2(C/T)AG_3...
Budding yeast (*Saccharomyces cerevisae*)	...$TG_{1\text{-}3}$...
Alga (*Chlamydomonas*)	...T_4AG_3...
Mammals (*Homosapiens sapiens*)	
Slime mold (*Physarum, Didymium*)	...T_2AG_3...
Filamentous fungus (*Neurospopra*)	
Kinetoplastid protozoa (*Trypanosoma, Crithidia*)	
Higher plant (*Arabidopsis*)	...T_3AG_3...
Fission yeast (*Saccharomyces prombe*)	...$T_2AC(A)G_{2\text{-}5}$...
Cellular slime mould (*Dictyostelium*)	...$AG_{1\text{-}8}$...

characterization of telomere is important for *better* understanding of life and it could have important implication for curing such diseases. Pursuing this goal many structures of telomeric DNA have been solved and they have added to the family of multistranded structures[21-25]. Proteins that interact with telomeric G-strand, C-strand and duplexes have been identified and could play a role in the relative stabilities of these structures[26]. A possible model for organization of duplex, *i*-motif and G-quadruplex in human telomeric DNA has been recently proposed[24]. Duplex-quadruplex inter-conversion has been studied for a number of different sequences[27-33].

Discovery of a number of proteins binding specifically to the quadruplex DNA[34-46] and cytosine rich sequences[47-51] lead to the belief that these higher order structures may be involved in several biological functions[23,52-54]. Elongation step of telomerase is known to get inactivated by quadruplex formation[55-56]. Analogues of a 2, 4, 6-triamino-1, 3, 5-triazine which stabilize the quadruplex structure help in arresting the growth of the cells. This provides indirect evidence for the *in vivo* formation of quadruplex and the compound has potential for use as anti-tumor drug[57]. Discovery of nuclease[37] and helicases[39,41,43,44,46] specific for quadruplex DNA provide strong evidence that these higher order structures may be playing regulatory functions in the cell. The current evidence on prokaryotic linear chromosomes and the eukaryotes that do not use telomerase and quadruplex DNA has led to the suggestion that quadruplex DNA might have played a role in the evolution of the protection mechanism of linear chromosomes rather than in overcoming the end replication problem[58].

A recent 1.8 Å crystal structure of a complex of a rat hepatocytes protein uqTBP25 and *Oxythricha* telomere d($G_4T_4G_4$) repeat DNA showed the formation of hairpin dimer quadruplex in addition to single stranded state. This quadruplex occupies a cavity created by three protein molecules in the crystal and interacts with each protein through either one of its two loops or one of its grooves[59]. The protein uqTBP25 is known to bind both unfolded, single-stranded telomeric DNA as well as the same DNA folded unimolecularly into a quadruplex[40].

NMR Studies of Telomeric DNA

Nuclear Magnetic Resonance (NMR) spectroscopy has been extensively used for *in vitro* structural characterization of DNA sequences[60]. Precise description of structure and dynamics of DNA in solution and difficulty in getting good crystals of nucleic acids gives NMR an added advantage over X-ray crystallography. This is evident from the number of DNA structures deposited in the Protein Data Bank (PDB). Though X-ray crystallography is a much older technique, there are almost the same number of DNA structures solved by NMR spectroscopy as by X-ray crystallography[61].

Exchangeable imino and amino protons resonances in the NMR spectra have characteristic shifts according to the structure present in the molecule under investigation. Imino proton chemical shifts are also indicative of Watson-Crick (>12 ppm), Hoogsteen (10.5-12.0 ppm), C-C^+ (16-17 ppm) type of base pairing. Multiple quantum NMR spectroscopy helps in identification of spin systems in oligonucleotides[62]. Uniform and selective ^{13}C and/or ^{15}N isotope labelling and heteronuclear experiments have been very useful for resonance assignment in complex DNA structures[23].

J correlated and ^{1}H-^{31}P experiments provide dihedral restraints[63], while NOESY provides distance restraints for structure calculations. Isotope labelling and development of new NMR techniques have also helped in direct identification of H-bond via measurement of scalar coupling constants across the H-bonds[64]. Dipolar coupling restraints are used to fix H-bond length[65] and refinement of structures[66-69]. Using these restraints structure of DNA is calculated by simulated annealing or distance geometry techniques followed by restrained molecular dynamics and energy minimization[70-72]. Relaxation matrix analysis is often used for structure refinements.

G-Quadruplex Structure and Variety in Telomere DNA Sequences

NMR has contributed immensely in delineating the details of structure and variety in G-quadruplex. G-quadruplex structures are formed by association of four strands of continuous G-stretches (Fig 1)[23,73-77]. This relies on the property of G-nucleotides that four G bases in a plane can form a G-tetrad via Hoogsteen type H-bonds between the adjacent bases as shown in Fig 1a. Whenever, there are several continuous G's in a sequence, such G-tetrads would stack on top of each other to give rise to a highly stable helical structure, namely, G-quadruplex. G-stretch associations can be inter-molecular (Fig 1b, 1c, 1d) or intra-molecular (Fig 1e) in nature. In intra-molecular structures, a single nucleotide chain having four stretches of G nucleotides folds into a unimolecular quadruplex. The chain folds three times to bring the four stretches of G nucleotides in proper alignment (Fig 1e). An inter-molecular quadruplex is formed by association of two hairpins having two G-stretches each (Fig 1d) or by linear association of four different strands forming either a parallel (Fig 1b) or an anti-parallel quadruplex (Fig 1c) depending on the orientation of the four strand. The structure of the G-quadruplex depends critically on the exact sequence, chain length, salt type and its concentration[78]. Effects of divalent and monovalent cations on the G-tetrad have been extensively studied[79-80] and recently an interesting Ca^{2+} induced structural transition from the antiparallel to parallel G-quadruplex, and finally G-wire formation has been also observed[81]. Theoretical calculations have been performed to understand the effects of cations in the formation of the tetrads[82]. Thermodynamic and kinetics of dissociation and assembly of quadruplexes has also been studied[83]. The biological aspects of quadruplex DNA has been recently reviewed[52-54].

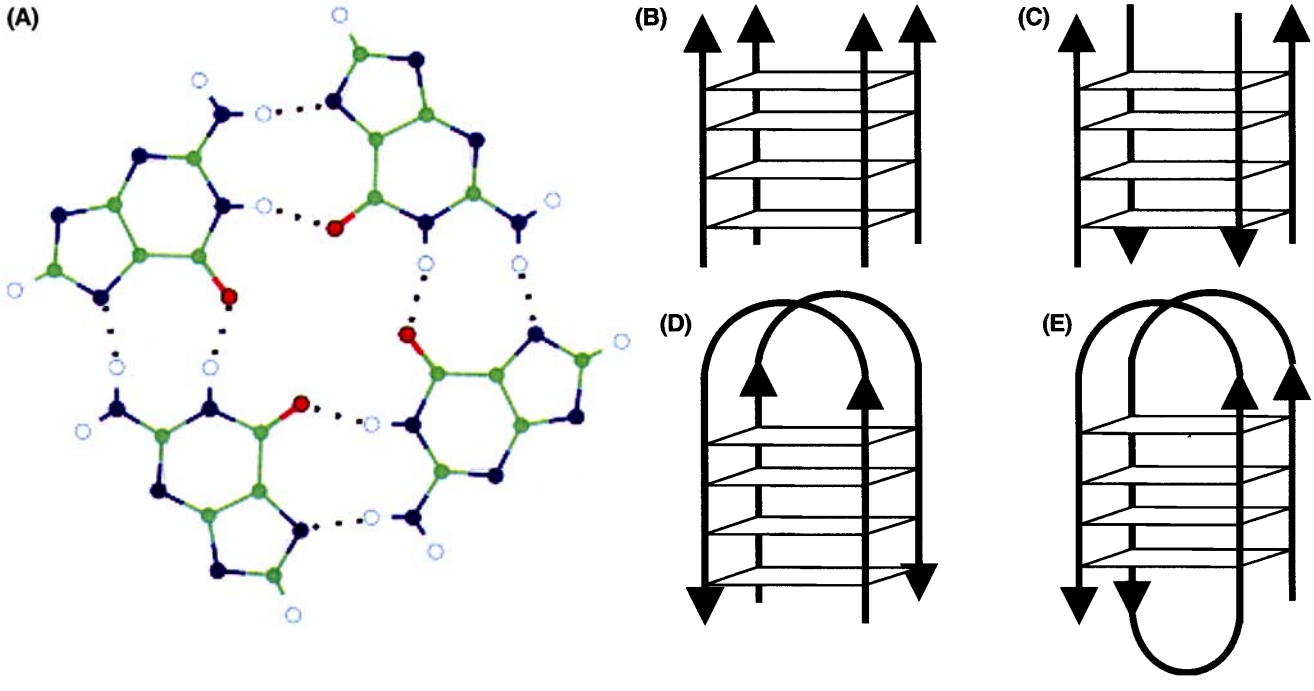

Fig. 1 The schematic representation of the H-bond pairing seen in G-tetrad (a) and different G-quadruplex structures. Intermolecular quadruplex can be formed by linear association of four different strands, either in parallel (b) or anti parallel (c) fashion or by dimerization of two hairpins (d) where as intermolecular quadruplex is formed by extensive folding of a single strand DNA having four streches of G-nucleotides (e). Each shaded plane represents a G-tetrad.

As can be seen from Table I, the telomeric repeat sequences contain also other nucleotides like A/T/C. Studies on these telomeric sequences have shown that while G nucleotides form the basis of G Quadruplex, T nucleotides participate in the loop structures. Various triad, tetrad, pentad, hexad and heptad structures have been reported in various designed DNA sequences, which have showed interesting H-bonding patterns between different nucleotides (reviewed in ref. [25]). The structural features of quadruplexes involving novel A/T/C tetrad are discussed below.

T-Tetrad

The role of T nucleotide has been investigated in the various structures solved for different telomeric repeat sequences[84-85]. In general, the T nucleotide takes part in the formation of a loop and various different kinds of loop structures in G quadruplexes have been extensively investigated (reviewed in ref.[23,25]). To further investigate the role of single T nucleotide, the structure of d-TGGTGGC containing two repeats of TG_2 sequence of *Saccharomyces cerevisiae* telomere DNA was solved at physiological pH by NMR methods[86]. Interestingly, in the NMR spectrum T4-imino and methyl proton were shifted quite upfield to 9.6 ppm and 0.78 ppm respectively indicating the strong shielding ring current shifts from adjacent guanine bases. This indicated proper stacking of T4 base over G tetrads. Strong NOE cross peaks from T4NH to T4Me and G5NH protons and melting curve spectra led to the conclusion about the presence of a symmetrical T tetrad with four Ts in a plane held together by H-bonds. Similar characteristic features were also observed for T1 at lower pH (4.8) and temperature (5° C). T tetrad formation can occur via O4-H3 H-bond. The formation of T tetrad caused slight underwinding of the helix in the middle (Fig. 2) causing a decrease in the intertwining of the helices as compared to other G quadruplex structures[87].

A-Tetrad

Telomeric DNA of human contains about 5 to 8 Kb tandem repeats of the sequence d-TTAGGG with a 3' overhang of 100-200 bases. Various structures solved for this repeat sequences have shown the formation of parallel/anti-parallel quadruplexes[88-90], *i*-motif or loop[24]. The solution structure of the sequence d-AGGGT at physiological pH showed the formation of a novel A-tetrad in the parallel stranded quadruplex[89,91]. Intense cross peaks from $A1NH_2$ to A1H8 and A1H2 protons in low mixing time NOESY spectrum provided strong evidence for an inter-strand interaction where the A-bases come close and arrange in a symmetrical fashion. At the end of structural calculation, the final structures were seen to be distributed between two distinct patterns of A alignment which were termed as N61 and N67 having $A1NH_2$ H-bond to N1 and N7 respectively (Fig. 3).

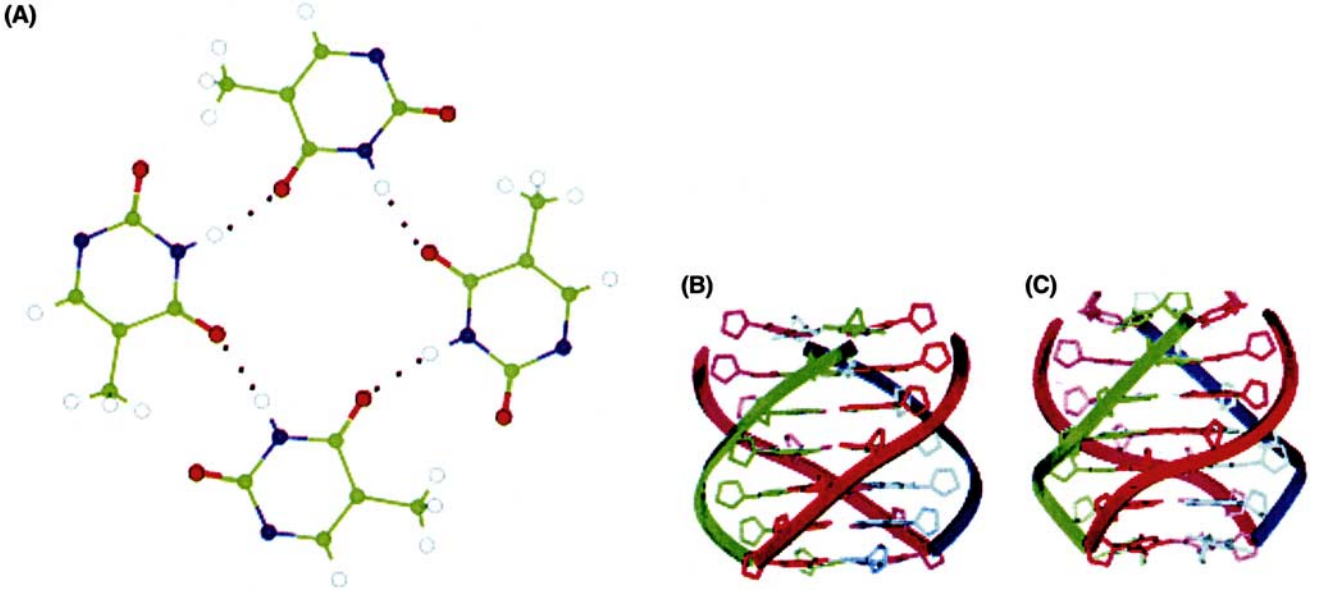

Fig. 2 The H-bonding pattern in the T4 nucleotides in the average structure of the sequence d-TGGTGGC (a). Comparison of quadruplex structures of the DNA sequence d-TGGTGGC, (b) and d-TTGGGGT, (c) determined in presence of 100mM KCl. The flattening of the back bone in the latter quadruplex can be seen.

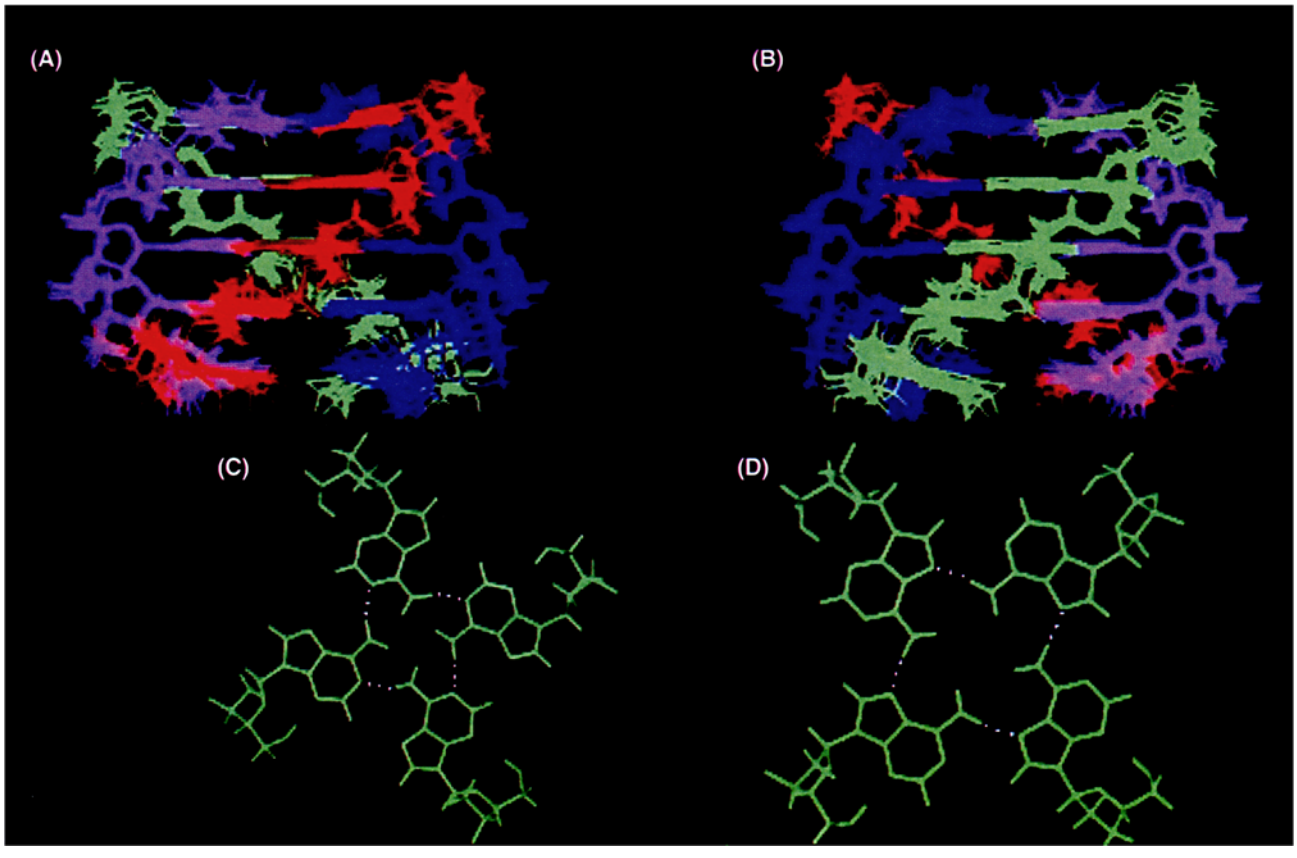

Fig. 3 Two distinct patterns of A-tetrad formation, N61 (c) and N67 (d) depending on the nitrogen atom on *A* accepting the amino hydrogen in each of the H-bonds observed in the structure of d-AGGGT with corresponding views of superpositions of 8 relaxation matrix refined structures (a and b).

Simultaneous observation of A1NH_2-A1H8 and A1NH_2-A1H2 NOEs and presence of only one resonance for each proton indicated the existence of dynamism in the structure. A rapid equilibrium between the two modes accounted for the chemical shift of 6.8 ppm, for both A1NH_2 protons, since both of them would be exchanging between H-bonded state (8.0-9.5 ppm) and a free state (5.0-6.5 ppm). Theoretical calculations on these tetrads have found N67 model to be more stable, which has a more planar A alignment[92-93].

C-Tetrad

Complementary strand of the G rich telomeric repeat sequence containing C nucleotide repeats has shown the formation tetrameric structures[94]. The formation a symmetrical C-tetrad stacked in parallel stranded G quadruplex was observed in the solution structures of the sequences d-TGGGCGGT, d-TGGCGGGT and d-TGGCGGC[31,95] (Fig. 4). A cross peak from CNH_2 to CH1' apart from CH5 indicated inter strand interaction. In the final structure symmetrical pairing of four Cs in a plane via NH_2-O_2 H-bond was seen (Fig. 4a). The formation of C-tetrad in the middle of the G quadruplex caused an under-winding of the helix. The cavity in the centre of the C-tetrad is larger than that in the G-tetrad. Formation of C-tetrad was also seen in the structure of the *Saccharomyces cerevisiae* telomere DNA repeat sequence d-TGGTGGC (Fig. 4b). A planar C-tetrad is formed at the 3' end with NH_2-N3 H-bonds. The cavity of this C-tetrad is smaller than the cavity in C-tetrad formed via NH_2-O_2 H-bond. Detailed theoretical calculations have been performed on possible arrangement of C-tetrad in presence and absence of different monovalent cations[82,93].

Interesting cation dependent duplex-qaudruplex conformation switch was observed in the sequence d-TG_2CG_2C. The sequence forms a parallel stranded quadruplex in the presence of K^+ ion while it forms a mismatch duplex in the presence of Na^+ ions at physiological pH. As the pH of the Na^+ sample is lowered to pH 2.2 the molecule transforms to a quadruplex[31].

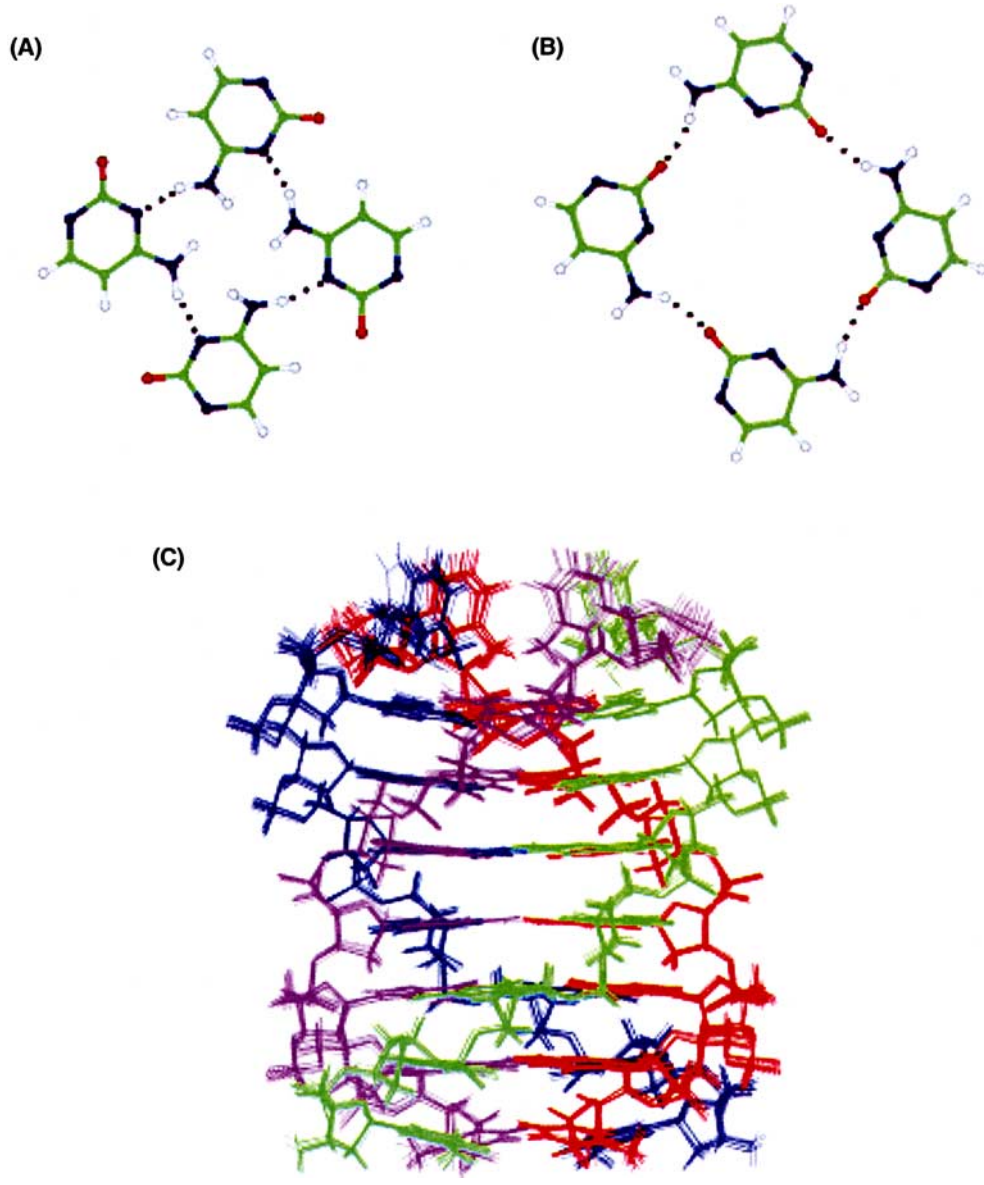

Fig. 4 Two distinct patterns of C-tetrad formation: (a) observed in the structure of the sequence d-TGGGCGGT, d-TGGCGGGT and d-TGGCGGC and, (b) observed in the structure of d-TGGTGGC at the 3' end, (c) superpositions of 10 relaxation matrix refined structures of d-TGGGCGGT

GCGC-Tetrad

A unique GCGC tetrad was observed in the structure of d(GGGCT$_4$GGGC)[96-98]. Simultaneous formation of ATAT and GCGC tetrads stabilized by inter subunit Watson-Crick AT and GC pairs was observed in d(GAGCAGGT) sequence[99]. These structures in telomere can form if there is any systematic mutation or there is involvement of the C rich strand of telomere in the formation quadruplex structures.

i-Motif

i-motif is a tetrameric DNA structure formed by the interdigitation of two parallel stranded duplexes containing C-C$^+$ base pairs running in opposite directions (Fig. 5b,c). The H-bonding pattern is shown in Fig. 5a[22,100,101]. This kind of association requires protonation of one of the cytosines at N3 position and consequently is better stabilized at acidic pH conditions[102]. This requirement of low pH, at first sight, suggests an unlikelyhood for *i*-motif occurrence *in vivo*. However it is conceivable that supercoiling of DNA may change the pKa values and likewise, binding of different proteins may be thought to provide low pH environments at localized positions[100,101,103]. As with G-quadruplex structures, *i*-motif structures may be formed by either intermolecular (Fig. 5b) or

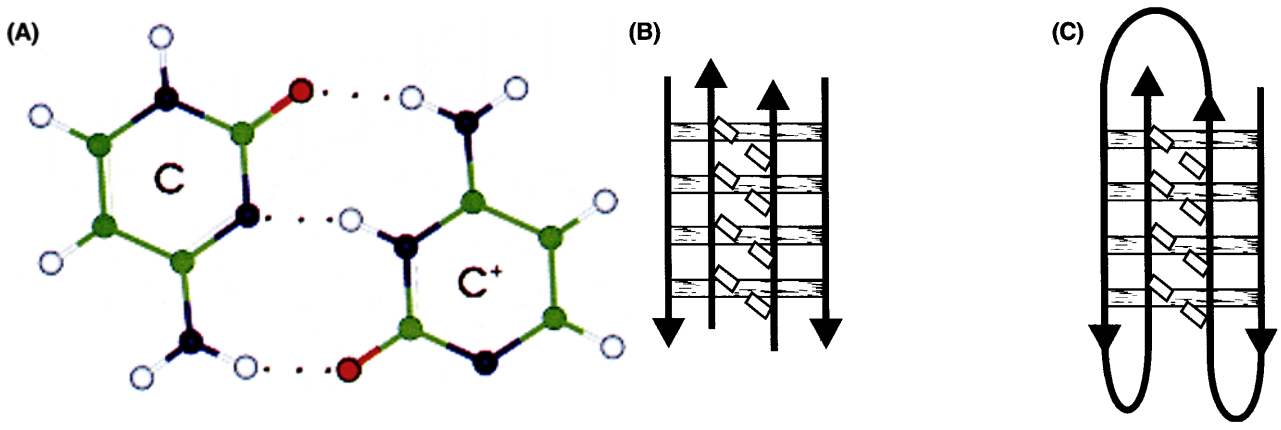

Fig. 5 (a) C-C$^+$ H-bond scheme seen in *i*-motif. (b,c) schematic drawing of intermolecular, (b) intramolecular, (c) *i*-motif. The shaded boxes indicate the base-pair formation between parallel strands.

intramolecular (Fig. 5c) associations depending on the exact DNA sequence and the length of the chain[104].

A novel *i*-motif topology was observed in the NMR structure of a human telomere repeating unit, d(CCCTAA). In this i-motif topology (T form) T4 is intercalated between C1 and C2 of the other duplex. Besides two other topologies: (i) where C1 is intercalated between C2 and C3 of the other parallel duplex, resulting in the non-stacking of T4 residues (R-form), and (ii) where C1 is stacked between C3 and T4 of the other duplex (S-form) have also been described[94]. Details about different *i*-motif structures have been reviewed in ref.[22].

Concluding Remarks

With the realization of biological and therapeutic significance of higher order telomeric DNA structures it has become very important to characterize these motifs in detail. The development of various experimental techniques and in particular the advent of newer NMR methodologies has provided a great impetus to the study of these structures and their interaction with different proteins at physiological conditions. pH and cation dependent conformational switches in telomeric repeats are important from regulation point of view. NMR has been very successful in structural characterization of many telomeric repeat sequences and it has a huge potential to provide high resolution information on dynamics and folding of telomeric DNA and its complexes with proteins and other ligands. Thus the structural and dynamics insight gained through all these investigations will be beneficial in understanding the various fundamental biological questions and will also be useful for designing drugs for many diseases. NMR is expected to provide deep insights into all these aspects in the years to come.

Acknowledgements

We thank the National Facility for High Field NMR at TIFR supported by Government of India. One of the authors (NSB) is recipient of the TIFR Alumni Association Scholarship for career development supported by TIFR endowment fund.

References

1 E H Blackburn, *J Biol Chem* **265** (1990) 5919
2 E H Blackburn and C W Greider (Eds.) *Telomeres* Cold Spring Harbor Laboratory, Plainview, New York (1995)
3 D Rhodes and R Giraldo *Curr Opin Struct Biol* **5** (1995) 311
4 V A Zakian *Annu Rev Genet* **23** (1989) 579
5 V A Zakian *Science* **270** (1995) 1601
6 E H Blackburn *Nature* **408** (2000) 53
7 E H Blackburn *Nat Struct Biol* **7** (2000) 847
8 K Collins *Curr Opin Struct Biol* **12** (2000) 378
9 M J McEachern, A Krauskopf and E H Blackburn *Annu Rev Genet***34** (2000) 331
10 D Rhodes, L Fairall, T Simonsson, R Court and L Chapman *EMBO Rep* **3** (2002) 1139
11 C W Greider and E H Blackburn *Nature* **337** (1985) 331
12 C W Greider and E H Blackburn *Cell* **43** (1985) 405
13 D Shippen-Lentz and E H Blackburn *Mol Cell Biol* **9** (1989) 2761
14 E H Blackburn *Molecular Structures of Telomeres in Lower Eukaryotes* (Ed. L Bogorad) Alan R Liss Inc (1986) 69
15 E R Henderson and E H Blackburn *Mol Cell Biol* **9** (1989) 345
16 M K Raghuraman, B J Brewer and W L Fangman *Science* **276** (1997) 806

17 D Sen and W Gilbert *Nature* **334** (1988) 364

18 J D Griffith, L Comeau, S Rosenfield, R M Stansel, A Bianchi, H Moss and T de Lange *Cell* **97** (1999) 503

19 D Shore *Curr Opin Genet Dev* **11** (2001) 189

20 H Cherif, J L Tarry, S E Ozanne and C N Hales *Nucleic Acids Res* **31** (2003) 1576

21 D E Gilbert and J Feigon *Curr Opin Struct Biol* **9** (1999) 305

22 M Guéron and J L Leroy *Curr Opin Struct Biol* **10** (2000) 326

23 M A Keniry *Biopolymer* **56** (2001) 123

24 A T Phan and J L Mergny *Nucleic Acids Res* **30** (2002) 4618

25 J Sühnel *Biopolymers* **61** (2002) 31

26 S M Kerwin *Curr Pharm Design* **6** (2000) 441

27 C C Hardin, T Watson, M Corregan and C Bailey *Biochemistry* **31** (1992) 833

28 T Miura and G J Thomas Jr *Biochemistry* **33** (1994) 7848

29 H Deng and W H Braunlin *Biopolymers* **35** (1995) 677

30 M Salazar, B D Thompson, S M Kerwin and L H Hurley *Biochemistry* **35** (1996) 16110

31 P K Patel, N S Bhavesh and R V Hosur *Biochem Biophys Res Commun* **278** (2000a) 833

32 A Rangan, O Y Fedoroff and LH Hurley *Jour Biol Chem* **276** (2001) 4640

33 W Li, P Wu, T Ohmichi and N Sugimoto *FEBS Lett* **526** (2002) 77

34 G Fang and T R Cech *Cell* **74** (1993) 875

35 Z Liu, J D Frantz, W Gilbert and B K Tye *Proc Natl Acad Sci USA* **90** (1993) 3157

36 P Weisman-Shomer and M Fry *Jour Biol Chem* **268** (1993) 3306

37 R Giraldo and D Rhodes *EMBO Jour* **13** (1994) 2411

38 R Giraldo, M Suzuki, L Chapman and D Rhodes *Proc Natl Acad Sci USA* **91** (1994) 7658

39 N Baran, L Pucshansky, Y Marco, S Benjamin and H Manor *Nucleic Acids Res* **25** (1997) 297

40 R Erlitzki and M Fry *Jour Biol Chem* **272** (1997) 15881

41 H Sun, J K Karow, I D Hickson and N Maizels *Jour Biol Chem* **273** (1998) 27587

42 L A Dempsey, H Sun, L A Hanakahi and N Maizels *Jour Biol Chem* **274** (1999) 1066

43 M Fry and L A Loeb *Jour Biol Chem* **274** (1999) 12797

44 P McGlynn and R G Lloyd *Nucleic Acids Res* **27** (1999) 3049

45 G Sarig, P Weisman-Shomer, R Erlitzki and M Fry *Jour Biol Chem* **272** (1997) 4474

46 H Sun, R J Bennett and N Maizels *Nucleic Acids Res* **27** (1999) 1978

47 J E Eid and B Sollner-Webb *Mol Cell Biol* **15** (1995) 389

48 E Marsich, A Piccini, L E Xodo and G Manzini *Nucleic Acids Res* **24** (1996) 4029

49 E Marsich, L E Xodo and G Manzini *Eur Jour Biochem* **258** (1998) 93

50 D Huertas, L Fanti, S Pimpinelli, F X Marsellach, B Pina and F Azorin *EMBO Jour* **18**(1999) 3820

51 L Lacroix, H Lienard, E Labourier, M Djavaheri-Mergny, J Lacoste, H Leffers, J Tazi, C Helene and J L Mergny *Nucleic Acids Res* **28** (2000) 1564

52 H Arthanari and P H Bolton *Chem Biol* **8** (2001) 221

53 S Neidle and M A Read *Biopolymers* **56** (2001) 195

54 R H Shafer and I Smirnov *Biopolymers* **56** (2001) 209

55 A M Zahler, J R Williamson, T R Cech and D M Prescott *Nature* **350** (1991) 718

56 J L Mergny, P Mailliet, F Lavelle, J F Riou, A Laoui and C Hélène *Anticancer Drug Des* **14** (1999) 327

57 J F Riou, L Guittat, P Mailliet, A Laoui, E Renou, O Petitgenet, F Mégnin-Chanet, C Hélène and J L Mergny *Proc Natl Acad Sci USA* **99** (2002) 2672

58 H Arthanari and P H Bolton *Mini Rev Med Chem* **3** (2003) 1

59 M P Horvath and S C Schultz *Jour Mol Biol* **310** (2001) 367

60 R V Hosur *Curr Sci* **55** (1986) 597

61 H M Berman, J Westbrook, Z Feng, G Gilliland, T N Bhat, H Weissig, I N Shindyalov and P E Bourne *Nucleic Acids Res,* **28** (2000) 235

62 A Majumdar and R V Hosur *Jour Magn Reson* **88** (1990) 150

63 A Majumdar and R V Hosur *Prog NMR Spectrosc* **24** (1992) 109

64 A Majumdar and D J Patel *Acc Chem Res* **35** (2002) 1

65 Z Wu, A Ono, M Kainosho and A Bax *Jour Biomol NMR* **19** (2001) 361

66 M R Hansen, M Rance and A Pardi *Jour Am Chem Soc***120** (1998) 11210

67 H Zhou, A Vermeulen, F M Jucker and A Pardi *Biopolymers* **52** (2000)168

68 O Mauffret, G Tevanian and S Fermandjian *Jour Biomol NMR* **24** (2002) 317

69 K McAteer and M A Kennedy *Jour Biomol Struct Dyn* **20** (2003) 487

70 R A Kumar, R V Hosur and G Govil *Jour Biomol NMR* **1** (1991) 363

71 G Varani, F Aboul-ela and F H T Allain *Prog NMR Spectrosc* **29** (1996) 51

72 U Schmitz, F J H Blocker, T L James *Oxford Handbook of Nucleic Acid Structure* (Ed. S Neidle) Oxford University Press Oxford (1999) 253

73 M Gellert, M N Lipsett and D R Davies *Proc Natl Acad Sci (USA)* **48** (1962) 2013

74 W Guschlbauer, J F Chantot and D Thiele *Jour Biomol Struct Dyn* **8** (1990) 491

75 J R Williamson *Annu Rev Biophys Biomol Struct* **23** (1994) 703

76 R H Shafer *Prog Nucleic Acid Res Mol Biol* **59** (1998) 55

77 T Simonsson *Biol Chem* **382** (2001) 621

78 T Miura and G J Thomas Jr *Biochemistry* **34** (1995) 9645

79 P Schultze, N V Hud, F W Smith and J Feigon *Nucleic Acids Res* **27** (1999) 3018

80 D Miyoshi, A Nakao and N Sugimoto *FEBS Lett* **496** (2001) 128

81 D Miyoshi, A Nakao and N Sugimoto *Nucleic Acids Res* **31** (2003) 1156

82 M Meyer and J Suhnel *Jour Biomol Struct Dyn* **20** (2003) 507

83 C C Hardin, A G Perry and K White *Biopolymers* **56** (2001) 147

84 Y Wang and D J Patel *Structure* **2** (1994) 1141

85 S Haider, G N Parkinson and S Neidle *Jour Mol Biol* **320** (2002) 189

86 P K Patel and R V Hosur *Nucleic Acids Res* **27** (1999) 2457

87 J Feigon, K M Koshlap and F W Smith *Methods Enzymol* **261** (1995) 225

88 Y Wang and D J Patel *Structure* **1** (1993) 263

89 P K Patel, A S R Koti and R V Hosur *Nucleic Acids Res* **27** (1999) 3836

90 G N Parkinson, M P Lee and S Neidle *Nature* **417** (2002) 876

91 P K Patel and R V Hosur *Curr Sci* **74** (1998) 902

92 J Gu, J Leszczynski and M Bansal *Chem Phys Lett* **335** (2001) 465

93 M Meyer, C Schneider, M Brandl and J Sulhnel *Jour Phys Chem A* **105** (2001) 11560

94 K Kanaori, N Shibayama, K Gohda, K Tajima and K Makino *Nucleic Acids Res* **29** (2001) 831

95 P K Patel, N S Bhavesh and R V Hosur *Biochem Biophys Res Commun* **270** (2000b) 967

96 A Kettani, R A Kumar and D J Patel *Jour Mol Biol* **254** (1998) 639

97 S Bouaziz, A Kettani and D J Patel *J Mol Biol* **282** (1998) 637

98 A Kettani, S Bouaziz, A Gorin, H Zhao, R A Jones and D J Patel *Jour Mol Biol* **282** (1998) 619

99 N Zhang, A Gorin, A Majumdar, A Kettani, N Chernichenko, E Skripkin and D J Patel *Mol Biol* **312** (2001) 1073

100 K Gehring, J L Leroy and M Gueron *Nature* **363** (1993) 561

101 L Chen, L Cai, X Zhang and A Rich *Biochemistry* **33** (1994) 13540

102 J L Leroy, K Gehring, A Kettani and M Guéron *Biochemistry* **32** (1993) 6019

103 S Ahmed, A Kintanar and E Henderson *Nature Struct Biol* **1** (1994) 83

104 A T Phan and J L Leroy *Jour Biomol Struct Dyn* **S2** (2000) 245