Chapter 4

Structural Status of a G-rich Sequence of MYH7 Gene Promoter

SECTION-A
(i) Structural Status of 23-mer G-rich sequences

DNA primarily exists as a double helix. However, DNA can also adopt other structural conformations that have the potential to play critical roles in a range of biological processes. One such structures is G-quadruplex DNA, which was discovered in the late 1980s when biochemical experiments demonstrated that oligodeoxynucleotides that contain four separated runs of two, three or four guanines (G-tracts) can spontaneously form four-stranded structures (Sen et al, 1988; Sundquist et al, 1989). Nucleic acid sequences rich in guanines in particular are predisposed to form higher-order structures because of the capacity of guanine bases to self-associate via Hoogsteen bonds to form planar G-quartets. Within a single G-rich strand, several consecutive G-quartets can form and stack upon each other, leading to a four-stranded structure, called a G-quadruplex. Such structures are further stabilized by the presence of various monovalent cations, especially potassium (Williamson et al, 1989; Sen et al, 1990). Various G-rich DNA sequences which form G-quadruplex structures have been identified in biologically important regions of the genome. More recently, G-quadruplex-forming sequences have been identified in the gene promoters of c-myc (Siddiqui-Jain et al, 2002; Grand et al, 2002) c-kit (Fernando et al, 2006; Phan et al, 2007; Todd et al, 2007; Hsu et al, 2009; Kuryavyi et al, 2010), hTERT (Lim et al, 2010), VEGF (Sun et al, 2005; 2008; 2010; 2011), RET (Guo et al, 2007), PDGF (Qin et al, 2010) and HIF1-a (De Armond et al, 2005) and REF2 (Skoblov et al, 2006). Evidence of G-quadruplex formation in the regulatory sequences of muscle-specific genes (Yafe et al, 2005) and at the 5’- end of the retinoblastoma gene (Xu et al, 2006) has also been reported.
Putative quadruplex-forming sequences seem to be prevalent in eukaryotic genomes. Genome-wide bioinformatic searches have revealed a significant number (~376000) of putative quadruplex-forming sequences in the human genome (Huppert et al, 2005; Todd et al, 2005). In particular, the promoter regions of genes are significantly enriched in quadruplex forming sequences relative to the rest of the genome, with > 40% of human gene promoters containing one or more quadruplex motifs (Huppert et al, 2007). It has been suggested that the potential for DNA quadruplex formation is correlated with gene function (Eddy et al, 2006). Guanine rich regions are highly enriched in the whole transcriptional regulatory regions of genes active in the lung, heart and brain including cerebellum peduncles and caudate nucleus. The high G4 enrichment in these tissues may indicate that these tissues have evolved certain mechanisms to utilize G4s to regulate gene transcription (i.e. providing binding sites for tissue restricted TFs (transcription factors) (Zhou et al, 2013).

With this information in mind, we adopted biophysical and biochemical approach to study two 23-mer G-rich sequences from promoter location of Myosin β gene (MYH7) of Human and Rabbit. The G-rich sequence of Rabbit (RM23) differs from the Human 23-mer G-rich sequence (HM23) in three bases at position 8, 10 and 23. HM23 and RM23 DNA sequences used for the study with their complementary sequences HM23C and RM23C respectively.

- 5′–TGGGTGC{CGT}{G}{T}TGGGGGTGGGGG{T}-3′ (HM23)
- 5′–ACCCCCA{ACCCCCA}{ACGGGCACCCA}-3′ (HM23C)
- 5′–TGGGTG{CA}{G}{G}TGGGGGTGGGGG{G}C-3′ (RM23)
- 5′–GCCCCCACCCCCACCTGCACCCA-3′ (RM23C)

This section of chapter 4 describes the study on structural status of the abovesaid G-rich sequences with their counterparts. Under varied solution conditions, their structure, molecularity and stability were determined by non-denaturating gel electrophoresis, UV-thermal denaturation and circular dichroism (CD) studies. Prior to conducting the experiments, the purity of oligonucleotide sequences was checked by performing gel electrophoresis under denaturing conditions.
4.1 Polyacrylamide Gel Electrophoresis (PAGE)

4.1.1 Denaturating Gel Electrophoresis

The denaturating electrophoretogram of the sequences under study i.e. PAL24, HM23, HM23C, RM23 and RM23C run in lane 1-5, is shown in Fig. 4.1. PAL24 is a 24- nucleotide (nt) control size marker used under denaturating conditions to compare the mobility of oligonucleotide sequences under study. On comparing their mobility, it was found that sequences in lane 2-5 migrate according to their corresponding size. It was inferred from their electrophoretic mobility that HM23, HM23C, RM23 and RM23C move as single strands under denaturating conditions.

4.1.2 Non-Denaturating Gel Electrophoresis

Non-denaturating gel electrophoresis was used for elucidating the structural status and molecularity of the sequence. Through native gel, it is easy to describe the exact molecularity of the sequence and type of structure formed. This method has been extensively used to monitor and differentiate various multistranded structures. It allowed us to evaluate the coexistence of different folded structures based on their mobilities. The mobility of HM23, HM23+HM23C, RM23, RM23+RM23C in 1:1 ratio at 15µM strand concentration was carried out through 15% native PAGE in 20mM sodium cacodylate buffer (pH 7.0) containing 0.1 M NaCl and 0.1 mM EDTA is shown in Fig 4.2 (a). To predict the molecularity of gel bands, PAL24, M35 and M35+M35C were used as control size marker. PAL24 is fully palindromic 24-nt sequence, which under native condition form perfect duplex and move as a 48-nt oligomer (lane 2). M35 is a 35-nt single strand size marker which migrates as 35-nt single strand (lane 1). M35C is complementary strand of M35, equimolar mixture of M35 and M35C migrates as duplex (lane 7) under native condition and move as 70-nt oligomer. The oligomer (TG₃TGCGTG₂G₃TG₃T) HM23 (lane 3) showed two bands, the upper band is much retarded in comparison to M35+M35C duplex. Comparing the mobility of upper retarded band of HM23 with M35+M35C duplex (lane 7) it can be said that it could be of tetramer of HM23. Formation of G-wires or frayed wires is ruled out because the end of sequence is protected due to presence of thymine (T) at
the end. G-wire formation takes place when the formation of these structures appears
to depend on the sequences having terminal tracts of consecutive guanines.
Oligonucleotides with the general sequence d(GxNy) or d(NyGx) where x ≥ 10 and y ≥ 6, where N is any base (but not G or C), spontaneously form these higher order
structures (Macgregor et al, 2000). The lower band of HM23 is migrating slightly upper than M35 duplex it could be assigned to three stranded structure (trimeric) of
HM23. Being guanine rich the mobility of this band is slightly retarded than M35
duplex. Equimolar mixture of HM23 and HM23C showed two bands (lane 4),
reflecting the presence of two structures. The upper band migrating equivalent to
PAL24, thus is the perfect duplex formed by HM23+HM23C. The lower band is
assigned to the unstructured single stranded HM23C. This observation confirms that
HM23+HM23C correspond to the 23-mer intermolecular duplex. RM23 (lane 5)
exhibited two bands. On comparing the mobility of bands it was found that the upper
band migrating corresponding to the lower band of HM23 (lane 3) whereas the lower
band migrated with PAL24 (lane 2). On comparing the lower band with PAL24, it can
be said that it could be of dimeric species of RM23. Again a single band is observed
for RM23+RM23C (lane 6) which is migrated equivalent to PAL24, which confirms
duplex formation by RM23+RM23C. The lower band is assigned to unstructured
single stranded RM23C.

Keeping in mind that the presence of K⁺ stabilizes the G-quadruplexes, further
gel studies were carried out in the presence of K⁺. Fig. 4.2 (b) illustrates the mobility
pattern of the said oligonucleotides and their combinations containing 0.1 M KCl. We
observed three bands for HM23 (lane 3). When compared to the mobility of gel
markers and individual strand, the mobility of these bands were self-explanatory and
dictates the exact molecularity and possible structural status of species present
therein. The upper band migrates slightly above the M35+M35C duplex (lane 7)
corresponds to the tetrameric species of HM23. The lower band is migrating with the
PAL24, is assigned to the dimeric species of HM23. On comparing mobility of the
middle band of HM23 with control size markers, it was observed that it is migrating
with M35+M35C duplex which is 70-nt duplex. On considering the strand
stoichiometry, it was inferred that this structure involve three strands in structure
formation. It can be said that it could be of trimeric (three stranded) structure of HM23. A report by Mergny et al has demonstrated the formation of tri-G-quadruplex involving three strands with antiparallel topology. The G-quadruplex was constructed by using conventional Watson-Crick duplexes as guides for spatial orientation of G-tracts (Mergeny et al, 2012). Duplex of HM23 is prepared by mixing its counterpart HM23C in 1:1 ratio and its structural status was confirmed by non-denaturating gel. Lane 4 exhibited three distinct bands, the upper intense band migrating equivalent to the band observed for HM23 (lane 3) corresponds to the tetrameric species whereas the lower band is assigned to the unstructured single stranded HM23C. The middle band with mobility equivalent to PAL24 is the duplex formed by HM23+HM23C. RM23 showed two bands (lane 5), the lower band has mobility equivalent to the PAL24, thus is assigned to the dimer of RM23. On comparing the middle band of RM23 with size control marker M35 duplex it was observed that it migrates with duplex of M35+M35C. On consideration of strand stoichiometry the mobility of middle band signify that three strands get associated to form a trimeric (three stranded) structure. Lane 6 comprises duplex of RM23+RM23C. It exhibited three bands; the upper band migrates with duplex of M35 which is assigned to the trimeric species. Lower band is again of unstructured single stranded RM23C. The middle band having mobility equivalent to PAL24 is assigned to duplex of RM23+RM23C. It is important to mention here that HM23 differs from RM23 in only three bases at position 8, 10 and 23. However, the information regarding exact molecularity and structure adopted by said oligonucleotide sequences and their stability cannot be predicated on the basis of gel studies only. Thermal denaturation studies were carried out to evaluate the thermal stability of the oligomeric structures and the same is discussed in the following section.

4.1.3 UV-Thermal Denaturation Studies

Nucleic acid structures are sensitive to temperature. The melting temperature (Tm) of each oligonucleotide was determined by calculating the first derivative of its melting profile, representing the temperature at which 50% DNA dissociates into random coils. The thermal stability of the oligomeric structures form by G-rich sequences was
determined using UV thermal denaturation in the presence of Na\(^+\) and K\(^+\). G-quadruplex monitored at 265 nm (A\(_{265}\)) shows a sigmoidal (hyperchromic) curve, whereas monitoring the absorbance change with temperature at 295 nm (A\(_{295}\)) displays a characteristics inverse sigmoidal hypochromic curve (Mergny et al, 1998). The wavelength 295 was monitored as a function of increasing temperature. First to check the formation of self-associated structure of the above oligonucleotides, we carried out UV-thermal melting experiments of single stranded HM23 and RM23 sequences. The absorbance versus temperature profile of HM23 and RM23 sequences at 295 nm are displayed by Fig. 4.3, 4.4, 4.5 and 4.6 in 20 mM sodium cacodylate buffer (pH 7.0), 0.1 mM EDTA containing either 0.1 M NaCl or 0.1 M KCl respectively. An inverse sigmoidal curve obtained in UV melting experiments at 295 nm supports the presence of the G-quartet containing structures adopted by abovesaid G-rich sequences used in the present study. The melting profile of HM23 was found to be biphasic in presence of Na\(^+\)(Fig. 4.3 (a)). The biphasic transition corresponds to the two oligomeric structures in the solution. The calculated melting temperatures for HM23 (Fig. 4.3 (b)) were found to be 44 °C and 58 °C. Tm result is in good agreement with the gel studies. Two bands were observed for HM23 (lane 3, Fig. 4.2 (a)). The higher temperature transition 58 °C is assigned to the trimeric structure involving three strands for the structure formation and the lower temperature transition can be assigned to the uncoiling of the four stranded structure of HM23. It is known that DNA molecules give rise to steps in melting curves (Beers et al, 1967) corresponding to different regions of the molecule melting at different temperatures. The biphasic melting curve displayed by RM23 in Fig.4.4 (a) has two well defined transitions corresponding to the concurrent melting of two structures with distinct conformations. The first derivative plot of the biphasic melting profile shows two peaks depicting a Tm value of 45 °C for the low temperature transition and 62 °C corresponding to the upper temperature transition (Fig. 4.4 (b)). The transition with Tm value 45 °C can be assigned to the trimeric structure adopted by RM23 sequence and higher temperature transition can be assigned to the dimeric structure of RM23.

It is evident from various reports that cations play a substantial role in the stabilization of G-rich structures (Sannohelel et al, 2010). Especially K\(^+\) ion stabilize G-quadruplex structure more efficiently (Chanot and Guschlbauer, 1969). Keeping this
in mind, thermal denaturation experiments were performed in presence of K\(^+\) to see the effect of K\(^+\) cation on the stability of structure adopted by used oligonucleotide sequences. Fig. 4.5 displayed the UV-melting transition of HM23 in 20 mM sodium cacodylate buffer (pH 7.0), 0.1 M KCl containing 0.1 mM EDTA. Oligonucleotides with repeat/inverted repeat/homopurines/homopyrimidines subset of oligonucleotides sequences can generate various sequence-specific structures involving diverse base-association schemes, and are usually manifested in multiphasic thermal denaturation profiles. A triphasic UV-melting transition is observed for HM23 (Fig. 4.5 (a)) displayed well defined triphasic transition corresponding to concurrent melting of three different structures with distinct conformations. The first derivative plot of the melting profile shows three peaks depicting a Tm value of 44 °C for the lower temperature transition, 58 °C corresponds to middle temperature transition with 72 °C for the upper temperature transition (Fig. 4.5 (b)). Lower and middle melting temperature was also observed in presence of Na\(^+\) for HM23 where low temperature transition is assigned to the tetrameric species of HM23 and 58 °C was assign to the trimeric structure. One additional higher transition is displayed in presence of K\(^+\). From the gel studies it is clear that HM23 formed tetrameric, trimeric and dimeric structures in presence of K\(^+\). The low temperature transition is assigned to the tetrameric species and the middle melting temperature could be of trimeric species. The upper transition was assigned to the dimer of HM23. Fig. 4.6 displayed the UV-melting profile of RM23 sequence. Biphasic transition was obtained for the RM23 in presence of K\(^+\) which goes well with gel studies. In native gel two bands were observed for RM23 in presence of Na\(^+\) as well as in K\(^+\). Analysis of melting curve revealed that low temperature transition with Tm 55 °C is assigned to trimeric structure of RM23 and the higher temperature transition with Tm value 80 °C is exhibited by dimeric structure of RM23. HM23 and RM23 being guanine rich behave differently in presence of Na\(^+\) and K\(^+\). It is evident from gel and thermal melting studies that in presence of K\(^+\), G-quadruplexes get more stabilize than Na\(^+\). Although the structure appears to be stabilized by a hydrogen bonding network involving N7:N2H and O6:N1H. Indeed the central core of the G-quartet facilitate a specific
geometric arrangement of lone pair of electrons from the four G-O6, which can coordinate a monovalent ion of the correct size such as Na\(^+\) or K\(^+\). Generally G-quadruplex structures do not form in the absence of these ions. The smaller Na\(^+\) ion sits in the plane formed by these atoms, whereas the larger K\(^+\) ion requires a non-polar component which is present between two such G-quartets. This allows additional coordination of K\(^+\) ion to satisfy hexacoordinate stereochemistry (Lane et al, 2008). Literature is rich in reports of stability and structural polymorphism of G-quadruplexes provided by K\(^+\) (Luu et al, 2006; Dai et al, 2007; Zhang et al, 2010; Palacky et al, 2012). Results of thermal denaturation profiles described here well correlated with the gel studies of previous section.

### 4.1.4 Circular Dichroism Studies

The CD spectroscopy is a valuable tool for mapping conformational properties of particular DNA molecule and for elucidating the conformations adopted by various nucleic acid secondary structures (Kypr et al, 2009). It is a fast and qualitative method to elucidate the formation of G-quadruplex structures. It also provides information of structural differences in oligomeric structures in Na\(^+\)/K\(^+\) cations. CD signatures of different conformations of DNA are well known (Kypr et al, 2009). CD measurements on HM23 and RM23 were performed to have information on the structural type formed by the sequences in native gel electrophoresis experiment and in order to better understand the different conformations adopted by the sequences. Fig. 4.7 displays the CD spectra of HM23 in 20 mM sodium cacodylate buffer (pH 7.0), 0.1 mM EDTA containing either 0.1 M NaCl or 0.1 M KCl. Spectra are characterized by strong positive peak at 263 nm and a negative peak centred at 241 nm with an additional positive peak at 210 nm in presence of Na\(^+\) as well as in K\(^+\). These CD spectra demonstrated well-defined characteristics of parallel G-quadruplexes. It is well known that parallel G-quadruplexes in which four strands with all G-glycosyl bonds in *anti* conformation run parallel to each other have a dominant CD positive band around 260-265 nm and a negative peak at 240 nm. On
the contrary, the spectra of antiparallel G-quadruplex, where guanines alternate syn and anti glycosyl conformation along each strand have a negative band around 260-265 nm with a positive peak centered around 290-295 nm (Balagurumoorthy et al, 1992; Lu et al, 1992). Both quadruplex types display an additional characteristic positive peak at 210 nm (Kypr et al, 2009).

Gel studies and Tm studies have concluded that in presence of Na\(^+\), HM23 forms tetramer and trimer structures while in presence of K\(^+\) it adopts tetramer, trimer and dimer form of G-quadruplexes. Importantly, CD showed a single strong positive peak at 263 nm with a negative peak at 241 nm in presence of Na\(^+\) as well as in K\(^+\) ion, which indicates a parallel topology for all the three structures of varied molecularity. Further, Fig. 4.8 represented the CD spectra of RM23 in 20 mM sodium cacodylaye buffer (pH 7.0), 0.1 mM EDTA with either 0.1 M NaCl or 0.1 M KCl. The spectra are characterized by strong positive peak at 262 nm with a small hump at 296 nm and a negative peak centered at 239 nm. Additional positive peak was also observed at 210 nm, which confirms the formation of G-quadruplex. RM23 displayed CD spectra of both parallel and antiparallel quadruplexes. It can be said that the parallel form dominates over a small population of antiparallel quadruplex. The CD results thus obtained establish nice correlation with the gel electrophoretic and thermal denaturation studies. In the proposed model (Fig. 4.10) for HM23 in presence of Na\(^+\)/K\(^+\) ion with parallel topology it is clearly shown that there is a possibility of 10 G·G hydrogen bonding in parallel trimeric G-quadruplex. These hydrogen bonds provide extra stability to the unusual trimeric structure.

There are numerous reports in literature regarding possibilities of G.G hydrogen bonding (Li et al, 1991; Ke et al, 1993; Evertsz et al, 1994; Suda et al, 1995; Dolinnaya et al, 1997; Kaushik et al, 2003) stating that G.T, G.G and G.A were among the most stable mismatches. Purine-purine mis-matches were generally more stable then pyrimidine-pyrimidine mispairs. The possible mispairing assumed is shown in Fig. 4.9. Trimeric structure of RM23 also possesses G.G hydrogen bonds. All the possible structures adopted by HM23 and RM23 in Na\(^+\) as well as in K\(^+\) are shown in Fig. 4.10 and Fig. 4.11.
Section (A)

(ii) **Effect of Mutation on Structural status of 23-mer G-rich (HM23) sequence**

As discussed in earlier section 4A (i) formation of an unusual three-stranded G-quadruplex structure by HM23 sequence was observed. Further to confirm that the involvement of G5 stretch is essential here to the formation of trimeric structure, we carried out a control experiment, where HM23 was mutated by replacing G by A at positions 3, 14 and 20 (TGAGTGCCGTTGGAGGTGGAGGT) [HM23Mut]. Oligonucleotide sequences used for the study are:

- 5'-TGGGTGCCGTTGGGGGTGGGGGT-3' (HM23)
- 5'-ACCCCCACCCCCAACCACCGCACCACCA-3' (HM23C)
- 5'-TGAGTGCAAGTGGAGGTGGAGGT-3' (HM23Mut)
- 5'-ACCTCCACCTCCACCTGCACCTCA-3' (HM23MutC)

### 4.2 Polyacrylamide Gel Electrophoresis

#### 4.2.1 Non-Denaturating Gel Electrophoresis

A 15% polyacrylamide native gel electrophoretic pattern in 20mM sodium cacodylate buffer (pH 7.0) containing 0.1 M NaCl and 0.1 mM EDTA is shown in **Fig. 4.12 (a)**. To predict the molecularity of structural species, four control size markers PAL12 (12-mer duplex, lane 1, lower band), PAL24 (24-mer duplex, lane 1, upper band), M35 (35-mer single strand, lane 8) and M35+M35C (35-mer duplex, lane 9) are used in gel assays. Their status has already been discussed in previous section. It was found that in presence of Na\(^+\), HM23 (lane 2) showed the same status as discussed in earlier section (**Fig. 4.2 (a)**, lane 3). Lane 3 comprises complementary strand of HM23 which showed a single band migrated with PAL12 (lane 1) which assign to the unstructured C-rich strand. Equimolar mixture of HM23+HM23C displayed an intense band migrated with PAL24 (24-mer duplex) confirms the formation of perfect duplex (lane 4) with a faint lower band of excess unstructured C-rich strand. The oligonucleotide HM23Mut showed two bands in Na\(^+\) ion (lane 5)
indicating the presence of two structural species in the solution. On comparing the mobility of upper band of HM23Mut (lane 5) with size markers, it is observed that its mobility is retarded than M35+M35C (35-mer duplex) (lane 1) which can be assigned to the tetrameric species of HM23Mut. Lower band of HM23Mut is migrated equivalent to the duplex of PAL12 is reflecting formation of folded compact structure, which could be of unimolecular species. A single band is observed for the complementary strand of HM23MutC (lane 6) which can be assigned to the unstructured C-rich strand. Lane 7 depicts a band from the equimolar ratio of HM23Mut and HM23MutC. It showed a single band migrating with PAL24 (24-mer duplex) further confirming the duplex formation.

After examining the structural status of HM23Mut in Na\(^+\), the same was further examined in K\(^+\) ion. Native gel electrophoresis pattern of HM23, HM23C, HM23Mut and HM23MutC in 0.1 M KCl is shown in Fig. 4.12 (b). HM23 (lane 2) and HM23+HM23C (lane 4) migrated in similar pattern as discussed earlier (Fig. 4.2 (b)), lane 3 & lane 4). HM23C (lane 3) also migrated as unstructured single strand. Interestingly, HM23Mut (lane 5) showed two bands migrating equivalent to PAL24 and PAL12. The upper band is migrating with PAL24 (24-mer duplex) displayed the formation of bimolecular species whereas the lower band is migrated with 12-mer duplex of PAL12 again reflecting the presence of compact folded structure in solution. A single band is observed for HM23Mut C (lane 6) sequence confirms the unstructured status of this oligonucleotide at neutral pH. Equimolar mixture of HM23Mut+HM23MutC (lane 7) showed a single band migrating with PAL24 reflecting the formation of perfect duplex by HM23Mut+HM23MutC. It is evident from the gel results that after introducing mutation in HM23 sequence trimeric structure formation is inhibited in Na\(^+\) as well as in K\(^+\) cation. Thus, gel results have confirmed the involvement of G5 stretch in the formation of three stranded G-quadruplex structure (trimeric) adopted by HM23 oligonucleotide.

### 4.2.2 UV-Thermal Denaturation Studies

After examining the structural status of HM23Mut under gel experiments, its thermal stability was investigated. Fig. 4.13, displayed the melting profile of HM23Mut in 20 mM sodium cacodylate buffer (pH 7.0) containing 0.1 M NaCl and
0.1 mM EDTA. An inverted melting profile is observed for HM23Mut at 295 nm confirming the G-quadruplex formation by the sequence. The biphasic melting profile is obtained for HM23Mut with melting temperatures 36°C and 67°C, which suggests unfolding of two structural species in solution. Earlier, gel assays of HM23Mut have shown two bands, corresponding to a tetramer and unimolecular species. The low temperature transition with 36 °C is attributed to the unimolecular species while the high temperature transition with Tm 67 °C can be assigned to the disordering of the tetramer.

Tm studies were also performed with the aim of predicting the stabilities of G-quadruplex structure formed by HM23Mut in K+ ion. The denaturation profile of HM23Mut again showed a biphasic transition in K+ ion (Fig 4.13). Tm values calculated by first derivative revealed the melting of two structural species. Since, the gel studies showed two bands of HM23Mut, reflecting two structural species in solution. Therefore, lower Tm 39 °C could be of unimolecular species whereas the higher temperature transition with Tm 76 °C can be assigned to the bimolecular species. Unimolecular species formation takes place in presence of Na+ as well as in K+ having approximately same Tm.

### 4.2.3 Circular Dichroism Studies

Circular dichroism (CD) spectroscopy method has been particularly proved to be informative in elucidating the conformation of nucleic acids secondary structures. A CD spectrum of HM23Mut in 20 mM sodium cacodylate buffer (pH 7.0) containing either 0.1 M NaCl or 0.1 M KCl and 0.1 mM EDTA is shown in Fig. 4.14. The CD profile of HM23Mut in presence of Na+ as well as in K+ exhibited two positive peaks at 258 nm and 278 nm with a negative peak at 236 nm. Based on existing knowledge from the literature, these peak positions have been interpreted in terms of parallel and antiparallel G-quadruplex structures (Balagurumoorthy et al, 1992). Gel assay of HM23Mut have also revealed formation of two structural species. Correlated with gel and Tm studies the positive peak at 258 nm can be assigned to the parallel tetrameric G-quadruplex and 278 nm to the antiparallel unimolecular G-
quadruplex in Na\(^+\) whereas 258 nm peak was depicted parallel unimolecular G-quadruplex structure in K\(^+\) and 278 nm peak is assign to the bimolecular antiparallel G-quadruplex species. This collective analysis of the gel, Tm and CD studies established a nice correlation and made us to propose structural models in terms of strand stoichiometry for the G-quadruplexs of varied molecularity. These proposed models are displayed in Fig. 4.15. According to the proposed parallel tetramer model in Na\(^+\) the higher Tm can be assigned to this structure due to the presence of more G-stacks than the unimolecular species which has only two G-quartets. Since, it is known that K\(^+\) promotes and stabilize parallel G-quadruplexes. The proposed model for HM23Mut in K\(^+\) shows that unimolecular structure belongs to parallel topology while the bimolecular assembly attains antiparallel conformation. Presence of four G-quartets in bimolecular structure displayed higher Tm value than unimolecular species consisting of only two G-quartets. It is in accordance with the report that stability of the G-quadruplex increases with increasing number of G-stacks (Rachwal et al, 2007).
Section (A)

(iii)  **Structural status of 23-mer G-rich sequences with flanking bases**

It is believed that G-Quadruplexes motif can not exist independently *invivo*, but are embedded in the chromosome and are linked by flanking sequence in genomic system. It is imperative to analyze the effect of these neighboring sequences on quadruplex stability. It has been widely documented that changes of only one base in the sequence at 3'- or 5'- end can drastically modify quadruplex topology (Ambrus et al, 2006). So, keeping this in mind, effect of flanking bases on the structure of 23-mer G-rich sequences HM23 and RM23 was investigated by extending the HM23 and RM23 to 34-mer sequence naturally flanked by six nt at 5'- side and five nt at 3'- side as follows:

- 5’–AGGGACTGGTGCCGTTGGGGGTGGGGGTGCCCT-3’ (HM34)
- 5’–AGGGACCCCCACCCCCAACGGACACCCGTCCCT-3’ (HM34C)
- 5’–AGGGACTGGGTGACCCAGGTGGGGGTGCCCT-3’ (RM34)
- 5’–AGGGCGCCCCACCCCCACCTGCACCCAGTCCCT-3’ (RM34C)

### 4.3 Polyacrylamide Gel Electrophoresis

#### 4.3.1 Non-Denaturating Gel Electrophoresis

In order to see the effect of flanking bases on the structural status of HM23 and RM23, gel electrophoretic studies were carried out on the 5'- and 3'- extended versions of these oligonucleotides, namely HM34 and RM34 having core sequences HM23 and RM23 respectively; HM34C and RM34C being their complementary sequences respectively. The electrophoretogram in Fig. 4.16 (a) presents 15% native PAGE pattern of HM34, HM34+HM34C, RM34, RM34+RM34C, at 1:1 ratio in 20 mM sodium cacodylate buffer (pH 7.0), 0.1 mM EDTA containing 0.1 M Na⁺. To predict the structural species formed by said oligonucleotides, PAL12 (12-mer duplex, lane 1, lower band), PAL24 (24-mer duplex, lane 2), and M35+M35C (35-mer duplex, lane 7) were used as control size markers in gel assays. In presence of Na⁺,
HM23 sequence showed two bands discussed in earlier section [Section A (i) Fig. 4.2 (a) lane 3]. A single band is observed for the extended version of HM23 sequence i.e. HM34 (lane 3) reflecting the presence of a single structural species in the solution. HM34, the 34-mer sequence migrated equivalent to the duplex of PAL12 (12-mer duplex) indicating the migration of HM34 as folded species. On the basis of mobility, it seems conceivable to assume that the gel band corresponds to the intramolecular G-quadruplex. Lane 4 shows the electrophoretic mobility pattern of 1:1 mixture of intermolecular duplex (HM34+HM34C). It exhibited a single band which migrates with the duplex of M35+M35C (lane 7) which shows the perfect duplex formation by HM34+HM34C. Interestingly, RM34 sequence, extended version of RM23 showed three bands in presence of Na\(^+\) ion (lane 5). It is important to mention here that core sequence RM23 differs in only three bases from HM23 sequence. Flanking bases are same in both HM34 as well as in RM34 sequence. The mobility of the upper band of RM34 is much retarded than the mobility of M35+M35C duplex (70-nt). The most probable structure can be an oligomeric tetramer. The middle band migrated equivalent to duplex of M35+M35C correspond to a structure comprising of two strands, which can be assign to the bimolecular species formed by RM34. Third and lowermost band is migrated equivalent to PAL12 (12-mer duplex, lane 1) indicating migration of folded compact structure. This indicates the formation of unimolecular species. Equimolar mixture of RM34 and RM34C displayed two bands (lane 6) there seems a fair possibility of incomplete hybridization and generation of self-associated structure by individual strand. The upper band can be assign to the tetrameric species formed by RM34. Lower band is migrated with the duplex of M35+M35C (lane 7), which could be of duplex formed by RM34+RM34C.

Further in order to see the effect of cations on the structure adopted by these extended sequences, native gel was performed in presence of K\(^+\) ion. Fig. 4.16 (b) shows the 15% native PAGE pattern of HM34, HM34+HM34C, RM34 and RM34+RM34C in 20 mM sodium cacodylate (pH 7.0), 0.1 mM EDTA containing 0.1 M K\(^+\). On analysis, it was found that HM34 (lane 3) migrated as single band with mobility slightly retarded than M35+M35C (lane 7) duplex. This indicates the possibility of a bimolecular structure. Being G-rich its mobility was slightly retarded
than M35 duplex (lane 7). Lane 4 comprises the equimolar mixture of HM34+HM34C. A single band was observed migrating with duplex of M35+M35C, which confirms the formation of perfect duplex by HM34+HM34C. Interestingly, RM34 (lane 5) exhibited two bands with mobility of upper band retarded than M35 duplex and lower band migrating with duplex of M35. Upper slow moving band was spotted to be less intense than the lower band indicating the presence of two structural species of different molecularity and proportion. This clearly shows that the upper band is due to the formation of tetramers as minor species and bimolecular structure as major species. RM34+RM34C (lane 6) form a duplex (lower intense band) which migrates with duplex of M35+M35C with a small population of tetrameric species (upper band). It is interesting to note that HM23 which is the core sequence of extended version HM34 adopts tetrameric and trimeric species in Na$^+$ and tetrameric, trimeric and dimeric structure in presence of K$^+$ while its extended version HM34 showed a single band in presence of both the cations with varied molecularity (unimolecular in Na$^+$ and bimolecular in K$^+$ cation). Surprisingly, RM23, the core sequence of RM34 forms trimeric and dimeric species in Na$^+$ as well as in K$^+$ ion while its extended version RM34 adopts three distinct structures with varied molecularity (tetramolecular, bimolecular and unimolecular). After assigning molecularity of said oligonucleotides, thermal denaturation experiments were performed to investigate the stability of the structures adopted by oligonucleotides. The same is discussed in the following section.

### 4.3.2 UV-Thermal Denaturation Studies

Information on the thermal stability is also an important aspect in elucidating the possibility of various proposed DNA structures. When the UV-absorbance is monitored at 295nm, thermal denaturation experiments allow determination of Tm of the G-quadruplexes. The thermal stability of structures formed by the oligonucleotide sequences under study was assessed by thermal denaturation.

**Fig. 4.17** shows the thermal denaturation profile of HM34 in the presence of 20 mM sodium cacodylate (pH 7.0), 0.1 mM EDTA containing 0.1 M NaCl. An inverted sigmoidal curve is obtained for HM34 confirms the formation of G-
quadruplex. Analysis of this melting curve revealed that HM34 displayed a broad monophasic transition with a Tm of 56 °C. This is in good agreement with the gel studies where a single band is exhibited by HM34 (Fig. 4.15 (a) lane 3). Thermal melting profile of RM34 is shown in Fig. 4.18. The melting curve of RM34 was found to be biphasic in presence of Na\(^+\) which reflects the presence of two structural species. Gel results have already demonstrated that RM34 can form three structures of varied molecularity. The melting temperatures were found to be 54 °C and 75 °C. Since the lower temperature transition of biphasic curve is broad it may reflect the melting of two species with small difference in Tm values resulting in overlapping of two transitions, giving a broad melting profile. The upper temperature transition with Tm 75 °C could be assigned to the unimolecular species of RM34. It is well known fact that intramolecular G-quadruplexes are more stable than intermolecular species because formation and folding kinetics of intramolecular assembly is very complex (Lane et al, 2008). So, it is easy to deduce the higher temperature transition to the unfolding of unimolecular species.

Cations play major role in stabilization of G-quadruplexes. Based on the reports that K\(^+\) can not only induce G-quadruplex structures but also stabilize G-quadruplexes (Scaria et al, 1992; Rachwal et al, 2007; Smargiasso et al, 2008; Bugaut et al, 2008), thermal denaturation of HM34 and RM34 were also performed in presence of 0.1 M KCl. Apparently broad biphasic melting profile was obtained for HM34 (Fig. 4.19 (a)) indicating the existence of two species, each showing distinct quadruplex melting transitions. To our surprise gel studies showed a single band in presence of K\(^+\) for HM34 (Fig. 4.15 (b), lane 3) which is assigned to the bimolecular species. Biphasic melting profile here reflects the possibility of formation of two structural species with identical molecularity. A lower temperature transition with a Tm value of 32 °C is followed by a higher temperature transition with a Tm 61 °C (Fig. 4.19 (b)). The lower transition can be assigned to the quadruplex species whereas the higher melting transition with Tm 61 °C could be of dimeric quadruplex species. This interpretation is in agreement with the literature that K\(^+\) ion promotes and stabilizes parallel G-quadruplexes. However, at this stage topology of the said G-quadruplexe structures can not be assigned.Tm experiments in presence of K\(^+\) also
were performed with RM34 which reflected a broad monophasic transition (Fig. 4.20 (a)). It was clear from the gel results that RM34 showed two bands, assigned to be tetrameric and bimolecular G-quadruplex (Fig. 4.15 (b), lane 5)). So, again a broad monophasic transition of RM34 might be due to overlap of two transitions, coming from two structural species. The calculated melting temperature is 53 °C (Fig. 4.20(b)).

Further, to characterize the secondary structure of HM34 as well as RM34 sequences and information on G-quadruplex topologies, circular dichroism (CD) spectroscopic studies were performed.

4.3.3 Circular Dichroism Studies

Based on the results obtained by UV-thermal denaturation and gel studies, we performed CD studies of HM34 and RM34 in the presence of Na\(^+\) and K\(^+\) to see the topology of G-quadruplexes formed in solution. Fig. 4.21 displays the CD spectra of HM34 in 20 mM sodium cacodylate buffer (pH 7.0), 0.1 mM EDTA containing either 0.1 M NaCl or 0.1 M KCl. HM34 in presence of Na\(^+\) is characterized by two positive peaks at 258 nm and 282 nm with a negative peak centred at 235 nm whereas in presence of K\(^+\) a remarkable shift in the positive peaks towards longer wavelength was observed. Hence, CD spectrum of HM34 in K\(^+\) exhibits two positive peaks at 264 nm and 287 nm with negative band at 235 nm. When compared with CD experiment in presence of Na\(^+\), it is evident from CD results that parallel quadruplex gets stabilized in presence of K\(^+\) ion. CD results are in good correlation with gel and thermal denaturation studies and conclude that HM34 adopts two unimolecular species with varied topology.

Fig. 4.22 displays CD spectra of RM34 in 20 mM sodium cacodylate (pH 7.0), 0.1 mM EDTA containing either 0.1 M NaCl or 0.1 M KCl. CD spectrum of RM34 in Na\(^+\), showed two positive peaks at 257 nm and 287 nm with a negative peak at 235 nm. Presence of K\(^+\) brought a red shift with increase in ellipticities at 270 nm and 287 nm. Thus, it is clear that presence of Na\(^+\), parallel quadruplex is the major species with a small population of antiparallel G-quadruplex as minor, whereas in K\(^+\)
ellipticity at 270 nm is less than 287 nm which shows that antiparallel being the major species with parallel quadruplex as minor species. To sum up, CD studies with gel and Tm studies it is evident that RM34 adopts parallel and antiparallel G-quadruplexes of varied molecularity. On the basis of gel, Tm and CD studies all the possible models of G-quadruplexes proposed for HM34 and RM34 under different ionic conditions are represented in Fig. 4.24 and Fig. 4.25. In the proposed model for HM34 in presence K\(^+\) ion (Fig. 4.24 (b)) with antiparallel and parallel topology it is clearly shown that there is a possibility of 8 G·G bonding in parallel dimer while four G·T mismatches were observed in case of antiparallel dimer. These hydrogen bonds provide extra stability to the structure. There are numerous reports in literature regarding possibilities of G.T and G.G hydrogen bonding already discussed in previous section of this chapter (Section A(i)) stating that G.T, G.G and G.A were among the most stable mismatches. Purine-purine mis-matches were generally more stable than pyrimidine-pyrimidine mismairs. The various possible mispairing assumed were shown in Fig. 4.23. The proposed models of tetramer, dimer and unimolecular RM34 with varied conformation are represented in Fig. 4.25.
**SECTION (B)**

*Effect of Varied G-tract and intervening Thymine (T) Residues on the G-quadruplex Formation*

DNA primarily exists as a double helix. However, depending on solution conditions it can also adopt various alternative structures that have the potential to play critical roles in a variety of biological processes. The utmost importance is given to guanine-rich regions which can adopt a non-canonical four stranded topology, known as G-quadruplex (Patel et al, 2007). In most eukaryotes, telomeric DNA consists of a short motif that includes consecutive guanines and may hence fold into G-quadruplexes. Besides telomeres, the G-quadruplex forming sequences also occur in other biologically relevant regions of the genome including the transcriptional regulatory regions and promoter regions (Bochman, et al, 2012). Several reports have shown that there is an interesting interplay between the length of G-tracts and loop form by the intervening bases in various types of quadruplexes. In general, the stability of the folded quadruplex increases with the length of the guanine runs. However, Rachwal et al, have shown an exceptional behaviour for the sequence d-(G₃T)₄. It was suggested that it is more stable than the corresponding d-(G₇T)₄. The literature is rich in exploring the effect of G-runs and bases present in loops on conformation of G-quadruplexes.

With this information sequences with varied G-tract length containing G5 stacks with T’s as intervening bases have been designed. Using Polyacrylamide Gel Electrophoresis (PAGE), UV-Thermal denaturation and Circular Dichroism, we investigated the stability, secondary structures and stoichiometry of different structural species formed by varying the number of G-tracts and T bases in loop. G-rich oligonucleotide sequences 13GT1, 14GT2, 15GT3, 16GT4, 25GT1, 28GT2, 31GT3 and 34GT4 were used for the study with their counterparts 13GT1C, 14GT2C, 15GT3C, 16GT4C, 25GT1C, 28GT2C, 31GT3C and 34GT4C respectively.
4.4 Polyacrylamide Gel Electrophoresis

4.4.1 Non-Denaturating Gel Electrophoresis

Oligonucleotide structure in the presence of sodium

Gel electrophoresis has been extensively used to monitor and differentiate various multistranded structures. It allowed us to evaluate the coexistence of different folded structures based on their mobilities. For information of various structures adopted by above mentioned G-rich sequences, non-denaturating gel electrophoresis was performed. Fig. 4.26 (a) depicts the gel mobility of G-rich sequences incubated in 20mM sodium cacodylate buffer (pH 7.0) containing 0.1 M NaCl and 0.1 mM EDTA. To predict the molecularity of structures PAI24 is used as control size marker. 13GT1 (13-mer sequence) show a single band, was displayed by (lane 2) which migrated almost equal to PAL24 reflecting the formation of a structure by 13GT1. The exact
The status of 13GT1 strands in terms of molecularity, when compared with that of the control oligomer band, on the basis its gel retardation was assumed to be tetrameric. Lane 3 comprises mixture of equimolar ratio of 13GT1 and 13GT1C. It showed three distinct bands. The upper retarded band again moves corresponding to 13GT1 band and can be assigned to the tetrameric species of 13GT1. The middle band moved faster than the marker PAL24 (24-mer duplex), corresponding to the duplex. The third band visible in the lane 3 corresponds to the unstructured C-rich strand (13GT1C). The 25-mer sequence, 25GT1 is the longer version of 13GT1 (Fig 4.26 (a), lane 4) exhibited a single band and migrated with the tetrameric species of 13GT1 sequence and slightly faster than PAL24. The electrophoretic mobility of 25GT1 (d-TG₅TG₅TG₅TG₅) indicates the formation of some folded compact structure. The exact status of 25GT1 in terms of molecularity, when compared with the control size marker PAL24, on the basis of extent of retardation, was assumed to be dimeric. To assume the compactness of the structure formed by 25GT1 we propose a dimeric structure can be formed by the involvement of two 25GT1 strands or it can be formed by two monomers placed on one above the other. A report by Quang (Quang et al, 2011) also proposed later kind of structure in their NMR study. Lane 5 comprises 25GT1 with its complementary strand in 1:1 ratio. The same shows three bands, where the migration of upper band correspond to PAL24 (24-mer duplex, lane 1), can be assigned to the duplex of 25GT1+25GT1C, the middle band can be assigned to the bimolecular species of 25GT1, while the lowermost band is due to the unreacted C-rich strand. Lanes 6 & 7 of the same gel contains 14GT2 (14-mer G-rich sequence). The status of this sequence alone and with its counterpart (14GT2C) was found identical to the status of 13GT1 and its counterpart. In lane 8, the longer stretch of 14GT2 sequence namely 28GT2, exhibited a single intense band which migrated faster in gel reflecting the formation of a compact folded unimolecular species. Lane 9 displayed three bands for the equimolar mixture of 28GT2+28GT2C. The upper intense band is migrated slightly retarded than PAL24 which can be assigned to the duplex (28GT2+28GT2C), and the lowermost band shows a small population of unimolecular species. The middle band reflects of unreacted C-rich strand.
**Oligonucleotide structure in the presence of potassium**

It is been known since long that Cations play a pivotal role in the structure formed by G-rich sequences (Sen and Gilbert, 1988). Keeping this in mind the said oligos and their combinations were studied by PAGE in K\(^+\) solution. 15% native gel electrophoresis was performed in 20 mM sodium cacodylate containing 0.1 M KCl and 0.1 mM EDTA. Fig. 4.26 (b) displayed the mobility pattern of the oligonucleotides. Lane 1 shows PAL24, 24-base pair control size marker which exhibited a single band. 13GT1 (lane 2), showed a single band reflecting the presence of single structural species which can be compared to the mobility of PAL24 (lane 1) and be assigned to a tetrameric structure. Lane 3 comprises of equimolar mixture of 13GT1+13GT1C sequences and their duplex status was found as similar was observed in Na\(^+\). However, it is noteworthy that the status of 25GT1 (lane 4), comes completely different in presence of K\(^+\). In presence of Na\(^+\) ion 25GT1 exhibited a single band corresponding to the tetrameric species of 13GT1, it migrates much faster than the PAL24 (24-mer duplex control size marker) in K\(^+\) solution. This fast migration clearly indicates the formation of unimolecular structure by 25GT1. The lane 5 showed three bands for 25GT1 and the mixture of 25GT1 and its complementary strand. The uppermost band migrating with PAL24 is assign to the duplex of 25GT1+25GT1C; the middle band assigned to the unstructured C-rich strand and lowermost fast migrating band is assigned to the folded compact unimolecular species. 14GT2 (lane 6), 14GT2+14GT2C (lane 7) showed the same results as in Na\(^+\), while the longer version 28GT2 (lane 8) exhibited a single intense band migrated below the 24-mer duplex (PAL24), and can be a folded compact structure. The fast migration of 28GT2 in gel reflects the formation of unimolecular structure. Lane 9 consists of 28GT2 with its complementary strand and displayed three intense bands. The upper less intense band migrating correspond to PAL 24, is assigned to the duplex of 28GT2, the middle band is assigned to the unstructured C-rich strand and the lowermost intense band is of unimolecular structure.

Fig. 4.27 (a) displayed the electrophoretic pattern of 15GT3, 16GT4, 31GT3 and 34GT4 with their counterparts 15GT3C, 16GT4C, 31GT3C and 34GT4C respectively. 15% native gel electrophoresis is performed in 20 mM sodium cacodylate containing 0.1 M KCl and 0.1 mM EDTA.
cacodylate containing 0.1 M NaCl and 0.1 mM EDTA. 35-mer duplex (M35+M35C) was used as control size marker for comparing the mobility of structural species adopted by the oligonucleotides. 15GT3 (lane 2) showed an intense band migrating faster than duplex of M35. On comparing the mobility of the band with control marker it can be assigned to a tetramer structure of 15GT3. Lane 3 comprises equimolar mixture of 15GT3+15GT3C and showed three bands. The upper band is shown to migrate with tetramer of 15GT3, the middle band could be of the intermolecular duplex formed by 15GT3+15GT3C. Lowermost band can be assigned to the unreacted C-rich single strand. The extended version of 15GT3 named 31GT3 exhibited a single band (lane 4). This fast moving band of 31GT3, reflects to the formation of a folded structure. On the basis of mobility, it seems conceivable to assume that the gel band corresponds to the intramolecular G-quadruplex. Lane 5 exhibited duplex of 31GT3 giving one intense band. On comparing mobility of the band with control marker M35+M35C duplex it is visible that band is migrated equivalent to M35 duplex. 16GT4 (lane 6) showed two bands, the upper band is migrated equivalent to the duplex of M35 size marker which could be of tetramer of 16GT4. The lower faint band showed mobility equivalent to unimolecular species of 31GT3 which reflects the formation of bimolecular species of 16GT4. Equimolar mixture of 16GT4 with its counterpart 16GT4C showed three distinct bands (lane 7). The upper band can be of tetrameric species, middle band reflects the formation of intermolecular duplex of 16GT4+16GT4C with the lower band can be of unreacted C-rich strand. The extended version of 16GT4 the 34GT4 (lane 8), displayed a fast moving single band. On comparing the mobility it was clear that it could be of folded compact intramolecular structure formed by 34GT4. Lane 9 comprises duplex of 34GT4, showing two bands. Intensity of upper band is more in comparison to the lower band, shows upper species as major one with lower as the minor one. On comparing the mobility of upper band with M35 duplex it was observed that its mobility is equivalent to the M35 duplex which confirms perfect duplex formation by 34GT4+34GT4C.

Further, to see the effect of K⁺ on the G-structures, gel studies is done in presence of KCl. Fig. 4.27 (b) displayed the 15% gel electrophoretogram in 20 mM
sodium cacodylate buffer (pH 7.0) containing 0.1 M KCl and 0.1 mM EDTA. 35-mer duplex is used as control size marker to compare the mobility of all the structural species formed by said oligonucleotides. Lane 2 consists of single strand of 15GT3, which showed two bands, the upper band is migrating slightly below than M35 duplex, which can be assigned to the tetramer of 15GT3. The lower band is assigned to the folded compact species. Lane 3 comprises duplex of 15GT3 with its counterpart 15GT3C, three bands were observed in this lane, the upper band is again of tetramer of 15GT3 middle band is migrating below than M35 duplex can be assigned to duplex formed by 15GT3 and its counterpart. The lowermost band is designated the excess C-strand. The extended version of this sequence i.e. 31GT3 (lane 4) exhibited two bands the upper band is moving a little bit lower than the duplex of M35 which might be of bimolecular species involving two strands of 31GT3, the lower fast migrating band is assigned to the unimolecular structural species. Lane 5 exhibited three bands for equimolar mixture of 31GT3+31GT3C, in which the upper band corresponds to the duplex of 31GT3 and its counterpart, the middle one assign to bimolecular species and the lowermost band is again of the unimolecular species formed by the single G-rich strand. The 16GT4 lane 6) sequence surprisingly exhibited four bands, the upper most band is migrating above the M35 duplex, which could not be of tetramer. G-rich DNA sequences may also form higher order structure like G-wires (Henderson et al, 1994) and DNA frayed wires, however, the formation of these structures appears to depend on the sequences having terminal tracts of consecutive guanines. Oligonucleotides with the general sequence d(GxNy) or d(NyGx) where x≥ 10 and y≥ 6, where N is any base (but not Gor C), spontaneously form these higher order structures (Macgregor et al, 2000). Our 16GT4, d(TG5T4G5T) does not have ability to form G-wires as its end are protected by T. Hence, the formation of G-wires is ruled out. A report by Gabelica et al (Gabelica et al, 2008) has proposed a model which readily accounts for the preferential multimer formation for by parallel G-quadruplexes since there is no loop impediment to stacking. They showed the step of formation of multimer starts with the formation of parallel intramolecular structures and then the assembly into multimers. So that, we proposed a dimer of tetramer like structure for the uppermost band displayed by 16GT4.Two tetramer may stack on one another. The second band of 16GT4 is migrating with the duplex of M35+M35C,
which is supposed to be the tetrameric species of 16GT4, below this the third band is migrating with the 31GT3 unimolecular species, based on its mobility it can be assigned to the dimeric species. The lowest band corresponds to unreacted G rich sequence. The 16GT4 with its complementary strand comprises five bands (lane 7), the other four structures are also exhibited by the 16GT4 single strand, and the fifth band is of duplex of 16GT4 + 16GT4C. The extended version of 16GT4, sequence 34GT4 (lane 8) consists of two bands, the upper band is moving slightly below than M35+M35C, it reflects the formation of bimolecular species. The second band is moving faster which can be assigned to a very compact intramolecular species; due to its compactness it moves faster in gel. Lane 9 exhibited three bands for equimolar mixture of 31GT3+31GT3C, in which the upper band corresponds to the duplex of 31GT3 and its counterpart, the middle one assign to bimolecular species and the lowermost band is again of the unimolecular species formed by the single G-rich strand.

So, it is clear from the gel studies that all the G-rich sequences adopt higher order structures which is also exhibited even in the presence of their complementary strands. The information regarding exact molecularity and topology of structures adopted by said oligonucleotide sequences and their stability can not be predicated on the basis of gel studies only. Further, UV- thermal denaturation and CD studies were carried out and are discussed in the following sections.

### 4.4.2 UV-Thermal Denaturation Studies

Nucleic acid structures are sensitive to temperature. Hence, UV-melting profiles can be used to determine the melting temperature (Tm) of cation- induced G-quadruplexes and temperature range in which the complexes are stable. The thermal stability of G-quadruplex DNA was determined using UV thermal denaturation in the presence of Na\(^+\) and K\(^+\). Quadruplex monitored at 265nm (A\(_{265}\)) shows a sigmoidal (hyperchromic) curve, whereas monitoring the absorbance change with temperature at 295nm (A\(_{295}\)) displays a characteristics inverse sigmoidal hypochromic curve.
Fig. 4.28 (a, b, c & d) shows the thermal denaturation profile and their derivatives for 13GT1, 14GT2, 15GT3 and 16GT4 in the presence of 20 mM sodium cacodylate buffer (pH 7.0) containing either 0.1 M NaCl or 0.1 M KCl and 0.1 mM EDTA. Inverse sigmoidal curve obtained in UV melting experiments at 295 nm by the said oligonucleotides supports the presence of the G-quartets containing structures. A biphasic transition is obtained for 13GT1 (Fig. 4.28(a)) in presence of Na\textsuperscript{+} reflecting the unfolding of two structural species with a Tm 47 °C and 72 °C. It is important to note that biphasic thermal melting profile of 13GT1 contradict their status in gel picture demonstrating the presence of a single tetrameric structural species. It can be hypothesized that population of other species might be less that is why it could not be detected in gel. 14GT2 displayed a broad monophasic transition shows presence of a single structure. On analysis of melting curve from the first derivative plot (Fig. 4.28(b)) a broad transition with 53 °C Tm was obtained which suggests unfolding of two structural species. Since the transition of melting curve is broad, it reflects that the Tm’s of two species are slightly closer to each other resulting in overlapping of two transitions giving a broad melting profile. Biphasic melting profile is obtained for 15GT3 and 16GT4 sequences corresponding to two oligomeric structures in solution. The calculated melting temperatures for 15GT3 (Fig. 4.28(b)) were found to be 40 °C and 54 °C and for 16GT4 were found to be 38 °C and 51 °C respectively. The lower temperature transition corresponds to the melting of dimeric G-quadruplex species whereas the higher temperature transition of the biphasic curve presumably is the melting of the tetramericquadruplex species.

After establishing the G-quadrulex structure in Na\textsuperscript{+} further structural status of the oligonucleotides was determined in K\textsuperscript{+} ion. Fig. 4.28 (c & d) displayed the thermal denaturation profile of 13GT1, 14GT2, 15GT3 and 16GT4 in same solution condition containing 0.1 M KCl. 13GT1 exhibited a monophasic transition with a Tm 70 °C which suggesting melting of a single structural species. It can be assigned to the tetramer of 13GT1. UV-melting profile of 14GT2 also showed a monophasic transition. Analysis of melting curve reveals the melting temperature with a Tm 63 °C and can be assigned to the tetrameric species of 14GT2. It is in well correlation with the gel studies where only a single band is observed as a tetramer. Biphasic transition
is obtained for 15GT3 reflecting the possibility of two structural species in solution which goes well with gel studies. Analysis of melting curve revealed that lower transition with Tm 49 °C is assigned to bimolecular structure of 15GT3 and the higher transition with Tm value 80 °C is exhibited by tetrameric structure of 15GT3. Surprisingly, a monophasic transition is displayed by 16GT4 with calculated Tm of 76 °C. The same is in contradiction with gel studies where four bands were exhibited by 16GT4 in presence of K⁺ ion. This Tm value corresponds to the major tetramer and dimer unfolding simultaneously due to less difference in Tm values.

Further Thermal denaturation experiments were also performed for the extended version 25GT1, 28GT2, 31GT3 and 34GT4 oligonucleotides in 20mM sodium cacodylate buffer (pH 7.0) containing either 0.1 M NaCl or 0.1 M KCl and 0.1 mM EDTA. Fig. 4.29 displayed melting profile of said oligonucleotides in Na⁺ as well as in K⁺ ion. Biphasic transition is obtained for 25GT1 (Fig.4.29(a)) in presence of Na⁺ ion reflecting two oligomeric species in solution. Analysis of melting curve revealed that lower transition is 44 °C with upper temperature transition with Tm 70 °C. It is clear from the gel studies that only a single band is exhibited by 25GT1 which migrated with tetramer of 13GT1 confirming the bimolecular structure in gel. It is possible that the bimolecular species in gel has same molecularity. Monophasic transition were observed for 28GT2, 31GT3 and 34GT4 (Fig. 4.29(a&b)) oligonucleotide with Tm 74 °C, 69 °C and 60 °C respectively which is in good agreement with the result of gel assay where a single band of unimolecular species was obtained for these three sequences. It is clearly seen here that as the no. of T's increases in loop, the Tm decreased subsequently. It can be concluded that increase in T's in loop destabilizes the quadruplex structures.

Thermal denaturation experiments were performed in K⁺ ion to investigate the role played by cation on G-quadruplex structure formation and stabilization. Fig. 4.29(c&d) showed melting profiles of 25GT1, 28GT2, 31GT3 and 34GT4 in presence of K⁺ ion. Monophasic transition is obtained for all the four oligonucleotides reflecting the unfolding of single structural species in solution. Analysis of melting profile for 25GT1, reveals the high temperature value which can not be detected by
first derivative plot. It is also evident from the gel studies where only a single band was exhibited by 25GT1 which could be assigned to the folded unimolecular species. There is a fair correlation between Tm and gel studies. It was also reported that K\(^+\) induce more stable inter/intramolecular G quadruplex than their Na\(^+\) counterparts (Keniry et al, 2001). Calculated melting temperature for the 28GT2, 31GT3 and 34GT4 are 94 °C, 93 °C and 86 °C respectively. Interestingly gel results in K\(^+\) revealed two distinct bands of unimolecular and bimolecular species for 28GT2, 31GT3 and 34GT4. So, there seems a fair possibility that unimolecular species being very compact structure, exhibited high Tm and could not melt in the given temperature range in denaturation experiment. Therefore the detected Tm can be assigned to the bimolecular structure of G-quadruplexes for 28GT2, 31GT3 and 34GT4 respectively. After examining the stability of various structural species adopted by G-rich oligonucleotides their secondary structure was established by using Circular dichroism studies in following section.

4.4.3 Circular Dichroism Studies

We see that CD spectroscopy is a valuable tool elucidating the conformations adopted by various multistranded structure of nucleic acid. It is a fast and qualitative method to elucidate the formation of G-quadruplex. It also provides information of structural differences in sodium and potassium cations. CD experiments were performed in 20 mM sodium cacodylate buffer (pH 7.0) containing either 0.1 M NaCl or 0.1 M KCl and 0.1 mM EDTA to confirm the structural type formed by the sequences in native gel electrophoresis experiment and in order to better understand the different topologies adopted by the sequences. It is observed that with increase in the number of intervening bases T from T\(_1\)→T\(_4\) in all the G-rich oligos under study showed two positive peaks at 261 nm and 298 nm with negative peak centred at 241 nm. In presence of Na\(^+\), 13GT1 with single T in given sequence shows a prominent positive peak at 261nm with negative peak centred at 241 nm (Fig. 4.30 (a)). A shoulder at 298 was also observed which reflects the presence of a minor population of antiparallel species. However, the sequences 14GT2, 15GT3 and 16GT4 showed two positive peaks, centered at 260 nm and the other positive peak at 298 nm with
negative peak centred at 238 nm. Since, it is known to form parallel G-quadruplex structures give a positive peak at \( \approx 265 \) nm and a negative band at \( \approx 240 \) nm while antiparallel G-quadruplex structures show a positive peak at 295 nm with a negative band at 263 nm. The oligomers 13GT1, 14GT1, 15GT3, 16GT4 are shown to form parallel quadruplex while 14GT2, 15GT3 and 16GT4 also showed signature of antiparallel quadruplex. Thus, these spectral features are mainly attributed to the specific guanine stacking in various kinds of G-quadruplex structures (Gray et al, 2008). Thus, these spectra are also typical of mixed parallel/antiparallel topologies (Ambrus et al, 2005; Dai et al, 2006). Parallel G-quadruplex being major species and antiparallel as the minor. To our surprise, in case of \( K^+ \) (Fig. 4.30 (b)), all the said G-rich sequences exhibit a single intense positive peak at 264 nm with a negative band at 242 nm, typically assign to parallel type of G-quadruplexes. Fig. 4.31 displayed CD spectra of extended versions 25GT1, 28GT2, 31GT3 and 34GT4 in 20 mM sodium cacodylate buffer containing 0.1 M KCl and 0.1 mM EDTA. 25GT1 shows identical CD in case of Na\(^+\) as well as in K\(^+\). In both the cases it showed a prominent positive peak at 262 nm, a small hump at 297nm with a negative band at 239nm. But as the length of G-stacks increases with increase in number of intervening T bases intramolecular structures are more favoured in presence of Na\(^+\). Thus, in the presence of Na\(^+\) all the sequences except 25GT1 (Fig. 4.31(a)), show a sharp positive peak at 297 nm with a negative band at 263 nm. This CD signature is assign to antiparallel quadruplexes. It can be seen, that in the presence of K\(^+\) ion (Fig. 4.31(b)) all the longer G-rich sequences displayed a major positive peak around 264 nm with a minimum around 241 nm with an additional positive signal at 295 nm. It is clear that the sequence adopt mixed parallel/antiparallel topologies. CD studies are in good agreement with native gel results. Gel, Tm and CD results for all the used oligonucleotides are summarized and represented in the tabulated form in Table 4.1 (in Na\(^+\)) and 4.2 (K\(^+\)).

These results confirm that all the oligonucleotides fold to form G-quadruplexes and suggest that the sequences are polymorphic. Interestingly, Kypr et al (Kypr et al, 2009) reported that both types of quadruplexes display an additional characteristic positive peak at 210 nm. It was also observed that the long-wavelength
CD spectrum of a parallel quadruplex is similar to the CD spectrum of the A-form DNA. This may suggest an A-type stacking in these conformations (Kypr et al, 2001; Kypr et al, 2002). The quadruplex, however, has a positive band at 210 nm, whereas the A-form displays a deep negative one at 210 nm. The CD spectra of studied (14GT2, 15GT3, 16GT4, 25GT1, 28GT2, 31GT3 and 34GT4) G-rich sequences show a differential behaviour in Na$^+$ and K$^+$. G-quadruplex with parallel and antiparallel topologies can be present in equilibrium with each other according to the nature of cations in solutions (Alberti et al, 2011). Although CD spectra are routinely used to assign DNA secondary structures (Chairs et al, 2005) the interpretation of optical properties such as hypochromicity or the shape and sign of CD bands can be controversial (Kypr et al, 2005). G-stacking, the polarity and the rotation angle between the stacks are determining factors in the intensity and shape of the CD spectrum and specifically. Explanation of CD spectra in the context of folding pattern of G-quadruplex depends upon the inherent postulation that a sequence folds into a single G-quadruplex structure (Spada et al, 2010). But in reality a single sequence can fold in more than one structure in the same solution. We observed this in gel studies for the sequences 16GT4, 28GT2, 31GT3 and 34GT4 in presence of K$^+$ for the sequences 16GT4, 28GT2, 31GT3 and 34GT4. The influence of G-tract length, loop length and loop sequence on G-quadruplex topology and stabilities has been studied and explored by various groups on various sequences (Wang et al, 1994; Ristiano et al, 2003; Hazel et al, 2004; Rachwal et al, 2007; Bugaut et al, 2008; Guedin et al, 2008; Lim et al, 2009; Hu et al, 2009; Mergny et al, 2010) but no universal rule has emerged yet. To sum up, based on results of Gel studies, melting studies and CD studies all the possible structure adopted by G-rich sequences used in study has been proposed (Fig 4.32 and 4.33).
Table 4.1: Structural Status of DNA sequences in Na\(^+\) (Gel Studies, Tm & CD Studies)

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Sequences</th>
<th>Gel Studies</th>
<th>Tm ((^\circ)C)</th>
<th>CD Studies</th>
<th>Conformations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>TGGGGGTGCGGGGT (13GT1)</td>
<td>Tetramer (One Band)</td>
<td>47, 72</td>
<td>+ve261nm, 298nm</td>
<td>Tetramer (Parallel)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- ve241nm</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>TGGGGGTGCGGGGT (14GT2)</td>
<td>Tetramer (One Band)</td>
<td>53</td>
<td>+ve260nm, 298nm, -ve 238nm</td>
<td>Tetramer (Parallel)</td>
</tr>
<tr>
<td>3.</td>
<td>TGGGGGTGCGGGGT (15GT3)</td>
<td>Tetramer &amp; Dimer (Two Bands)</td>
<td>40, 54</td>
<td>+ve260nm, 298nm, -ve238nm</td>
<td>Tetramer (Parallel) &amp; Dimer (Antiparallel)</td>
</tr>
<tr>
<td>4.</td>
<td>TGGGGGTGCGGGGT (16GT4)</td>
<td>Tetramer &amp; Dimer (Two Bands)</td>
<td>38, 51</td>
<td>+ve 260nm, 298nm, -ve238nm</td>
<td>Tetramer (Parallel) &amp; Dimer (Antiparallel)</td>
</tr>
<tr>
<td>5.</td>
<td>TGGGGGTGCGGGGTGCGG (25GT1)</td>
<td>Dimer (One Band)</td>
<td>44, 70</td>
<td>+ve262nm, 297nm, -ve 239nm</td>
<td>Mixed (Parallel &amp; Antiparallel)</td>
</tr>
<tr>
<td>6.</td>
<td>TGGGGGTGCGGGGTGCGG (28GT2)</td>
<td>Unimolecular (One Band)</td>
<td>74</td>
<td>+ve297nm, -ve263nm</td>
<td>Intramolecular (Antiparallel)</td>
</tr>
<tr>
<td>7.</td>
<td>TGGGGGTGCGGGGTGCGG (31GT3)</td>
<td>Unimolecular (One Band)</td>
<td>69</td>
<td>+ve297nm, -ve263nm</td>
<td>Intramolecular (Antiparallel)</td>
</tr>
<tr>
<td>8.</td>
<td>TGGGGGTGCGGGGTGCGG (34GT4)</td>
<td>Unimolecular (One Band)</td>
<td>61</td>
<td>+ve297nm, -ve263nm</td>
<td>Intramolecular (Antiparallel)</td>
</tr>
</tbody>
</table>
Table 4.2: Structural Status of DNA sequences in K⁺ (Gel Studies, Tm & CD Studies)

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Sequences</th>
<th>Gel Studies</th>
<th>Tm (°C)</th>
<th>CD Studies</th>
<th>Conformations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>TGGGGGTG GGGGT (13GT1)</td>
<td>Tetramer (One Band)</td>
<td>70</td>
<td>+ve 264nm, -ve 242nm</td>
<td>Tetramer (Parallel)</td>
</tr>
<tr>
<td>2.</td>
<td>TGGGGGTGGGGGT (14GT2)</td>
<td>Tetramer (One Band)</td>
<td>63</td>
<td>+ve 264nm, -ve 242nm</td>
<td>Tetramer (Parallel)</td>
</tr>
<tr>
<td>3.</td>
<td>TGGGGGTTGGGGGT (15GT3)</td>
<td>Tetramer, Dimer (Two Bands)</td>
<td>49, 80</td>
<td>+ve 264nm, -ve 242nm</td>
<td>Tetramer (Parallel)</td>
</tr>
<tr>
<td>4.</td>
<td>TGGGGGTTGGGGGT (16GT4)</td>
<td>Dimer of Tetramer, Tetramer &amp; Dimer (Four Bands)</td>
<td>75</td>
<td>+ve 264nm, -ve 242nm</td>
<td>Tetramer, Dimer of Tetramer, Dimer (Parallel)</td>
</tr>
<tr>
<td>5.</td>
<td>TGGGGGTGGGGGTGGGGGTGGGTTG (25GT1)</td>
<td>Unimolecular (One Band)</td>
<td>nd*</td>
<td>+ve 264nm, 295nm, -ve 241nm</td>
<td>Mixed (Parallel &amp; Antiparallel)</td>
</tr>
<tr>
<td>6.</td>
<td>TGGGGGTTGGGGGTGGGGGT (28GT2)</td>
<td>Dimer &amp; Intramolecular (Two Bands)</td>
<td>94</td>
<td>+ve 264nm, 295nm, -ve 241nm</td>
<td>Mixed (Parallel &amp; Antiparallel)</td>
</tr>
<tr>
<td>7.</td>
<td>TGGGGGTTGGGGGTGGGGGTGGGGGT (31GT3)</td>
<td>Dimer &amp; Intramolecular (Two Bands)</td>
<td>93</td>
<td>+ve 264nm, 295nm, -ve 241nm</td>
<td>Mixed (Parallel &amp; Antiparallel)</td>
</tr>
<tr>
<td>8.</td>
<td>TGGGGGTTGGGGGTGGGGGT (34GT4)</td>
<td>Dimer &amp; Intramolecular (Two Bands)</td>
<td>86</td>
<td>+ve 264nm, 295nm, -ve 241nm</td>
<td>Mixed (Parallel &amp; Antiparallel)</td>
</tr>
</tbody>
</table>

(nd* not detected)