CHAPTER 2

G-QUADRUPLEX COUPLED (QC) KINETICS
2.1 Introduction

Expression of the oncogene c-MYC is associated with cellular proliferation and control of differentiation. As a result, loss of regulation resulting in overexpression of c-MYC is correlated to a large number of human and animal cancers (81,82,83,84). Antisense oligonucleotide mediated transcription silencing has been observed to induce differentiation in myelocytes indicating the role of aberrant c-MYC overexpression in differentiation (85,86). Transcription regulation of c-MYC is complex and involves multiple promoters, P1 and P2 being prominent among them (for reviews see (82,87). The nuclease hypersensitive element (NHE) III, corresponding to -147 to -117 bases relative to P1 transcription initiation site controls more than 80% of c-MYC transcription and hence is an important anti-cancer target (88,89,90,91). It has been observed that the purine-rich anti-sense strand of the NHE adopts a G-quadruplex conformation and it was recently shown that the structure could be a regulatory switch for c-MYC (13,92). Based on this and various other observations, postulated models of regulatory control entail a switch between the G-quadruplex and duplex DNA, which could be central in elucidation of the mechanism of c-MYC transcription and design of antisense therapy (88,89,91,93). The orchestration of the structural transitions driving this quadruplex-duplex competition is poorly understood.

The G-quadruplex constitutes a four-strand fold-back structure of stacked guanine-tetrads. These tetrads are coplanar arrangement of four guanines held together by Hoogsteen hydrogen bonds (17,53). Apart from the promoter region of c-MYC, sequences that form G-quadruplex in vitro have been found in the telomeres (94) and within the switch regions of immunoglobin heavy-chain genes (95). Interestingly, recent evidence implicates these unusual DNA structures as ‘at risk motifs’ (96) due to their involvement in genome rearrangements induced by polymerase slippage events in the nematode Caenorhabditis elegans on inactivation of a putative helicase, DOG-1 (97). In a genomic context, formation of G-quadruplex competes with duplex formation and thus the kinetics and thermodynamics of the structural transitions would be the underlying factors determining its functional role.
Determination of the competing rate constants (G-quadruplex folding and hybridization) requires simultaneous determination of the folding/unfolding rates and the duplex formation rates. Nanomotors have been designed based on the folding/unfolding of G-quadruplex motifs, which were demonstrated using FRET (98). The rates of folding/unfolding determined the efficiency of the nanomachine and could be regulated using a duplex trap. A FRET-based study has been used to observe the quadruplex folding constants in the presence of a PNA trap, where the PNA strand concentration was maintained such that hybridization was very fast (99). However, in vivo extrapolations can be made only when the strand concentrations are equimolar and very low. A recent report addresses this problem using human telomeric G-quadruplex hybridization on an optical biosensor based on surface plasmon resonance (SPR) and suggests a possible quadruplex-duplex competition mechanism at low equimolar concentration of the complementary strand (57).

Based on DNase I hypersensitivity, it was reported that the major regulatory element of c-MYC exists in a strand-separated form rendering this location as a nuclease hypersensitive element (NHE III) (88). We hypothesized that the underlying inherent kinetics of duplex formation may play a significant role, in conjunction with other cellular factors, which could be important for the crucial regulatory mechanism. In this study, we used SPR-based biosensor to observe competing hybridization versus G-quadruplex formation in the c-MYC regulatory region at physiological conditions. Using an analytical component resolution method described herein for the first time, we could not only simultaneously determine the individual rate constants of folding/unfolding (of G-quadruplex) and association/dissociation (of hybridization) but we also separated the two components of the hybridization reaction. One resulting from hybridization with pre-equilibrated unstructured oligonucleotides present on sensor surface and the other due to hybridization with immobilized molecules unfolding in the presence of the complementary strand (during injection). We observed that both the folded and unfolded forms have short half-lives of less than 90 sec and our results further indicated that the rate-limiting step changes as a result of complementary strand concentration. At low strand concentration hybridization is slow and determines the overall rate while with
increasing concentration motif transition becomes rate determining. Based on our results we conclude that G-quadruplex may be the predominant state at the low intracellular strand concentrations because duplex formation is kinetically unfavorable.

2.2 Material and Methods

2.2.1 Oligonucleotides Used in this Study

All oligodeoxynucleotides used in this study (Table 2.1, below) were obtained in HPLC purified form from Sigma Genosys and dissolved in MilliQ purified water.

Table 2.1

<table>
<thead>
<tr>
<th></th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>5'-CCCCACCTTCCCCACCTCCCCACCTCCCCC-3'</td>
</tr>
<tr>
<td>G1</td>
<td>5'-GGGGAGGGTGGGGAGGGTGGGGAAGGTGGGG-3'</td>
</tr>
<tr>
<td>C1B</td>
<td>biotin-5'-acgtacgtCCCTTTTCCCCACCTCCCCACCTCCCCACCTCCCC-3'</td>
</tr>
<tr>
<td>G1B</td>
<td>biotin-5'-acgtacgtGGGGAGGGGTGGGGAGGGGTGGGGAAGGTGGGG-3'</td>
</tr>
<tr>
<td>M1</td>
<td>5'-CCCTACCTTCCCCACCTCCCCACCTCCCCACCTCCCC-3'</td>
</tr>
<tr>
<td>M2</td>
<td>5'-CCCCACCTTCCCCACCTCCCCACCTCCCCACCTCCCC-3'</td>
</tr>
<tr>
<td>M3</td>
<td>5'-CCCCACCTTCCCCACCTCCCCACCTCCCCACCTCCCC-3'</td>
</tr>
</tbody>
</table>

The 31-mer oligonucleotides C1 and G1 constitute the nuclease hypersensitive element the in c-MYC promoter region corresponding to -147 to -117 bases relative to P1 transcription initiation site. Biotinylated oligonucleotides C1B and G1B with 8-mer spacers (subscript) at the 5'-end were used for immobilization on the sensor surface. M1, M2 and M3 have single base substitutions (bold) with respect to C1

2.2.2 CD Spectroscopy

Circular Dichroism (CD) measurements were performed on a Jasco Spectropolarimeter (model J 715) equipped with a thermostat controlled cell holder with a cell path length of 1 cm as described previously (100). 1.28 µM of G1 in buffer (10 mM HEPES, 150 mM NaCl, 3 mM MgCl₂, pH 7.4 adjusted with NaOH) was heated at 95 °C for 10 min before slowly cooling to 25 °C and mixed with C1 (1.28, 2.56 or
3.84 µM). Samples were equilibrated for at least 2 hr after each addition before recording CD spectra from 230 to 330 nm with an averaging time of 3 sec.

2.2.3 Poly-Acrylamide Gel Electrophoresis (PAGE)

Nondenaturing gel electrophoresis experiments were performed with labeled C1 (5’-end labeled with T4 polynucleotide kinase (New England Biolabs)) and 32P ATP. Free ATP was removed by denaturing gel electrophoresis. Experiments were done in 10 mM HEPES, 150 mM NaCl, 3 mM MgCl₂, pH 7.4 (adjusted with NaOH) using radiolabeled 10 nM C1 (in the presence of 0.5 µM unlabeled C1) after heating at 95 °C for 10 min followed by incubation at 4 °C for 18 hr with G1 (0.5 or 3 µM) in the absence or presence (15 min) of DNase I (5 U) before loading on a nondenaturing 20% polyacrylamide gel. Electrophoresis was performed in 0.5X Tris-Borate-EDTA (TBE) buffer (pH 8.0) in a thermostated apparatus (SE 600, Hoefer Scientific) run at 4 °C for 6-8 hrs at 90 V. Gels were vacuum dried and analyzed on a phosphorimager (Fujifilm 2000).

2.2.4 Hybridization Kinetics using Surface Plasmon Resonance

SPR measurements were performed with BIAcore 2000 (BIAcore Inc.) system using streptavidin-coated sensor chips (Sensor chip SA, BIAcore Inc.). The 39-mer 5’-biotinylated sequence G1B and C1B were immobilized on flow cells as described before (101). Briefly, the streptavidin sensor-chip surface was prepared for DNA immobilization with three consecutive 1 min injections of 1 M NaCl in 50 mM NaOH followed by extensive washing with buffer. The 5’ biotinylated oligomer was heated at 95 °C for 5 min and slowly cooled to room temperature in HBS buffer (filtered and degassed 10 mM HEPES, 150 mM NaCl, 3 mM MgCl₂, pH 7.4 and 0.005% surfactant P20 from Biacore) before noncovalent immobilization on flow cell 2. Flow cell 1 was left blank as control to account for any signal generated due to bulk solvent effect or any other effect not specific to the DNA interaction, which was subtracted from the signal obtained in flow cell 2 and 3. All experiments were performed at 25 °C using running buffer (filtered and degassed 10 mM HEPES with 150 mM of either KCl, NaCl or LiCl and 0.005 % surfactant IGEPAL) at pH 7.4 (adjusted with the respective base
KOH, NaOH or LiOH, respectively). Oligonucleotide immobilized surface was exposed to the running buffer for at least 2 hrs at a flow rate of 5 µl/min for attaining baseline stability. Analyte (C1, M1, M2, M3 or G1) solutions at different concentrations (16 – 1024 nM) in the running buffer were injected at 20 µl/min for 180 sec in random series to avoid any systematic error, using an automated protocol. Following this, dissociation from the surface was monitored for 300 sec in running buffer. Seven different concentrations were used for each analyte and each injection was run in duplicate before fitting to respective equations in order to extract kinetic parameters. Regeneration was done using 1 M NaCl in 50 mM NaOH, as the running buffer could not completely dissociate the complex from the surface. Mass transfer analysis done at varying flow rates (5, 20, 50 µl/min) showed no significant difference in association rates. The BIAevaluation 3.1.1 software supplied by manufacturer was used to compile the quadruplex-coupled hybridization model. BIAevaluation 3.1.1 was used as it is for fitting the simple-hybridization model. The dissociation phase was used to determine \( k_d \), which was used in the association phase to extract rest of the kinetic parameters.

2.3 G-quadruplex Coupled (QC) Hybridization Model

The method is based on surface plasmon resonance (SPR) applied to an optical biosensor (from BIACore Inc.), that enables real time detection of molecular association and dissociation reactions by monitoring change in refractive index due to alteration in mass on the optical sensor (102,103). The refractive-index change is represented as response units (RU) in a sensorgram, which is proportional to the amount of analyte (injected in mobile phase) binding to ligand immobilized on the sensor surface. The basic method and the theoretical background have been described in detail previously (104,105,106). In SPR biosensor, the interaction between the surface immobilized ligand (A) and the solution phase analyte (B) can be described as,

\[
A + B \xrightleftharpoons[k_d]{k_a} AgB
\]
where $k_a$ and $k_d$ are the association and dissociation rate constants respectively and AB represents the bound complex formed on the biosensor surface. The kinetic expression for the above type of general ligand-analyte interaction has been discussed before and is given by (105)

$$R_{t(\text{assoc})} = \frac{Ck_aR_{\text{max}}(1 - e^{-(k_aC+k_d)(t'}-t_0)})}{(k_aC+k_d)} + R_i$$  \hspace{1cm} (2.2)$$

for the association phase and

$$R_{t(\text{diss})} = R_{\text{amp}}e^{-kd}t + R_\text{t}(t\rightarrow \infty)$$  \hspace{1cm} (2.3)$$

for the dissociation part of the sensorgram, where $R_i$ is the actual observed signal expressed in response units (RU) at any time $t$, $R_{\text{max}}$ is the expected maximal response proportional to the immobilized ligand, $C$ is the constant concentration of analyte in solution, $R_i$ is a fitting parameter equivalent to the signal at the point of injection of analyte ($t=0$) and accounts for any change in the running buffer composition, $R_{\text{amp}}$ is defined as the amplitude of the dissociation curve and $R_\text{t}(t\rightarrow \infty)$ is the response value after infinite time and represents complete dissociation of the complex. The rate constants, $k_a$ and $k_d$ are determined by fitting the sensorgrams to equations (2.2) and (2.3).

Although the above equations (2.2) and (2.3) hold good for any general 1:1 interaction, they cannot account for the presence of an additional equilibrium on the sensor surface. We propose a coupled-kinetic model for this purpose, which simultaneously determines the rate constants for the equilibrium between different conformations of the surface-attached ligand and its interaction with the analyte in mobile phase. We have considered the surface equilibrium as inter-convolutions between two possible conformations of the ligand in its immobilized form, where only one of the conformations can interact with the analyte. For simplicity of the principle and theoretical aspects of rate law expressions, we classified the whole sensorgram into three major parts, as also shown in Figure 2.1 below,
2.3.1 Pre equilibrium ('b - c' in Figure 2.1)

It is considered that the surface immobilized ligand is present in two structural isoforms, one folded conformation (F) representing the G-quadruplex, while the other being the unstuctured conformation (U) of the guanine-rich sequence G1B under the given buffer conditions. So, after regeneration ('a - b' in Figure 2.1) the guanine-rich oligonucleotide would establish an equilibrium between U and F in time $t_e$ ('b - c' in Figure 2.1) which can be represented as follows,

$$ F \overset{k_f}{\underset{k_{f}}{\rightleftharpoons}} U $$

(2.4)

at time, $t=0$ $R_{\text{max}}$ 0

at time, $t=t_e$ $(R_{\text{max}} - U_e)$ $U_e$

Figure 2.1  Binding of immobilized ligand to analyte molecules results in increase of RU (association) during injection followed by decrease in RU (dissociation) after injection stop.
where, $k_u$ and $k_f$ are the unfolding and folding rate constants of the immobilized ligand, $R_{\text{max}}$ is the concentration of the total immobilized ligand on the chip surface and $U_0$ is the concentration of the unfolded conformation at time $t=t_e$, which is the concentration of $U$ available for hybridization at the initiation of analyte injection. The high monovalent ion concentration used during regeneration (1 M NaCl) ensures maximal folded structures whereby at time $t = 0$, $U$ has been taken to be 0. The differential rate expression of the above equation can be written as,

$$
\frac{dU_e}{dt} = k_u (R_{\text{max}} - U_e) - k_f U_e
$$

(2.5)

Integrating the above equation, we have,

$$
U_e = \frac{k_u R_{\text{max}} (1 - e^{-(k_u + k_f) t_e})}{(k_u + k_f)}
$$

(2.6)

From equation (2.6), it can be easily noted that since $k_u$, $k_f$ and $R_{\text{max}}$ are constant parameters, the concentration of unfolded conformation ($U_e$) would be fully dependent on the equilibration time $t_e$.

**2.3.2 Association (‘c – d’ in Figure 2.1)**

In the quadruplex-coupled model we consider that the total response observed (RU) in a sensorgram is due to the summation of two components. Due to hybridization of analyte with (a) unfolded species already present at the time of injection and (b) unfolded species forming during injection under the influence of the complementary strand – i.e. due to coupling of the hybridization equilibrium with the quadruplex folding/unfolding equilibrium.

The first component can be represented as:
\[ U + C \xrightleftharpoons[k_d]{k_u} D \]  
(2.7)

at time, \( t=0 \) \( U_c \quad 0 \)
at any time (t), after injection \( (U_c - D_c) \quad D_c \)
where, \( D_c \) is the duplex formed from the pre-existing unfolded ligand isoform, \( U_e \).
Rate expression of the above equation would be,

\[
\frac{dD_c}{dt} = k_u C(U_c - D_c) - k_d D_c
\]
(2.8)

On integration,

\[
D_c = \frac{k_u U_c C (1 - e^{-(k_u C + k_d) (t - t_o)})}{(k_u C + k_d)}
\]
(2.9)

where ‘\( t \)’ represents the time elapsed after the analyte injection.
Substituting the value of \( U_c \) from equation (2.6) we have:

\[
D_c = \frac{k_u k_u R_{max} C (1 - e^{-(k_u C + k_d) (t - t_o)}) (1 - e^{-(k_u + k_f) t_o})}{(k_u C + k_d)(k_u + k_f)}
\]
(2.10)

The second component of duplex formation arising out of hybridization coupled with quadruplex unfolding, may be represented by considering the two equilibria simultaneously:

\[
F \xrightleftharpoons[k_f]{k_u} U
\]
\[ U + C \xrightleftharpoons[k_d]{k_u} D \]  
(2.11)
In case of quadruplexes, $k_a >> k_f$, i.e. the rate constant of hybridization is fast compared to folding (as observed in a number of cases (54,57,99,107,108)), thus the above expression can be rewritten as:

$$
\begin{align*}
F & \xrightarrow{k_u} U \\
U + C & \xrightarrow{k_f} D \\
\end{align*}
$$

at time, $t=t_o$  
\[(R_{max} - U_e) \quad 0\]

and at any time = $t$, \[(R_{max} - U_e - D_t) \quad D_t\]

where, \((R_{max} - U_e)\) is the concentration of folded molecules after the equilibration phase (quadruplex folding/unfolding equilibrium before injection) and $D_t$ is duplex formed from hybridization with molecules unfolding during injection. Since, the slowest (rate-limiting) step in the above expression is ligand unfolding $k_u$, the formation of $D_t$ can be expressed as follows,

$$
\frac{dD_t}{dt} = k_u(R_{max} - U_e - D_t) 
$$

(2.13)

On integration we have,

$$
D_t = (R_{max} - U_e)(1 - e^{-k_u(t-t_o)}) 
$$

(2.14)

Substituting for $U_e$ from equation (6),

$$
D_t = \frac{R_{max}\left[k_f + (k_u e^{(k_u + k_f)t_o})\right]}{(k_u + k_f)}(1 - e^{-k_u(t-t_o)}) 
$$

(2.15)
So finally, as discussed earlier, the real-time response observed in a quadruplex-coupled sensorgram can be mathematically expressed as summation of two components, \[ D = D_c + D_t + R_i \] (2.16)

where, \( R_i \) is a fitting parameter which is equivalent to any response change due to the alteration in bulk refractive index between running buffer and analyte injection buffer and has been discussed previously (105). Thus from equation (2.10) and (2.15) the total duplex formed at any time \( t \) during the association phase of the sensorgram is given by,

\[
D = k_u k_u R_{\text{amp}} C \left(1 - e^{-(k_u C + k_d) t_{\text{amp}}}ight) \left(1 - e^{-(k_u + k_f) t_{\text{amp}}}ight) + \frac{R_{\text{max}} \left(k_u + (k_u e^{(k_u + k_f) t_{\text{amp}}}) \right) \left(1 - e^{-(k_u + k_f) t_{\text{amp}}}ight)}{(k_u + k_f)} + R_i
\]

(2.17)

2.3.3 Dissociation ('d – e' in Figure 2.1)

The dissociation phase (after ‘injection stop’ in Figure 2.1) can be fitted with

\[
R_{t(\text{diss})} = R_{\text{amp}} e^{-k_d t} + R_{t\rightarrow\infty}
\]

(2.18)

where, \( R_i \) is the actual signal observed expressed in Response Units (RU) at any time \( t \), \( R_{\text{amp}} \) is defined as the amplitude of the dissociation curve and \( R_{t\rightarrow\infty} \) is the response value after infinite time and represents complete dissociation of the complex.

2.4 Results

2.4.1 Hybridization Influenced by G-quadruplex Formation on Sensor Surface

Figure 2.2 shows that the purine-rich strand (G1) from c-MYC NHE forms parallel G-quadruplex under our experimental conditions as indicated by the
characteristic positive and negative maxima at 262 nm and 236 nm, respectively (53). Many previous studies have observed G-quadruplex formation in vitro by this sequence under various conditions (13,92,101,109,110). Duplex formation was observed on titration with C1 (with 1:1 molar ratio; positive peak shifts to 268 nm) however, no triplex formation (expected positive peak at 282 nm (111)) could be observed on using excess C1. At molar excess of C1, the observed CD profile is characteristic of a mixture of unstructured single strand C1 (positive maxima at 277 nm (93)) and duplex DNA. The reason for not observing any intramolecular \textit{i}-tetraplex formation (with positive CD peak at 285 nm) at molar excess of C1 may be the slightly acidic conditions required for \textit{i}-tetraplex formation (93,112).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2.2.png}
\caption{Hybridization of C1 and G1 monitored by CD. 2 \textmu M G1 (black line) was treated with 1 \textmu M (red line), 2 \textmu M (red dash) or 3 \textmu M (black dash) C1 in 10 mM HEPES, 150 mM NaCl, 3 mM MgCl\textsubscript{2}, pH 7.4. Spectra were recorded at 25 °C, 2 hrs after each addition.}
\end{figure}

In an earlier study using specific binding of Hoechst 33258 to the c-MYC G-quadruplex we obtained evidence of quadruplex formation by G1B on the sensor chip surface (101). Herein, G1B was immobilized on sensor and hybridized with increasing concentration of the complementary strand C1 (or M1, M2, M3).
Sensorgrams obtained on hybridization were fitted with the quadruplex-coupled (QC) model (equation (2.17) and (2.18)) to obtain very good fits in both K⁺ and Na⁺ (Figure 2.3). Sensorgrams were obtained by hybridization with immobilized G1 (A, B, C and E) or C1 (D) using 16, 32, 64, 128, 256, 512 or 1024 nM of the respective complementary strand in the mobile phase. Fitted curves (red) were obtained by fitting the sensorgrams (black) with either the quadruplex-coupled model using Equation 17 (A and B) or simple-hybridization model using Equation 2 (C, D and E) as in Material and Methods. Sensorgrams were obtained at 25 °C and pH 7.4 in either 150 mM KCl (A) or NaCl (B, C, D). (E) Hybridization in the presence of 150 mM Li⁺. The fitted parameters were $k_a$ ($2.8 \times 10^5$ M$^{-1}$s$^{-1}$ (± 2.41%)) and $k_d$ ($5.4 \times 10^{-5}$ s$^{-1}$ (± 1.98%)) respectively giving an equilibrium dissociation constant $K_D = 0.19$ nM. All experimental conditions were as of above set except that LiOH was used to adjust buffer pH.
Results are listed in Table 2.2. The sensorgrams could not be adequately represented using the two-state association model (equation (2.2) and (2.3)). This is evident by a comparison of Figure 2.3B and 2.3C wherein the same set of sensorgrams obtained in Na+ was fitted using either the QC model or the two-state model, respectively. A similar effect was also observed in K+ indicating secondary structure formation by the c-MYC purine-rich sequence on the sensor in both K+ and Na+. The complementary pyrimidine-rich strand does not adopt secondary structure at pH 7.4 as seen in Figure 2.2 with molar excess of C1, as 2.3A and 2.3B, respectively). Results are listed in Table 2.2. The sensorgrams could not be adequately represented using the two-state association model (equation (2.2) and (2.3)). This is evident by a comparison of Figure 2.3B and 2.3C wherein the same set of sensorgrams obtained in Na+ was fitted using either the QC model or the two observed earlier (93,112). We used biotinylated C1 (C1B) as a control as it was expected to follow simple two-state hybridization in the absence of additional folding/unfolding equilibrium on sensor like G1B. Figure 2.3D shows that the sensorgrams obtained with G1 in the mobile phase complied with a two-state hybridization model. Similarly, G-quadruplex formation in the presence of Li+ is unlikely whereby we expected duplex formation to follow a simple hybridization model. Sensorgrams obtained in 150 mM Li+ could be fitted satisfactorily without using a QC model (Figure 2.3E). Taken together, these observations indicate that the effect of secondary structure formation on hybridization is not an artifact. An 8-mer spacer separating biotin from the actual NHE sequence was used such that it tethers off the surface and minimizes the effect of the carboxymethyl dextran surface on the interaction.

Before using the QC model, it was important to ascertain whether triplex formation was occurring at high analyte concentration as this could contribute to the biphasic transition considered in our model. We did not observe triplex formation by CD (Figure 2.2). This was further confirmed using a non-denaturing PAGE with labeled C1 in the presence of excess G1, which showed duplex but no triplex formation as
confirmed by DNase I cleavage (Figure 2.4A). Table 2.2 summarizes the kinetic
parameters obtained using the QC model for hybridization with C1 and three other
single base mutant oligonucleotides (M1, M2 and M3) in the mobile phase. In case of
C1, as expected, the folded form of the G-quadruplex attached to the sensor was
more stable in K⁺ than Na⁺ (as observed from the corresponding folding half-lives ($t_{1/2}$) and
equilibrium folding constants ($K_r$)) (Table 2.2). This resulted in a relatively higher
amount of unfolded form on the sensor surface in case of Na⁺ and was reflected in
the increased amount of hybridization observed in Na⁺. The equilibrium RU observed
in K⁺ (Figure 2.3A) was lower than in Na⁺ (Figure 2.3B) by almost four-folds as
expected from the difference in respective $K_r$ values. Our results are consistent with
previous reports indicating stabilizing effect of K⁺ on G-quadruplex folding in general
(113,114). Similar observations have also been made in a recent study on

Figure 2.4 Hybridization of C1 and G1 shows no triplex formation (A). Reactions
had 10 nM of 5'-end labeled and 0.5 µM unlabeled C1 (lane 1) with either, 0.5 µM
G1 (lane 2) and 5U of DNase I (lane 3) or 1.5 µM G1 (lane 4) and 5U of DNase I
(lane 5). Samples were incubated at 4 °C for 18 hrs before 15 min DNase I
treatment (lanes 3 and 5). G1 shows multiple folded conformations (B). 10 nM of
5'-end labeled and 0.5 µM unlabeled G1 (lane 1), C1 (lane 2) and 31-mer control
dT31 (lane 3) were incubated for 4 hours at 25 °C before separation. Both (A) and
(B) were incubated in 10 mM HEPES, 150 mM NaCl, 3 mM MgCl₂, pH 7.4. Bands
were separated in a 20 % non-denaturing polyacrylamide gel electrophoresis in
0.5x TBE buffer (pH 8.0) at 4 °C for 6 hrs at 90 V and visualized using
autoradiography on phosphorimag er (Fujifilm 2000).
folding/unfolding of the telomeric G-quadruplex (57). Bimolecular hybridization yielded an equilibrium dissociation constant of $3.61 \times 10^{-10}$ M in $K^+$.  

**Table 2.2.** Kinetic parameters for quadruplex folding/unfolding and hybridization of the NHE in the c-MYC promoter $^a$

<table>
<thead>
<tr>
<th></th>
<th>150 mM K$^+$</th>
<th>150 mM Na$^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C1$^b$</td>
<td>M1$^b$</td>
</tr>
<tr>
<td>$k_u$ (s$^{-1}$)</td>
<td>$7.90 \times 10^{-3}$ (±1.52%)</td>
<td>$1.56 \times 10^{-2}$ (±1.52%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_f$ (s$^{-1}$)</td>
<td>$1.65 \times 10^{-2}$ (±1.84%)</td>
<td>$8.34 \times 10^{-3}$ (±2.25%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_{1/2}$ (s)</td>
<td>87.7</td>
<td>45.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_{u1/2}$ (s)</td>
<td>42.0</td>
<td>83.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_F$</td>
<td>2.09</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_a$ (M$^{-1}$s$^{-1}$)</td>
<td>$1.37 \times 10^{6}$ (±1.66%)</td>
<td>$5.15 \times 10^{5}$ (±0.87%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_D$ (M)</td>
<td>$3.61 \times 10^{-10}$</td>
<td>$0.97 \times 10^{-10}$</td>
</tr>
</tbody>
</table>

$^a$ Sensorgrams were obtained in 150 mM K$^+$ or Na$^+$ at 25 °C and fitted to the quadruplex-coupled hybridization model. $K_F$ is the equilibrium constant for quadruplex formation calculated from $k_u/k_f$; $K_D$ denotes the equilibrium dissociation constant for duplex formation, given by $k_d/k_a$. The half-lives, $t_{1/2} (= \ln 2/k_u)$ and $t_{u1/2} (= \ln 2/k_f)$ are for the folded and unfolded forms, respectively. Numbers in parentheses are for standard errors. $^b$ Sensor surface was immobilized with 1165 RU of the G-rich oligonucleotide G1B before using the respective oligonucleotides C1 (in K$^+$ or Na$^+$) or M1, M2 and M3 in Na$^+$ as analyte in the mobile phase. Kinetic parameters were extracted using the QC model. $^c$ Sensor surface was immobilized with 1050 RU of the C-rich oligonucleotide C1B; in accordance with the folding constant $K_F$ (0.54) obtained for G1B (equilibrium on surface) ~65% of the injected G1 concentration, which was expected to be unfolded in solution was used for fitting. Kinetic parameters were extracted using a simple hybridization model (1:1 Langmuir association from BIAevaluation 3.1.1).
(and 0.97 x 10^{-10} M in Na^+), which is of the same order of magnitude as observed by others using SPR under similar conditions (57,108). The observed higher duplex binding affinity in Na^+ with respect to K^+ was primarily due to increased (~3.7 fold) duplex association in the presence of Na^+ as the dissociation constants in both cases were similar. The mutants (M1, M2 and M3) were used to ascertain the QC model as they were expected to change the hybridization rates (k_a/k_d) without affecting the folding/unfolding (k_u/k_f) of G1B. We obtained very similar k_u and k_f parameters for C1 and the mutants (Table 2.2). The association rates (k_a) observed for the mutants were lower than the C1 while the dissociation rate was observed to be somewhat higher only for M1. This is consistent with earlier reports of hybridization observed by SPR using single base mismatches (107). On the other hand, on using G1 in the mobile phase, a two-state hybridization reaction gave k_a and k_d values within 5%. In this case association (k_a) was almost 10-fold lower than association with immobilized G1 and dissociation was about 2.6-fold faster yielding a K_D, which was higher than 40-fold. In accordance with the equilibrium folding constant K_F (0.54) obtained for G1B in Na^+ (on sensor surface), ~65% of the injected G1 concentration, which was expected to be unfolded in solution, was used for fitting. We hypothesized that the observed discrepancy could be due to additional secondary structure formation by G1 that does not participate in the hybridization reaction. In order to confirm this, we performed non-denaturing PAGE and found that G1 formed multiple folded conformations while C1 did not show alternative conformations (Figure 2.4B).

2.4.2 G-quadruplex Coupled (QC) Hybridization Results and Analysis

G1B molecules attached to sensor surface present two different modes for hybridization – (a) unfolded strand due to the pre-equilibration phase and (b) G1B unfolding in the presence of C1 during injection. The latter mode is important in the context of genomic sequences where most transitions are bound to occur in the
presence of the complementary strand. We attempted to resolve these components. Using an association model after incorporating the secondary structure unfolding (QC model) allowed us to simulate the different components of the bimolecular association reaction independently. Figure 2.5 (A - F) shows the sensorgrams

Figure 2.5. Resolution of two modes of duplex formation on sensor surface using quadruplex-coupled (QC) hybridization model as shown in (G). Relationship between $D_e$ and $D_t$, as percentage contribution towards total duplex formation as a function of strand concentration is shown in (H). All experiments were done in 150 mM NaCl as described in Figure 2.2
obtained in Na⁺ at increasing C1 concentration (in mobile phase) along with the theoretically obtained simulations for Dₑ and Dᵳ. Experiments in Na⁺ are shown as the higher observed amplitude (equilibrium RU) relative to hybridization in K⁺ gives more clarity to our observations; similar results were obtained with K⁺ also. Dₑ denotes duplex formation due to hybridization of C1 with the pre-existing unfolded G1B (as a result of the quadruplex folding/unfolding equilibrium on sensor surface) and Dᵳ denotes association with G1B molecules, which unfold under the influence of the complementary strand during injection. In all cases, as expected, both components together constitute the fitted curve shown in red. At low strand concentration, hybridization with pre-existing unfolded molecules (Dₑ) on the sensor surface was slow and the major component of association was apparently from the molecules, which unfold during the injection time (Dᵳ). At higher strand concentrations (above 64 nM) we observed a trend reversal – higher component of the association was from a very fast saturating hybridization with pre-equilibrated unfolded oligonucleotides.

These observations can be explained by considering two competing equilibria on the sensor surface - intramolecular quadruplex folding/unfolding (kᵤ/kᵢ) and bimolecular hybridization (kᵤ/kᵢ). As expected for such a coupled reaction, a plot of kᵤ (G1B unfolding) and kₒₚₛ ((kᵢC + kᵢ), where C is analyte concentration) of hybridization versus strand concentration (Figure 2.5G), generated using kinetically extracted parameters, showed that kᵤ remained constant whereas kₒₚₛ increased linearly with the concentration of analyte. Thus at low strand concentration, rate of duplex formation was slower. The relatively fast unfolding kinetics in this case contributes more to the overall duplex formation as exemplified by a higher Dᵳ component in Figure 2.5A, 2.5B and 2.5C. On the other hand, it also implies that at low C1 concentration the association reaction would follow a simple 1:1 Langmuir model as hybridization is the rate-limiting step. We tested this possibility using very low C1 concentration (1 – 16 nM) and all the sensorgrams could be fitted using simple
association kinetics (Figure 2.6), without considering the coupled model. With increase in strand concentration, rate of hybridization increases and duplex formation with the pre-equilibrated unfolded form \( (D_e) \) was predominant (Figure 2.5D, 2.5E and 2.5F). A sharp transition between \( D_e \) and \( D_t \) profiles was observed above 64 nM (Figure 2.5C). This is shown in Figure 2.5H, where the contribution of \( D_e \) and \( D_t \) components as a percentage of total duplex formation was plotted versus strand concentration.

2.4.3 G-quadruplex formation is Kinetically Favored at Low Duplex Concentration

It was observed that the promoter element regulating c-MYC is sensitive to S1 nuclease cleavage designating it a nuclease hypersensitive element (NHE III) (88,115). It has been speculated that G-quadruplex formation in this region might be the reason for nuclease sensitivity (92). Using our model we attempted to explore whether the existence of a G-quadruplex was kinetically feasible. Figure 2.5G and Figure 2.5H together indicated a change in the rate-determining step in the range of

![Figure 2.6](image-url)

**Figure 2.6** Hybridization at very low complementary strand concentration follows 1:1 Langmuir association kinetics. C1 (1, 2, 4, 8 and 16 nM) was used as analyte with G1B immobilized on sensor surface. Sensorgrams (black line) were fitted (red line) using Langmuir 1:1 model (BIASvaluation 3.1); the fitted parameters were \( k_a \) and \( k_d \) of 2.45 x 10^5 M^{-1}s^{-1} (± 1.2 %) and 4.9 x 10^5 s^{-1} (± 1.98%), respectively. All experimental conditions were done in 150 mM Na^+ and 25 °C.
30 -100 nM C1 concentration suggesting that below 100 nM, duplex formation was kinetically unfavorable. An estimate of the concentration of immobilized G1B indicates it to be around ~82.6 nM (surface concentration calculation was done assuming monolayer formation on sensor). This suggests that the cross-over in the rate-determining step occurred in the region of equimolar strand concentration.

2.5 Discussion

Recent evidence directly implicates quadruplexes in various biological processes including regulation of the oncogene c-MYC (6,13,92,96,97,109). We recently observed that mutations that increase c-MYC expression affect quadruplex/duplex competition in the promoter region of c-MYC by enhancing duplex formation in vitro (100). However, though many reports have studied thermodynamics of G-quadruplex formation, only a few have studied the kinetics (57,99). In biological context, it would be of interest to observe both folding/unfolding of quadruplex and hybridization simultaneously, in the presence of low and equimolar concentration of both strands. Herein, we report the development of a SPR-based method, which allows resolution of two simultaneous and competing equilibria at very low concentrations. Using the sequence from the NHE III in the promoter region of c-MYC we demonstrate that at low equimolar strand concentration duplex formation is unfavorable.

While this work was in progress, a coupled-hybridization model was reported, which studied folding versus hybridization of the quadruplex formed by the telomeric repeat (TTAGGG)₄ (57). They observed using a series of concentrations that simple two-state hybridization could not simulate the observed sensorgrams when the immobilized molecules could adopt secondary structure, indicating the presence of an additional equilibrium. However, the possibility of a second equilibrium arising out of triplex formation at high analyte strand concentration, which could contribute to a non-two-state model was not ruled out. The authors applied their model to equilibrium conditions and showed that at low and equimolar strand concentration quadruplex formation is favored over hybridization. We have used an alternative approach from first principles to derive a quadruplex-coupled model and applied this to demonstrate that quadruplex-duplex competition can be observed without considering the equilibrium
approximations or the equimolar conditions (Figure 2.3), as done in the previous study. Our general solution clearly shows that quadruplex-duplex competition can be controlled by complementary strand concentration, which changes the rate determining steps involved in quadruplex-folding vis-à-vis hybridization. This results in a mechanistic change with increasing complementary strand concentration and the ‘cross-over’ region is at near equimolar concentration.

The quadruplex folding/ unfolding constants obtained by us are within the same order of magnitude as observed before for telomeric sequences from human (57,99) and Oxytricha (54) (Table 2.3). It was interesting to note that in the presence of K+, half-life of folded c-MYC quadruplex \( t_{1/2} = 87.7 \text{ sec} \) was six-fold less than that of the telomeric sequence \( t_{1/2} = 533 \text{ sec} \) (57) suggesting that the c-MYC quadruplex was thermodynamically unfavorable relative to the telomeric quadruplex. The contrasting characteristics are interesting, considering the fact that the number of ‘G-tetrad’ units (which are believed to impart stability to the quadruplex moiety) are three in the telomeric quadruplex in comparison to four possible in the c-MYC quadruplex. Thus factors other than tetrad stability, e.g. loop constitution, may be important in rendering stability to the quadruplex moiety (113).

We also noted the hybridization rates observed for several sequences using SPR and compared them with the ones obtained by our 31-mer sequence (Table 2.3). The association constant reported by Zhao et al. (57), for the telomere sequence studied under analogous conditions (in K+) was higher by an order of magnitude. The reason for this substantial difference is not very clear to us. However, some discrepancy may result in the comparison of the parameters as ~75% standard error was associated with the fitted dissociation constant, which was used for fitting the association curve in the sensorgram, in the previous study (57,100). In an earlier study (100), we determined hybridization constants for G1 and C1 at pH 6.6 and obtained \( k_a = 3.2 \times 10^4 \text{ M}^{-1}\text{s}^{-1} \) and \( k_d = 3.5 \times 10^{-3} \text{ s}^{-1} \), which are an order of magnitude different from the ones observed here. A difference in terms of lower rates is expected at pH 6.6 relative to pH 7.4. However, the QC model was not considered in the earlier case, which may be the reason for observed errors in fitting (16- 23 %) and the large difference in rates observed in comparison with the current
study. Several previous reports have determined folding/unfolding constants ($K_F$) of quadruplexes by UV-melting and compared or used them for extraction of other kinetic parameters (57,99,107). The high melting point of G1 under our conditions (more than 92 °C) precluded this.

It must be noted that multiple folding/unfolding rates may result from the presence of more than one folded motif on the sensor surface (as evident from Figure 2.2F). Therefore the reported kinetic folding/unfolding parameters are likely to represent average apparent values. One of the limitations of fitting multiple parameters to a single equation (as in equation (2.17)) is that it could potentially give several minima, i.e., several sets of optimal values for the parameters, which equally fit the equation. While such a possibility cannot be completely ruled out, typically, in such cases when optimization of all parameters is carried out simultaneously, large standard deviations result. We further checked the effect of each parameter on the fitting by perturbing (both increase and decrease) each parameter at a time and observed its effect on the other variables. Chi-square values were observed to progressively increase indicating deviation from the correct solution in all cases (Table 2.3). Additionally, in our case all folded forms are likely to result from the

<table>
<thead>
<tr>
<th>Change in $k_u$</th>
<th>Chi(Sq)</th>
<th>Change in $k_f$</th>
<th>Chi(Sq)</th>
<th>Change in $K_A$</th>
<th>Chi(Sq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5X$^b$</td>
<td>10.2</td>
<td>0.5X</td>
<td>7.4</td>
<td>0.5X</td>
<td>38.9</td>
</tr>
<tr>
<td>0.67X</td>
<td>5.5</td>
<td>0.67X</td>
<td>2.5</td>
<td>0.67X</td>
<td>14.5</td>
</tr>
<tr>
<td>0.83X</td>
<td>2.6</td>
<td>0.83X</td>
<td>2.2</td>
<td>0.83X</td>
<td>4.2</td>
</tr>
<tr>
<td>1X</td>
<td>1.7</td>
<td>1X</td>
<td>1.7</td>
<td>1X</td>
<td>1.7</td>
</tr>
<tr>
<td>1.2X</td>
<td>3.1</td>
<td>1.2X</td>
<td>2.2</td>
<td>1.2X</td>
<td>3.9</td>
</tr>
<tr>
<td>1.5X</td>
<td>37.6</td>
<td>1.5X</td>
<td>4.4</td>
<td>1.5X</td>
<td>10.4</td>
</tr>
<tr>
<td>2X</td>
<td>40.7</td>
<td>2X</td>
<td>9.5</td>
<td>2X</td>
<td>20.3</td>
</tr>
</tbody>
</table>

$^a$ Representative values from fitting observed with 512 nM C1 in mobile phase with G1 immobilized. $^b$ X indicates times changed
parallel form of the motif (as observed in the CD spectra in Figure 2.1 and high $T_m$ in UV melting experiments). Therefore the difference between individual kinetic rates may not be too high, which may be a reason for observed low standard errors for the parameters.

Our analysis with the c-MYC sequence suggests that at the low intracellular concentrations (lower than $10^{-11}$ M (57)), the chances of favorable quadruplex formation cannot be ruled out though it is thermodynamically more expensive than duplex formation. This is interesting in the context of previous reports, which have observed that the promoter element of c-MYC harboring this sequence exists in a strand-separated form in vivo and has been designated as a nuclease hypersensitive element for this reason (88,115). However, it may not be true for all potential quadruplex-forming sequences as the stability of the motif ($k_u/k_f$ equilibrium) will play a significant role in the competition. Analytical component resolution ($D_e$ and $D_t$ modes of hybridization, Figure 2.3) is important with respect to chromosomal sequence where unfolding of the quadruplex is mostly in presence of the complementary strand, which is represented by the $D_t$ component. Thus $D_t$ is expected to more appropriately represent in vivo situations allowing extrapolations from in vitro studies whereby $D_t$ may be used to examine the effect of ligands and other factors on the structural transitions in the context of triplex formation or G-quadruplex stabilizing molecules. However, it must be noted that the short immobilized oligonucleotide on the sensor surface and the complementary strand in the mobile phase do not effectively replicate an in vivo situation wherein the duplex state of the flanking regions are bound to have an effect on the kinetic parameters. It may be contemplated that a ‘zipping-like’ mechanism may help duplex formation. On the other hand, intracellular molecules may help stabilization of the quadruplex.

Proposed models of NHE III, controlled c-MYC expression invoke G-quadruplex formation in the NHE as a negative regulator (92,110). This is primarily based on observations that a single base mutation, which destabilizes the quadruplex, increases c-MYC expression while stabilization of the motif decreases c-MYC expression (92). The paranemic quadruplex form is converted to unstructured single
strand form before c-MYC activation – possibly by intervention of the transcription factor NM23-H2 (116), which may bind to both forms of the NHE. CNBP and hnRNP K are also known to play a role in c-MYC transcriptional activation by binding to the purine- and pyrimidine-rich strands of the duplex NHE directly (89,117,118). These collectively suggest that both the folded and the unfolded form of the NHE are significant components of c-MYC transcription control, which may be orchestrated by presentation of different molecular topology as transcription factor binding sites. The intrinsic properties of these topologies conferring different molecular recognition properties vis-à-vis duplex DNA make them attractive targets for selectively intervening oncogene expression.