ARTICLE

Dominant driving forces in human telomere quadruplex binding-induced structural alterations

Matjaž Bončina,[a] Florian Hamon,[b] Barira Islam,[c] Marie-Paule Teulade-Fichou,[b] Gorazd Vesnaver,[a] Shozeb Haider,[c,d] and Jurij Lah[a,]*

[a] Faculty of Chemistry and Chemical Technology, University of Ljubljana, Večna pot 113, SI-1000 Ljubljana, Slovenia.
[b] Institut Curie, CNRS UMR-176, Centre Universitaire d'Orsay, Paris-Sud 91405 Orsay Cedex France.
[c] Centre for Cancer Research and Cell Biology, Queen's University of Belfast, Belfast BT9 7BL, UK.
[d] UCL School of Pharmacy 29-39 Brunswick Square, Bloomsbury, London WC1N 1AX, UK.
[*] Corresponding Author: Jurij Lah, Faculty of Chemistry and Chemical Technology, University of Ljubljana, SI–1000 Ljubljana, Slovenia; E-mail: jurij.lah@fkkt.uni-lj.si.

Abstract

Recently various pathways of human telomere (ht) DNA folding into G-quadruplexes and of ligand binding to these structures have been proposed. However, the key issue as to the nature of forces driving the folding and recognition processes remains unanswered. In this study structural changes of 22 mer ht-DNA fragment (Tel22), induced by binding of ions (K^+, Na^+) and specific bisquinolinium ligands, were monitored by calorimetric and spectroscopic methods and by gel electrophoresis. Using the global model analysis of a wide variety of experimental data we were able to characterize the thermodynamic forces that govern the formation of stable Tel22 G-quadruplexes, folding intermediates and ligand-quadruplex complexes and to predict Tel22 behavior as a function of temperature, salt concentration and ligand concentration. On the basis of the above, we believe that the presented work sets the framework for better understanding the heterogeneity of ht-DNA folding and binding pathways and its structural polymorphism.

Keywords: Thermodynamics of G-quadruplex folding, Thermodynamics of G-quadruplex ligand binding, Dissection of energetics, G-quadruplex specific ligands, ITC, DSC

Introduction

Guanine-rich DNA sequences in the presence of cations can fold into four-stranded structures called G-quadruplexes. The existence of potential quadruplex sequences in key regions of the eukaryotic genome, including the immunoglobulin heavy chain switch region, promoter regions, ribosomal DNA, oncogenes and telomeres, suggests that they may play an important role in the mechanism and control of several cellular processes (1,2,3). Therefore, G-quadruplexes are relevant targets of small molecules that can potentially modulate their biological functions, gene expression, and protein synthesis (4,5).

Quadruplex topologies may differ in glycosidic bond angles, strand orientation, connecting loop regions, and molecularity leading to conformational heterogeneity of G-quadruplex structures. This is well exemplified by guanine-rich human telomeric (ht) repeat sequences, which are capable of adopting multiple topologies. For example, monomeric human

1

telomeric (ht) quadruplexes containing the core sequence d(AGGG(TTAGGG)$_3$) (Tel22) can adopt several distinct quadruplex topologies. X-ray crystallography reveals that in the presence of K$^+$ ions, Tel22 shows all-parallel strand orientation (6) while in K$^+$ solutions it adopts, according to NMR and other biophysical techniques, a (3+1) hybrid-type topology (denoted as H$_K$+) (7,8,9,10). By contrast, in Na$^+$ solutions Tel22 adopts a conformation with antiparallel strand orientation (denoted as A$_{Na}$+) (11). As shown recently, binding of ligands capable of inducing binding-coupled conformational transitions of G-quadruplexes may be additional cause of the observed ht-quadruplex polymorphism (12,13,14,15,16).

Detailed computational and experimental (thermodynamic, kinetic) investigation of folding and ligand binding pathways of quadruplex DNA has begun very recently (17,18,19,20,21,22). These studies leave no doubt that in solution the Tel22 folding/unfolding process may consist of several structural steps meaning that such solutions should be considered as equilibrium mixtures of Tel22 molecules in their folded (Q), unfolded (U) and intermediate (I) states. Since in the reported studies of ligand binding to Tel22, as a rule, the involvement of Tel22 folding intermediates has been neglected one may expect the reported thermodynamics of binding to be incomplete. Thus, the aim of this work was besides investigating the nature and the relative importance of forces that govern folding of Tel22 and its recognition by ligands, to estimate how important in controlling these binding events may be the involvement of the Tel22 intermediates. We believe that such approach will lead to an improvement in our understanding of the heterogeneity observed in the ht-quadruplex folding and binding processes and of the complex interplay between the enthalpic and entropic contributions to the free energy changes accompanying these events.

Materials and Methods

Sample Preparation

HPLC pure oligonucleotide 5'-AGGGTTAGGGTTAGGGTTAGGG-3' (Tel22) was obtained from Midland Co., U.S.A. The buffer solutions used in our experiments consisted of 20 mM cacodylic acid, 1 mM EDTA and various concentrations of Na$^+$ or K$^+$ ions. NaOH (KOH) was added to cacodylic acid to reach pH = 6.9. Then, NaCl (KCl) was added to obtain the desired concentration of Na$^+$ (K$^+$) ions (100 or 200 mM Na$^+$ (K$^+$) in cacodylic buffer). DNA was first dissolved in water and then extensively dialyzed against the buffer using a dialysis tube Float-A-Lyser (Spectrum Laboratories, USA, M_w cutoff 500 - 1000 Da). The starting solution of oligonucleotide was first heated up to 95 °C in an outer thermostat for 5 min to make sure that all DNA transforms into the unfolded form, cooled down to 5 °C at the cooling rate of 0.05 °C min^{-1} to allow DNA to adopt quadruplex structure(s) and then used in the experiments. Concentration of the DNA in the buffer solution was determined at 25 °C spectrophotometrically using Cary 100 BIO UV/Visible Spectrophotometer (Varian Inc.) equipped with a thermoelectric temperature controller. Tel22 concentrations at 25 °C were obtained from the melting curves monitored at wavelength λ = 260 nm. For the extinction coefficient of Tel22 unfolded form at 25 °C we used the value ε_{260} = 228500 M^{-1} cm^{-1} estimated from the nearest-neighbor data of Cantor et al (23).

The two bisquinolinium derivatives, Phen-DC3 (M = 848.12 g mol^{-1}; Fig. S1 in the Supporting Information), and 360A-Br (M = 780.90 g mol^{-1}; Fig. S1), are poorly soluble in aqueous solution, therefore they were first dissolved in DMSO and then transferred into buffer solution. The lowest possible content of the DMSO in the buffer solution sufficient for ligand solubility (μM range) was 3 %. Therefore, all the ligand and Tel22 were dissolved in buffer solution with 3% DMSO. Concentrations of the ligands were determined by measuring absorbance at 25 °C ($\varepsilon_{Phen-DC3, 350\ nm}$ = 6200 M^{-1} cm^{-1} and $\varepsilon_{360A-Br, 370\ nm}$ = 5980 M^{-1} cm^{-1}).

CD Spectroscopy

CD spectra of DNA and ligand-DNA complexes were recorded at 25 °C in a 1.0 cm cuvette in the wavelength range between 215 and 350 nm. CD titrations were conducted at 25 °C by titrating DNA solution ($c_{DNA} \approx$ 70 μM) into a 600 μL ligand solution ($c_L \approx$ 20 μM). Ellipticity, Θ,

was measured at 293 nm in a 1.0 cm cuvette with signal averaging time of 30 s and 5 nm bandwidth. Temperature dependence of CD spectra of Tel22 was collected between 215 and 320 nm in a 0.25 mm cuvette with a signal averaging time of 10 s and 5 nm bandwidth. Experiments were performed using AVIV CD Spectrophotometer 62A DS (Aviv Biomedical, Lakewood, NJ, USA) equipped with a thermoelectric temperature controller.

Fluorimetry (FL)

FL titrations were conducted at 25 °C by titrating DNA solution ($c_{DNA} \approx 80$ µM) into a 1000 µL 360A-Br ligand solution ($c_L \approx 10$ µM). Emission spectra were recorded between 340 – 570 nm (excitation wavelength $\lambda_{ex} = 330$ nm) in a 1.0 cm cuvette with scanning speed of 50 nm min^{-1}. Phen-DC3 exhibits very weak induced fluorescence, therefore we were unable to obtain reliable experimental data for further model analysis of its binding. Experiments were performed using Perkin Elmer LS 55 luminescence spectrometer (Perkin Elmer, Waltham, MA, USA) equipped with a thermally controlled cell holder.

Isothermal Titration Calorimetry (ITC)

ITC experiments were performed between 15 and 35 °C by titrating a solution of DNA ($c_{DNA} \approx 80$ µM) into a ligand solution ($c_L \approx 10$ µM, $V = 1.386$ mL) using a VP-ITC isothermal titration calorimeter from Microcal Inc. (Northampton, MA, USA). The area under the peak following each injection of DNA solution was obtained by integration of the raw signal, corrected for the corresponding heat of dilution (blank titration of DNA into buffer solution) and expressed per mole of added DNA per injection, to give the enthalpy of interaction, (ΔH_T).

Differential Scanning Calorimetry (DSC)

DSC experiments were performed at the DNA concentration of about 0.3 mM in the temperature range between 1 and 95 °C at the heating and cooling rate of 1.0 °C min^{-1}. The corresponding baseline (buffer−buffer) thermograms were subtracted from the heating thermograms and the obtained differences were normalized to 1 mole of DNA to obtain the partial molar heat capacity of DNA as a function of temperature. Data were analyzed in the same way as described in ref 24. DSC experiments were performed using Nano DSC instrument (TA Instruments, New Castle, DE, U.S.).

Results and Discussion
Thermodynamic analysis of binding-induced structural transitions
We investigated thermodynamics of Tel22 structural transitions in solution at various temperatures, concentrations of K$^+$ or Na$^+$ ions and in the presence or absence of bisquinolinium ligands Phen-DC3 and 360A-Br (Fig. S1), known for their high affinity for G-quartets most likely due to strong π-π stacking interactions (25,26,27,28). The analysis of calorimetric (DSC, ITC) and spectroscopic (CD, FL) data obtained in solutions with K$^+$ (Figs. 1, S5, S7) or Na$^+$ ions (Figs. S3, S6, S8) suggests that the observed binding-coupled folding processes may be described by the model mechanism that involves five macroscopic states (see Scheme 1). Reversibility of folding/unfolding of Tel22 in the absence of ligands and in the presence of K$^+$ or Na$^+$ ions (U $\leftrightarrow$ I $\leftrightarrow$ Q; Fig. S3) was verified by DSC while the reversibility of ligand binding (Q + L $\leftrightarrow$ Q'L + L $\leftrightarrow$ Q''L$_2$) was verified by gel electrophoresis (Fig. S11).
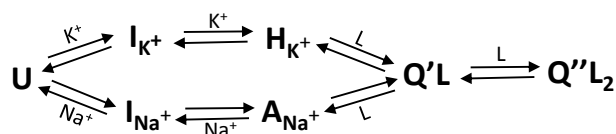
SCHEME 1 Mechanism of Tel22 structural transitions induced by binding of K$^+$ or Na$^+$ ions and bisquinolinium ligand (L = Phen-DC3 or 360A-Br) consistent with all the presented experimental data.

For a given model mechanism (Scheme 1) one can write a general equation for the spectroscopic and calorimetric properties of the ligand-DNA solution, measured in the presence of K^+ or Na^+ ions (Eq.1), from which the corresponding contributions of the buffer are subtracted. Such property, X, can be presented as a linear combination of molar ratios, $\alpha_j = c_j/c_{DNA}$ of all species, j, predicted by the model mechanism (j = L, U, I, Q, Q'L, Q''L$_2$).

$$X = X_L\alpha_L + X_U\alpha_U + X_I\alpha_I + X_Q\alpha_Q + X_{Q'L}\alpha_{Q'L} + X_{Q''L_2}\alpha_{Q''L_2}$$

(1)

In Eq. 1 X_j represents the property of the solute j at a given pressure, temperature, T, salt type (KCl, NaCl), salt concentration and the total DNA concentration, c_{DNA}. Since according to the model $\alpha_U + \alpha_I + \alpha_Q + \alpha_{Q'L} + \alpha_{Q''L_2} = 1$ and $\alpha_L + \alpha_{Q'L} + 2\alpha_{Q''L_2} = r$ where $r = c_{L,tot}/c_{DNA} = n_{L,tot}/n_{DNA}$ ($c_{L,tot}$ is the total ligand concentration) one obtains

$$X = X_L r + (\Delta X_{IU} + \Delta X_{QI})\alpha_U + \Delta X_{QI}\alpha_I + X_Q + \Delta X_{Q'L}\alpha_{Q'L} + (\Delta X_{Q'L} + \Delta X_{Q''L_2})\alpha_{Q''L_2}$$

(2)

where the changes $\Delta X_{IU} = X_U - X_I$, $\Delta X_{QI} = X_I - X_Q$, $\Delta X_{Q'L} = X_{Q'L} - X_Q - X_L$, $\Delta X_{Q''L_2} = X_{Q''L_2} - X_{Q'L} - X_L$ refer to each step in the model mechanism presented in Scheme 1.

From eq. 2 are derived various model functions (SI Eqs. 3-6, 9) expressed in terms of a set of adjustable parameters that describe the CD ($X = [\Theta] = \Theta/(c_{DNA}l)$ = normalized ellipticity at given λ), FL ($X = [F] = I_F/(c_{DNA}l)$ = normalized emitted fluorescence measured at given λ) and DSC and ITC ($X = H$ = enthalpy of solution per mol of DNA). Each step in the suggested mechanism is described in terms of the corresponding changes of three standard thermodynamic parameters that are independent on the Na^+ or K^+ concentration. Two of them, $\Delta G^o_{i(T_0)}$ and $\Delta H^o_{i(T_0)}$, depend on temperature and are thus determined at the reference temperature $T_0 = 298.15$ K while the third parameter, $\Delta C^o_{P,i}$ is assumed to be temperature independent. These three parameters define the standard free energy and enthalpy of folding or binding, $\Delta G^o_{i(T)}$ and $\Delta H^o_{i(T)}$, at any T through the Gibbs-Helmholtz relation $[\partial(\Delta G^o_{i(T)}/T)/\partial T]_P = -\Delta H^o_{i(T)}/T^2$ and the Kirchhoff's law $[\partial\Delta H^o_{i(T)}/\partial T]_P = \Delta C^o_{P,i}$. Usually, each step is described in terms of the apparent $\Delta G^o_{i(T,X^+)}$ (X^+ stands for Na^+ or K^+) which depends on the salt concentration. Its relation with the true thermodynamic $\Delta G^o_{i(T)}$ is given by

$$\Delta G^o_{i(T,X^+)} = \Delta G^o_{i(T)} + \Delta n_i RT\ln[X^+] =$$
$$= \Delta G^o_{i(T_0)}T/T_0 + \Delta H^o_{i(T_0)}[1 - T/T_0] + \Delta C^o_{P,i}[T - T_0 - T\ln(T/T_0)] + \Delta n_i RT\ln[X^+]$$

(3)

where parameter Δn_i represents the number of ions released or uptaken in the transition step i and is assumed to be independent on T (29). Note that equilibrium concentration of unbound X^+, $[X^+]$, appearing in Eq. 3 is normalized to 1 M concentration in the reference (standard) state. Four thermodynamic parameters ($\Delta G^o_{i(T_0)}$, $\Delta H^o_{i(T_0)}$, $\Delta C^o_{P,i}$, Δn_i for each step i in the suggested mechanism) define each equilibrium constant, $K_{i(T,X^+)} = \exp(-\Delta G^o_{i(T,X^+)}/RT)$, appearing in the proposed model presented in Scheme 1. In other words, 16 parameters specify the populations of species U, I, Q, Q'L, Q''L$_2$ in the solution at any T, ligand, salt and DNA concentration ($K_{i(T,X^+)} = f(\alpha_{j(T,X^+)}; j = I, Q, Q'L, Q''L_2), \sum_j \alpha_{j(T,X^+,r,c_{DNA})} = 1$) and consequently also the

corresponding CD, FL, ITC and DSC model functions (right-hand side of SI Eqs. 3, 4, 5, 6, 9). Global fitting of the model functions to the experimental CD, FL, ITC and DSC data was based on the non-linear Levenberg-Marquardt χ^2 regression procedure. The first step in the global fitting procedure was a semi-global model analysis of DSC and CD folding/unfolding data of the ligand-free Tel22 (U $\leftrightarrow$ I $\leftrightarrow$ Q equilibrium) in the presence of K$^+$ (Fig. 1a,b) or Na$^+$ (Fig. S3) ions in which 8 adjustable parameters were used (Table S4). In the second step, these parameters were used in the description of the binding isotherms (Figs. 1c,d, S5-S8) as fixed values. Moreover, due to high correlation between the adjustable parameters describing the Q'L + L $\leftrightarrow$ Q''L$_2$ step we were forced to reduce the no. of the adjustable parameters. In this light we assumed that the heat capacity and the number of ions released is the same for Q + L $\leftrightarrow$ Q'L and Q'L + L $\leftrightarrow$ Q''L$_2$ step (Table S2). This assumption may be supported by the observed structural features of a number of (aromatic) ligand-quadruplex complexes in which the interacting surfaces of the first and the second aromatic ligand molecule, bound at the opposite ends of the G-quadruplex, are similar (refs XX). Similar surface areas buried upon binding of the ligand in each step should results in similar changes in heat capacities, and the accompanying number of released ions should be, due to similar electrostatic interactions, approximately the same. Taken together, 9 adjustable parameters were used in the fitting procedure in the case of Phen-DC3 (ITC and CD data; Table S5) and 11 in the case of 360A-Br (ITC, CD and FL data).

In the absence of ligand, the folding/unfolding steps were described as a three-state process involving U, I and Q (24). In K$^+$ or Na$^+$ solutions I can be considered to be a mixture of so called G-triplex conformations (I$_{K^+}$ or I$_{Na^+}$) (19,30), which according to the measured CD spectra exhibit structural properties similar to H$_{K^+}$ or A$_{Na^+}$ (Fig. 2). Each stage of folding is in the presence of K$^+$ [U $\rightarrow$ I$_{K^+}$ $\rightarrow$ H$_{K^+}$] and Na$^+$ [U $\rightarrow$ I$_{Na^+}$ $\rightarrow$ A$_{Na^+}$], characterized by an extensive enthalpy-entropy compensation (31) (Fig. 3, Table S1) and a large negative change in the heat capacity accompanying the first transition step [U $\rightarrow$ I, $\Delta C°_P \approx$ -400 cal mol^{-1} K^{-1}]. These thermodynamic parameters are comparable with those reported for the thrombin binding aptamer folding/unfolding transition (32). Q is more thermodynamically stable ($\Delta G°_{U \rightarrow Q}$ is lower) in solutions with K$^+$ than with Na$^+$ ions, which is a general characteristics of the G-quadruplex stability (33,34).

Binding of Phen-DC3 and 360A-Br to Tel22 is successfully described by the sequential binding mechanism Q + L $\leftrightarrow$ Q'L + L $\leftrightarrow$ Q''L$_2$ (Figs. 1, S5-S8) which assumes that Tel22 unfolded and intermediate states contain no binding sites. All the ligand binding experiments were performed at conditions at which the model analysis of Tel22 melting curves (Figs. 1, S3) predicts the presence of the folding intermediates I. The fact that the global analysis of the measured binding data is appropriate (good quality of fit, reasonable values of thermodynamic parameters) only when the model predicted population of I is taken into account, supports the suggested linkage between the folding and binding processes (Scheme 1, Figs. 1a,b, S5-S8). It should be noted that this linkage could be better supported by thermal unfolding experiments performed with ligand-DNA complexes using DSC or CD (35). Our attempts to perform such experiments and the corresponding data analysis were not successful. Namely, the detection of influence of ligand binding on melting transitions, monitored by DSC and CD spectroscopy, requires ligand concentrations that are much higher than those used in our titration experiments.

In other words, due to the low solubility of ligands in 100 mM K$^+$ or Na$^+$ solutions the melting experiments cannot be performed at "reversible" conditions. Moreover, aggregation of ligand-G-quadruplex complexes prevents adequate thermodynamic analysis of titration experiments performed at very low K$^+$ or Na$^+$ concentrations where the model analysis predicts the presence of U and I.

Taken together, Scheme 1 represents the simplest model (Model 1) of ion and ligand binding-induced structural alterations in the presence of K$^+$ and Na$^+$ ions, consistent with all

the presented experimental data, whose fitting gives reliable values of thermodynamic parameters. To provide evidence that slightly simpler models fit worse the data we present in Fig. S4 some characteristics of the best global fit of the U ↔ Q + L ↔ Q'L + L ↔ Q''L$_2$ model (Model 2), which assumes that the state I is not populated, and the model that takes into account the U ↔ I ↔ Q equilibrium and assumes that L binds to two equivalent independent binding sites on Q (Model 3). Fig. S4 shows that the Model 2 cannot describe the DSC thermograms and ITC data measured at 35 °C, while the Model 3 fails to describe CD titration data. On the other hand, more complex models involve too many adjustable parameters that are highly correlated and thus cannot be determined with sufficient accuracy. Our analysis emphasizes an important advantage of the global fitting over the traditional fitting of the model to limited datasets (29). For example, ITC data alone (measured at T < 30 °C) can be successfully described by the simplified Model 3 (no I present in the solution), however, according to other available experimental data (DSC, CD titration), such analysis results in thermodynamic binding parameters that have no physical meaning.
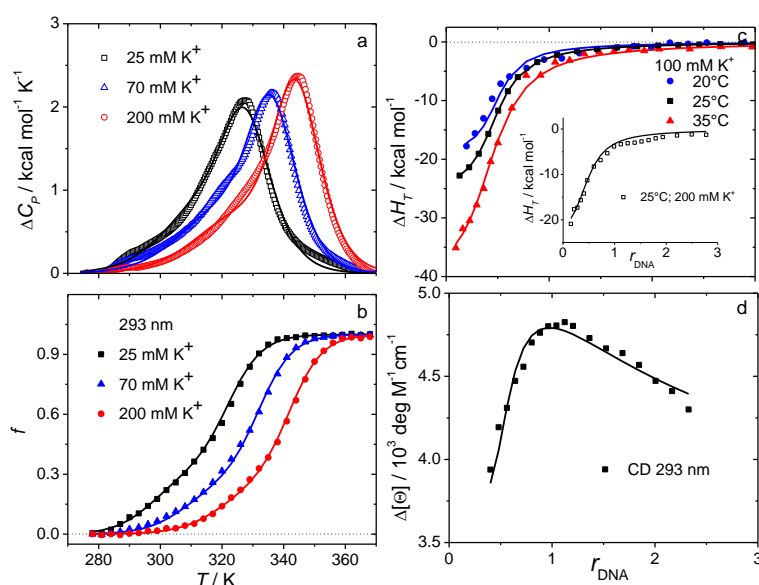

FIGURE 1 Model analysis of experimental data. Best-fit global model functions (lines) show good agreement with experimental data (symbols represent unfolding data obtained by DSC and CD spectroscopy (a, b), and ligand (Phen-DC3) binding data obtained by ITC and CD spectroscopy (c, d) measured as a function of temperature, K$^+$ ion concentration and DNA/ligand molar ratio $r_{DNA} = c_{DNA}/c_{L,tot}$ (see SI for details).

Thermodynamics and structural features

CD spectra (Figs. 2, S9) suggest for both ligands (Phen-DC3, 360A-Br) that their binding is accompanied by quadruplex conformational changes and that the resulting complexes (Q'L, Q''L$_2$) have similar structures in solutions with K$^+$ and Na$^+$ ions. CD spectra of Q'L and Q''L$_2$ complexes show characteristics of A$_{Na+}$ spectrum (15). Interestingly, difference CD spectra (Fig. S10) corresponding to Q + L ↔ Q'L and Q'L + L ↔ Q''L$_2$ binding events are almost mirror images, suggesting that binding of the first ligand molecule to one end of the quadruplex induces changes in CD spectrum that are opposite to those induced by binding of the second ligand molecule to the other end of the quadruplex.

The influence of the bound ligand on the conformation of the folded quadruplexes was examined also by gel electrophoresis experiments, which show that quadruplexes complexed with the ligand are electrophoretically faster than the ligand-free quadruplexes (Fig. S11). According to our model analysis, binding of dicationic ligand displaces only about

one (nonspecifically) bound cation, which means that the net (negative) charge of the ligand-quadruplex complexes is lower than the one of the ligand-free quadruplexes. Thus, if ligand binding to the quadruplex is a rigid-body association the surface net charge of the ligand-quadruplex complex should be lower than that of the ligand-free quadruplex, and consequently the studied complexes should exhibit lower gel-mobility than the ligand-free quadruplexes (ref XX). By contrast, our results (Fig. S11) show an opposite effect, that is, an increase in mobility of the ligand-quadruplex complexes which may be, in principle, ascribed to increased hydrodynamical compactness and/or increased surface net charge of the ligand-quadruplex complexes. Since both of the two effects can arise from the increased compactness of the ligand-quadruplex complex (in spite of lower net (negative) charge its surface net charge can increase), we believe that the observed increased mobility of the complexes results, very likely, from the ligand-induced conformational changes of the quadruplexes. In addition, electrophoresis results suggest that, Q' and Q'' structures differ from A_{Na}+ and H_K+ and also from the possible all parallel Tel22 quadruplex conformation that appears to be gel-electrophoretically slower than A_{Na}+ and H_K+ (36).

The observed thermodynamic characteristics of ligand binding to Tel22 in K^+ and Na^+ solutions are very similar. For both ligands, the Q binding affinity for the first ligand molecule is higher than for the second one (Table S2). Both steps are enthalpy driven, accompanied by negative change in entropy and heat capacity (Fig. 3, Table S2). This suggests that ligand binding is driven mainly by ligand-quadruplex π - π stacking ($\Delta H° < 0$) and by displacement of water from the ligand-quadruplex binding interface ($\Delta C°_P < 0$). Moreover, for both ligands the overall thermodynamics of binding-coupled folding (U + L → Q'L or U + 2L → Q''L_2) is in the K^+ and Na^+ environment nearly the same (Fig. 3, Table S2). This supports our suggestion that in solutions with either of the two ions the ligand-bound structures are similar. In addition, H + L → Q'L (K^+) and A + L → Q'L (Na^+) (and H + 2L → Q''L_2 and A + 2L → Q''L_2) events are also accompanied by similar energetic contributions (Fig. 3, Table S2). Since according to thermodynamic parameters that characterize unfolding of A and H (Table S1), the A and H conformations are energetically very similar, this observation is consistent with the suggested similarity of the ligand-quadruplex structures in the presence of K^+ and Na^+.
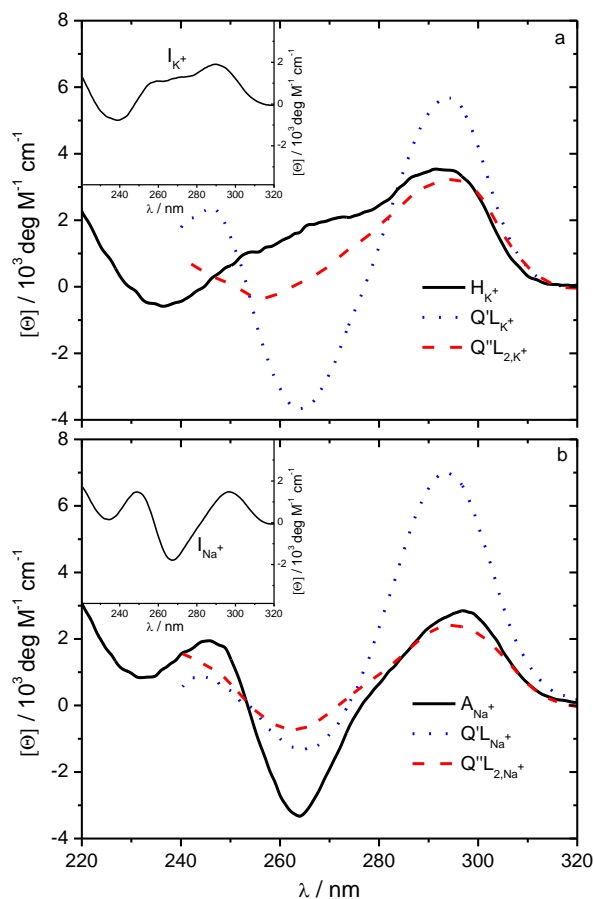
FIGURE 2 Structural features monitored by CD spectroscopy. Spectra corresponding to hybrid (H_K+) and antiparallel (A_{Na}+) structures, complexes with one (Q'L) and two bound ligand (L = Phen-DC3) molecules (Q''L_2) and folding intermediates (I_K+, I_{Na}+, see the inset) at 25 °C in the presence of 100 mM K^+ (panel a) or Na^+ (panel b) ions. CD spectra of intermediates and complexes were estimated by deconvolution of the measured spectra based on the model-predicted populations of species (Figs. S5-S8) and spectrum of Q (A_{Na}+ or H_K+) form.

Driving forces of binding-induced structural alterations

Several recent papers discuss the thermodynamic forces that may control folding of G-quadruplexes (37,38,39). We present here, to the best of our knowledge, the first attempt of quantitative dissection of $\Delta G°$ accompanying the ligand binding-coupled structural transitions of ht-DNA to more fundamental contributions (Fig. 3). Following the additivity approach (40,41,42), $\Delta G°$ can be treated as a sum of the following main contributions:

$$\Delta G° = \Delta G°_{solv} + \Delta G°_{int} + \Delta G°_{rt} + \Delta G°_{conf} \tag{4}$$

in which the first contribution, $\Delta G°_{solv}$, ascribed to the solvation effects (reorganization of water molecules surrounding Tel22, ions and ligands) may be further expressed as $\Delta G°_{solv} = \Delta G°_{ion} + \Delta G°_{hyd}$. The $\Delta G°_{ion}$ contribution reflects the dehydration of K^+ or Na^+ ions accompanying their coordinative binding within Tel22. Its enthalpic and entropic origin has been well characterized by Marcus (43) and results in the estimates of $\Delta G°_{ion,K}+ = 72.6$ kcal mol^{-1} and $\Delta G°_{ion,Na}+ = 89.6$ kcal mol^{-1}. The $\Delta G°_{hyd}$ contribution is ascribed to the desolvation/solvation of Tel22 and ligand molecules and may be interpreted mainly as a hydrophobic contribution to the overall $\Delta G°$ of folding and/or binding (40,41,44). At 25 °C it may be estimated by an empirical relation as $\Delta G°_{hyd} = \Delta C°_P \cdot 80(\pm 10)K$ (45,46), where $\Delta C°_P$ is the corresponding heat capacity change determined by the global model analysis of

experimental data (Fig. 1, SI) combined with the careful treatment of the corresponding baselines (DSC, ITC). The second main contribution, $\Delta G°_{int}$, reflects the specific intra- and inter-molecular interactions (van der Waals (base stacking), H-bonds, Coulombic interactions, cation coordination) that stabilize a particular Tel22 conformation. It can be considered mainly as an enthalpic contribution and estimated as the difference between the measured enthalpy change, $\Delta H°$, and the corresponding enthalpy of ion dehydration, $\Delta H°_{ion}$ (43). Thus, $\Delta G°_{int} \approx \Delta H° - \Delta H°_{ion}$. The third main contribution, $\Delta G°_{rt}$, is considered as an entropic contribution due to changes of rotational and translational freedom of ligand and Tel22 (lost upon ligand binding): $\Delta G°_{rt} = -T\Delta S°_{rt}$. Similarly, the fourth main contribution $\Delta G°_{conf}$ is interpreted as an entropic contribution accompanying the changes of conformational freedom of Tel22, $\Delta G°_{conf} = -T\Delta S°_{conf}$. It may be estimated as: $\Delta G°_{conf} = \Delta G° - \Delta G°_{solv} - \Delta G°_{int} - \Delta G°_{rt}$. At this point we would like to mention that in our dissection of $\Delta G°$ the free energy contribution due to the release of the non-specifically bound cations, that accompanies unfolding of Tel22 and ligand binding to it, has been neglected. Namely, according to the prediction of polyelectrolyte theory (ref xx) this electrostatic contribution is smaller than the estimated error in each of the suggested contributions in Eq. 4.

We are well aware that due to approximations involved in the additivity approach presented above the interpretation of $\Delta G°$ contributions should be taken with great care. It should be emphasized that the main aim of using Eq. 4 is, to show the importance of various types of interactions and conformational changes in the formation of folding intermediates and ligand-quadruplex complexes. To demonstrate this, the $\Delta G°$ contributions do not need to be specified with high precision (see below). In this light, the described dissection of energetics (Fig. 3) enables us to characterize the dominant driving forces involved in the Tel22 binding-coupled structural transitions in the following way:
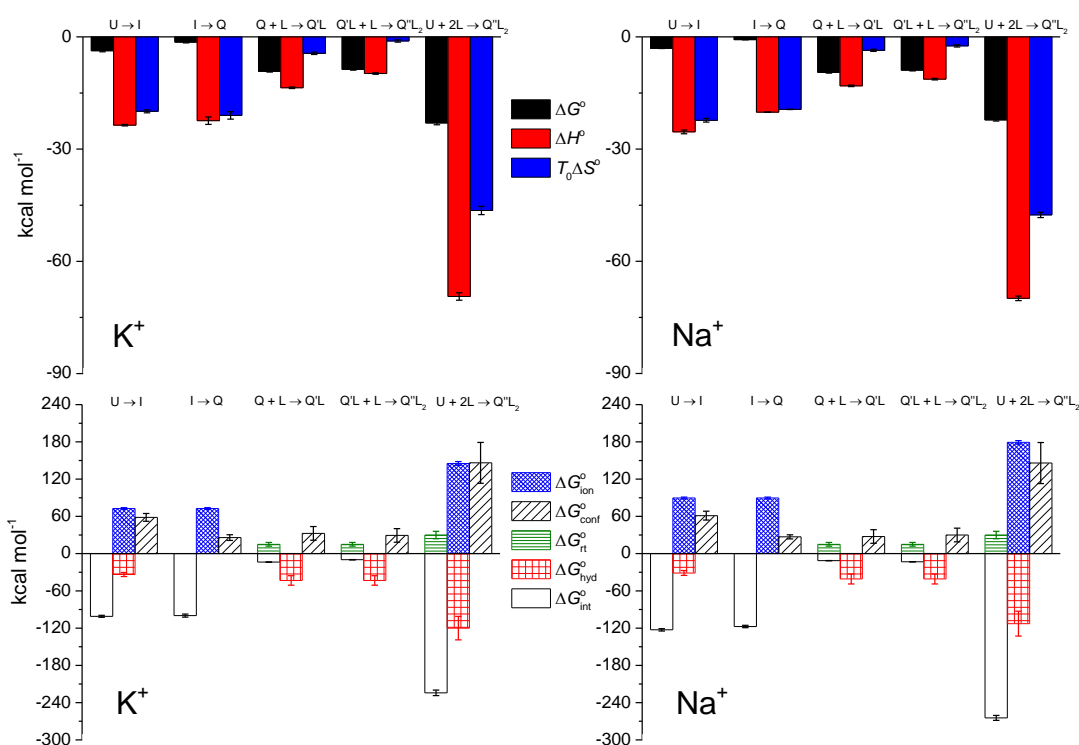

FIGURE 3 Thermodynamic profiles (top) and Gibbs free energy contributions (driving forces) for each step in the Tel22 binding-coupled folding mechanism in the presence of ligand

(Phen-DC3) and 100 mM K^+ or Na^+ (bottom) at 25 °C. The errors of $\Delta G°$ contributions (Eq. 4) were calculated combining the errors of experimental quantities ($\Delta G°$, $\Delta H°$, $\Delta C°_P$; global model analysis) and errors reported in the literature ($\Delta G°_{ion}$, $\Delta G°_{hyd}$, $\Delta G°_{rt}$) (43,45).

(i) U → I transition accompanied by specific binding (dehydration) of one cation appears to be driven by specific and hydrophobic interactions since [– ($\Delta G°_{int}$ + $\Delta G°_{hyd}$ > ($\Delta G°_{conf}$ + $\Delta G°_{ion}$)]. It seems that even though base stacking, H-bonding, and cation coordination are needed for the early stage of Tel22 secondary structure formation, intermediate states would not be significantly populated without being stabilized by hydrophobic interactions (if $\Delta G°_{hyd} \approx$ 0 ⇒ $-\Delta G°_{int}$ < ($\Delta G°_{conf}$ + $\Delta G°_{ion}$) ⇒ $\Delta G°_{U \to I}$ > 0; I formation unfavorable). Although I_{K^+} and I_{Na^+} are structurally different (Fig. 2), the observation that $\Delta G°_{hyd,Na^+} \approx \Delta G°_{hyd,K^+}$ suggests that hydrophobic interactions are equally important for driving the first step of Tel22 folding in Na^+ and K^+ solutions.

(ii) I → Q transition accompanied by specific binding (dehydration) of one cation appears to be driven entirely by specific interactions that overcome loss of conformational freedom and unfavorable ion dehydration [$-\Delta G°_{int}$ > ($\Delta G°_{conf}$ + $\Delta G°_{ion}$); $\Delta G°_{hyd} \approx$ 0]. The conformational entropy loss is about two times lower than in the case of U → I step which is in accordance with structural properties of I that are closer to Q than to U (Fig. 3). The observation that for U → I and I → Q step $\Delta G°_{int,Na^+}$ < $\Delta G°_{int,K^+}$ suggests that specific interactions are more favorable for stabilizing I and Q in the presence of Na^+ ions which can be ascribed to energetically more favorable coordination of Na^+ compared to K^+ ions.

(iii) Q + L → Q'L and Q'L + L → Q''L_2 steps resulting in the formation of Q'L and Q''L_2 conformations that differ from the corresponding ligand-free structures Q are driven predominantly by hydrophobic interactions (removal of water from ligand-Tel22 interacting surface) which together with specific (ligand-Tel22 stacking) interactions compensate unfavorable loss of conformational, translational and rotational freedom. Even if we overestimate the magnitude of $\Delta G°_{rt}$, by taking as an approximation $\Delta G°_{rt} \approx$ 15 kcal mol^{-1} (47,48), the $\Delta G°_{conf}$ is still very high (Fig. 3). This means that the Q + L → Q'L and Q'L + L → Q''L_2 steps are accompanied by a large Tel22 conformational entropy loss which is in accordance with the observed ligand-induced conformational changes (Figs. 2 and S9).

Conclusions
The global model analysis of a wide variety of experimental data enabled us to the describe the ht-DNA (Tel22) behavior at various conditions (temperature, salt and ligand concentration; Fig. 4). It resulted in the proposed hierarchy of forces that drive the binding-coupled folding of Tel22 in the presence of G-quadruplex specific ligands, Phen DC-3 and 360A-Br. We believe that such approach can generally be used in discussing the heterogeneity of both, the ht-quadruplex structural polymorphism and its folding and binding pathways.
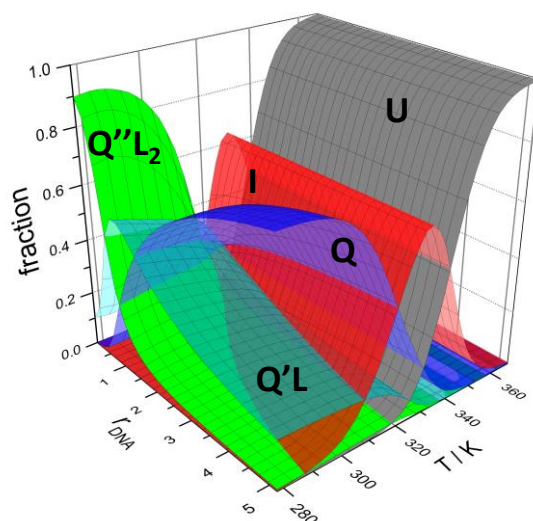
FIGURE 4 Fractions of Tel22 species presented as a function of DNA/ligand molar ratio and temperature in the presence of ligand Phen-DC3 in 100 mM K⁺ solution predicted by global thermodynamic analysis of data in terms of the proposed mechanism (Scheme 1).

Supporting Information
Methods: Model analysis of data. Results: Figures and tables describing ITC, DSC, CD, FL and PAGE electrophoresis experiments.

Author Contributions
Designed research: Jurij Lah, Gorazd Vesnaver;
Performed research: Matjaž Bončina, Florian Hamon, Barira Islam;
Contributed new reagents: Marie-Paule Teulade-Fichou, Florian Hamon;
Analyzed data: Matjaž Bončina, Jurij Lah, Shozeb Haider;
Wrote the paper: Jurij Lah, Matjaž Bončina, Gorazd Vesnaver, Shozeb Haider

References

1. Biffi, G., D. Tannahill, J. McCafferty, S. Balasubramanian. 2013. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Commun.* 5:182-186.
2. O'Sullivan, R.J., J. Karlseder. 2010. Telomeres: protecting chromosomes against genome instability. *Nat. Rev. Mol. Cell Biol.* 11:171-181.
3. Siddiqui-Jain, A., C.L. Grand, D.J. Bearss, L.H. Hurley. 2002. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. USA.* 99:11593-11598.
4. Zahler, A.M., J.R. Williamson, T.R. Cech, D.M. Prescott. 1991. Inhibition of telomerase by G-quartet DNA structures. *Nature.* 350:718-720.
5. Balasubramanian, S., L.H. Hurley, S. Neidle. 2011. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.* 10:261-275.
6. Parkinson, G. N., M.P.H. Lee, S. Neidle. 2002. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature.* 417:876-880.

7. Ambrus, A., D. Chen, J. Dai, T. Bialis, R.A. Jones, D. Yang. 2006. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.* 34:2723-2735.

8. Phan, A.T., V. Kuryavyi, K.N. Luu, D.J. Patel. 2007. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K+ solution. *Nucleic Acids Res.* 35: 6517-6525.

9. Gray, R.D., L. Petraccone, J.O. Trent, J.B. Chaires. 2010. Characterization of a K+-Induced Conformational Switch in a Human Telomeric DNA Oligonucleotide Using 2-Aminopurine Fluorescence. *Biochemistry*. 49:179-194.

10. Luu, K.N., A.T. Phan, V. Kuryavyi, L. Lacroix, D.J. Patel. 2006. Structure of the Human Telomere in K+ Solution: An Intramolecular (3 + 1) G-Quadruplex Scaffold. *J. Am. Chem. Soc.* 128:9963-9978.

11. Wang, Y., D.J. Patel. 1993. Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure*. 1:263-282.

12. Nicoludis, J.M., S.P. Barrett, J.L. Mergny, L.A. Yatsunyk. 2012. Interaction of human telomeric DNA with N-methyl mesoporphyrin IX. *Nucleic Acids Res.* 40:5432-5447.

13. Rahman, K.M., A.P. Reszka, M. Gunaratnam, S.M. Haider, P.W. Howard, K.R. Fox, S. Neidle, D.E. Thurston. 2009. Biaryl polyamides as a new class of DNA quadruplex-binding ligands. *Chem. Commun*. 27:4097-4099.

14. Granzhan, A., H. Ihmels. K. Jäger. 2009. Diazonia- and tetraazoniapolycyclic cations as motif for quadruplex-DNA ligands. *Chem. Commun*. 10:1249-1251.

15. Garner, T.P., H.E.L. Williams, K.I. Gluszyk, S. Roe, N.J. Oldham, M.F.G. Stevens, J.E. Moses, M.S. Searle. 2009. Selectivity of small molecule ligands for parallel and anti-parallel DNA G-quadruplex structures. *Org. Biomol. Chem.* 7:4194-4200.

16. Chen, M., G. Song, C. Wang, D. Hu, J. Ren, X. Qu. 2009. Small-Molecule Selectively Recognizes Human Telomeric G-Quadruplex DNA and Regulates Its Conformational Switch. *Biophys. J.* 97:2014-2023.

17. Chu J.F.,T.C. Chang, H.W. Li. 2010. Single-Molecule TPM Studies on the Conversion of Human Telomeric DNA. *Biophys. J.* 98:1608-1616.

18. Čeru, S., P. Šket, I. Prislan, J. Lah, J. Plavec. 2014. A New Pathway of DNA G-Quadruplex Formation. *Angew. Chem. Int. Ed.* 53:4881-4884.

19. Koirala, D., T. Mashimo, Y. Sannohe, Z. Yu, H. Mao, H. Sugiyama. 2012. Intramolecular folding in three tandem guanine repeats of human telomeric DNA. *Chem. Commun*. 48:2006-2008.

20. Gray, R.D., R. Buscaglia, J.B. Chaires. 2012. Populated intermediates in the thermal unfolding of the human telomerc quadruplex. *J. Am. Chem. Soc.* 134:16834-16844.

21. Stadlbauer, P., L. Trantírek, T.E. Cheatham, J. Koča, J. Sponer. 2014. Triplex intermediates in folding of human telomeric quadruplexes probed by microsecond-scale molecular dynamics simulations. *Biochimie*. 105:22-35.

22. Dettler, J.M., R. Buscaglia, V.H. Le, E.A. Lewis. 2011. DSC Deconvolution of the Structural Complexity of c-MYC P1 Promoter G-Quadruplexes. *Biophys. J.* 100:1517-1525.

23. Cantor, C. R., M.M. Warshaw, H. Shapiro. 1970. Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers.* 9:1059-1077.

24. Bončina, M., J. Lah, I. Prislan, G. Vesnaver. 2012. Energetic basis of human telomeric DNA folding into G-quadruplex structures. *J. Am. Chem. Soc.* 134:9657-9663.

25. De Cian, A., E. DeLemos, J.L. Mergny, M.P. Teulade-Fichou, D. Monchaud. 2007. Highly Efficient G-Quadruplex Recognition by Bisquinolinium Compounds. *J. Am. Chem. Soc.* 129:1856-1857.

26. Monchaud, D., M.P. Teulade-Fichou. 2008. A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem.* 6:627-636.

27. Sidibe, A., F. Hamon, E. Largy, A. Guédin, D. Gomez, M.P. Teulade-Fichou, C. Trentesaux, J.F. Riou. 2012. Effects of a halogenated G-quadruplex ligand from the pyridine dicarboxamide series on the terminal sequence of XpYp telomere in HT1080 cells. *Biochimie*. 12:2559-2568.

28. Chung, W.J., B. Heddi, F. Hamon, M.P. Teulade-Fichou, A.T. Phan. 2014. Solution Structure of a G-quadruplex Bound to the Bisquinolinium Compound Phen-DC3. *Angew. Chem. Int. Ed.* 53:999-1002.

29. Drobnak, I., G. Vesnaver, J. Lah. 2010. Model-based thermodynamic analysis of reversible unfolding processes. *J. Phys. Chem. B.* 114: 8713-8722.

30. Mashimo, T., H. Yagi, Y. Sannohe, A. Rajendran, H. Sugiyama. 2010. Folding Pathways of Human Telomeric Type-1 and Type-2 G-Quadruplex Structures. *J. Am. Chem. Soc.* 132:14910-14918.

31. Olsen, C.M., W.H. Gmeiner, L.A. Marky. 2006. Unfolding of G-Quadruplexes: Energetic, and Ion and Water Contributions of G-Quartet Stacking. *J. Phys. Chem. B.* 110:6962-6969.

32. Limongelli, V., S. De Tito, L. Cerofolini, M. Fragai, B. Pagano, R. Trotta, S. Cosconati, L. Marinelli, E. Novellino, I. Bertini, A. Randazzo, C. Luchinat, M. Parrinello. 2013. The G-Triplex DNA. *Angew. Chem. Int. Ed.* 52:2269-2273.

33. Gray, R.D., J. Li, J.B. Chaires. 2009. Energetics and Kinetics of a Conformational Switch in G-Quadruplex DNA. J*. Phys. Chem. B.* 113:2676-2683.

34. Mergny, J.L., A.T. Phan, L. Lacroix. 1998. Following G-quartet formation by UV-spectroscopy*. FEBS Lett.* 435:74-78.

35. Brandts, J.F., L.-N. Lin. 1990. Study of strong to ultratight protein interactions using differential scanning calorimetry. *Biochemistry.* 29:6927-6940.

36. Buscaglia, R., M.C. Miller, W.L. Dean, R.D. Gray, A.N. Lane, J.O. Trent, J.B. Chaires. 2013. Polyethylene glycol binding alters human telomere G-quadruplex structure by conformational selection. *Nucleic Acids Res.* 41:7934-7946.

37. Lane, A.N., J.B. Chaires, R.D. Gray, J.O. Trent. 2008. Stability and kinetics of G-quadruplex structures. *Nucleic Acids Res.* 36:5482-5515.

38. Smirnov, I.V., R.H. Shafer. 2007. Electrostatics dominate quadruplex stability. *Biopolymers.* 85:91-101.

39. Gray, R.D., J.B. Chaires. 2011. Linkage of cation binding and folding in human telomeric quadruplex DNA. *Biophys. Chem.* 159:205-209.

40. Haq, I., J.E. Ladbury, B.Z. Chowdhry, T.C. Jenkins, J.B. Chaires. 1997. Specific Bionding of Hoechst 33258 to the d(CGCAAATTTGCG)2 Duplex: Calorimetric and Spectroscopic Studies. *J. Mol. Biol.* 271:244-257.

41. Lah, J., G. Vesnaver. 2004. Energetic Diversity of DNA Minor-groove Recognition by Small Molecules Displayed Through Some Model Ligand-DNA Systems. *J. Mol. Biol.* 342:73-89.

42. Lah, J., I. Drobnak, M. Dolinar, G. Vesnaver. 2008. What drives the binding of minor groove-directed ligands to DNA hairpins? *Nucleic Acids Res.* 36:897-904.

43. Marcus, Y.A. 1986. *Ion solvation*; John Wiley & Sons Ltd.

44. Majhi, P. R., J. Qi, C.F. Tang, R.H. Shafer. 2008. Heat Capacity Changes Associated with Guanine Quadruplex Formation: An Isothermal Titration Calorimetry Study. *Biopolymers.* 89:302-309.

45. Baldwin, R.L. 1986. Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl. Acad. Sci. USA.* 83:8069-8072.

46. Spolar, R.S., M.T. Record. 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science.* 263:777-784.

47. Finkelstein, A. V., J. Janin. 1989. The price of lost freedom - entropy of bimolecular complex-formation. *Protein Eng.* 3:1-3.

48. Dolenc, J., R. Baron, C. Oostenbrink, J. Koller, W.F. van Gunsteren. 2006. Configurational Entropy Change of Netropsin and Distamycin upon DNA Minor-Groove Binding. *Biophys. J.* 91:1460-1470.