G-Quadruplexes

International Edition: DOI: 10.1002/anie.201605350
German Edition: DOI: 10.1002/ange.201605350

Unraveling the Thermodynamics of the Folding and Interconversion of Human Telomere G-Quadruplexes

Matjaž Bončina, Gorazd Vesnaver, Jonathan Brad Chaires, and Jurij Lah**

Dedicated to Professor Vojko Vlachy on the occasion of his 70th birthday

Abstract: *Why human telomere DNA fragments fold into different G-quadruplex structures with parallel, hybrid, and antiparallel strand orientations depending on the temperature and concentration of co-solutes remains poorly understood. Similarly, the formation of intermediate structures along the folding or interconversion pathways is not well understood. Herein, we address these questions by introducing a conceptual framework, based on the global thermodynamic analysis of DSC and CD spectroscopy data, which led to a detailed description of the topological phase space (phase diagram) of the stability of the human telomere fragment 5′-AGGG-(TTAGGG)$_3$-3′ (Tel22). This framework clarifies the driving forces of quadruplex folding and interconversion processes over a wide range of temperatures and ion (K^+, Na^+) and polyethylene glycol (PEG) concentrations and demonstrates their linkage to the human telomere DNA structural features.*

Guanine-rich human telomere (ht) repeat DNA sequences fold into various G-quadruplex structures by complex mechanisms involving multiple intermediates.[1] Key questions about quadruplex formation relate to how the ht quadruplexes fold and interconvert, and why they fold in a specific way. Several studies dealing with the mechanisms and kinetics of folding and interconversion have attempted to answer the "how" question.[2] On the other hand, the "why" question, related to the thermodynamic driving forces, remains mostly unanswered owing to the complexity of a thermodynamic analysis that has to take into account multiple contributions to the stability of the intermediates and folded structures. Herein, we attempt to address this question by introducing a conceptual thermodynamic framework that enables the description of the topological phase space of Tel22, from which estimates of the driving forces of its folding and interconversion processes can be made over a wide range of temperatures (T) and ion (K^+, Na^+) and PEG concentrations.

[*] Dr. M. Bončina, Prof. Dr. G. Vesnaver, Prof. Dr. J. Lah
Faculty of Chemistry and Chemical Technology
University of Ljubljana
Večna pot 113, Ljubljana (Slovenia)
E-mail: jurij.lah@fkkt.uni-lj.si
Prof. Dr. J. B. Chaires
James Graham Brown Cancer Center
University of Louisville
505 S. Hancock St., Louisville, KY 40202 (USA)
E-mail: j.chaires@louisville.edu
Supporting information and the ORCID identification number for an author of this article can be found under http://dx.doi.org/10.1002/anie.201605350.

In different solutions, Tel22 forms an ensemble of quadruplex structures with a variety of strand-segment orientations and loop topologies, including antiparallel "basket" forms (designated as A, corresponding to Na^+ structure 143D),[3] "hybrid" (H) forms,[4] and a parallel "propeller" (P, corresponding to the K^+ crystal structure 1KF1)[5] form. These forms are in equilibrium with one another and with stable folding intermediate species, resulting in a complex phase space. We previously described the structural features and stability of Tel22 conformations in K^+ and Na^+ solutions,[1a,6] and herein, we present new thermodynamic data collected in the presence of the co-solute PEG (Figure 1). The availability of such an extensive thermodynamic dataset raised two questions that are crucial for our understanding of the ht-DNA fragment (un)folding and interconversion processes: 1) What is the simplest model mechanism that can describe all of the available experimental data simultaneously? 2) How can thermodynamic parameters be used to dissect and correlate the molecular contributions to folding and interconversion to changes in structural features? Answering these questions will enable us to achieve our fundamental goals of predicting the population of Tel22 conformations (phases) as a function of T, $[K^+]$, $[Na^+]$, and [PEG] and estimating the dominant driving forces of Tel22 folding and interconversion processes.

To address question (1), we monitored the thermal unfolding of Tel22 in solutions with $[K^+] = 25$ mM and various [PEG] by differential scanning calorimetry (DSC;[7] Figure 1a). Furthermore, we recorded circular dichroism (CD) spectra[8] of Tel22 in solution at $[K^+] = 25$ mM and different [PEG] (Figure 1b) or at [PEG] = 60 mM with different Na^+ concentrations (Figure 1c). Global model analysis (see the Supporting Information for details) suggests that the simplest mechanism that can describe all experimental data involves seven macroscopic states (Scheme 1). A transition between any of these states was described by the corresponding apparent standard Gibbs free energy change, $\Delta G_{(T,K,Na,PEG)}$ [see the Supporting Information, Eq. (S4)], which is defined by seven parameters, including the standard enthalpy (ΔH), entropy (ΔS), and heat capacity (ΔC_P) changes. They define the corresponding model functions [Eq. (S2) and (S3)] that fit very well to the experimental data (Figure 1) and the thermodynamic profiles (Tables S1 and S2).

To address question (2), we describe the ht-DNA folding and interconversion between the species shown in Scheme 1 at $T = 25$ °C, $[K^+] = [Na^+] = 25$ mM, and [PEG] = 0 mM in terms of more fundamental driving forces. First, we dissected the corresponding apparent $\Delta G_{(T,K,Na,PEG)}$ of folding into its

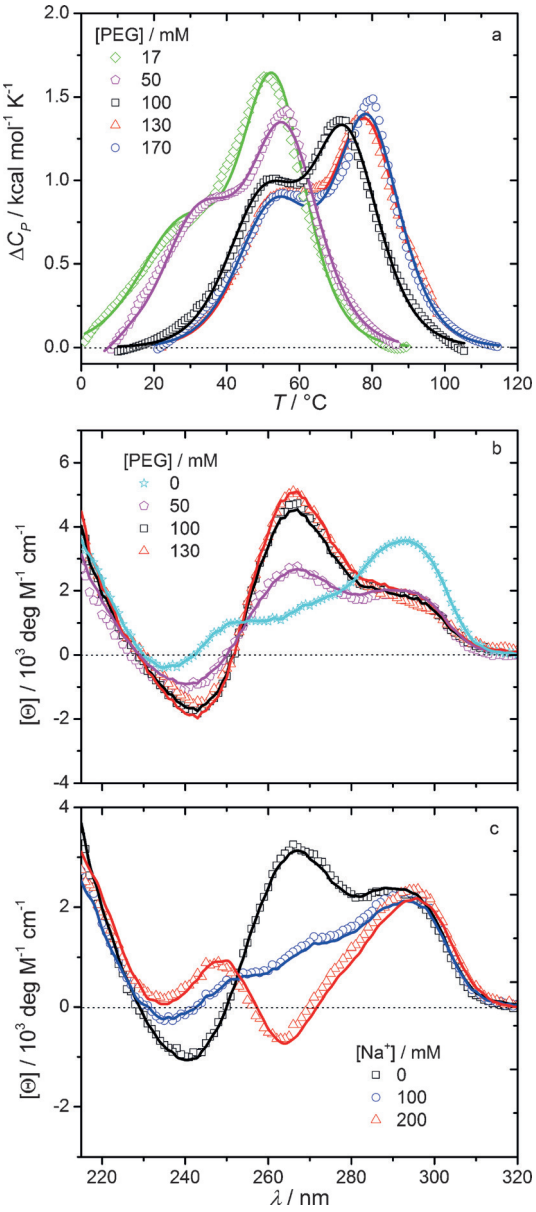
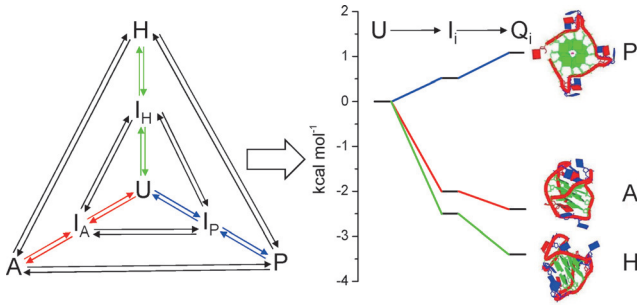


Figure 1. Tel22 unfolding and interconversion at [K$^+$] = 25 mM monitored by DSC at different [PEG] (a), and by CD spectroscopy at T = 25 °C at varying [PEG] (b) or [Na$^+$] at [PEG] = 60 mM (c). The globally fitted model functions [Eq. (S3) and (S4); shown as solid lines] derived on the basis of the model presented in Scheme 1 show good agreement with the experimental data (symbols).



Scheme 1. The simplest model mechanism (left) that successfully describes (un)folding and structural interconversion processes of Tel22 at various T, [K$^+$], [Na$^+$], and [PEG] includes hybrid (H), antiparallel (A), and parallel (P) G-quadruplexes (Q), intermediates (I_H, I_A, I_P), and the unfolded structure (U). A free energy diagram for the folding of the quadruplex forms under specific solution conditions (T = 25 °C, [K$^+$] = [Na$^+$] = 25 mM, and [PEG] = 0 mM) is shown on the right.

enthalpy and entropy components (Table S1). These data show that under these conditions, Tel22 folding into P is unfavorable and becomes favorable only at high [PEG] (Figure S1). Each stage of the U→I_P→P folding process is, just like the U→I_H→H and U→I_A→A transitions, characterized by extensive enthalpy–entropy compensation. It should be emphasized that our calorimetric data and global analysis uniquely provide a reasonable estimate of ΔC_P, a fingerprint of changes in hydration. ΔC_P correlates with changes in the solvent-accessible surface area and may be translated into the ΔG contribution (ΔG_{hyd}), which, at $T \approx 25$ °C, reflects mainly the entropy of the (de)hydration of hydrophobic groups. Figure 2a shows that most of the ΔG_{hyd} change upon folding occurs in the formation of stable intermediates, suggesting that hydrophobic dehydration is the primary driving force of these processes. By contrast, specific ion binding, base stacking, and hydrogen bonding appear to be the main driving forces for I_P→P, I_H→H, and I_A→A folding. Taken together, the hydrophobic dehydration effects are significant in the initial steps of ht quadruplex folding, a process that may be thought of as being analogous to molten globule formation in the folding of globular proteins.

Next, we used thermodynamic cycles to estimate the thermodynamic parameters for the interconversion of folded structural forms (Figure 2; see also Table S2). They show that the P→H conversion is strongly favored, driven by a large favorable entropy contribution. The A→H conversion is energetically favorable, but with a ΔG value of only −1 kcal mol^{-1}. The free energy changes in this case result from equal enthalpy and entropy contributions. Furthermore, we attempted to estimate the four main contributions to the observed free energies of interconversion under the specific solution conditions, $\Delta G = \Delta G_{hyd} + \Delta G_{ion} + \Delta G_{int} + \Delta G_{conf}$. (Changes in Tel22 and ion translational and rotational freedom may be neglected as the starting and end states are similarly ligated.) ΔG_{hyd} has already been mentioned above. ΔG_{ion} is the contribution of ion (K$^+$, Na$^+$) (de)hydration, ΔG_{conf} represents the entropic contribution that is due to changes in the conformational freedom, and ΔG_{int} is the contribution that is due to specific interactions arising mainly from differences in the potential energy of van der Waals (base stacking), hydrogen bonding, and cation coordination interactions (see the Supporting Information for details on the estimation of these contributions). We are well aware that owing to the approximations involved in this dissection, the estimated values are tenuous, and the ΔG contributions should be interpreted with great care. Nevertheless, we believe that for the interconversions, the ΔG contributions are reliable enough to characterize their driving forces. The contributions describing the A→H transition (Figure 2b) show that the H conformation is more populated than the A conformation, which is mainly due to a favorable ion

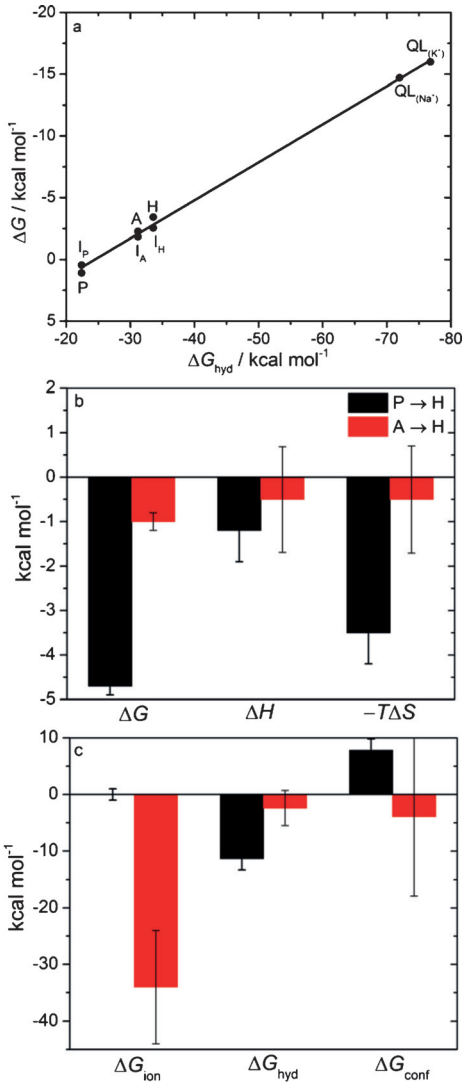


Figure 2. Thermodynamic profiles of Tel22 ($T = 25\,°C$, $[K^+] = [Na^+] = 25$ mM, [PEG] = 0 mM). a) Free energy of the Tel22 folding into intermediate, folded, and ligand (L = Phen-DC3) bound[6b] folded forms as a function of the "dehydration" contribution estimated as $\Delta G_{hyd} = \Delta C_P \times 80$ K.[13] b) Energetic contributions to the P→H (black) and A→H (red) interconversions; free energy, enthalpy, and entropy contribution values. c) Best estimates for the ΔG_{ion}, ΔG_{hyd}, and ΔG_{conf} contributions (see Table S2).

(de)hydration effect (Na^+ hydration, K^+ dehydration; $\Delta G_{ion} \ll 0$), which is in agreement with the generally accepted interpretation of Na^+ and K^+ influence on quadruplex stability.[9] The H conformation is slightly favored over the A conformation also by other hydration contributions ($\Delta G_{hyd} < 0$) and higher conformational freedom ($\Delta G_{conf} < 0$). All of these contributions together overcompensate the highly unfavorable ΔG_{int} value (which is mainly due to Na^+ decoordination from the quadruplex core). The relatively small ΔG_{hyd} and ΔG_{conf} values are in accordance with the similar exposure of surfaces to water in the A and H structures (Table S3) and with their similar compactness.[5d]

The driving forces for P→H conversion differ significantly from those for A→H conversion. P is not significantly populated because ΔG_{hyd}, which strongly favors the H form, overcompensates the positive ΔG_{conf}. This interpretation is in accordance with the P and H structural features, which indicate that the H structure is more compact than the P structure, and that the P form has a larger water-exposed (hydrophobic) surface area than H.[5d] This is supported by the ΔC_P value estimated from the corresponding change in the solvent-accessible surface area (Table S3), which is in accordance with our experimentally determined ΔC_P value. Lowering the exposure of hydrophobic groups in Tel22 may thus increase the P type population. One way to achieve this is to add PEG.[10] The hypothesis that the H→P transition becomes favorable in the presence of PEG (Figure S1) owing to selective PEG binding to the P conformation[5e] and not because of crowding is in full agreement with the analysis of the P↔H conversion presented here (see the Supporting Information). The hydrophobicity may also be reduced by replacing the thymines in the Tel22 sequence with less hydrophobic uracils.[11] Indeed, the shape of the CD spectrum of the modified sequence (Figure S5) suggests the coexistence of P- and H-type conformations. Lowering the hydrophobicity may thus be an important factor governing the formation of the parallel structure from RNA Tel22.[12]

The thermodynamic parameters used for the estimation of the driving forces enable also the calculation of the species populations at any T, $[K^+]$, $[Na^+]$, and [PEG] (Figure S4) so that Tel22 phase diagrams can be constructed (Figure 3). These diagrams indicate the possible macroscopic pathways of Tel22 (un)folding and interconversion. For example, Figure 3a shows that Tel22 thermal unfolding at low [PEG] follows mainly a H→I_H→U pathway, but at high [PEG], a P→I_P→U pathway predominates. The Tel22 conformational change at 25 °C induced by increasing [PEG] follows a H→P pathway (Figure 3a), but follows a P→A pathway upon increasing $[Na^+]$ at [PEG] = 60 mM (Figure 3b). The $[Na^+]$ versus $[K^+]$ diagrams (Figure 3c,d) show that at 25 °C, phase A is accessible to Tel22 only at $[Na^+] \gg [K^+]$.

The phase diagrams also show that in the absence of PEG, the intermediates appear mainly at $[Na^+]$ or/and $[K^+]$ values between 1 mM and 10 mM whereas at high [PEG], the I_P phase appears at $[K^+] < 1$ mM. This analysis suggests that the U phase collapses into phase I under conditions at which the Tel22 chain is able to coordinate specifically one K^+ or Na^+ ion and simultaneously remove a significant amount of hydrophobic surface from water. Tel22 forms in phase I appear to be structured as their measured CD spectra are similar to those of the folded quadruplexes (Figure S3). These intermediates may thus be considered to be analogous to the molten globule structures in proteins. Recent thermal unfolding studies on the same and similar ht-DNA sequences[1a,b] suggest that at specific conditions ($[PEG] = [Na^+] = 0$ mM, $[K^+] > 0$ mM), phase I may consist of a greater number of stable intermediate species than predicted in Scheme 1. Despite this difference, which is a consequence of the different sensitivities of the biophysical methods used to monitor the unfolding pathway, the predicted T range at which phase I is stable (Figure 3a) is the same as the one estimated on the basis of the melting temperature data reported for several ht-DNA sequence variants.[1b] Taken

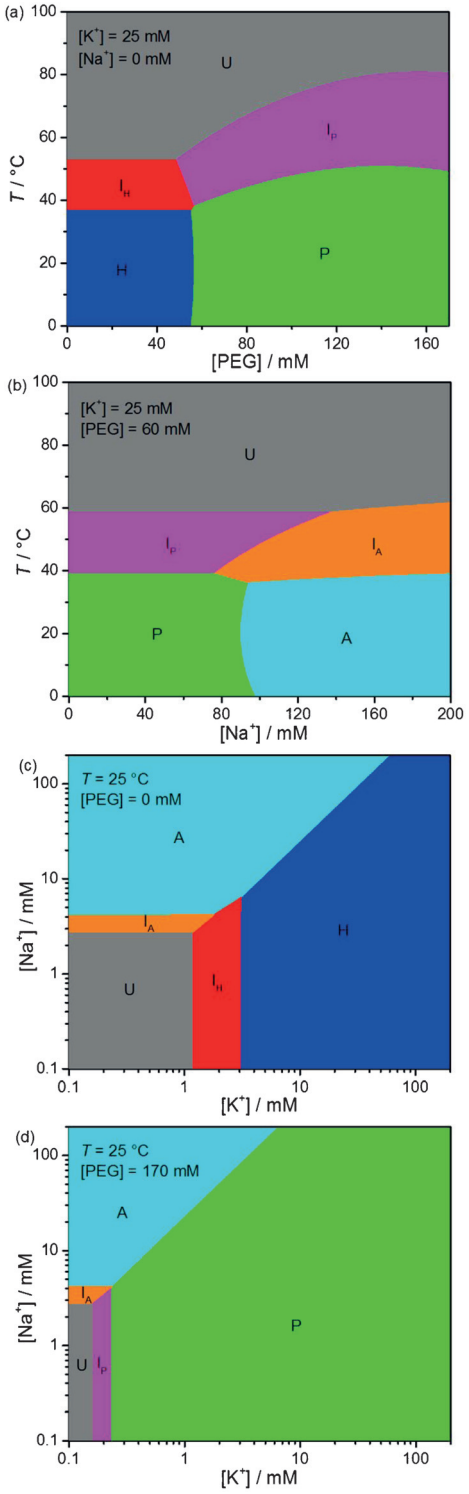


Figure 3. Phase diagrams showing the most populated Tel22 species (phases) at a given *T*, [K^+], [Na^+], and [PEG].

together, the presented phase diagrams emphasize that ht-DNA adopts an ensemble of structures in equilibrium, an equilibrium that can be easily perturbed by changes in the solution conditions, co-solutes, and quadruplex-binding ligands.

In summary, by global thermodynamic analysis of the structural transitions of a human telomere fragment, we have described in detail its topological phase space and suggested why certain intermediate and folded quadruplex states are populated at given temperatures and co-solute concentrations. We have shown that hydrophobic desolvation represents a crucial driving force for the formation of stable compact folding intermediates that could be, in analogy with protein folding, characterized as human telomere quadruplex molten globules. Significantly, this thermodynamic analysis appears to give reliable predictions of the driving forces for human telomere quadruplex interconversions, which are consistent with the observed structural features. Therefore, we believe that this work represents an important step in elucidating the thermodynamic principles by which human telomere G-quadruplexes fold and interconvert into different structures.

Acknowledgements

This work was supported by grant P1-0201 of the Slovenian Research Agency (M.B., G.V., J.L.) and by grants CA35635 and GM077422 (J.B.C.).

Keywords: differential scanning calorimetry · DNA structures · G-quadruplexes · human telomere repeat · thermodynamics

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 10340–10344
Angew. Chem. **2016**, *128*, 10496–10500

[1] a) R. D. Gray, R. Buscaglia, J. B. Chaires, *J. Am. Chem. Soc.* **2012**, *134*, 16834–16844; b) R. Buscaglia, R. D. Gray, J. B. Chaires, *Biopolymers* **2013**, *99*, 1006–1018; c) P. Stadlbauer, L. Trantírek, T. E. Cheatham III, J. Koča, J. Šponer, *Biochimie* **2014**, *105*, 22–35; d) A. Rajendran, M. Endo, K. Hidaka, H. Sugiyama, *Angew. Chem. Int. Ed.* **2014**, *53*, 4107–4112; *Angew. Chem.* **2014**, *126*, 4191–4196; e) W. Li, X. M. Hou, P. Y. Wang, X. G. Xi, M. Li, *J. Am. Chem. Soc.* **2013**, *135*, 6423–6426.

[2] a) I. Bessi, H. R. A. Jonker, C. Richter, H. Schwalbe, *Angew. Chem. Int. Ed.* **2015**, *54*, 8444–8448; *Angew. Chem.* **2015**, *127*, 8564–8568; b) A. Y. Q. Zhang, S. Balasubramanian, *J. Am. Chem. Soc.* **2012**, *134*, 19297–19308; c) R. D. Gray, J. B. Chaires, *Nucleic Acids Res.* **2008**, *36*, 4191–4203; d) M. W. da Silva, *Chem. Eur. J.* **2007**, *13*, 9738–9745; e) E. Largy, A. Marchand, S. Amrane, V. Gabelica, J. L. Mergny, *J. Am. Chem. Soc.* **2016**, *138*, 2780–2792; f) S. Čeru, P. Šket, I. Prislan, J. Lah, J. Plavec, *Angew. Chem. Int. Ed.* **2014**, *53*, 4881–4884; *Angew. Chem.* **2014**, *126*, 4981–4984.

[3] Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263–282.

[4] A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones, D. Yang, *Nucleic Acids Res.* **2006**, *34*, 2723–2735.

[5] a) G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876–880; b) Y. Xue, J. Liu, K. Zheng, Z. Kan, Y. Hao, Z. Tan, *Angew. Chem. Int. Ed.* **2011**, *50*, 8046–8050; *Angew. Chem.* **2011**, *123*, 8196–8200; c) B. Heddi, A. T. Phan, *J. Am. Chem. Soc.* **2011**, *133*, 9824–9833; d) M. C. Miller, R. Buscaglia, J. B. Chaires, A. N. Lane, J. O. Trent, *J. Am. Chem. Soc.* **2010**, *132*, 17105–17107; e) R. Buscaglia, M. C. Miller, W. L. Dean, R. D. Gray, A. N. Lane, J. O. Trent, J. B. Chaires, *Nucleic Acids Res.* **2013**, *41*, 7934–7946.

[6] a) M. Bončina, J. Lah, I. Prislan, G. Vesnaver, *J. Am. Chem. Soc.* **2012**, *134*, 9657–9663; b) M. Bončina, F. Hamon, B. Islam, M. P.

Teulade-Fichou, G. Vesnaver, S. Haider, J. Lah, *Biophys. J.* **2015**, *108*, 2903 – 2911.
[7] B. Pagano, A. Randazzo, I. Fotticchia, E. Novellino, L. Petraccone, C. Giancola, *Methods* **2013**, *64*, 43 – 51.
[8] A. I. Karsisiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba da Silva, *Angew. Chem. Int. Ed.* **2011**, *50*, 10645 – 10648; *Angew. Chem.* **2011**, *123*, 10833 – 10836.
[9] a) N. V. Hud, F. W. Smith, F. A. L. Anet, J. Feigon, *Biochemistry* **1996**, *35*, 15383 – 15390; b) E. Largy, J.-L. Mergny, V. Gabelica, *Met. Ions Life Sci.* **2016**, *16*, 203 – 258.
[10] D. Miyoshi, A. Nakao, N. Sugimoto, *Biochemistry* **2002**, *41*, 15017 – 15024.
[11] C. M. Olsen, L. A. Marky, *J. Phys. Chem. B* **2009**, *113*, 9 – 11.
[12] A. Arora, S. Maiti, *J. Phys. Chem. B* **2009**, *113*, 10515 – 10520.
[13] R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 8069 – 8072.

Received: June 1, 2016
Published online: August 2, 2016

