Supplementary Information

Kinetics and mechanism of K^+- and Na^+-induced folding of models of human telomeric DNA into G-quadruplex structures

Robert D. Gray and Jonathan B. Chaires

James Graham Brown Cancer Center, University of Louisville, 529 S. Jackson St., Louisville, KY 40202, USA and Department of Biochemistry & Molecular Biology, University of Louisville, Louisville, KY 40292

Summary of the singular value decomposition (SVD) method of analysis of multi-wavelength spectroscopic data sets

Some key features of SVD analysis are briefly summarized below. The reader is referred to reference S1 for a comprehensive review of the SVD method and to reference S2 for an application of SVD analysis to detection of spectroscopic intermediates in polynucletodide melting.

Spectrophotometric titration data such as those shown in Figures 2 and S1 or kinetic data such as that in Figures 4, S5 and S6 can be arranged in matrix form. One arrangement is to have absorbance values measured at discrete wavelengths in columns (wavelength vectors) and ligand concentrations or times are arranged in rows (concentration or time vectors). Thus each element $A_{i,j}$ of the data matrix **A** corresponds to an absorbance value determined at ligand concentration *i* or time *i* and wavelength *j*. **A** can be decomposed into the product of three matrices such that $\mathbf{A = USV^T}$, where $\mathbf{V^T}$ is the transform of matrix **V**, **S** is a diagonal matrix that contains the relative weights of the signals contributing to the spectra and **V** contains information related to the relative concentrations of the spectroscopically distinct species as a function of ligand concentration or time. The product matrix **US** defines basis spectra that reflect the relative weights of the component absorption spectra. It is generally noted that the vectors in **US** do not represent the actual absorption spectra of the individual components. The true absorption spectra of the components can be determined by least-squares fitting the data matrix to an analytical expression corresponding to a specific chemical mechanism or model reaction. A second important characteristic of the SVD method is that it allows construction of a smoothed data matrix **A'** with reduced noise from the expression $\mathbf{A' = USV^T}$ by including only significant (non-noise) elements of **S**.

The relative significance of the spectroscopically distinct species can be assessed by several criteria including: (a) the relative magnitudes of successive singular values in **S**; (b) the non-random wavelength dependence of the **US** matrix elements (e.g. similarity to absorption spectra); (c) the concentration dependence of the **V** matrix elements (e.g. similarity to a titration or kinetic progress curve); and (d) a positive value of the autocorrelation coefficients of successive **U** and **V** matrix elements that approach the theoretical maximum value of +1 for fully correlated elements of the matrix. Our criteria for significance in the cation titrations is a value of the autocorrelation coefficient $C_i[V] > 0.5$.

SVD analysis of titration data

SVD analysis of the difference spectra in Figures 2 and S1 for K^+ binding to ODN1, ODN2 and ODN3 gave **US**-wavelength plots and **V**-[cation] plots shown in Panels A and B of Figures S2, S3 and S4, respectively. Similar data are shown in Figures S2-S4, panels C and D for Na^+-driven folding of these sequences. Only data above 250 nm were included in these analyses because of slow spectral changes noted at wavelengths <250 nm. The most significant basis spectra (dotted lines in Panels A and C of Figures S2-S4) resemble the experimental difference spectra $\Delta\varepsilon_U - \Delta\varepsilon_F$ in Figures 2

and S1, exhibiting pronounced differences at ~295 nm and ~270 nm. The relative weights of species 2 and 3 as indicated by their singular values (Table S1) is generally small (<10%), suggesting that either these species are present in small concentration and/or that their UV absorption spectra differ only slightly from that of the unfolded oligonucleotide. The basis spectra for these less significant species have shapes that suggest only slight differences in the UV absorption by these species compared to the fully unfolded oligonucleotides.

The dependence of the **V** coefficients on cation concentration (which resembles titration curves) indicates the relative amplitudes and concentration profiles of the significant species during the titration. The sequential nature of the three curves suggests that these species are true titration intermediates whose concentrations change during the titration as they are interconverted, e.g. the less significant components tend to appear earlier in the titration with component 2 (red curves in Figures S3-S5) preceding component 3 (blue curves).

Based on the **V** autocorrelation coefficients summarized in Table S1, we suggest that three intermediate species are required to represent the titration data for K^+ driven folding of ODN1 and ODN2 as well as Na^+-dependent folding of ODN2. Based on this criterion, evidence of two intermediates is apparent for Na^+-driven folding of ODN1 as well as folding of ODN3 in the presence of either cation.

SVD analysis of kinetic data

The SVD analysis of the kinetic progress curves is not reproduced here because the least squares fitting procedures outlined in the published manuscript suffice to indicate the analytical basis for number of kinetically important species and their difference spectra with respect to the folded oligonucleotides (see e.g. Figures 5 and S7).

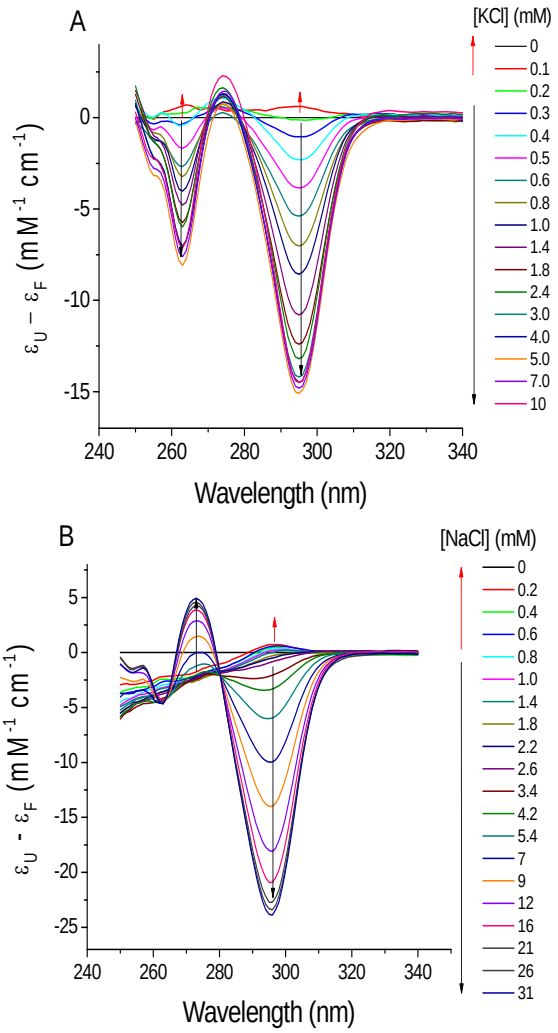

Figure S1. Difference spectra showing titration of ODN3 with KCl (A) and NaCl (B). Data are presented as the difference in absorptivity between the unfolded oligonucleotide (ε_U) and the folded oligonucleotide (ε_F). The vertical arrows correlate increases or decreases in $\Delta\varepsilon$ with increasing cation concentrations. Conditions: initial [ODN3] = 7.2 μM in folding buffer (10 mM tetrabutyl ammonium phosphate, 1 mM EDTA, pH 6.97). The spectra are corrected for sample dilution resulting from serial additions of cation.

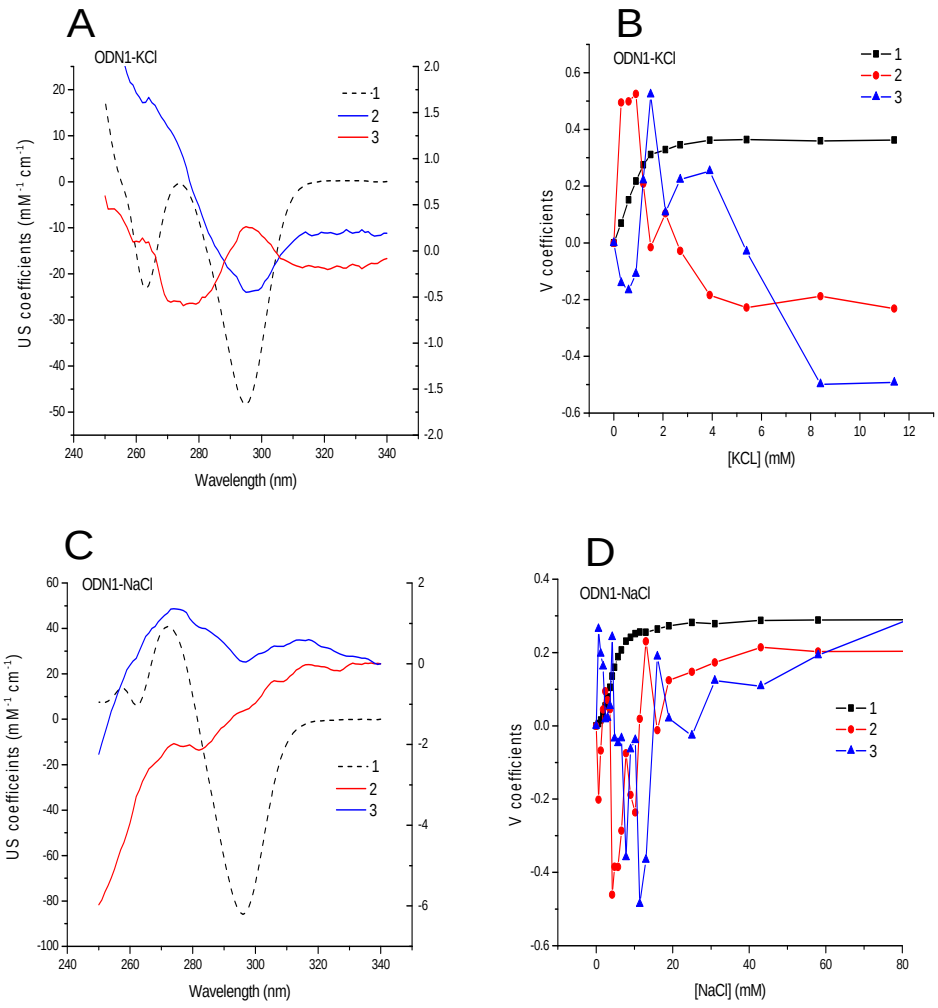

Figure S2. Basis spectra (panels A and C) and amplitude vectors (panels B and D) for the three most significant components in the titrations of ODN1 with KCl and NaCl. The data in the figures was derived by SVD analysis of the difference spectra in Figure 2.

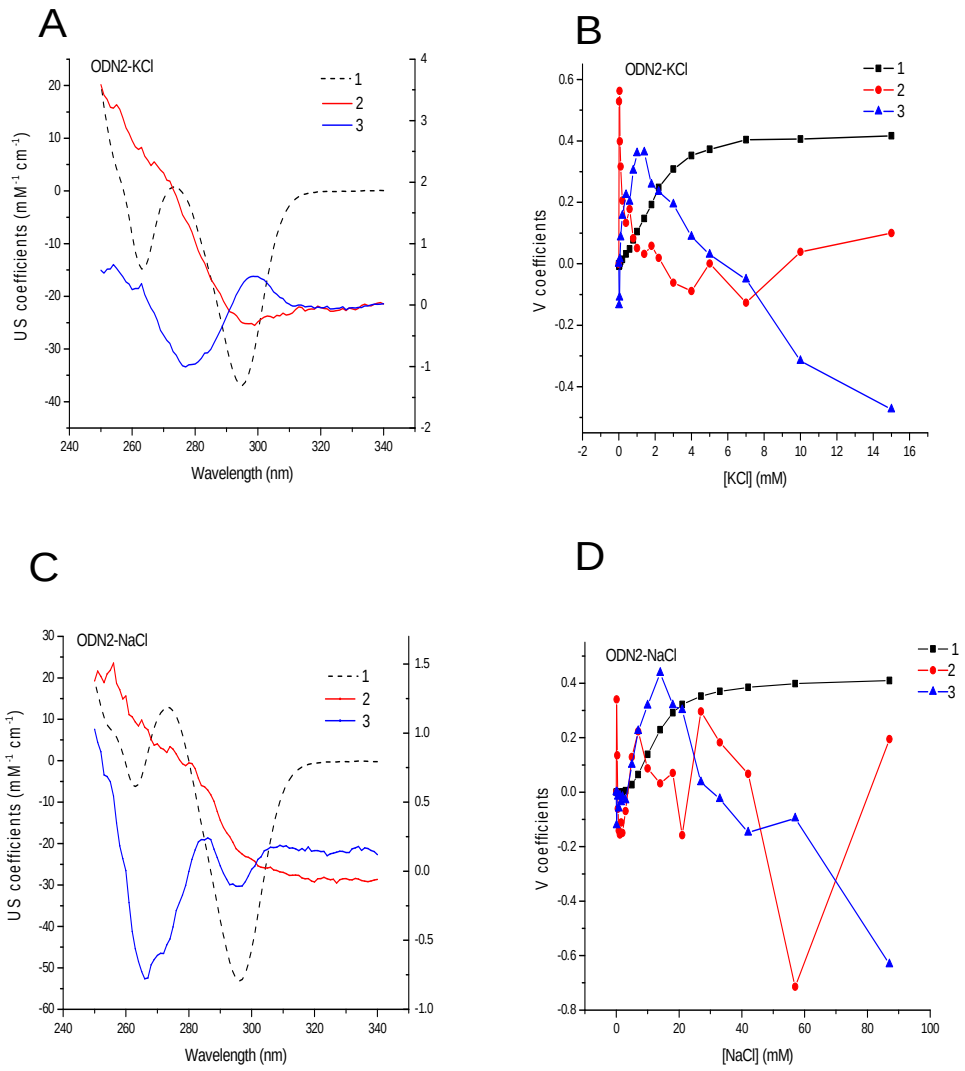

Figure S3. Basis spectra (panels A and C) and amplitude vectors (panels B and D) for the three most significant components in the titrations of ODN2 with KCl and NaCl. The data in the figures was derived by SVD analysis of the difference spectra in Figure 2.

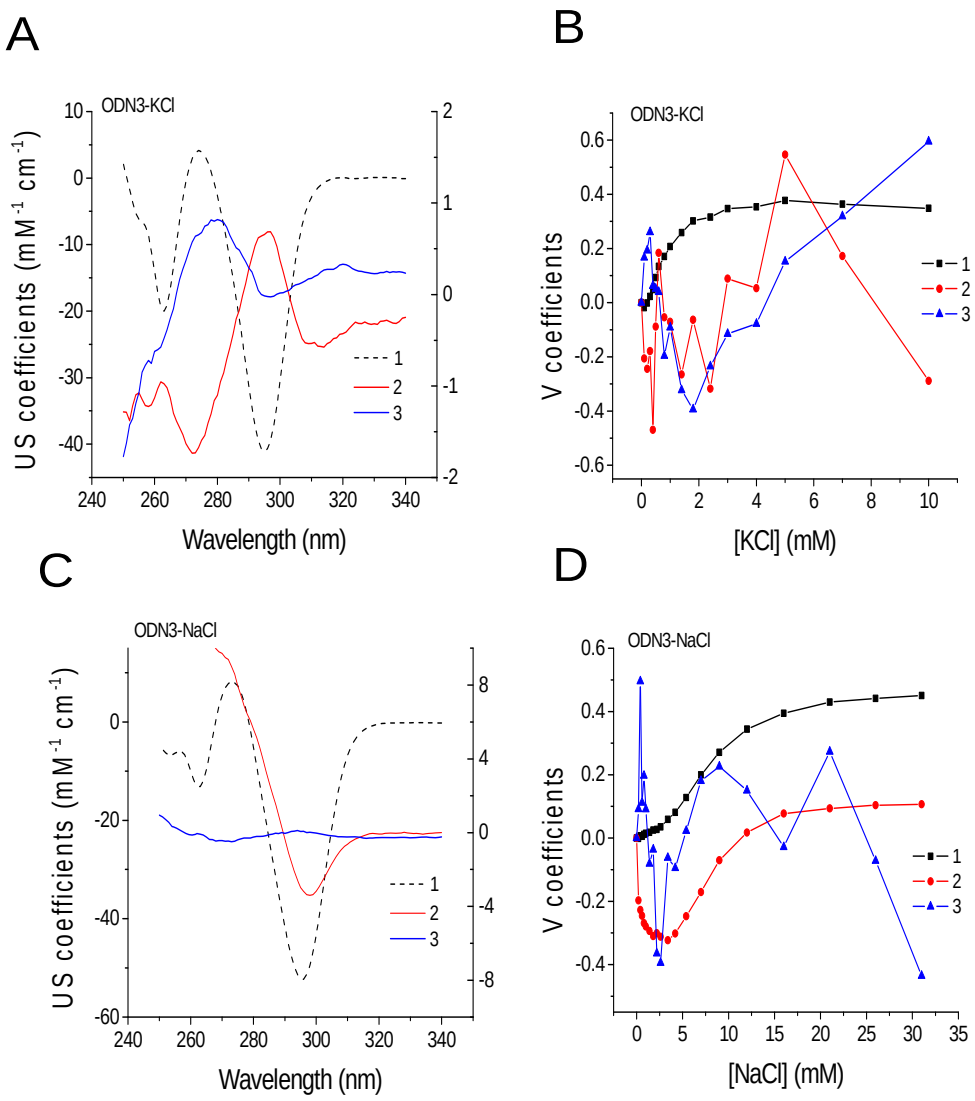

Figure S4. Basis spectra (panels A and C) and amplitude vectors (panels B and D) the three most significant components in the titrations of ODN3 with KCl and NaCl. The data in the figures was derived by SVD analysis of the difference spectra in Figure S1.

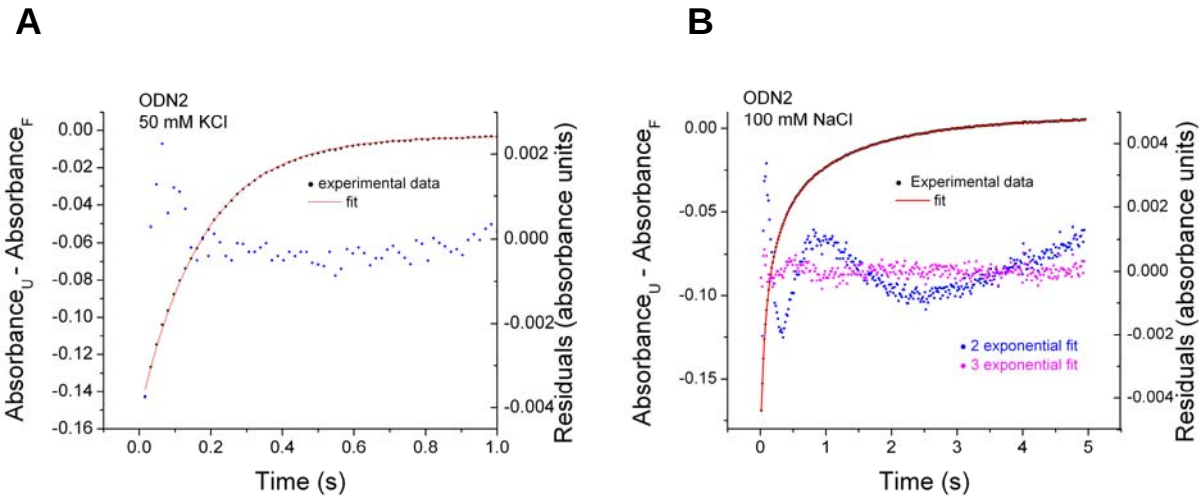

Figure S5. Panel A**:** comparison of the experimental reaction progress curve measured at 295 nm for KCl-induced folding of ODN2 (4.8 µM) with the simulated curve generated by global fitting of the wavelength-time course data matrix to a single exponential process. The residual plot indicates that a single exponential fit adequately reproduces the experimental data. The best fitting rate constant for the experiment shown is k_1 = 4.6 ± 0.2 s^{-1}. Panel B: comparison of the experimental reaction progress curve measured at 295 nm with the simulated curve generated by global fitting of the wavelength-time course data matrix to two- and three-consecutive exponential processes for NaCl-induced folding of ODN2 (4.8 µM). The best fit of the experimental data requires three exponentials as illustrated by the residual plots. The optimized values of the rate constants for this data set are: k_1 =14.5 ± 0.6 s^{-1}, k_2 = 3.0 ± 0.2 s^{-1} and k_3 = 0.7 ± 0.1 s^{-1}. Temperature = 25 ºC.

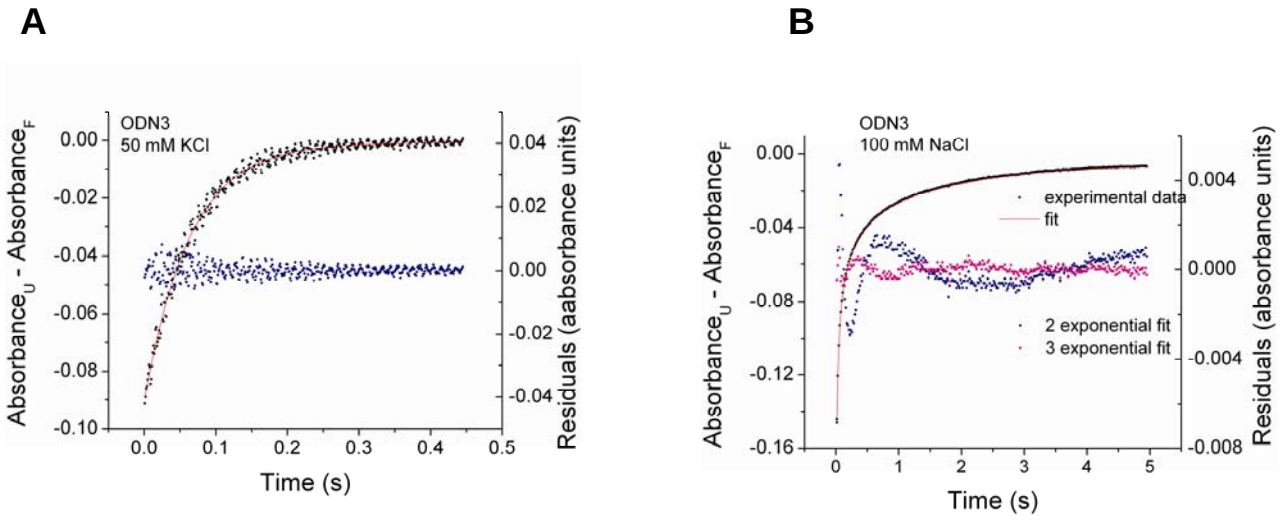

Figure S6. Panel A**:** comparison of the experimental reaction progress curve measured at 295 nm for KCl-induced folding of ODN3 (4.0 μM) with the simulated curve generated by global fitting of the wavelength-time course data matrix to a single exponential process. The best fitting rate constant for the experiment shown is k_1 = 15.8 ± 0.3 s^{-1}. Panel B: comparison of the experimental reaction progress curve measured at 295 nm with the simulated curve generated by global fitting of the wavelength-time course data matrix to two- and three-consecutive exponential processes for NaCl-induced folding of ODN3 (4.0 μM) . The best fit of the experimental data requires three exponentials as illustrated by the residual plots. The optimized values of the rate constants for this data set are: k_1 =26.2 ± 1.4 s^{-1}, k_2 = 4.0 ± 0.3 s^{-1} and k_3 = 0.67 ± 0.03 s^{-1}. Temperature = 25 ^{o}C.

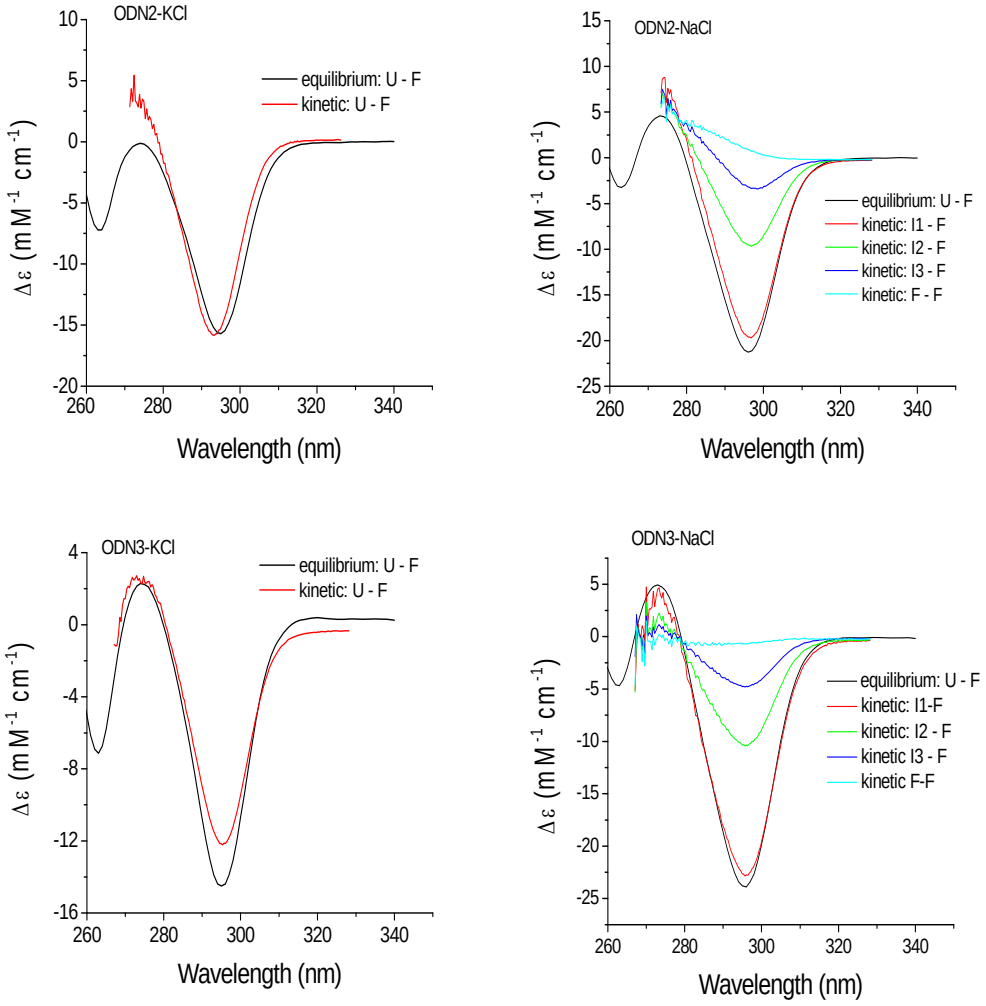

Figure S7. Comparison of absorbance changes associated with cation-dependent G-quadruplex folding as determined by equilibrium titrations (black lines) and by global fitting of the kinetic data to single (K^+) and consecutive triple (Na^+) exponentials using the program Specfit32. The difference spectra were derived from the data sets in Figures S5 and S6.

Table S1. Analysis of cation titrations of human telomere model oligonucleotides by singular value decomposition

ODN	Cation	Component	Singular values	C_i [U]	C_i[V]
1	K^+	1	182.793	0.9880	0.9235
		2	9.714	0.9640	0.7441
		3	2.822	0.9643	0.5234
		4	1.601	0.9566	-0.2355
2	K^+	1	138.337	0.9813	0.9052
		2	13.979	0.9658	0.7904
		3	4.174	0.9801	0.7920
		4	1.688	0.9688	-0.1205
3	K^+	1	150.72	0.9922	0.9318
		2	8.19	0.9826	0.3012
		3	5.21	0.9336	0.6252
		4	1.95	0.9074	0.0907
1	Na^+	1	328.363	0.993	0.9549
		2	21.453	0.9602	0.6807
		3	7.127	0.9425	0.3800
		4	3.492	0.9629	0.1255
2	Na^+	1	197.454	0.9887	0.9049
		2	6.550	0.9764	-0.0214
		3	3.367	0.9361	0.5592
		4	1.744	0.8991	0.7843
3	Na^+	1	197.46	0.9941	0.8870
		2	64.59	0.9633	0.9580
		3	2.77	0.9237	0.1004
		4	2.06	0.9465	0.2128

C_i [U] = autocorrelation coefficient of columns of the U matrix
C_i [V] = autocorrelation coefficient of columns of the V matrix
C_i [V] values > 0.5, the cutoff value chosen for significance in this study, are given in blue typeface. The values were derived from SVD analysis of the data sets shown in Figure 2 and Figure S1.

REFERENCES

1. Hendler,R.W. and Shrager,R.I. (1994) Deconvolutions based on singular value decomposition and the pseudoinverse: a guide for beginners. *J. Biochem. Biophys. Methods*, 28, 1-33.

2. Sheardy,R.D., Suh,D., Kurzinsky,R., Doktycz,M.J., Benight,A.S. and Chaires,J.B. (1993) Sequence dependence of the free energy of B-Z junction formation in deoxyoligonucleotides. *J. Mol. Biol.*, 231, 475-488.