Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry

Polymorphism and resolution of oncogene promoter quadruplex-forming sequences.

M. Clarke Miller, Huy T. Le, William L. Dean, Patrick A. Holt, Jonathan B. Chaires, and John O. Trent*

Department of Medicine, Clinical Translational Research Building, University of Louisville, 505 Hancock St., Louisville, KY 40202, USA.

Email: john.trent@louisville.edu

Supplementary Materials:

Contents:

Methods:

Preparation of samples. Oligos were obtained from IDT as dry, de-salted pellets. The purified oligonucleotides were dissolved in water and dialyzed into buffer containing 185 mM KCl, 8 mM KH_2PO_4, 2 mM K_2HPO_4 (pH= 7.0) 1mM EDTA using Slide-A-Lyzer dialysis cassettes (MWCO 2,000, Thermo Scientific) and diluted to a final concentration of approximately 200 µM. Samples were then equilibrated for ten minutes at 100°C in a water bath, followed by gradual cooling to room temperature overnight. The DNA samples were then refrigerated until use.

Size Exclusion Chromatography. A Superdex 75 10/300 column (GE Healthcare) was installed on an Waters 600 pump equipped with 2998 photodiode array detector, 2707 autosampler and a Waters Fraction Collector III. The column was prepared with a mobile phase consisting of 100 mM KCl, 25 mM K_2HPO_4 (pH=8.0) and calibrated as previously described[1]. For DNA separation, the flow rate was altered to 0.05 ml/min. DNA was injected at 10 to 100 µl aliquots.

Analytical Ultracentrifugation. Analytical ultracentrifugation was carried out in a Beckman Coulter ProteomeLab XL-A analytical ultracentrifuge at 20^o C overnight at 50,000 rpm in standard 2 sector cells. Data were analyzed using both DCDT+ (version 2.3.2, John Philo, Thousand Oaks CA) and Sedfit (free software: www.analyticalultracentrifugation.com) software packages. Excellent agreement between the two software packages for s20,w values and species distributions was achieved. Buffer density was determined on a Mettler/Paar Calculating Density Meter DMA 55A at 20.0^o C and buffer viscosity was calculated using Sednterp software (free software: www.jphilo.mailway.com).

NMR sample preparation and spectroscopy. The c-kit 2kj2 quadruplex was separated with a mobile phase consisting of 20 mM KCL and 5 mM K_2HPO_4 (pH=8.0). Fractions were collected at 0.05 ml intervals. The five fractions corresponding to the maxima for the major peak were collected and combined over five separate runs. HPLC fractions were kept frozen at -80 C between runs. This material was then frozen in liquid nitrogen and lyophilized. The lyophilized material was then re-dissolved in 330 µl water with ≈30 µM DSS and 10% D_2O and loaded into a 5 mm Shigemi NMR tube. The un-separated c-kit 2kj2 sample was prepared by adding 10% D_2O and DSS to 100 µM to the sample produced by the preparation described above and loading into a 5 mm Shigemi NMR tube. NMR spectra were recorded using a 5 mm inverse triple resonance (HCN) probe on Varian Inova spectrometer at 14.1 T using a cold probe.

Figure S1: AUC data for the promoter sequences chosen for this study. Each sequence shows significant polymorphism. A. c-kit 2KJ2, B. c-kit 2O3M, C. HIF-1α, D. HIF-1α without flanking sequences, E. Retinoblastoma, F. c-myc, G. Her2, H. bcl-2, I. VEGF, and J. KRAS. Each sequence results in a mixture of species as evidenced by broadened or multiple peaks.

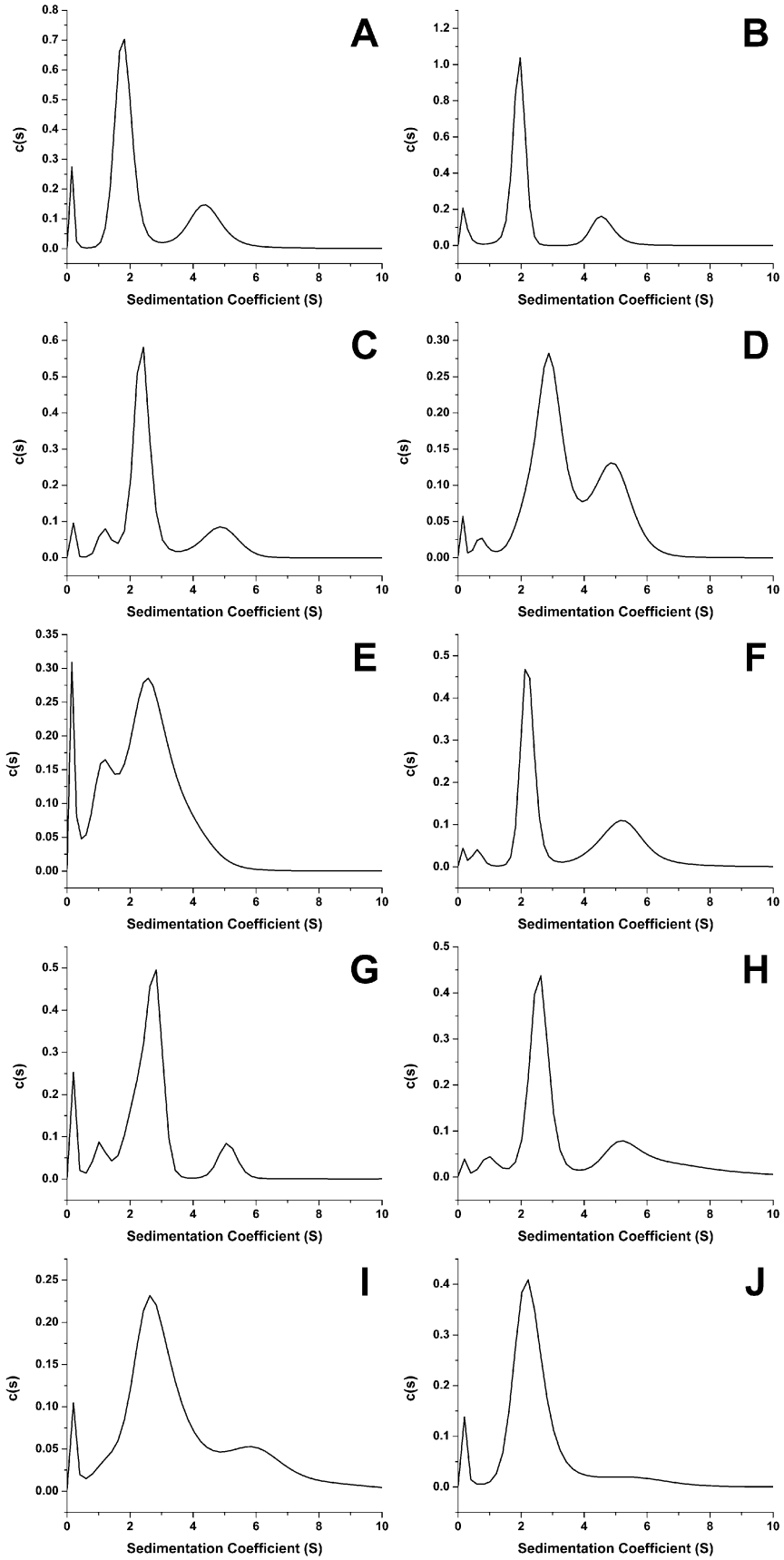


Figure S2: c-kit 2KJ2 fraction 2 re-injected and compared to the original separation data. After concentration by lyophilization, re-dissolving, and use for collecting NMR and AUC data, the samples of c-kit fractions 1, 2, and 3 (RED) were re-injected to verify that they possessed the same elution characteristics as seen in the original separation of the parent mixture (BLACK). Fraction 1 and 2 showed the same overall elution characteristics with only a small amount of contamination by species that eluted at a different volume. Fraction 3 showed the most deviation from the parent mixture. The peak eluting corresponding to fraction 3, approximately 13.66 ml, shows a broadened profile and a significant amount of re-equilibration to the form putatively found in fraction 2. This is supported by the results of the AUC analysis (Figure 3) Conversely, the shift from Fraction 2 to Fraction 3 seems to be much less favored.

.

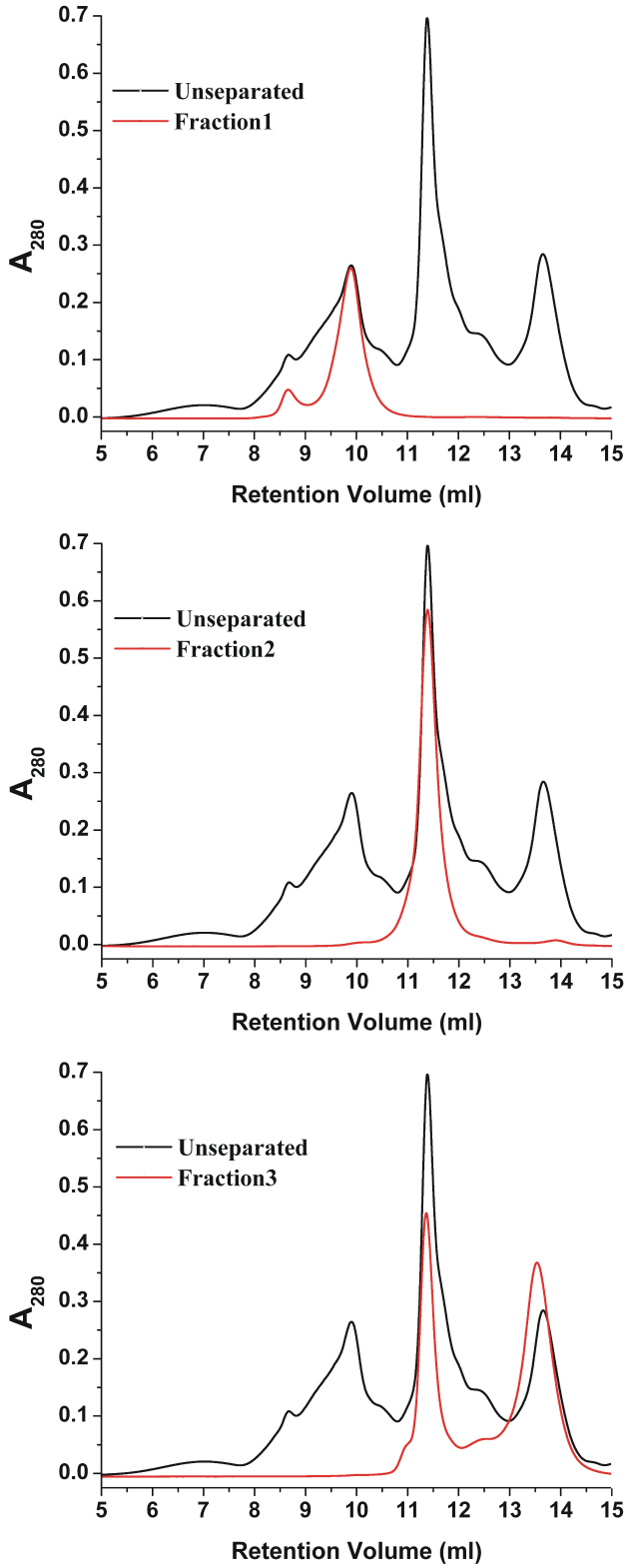
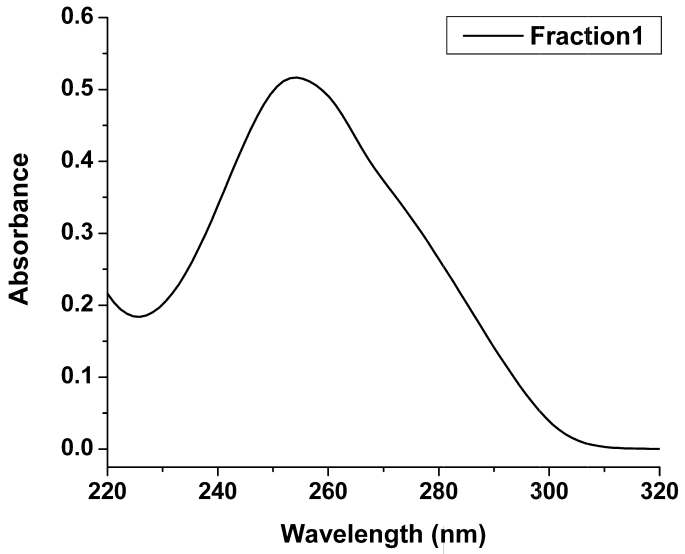
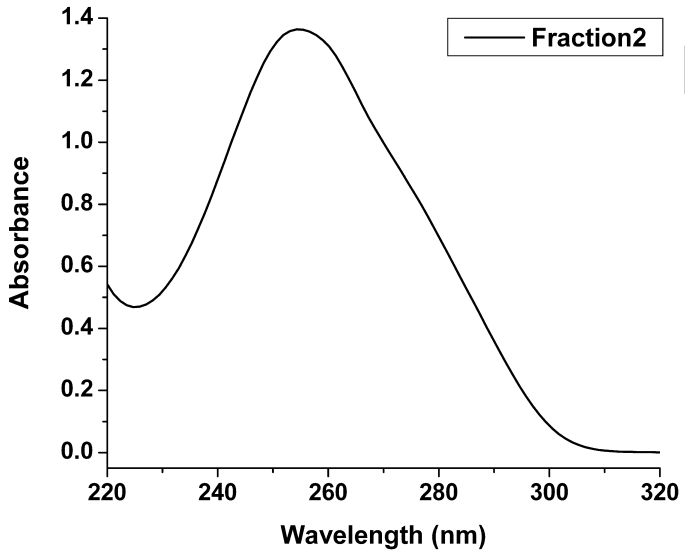


Figure S3: Comparison of UV-vis spectra of each c-kit 2KJ2 fraction. Each fraction has a similar UV-vis signature. The UV data was extracted directly from the HPLC data.





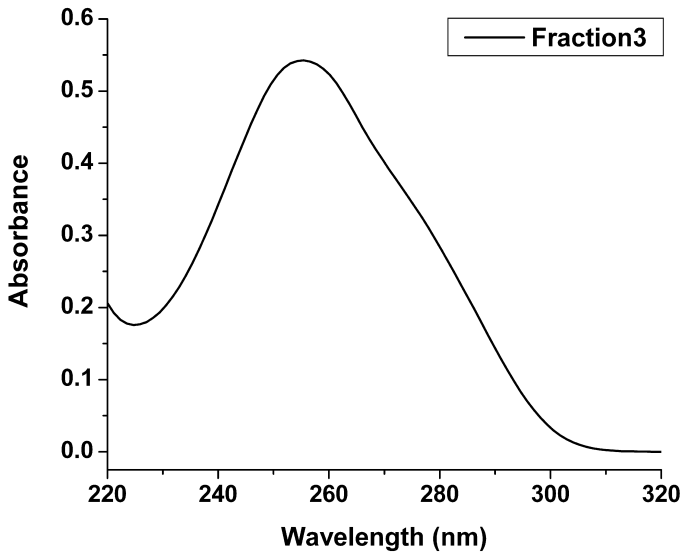


Figure S4: CD data for each of the isolated fractions of c-kit. Each CD spectra is similar, yet shows a slightly different maximum near 260.

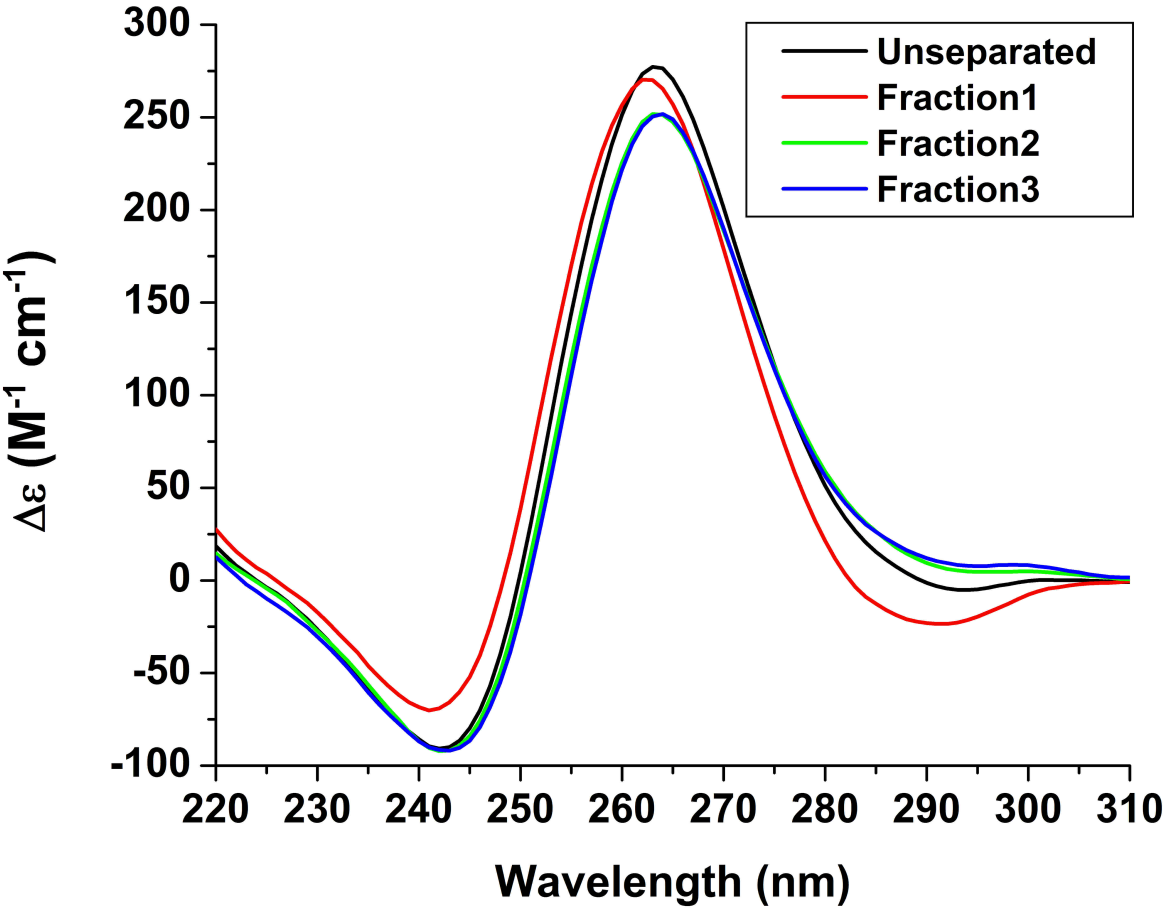


Figure S5: Calibration of the Superdex 75 10/300 column. LEFT: A mixture of proteins of known molecular weights resolved on the column. The protein standards were: 1. conalbumin (75 kDa, V_R=9.36), 2. ovalbumin (43 kDa, V_R=10.19), 3. carbonic anhydrase (29 kDa, V_R=11.32), 4. ribonuclease A (13.7 kDa, V_R=12.84), 5. bovine lung aprotonin (6.5 kDa, V_R=14.66). RIGHT: The linear calibration curve developed by graphing V_R/V_0 vs. Log MW. V_0 was determined to be 7.61 ml by injection of blue dextran (MW = 2,000,000 Da). The result of the linear fit is at the bottom left of the graph. If the column is calibrated by comparing R_S with V_R/V_0 then the linear fit becomes: y = -42.144x + 89.124, R^2 = 0.9617.

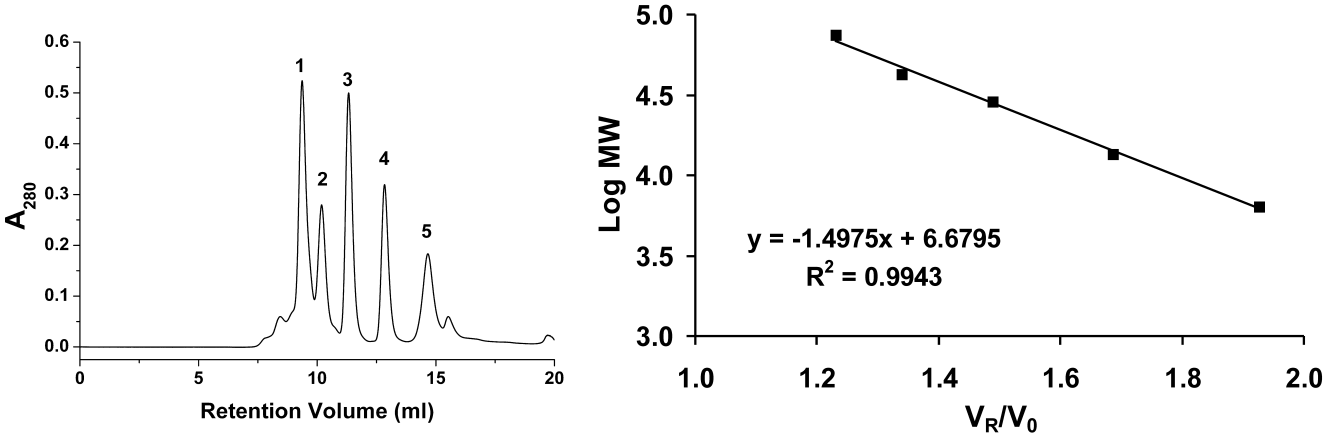


Figure S6: Size exclusion chromatography of several human telomere sequences. A. wtTel22, B. Tel26 (Hybrid 1), and C. wtTel26 (Hybrid 2). The human telomere sequences examined were not resolvable via SEC. Even sequences that are known to form 2 species in solution were irresolvable.

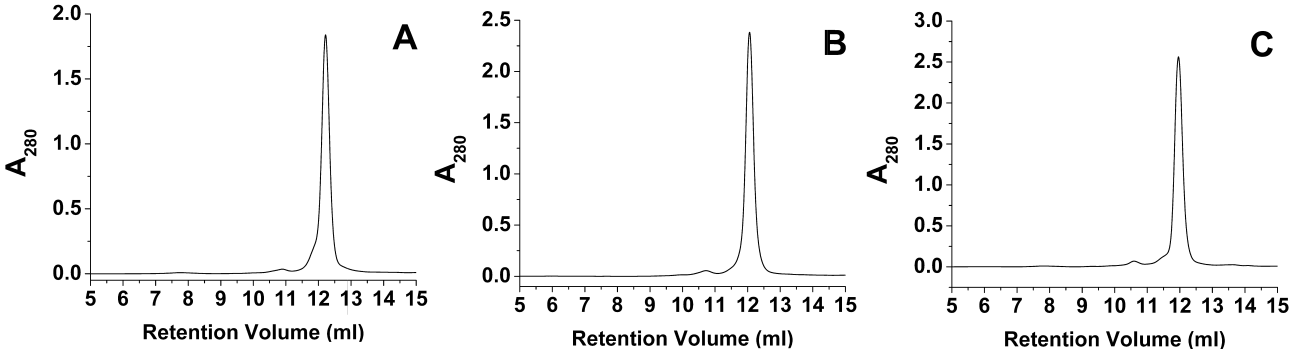


Table S1: Calculated Stokes Radius (R_S) for several quadruplex structures. The R_S was calculated using the available structural information and Hydropro(REF).

Sequence Name:	Sequence:	R_S (Å)	PDB ID:	Reference:
wtTel22	AGGGTTAGGGTTAGGGTTAGGG	15.47	143D	[2]
mutTel24	TTGGGTTAGGGTTAGGGTTAGGGA	15.85	2GKU	[3]
Tel26	AAAGGGTTAGGGTTAGGGTTAGGGAA	15.23	2HY9	[4]
wtTel26	TTAGGGTTAGGGTTAGGGTTAGGGTT	15.13	2JPZ	[5]
wtTel25b	TAGGGTTAGGGTTAGGGTTAGGGTT	15.92	2JSL	[6]
wtTel23	TAGGGTTAGGGTTAGGGTTAGGG	15.65	2JSM	[6]
c-kit	CGGGCGGGCGCGAGGGAGGGT	16.67	2KQH	[7]
c-kit	AGGGAGGGCGCTGGGAGGAGGG	15.57	2O3M	[8]

Table S2: Stokes Radius (R_S) of molecular weight standards used for column calibration.

Protein:	Molecular Weight (g/mole):	R_S (Å)	Reference:
Aprotinin	6500	9.83	[9]
Ribonuclease A	13700	16.4	[10]
Carbonic Anhydrase	29000	25.5	[11]
Ovalbumin	43000	30.5	[10]
Conalbumin	75000	40.4	[11]

1. M. M. Dailey, M. C. Miller, P. J. Bates, A. N. Lane and J. O. Trent, Nucleic Acids Res., 2010, 38, 4877-4888.
2. Y. Wang and D. J. Patel, Structure, 1993, 1, 263-282.
3. K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix and D. J. Patel, J Am Chem Soc, 2006, 128, 9963-9970.
4. J. Dai, C. Punchihewa, A. Ambrus, D. Chen, R. A. Jones and D. Yang, Nucleic Acids Res, 2007, 35, 2440-2450.
5. J. Dai, M. Carver, C. Punchihewa, R. A. Jones and D. Yang, Nucleic Acids Res, 2007, 35, 4927-4940.
6. A. T. Phan, V. Kuryavyi, K. N. Luu and D. J. Patel, Nucleic Acids Res., 2007, 35, 6517-6525.
7. S. T. D. Hsu, P. Varnai, A. Bugaut, A. P. Reszka, S. Neidle and S. Balasubramanian, Journal of the American Chemical Society, 2009, 131, 13399-13409.
8. A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle and D. J. Patel, J Am Chem Soc, 2007, 129, 4386-4392.
9. J. C. T. Carlson, A. Kanter, G. R. Thuduppathy, V. Cody, P. E. Pineda, R. S. McIvor and C. R. Wagner, Journal of the American Chemical Society, 2003, 125, 1501-1507.
10. G. D. Fasman, *Practical handbook of biochemistry and molecular biology*, CRC Press, Boca Raton, Fla., 1989.
11. W. M. Southerland, D. R. Winge and K. V. Rajagopalan, Journal of Biological Chemistry, 1978, 253, 8747-8752.