Electronic Supplementary Material (ESI) for ChemComm.

Applicability of Nearest-neighbour Model for Pseudoknot RNAs

Sagar Satpathi,[1] Tamaki Endoh[1] and Naoki Sugimoto[1,2,*]

[1]Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 7-1-20 Minatojima-minamimachi, Kobe, 650-0047, Japan
[2]Graduate School of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, 7-1-20 Minatojima-minamimachi, Kobe, 650-0047, Japan

*To whom correspondence should be addressed. Tel: +81-78-303-1416; Fax: +81-78-303-1495; Email: sugimoto@konan-u.ac.jp

Materials and Methods

Oligonucleotides and chemical reagents

All RNA oligonucleotides (HPLC purification grade) listed in Table S1 were purchased from Fasmac Co., Ltd. Concentrations of the RNA oligonucleotides were calculated from the absorbance at 260 nm in the temperature range of 80–90 °C, utilising their single-strand molar extinction coefficient. Sodium chloride (NaCl) and disodium hydrogen phosphate (Na_2HPO_4) were purchased from FUJIFILM Wako Pure Chemical Corporation (Osaka, Japan). Disodium ethylenediaminetetraacetate (Na_2EDTA) was purchased from Dojindo Molecular Technologies (Kumamoto, Japan). The pH of the Na_2HPO_4 buffer was adjusted to 7.0 using HCl purchased from FUJIFILM Wako Pure Chemical Corporation (Osaka, Japan).

Circular dichroism (CD) measurements

Circular dichroism (CD) was analysed using a JASCO J-1500 spectropolarimeter equipped with a temperature controller. CD spectra of RNA oligonucleotides at 20 μM were collected in a buffer containing 10 mM Na_2HPO_4 (pH 7.0), 0.1 M NaCl, and 1 mM Na_2EDTA in 0.1 cm path-length cuvettes at a scan rate of 100 nm/min. For CD melting studies, RNA oligonucleotides were incubated at 70 °C for 5 min and cooled to 0 °C at 1 °C min^{-1} to refold their structure. Melting profiles of the RNA oligonucleotides at 210 nm were obtained by heating the samples from 0 to 95 °C at a rate of 0.2 °C min^{-1}.

UV melting measurements

UV absorbance at 260 nm of RNA oligonucleotides (2 μM) was measured in a buffer containing 10 mM Na_2HPO_4 (pH 7.0), 0.1 M NaCl, and 1 mM Na_2EDTA, using a Shimadzu 1800 UV/Vis spectrophotometer equipped with a temperature controller. The melting profiles

were collected by heating the samples from 0 to 95 °C at a rate of 0.2 °C min^{-1}. RNA oligonucleotides were refolded from 70 to 0 °C before collecting the melting profiles. The condensation of water on the cuvette exterior at lower temperatures was avoided by blowing a constant stream of dry N_2 gas.

Thermodynamic analysis

To extract the melting profile of the S2 stem region, the CD and/or UV melting profiles of the control HP were subtracted from the respective PK melting profiles. The sigmoidal melting profiles obtained were normalised in the temperature range from 0 to 80 °C and fitted to obtain thermodynamic parameters for the S2 stem region in PKs using Equation (1).

$$Signal_{(normalized)} = (a \times (T + 273.15) + b) + \frac{((c \times (T + 273.15) + d) - (a \times (T + 273.15) + b))}{1 + e^{\left(\frac{\Delta H^{\circ}}{R \times (T + 273.15)} - \frac{\Delta S^{\circ}}{R}\right)}} \quad (1)$$

where 'a' and 'b' are the slope and intercept of the lower baseline for the melting curve, respectively, 'c' and 'd' represent the slope and intercept of the upper baseline for the melting curve, respectively, *T* is the temperature in Kelvin, *R* is gas constant, and Δ*H*º and Δ*S*º are the enthalpy and entropy changes associated with the S2 region melting, respectively.

The calculated Δ*H*º and Δ*S*º were utilised to obtain the free energy change at 37 °C (Δ*G*°$_{37}$) using Equation (2).

$$\Delta G_{37}^{\circ} = \Delta H^{\circ} - T\Delta S^{\circ} \quad (2)$$

where *T* is 310.15 K.

The melting temperature in °C (T_m) was obtained using Equation (3).

$$T_m = \frac{\Delta H^{\circ}}{\Delta S^{\circ}} - 273.15 \quad (3)$$

Prediction of Δ*G*°$_{37}$ for the RNA duplex at 100 mM NaCl condition

Predicted Δ*G*°$_{37}$ value of RNA duplex at 1000 mM NaCl condition (Predicted Δ*G*°$_{37}$ (1000 mM)) was calculated following the nearest-neighbour (NN) parameters, in which parameters were obtained in a buffer containing 1000 mM NaCl, 10 or 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7.0 .[2] Based on the linear correlation between the thermodynamic stabilities of RNA duplexes in the presence of 1000 mM and 100 mM NaCl that has been experimentally

demonstrated,[3] the predicted $\Delta G°_{37}$ value at condition with 100 mM NaCl (Predicted $\Delta G°_{37}$ (100 mM)) was recalculated by using an equation of $\Delta G°_{37}$ (100 mM) = 0.630 $\Delta G°_{37}$ (1000 mM) -1.667.[3] The process to calculate both Predicted $\Delta G°_{37}$ (1000 mM) and Predicted $\Delta G°_{37}$ (100 mM) are shown in Scheme S2.

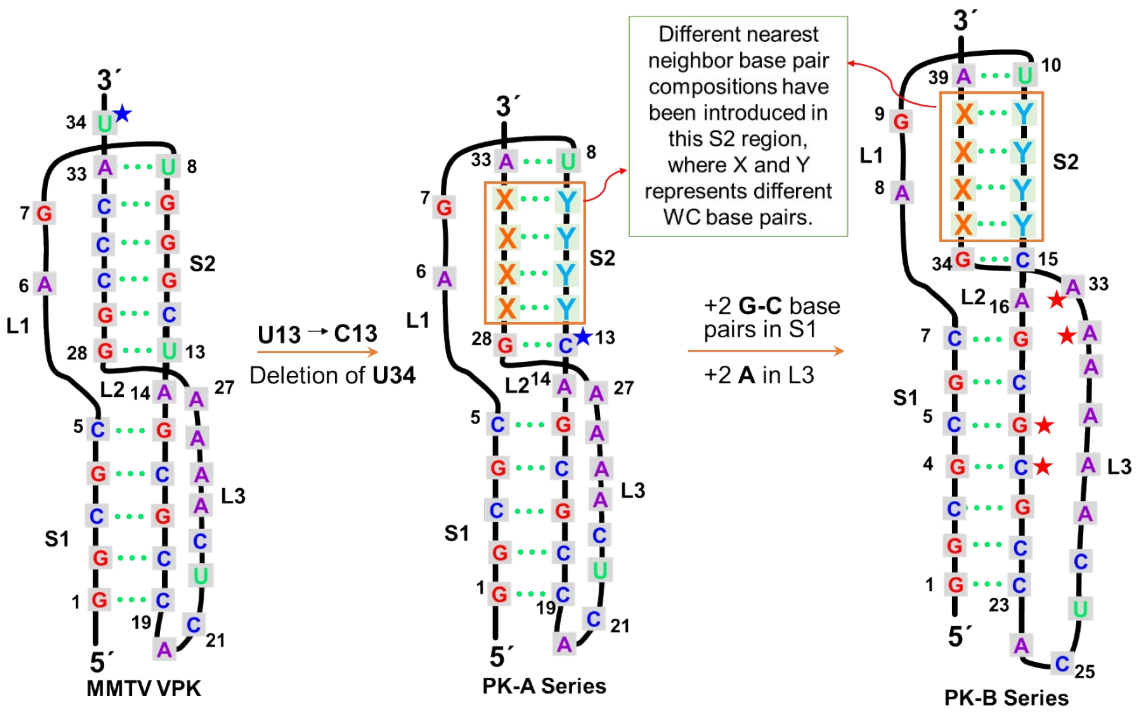


Figure S1. Schematic representation for designing the analysed PKs. Designed PKs were derived from MMTV VPK (left). The PK-A series was designed to have different base pair compositions at the S2 stem region of MMTV VPK with the replacement of U13 to C13 (blue asterisk) and removal of U34 (blue asterisk). The PK-B series was designed to have a long S1 stem. Base pairs inserted in the S1 stem and L3 regions are indicated by red asterisks.

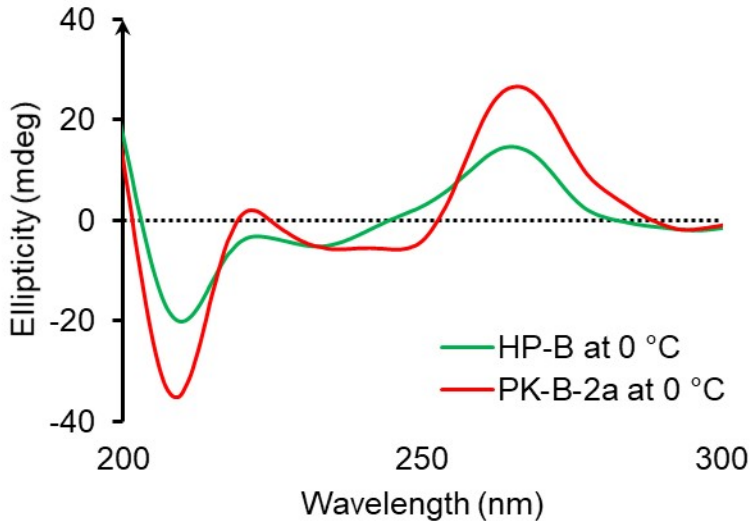


Figure S2. CD spectra of PK-B-2a (red) and HP-B (green) at 0 °C in a buffer containing 10 mM Na_2HPO_4 (pH 7), 100 mM NaCl, and 1 mM Na_2EDTA. Oligonucleotides were analysed at a concentration of 20 μM.

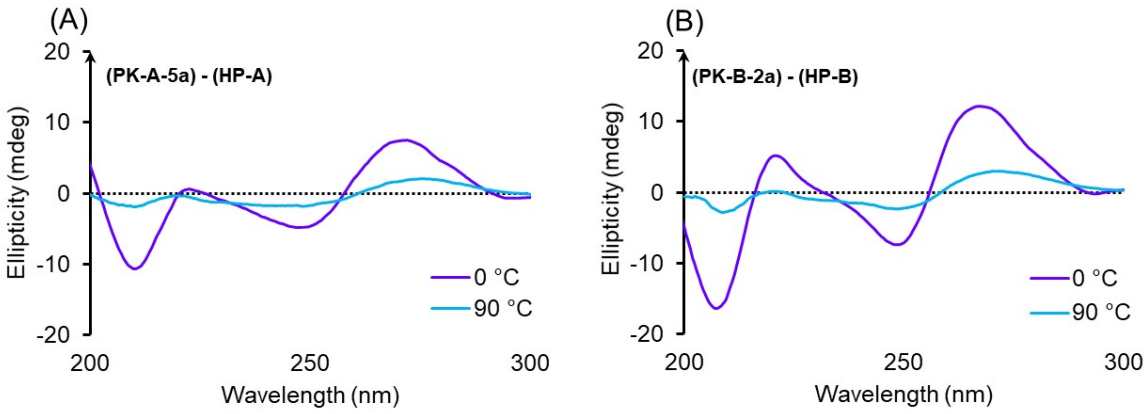
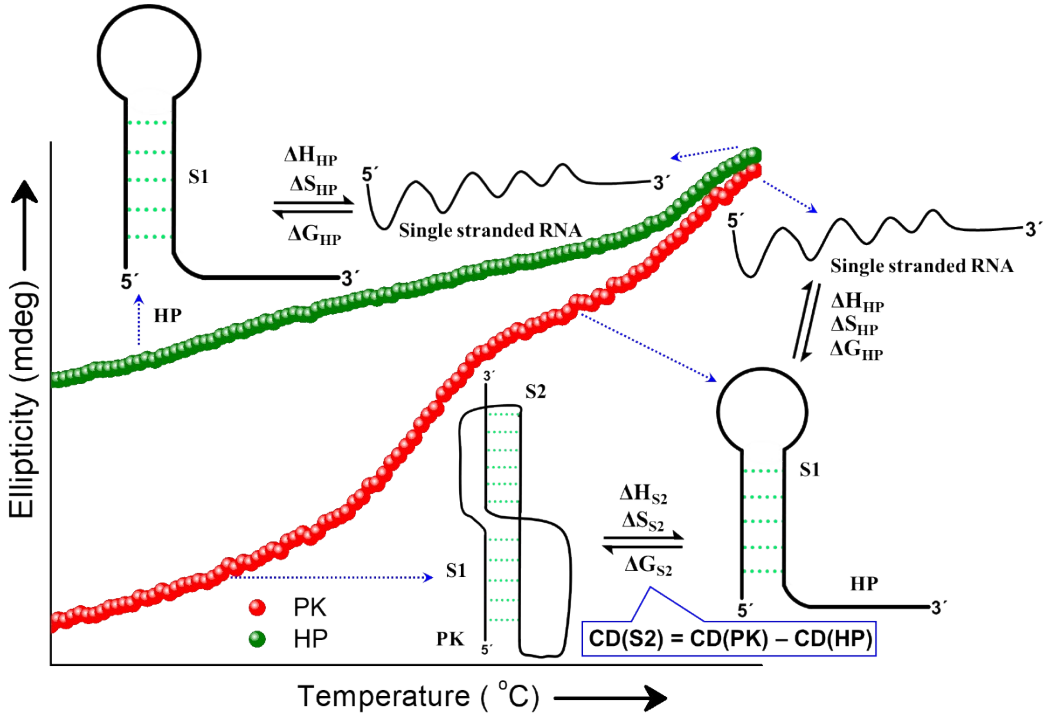


Figure S3. Differential CD spectra of (A) PK-A-5a and (B) PK-B-2a from HP-A and HP-B, respectively. CD spectra were measured at 0 ºC (violet) and 90 ºC (cyan) in a buffer containing 10 mM Na_2HPO_4 (pH 7), 100 mM NaCl, and 1 mM Na_2EDTA.



Scheme S1. Schematic representation of different states undergoing the CD melting of PK and HP.

ΔH_{S2}, ΔS_{S2}, and ΔG_{S2} are the enthalpy, entropy, and free energy change associated with the transition between the PK and HP structures, respectively.

ΔH_{HP}, ΔS_{HP}, and ΔG_{HP} are the enthalpy, entropy, and free energy change associated with the transition between HP and single-stranded RNA structures, respectively.

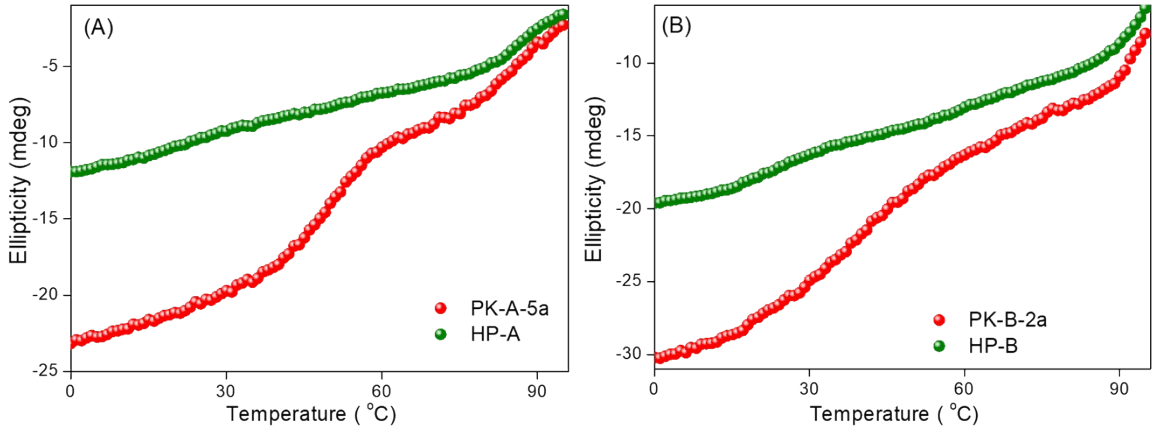


Figure S4. CD melting profiles at 210 nm of (A) PK-A-5a (red) and HP-A (green) and (B) PK-B-2a (red) and HP-B (green) in a buffer containing 10 mM Na_2HPO_4 (pH 7), 100 mM NaCl, and 1 mM Na_2EDTA. Oligonucleotides were analysed at 20 μM.

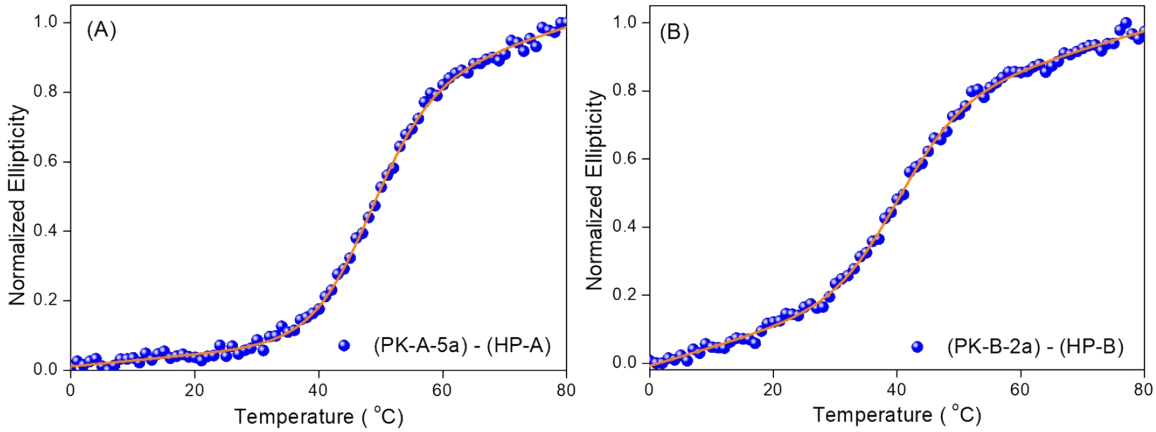


Figure S5. Normalised CD melting profiles at 210 nm of (A) PK-A-5a after subtracting the HP-A signal and (B) PK-B-2a after subtracting the HP-B signal. Orange lines in the figures correspond to the fitted line obtained using Equation (1).

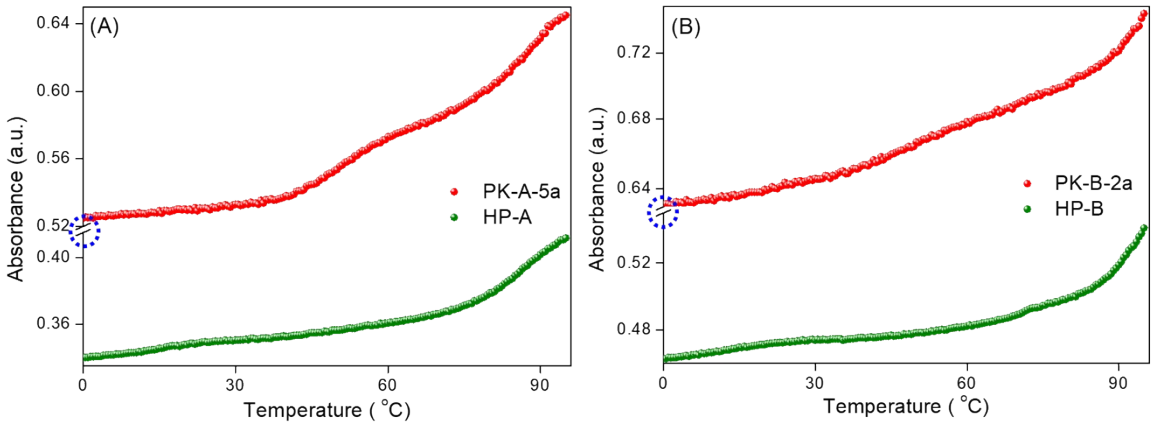


Figure S6. UV melting profiles at 260 nm of (A) PK-A-5a (red) and HP-A (green) and (B) PK-B-2a (red) and HP-B (green) in a buffer containing 10 mM Na_2HPO_4 (pH 7), 100 mM NaCl, and 1 mM Na_2EDTA. Oligonucleotides were analysed at 2 μM. The blue circles in both the figures indicate the breaking point in the Y-axis.

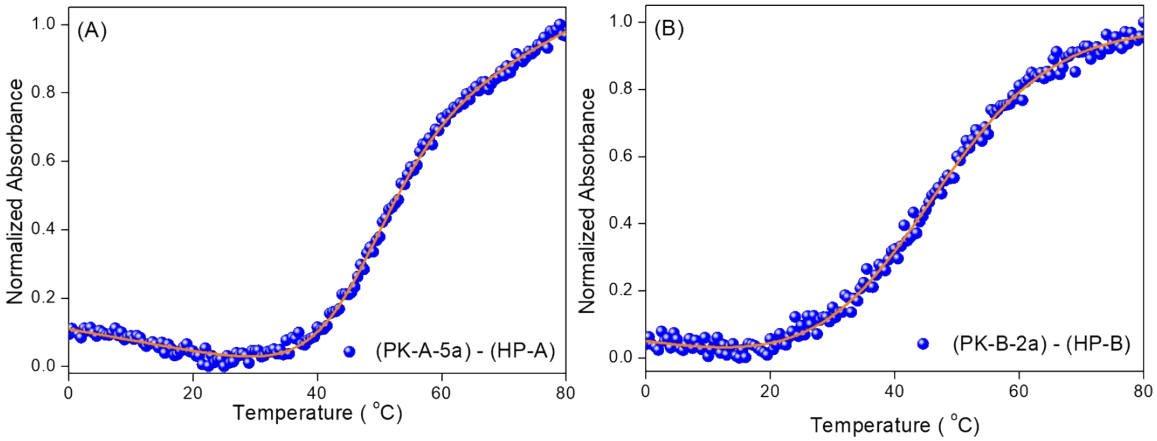
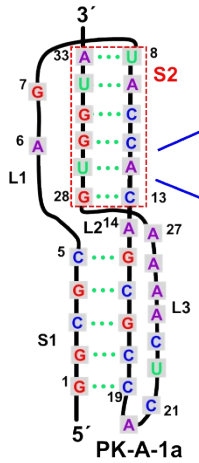


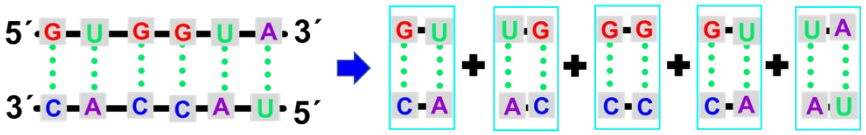
Figure S7. Normalised UV melting profiles at 260 nm of (A) PK-A-5a after subtracting the HP-A signal and (B) PK-B-2a after subtracting the HP-B signal. Orange lines in the figures correspond to the fitted line obtained using Equation (1).



Calculated the measured free energy change, ΔG°_{37} from CD and UV melting studies at 100 mM NaCl.

Measured ΔG°_{37} at 100 mM NaCl

If we consider the same S2 region is forming the conventional RNA duplex, then we can predict the free energy change of that RNA duplex i.e., **Predicted ΔG°_{37}**. A comparison between **Measured and Predicted ΔG°_{37}** highlights the difference and correlation between a duplex in PK and conventional canonical structure.



Predicted ΔG^{o}_{37} (1000 mM)

$$= \sum_j n_j \Delta G^{o}_{37,\,NN(j)} + \Delta G^{o}_{init} + m_{A\text{-}U\ terminal}\,\Delta G^{o}_{A\text{-}U\ terminal} + \Delta G^{o}_{37,\,sym}$$

$$= 2 \times \Delta G^{o}_{37,\,(GU/CA)} + \Delta G^{o}_{37,\,(UG/AC)} + \Delta G^{o}_{37,\,(GG/CC)} + \Delta G^{o}_{37,\,(UA/AU)} + \Delta G^{o}_{init} + m_{A\text{-}U\ terminal}\,\Delta G^{o}_{A\text{-}U\ terminal} + \Delta G^{o}_{37,\,sym}$$

= -6.64 kcal/mol

where $\Delta G^{o}_{37,\,NN(j)}$ is the free energy change for base pair formation between two adjacent pairs, ΔG^{o}_{init} is the free energy change for helix initiation, $\Delta G^{o}_{A\text{-}U\ terminal}$ is the free energy change for the terminal A-U base pair and $\Delta G^{o}_{37,\,sym}$ is the symmetry correction which is 0.43 kcal/mole and zero for self-complementary and non-self-complementary duplexes, respectively. $m_{A\text{-}U}$ is the number of A-U terminals in the duplex.

ΔG^{o}_{37} at 100 mM NaCl was calculated using the following equation:

Predicted ΔG^{o}_{37} (100 mM) = 0.630 × Predicted ΔG^{o}_{37} (1000 mM) -1.667

Scheme S2. Schematic representation for the calculation of predicted free energy change of the same S2 region in PK considering the same sequence forming the conventional RNA duplex. Predicted ΔG°_{37} at 100 and 1000 mM NaCl have been calculated using the established NN parameters of the RNA duplex.[2,3]

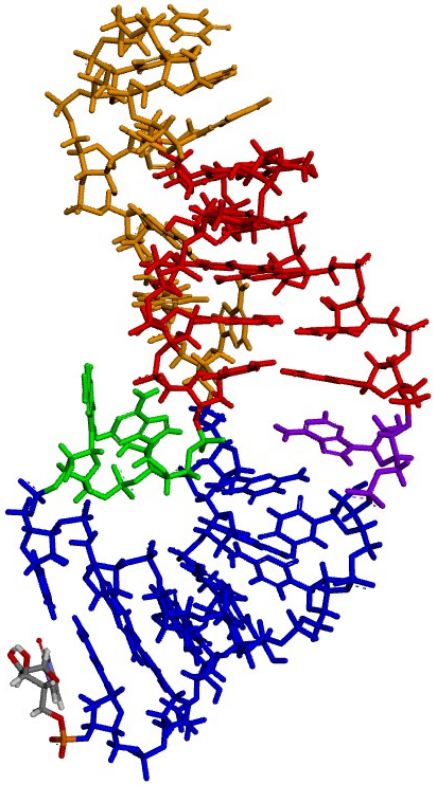


Figure S8. Tertiary structure of MMTV VPK reported in PDB ID - 1RNK.[1] Different colours have been assigned to the different regions of PK: S1 stem (red), L1 loop (green), S2 stem (blue), L2 loop (purple), and L3 loop (yellow).

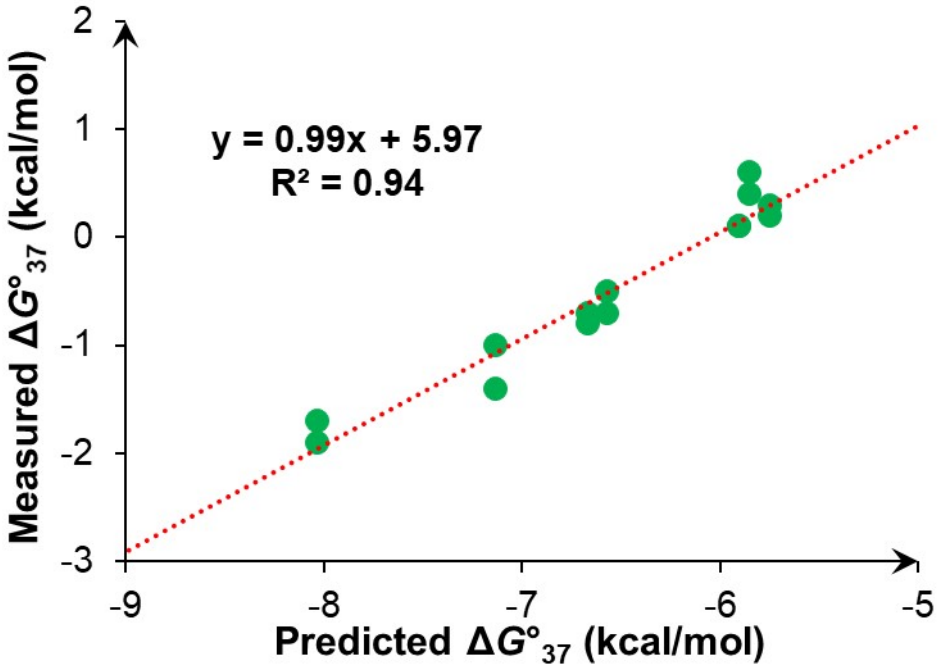


Figure S9. Linear correlation between predicted and experimentally obtained ΔG°_{37}. ΔG°_{37} of the S2 stem of the respective PK structures were plotted with the predicted ΔG°_{37} of the same S2 stems that formed the canonical RNA duplex in the presence of 100 mM NaCl.

Table S1. RNA sequences used in the study

RNA Name [a]	RNA Sequences (5´ to 3´) [b]
MMTV VPK	GGCGCAGUGGGCUAGCGCCACUCAAAAGGCCCAU
PK-A-1a	GGCGCAGUACCACAGCGCCACUCAAAAGUGGUA
PK-A-1b	GGCGCAGUGGUACAGCGCCACUCAAAAGUACCA
PK-A-2a	GGCGCAGUUCCGCAGCGCCACUCAAAAGCGGAA
PK-A-2b	GGCGCAGUUCGGCAGCGCCACUCAAAAGCCGAA
PK-A-3a	GGCGCAGUAGGUCAGCGCCACUCAAAAGACCUA
PK-A-3b	GGCGCAGUCCUACAGCGCCACUCAAAAGUAGGA
PK-A-4a	GGCGCAGUUAGGCAGCGCCACUCAAAAGCCUAA
PK-A-4b	GGCGCAGUAAGCCAGCGCCACUCAAAAGGCUUA
PK-A-5a	GGCGCAGUCACCCAGCGCCACUCAAAAGGGUGA
PK-A-5b	GGCGCAGUGGUCCAGCGCCACUCAAAAGGACCA
PK-A-6a	GGCGCAGUGCGUCAGCGCCACUCAAAAGACGCA
PK-A-6b	GGCGCAGUGUCGCAGCGCCACUCAAAAGCGACA
PK-A-7a	GGCGCAGUCGGGCAGCGCCACUCAAAAGCCCGA
PK-A-7b	GGCGCAGUCCCGCAGCGCCACUCAAAAGCGGGA
HP-A	GGCGCAGUGGGCCAGCGCCACU
PK-B-1a	GGCGCGCAGUACCACAGCGCGCCACUCAAAAAAGUGGUA
PK-B-1b	GGCGCGCAGUGGUACAGCGCGCCACUCAAAAAAGUACCA
PK-B-2a	GGCGCGCAGUUCCGCAGCGCGCCACUCAAAAAAGCGGAA
PK-B-2b	GGCGCGCAGUUCGGCAGCGCGCCACUCAAAAAAGCCGAA
PK-B-3a	GGCGCGCAGUAGGUCAGCGCGCCACUCAAAAAAGACCUA
PK-B-3b	GGCGCGCAGUCCUACAGCGCGCCACUCAAAAAAGUAGGA
PK-B-4a	GGCGCGCAGUUAGGCAGCGCGCCACUCAAAAAAGCCUAA
PK-B-4b	GGCGCGCAGUAAGCCAGCGCGCCACUCAAAAAAGGCUUA
HP-B	GGCGCGCAGUGGGCCAGCGCGCCACU

a) 'PK-X-Ya' and 'PK-X-Yb' represents the pair of PKs with identical nearest neighbour base pairs at the S2 stem.

b) Red and blue colours denote the S1 and S2 stems, respectively.

Table S2. Thermodynamic parameters for S2 formation in various PK-A and PK-B analysed from UV melting at 260 nm.

RNA sequences	ΔH^o (kcal/mol) [a]	$T\Delta S^o$ (kcal/mol) [a]	ΔG^o_{37} (kcal/mol) [a]	T_m (ºC) [a]
MMTV VPK	−−31.3 ± 2.7	−30.0 ± 2.7	−1.3 ± 0.1	50.8 ± 1.5
PK-A-1a	−38.1 ± 2.4	−38.8 ± 2.5	0.7 ± 0.1	31.3 ± 0.8
PK-A-1b	−41.0 ± 5.7	−42.5 ± 5.9	1.5 ± 0.2	26.0 ± 1.5
PK-A-2a	−33.6 ± 2.5	−33.2 ± 2.9	−0.5 ± 0.4[b]	41.8 ± 3.6
PK-A-2b	−33.6 ± 2.9	−33.6 ± 2.9	0.1 ± 0.4[b]	36.4 ± 3.3
PK-A-3a	−37.3 ± 3.2	−37.7 ± 3.4	0.4 ± 0.2	33.8 ± 1.2
PK-A-3b	−31.5 ± 3.0	−31.9 ± 3.1	0.3 ± 0.1[b]	33.8 ± 0.8
PK-A-4a	−34.8 ± 2.0	−35.4 ± 2.0	0.6 ± 0.1	31.6 ± 0.8
PK-A-4b	−39.7 ± 2.2	−40.1 ± 2.3	0.4 ± 0.1	34.1 ± 0.6
PK-A-5a	−33.2 ± 2.8	−32.1 ± 2.7	−1.1 ± 0.1	48.1 ± 0.9
PK-A-5b	−32.5 ± 5.2	−31.7 ± 5.1	−0.8 ± 0.1	44.9 ± 0.7
PK-A-6a	−34.5 ± 2.2	−33.9 ± 2.3	−0.7 ± 0.3[b]	43.4 ± 2.6
PK-A-6b	−30.1 ± 5.1	−29.4 ± 5.4	−0.7 ± 0.2	44.6 ± 3.5
PK-A-7a	−25.5 ± 2.7	−24.1 ± 2.7	−1.4 ± 0.1	55.4 ± 3.5
PK-A-7b	−33.4 ± 3.8	−31.3 ± 3.7	−2.0 ± 0.1[b]	57.4 ± 1.5
PK-B-1a	−30.5 ± 9.9	−31.3 ± 10.2	0.8 ± 0.7	29.5 ± 8.0
PK-B-1b	−32.9 ± 6.1	−32.8 ± 6.2	−0.1 ± 0.1	37.7 ± 1.1
PK-B-2a	−22.4 ± 2.6	−21.8 ± 2.6	−0.5 ± 0.2[b]	44.8 ± 2.5
PK-B-2b	−20.2 ± 2.3	−19.8 ± 2.5	−0.4 ± 0.4	43.3 ± 6.4
PK-B-3a	−28.8 ± 4.2	−29.0 ± 4.4	0.2 ± 0.2	35.5 ± 1.8
PK-B-3b	−18.7 ± 1.7	−18.9 ± 1.7	0.2 ± 0.1	33.7 ± 2.2
PK-B-4a	−30.5 ± 4.7	−30.9 ± 4.8	0.4 ± 0.2	33.4 ± 2.0
PK-B-4b	−30.6 ± 2.3	−30.6 ± 2.3	0.0 ± 0.1	37.3 ± 1.3

[a] Values and errors are the average ± SD of at least five or more experiments.

[b] Averaged $\Delta G°37$ values have a slight deviation from a value calculated from $\Delta H°$ and $T\Delta S°$ values in the left columns due to the rounding off process of each replicated data.

Table S3. Predicted free energy change of RNA duplexes consisting of the same base pairs with S2 stem region

Corresponding PK	Duplex sequence	Predicted ΔG^{o}_{37} (kcal/mol) at 1000 mM NaCl	Predicted ΔG^{o}_{37} (kcal/mol) at 100 mM NaCl
PK-A-1a	5′-UACCAC-3′ 3′-AUGGUG-5′	−6.64	−5.85
PK-A-1b	5′-UGGUAC-3′ 3′-ACCAUG-5′		
PK-A-2a	5′-UUCCGC-3′ 3′-AAGGCG-5′	−7.78	−6.57
PK-A-2b	5′-UUCGGC-3′ 3′-AAGCCG-5′		
PK-A-3a	5′-UAGGUC-3′ 3′-AUCCAG-5′	−6.72	−5.90
PK-A-3b	5′-UCCUAC-3′ 3′-AGGAUG-5′		
PK-A-4a	5′-UUAGGC-3′ 3′-AAUCCG-5′	−6.48	−5.75
PK-A-4b	5′-UAAGCC-3′ 3′-AUUCGG-5′		
PK-A-5a	5′-UCACCC-3′ 3′-AGUGGG-5′	−8.68	−7.13
PK-A-5b	5′-UGGUCC-3′ 3′-ACCAGG-5′		
PK-A-6a	5′-UGCGUC-3′ 3′-ACGCAG-5′	−7.94	−6.67
PK-A-6b	5′-UGUCGC-3′ 3′-ACAGCG-5′		
PK-A-7a	5′-UCGGGC-3′ 3′-AGCCCG-5′	−10.11	−8.04
PK-A-7b	5′-UCCCGC-3′ 3′-AGGGCG-5′		

of PKs following the NN model.

References

1. L. X. Shen and J. I. Tinoco, *J. Mol. Biol.*, 1995, **247**, 963-978.
2. T. Xia, J. SantaLucia, M. E. Burkard, R. Kierzek, S. J. Schroeder, X. Jiao, C. Cox and D. H. Turner, *Biochemistry*, 1998, **37**, 14719-14735.
3. S.-i. Nakano, M. Fujimoto, H. Hara and N. Sugimoto, *Nucleic Acids Res.*, 1999, **27**, 2957-2965.