Supplementary Information for "Free energy landscape of RNA hairpins constructed via dihedral angle principal component analysis"

Laura Riccardi, Phuong H. Nguyen, and Gerhard Stock*

Institute of Physical and Theoretical Chemistry, Goethe University, Max-von-Laue-Str. 7, D-60438 Frankfurt, Germany

* New address: Institute of Physics, Albert Ludwigs University, Hermann-Herder-Strasse. 3, 79104 Freiburg, Germany. E-mail: stock@physik.uni-freiburg.de

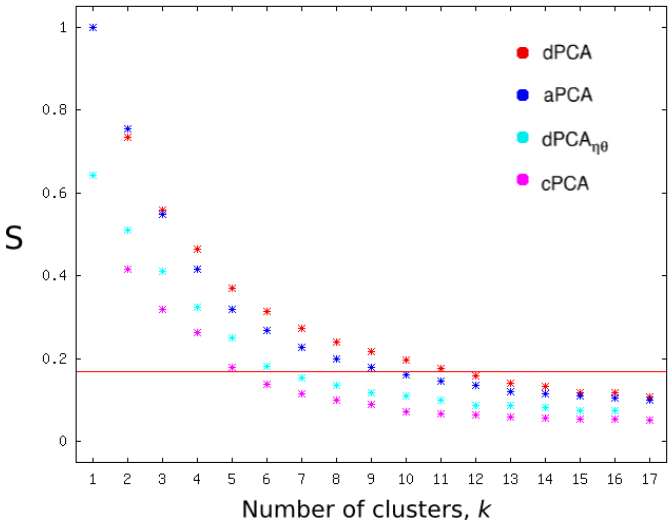

FIG. 1: Normalized sum of k-means intracluster RMSDs, S, plotted as a function of the number of clusters, k. The k-means results were obtained from the dPCA (red), aPCA (blue), dPCA$_{\eta\theta}$ (cyan) and cPCA (pink) of the UUCG hairpin, respectively. To determine an appropriate value of k, we choose k as the maximal number of clusters which satisfies $S(k) \geq S_0$, where $S_0 = 0.17$ is a preset threshold value.

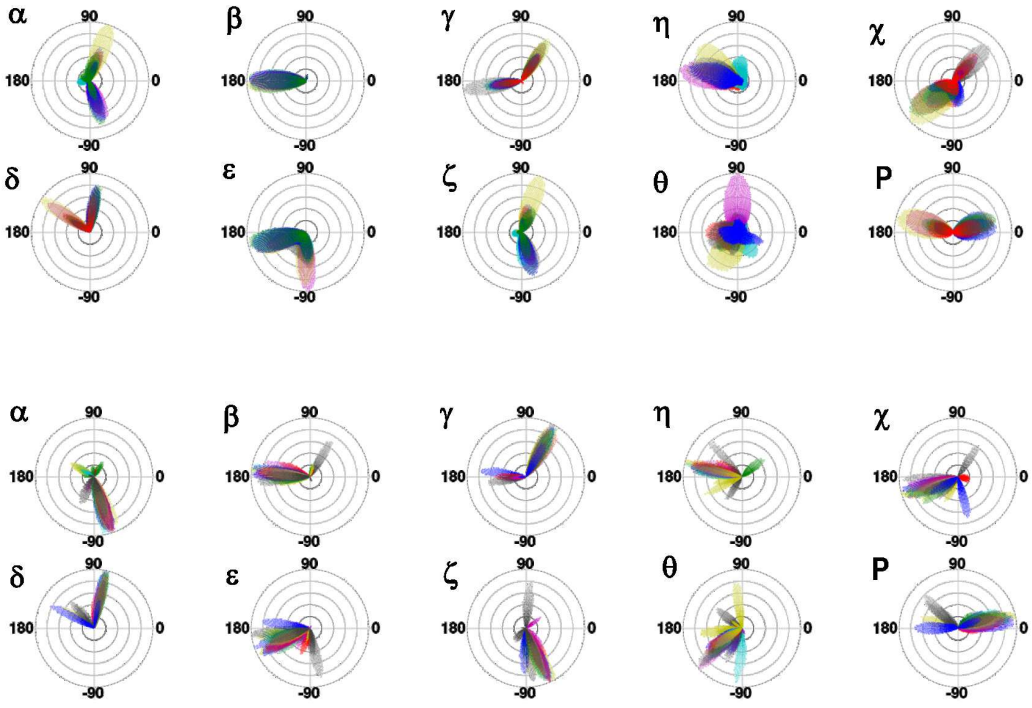

FIG. 2: Distribution of RNA dihedral angles as obtained from MD simulations of the UUCG hairpin at 460 K (top) and of the SL1m 36-mer RNA (bottom).

cluster id	population [%]	stacking interactions	total # of HB	Watson&Crick 1C-10G	2G-9C	3C-8G	Wobble 4U-7G	other hydrogen bonds between bases	other hydrogen bonds involving backbone
1	24.0	2-10 4-6 8-9	11	2/3	y	y	1/2		**O2'$_{4U}$:O6$_{7G}$, N4$_{C6}$:O2P$_{5U}$**
2	22.0	1-2 4-6	16	2/3	y	y	-	N4$_{1C}$:O6$_{2G}$,N2$_{2G}$:N3$_{9C}$ N3$_{4U}$:O6$_{7G}$,N1$_{7G}$:O4$_{4U}$	**N4$_{6C}$:O2P$_{5U}$, O2'$_{1C}$:O4'$_{2G}$,** O2'$_{2G}$:O4'$_{3C}$, O2'$_{9C}$:O4'$_{10G}$
3	12.0	4-6 8-9	9	-	2/3	y	-	**N2$_{7G}$:O2$_{4U}$**	**N4$_{6C}$:O2P$_{5U}$, O2'$_{1C}$:O5'$_{2G}$,** O2'$_{9C}$:O5'$_{10G}$
4	11.8	1-2 4-6 8-9	14	y	y	y	1/2	N2$_{7G}$:O2$_{6C}$	**O2'$_{4U}$:O6$_{7G}$, N4$_{6C}$:O2P$_{5U}$** O2'$_{7G}$:O1P$_{8G}$
5	6.5	1-2 4-8 9-10	15	y	y	y	1/2	**N2$_{7G}$:O2$_{4U}$**	O2'$_{1C}$:O4'$_{2G}$, O2'$_{2G}$:O5'$_{3C}$ O2'$_{9C}$:O4'$_{10G}$, O2'$_{5U}$:6C$_{O1P}$
6	5.6	8-9 9-10	13	y	y	2/3	-	**N2$_{7G}$:O2$_{4U}$**	O2'$_{6C}$:O1P$_{7G}$, O2'$_{4U}$:O5'$_{5U}$ O2'$_{4U}$:O2$_{5U}$, N2$_{7G}$:O3'$_{6C}$
7	5.3	3-7	12	-	y	y	-	N2$_{7G}$:O2$_{5U}$	O2'$_{2G}$:O5'$_{3C}$, O2'$_{4U}$:O5'$_{5U}$ O2'$_{8G}$:O5'$_{9C}$
8	4.6	4-6	16	y	y	y	-	N3$_{4U}$:N7$_{7G}$	O5'$_{1C}$:O2$_{1C}$, O2'$_{2G}$:O5'$_{3C}$ O2'$_{2G}$:O5'$_{10G}$, O2'$_{3C}$:O5'$_{4U}$ O2'$_{8G}$:O4'$_{9C}$, O2'$_{5U}$:O2$_{4U}$
9	4.0	8-9 9-10	13	-	y	y	1/2		**O2'$_{4U}$:O6$_{7G}$, N4$_{6C}$:O2P$_{5U}$** O2'$_{9C}$:O5'$_{10G}$, N2$_{10G}$:O2P$_{10G}$ O2'$_{1C}$:O4'$_{2G}$, O2'$_{7G}$:O1P$_{8G}$
10	2.3	1-10 2-4 4-6 7-9	13	-	y	-	y		**N4$_{6C}$:O2P$_{5U}$, N2$_{10G}$:O2P$_{10G}$** O2'$_{3C}$:O2P$_{4U}$, N4$_{6C}$:O2P$_{4U}$ O5'$_{1C}$:O2$_{1C}$, O2'$_{6C}$:O1P$_{7G}$ N1$_{7G}$:O1P$_{6C}$, O2'$_{7G}$:O1P$_{8G}$
11	1.7	2-3 4-6 7-9 9-10	13	-	y	y	1/2	**N2$_{7G}$:O2$_{4U}$**	**O2'$_{4U}$:O6$_{7G}$, O2'$_{3C}$:O5'$_{4U}$** N2$_{10G}$:O2P$_{10G}$, O2'$_{5U}$:O1P$_{6C}$ O2'$_{7G}$:O1P$_{8G}$

TABLE I: Hydrogen bonding pattern of the dPCA conformational states of the UUCG hairpin. "y" indicates a stable base-pair, fractions indicate the number of stable hydrogen bonds (i.e., 2/3 means two out of three). "Native" hydrogen bonds are shown in bold face. Stacking interactions are shown indicating the number of the involved nucleotides.

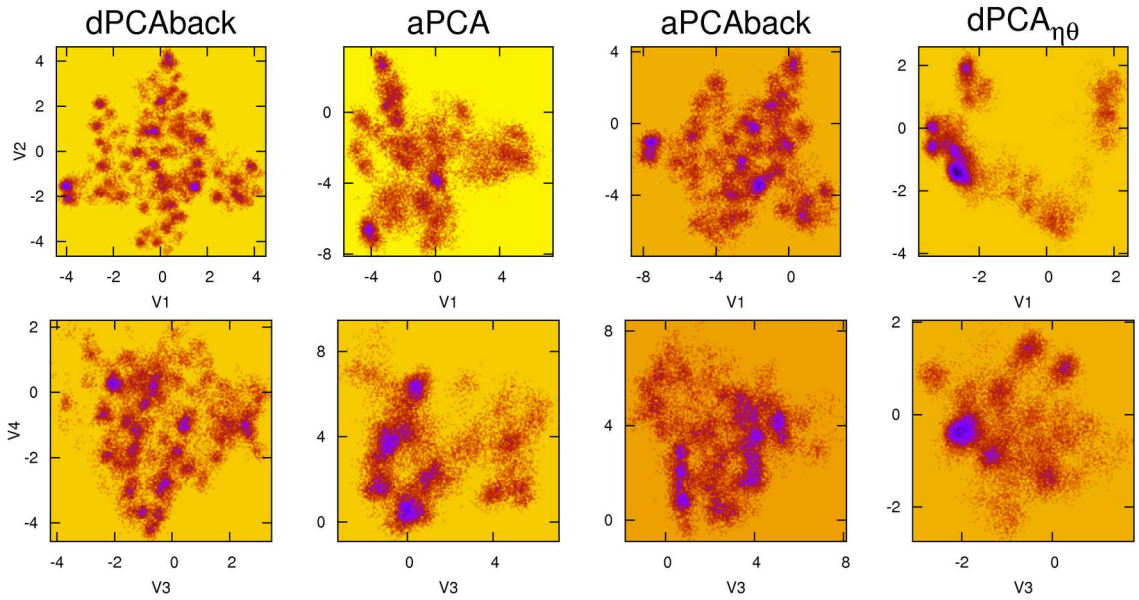

FIG. 3: Two-dimensional representations of the free energy landscape ΔG of SL1m as obtained by dPCA (considering only backbone dihedral angles), aPCA (backbone dihedral angles only), and dPCA$_{\eta\theta}$.

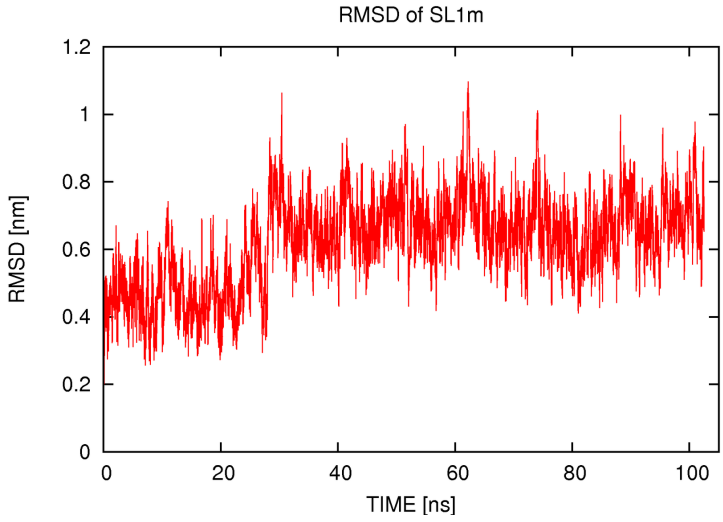

FIG. 4: Time evolution of the all-atom RMSD of SL1m with respect to its initial structure.

dPCA	angle
1+9(45%)	1
2	3+4
3	5
4+9(37%)	2
5	6
6	7
7	9
8	8
10	11
11	10

TABLE II: Comparison of conformational clusters of the UUCG hairpin as obtained from the dPCA and from a clustering on the original angle data (i.e., without PCA but re-centering all angle values around their angular mean, see Methods).