Supporting Information

Modeling noncanonical RNA base pairs by coarse-grained IsRNA2 model

Dong Zhang[a], Shi-Jie Chen[b*], and Ruhong Zhou[a,*]

[a] College of Life Sciences and Institute of Quantitative Biology, Zhejiang University, Hangzhou 310058, China

[b] Department of Physics, Department of Biochemistry, and Institute of Data Science and Informatics, University of Missouri, Columbia, Missouri, 65211, USA

Formulas of energy functions in IsRNA2 model

The energy functions for three local covalent energy terms in the coarse-grained (CG) force field (Eq. 1 in the main text) have the following functional forms:

$$E_{bond}(b) = K_b(b - b_0)^2 + \frac{H_b}{\sqrt{2\pi}\sigma_b} e^{-0.5(b-b_1)^2/\sigma_b^2} \quad \text{(S1)}$$

$$E_{angle}(\theta) = K_a(\theta - \theta_0)^2 + \frac{H_a}{\sqrt{2\pi}\sigma_a} e^{-0.5(\theta-\theta_1)^2/\sigma_a^2} \quad \text{(S2)}$$

$$E_{torsion}(\phi) = \sum_{n=1}^{4} K_n[1 + \cos(n\phi - \phi_n)] \quad \text{(S3)}$$

where b, θ and ϕ denote the bond stretching length between two connected beads, the bond angle between two adjacent bonds, and the torsion angle between three successive bonds, respectively. Parameters K_b, b_0, H_b, σ_b, and b_1 in Eq. S1 describe the strength of bond stretching and the equilibrium bond length. Parameters K_a, θ_0, H_a, σ_a, and θ_1 in Eq. S2 characterize the bending energy of the bond angle and the equilibrium bond angle. K_n and ϕ_n (n = 1, 2, 3, 4) in Eq. S3 are related to the local minima on the profile of the torsional energy. Those energy parameters were determined using the iterative simulated reference state approach, and their detailed values are given in Table S1~S3, respectively. Compared to the widely used harmonic potentials for bond stretching and bond angle terms in previous CG models, extra Gaussian terms are introduced in Eqs. S1 and S2 in order to capture more precise details of the statistical potential and to enable broader sampling of the backbone conformations in simulations.

Table S1 Energy parameters for 21 types of bond stretching energies $E_{bond}(b)$ in the IsRNA2 force field, in which the last 6 types are for the canonical base pairing (GC, AU, and GU) interactions. The unit for potential energy in the IsRNA2 model is e.u.

Type	K_b (e.u./Å^2)	b_0 (Å)	H_b (e.u.*Å)	b_1 (Å)	σ_b (Å)
P-S	6.508	3.460	-0.978	3.899	0.114
S-R1	22.753	4.207	5.443	4.040	0.343
S-Y1	3.954	4.047	-0.723	3.821	0.169
R1-A1	173.453	2.964	-0.172	2.983	0.031
R1-A2	120.572	2.864	-0.227	2.878	0.029
R1-G1	140.671	2.883	-0.211	2.885	0.032
R1-G2	95.093	3.560	-0.250	3.569	0.035
A1-A2	216.486	2.305	-0.114	2.310	0.021
G1-G2	106.309	3.002	-0.217	3.021	0.035
Y1-Y2	245.302	2.632	-0.164	2.638	0.030
Y1-C1	420.063	2.410	-0.118	2.417	0.021
Y1-U1	632.051	2.437	-0.061	2.440	0.018
Y2-C1	252.14	2.941	-0.090	2.948	0.021
Y2-U1	175.06	3.029	-0.146	3.035	0.027
S-P	5.924	3.391	-1.028	3.888	0.132
G1-C1	2.025	3.486	-0.762	3.272	0.203
G2-Y2	2.329	3.599	-0.981	3.361	0.229
A1-U1	2.054	3.72	-0.653	3.437	0.197
A2-Y2	3.369	3.678	-0.283	3.532	0.201
G1-Y2	1.819	3.815	-3.058	3.205	0.441
G1-U1	0.814	3.817	-1.201	3.875	0.36

Table S2 Energy parameters for 16 types of bond angle potentials $E_{angle}(\theta)$ in the IsRNA2 force field, in which the last 3 types are for canonical base pairing (GC, AU, and GU) interactions. The unit for potential energy in the IsRNA2 model is e.u.

Type	K_a (e.u./radian^2)	θ_0 (degree)	H_a (e.u.*radian)	θ_1 (degree)	σ_a (radian)
P-S-P	1.958	120.02	-0.504	89.94	0.173
P-S-R1	2.060	115.94	-0.634	85.04	0.152
P-S-Y1	2.068	115.28	-0.881	73.57	0.171
S-P-S	1.723	119.22	-0.227	103.95	0.079
S-R1-A1	17.825	163.44	0.277	153.21	0.092
S-R1-A2	10.579	115.39	3.012	119.92	0.248
S-R1-G1	9.936	164.61	0.973	149.77	0.163
S-R1-G2	2.652	59.30	-0.344	127.25	0.093
S-Y1-Y2	3.877	34.92	-1.127	96.66	0.160
S-Y1-C1	9.927	153.72	1.695	141.06	0.213
S-Y1-U1	9.983	155.33	1.397	142.38	0.205
R1-S-P	8.012	111.88	-0.319	91.53	0.088
Y1-S-P	5.219	117.43	-0.310	94.99	0.093
G2-Y2-C1	6.228	91.03	-0.183	89.26	0.100
A2-Y2-U1	1.849	70.32	-0.434	68.92	0.135
R1-G1-U1	2.891	168.28	-0.117	159.75	0.104

Table S3 Energy parameters for 38 types of torsional angle energies $E_{torsion}(\phi)$ in the IsRNA2 force field, in which the last 14 types are for canonical base pairing (GC, AU, and GU) interactions. The unit for potential energy in the IsRNA2 model is e.u.

Type	K_1 (e.u.)	ϕ_1 (degree)	K_2 (e.u.)	ϕ_2 (degree)	K_3 (e.u.)	ϕ_3 (degree)	K_4 (e.u.)	ϕ_4 (degree)
P-S-P-S	0.370	45.40	0.054	10.38	0.187	92.90	0.069	30.90
P-S-R1-A1	-2.166	-1.89	0.876	-2.38	-0.108	36.12	-0.220	23.19
P-S-R1-A2	0.721	11.31	-0.349	77.79	-0.131	75.39	-0.135	-53.66
P-S-R1-G1	-0.703	32.97	0.537	12.95	-0.144	-6.52	0.004	-18.97
P-S-R1-G2	0.594	-3.96	0.590	-89.18	0.290	-104.95	-0.289	-15.71
P-S-Y1-Y2	0.523	-24.58	-0.573	49.92	0.091	-98.30	0.173	108.37
P-S-Y1-C1	0.210	-57.55	0.477	12.11	-0.147	57.21	-0.075	1.91
P-S-Y1-U1	0.094	-27.73	0.347	8.28	-0.097	17.29	-0.076	22.17
S-P-S-P	0.383	20.78	-0.367	17.06	0.139	-39.28	-0.072	-20.61
S-P-S-R1	0.288	0.636	0.173	-63.02	0.104	42.10	-0.051	15.09
S-P-S-Y1	-0.109	-57.63	0.188	-102.03	0.146	49.74	0.137	92.82
S-R1-A1-A2	-0.418	49.58	0.176	31.67	0.337	-57.75	-0.193	95.95
S-R1-G1-G2	0.436	-25.91	0.145	52.75	0.170	39.80	0.140	-63.86
S-Y1-C1-Y2	-1.458	15.27	-0.657	38.38	0.187	-47.90	0.260	17.53
S-Y1-U1-Y2	-2.135	11.53	0.751	-152.89	0.309	11.87	-0.465	-149.58
R1-S-P-S	0.186	71.39	0.080	-45.30	0.156	74.49	0.072	62.60
A1-R1-S-P	1.613	41.44	0.907	76.91	0.191	-1.14	0.298	-5.14
A2-R1-S-P	0.650	-88.30	-0.292	95.94	0.089	96.15	0.101	28.55
G1-R1-S-P	0.093	-56.96	0.136	-29.09	0.371	147.25	-0.176	25.71
G2-R1-S-P	-0.623	81.12	-0.550	134.66	-0.053	21.21	-0.333	-135.70
Y1-S-P-S	0.142	-36.15	0.074	-20.04	0.174	37.52	0.021	-3.53
Y2-Y1-S-P	0.892	-112.77	-0.361	98.75	0.209	108.25	-0.209	-25.16
C1-Y1-S-P	0.440	-12.54	0.049	71.60	0.211	-60.62	-0.158	13.35

U1-Y1-S-P	0.257	-0.81	-0.071	-6.14	0.204	-28.86	-0.141	-43.15
R1-G1-G2-Y2	1.605	6.61	1.376	-164.81	0.761	13.55	-0.587	-4.77
G1-G2-Y2-C1	-1.491	-7.11	1.047	165.41	0.535	156.73	-0.149	-38.33
G2-Y2-C1-Y1	2.180	6.10	1.238	-162.66	0.600	22.93	-0.095	22.42
R1-G1-C1-Y1	-0.811	-11.62	-0.055	-93.01	-0.077	-59.59	0.024	-75.17
R1-G2-Y2-Y1	0.760	-35.36	-0.837	155.48	0.056	-46.76	0.275	121.88
R1-A1-A2-Y2	1.654	7.85	-1.204	15.15	0.637	14.04	-0.373	-3.72
A1-A2-Y2-U1	-1.088	-3.56	-0.652	-7.54	-0.215	-14.22	0.029	29.94
A2-Y2-U1-Y1	1.743	2.05	-1.214	15.32	0.650	17.65	-0.486	-7.23
R1-A1-U1-Y1	-0.631	2.70	0.182	6.66	-0.297	-23.95	0.158	-15.20
R1-A2-Y2-Y1	0.386	-9.86	0.233	-36.94	-0.088	-2.17	0.016	59.39
R1-G2-G1-Y2	-1.060	-11.41	0.310	-55.89	0.552	-5.21	0.575	150.99
G2-G1-Y2-U1	1.155	7.14	0.237	-119.78	-0.876	155.18	0.164	75.47
G1-Y2-U1-Y1	1.617	-5.28	1.084	-170.28	0.607	15.23	-0.511	4.01
R1-G1-U1-Y1	0.941	-13.70	0.212	-137.28	0.163	97.70	-0.181	3.23

Table S4 Energy parameters for 66 types of pairwise interactions $E_{pair}(r)$ in the IsRNA2 force field. The unit for potential energy in the IsRNA2 model is e.u.

Type	σ (Å)	D_0 (e.u.)	α (Å^{-1})	r_0 (Å)	H_1 (e.u.*Å)	r_1 (Å)	σ_1 (Å)	H_2 (e.u.*Å)	r_2 (Å)	σ_2 (Å)	r_{cut} (Å)
P--P	3.7	0.061	0.177	13.46	-0.674	5.89	0.573	-0.632	9.36	2.007	13.5
P--S	2.7	0.317	0.483	4.41	0.342	5.53	0.407	0.485	7.98	1.013	10.0
P--R1	3.5	0.001	0.314	15.54	-0.451	4.44	0.237	-0.090	9.17	0.204	10.1
P--A1	3.4	0.261	0.573	5.21	0.198	5.27	0.201	0.476	6.70	1.034	12.1
P--A2	3.6	0.001	1.174	6.54	-0.180	6.20	0.372	-0.223	9.04	0.351	10.2
P--G1	3.3	0.697	0.679	4.50	0.785	5.48	0.530	1.045	6.49	1.075	9.2
P--G2	3.1	0.829	0.861	4.21	0.819	6.23	0.887	-0.341	8.67	0.314	9.8
P--Y1	3.4	0.001	1.501	6.05	0.418	6.46	0.437	0.082	8.71	0.260	9.4
P--Y2	3.7	0.301	1.084	4.46	0.569	5.65	0.729	-0.144	8.76	0.168	9.4
P--C1	3.6	0.650	0.676	4.76	0.627	6.12	0.526	0.056	9.03	0.233	11.5
P--U1	3.3	0.323	1.021	4.64	-0.777	3.99	0.363	-0.174	7.71	0.445	9.0
S--S	3.1	0.282	0.972	4.46	0.643	5.87	0.879	-0.263	6.30	0.316	8.0
S--R1	3.3	0.023	0.902	5.72	-1.594	3.83	0.536	-0.125	7.09	0.308	8.0
S--A1	3.2	0.001	0.456	9.37	0.077	5.44	0.234	-0.092	6.18	0.328	7.1
S--A2	3.1	0.564	0.828	4.29	0.624	4.79	0.605	0.194	7.60	0.446	8.7
S--G1	3.1	0.066	0.466	6.49	0.554	5.68	0.755	-0.098	7.34	0.284	10.1
S--G2	2.9	0.263	0.614	4.88	0.291	7.20	0.419	-0.116	8.33	0.340	9.5
S--Y1	3.0	0.961	0.999	3.84	-0.467	3.88	0.330	0.048	7.31	0.237	8.9
S--Y2	3.1	0.071	3.699	3.54	-0.151	5.57	0.225	-0.117	8.31	0.298	9.1
S--C1	3.3	0.011	0.293	11.85	0.118	5.47	0.206	-0.382	7.45	0.383	10.0
S--U1	3.1	0.001	0.672	8.66	0.669	5.09	1.066	-0.183	7.53	0.331	8.6
R1--R1	3.2	0.618	0.792	4.38	0.882	5.17	0.578	0.295	6.35	0.269	9.3
R1--A1	2.9	0.532	1.246	3.63	0.527	5.94	0.472	0.302	8.84	0.588	10.7
R1--A2	2.9	0.514	1.265	3.72	0.538	5.52	0.840	0.321	9.34	0.690	11.2
R1--G1	2.9	0.672	1.301	3.65	0.698	4.95	1.128	-0.132	6.95	0.234	8.0
R1--G2	2.9	0.541	1.230	3.68	0.391	6.09	0.460	-0.355	7.37	0.443	8.6
R1--Y1	3.1	0.225	0.854	4.79	-0.542	3.73	0.333	0.587	5.87	0.417	8.7
R1--Y2	2.9	0.105	0.626	5.25	-0.667	4.04	0.398	-0.105	6.76	0.177	7.6
R1--C1	2.9	0.494	1.771	3.52	0.290	6.18	0.454	-0.145	6.90	0.359	8.0
R1--U1	3.0	0.569	0.554	4.45	1.240	5.73	0.761	0.283	9.14	0.507	10.4
A1--A1	2.5	0.846	0.467	4.52	2.237	5.37	0.762	0.754	8.59	0.982	10.7
A1--A2	2.7	1.531	0.526	4.16	3.128	5.13	1.173	-0.320	6.73	0.585	11.6
A1--G1	2.7	0.971	1.463	3.40	0.999	5.45	1.088	-0.831	6.69	0.557	8.3
A1--G2	2.7	0.484	0.699	4.32	0.844	7.04	1.664	-0.441	7.62	0.725	10.6
A1--Y1	2.9	0.719	1.611	3.51	0.340	4.98	0.605	0.103	6.48	0.208	7.1
A1--Y2	2.9	0.955	1.360	3.52	-0.091	5.19	0.261	-0.422	7.47	0.535	9.2
A1--C1	2.9	0.639	1.887	3.40	0.824	5.88	0.734	-0.517	6.73	0.355	7.8
A1--U1	2.6	0.980	1.198	3.44	-0.672	6.90	0.535	-0.500	9.69	1.103	12.1

A2--A2	2.7	0.570	1.674	3.56	0.679	5.90	0.622	-0.314	7.60	0.706	9.2
A2--G1	2.9	0.284	1.873	3.53	0.250	5.76	0.679	-0.268	6.82	0.414	8.0
A2--G2	2.9	0.307	1.287	3.77	0.256	5.45	0.486	0.197	9.36	0.735	11.0
A2--Y1	2.9	0.414	1.318	3.80	0.779	5.73	0.547	0.553	9.25	0.642	11.1
A2--Y2	2.5	0.616	0.558	4.42	0.775	5.04	0.517	0.335	6.17	0.362	10.7
A2--C1	2.7	0.535	0.785	4.06	0.964	5.48	0.671	-0.186	6.86	0.321	8.6
A2--U1	2.7	0.318	0.516	4.73	1.367	5.57	0.738	0.743	8.55	0.766	10.4
G1--G1	2.7	0.616	1.183	3.62	0.454	5.27	0.610	-0.434	7.01	0.402	8.3
G1--G2	2.9	0.296	1.404	3.59	-0.215	7.43	0.590	0.077	9.28	0.371	10.0
G1--Y1	2.9	0.867	1.262	3.58	0.468	4.80	0.414	-0.039	7.15	0.160	7.9
G1--Y2	2.9	0.437	1.165	3.40	0.247	4.63	0.671	-0.046	5.81	0.256	6.9
G1--C1	2.9	0.639	2.164	3.34	0.449	5.03	0.407	-0.326	6.74	0.268	7.8
G1--U1	2.7	0.516	1.251	3.62	0.683	5.03	0.614	-0.382	6.83	0.482	8.4
G2--G2	2.7	0.220	1.011	4.11	0.191	6.56	0.280	0.153	7.93	0.231	8.8
G2--Y1	3.0	0.544	1.051	3.94	1.196	5.44	0.995	-0.322	7.02	0.292	8.1
G2--Y2	2.5	0.760	0.674	3.97	0.979	5.23	0.949	0.177	9.44	0.404	10.6
G2--C1	2.7	0.446	0.690	4.25	0.684	5.34	0.558	0.265	8.62	0.475	10.0
G2--U1	2.7	0.648	0.524	4.49	1.324	5.31	0.815	0.714	7.45	1.552	10.2
Y1--Y1	3.1	0.394	1.159	4.18	0.360	6.19	0.414	-0.199	8.28	0.521	9.7
Y1--Y2	2.7	0.012	0.286	11.02	-2.363	4.17	0.799	0.268	9.48	0.519	10.8
Y1--C1	2.8	0.315	1.131	3.58	0.138	3.89	0.264	-0.098	7.06	0.191	8.0
Y1--U1	3.1	0.335	1.646	3.63	0.020	6.67	0.124	0.121	9.36	0.362	10.4
Y2--Y2	2.9	0.134	1.410	3.85	0.253	6.08	0.284	0.160	7.78	0.393	8.7
Y2--C1	2.7	0.166	0.422	5.30	0.332	6.31	0.382	0.178	9.49	0.324	10.4
Y2--U1	2.7	0.190	0.810	4.58	-0.406	3.10	0.260	0.428	5.78	0.733	7.6
C1--C1	2.5	0.422	0.751	4.16	0.988	5.23	0.738	-0.317	6.82	0.401	8.6
C1--U1	2.9	0.121	0.531	6.63	-5.777	3.03	0.614	0.326	5.79	0.288	8.6
U1--U1	2.7	0.805	1.376	3.50	0.519	4.85	0.444	-0.288	6.95	0.383	8.1