Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.
This journal is © the Owner Societies 2017

Electronic Supplementary Information for "Structural and Dynamic Properties of the C-terminal Region of the *Escherichia Coli* RNA Chaperone Hfq: Integrative Experimental and Computational Studies"

Bin Wen, Weiwei Wang[$], Jiahai Zhang, Qingguo Gong, Yunyu Shi, Jihui Wu* and Zhiyong Zhang[*]

Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, P. R. China

*To whom correspondence should be addressed. Tel: +86 551 63600854; Fax: +86 551 63600374; Email: zzyzhang@ustc.edu.cn

Correspondence may also be addressed to Jihui Wu. Tel: +86 551 63600394; Email: wujihui@ustc.edu.cn

[$]Present Address: Laboratory of Molecular Neurobiology and Biophysics, Rockefeller University, Howard Hughes Medical Institute, New York, NY 10021, USA

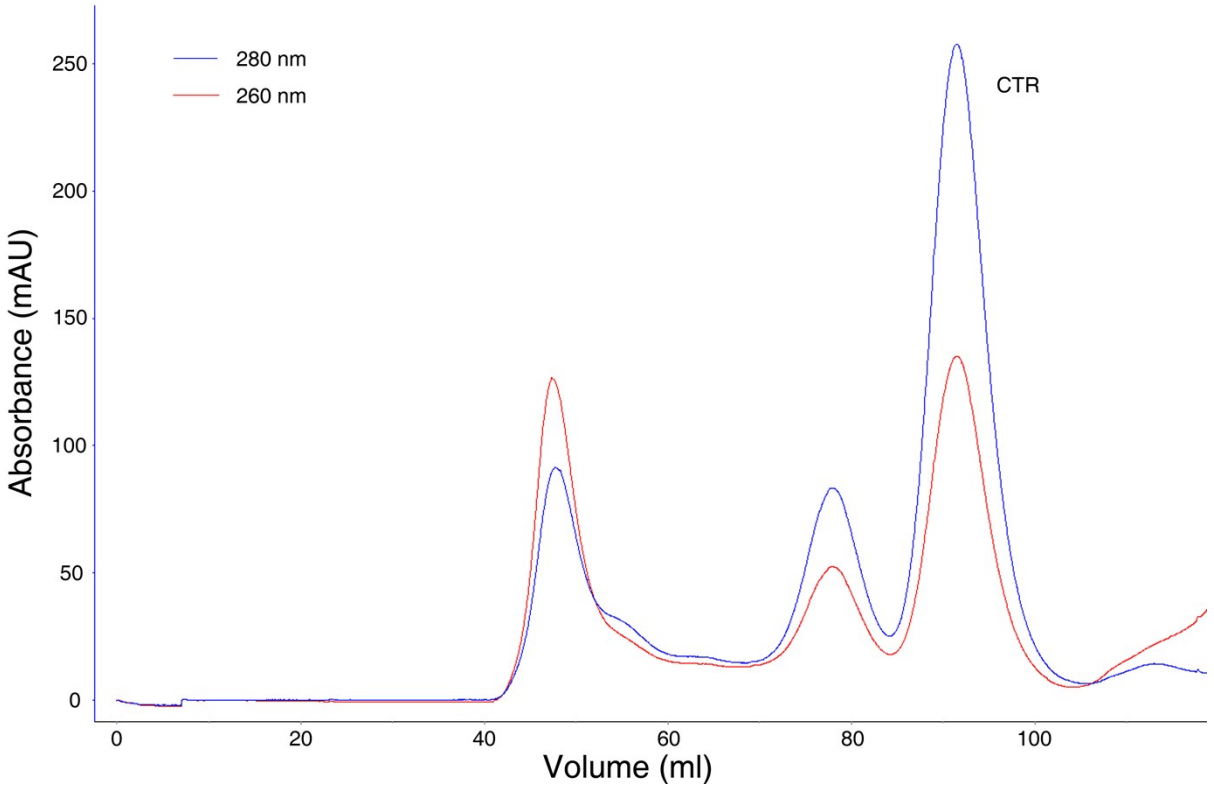

Figure S1. Purification of the CTR on a Hiload 16/60 Superdex 75 column. The elution column of the CTR is about 93 ml, which is slightly smaller than a standard globular protein with a molecular weight of 4 kDa. This may be because the CTR is an intrinsically disordered protein, resulting in a higher apparent molecular weight upon size exclusion chromatography. Nevertheless, the gel filtration chromatogram can confirm that the purified protein is not degraded.

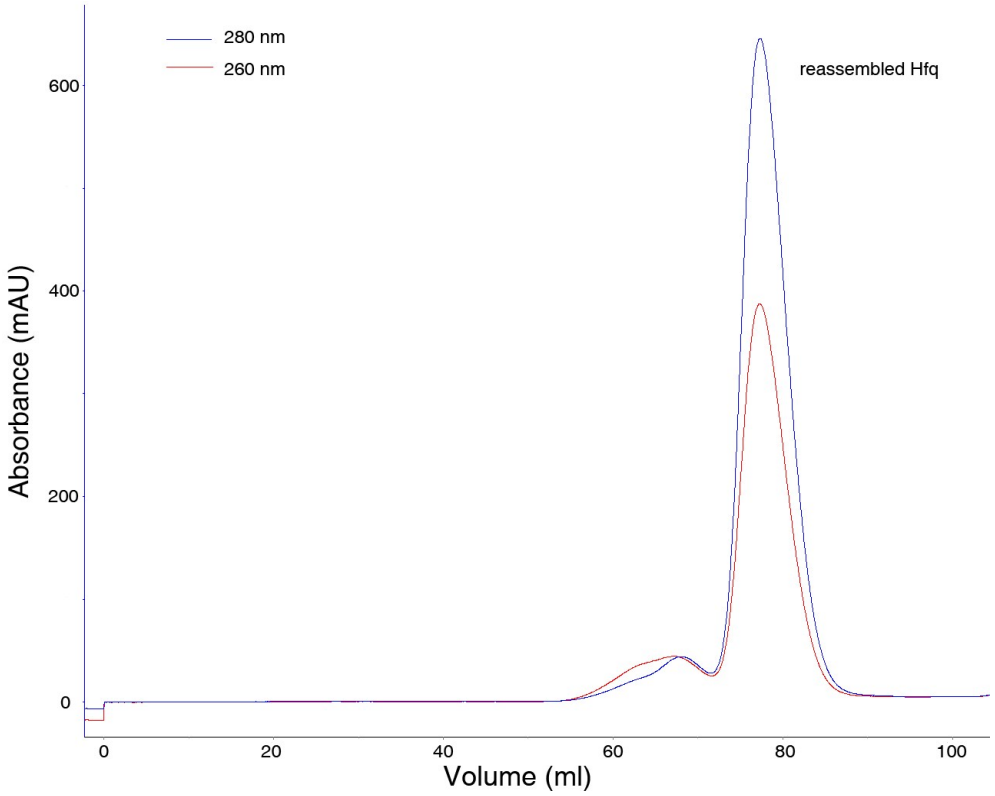

Figure S2. Gel filtration chromatogram of the reassembled Hfq. The state of the refolded protein was detected on a Superdex 200 column. The single peak indicates that the reassembled Hfq is a stable homohexamer.

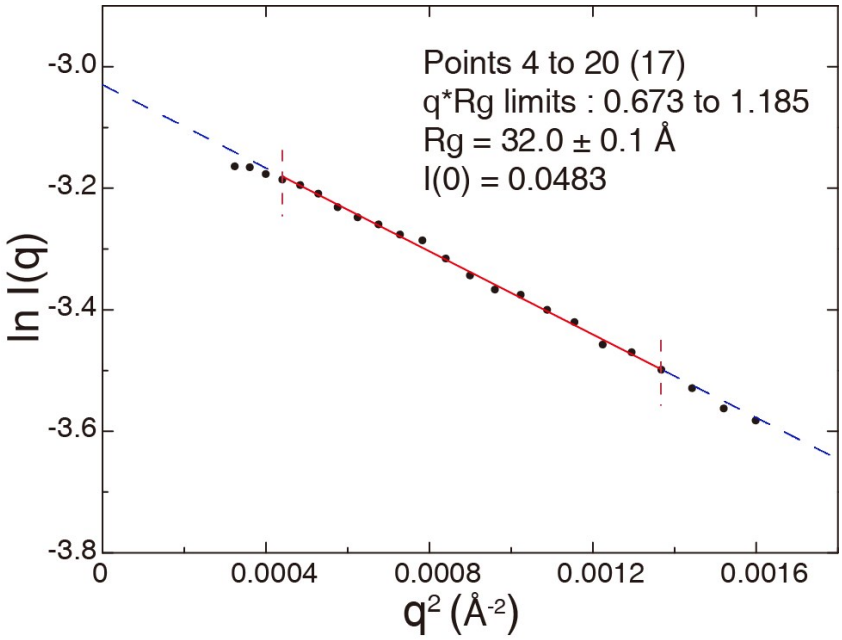

Figure S3. Guinier plot of the SAXS curve for Hfq. Rg was estimated to be 32.0 ± 0.1 Å by Guinier analysis using autoRg.

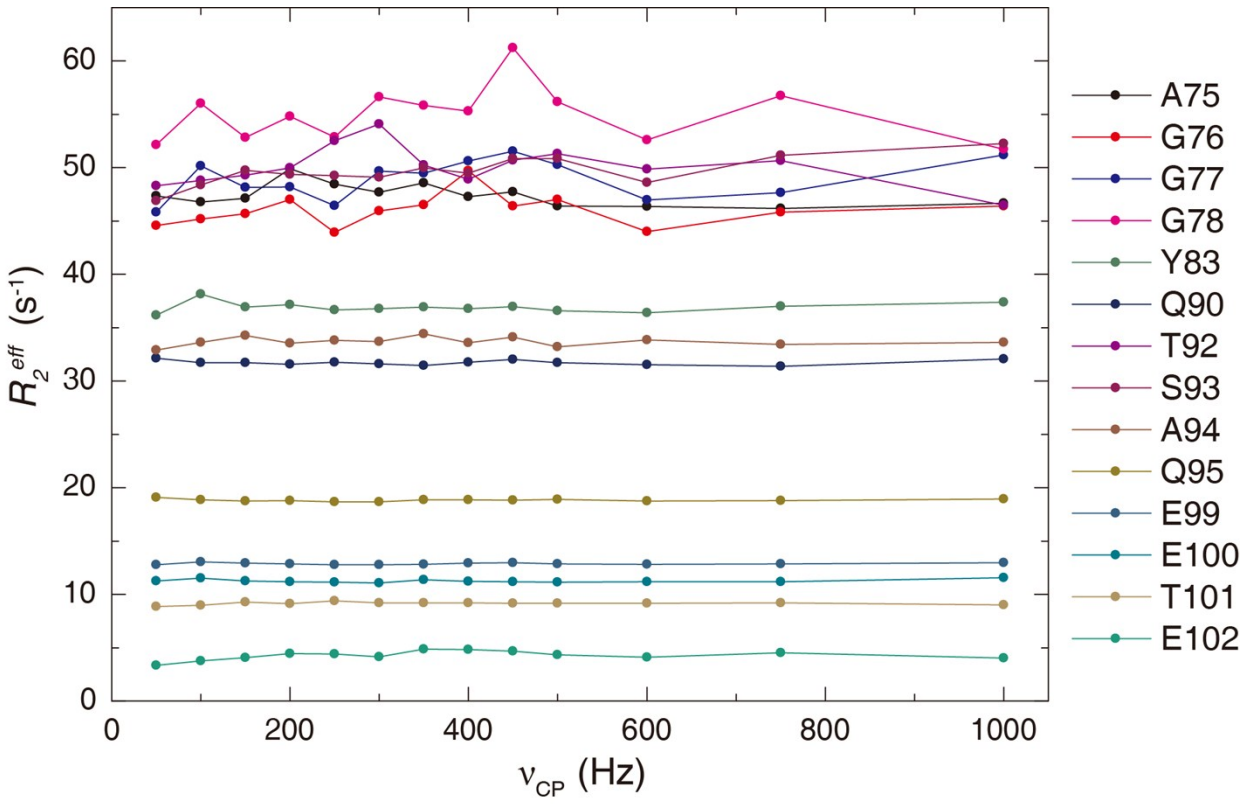

Figure S4. Relaxation dispersion data of HfqCTR recorded at 298 K and 800 MHz. Effective transverse ^{15}N relaxation rates, R_2^{eff}, are plotted against the CPMG field strength, ν_{CP}, of residues in HfqCTR.

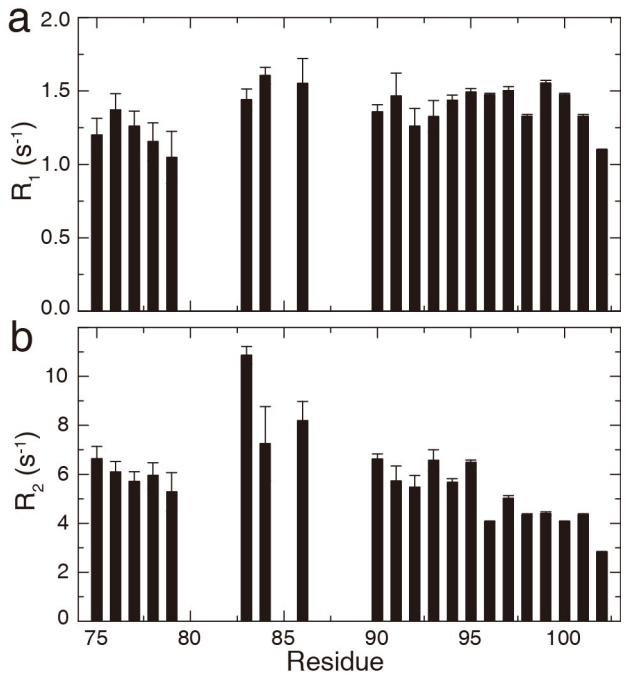

Figure S5. (a) Longitudinal and (b) transverse ^{15}N relaxation rates of HfqCTR.

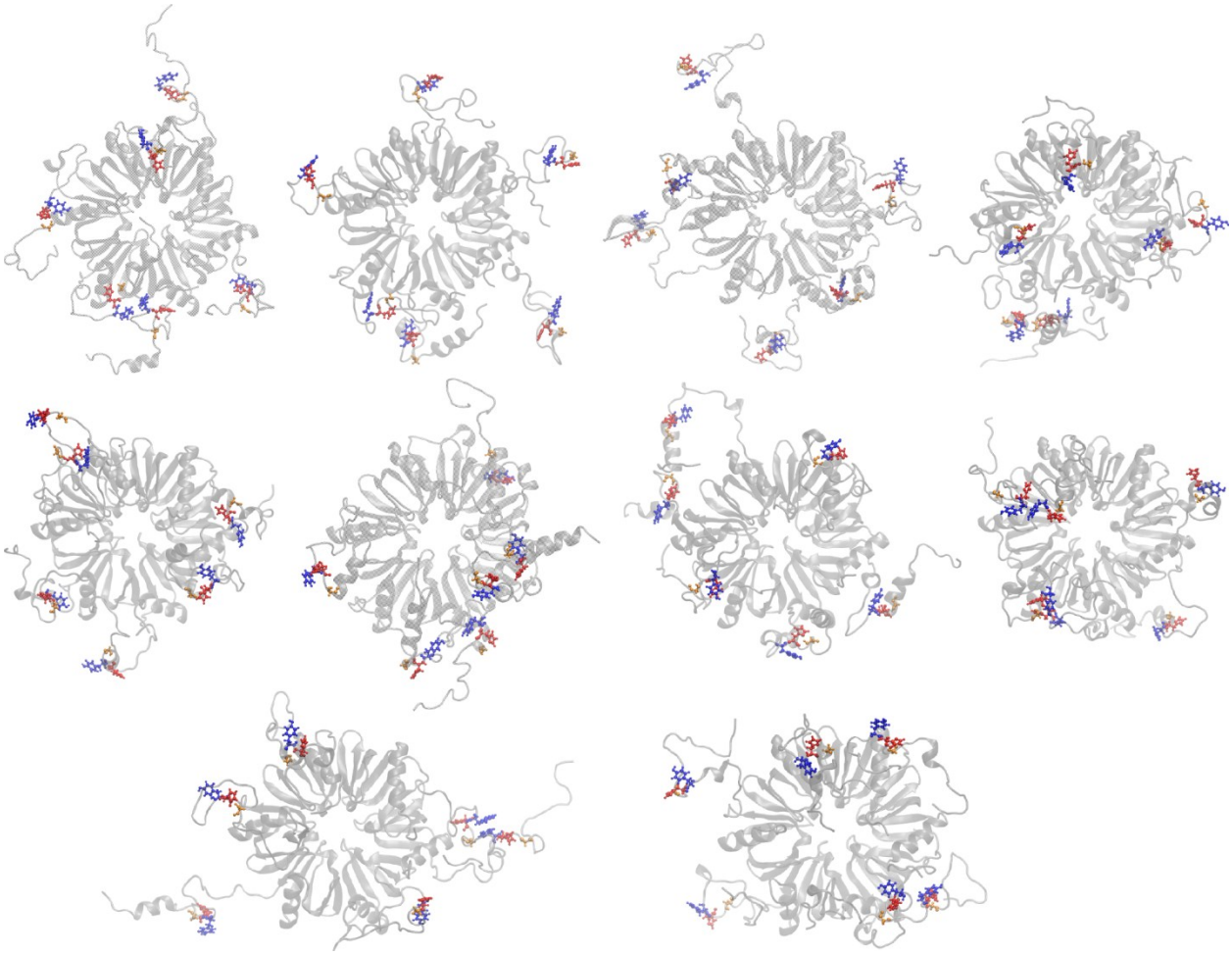

Figure S6. Ten conformations of Hfq (the same as those in Figure 4a), representing the final snapshots of the ten independent MD simulations, are displayed individually to show the local structures of several residues in HfqCTR. Y83 is indicated in blue, H84 is indicated in red, and G86 is indicated in orange.

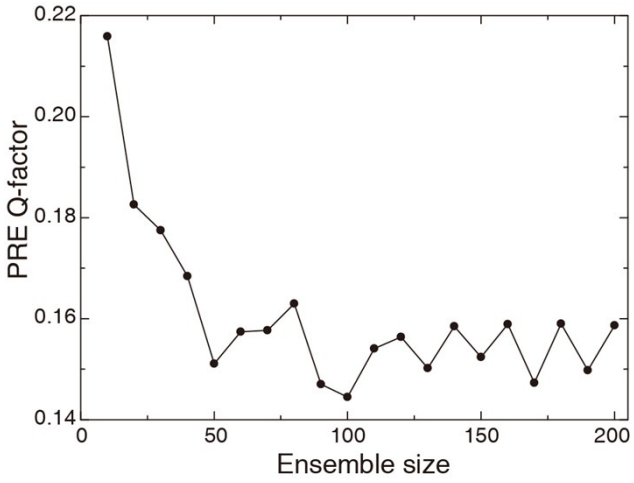

Figure S7. PRE Q-factor as a function of the ensemble size.

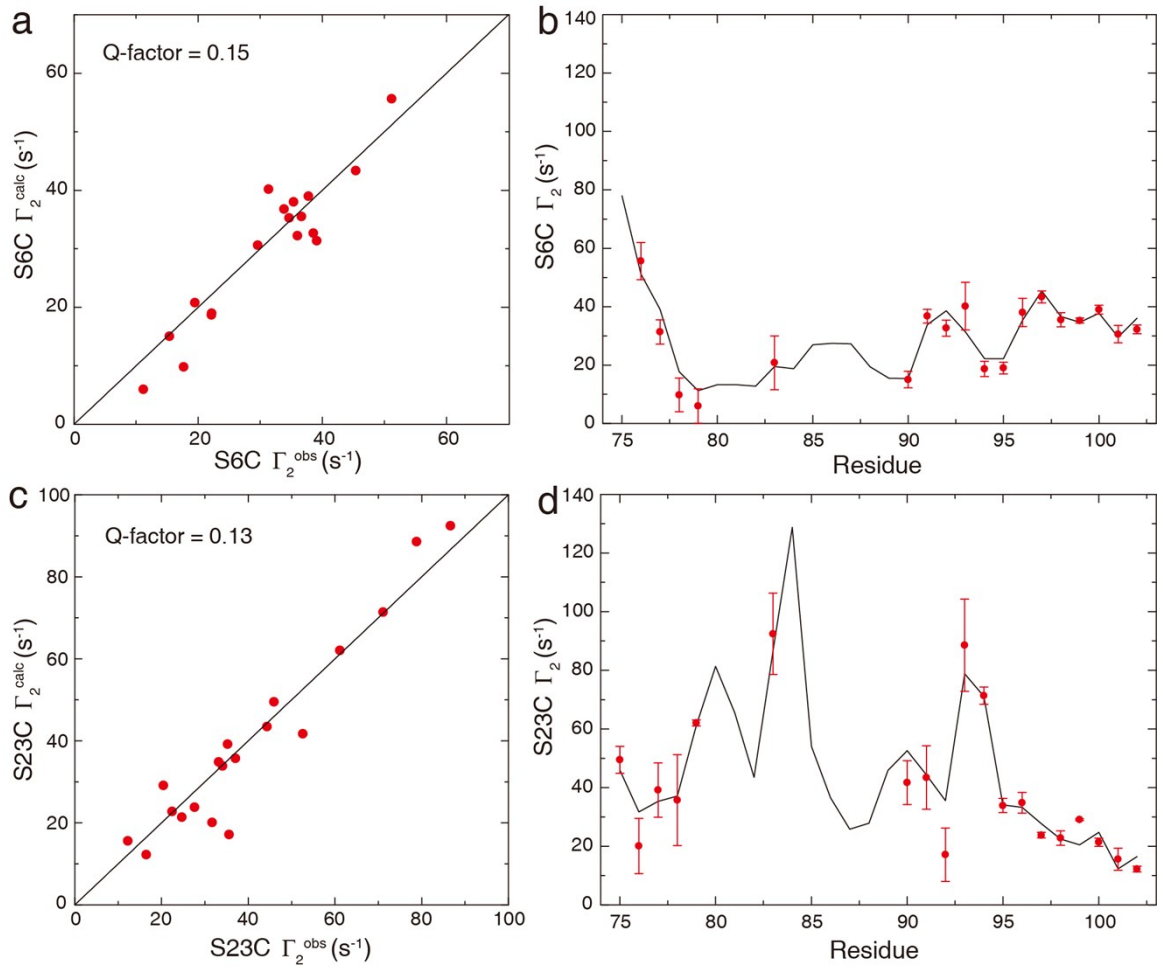

Figure S8. PRE measurements for the reassembled Hfq with a spin label conjugated to S6C and S23C. Correlations between the observed and calculated PREs are shown in (a) S6C and (c) S23C. Comparison of the observed PRE profiles (red circles) with the back-calculated values (averaged from the structure ensemble) are shown in (b) S6C and (d) S23C.

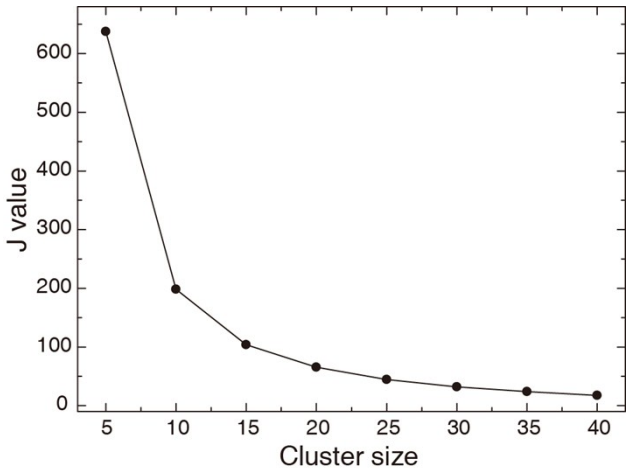

Figure S9. Cluster analysis of the PRE-based structure ensemble. The ensemble determined by the PRE data contains 100 conformers of the Hfq monomer. Each structure is represented by a vector comprising a list of C_α-C_α distances between three residues (S6, R17, and S23) in the core and those in HfqCTR. Therefore, the similarity between any two structures can be measured by the distance between the two corresponding vectors. Clustering was performed using the standard *k*-means algorithm using the above metrics, which aims to divide the 100 conformers into *k* clusters by minimizing a distortion function $J = \sum_{c=1}^{k} \sum_{X_t \in c} \left\| X_t - \mu_c \right\|^2$, where μ_c is the centroid of the cluster *c* (from 1 to *k*) that the structure X_t belongs to. *J* represents the within-cluster sum of squared distances between structures and cluster centroids. To determine a suitable cluster size, the cluster number varied from 5 to 40 with an increment of 5, and the corresponding J value is shown. A cluster size of 20 was chosen.

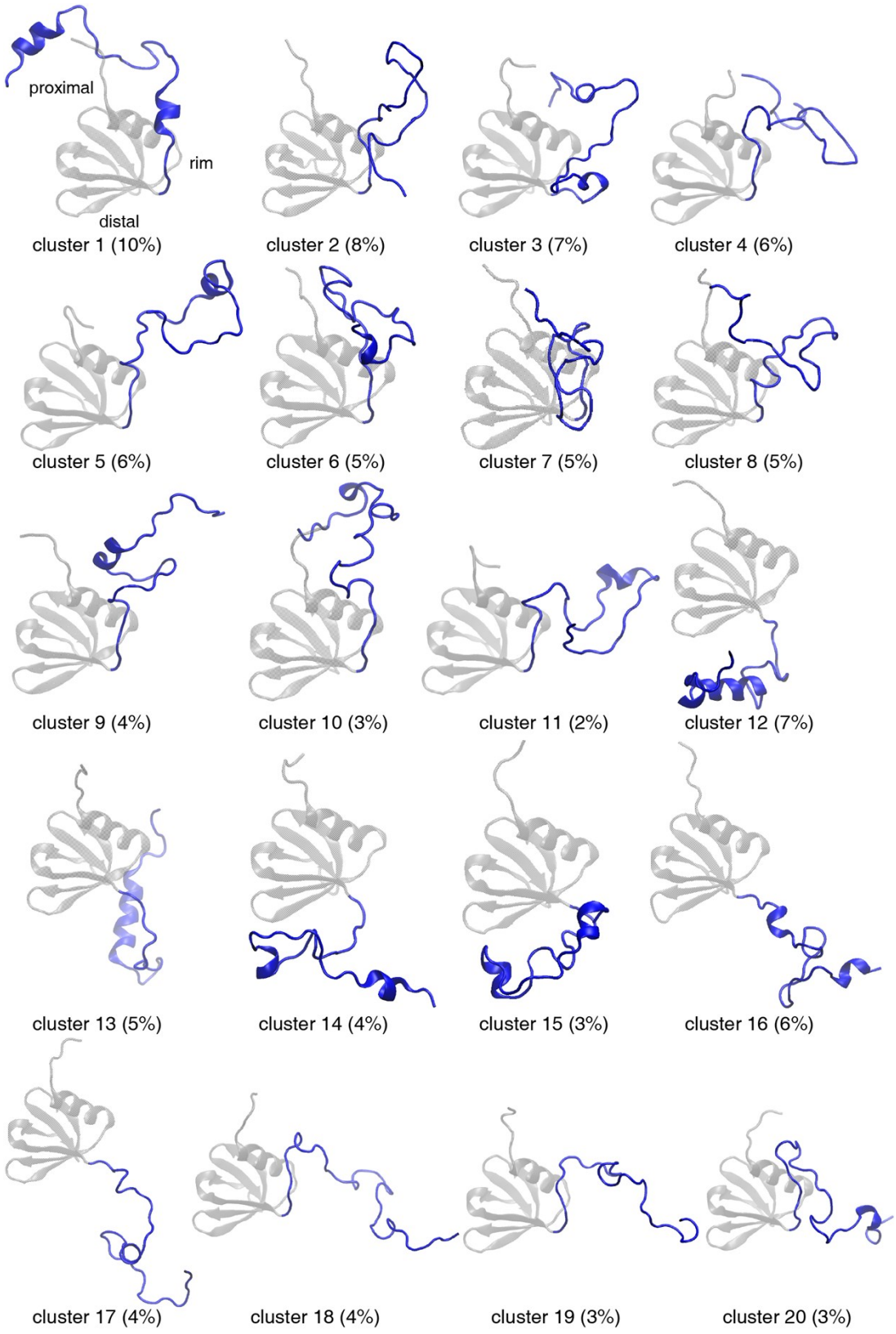

Figure S10. Cluster analysis of the structure ensemble based on PRE data. The representative conformer of each cluster is shown, and its population is given in parentheses.

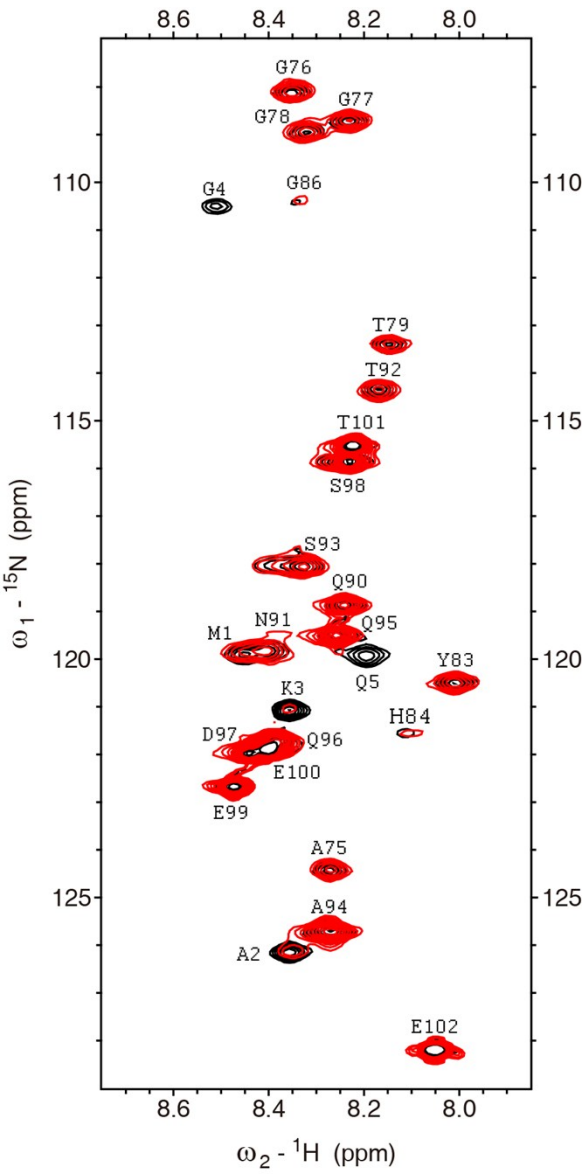

Figure S11. Superposition of a selected region of the ^{1}H-^{15}N HSQC spectra of Hfq in the free state (black) and in the presence of AU_6A (red). No significant chemical shift change is observed for residues in HfqCTR upon the binding of AU_6A.

Table S1. Time points of binding between the protein (Hfq or Hfq65) and the RNA (AU_6A or A_7) during CG simulations.

Ind[a]	Hfq-AU_6A	Hfq65-AU_6A	Hfq-A_7	Hfq65-A_7
1	7717000[b]	685000	**7712000**[c]	**8928000**
2	7133000	670000	**4518000**	**7715000**
3	5027000	609000	**3911000**	**6191000**
4	**458000**	**470000**	**2086000**	**5761000**
5	450000	341000	**1614000**	**4365000**
6	417000	339000	**1271000**	**3289000**
7	369000	275000	**1202000**	**2491000**
8	360000	252000	**870000**	**1820000**
9	284000	224000	**638000**	**1102000**
10	**180000**	**216000**	**619000**	**1046000**
11	172000	168000	**461000**	**1044000**
12	163000	162000	**373000**	**892000**
13	156000	118000	**216000**	**807000**
14	151000	104000	**198000**	**383000**
15	137000	104000	**153000**	**350000**
16	115000	89000	**139000**	**297000**
17	108000	77000	**113000**	**292000**
18	98000	53000	**99000**	**156000**
19	70000	46000	**77000**	**154000**
20	51000	33000	**64000**	**143000**

[a]The 20 CG simulations are indexed in descending order of the time points of protein-RNA binding

[b]Time step

[c]**Bold** indicates that Hfq is faster at binding RNA than Hfq65