Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

SI Appendix

Lingci Zhao[a,b], Rodolfo M. Rasia[c], Irina P. Suarez[c], Diego F. Gauto[c], Jin Wang[a,b,d,*]

[a]College of Physics, Jilin University, Changchun, Jilin 130012, People's Republic of China; [b]State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, Jilin 130022,People's Republic of China;
[c]Instituto de Biología Molecular y Celular de Rosario. 27 de Febrero 210 bis, predio CCT, 2000 Rosario, Argentina;
[d]Department of Chemistry and Physics, The State University of New York, Stony Brook, NY 11794-3400, United States of America

E-mail: jin.wang.1@stonybrook.edu

Simulation details

All simulations were carried out with Gromacs 4.5.4 (1), integrated by Langevin equation with constant friction coefficient γ = 1.0. The cut-off for nonbonded terms was set to 3.0 *nm*, and all bonds were constrained by LINCS algorithm to ensure the MD time step of 2 *fs*. The complex was placed at the center of a 50 *nm* * 50 *nm* * 50 *nm* cubic box. A strong harmonic potential was added if the distance between the center of mass of the two chains of complex is farther than 8 *nm* to enhance the sampling of binding.

We performed replica exchange molecular dynamics (REMD)(2) simulations to determine the binding transition temperature of the DCL1-A and dsRNA complex. 48 replicas covering a temperature range from 55 degree to 190 degree ensure an efficient sampling. Each replica was performed for 5 * 10^8 MD steps, and the neighbor replica attempted to exchange at every 2500 MD steps. The averages of exchange rates are from 18% to 42%, indicating an efficient sampling. Finally, we obtained the heat capacity curve at different temperatures, and we defined the binding transition temperature as T = 160.90 K at the peak of heat capacity. Thus our thermodynamic simulation temperature was set to T_s = 160.90 K, so that we can achieve the most binding/unbinding transitions in the limited simulation time.

Estimation of the binding affinity

To match the binding affinity according to the experimental measurement, we introduced ε_b to rescale the energy of intermolecular interactions between DCL1-A and dsRNA. To estimate the equilibrium disassociation constant K_D reflecting the binding affinity, we approximately consider the binding process as two-state kinetics for the formation of DCL1-A:dsRNA complex from free IDP and dsRNA at equilibrium, which may be expressed as,

$$DCL1\text{-}A + RNA \underset{K_{off}}{\overset{K_{on}}{\rightleftharpoons}} DCL1\text{-}A{:}RNA$$

where K_{on} is the second-order rate constant as the association rate, while K_{off} is the first-order rate constant for disassociation. The ratio between the two rate constants yields the equilibrium disassociation constant K_D (that has unit of concentration) given by,

$$K_D = \frac{K_{off}}{K_{on}} = \frac{[D]*[R]}{[C]} = \frac{P_u^2}{1-P_u} * [D_0] = \frac{P_u^2}{1-P_u} * \frac{397}{R^3}[mol/L]$$

We define the concentration of free DCL1-A [D_f], free RNA [R_f], and the DCL1-A:dsRNA complex [C]. Thus the total DCL1-A concentration [D_0] = [D_f] + [C], and total RNA concentration [R_0] = [R_f] + [C]. P_u is the population of unbound DCL1-A. In the present work, [D_0] = [R_0], so [D_f] = [R_f]. The concentration of DCL1-A is controlled by the wall potential. R in the unit of Å is the radius of the effective simulation spherical box whose size is determined by the wall potential. $\frac{397}{R^3}$ is the result of unit conversion from $1 molecule/Å^3$ to $1 mol/L$.

In our simulations, we set ε_b=0.9, resulting in a disassociation constant K_D=3.39μM, which is about ten times of the experimental measurement. In fact, the high affinity at room temperature and physiological conditions disfavors bind/unbinding transitions. Thus a higher and close to transition temperature and concentrated condition is widely used to obtain more transitions. Although the thermodynamic analysis was based on a slightly higher temperature than that in the experimental condition, the underlying mechanism is supposed to be robust.

The energy of intramolecular contacts within DCL1-A were rescaled by altering ε_f to match the helical content in the experiments. The fraction of helix was calculated from the number of consecutive torsions. A residual helix requires at least 3 consecutive torsions between 30° and 120°, similar to the settings in previous simulations (3, 4). The alpha helical content of DCL1-A in complex is at about 40%, which is consistent with the experimental data (about 45%). Finally, we calibrated ε_f to 1.05 to acquire the 8.1% alpha helical content of DCL1-A in free form (experimentally 3%). Considering that the structure-based models always overestimate the population of helical content in free form, the simulation results based on our model are reliable.

Analysis of the rates of binding

We calculated ln (binding rate), ln (capture rate) and ln (evolution rate) based on FPT_{on}, MPT_{cap} and FPT_{evo}, respectively,

$$lnk = \frac{1}{n}\sum_{i}^{n} ln\left(\frac{1}{tau_i}\right)$$

where k represents binding rate, capture rate and evolution rate, n is the 200 trajectories at each salt concentration and tau_i refers to FPT_{on}, MPT_{cap}, and FPT_{evo}.

Cut-off algorithm to count non-native contacts

Considering that the non-native contacts play an important role in the formation of DCL1-A:dsRNA complex, it is necessary to count both native and non-native contacts in a consistent manner. Here, we used a cut-off algorithm to describe the intermolecular contacts between the two chains. In the native structure, we find that the average distance between the two C_α atoms in a native intermolecular contact is 7.69 Å, while the minimum of the distance is 4.16 Å. Therefore, we define a contact value using a dual-radius cut-offs. We consider a contact between DCL1-A and dsRNA as fully formed (contact value 1.0) when the distance between the two Cα atoms is shorter than 5.0 Å (which is about 1.2 times 4.16 Å); while the contact is considered as partially formed (contact value 0.5) if the distance is greater than 5.0 Å and under 9.5 Å (about 1.2 times the mean distance 7.69 Å). Using this algorithm to describe the native complex, we achieve a similar counting of both native and non-native contacts.

Wall potential

We also introduced a wall potential, whose functional form can be expressed as,

$$V_{wall}(R) = K_w * \left(|R - R_L| + |R - R_H| - |R_L - R_H|\right)^4$$

where K_w is an energy constant 40 *kJ/mol*. R is the distance of the center of mass between DCL1-A and dsRNA. R_L is the lowest limit distance, and R_H is the highest limit distance. The wall potential begins to function when the distance of center of mass between DCL1-A and dsRNA is greater than R_H or lower than R_L. With the help of this wall potential, we can control the movement of DCL1-A in a convenient way, such as keeping it in the unbound state or increasing the protein concentrations. Here, we set R_H as 8.0*nm* to construct a concentrated condition.

1. Hess B, Kutzner C, van der Spoel D, & Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* 4(3):435-447.
2. Sugita Y & Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314(1-2):141-151.
3. Chu WT, Clarke J, Shammas SL, & Wang Je (2017) Role of non-native electrostatic interactions in the coupled folding and binding of PUMA with Mcl-1. *Plos Computational Biology* 13(4).
4. Rogers JM*, et al.* (2014) Interplay between partner and ligand facilitates the folding and binding of an intrinsically disordered protein. *P Natl Acad Sci USA* 111(43):15420-15425.