

A. Appendix: The MM-PBSA method

The MM-PBSA method has been shown to be useful in many applications to compute binding free energies for protein-protein and protein-nucleic acid interactions²⁵ and recently for small organic ligand – RNA aptamer interactions.³²⁻³³ The details of the calculations of individual contributions to the binding free energy are given elsewhere,^{24-26,32-33}

The binding free energy is defined as follows:

$$\Delta G_{binding} = G_{complex} - G_{RNA} - G_{Ligand}$$

where ligand is either MAG or TMR molecule, and G_i is the free energy of molecule i computed by averaging its conformational free energies over the 2.5 ns trajectory. The free energy G_i term is computed as follows:

$$G_i = E_{mm} + E_{PB} + E_{NP} - TS$$

where E_{mm} , E_{PB} , E_{NP} , T , and S denote the conformational energy (in vacuum) contribution to the free energy expression, electrostatic component of the solvation energy, non-polar component of the solvation energy, temperature and conformational entropy, respectively. Reference 26 gives additional computational details of parameters used in this work.

Table A1 lists individual contribution to free energy of binding of both MAG-RNA and TMR-RNA aptamer complexes whose overall binding free energies are reported in the text. Table A2 details individual contribution to the conformational energy (the force field) term in vacuum.

B. Appendix: Statistical inefficiency

In conformational sampling using molecular dynamics, successively generated configurations in the trajectory are highly correlated. In order to calculate meaningful averaged properties, it is essential to know the relaxation time for which a newly generated configuration is configurationally independent from previous ones. The time relaxation is useful for assessing the statistical independence and sampling space of a particular trajectory. At the same time, it is also useful for assessing the stability or persistency of a particular mode of interactions, since a long relaxation time correlates with the stability of the interaction over the entire trajectory, while a short relaxation time is the opposite.

The relaxation time can be computed using the statistical inefficiency method developed by Jacucci and Rahman³⁰. Here we summarize the essence of this method.

Given a set of n-sequential data points:

$$A = \{a_1, a_2, \dots, a_n\} \quad (\text{Eq. B1})$$

breaking \mathbf{A} into \mathbf{b} blocks A_1, A_2, \dots, A_b of equal length \mathbf{n}_b (block size). We compute the block average for some property $\langle A \rangle_i$ for each block \mathbf{A}_i and the average $\langle A \rangle$ for all data points covered in the blocks:

$$\langle A \rangle_i = \frac{1}{n_b} \sum_{j=1}^{n_b} a_j \quad \text{where } a_j \in \text{block } A_i \text{ for } i = 1..b \quad (\text{Eq. B2})$$

$$\langle A \rangle = \frac{1}{n} \sum_{i=1}^n a_i \quad (\text{Eq. B3})$$

Block variance for a given block size \mathbf{n}_b ($\sigma^2(\langle A \rangle_{n_b})$) and variance for each individual data points covered in the blocks ($\sigma^2(\langle A \rangle)$) are computed as follows:

$$\sigma^2(\langle A \rangle_{n_b}) = \frac{1}{b} \sum_{i=1}^b (\langle A \rangle_i - \langle A \rangle)^2 \quad (\text{Eq. B4})$$

$$\sigma^2(\langle A \rangle) = \frac{1}{n} \sum_{i=1}^n (a_i)^2 - \langle A \rangle^2 \quad (\text{Eq. B5})$$

The statistical inefficiency, a function of \mathbf{n}_b , is then defined as follows:

$$s(n_b) = \lim_{n_b \rightarrow \infty} \frac{n_b \sigma^2(\langle A \rangle_{n_b})}{\sigma^2(\langle A \rangle)} \quad (\text{Eq. B6})$$

The relaxation time can be derived by plotting the statistical inefficiency s versus the block size \mathbf{n}_b . It is then the value at which a plateau (asymptotically approached) occurs for the large value of block size \mathbf{n}_b . If each element of a given set of data is independent (random) from each other, a plateau should be reached at a small value of s . On the other hand, a large value of s will result if each successive element of a given set of data is strongly correlated. Figure B1 illustrates these two extremes. In Figure B1a, all 5000 elements of the data set are randomly generated from the normal distribution; whereas in Figure B1b, all 5000 elements are generated sequentially from 1 to 5000 with an equal interval of length 1. In the former case, a plateau occurs at $s \sim 1.0$. In the latter case, the plateau would never occur, which gives infinite relaxation time. For a trajectory generated from molecular dynamics or from a set of configurations generated by the Monte-Carlo method, the behavior of s should lie between these two extreme cases.

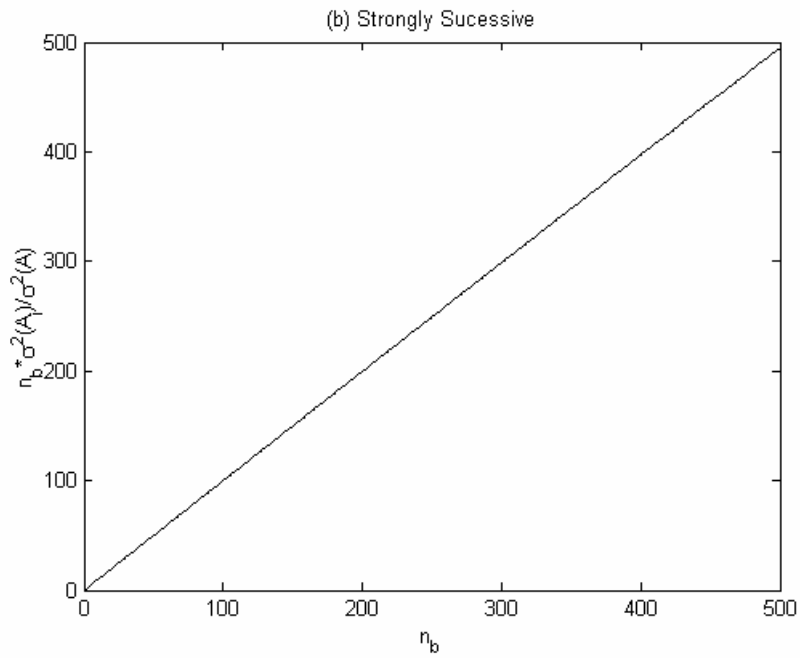
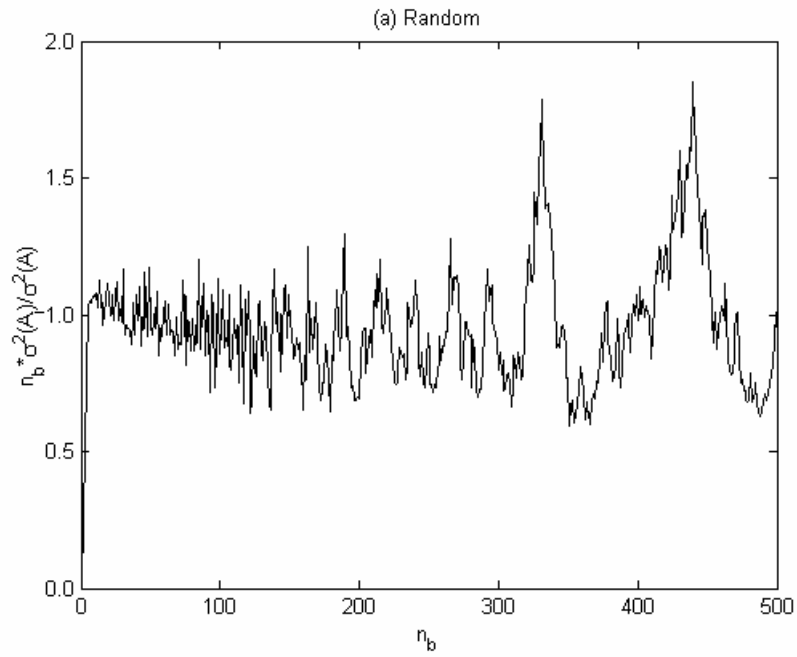


Table A1: Individual Contribution to Free Energy of Binding

	E_{mm}		E_{PB}		E_{NP}		Sub-total		Entropy (+TS)		Total
	Average	St. Err.	Average	St. Err.	Average	St. Err.	Average	St. Err.	Average	St. Err.	
MAG-RNA Complex	6107.5	7.0	-14533.4	6.7	40.7	0.0	-8385.3	0.8	1027.5	6.1	-9412.8
TMR-RNA Complex	5969.2	3.5	-14459.6	3.3	40.9	0.0	-8449.5	0.8	1025.2	4.8	-9474.7
TMR	74.4	0.1	-40.2	0.0	4.7	0.0	38.9	0.1	53.8	0.0	-14.9
MAG	126.6	0.1	-39.5	0.0	4.8	0.0	91.8	0.1	54.3	0.0	37.6
RNA	6928.6	8.7	-15417.1	8.5	41.2	0.0	-8447.3	0.7	994.9	5.3	-9442.2
Binding Energy of MAG	-947.7		923.2		-5.3		-29.8		-21.6		-8.2
Binding Energy of TMR	-1033.8		997.7		-5.0		-41.1		-23.5		-17.6

Table A2: Individual Contribution to the E_{mm} Term

	$E_{v.d.w}$		$E_{e.e}$		E_{bond}		E_{mm}	
	Average	St. Err.	Average	St. Err.	Average	St. Err.	Average	St. Err.
MAG-RNA Complex	-407.9	0.4	4496.6	7.1	2018.7	0.7	6107.5	7.0
TMR-RNA Complex	-403.8	0.4	4365.1	3.5	2007.9	0.8	5969.2	3.5
TMR	16.2	0.1	-12.0	0.0	70.2	0.1	74.4	0.1
MAG	16.8	0.1	34.8	0.0	74.9	0.1	126.6	0.1
RNA	-382.4	0.4	5373.7	8.8	1937.4	0.7	6928.6	8.7
Binding Energy of MAG	-42.3		-911.9		6.5		-947.7	
Binding Energy of TMR	-37.6		-996.6		0.4		-1033.8	