Supporting Information

Capture and quality control mechanisms for ATP binding

Li Li, Susan A. Martinis and Zaida Luthey-Schulten

Supporting text.

1 Determination of the coordination and protonation state of ATP

To determine the coordination state of ATP, we first inspected the crystal structure, and noticed that one oxygen from each of the three phosphate groups coordinate with Mg^{2+}, with the rest of oxygen atoms acting as hydrogen bond acceptors. This suggests that ATP has a protonation state as $[ATP]^{-4}$ and a coordination state of $[ATP{:}Mg]^{2-}$. This assignment is consistent with the result that $[ATP{:}Mg]^{2-}$is the predominant species in the solution of pH 6.5 and 5 mM Mg^{2+}(the crystallization condition of the ternary complex [1]), using the equilibrium constant measured by Sigel *et al* [2]. In addition, three water molecules were coordinating with the Mg^{2+}in the crystal structure, resulting a stable octahedral coordination state of Mg^{2+}.

2 Free energy reconstruction from SMD simulations

In SMD simulation, the Hamiltonian of the system, $\tilde{H}(r,p;\lambda)$, was the sum of its original Hamiltonian, $H(r,p)$, and a guiding potential:

$$\tilde{H}(r,p;\lambda) = H(r,p) + \frac{k}{2}[x(r)-\lambda]^2, \tag{1}$$

where λ is the position of an imaginary particle moving at a constant speed v: $\lambda_t = \lambda_0 + vt$. Following the Jarzynski identity [3], the change of this perturbed system's free energy, $\tilde{F}$, can be estimated by

$$\exp\left[-\beta(\tilde{F}(\lambda_t) - \tilde{F}(\lambda_0))\right] = \langle \exp\left(-\beta W_{0\to t}\right)\rangle_{\text{traj}}, \tag{2}$$

where $W_{0\to t}$ is the work done on the perturbed system during the time interval 0 to t. For a finite-size sampling, an unbiased estimate of the ensemble average of $\exp(-\beta W)$ is given by

the second cumulant expansion, and after some manipulations of Equation 2, we have

$$\tilde{F}(\lambda_t) - \tilde{F}(\lambda_0) = \langle W_{0\to t}\rangle - \frac{\beta}{2}\left(\langle W_{0\to t}^2\rangle - \langle W_{0\to t}\rangle^2\right). \tag{3}$$

Finally, by evoking stiff-spring approximation [4], the free energy change of the unbiased system (F) can be approximated by

$$F(\lambda) \approx \tilde{F}(\lambda). \tag{4}$$

3 Comments on the unbound state in the metadynamics simulations

Strictly speaking, the center-of-mass distance between ATP and GluRS in the unbound state should be infinite. In the metadynamics simulations, the true unbound state was approximated by a state that separated ATP and GluRS by 20 Å. Additional SMD simulations have been performed to pull the ATP further away by 10 Å. This ensures the minimum distance between ATP and any atoms in GluRS is 18 Å, which is longer than the cutoff distance for the nonbonded interactions in MD simulations. Therefore, in the end state of SMD run, there is no interaction between ATP and GluRS, which is, energetically, equivalent to an "infinite" distance. As such, the free energy difference between the start and end states of the SMD run equals to the difference between the unbound states in metadynamics run and the theoretical unbound states. By using Jarzynski identity, the free energy difference was calculated to be merely -0.3 kcal/mol, therefore justifying the assignment of the unbound state.

References

[1] S. Sekine, O. Nureki, D.Y. Dubois, S. Bernier, R. Chênevert, J. Lapointe, D.G. Vassylyev, and S. Yokoyama. ATP binding by glutamyl-tRNA synthetase is switched to the

productive mode by tRNA binding. *EMBO J.*, 22:676–688, Feb 2003.

[2] Helmut Sigel, Roger Tribolet, Raman Malini-Balakrishnan, and R. Bruce Martin. Comparison of the stabilities of monomeric metal ion complexes formed with adenosine 5'-triphosphate (ATP) and pyrimidine-nucleoside 5'-triphosphate (CTP, UTP, TTP) and evaluation of the isomeric equilibria in the complexes of ATP and CTP. *Inorganic Chemistry*, 26(13):2149–2157, 1987.

[3] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78(14):2690–2693, Apr 1997.

[4] S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten. Free energy calculation from steered molecular dynamics simulations using Jarzynskis equality. *J. Chem. Phys.*, 119(6):3559–3566, Aug 2003.

The complete reference 33 in the main text is:

A. D. MacKerell, D. Bashford, Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102(18):3586–3616, 1998.

Supporting figures

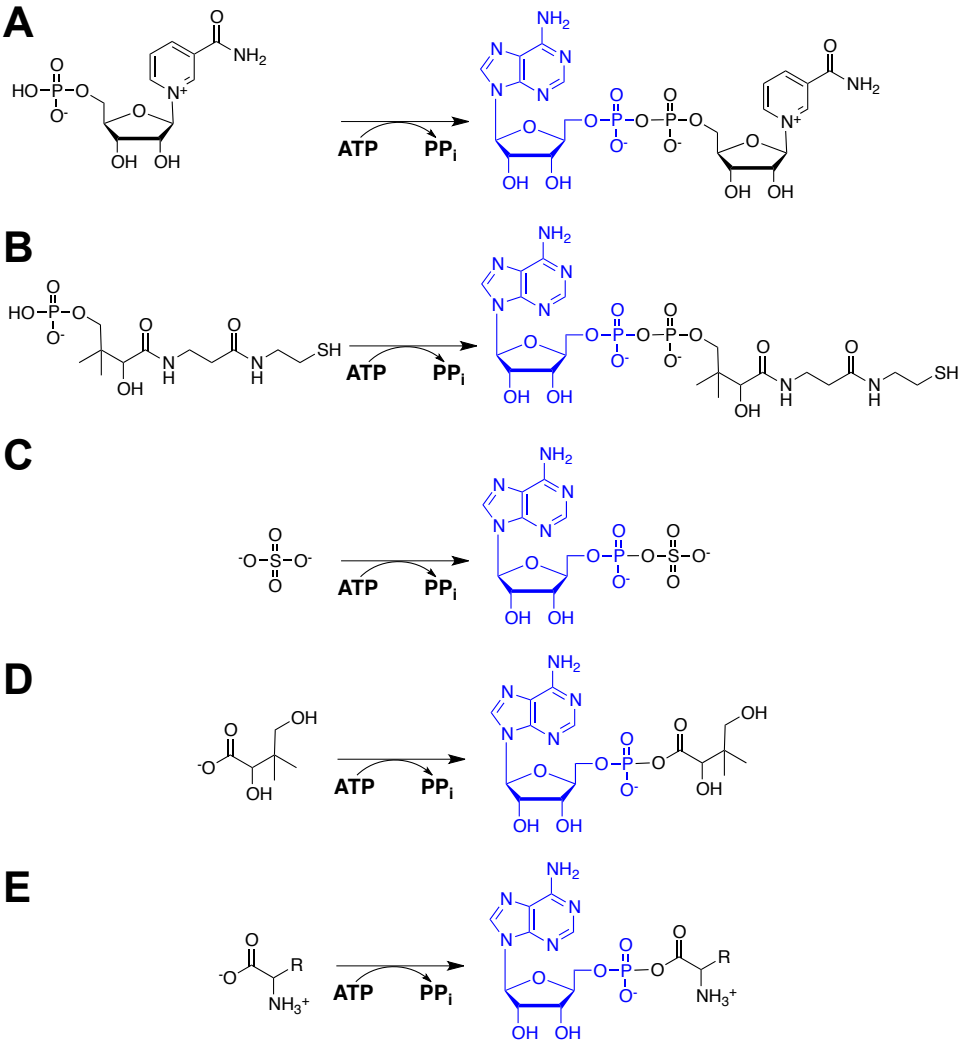

Fig. S1: **Reaction schemes for nucleotidylyl-transferase superfamily.** A. The AMP moiety of ATP is linked to nicotinamide mononucleotide (NMN) by NMN adenylyltransferase. B. Phosphopantetheine adenylyltransferase catalyzes the transfer of an adenylyl group from ATP to 4′-phosphopantetheine, yielding 3′-dephospho-CoA. C. ATP sulfurylase activates sulfate with ATP to give adenosine-5′-phosphosulfate. D. Pantothenate synthetase activates pantoate with ATP to form pantoyl-adenylate intermediate, which further reacts with β-alanine to form pantothenate. E. Amino acid is activated by ATP to form aminoacyl-adenylate, which subsequently transfers its aminoacyl moiety to tRNA. In all five reactions, the transferred adenylate moiety is shown in blue.

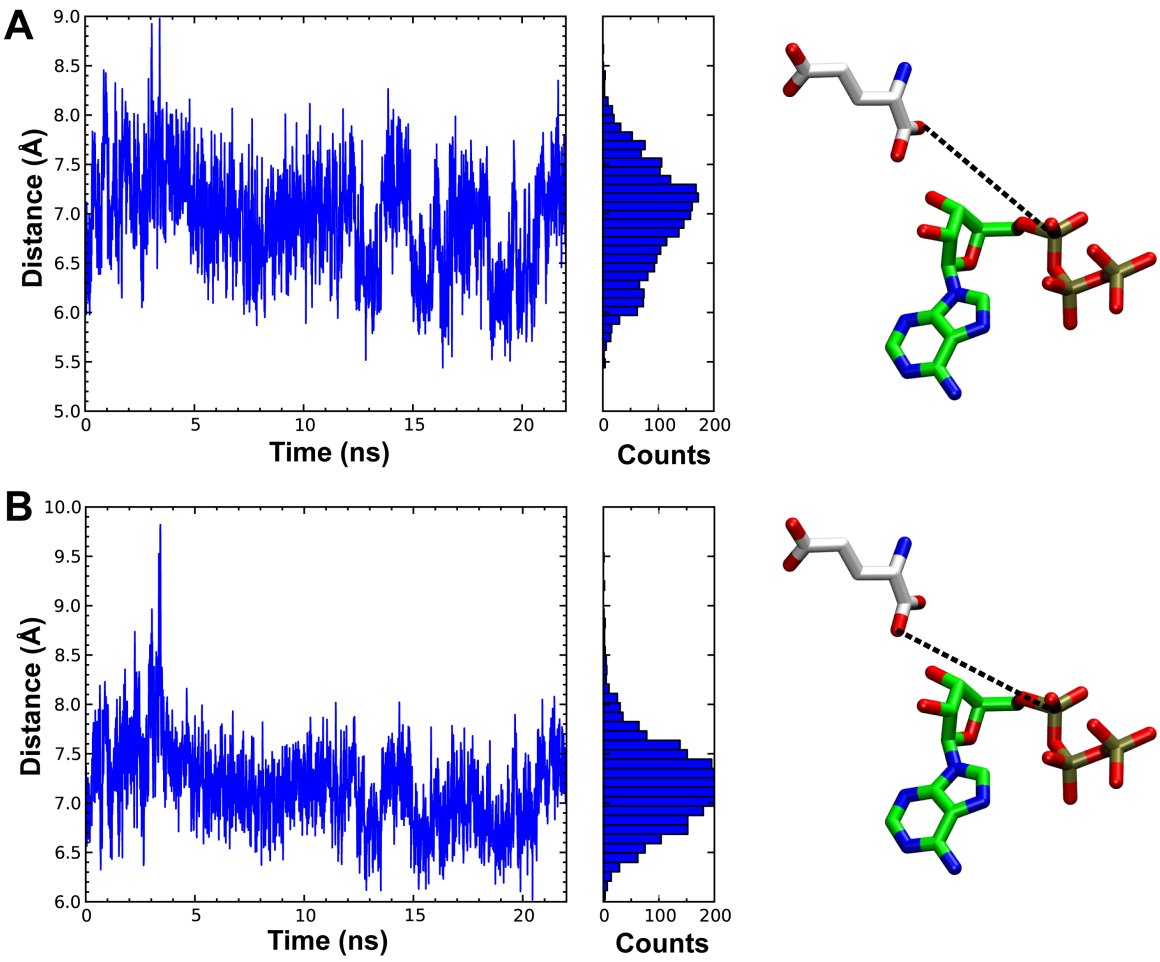

Fig. S2: **ATP adopts an inactive conformation in 22-ns unbiased MD.** The distances between P_α in ATP (green) and oxygens in the backbone of glutamate (white) are plotted as a function of simulation time. The distributions of the distances during the simulations are shown in the histograms.

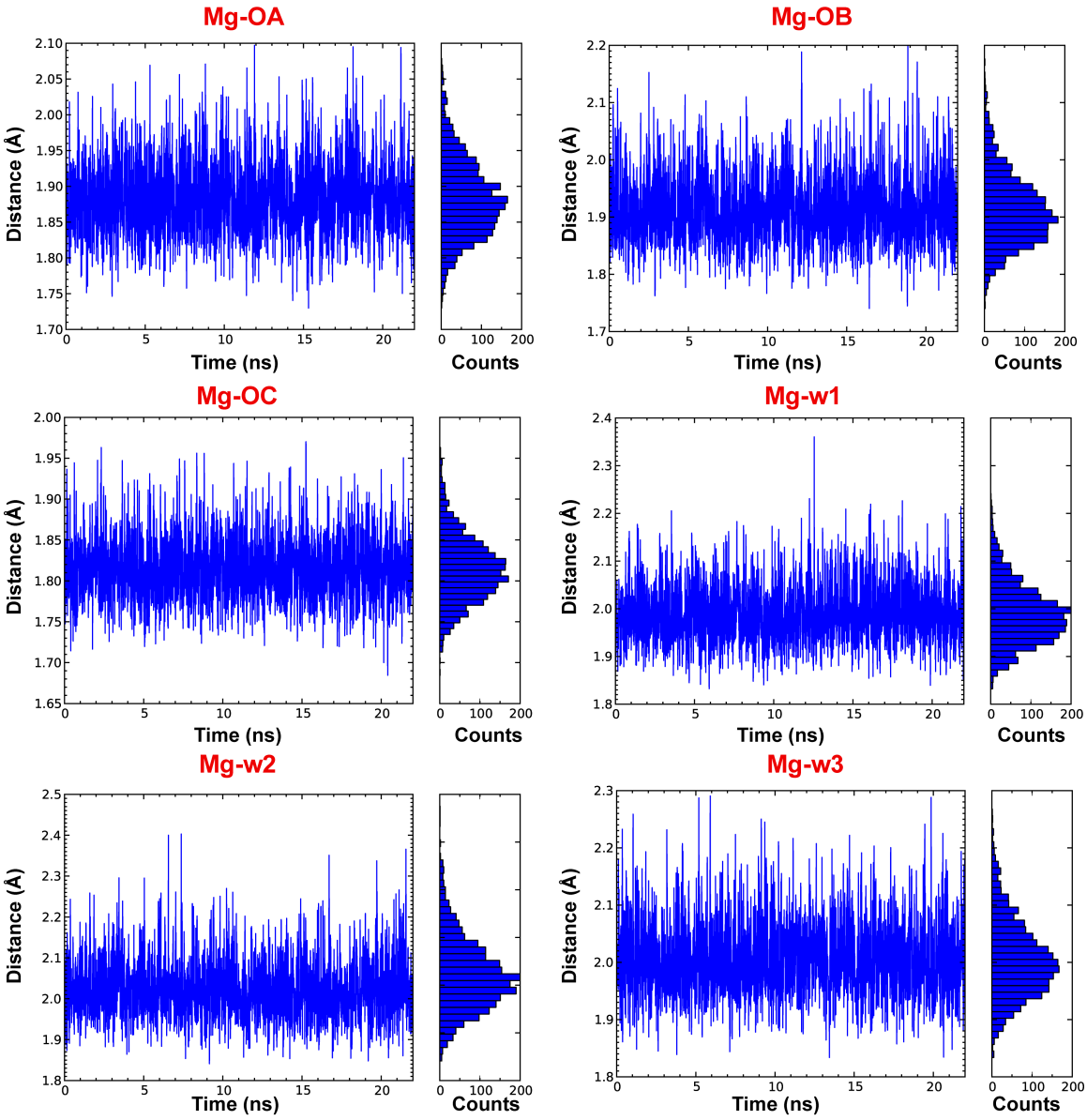

Fig. S3: **$[ATP{\cdot}Mg{\cdot}3H_2O]^{2-}$ complex remains stable in GluRS active site during a 22-ns unbiased MD simulation.** The distance between Mg^{2+} and each coordinating oxygen atom is plotted as a function of simulation time. The distributions of the distances during the simulations are shown in the histograms.

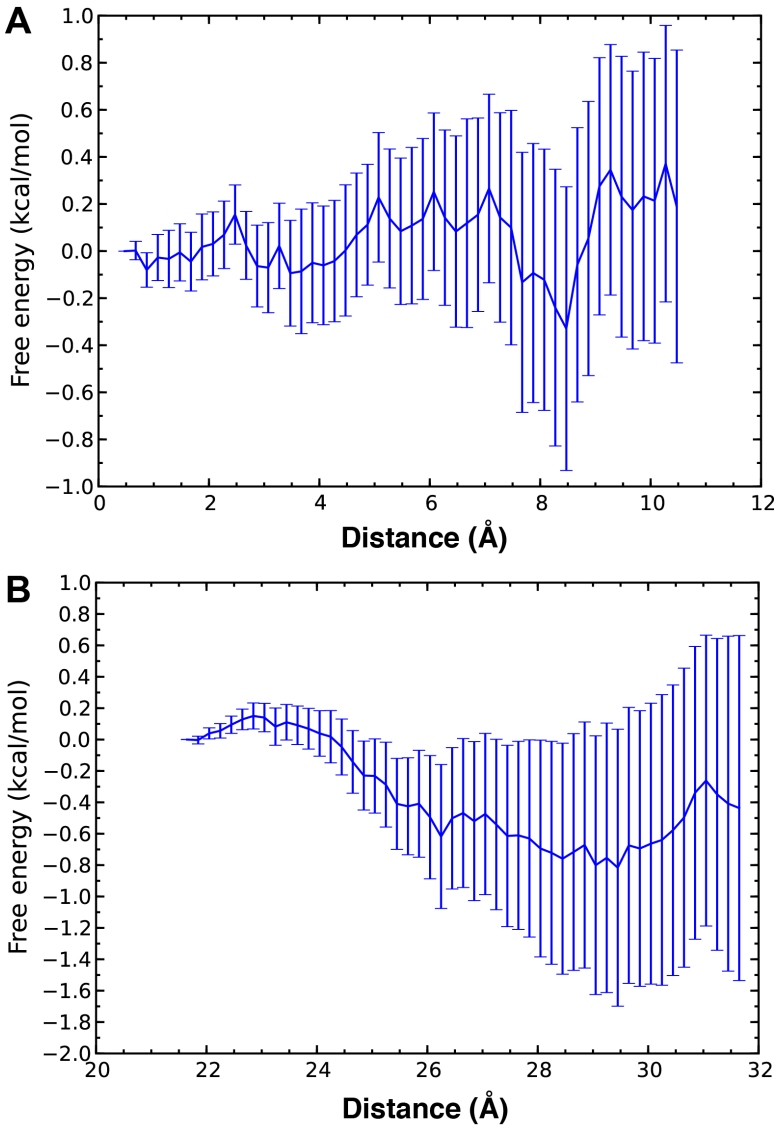

Fig. S4: **Free energy profile of ATP pulling.** Free energy profile of ATP pulling in water (A) or from the unbound state (B) is plotted as a function of the center-of-mass distance between ATP and the origin of the Cartesian coordinate (A) or the center-of-mass distance between ATP and GluRSs' active site (B). The average free energy was calculated from 50 and 120 independent 1-ns SMD simulations, respectively. The error bar indicated the standard deviation calculated from 200 bootstrapping.

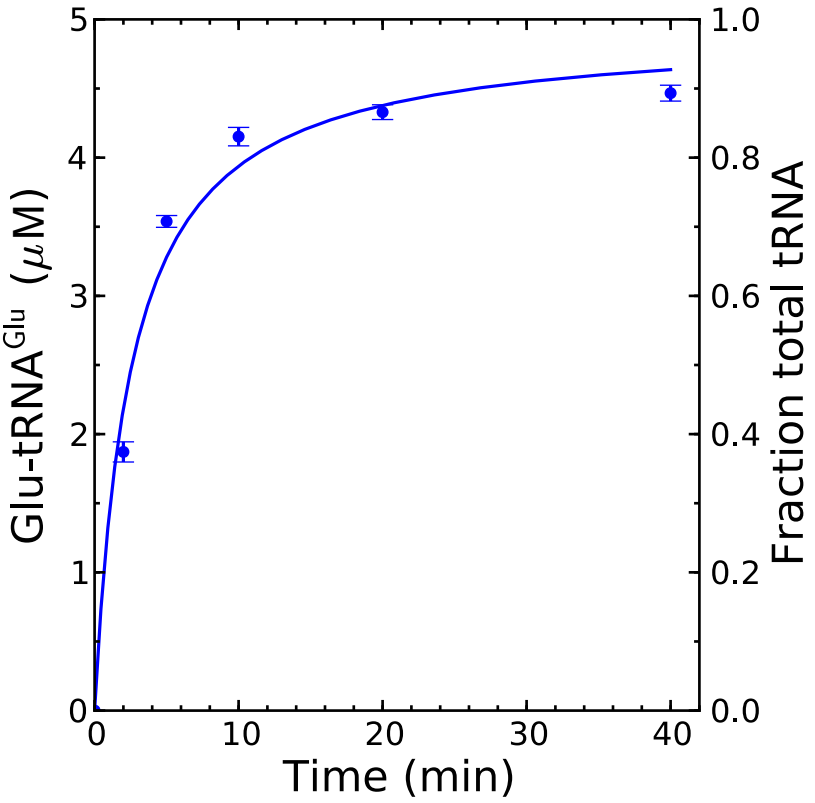

Fig. S5: **Recombinant *T. thermophilus* GluRS aminoacylates transcribed *T. thermophilus* tRNA^{Glu}.** The aminoacylation assay was performed at 37 °C, pH 7.5, and the reaction buffer included 40 μM [^{14}C]-glutamate, 2 mM ATP, 5 μM tRNA^{Glu} and 100 nM GluRS. Error bars represent standard deviations from triplicated reactions.

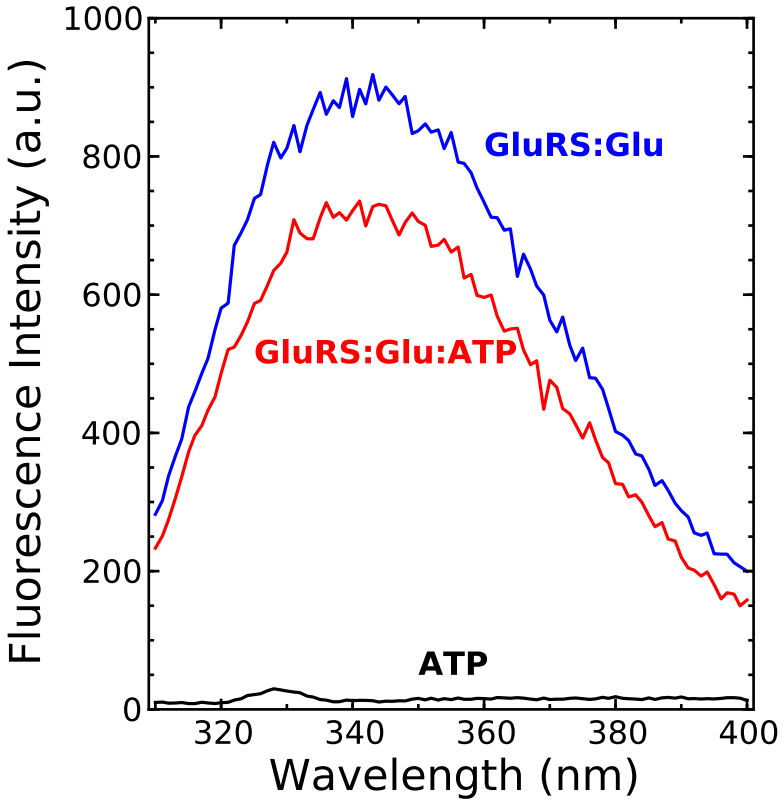

Fig. S6: **Fluorescence emission spectra for *T. thermophilus* GluRS.** The fluorescence emission spectra for ATP, *T. thermophilus* GluRS:Glu and GluRS:Glu:ATP are shown (λ_{ex} = 295 nm, λ_{em} = 310-400 nm). Emission spectra were measured in a buffer with 60 mM tris(hydroxymethyl)aminomethane (Tris) pH 7.5, 10 mM $MgCl_2$ and 1 mM dithiothreitol (DTT). For GluRS:Glu and GluRS:Glu:ATP, the concentrations of GluRS and Glu were 0.5 μM and 1 mM, respectively. For ATP and GluRS:Glu:ATP, concentration of ATP was 1 mM.

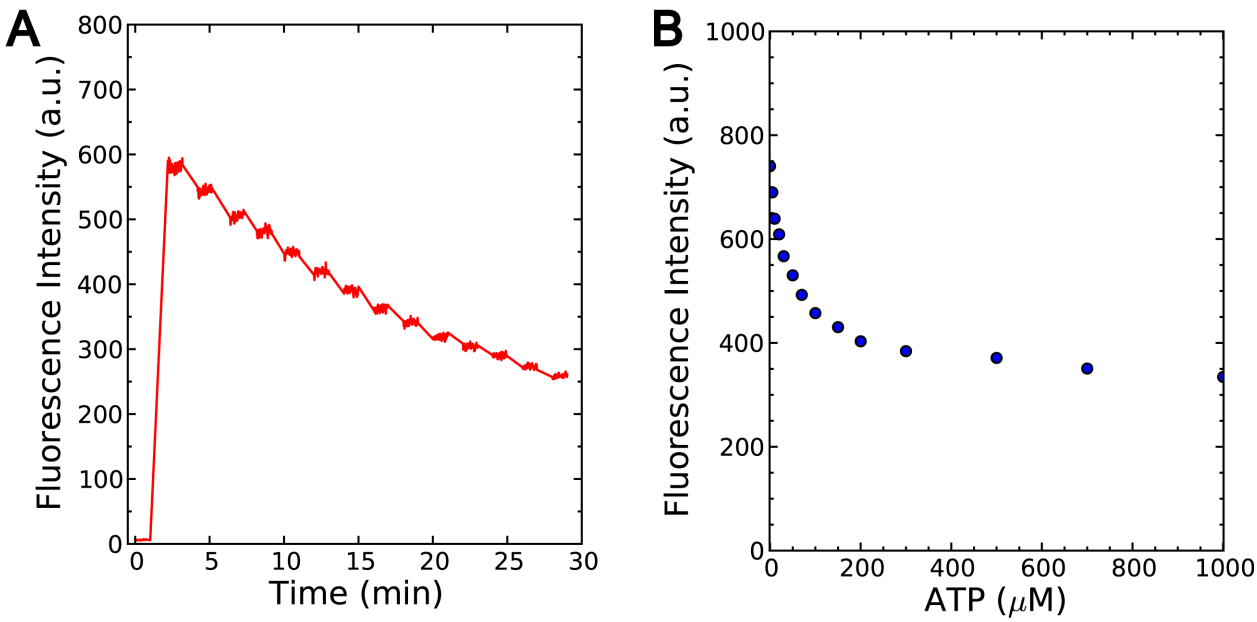

Fig. S7: **Titration of GluRS:Glu with ATP.** A. A typical time-course of the fluorescence titration of ATP to 0.25 μM GluRS:Glu complex. B. The fluorescence intensity after the correction of inner filter effect.

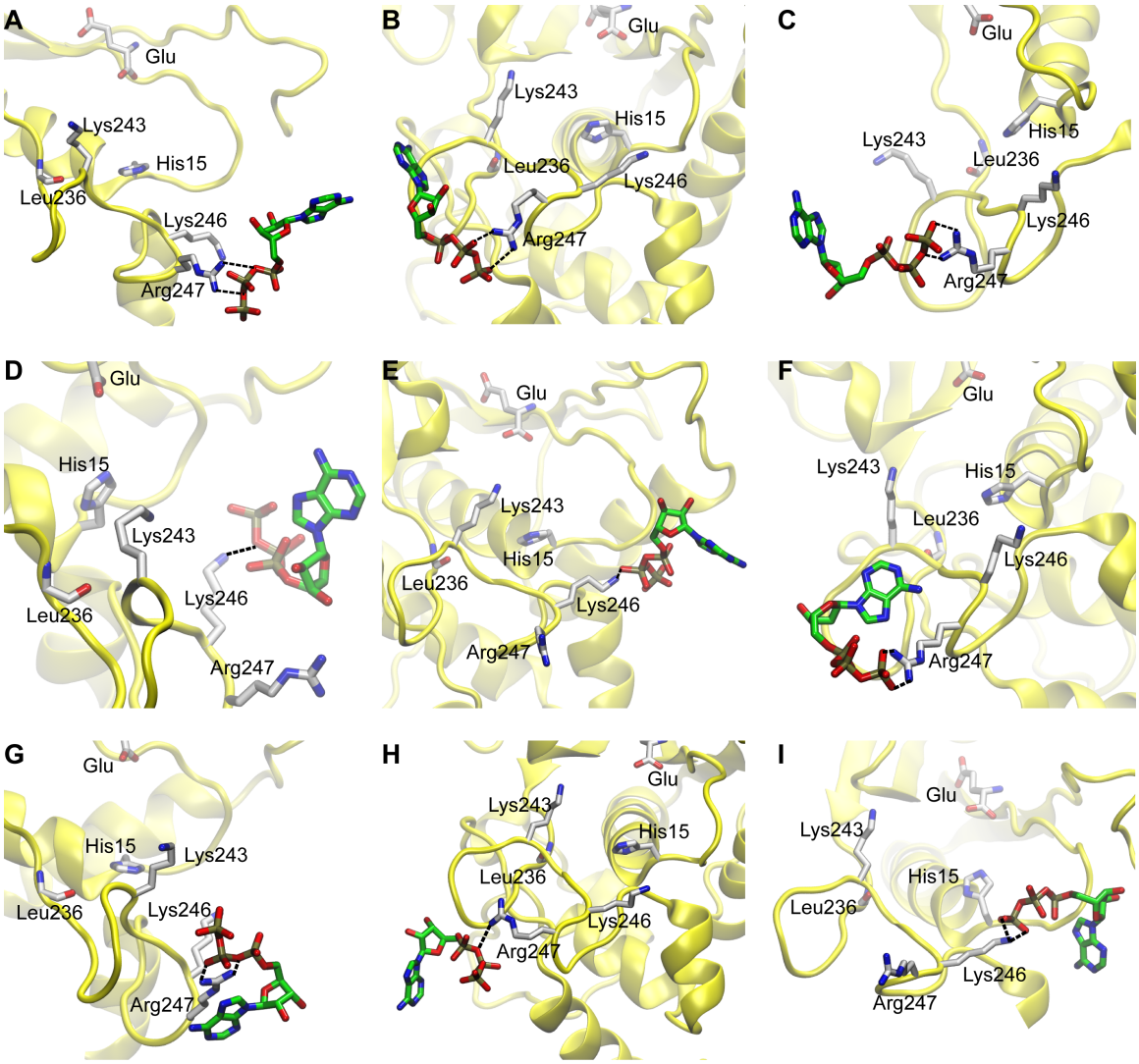

Fig. S8: **Encounter complex ensemble of ATP binding identified by unbiased MD simulations.** From sixteen unbiased simulations, we identified nine putative states of the encounter complex ensemble (A-I). The ATP and its interacting residues were shown in licorice, while the protein was shown in yellow cartoon. Key interactions that stabilize the encounter complexes are indicated by black dashed lines. In state G, the π-cation interactions between Arg247 and adenine also contributed to stabilize the encounter complex. The hydrogen atoms were not displayed for clarity.

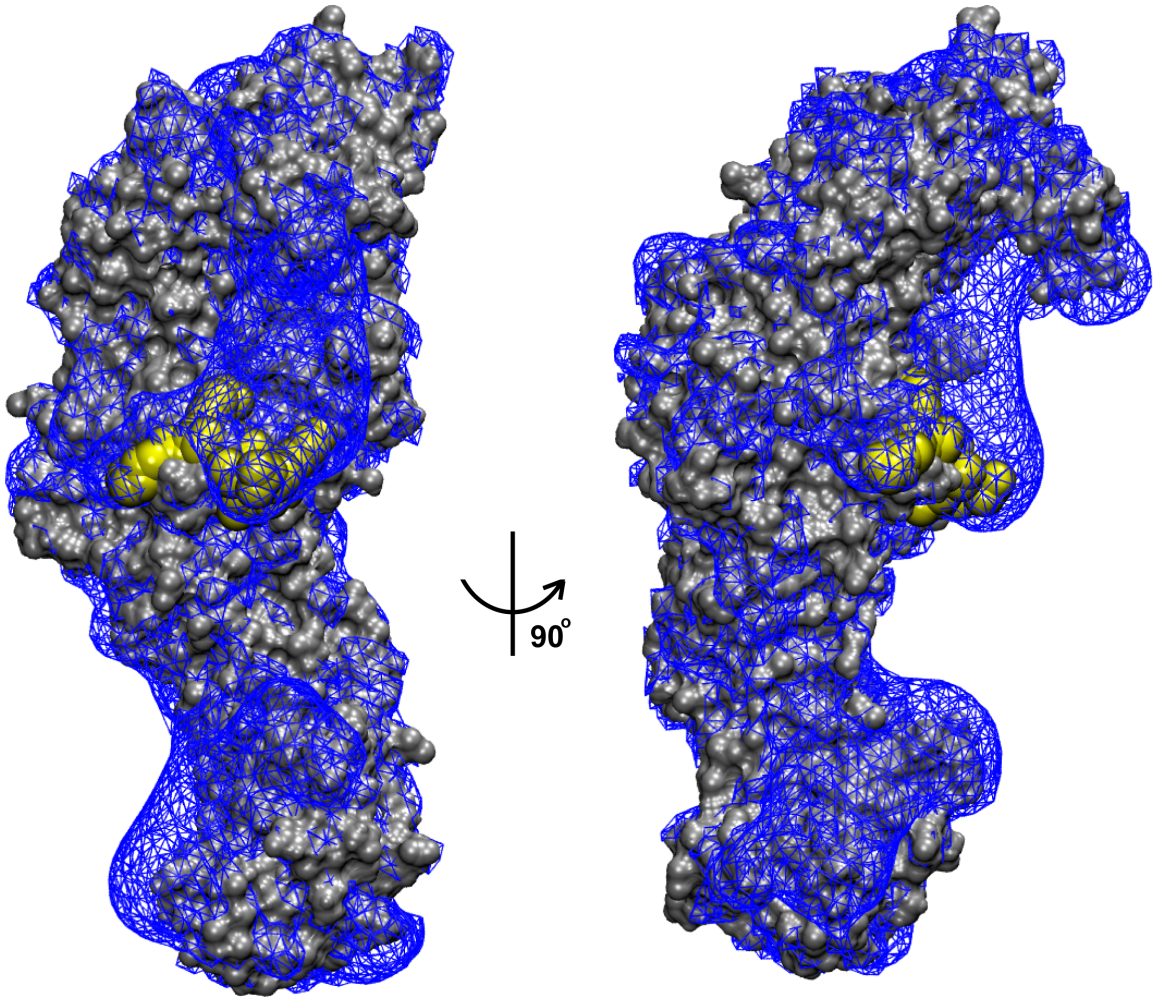

Fig. S9: **Electrostatic properties of GluRS:Glu complex.** Molecular surfaces of GluRS:Glu in grey with its electrostatic potential isosurface, calculated at 1 k_BT, displayed in blue meshes. The KMSK loop is shown in yellow with VDW representations.

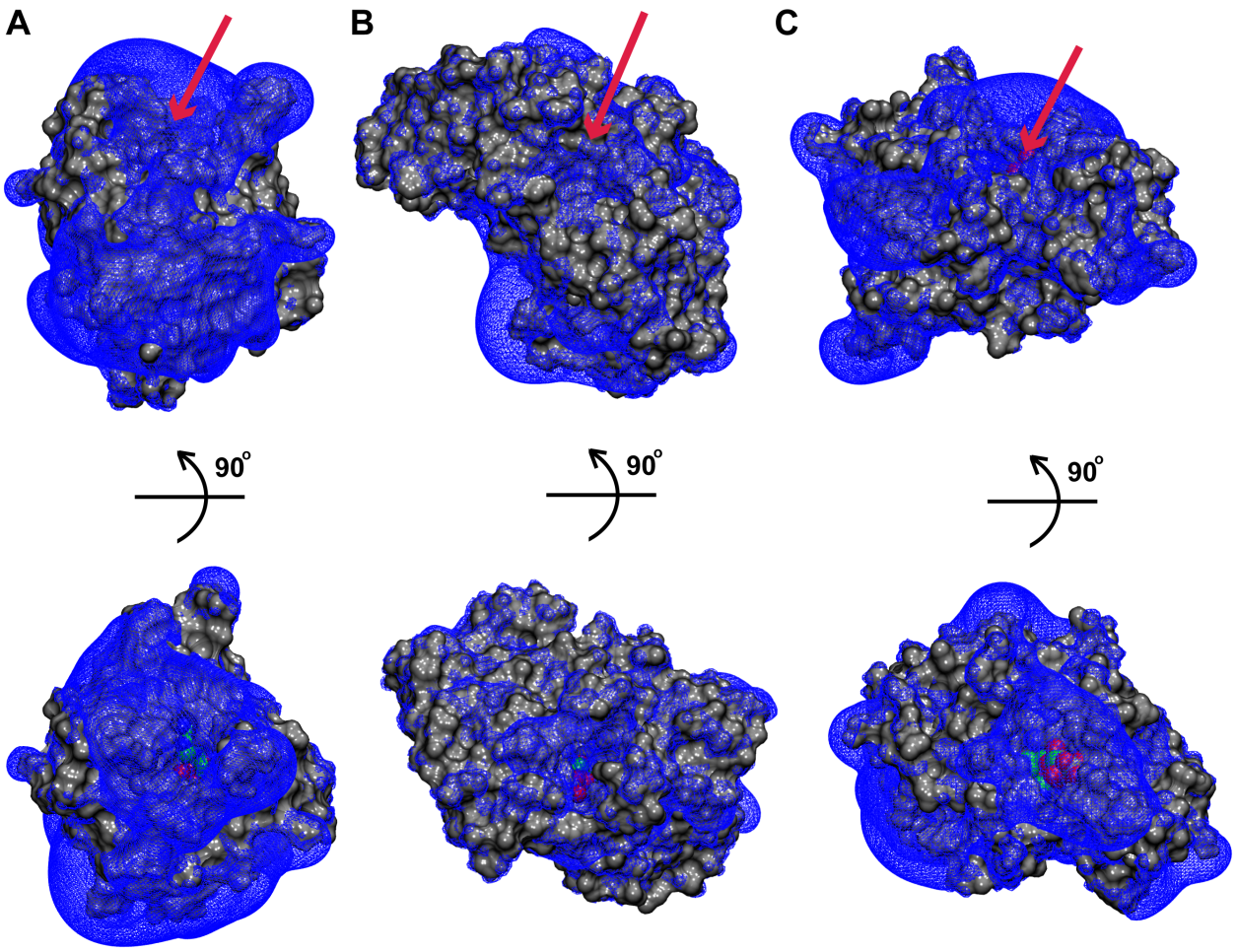

Fig. S10: **Electrostatic surfaces of nucleotidylyl transferase superfamily members.** The catalytic cores of *Saccharomyces cerevisiae* ATP sulfurylase (PDB 1G8G; A), *Mycobacteria tuberculosis* pantothenate synthetase (PDB 1N2B; B) and human cytosolic NMN/NaMN adenylyltransferase (PDB 1NUS; C) are shown in grey surface representation and their 1 k_BT electrostatic potential isosurface ares displayed in blue meshes. In each panel, ATP is shown in spheres and its binding pocket is indicated by a red arrow.

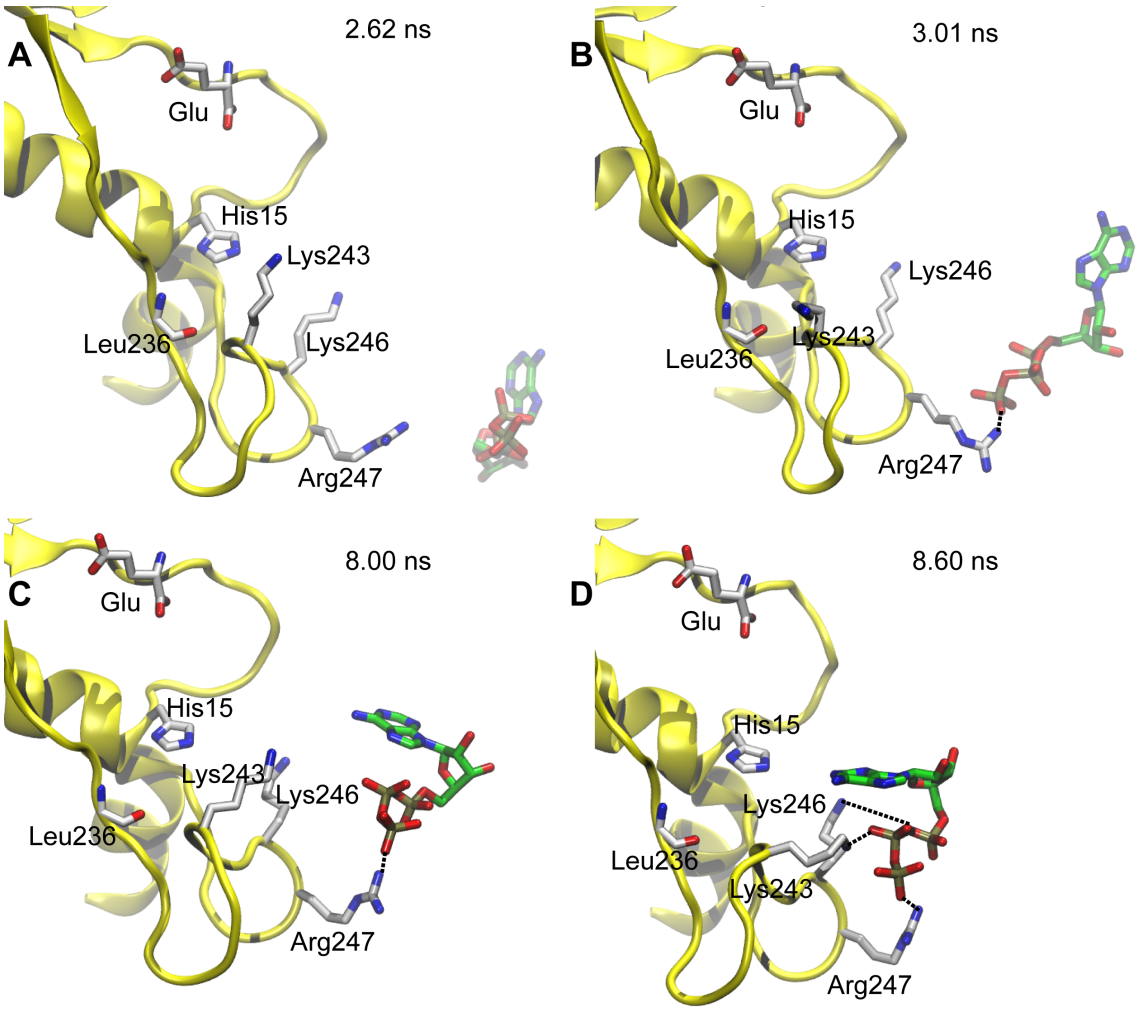

Fig. S11: **Binding of the triphosphate group of ATP to GluRS.** Four snapshots were taken along the 10-ns trajectory that captures the binding of triphosphate group to GluRS. The ATP and its interacting residues were shown in licorice, while the protein was shown in yellow cartoon. Key interactions that stabilize the encounter complexes are indicated by black dashed lines. The hydrogen atoms were not displayed for clarity. The side chain of Arg247 gradually folds back as ATP binds, reminiscent of the "fly-casting" mechanism of DNA-protein binding.

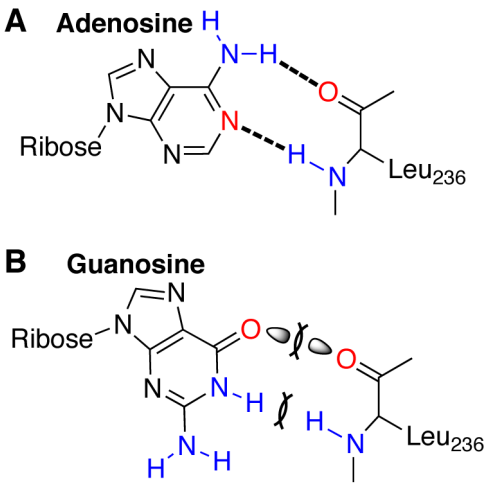

Fig. S12: **Proposed selection mechanism of ATP over GTP in GluRS active site.** A. Adenosine binds to the backbone of Leu236 via two hydrogen bonds (black dashed lines). B. Guanosine is rejected by repulsive electrostatic interactions. The hydrogen-bond acceptor and donor groups are colored by red and blue, respectively. The clashes between electrons (grey lobes) or hydrogens are indicated by crossed parenthesis.

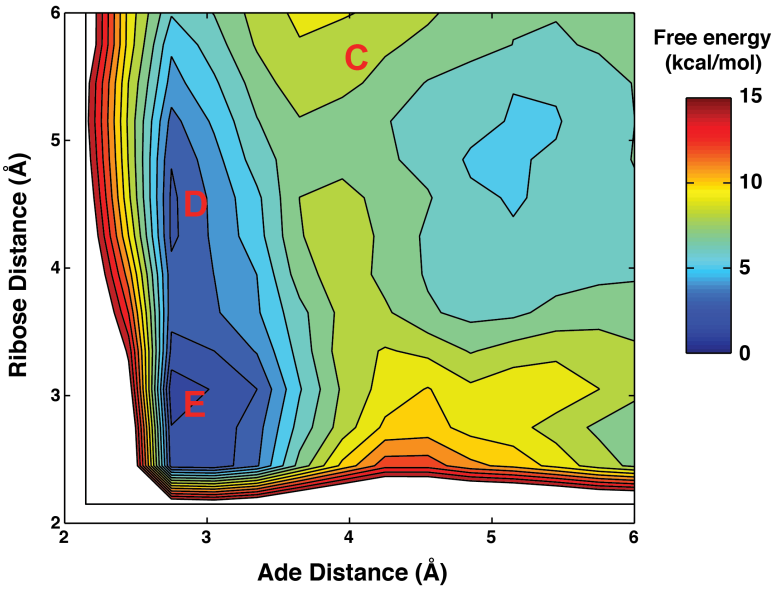

Fig. S13: **Free energy surface of adenine and ribose binding as a function of two distance CVs.** The contours were drawn every 1 kcal/mol. The important states for ATP binding (C, D and E) identified from the previous calculation (Figure 4) were labeled. In contrast to previous run, the metadynamics was initiated from the intermediate state C, rather than the bound state E.

Supporting tables

Table S1: **Summary of MD simulation systems.**

System name	Simulation method	Simulation time (ns)
Unbound	Unbiased	25
Bound	Unbiased	25
Base-stacked intermediate state	Unbiased	30
ATP binding I	Unbiased	10×4
ATP binding II	Unbiased	10×4
ATP binding III	Unbiased	10×4
ATP binding IV	Unbiased	10×4
Short ATP binding	Unbiased	0.01×1,000
Short GTP binding	Unbiased	0.01×1,000
ATP pulling in water	SMD	1×50
ATP pulling from unbound state	SMD	1×120
ATP binding using RMSD and distance as collective variable	Metadynamics	60
Nucleoside binding	Metadynamics	60+20
Total		570

Table S2: **Lists of atoms defining the center of ATP binding site in GluRS.**

Residue	**Atom**
His15	NE2
Leu236	N
Leu236	O
Lys243	NZ
Ser245	OG
Lys246	NZ
Arg247	CZ
Glu	OT1
Glu	OT2

Table S3: **Oligonucleotide primer sequences.**

Name	**Sequence 5′→ 3′(direction)**
H15Af	GGC GAC CCC GCC GTG GGC ACG
H15Ar	CGT GCC CAC GGC GGG GTC GCC
K246Af	ACC AAG ATC TCC GCG CGC AAA AGC CAC
K246Ar	GTG GCT TTT GCG CGC GGA GAT CTT GGT
R247Af	CAA GAT CTC CAA GGC CAA AAG CCA CAC C
R247Ab	GGT GTG GCT TTT GGC CTT GGA GAT CTT G