Confluence of Theory and Experiment Reveal the Catalytic Mechanism of the Varkud Satellite Ribozyme

Abir Ganguly,[*,†] Benjamin P. Weissman,[¶] Timothy J. Giese,[*,†] Nan-Sheng Li,[¥] Shuichi Hoshika,[§] Saieesh Rao,[¶] Steven A. Benner,[§] Joseph A. Piccirilli,[¶,¥] Darrin M. York[*,†,‡]

[*]*Laboratory for Biomolecular Simulation Research, Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8087, USA*
[†]*Institute for Quantitative Biomedicine, Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8087, USA*
[‡]*Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8087, USA*
[¶]*Department of Chemistry, The University of Chicago, Chicago, IL 60637, USA*
[¥]*Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637, USA*
[§]*Foundation for Applied Molecular Evolution, Firebird Biomolecular Sciences LLC, 13709 Progress Boulevard, Alachua, FL 32615, USA*

Correspondence: Darrin.York@rutgers.edu, jpicciri@uchicago.edu

Supplementary Information

Table of Contents

Computational Methods

Molecular Dynamics (MD) Simulations

MD simulations performed in this work were based on the available crystal structures of the VS ribozyme, specifically the A756G mutant (PDB ID: 4R4P),[1] G638A mutant (PDB ID: 4R4V),[1] and the most recent G638A mutant without the C634G mutation (PDB ID: 5V3I) [2]. In each case, the mutated catalytic nucleotides were replaced by their corresponding wildtype counterparts. Simulations were performed in truncated octahedral boxes containing TIP4P-Ew water molecules,[3] having a buffer distance of 12 Å with the solute. Monovalent ions (Na^+ and Cl^-) were added to neutralize the boxes and provide them with ionic strengths of 150 mM. Simulations were run with a time step of 1fs with full periodic boundary conditions. The smooth particle mesh Ewald (PME) method[4] with a 12Å cutoff was used to calculate long range electrostatic interactions and the SHAKE algorithm[5] was used to constrain all bonds involving hydrogen. Langevin dynamics with a collision frequency (γ) of 5.0 was employed to maintain constant temperature during simulations. For constant pressure and temperature simulations, the Berendsen barostat[6] with a pressure relaxation time of 1ps and isotropic pressure scaling was used along with Langevin dynamics (γ=5.0).

All simulations were performed on truncated models of the VS ribozyme that were constructed by systematically reducing the dimeric ribozymes under the constraint that the important base-pairings and tertiary interactions surrounding a single active site remain conserved (Figure S2). The putative general base G638 and general acid A756 were kept deprotonated and protonated, respectively, at their N1 positions. Parameters for nonstandard nucleobases, namely deprotonated G638 and protonated A756 were derived from Restrained Electrostatic Potential (RESP)[7,8] calculations and were adopted from a previous study.[9] The truncated models were validated by comparing several model simulations to simulations of the full dimeric ribozyme and ensuring that the model simulations reproduce the essential structural and dynamical features observed in the dimer simulations. Specifically, the root mean squared deviations (RMSDs) and root mean squared fluctuations (RMSFs) of the various residues calculated from the model simulations were found to be qualitatively similar to those calculated from the dimer simulations. These simulations will be discussed in more detail in follow up technical study, the manuscript of which is currently under preparation.

Before initiating production trajectories, the MD boxes were subjected to a series of rigorous equilibration steps that are summarized below:

1) During the first part of equilibration, the solute (RNA) is kept fixed to its crystal coordinates by imposing harmonic restraints with force constants of 100 kcal/mol-$Å^2$ on solute heavy atoms, and following steps were performed -
    a. Energy minimization of the solvent (water molecules + ions)
    b. Simulated annealing (NVT ensemble) of solvent according to
        i. Increase temperature from 0K to 298K in 298 ps, MD at 298 K for 500 ps
        ii. Increase temperature from 298K to 600K in 302 ps, MD at 600K for 500 ps
        iii. Decrease temperature from 600K to 298K in 302 ps, MD at 298K for 1500 ps

c. MD at 298K for 3000 ps (NPT ensemble) of solvent
d. Repeat of Step b
e. Repeat of Step c

2) The equilibrated solvent configuration is kept fixed and energy minimization of the solute is performed.
3) With the solvent free to move, the solute is relaxed in stages by slowly releasing the harmonic restraints on the solute heavy atoms according to –
   a. MD at 298K for 100 ps (NPT ensemble) with a force constant of 50 kcal/mol-Å^2
   b. MD at 298K for 100 ps (NPT ensemble) with a force constant of 25 kcal/mol-Å^2
   c. MD at 298K for 100 ps (NPT ensemble) with a force constant of 10 kcal/mol-Å^2
   d. MD at 298K for 100 ps (NPT ensemble) with a force constant of 5 kcal/mol-Å^2
   e. MD at 298K for 100 ps (NPT ensemble) with a force constant of 2 kcal/mol-Å^2
4) With both solute and solvent free to move, MD at 298K for 5000ps (NPT ensemble).

All simulations were performed with the AMBER18 software package,[10] using the pmemd.cuda program[11] and employing the Amber parm99 force field[12,13] with the *α/γ* corrections(parmbsc0)[14] and *χ* modifications(OL3)[15] for nucleic acids that has been shown to result in improved descriptions of nucleic acid dynamics.[16] Monovalent ion parameters optimized for the TIP4P-Ew water model were employed,[17,18] while Mg^{2+} ions were described using a recently developed m12-6-4 model,[19] a modified 12-6-4 potential model[20] that has been shown to produce accurate and balanced nucleic acid–Mg^{2+} interactions.

3D-RISM Calculations

The three-dimensional reference interaction site model (3D-RISM) has been shown to be a useful method for predicting solvation structure and thermodynamics of biomolecules.[21-24] 3D-RISM is a molecular solvation theory-based approach in which the solution particle density distributions, $\rho_\gamma(\boldsymbol{r})$, is expressed in terms of the total correlation functional, $h_\gamma(\boldsymbol{r}) = \frac{\rho_\gamma(\boldsymbol{r})}{\rho_\gamma^{bulk}} - 1$, and direct correlational functional, $c_\gamma(\boldsymbol{r})$, which are inter-dependent according to $h_\gamma(\boldsymbol{r}) = \sum_\alpha \int c_\alpha(\boldsymbol{r} - \boldsymbol{r}')\chi_{\alpha\gamma}(\boldsymbol{r}')d\boldsymbol{r}'$, where $\chi_{\alpha\gamma}(r)$ is the site-site solvent susceptibility of solvent sites α and γ and contains the bulk properties of the solvent.[25-27] $\chi_{\alpha\gamma}$ can be computed separately using dielectrically consistent RISM.[28,29] To arrive at a unique solution for h, a second equation, referred to as a closure relation, relating h and c is necessary, that has a major impact on the accuracy of the 3D-RISM results. Herein, we have employed the PSE2 closure which provides a good balance between speed and accuracy.[30]

The 3D-RISM method was employed to explore probable Mg^{2+} ion binding sites at the VS ribozyme active site. 3D-RISM calculations were performed on the truncated VS ribozyme model in several configurations, including the 4R4P, 4R4V, and 5V3I crystal

coordinates and average configurations obtained from MD trajectories. For the latter, the configurations were stripped of all solvent molecules and ions. The calculations were performed on the configurations using a 0.5 Å spaced 3D grid and a 1D solvent susceptibility value calculated separately for SPC/E water[31] (55.5 M) containing 150mM NaCl and 100mM $MgCl_2$. Figure 1D illustrates the results of a 3D-RISM calculation performed on a configuration obtained by averaging over all "Active" configurations reported in Figure 1G. All 3D-RISM and 1D-RISM calculations were performed using the AMBER18 software package.[10]

QM/MM free energy simulations

To compute the VS ribozyme catalytic reaction pathway, we first performed multi-dimensional QM/MM simulations using the finite temperature string (FTS) method [32,33] employing the semi-empirical Hamiltonian AM1/d-PhoT, [34] followed by *ab initio* QM/MM refinement. The semi-empirical simulations provide an overall reliable representation of the free energy surface from which a reasonable estimate for the 1D reaction pathway can be determined.[35] This 1D reaction path is then used as a departure point for path and energy refinement using *ab initio* QM/MM, which is more accurate but also more computationally intensive by a factor of roughly10^2-10^3. *Ab initio* QM/MM methods[36,37] have been used to examine phosphoryl transfer mechanisms in complex systems,[38] and when used together with the FTS method have been shown to be a powerful tool in determining favorable reaction pathways in multidimensional phase space.[39,40]

Semi-empirical (AM1d/Pho-T) Finite Temperature String (FTS) Simulations

The mathematical framework underlying the FTS method has been described in previous studies;[39,40] here, we summarize the technique qualitatively. In this method, a reaction pathway is approximated as a string (a curve) in the space of a handful of chosen coordinates that are relevant to the reaction. The string, initially along a guess reaction pathway, is discretized into a series of equally spaced images, and umbrella sampling simulations are performed at the image positions with the reaction coordinates harmonically restrained to the image centers. From these umbrella sampling simulations, average values of the reaction coordinates are calculated, which are then used to construct a new string. The new string is discretized again, and umbrella sampling simulations are performed along its length. This iterative process of updating the curve based on restrained simulations is continued until the curve is converged, that is the displacement of curves between successive iterations is negligible. The restrained simulations from all the iterations can then be jointly unbiased to obtain a multidimensional free energy surface, on which the converged string represents the minimum free energy path (MFEP) of the reaction.

The AM1d-PhoT FTS simulations were performed with sugar pucker corrections using three reaction coordinates: (1) the difference between G620:O2'-G620:HO2' distance and G638:N1-G620:HO2', that tracks the general base proton transfer; (2) the difference between G620:O2'-A621:P distance and A621:P-A621:O5' distance, that tracks the phosphorus oxygen bond formation/cleavage; and (3) the difference between A621:O5'-A756:N1 distance and A756:N1-A756:H1 distance, that tracks the general acid proton transfer. The QM region consisted of 123 atoms and is illustrated in Figure S8.

The initial pathway was generated by taking configurations from MD simulations, running short (~1-5 ps) QM/MM equilibration, and doing a linear interpolation along the three reaction coordinates described above. The string was divided into 50 images; at each image 0.3 ps of restrained dynamics were performed with harmonic force constant values of 100 kcal/mol/Å^2 before each string update. In total 50 string iterations were performed. To obtain the free energy surface, the restrained simulations at the various image positions from all the iterations are jointly unbiased using a multidimensional implementation of the weighted histogram analysis method (WHAM).[41]

The AM1/d-PhoT FTS simulations suggested a sequential catalytic pathway in which the 2'O nucleophile deprotonation by G638 occurs initially, as a separate step, followed by concerted nucleophilic attack and leaving group departure; the associated free energy profile along the pathway suggested the latter to be rate-limiting (Supplementary Figure 9). This result is consistent with Brønsted analysis of the G638 analogs, and together support a mechanism where the proton transfer from the 2'O-hydroxyl nucleophile occurs in a pre-equilibrium step. Therefore, we embarked on a more detailed investigation of the chemical step of the reaction starting from a configuration in which the 2'OH nucleophile has fully transferred its proton to G638:N1 using *ab initio* QM/MM simulations.

Ab initio 2D and 1D Umbrella Sampling (US) Simulations.

In order to validate the transition state predicted by the converged AM1/d-PhoT FTS reaction pathway, and verify that this pathway is the only significant minimum free energy pathway in the space of the important reaction coordinates, we calculated the free energy surface underlying the reaction using extensive *ab initio* 2D umbrella sampling simulations in the space of the phosphorus bond breaking/forming and the general acid proton transfer coordinates described earlier. The 2D QM/MM simulations consisted of 12 windows along each coordinate. At each window, 1-2 ps of restrained dynamics were performed with force constant values of 100 kcal/mol/Å^2. Simulation data from all windows were unbiased using the 2D variational free energy perturbation (vFEP) method developed in our laboratory.[42,43]

Further, in order to obtain better resolution of free energy along the converged AM1/d-PhoT FTS reaction pathway, we performed additional *ab initio* 1D umbrella sampling simulations along the pathway. The converged pathway was divided into 30 windows, and 1-3 ps of US simulations were performed at each window with harmonic force constant values of 100 kcal/mol/ Å^2. The 1D free energy profiles along the pathways were calculated by unbiasing these restrained simulations using an in-house implementation of multistate Bennett acceptance ratio (MBAR) technique.[44]

It should be noted that efforts to examine the effect of Mg^{2+} in providing direct electrostatic stabilization of the transition state (e.g., by replacing Mg^{2+} with Na^{+} in the active state) were complicated by the critical role played by the divalent ion in organizing the active site, and in particular maintaining the position of the general base as discussed previously, as well as stabilizing the electrostatic repulsions between the A621 and A622 phosphates. Further theoretical work is needed in order to decompose these contributions in such a way as to provide meaningful insight specifically into the role of Mg^{2+} in the chemical steps of the reaction.

The *ab initio* QM/MM simulations were run using the PBE0/6-31G** Hamiltonian with a time step of 1fs with full periodic boundary conditions. Long-range electrostatics were treated with the Ambient-Potential Composite Ewald (CEw) method.[45] A 12Å real-space cutoff was used, and the reciprocal-space energy was solved using a 1pt/Å uniform fast Fourier transform grid. The wavefunction was converged to within a commutator error of 10^{-7} au at each step of dynamics. The SHAKE algorithm[5] was used to constrain all bonds involving hydrogen that are part of the MM region and not in the QM region. The simulations were run in the NPT ensemble using Langevin dynamics along with the Berendsen barostat[6] to maintain constant temperature and pressure. The simulations are performed using a code that is implemented within a developmental version of AmberTools 15 and interfaced to the Sander MD program.[46]

Experimental methods

RNA oligonucleotides and constructs

The VS enzyme strand was prepared by ligation of two RNA oligomers: one generated from solid-phase synthesis and the other by T7 transcription. The 115-nt 5' fragment (sequence: 5'- GGC GGU AGU AAG CAG GGA ACU CAC CUC CAA UUU CAG UAC UGA AAU UGU CGU AGC AGU UGA CUA CUG UUA UGU GAU UGG UAG AGG CUA AGU GAC GGU AUU GGC GUA AGU CAG UAU U) was generated by T7 transcription under standard conditions from a double-stranded DNA template, which was prepared as follows. A pUC19 plasmid containing 3' ligation fragment was transformed into competent DH5α *Escherichia coli* cells and amplified, and the sequence was confirmed by sequencing, (The University of Chicago Core Facility). The plasmid was digested by EarI (New England Biolabs) to generate a double stranded DNA template that contains the T7 promotor sequence and a blunt end at the 3'-end of the ribozyme sequence. Following transcription, transcripts were purified first by phenol-chloroform-isoamyl alcohol extraction then by 6% denaturing polyacrylamide gel electrophoresis (dPAGE). The transcripts were visualized under UV light, excised from the gel and eluted overnight in TEN buffer (10mM Tris pH = 8, 1mM EDTA, 300mM NaCl), then concentrated on centrifugal filter units for use. The 30-nt 3' ligation fragments (5'- GCA GAC CAG CAC AAG CCC GCU UGC GAG AAU – where the underlined A corresponds to A756) were synthesized in-house by solid-phase synthesis on a 1-µmol scale using an Expedite Nucleic Acid Synthesis System (8900) by following standard RNA synthesis protocols with either commercially available phosphoramidites (ChemGenes and Glen Research) or phosphoramidites synthesized in-house (see synthetic methods). The oligonucleotides were released from solid support with 3:1 NH_4OH/EtOH at 55 °C for 8 h, desilylated with 300 µL 6:3:4 N-methylpyrrolidinone/triethylamine/triethylamine-3HF at 65 °C for 2 h and precipitated by n-BuOH. The oligomers were further purified by dPAGE, collected in pure water and stored at −80 °C until further use.

The 3' ligation fragment was 5′-end phosphorylated with T4 polynucleotide kinase (New England Biolabs, NEB) in the presence of excess ATP at 37 °C for about 1 h followed by heat inactivation of the enzyme at 60 °C for 20 min. For the ligation reaction, 3' and 5' ligation fragments (1.5:1 molar ratio) were incubated in the presence of annealing buffer (20mM Tris pH = 7.5, 200nM NaCl, 0.2mM EDTA) at 90ºC for 2 minutes and slowly cooled to 10º over 1.5 hours. After annealing, T4 RNA ligase buffer (50 mM Tris pH 7.5, 2 mM $MgCl_2$, 1 mM DTT, 400 µM ATP) and T4 RNA ligase 2 (NEB) were added, and the resulting reaction mixture was incubated at 16ºC for 18 to 24 hours. Ligation reaction mixtures were purified by 10% dPAGE as described above and concentrations for pure oligos were determined using Nanodrop (Thermo Scientific).

The 24-nt VS substrates (5'- GCG CG*A AGG GCG UCG UCG GCC CGA, where the underlined G denotes G623 and the asterisk denotes the scissile phosphate, G620) strands were synthesized by solid-phase synthesis following the same procedures for described above. Substrates containing phosphorothioate mutations were further separated by 4.6 x 250mm C18 reversed-phase high-performance liquid chromatography (HPLC) through an AcclaimTM 120 4.6 x 250mm C18 column (Dionex) with a 2795 Separations module (Waters). After a 10 minutes column equilibration step, a linear gradient from 9 to 10.6% acetonitrile over 35 minutes in a constant background of 0.1 M

TEAA. Typically, this procedure resulted in clean separation of diastereomers with elution times around 15 and 19 minutes after the initiation of the gradient. The identities of the stereoisomers were identified by digestion with snake venom phosphodiesterase and nuclease P1 revealing that the first peak corresponded to S_p thio isomer while the second peak corresponded to the R_p thio isomer. Following HPLC purification, separated diastereomers were dried overnight under vacuum, resuspended in 300µL ddH_2O, ethanol precipitated, collected in pure water and stored at -80 °C until further use. Substrate strands were all 5′-end radiolabeled with T4 polynucleotide kinase (New England BioLabs) with [γ-^{32}P]ATP (PerkinElmer) prior to their use is kinetic reactions.

Ribozyme kinetic assays and data analysis

"Standard" conditions for ribozyme kinetic reactions are as follows: ribozyme RNA was incubated in water at 70º for 2 minutes to denature and allowed to renature at room temperature for 10 minutes in the presence of reaction buffer (25mM TAPS, 25mM HEPES, 10mM $MgCl_2$, 25mM KCl, 2mM spermidine, pH = 8). The reaction was then initiated by the addition of a trace amount of radiolabeled substrate RNA to a final concentration of 500nM ribozyme RNA (control reactions with four-fold higher ribozyme concentrations show a concentration of 500nM ribozyme to be near saturating) in a volume of 10µL. For pH dependence, appropriately pH-adjusted buffers containing either MES and HEPES (pH 5.5-8) or TAPS and HEPES (pH 7.5-9.5) were used. Aliquots of the reaction mixture were taken at appropriate time points and quenched in stop solution (97% formamide, 15mM EDTA, 0.1% w/v bromophenol blue and xylene cyanol) and rapidly chilled on dry ice. Cleavage products were separated from uncleaved substrate on a denaturing 20% polyacrylamide/7M urea gel and quantified using a Typhoon PhosphorImager with ImageQuant software (Molecular Dynamics). All rate constants presented are the average of at least three independent time courses conducted over at least two separate days.

Data for cleavage reactions were fit in KaleidaGraph (Synergy Software) to equation 1 where *f* is the fraction cleaved, A is the amplitude or endpoint of the reaction (typically ~70% for HDV cleavage reactions), k_{obs} is the observed rate constant and *t* is the time.

$$f = A(1- e^{- k_{obs} t}) \quad (1)$$

pH rate profiles were fit to a model for double ionization (equation 2) assuming a requirement for one protonated and one deprotonated group, where pK_{a1} corresponds to the rate stimulating titration, pK_{a2} corresponding to the rate-inhibiting titration and k_{cat} is the intrinsic rate constant or the reaction. Both pK_as and k_{cat} were fitted for pH rate profiles of the wild type ribozyme and ribozymes with A756 analogs, but pK_{a1} was fixed at 5.8 for fitting of G analog pH rate profiles.

$$k_{obs} = \frac{k_{cat}}{1+10^{pH- pK_{a1}} +10^{pK_{a2}- pH} +10^{pK_{a2}- pK_{a1}}} \quad (2)$$

The pH dependence for cleavage of the G638P* substrate was fit to the model for a single ionization in equation 3.

$$k_{obs} = \frac{k_{cat}}{10^{pH- pK_{a1}}} \quad (3)$$

Assuming that VS ribozyme reacts through a one-channel mechanism that depended on protonated A756 and deprotonated G638 residues, the intrinsic rate constant for cleavage can be modeled by equation 4 where α and β are the Brønsted coefficients for A756 and G638, respectively, and C is a constant for the reaction.

$$k_{cat} = b(pK_{a,G638}) - a(pK_{a,A756}) + C \quad (4)$$

The Brønsted coefficients describe the extent of proton transfer in the transition state with possible values between 0 and 1. At one extreme, a value of zero for either coefficient would signify no proton transfer from the corresponding nucleobase and alteration of that nucleobase should have no impact on k_{cat}. At the other extreme, a value of one would signify complete proton transfer in the transition state. To determine the Brønsted coefficients for G638 and A756 in VS catalysis, we plotted the k_{cat} and pK_a inferred from the pH-rate dependence profiles described above and fit the resulting data linearly in KaleidaGraph (Synergy Software)

Thio effects and mutational and metal-ion rescue

The effect of substituting a sulfur at a non-bridging position of a phosphate is referred to as a thio effect (equation 5) where k_O is the rate constant for cleavage of the unsubstituted substrate and k_S is the rate constant for cleavage of the thio substrate.

$$\text{Thio effect} = \frac{k_O}{k_S} \quad (5)$$

When thio substitution reduces the observed rate constant for cleavage, the thio effect is larger than one and referred normal. For reactions where thio substation increases the observed rate constant, the thio effect will be less than one and referred as inverse.

To test whether thio effects arise from a loss of interaction with a specific chemical group, we observed thio effects in the presence and absence of that group. For instance, to test whether any of the observed thio effects arise from a loss of interaction with G638:N2, we test thio effects in the context of wild type and G638I substrates that lacked G638:N2. The mutational rescue can be quantified in equation 6 where numerator is the thio effect in the wild type background and the denominator is the thio effect in the mutated background.

$$\text{Mutational Rescue} = \frac{\left(\frac{k_O}{k_s}\right)^{WT}}{\left(\frac{k_O}{k_s}\right)^{Mut}} \quad (6)$$

A mutational rescue of one indicates that the mutation has no impact on the thio effect, but when the a thio effect is reduced in the mutant background the mutational rescue will be larger than one indicative of a functional linkage between the mutation and the thio substitution.

If thio effects arise due to a loss of an interaction with a divalent metal ion, then addition of a thiophilic metal ion should “rescue” the lost interaction and ameliorate the thio effect [47]. Thus, we observed thio effect under standard conditions (10mM $MgCl_2$) and

in the presence of a thiophilic metal ion (10mM $MgCl_2$ and 20μM $CdCl_2$). Metal ion rescue effect can be quantified following equation 7 where the numerator if the thio effect under standard conditions and the denominator is the thio effect in the presence of thiophilic ions.

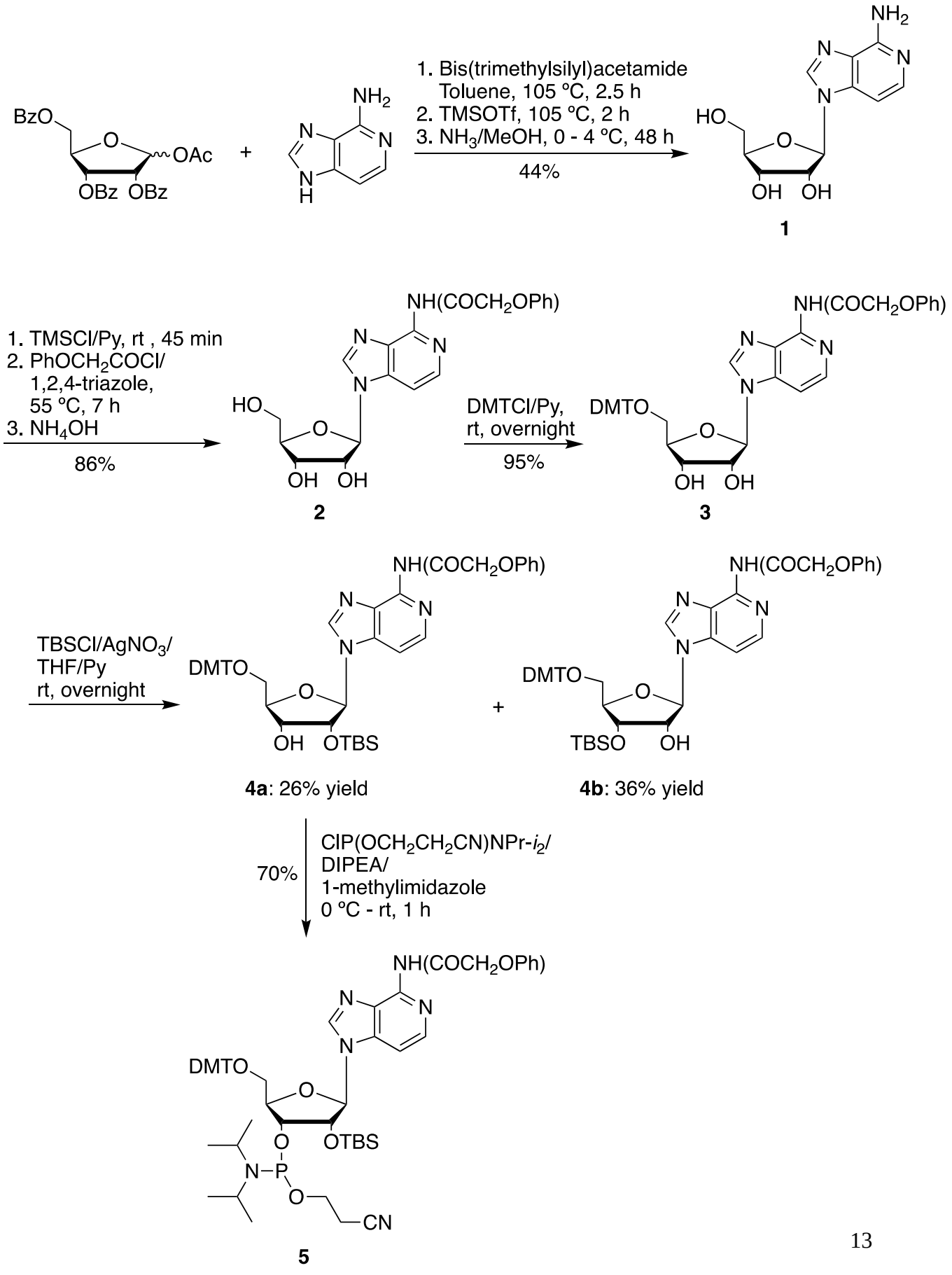
$$\text{Metal ion Rescue} = \frac{\left(\frac{k_O}{k_s}\right)^{Mg^{2+}}}{\left(\frac{k_O}{k_s}\right)^{Cd^{2+}}} \quad (7)$$

A value greater one will be observed when the thio effect is reduced (or inverted) in the presence of thiophilic metal ions, thereby suggesting an interaction between the thio-substituted atom and the thiophilic ion. We conducted metal ion rescue under the (10mM $MgCl_2$ and 20μM $CdCl_2$) at that concentration of $CdCl_2$ has essentially no impact on the rate of unsubstituted substrates, but fully stimulates thio substituted substrates.

Synthetic Methods

3-Deazaadenosine (1): Following a synthesis adapted from Micura and co-workers [48], *N*,*O*-bis(trimethylsilyl)acetamide (3.20 mL, 13.1 mmol) was added to a mixture of 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-D-ribofuranose (1.88 g, 3.73 mmol) and 3-deazaadenine (500 mg, 3.73 mmol) in dry toluene (80 mL) under an argon atmosphere. The suspension was heated to 105ºC and stirred for 2.5 hours and then cooled to room temperature. Trimethylsilyl trifluoromethanesulfonate (2.03 mL, 11.2 mmol) was added and the solution was stirred at 105 °C for an additional 2 hours. Following cooling of the reaction mixture and evaporation of the solvent, the residue was diluted in dichloromethane and washed with saturated $NaHCO_3$ solution. The combined organic layers were dried over $MgSO_4$. After the solution was filtered and evaporated, the residue was purified by flash silica gel chromatography eluting with 1% methanol in dichloromethane containing 1% trimethylamine to give the glycosylation product as beige foam. The glycosylation product was dissolved into methanol (20 mL) and the solution was saturated with ammonia at 0 ºC for 30 minutes and stored at 4 ºC for 2 days. The solvent was removed and the residue was isolated by silica gel chromatography eluting with 25% methanol in dichloromethane to give **1** as a white foam [48-51] (440 mg, 44% yield). ^{1}H NMR (DMSO-d_6) δ 8.40 (s, 1H), 7.67 (d, 1H, *J* = 6.0 Hz), 7.05 (d, 1H, *J* = 6.0 Hz), 6.78 (brs, 2H), 5.79 (d, 1H, *J* = 6.4 Hz), 5.55 (d, 1H), 5.28 (d, 1H), 5.18 (t, 1H), 4.30 (m, 1H), 4.10 (m, 1H), 3.97 (m, 1H), 3.63 (m, 2H); ^{13}C NMR (DMSO-d_6) δ 151.5, 141.0, 138.1, 137.8, 126.8, 98.1, 88.9, 85.9, 74.3, 70.2, 61.3.

***N*^{6}-Phenoxyacetyl-3-deazaadenosine (2):** TMSCl (3.30 mL, 26.4 mmol) was added dropwise to the suspension of **1** (914 mg, 3.52 mmol) in anhydrous pyridine (20 mL), the reaction mixture was stirred at room temperature for 45 minutes. Meanwhile phenoxyacetyl chloride (0.73 mL, 5.3 mmol) was added dropwise to 1,2,4-triazole (0.37 g, 5.3 mmol) in acetonitrile (7.0 mL) and pyridine (7.0 mL). The acylating reagent, formed *in situ*, was transferred by cannula needle to the flak containing transiently protected nucleoside. The reaction mixture was heated up and stirred at 55 ºC for 7 hours. After it was cooled down to room temperature, water (3.0 mL) was added, followed by 28% ammonium hydroxide (1.7 mL). The solution was concentrated to remove pyridine, redissolved into water and extracted with dichloromethane. The solid was precipitated out from aqueous layer, filtered, washed with dichloromethane and dried over vacuum to give **2** (1.218 g, 86% yield). ^{1}H NMR (DMSO-d_6) δ 10.37 (brs, 1H), 8.59 (s, 1H), 8.13 (d, 1H, *J* = 5.6 Hz), 7.70 (d, 1H, *J* = 6.0 Hz), 7.30 (m, 2H), 7.05-6.90 (m, 3H), 5.91 (d, 1H, *J* = 6.0 Hz), 5.60 (d, 1H, *J* = 6.4 Hz), 5.30 (d, 1H, *J* = 4.4 Hz), 5.22 (t, 1H, *J* = 5.0 Hz), 4.90 (s, 2H), 4.37 (m, 1H), 4.13 (m, 1H), 4.01 (m, 1H), 3.66 (m, 2H); ^{13}C NMR (DMSO-d_6) δ 167.0, 157.9, 143.0, 140.0, 139.1, 132.8, 129.5, 121.2, 114.7, 114.4, 105.4, 88.9, 86.0, 74.0, 70.2, 67.1, 61.2; HRMS calcd for $C_{19}H_{21}N_4O_6$ [MH^+]: 401.1461, found: 401.1461.


1. Bis(trimethylsilyl)acetamide
Toluene, 105 °C, 2.5 h
2. TMSOTf, 105 °C, 2 h
3. NH3/MeOH, 0 - 4 °C, 48 h
44%
1
1. TMSCl/Py, rt , 45 min
2. PhOCH2COCl/
1,2,4-triazole,
55 °C, 7 h
3. NH4OH
86%
2
DMTCl/Py,
rt, overnight
95%
3
TBSCl/AgNO3/
THF/Py
rt, overnight
4a: 26% yield
4b: 36% yield
70%
ClP(OCH2CH2CN)NPr-i2/
DIPEA/
1-methylimidazole
0 °C - rt, 1 h
5

Supplementary Scheme 1. Synthesis of 3-deazaadenosine phosphoramiditie

5´-*O*-DMT-*N*^{6}-phenoxyacetyl-3-deazaadenosine (3): Under an argon atmosphere DMTCl (400 mg, 1.18 mmol) and DMAP (20 mg, 0.16 mmol) were added into the solution of **2** (236 mg, 0.59 mmol) in pyridine (5 mL). After the mixture was stirred at room temperature overnight, the reaction was quenched with methanol (1 mL). The solvent was removed under vacuum, then residue was redissolved into dichloromethane, washed with saturated aqueous $NaHCO_3$ and brine and dried over $MgSO_4$. The solvent was again removed, and the residue was isolated by silica gel chromatography, eluting with 5% methanol in dichloromethane to give **5** (394 mg, 95% yield) as a pale-yellow foam. ^{1}H NMR ($CDCl_3$/TMS) δ 9.47 (brs, 1H), 8.04 (s, 1H), 7.58 (d, 1H, *J* = 5.6 Hz), 7.39 (d, 2H, *J* = 7.2 Hz), 7.35-7.10 (m, 10H), 7.05-6.95 (m, 3H), 6.76 (m, 4H), 5.84 (d, 1H, *J* = 6.0 Hz), 4.70-4.55 (m, 3H), 4.50 (m, 1H), 4.30 (m, 1H), 3.70 (s, 3H), 3.69 (s, 3H), 3.44 (m, 2H); ^{13}C NMR ($CDCl_3$) δ 158.7, 158.6, 157.1, 144.5, 142.2, 138.8, 135.6, 135.5, 130.24, 130.20, 129.9, 128.3, 128.0, 127.1, 122.4, 115.1, 113.3, 105.0, 89.9, 86.8, 84.7, 74.5, 71.2, 67.9, 63.6, 55.3; HRMS calcd for $C_{40}H_{39}N_4O_8$ [MH^+]: 703.2768, found: 703.2769.

5´-*O*-DMT-2´-*O*-TBS-*N*^{6}-phenoxyacetyl-3-deaza-adenosine (4a) and 5´-*O*-DMT-2´-*O*-TBS-*N*^{6}-phenoxyacetyl-3-deazaadenosine (4b): Under an argon atmosphere to the solution of **5** (433 mg, 0.616 mmol) in dry THF (10 mL), pyridine (0.25 mL, 3.1 mmol), silver nitrate (209 mg, 1.23 mmol) and TBSCl (139 mg, 0.92 mmol) were added. The reaction mixture was stirred at room temperature overnight. The next morning, the solid was filtered off, the solution was evaporated and the residue was redissolved into dichloromethane, washed with water, saturated $NaHCO_3$, brine and dried over $MgSO_4$. The solvent was again removed, and the residue was isolated by silica gel chromatography, eluting with 2% methanol in dichloromethane to give the up spot on TLC, **4a** (131 mg, 26% yield) and the low spot on TLC, **4b** (0.183 g, 22% yield) as white films.

4a: ^{1}H NMR ($CDCl_3$/TMS) δ 9.55 (brs, 1H), 8.10 (s, 1H), 7.92 (d, 1H, *J* = 5.6 Hz), 7.50-7.20 (m, 12H), 7.15-7.00 (m, 3H), 6.85-6.72 (m, 4H), 5.83 (d, 1H, *J* = 7.2 Hz), 4.81 (s, 2H), 4.75 (dd, 1H, *J* = 5.2, 1.6 Hz), 4.33 (m, 1H), 3.78 (s, 3H), 3.77 (s, 3H), 3.60-3.45 (m, 2H), 0.82 (s, 9H), -0.11 (s, 3H), -0.35 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 158.8, 157.3, 144.3, 143.5, 141.8, 141.3, 138.7, 135.2, 135.1, 130.3, 129.9, 128.3, 128.2, 127.3, 122.3, 115.1, 113.4, 104.9, 89.2, 87.3, 84.7, 75.7, 71.7, 68.1, 63.5, 55.4, 25.6, 17.9, -5.20, -5.22; HRMS calcd for $C_{46}H_{53}N_4O_8Si$ [MH^+]: 817.3633, found: 817.3635.

4b: ^{1}H NMR ($CDCl_3$/TMS) δ 8.21 (s, 1H), 8.06 (d, 1H, *J* = 5.6 Hz), 7.46 (d, 1H, *J* = 5.6 Hz), 7.43-6.90 (m, 14H), 6.85-6.75 (m, 4H), 5.83 (d, 1H, *J* = 4.8 Hz), 4.81 (s, 2H), 4.45-4.40 (m, 2H), 4.18 (m, 1H), 3.78 (s, 3H), 3.77 (s, 3H), 3.55 (m, 1H), 3.36 (m, 1H); ^{13}C NMR ($CDCl_3$) δ 158.8, 157.4, 144.1, 143.2, 141.4, 141.1, 138.8, 135.3, 135.2, 130.2, 129.8, 129.2, 128.2, 128.1, 122.2, 115.1, 113.4, 104.9, 90.9, 87.0, 84.9, 74.9, 71.7, 68.1, 62.5, 55.3, 25.8, 18.1, -4.53, -4.78; HRMS calcd for $C_{46}H_{53}N_4O_8Si$ [MH^+]: 817.3633, found: 817.3634.

5´-*O*-DMT-2´-*O*-TBS-*N*^{6}-phenoxyacetyl-3-deazaadenosine 3´-*N,N*-Diisopropyl(cyanoethyl)phosphoramidite (5): To the solution of **4a** (116 mg, 0.14

mmol) and *i*-Pr_2NEt (122 μL, 0.70 mmol) in anhydrous dichloromethane (5.0 mL) at 0 ºC, $ClP(NPr\text{-}i_2)OCH_2CH_2CN$ (95 μL, 0.42 mmol) was added, followed by the addition of 1-methylimidazole (6.0 μL, 0.07 mmol). The reaction mixture was stirred at room temperature for 1 hour, then quenched with methanol (1.0 mL). The solvent was removed under vacuum and the residue was purified by silica gel chromatography, eluting with 0.5% Et_3N in CH_2Cl_2 to give **5** as a yellow foam (100 mg, 70% yield). ^{31}P NMR (CD_3CN) δ 151.6, 149.1 ppm; HRMS calcd for $C_{55}H_{71}N_6O_9PSi$ [MH^+]: 1018.4789, found: 1018.4778.

3-Deazapurine (6): Under an argon atmosphere, to the mixture of 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-D-ribofuranose (4.24 g, 8.40 mmol) and 3-deazapurine (1000 mg, 8.40 mmol) in dry toluene (100 mL), *N*,*O*-bis(trimethylsilyl)acetamide (7.20 mL, 29.4 mmol) was added. The suspension was heated up to 105 °C, stirred for 1.5 hours, then cooled to room temperature before the addition of trimethylsilyl trifluoromethanesulfonate (4.56 mL, 25.2 mmol). The reaction mixture was then stirred at 105 °C overnight. The next morning, the mixture was cooled to room temperature and the solvent was removed under vacuum, then the residue was diluted in dichloromethane and washed with saturated $NaHCO_3$ solution. The combined organic layers were dried over $MgSO_4$. After the solution was filtered and evaporated, the residue was purified by flash silica gel chromatography eluting with 4-8% methanol in dichloromethane to give the glycosylation product. The glycosylation product was dissolved into methanol (20 mL) and the solution was saturated with ammonia at 0 ºC for 30 minutes and stored at 4 ºC for 24 hours. The solvent was removed and the reside was isolated by silica gel chromatography eluting with 15% methanol in ethyl acetate to give **6** [52] as a white foam (312 mg, 15% yield). ^1H NMR (CD_3OD) δ 8.89 (brs, 1H), 8.63 (s, 1H), 8.31 (d, 1H, *J* = 5.6 Hz), 7.84 (d, 1H, *J* = 5.6 Hz), 5.99 (d, 1H, *J* = 6.0 Hz), 4.45 (t, 1H, *J* = 5.6 Hz), 4.28 (m, 1H), 4.14 (dd, 1H, *J* = 6.0, 3.2 Hz), 3.85 (dd, 1H, *J* = 12.2, 3.0 Hz), 3.78 (dd, 1H, *J* = 12.2, 3.4 Hz); ^{13}C NMR (CD_3OD) δ 145.5, 142.6, 142.5, 141.9, 139.6, 108.9, 91.0, 87.6, 76.2, 71.9, 62.6; HRMS calcd for $C_{11}H_{14}N_3O_4$ [MH^+]: 252.0984, found: 252.0982.

5´-*O*-DMT-3-deazapurine (7): Under an argon atmosphere DMTCl (756 mg, 2.23 mmol) was added to a solution of **6** (312 mg, 1.24 mmol) in pyridine (5 mL). After the mixture was stirred at room temperature overnight, the solvent was removed, the residue was dissolved in dichloromethane, washed with saturated aqueous $NaHCO_3$, brine and dried over $MgSO_4$. The solvent was removed, and the residue was isolated by silica gel chromatography, eluting with 5% methanol in dichloromethane to give **7** (322 mg, 47% yield) as a pale yellow foam. ^1H NMR ($CDCl_3$/TMS) δ 8.60 (s, 1H), 8.10 (s, 1H), 7.67 (d, 1H, *J* = 5.6 Hz), 7.40 (d, 2H, *J* = 7.6 Hz), 7.35 (d, 1H, *J* = 5.6 Hz), 7.30 (d, 4H, *J* = 8.6 Hz), 7.21 (m, 2H), 7.14 (t, 1H, *J* = 7.2 Hz), 6.80-6.70 (m, 4H), 5.86 (d, 1H, *J* = 6.4 Hz), 4.59 (m, 1H), 4.51 (dd, 1H, *J* = 5.2, 2.4 Hz), 4.35 (m, 1H), 3.69 (s, 6H), 3.48 (m, 2H); ^{13}C NMR ($CDCl_3$) δ 158.7, 144.4, 143.6, 141.7, 141.1, 140.9, 137.6, 135.5, 135.4, 130.22, 130.20, 128.2, 128.1, 127.1, 113.3, 107.7, 90.1, 86.9, 84.8, 73.9, 71.2, 63.7, 55.3; HRMS calcd for $C_{32}H_{32}N_3O_6$ [MH^+]: 554.2291, found: 554.2292.

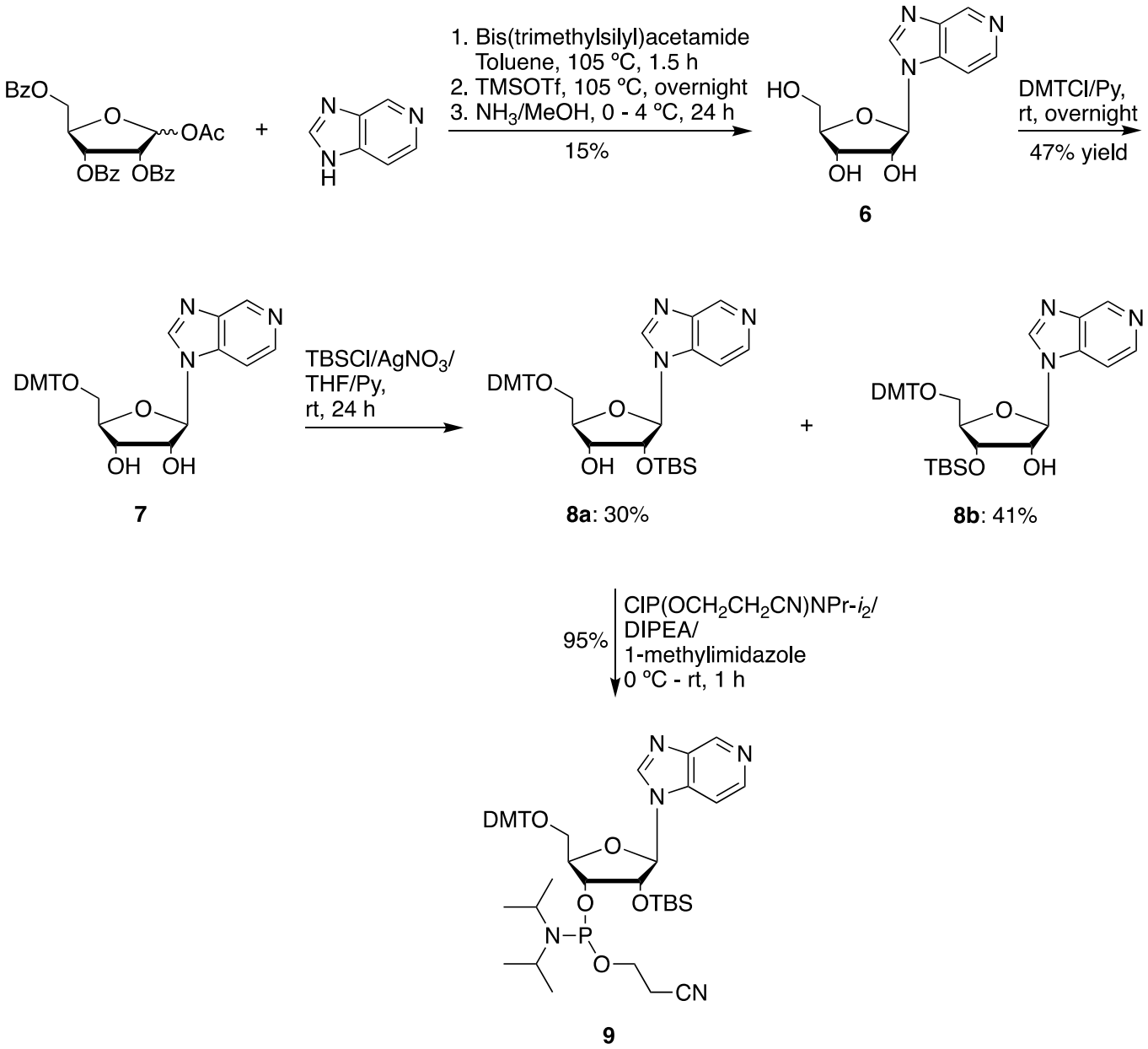
5´-*O*-DMT-2´-*O*-TBS-3-deazapurine (8a) and 5´-*O*-DMT-3´-*O*-TBS-3-deazapurine (8b): Compound **7** (0.322 g, 0.582 mmol) was coevaporated with dry pyridine (4 mL), and then

dried over vacuum for 30 minutes and redissolved in dry THF (10 mL). To that solution, pyridine (0.24 mL, 2.91 mmol), silver nitrate (197 mg, 1.16 mmol) and TBSCl (106 mg, 0.70 mmol) were added and the mixture was stirred at room temperature overnight. The next morning, TLC showed unreacted **7**, so additional TBSCl (26 mg, 0.17 mmol) was added and the mixture was stirred at room temperature for an additional 2 hours. The solid was filtered off, the filtrate was diluted with dichloromethane, washed with saturated $NaHCO_3$, brine and dried over $MgSO_4$. The solvent was removed, and the residue was isolated by silica gel chromatography, eluting with 2 - 3% methanol in dichloromethane to give the up spot on TLC, **8a** (117 mg, 30% yield) and the low spot on TLC, **8b** (161 mg, 41% yield) as white foams.

8a: ^{1}H NMR ($CDCl_3$/TMS) δ 9.11 (s, 1H), 8.14 (s, 1H), 8.12 (d, 1H, *J* = 5.0 Hz), 7.58 (d, 1H, *J* = 5.5 Hz), 7.43 (d, 2H, *J* = 7.0 Hz), 7.38-7.20 (m, 7H), 6.90-6.75 (m, 4H), 5.86 (d, 1H, *J* = 7.5 Hz), 4.76 (dd, 1H, *J* = 7.0, 5.0 Hz), 4.38 (m, 1H), 3.79 (s, 3H), 3.78 (s, 3H), 3.56 (dd, 1H, J = 10.8, 2.5 Hz), 3.50 (dd, 1H, J = 10.8, 3.0 Hz), 0.81 (s, 9H), -0.12 (s, 3H), -0.38 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 158.85, 158.83, 144.4, 144.4, 143.7, 142.9, 142.6, 137.7, 135.3, 135.2, 130.30, 130.27, 128.25, 128.15, 127.3, 113.4, 107.1, 89.0, 87.2, 84.7, 75.8, 71.8, 63.6, 55.4, 25.6, 17.9, -5.24, -5.30; HRMS calcd for $C_{38}H_{46}N_3O_6Si$ [MH^+]: 668.3156, found: 668.3157.

8b: ^{1}H NMR ($CDCl_3$/TMS) δ 9.12 (s, 1H), 8.25 (d, 1H, *J* = 5.6 Hz), 8.23 (s, 1H), 7.62 (d, 1H, *J* = 5.6 Hz), 7.50-7.20 (m, 9H), 6.90-6.75 (m, 4H), 5.85 (d, 1H, *J* = 5.2 Hz), 4.50-4.40 (m, 2H), 4.19 (dd, 1H, *J* = 6.0, 3.2 Hz), 3.78 (s, 3H), 3.77 (s, 3H), 3.55 (dd, 1H, J = 10.8, 2.8 Hz), 3.35 (dd, 1H, J = 10.8, 3.6 Hz), 0.90 (s, 9H), 0.087 (s, 3H), 0.007 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 158.8, 144.2, 143.5, 142.5, 142.4, 141.5, 137.7, 135.33, 135.28, 130.18, 130.15, 128.2, 128.1, 127.3, 113.4, 107.0, 90.7, 87.0, 84.9, 74.8, 71.9, 62.7, 55.4, 25.8, 18.1, -4.50, -4.76; HRMS calcd for $C_{38}H_{46}N_3O_6Si$ [MH^+]: 668.3156, found: 668.3158.

5´-*O*-DMT-2´-*O*-TBS-3-deazapurine 3´-*N,N*-Diisopropyl(cyanoethyl)phosphoramidite (9): To the solution of **8a** (59 mg, 0.088 mmol) and *i*-Pr_2NEt (150 μL, 0.88 mmol) in anhydrous dichloromethane (5.0 mL) at 0 ºC, ClP(NPr-*i*$_2$)OCH_2CH_2CN (59 μL, 0.264 mmol) was added, followed by the addition of 1-methylimidazole (3.5 μL, 0.044 mmol). The reaction mixture was stirred at room temperature for 1 hour then quenched with methanol (1.0 mL). The solvent was removed, the residue was purified by silica gel chromatography, eluting with 0.5% Et_3N in CH_2Cl_2 to give **9** as a white foam (73 mg, 95% yield). ^{31}P NMR (CD_3CN) δ 151.6, 149.0 ppm; HRMS calcd for $C_{47}H_{63}N_5O_7PSi$ [MH^+]: 868.4234, found: 868.4230.



Supplementary Scheme 2. Synthesis of 3-deazapurine phoshphoramidite

2´,3´,5´-Tri-*O*-acetyl-8-azaquanosine (10): Following a procedure adapted from Rosemeyer and collogues [53], $SnCl_4$ (2.58 mL, 22.1 mmol) was added to a suspension of 8-azaadenine (1.00 g, 7.35 mmol) and 1,2,3,5-tetra-*O*-acetyl-β-D-ribofuranose (2.34 g, 7.35 mmol) in dry acetonitrile (30 mL) under an argon atmosphere and stirred at room temperature for 24 hours. The reaction mixture was poured into saturated $NaHCO_3$ (100 mL). The precipitate was filtered off, and washed with water (2 × 30 mL). The filtrate was extracted with dichloromethane. The organic layers were combined and dried over $MgSO_4$. The solvent was removed, and the residue was isolated by silica gel chromatography, eluting with 5% methanol in dichloromethane to collect the faster migrating spot on TLC, **10** (808 mg, 28% yield) as a white foam. ^{1}H NMR ($CDCl_3$/TMS) δ 8.41 (s, 1H), 7.30 (s, 2H), 6.56 (d, 1H, *J* = 4.0 Hz), 6.18 (m, 1H), 5.86 (m, 1H), 4.52-4.45

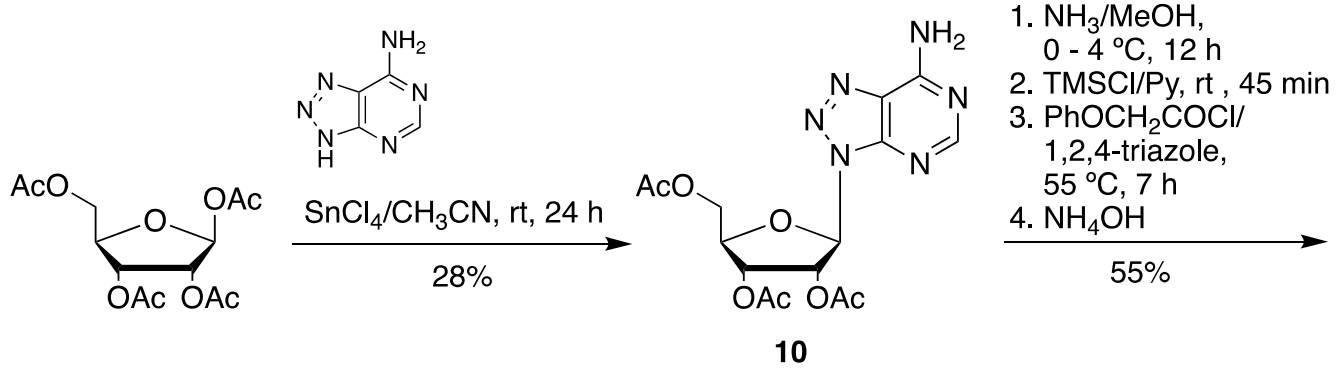
(m, 2H), 4.25 (m, 1H), 2.13 (s, 3H), 2.09 (s, 3H), 2.04 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 170.7, 169.7, 169.5, 157.3, 156.2, 149.5, 124.5, 87.4, 80.7, 73.3, 71.0, 63.1, 20.8, 20.6, 20.5.

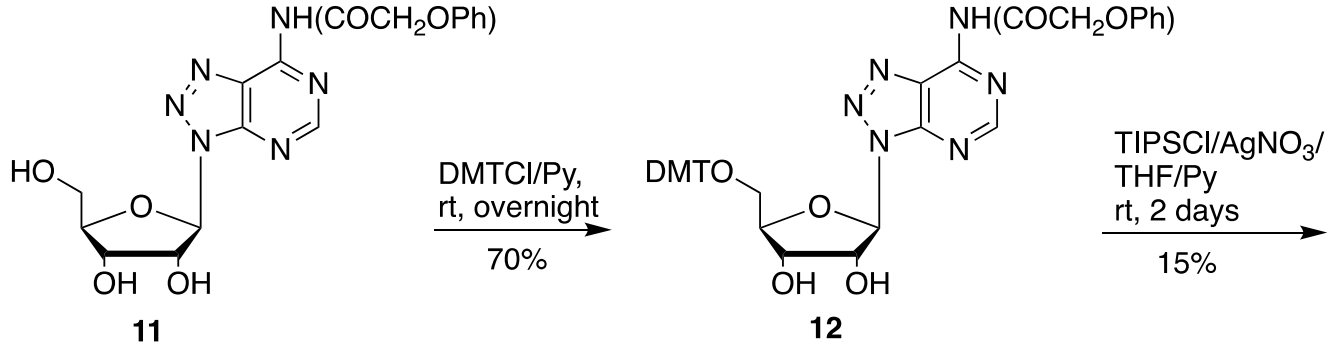
***N*^{6}-Phenoxyacetyl-8-azaadenosine (11):** The solution of **10** (936 mg, 2.37 mmol) was saturated in methanol (20 mL) and stored at 0 ºC for 30 minutes, then 4 ºC for 12 hours. The solvent was removed, the residue was dried over vacuum for 30 minutes and redissolved into dry pyridine (15 mL) under an argon atmosphere. To this solution, TMSCl (2.20 mL, 17.8 mmol) was added and the mixture was stirred at room temperature for 45 minutes. Meanwhile, the acylating agent was formed *in situ* by the dropwise addition of phenoxyacetyl chloride (0.50 mL, 3.6 mmol) to 1,2,4-triazole (250 mg, 3.6 mmol) in acetonitrile (5.0 mL) and pyridine (5.0 mL). The acylating reagent was then transferred by cannula needle to the flak containing transiently protected nucleoside. The reaction mixture was heated to 55ºC and stirred for 7 hours. After the mixture cooled to room temperature, water (2.0 mL) was added, followed by 28% ammonium hydroxide (1.1 mL). The solution was concentrated to remove pyridine, redissolved into water (20 mL), extracted with dichloromethane (16 mL). The organic layer was concentrated and purified by silica gel chromatography eluting with 5% methanol in dichloromethane to to give **11** as a white foam (528 mg, 55% yield). ^{1}H NMR (DMSO-d_6) δ 11.82 (brs, 1H), 8.92 (s, 1H), 7.30 (m, 2H), 7.00-6.95 (m, 3H), 6.29 (d, 1H, *J* = 4.8 Hz), 5.67 (d, 1H, *J* = 5.6 Hz), 5.36 (d, 1H, *J* = 5.6 Hz), 5.13 (s, 2H), 4.91 (dd, 1H, *J* = 10.0, 4.8 Hz), 4.86 (t, 1H, *J* = 6.0 Hz), 4.38 (dd, 1H, *J* = 9.6, 4.8 Hz), 4.05 (dd, 1H, *J* = 9.6, 4.8 Hz), 3.64 (m, 1H), 3.50 (m, 1H); ^{13}C NMR (DMSO-d_6) δ 168.0, 157.8, 156.3, 150.8, 150.3, 129.5, 126.0, 121.1, 114.6, 89.6, 86.3, 73.1, 70.6, 67.3, 61.7; HRMS calcd for $C_{17}H_{19}N_6O_6$ [MH^+]: 403.1366, found: 403.1361.

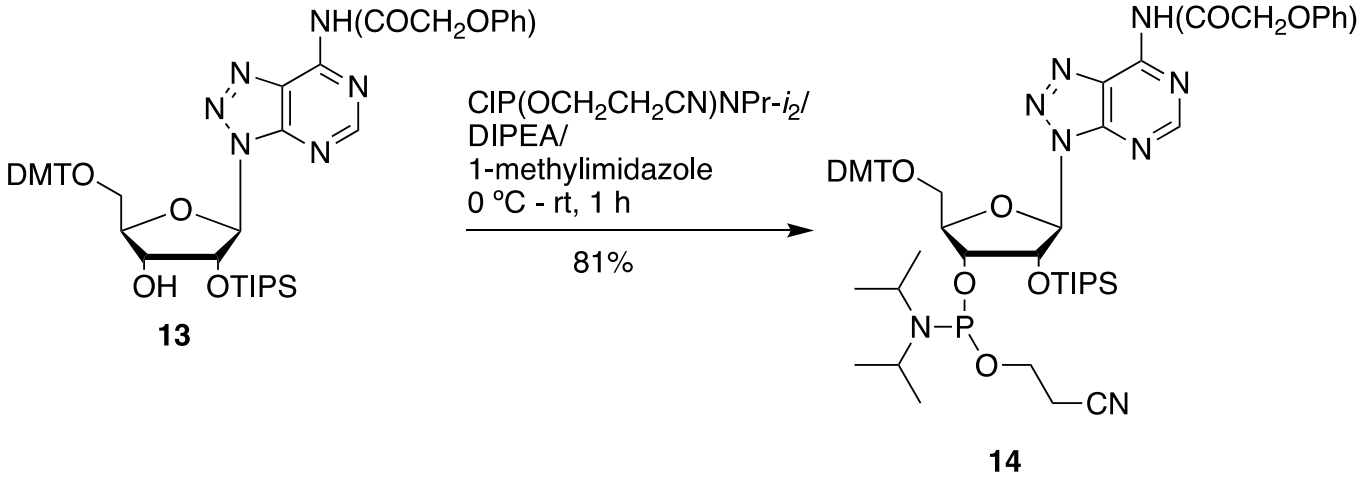
5´-*O*-DMT-*N*^{6}-phenoxyacetyl-8-azaadenosine (12): Under an argon atmosphere DMTCl (480 mg, 1.41 mmol) was added into the solution of **11** (476 mg, 1.18 mmol) in pyridine (10 mL). After the mixture was stirred at room temperature overnight, the reaction was quenched with methanol (5 mL). The solvent was removed, the residue was redissolved into dichloromethane, washed with saturated aqueous $NaHCO_3$, brine and dried over $MgSO_4$. The solvent was removed, and the residue was isolated by silica gel chromatography, eluting with 3% methanol in dichloromethane to give **12** (579 mg, 70% yield) as a pale yellow foam. ^{1}H NMR ($CDCl_3$/TMS) δ 9.76 (brs, 1H), 8.74 (s, 1H), 7.35-6.95 (m, 14H), 6.66 (m, 4H), 6.54 (d, 1H, *J* = 3.6 Hz), 5.22 (m, 1H), 4.81 (brs, 3H), 4.50 (brs, 1H), 4.34 (dd, 1H, *J* = 9.2, 4.8 Hz), 3.91 (brs, 1H), 3.67 (s, 3H), 3.66 (s, 3H), 3.41 (dd, 1H, *J* = 10.4, 3.6 Hz), 3.26 (dd, 1H, *J* = 10.4, 4.8 Hz); ^{13}C NMR ($CDCl_3$) δ 167.2, 158.4, 156.8, 156.4, 150.7, 149.4, 144.6, 135.9, 135.8, 130.08, 130.04, 129.96, 128.2, 127.8, 126.8, 125.7, 122.7, 114.9, 113.0, 105.0, 90.0, 86.3, 84.5, 74.1, 71.7, 68.0, 63.5, 55.2; HRMS calcd for $C_{38}H_{37}N_6O_8$ [MH^+]: 705.2673, found: 705.2654.

5´-*O*-DMT-2´-*O*-TIPS-*N*^{6}-phenoxyacetyl-8-aza-adenosine (13): Under an argon atmosphere to the solution of **12** (579 mg, 0.82 mmol) in dry THF (10 mL), pyridine (0.66 mL, 8.2 mmol), silver nitrate (280 mg, 1.64 mmol) and triisopropylsilyl chloride (TIPSCl) (220 μL, 1.03 mmol) were added. The reaction mixture was stirred at room temperature for 2 days. The solid was filtered off, the solution was evaporated, and the residue was isolated by silica gel chromatography, eluting with 20% ethyl acetate in hexane to give **13**

(106 mg, 15% yield) as a white foam. ^{1}H NMR ($CDCl_3$/TMS) δ 9.71 (brs, 1H), 8.89 (s, 1H), 7.45-7.05 (m, 14H), 6.76 (m, 4H), 6.53 (d, 1H, *J* = 4.5 Hz), 5.48 (t, 1H, *J* = 5.0 Hz), 4.90 (s, 2H), 4.54 (m, 1H), 4.32 (m, 1H), 3.76 (s, 6H), 3.48 (dd, 1H, *J* = 10.5, 4.0 Hz), 3.23 (dd, 1H, *J* = 10.5, 4.5 Hz), 2.82 (d, 1H, *J* = 5.0 Hz), 1.05-0.85 (m, 21H); ^{13}C NMR ($CDCl_3$) δ 166.5, 158.5, 156.89, 156.86, 151.0, 149.8, 144.8, 136.1, 135.9, 130.23, 130.21, 130.1, 128.3, 127.9, 126.8, 126.0, 122.9, 115.1, 113.17, 113.15, 90.2, 86.5, 85.1, 75.1, 72.3, 68.1, 63.5, 55.3, 17.8, 17.7, 12.1; HRMS calcd for $C_{47}H_{57}N_6O_8Si$ [MH^+]: 861.4007, found: 861.4002.







Supplementary Scheme 3. Synthesis of 8-azaadenosine phosphoramidite

5´-*O*-DMT-2´-*O*-TIPS-N^6-phenoxyacetyl-8-azaadenosine 3´-*N,N*-Diisopropyl(cyanoethyl)phosphoramidite (14): To the solution of **13** (100 mg, 0.116 mmol) and *i*-Pr_2NEt (200 μL, 1.16 mmol) in anhydrous dichloromethane (5.0 mL) at 0 ºC, $ClP(NPr\text{-}i_2)OCH_2CH_2CN$ (78 μL, 0.348 mmol) was added, followed by the addition of 1-methylimidazole (4.6 μL, 0.058 mmol). The reaction mixture was stirred at room temperature for 1 hour. The reaction was quenched with methanol (1.0 mL). The solvent was removed, the residue was purified by silica gel chromatography, eluting with 0.25% Et_3N in CH_2Cl_2 to give **14** (100 mg, 81% yield). ^{31}P NMR (CD_3CN) δ 151.7, 149.4 ppm; HRMS calcd for $C_{56}H_{74}N_8O_9PSi$ [MH^+]: 1061.5086, found: 1061.5084.

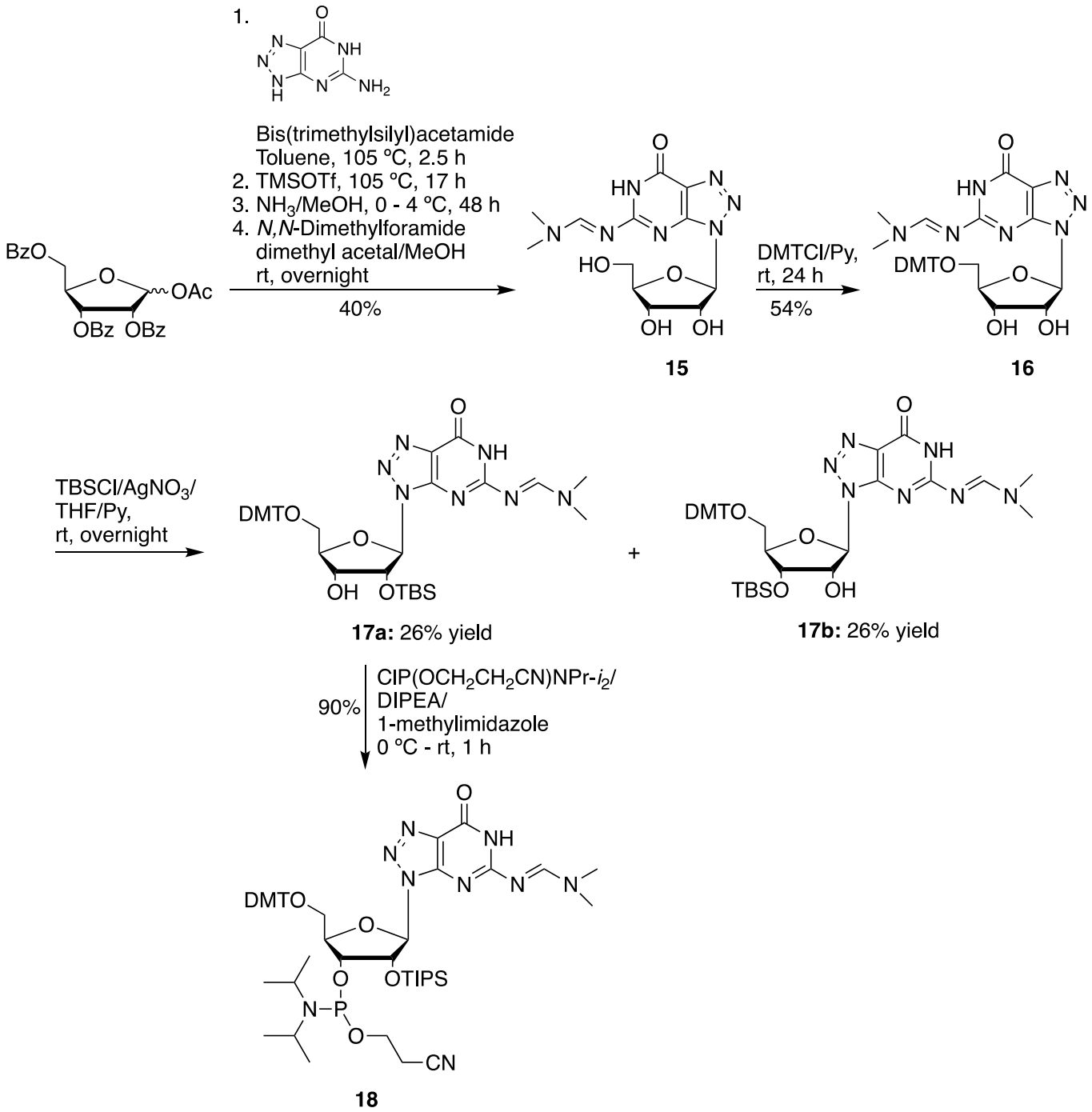
***N^2*-Dimethylaminomethylene-8-azaguanosine (15):** Under an argon atmosphere, to the mixture of 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-D-ribofuranose (504 mg, 1.0 mmol) and 8-azaguanine (152 mg, 3.73 mmol) in dry toluene (25 mL), *N*,*O*-bis(trimethylsilyl)acetamide (0.855 mL, 3.5 mmol) was added. The suspension was heated up and stirred at 105 °C for 2.5 hours and then cooled to room temperature. Trimethylsilyl trifluoromethanesulfonate (0.54 mL, 3.0 mmol) was added and the solution was stirred at 105 °C for 17 hours. After the reaction cooled to room temperature, the solvent was removed under vacuum and the residue was redissolved in dichloromethane and washed with saturated $NaHCO_3$ solution. The combined organic layers were dried over $MgSO_4$, filtered and finally evaporated under vacuum. The residue was dissolved into methanol (20 mL) and the solution was saturated with ammonia at 0 ºC for 30 minutes then stored at 4 ºC for 2 days. The residue was dried over vacuum overnight and redissolved in dry methanol (15 mL), then *N,N*-dimethylformamide dimethyl acetal (1.33 mL, 10.0 mmol) was added and the mixture was stirred at room temperature overnight. The solvent was removed, the residue was isolated by silica gel chromatography, eluting with 10% methanol in dichloromethane to give **15** (134 mg, 40% yield) as white solid. ^1H NMR (DMSO-d_6) δ 11.72 (brs, 1H), 8.66 (s, 1H), 6.01 (d, 1H, *J* = 5.0 Hz), 5.58 (d, 1H, *J* = 5.0 Hz), 5.31 (d, 1H, *J* = 4.0 Hz), 4.88 (t, 1H, *J* = 5.5 Hz), 4.74 (m, 1H), 4.26 (m, 1H), 3.96 (m, 1H), 3.65-3.50 (m, 2H), 3.20 (s, 3H), 3.07 (s, 3H); ^{13}C NMR (DMSO-d_6) δ 160.2, 159.2, 156.6, 151.0, 126.6, 88.5, 85.9, 73.2, 70.9, 62.1, 41.2, 35.1; HRMS calcd for $C_{12}H_{18}N_7O_5$ [MH^+]: 340.1369, found: 340.1369.

5´-*O*-DMT-*N^2*-dimethylaminomethylene-8-azaguanosine (16): Under an argon atmosphere DMTCl (1.65 g, 4.88 mmol) was added into the solution of **15** (1.382 g, 4.07 mmol) in pyridine (35 mL). After the mixture was stirred at room temperature for 24 h, the reaction was quenched with methanol (5 mL). The solvent was removed, the residue was redissolved into dichloromethane, washed with saturated aqueous $NaHCO_3$, brine and dried over $MgSO_4$. The solvent was removed, and the residue was isolated by silica gel chromatography, eluting with 3% methanol in dichloromethane to give **16** (1.402 mg, 54% yield) as a pale-yellow foam. ^1H NMR ($CDCl_3$/TMS) δ 10.08 (brs, 1H), 8.30 (s, 1H), 7.38 (d, 2H, *J* = 7.8 Hz), 7.26 (d, 4H, *J* = 8.4 Hz), 7.16 (t, 2H, *J* = 7.4 Hz), 7.09 (t, 1H, *J* = 7.0 Hz), 6.71 (d, 4H, *J* = 8.4 Hz), 6.25 (d, 1H, *J* = 3.6 Hz), 5.10 (m, 1H), 4.61 (m, 1H), 4.30 (m, 1H), 3.69 (s, 6H), 3.38 (m, 1H), 3.27 (m 1H), 2.89 (s, 3H), 2.82 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 159.3, 159.0, 158.4, 157.6, 151.5, 144.9, 136.1, 136.0, 130.2, 128.3, 127.8,

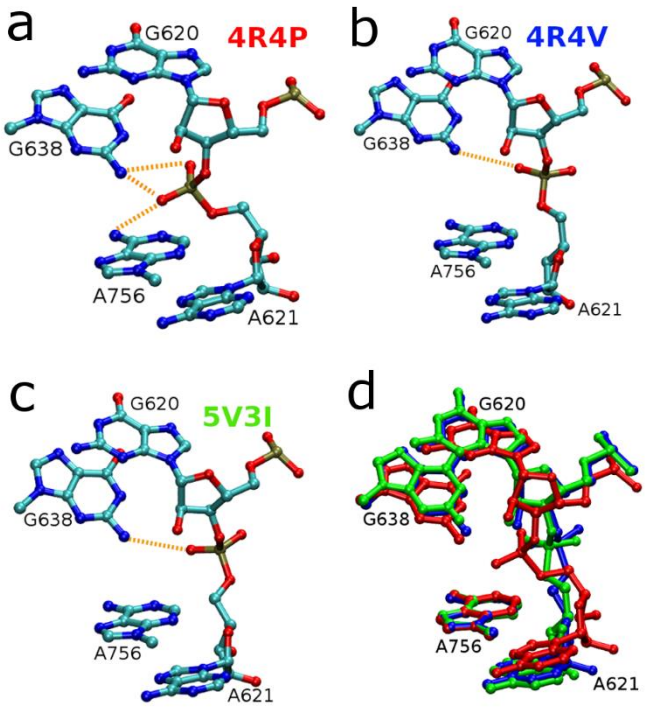
126.8, 126.3, 113.1, 88.5, 86.3, 84.2, 73.9, 71.8, 64.1, 55.3, 41.7, 35.5; HRMS calcd for $C_{33}H_{36}N_7O_7$ [MH^+]: 642.2676, found: 642.2667.

5´-*O*-DMT-2´-*O*-TBS-N^2-dimethylaminomethylene-8-azaguanosine (17a) and 5´-*O*-DMT-3´-*O*-TBS-N^2-dimethylaminomethylene-8-azaguanosine (17b): Under an argon atmosphere to the solution of **16** (925 mg, 1.44 mmol) in dry THF (25 mL), pyridine (1.16 mL, 11.7 mmol), silver nitrate (414 mg, 2.88 mmol) and TBSCl (326 mg, 2.16 mmol) were added. The reaction mixture was stirred at room temperature overnight. The solid was filtered off, the solution was evaporated and redissolved into dichlormethane, washed with saturated $NaHCO_3$, brine, and dried over $MgSO_4$. The solvent was removed and the residue was isolated by silica gel chromatography, eluting with 1-2% methanol in dichloromethane. TLC showed two products with the faster migrating spot corresponding to **17**a (278 mg, 26% yield) and the slower migrating spot to **17b** (278 mg, 26% yield) as light-yellow foams.
17a:^{1}H NMR ($CDCl_3$/TMS) δ 10.04 (brs, 1H), 8.76 (s, 1H), 7.60-7.25 (m, 9H), 6.89 (d, 4H, *J* = 8.8 Hz), 6.29 (d, 1H, *J* = 5.6 Hz), 5.40 (t, 1H, *J* = 5.2 Hz), 4.47 (m, 1H), 4.37 (m, 1H), 3.87 (s, 6H), 3.54 (dd, 1H, *J* = 10.4, 3.6 Hz), 3.32 (dd, 1H, *J* = 10.4, 4.4 Hz), 3.21 (s, 3H), 3.17 (s, 3H), 2.90 (d, 1H, *J* = 4.0 Hz), 0.96 (s, 9H), 0.17 (s, 3H), 0.003 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 159.7, 159.3, 158.5, 157.0, 152.0, 144.8, 136.1, 135.9, 130.2, 128.3, 127.9, 127.1, 126.8, 113.2, 87.8, 86.4, 84.6, 74.7, 72.2, 63.8, 55.3, 41.6, 35.6, 25.7, 18.0, -4.78, -5.02; HRMS calcd for $C_{39}H_{50}N_7O_7Si$ [MH^+]: 756.3541, found: 756.3538.
17b: ^{1}H NMR ($CDCl_3$/TMS) δ 9.85 (brs, 1H), 8.72 (s, 1H), 7.55-7.20 (m, 9H), 6.85 (m, 4H), 6.41 (d, 1H, *J* = 2.4 Hz), 4.94 (m, 1H), 4.89 (m, 1H), 4.34 (m, 1H), 3.85 (s, 6H), 3.46 (dd, 1H, *J* = 10.4, 3.0 Hz), 3.32 (d, 1H, *J* = 4.0 Hz), 3.26 (dd, 1H, *J* = 10.4, 5.6 Hz), 3.23 (s, 3H), 3.18 (s, 3H), 0.97 (s, 9H), 0.18 (s, 3H), 0.07 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 159.6, 159.2, 158.4, 156.8, 151.4, 144.7, 136.0, 135.9, 130.1, 130.0, 128.2, 127.8, 126.8, 126.7, 113.1, 88.1, 86.2, 84.2, 74.5, 72.4, 63.5, 55.3, 41.6, 35.4, 25.8, 18.0, -4.72, -4.80; HRMS calcd for $C_{39}H_{50}N_7O_7Si$ [MH^+]: 756.3541, found: 756.3542.

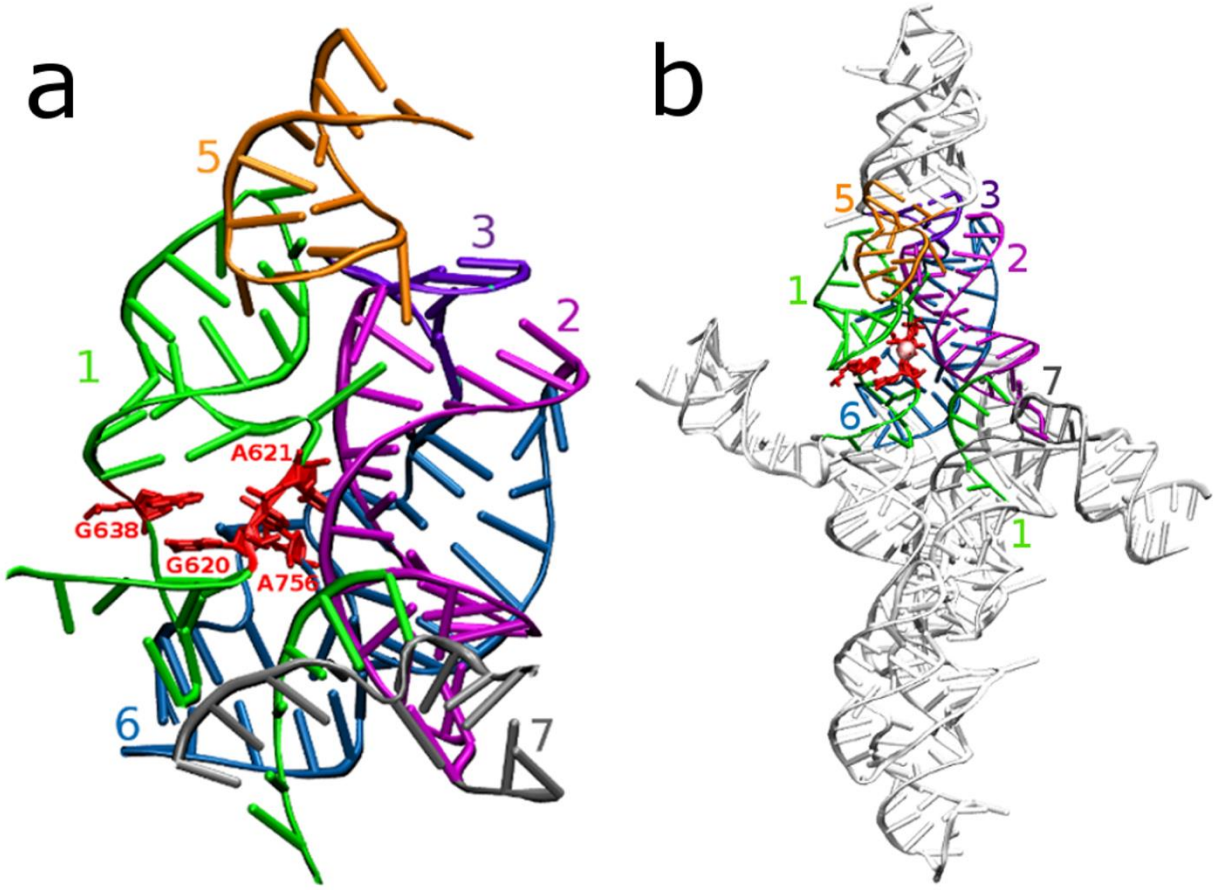
5´-*O*-DMT-2´-*O*-TBS-N^2-dimethylaminomethylene-8-azaguanosine 3´-*N,N*-Diisopropyl(cyanoethyl)phosphoramidite (18): To the solution of **17a** (182 mg, 0.241 mmol) and *i*-Pr_2NEt (420 μL, 2.41 mmol) in anhydrous dichloromethane (10.0 mL) at 0 ºC, $ClP(NPr\text{-}i_2)OCH_2CH_2CN$ (161 μL, 0.723 mmol) was added, followed by the addition of 1-methylimidazole (9.6 μL, 0.121 mmol). The reaction mixture was stirred at room temperature for 1 hour. The reaction was quenched with methanol (1.0 mL). The solvent was removed, the residue was purified by silica gel chromatography, eluting with 0.25% Et_3N in CH_2Cl_2 to give **18** (207 mg, 90% yield) as a white foam. ^{31}P NMR (CD_3CN) δ 150.5, 149.7 ppm; HRMS calcd for $C_{48}H_{67}N_9O_8PSi$ [MH^+]: 956.4620 found: 956.4625.



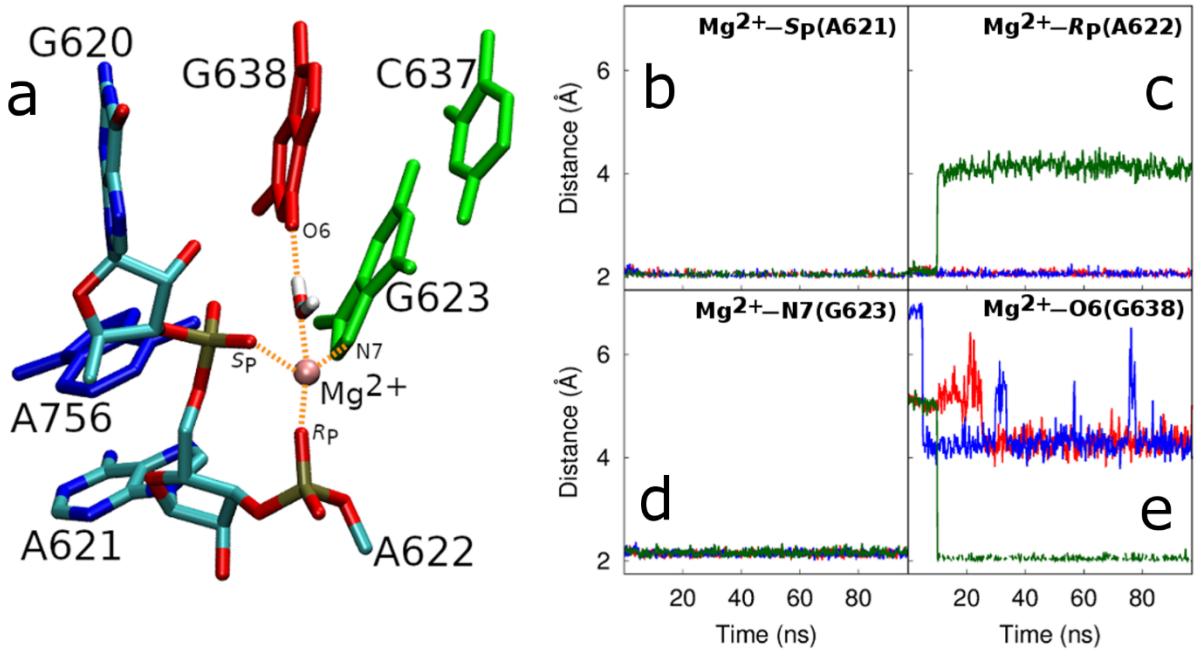
Supplementary Scheme 4. Synthesis of 8-azaguanosine phosphoramidite



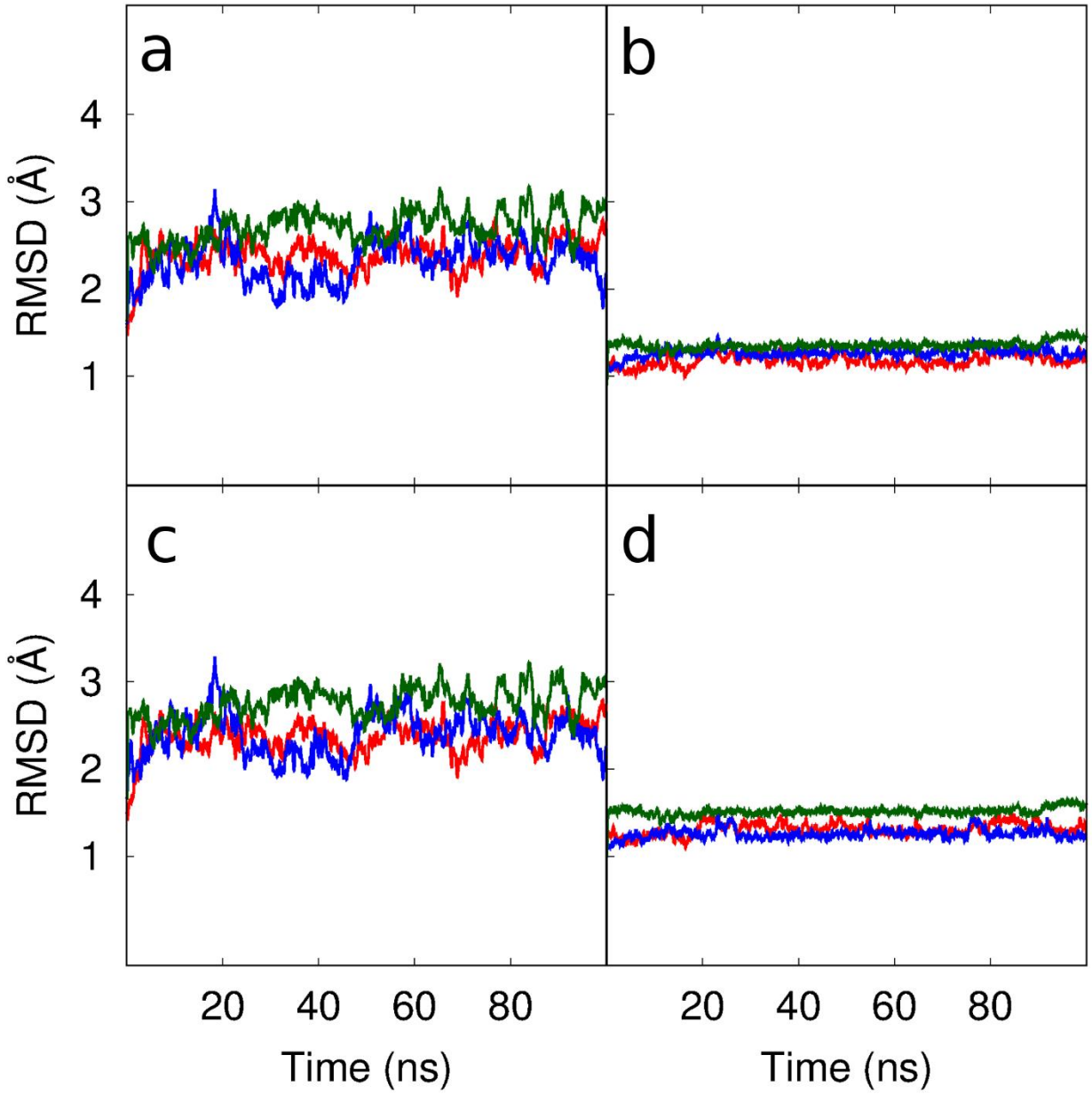
Supplementary Figure 1. Catalytic site of the VS ribozyme. Active sites in the (a) 4R4P, (b) 4R4V, (c) 5V3I crystal structures after the respective mutated residues are replaced with the corresponding wildtype nucleobases. (d) Superimposition of the three different active sites (4R4P – red, 4R4V – green, 5V3I – blue). 4R4P was obtained by mutating the putative general acid, A756, to a G, and 4R4V was obtained by mutating the putative general base, G638, to an A. The 5V3I structure also corresponds to a G638A mutant but does not contain the C634G mutation that was introduced in the 4R4P and 4R4V structures to facilitate substrate binding. The active sites of these crystal structures do not clearly exhibit the key catalytic features that are laid out in Breaker's framework for discussing RNA catalysis, possibly because of distortions caused by mutations to the catalytic nucleobases. Specifically, in each of the three structures, the A756 and G638 residues are positioned too far (≥4 Å) from the 2'O nucleophile and 5'O leaving group to participate in ɣ and δ catalysis, respectively. The angle formed between the O2' oxygen, scissile phosphorus atom, and O5' oxygen (τ) ranges from 97-129°, far from the ideal the in-line geometry (τ = 180°) optimal for α catalysis. In the 4R4P structure, the G638 exocyclic amine is within interacting distance to both NPOs, while in the 4R4V and 5V3I structures, it is positioned to interact with the pro-R_P NPO alone. The 4R4P structure also has the A756 exocyclic amine within interacting distance of the pro-R_P NPO; the residue being further away from the scissile phosphate in the other two structures. Thus, the functional relevance of these NPO interactions for β catalysis is unclear. Although the crystal structures are unable to offer an unambiguous mechanistic interpretation, they provide a critical starting point from which to computationally derive and experimentally test detailed dynamical models of the active state ensemble in solution, allowing specific catalytic strategies to be revealed.



Supplementary Figure 2. Illustration of the truncated model of the ribozyme used in this study. (a) Tertiary structure of the model, with the different helical elements numbered and colored according to ref [1]. The cleavage site along with the catalytic nucleobases are shown as red sticks. (b) The model is depicted in context of the dimeric ribozyme by showing the dimer in white and residues that belong to the model in various colors.

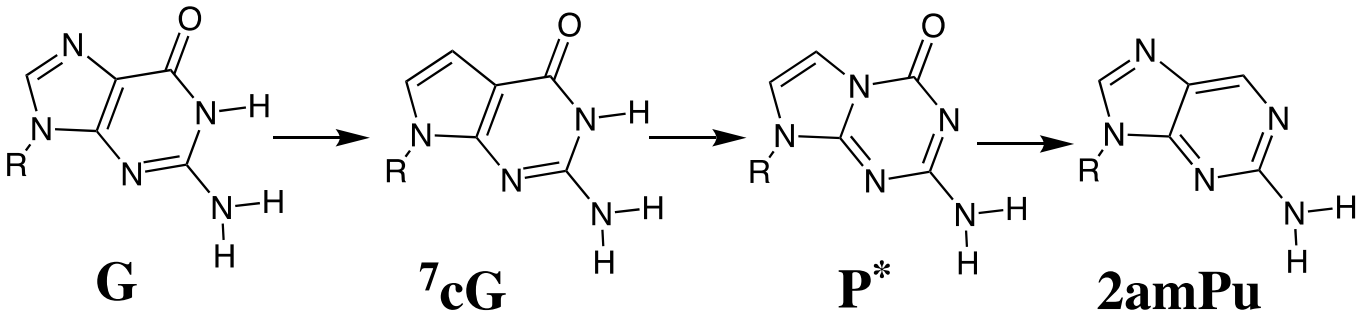


Supplementary Figure 3. Proposed Mg^{2+} ion binding site in the active site. (a) Cartoon depiction of the Mg^{2+} ion binding site. (b)-(e) illustrate the various Mg^{2+} coordination distances during the MD simulations.

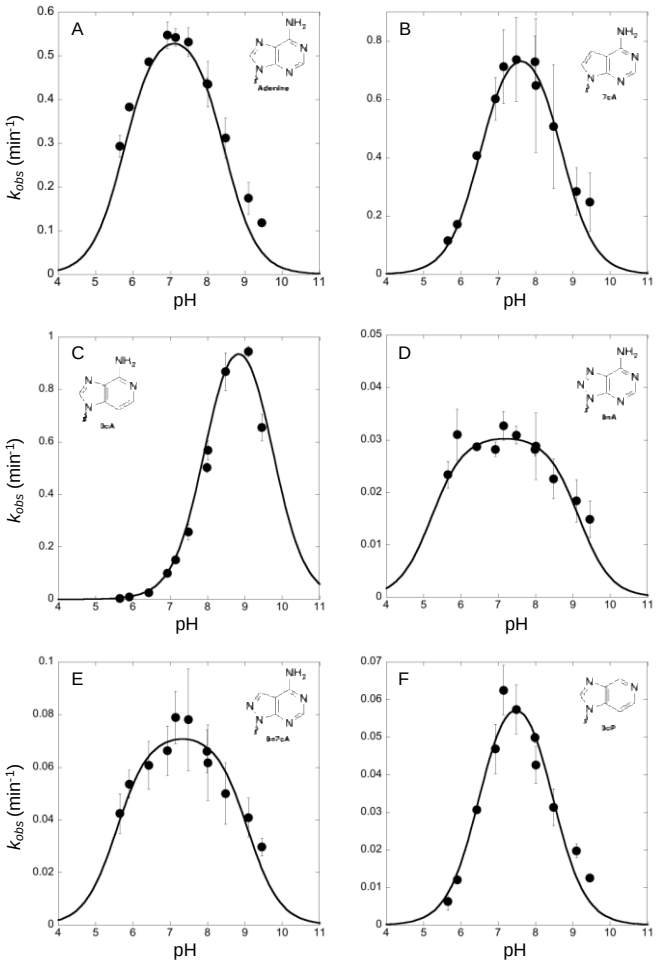


Supplementary Figure 4. Convergence of MD simulations. Root mean squared deviations (RMSDs) of (a) all residues belonging to the ribozyme model, and (c) residues belonging to the active site, are reported along 100 ns MD simulations. Red, blue, and green colored lines correspond to independent simulations based on the three different crystal structures with PDB IDs 4R4P, 4R4V, and 53VI, respectively. These crystal structures are all very similar (within 1.1 Å RMSD heavy-atom backbone). RMSDs in panels [a-b] are calculated with respect to a single structure obtained from averaging the 4R4P, 4R4V, and 53VI crystal structures. The average structure was obtained by averaging the heavy atom positions of the residues in the three crystal structures after aligning the structures by RMSD fitting. Panels [c-d] are analogous to [a-b], except in [c-d] the RMSDs of all residues (c) and active site residues (d) are calculated with respect to the crystal structure from which that particular simulation departed. According to panel (a), during the entire length of the MD simulations, the ribozyme does not undergo large

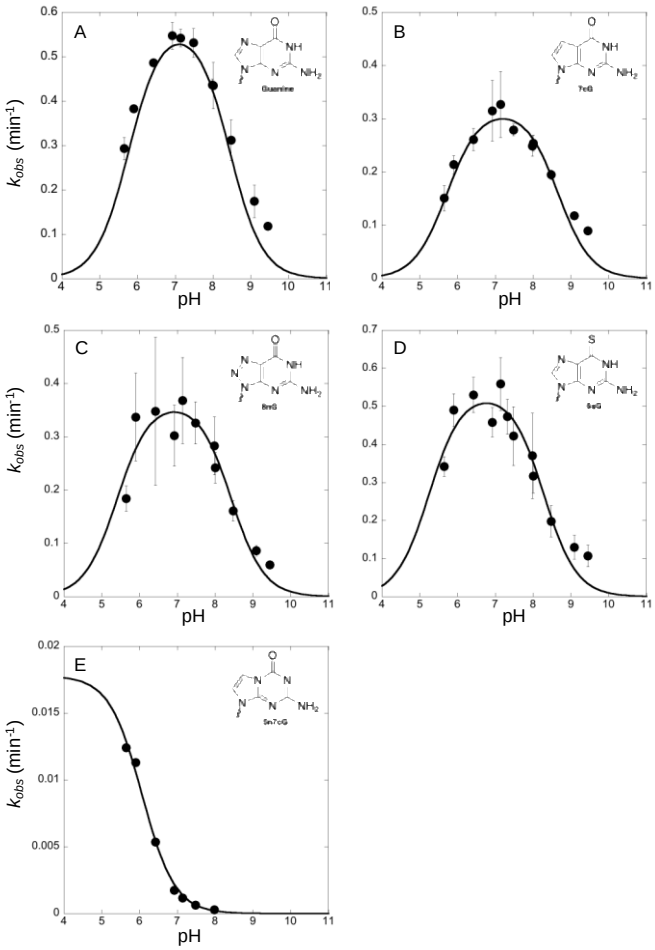
scale conformational changes and remains close to the average crystal structure reference with RMSDs typically between 2-3 Å. According to panel (b), the active site configurations in all three simulations departing from the three crystal structures are very similar and also remain very close to the average crystal structure reference with RMSDs typically around 1 Å. According to panels [c-d], RMSDs calculated with respect to the crystal structures from which the different simulations departed are very similar to RMSDs calculated with respect to the single average crystal structure reference. In the MD simulations reported in this figure, a Mg^{2+} was bound at the pro-S_P NPO of the scissile phosphate.



Supplementary Figure 5. Atomic mutations series to assess functional contributions from G's keto group.



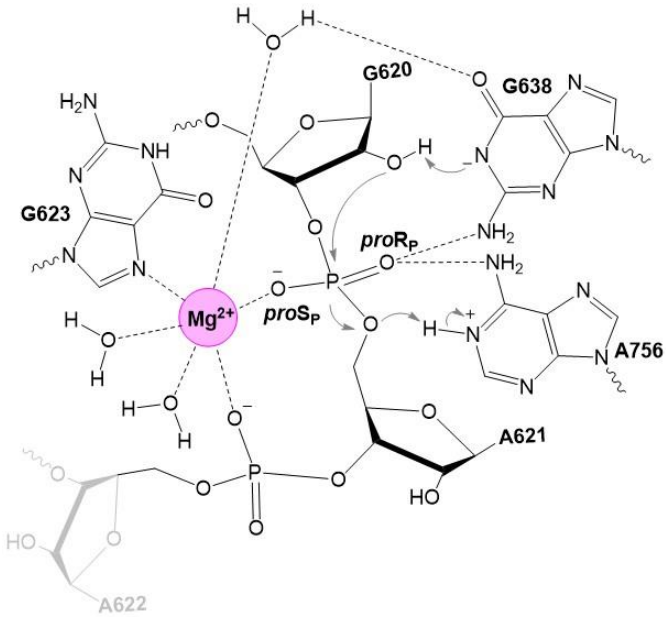
Supplementary Figure 6. Cleavage rate over a range of pH for VS ribozymes variants with A756 mutations. A) Wild type/Adenine, B) 7-deazaadenine (7cA), C) 3-deazaadenine (3cA), D) 8-azaadenine (8nA), E) 8-aza-7-deazaadenine (8n7cA), F) 3-deazapurine (3cP). All pH-profiles were fit to a model for double ionization the results of which are summarized in supplemental table 1.



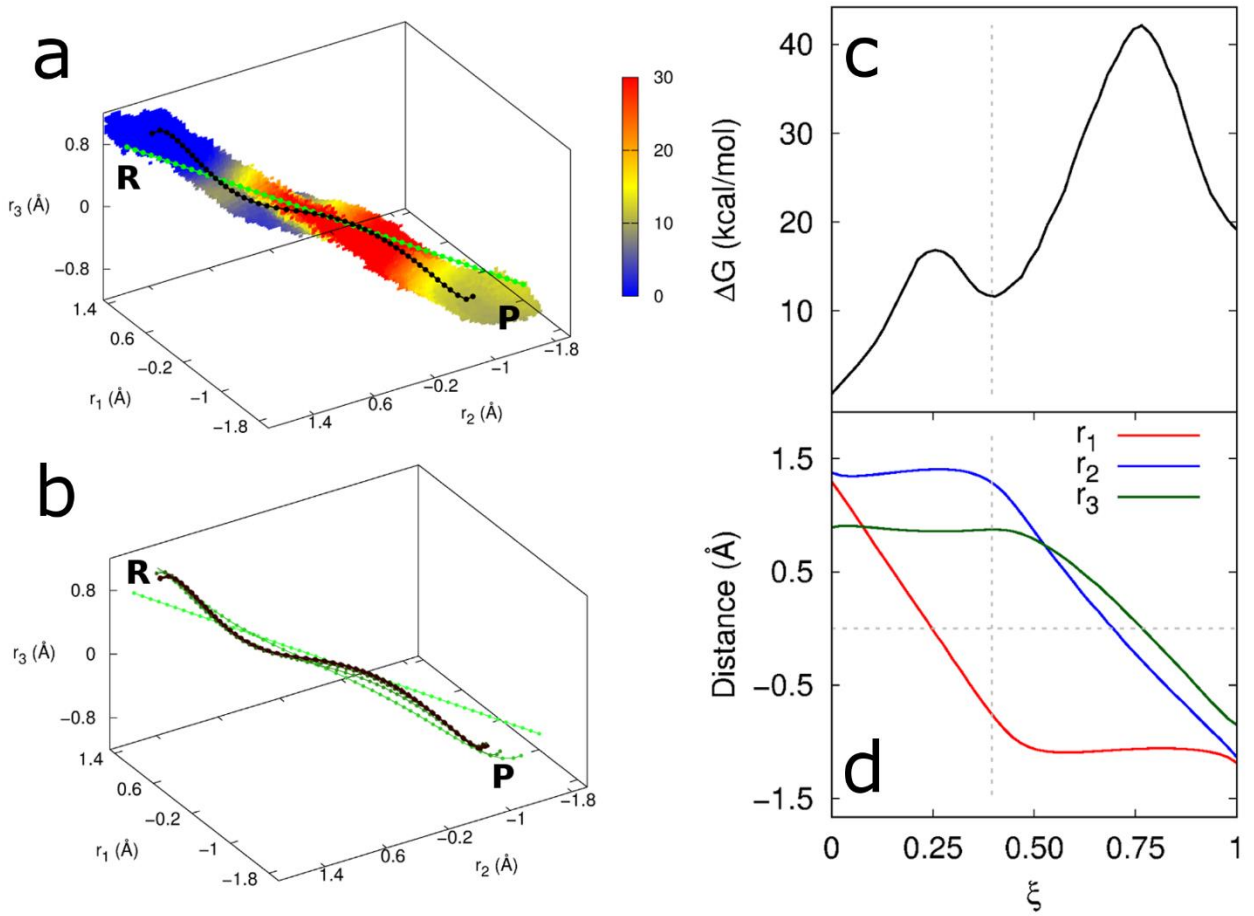
Supplementary Figure 7. Cleavage rate over a range of pH for VS ribozymes variants with G638 mutations. A) Wild type/Guanine, B) 7-deazaguanine (7cG), C) 8-azaguanine (8nG), D) 6-thioguinine (6sG), E) 5-aza-7-deazaguanine (5n7cG). All pH-profiles were fit to a model for double ionization with the exception of 5n7cG/P*, which was fit to a model for a single ionization. The results of a fits are summarized in supplemental table 1.

Ribozyme	Substrate	k_{cat} (min^{-1})	p$K_{a,1}$	p$K_{a,2}$
Wild Type	Wild Type	270 ± 120	5.8 ± 0.1	8.4 ± 0.1
A756(7cA)	Wild Type	120 ± 50	6.5 ± 0.1	8.7 ± 0.1
A756(3cA)	Wild Type	70 ± 40	8.0 ± 0.1	9.7 ± 0.2
A756(8nA)	Wild Type	260 ± 160	5.2 ± 0.2	9.1 ± 0.1
A756(8n7aA)	Wild Type	230 ± 110	5.6 ± 0.1	9.1 ± 0.1
A756(3cP)	Wild Type	6 ± 3	6.5 ± 0.1	8.4 ± 0.1
Wild Type	G638(8nG)	120 ± 40	5.8 ± *n.a*	8.3 ± 0.1
Wild Type	G638(7cG)	270 ± 60	5.8 ± *n.a*	8.7 ± 0.1
Wild Type	G638(6sG)	100 ± 30	5.8 ± *n.a*	8.0 ± 0.2
Wild Type	G638(P*)	0.018 ± 0.001	< 6.1	

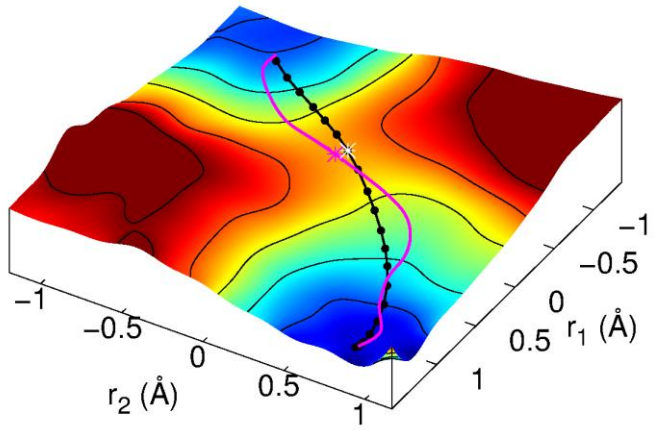
Supplementary Table 1. Fitted parameters of from pH-profiles in supplemental figures S5 and S6. All profiles were fit to a model for double ionization with the exception of the G638P* mutant substrate, which was fit to a model for a single ionization. Values with errors given as “n,a.” were fixed in data fitting.



Supplementary Figure 8. QM region considered in the QM/MM simulations. The QM region in simulations with the active site Mg^{2+} ion consisted of 123 atoms and is depicted in the figure in dark black color.



Supplementary Figure 9. Finite temperature string simulations of the VS ribozyme catalytic pathway in presence of the active site Mg^{2+} ion using AM1-d/pho-T Hamiltonian. (a) Free energy surface with initial and final strings. (b) Convergence of the strings. (c) 1D free energy profile along the converged string. (d) Evolution of reaction coordinates along converged string. In (c) and (d) the X-axis, ξ, represents the reduced length of the converged string ($0 \leq \xi \leq 1$). The reactant and product states correspond to the values of the string at $\xi = 0$, and $\xi = 1$, respectively.



Supplementary Figure 10. Comparison of the converged pathway obtained from FTS simulations using the AM1-d/pho-T Hamiltonian with the analytic minimum free energy path on the 2D free energy surface calculated from *ab initio* umbrella sampling simulations. The converged pathway obtained from the FTS simulations is shown in black. The analytic pathway, shown in magenta, is calculated by first identifying the reactant and product minima on the 2D surface. Next, a series of linearly interpolated points are generated between the two minima and a path is constructed passing through the points and connecting the minima. This path is then iteratively refined until convergence by performing local free energy minimization at the various points and reconstructing the path. The "*" symbols shown in white and magenta denote the transition states of the pathway obtained from FTS simulations and the analytic pathway, respectively. The figure illustrates that the transition states obtained from the two methods have very similar configurations.

Extended discussions

This section is intended to provide a slightly elaborate discussion of certain sections of the main text that had to kept brief because of length restrictions.

Phosphorothioate substitution and mutational analysis of A756 and G638 reveal catalytic interactions in the active site

In an effort to better understand the interactions of the scissile phosphate NPOs (mechanism(s) of β catalysis) in VS ribozyme, we tested the effect of sulfur substitution at the non-bridging positions of the scissile phosphate. Earlier data demonstrated the deleterious effect of phosphorothioate substitution at the scissile phosphate; however, interpretation of these data is complicated by the fact that one study employed a mixture of diastereomers[54] and another observed the effect of thio substitution on the reverse reaction[55]. In this study, the R_P and S_P thio diastereomers were synthesized as a diastereomeric mixture, separated by HPLC and identified by selective endonuclease digestion (see Methods). The results from the phosphorothioate experiments are summarized in Table 1. Under standard conditions (10mM $MgCl_2$, 25mM KCl, 2mM spermidine, pH = 8), the wild type ribozyme cleaved with k_{obs}= 0.32 min^{-1}. In contrast, the A621 R_P and S_P thio substrates cleaved with rates below the limit of detection (here, 10^{-4} min^{-1}) corresponding to thio effects of at least 3000-*fold*. These data suggest that critical catalytic interactions are present at both NPOs of the scissile phosphate. Later, we show that the S_P thio substrate exhibits metal ion rescue effects, whereas the R_P thio substrate does not.

We used chemically precise nucleobase modifications to pinpoint key stereospecific functional interactions in the active site with the pro-R_P position of the substrate scissile phosphate. Inspection of the crystal structures showed that the exocyclic amines of G638 and A756 reside spatially close to the scissile phosphate NPOs. Previous mutational studies[56,57] of these nucleobases have hinted at the importance of their exocyclic amines. However, the interpretation of these studies was complicated by the fact that the mutations made to remove the exocyclic amines at the 638 and 756 positions alter the pK_as of these residues, which consequently could affect their putative roles in proton transfer during the reaction (see section 8 in the main paper). Moreover, these mutations were not analyzed in combination with phosphorothioate substitution. To identify potential interactions of the G638 and A756 exocyclic amines to the NPOs, we observed cleavage for ribozymes with the G638 mutated to inosine (G638I), A756 mutated to 3-deazapurine (A756(3cP)), or both (referred to here as “double mutant”) in oxo and phosphorothioate backgrounds. The use of 3-deazapurine provides a strategy to test removal of A’s exocyclic amine without major perturbation to the N1 pK_a.

Independent MD simulations departing from different crystal structures converge on a functionally inactive state of the ribozyme in the absence of Mg^{2+}

The MD simulations are represented in Figure 2 in the form of interaction maps that report on the probability of observing specific active site interactions that are deemed to be important from functional data and are associated with each of the catalytic strategies (α-δ). In simulations without active site Mg^{2+} ion (Figure 2a), the phosphate reaction center seldom achieves an in-line conformation (blue), suggesting inadequate α catalysis. Among the functionally implicated interactions that affect the stabilization of negative charge developed at scissile NPOs (β catalysis, shown in green), only the A756 exocyclic amine engages in a steady interaction with the pro-R_P NPO, consistent with our experimental data; in contrast, G638 is mobile, its exocyclic amine forms only weak transient interaction with the functionally implicated pro-R_P NPO, and a Na^+ ion coordinates to the pro-S_P NPO with partial occupancy. Moreover, the position of the putative general base G638:N1 very rarely interacts with the nucleophilic 2'-hydroxyl group (red), and thus the simulation model does not support activation of O2' nucleophile (γ catalysis). A756:N1 is observed to interact with the O5' oxygen of the leaving group (purple), albeit not consistently, suggesting moderate participation in stabilization of the leaving group (δ catalysis). Thus, the MD simulations based on the crystal structures suggest that each of the key elements of catalysis (α, β, γ, and δ) are either absent or not present in full strength to support efficient catalysis.

In simulations with active site Mg^{2+} ion (Figure 2f), The phosphate reaction center is observed to be in-line during the entire length of these simulations (blue) indicating strong in-line fitness (α catalysis). Stabilization of scissile NPOs (β catalysis, shown in green) is supported by stable interactions between the exocyclic amines of G638 and A756 and the pro-R_P NPO, in agreement with the mutational rescue data, and between the Mg^{2+} ion and the pro-S_P NPO, in accord with the metal ion rescue data. Persistent interactions are also observed between G638:N1 and the 2'-hydroxyl group (red), and A756:N1 and the 5'O leaving group (purple), consistent with their anticipated involvement in O2' nucleophile activation (γ catalysis) and leaving group O5' stabilization (δ catalysis), respectively (see below). The high color intensity of these interaction maps reflects preservation of the active site interactions throughout the simulations, supporting each of the key catalytic strategies (α, β, γ, and δ).

LFERs suggest proton transfer is nearly complete for the general base and partial for the general acid in the transition state

In general acid-base catalysis, the intrinsic rate constant should depend log-linearly on the pK_a of the general acid and base. Plotting log(k_{cat}) against pK_a gives a so-called Brønsted plot,[58] the slope of which corresponds to a Brønsted coefficient (α or β), which is a measurement of the sensitivity of the reaction to acid or base strength and reflects the extent of proton transfer [59]. In practice, measurement of Brønsted coefficients in ribozyme (or enzyme) catalysis is complicated by the fact that alterations to nucleobases that affect pK_a often also impact other aspects of catalysis. For instance, mutation of guanosine to inosine results in an omission of an exocyclic amine which interacts with the pro-R_P NPO of the scissile phosphate, thus such a mutation confounds Brønsted analysis. To avoid this problem, we employed a series of purine analogs with endocyclic nitrogen perturbations that exhibit a broad range of pK_as at the N1 position,

but differ structurally only with respect to the insertion and/or deletion of endocyclic nitrogens. These “isofunctional” purine analogs thereby leave the Watson-Crick face and exocyclic substituents of the nucleobase intact, but provide a range of nucleobase pK_as. If the charge on the nucleobase under investigation changed between the ground state and transition state of the reaction consistent with proton transfer, then k_{cat} for the analogs would be expected to vary log-linearly with pK_a, giving a slope that reflects the change in charge. Alternatively, if the charge on the nucleobase under investigation does not change over the course of the reaction, or if the perturbations of the endocyclic nitrogens affect catalysis through some mechanism other than pK_a shifting, then the k_{cat} will remain constant, or its relationship with pK_a will diverge from log-linearity.

1 Suslov, N. B. *et al.* Crystal structure of the Varkud satellite ribozyme. *Nat Chem Biol* **11**, 840-846, doi:10.1038/nchembio.1929 (2015).
2 DasGupta, S., Suslov, N. B. & Piccirilli, J. A. Structural Basis for Substrate Helix Remodeling and Cleavage Loop Activation in the Varkud Satellite Ribozyme. *J Am Chem Soc* **139**, 9591-9597, doi:10.1021/jacs.7b03655 (2017).
3 Horn, H. W., Swope, W. C. & Pitera, J. W. Characterization of the TIP4P-Ew water model: Vapor pressure and boiling point. *J Chem Phys* **123**, 194504, doi:10.1063/1.2085031 (2005).
4 Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J Chem Phys* **98**, 10089-10092, doi:10.1063/1.464397 (1993).
5 Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comp Phys* **23**, 327-341, doi:10.1016/0021-9991(77)90098-5 (1977).
6 Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J Chem Phys* **81**, 3684-3690, doi:10.1063/1.448118 (1984).
7 Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* **97**, 10269-10280, doi:10.1021/j100142a004 (1993).
8 Cieplak, P., Cornell, W. D., Bayly, C. & Kollman, P. A. Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins. *J Comp Chem* **16**, 1357-1377, doi:10.1002/jcc.540161106 (1995).
9 Gaines, C. S. & York, D. M. Model for the Functional Active State of the TS Ribozyme from Molecular Simulation. *Angew Chem Int Ed Engl* **56**, 13392-13395, doi:10.1002/anie.201705608 (2017).
10 Lee, T.-S. *et al.* GPU-Accelerated Molecular Dynamics and Free Energy Methods in Amber18: Performance Enhancements and New Features. *J Chem Inf Model* **58**, 2043-2050, doi:10.1021/acs.jcim.8b00462 (2018).
11 Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J Chem Theory Comput* **9**, 3878-3888, doi:10.1021/ct400314y (2013).
12 Cornell, W. D. *et al.* A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules J. Am. Chem. Soc. 1995, 117, 5179−5197. *J Am Chem Soc* **118**, 2309-2309, doi:10.1021/ja955032e (1996).
13 Wang, J., Cieplak, P. & Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comp Chem* **21**, 1049-1074, doi:10.1002/1096-987X(200009)21:12<1049::AID-JCC3>3.0.CO;2-F (2000).
14 Pérez, A. *et al.* Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys J* **92**, 3817-3829, doi:10.1529/biophysj.106.097782 (2007).

15 Zgarbová, M. *et al.* Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J Chem Theory Comput* **7**, 2886-2902, doi:10.1021/ct200162x (2011).
16 Banáš, P. *et al.* Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J Chem Theory Comput* **6**, 3836-3849, doi:10.1021/ct100481h (2010).
17 Joung, I. S. & Cheatham, T. E., 3rd. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* **112**, 9020-9041, doi:10.1021/jp8001614 (2008).
18 Li, P., Roberts, B. P., Chakravorty, D. K. & Merz, K. M., Jr. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J Chem Theory Comput* **9**, 2733-2748, doi:10.1021/ct400146w (2013).
19 Panteva, M. T., Giambaşu, G. M. & York, D. M. Force Field for Mg^{2+}, Mn^{2+}, Zn^{2+}, and Cd^{2+} Ions That Have Balanced Interactions with Nucleic Acids. *J Phys Chem B* **119**, 15460-15470, doi:10.1021/acs.jpcb.5b10423 (2015).
20 Li, P. & Merz, K. M. Taking into Account the Ion-Induced Dipole Interaction in the Nonbonded Model of Ions. *J Chem Theory Comput* **10**, 289-297, doi:10.1021/ct400751u (2014).
21 Giambaşu, George M., Luchko, T., Herschlag, D., York, Darrin M. & Case, David A. Ion Counting from Explicit-Solvent Simulations and 3D-RISM. *Biophys J* **106**, 883-894, doi:10.1016/j.bpj.2014.01.021 (2014).
22 Lipfert, J., Doniach, S., Das, R. & Herschlag, D. Understanding Nucleic Acid–Ion Interactions. *Ann Rev Biochem* **83**, 813-841, doi:10.1146/annurev-biochem-060409-092720 (2014).
23 Giambaşu, G. M. *et al.* Competitive interaction of monovalent cations with DNA from 3D-RISM. *Nucleic Acids Res* **43**, 8405-8415, doi:10.1093/nar/gkv830 (2015).
24 Giambaşu, G. M., Case, D. A. & York, D. M. Predicting Site-Binding Modes of Ions and Water to Nucleic Acids Using Molecular Solvation Theory. *J Am Chem Soc* **141**, 2435-2445, doi:10.1021/jacs.8b11474 (2019).
25 Beglov, D. & Roux, B. An Integral Equation To Describe the Solvation of Polar Molecules in Liquid Water. *J Phys Chem B* **101**, 7821-7826, doi:10.1021/jp971083h (1997).
26 Kovalenko, A., Ten-no, S. & Hirata, F. Solution of three-dimensional reference interaction site model and hypernetted chain equations for simple point charge water by modified method of direct inversion in iterative subspace. *J Comp Chem* **20**, 928-936, doi:10.1002/(SICI)1096-987X(19990715)20:9<928::AID-JCC4>3.0.CO;2-X (1999).
27 Cortis, C. M., Rossky, P. J. & Friesner, R. A. A three-dimensional reduction of the Ornstein–Zernicke equation for molecular liquids. *J Chem Phys* **107**, 6400-6414, doi:10.1063/1.474300 (1997).
28 Perkyns, J. S. & Montgomery Pettitt, B. A dielectrically consistent interaction site theory for solvent—electrolyte mixtures. *Chem Phys Lett* **190**, 626-630, doi:10.1016/0009-2614(92)85201-K (1992).
29 Perkyns, J. & Pettitt, B. M. A site–site theory for finite concentration saline solutions. *J Chem Phys* **97**, 7656-7666, doi:10.1063/1.463485 (1992).

30 Kast, S. M. & Kloss, T. Closed-form expressions of the chemical potential for integral equation closures with certain bridge functions. *J Chem Phys* **129**, 236101, doi:10.1063/1.3041709 (2008).
31 Berendsen, H. J. C., Grigera, J. R. & Straatsma, T. P. The missing term in effective pair potentials. *The Journal of Physical Chemistry* **91**, 6269-6271, doi:10.1021/j100308a038 (1987).
32 E, W., Liu, D. & Vanden-Eijnden, E. Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates. *J Chem Phys* **123**, 194107, doi:10.1063/1.2109987 (2005).
33 Vanden-Eijnden, E. & Venturoli, M. Revisiting the finite temperature string method for the calculation of reaction tubes and free energies. *J Chem Phys* **130**, 194103, doi:10.1063/1.3130083 (2009).
34 Nam, K., Cui, Q., Gao, J. & York, D. M. Specific Reaction Parametrization of the AM1/d Hamiltonian for Phosphoryl Transfer Reactions: H, O, and P Atoms. *J Chem Theory Comput* **3**, 486-504, doi:10.1021/ct6002466 (2007).
35 Panteva, M. T. *et al.* in *Methods in Enzymology* Vol. 553 (eds Shi-Jie Chen & Donald H. Burke-Aguero) 335-374 (Academic Press, 2015).
36 Gao, J. *et al.* Mechanisms and Free Energies of Enzymatic Reactions. *Chem Rev* **106**, 3188-3209, doi:10.1021/cr050293k (2006).
37 Garcia-Viloca, M., Gao, J., Karplus, M. & Truhlar, D. G. How Enzymes Work: Analysis by Modern Rate Theory and Computer Simulations. *Science* **303**, 186, doi:10.1126/science.1088172 (2004).
38 Tse, C. K. M. *et al.* Intrinsic cleavage of RNA polymerase II adopts a nucleobase-independent mechanism assisted by transcript phosphate. *Nature Catalysis*, doi:10.1038/s41929-019-0227-5 (2019).
39 Rosta, E., Nowotny, M., Yang, W. & Hummer, G. Catalytic mechanism of RNA backbone cleavage by ribonuclease H from quantum mechanics/molecular mechanics simulations. *J Am Chem Soc* **133**, 8934-8941, doi:10.1021/ja200173a (2011).
40 Ganguly, A., Thaplyal, P., Rosta, E., Bevilacqua, P. C. & Hammes-Schiffer, S. Quantum mechanical/molecular mechanical free energy simulations of the self-cleavage reaction in the hepatitis delta virus ribozyme. *J Am Chem Soc* **136**, 1483-1496, doi:10.1021/ja4104217 (2014).
41 Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comp Chem* **13**, 1011-1021, doi:10.1002/jcc.540130812 (1992).
42 Lee, T. S., Radak, B. K., Pabis, A. & York, D. M. A New Maximum Likelihood Approach for Free Energy Profile Construction from Molecular Simulations. *J Chem Theory Comput* **9**, 153-164, doi:10.1021/ct300703z (2013).
43 Lee, T. S., Radak, B. K., Huang, M., Wong, K. Y. & York, D. M. Roadmaps through free energy landscapes calculated using the multi-dimensional vFEP approach. *J Chem Theory Comput* **10**, 24-34, doi:10.1021/ct400691f (2014).
44 Shirts, M. R. & Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* **129**, 124105, doi:10.1063/1.2978177 (2008).
45 Giese, T. J. & York, D. M. Ambient-Potential Composite Ewald Method for ab Initio Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulation. *J Chem Theory Comput* **12**, 2611-2632, doi:10.1021/acs.jctc.6b00198 (2016).

46 D.A. Case *et al.* AMBER 2015, University of California, San Francisco. (2015).
47 Frederiksen, J. K. & Piccirilli, J. A. Identification of catalytic metal ion ligands in ribozymes. *Methods* **49**, 148-166, doi:10.1016/j.ymeth.2009.07.005 (2009).
48 Mairhofer, E., Fuchs, E. & Micura, R. Facile synthesis of a 3-deazaadenosine phosphoramidite for RNA solid-phase synthesis. *Beilstein J Org Chem* **12**, 2556-2562, doi:10.3762/bjoc.12.250 (2016).
49 Erlacher, M. D. *et al.* Efficient ribosomal peptidyl transfer critically relies on the presence of the ribose 2'-OH at A2451 of 23S rRNA. *J Am Chem Soc* **128**, 4453-4459, doi:10.1021/ja0588454 (2006).
50 Montgomery, J. A., Shortnacy, A. T., Arnett, G. & Shannon, W. M. 2-Substituted derivatives of 9-alpha-D-arabinofuranosyladenine and 9-alpha-D-arabinofuranosyl-8-azaadenine. *J Med Chem* **20**, 401-404 (1977).
51 Minakawa, N. & Matsuda, A. Nucleosides and nucleotides. 116. Convenient syntheses of 3-deazaadenosine, 3-deazaguanosine, and 3-deazainosine via ring closure of 5-ethynyl-1-β-D-ribofuranosylimidazole-4-carboxamide or -carbonitrile. *Tetrahedron* **49**, 557-570, doi:10.1016/S0040-4020(01)86259-1 (1993).
52 Krenitsky, T. A. *et al.* Imidazo[4,5-c]pyridines (3-deazapurines) and their nucleosides as immunosuppressive and antiinflammatory agents. *J Med Chem* **29**, 138-143 (1986).
53 Seela, F., Münster, I., Lüchner, U. & Rosemeyer, H. 8-Azaadenosine and Its 2′-Deoxyribonucleoside: Synthesis and oligonucleotide base-pair stability. *Helv Chim Acta* **81**, 1139-1155, doi:doi:10.1002/hlca.19980810527 (1998).
54 Kovacheva, Y. S., Tzokov, S. B., Murray, I. A. & Grasby, J. A. The role of phosphate groups in the VS ribozyme-substrate interaction. *Nucleic Acids Res* **32**, 6240-6250, doi:10.1093/nar/gkh957 (2004).
55 Zamel, R. & Collins, R. A. Rearrangement of substrate secondary structure facilitates binding to the Neurospora VS ribozyme. *J Mol Biol* **324**, 903-915, doi:10.1016/S0022-2836(02)01151-8 (2002).
56 Lafontaine, D. A., Wilson, T. J., Zhao, Z. Y. & Lilley, D. M. J. Functional group requirements in the probable active site of the VS ribozyme. *J Mol Biol* **323**, 23-34, doi:10.1016/S0022-2836(02)00910-5 (2002).
57 Wilson, T. J., McLeod, A. C. & Lilley, D. M. J. A guanine nucleobase important for catalysis by the VS ribozyme. *Embo J* **26**, 2489-2500, doi:10.1038/sj.emboj.7601698 (2007).
58 Jencks, W. P. A primer for the Bema Hapothle. An empirical approach to the characterization of changing transition-state structures. *Chem Rev* **85**, 511-527, doi:10.1021/cr00070a001 (1985).
59 Anslyn, E. V. & Dougherty, D. A. *Modern physical organic chemistry*. (University Science, 2006).