HHS Public Access

Author manuscript

Methods Enzymol. Author manuscript; available in PMC 2016 March 04.

Published in final edited form as:

Methods Enzymol. 2015 ; 553: 215–234. doi:10.1016/bs.mie.2014.10.061.

Integrating molecular dynamics simulations with chemical probing experiments using SHAPE-FIT

Serdal Kirmizialtin[1,2,*], **Scott P. Hennelly**[1,2], **Alexander Schug**[3], **Jose N. Onuchic**[4,5], and **Karissa Y. Sanbonmatsu**[1,2,*]

[1] New Mexico Consortium, 4200 West Jemez Road, Suite 301, Los Alamos, New Mexico 87544

[2] Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, Los Alamos, MS K710, NM 87545

[3] Steinbuch Centre for Computing, Karlsruhe Institute of Technology, Germany

[4] Center for Theoretical Biological Physics and Department of Physics and Astronomy, Chemistry, and Biosciences, Rice University, Houston, Texas

[5] Department of Chemistry and Department of Biochemistry and Cell Biology, Rice University, Houston, Texas

Abstract

Integration and calibration of molecular dynamics simulations with experimental data remains a challenging endeavor. We have developed a novel method to integrate chemical probing experiments with molecular simulations of RNA molecules by using a native structure-based model. Selective 2'-hydroxyl acylation by primer extension (SHAPE) characterizes the mobility of each residue in the RNA. Our method, SHAPE-FIT, automatically optimizes the potential parameters of the forcefield according to measured reactivities from SHAPE. The optimized parameter set allows simulations of dynamics highly consistent with SHAPE probing experiments. Such atomistic simulations, thoroughly grounded in experiment, can open a new window on RNA structure-function relations.

1 Introduction

Molecular dynamics simulations enable studies of biomolecules in atomic resolution. Over the past few decades, the predictive capability of this method has improved significantly due to the advances in hardware technologies [1] and novel computational methods [2-4][5]. These advances invite the development of more accurate forcefields for biomolecular simulations. However, the development of highly accurate force field potential functions remains a challenge of molecular simulation. Many studies have been successful in producing dynamics consistent with NMR spectroscopy studies[6-8][9]. In addition to NMR studies the development of nucleotide resolution chemical probing assays in the RNA community presents a new source of experimental data that can be used to benchmark and improve molecular simulation force fields. [10][11]

*Corresponding authors: serdal@lanl.gov, kys@lanl.gov.

From a biochemical perspective, RNA has the advantage over proteins in being amenable to reverse transcription readout assays, yielding information at nucleotide resolution. These assays were used extensively in ribosome studies to determine the ribosome secondary structure, binding sites and conformational changes[12-14]. The development of in-line probing in the riboswitch community by Breaker and co-workers enabled readout of backbone mobility [10]. Selective 2'-hydroxyl acylation by primer extension (SHAPE) was developed by Weeks and co-workers [11]. This method is a rapid assay capable of backbone mobility readout at nucleotide resolution for a variety of environmental conditions (*e.g.*, magnesium titration). While NMR spectroscopy studies produce superb data sets monitoring RNA mobility, [15][16-25] SHAPE allows one to obtain mobility information in experiments over the course of a few days and also for very large RNA systems (Fig.1). This technique has opened the door to studies using a wide variety of environmental conditions, mutation sequences, and system sizes [26]. This technique is a powerful, widespread method in the RNA community that has produced important experimental datasets for comparison with molecular simulations. Weeks and co-workers have used SHAPE probing to generate three-dimensional structural models of the tRNA based on a three-bead model. Here, we investigated dynamics and calibrate dynamics with chemical probing reactivity measurements [27].

From the perspective of RNA molecular simulations, important advances have been made in recent years regarding force field parameters for all-atom explicit solvent molecular dynamics simulations[28, 29]. Few studies have compared RNA simulation with experiment in a detailed manner including a recent PreQ riboswitch study[17, 30-32] and studies of Small Angle X-Ray Scattering [33, 34]. While these studies are essential for improving forcefields, their high computational costs limits their sampling capability and therefore affect the accuracy of the entropic component of the free energy. Specifically, the functional dynamics of many RNA systems occurs on the time scale of hundreds of milliseconds to seconds [35, 36]. While large-scale simulations have produced millisecond simulations of small proteins[37] and microsecond simulations of large systems [38], current computing capabilities prevent all-atom explicit solvent molecular dynamics simulations from accessing the physiological time scales of 100 ms – 1 s.

To improve molecular simulation sampling, structure-based potentials have been used [39-44][45, 46]. This potential is defined by the crystallographic structure and has the advantage of preserving stereochemistry in the crystallographic structure while sampling hundreds of milliseconds. The method allows reproducibly folding and unfolding small to medium size protein and nucleic acid structures hence dramatically improving sampling and therefore the accuracy of the entropic component of the free energy. An additional advantage is that the potential is robust to changes in parameters, enabling calibration to experimental data while leaving the stereochemistry intact.

In this paper, we present SHAPE-FIT, a novel technique to automatically calibrate molecular simulations to RNA chemical probing experiments. We demonstrate this method on the *T. tengcongensis metF* SAM-I riboswitch aptamer domain (Fig. 2), a useful test system that has previously been studied using a variety of experimental and computational techniques. Our method is easily extendable to large RNA systems. The approach can also

be combined with explicit water all atom simulations. SHAPE data integrated with molecular simulations improve the forcefield and produce mechanistic studies of RNA systems grounded in experimental data.

2 Materials and Methods

2.1 Computation of SHAPE reactivity

SHAPE reactivity is inversely correlated with the base stability. The acylation reaction rate is higher if the nucleotide is mobile or easily accessible to the probing molecule (Fig.1). However the exact relationship between the SHAPE reactivity and stability is not yet known. To account for the realtionship between the base stability and SHAPE reactivity Deigan et al. [47] proposed a pseudo-energy function. Here, the stability of nucleotide *k* is given by the relation $\Delta G(k)=m\ln(a(k)+1)+n$ where $a(k)$ is the normalized SHAPE reactivity of nucleotide *k*, *m* and *n* are fitting parameters fitted to *m*=2.6kcal/mol and *n*=0.8kcal/mol using Ecoli 16S ribosomal RNA SHAPE data. Here in our study we will use the same pseudo-energy term and the coefficients.

To account for the base stability and backbone mobility we use the fluctuations in the angle between the 2'-hydroxyl group, the phosphate, and the adjacent O5' group (O2'-P-O5') (see Fig.1). This angle characterizes the mobility of the 2'-hydroxyl atom and is relevant to both SHAPE and in-line probing reaction geometry. Effective stability is characterized by the fluctuations as:

$$\Delta G_{SIM}(k) = k_B T \quad \ln \frac{\left\langle \theta_k^2 \right\rangle - \langle \theta_k \rangle^2}{\sum_{i=1}^{N} \left(\left\langle \theta_i^2 \right\rangle - \langle \theta_i \rangle^2 \right) / N} \quad (1)$$

Here, the stability of nucleotide *k* is computed from the fluctuation of this angle where $\langle \ldots \rangle$ represents the ensemble average computed from the time series of the simulation trajectory. The value is normalized with the average fluctuation of the RNA chain of length *N* nucleotides.

Combining the pseudo-energy term above with Eq. 1 we obtain the computed reactivity as:

$$a_{SIM}(k) = \exp\left(\frac{\Delta G_{SIM}(k) - n}{m} \right) - 1 \quad (2)$$

Note that neither the choice of the order parameter nor the formulation of the stability are unique.

2.2 Optimization of Potential Energy Function

To integrate SHAPE reactivity into the structure-based potential we optimize the native structure-based potential (SBM) [39-44][45, 46] by steepest-descent search in parameter space. The basic function of the previously defined SBM potential (Eq. 3) is a summation of harmonic potentials restraining the bond-lengths, bond-angles and dihedral angles to the native state structure that is given a priori by X-ray or NMR studies:

$$E=\sum_{bonds} K_r(r-r_0)^2+\sum_{angles} K_\theta(\theta-\theta_0)^2+\sum_{dihedrals} K_\phi^{(n)}\left[1-\cos\left(n\times(\phi-\phi_0)\right)\right]$$
$$+\sum_{impropers/planars}(\chi_1-\chi_0)^2+\sum_{i<j-3}\left\{\epsilon(i,j)\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-2\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]+\epsilon_2(i,j)\left(\frac{\sigma_{ij}}{r_{ij}}\right)\right)^{12}\Bigg\} \quad (3)$$

The interactions between atoms that are not bonded are represented by a repulsive term that accounts for excluded volume of the polymer and an attractive term is used to account for the native interactions dictated by the structure. Native interactions are defined as contact pairs with a simple cut-off distance of $r \le r_c = 4$ Å for nucleic acids and 6 Å for proteins. We emphasize that while more elaborate definitions for contacts exist [48], the exact contact definition will not significantly affect local dynamics in this study. This choice is more critical for large-scale conformational transitions such as those found in protein folding. The functional form for native contacts is a Lennard-Jones potential where the minimum is set to reproduce the native structure [44]. The barrier height of the dihedral potential $K_\phi^{(n)}$ and the strength of the non-bonding native interactions between atom pairs $\epsilon(i,j)=(\epsilon_i\epsilon_j)^{1/2}$ assume a uniform weight in such a way that the ratio of the total non-bonded native interaction to the sum of all torsional angle contributions, $\sum\epsilon(i,j)/\sum K_\phi^{(n)}$ is set to 4.

The results of simulations using the SBM potential (Eq. 3) depend on the choice of the set of parameters $\left\{\{K_r\},\{K_\theta\},\left\{K_\phi^{(n)}\right\},\{\chi_i\},\{\epsilon\},\{\epsilon_2\}\right\}$. Here the most sensitive parameters for dynamics and at the same time the least known parameters to us are torsional angle parameters and non-bonded native interaction terms. For that purpose we search for the parameter space for these set of parameters. Hence our sequence depended parameter space is defined as $\pi=\left\{\left\{K_\phi^{(n)}\right\},\{\epsilon\}\right\}$

Following [49], we search the parameter space to minimize the difference between the experimental and simulation result of the same observable. The target function is defined as the distance between the two data sets:

$$\Psi(\pi)=1/N\sum_{k=1}^{N}\left[\ln a_{SIM}{}^{\pi}(k)-\ln a_{EXP}(k)\right]^2 \quad (4)$$

Here $a_{SIM}{}^{\pi}(k)$ denotes the computed SHAPE reactivity of nucleotide *k* with respect to the parameter set π, while a_{EXP} (*k*) is the reactivity of the same nucleotide in experiment. We aim to minimize the target function. Typically achieving this goal can be difficult. In some cases, oversimplification of the physical interactions by the potential makes it difficult to find a parameter set that reproduces the experimental data. In other cases, the sought after parameter space becomes too large for an exhaustive search. Here, we tested the functional form of the SBM potential against SHAPE experiments. We reduce the paramter space by focusing only the torsional and native interactions. To effectively search this reduced space we introduced the steepest descent minimization algorithm by iteratively solving the following differential equation (see Eq. 5):

$$\pi_i = \pi_{i-1} - \alpha \nabla_\pi \Psi(\pi_{i-1}) \quad (5)$$

where, π_i is the parameter set in iteration *i* and α is a scalar that determines the strength of the target function. We note that the choice of α can be optimized to reduce the computational cost [50]. We use constant α at values of $0.01 \leq \alpha \leq 0.1$. The iterative procedure is terminated when both $\Psi_{(\pi_i)} < \gamma_1$ and $|\Psi(\pi_{i-1}) - \Psi(\pi_i)| < \gamma_2$ are satisfied, where γ_1 and γ_2 are preselected positive numbers chosen as 0.2 and 0.05 respectively in our study.

2.3 Molecular Simulations

We studied wild type sequences of the *T. tengcongensis met* riboswitch aptamer domain (residues 1-101) in two different solution conditions: (i) riboswitch in the presence of cognate ligand S-adenosylmethionine (SAM-I) and (ii) in the absence of the ligand. Because crystallographic structures of these sequences were not available, homology models were constructed using our previously published RNA homology modeling techniques used to model the ribosome [51-53]. Our model of the SAM-bound aptamer was derived directly from the crystallographic structure published by Batey and co-workers [54-56]. To obtain the atomic model of the SAM-I free configuration, we used the Kratky plots published in earlier studies [55].

We performed all-atom simulations of the SAM-I riboswitch aptamer by using structure-based potential [39, 42-44]. Simulations were performed using GROMACS 4.6 suit of programs [57]. We used leap-frog stochastic integrator with an inverse friction coefficient of 1ps with a reference temperature of 77K and a time step of 1fs throughout our simulations. VdW interactions are computed with twin range cut-offs with a cut-off distance of 12 Å. To compute the SHAPE reactivity we simulate each system and compute the ensemble average from the time trace of trajectories. Ideally statistical errors decrease as the length of the simulation increases; however the computational cost also increases. Here, we determine the optimal length of the simulation by computing the average effective stability (Eq. 1) of the RNA chain as a function of simulation length

$L, \quad C(L) = (1/LN) \int_0^L \sum_{k=1}^{N} \Delta G_k^{SIM}(t)\, dt$ and compute the standard error for the observable *C*. Our analysis indicated that $L=5.10^6$ steps gives less than %1 error in the *C* estimate. Therefore 5.10^6 steps of simulation length is used to obtain a statistically converged ensemble averages in our study. SHAPE reactivity from simulations is computed using Equations 1-2.

2.4 SHAPE probing experiments

We use SHAPE probing and in-line probing to follow changes in secondary and tertiary structure occurring upon ligand binding [26]. We study these changes for *T. tengcongensis metF* aptamers. In-line probing measures the ability of more mobile bases to spontaneously cleave [58]. SHAPE probing measures the ability of more mobile bases to react with the 1M7 reagent (Fig. 1). While the methods measure the same basic properties, we have found in-line probing to be less ambiguous in some structures. 1M7 is relatively insensitive to temperature and Mg^{2+} concentration.

2.5 Preparation of RNA systems

Synthetic DNA oligonucleotide (IDT, Coralville, IA) templates were PCR amplified for use as *in vitro* transcription templates. Transcription reactions were performed using Ampliscribe (EPICENTRE Biotechnologies, Madison, WI) high yield T7 RNA polymerase transcription kits as per instructions.

2.6 SHAPE chemical probing

The aptamer domain RNA was folded at a concentration of 5 nM in 1X HMK buffer (50 mM HEPES-KOH pH 8.0, 2 mM MgCl2, 100 mM KCl) and various concentrations of S-Adenosylmethionine (SAM). 30 μls of 60 mM 1-methyl-7-nitroisatoic anhydride (1M7) in DMSO was added to a 300 μl volume of RNA. The reaction proceeded for 5 minutes at 25 °C and was then precipitated by the addition of 0.1 volumes 3M Na:Acetate pH 6.5, 75 μg glycogen and 3 volumes EtOH. The recovered RNA was then subjected to reverse transcription analysis as described above. Capillary electrophoresis data was integrated by simultaneously fitting Gaussian curves to the whole trace using in-house scripts. The traces were normalized using residues whose reactivity does not change in response to SAM.

2.7 In-line chemical probing

All chemical probing reactions were performed a minimum of three times with and without S-adenosylmethionine. SAM (NEB, Ipswich, MA.) was added to a final concentration of 10 μM. In-line probing was performed as previously described[26] in in-line probing buffer (50 mM Tris-HCl pH 8.3, 20mM MgCl2, 100 mM KCl) with the following modifications: RNAs were unfolded by heating to 90 °C in water for 2 minutes and then crash cooled on ice for 2 minutes followed by addition of buffer with or without SAM. In-line probing reactions were performed at RNA concentrations of 0.3 μM for 40 hrs at 25 °C. Reactions were purified by precipitation with 3 volumes EtOH and 50 μg RNase free glycogen (Ambion) and then analyzed by capillary electrophoresis.

2.8 Analysis of chemical probing reactions

In-line probing and SHAPE probing reactions were analyzed by fluorescently labeling of the RNA on the 3’ terminus. The labeled RNA was then diluted 1-5 μl into 20 μl of formamide, depending on recovery and labeling efficiency, followed by heating to 90 °C for 2 minutes prior to loading on an ABI Prism 310 genetic analyzer (Applied Biosystems) capillary electrophoresis system equipped with a laser induced fluorescence detector. The reactions were electrokinetically injected at 12.5 kVolts for 30 seconds and run in POP-6 polymer (ABI) at 70 °C for 1 hour. Sequencing reactions were performed by transcribing the RNA constructs in the presences of a low level of α-phosphorothioate nucleotides followed by iodine cleavage and capillary electrophoresis.

3 Results

We performed molecular dynamics simulations of the SAM-I riboswitch aptamer wild type sequence in the presence of SAM (Figure 2a-b) with the unmodified potential energy function. Computed SHAPE reactivity is then used to calculate the distance function. We modify the potential as detailed in methods section to minimize the difference between the

experimental and computed SHAPE data. Our iterative procedure effectively decreases the target function in each iteration. Figure 3 shows the change in the target function as a function of the number of iterations. The procedure is stopped when a convergence in the distance function is achieved, as there was no more improvement in the target function.

Figure 4a displays the reactivities based on fluctuations from the simulations with the optimized potential and normalized reactivity based on in-line probing experiments. We achieve excellent agreement between simulation and experiment. Comparison of the reactivities with the unmodified potential (see Figure 5a) on the other hand significantly differs from the experiment, showing the success of our automatic procedure. We repeated the optimization procedure for the riboswitch in the absence of SAM-I. Our results are again in agreement (Fig. 4b) with the experiment and the data is quite different with the unmodified simulation results (see Figure 5b). Despite the excellent agreement there are still remaining questions: What is the reason for a nucleotide to show high reactivity while the other do not? What is the difference between liganded and unliganded states that leads to a significant change in the SHAPE spectra? What is the mode of dynamics in the region of high reactivities?

To accurately address the differences in the reactivities and to better understand the changes in mobility between the liganded and unliganded states we used the trajectories from our optimized potentials. Figure 6a-b shows a conformation with bound SAM-I ligand. We report here the most populated conformation in the ensemble color coded according to the SHAPE reactivities. The most populated structure is obtained by clustering the trajectory with RMSD as distance measure. The cluster with the highest weight in the ensemble is the most populated state. We choose the cluster center configuration to be representative for the respective state. We observe that SHAPE reactivity strongly correlates with riboswitch topology. Unpaired regions are, in general, significantly more reactive relative to base paired regions. Regions on the exterior of the molecule tend to be more reactive than those in the core. Simulations that incorporate SHAPE data show the dynamics in highly reactive regions. For example, P1, which is close to the SAM-I ligand remains formed during our simulations yet its high mobility is due to the rotational motion of the domain in solution (see Figure 6a-b). In contrast, P4, another highly reactive domain, goes back and forth towards P2 making contacts with it which results in exposing the buried residues to the solvent in the off state and hence increases the SHAPE reactivity.

A detailed comparison of the structure and dynamics in the *absence* of the ligand on the other hand depicts that the P1 domain is now unfolded (Figure 6c-d). The high reactivity in this region is due to the resulting high mobility of the backbone chain as well as due to unpaired nucleotides exposed to the solution. The P4 domain, which also shows high reactivity is now an extended helix. The structure that is dominated in our simulation with optimized potential is similar to that constructed by Lilley and co-workers [59]. Our structure also reveals a lack of pseudo-knot tertiary interactions in the absence of the ligand. The lack of pseudo-knot contacts leads to more extended conformations in the conformational ensemble increasing the radius of gyration by ~10%. This observation is also in accord with previous SAXS studies [55]. Our simulation results are consistent with

nucleotide analog interference modification studies (NAIM) and suggest that ligand promotes a global collapse of the aptamer domain by stabilizing tertiary contacts [55].

4 Discussions

We have developed a systematic method to integrate molecular dynamics simulations with chemical probing experiments. The method produces simulated RNA fluctuations consistent with experiment at each nucleotide. The method called SHAPE-FIT hence offers a solution to fine-tune an energy function toward better agreement with experiment with applications in RNA structure and dynamics. While SHAPE probing has been extremely successful in predicting secondary structures and evaluating the mobility of RNA molecules, the exact relation between SHAPE reactivity, nucleotide mobility, and solvent exposure is not entirely understood. The order parameter we choose to monitor has successfully recapitulated the experimental data. Improvements may result from a clearer understanding of the SHAPE reaction. This may be achieved by, for example, via quantum chemical calculations of acylation reaction or understanding of RNA polymerase transcription mechanism.

Here in our study we use the aptamer domain of the SAM-I riboswitch that is designed to fold to a single RNA native state in solution. This may not be the general case for the complete sequence that switches between on and off-states. It may be also difficult to justify a singe native state in many RNA systems as significant populations of alternative folds known to present in solution. Still, native interactions have been found to dominate RNA-folding[60]. A future direction would be to account for multiple native states in our potential. Multiple basin Go models have been successfully applied in protein folding [61, 62]. Application of this model to SHAPE-FIT will be a future direction. Extensions to the method will also include non-native interactions, explicit electrostatic interactions and additional calibration of our model with thermodynamic melting studies.

While our method achieved excellent agreement with experimental data, our use of a local minimization procedure leads to a single solution of the parameter set. One could easily imagine that multiple solutions are possible. These different solutions may lead to different interpretations of the same data. Modeling multiple simultaneous solutions is an opportunity for future research. Another direction is to add multiple experiments to the target function in addition to SHAPE such as SAXS, melting temperatures, etc. This would reduce the parameter space and result in a more robust potential across experimental data sets.

Overall, our method calibrates the structure-based potential to our experimental data, producing simulations consistent with chemical probing experiments. We note that our method is easily scalable to large systems. Thus, mechanistic studies of the folding and function of RNAs using these potentials will be consistent with SHAPE measurements, yielding conclusions grounded in experiment. Accordingly, predictions from such studies have the potential to be more precise. At the same time, these simulations would help interpret SHAPE measurements. Dynamics at each residue can be used for mechanistic interpretation of the SHAPE data as well as providing an atomic level interpretation of the reactivities based on fluctuations. The method has the potential to provide new mechanistic

insights into the interplay between structure and dynamics, tightly integrating experiment and atomistic simulation to obtain an atomic level picture of RNA function.

Acknowledgements

This work was supported by the Center for Theoretical Biological Physics sponsored by the National Science Foundation (Grant PHY-1427654), by the NSF (Grant MCB-1214457) and by the Human Frontiers Science Program. J.N.O. is a CPRIT Scholar in Cancer Research sponsored by the Cancer Prevention and Research Institute of Texas.

References

1. Shaw DE, et al. Anton, a special-purpose machine for molecular dynamics simulation. Commun. ACM. 2008; 51(7):91–97.
2. Kirmizialtin S, Elber R. Revisiting and Computing Reaction Coordinates with Directional Milestoning. Journal of Physical Chemistry A. 2011; 115(23):6137–6148.
3. Laio A, Parrinello M. Escaping free-energy minima. Proceedings of the National Academy of Sciences. 2002; 99(20):12562–12566.
4. Sugita Y, Okomoto Y. Replica-exchange molecular dynamics method for protein folding. Chem Phys. Lett. 1999; 314(1-2):141–151.
5. Adcock SA, McCammon JA. Chem Rev. 2006; 106(5):1589–1615. [PubMed: 16683746] Molecular dynamics: Survey of methods for simulating the activity of proteins. Chem. Rev. 2006; 106(5): 1589–1615. [PubMed: 16683746]
6. Lipari G, Szabo A, Levy RM. Protein dynamics and NMR relaxation: comparison of simulations with experiment. Nature. 1982; 300(5888):197–198.
7. Maragakis P, et al. Microsecond molecular dynamics simulation shows effect of slow loop dynamics on backbone amide order parameters of proteins. The Journal of Physical Chemistry B. 2008; 112(19):6155–6158. [PubMed: 18311962]
8. Zagrovic B, Gunsteren F.v.W. Comparing Atomistic Simulation Data With the NMR Experiment: How Much Can NOEs Actually Tell Us? Proteins. 2006; 63:210–218. [PubMed: 16425239]
9. Lange OF, et al. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. Science. 2008; 320(5882):1471–1475. [PubMed: 18556554]
10. Soukup GA, Breaker RR. Relationship between internucleotide linkage geometry and the stability of RNA. RNA. 1999; 5(10):1308–25. [PubMed: 10573122]
11. Merino EJ, et al. RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). J Am Chem Soc. 2005; 127(12):4223–31. [PubMed: 15783204]
12. Woese CR, et al. Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence. Nucleic acids research. 1980; 8(10):2275–93. [PubMed: 6159576]
13. Moazed D, Noller HF. Transfer RNA shields specific nucleotides in 16S ribosomal RNA from attack by chemical probes. Cell. 1986; 47(6):985–94. [PubMed: 2430725]
14. Moazed D, Noller HF. Intermediate states in the movement of transfer RNA in the ribosome. Nature. 1989; 342(6246):142–8. [PubMed: 2682263]
15. Fourmy D, et al. Structure of the A site of Escherichia coli 16S ribosomal RNA complexed with an aminoglycoside antibiotic. Science. 1996; 274(5291):1367–1371. [PubMed: 8910275]
16. Blanchard SC, Puglisi JD. Solution structure of the A loop of 23S ribosomal RNA. Proc Natl Acad Sci U S A. 2001; 98(7):3720–5. [PubMed: 11259644]
17. Eichhorn CD, et al. Unraveling the structural complexity in a single-stranded RNA tail: implications for efficient ligand binding in the prequeuosine riboswitch. Nucleic acids research. 2012; 40(3):1345–55. [PubMed: 22009676]
18. Chen B, et al. Multiple conformations of SAM-II riboswitch detected with SAXS and NMR spectroscopy. Nucleic acids research. 2011

19. Hall KB. RNA in motion. Curr Opin Chem Biol. 2008; 12(6):612–8. [PubMed: 18957331]
20. Gherghe CM, et al. Strong correlation between SHAPE chemistry and the generalized NMR order parameter (S2) in RNA. J Am Chem Soc. 2008; 130(37):12244–5. [PubMed: 18710236]
21. Davis JH, et al. Role of metal ions in the tetraloop-receptor complex as analyzed by NMR. RNA. 2007; 13(1):76–86. [PubMed: 17119098]
22. Zhang Q, et al. Comparison of solution and crystal structures of preQ1 riboswitch reveals calcium-induced changes in conformation and dynamics. Journal of the American Chemical Society. 2011; 133(14):5190–3. [PubMed: 21410253]
23. Proctor DJ, et al. Folding thermodynamics and kinetics of YNMG RNA hairpins: specific incorporation of 8-bromoguanosine leads to stabilization by enhancement of the folding rate. Biochemistry. 2004; 43(44):14004–14. [PubMed: 15518549]
24. Clore G, Kuszewski J. Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. JOURNAL OF THE AMERICAN CHEMICAL SOCIETY. 2003; 125(6):1518–1525. [PubMed: 12568611]
25. Showalter SA, Hall KB. A functional role for correlated motion in the N-terminal RNA-binding domain of human U1A protein. J Mol Biol. 2002; 322(3):533–42. [PubMed: 12225747]
26. Hennelly SP, Sanbonmatsu KY. Tertiary contacts control switching of the SAM-I riboswitch. Nucleic Acids Res. 2011; 39(6):2416–31. [PubMed: 21097777]
27. Gherghe CM, et al. Native-like RNA tertiary structures using a sequence-encoded cleavage agent and refinement by discrete molecular dynamics. J Am Chem Soc. 2009; 131(7):2541–6. [PubMed: 19193004]
28. Hart K, et al. Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. Journal of Chemical Theory and Computation. 2012; 8(1):348–362. [PubMed: 22368531]
29. Zgarbova M, et al. Toward Improved Description of DNA Backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. Journal of Chemical Theory and Computation. 2013; 9(5):2339–2354. [PubMed: 24058302]
30. Feng J, Walter NG, Brooks CL 3rd. Cooperative and directional folding of the preQ1 riboswitch aptamer domain. Journal of the American Chemical Society. 2011; 133(12):4196–9. [PubMed: 21375305]
31. Sarkar K, et al. Fast folding of an RNA tetraloop on a rugged energy landscape detected by a stacking-sensitive probe. Biophysical journal. 2009; 97(5):1418–27. [PubMed: 19720030]
32. Sarkar K, Nguyen DA, Gruebele M. Loop and stem dynamics during RNA hairpin folding and unfolding. RNA. 2010; 16(12):2427–34. [PubMed: 20962040]
33. Kirmizialtin S, et al. RNA and Its Ionic Cloud: Solution Scattering Experiments and Atomically Detailed Simulations. Biophysical Journal. 2012; 102(4):819–828. [PubMed: 22385853]
34. Meisburger SP, et al. Polyelectrolyte Properties of Single Stranded DNA Measured Using SAXS and Single-Molecule FRET: Beyond the Wormlike Chain Model. Biopolymers. 2013; 99(12):1032–1045. [PubMed: 23606337]
35. Blanchard SC. Single-molecule observations of ribosome function. Curr Opin Struct Biol. 2009; 19(1):103–9. [PubMed: 19223173]
36. Al-Hashimi HM, Walter NG. RNA dynamics: it is about time. Current opinion in structural biology. 2008; 18(3):321–9. [PubMed: 18547802]
37. Shaw DE, et al. Atomic-level characterization of the structural dynamics of proteins. Science. 2010; 330(6002):341–6. [PubMed: 20947758]
38. Whitford PC, et al. Connecting the kinetics and energy landscape of tRNA translocation on the ribosome. PLoS Comput Biol. 2013 This manuscript has received favorable reviews by the PLoS Computation Biology reviewers.
39. Ratje AH, et al. Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites. Nature. 2010; 468(7324):713–6. [PubMed: 21124459]
40. Whitford PC, Onuchic JN, Sanbonmatsu KY. Connecting energy landscapes with experimental rates for aminoacyl-tRNA accommodation in the ribosome. J Am Chem Soc. 2010; 132(38):13170–1. [PubMed: 20806913]

41. Noel JK, et al. SMOG@ctbp: simplified deployment of structure-based models in GROMACS. Nucleic Acids Res. 2010; 38:W657–61. Web Server issue. [PubMed: 20525782]
42. Whitford PC, et al. Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. RNA. 2010; 16(6):1196–204. [PubMed: 20427512]
43. Whitford PC, et al. Nonlocal helix formation is key to understanding S-adenosylmethionine-1 riboswitch function. Biophys J. 2009; 96(2):L7–9. [PubMed: 19167285]
44. Whitford PC, et al. An all-atom structure-based potential for proteins: bridging minimal models with all-atom empirical forcefields. Proteins. 2009; 75(2):430–41. [PubMed: 18837035]
45. Lutz B, et al. Differences between cotranscriptional and free riboswitch folding. Nucleic Acids Research. 2014; 42(4):2687–2696. [PubMed: 24275497]
46. Lutz B, et al. eSBMTools 1.0: enhanced native structure-based modeling tools. Bioinformatics. 2013; 29(21):2795–2796. [PubMed: 24021379]
47. Deigan KE, et al. Accurate SHAPE-directed RNA structure determination. Proc Natl Acad Sci U S A. 2009; 106(1):97–102. [PubMed: 19109441]
48. Noel JK, Whitford PC, Onuchic JN. The Shadow Map: A General Contact Definition for Capturing the Dynamics of Biomolecular Folding and Function. Journal of Physical Chemistry B. 2012; 116(29):8692–8702.
49. Di Pierro M, Elber R. Automated Optimization of Potential Parameters. Journal of Chemical Theory and Computation. 2013; 9(8):3311–3320. [PubMed: 24015115]
50. Kirk, DE. Dover Books on Electrical Engineering; 2004. Optimal Control Theory: An Introduction.
51. Tung CS, Joseph S, Sanbonmatsu KY. All-atom homology model of the Escherichia coli 30S ribosomal subunit. Nat Struct Biol. 2002; 9(10):750–5. [PubMed: 12244297]
52. Tung CS, Sanbonmatsu KY. Atomic model of the Thermus thermophilus 70S ribosome developed in silico. Biophys J. 2004; 87(4):2714–22. [PubMed: 15454463]
53. Korostelev A, et al. Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. Cell. 2006; 126(6):1065–77. [PubMed: 16962654]
54. Montange RK, Batey RT. Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. Nature. 2006; 441(7097):1172–5. [PubMed: 16810258]
55. Stoddard CD, et al. Free state conformational sampling of the SAM-I riboswitch aptamer domain. Structure. 2010; 18(7):787–97. [PubMed: 20637415]
56. Lu C, et al. SAM recognition and conformational switching mechanism in the Bacillus subtilis yitJ S box/SAM-I riboswitch. J Mol Biol. 2010; 404(5):803–18. [PubMed: 20951706]
57. Hess B, et al. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. Journal of Chemical Theory and Computation. 2008; 4(3):435–447. [PubMed: 26620784]
58. Regulski EE, Breaker RR. In-line probing analysis of riboswitches. Methods Mol Biol. 2008; 419:53–67. [PubMed: 18369975]
59. Lemay JF, et al. Folding of the adenine riboswitch. Chem Biol. 2006; 13(8):857–68. [PubMed: 16931335]
60. Behrouzi R, et al. Cooperative Tertiary Interaction Network Guides RNA Folding. Cell. 2012; 149(2):348–357. [PubMed: 22500801]
61. Okazaki K, et al. Multiple-basin energy landscapes for large-amplitude conformational motions of proteins: Structure-based molecular dynamics simulations. Proceedings of the National Academy of Sciences of the United States of America. 2006; 103(32):11844–9. [PubMed: 16877541]
62. Schug A, et al. Mutations as trapdoors to two competing native conformations of the Rop-dimer. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104(45): 17674–17679. [PubMed: 17968016]

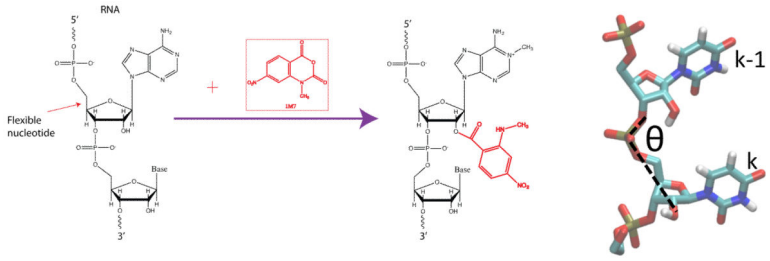

Figure 1.

Detecting nucleotide mobility experimentally and computationally. (a) Schematic for the acylation reaction and the 2'-hydroxyl group of an RNA nucleotide with the SHAPE reagent (NMIA). The acylation reaction is more probable when backbone is mobile and base is unpaired (b) Mobility of the 2'-hydroxyl group is characterized in molecular dynamics simulations using the RMS fluctuations of the angle between the 2'-hydroxyl group, phosphate group, and the 5' oxygen.

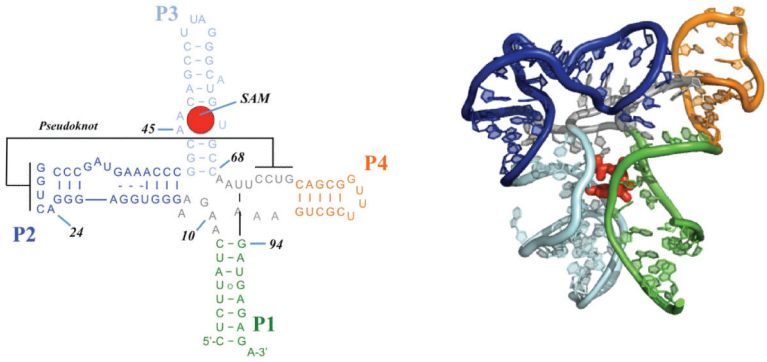

Figure 2.

The *T. tengcongensis metF* SAM-I riboswitch aptamer domain in the off-state. (a) Secondary structure of the aptamer domain with different colors representing secondary structure elements. (b) Tertiary structure of the sequence in the presence of S-adenosylmethionine (SAM) ligand.

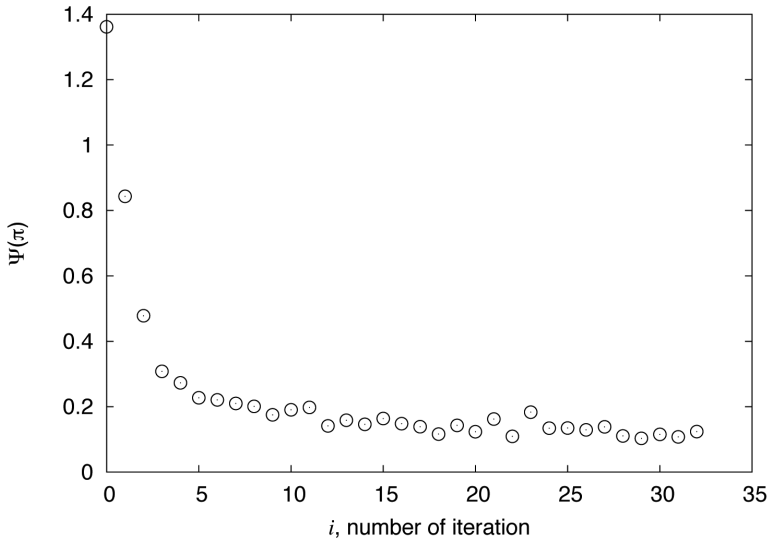

Figure 3.

Evolution of the distance function $\Psi(\Pi)$, quantifying agreement between simulated and measured SHAPE reactivities, as a function of the iteration number.

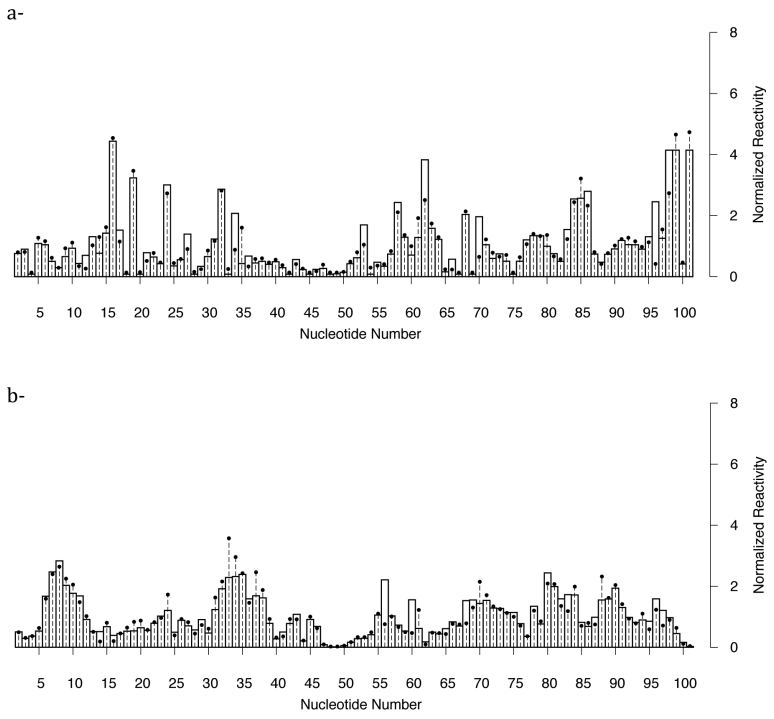

Figure 4.

Comparison between experiment and simulation of SHAPE probing data of SAM-I riboswitch aptamer domain. Dashed horizontal lines with points are normalized activity from SHAPE reactivity measurements. Bars are activities computed from structure-based molecular dynamics simulations after optimization of the potential. Close agreement is achieved *after* the optimization of the potential (a)- in the presence of SAM and (b)- in the absence of it.

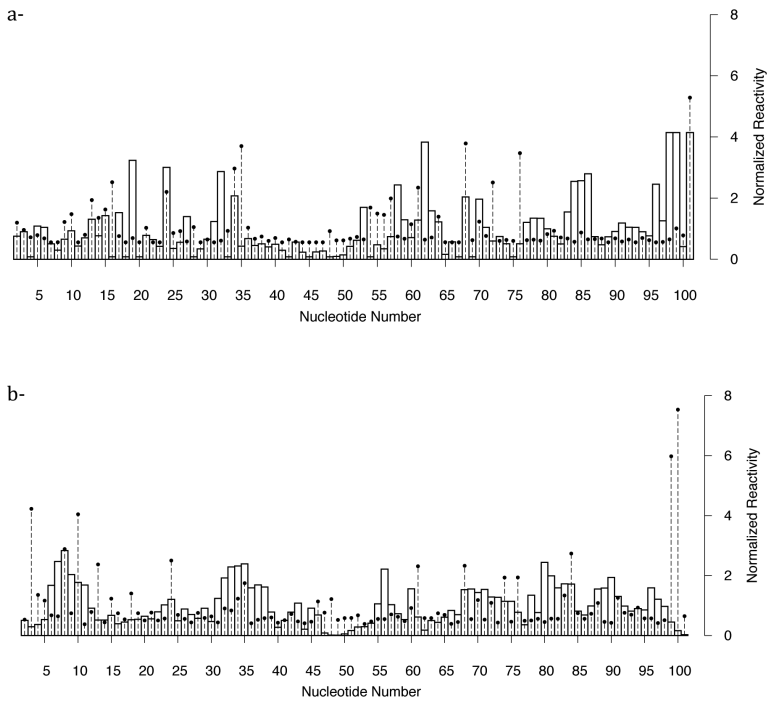

Figure 5.

Comparison between experiment and simulation of SHAPE probing data of SAM-I riboswitch aptamer domain. Dashed horizontal lines with points are normalized activity from SHAPE reactivity measurements. Bars are activities computed from structure-based molecular dynamics simulations *before* optimization of the potential. (a)- in the presence of SAM and (b)- in the absence of it.

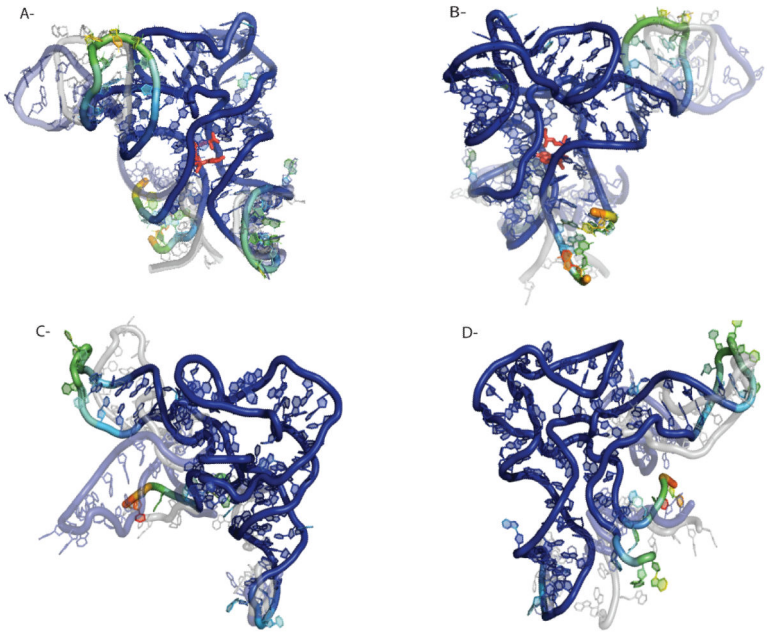

Figure 6.

Comparison of SHAPE reactivities of SAM-I riboswitch aptamer domain with and without SAM-I. Residues are colored according to their SHAPE reactivity, blue is unreactive and red is highly reactive. The type of motion in reactive regions are depicted by overlying conformations (partially transparent). (a) Shows the dynamics of riboswitch with bound ligand and (b) is the same molecule rotated by 180 degrees. (c)-(d) show the dynamics of the riboswitch unliganded.