Exploring RNA structure and dynamics through enhanced sampling simulations

Exploring RNA structure and dynamics through enhanced sampling simulations

Vojtěch Mlýnský, Giovanni Bussi*

Scuola Internazionale Superiore di Studi Avanzati, SISSA,

via Bonomea 265, 34136 Trieste, Italy

Abstract

RNA function is intimately related to its structural dynamics. Molecular dynamics simulations are useful for exploring biomolecular flexibility but are severely limited by the accessible timescale. Enhanced sampling methods allow this timescale to be effectively extended in order to probe biologically-relevant conformational changes and chemical reactions. Here, we review the role of enhanced sampling techniques in the study of RNA systems. We discuss the challenges and promises associated with the application of these methods to force-field validation, exploration of conformational landscapes and ion/ligand-RNA interactions, as well as catalytic pathways. Important technical aspects of these methods, such as the choice of the biased collective variables and the analysis of multi-replica simulations, are examined in detail. Finally, a perspective on the role of these methods in the characterization of RNA dynamics is provided.

Introduction

Ribonucleic acids (RNA) play fundamental roles in the cell, ranging from catalysis [1] to control of gene expression [2]. RNA function is often linked to its three-dimensional structure, typically obtained using X-ray crystallography [3] or nuclear magnetic resonance (NMR) [4]. However, RNA molecules are not static and might exhibit a multitude of accessible functional structures in a narrow energetic range. Many examples have been reported, ranging from flexible RNA motifs [5] to excited states [6] and, in the extreme case, riboswitches [7]. In addition, RNA catalysis is initiated by a pre-organization of the active site, and transition states (TSs) need to be stabilized by neighboring groups [8]. The mentioned events might require timescales ranging from microseconds to seconds or hours to be observed in an experimental setup.

Molecular dynamics (MD) simulations, both using empirical force fields and quantum mechanics/molecular mechanics approaches (QM/MM), are in principle a powerful tool to access RNA flexibility. However, they are limited to timescales of a few microseconds (for empirical force fields) or a few hundreds of picoseconds (for QM/MM-MD approaches). In order to address the conformational transitions and chemical reactions mentioned above, they should be complemented with enhanced sampling methods. Even dedicated machines capable to perform millisecond-long classical MD need enhanced sampling methods in order to access biologically relevant timescales [9].

*Corresponding author

Email addresses: vojtech.mlynsky@sissa.it (Vojtěch Mlýnský), bussi@sissa.it (Giovanni Bussi)

We here present a survey on the recent applications of enhanced sampling techniques to atomistic MD simulations of RNA systems. Many recent reviews discuss in detail enhanced sampling methods [10, 11] and MD simulations of RNA [12, 13, 14, 15, 16]. We opted for proceeding in an orthogonal direction, highlighting which enhanced sampling methods have been recently applied to RNA systems and, at the same time, underlying which aspects of RNA dynamics can benefit of enhanced sampling methods. In order to take a picture of the current state of the art for the application of these techniques to RNA systems, we deliberately limited the survey to the past two years. In addition, we only considered atomistic explicit solvent simulations where hydrogen atoms and water molecules are explicitly included.

Basic Assumptions. A fundamental issue in MD simulations is the choice of an appropriate model to compute the interatomic forces. This is done using empirical force fields (see [14, 15, 16] and references therein) and/or QM methods. In the latter case, a compromise between accuracy and computational cost should be found, choosing between fast but approximate semi-empirical (SE) methods and more accurate but computationally demanding density functional theory (DFT) methods (see [12, 13] and references therein).

Enhanced Sampling. A central idea of all enhanced sampling methods is to alter the system's dynamics to characterize specific events that would otherwise require significantly longer simulation timescales. Generally speaking, this can be done in two ways (Figure 1): (i) by changing the probability distribution of a limited number of selected degrees of freedom, so called collective variables (CVs), deemed to be important for the investigated conformational transition; (ii) by acting on the total energy or, equivalently, on the temperature of the system. Prototypical methods for these two approaches are (i) umbrella sampling (US) and (ii) temperature replica-exchange molecular dynamics (T-REMD), respectively (see [10, 11] and references therein). Methods based on CVs are extremely efficient when the chosen CVs identify correctly the kinetically relevant states of the system, including metastable and TSs. Methods based on tempering are more computationally demanding, but usually require less *a priori* information. CV-based and tempering methods can be combined and methods at the boundary between these two classes have been proposed as well. We note that the usage of replicas is not necessarily a distinctive trait of tempering methods. US can indeed be performed in a replica-exchange scheme, as it is discussed below. Conversely, temperature methods using a single simulation are used as well. Alchemical approaches such as the free-energy perturbation method, where transitions are enforced through a non-physical path involving changes in particle number and/or identity [10], can be considered as a special case of CV-based methods. Other approaches using unbiased simulations to analyze and reconstruct long-time kinetics, as well as non-dynamical methods where energies of individual structures are calculated and compared, are not considered here.

Applications of Enhanced Sampling Methods to RNA Systems

Table 1 reports an extensive list of publications in the last two years where enhanced sampling methods were applied to RNA systems. We arbitrarily classified them in groups according to the presented application, although some of them could be assigned to more than one group.

Refinements and Validations of Force Fields. Historically, force fields have been validated by analyzing plain MD simulations starting from the native structure. Taking advantage of enhanced sampling techniques, small RNA motifs can be sampled until statistical convergence is reached in

Enhanced Sampling Method	CVs	RNA Systems	QM/MM	Total Timescale (μs)	Ref.
Refinements and Validations of Force Fields					
M-REMD	-	UUCG-TL (10,14-mers), CCCC, GACC	MM	∼7663	[17]
M-REMD	-	CCCC, GACC	MM	∼1382.4	[18]
H-REMD	-	single strands (6-mers), duplexes (12-mers)	MM	1.68	[19]
T-REMD	-	AAAA, CAAU, CCCC, GACC, UUUU	MM	264	[20]
METAD, T-REMD, RECT	α,β,γ,ϵ,ζ,Z_x,χ, distance (COM)	CC, AA, CA, AC, GACC, CCCC, AAAA	MM	∼206	[21]
METAD, T-REMD, REST2	H-bonds, RMSD	GAGA-TL (8,10-mers)	MM	966	[22]
T-REMD+METAD	ϵRMSD	GAGA-TL, UUCG-TL (6,8-mers)	MM	96	[23]
RECT	α,β,γ,ϵ,ζ,Z_x,Z_y,χ, distance (COM)	A, C, AA, AC, CA, CC	MM	∼35	[24]
RECT, T-REMD+METAD	α,β,γ,ϵ,ζ,Z_x,Z_y,χ, distance (COM), ϵRMSD	GACC, CCCC, AAAA, CAAA, UUCG-TL (8-mer)	MM	∼504	[25]
US	α,β,γ,ϵ,ζ,χ	16 dinucleotides	MM	16.128	[26]
Conformational Landscapes					
RECT, H-REMD, T-REMD	α,β,γ,δ,ϵ,ζ,Z_x,Z_y,χ, distance (COM)	GACC	MM	∼14.4	[27]
METAD	H-bonds, RGyr, RMSD, χ	GAGA-TL, UUCG-TL (10-mers)	MM	4.44	[28]
T-REMD	-	SAM-II riboswitch	MM	6	[29]
T-REMD, SMD	distance (COM)	(cgauUCUaugc) duplex (22-mer)	MM	∼5	[30]
T-REMD	-	SVL loop (17-mer)	MM	57.6	[31]
A-REMD	Z_x, Z_y, χ	U nucleoside	QM/MM	∼0.1	[32]
BE-METAD	H-bonds, RGyr, energy	gene32 mRNA pseudoknot (32-mer)	MM	3	[33]
US	distance (COM)	*add* riboswitch	MM	1.177	[34]
US	χ	U and 2-thio-U nucleosides	MM	0.144	[35]
US	χ, pseudodihedral (COM)	A, G, U, and C nucleosides, duplex with CUG (18-mer)	MM	6.016	[36]
US+pseudo-spring method	distance (COM)	duplex (32-mer)	MM	∼0.3	[37]
T-REMD	-	GCAA -TL (8-mer)	MM	448	[38]
T-REMD	-	pT181 RNA hairpins (48-mers)	MM	17.16	[39]
RAM	H-bonds, distance (COM)	TAR (29-mer) in RNA:peptide complex	MM	0.8	[40]
REMD+US	distance, RMSD	U-singlestrand (5-mer) in RNA:protein complex	MM	3.6	[41]
US	distance (COM)	GTPase center of rRNA	MM	4.494	[42]
RAM	α,β,δ,ϵ,ζ, H-bonds	UUCG-TL (14-mer)	MM	1.04	[43]
T-REMD	-	theophylline-binding aptamer (33-mer)	MM	1.6	[44]
GaMD, TMD	RMSD	CRISPR-Cas9 RNA complex	MM	∼12	[45]
METAD	distance (COM), stacking	PNAs (6-mers), PNA:RNA duplex (12-mer)	MM	∼1.2	[46]
T-REMD+METAD	H-bonds, RGyr	polio viral RNA hairpin (22-mer)	MM	16	[47]
US	pseudodihedral (COM)	hairpin from group-II intron (35-mer)	MM	1.008	[48]
METAD	χ, pseudodihedral (COM)	duplexes with A-A mismatches (18-mers)	MM	∼0.6	[49]
T-REMD	-	gcGCAAgc-TL (8-mer)	MM	356	[50]
Ion Interactions and Ion/ligand Induced Conformational Changes					
METAD	distance (COM), hydrophobic contacts, H-bonds	duplex with A-A mismatches (20-mer)	MM	1.4	[51]
US+TI	distance	mononucleotides, hammerhead ribozyme	MM	∼3.5	[52]
SMD	distance	preQ$_1$-III riboswitch	MM	0.1	[53]
BE-METAD	distance, coordination	nucleosides, GpG dinucleotide, GC duplex (8-mer)	MM	82	[54]
GCMC-MD	distance, coordination	BWYV pseudoknot, VS ribozyme, 23S rRNA, Mg^{2+} riboswitch	MM	4	[55]
US	distance (COM)	GTPase center of rRNA	MM	20.488	[56]
TI	distance	guanine riboswitch	MM	0.576	[57]
T-REMD, METAD	distance (COM)	siRNA duplex (42-mer)	MM	28.8	[58]
Reactivity and Catalysis					
A-REMD	distance	HDV ribozyme	QM/MM	∼0.235	[59]
US+string method	distance	*glmS* riboswitch	QM/MM	0.00225, ∼0.00015	[60]
T-REMD	-	hammerhead ribozyme	MM	5	[61]
US+string method	distance, angle	HDV ribozyme	MM, QM/MM	0.0465, ∼0.00019	[62]
US, TI	distance	twister ribozyme	MM	0.525	[63]
US+string method	distance	*glmS* riboswitch	QM/MM	∼0.000098	[64]
TI	distance	group-II introns	QM/MM	∼0.0001	[65]
METAD	distance	GAAA-TL, UUCG-TL (10-mers), GC duplex (16-mer)	QM/MM	0.72	[66]
US+string method	distance	*glmS* riboswitch	QM/MM	∼0.0072	[67]
T-REMD	-	hairpin ribozyme	MM	25.6	[68]
US, TI	distance	hammerhead ribozyme	MM, QM/MM	∼0.33, ∼0.0004	[69]

Table 1: Summary of recent application of enhanced sampling methods to RNA systems. Explanation of the acronyms and additional information are included in Table S1 in Supporting Information. The online version of this table will be kept up to date at `https://github.com/srnas/rna-enhanced-sampling`.

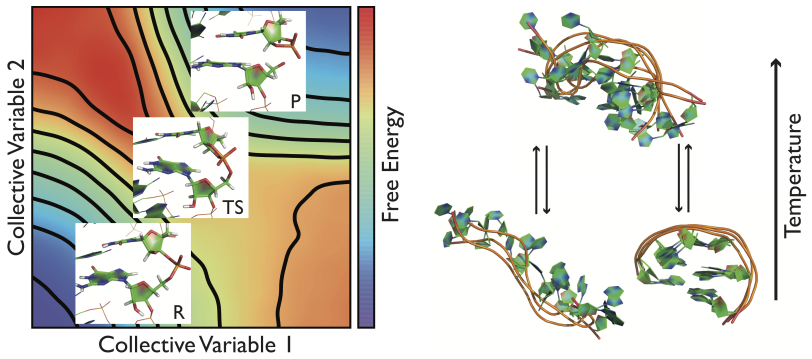

Figure 1: Scheme representing distinctive features of enhanced sampling methods of different classes. Methods based on collective variables (CVs), where a small number of CVs capable to describe the important free-energy minima (*e.g.*, reactant (R), transition (TS), and product (P) states), are identified *a priori* (panel on the left). These variables are then exploited to enhance sampling. Methods based on tempering, where the temperature of the system is repeatedly increased and decreased (panel on the right). The increased conformational dynamics at high temperature allows more conformations to be explored also at low temperature.

order to expose all the potential force field limitations. In several studies listed in this category, the population of the native structure (or the relative population of a number of structures) was computed and compared with experiments [17, 18, 22, 23, 25, 26]. The direct comparison with raw experimental data is more suitable for small unstructured RNAs [21, 23, 24]. Ref. [19] encouragingly predicted the effect of simple mutations on the dimerization energy of a duplex in reasonable agreement with experimental data. Refs. [21, 24] used enhanced sampling methods during the construction of the force field rather than just to validate it. Some works of this category [17, 22, 23] suggest that none of the available force fields yet allows predictive folding of RNA hairpin loops or larger systems to be performed reliably. One may wonder that the requirement to obtain a stable folded structure with all native hydrogen bonds simultaneously formed might be too restrictive for some motifs. However, a direct comparison with NMR data in ref. [23] emphasized that, at least for one of the investigated tetraloops, using a too loose criterion would lead to structures that are not compatible with solution experiments being reported as correct.

Conformational Landscapes. This is the wider group considered and includes papers discussing the conformational landscape of systems ranging from individual nucleosides [32, 35, 36], small loops [27, 28, 38, 50], duplexes [30, 46, 49], stem-loops [31, 36, 37, 39, 40, 43, 47, 48], and pseudoknots [33], up to larger aptamers [42, 44], riboswitches [29, 34], RNA:peptide [40], and RNA:protein complexes [41, 45]. None of the applications to large RNA systems is designed to sample extensively the conformational space, which would be extremely expensive and probably counterproductive considering the force-field limitations discussed above. However, local excitations can provide a wealth of information that can be compared with experiment. Interestingly, in some of these works the simulation is complemented with experimental data in order to improve the accuracy of the resulting structural ensemble [40, 43].

Ion Interactions and Ion/ligand Induced Conformational Changes. Divalent cations are crucial for RNA folding and catalysis. However, the typical timescale for direct binding of divalent cations on RNA is on the order of the millisecond, and should thus be simulated using enhanced sampling

methods. Some of the works of this section focus indeed on interactions between Mg^{2+} ions and individual nucleosides or RNA structural motifs [52, 54, 55, 56]. Other studies address structural reconformations induced by the presence of a (usually) small rigid molecule (ligand) in a binding pocket and the related problem of computing the affinity between ligands and RNA motifs [51, 53, 57, 58].

Reactivity and Catalysis. Small self-cleaving ribozymes are interesting model systems for probing general principles of RNA catalysis. However, the rugged free-energy landscapes of phosphodiester cleavage reactions present a significant obstacle in a consistent identification of feasible reaction pathways in catalytic systems as well as in general 'noncatalytic' RNA motifs [66]. Application of enhanced sampling methods helped to characterize reaction mechanisms in hammerhead [69], hepatitis delta virus (HDV) [59, 62], twister [63], group-II intron ribozymes [65] and glucosamine-6-phosphate synthase (*glmS*) riboswitch [60, 64, 67], where charged nucleobases, Mg^{2+} ions, water molecules and/or other ligands are involved. Other works reported how changes in external conditions, *i.e.*, interaction with a mineral surface [61] or high pressure [68], affect the pre-organization of the active site required for catalysis.

General Considerations. The simulations performed with empirical force fields are typically sampling timescales ranging from approximately 1 μs to a few tens of μs, which corresponds to a few days up to a few weeks of computational time considering current hardware and software. Remarkable efforts have also been reported, such as the extensive benchmark of force fields performed by Bergonzo *et al.* [17, 18] and by Kuhrova *et al.* [22], which reached or even surpassed the millisecond timescale of aggregated time. QM methods are considerably slower and accurate DFT methods are typically used on the ps time scale. An exception is represented by SE potentials that can probe the ns timescale. The AMBER force field is by far the most adopted empirical force field. In some works, alternative modifications were tested, including modified non-bonded parameters [17, 18, 22, 25, 38, 39] or dihedral reparametrizations [24, 26, 39, 53]. A limited number of applications used the CHARMM force field, either for very short simulations [35] or for simulations where RNA backbone was constrained [55]. The CHARMM force field has been already reported to lead to unstable trajectories in plain MD simulations. However, there is a significant chance to observe spurious states with any of the current force fields when applying enhanced sampling methods. The quality of the force field in the MM part of QM/MM simulations is probably less critical due to the short simulated timescales and to the fact that usually enhanced sampling methods are employed to accelerate events in the QM part. DFT functionals are mostly used for QM calculations because they offer sufficient accuracy for estimating reaction barriers. Some works benefited from the usage of efficient SE methods that allow extensive sampling at the expense of some tuning and external corrections [32, 59, 66].

Enhanced Sampling Methods. A popular enhanced sampling method in this community is T-REMD, probably also thanks to its wide availability and ease of use [20, 21, 22, 23, 25, 27, 29, 30, 31, 38, 39, 44, 47, 50, 61, 68]. Other tempering methods are also notable. In particular, multi-dimensional REMD schemes (M-REMD) where dihedral torsional potentials are in addition scaled were successfully used for sampling unstructured oligonucleotides [17, 18], and replica exchange with solute tempering was used to fold a tetraloop [22]. Special caution should be taken into account when analyzing replica-exchange simulations, since continuous trajectories should display transitions between the relevant substates (Figure 2) [20, 22, 27]. Many other applications take

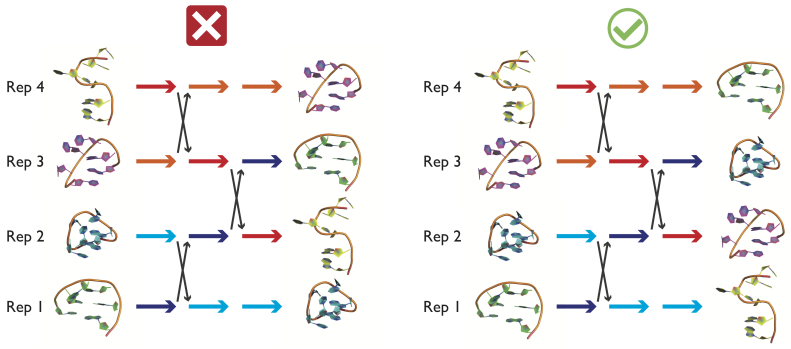

Figure 2: Scheme representing two possible behaviors during replica-exchange simulations. The four different molecular structures schematically represent four different conformations. Horizontal axis represents time, and vertical axis the replica index. The initial and final structures are different in each replica within both panels. However, conformational changes illustrated in the left panel are only due to replica exchanges (black arrows), and the continuous trajectories obtained following the arrows with respective colors are stuck in a single conformation. In this pathological case, the resulting populations would be affected by a significant systematic error. This analysis applies to all replica-based methods, including both methods based on CVs (such as replica-exchange US) and methods based on tempering (such as T-REMD).

advantage of CV-based methods. The most popular choice here is the US method used in its implementation where multiple restraints are used to gradually bring the system from one state to another and the resulting trajectories are combined using the weighted-histogram analysis method [26, 34, 35, 36, 37, 41, 42, 48, 52, 56, 60, 62, 63, 64, 67, 69]. In the case of complex conformational transitions that are not sufficiently described by the employed CVs, results of multiple-restraints US could be highly dependent on the protocol used to initialize the simulations [34] introducing systematic errors that are difficult to detect. A more robust alternative is provided by replica-exchange US simulations, used for instance in Refs. [32, 41, 59], at least if continuous trajectories are analyzed and transition events are detected as discussed above for T-REMD (Figure 2). We notice that several works related to catalysis took advantage of the string method in order to sample a reactive pathway [60, 62, 64, 67]. This approach allows multiple CVs to be combined at the price of limiting the exploration to a reaction tube. Another popular method is metadynamics, where an adaptive bias potential is constructed iteratively so as to induce conformational transitions in a small number of preselected CVs [22, 28, 46, 49, 51]. Metadynamics can be run with multiple replicas in order to accelerate convergence [66], and can be used with a larger number of CVs either by biasing them one-at-a-time, as in bias-exchange metadynamics [33, 54], or concurrently, as in replica exchange with CV tempering [24, 25, 27]. Finally, metadynamics and T-REMD can be combined, as done for instance in Refs. [23, 25, 47].

Employed CVs. The success of CV-based methods depends heavily on the chosen CVs. Many of the works discussed here used simple geometric CVs such as distances between atoms or atom groups. Chemical reactions are typically accomplished by biasing a combination of distances, where some describe newly formed or expired contacts and others enforce related proton transfers [59, 60, 64, 65, 66, 67, 69]. Dihedral angles can be used to enforce the exploration of multiple rotamers [21, 24, 26, 27, 32, 43]. In some cases, the barriers associated to sugar repuckering were accelerated using pseudodihedrals [21, 24, 25, 27, 32]. Some other works used root-mean-square deviation (RMSD) from the native structure [22, 28, 41, 45]. RMSD is known to be a poor descriptor, in

particular for RNA systems. Refs. [23, 25] report cases where an RNA-dedicated metric (ϵRMSD) was utilized to distinguish native from non-native structures and biased. Whereas measures such as the ϵRMSD distinguish structures using the entire map of observed and non-observed contacts, variables such as the number of native contacts [22, 28, 40, 43, 47] are unaffected by the presence or absence of competing non-native contacts. As a general consideration, it should be taken into account that, for intrinsically high-dimensional free-energy landscapes, identifying a small number of CVs capable to describe all the relevant substates might be difficult or even virtually impossible.

Discussion and Perspectives

In this Review we surveyed the enhanced sampling methods recently applied to study RNA structural dynamics. In the following, we summarize our recommendations and perspectives.

Different Methods for the Same Problem. Different groups used different methods to tackle very similar problems. An example can be seen by comparing three works where the affinities of divalent ions for specific sites in RNA motifs were computed, ranging from classical US [52] through a novel grand-canonical Monte Carlo/MD approach [55] to a bias-exchange-like metadynamics procedure [54]. It would be interesting to compare these three methods on identical setups in order to assess their computational efficiency. Similarly, related catalytic reactions were tackled by different authors using US combined with string method [60, 62, 64, 67] or thermodynamic integration (TI) [65]. Albeit employed on different systems, both approaches used comparable QM/MM setups with similar sized QM regions described by the same DFT functional. String method calculations might allow a more precise definition of TSs, although the employed iterative procedure is expensive. The TI approach exploits a monodimensional pathway and required a verification that the related proton transfer was in fact induced in a reversible manner. Whereas the rearrangements associated to phosphodiester cleavage reactions are significantly simpler than those associated to RNA folding, we suggest that replica-exchange procedures where coordinates are swapped between consecutive windows might be beneficial in performing cleavage simulations, allowing reactive events to be observed in continuous trajectories.

Recommendations for Enhanced Sampling Simulations. At variance with plain MD simulations, which are often analyzed in a qualitative fashion, enhanced sampling simulations are usually employed to report thermodynamic averages to be directly compared with experiments. For a quantitative interpretation, it is however necessary that statistical errors are reported together with the results. Since the estimate of the error itself is sometimes non trivial, we suggest that authors should explain clearly how the errors were computed. We would like to reiterate that enhanced sampling methods can give statistically reliable results only if multiple transitions are observed in continuous trajectories. In addition, given the difficulty in reproducing this kind of calculations, authors should share input parameters and, when feasible, generated trajectories. For instance, the protocols introduced in [23] and [27] were used after a very short time by an independent group [25]. Finally, CV-based methods usually require a significant number of trial and errors in order to identify proper CVs. Sharing the non-working setups, perhaps in the form of supporting information, could speed up the progress in the field avoiding other groups to repeat similar mistakes.

Recommendations for RNA Simulations. In the last two years, mostly thanks to the publication of extensive benchmarks using enhanced sampling techniques, it was suggested that current empirical force fields are not yet accurate enough for blind prediction of some RNA native structures and for

reliably reproducing the conformational ensembles of small unstructured RNAs. Nevertheless, even without predictive accuracy, MD simulations are able to provide significant insights on experiments, mostly thanks to their spatial and temporal resolution. However, we feel that the community should work in the direction of improving force fields, taking advantage of enhanced sampling techniques in order to validate them. In this respect, we believe it is crucial that researchers continue sharing benchmarks and negative results. In addition, predictive simulations should be validated with care against independent experimental data. In this respect, solution-phase experiments such as NMR are particularly useful since they provide ensemble averages that can be directly compared with MD simulations and that account for RNA dynamics. A promising growing field is based on the idea of simultaneously applying enhanced sampling simulations and restraints obtained from experiments [24, 40, 43].

Perspectives. The importance of RNA structural dynamics in molecular biology is steadily growing. Structures of new RNA catalytic systems are being discovered at a constant pace, and hypotheses on their mechanism of action benefit from explicit modeling of the corresponding reaction pathways. In addition, local dynamics of flexible RNA motifs, especially in relation to their capability to bind ions, small ligands, proteins, and other RNA molecules, is receiving an increasing attention. We thus predict the role of enhanced sampling techniques in the RNA community to increase in the next years.

Acknowledgments

Alejandro Gil-Ley, Sandro Bottaro, Carlo Camilloni, Angel E. Garcia, Alex MacKerell, and Alessandra Magistrato are acknowledged for reading the manuscript and providing useful suggestions. This work has been supported European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement n. 306662, S-RNA-S.

[1] D. M. Lilley, How RNA acts as a nuclease: Some mechanistic comparisons in the nucleolytic ribozymes, Biochem. Soc. Trans. 45 (3) (2017) 683–691.

[2] K. V. Morris, J. S. Mattick, The rise of regulatory RNA, Nat. Rev. Genet. 15 (6) (2014) 423.

[3] E. Westhof, Twenty years of RNA crystallography, RNA 21 (4) (2015) 486–487.

[4] J. Rinnenthal, J. Buck, J. Ferner, A. Wacker, B. Fürtig, H. Schwalbe, Mapping the landscape of RNA dynamics with NMR spectroscopy, Acc. Chem. Res. 44 (12) (2011) 1292–1301.

[5] D. M. Lilley, The structure and folding of kink turns in RNA, Wiley Interdiscip. Rev. RNA 3 (6) (2012) 797–805.

[6] E. A. Dethoff, K. Petzold, J. Chugh, A. Casiano-Negroni, H. M. Al-Hashimi, Visualizing transient low-populated structures of RNA, Nature 491 (7426) (2012) 724–728.

[7] A. Serganov, E. Nudler, A decade of riboswitches, Cell 152 (1) (2013) 17–24.

[8] G. M. Emilsson, S. Nakamura, A. Roth, R. R. Breaker, Ribozyme speed limits, RNA 9 (8) (2003) 907–918.

[9] A. C. Pan, T. M. Weinreich, S. Piana, D. E. Shaw, Demonstrating an order-of-magnitude sampling enhancement in molecular dynamics simulations of complex protein systems, J. Chem. Theory Comput. 13 (3) (2016) 1360–1367.

[10] R. C. Bernardi, M. C. Melo, K. Schulten, Enhanced sampling techniques in molecular dynamics simulations of biological systems, Biochim. Biophys. Acta, Gen. Subj. 1850 (5) (2015) 872–877.

[11] O. Valsson, P. Tiwary, M. Parrinello, Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint, Annu. Rev. Phys. Chem. 67 (2016) 159–184.

[12] J. Šponer, J. E. Šponer, A. Mládek, P. Banáš, P. Jurečka, M. Otyepka, How to understand quantum chemical computations on DNA and RNA systems? A practical guide for non-specialists, Methods 64 (1) (2013) 3–11.

[13] M. Huang, T. J. Giese, D. M. York, Nucleic acid reactivity: Challenges for next-generation semiempirical quantum models, J. Comput. Chem. 36 (18) (2015) 1370–1389.

[14] J. Šponer, M. Krepl, P. Banáš, P. Kührová, M. Zgarbová, P. Jurečka, M. Havrila, M. Otyepka, How to understand atomistic molecular dynamics simulations of RNA and protein–RNA complexes?, Wiley Interdiscip. Rev. RNA 8 (3) (2017) e1405.

[15] S. Vangaveti, S. V. Ranganathan, A. A. Chen, Advances in RNA molecular dynamics: A simulator's guide to RNA force fields, Wiley Interdiscip. Rev. RNA 8 (2) (2017) e1396.

[16] L. G. Smith, J. Zhao, D. H. Mathews, D. H. Turner, Physics-based all-atom modeling of RNA energetics and structure, Wiley Interdiscip. Rev. RNA 8 (5) (2017) e1422.

[17] C. Bergonzo, N. M. Henriksen, D. R. Roe, T. E. Cheatham III, Highly sampled tetranucleotide and tetraloop motifs enable evaluation of common RNA force fields, RNA 21 (9) (2015) 1578–1590,
•• Extensive investigation on several tetraloops and tetranucleotides using a multidimensional replica-exchange scheme. For tetranucleotides, this work provides the first virtually converged simulation, showing that the population of intercalated structures is larger than expected from NMR data .

[18] C. Bergonzo, T. E. Cheatham III, Improved force field parameters lead to a better description of RNA structure, J. Chem. Theory Comput. 11 (9) (2015) 3969–3972.

[19] S. Sakuraba, K. Asai, T. Kameda, Predicting RNA duplex dimerization free-energy changes upon mutations using molecular dynamics simulations, J. Phys. Chem. Lett. 6 (21) (2015) 4348–4351.

[20] S. Bottaro, A. Gil-Ley, G. Bussi, RNA folding pathways in stop motion, Nucleic Acids Res. 44 (12) (2016) 5883–5891.

[21] A. Gil-Ley, S. Bottaro, G. Bussi, Empirical corrections to the AMBER RNA force field with target metadynamics, J. Chem. Theory Comput. 12 (6) (2016) 2790–2798.

[22] P. Kührová, R. B. Best, S. Bottaro, G. Bussi, J. Šponer, M. Otyepka, P. Banáš, Computer folding of RNA tetraloops: Identification of key force field deficiencies, J. Chem. Theory Comput. 12 (9) (2016) 4534–4548,

•• Extensive investigation on a tetraloop using several enhanced sampling methods including temperature replica exchange, solute tempering, and metadynamics. Different methods result in qualitatively consistent results. The most important current force fields are compared, and none of them can reproduce a significant population of the known native structure.

[23] S. Bottaro, P. Banáš, J. Šponer, G. Bussi, Free energy landscape of GAGA and UUCG RNA tetraloops, J. Phys. Chem. Lett. 7 (20) (2016) 4032–4038,
•• Converged free-energy landscapes for two RNA tetraloops including both native and non-native structures are here reported for the first time. The result is obtained exploiting a RNA-specific structural metrics as biased collective variable. This procedure allows to quantify the (in)stability of the native structure. In addition, a direct comparison with available NMR data is performed.

[24] A. Cesari, A. Gil-Ley, G. Bussi, Combining simulations and solution experiments as a paradigm for RNA force field refinement, J. Chem. Theory Comput. 12 (12) (2016) 6192–6200.

[25] C. Yang, M. Lim, E. Kim, Y. Pak, Predicting RNA structures via a simple van der Waals correction to an all-atom force field, J. Chem. Theory Comput. 13 (2) (2017) 395–399.

[26] A. H. Aytenfisu, A. Spasic, A. Grossfield, H. A. Stern, D. H. Mathews, Revised RNA dihedral parameters for the AMBER force field improve RNA molecular dynamics, J. Chem. Theory Comput. 13 (2) (2017) 900–915.

[27] A. Gil-Ley, G. Bussi, Enhanced conformational sampling using replica exchange with collective-variable tempering, J. Chem. Theory Comput. 11 (3) (2015) 1077–1085,
•• A novel enhanced sampling method is introduced that combines metadynamics and replica exchange. The method accelerates simultaneously many collective variables and is particularly suitable to accelerate transitions between different RNA rotamers.

[28] S. Haldar, P. Kührová, P. Banáš, V. Spiwok, J. Šponer, P. Hobza, M. Otyepka, Insights into stability and folding of GNRA and UNCG tetraloops revealed by microsecond molecular dynamics and well-tempered metadynamics., J. Chem. Theory Comput. 11 (8) (2015) 3866.

[29] X. Xue, W. Yongjun, L. Zhihong, Folding of SAM-II riboswitch explored by replica-exchange molecular dynamics simulation, J. Theor. Biol. 365 (2015) 265–269.

[30] H. Park, À. L. González, I. Yildirim, T. Tran, J. R. Lohman, P. Fang, M. Guo, M. D. Disney, Crystallographic and computational analyses of AUUCU repeating RNA that causes Spinocerebellar Ataxia type 10 (SCA10), Biochemistry 54 (24) (2015) 3851.

[31] C. Bergonzo, K. B. Hall, T. E. Cheatham III, Stem-Loop V of Varkud Satellite RNA exhibits characteristics of the Mg^{2+} bound structure in the presence of monovalent ions, J. Phys. Chem. B 119 (38) (2015) 12355–12364.

[32] B. K. Radak, M. Romanus, T.-S. Lee, H. Chen, M. Huang, A. Treikalis, V. Balasubramanian, S. Jha, D. M. York, Characterization of the three-dimensional free energy manifold for the uracil ribonucleoside from asynchronous replica exchange simulations, J. Chem. Theory Comput. 11 (2) (2015) 373–377.

[33] Y. Bian, J. Zhang, J. Wang, J. Wang, W. Wang, Free energy landscape and multiple folding pathways of an H-type RNA pseudoknot, PloS One 10 (6) (2015) e0129089.

[34] F. Di Palma, S. Bottaro, G. Bussi, Kissing loop interaction in adenine riboswitch: Insights from umbrella sampling simulations, BMC Bioinformatics 16 (Suppl 9) (2015) S6,
•• Umbrella sampling is used to investigate the ligand-dependent stabilization of a kissing loop in the adenine riboswitch, obtaining a qualitative agreement with experiment. Importantly, the dependence of the results on the initialization protocol is discussed in detail.

[35] A. T. Larsen, A. C. Fahrenbach, J. Sheng, J. Pian, J. W. Szostak, Thermodynamic insights into 2-thiouridine-enhanced RNA hybridization, Nucleic Acids Res. 43 (16) (2015) 7675–7687.

[36] I. Yildirim, D. Chakraborty, M. D. Disney, D. J. Wales, G. C. Schatz, Computational investigation of RNA CUG repeats responsible for myotonic dystrophy 1, J. Chem. Theory Comput. 11 (10) (2015) 4943–4958.

[37] Y.-Y. Wu, Z.-L. Zhang, J.-S. Zhang, X.-L. Zhu, Z.-J. Tan, Multivalent ion-mediated nucleic acid helix-helix interactions: RNA versus DNA, Nucleic Acids Res. 43 (12) (2015) 6156–6165.

[38] J. C. Miner, A. A. Chen, A. E. García, Free-energy landscape of a hyperstable RNA tetraloop, Proc. Natl. Acad. Sci. U.S.A. 113 (24) (2016) 6665–6670.

[39] M. K. Takahashi, K. E. Watters, P. M. Gasper, T. R. Abbott, P. D. Carlson, A. A. Chen, J. B. Lucks, Using in-cell SHAPE-seq and simulations to probe structure–function design principles of RNA transcriptional regulators, RNA 22 (6) (2016) 920–933.

[40] A. N. Borkar, M. F. Bardaro, C. Camilloni, F. A. Aprile, G. Varani, M. Vendruscolo, Structure of a low-population binding intermediate in protein–RNA recognition, Proc. Natl. Acad. Sci. U.S.A. 113 (26) (2016) 7171–7176,
•• The structural dynamics of an RNA:peptide complex formed by TAR RNA and a cyclic peptide is characterized. By combining enhanced sampling simulations with NMR data the authors are able to identify low population structures of the complex that are then validated experimentally. This work shows how to effectively combine experimental data and enhanced sampling simulations.

[41] L. Vukovic, C. Chipot, D. L. Makino, E. Conti, K. Schulten, Molecular mechanism of processive 3' to 5' RNA translocation in the active subunit of the RNA exosome complex, J. Am. Chem. Soc. 138 (12) (2016) 4069–4078.

[42] H. S. Hayatshahi, C. Bergonzo, T. E. Cheatham III, Investigating the ion dependence of the first unfolding step of GTPase-associating center ribosomal RNA, J. Biomol. Struct. Dyn. (2017) 1–11.

[43] A. N. Borkar, P. Vallurupalli, C. Camilloni, L. E. Kay, M. Vendruscolo, Simultaneous NMR characterisation of multiple minima in the free energy landscape of an RNA UUCG tetraloop, Phys. Chem. Chem. Phys. 19 (4) (2017) 2797–2804.

[44] B. M. Warfield, P. C. Anderson, Molecular simulations and Markov state modeling reveal the structural diversity and dynamics of a theophylline-binding RNA aptamer in its unbound state, PloS One 12 (4) (2017) e0176229.

[45] G. Palermo, Y. Miao, R. C. Walker, M. Jinek, J. A. McCammon, CRISPR-Cas9 conformational activation as elucidated from enhanced molecular simulations, Proc. Natl. Acad. Sci. U.S.A. 114 (28) (2017) 7260–7265.

[46] M. D. Verona, V. Verdolino, F. Palazzesi, R. Corradini, Focus on PNA flexibility and RNA binding using molecular dynamics and metadynamics, Sci. Rep. 7 (2017) 42799.

[47] A. K. Pathak, T. Bandyopadhyay, Water isotope effect on the thermostability of a polio viral RNA hairpin: A metadynamics study, J. Chem. Phys. 146 (16) (2017) 165104.

[48] Z. Sun, X. Wang, J. Z. Zhang, Protonation-dependent base flipping in the catalytic triad of a small RNA, Chem. Phys. Lett. 684 (2017) 239–244.

[49] F. Pan, V. H. Man, C. Roland, C. Sagui, Structure and dynamics of DNA and RNA double helices of CAG and GAC trinucleotide repeats, Biophys. J. 113 (1) (2017) 19–36.

[50] J. C. Miner, A. E. García, Equilibrium denaturation and preferential interactions of an RNA tetraloop with urea, J. Phys. Chem. B 121 (15) (2017) 3734–3746.

[51] A. Bochicchio, G. Rossetti, O. Tabarrini, S. Krauβ, P. Carloni, Molecular view of ligands specificity for CAG repeats in anti-Huntington therapy, J. Chem. Theory Comput. 11 (10) (2015) 4911–4922.

[52] M. T. Panteva, G. M. Giambasu, D. M. York, 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 (50) (2015) 15460–15470.

[53] J. A. Liberman, K. C. Suddala, A. Aytenfisu, D. Chan, I. A. Belashov, M. Salim, D. H. Mathews, R. C. Spitale, N. G. Walter, J. E. Wedekind, Structural analysis of a class III preQ1 riboswitch reveals an aptamer distant from a ribosome-binding site regulated by fast dynamics, Proc. Natl. Acad. Sci. U.S.A. 112 (27) (2015) E3485–E3494.

[54] R. A. Cunha, G. Bussi, Unraveling Mg^{2+}–RNA binding with atomistic molecular dynamics, RNA 23 (5) (2017) 628–638.

[55] J. A. Lemkul, S. K. Lakkaraju, A. D. MacKerell Jr, Characterization of Mg^{2+} distributions around RNA in solution, ACS omega 1 (4) (2016) 680–688.

[56] H. S. Hayatshahi, D. R. Roe, R. Galindo-Murillo, K. B. Hall, T. E. Cheatham III, Computational assessment of potassium and magnesium ion binding to a buried pocket in GTPase-associating center RNA, J. Phys. Chem. B 121 (3) (2017) 451–462.

[57] G. Hu, A. Ma, J. Wang, Ligand selectivity mechanism and conformational changes in guanine riboswitch by molecular dynamics simulations and free energy calculations, J. Chem. Inf. Model. 57 (4) (2017) 918–928.

[58] G. Grasso, M. A. Deriu, V. Patrulea, G. Borchard, M. Möller, A. Danani, Free energy landscape of siRNA-polycation complexation: Elucidating the effect of molecular geometry, polymer flexibility, and charge neutralization, PloS One 12 (10) (2017) e0186816.

[59] B. K. Radak, T.-S. Lee, M. E. Harris, D. M. York, Assessment of metal-assisted nucleophile activation in the hepatitis delta virus ribozyme from molecular simulation and 3D-RISM, RNA 21 (9) (2015) 1566–1577.

[60] S. Zhang, A. Ganguly, P. Goyal, J. L. Bingaman, P. C. Bevilacqua, S. Hammes-Schiffer, Role of the active site guanine in the *glms* ribozyme self-cleavage mechanism: Quantum mechanical/molecular mechanical free energy simulations, J. Am. Chem. Soc. 137 (2) (2015) 784–798,
•• QM/MM free-energy calculations are here used to compare various mechanisms of phospodiester cleavage. The main focus is on the role of guanine nucleotide, which is present in the active sites of several small self-cleaving ribozymes. The effective usage of string method is also described in detail.

[61] J. B. Swadling, D. W. Wright, J. L. Suter, P. V. Coveney, Structure, dynamics, and function of the hammerhead ribozyme in bulk water and at a clay mineral surface from replica exchange molecular dynamics, Langmuir 31 (8) (2015) 2493–2501.

[62] P. Thaplyal, A. Ganguly, S. Hammes-Schiffer, P. C. Bevilacqua, Inverse thio effects in the hepatitis delta virus ribozyme reveal that the reaction pathway is controlled by metal ion charge density, Biochemistry 54 (12) (2015) 2160–2175.

[63] C. S. Gaines, D. M. York, Ribozyme catalysis with a twist: Active state of the twister ribozyme in solution predicted from molecular simulation, J. Am. Chem. Soc. 138 (9) (2016) 3058–3065.

[64] S. Zhang, D. R. Stevens, P. Goyal, J. L. Bingaman, P. C. Bevilacqua, S. Hammes-Schiffer, Assessing the potential effects of active site Mg^{2+} ions in the *glms* ribozyme–cofactor complex, J. Phys. Chem. Lett. 7 (19) (2016) 3984–3988.

[65] L. Casalino, G. Palermo, U. Rothlisberger, A. Magistrato, Who activates the nucleophile in ribozyme catalysis? An answer from the splicing mechanism of group II introns, J. Am. Chem. Soc. 138 (33) (2016) 10374–10377,
•• QM/MM free-energy calculations are here used to resolve the mechanism of phoshodiester hydrolysis in a large ribozyme. The authors explore a novel mechanism, where water molecules not only from the first coordination shell of Mg^{2+} but also from the bulk participate in the cleavage reaction. Expensive CPMD simulations, implementing a fully Hamiltonian coupling scheme for QM/MM, are used.

[66] V. Mlýnský, G. Bussi, Understanding in-line probing experiments by modeling cleavage of nonreactive RNA nucleotides, RNA 23 (5) (2017) 712–720.

[67] J. L. Bingaman, S. Zhang, D. R. Stevens, N. H. Yennawar, S. Hammes-Schiffer, P. C. Bevilacqua, The GlcN6P cofactor plays multiple catalytic roles in the *glms* ribozyme, Nat. Chem. Biol. 13 (4) (2017) 439–445.

[68] C. Schuabb, N. Kumar, S. Pataraia, D. Marx, R. Winter, Pressure modulates the self-cleavage step of the hairpin ribozyme., Nat. Commun. 8 (2017) 14661.

[69] H. Chen, T. J. Giese, B. L. Golden, D. M. York, Divalent metal ion activation of a guanine general base in the hammerhead ribozyme: Insights from molecular simulations, Biochemistry 56 (24) (2017) 2985–2994.