. SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI

SISSA Digital Library

Integrating experimental data with molecular simulations to investigate RNA structural dynamics

Original
Integrating experimental data with molecular simulations to investigate RNA structural dynamics / Bernetti, Mattia; Bussi, Giovanni. - In: CURRENT OPINION IN STRUCTURAL BIOLOGY. - ISSN 0959-440X. - 78:(2023). [10.1016/j.sbi.2022.102503]

Availability:
This version is available at: 20.500.11767/130351 since: 2025-01-29T21:01:58Z

Publisher:

Published
DOI:10.1016/j.sbi.2022.102503

Terms of use:

Testo definito dall'ateneo relativo alle clausole di concessione d'uso

Publisher copyright

note finali coverpage

(Article begins on next page)

01 May 2025

Integrating experimental data with molecular simulations to investigate RNA structural dynamics

Mattia Bernetti[a,b], Giovanni Bussi[c,*]

[a] *Computational and Chemical Biology, Italian Institute of Technology, 16152 Genova, Italy*
[b] *Department of Pharmacy and Biotechnology, University of Bologna, 40126 Bologna, Italy*
[c] *Scuola Internazionale Superiore di Studi Avanzati, via Bonomea 265, 34136, Trieste, Italy*

Abstract

Conformational dynamics is crucial for ribonucleic acid (RNA) function. Techniques such as nuclear magnetic resonance, cryo-electron microscopy, small- and wide-angle X-ray scattering, chemical probing, single-molecule Förster resonance energy transfer or even thermal or mechanical denaturation experiments probe RNA dynamics at different time and space resolutions. Their combination with accurate atomistic molecular dynamics (MD) simulations paves the way for quantitative and detailed studies of RNA dynamics. First, experiments provide a quantitative validation tool for MD simulations. Second, available data can be used to refine simulated structural ensembles to match experiments. Finally, comparison with experiments allows for improving MD force fields that are transferable to new systems for which data is not available. Here we review the recent literature and provide our perspective on this field.

Keywords: RNA, Molecular dynamics simulations, Solution experiments

©2022. This manuscript version is made available under the CC-BY-NC-ND 4.0 license https://creativecommons.org/licenses/by-nc-nd/4.0/. This is an accepted manuscript (postprint). The published journal article is available at `https://doi.org/10.1016/j.sbi.2022.102503`.

Introduction

Ribonucleic acid (RNA) structure is often necessary for function [1]. However, RNA structures must not be considered as static pictures. RNA dynamics, i.e., the capability to explore multiple conformations, is indeed crucial as it is required for (and modulated by) binding with small molecules, either metabolites [2] or drug-like molecules [3], or macromolecules [4] (see Figure 1a–b). Many experimental techniques have been developed to probe RNA structure. X-ray crystallography [6] provides very detailed pictures but restricts dynamics through crystal packing effects, thus with limited applicability when dealing with flexible RNAs. Conversely, solution methods such as nuclear magnetic resonance (NMR) [7] and cryo-electron microscopy [8] allow for probing RNA dynamics, at the price of a lower atomic resolution. Additionally, coarser-grained experimental techniques, such as small- (or wide-) angle X-ray scattering (SAXS/WAXS) [9], chemical probing [10], single-molecule Förster resonance energy transfer [11] or, to some extent, thermal [12] or mechanical [13] denaturation experiments are also capable to access RNA dynamics. A common trait of these techniques is that they report possibly noisy ensemble averages over the set of accessible structures (see Figure 1c–e). The dynamical information is thus present but must be deconvolved. Molecular dynamics (MD) simulations [14] enable RNA dynamics to be predicted and characterized at the atomistic scale. Limitations on the accessible timescales can be alleviated by taking advantage of enhanced sampling methods [15]. However, despite recent improvements, the accuracy of the models used to describe interactions (force fields) for RNA is still limited, in particular for some structural motifs [16]. Nevertheless, MD simulations can be fruitfully applied also when current force fields are not predictive, by suitable integration of experimental data (see Figure 1f).

In this review, we discuss how MD simulations and experimental data can be combined to investigate the conformational dynamics of RNA molecules, covering

*Corresponding author
Email addresses: `Mattia.Bernetti@iit.it` (Mattia Bernetti), `bussi@sissa.it` (Giovanni Bussi)

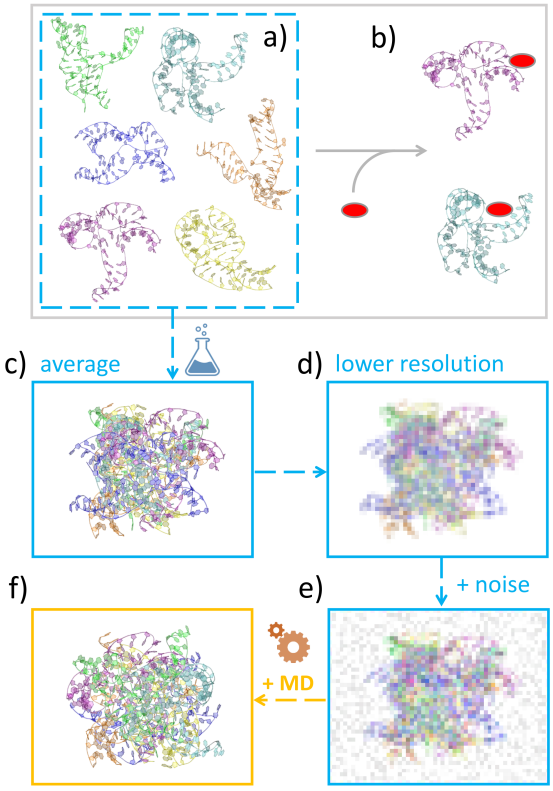

Figure 1: RNA dynamics. RNA molecules in solution sample multiple conformations (a). Upon binding with a small ligand or another macromolecule (in red), a subset of these conformations is selected (conformational selection) or a few additional ones are induced (induced fit) (b). Being able to characterize the unbound and bound structural ensembles is central in studying interactions. Solution experiments usually report information that is averaged over the conformational ensemble (c) and only contains a lower resolution description (d), usually affected by noise (e). Accurate modeling tools are required to deconvolve dynamics and reconstruct the original ensembles (f). Representative RNA structures were adapted from those generated in Ref. [5].

recent studies where simulations are either quantitatively validated on solution experiments or integrated with experimental information. We will not discuss structure prediction methods, whose typical goal is to find a single relevant structure. The survey will be limited to the past few years.

Combining molecular simulations and solution experiments

Integration strategies

Multiple strategies can be used to integrate MD and experimental information (see [17] for a recent review and Figure 2 for a schematic).

First, experiments can be used to quantitatively *validate* MD-generated ensembles (Figure 2c). For instance, testing multiple force fields, it is possible to determine which is the most trustable one, thus generating information transferable to other systems.

Second, experiments can be used to *improve generated ensembles* using either:

- Qualitative methods, where data are used to generate initial structural models or to restrain the simulations without explicitly modeling the experimental dynamics (Figure 2d).
- Quantitative methods, ranging from maximum parsimony, such as sample-and-select (Figure 2e), to maximum entropy (Figure 2f); see Ref. [18] for an overview. Results are not transferable to other systems but can be used to (a) translate the experiment to detailed structural information and (b) predict new experiments on the same system. Methods can be used both for reweighting/analyzing existing trajectories or to restrain them on-the-fly, the latter approach being more efficient when the original simulated ensembles are farther from the experiment.

Finally, experiments can be used to *improve force fields* [19] (Figure 2g). Force-field parameters can be first trained using reweighting techniques that enable predicting the impact of parameter modifications, and then validated with new simulations. Results are expected to be transferable to other systems for which the experiments are not available yet. To enhance transferability, one should train the force field on a large and heterogeneous set of systems.

Critical issues

Several critical issues should be taken into account when comparing simulations and experiments. First,

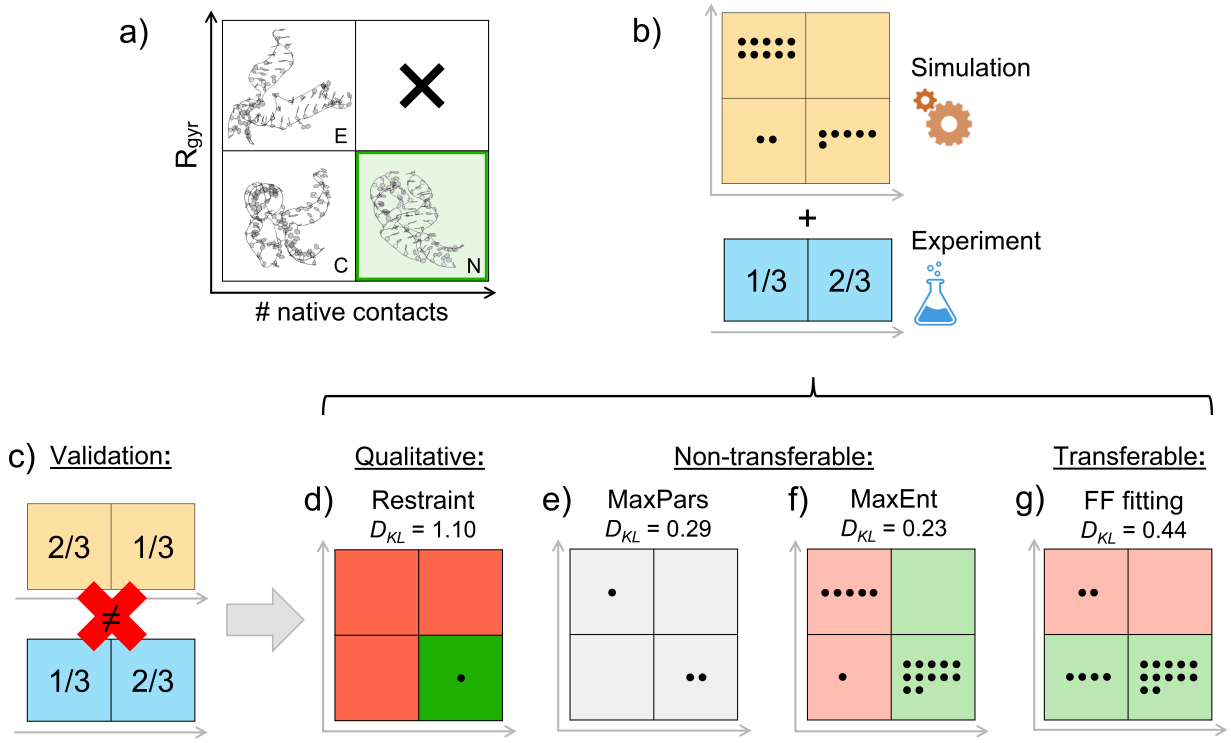

Figure 2: Integrating simulations and experiments. We consider a molecule with three available conformations (a): extended (E, large gyrations radius, R_{gyr}, and no native contacts formed), compact (C, small R_{gyr} and no native contacts formed), and native (N, small R_{gyr} and many native contacts formed, highlighted by the green shade and box contour). Native contacts are assumed to be those present in a reference structure. An MD simulation results in populations of the three states as reported in the yellow square, proportional to the number of dots, and the experiment suggests that the population of the native state is 2/3 (b). The comparison of the MD simulation with the experiment indicates that the MD ensemble is incorrect (c). One could qualitatively restrain the simulation to native (d). More quantitatively, one can choose a minimal number of conformations that agree with the experiment (maximum parsimony, MaxPars, e), or minimally modify the MD ensemble (maximum entropy, MaxEnt, f). Both these approaches are not transferable in that they only allow extracting information about the system for which experiments were performed. Finally, a force-field (FF) fitting strategy could be used to obtain transferable corrections (g). In this example, we assumed that the water model or the torsional parameters are modified, impacting the relative population of the extended state. Green and red shades indicate which states are favored and disfavored by the experimental data, respectively. Kullback-Leibler divergences (D_{KL}) from the original MD ensemble, computed using the populations of the three states, are reported for the four generated ensembles. Representative RNA structures were adapted from those generated in Ref. [5].

the magnitude of the experimental error should be considered. An agreement between simulation and experiment that is better than the uncertainty of the experiment itself should be seen as an indication of overfitting. Second, the *forward models*, i.e. the equations to back-calculate the experiment from the MD-simulated structures, are often empirically parametrized and thus subject to systematic errors. Then, back-calculated experiments are affected by statistical errors due to the finite length of the simulations. Comparing experiments with non-converged simulations might lead to incorrect assessments. When multiple structures are required to explain experimental results, visiting them in an accessible simulation time almost invariably requires enhanced sampling techniques. Finally, no experimental or computational technique is fool-proof, and experience is required to properly assess the relative merit of different approaches.

Recent applications to RNA dynamics

Here we review recent applications where MD and experiments were combined to study RNA dynamics. All the discussed applications are summarized in Table 1.

Using experimental data for validation or force-field selection

In several papers, quantitative comparisons with reference experimental data have been used to assess different force fields. Analyses were based on NMR data for dinucleoside monophosphates [20], tetranucleotides [21], or hexanucleotides [21, 22] and involved a couple of force fields, typically a widely adopted one, used as a reference, and a newer variant tested. NMR data on tetramers were used in Ref. [23] to scan a wide range of parameters for hydrogen-bond fixes. Two-way junctions, as well as single-stranded RNAs, were used to test corrections to non-bonded terms where RNA-RNA and RNA-solvent interactions were tuned separately [24]. Here, specific sets of parameters were compared to find those better reproducing SAXS data and then validated against NMR data. The tests reported in Refs. [23, 24] were done in a spirit that, although based on enumeration rather than on optimization, resembles that of methods discussed in Section *Integrating experimental data and simulations to improve force fields*. In other papers, alchemical calculations have been validated with binding affinities [25] or denaturation free energies [26], testing a single set of force-field parameters against experiments.

Integrating experimental data and simulations to improve ensembles

Two recent works used restraints based on the secondary structure obtained from either NMR [27] or chemical probing [2] data, employing a two-dimensional replica-exchange method. Here, replicas in which the experimental information is used to enforce known base pairings are coupled with unbiased replicas. In this manner, the correct base pairing is encouraged but not enforced, and dynamics is preserved. Base pairing information obtained from chemical probing and NMR data was also used to construct initial structures in Ref. [28]. Enhanced sampling simulations were then conducted without restraints, but the experimental base pairing was preserved by choosing a suitable replica-exchange approach. Contacts predicted by nuclear Overhauser effect (NOE) data have been qualitatively used as instantaneous restraints when equilibrating protein-RNA complexes [29, 30], followed by unrestrained MD simulations with the exception of a few simulations employing soft NOE restraints with a limited impact on the dynamics. In Ref. [29] the authors validated quantitatively the results against experimental RNA-protein affinities and effects of chemical modifications. In all these works, for an infinitely long simulation, one should expect the result to be unaffected by the presence of the experimental information. Hence, these methods should be better seen as sampling or modeling tools rather than integrative methods. Experimentally derived secondary structures were used also in Ref. [31]. Here, structures obtained from co-transcriptional chemical-probing experiments were enforced during the simulation, using an *ad hoc* procedure to avoid the system being stuck in non-productive pathways. In this case, the experimentally-derived secondary structure was enforced in the MD-generated structures. However, dynamics was not quantitatively controlled.

Cryo-electron-microscopy experiments have been used in so-called flexible-fitting MD simulations to refine structures [32, 33]. In this setting, MD simulations are used as a modeling tool to generate structures compatible with experiments, but the quantitative degree of dynamics compatible with the experimental density maps is not explicitly controlled.

Several works used the maximum entropy principle to enforce ensemble averages quantitatively. The simulated ensemble of a UUCG tetraloop was reweighted to match a set of NMR data, including NOEs, exact NOEs, ^3J scalar couplings, and solvent paramagnetic relaxation enhancement [34]. The MD simulation suggested an alternative structure for the loop. Including available

data suggested a lower but not negligible population of this alternative structure. Interestingly, different data sets led to different results. A similar approach was used in Ref. [35], where reweighting was conducted using either NMR or SAXS data on ensembles generated for an RNA hexamer. Agreement with NMR did not necessarily result in agreement with SAXS, and vice versa, suggesting that multiple, independent experimental observables are important to assess the accuracy of heterogeneous structural ensembles. Multiple force-field parameters for atomistic MD were compared. NOE data were also used in Ref. [36] to construct heterogeneous ensembles for an RNA hairpin found in a non-coding RNA. In this case, reweighting was not sufficient and experimental data were thus enforced during the simulation. A maximum parsimony approach was then employed to generate ensembles compatible with the experiment but composed of structures belonging to a limited number of structural clusters. This approach has the advantage that it can be used to obtain ensembles that are simpler to describe, at the price of decreasing dynamics and modifying more the ensemble relative to the one simulated with MD. SAXS data were used in Ref. [5] to quantify the population of compact and extended conformations in a larger structured RNA. Here, different forward models for SAXS were tested, showing that both solvent and dynamical effects are crucial to match experimental data. In this case, enhanced sampling allowed to have both compact and extended structures in the pool of conformations used for reweighting.

Other available methods for ensemble refinement are based on the maximum parsimony principle, and include for instance the popular sample-and-select and ensemble optimization approaches. In Ref. [37], WAXS data were used to select conformations from MD simulations of RNA and DNA duplexes. Simulations performed at a higher temperature were used to enhance sampling in cases where conventional MD at room temperature was not providing conformations capable of rationalizing the experimental spectrum. A similar approach was used in Ref. [38] to analyze RNA triplexes. Here, WAXS-driven MD was used to tackle cases where high-temperature sampling was not sufficient. WAXS data were thus used both on-the-fly and a posteriori. The sample-and-select method was also used to analyze MD simulations of TAR RNA to match measured residual dipolar couplings [39]. Interestingly, in this case, the authors reported that an ensemble generated with a fragment assembly method (FARFAR) was matching experiments better than atomistic MD. It is unclear if this is due to the higher accuracy of FARFAR or to better sampling. Furthermore, these results might be contrasted with the opposite finding in two papers discussed above [28, 35].

Integrating experimental data and simulations to improve force fields

The quantitative mismatch between simulation and experiment has been also used to refine force-field parameters. In Ref. [40], simulations of a tetramer were compared with NMR experimental data. A minimization strategy was then used to optimize hydrogen-bond corrections, additionally including in the cost function the comparison between tetraloop simulations and static information from X-ray crystallography. The resulting force field was capable to stabilize the difficult UUCG tetraloop and reproduce solution data for additional tetramers that were not used during training.

A torsional potential and, more notably, partial charges were fine-tuned in Ref. [41] for a modified nucleobase. Parameters were tuned with an optimization procedure to match $\Delta\Delta G$s measured in thermal denaturation experiments. Whereas modifying charges leads to long-range effects that might make the reweighting procedure not efficient, tiny modifications were sufficient to significantly improve the agreement with the experiments. This work shows that it is possible to significantly improve agreement with experiment by tuning partial charges without leaving the fixed-charge approximation. However, the explicit modeling of polarization effects might be necessary to enable a further step forward in the accuracy and transferability of RNA force fields [42].

Reference [43] used a two-dimensional correction on the α and ζ backbone torsional angles to decrease the population of intercalated structures for RNA tetramers. The procedure was carried out with reweighting, to virtually test a large number of parameter sets, but without explicitly computing derivatives of the cost function with respect to the parameters. It is worth noting that a similar effect can be partly obtained using a completely different approach, namely switching to a different water model. Indeed, force-field refinement strategies do not lead to unique results and require manual intervention to decide which parameters should be optimized.

Discussion and perspective

The integration of experimental data and atomistic MD simulations is a growing field of research. Herein, we gave an overview of the recent works done on this subject in the context of RNA molecules, summarized in Table 1. A significant fraction of the papers discussed

Table 1: Recent applications where MD and experiments were combined to study RNA dynamics. For each of the discussed articles, the table reports which method was used to integrate MD simulations and experiments, which type of experiment was performed, which is the corresponding structural information, and the references to the simulation work discussed here and to the primary source for experimental data. Abbreviations: maximum entropy (MaxEnt), maximum parsimony (MaxPars), nuclear magnetic resonance (NMR), nuclear Overhauser effect (NOE), exact nuclear Overhauser effect (eNOE), residual dipolar coupling (RDC), solvent paramagnetic relaxation enhancement (sPRE), small-angle X-ray scattering (SAXS), wide-angle X-ray scattering (WAXS), cryo-electron microscopy (cryo-EM).

Integration method	Data	Information provided	Simulation	Experiment
Validation against experiment or force field selection				
Validation	^3J couplings	Sugar conformation	[20]	[44]
Validation	^3J couplings and NOEs	Sugar and backbone conformation, contacts	[21]	[21, 45]
Validation	^3J couplings and NOEs	Sugar and backbone conformation, contacts	[22]	[22]
Validation	^3J couplings and NOEs	Sugar and backbone conformation, contacts	[23]	[45–47]
Validation	^3J couplings and SAXS	Sugar and backbone conformation, inter-helical orientation	[24]	[45, 48]
Validation	Equilibrium filtration	Binding affinities	[25]	[49]
Validation	Thermal denaturation	Mutational free energies	[26]	[50]
Qualitative integration of experimental dynamics				
Secondary structure 2D replica exchange	NMR	Secondary structure	[27]	[27]
Secondary structure 2D replica exchange	Chemical probing	Secondary structure	[2]	[2]
Restraints during equilibration	NMR and chemical probing	Secondary structure	[28]	[51]
Restraints during equilibration	NOEs	Contacts	[29]	[29]
Restraints during equilibration	NOEs	Contacts	[30]	[30]
Restraints	Chemical probing	Secondary structure	[31]	[31]
Flexible fitting	Cryo-EM	Density map	[32]	[32]
Flexible fitting	Cryo-EM	Density map	[33]	[33]
Quantitative integration of experimental dynamics (non-transferable ensemble refinement)				
MaxEnt reweighting	^3J coupling, NOEs, eNOEs, RDCs, sPRE	Sugar and backbone conformation, contacts, dipole orientations, solvent exposure	[34]	[52–55]
MaxEnt reweighting	Chemical shifts, ^3J coupling, NOEs, SAXS	Sugar and backbone conformation, contacts, structure compactness	[35]	[21, 35]
MaxEnt and MaxPars on-the-fly + reweighting	NOEs	Contacts	[36]	[56]
MaxEnt reweighting	SAXS	Structure compactness	[5]	[57]
Sample-and-select	WAXS	Pairwise distance correlations	[37]	[37]
Sample-and-select + WAXS driven	WAXS	Pairwise distance correlations	[38]	[38]
Sample-and-select	RDCs	Dipole orientations	[39]	[58, 59]
Quantitative integration of experimental dynamics (transferable force-field refinement)				
Force-field refinement (hydrogen bonds)	^3J coupling, NOEs	Sugar and backbone conformation, contacts	[40]	[45, 46]
Force-field refinement (charges and torsion)	Thermal denaturation	Mutational free energies	[41]	[50, 60]
Force-field refinement (torsions)	NMR	Population of intercalated structures	[43]	[45]

here were based on experimental data previously made available by other authors, also listed in Table 1. In this respect, the effort to measure and publish data for systems that sometimes are not even biologically relevant is of great value for groups developing and applying integrative methods. Moreover, we stress that the interpretation of computational and experimental work is not easy, and that collaboration between computational and experimental groups is often required. In the following, we outline our perspective for this field, summarized in Box 1.

An important limitation of the current approaches based on explicit solvent MD simulations is their significant computational cost. Enhanced sampling methods are thus increasingly adopted in the community. However, current applications are either addressing unstructured oligomers or larger but relatively structured RNAs. The simulation of highly flexible long noncoding RNAs might result in further challenges and require new methodological improvements. The limited accuracy of existing force fields makes it necessary to include many data points to construct realistic ensembles. Experimental groups have been traditionally employing implicit-solvent modeling for structure determination and optionally short explicit-solvent refinements which can only account for minor conformational dynamics (see, e.g., Refs. [50, 52, 54–56], for NMR, and Refs. [32, 33], for cryo-electron microscopy). Atomistic force fields are likely more accurate, and a transition to using extensive explicit-solvent MD to interpret their data is already undergoing in several experimental laboratories.

Whereas most of the applications to date are related to non-transferable methods, where MD is used to deconvolve available experimental data, a few groups are trying to use experimental data to systematically improve force fields in a transferable manner. In this sense, force fields can be seen as energy-based machine-learning models [61]. Training should be done with examples that are representative of what one wants to reproduce. If the aim is to reproduce and predict experimental observables, experimental data should be used during training. The physics-based architecture of the force-field model, though limited by the lack of important aspects such as multi-body and polarization effects, can improve model transferability. However, to ensure the correctness of the microscopic models and to obtain force fields that properly describe conformations not yet seen in an experiment, a suitable combination of solution data, crystallographic information, and bottom-up data from quantum chemistry is required. Differences in experimental settings, such as ionic strength, temperature, and crystal packing, should be properly taken into account by performing reference simulations in appropriate conditions or even in crystalline environments [62]. The ambitious Open Force Field initiative is trying to use these ideas on large databases of heterogeneous systems, but to date this strategy has only been used to parametrize small molecules [63]. Experimental data on macromolecules will be required to push this approach forward to protein and RNA systems. Importantly, more commonly used non-transferable approaches such as maximum entropy or maximum parsimony integration methods will benefit from improved force fields used to generate prior ensembles. Kinetic data might also provide an extremely valuable source of information as they can probe transition states [64]. In theory, they would give access to free-energy barriers that are not seen in equilibrium experiments. However, integrating them in an ensemble or force-field refinement is non-trivial and is an active field of research (see, e.g., [65]).

Another critical issue is the parametrization of forward models used to connect ensembles and experiments. For example, several sets of Karplus equations have been proposed to back-calculate ^3J scalar couplings (see, e.g., the supporting information of Ref. [45]), which would lead to different results. Some of these models were developed considering reference static structures and might benefit from a revisiting with

Box 1: Our wish list for the next few years.

- More systematic use of explicit solvent and enhanced sampling simulations in combination with solution experiments.
- Further improvements in RNA force fields trained to reproduce solution experiments.
- Integration of kinetic data in ensemble or force-field refinement.
- Systematic optimization of forward models to back-calculate experimental observables.
- Development of new forward models to link chemical-probing experiments and three-dimensional dynamics.
- Explicit reconstruction of RNA dynamics from cryo-electron microscopy data.
- Simultaneous integration of experiments that probe RNA dynamics at different levels.
- Combination of experiments and deep-learning structure-prediction methods.

dynamics in mind [66]. Current Karplus equations, with their uncertainty, can thus be used to validate structures but might not be accurate enough for extracting quantitative populations. For SAXS experiments, different models include solvent effects to a different degree [67]. Chemical probing experiments are extremely cheap and scalable. Although several attempts to rationalize their signal in terms of atomistic structure and dynamics have been done in the past years [68–72], quantitative forward models are not available yet so their inclusion in MD simulations is usually done in terms of secondary-structure constraints [2, 28, 31]. Thermal denaturation experiments are also cheap and scalable, and can provide quantitative reference populations. Single-molecule Förster resonance energy transfer (smFRET) experiments provide access to dynamics. Notably, there has been some recent improvement in the back-calculation of smFRET from RNA ensembles [11], with an application reported for an RNA tetraloop receptor [73]. Cryo-electron-microscopy data have been mostly used for RNA structure determination so far. Methods explicitly modeling conformational dynamics in the back-calculation of density maps have been used in protein systems [74] and could enable a more accurate description of RNA dynamics.

Some of the mentioned experimental techniques can probe dynamics exclusively at a local or at a global scale, e.g., chemical probing and ^3J scalar couplings vs residual-dipolar-couplings, smFRET and SAXS. Given the ubiquitous risk of overfitting, we suggest that, whenever possible, complementary information should be combined so as to enhance the dynamical description at multiple scales, as it has been done for instance in Refs. [24, 34, 35].

Importantly, there are connections between the fields of RNA dynamics and RNA structure prediction. Protein structure prediction has been recently revolutionized by the advent of AlphaFold [75], which has also been suggested to be useful in dissecting protein dynamics [76]. The use of deep-learning methods in RNA structure prediction has been limited to date [77]. All the techniques developed for integrating MD simulations and solution experiments might be transferred to such tools when/if they will be widely available and tested.

In conclusion, the combination of MD simulations and experimental data can be seen from a dual perspective. From the computational side, experiments increase the quality of the generated dynamical ensembles. From the experimental side, simulations are a tool for interpreting the averaged signals obtained from solution measurements. Overall, MD simulations and experimental information can be synergistically combined filling their respective gaps, giving access to detailed and accurate descriptions of RNA dynamics that are instead unreachable using a single approach.

Acknowledgements

Sandro Bottaro and Serdal Kirmizialtin are acknowledged for reading the manuscript and providing useful suggestions.

References

[1] Q. Vicens, J. S. Kieft, Thoughts on how to think (and talk) about RNA structure, Proc. Natl. Acad. Sci. U.S.A. 119 (17) (2022) e2112677119,
•This recent perspective article provides an insightful discussion on RNA structure and dynamics and the relationship to function. It is an excellent introduction to the major RNA structural features.

[2] L. Cheng, E. N. White, N. L. Brandt, A. M. Yu, A. A. Chen, J. B. Lucks, Cotranscriptional RNA strand exchange underlies the gene regulation mechanism in a purine-sensing transcriptional riboswitch, Nucleic Acids Res. (2022) https://doi.org/10.1093/nar/gkac102.

[3] L. R. Ganser, J. Lee, A. Rangadurai, D. K. Merriman, M. L. Kelly, A. D. Kansal, B. Sathyamoorthy, H. M. Al-Hashimi, High-performance virtual screening by targeting a high-resolution RNA dynamic ensemble, Nat. Struct. Mol. Biol. 25 (5) (2018) 425–434.

[4] S. C. Abeysirigunawardena, H. Kim, J. Lai, K. Ragunathan, M. C. Rappé, Z. Luthey-Schulten, T. Ha, S. A. Woodson, Evolution of protein-coupled RNA dynamics during hierarchical assembly of ribosomal complexes, Nat. Commun. 8 (1) (2017) 1–9.

[5] M. Bernetti, K. B. Hall, G. Bussi, Reweighting of molecular simulations with explicit-solvent SAXS restraints elucidates ion-dependent RNA ensembles, Nucleic Acids Res. 49 (14) (2021) e84–e84,
••Combination of MD simulations and SAXS experimental data using the maximum entropy principle. In this work, the experimental spectra were quantitatively used to obtain the relative weight of compact and extended confromations of a complex and relatively large (around 60 nucleotides) RNA structure.

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

[7] B. Liu, H. Shi, H. M. Al-Hashimi, Developments in solution-state NMR yield broader and deeper views of the dynamic ensembles of nucleic acids, Curr. Opin. Struct. Biol. 70 (2021) 16–25.

[8] K. Kappel, K. Zhang, Z. Su, A. M. Watkins, W. Kladwang, S. Li, G. Pintilie, V. V. Topkar, R. Rangan, I. N. Zheludev, et al., Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures, Nat. Methods 17 (7) (2020) 699–707.

[9] Y. Chen, L. Pollack, SAXS studies of RNA: structures, dynamics, and interactions with partners, Wiley Interdiscip. Rev. RNA 7 (4) (2016) 512–526.

[10] K. M. Weeks, SHAPE directed discovery of new functions in large RNAs, Acc. Chem. Res. 54 (10) (2021) 2502–2517.

[11] F. D. Steffen, R. K. Sigel, R. Börner, FRETraj: integrating single-molecule spectroscopy with molecular dynamics, Bioinformatics 37 (21) (2021) 3953–3955.

[12] M. Andronescu, A. Condon, D. H. Turner, D. H. Mathews, The determination of RNA folding nearest neighbor parameters, Vol. 1097 of Methods Mol. Biol., Humana Press, Totowa, NJ, 2014, pp. 45–70.

[13] P. Rissone, C. V. Bizarro, F. Ritort, Stem–loop formation drives RNA folding in mechanical unzipping experiments, Proc. Natl. Acad. Sci. U.S.A. 119 (3) (2022) e2025575119.

[14] J. Šponer, G. Bussi, M. Krepl, P. Banáš, S. Bottaro, R. A. Cunha, A. Gil-Ley, G. Pinamonti, S. Poblete, P. Jurečka, N. G. Walter, M. Otyepka, RNA structural dynamics as captured by molecular simulations: a comprehensive overview, Chem. Rev. 118 (8) (2018) 4177–4338.

[15] J. Hénin, T. Lelièvre, M. R. Shirts, O. Valsson, L. Delemotte, Enhanced sampling methods for molecular dynamics simulations, arXiv preprint arXiv:2202.04164 (2022).

[16] K. Mráziková, V. Mlỳnskỳ, P. Kührová, P. Pokorná, H. Kruse, M. Krepl, M. Otyepka, P. Banáš, J. Šponer, UUCG RNA tetraloop as a formidable force-field challenge for MD simulations, J. Chem. Theory Comput. 16 (12) (2020) 7601–7617.

[17] S. Orioli, A. H. Larsen, S. Bottaro, K. Lindorff-Larsen, How to learn from inconsistencies: Integrating molecular simulations with experimental data, Prog. Mol. Biol. Transl. Sci. 170 (2020) 123–176.

[18] D. Medeiros Selegato, C. Bracco, C. Giannelli, G. Parigi, C. Luchinat, L. Sgheri, E. Ravera, Comparison of different reweighting approaches for the calculation of conformational variability of macromolecules from molecular simulations, ChemPhysChem 22 (1) (2021) 127–138.

[19] T. Fröhlking, M. Bernetti, N. Calonaci, G. Bussi, Toward empirical force fields that match experimental observables, J. Chem. Phys. 152 (23) (2020) 230902.

[20] A. Taghavi, I. Riveros, D. J. Wales, I. Yildirim, Evaluating geometric definitions of stacking for RNA dinucleoside monophosphates using molecular mechanics calculations, J. Chem. Theory Comput. 18 (2022) 3637–3653.

[21] J. Zhao, S. D. Kennedy, K. D. Berger, D. H. Turner, Nuclear magnetic resonance of single-stranded RNAs and DNAs of CAAU and UCAAUC as benchmarks for molecular dynamics simulations, J. Chem. Theory Comput. 16 (3) (2020) 1968–1984,
•• This work reports nuclear magnetic resonance reference data for RNA and DNA hexamers. The study, which is part of a series of works from the same lab, reports data extremely valuable for benchmark and tuning of molecular dynamics models on unstructured oligomers.

[22] J. Zhao, S. D. Kennedy, D. H. Turner, Nuclear magnetic resonance spectra and AMBER OL3 and ROC-RNA simulations of UCUCGU reveal force field strengths and weaknesses for single-stranded RNA, J. Chem. Theory Comput. 18 (2) (2022) 1241–1254.

[23] V. Mlỳnskỳ, P. Kührová, T. Kühr, M. Otyepka, G. Bussi, P. Banáš, J. Šponer, Fine-tuning of the AMBER RNA force field with a new term adjusting interactions of terminal nucleotides, J. Chem. Theory Comput. 16 (6) (2020) 3936–3946.

[24] W. He, N. Naleem, D. Kleiman, S. Kirmizialtin, Refining the RNA force field with small-angle X-ray scattering of helix–junction–helix RNA, J. Phys. Chem. Lett. 13 (2022) 3400–3408.

[25] Y. Tanida, A. Matsuura, Alchemical free energy calculations via metadynamics: Application to the theophylline-RNA aptamer complex, J. Comput. Chem. 41 (20) (2020) 1804–1819.

[26] T. Hurst, S.-J. Chen, Deciphering nucleotide modification-induced structure and stability changes, RNA Biol. 18 (11) (2021) 1920–1930.

[27] L. Baronti, I. Guzzetti, P. Ebrahimi, S. Friebe Sandoz, E. Steiner, J. Schlagnitweit, B. Fromm, L. Silva, C. Fontana, A. A. Chen, et al., Base-pair conformational switch modulates miR-34a targeting of Sirt1 mRNA, Nature 583 (7814) (2020) 139–144,
••The authors used MD simulations with restraints based on the secondary structure information obtained from NMR experimental data, so as to enforce a qualitative agreement with the experiments. The approach exploits two-dimensional replica-exchange MD. In this work, the method was used to study the interaction between a miRNA and its mRNA target.

[28] S. Bottaro, G. Bussi, K. Lindorff-Larsen, Conformational ensembles of noncoding elements in the SARS-CoV-2 genome from molecular dynamics simulations, J. Am. Chem. Soc. 143 (22) (2021) 8333–8343.

[29] M. Krepl, F. F. Damberger, C. von Schroetter, D. Theler, P. Pokorná, F. H.-T. Allain, J. Šponer, Recognition of N6-methyladenosine by the YTHDC1 YTH domain studied by molecular dynamics and NMR spectroscopy: the role of hydration, J. Phys. Chem. B 125 (28) (2021) 7691–7705.

[30] A. Cléry, M. Krepl, C. K. Nguyen, A. Moursy, H. Jorjani, M. Katsantoni, M. Okoniewski, N. Mittal, M. Zavolan, J. Šponer, et al., Structure of SRSF1 RRM1 bound to RNA reveals an unexpected bimodal mode of interaction and explains its involvement in SMN1 exon7 splicing, Nat. Commun. 12 (1) (2021) 1–12.

[31] A. M. Yu, P. M. Gasper, L. Cheng, L. B. Lai, S. Kaur, V. Gopalan, A. A. Chen, J. B. Lucks, Computationally reconstructing cotranscriptional RNA folding from experimental data reveals rearrangement of non-native folding intermediates, Mol. Cell 81 (4) (2021) 870–883.

[32] S. L. Bonilla, M. E. Sherlock, A. MacFadden, J. S. Kieft, A viral RNA hijacks host machinery using dynamic conformational changes of a tRNA-like structure, Science 374 (6570) (2021) 955–960.

[33] S. L. Bonilla, Q. Vicens, J. S. Kieft, Cryo-EM reveals an entangled kinetic trap in the folding pathway of a catalytic RNA, Sci. Adv. 8 (34) (2022) eabq4144.

[34] S. Bottaro, P. J. Nichols, B. Vögeli, M. Parrinello, K. Lindorff-Larsen, Integrating NMR and simulations reveals motions in the UUCG tetraloop, Nucleic Acids Res. 48 (11) (2020) 5839–5848,
••Combination of MD simulations and a large number of heterogeneous NMR experimental data using the maximum entropy principle through a reweighting procedure. A non-negligible population for an alternative conformation of the UUCG tetraloop was identified. The work includes a detailed and clear discussion on how to conduct reweighting using the maximum entropy principle.

[35] C. Bergonzo, A. V. Grishaev, S. Bottaro, Conformational heterogeneity of UCAAUC RNA oligonucleotide from molecular dynamics simulations, SAXS, and NMR experiments, RNA 28 (7) (2022) 937–946,
•Combination of MD simulations and experimental data using the maximum entropy principle through a reweighting procedure.

[36] S. Reißer, S. Zucchelli, S. Gustincich, G. Bussi, Conformational ensembles of an RNA hairpin using molecular dynamics and sparse NMR data, Nucleic Acids Res. 48 (3) (2020) 1164–1174.

[37] W. He, Y.-L. Chen, L. Pollack, S. Kirmizialtin, The structural plasticity of nucleic acid duplexes revealed by WAXS and MD, Sci. Adv. 7 (17) (2021) eabf6106,
••Combination of MD simulations and WAXS experimental data using the sample-and-select approach. The work allowed reconstructing conformational ensembles of RNA and DNA duplexes.

[38] Y.-L. Chen, W. He, S. Kirmizialtin, L. Pollack, Insights into the structural stability of major groove RNA triplexes by WAXS-

guided MD simulations, Cell Rep. Phys. Sci. 3 (7) (2022) 100971.

[39] H. Shi, A. Rangadurai, H. Abou Assi, R. Roy, D. A. Case, D. Herschlag, J. D. Yesselman, H. M. Al-Hashimi, Rapid and accurate determination of atomistic RNA dynamic ensemble models using NMR and structure prediction, Nat. Commun. 11 (1) (2020) 1–14,
••Combination of the FARFAR method and NMR experimental data using the sample-and-select approach. The conformational ensemble reconstructed for the HIV-1 TAR RNA was reported to be more accurate when using FARFAR- than MD-generated structures.

[40] T. Fröhlking, V. Mlỳnskỳ, M. Janeček, P. Kührová, M. Krepl, P. Banáš, J. Šponer, G. Bussi, Automatic learning of hydrogen-bond fixes in an AMBER RNA force field, J. Chem. Theory Comput. 18 (7) (2022) 4490–4502,
•This work integrates MD simulations and experimental data from NMR and X-ray reference structures to improve force field parameters. The procedure was based on a minimization strategy to optimize hydrogen-bond corrections. Several tetramers and tetraloops were taken into account in the study, including the challenging UUCG tetraloop.

[41] V. Piomponi, T. Fröhlking, M. Bernetti, G. Bussi, Molecular simulations matching denaturation experiments for N6-methyladenosine, ACS Cent. Sci 8 (8) (2022) 1218–1228,
••The study integrates MD simulations and thermodynamic data measured in denaturation experiments to improve force field parameters. In particular, partial charges and a torsional potential of the methylated adenosine (m6A) were optimized over experimental ΔΔGs for a number of RNA duplexes. Notably, minimal modifications to partial charges allowed for a significantly improved agreement with experiment.

[42] A. M. Salsbury, J. A. Lemkul, Recent developments in empirical atomistic force fields for nucleic acids and applications to studies of folding and dynamics, Curr. Opin. Struct. Biol. 67 (2021) 9–17.

[43] J. Chen, H. Liu, X. Cui, Z. Li, H.-F. Chen, RNA-specific force field optimization with CMAP and reweighting, J. Chem. Inf. Model. 62 (2022) 372–385.

[44] F. S. Ezra, C.-H. Lee, N. S. Kondo, S. S. Danyluk, R. H. Sarma, Conformational properties of purine-pyrimidine and pyrimidine-purine dinucleoside monophosphates, Biochemistry 16 (9) (1977) 1977–1987.

[45] D. E. Condon, S. D. Kennedy, B. C. Mort, R. Kierzek, I. Yildirim, D. H. Turner, Stacking in RNA: NMR of four tetramers benchmark molecular dynamics, J. Chem. Theory Comput. 11 (6) (2015) 2729–2742.

[46] I. Yildirim, H. A. Stern, J. D. Tubbs, S. D. Kennedy, D. H. Turner, Benchmarking AMBER force fields for RNA: Comparisons to NMR spectra for single-stranded r(GACC) are improved by revised χ torsions, J. Phys. Chem. B 115 (29) (2011) 9261–9270.

[47] J. D. Tubbs, D. E. Condon, S. D. Kennedy, M. Hauser, P. C. Bevilacqua, D. H. Turner, The nuclear magnetic resonance of CCCC RNA reveals a right-handed helix, and revised parameters for AMBER force field torsions improve structural predictions from molecular dynamics, Biochemistry 52 (6) (2013) 996–1010.

[48] Y.-L. Chen, T. Lee, R. Elber, L. Pollack, Conformations of an RNA helix-junction-helix construct revealed by SAXS refinement of MD simulations, Biophys. J. 116 (1) (2019) 19–30.

[49] R. D. Jenison, S. C. Gill, A. Pardi, B. Polisky, High-resolution molecular discrimination by RNA, Science 263 (5152) (1994) 1425–1429.

[50] C. Roost, S. R. Lynch, P. J. Batista, K. Qu, H. Y. Chang, E. T. Kool, Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification, J. Am. Chem. Soc. 137 (5) (2015) 2107–2115.

[51] A. Wacker, J. E. Weigand, S. R. Akabayov, N. Altincekic, J. K. Bains, E. Banijamali, O. Binas, J. Castillo-Martinez, E. Cetiner, B. Ceylan, et al., Secondary structure determination of conserved SARS-CoV-2 RNA elements by NMR spectroscopy, Nucleic Acids Res. 48 (22) (2020) 12415–12435.

[52] S. Nozinovic, B. Fürtig, H. R. Jonker, C. Richter, H. Schwalbe, High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA, Nucleic Acids Res. 38 (2) (2010) 683–694.

[53] 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.

[54] C. Hartlmueller, J. C. Guenther, A. C. Wolter, J. Woehnert, M. Sattler, T. Madl, RNA structure refinement using NMR solvent accessibility data, Sci. Rep. 7 (1) (2017) 1–10.

[55] P. J. Nichols, M. A. Henen, A. Born, D. Strotz, P. Güntert, B. Vögeli, High-resolution small RNA structures from exact nuclear Overhauser enhancement measurements without additional restraints, Commun. Biol. 1 (1) (2018) 1–11.

[56] P. Podbevšek, F. Fasolo, C. Bon, L. Cimatti, S. Reißer, P. Carninci, G. Bussi, S. Zucchelli, J. Plavec, S. Gustincich, Structural determinants of the SINE B2 element embedded in the long noncoding RNA activator of translation AS Uchl1, Sci. Rep. 8 (1) (2018) 1–13.

[57] R. Welty, S. A. Pabit, A. M. Katz, G. D. Calvey, L. Pollack, K. B. Hall, Divalent ions tune the kinetics of a bacterial GTPase center rRNA folding transition from secondary to tertiary structure, RNA 24 (12) (2018) 1828–1838.

[58] Q. Zhang, A. C. Stelzer, C. K. Fisher, H. M. Al-Hashimi, Visualizing spatially correlated dynamics that directs RNA conformational transitions, Nature 450 (7173) (2007) 1263–1267.

[59] L. Salmon, G. Bascom, I. Andricioaei, H. M. Al-Hashimi, A general method for constructing atomic-resolution RNA ensembles using NMR residual dipolar couplings: the basis for interhelical motions revealed, J. Am. Chem. Soc. 135 (14) (2013) 5457–5466.

[60] E. Kierzek, X. Zhang, R. M. Watson, S. D. Kennedy, M. Szabat, R. Kierzek, D. H. Mathews, Secondary structure prediction for RNA sequences including N6-methyladenosine, Nat. Commun. 13 (1) (2022) 1–10.

[61] Y. Du, I. Mordatch, Implicit generation and modeling with energy-based models, in: H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alche Buc, E. Fox, R. Garnett (Eds.), Adv. Neural. Inf. Process. Syst., Vol. 32, 2019.

[62] D. S. Cerutti, D. A. Case, Molecular dynamics simulations of macromolecular crystals, Wiley Interdiscip. Rev. Comput. Mol. Sci 9 (4) (2019) e1402.

[63] Y. Qiu, D. G. Smith, S. Boothroyd, H. Jang, D. F. Hahn, J. Wagner, C. C. Bannan, T. Gokey, V. T. Lim, C. D. Stern, et al., Development and benchmarking of Open Force Field v1.0.0—the Parsley small-molecule force field, J. Chem. Theory Comput. 17 (10) (2021) 6262–6280.

[64] G. Han, Y. Xue, Rational design of hairpin RNA excited states reveals multi-step transitions, Nat. Commun. 13 (1) (2022) 1–12,
•This study exploited NMR relaxation dispersion experiments to study the kinetics underlying transient states of small hairpin RNAs. The work highlights how kinetic information on RNA molecules can be obtained from experiments, which could possibly be integrated with MD simulations to study RNA conformational dynamics.

[65] Z. F. Brotzakis, M. Vendruscolo, P. G. Bolhuis, A method of incorporating rate constants as kinetic constraints in molecular dynamics simulations, Proc. Natl. Acad. Sci. U.S.A. 118 (2) (2021) e2012423118.

[66] K. Lindorff-Larsen, R. B. Best, M. Vendruscolo, Interpreting dynamically-averaged scalar couplings in proteins, J. Biomol. NMR 32 (4) (2005) 273–280.

[67] M. Bernetti, G. Bussi, Comparing state-of-the-art approaches to back-calculate SAXS spectra from atomistic molecular dynamics simulations, Eur. Phys. J. B 94 (9) (2021) 1–8.

[68] G. Pinamonti, S. Bottaro, C. Micheletti, G. Bussi, Elastic network models for RNA: a comparative assessment with molecular dynamics and shape experiments, Nucleic Acids Res. 43 (15) (2015) 7260–7269.

[69] V. Mlỳnskỳ, G. Bussi, Molecular dynamics simulations reveal an interplay between SHAPE reagent binding and RNA flexibility, J. Phys. Chem. Lett. 9 (2) (2018) 313–318.

[70] T. Hurst, X. Xu, P. Zhao, S.-J. Chen, Quantitative understanding of SHAPE mechanism from RNA structure and dynamics analysis, J. Phys. Chem. B 122 (18) (2018) 4771–4783.

[71] E. Frezza, A. Courban, D. Allouche, B. Sargueil, S. Pasquali, The interplay between molecular flexibility and RNA chemical probing reactivities analyzed at the nucleotide level via an extensive molecular dynamics study, Methods 162 (2019) 108–127.

[72] T. Hurst, S.-J. Chen, Sieving RNA 3D structures with SHAPE and evaluating mechanisms driving sequence-dependent reactivity bias, J. Phys. Chem. B 125 (4) (2021) 1156–1166.

[73] F. Erichson, F. D. Steffen, R. Börner, FRET-assisted structural model of the GAAA RNA tetraloop receptor, no. 002 in 26. Interdisziplinäre Wissenschaftliche Konferenz Mittweida, Hochschule Hochschule Mittweida, Mittweida, 2021, pp. 230–233. doi:10.48446/opus-12283.

[74] R.-A. A. Garibsingh, E. Ndaru, A. A. Garaeva, Y. Shi, L. Zielewicz, P. Zakrepine, M. Bonomi, D. J. Slotboom, C. Paulino, C. Grewer, et al., Rational design of ASCT2 inhibitors using an integrated experimental-computational approach, Proc. Natl. Acad. Sci. U.S.A. 118 (37) (2021) e2104093118.

[75] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al., Highly accurate protein structure prediction with AlphaFold, Nature 596 (7873) (2021) 583–589.

[76] N. J. Fowler, M. P. Williamson, The accuracy of protein structures in solution determined by AlphaFold and NMR, Structure 30 (7) (2022) 925–933.

[77] R. J. Townshend, S. Eismann, A. M. Watkins, R. Rangan, M. Karelina, R. Das, R. O. Dror, Geometric deep learning of RNA structure, Science 373 (6558) (2021) 1047–1051.