FRET-guided selection of RNA 3D structures

Mirko Weber[1], Felix Erichson[1], Maciej Antczak[2,3], Tomasz Zok[3], Fabio D. Steffen[4], Marta Szachniuk[2,3], and Richard Börner[1,*]

[1]Laserinstitut Hochschule Mittweida, University of Applied Sciences Mittweida, Technikumplatz 17, 09648, Mittweida, Germany

[2]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704, Poznan, Poland

[3]Institute of Computing Science, Poznan University of Technology, Piotrowo 2, 60-965, Poznan, Poland

[4]Department of Oncology, University of Zurich, University Children's Hospital, 8008, Zurich, Switzerland

[*]Corresponding author: `richard.boerner@hs-mittweida.de` - Tel: +49 3727 58 1009

September 24, 2025

Abstract

Integrative biomolecular modeling of RNA relies on structural refined collections and accurate experimental data that reflect binding and folding behavior. However, the prediction of such collections remains challenging due to the rugged energy landscape and extensive conformational heterogeneity of large RNAs. To overcome these limitations, we applied a FRET-guided strategy to identify RNA conformational states consistent with single-molecule FRET (smFRET) experiments. We predicted 3D structures of a ribosomal RNA tertiary contact comprising a GAAA tetraloop and a kissing loop using three popular RNA 3D modeling tools namely RNAComposer, FARFAR2, and AlphaFold3, yielding a collection of candidate conformations. These models were structurally validated based on Watson-Crick base-pairing patterns and filtered using an eRMSD threshold. For each retained structure, we computed the accessible contact volume (ACV) of the sCy3/sCy5 dye pair using FRETraj to predict FRET distributions. These distributions were then compared and weighted against experimental smFRET data to identify conformational states compatible with the observed FRET states. Our results demonstrate that experimental transfer efficiencies can be reproduced using *in silico* predicted RNA 3D structures. This FRET-guided workflow, combined with structural validation, lays the foundation for capturing the highly diverse conformational states characteristic of flexible RNA motifs.

1 Introduction

Three-dimensional (3D) structure prediction and molecular dynamics (MD) simulations have been instrumental in resolving biomolecular structures at atomic resolution [1]. While these methods have been particularly impactful in protein structure prediction and dynamics [2, 3, 4], attention is increasingly shifting towards RNA and RNA-protein complexes, due to their structural complexity and functional diversity [5, 6]. In response, RNA 3D structure prediction has become a rapidly evolving field, with sustained efforts aimed at enhancing accuracy, reliability, and robustness. A key development in this area was the launch of RNA-Puzzles [7], a community-wide initiative that facilitates blind benchmarking of RNA 3D prediction tools against experimentally determined reference structures [8, 9, 10, 11]. This effort has significantly accelerated the field by enabling objective performance evaluation and encouraging methodological advancements [12]. Despite a growing range of computational strategies, modeling RNA structures remain dependent on experimental data from high-resolution techniques such as X-ray crystallography [13], NMR spectroscopy [14], or cryo-EM [15]. As the demand for accurate modeling of flexible and functionally relevant RNA conformations continues to grow, the integration of predictive modeling with experimental validation, such as FRET, remains a cornerstone of progress in RNA structural biology.

Although 3D structure prediction and MD simulations provide critical insights into the three-dimensional organization of RNA, they often focus on identifying a single, energetically favorable state or sample a limited subspace of the full conformational landscape. This single-state representation is inadequate to capture the structural heterogeneity observed in many functional RNAs, where multiple transiently populated conformations co-exist. [16]. To address these limitations, in-solution techniques such as SAXS [17, 18], EPR [19, 20], and FRET [21, 22] serve as complementary approaches for probing conformationally heterogeneous RNA structure collections. These methods enable structural characterization and are particularly well-suited for detecting dynamic behavior [23, 24]. EPR and single-molecule FRET (smFRET), in particular, provide access to highly precise, site-specific distance information that is critical for resolving low-populated or transient states. Within integrative or hybrid modeling approaches [25, 26], such distance constraints offer an experimental benchmark for selecting and validating conformations from predicted structure collections. Among available experimental techniques, smFRET combines nanometer-scale resolution with single-molecule sensitivity, enabling direct comparison between predicted and experimental distances or whole distance distributions [27], which is an ideal technique for detecting conformational subpopulations and conformational changes to capture the dynamic behavior of RNAs.

RNA folding is a dynamic, hierarchical process, where proteins assist folding by stabilizing native-like intermediates, and tertiary interactions are essential for compact RNA architecture [28]. In particular, tertiary contacts between distant secondary structure elements frequently form only transiently or in the presence of stabilizing factors such as Mg (II) ions or RNA-binding proteins [29, 30, 31]. This structural plasticity, i.e., the intrinsic flexibility of RNA, defines a collection of unbound states. Characterizing this conformational collection experimentally remains challenging, due to its high structural heterogeneity. Single-molecule techniques such as smFRET offer a powerful approach to probe this heterogeneity [22], as they can resolve subpopulations within the collection and guide structure prediction by identifying conformations consistent with observed distance constraints [32, 27, 33].

Here, we apply FRET-guided integrative modeling to an ribosomal RNA model construct, featuring a kissing loop (KL) and highly flexible GAAA tetraloop (TL_{GAAA}) domain, which can serve as a tertiary contact binding to the KL [31], that we aim to characterize structurally in its unbound state. Using RNAComposer [34, 35], FARFAR2 [36], and AlphaFold3 [2] we generated a

collection of RNA 3D models based on diverse structural prediction strategies. For model validation, we calculated multiple ACV [37, 38, 27] to predict *in silico* FRET from the dye-labeled RNA models and used these to identify an initial pool of candidate structures matching the experimental smFRET distribution. To ensure structural plausibility, we applied established validation metrics, including Watson–Crick (WC) base-pairing [29] analysis and the eRMSD score [39, 40], yielding a refined collection suitable for smFRET-guided conformer selection.

Ultimately, we used the experimentally measured smFRET distribution to guide structure selection from the validated collections. Instead of comparing FRET values post hoc, we applied the distribution as a filter to extract conformers whose ACV-based predictions matched the experimental data. By enabling targeted sampling within the conformational landscape resolved by smFRET, this approach advances RNA structural modeling beyond static representations toward
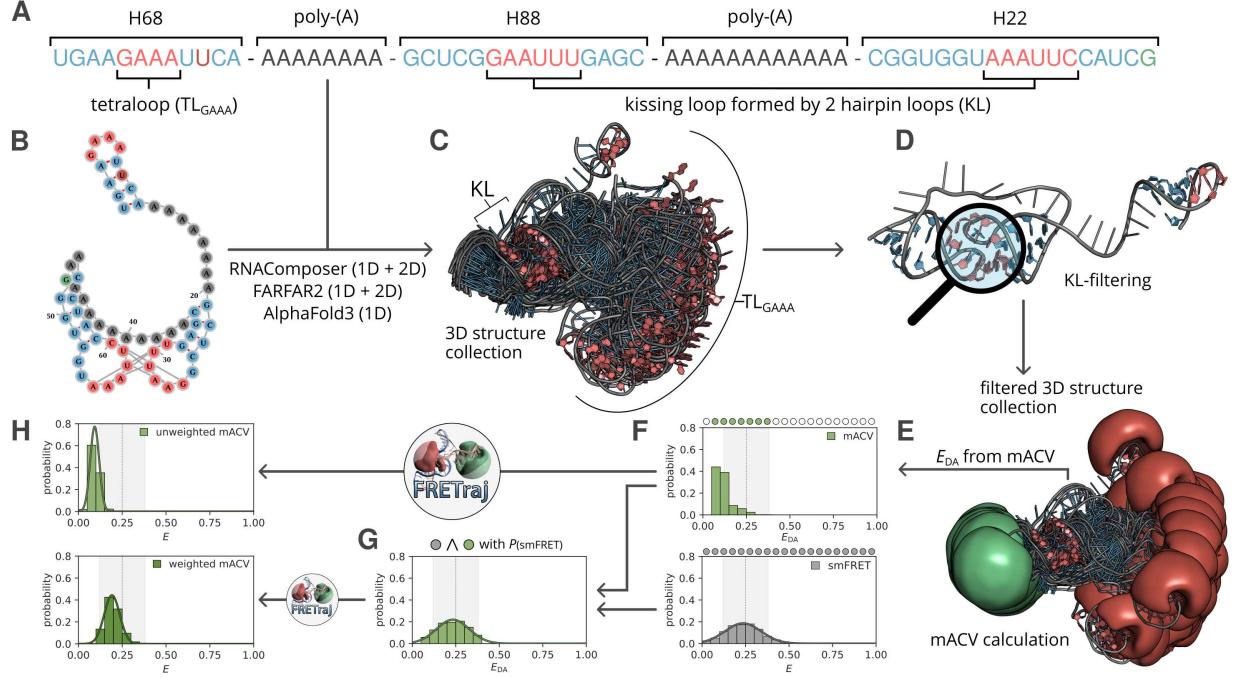

Figure 1: Representation of the workflow from sequence and secondary structure to the modeling of a final 3D structure collection that best represents the experimental smFRET data under the condition of a correctly folded KL. (**A**) Sequence showing all structure elements of the investigated model construct. (**B**) Secondary structure representation of the construct. (**C**) A complete structure collection of 1,000 structures predicted by RNAComposer. (**D**) Evaluation of predicted structures with Barnaba to ensure correctly folded KLs in each structure with accurate WC base-pairing and the calculation of the eRMSD to the reference structure. (**E**) For each filtered structure, the mACVs of the dyes sCy3 and sCy5 were predicted, and the mean FRET efficiency was calculated. (**F**) Distribution of the E_{DA} values calculated from the mACVs alongside the experimental smFRET distribution, each represented with their respective bin populations. (**G**) Reweighting of the structure collection, ensuring that structures are selected from each populated bin with the same probability as observed in the smFRET distribution. (**H**) The comparison of FRET efficiencies after photon sampling with FRETraj for both the unweighted E_{DA} distribution and the weighted distribution with the probabilities from the experiment.

a dynamic, experimentally grounded collection.

2   Methods

2.1   Probing the unbound state with smFRET

smFRET measurements of the Cy3/Cy5-labeled ribosomal RNA model construct, designed to probe KL formation and potential TL_{GAAA} binding, were performed and analyzed according to standard protocols [27, 23]. The construct was prepared as described in [31] (Figure 1A and B). For molecular sorting POE-type smFRET experiments were performed in standard buffer containing $[\text{K}^+] = 116\,\text{mmol}\,\text{L}^{-1}$ and EDTA to allow KL formation without TL_{GAAA} binding. The smFRET data was further corrected for background, bleed through, and direct excitation. γ-correction was applied to account for differences in detection efficiency and quantum yield of the FRET pair. These corrections are essential for the comparability of the experimental data with FRET distributions derived from *in silico* structure collections using FRETraj [38], together with the experimental burst size distribution (Supplementary Figure 3), i.e., the sum of donor and acceptor intensities (Figure 1F).

2.2   *In silico* RNA 3D structure prediction and MD simulation

Structure collections consisting of 10,000 structures each were predicted using RNAComposer, FARFAR2 and AlphaFold3 (Supplementary Methods). Additionally, we performed six independent 1 µs MD simulations with different initial seed structures chosen to represent structurally diverse conformations (Supplementary Methods). An equal number of structures was sampled from each of the six MD simulations for comparison with the 3D prediction tools and for subsequent use in the FRET-guided reweighting. ACVs of both dyes were calculated for each structure predicted by the 3D prediction tools as well as every 100 ps along the MD trajectories. Photon emission events were simulated by FRETraj [38, 41, 42]. The FRET distributions were also analyzed both individually for each MD simulation (Supplementary Figure 2) and as a combined collection (Figure 4A and B - MD simulation). All MD simulations were annotated every 100 ps based on Watson–Crick (WC) base-pairing using the Barnaba toolbox [40], resulting in six sets of 10,000 annotated structures representing the unbound RNA conformation (averaged in Figure 3A right).

2.3   Estimating structure collection size with the Kullback-Leibler divergence

To estimate the minimum number of structures required for our structure collections to be representative of a collection of $N = 10,000$ structures, we compared the full structure collections probability distribution P of the mean transfer efficiencies $\langle E_{\text{DA}} \rangle$ calculated via ACVs with sub-distributions sampled Q_{x_i} of each prediction tool. For each i, where $i \in \{10, 20, 30, \ldots, 9990\}$, x_i represents the number of samples drawn from the structure collection of size N and for each x_i the structures were randomly sampled from N. To compare the probability distribution we applied the Kullback-Leibler divergence (KLD) [43], thus for each repetition it is computed as:

$$D_{\text{KLD}}(P \| Q_{x_i}^{(r)}) = \sum_{j=1}^{m} p_j \log \left(\frac{p_j}{q_j^{(r)}} \right),$$

where m is the total number of bins, p_j is the true probability of bin j, and $q_j^{(r)}$ is the corresponding probability of the bin from the sampled distribution. The mean KLD across all repetitions for a

given x_i is then defined as:

$$\overline{D_{\mathrm{KLD}}}(x_i) = \frac{1}{k_i} \sum_{r=1}^{k_i} D_{\mathrm{KLD}}(P \| Q_{x_i}^{(r)}).$$

where k_i was increased until the standard deviation of the corresponding KLD values was less than 1% of their mean. Based on this information, we determined the minimum number of structures required for each structure collection to achieve an accurate representation (Figure 2 and Supplementary Figure 1).

2.4 Validation of the KL formation and filtering of the structure collections

To ensure formation of the KL as observed in the reference structure (PDB-ID: 3JCT and Figure 3A), we calculated the WC base-pairs annotated by Barnaba [40] across all predicted structures with those from 10,000 MD-derived states. This served to validate the structural preservation of the KL motif. Subsequently, WC base-pairs were annotated for all structures from RNAComposer, FARFAR2, and AlphaFold3. Structures were classified as exhibiting preserved KL formation if all relevant base-pairs were canonical, except U27–A59, which was allowed to vary consistent with the reference structure. Any deviation led to annotation as altered KL formation (Figure 3C). In parallel, the structural similarity of each model to the reference was quantified using the eRMSD [39]. Models with eRMSD $\leq$ 0.8 and preserved base-pairing were selected for further analysis (Figure 3B left). To assess structural stability and confirm base-pair preservation after refinement, each filtered model was subjected to a short 1 ns MD simulation, followed by re-annotation and recomputation of WC base-pairs and eRMSD (Figure 3B right).

2.5 Reweighting of FRET efficiencies from mACVs

For the comparison of the structure collections obtained from the investigated 3D prediction tools and the described MD simulations with an experimental smFRET distribution, photon emission events were simulated using FRETraj based on two different structure sampling strategies. In the first approach, each structure was selected exactly once, thereby preserving the original distribution of predicted mean FRET efficiencies E_{DA}. Consequently, the relative frequency of structures in each bin B_i reflected the number of structures falling within the corresponding E_{DA} range, i.e.,

$$P_{\mathrm{unweighted}}(i) = \frac{|B_i|}{N},$$

where $|B_i|$ denoted the number of structures in bin i, and N was the total number of structures in the filtered collection (Figure 1H - unweighted mACV).

In the second approach, structures were sampled such that the resulting distribution approximated the experimental smFRET distribution p^{smFRET} (Figure 1H - weighted mACV):

$$P_{\mathrm{weighted}}(i) = \begin{cases} p_i^{\mathrm{smFRET}} & \text{if } |B_i| > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where p_i^{smFRET} denoted the smFRET probability of bin i, provided that B_i was not empty.

Based on this, two scenarios were distinguished for each bin:

Scenario 1: If the relative number of structures in B_i exceeded the experimental probability p_i^{smFRET}, a subset of structures was randomly sampled from B_i to match p_i^{smFRET}.

5

Scenario 2: If the number of structures in B_i was insufficient to satisfy p_i^{smFRET}, all structures in B_i were selected repeatedly until the required number was reached. Any remaining fraction was filled by randomly sampling additional structures from B_i. Each repetition was randomly permuted to construct an expanded set $\bar{B}_i$, from which sampling was performed to match p_i^{smFRET}.

3 Results

3.1 1,000 RNA 3D structures of the KL-TL$_{\text{GAAA}}$ are sufficient to capture experimental FRET distributions

Today, a wide range of *in silico* RNA 3D structure prediction tools is available, each employing different methodologies for structure modeling. However, they all share a key characteristic: the ability to generate RNA 3D structures with exceptional speed and efficiency. Nevertheless, it remains unknown how many structures are actually needed to accurately represent a diverse structure collection, particularly in the context of FRET-guided integrative modeling.

We compared three widely used tools, RNAComposer, FARFAR2, and AlphaFold3, and evaluated their ability to predict the unbound state of the KL-TL$_{\text{GAAA}}$ model construct. Each tool was used to generate 10,000 structures, serving as a baseline for evaluating the structural diversity of the unbound state (Figure 2). We generated subsets of varying sizes from each baseline dataset and assessed their similarity to the full E_{DA} distribution using the KLD.

Therein, the bin size of the FRET distribution is crucial. The more precisely we represent the
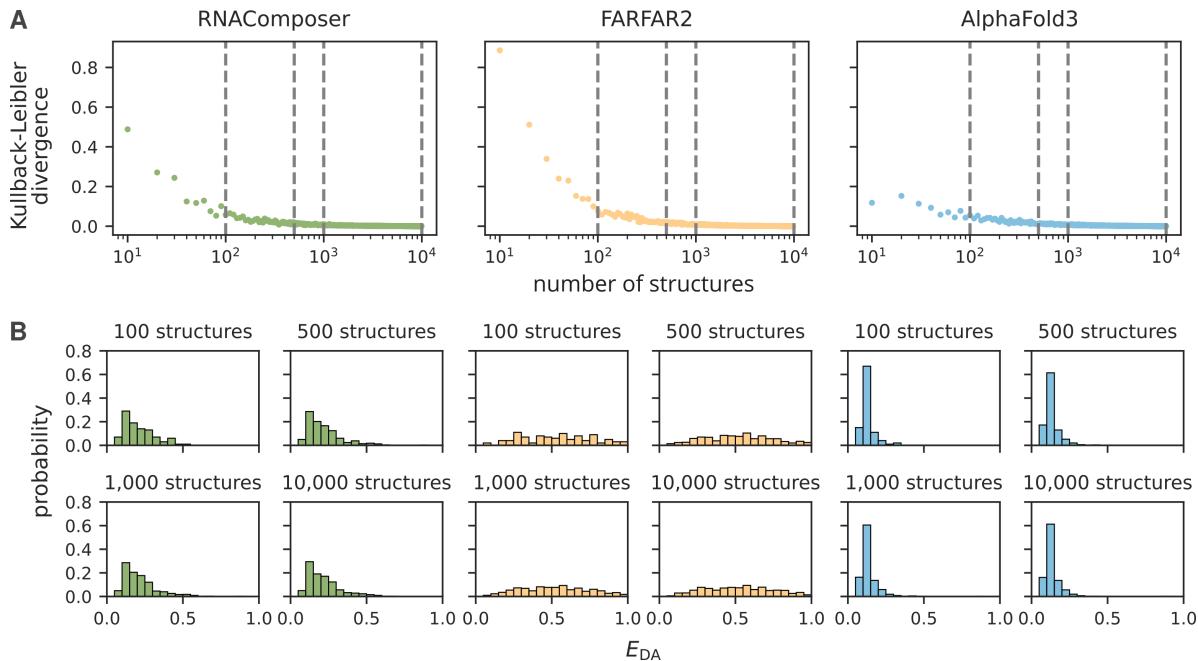

Figure 2: Distribution of FRET efficiencies calculated from mean donor–acceptor distances via ACVs for RNAComposer, FARFAR2, and AlphaFold3, shown alongside the initial reference distribution (10,000 structures). **(A)** KLD between subsets of 10 structures and the full 10,000 structure collections for all three tools. Dashed vertical lines indicate the subset sizes used for the E_{DA} distributions shown in B. **(B)** The distributions based on 10,000 structures denote the reference probability distributions.

FRET values the more structures are required to accurately resemble the baseline. A FRET bin size of 0.05 was chosen based on a benchmark study demonstrating an experimental accuracy of ΔE=$\pm$0.05 [23], and is used as a criterion throughout this work. Our results demonstrat, that the KLD converges toward zero across all E_{DA} bin sizes tested (Supplementary Figure 1). The minimum feasible subset is determined by the diversity of the original E_{DA} distribution. As a result, FARFAR2 requires more structures compared to RNAComposer and AlphaFold3 to fulfill the FRET criterion and to reproduce the baseline (Figure 2A). RNAComposer and AlphaFold3 show a starting convergence at a subset size of around 100 structures, while FARFAR2 requires a slightly larger subset due to a wider range of predicted E_{DA} values (Figure 2A and B). To ensure a sufficiently large set of structures for subsequent validation while also retaining enough conformers to fully represent the FRET distribution, we chose a structure collection size of 1,000 for each tool. This number is sufficient to model FRET in the unbound (low-FRET) state of the investigated construct but may differ for RNAs with greater structural complexity.

3.2 Tertiary structure annotations are critical for selecting reliable KLs

Since the proper fold of the tertiary contact in our construct depends on a properly foleded KL, we specifically analyzed our structure collections to ensure the accurate formation of the KL in accordance with a cryo-EM reference structure of the ribosomal RNA (PDB-ID: 3JCT). We first assessed the folding of the KL by analyzing the WC base-pairing pattern observed in the reference cryo-EM structure [44]. Our analysis revealed that the hydrogen-bonding distance between bases A27 and U59 exceeded the favorable range for stable WC base-pairing and therefore did not fall into any canonical base-pair classification. To investigate the stability and variability of this motif, we used all six 1 µs MD simulations of our model construct (Figure 3A). In contrast, these simulations show that all base pairings, including A27–U59, adopt canonical WC geometry, indicating improved pairing consistency in the MD collection. For subsequent classifications of the KL integrity, we established a criterion requiring WC base-pairings between all bases except A27-U59, for which we also allowed spatially proximal representations not classified under a specific WC base-pairing category.

To complement WC base-pairing analysis and refine structural characterization of the KL, we additionally computed the eRMSD for all predicted structures relative to the cryo-EM reference (PDB-ID: 3JCT). The eRMSD quantifies geometric differences between base pairs and stacking interactions, making it a robust metric for assessing RNA motif similarity. Across all prediction tools, a substantial number of structures clustered around an eRMSD of approximately 0.5 (Figure 3B left and Supplementary Figure 4 and 5 left). RNAComposer and FARFAR2 showed smoother transitions to higher eRMSD values, whereas AlphaFold3 exhibited a distinct cutoff. To introduce minor conformational variation and assess structural consistency, each structure was subjected to a short 1 ns MD simulation, followed by re-annotation of base pairings and recalculation of the eRMSD (Figure 3B right and Supplementary Figure 4 and 5 right). These simulations introduced local fluctuations in the KL, shifting some conformers above an eRMSD = 0.8 and increasing the number of non WC base-pairs.

Based on WC base-pairing analysis and eRMSD evaluation, we find that neither metric alone provides a sufficient criterion for verifying a preserved KL formation in our construct based on the reference structure. Using only the threshold of eRMSD $\leq$ 0.8 would retain structures that deviate from the canonical WC pattern and thus do not represent a preserved KL. Conversely, relying solely on WC base-pairing can include structures with significantly altered backbone geometries. However, applying both filters in combination consistently selects folded KL formations adapting WC base-pairing across all three structure prediction tools (Figure 3C and Supplementary Figure
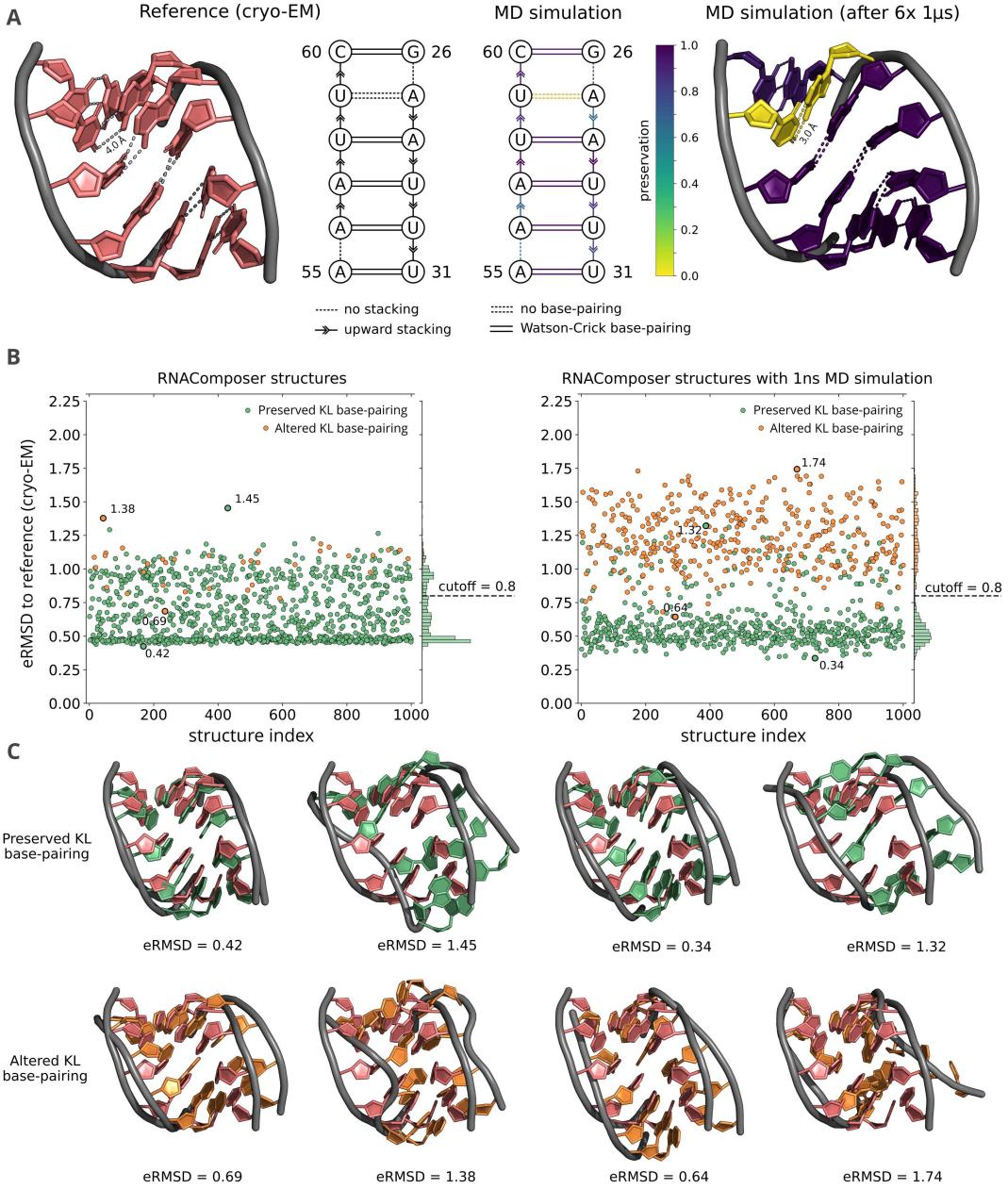
4C and 5C).

Figure 3: Reference KL fold and its comparison with structure collections. **(A)** The averaged preservation of the KL over the six sets of 1 µs MD simulations relative to the reference cryo-EM structure. The results demonstrate conserved WC base-pairings, except for bases A27 and U59, where the conformation observed in the cryo-EM structure is in close spatial proximity but not forming any canonical WC base-pairs. **(B)** The eRMSD of all structures from the RNAComposer structure collection (left) and RNAComposer structure collection after 1 ns MD simulation (right) compared to the reference cryo-EM structure, exclusively focusing on the KL region. **(C)** The structures classified as preserved KL with the lowest and highest eRMSD, as well as those classified as altered KL, also with the lowest and highest eRMSD with and without MD simulation.

8

3.3 The experimental smFRET distribution guides the selection of structures from the predicted collections

Building on the filtered structure collections, we can now probe their conformational landscape and guide further refinement using the experimental smFRET distribution. Even without photon simulations, the initial E_{DA} distributions derived from ACVs of all filtered structures already reveal distinct conformational profiles across the different prediction strategies (Figure 4A). FARFAR2 and the combined MD simulations sample a broad range of inter-dye distances, covering not only the experimental smFRET range but the full theoretical FRET range from 0 to 1. This is also evident in the structural visualizations, which display an almost globular distribution of different $\mathrm{TL}_{\mathrm{GAAA}}$ positions. In contrast, RNAComposer and AlphaFold3 predominantly yield conformations in the low-FRET region, suggesting more compact or structurally constrained collections, which are likely due to similarly folded poly(A)-linkers that limit the accessible positions of the $\mathrm{TL}_{\mathrm{GAAA}}$. Interestingly, these initial distributions already indicate that none of the tools can fully reproduce the experimentally observed broad low-FRET range distribution.

To directly compare our experimental smFRET distribution with the predicted structure collections, we performed photon burst simulations for each structure collection, including all RNA 3D prediction tools and MD trajectories. This approach includes experimental conditions such as shot noise broadening, direct excitation, and gamma correction of the underlying energy transfer efficiency E_{DA} distribution resulting in an *in silico* FRET distribution. Given the assumption of fast dynamics of the RNA model construct in the FRET experiment, bursts were determined by averaging rather than segmenting them into trajectory splits of individual MD runs or subsamples of predicted structures. The resulting FRET distributions are thus uniformly sampled from all structures (Figure 4B - unweighted and Table 1 - w/ photon sampling unweighted). Notably, despite partial overlap, none of the *in silico* FRET distributions generated from RNA 3D prediction tools or MD simulations again fully recapitulate the experimental smFRET distribution.

Table 1: FRET measurements for all approaches

	w/o photon sampling	w/ photon sampling	
		unweighted	weighted
RNAComposer	0.12 ± 0.05	0.10 ± 0.03	0.20 ± 0.04
FARFAR2	0.50 ± 0.22	0.58 ± 0.10	0.24 ± 0.05
AlphaFold3	0.12 ± 0.06	0.10 ± 0.03	0.22 ± 0.05
MD simulation	0.48 ± 0.27	0.53 ± 0.09	0.23 ± 0.05
smFRET	-	0.25 ± 0.13	

Therefore, as the initially predicted E_{DA} distributions without photon sampling appeared promising (Figure 4A), we investigated an alternative sampling approach in which structures are no longer uniformly sampled, but directly selected based on their probability of appearing in the experimental smFRET distribution. The resulting FRET distributions (Figure 4B - weighted) now closely match the range of the experimental smFRET distribution across all prediction tools and MD simulations (Table 1, w/ photon sampling - weighted). However, the experimental width of the FRET distribution could not be fully reproduced, despite accounting for shot noise broadening of the experimental burst size distribution and applying gamma correction.

Although our weighted sampling approach successfully reproduced the experimental smFRET distribution, we observed substantial differences in the structural diversity of the selected collec-
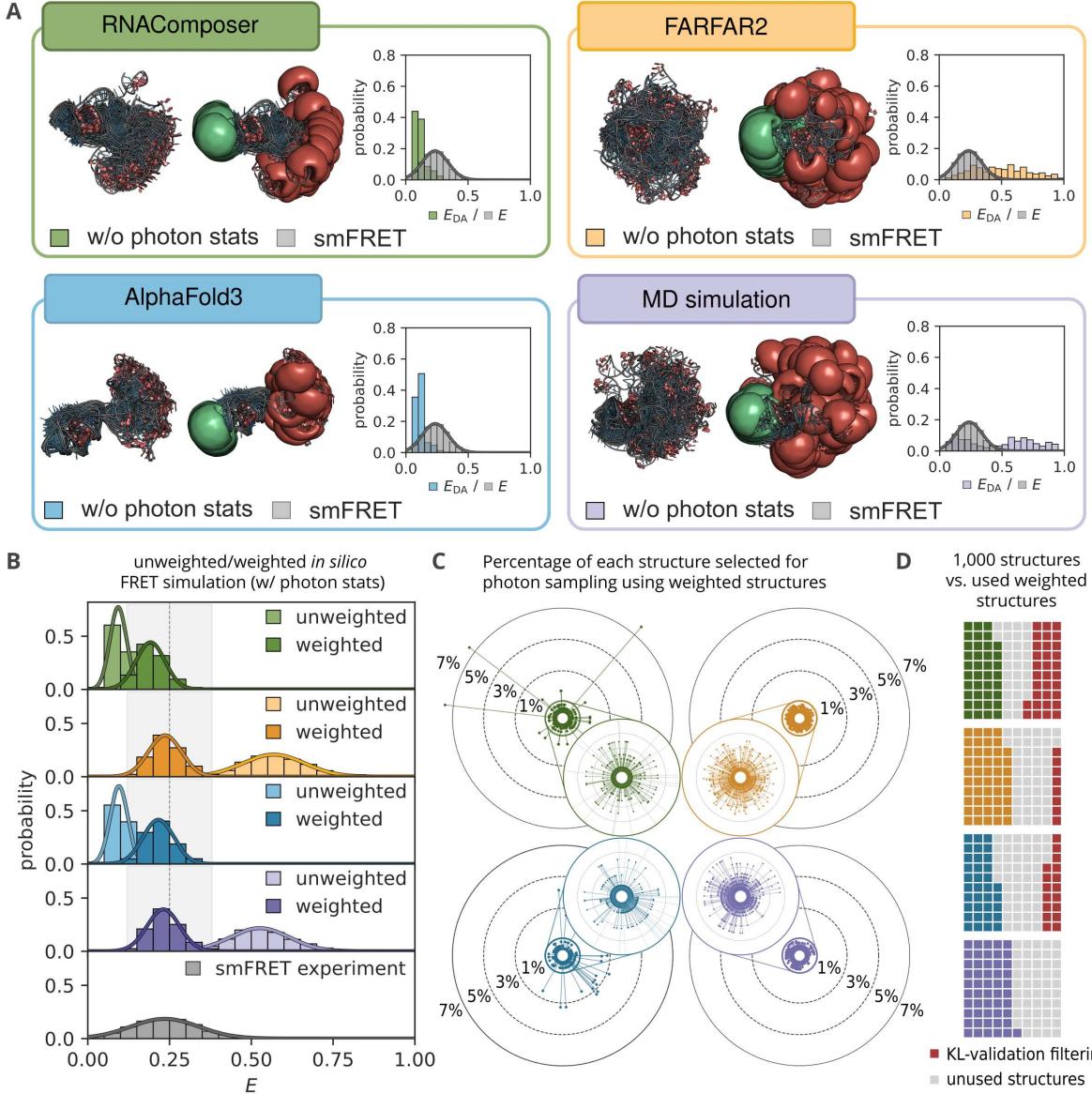
Figure 4: Comparison of all 3D structure prediction tools and MD simulations to the smFRET experiment. **(A)** The FRET distribution of all approaches without photon sampling, as well as the final 3D structure collection after filtering with and without ACVs. **(B)** Representation of all FRET distributions with photon statistics using FRETraj. "Unweighted" describes a uniform distribution of each structure for the burst calculation, whereas "weighted" denotes the sampling of structures with the same probabilities as in the smFRET experiment. **(C)** The contribution of all structures for each tool used for the new weighted approach. FARFAR2 has many different structures contributing to the resampled collection, while RNAComposer and AlphaFold3 exhibit few, highly represented structures in the refined collection. **(D)** The number of lost structures during filtering for validated KLs and, ultimately, how many structures are selected for the weighting approach against the experimental data. KL filtering was not applied to MD simulation structures.

tions. Particularly for RNAComposer and AlphaFold3, specific FRET ranges are represented by only a few structures. To match the bin probabilities of the experimental distribution, these structures are selected disproportionately often, with some individuals contributing over 7% to the final weighted RNAComposer collection (Figure 4C). While this enables an accurate match to the experimental FRET profile, it comes at the cost of reduced conformational diversity. The number of structures ultimately contributing to the weighted distribution depends on how many initially fall within the experimental smFRET range and how frequently each FRET bin is populated. FAR-FAR2 and the MD-based collection show broader structural representation, while AlphaFold3 and RNAComposer rely heavily on a few dominant conformers. The overall filtering process, from the initial structure count through the KL divergence filtering to the final number of structures used in the weighted sampling, is summarized in Figure 4D (Supplementary Table for exact values).

4 Discussion

FRET-guided integrative modeling has often relied on MD simulations to generate structure collections for comparison with experimental data [25, 45, 27]. In this study, we have addressed whether RNA 3D prediction tools can serve as a practical alternative, how many structures are actually needed to represent low-FRET collections adequately, and how diverse and reliable the resulting *in silico* FRET predictions and structures are.

We first find that RNAComposer and AlphaFold3 model the unbound state in an extended conformation, leading to a narrow distribution of low energy transfer efficiencies and limited overlap with the experimental smFRET data. In this construct, the TL_{GAAA} is connected to the KL via a poly(A)-linker and does not form interactions with either element [31]. The observed dye distance and the according FRET efficiency is therefore likely influenced by the conformational flexibility of the linker [46, 47, 48]. Poly(A)-linkers are known to adopt stacked base conformations in low energy states [49, 50], which may restrict the TL_{GAAA} range of motion, consistent with the conformations predicted by RNAComposer and AlphaFold3.

In contrast, FARFAR2 generates a highly diverse structure collection spanning the full FRET range, overlapping the experimental smFRET distribution. This includes poly(A)-linker conformations that allow greater spatial freedom of the TL_{GAAA} and therefore higher FRET efficiencies. This indicates that, despite the initial conditions, the poly(A)-linker explores a wide range of conformations on the nanosecond timescale (Supplementary Figures 2 and 3). Similarly, the six $1\,\mu s$ MD trajectories, each sampled from diverse seed conformations, produce a broad FRET distribution. However, since states were randomly selected across all trajectories, the resulting collection is highly dependent on the manually chosen seed structures to ensure representative sampling.

To evaluate how well predicted collections reflect experimentally observed FRET signals under realistic conditions, we simulated photon bursts using FRETraj while accounting for experimental conditions. These corrections ensure that the simulated FRET values reflect what would be observed in a real experiment. FRETraj treats predicted structures, whether from MD simulations or RNA 3D prediction tools like FARFAR2, as if they were rapidly sampled during the photon detection time. This means the method assumes that the RNA explores all conformations on the nanosecond timescale and averages the resulting FRET signals accordingly. As a result, the simulated FRET distributions from both MD and FARFAR2 converge on a similar mean FRET value of approximately 0.5 [38, 27, 42]. This convergence indicates, that both structural diversity (static heterogeneity) and fast conformational motion (dynamic averaging) can lead to the same average FRET signal if the timescale of motion is fast compared to the burst detection time window. However, this behavior does not match the experimental FRET distribution, which shows a lower

mean and broader width, suggesting that the actual RNA dynamics occur on slower timescales, from microseconds to milliseconds [51, 52, 53], which are not captured by the current simulation approach.

We introduced a FRET-guided reweighting strategy, that was implemented to refine the predicted structure collections based on their agreement with the experimental FRET distribution. This procedure yields structure collections, that are not only consistent with the experimental mean (Table 1 w/ photon sampling - weighted) but also reflect the relative occupancy of FRET states. Following reweighting, RNAComposer and AlphaFold3 yield FRET distributions with mean values and widths comparable to those of FARFAR2 and MD, although all remain slightly shifted toward lower efficiencies relative to the experimental distribution. Among these, FARFAR2 best reproduces the experimental mean of 0.25, though the distribution remains narrower than in the experimental data. These findings give rise to two important questions: first, what accounts for the substantial exclusion of structures, particularly from the extensive MD and FARFAR2 collections, during the reweighting process; and second, why does the weighted collection, despite improved agreement in mean FRET efficiency, still fail to recapitulate the full width of the experimental distribution?

The weighting procedure inherently filters out structures with FRET values outside the narrow experimental range, particularly compact conformations with high donor–acceptor proximity. FARFAR2 and MD simulations generate broad collections that include many such structures, which are systematically excluded. The same applies to RNAComposer and AlphaFold3, which contain a large number of low-FRET structures, that despite offering some structural diversity within specific FRET bins, provide limited coverage of the full experimental distribution. It is worth noting that structural properties beyond global distance metrics, such as topological features, may further influence the effective usability of predicted models. In this context, structural entanglements were found to be common across all methods, potentially contributing to the presence of sterically or topologically invalid conformations (Supplementary Tables).

We showed that none of the validated structure collections, including those derived from different MD states, will reproduce the low FRET distribution observed in the smFRET experiment without weighting. To capture the highly dynamic nature of RNA, which is essential for understanding its full conformational landscape [54], thus reconciling the structure heterogeneity across different methodologies and timescales, structure prediction tools must deliver a diverse collection of structures. The latter enables sampling and selecting those structures present in the experimental FRET distribution.

FRET integrative modeling plays a crucial role in elucidating structure dynamics and conformational heterogeneity in biological systems. We advocate that 3D prediction tools, when combined with structural validation, offer a fast alternative to full-atom MD simulations. Furthermore, we demonstrate how FRET can be used to tailor structure collections to experimental observations, allowing for an adaptive adjustment of the initial number of structures depending on the level of detail in the subsequent structure validation. This approach will help shift the perspective from a static reference structure to a dynamic representation of RNA conformational landscapes.

Contributions
R.B. conceived the research. M.W., F.E. and M.A. generated structure collections of RNA 3D structures. MA and T.Z. analyzed RNAComposer-predicted models for entanglements. M.W. analyzed and compared the structure collections. M.W., F.E., and F.D.S. designed and implemented the software pipeline including *in silico* labeling, MD simulations, FRET predictions incl. mACV dye models, and the software development on Github. R.B. performed and analyzed single-molecule/ensemble fluorescence measurements. M.S. and R.B. supervised the project. M.W. and R.B. wrote the initial draft of the manuscript. All authors contributed to the finalization of the manuscript.

Data availability
The source code of FRETraj is available on GitHub https://github.com/RNA-FRETools/fretraj and Zenodo https://doi.org/10.5281/zenodo.15041891. Documentation can be found at https://rna-fretools.github.io/software/. All 10,000 initial structures predicted by RNAComposer, FARFAR2, and AlphaFold3, the downsampled sets of 1,000 structures used for FRET analysis, the six 1 μs MD trajectories initiated from selected seed structures, and all associated ACV files (.pkl) generated via FRETraj are available at Zenodo: https://doi.org/10.5281/zenodo.15971163.

Additional source material is available from the corresponding author upon request.

Conflict of interest
None declared.

References

[1] M. Karplus and J. Andrew McCammon. "Molecular dynamics simulations of biomolecules". In: *Nat. Struct. Biol.* 9 (2002), pp. 646–652. DOI: 10.1038/nsb0902-646.

[2] J. Abramson et al. "Accurate structure prediction of biomolecular interactions with AlphaFold 3". In: *Nature* 630 (2024), pp. 493–500. DOI: 10.1038/s41586-024-07487-w.

[3] Kwangho Nam and Magnus Wolf-Watz. "Protein dynamics: The future is bright and complicated". In: *Struct. Dyn.* 10.1 (2023), p. 014301. DOI: 10.1063/4.0000179.

[4] Adam Hospital et al. "Molecular dynamics simulations: advances and applications". In: *Adv. Appl. Bioinform. Chem.* 8 (2015), pp. 37–47. DOI: 10.2147/AABC.S70333.

[5] Luisa Statello et al. "Gene regulation by long noncoding RNAs and its biological functions". In: *Nat. Rev. Mol. Cell Biol.* 22.2 (2021), pp. 96–118. DOI: 10.1038/s41580-020-00315-9.

[6] Fabrizio Pucci and Alexander Schug. "Shedding light on the dark matter of the biomolecular structural universe: Progress in RNA 3D structure prediction". In: *Methods* 162-163 (2019), pp. 68–73. DOI: 10.1016/j.ymeth.2019.04.012.

[7] José Almeida Cruz et al. "RNA-Puzzles: A CASP-like evaluation of RNA three-dimensional structure prediction". In: *RNA* 18.4 (2012), pp. 610–625. DOI: 10.1261/rna.031054.111.

[8] Z. Miao et al. "RNA-Puzzles Round II: assessment of RNA structure prediction programs applied to three large RNA structures". In: *RNA* 21.6 (2015), pp. 1066–1084. DOI: 10.1261/rna.049502.114.

[9] Z. Miao et al. "RNA-Puzzles Round III: 3D RNA structure prediction of five riboswitches and one ribozyme". In: *RNA* 23.5 (2017), pp. 655–672. DOI: 10.1261/rna.060368.116.

[10] Z. Miao et al. "RNA-Puzzles Round IV: 3D structure predictions of four ribozymes and two aptamers". In: *RNA* 26.8 (2020), pp. 982–995. DOI: 10.1261/rna.075341.120.

[11] F. Bu et al. "RNA-Puzzles Round V: blind predictions of 23 RNA structures". In: *Nat. Methods* 22 (2024), pp. 399–411. DOI: 10.1038/s41592-024-02543-9.

[12] Clément Bernard et al. "State-of-the-RNArt: benchmarking current methods for RNA 3D structure prediction". In: *NAR Genom. Bioinform.* 6.2 (2024), lqae048. DOI: 10.1093/nargab/lqae048.

[13] Simona Galli. "X ray crystallography: One century of Nobel prizes". In: *J. Chem. Educ.* 91.12 (2014), pp. 2009–2012. DOI: 10.1021/ed500343x.

[14] R. C. Ferguson and W. D. Phillips. "High resolution nuclear magnetic resonance spectroscopy. Advances in instrumentation in this field are leading to new applications in chemistry and biology". In: *Science* 157.3786 (1967), pp. 257–267. DOI: 10.1126/science.157.3786.257.

[15] Sriram Subramaniam and Gerard J. Kleywegt. "A paradigm shift in structural biology". In: *Nat. Methods* 19.1 (2022), pp. 20–23. DOI: 10.1038/s41592-021-01361-7.

[16] Jiri Sponer et al. "RNA structural dynamics as captured by molecular simulations: A comprehensive overview". In: *Chem. Rev.* 118.8 (2018), pp. 4177–4338. DOI: 10.1021/acs.chemrev.7b00427.

[17] Dmitri I. Svergun and Michel H. J. Koch. "Small angle scattering studies of biological macromolecules in solution". In: *Rep. Prog. Phys.* 66.10 (2003), pp. 1735–1782. DOI: 10.1088/0034-4885/66/10/R05.

[18] Jan Niklas Tants and Andreas Schlundt. "Advances, applications, and perspectives in small angle X ray scattering of RNA". In: *Chembiochem.* 24.17 (2023), e202300110. DOI: 10.1002/cbic.202300110.

[19] Almudena Ponce-Salvatierra et al. "Computational modeling of RNA 3D structure based on experimental data". In: *Biosci. Rep.* 39.2 (2019), BSR20180430. DOI: 10.1042/BSR20180430.

[20] G. Jeschke. "MMM: A toolbox for integrative structure modeling". In: *Protein Sci.* 27.1 (2018), pp. 76–85. DOI: 10.1002/pro.3269.

[21] Th Förster. "Zwischenmolekulare energiewanderung und fluoreszenz". In: *Ann. Phys.* 437.1–2 (1948), pp. 55–75. DOI: 10.1002/andp.19484370105.

[22] E. Lerner et al. "FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices". In: *eLife* 10 (2021). DOI: 10.7554/eLife.60416.

[23] B. Hellenkamp et al. "Precision and accuracy of single molecule FRET measurements: a multi laboratory benchmark study". In: *Nat. Methods* 15.9 (2018), pp. 669–676. DOI: 10.1038/s41592-018-0085-0.

[24] Martin F. Peter et al. "Cross validation of distance measurements in proteins by PELDOR DEER and single molecule FRET". In: *Nat. Commun.* 13.1 (2022), p. 4396. DOI: 10.1038/s41467-022-31945-6.

[25] M. Dimura et al. "Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems". In: *Curr. Opin. Struct. Biol.* 40 (2016), pp. 163–185. DOI: 10.1016/j.sbi.2016.11.012.

[26] B. Vallat et al. "Development of a prototype system for archiving integrative hybrid structure models of biological macromolecules". In: *Structure* 26.6 (2018), 894–904.e2. DOI: 10.1016/j.str.2018.03.011.

[27] F. D. Steffen et al. "FRET-guided modeling of nucleic acids". In: *Nucleic Acids Res.* 52.13 (2024), e59. DOI: 10.1093/nar/gkae496.

[28] Sarah A. Woodson. "RNA Folding Pathways and the Self-Assembly of Ribosomes". In: *Accounts of Chemical Research* 44.12 (2011). PMID: 21714483, pp. 1312–1319. DOI: 10.1021/ar2000474. eprint: https://doi.org/10.1021/ar2000474. URL: https://doi.org/10.1021/ar2000474.

[29] N. B. Leontis and E. Westhof. "Geometric nomenclature and classification of RNA base pairs". In: *RNA* 7.4 (2001), pp. 499–512. DOI: 10.1017/S1355838201002515.

[30] Sarah A. Woodson. "Taming free energy landscapes with RNA chaperones". In: *RNA Biology* 7.6 (2010). PMID: 21045544, pp. 677–686. DOI: 10.4161/rna.7.6.13615. eprint: https://doi.org/10.4161/rna.7.6.13615. URL: https://doi.org/10.4161/rna.7.6.13615.

[31] S. Gerhardy et al. "Puf6 primes 60S pre-ribosome nuclear export at low temperature". In: *Nat. Commun.* 12.1 (2021), p. 4696. DOI: 10.1038/s41467-021-24964-2.

[32] M. Dimura et al. "Automated and optimally FRET-assisted structural modeling". In: *Nat. Commun.* 11.1 (2020), p. 5394. DOI: 10.1038/s41467-020-19023-1.

[33] Christian A. Hanke et al. "Making fluorescence-based integrative structures and associated kinetic information accessible". In: *Nature methods* 21.11 (2024), pp. 1970–1972. ISSN: 1548-7105. DOI: 10.1038/s41592-024-02428-x.

[34] M. Popenda et al. "Automated 3D structure composition for large RNAs". In: *Nucleic Acids Res.* 40.14 (2012), e112. DOI: 10.1093/nar/gks339.

[35] J. Sarzynska et al. "RNA tertiary structure prediction using RNAComposer in CASP15". In: *Proteins* 91.12 (2023), pp. 1790–1799. DOI: 10.1002/prot.26578.

[36] A. M. Watkins, R. Rangan, and R. Das. "FARFAR2: Improved de novo Rosetta prediction of complex global RNA folds". In: *Structure* 28.8 (2020), 963–976.e6. DOI: 10.1016/j.str.2020.05.011.

[37] F. D. Steffen, R. K. O. Sigel, and R. Börner. "An atomistic view on carbocyanine photophysics in the realm of RNA". In: *Phys. Chem. Chem. Phys.* 18.42 (2016), pp. 29045–29055. DOI: 10.1039/C6CP04277E.

[38] F. D. Steffen, R. K. O. Sigel, and R. Börner. "FRETraj: integrating single-molecule spectroscopy with molecular dynamics". In: *Bioinformatics* 37.21 (2021), pp. 3953–3955. DOI: 10.1093/bioinformatics/btab615.

[39]   S. Bottaro, F. Di Palma, and G. Bussi. "The role of nucleobase interactions in RNA structure and dynamics". In: *Nucleic Acids Res.* 42.21 (2014), pp. 13306–13314. DOI: 10.1093/nar/gku972.

[40]   S. Bottaro et al. "Barnaba: software for analysis of nucleic acid structures and trajectories". In: *RNA* 25.2 (2019), pp. 219–231. DOI: 10.1261/rna.067678.118.

[41]   M. Hoefling et al. "Structural heterogeneity and quantitative FRET efficiency distributions of polyprolines through a hybrid atomistic simulation and Monte Carlo approach". In: *PLoS One* 6.5 (2011), e19791. DOI: 10.1371/journal.pone.0019791.

[42]   M. Hoefling and H. Grubmüller. "In silico FRET from simulated dye dynamics". In: *Comput. Phys. Commun.* 184.3 (2013), pp. 841–852. DOI: 10.1016/j.cpc.2012.10.018.

[43]   S. Kullback and R. A. Leibler. "On Information and Sufficiency". In: *Ann. Math. Stat.* 22.1 (1951), pp. 79–86. DOI: 10.1214/aoms/1177729694.

[44]   N. B. Leontis and C. L. Zirbel. "Nonredundant 3D Structure Datasets for RNA Knowledge Extraction and Benchmarking". In: *Nucleic Acids and Molecular Biology*. Nucleic Acids and Molecular Biology. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 281–298. DOI: 10.1007/978-3-642-25740-7_13.

[45]   F. Erichson. "FRET-assisted structural model of the GAAA RNA tetraloop receptor". In: *Hochschuldruckerei Hochschule Mittweida*. 2021. DOI: 10.48446/opus-12283.

[46]   J. H. Hodak et al. "Docking kinetics and equilibrium of a GAAA tetraloop-receptor motif probed by single-molecule FRET". In: *Proc. Natl. Acad. Sci. U. S. A.* 102.30 (2005), pp. 10505–10510. DOI: 10.1073/pnas.0408645102.

[47]   S. Sindbert et al. "Accurate distance determination of nucleic acids via Förster resonance energy transfer: Implications of dye linker length and rigidity". In: *J. Am. Chem. Soc.* 133.8 (2011), pp. 2463–2480. DOI: 10.1021/ja105725e.

[48]   A. Barth et al. "Unraveling multi-state molecular dynamics in single-molecule FRET experiments. I. Theory of FRET-lines". In: *J. Chem. Phys.* 156.14 (2022), p. 141501. DOI: 10.1063/5.0089134.

[49]   C. D. Downey et al. "Metal ion dependence, thermodynamics, and kinetics for intramolecular docking of a GAAA tetraloop and receptor connected by a flexible linker". In: *Biochemistry* 45.11 (2006), pp. 3664–3673. DOI: 10.1021/bi0520941.

[50]   A. M. Willemsen and G. A. Van Os. "Interaction of magnesium ions with poly A and poly U". In: *Biopolymers* 10.6 (1971), pp. 945–960. DOI: 10.1002/bip.360100602.

[51]   B. Schuler and H. Hofmann. "Single-molecule spectroscopy of protein folding dynamics–expanding scope and timescales". In: *Curr. Opin. Struct. Biol.* 23.1 (2013), pp. 36–47. DOI: 10.1016/j.sbi.2012.10.008.

[52]   E. Nir et al. "Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments". In: *J. Phys. Chem. B.* 110.44 (2006), pp. 22103–22124. DOI: 10.1021/jp063483n.

[53]   S. Kalinin et al. "On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits". In: *J. Phys. Chem. B.* 114.18 (2010), pp. 6197–6206. DOI: 10.1021/jp100025v.

[54]   L. R. Ganser et al. "The roles of structural dynamics in the cellular functions of RNAs". In: *Nat. Rev. Mol. Cell Biol.* 20.8 (2019), pp. 474–489. DOI: 10.1038/s41580-019-0136-0.

Supplementary Information
FRET-guided selection of RNA 3D structures

Mirko Weber[1], Felix Erichson[1], Maciej Antczak[2,3], Tomasz Zok[3], Fabio D. Steffen[4], Marta Szachniuk[2,3], and Richard Börner[1,*]

[1]Laserinstitut Hochschule Mittweida, University of Applied Sciences Mittweida, Technikumplatz 17, 09648, Mittweida, Germany
[2]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704, Poznan, Poland
[3]Institute of Computing Science, Poznan University of Technology, Piotrowo 2, 60-965, Poznan, Poland
[4]Department of Oncology, University of Zurich, University Children's Hospital, 8008, Zurich, Switzerland
[*]Corresponding author: richard.boerner@hs-mittweida.de - Tel: +49 3727 58 1009

September 24, 2025

Supplementary Methods

smFRET

To ensure kissing loop formation of the KL-TL$_{GAAA}$ rRNA model construct, the fluorescently labelled RNA (purchased from IBA) was imaged free in solution in presence of 116 mM KCl, buffered in 50 mM Tris-HCl at pH 7.5 and room temperature (46). The single-molecule measurement was performed on a home-built confocal microscope equipped with a $N_A = 1.2$ water-immersion objective (Nikon), a Timeharp 260 TSCPC card (Picoquant), a 532 nm cw and a 638 nm pulsed laser source. The latter uses a 10 MHz repetition rate for pulsed overleaved excitation (POE) allowing molecular sorting.(**Wahl.2013**) Photons were spectrally separated donor 582/64 and acceptor channel 690/70 band pass filters and detected onto two avalanche photodiodes (Perkin Elmer). The data was analysed with a sliding time-window with two-color, all photon burst search (APBS) identified bursts with a total count of at least 40 photons ($N_D + N_A$). Double labeled molecules were selected by applying an intensity threshold of red photons after acceptor excitation ($N_{A,A} > 40$) to remove the donor-only population and a stoichiometry limit $S > 0.2$ to eliminate any acceptor-only species (**Kapanidis.2004**). FRET histograms were corrected according to standard protocols(**Lee.2005**, 21). Correction factors are summarized in table 1.

Structure collection prediction with RNAComposer

To obtain 10,000 unique *in silico* models of the KL-TL$_{GAAA}$ construct (46), we used the RNAComposer web server in batch mode (available to logged-in users) at https://rnacomposer.cs.put.poznan.pl/ (26). We submitted twelve identical batches ($\{batch\_no\}\, i = 1 \ldots 12$), each consisting of $j = 1 \ldots 10$

identical tasks. Each task was defined by a single RNA sequence and ten identical secondary structures $k = 1 \ldots 10$. An example batch is shown below:

```
>KLTL{batch_no}i
#j = 1
UGAAGAAAUUCAAAAAAAAAGCUCGGAAUUUGAGCAAAAAAAAAAAACGGUGGUAAAUUCCAUCG
#k = 1
((((....))))........((((.[[[[[[))))............(((((..]]]]]])))))
...
#k = 10
((((....))))........((((.[[[[[[))))............(((((..]]]]]])))))


...


#j = 10
UGAAGAAAUUCAAAAAAAAAGCUCGGAAUUUGAGCAAAAAAAAAAAACGGUGGUAAAUUCCAUCG
#k = 1
((((....))))........((((.[[[[[[))))............(((((..]]]]]])))))
...
#k = 10
((((....))))........((((.[[[[[[))))............(((((..]]]]]])))))
```

The input was submitted 1,200 times across all batches (10 inputs per task $\times$ 10 tasks per batch $\times$ 12 batches). For each such input, RNAComposer generated a family of ten 3D structure collections. The first model in each family was constructed using the highest-scoring structural elements selected deterministically based on RNAComposer's internal scoring function. The remaining nine models were generated by randomly assembling well-scoring alternative elements (where sequence similarity is still highest, but we allow for, e.g., lower experimental resolution), making the procedure non-deterministic. Consequently, for two identical inputs, the first predicted model would be the same, while the remaining nine would, with high probability, differ. As a result of processing 1,200 identical inputs, we obtained 12,000 RNA 3D models, of which 1,200 were identical (the first models from each family) and 10,800 were structurally diverse. The corresponding ZIP files containing all 12,000 models were downloaded from the RNAComposer web server workspace and stored in the 'models' folder. Filtering out identical structures was performed at a later stage of the processing pipeline. From the downloaded archives, we extracted only the PDB files and then removed the ZIP files using the following command:

```
find . -name "*.zip" -exec unzip -d . -j {} "*/*.pdb" \;; rm *.zip
```

We then split multi-model PDB files into individual model files using the following AWK script:

```
BEGIN { filename = ARGV[1]; gsub(/" "/, "-",filename);
split(filename, elems, ".");}
$1 == "MODEL" {file = elems[1] "#" $2 ".pdb";content="";}
$1 == "ATOM" || $1 == "HETATM" || $1 == "TER" {content = content $0 "\n";}
$1 == "ENDMDL"  {printf("%s",content) > file;}
```

The script was saved as `split` and applied within the 'models' folder using the command:

```
for i in *.pdb; do awk -f split "$i"; done;
find . -type f ! -name "*#*.pdb" -exec rm {} \;
```

To remove duplicate 3D structures from the prediction set, we used the `fdupes` tool. First, we created a copy of the full set of models in a new folder:

```
cp models models-without-duplicates
```

Then, we applied `fdupes` to identify and remove redundant structures:

```
fdupes -rdN models-without-duplicates > ./fdupes.log
```

This procedure reduced the dataset to 10,801 unique 3D models, which were stored in the *models-without-duplicates* folder. From this set, we randomly selected 10,000 structures for further analysis and copied them to the final folder named *rnacomposer-kl-tlgaaa* using the following commands:

```
mkdir rnacomposer-kl-tlgaaa;
find models-without-duplicates -mindepth 1 -maxdepth 1 -name '*.pdb' -print0 |
shuf -n 10000 -z | xargs -r0 cp -t rnacomposer-kl-tlgaaa;
```

These 10,000 unique RNA 3D models predicted by RNAComposer were used in the analyses described in this study.

Structure collection prediction with FARFAR2

For 3D RNA structure prediction of the KL-TL$_{GAAA}$ model construct, we used FARFAR2 ([27]) from the Rosetta Commons toolbox (https://rosettacommons.org/), incorporating sequence and secondary structure constraints to generate 10,000 individual structures. The structure generation process was adapted from the Supplementary Information of Steffen et al. ([49]). The following command was executed using `rna_denovo`:

```
rna_denovo.linuxgccrelease -nstruct 10000 -fasta kltlgaaa_sequence.fasta
-secstruct_file kltlgaaa_secstruct.txt -silent kltlgaaa.out
-minimize_rna true -cycles 20000
```

The contents of `kltlgaaa_sequence.fasta` and `kltlgaaa_secstruct.txt` files were as follows:

```
> KL-TLGAAA construct
ugaagaaauucaaaaaaaaagcucggaauuugagcaaaaaaaaaaaacgguggguaaauuccaucg
```

```
(((((....)))).......((((.[[[[[[)))).............(((((..]]]]]])))))
```

The 10,000 structures were subsequently extracted using `rna_denovo` and merged into a single PDB file with the following commands:

```
grep "^SCORE:" kltlgaaa.out | grep -v description | awk '{print $NF ": " $2}' |
tee all_models.txt
```

```
extract_pdbs.linuxgccrelease -in:file:silent kltlgaaa.out -tags
`cat all_models.txt | cut -f1 -d':'`
```

```
for pdb in `cat all_models.txt | cut -f1 -d':'`; do
  echo MODEL $i >> kltlgaaa.pdb
  echo TITLE "$pdb" >> kltlgaaa.pdb
  cat "$pdb".pdb >> kltlgaaa.pdb
  echo -e "ENDMDL\n\n" >> kltlgaaa.pdb
  i=$(($i+1))
done
```

Structure collection prediction with AlphaFold3

For structure prediction using AlphaFold3 (28) (https://github.com/google-deepmind/alphafold3),
10,000 structures were computed locally using Docker, utilizing the model weights distributed
by AlphaFold for non-commercial use (see installation instructions: https://github.com/google-
deepmind/alphafold3/blob/main/docs/installation.md). The following input JSON file was used
to generate 10,000 predicted structures with AlphaFold3, where `modelSeeds` was defined as a list
ranging from 1 to 10,000.

```
{
  "name": "KLTL_GAAA",
  "modelSeeds": [1, ..., 10000],
  "sequences": [
    {
      "rna": {
        "id": "A",
        "sequence": "
            UGAAGAAAUUCAAAAAAAAAGCUCGGAAUUUGAGCAAAAAAAAAAAACGGUGGUAAAUUCCAUCG"
      }
    }
  ],
  "dialect": "alphafold3",
  "version": 2
}
```

MD simulations

Molecular dynamics (MD) simulations were performed using GROMACS 2024.2 (**Abraham2015**)
with the AMBER force field (**Cornell1995**), incorporating parmbsc0 (**Perez2007**) and χOL3
(**Banas2010**, **Zgarbova2011**) corrections for RNA. The RNA was solvated in a dodecahedral
water box with TIP3P or TIP4P (**Abascal2005**), charge-neutralized with KCl, and equilibrated
at 298 K and 1 bar. The individual helices of the model construct (H22, H68, H88) were extracted
from the reference structure (PDB: 3JCT (**Wu2016**)) and connected *in silico* via poly-A linkers
using PyMOL (https://www.pymol.org/).

In total, six MD simulations were conducted using TIP4P water, each lasting 1 μs. All MD
simulations were used for analyzing the kissing loop of the model construct using BARNABA,
specifically assessing the preservation of WC base-pairings. The seed structures for the six MD
simulations were derived from the cryo-EM reference structure (PDB ID: 3JCT), truncated to the
tertiary contact core region comprising helices H22, H68, and H88. Nucleotides within the KL
region that deviated from the target sequence were mutated using PyMOL (**PyMOL**).

A poly-A linker connecting H22 and H88 was constructed using PyMOL builder and manually
integrated into the model so that it extended behind the KL receptor site. The tetraloop from
the reference structure was manually placed at six distinct positions, resulting in six individual

structures with identical KL domains. The poly-A linker between H88 and H68 was also generated using the PyMOL Builder and subsequently adjusted to connect each tetraloop configuration to its corresponding KL domain.

FRETraj Dye parameters

The following FRETraj (48) dye parameters were used to compute the mACVs for all structures, serving as the basis for calculating E_{DA} values and generating the corresponding FRET efficiency histograms. Atom IDs had to be adjusted for each 3D prediction tool due to the different numbering of atoms in the PDB files.

```
{"Position":
                {"Cy3-65-O3'":
                     {"attach_id": <ATOM-ID in PDB>,
                      "mol_selection": "all",
                      "linker_length": 20,
                      "linker_width": 3.5,
                      "dye_radius1": 8,
                      "dye_radius2": 3,
                      "dye_radius3": 1.5,
                      "cv_fraction": 0.25,
                      "cv_thickness": 3,
                      "use_LabelLib": false,
                      "grid_spacing": 1.0,
                      "simulation_type": "AV3",
                      "state": 1,
                      "frame_mdtraj": 0,
                      "contour_level_AV": 0,
                      "contour_level_CV": 0.7,
                      "b_factor": 100,
                      "gaussian_resolution": 2,
                      "grid_buffer": 2.0,
                      "transparent_AV": true
                     },
                 "Cy5-10-C5":
                     {"attach_id": <ATOM-ID in PDB>,
                      "mol_selection": "all",
                      "linker_length": 20,
                      "linker_width": 3.5,
                      "dye_radius1": 9.5,
                      "dye_radius2": 3,
                      "dye_radius3": 1.5,
                      "cv_fraction": 0.25,
                      "cv_thickness": 3,
                      "use_LabelLib": false,
                      "grid_spacing": 1.0,
                      "simulation_type": "AV3",
                      "state": 1,
                      "frame_mdtraj": 0,
                      "contour_level_AV": 0,
                      "contour_level_CV": 0.7,
                      "b_factor": 100,
                      "gaussian_resolution": 2,
                      "grid_buffer": 2.0,
                      "transparent_AV": false}
                 },
            "Distance": {"Cy3-Cy5":
                            {"R0": 54,
                             "n_dist": 1000000}
                        }
           }
```

FRETraj photon sampling parameters

The following parameters were used for photon sampling with FRETraj across all structure collections to generate both unweighted and weighted FRET distributions, using the corresponding rkappa files.

```
{
    "dyes": {
        "tauD": 1.4,
        "tauA": 1.12,
        "QD": 0.2,
        "QA": 0.35,
        "etaA": 1,
        "etaD": 0.37,
        "dipole_angle_abs_em": 0
    },
    "sampling": {
        "nbursts": 20000,
        "skipframesatstart": 0,
        "skipframesatend": 0,
        "multiprocessing": true
    },
    "fret": {
        "R0": 54,
        "kappasquare": 0.6666,
        "gamma": true,
        "quenching_radius": 1
    },
    "species": {
        "name": ["all"],
        "unix_pattern_rkappa": ["<unweighted> or <weighted> r_kappa file"],
        "unix_pattern_don_coords": [],
        "unix_pattern_acc_coords": [],
        "probability": [1],
        "n_trajectory_splits": null
    },
    "bursts": {
        "lower_limit": null,
        "upper_limit": null,
        "lambda": null,
        "QY_correction": false,
        "averaging": "all",
        "burst_size_file": "<experimental burst sizes>"
    }
}
```

Supplementary Figures
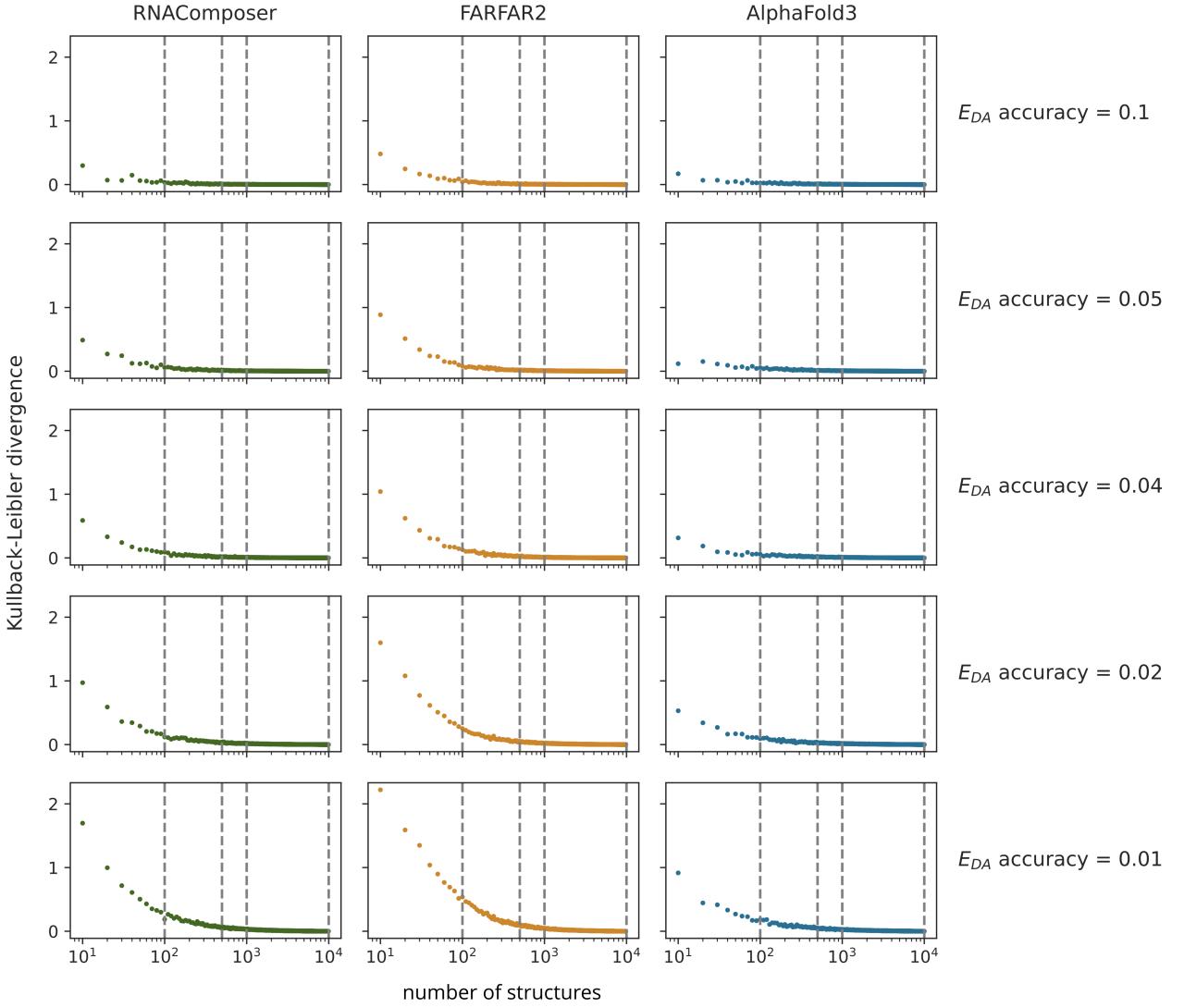
Figure 1: Dependence of Kullback-Leibler divergence (KLD) on the number of structures in the structure collections generated by RNAComposer, FARFAR2, and AlphaFold3. The E_{DA} accuracy reflects the quality of FRET-based predictions and is affected by the bin size used for computing KLD. At an E_{DA} accuracy of 0.1, as few as 100 structures suffice to represent a full structure collection of 10,000 structures across all tools, whereas an accuracy of 0.01 requires around 1,000 structures.
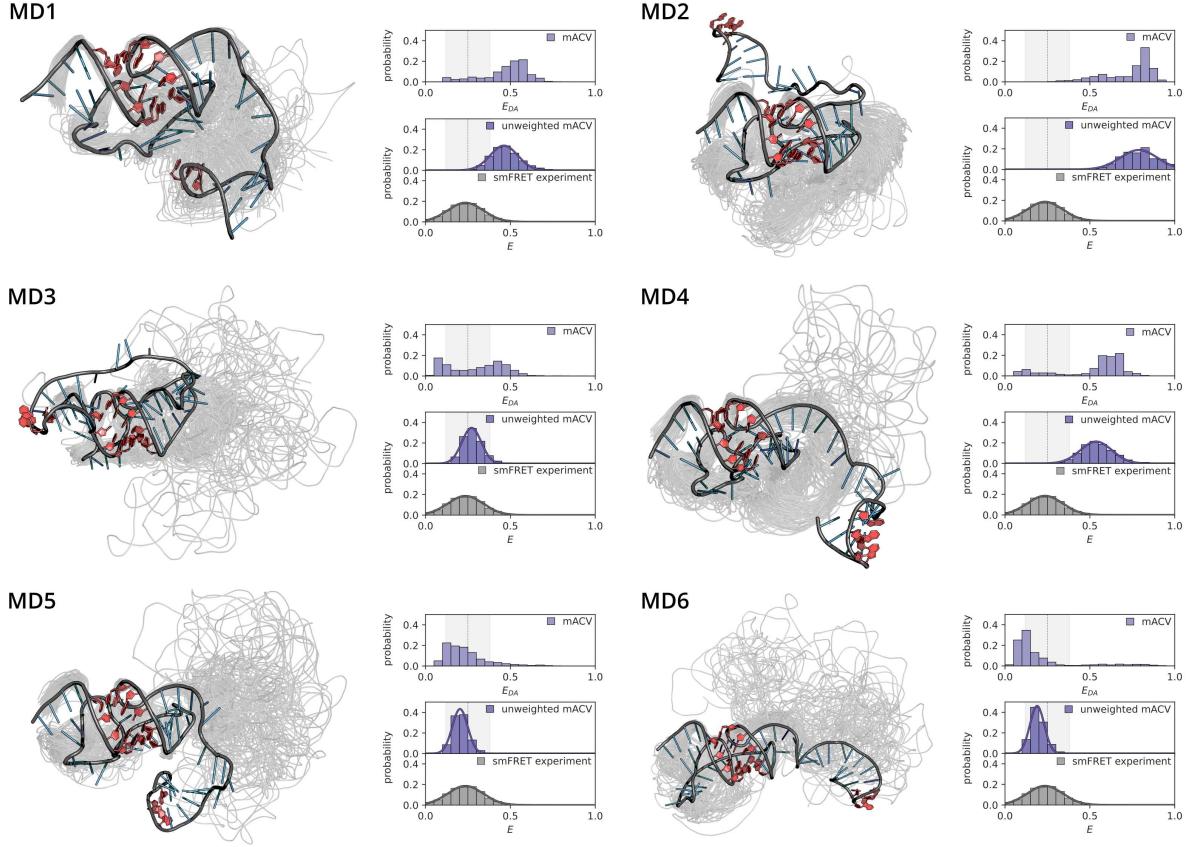
Figure 2: Seed structures used as starting points for the MD simulations. For each simulation, the input structure and ribbon conformations sampled every 10 ns are shown. Additionally, the corresponding E_{DA} distributions and unweighted FRET distributions obtained after photon sampling are plotted and compared with the experimental smFRET histogram.
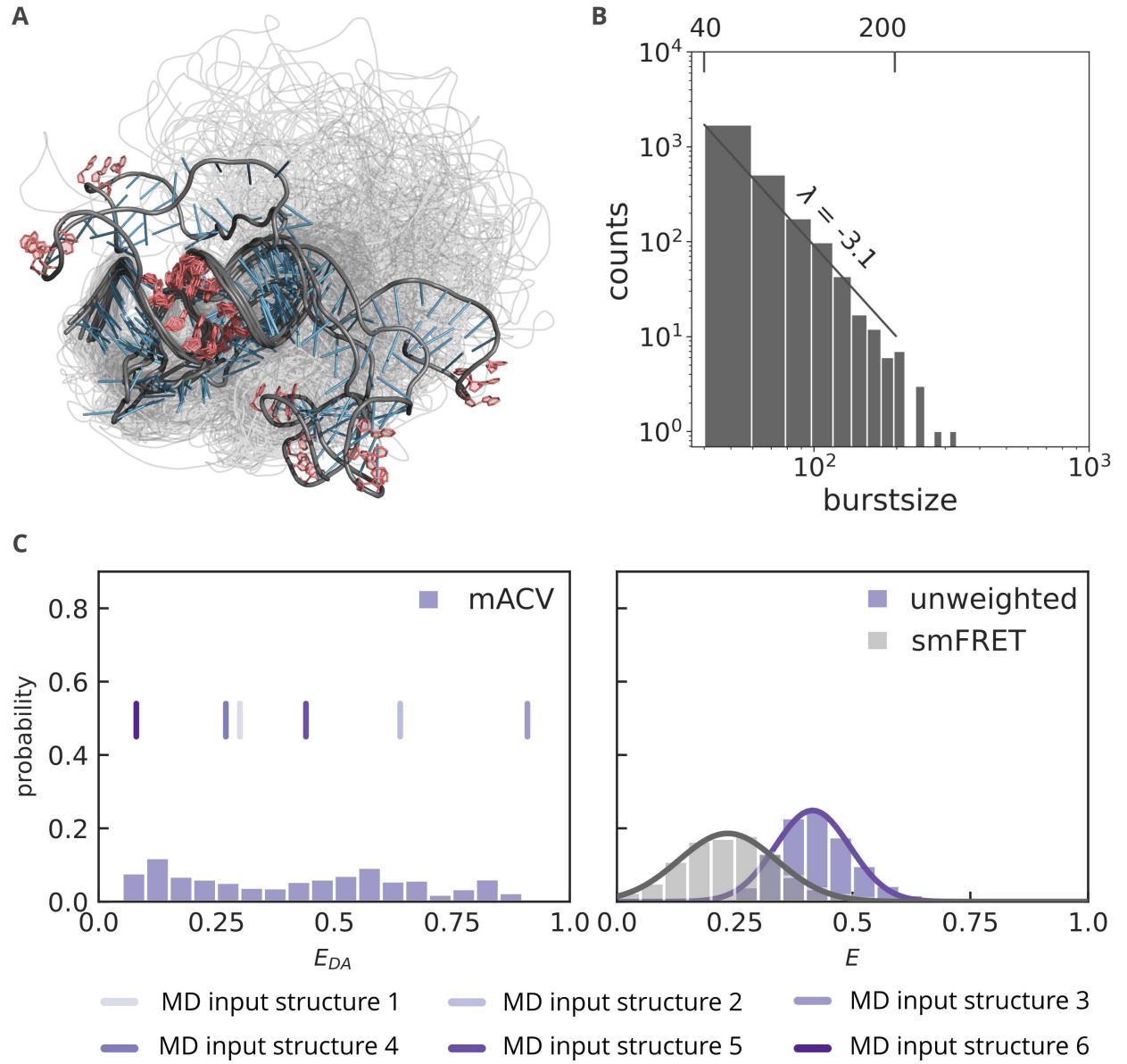
Figure 3: Six seed structures treated as one structure collection, with corresponding simulated FRET distributions. **(A)** shows the seed structures as well as the states, which were randomly chosen from all six simulations. **(B)** displays the experimental burst size distribution used for the calculation of the *in silico* burst simulation with FRETraj. **(C)** shows the E_{DA} histogram for the combined MD simulation states as well as the six input E_{DA} values of the seed structures on the left side. The right side shows the E histogram (with photon sampling) plotted against the experimental smFRET histgoram.
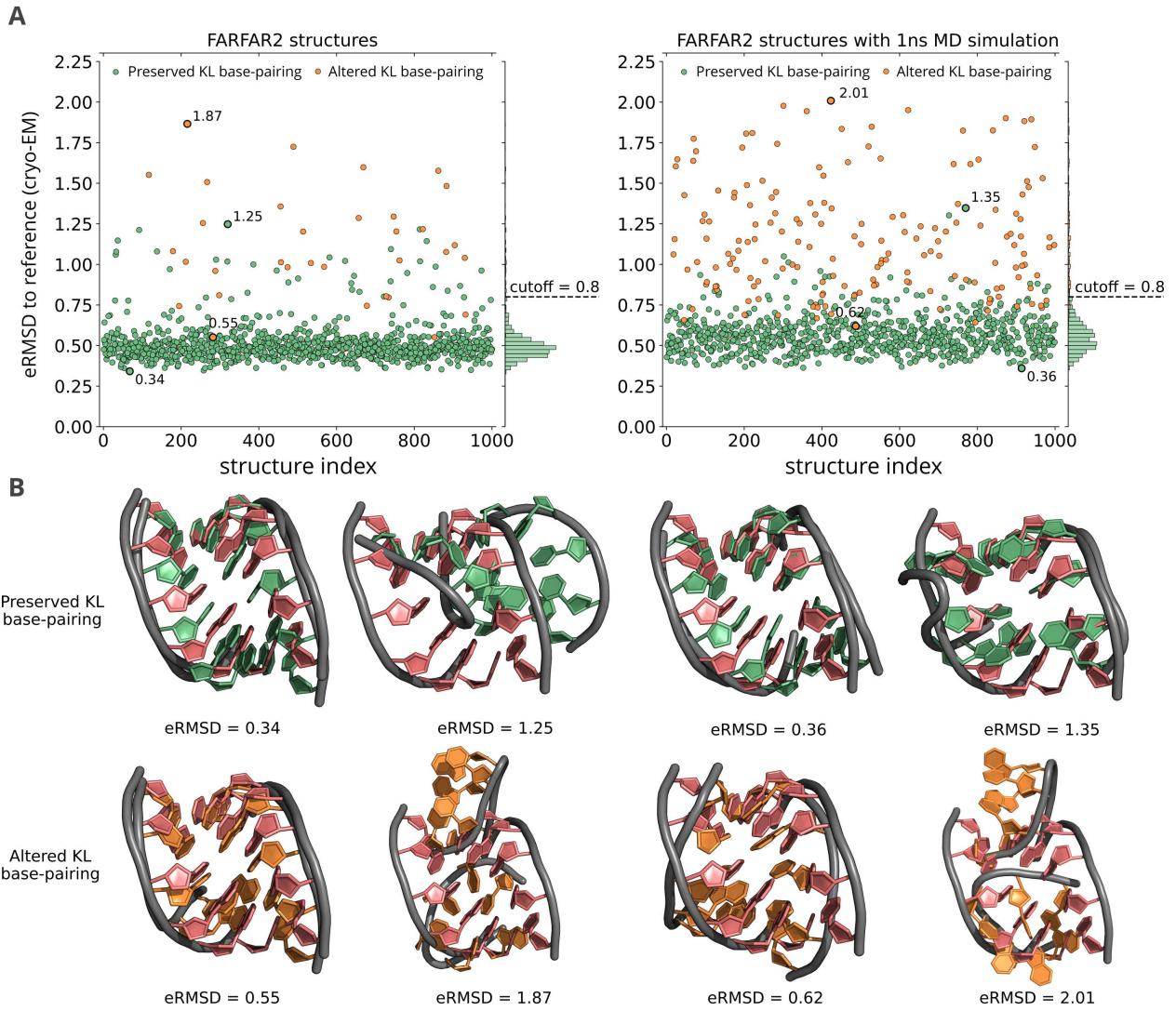
Figure 4: Correct kissing loop formation and eRMSD values for the FARFAR2 structure collection before and after 1ns MD simulation. **(A)** eRMSD of all structures from the FARFAR2 collection before (left) and after (right) 1 ns MD simulation, relative to the reference cryo-EM structure, focusing on the kissing loop region. **(B)** Representative structures classified as correctly forming the kissing loop (with the lowest and highest eRMSD), and incorrectly forming the kissing loop (also with the lowest and highest eRMSD), both before and after MD simulation.

11
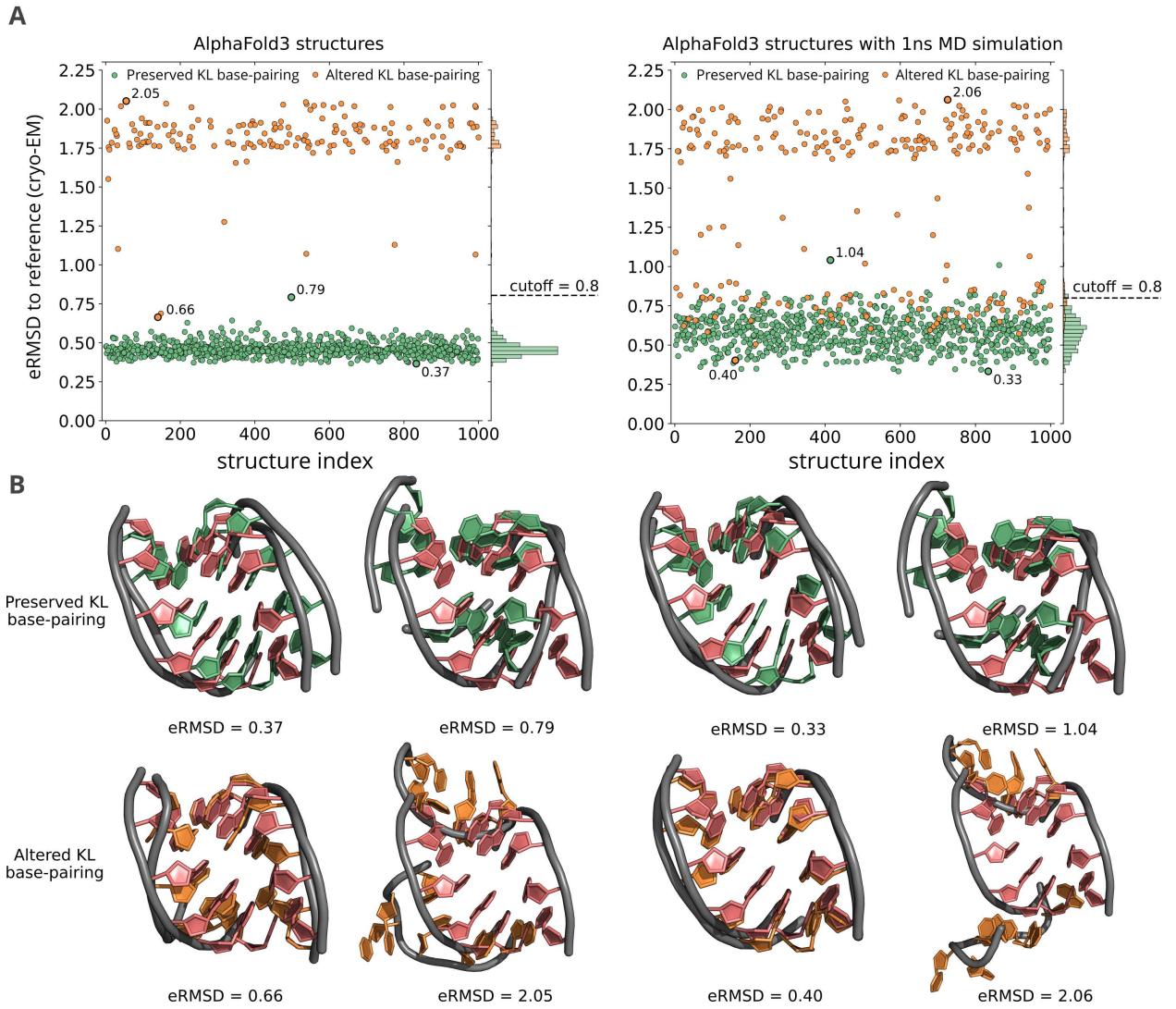
Figure 5: Correct kissing loop formation and eRMSD values for the AlphaFold3 structure collection before and after 1ns MD simulation. **(A)** eRMSD of all structures from the AlphaFold3 collection before (left) and after (right) 1 ns MD simulation, relative to the reference cryo-EM structure, focusing on the kissing loop region. **(B)** Representative structures classified as correctly forming the kissing loop (with the lowest and highest eRMSD), and incorrectly forming the kissing loop (also with the lowest and highest eRMSD), both before and after MD simulation.
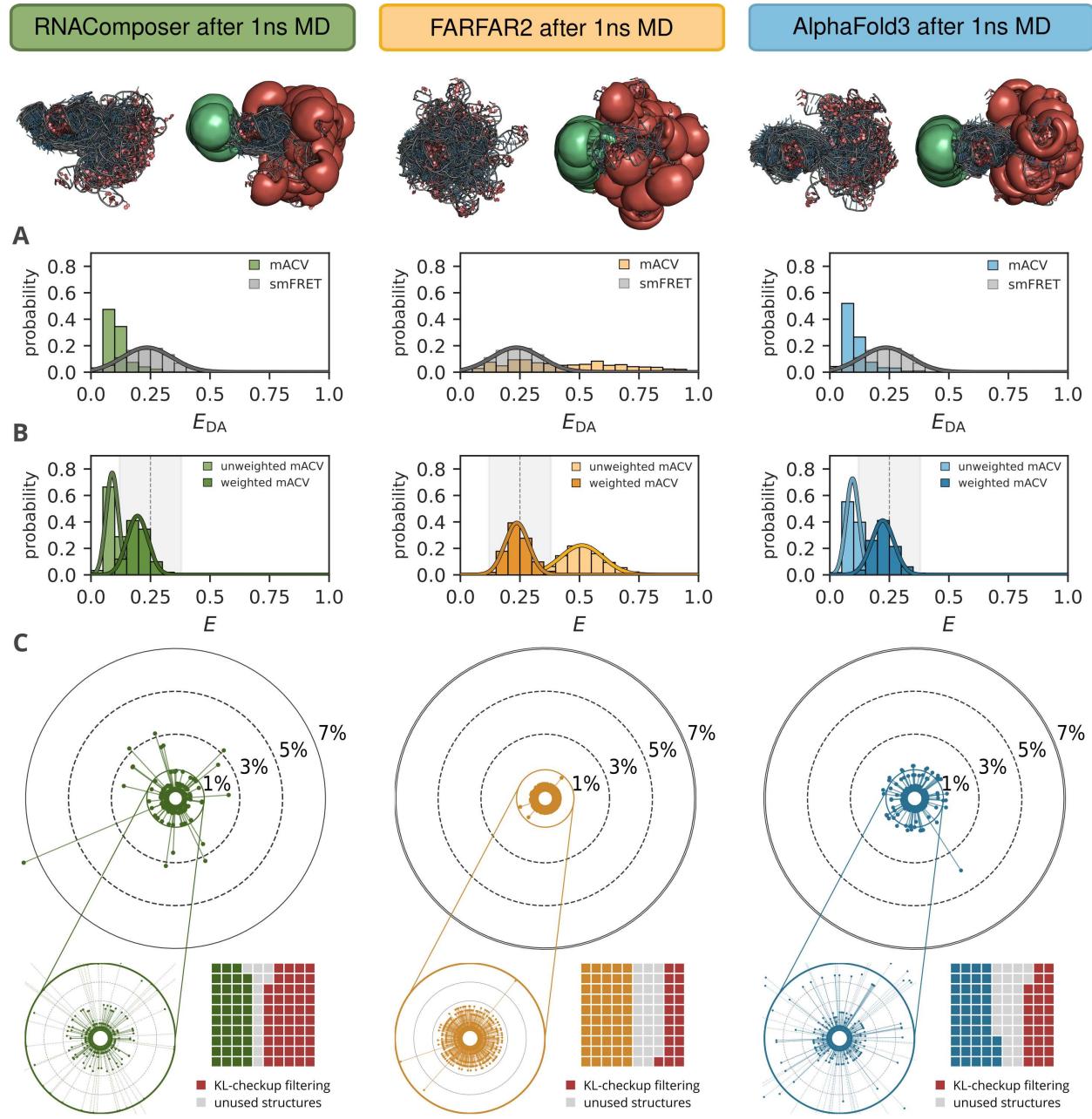
Figure 6: FRET distributions and structural composition of model collections after 1 ns MD simulation. **(A)** FRET distributions obtained from all methods without photon sampling. **(B)** FRET distributions calculated with photon statistics using FRETraj (48). "Unweighted" refers to uniform structure contributions during burst calculation; "weighted" reflects sampling based on experimental smFRET probabilities. **(C)** Structure contributions of each tool in the weighted approach. More uniform structure proportions indicate higher structural diversity and a greater number of required models. The waffle chart illustrates the reduction in structure numbers after filtering for validated kissing loops and the final number of structures used in the weighting analysis.

13

Supplementary Tables

Table 1: FRET correction factors for spectral crosstalk including bleed-through (bt) and direct excitation (dE), as well as the γ-correction factor with detection efficiency η and quantum yield Q.

Category	Correction factor	Donor	Acceptor
Bleed-through	bt	9.1%	0.3%
Direct excitation	dE	0.2%	4.2%
Detection efficiency	η	$\eta_A/\eta_D = 2.72$	
Quantum yield	Q	0.20	0.35

Table 2: Overview of structure loss due to kissing loop filtering and entanglements found in the remaining structures

	Initial collection	After KL check	w/o entanglements
RNAComposer	1,000	675	338
RNAComposer (after 1ns MD)	1,000	527	269
FARFAR2	1,000	935	139
FARFAR2 (after 1ns MD)	1,000	803	96
AlphaFold3	1,000	829	17
AlphaFold3 (after 1ns MD)	1,000	716	52

Table 2 highlights the importance of filtering implausible artificially generated 3D structures across the three tools. Although RNAComposer loses a large fraction of its models due to eRMSD and Watson–Crick base-pairing filters, only about half of the remaining structures exhibit entanglements. In contrast, FARFAR2 and AlphaFold3 retain most structures after our initial filtering, but nearly all are removed when additionally filtering for entanglements detected by RNAspider (50).

References

1. Jumper, J., Evans, R., Pritzel, A., *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589. DOI: 10.1038/s41586-021-03819-2

2. Nam, K., Wolf-Watz, M. (2023) Protein dynamics: The future is bright and complicated! *Struct Dyn*, **10**, 014301. DOI: 10.1063/4.0000179

3. Hospital, A., Goñi, J.R., Orozco, M., *et al.* (2015) Molecular dynamics simulations: advances and applications. *Adv Appl Bioinform Chem*, **8**, 37–47. DOI: 10.2147/AABC.S70333

4. Fabbri, M., Girnita, L., Varani, G., *et al.* (2019) Decrypting noncoding RNA interactions, structures, and functional networks. *Genome Res*, **29**, 1377–1388. DOI: 10.1101/gr.247239.118

5. Shin, Y., Qayyum, M.Z., Pupov, D., *et al.* (2021) Structural basis of ribosomal RNA transcription regulation. *Nat Commun*, **12**, 528. DOI: 10.1038/s41467-020-20776-y

6. Berg, M.D., Brandl, C.J. (2021) Transfer RNAs: diversity in form and function. *RNA Biol*, **18**, 316–339. DOI: 10.1080/15476286.2020.1809197

7. Sebastian-delaCruz, M., Gonzalez-Moro, I., Olazagoitia-Garmendia, A., *et al.* (2021) The role of lncRNAs in gene expression regulation through mRNA stabilization. *Noncoding RNA*, **7**, 3. DOI: 10.3390/ncrna7010003

8. Mattick, J.S., Amaral, P.P., Carninci, P., *et al.* (2023) Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat Rev Mol Cell Biol*, **24**, 430–447. DOI: 10.1038/s41580-022-00566-8

9. Cruz, J.A., Blanchet, M.F., Boniecki, M., *et al.* (2012) RNA-Puzzles: A CASP-like evaluation of RNA three-dimensional structure prediction. *RNA*, **18**, 610–625. DOI: 10.1261/rna.031054.111

10. Bernard, C., Postic, G., Ghannay, S., *et al.* (2024) State-of-the-RNArt: benchmarking current methods for RNA 3D structure prediction. *NAR Genom Bioinform*, **6**, lqae048. DOI: 10.1093/nargab/lqae048

11. Wang, X., Yu, S., Lou, E., *et al.* (2023) RNA 3D structure prediction: Progress and perspective. *Molecules*, **28**, DOI: 10.3390/molecules28145532

12. Galli, S. (2014) X-ray crystallography: One century of Nobel prizes. *J Chem Educ*, **91**, 2009–2012. DOI: 10.1021/ed500343x

13. Ferguson, R.C., Phillips, W.D. (1967) High-resolution nuclear magnetic resonance spectroscopy. Advances in instrumentation in this field are leading to new applications in chemistry and biology. *Science*, **157**, 257–267. DOI: 10.1126/science.157.3786.257

14. Subramaniam, S., Kleywegt, G.J. (2022) A paradigm shift in structural biology. *Nat Methods*, **19**, 20–23. DOI: 10.1038/s41592-021-01361-7

15. Svergun, D.I., Koch, M.H.J. (2003) Small-angle scattering studies of biological macromolecules in solution. *Rep Prog Phys*, **66**, 1735–1782. DOI: 10.1088/0034-4885/66/10/R05

16. Tants, J.-N., Schlundt, A. (2023) Advances, applications, and perspectives in small-angle X-ray scattering of RNA. *Chembiochem*, **24**, e202300110. DOI: 10.1002/cbic.202300110

17. Förster, T. (1948) Zwischenmolekulare energiewanderung und fluoreszenz. *Ann Phys*, **437**, 55–75. DOI: 10.1002/andp.19484370105

18. Lerner, E., Barth, A., Hendrix, J., *et al.* (2021) FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *Elife*, **10**, DOI: 10.7554/eLife.60416

19. Ponce-Salvatierra, A., Astha, Merdas, K., *et al.* (2019) Computational modeling of RNA 3D structure based on experimental data. *Biosci Rep*, **39**, BSR20180430. DOI: 10.1042/BSR20180430

20. Jeschke, G. (2018) MMM: A toolbox for integrative structure modeling. *Protein science : a publication of the Protein Society*, **27**, 76–85. DOI: 10.1002/pro.3269

21. Hellenkamp, B., Schmid, S., Doroshenko, O., *et al.* (2018) Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study. *Nat Methods*, **15**, 669–676. DOI: 10.1038/s41592-018-0085-0

22. Peter, M.F., Gebhardt, C., Mächtel, R., *et al.* (2022) Cross-validation of distance measurements in proteins by PELDOR/DEER and single-molecule FRET. *Nat Commun*, **13**, 4396. DOI: 10.1038/s41467-022-31945-6

23. Dimura, M., Peulen, T.O., Hanke, C.A., *et al.* (2022) Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Curr Opin Struct Biol*, **40**, 163–185. DOI: 10.1016/j.sbi.2016.11.012

24. Vallat, B., Webb, B., Westbrook, J.D., *et al.* (2018) Development of a prototype system for archiving integrative/hybrid structure models of biological macromolecules. *Structure*, **26**, 894–904.e2. DOI: 10.1016/j.str.2018.03.011

25. Popenda, M., Szachniuk, M., Antczak, M., *et al.* (2012) Automated 3D structure composition for large RNAs. *Nucleic Acids Res*, **40**, e112. DOI: 10.1093/nar/gks339

26. Sarzynska, J., Popenda, M., Antczak, M., *et al.* (2023) RNA tertiary structure prediction using RNAComposer in CASP15. *PROTEINS: Struct Funct Bioinform*, **91**, 1790–1799. DOI: 10.1002/prot.26578

27. Watkins, A.M., Rangan, R., Das, R. (2020) FARFAR2: Improved DE Novo Rosetta prediction of complex global RNA folds. *Structure*, **28**, 963–976.e6. DOI: 10.1016/j.str.2020.05.011

28. Abramson, J., Adler, J., Dunger, J., *et al.* (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, **630**, 493–500. DOI: 10.1038/s41586-024-07487-w

29. Güntert, P., Mumenthaler, C., Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program Dyana. *J Mol Biol*, **273**, 283–298. DOI: 10.1006/jmbi.1997.1284

30. Schwieters, C.D., Kuszewski, J.J., Tjandra, N., *et al.* (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson*, **160**, 65–73. DOI: 10.1016/s1090-7807(02)00014-9

31. Das, R., Karanicolas, J., Baker, D. (2010) Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat Methods*, **7**, 291–294. DOI: 10.1038/nmeth.1433

32. Watkins, A.M., Geniesse, C., Kladwang, W., *et al.* (2018) Blind prediction of noncanonical RNA structure at atomic accuracy. *Sci Adv*, **4**, eaar5316. DOI: 10.1126/sciadv.aar5316

33. Vaswani, A., Shazeer, N., Parmar, N., *et al.* (2017) Attention is all you need., DOI: 10.48550/arXiv.1706.03762

34. Ho, J., Jain, A., Abbeel, P. (2020) Denoising Diffusion Probabilistic Models., DOI: 10.48550/arXiv.2006.11239

35. Erichson, F. "FRET-assisted structural model of the GAAA RNA tetraloop receptor". In: Hochschuldruckerei Hochschule Mittweida, 2021. DOI: 10.48446/opus-12283.

36. Bu, F., Adam, Y.A.I., Adamiak, R., *et al.* (2024) RNA-Puzzles Round V: blind predictions of 23 RNA structures. *Nature Methods*, **22**, 399–411. DOI: 10.1038/s41592-024-02543-9

37. Shi, H., Rangadurai, A., Abou Assi, H., *et al.* (2020) Rapid and accurate determination of atomistic RNA dynamic ensemble models using NMR and structure prediction. *Nat Commun*, **11**, 5531. DOI: 10.1038/s41467-020-19371-y

38. Peulen, T.-O., Hengstenberg, C.S., Biehl, R., *et al.* (2023) Integrative dynamic structural biology unveils conformers essential for the oligomerization of a large GTPase. *Elife*, **12**, DOI: 10.7554/eLife.79565

39. Carrascoza, F., Carrascoza, F., Antczak, M., *et al.* (2021) Evaluation of the stereochemical quality of predicted RNA 3D models in the RNA-Puzzles submissions. *RNA*, **28**, 250–262. DOI: 10.1261/rna.078685.121

40. Bottaro, S., Di Palma, F., Bussi, G. (2014) The role of nucleobase interactions in RNA structure and dynamics. *Nucleic Acids Res*, **42**, 13306–13314. DOI: 10.1093/nar/gku972

41. Leontis, N.B., Westhof, E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499–512. DOI: 10.1017/s1355838201002515

42. Antczak, M., Popenda, M., Zok, T., *et al.* (2018) New algorithms to represent complex pseudoknotted RNA structures in dot-bracket notation. *Bioinformatics*, **34**, 1304–1312. DOI: 10.1093/bioinformatics/btx783

43. Steffen, F.D., Khier, M., Kowerko, D., *et al.* (2020) Metal ions and sugar puckering balance single-molecule kinetic heterogeneity in RNA and DNA tertiary contacts. *Nat Commun*, **11**, 104. DOI: 10.1038/s41467-019-13683-4

44. Nguyen, H.T., Hori, N., Thirumalai, D. (2019) Theory and simulations for RNA folding in mixtures of monovalent and divalent cations. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 21022–21030. DOI: 10.1073/pnas.1911632116

45. Fischer, N.M., Polêto, M.D., Steuer, J., *et al.* (2018) Influence of Na+ and Mg2+ ions on RNA structures studied with molecular dynamics simulations. *Nucleic Acids Res*, **46**, 4872–4882. DOI: 10.1093/nar/gky221

46. Gerhardy, S., Oborská-Oplová, M., Gillet, L., *et al.* (2021) Puf6 primes 60S pre-ribosome nuclear export at low temperature. *Nat Commun*, **12**, 4696. DOI: 10.1038/s41467-021-24964-2

47. Steffen, F.D., Sigel, R.K.O., Börner, R. (2016) An atomistic view on carbocyanine photophysics in the realm of RNA. *Phys Chem Chem Phys*, **18**, 29045–29055. DOI: 10.1039/C6CP04277E

48. Steffen, F.D., Sigel, R.K.O., Börner, R. (2021) FRETraj: integrating single-molecule spectroscopy with molecular dynamics. *Bioinformatics*, **37**, 3953–3955. DOI: 10.1093/bioinformatics/btab615

49. Steffen, F.D., Cunha, R.A., Sigel, R.K.O., *et al.* (2024) FRET-guided modeling of nucleic acids. *Nucleic Acids Res*, **52**, e59. DOI: 10.1093/nar/gkae496

50. Luwanski, K., Hlushchenko, V., Popenda, M., *et al.* (2022) RNAspider: a webserver to analyze entanglements in RNA 3D structures. *Nucleic Acids Res*, **50**, W663–W669. DOI: 10.1093/nar/gkac218

51. Popenda, M., Zok, T., Sarzynska, J., *et al.* (2021) Entanglements of structure elements revealed in RNA 3D models. *Nucleic Acids Res*, **49**, 9625–9632. DOI: 10.1093/nar/gkab716

52. Gren, B.A., Antczak, M., Zok, T., *et al.* (2024) Knotted artifacts in predicted 3D RNA structures. *PLoS Computational Biology*, **20**, e1011959. DOI: 10.1371/journal.pcbi.1011959)

53. Hoefling, M., Lima, N., Haenni, D., *et al.* (2011) Structural heterogeneity and quantitative FRET efficiency distributions of polyprolines through a hybrid atomistic simulation and Monte Carlo approach. *PLoS One*, **6**, e19791. DOI: 10.1371/journal.pone.0019791

54. Hoefling, M., Grubmüller, H. (2013) In silico FRET from simulated dye dynamics. *Comput Phys Commun*, **184**, 841–852. DOI: 10.1016/j.cpc.2012.10.018

55. Kullback, S., Leibler, R.A. (1951) On Information and Sufficiency. *Ann Math Stat*, **22**, 79–86. DOI: 10.1214/aoms/1177729694

56. Bottaro, S., Bussi, G., Pinamonti, G., *et al.* (2019) Barnaba: software for analysis of nucleic acid structures and trajectories. *RNA*, **25**, 219–231. DOI: 10.1261/rna.067678.118

57. Leontis, N.B., Zirbel, C.L. "Nonredundant 3D Structure Datasets for RNA Knowledge Extraction and Benchmarking". In: *Nucleic Acids and Molecular Biology*. Nucleic acids and molecular biology. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, 281–298. DOI: 10.1007/978-3-642-25740-7_13.

58. Miao, Z., Adamiak, R.W., Blanchet, M.F., *et al.* (2015) RNA-Puzzles Round II: assessment of RNA structure prediction programs applied to three large RNA structures. *RNA*, **21**, 1066–1084. DOI: 10.1261/rna.049502.114

59. Miao, Z., Adamiak, R.W., Antczak, M., *et al.* (2017) RNA-Puzzles Round III: 3D RNA structure prediction of five riboswitches and one ribozyme. *RNA*, **23**, 655–672. DOI: 10.1261\rna.060368.116

60. Miao, Z., Adamiak, R.W., Antczak, M., *et al.* (2020) RNA-Puzzles Round IV: 3D structure predictions of four ribozymes and two aptamers. *RNA*, **26**, 982–995. DOI: 10.1261/rna.075341.120

61. Kryshtafovych, A., Antczak, M., Szachniuk, M., *et al.* (2023) New prediction categories in CASP15. *Proteins*, **91**, 1550–1557.

62. Dimura, M., Peulen, T.-O., Sanabria, H., *et al.* (2020) Automated and optimally FRET-assisted structural modeling. *Nat Commun*, **11**, 5394. DOI: 10.1038/s41467-020-19023-1

63. Hollar, A., Bursey, H., Jabbari, H. (2023) Pseudoknots in RNA structure prediction. *Curr Protoc*, **3**, e661. DOI: 10.1038/s42003-024-05952-w

64. Hodak, J.H., Fiore, J.L., Nesbitt, D.J., *et al.* (2005) Docking kinetics and equilibrium of a GAAA tetraloop-receptor motif probed by single-molecule FRET. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10505–10510. DOI: 10.1073/pnas.0408645102

65. Sindbert, S., Kalinin, S., Nguyen, H., *et al.* (2011) Accurate distance determination of nucleic acids via Förster resonance energy transfer: Implications of dye linker length and rigidity. *J Am Chem Soc*, **133**, 2463–2480. DOI: 10.1021/ja105725e

66. Barth, A., Opanasyuk, O., Peulen, T.-O., *et al.* (2022) Unraveling multi-state molecular dynamics in single-molecule FRET experiments. I. Theory of FRET-lines. *J Chem Phys*, **156**, 141501. DOI: 10.1063/5.0089134

67. Downey, C.D., Fiore, J.L., Stoddard, C.D., *et al.* (2006) Metal ion dependence, thermodynamics, and kinetics for intramolecular docking of a GAAA tetraloop and receptor connected by a flexible linker. *Biochemistry*, **45**, 3664–3673. DOI: 10.1021/bi0520941

68. Willemsen, A.M., Van Os, G.A. (1971) Interaction of magnesium ions with poly A and poly U. *Biopolymers*, **10**, 945–960. DOI: 10.1002/bip.360100602

69. Schuler, B., Hofmann, H. (2013) Single-molecule spectroscopy of protein folding dynamics–expanding scope and timescales. *Curr Opin Struct Biol*, **23**, 36–47. DOI: 10.1016/j.sbi.2012.10.008

70. Nir, E., Michalet, X., Hamadani, K.M., *et al.* (2006) Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. *Journal of Physical Chemistry B*, **110**, 22103–22124. DOI: 10.1021/jp063483n

71. Kalinin, S., Sisamakis, E., Magennis, S.W., *et al.* (2010) On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits. *Journal of Physical Chemistry B*, **114**, 6197–6206. DOI: 10.1021/jp100025v

72. Ganser, L.R., Kelly, M.L., Herschlag, D., *et al.* (2019) The roles of structural dynamics in the cellular functions of RNAs. *Nat Rev Mol Cell Biol*, **20**, 474–489. DOI: 10.1038/s41580-019-0136-0