Pulsed EPR Methods in the Angstrom to Nanometre Scale Shed Light on the Conformational Flexibility of a Fluoride Riboswitch

Laura Remmel,[a,b] Andreas Meyer*,[a,c] Katrin Ackermann,[b] Gregor Hagelueken,[d] Marina Bennati*,[a,c] Bela E. Bode*[b]

[a]   Dr Laura Remmel, Dr Andreas Meyer, Prof. Dr Marina Bennati
        Research Group EPR Spectroscopy
        Max Planck Institute for Multidisciplinary Sciences
        Am Fassberg 11, 37077 Göttingen, Germany
        E-mail: andreas.meyer@mpinat.mpg.de, marina.bennati@mpinat.mpg.de
[b]   Dr Laura Remmel, Dr Katrin Ackermann, Dr Bela E. Bode
        EaStCHEM School of Chemistry, Biomedical Sciences Research Complex and Centre of Magnetic Resonance
        University of St Andrews
        North Haugh, St Andrews, KY16 9ST, United Kingdom
        E-mail: beb2@st-andrews.ac.uk
[c]   Dr Andreas Meyer, Prof. Dr Marina Bennati
        Institute of Physical Chemistry
        Georg-August University
        Tammannstraße 6, 37077 Göttingen, Germany
[d]   Dr Gregor Hagelueken
        Institute of Structural Biology
        University of Bonn
        Venusberg-Campus 1, 53127 Bonn, Germany

        Supporting information for this article is given via a link at the end of the document.

Abstract: Riboswitches control gene regulation upon external stimuli such as environmental factors or ligand binding. The fluoride sensing riboswitch from *Thermotoga petrophila* is a complex regulatory RNA proposed to be involved in resistance to F$^-$ cytotoxicity. The details of structure and dynamics underpinning the regulatory mechanism are currently debated. Here we demonstrate that a combination of pulsed electron paramagnetic resonance (ESR/EPR) spectroscopies, detecting distances in the angstrom to nanometre range, can probe distinct regions of conformational flexibility in this riboswitch. PELDOR (pulsed electron-electron double resonance) revealed a similar preorganisation of the sensing domain in three forms, i.e. the free aptamer, the Mg^{II}-bound *apo*, and the F$^-$-bound *holo* form. ^{19}F ENDOR (electron-nuclear double resonance) was used to investigate the active site structure of the F$^-$-bound *holo* form. Distance distributions without *a priori* structural information were compared with *in silico* modelling of spin label conformations based on the crystal structure. While PELDOR, probing longer distances, revealed varying conformational flexibility of the RNA backbone, ENDOR indicated low structural heterogeneity at the ligand binding site. Overall, the combination of PELDOR and ENDOR with sub-angstrom precision gave insight into structural organisation and flexibility of a riboswitch, not easily attainable by other biophysical techniques.

Introduction

Protein structure and function underpins all aspects of life from the structural framework of cells, enzymatic activity, cellular transport, to immune functions.[1] The production of the individual proteins is controlled through the expression of the corresponding genes. Riboswitches are non-coding RNAs mediating the regulation of genes[2–5] by initiating or terminating transcription or translation. Here, gene expression is being triggered by environmental factors such as temperature or pH or through ligands like ions or small molecules.[6–8] Ion or metabolite sensing riboswitches form highly selective binding pockets for their specific target.[9]

Fluoride sensing riboswitches are present in both, bacteria and archaea.[4,10–12] Increasing F$^-$ concentrations can inhibit cell growth and become acutely toxic to cells.[12–16] This antimicrobial effect has been exploited to eliminate harmful microorganisms in fermentation processes.[17] Fluoride riboswitches are involved in a cellular defence mechanism, e.g., initiating the production of F$^-$ exporters.[10,12,14]

The first fluoride sensing riboswitch was found in the *crcB* motif of *Pseudomonas syringae* in 2012.[14] In the same year, the crystal structure of the sensing domain of the fluoride sensing riboswitch from the bacterium *Thermotoga petrophila* in its F$^-$-bound form was reported by Ren *et al.* (Fig. 1a).[4] The tertiary structure of the aptamer comprises two stems, a pseudoknot, and reversed Watson-Crick (A6·U38) and Hoogsteen (A40·U48) base pairs (Fig. 1b). The riboswitch coordinates a cluster of three Mg^{II} ions which in turn encapsulates the F$^-$ ion (Fig. 1c). In this way the polyanionic RNA avoids electrostatic repulsion with the bound F$^-$. This motif composed of Mg^{II}, F$^-$, phosphate, and water was
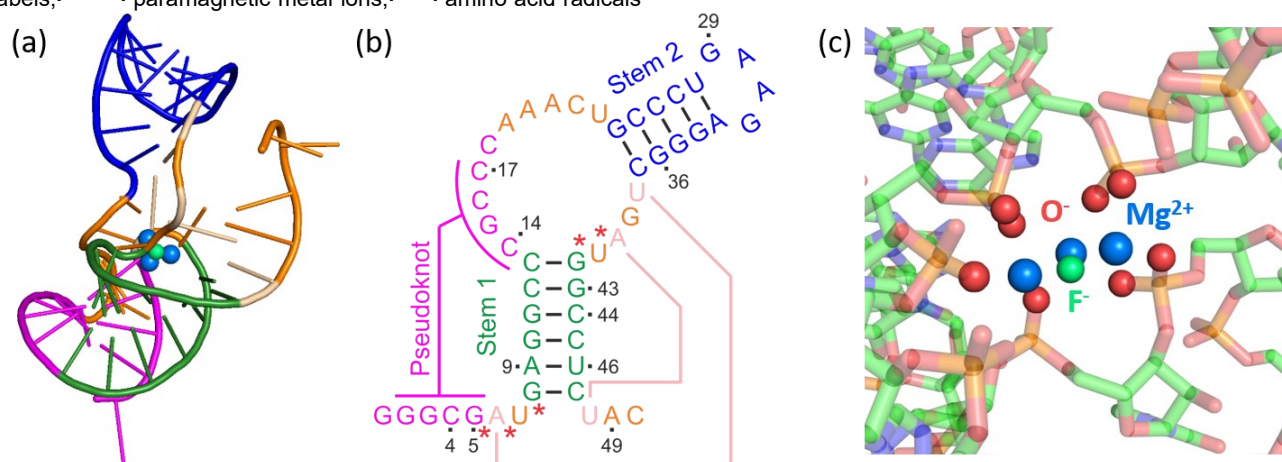
computationally found to be stable largely based on the electrostatic interactions between F- and Mg^{II} [18] A recent computational study proposes a stepwise assembly of the cluster with two Mg ions bound in the *apo* form but the third Mg ion only being incorporated together with F-.[19] The riboswitch is strongly selective for F- discriminating against other halides.[4]

Using solution NMR spectroscopy, the fluoride binding riboswitch from *Bacillus cereus* was shown to adopt a highly similar fold.[10] In *B. cereus,* the two analogous stem structures to *T. petrophila* form already in absence of Mg^{II} and F-,[10] whereas the formation of the pseudoknot requires presence of Mg^{II}.[20,21] Further addition of F- does not result in a significant change of the tertiary structure but suppresses dynamics involving a lowly populated excited *apo* state responsible for transcription termination.[10,22]

Pulsed electron-electron double resonance (PELDOR,[23,24] aka DEER for double electron-electron resonance[25]) spectroscopy permits measuring distances ranging from 20 Å to 100 Å and beyond[26–28] between paramagnetic centres such as spin labels,[26,29–31] paramagnetic metal ions,[32–37] amino acid radicals

and radical cofactors.[38–40] This provides a unique opportunity for investigating conformational ensembles through their spin-spin distance distributions. Complementary to PELDOR, ^{19}F electron-nuclear double resonance (ENDOR) is an emerging technique to detect distances in a shorter range of about 5 - 20 Å between nitroxide spin labels and ^{19}F nuclei.[41–44] This has recently been expanded to distances between fluorine nuclei and other spin centres such as triarylmethyl (TAM) radicals,[45,46] tyrosyl radicals,[47] Cu^{II},[48] or Gd^{III}.[49–51]

In this work, we have investigated the *T. petrophila* riboswitch in solution by magnetic resonance methods. ^{1}H NMR spectroscopy allowed monitoring the formation of base pairs through signals in the imino region of the free RNA, the Mg^{II}-bound *apo* riboswitch, and the F--bound *holo* riboswitch. We have further studied the solution-state preorganisation of the tertiary fold in the free and *apo* riboswitches by PELDOR and we have employed ^{19}F ENDOR to investigate the structure of the binding pocket of the *holo* riboswitch.

Figure 1. (a) Graphical representation of the crystal structure of the sensing domain of the fluoride binding riboswitch from *T. petrophila* in the *holo* form (PDB: 4ENC)[4] with the F- (green sphere) encapsulated by three Mg^{II} (blue spheres); (b) schematic representation of the secondary structure of the 50 nucleotide construct of the fluoride binding riboswitch used here, with the two stem structures (green and blue), the pseudoknot structure (magenta), and the reversed Watson-Crick (A6·U38) and reversed Hoogsteen (A40·U48) base pairs (salmon). Residue numbers for spin-labelling sites are indicated. The residues whose O- are involved in the coordination of the Mg^{II} are highlighted with red asterisks; (c) Graphical representation of the binding pocket of the fluoride binding riboswitch[4] with the F- (green sphere) encapsulated by three Mg^{II} (blue spheres). O- involved in the coordination are represented as red spheres. Additional ions not involved in the fluoride encapsulation and water molecules have been omitted for clarity throughout.

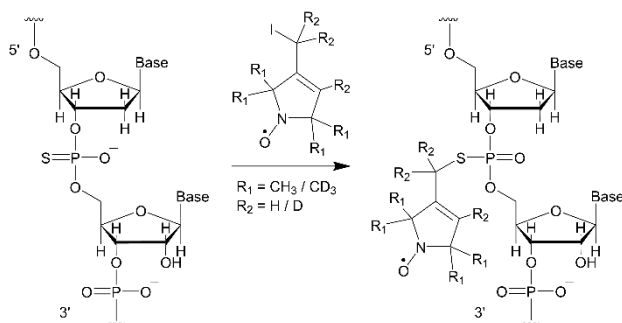## Results and Discussion

Design of spin labelled RNAs

PELDOR and ^{19}F ENDOR measurements require the presence of two or one paramagnetic spin labels, respectively. Abundant methods for the site-directed spin labelling of nucleic acids at the base, ribose, or phosphate exist.[52,53] Custom synthesised RNA sequences with phosphorothioate modification are commercially available as is the spin label precursor for labelling at the phosphate backbone (Scheme 1) making this labelling approach readily accessible.[54,55] This procedure leads to the introduction of a diastereomeric pair. The spin labelled RNA is therefore expected to display both diastereomers of the phosphothiotriester.

All constructs were designed based on the crystal structure of the *T. petrophila* riboswitch.[4] The labelling sites were screened using *in silico* labelling with MtsslSuite[56] and MMM[57] to predict distances between pairs of spin labels or spin labels and F-. Furthermore, ensembles of spin label rotamers can also be extracted allowing analysis of orientation selection in the anisotropic EPR spectrum.[25]

The following general criteria were applied: (i) nucleotides involved in the formation of the cluster, i.e. residues coordinating Mg^{II} were excluded (red asterisks in Fig. 1b); (ii) distances in the range between 18 – 45 Å were considered for PELDOR, and distances below 20 Å for ^{19}F ENDOR, respectively.

For PELDOR measurements four different constructs with two spin labels each were prepared to validate the preorganisation of the two stems and the pseudoknot, i.e. at pairs G29/G36 within stem 2 and the corresponding hairpin (blue), at C14/C44 located

2

in the region between the pseudoknot and stem 1 (green), at A9/A49 across stem 1 and the 3' end of the RNA (magenta), and at C4/C17 across the pseudoknot (Fig. S5 and Fig. S6).

Scheme 1. RNA labelling at the phosphorothioate modification of the backbone using either a protonated or a per-deuterated spin label. The sugar 5' to the labelling site is modified to a 2'-deoxyribose to avoid strand scission.[55]

For ^{19}F ENDOR, three constructs with one spin label each were chosen to result in linearly independent distance vectors between the respective label and the F^-, allowing the trilateration of the F^- position within the RNA structure (Fig. S7). G5 and G43 gave respective calculated mean distances (R_{model}) of 8.5 Å and 13.5 Å to the F^-, well resolvable below the 20 Å threshold. For U46 however, the *in silico* labelling displayed a bimodal distance distribution. The shorter $R_{model,1}$ of 15.5 Å was within the resolvable distance limit for ^{19}F ENDOR measurements with nitroxide labels. The second $R_{model,2}$ was predicted to be at about 21.5 Å, likely too long to be resolved by ^{19}F ENDOR.[41]
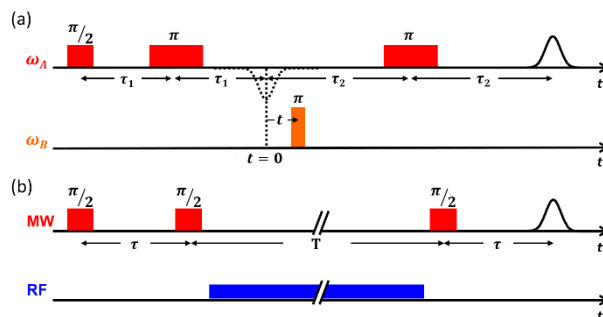
To ensure reproducible folding of the riboswitch and infer any perturbation of folding by the spin label, we established a folding protocol (see SI section 1) using ^{1}H NMR spectroscopy. This allowed to identify the free aptamer in absence of Mg^{II} and F^-, the Mg^{II}-bound *apo* riboswitch, and the F^--bound *holo* riboswitch in analogy to results for the *B. cereus* fluoride riboswitch.[10]

Comparison of ^{1}H NMR spectra in the imino regions of the free and *holo* aptamer for both, unlabelled and labelled RNA showed indistinguishable signals indicating appropriate folding after labelling (Fig. S10-16).

PELDOR measurements

PELDOR distance measurements using the 4-pulse DEER sequence (Scheme 2 (a)) are particularly well suited to study structural or conformational changes.[58–61] To map overall structure and preorganisation, PELDOR time traces were measured for the free, *apo* and *holo* aptamers for each of the four doubly spin labelled constructs. Sample preparation and experimental conditions for PELDOR are described in SI section 1. All spectra were recorded in frozen solution at 50 K. The time traces (Fig. 2) displayed differently pronounced dipolar modulations, which indicated different widths of the respective distance distributions. The time traces were analysed using Tikhonov regularisation implemented in the software DeerAnalysis (Fig. 2).[62] All obtained distance distributions, despite being broad, showed a distinct most probable distance, but did not show significant differences between the free, *apo* and *holo* aptamers for the C14/C44, G29/G36, and A9/A49 constructs.
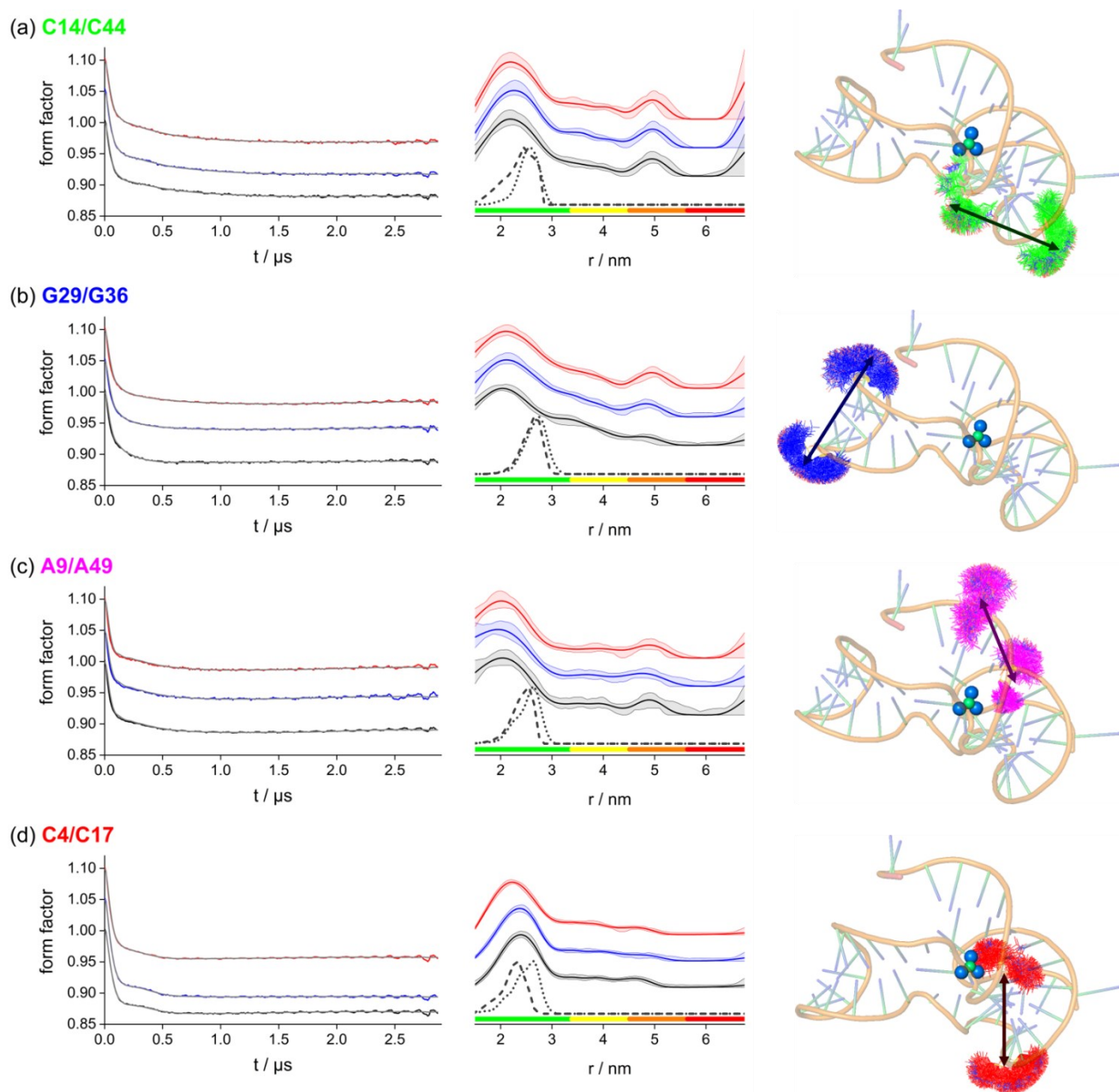
The C4/C17 construct displayed the most pronounced modulation of the time trace resulting in the narrowest distribution with the lowest uncertainty. Here, the free and *apo* forms provide no difference within uncertainty but the *holo* form yielded a slightly reduced most probable distance. This could be confirmed by an alternative processing with a different implementation of Tikhonov regularisation in the software DeerLab[63] and by neuronal network analysis with the software DeerNet[64] (Fig. S17). Interestingly, all distance distributions displayed significant distance probability up to 50 Å and beyond in all three processing approaches.

Scheme 2. Pulse sequences for 4-pulse DEER (a) and the Mims ENDOR (b) experiments with delays (t, τ), microwave (MW) pulses (π, $\pi/2$, red and orange), and a radiofrequency (RF) pulse (blue).

Comparing the experimental distance distributions with the modelling performed to select these constructs revealed substantially broader distributions in the former. Nevertheless, all modelled distance probability is fully covered by the experimental distributions. Particularly, the DeerLab analysis hinted at the presence of multiple distance peaks corresponding to multiple conformations. This illustrated a conformational flexibility of the RNA backbone that was not reproduced by rotamer modelling on the single conformer in the crystal structure. Interestingly, for all cases but C4/C17 the most probable distances (i.e., maxima of the experimental distributions) were shorter than the maximum of the modelled distributions.

The identical PELDOR data between the free, *apo*, and *holo* forms of three constructs confirmed the preorganisation of stems 1 and 2 in the free RNA and their retention in the *apo* and *holo* aptamers. The shortening of the C4/C17 distance in the *holo* compared to the free and *apo* forms was consistent with an increased population of aptamers forming the pseudoknot. In contrast to the findings for the *B. cereus* fluoride riboswitch only requiring the addition of Mg^{II} for pseudoknot formation[21] our PELDOR data revealed also F^- in the *holo* form being required for advancing the pseudoknot formation in our constructs. The findings of shorter most probable distances in the experiment compared to modelling for three constructs indicated the presence of more compact structural arrangements in solution than in the crystal structure. The long distances tailing up to 50 Å and beyond might be explained by an equilibrium also involving a disordered or unfolded form of the aptamer. This would be consistent with earlier fluoride binding ITC data that had been fitted to 0.75 and 0.87 binding sites.[4] Thus, 13-25% of aptamers would not be binding F^- and contributing very broad peaks to the distance distribution in all three forms of the riboswitch.

3

Figure 2. Background corrected PELDOR time traces (left) measured at Q-band MW frequency and 50 K and analysed using Tikhonov regularisation for the free (black), *apo* (blue), and *holo* (red) aptamer labelled at positions C14/C44 (a), G29/G36 (b), A9/A49 (c), and C4/C17 (d). Sample concentrations were adjusted to 10 μM RNA (a-c) and 20 μM RNA (d). The fit of the experimental time trace is given as grey line. Corresponding distance distributions (middle) with the respective 2σ confidence estimates. The colour bar is showing the reliability ranges (green: shape reliable, yellow: mean and width reliable, orange: mean reliable, red: no quantification possible) of the determined distributions. A vertical offset was introduced for both, the time traces and the distance distributions for clarity. The distance distributions obtained from *in silico* labelling with MtsslSuite (dashed line) and MMM (dotted line) are given for comparison. Graphical representation of the *in silico* spin labelled RNA constructs with arrows indicating the spin-spin distance assed by PELDOR measurements.

¹⁹F ENDOR spectroscopy

The F⁻ binding riboswitch provides a unique opportunity to measure inter-spin distances to the active site in the *holo* aptamer, containing an endogenous F⁻, by ¹⁹F ENDOR spectroscopy. For this, ¹⁹F Mims ENDOR (pulse sequence in Scheme 2b) was performed with three constructs containing the label at positions G5, G43, and U46 (Fig. 1b). ¹⁹F ENDOR spectra of G5 and G43 were measured at W-band (94 GHz) and four different excitation positions in the EPR spectrum (Fig. S19, left). Experimental Details are given in SI section 1. The orientationally selected ENDOR spectra were summed up by considering the respective EPR intensities at the excitation positions (see section 1 in SI) and the sum spectrum was analysed.

For construct U46, spectra were measured at Q-band (34 GHz) as the expected distance was longer and improved concentration sensitivity turned out to be advantageous (Fig. S20). Similar findings concerning sensitivity were recently observed employing trityls as spin labels.[46] At Q-band, ¹H ENDOR signals originating

4

from the coupling to the spin label protons and overlapping with the ^{19}F ENDOR signals (Fig. S21) were supressed by introducing a deuterated nitroxide spin label (see section 1 in SI). Furthermore, this construct showed the slowest echo decay (Fig. S22) contributing to improved sensitivity. At 34 GHz ^{19}F ENDOR spectra were measured only at two different spectral positions (Fig. S19, right), and no significant effect of orientation selection could be observed within the achievable resolution and signal-to-noise ratio.
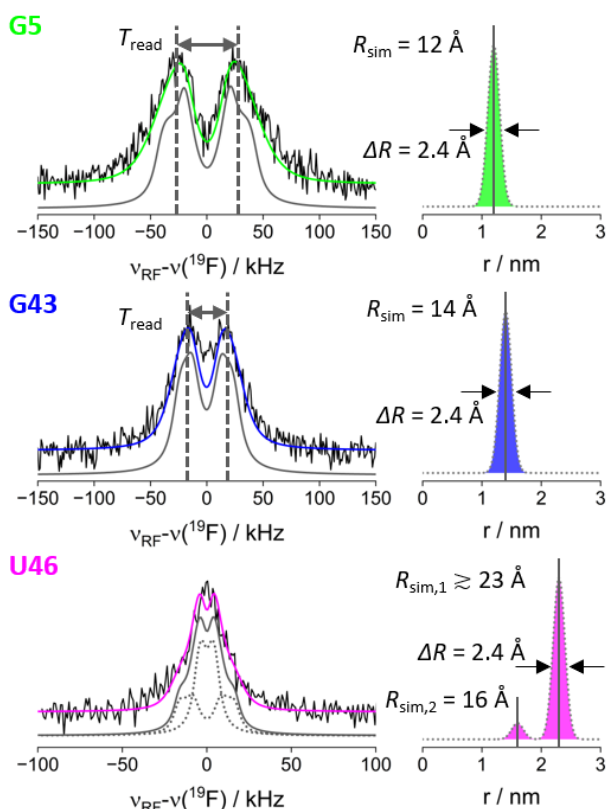
Figure 3. Mims ^{19}F ENDOR sum spectra (left column) of the *holo* aptamer and simulated spectra (coloured lines) using Gaussian distributions (right column) or based on T_{read} (grey lines, left). In case of U46 the simulated spectrum contains contributions of two distance components (dotted grey). Samples were composed of 100 µM, 150 µM, and 235 µM RNA for the RNA labelled in position G5, G43, and U46, respectively. Measurements were performed at W-band (G5 and G43) and Q-band (U46). For W-band measurements, a protonated spin label and for Q-band measurements a deuterated spin label was used.

The three resulting ^{19}F ENDOR spectra of the *holo* riboswitch are displayed in Fig. 3. For constructs G5 and G43, a well-resolved hyperfine splitting arising from the maxima of a Pake pattern[41] was observed. From this splitting, called T_{read}, the perpendicular component $T_\perp$ of the hyperfine tensor was estimated, giving first insight into the magnitude of the dipolar coupling and the corresponding inter-spin distance. Using the point-dipolar approximation the point-dipole distance (R_{read}) can be estimated as:[41]

$$T_{read} \approx T_\perp = \frac{\mu_0}{4\pi h}\left(\frac{g_e g_n \mu_B \mu_N}{R_{read}^3}\right) = \frac{C}{R^3} \qquad \text{Eq. 1}$$

with the vacuum permeability μ_0, the Planck constant h, the g-values of the electron and fluorine nucleus g_e and g_n, respectively, the Bohr magneton μ_B and the nuclear magneton μ_N. With these constants C amounts to C = 74.52 MHz Å^3.

For G5, a coupling constant T_{read} = 50 ± 7 kHz was read off that corresponds (Eq. 1) to a point-dipole distance of R_{read} = 11.5 ± 0.5 Å. The error was estimated from the width of the maxima of the peaks. For G43, a smaller splitting from a coupling constant T_{read} = 30 ± 4 kHz was observed, corresponding to a point-dipole distance of 13.5 ± 0.6 Å. For U46 no splitting could be observed, however a clear peak was still detected. To ensure that this signal of U46 did not arise from free F$^-$ interacting with the nitroxide label, a control experiment was performed. ^{19}F ENDOR spectra of the *free* aptamer labelled in position G43 and U46 were measured in the presence of F$^-$ at W-band and Q-band, respectively. In absence of Mg^{II} the F$^-$ binding site cannot form and indeed no ENDOR effect has been observed (Fig. S23). Thus, all ^{19}F ENDOR signals observed in the *holo* form arise from bound F$^-$.

To extract more information on conformational distributions from the ENDOR line shape we performed a spectral analysis in two steps. First, we simulated the spectra using our fast simulation routine SimSpec[43] (see section 1 in SI) without any prior model, just assuming a Gaussian distribution of distances centred at R_{read}. The centre and the width of the Gaussian distribution were manually adjusted to minimize the residual between experiment and simulated spectrum (Fig. S25 and S26). The simulations are displayed in Fig. 3 (left). The Gaussian distance distribution reproduced the line shapes of G5 and G43 very well using a broadening (full width at half maximum, FWHM) corresponding to 2.4 Å. The simulated T_{sim} component resulted in R_{sim} values as given in Fig. 3 (right) and in Table 1. For U46 a series of simulations was performed for Gaussian distributions of constant width 2.4 Å but varying R_{sim} between 18 and 28 Å. Examination of the residual (Fig. S27) suggested that $R_{sim} \gtrsim 20$ Å can be considered a lower boundary for the distance, when no splitting occurs. Indeed, the resolution of the splitting depended on the choice of the line width parameter. Our choice of a Lorentzian line of 7 kHz (slightly larger than the power broadening of the RF pulse) is based on an ongoing investigation of intrinsic ENDOR line widths. This resulted in $R_{sim} \gtrsim 23$ Å. Using a larger line width of 12 kHz, resulted in a shorter limit of $R_{sim} \gtrsim 21$ Å (Fig. S28). Moreover, a single distance above 20 Å was not sufficient for reproducing the broad base of the ENDOR spectrum. This feature could be simulated by introducing an additional, shorter distance centred at about 16 Å and in a much smaller population (only 10% of the overall distance distribution). We note that the Mims ENDOR experiment enhances the contribution of shorter distances, making such a small contribution discernible in the spectrum. In order to examine to what extent conformational distributions affect the spectra, we also performed simulations of rigid Pake patterns based on a single R_{sim} value convoluted with the estimated intrinsic ENDOR line width (Fig. 3). Comparison of spectral simulations (Fig. 3, Fig. S29) showed indeed additional broadening caused by a distance distribution between F$^-$ and spin label, which was most pronounced for the shortest distance in G5.

5

In the second step of analysis, we estimated the contribution of the spin label to the ENDOR line width by rotamer modelling of label orientations in MtsslSuite and MMM (Fig. S8), based on the available crystal structure (PDB 4ENC, see section 1 in SI). This modelling produces rotamer ensembles of spin label orientations as represented in Fig. 4b. For each rotamer, distance and orientation were considered. The distance was taken from the midpoint of the NO-bond to the F^- while the orientation was defined through the Euler angles α and β between the dipolar tensor and the spin label g-tensor, as illustrated in Fig. 4a. The Mims ENDOR spectrum for each of up to 400 rotamers was computed in SimSpec[43] (see section 1 in SI) and then added to generate the sum ENDOR spectrum (Fig. 4b). Both diastereomers were considered, however the modelling does not predict their relative weight. For G5, both diastereomers were considered in equal weight, while for G43 only one, as the second led to steric clashes. For U46, the weight of the diastereomer leading to the shorter distance was adjusted to reproduce the ENDOR spectrum.

Starting from the construct labelled in position G5 we obtained a substantial deviation of 3 Å between the predicted R_{model} and the observed R_{sim} distance. However, the distance distribution width with ΔR = 2.1 Å was very close to the simulated value of 2.4 Å (Fig. 4b). For construct G43 we found a good agreement between the Gaussian and the predicted distance distribution. For U46 a small deviation in the long distance ($R_{model,2}$) could be observed while the width of the distribution contribution was in good agreement with the Gaussian.

For all three constructs, we found that the width of the distance distributions predicted by the rotamer modelling, when keeping a rigid body of the RNA, was consistent with the width extracted from the model-free Gaussian distributions. This means that no significant broadening contributions from the RNA backbone around the active side were required to simulate the ENDOR spectra. This is a significant result, which differed from the PELDOR findings, where the obtained distance distributions were largely exceeding the predicted rotamer distributions. Instead,

considering the mean distances, we observed good agreement only for G43 and U46.

The significant deviation observed for G5, exceeding the one derived from rotamer modelling, can be rationalised by crystal contacts in the 5' region not being present in solution (*vide infra*). We conclude that the ENDOR data, although sensitive to small distance shifts of a few angstroms, do not report structural heterogeneity of the *holo* aptamer in the vicinity of the binding pocket. This indicates that the RNA structure around the active site is well-defined in the *holo* form.

The observed deviations between the modelled and observed ^{19}F ENDOR distances around G5 prompted us to investigate trilateration of the F^- position. This places the F^- in a position between the pseudoknot and the unpaired bases between the pseudoknot and stem 2 (Fig. S31a). In this position a coordination of the F^--Mg^{II} cluster as seen in the crystal structure would not be possible as the F^- is located between the bases instead of the phosphate backbone (Fig. S31b), strongly indicating a different structure or conformation around G5 well beyond the error of the ^{19}F ENDOR measurement. The crystal structure shows significant crystal contacts (Fig. S32) for the first two bases from the 5' end, suggesting a different solution structure in absence of these contacts. Nevertheless, a high similarity between five crystal structures of the riboswitch in presence of various ions[4] as well as with the NMR derived structure of the *B. cereus* riboswitch[10] has been reported (Fig. S33). The main deviation between crystal and NMR derived structures[10] consists of a slightly different average conformation of the riboswitch in liquid solution. The deviation in the distance between the nitroxide spin label and the F^- at G5 could be explained by a different orientation of the first five nucleotides at the 5' end. The distance between the G5 phosphate group and the F^- is 6.3 Å in the crystal structure. The spin label rotamers have a distance of 7 - 7.5 Å between phosphate and NO group. Thus, the 12 Å distance seen in the ^{19}F ENDOR measurement (Fig. 3) could arise from a differently oriented phosphate group.

Table 1: Summary of the dipolar T, R, and ΔR parameters obtained for samples labelled at G5, G43 and U46 from the read-off spectra, the simulation with Gaussian distributions of distances, and simulations with the rotamer model as well as the line width (LW) used for simulations.

Sample	T_{read} / kHz	T_{sim} / kHz	R_{read} / Å	R_{sim} / Å	ΔR_{sim} / Å	R_{model} / Å	ΔR_{model} / Å	LW / kHz
G5	50 ± 7	43 ∈ [41, 53]	11.5 ∈ [10.9, 12.0]	12 ∈ [11.2, 12.2]	2.4	8.5	2.1	14
G43	30 ± 4	27 ∈ [25, 30]	13.5 ∈ [13.0, 14.2]	14 ∈ [13.6, 14.4]	2.4	13.5	2.2	14
U46	≲20	≲6	> 15	≳23	2.4	21.5	2.4	7

Abstract: Riboswitches control gene regulation upon external stimuli such as environmental factors or ligand binding. The fluoride sensing riboswitch from *Thermotoga petrophila* is a complex regulatory RNA proposed to be involved in resistance to F^- cytotoxicity. The details of structure and dynamics underpinning the regulatory mechanism are currently debated. Here we demonstrate that a combination of pulsed electron paramagnetic resonance (ESR/EPR) spectroscopies, detecting distances in the angstrom to nanometre range, can probe distinct regions of conformational flexibility in this riboswitch. PELDOR (pulsed electron-electron double resonance) revealed a similar preorganisation of the sensing domain in three forms, i.e. the free

aptamer, the Mg^{2+}-bound *apo*, and the F^--bound *holo* form. ^{19}F ENDOR (electron-nuclear double resonance) was used to investigate the active site structure of the F^--bound *holo* form. Distance distributions without *a priori* structural information were compared with *in silico* modelling of spin label conformations based on the crystal structure. While PELDOR, probing the periphery of the RNA fold, revealed conformational flexibility of the RNA backbone, ENDOR indicated low structural heterogeneity at the ligand binding site. Overall, the combination of PELDOR and ENDOR with sub-angstrom precision gave insight into structural organisation and

6

flexibility of a riboswitch, not easily attainable by other biophysical techniques.

Introduction

Protein structure and function underpins all aspects of life from the structural framework of cells, enzymatic activity, cellular transport, to immune function.[1] The production of the individual proteins is controlled through the expression of the corresponding genes. Riboswitches are non-coding RNAs mediating the regulation of genes[2–5] by initiating or terminating transcription or translation. Here, gene expression is being triggered by environmental factors such as temperature or pH or through ligands like ions or small molecules.[6–8] Ion or metabolite sensing riboswitches form highly selective binding pockets for their specific target.[9]

Fluoride sensing riboswitches are present in both, bacteria and archaea.[4,10–12] Increasing F^- concentrations can inhibit cell growth and become acutely toxic to cells.[12–16] This antimicrobial effect has been exploited to eliminate harmful microorganisms in fermentation processes.[17] Fluoride riboswitches are involved in cellular defence mechanisms, such as initiating the production of F^- exporters.[10,12,14]
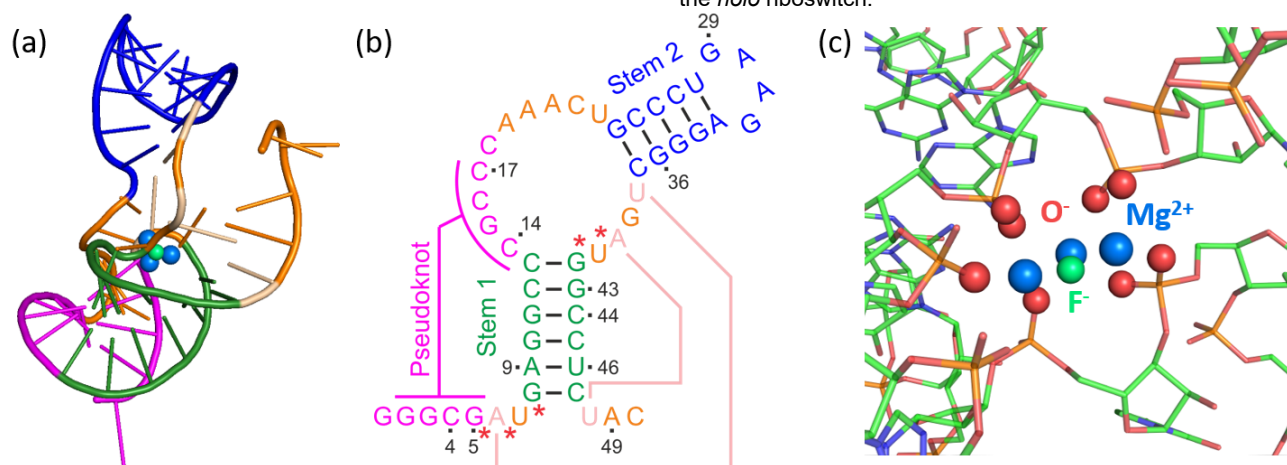
The first fluoride sensing riboswitch was found in the *crcB* motif of *Pseudomonas syringae* in 2012.[14] In the same year, the crystal structure of the sensing domain of the fluoride sensing riboswitch from the bacterium *Thermotoga petrophila* in its F^--bound form was reported by Ren *et al.* (Fig. 1a).[4] The tertiary structure of the aptamer comprises two stems, a pseudoknot, and reversed Watson-Crick (A6·U38) and Hoogsteen (A40·U48) base pairs (Fig. 1b). The riboswitch coordinates a cluster of three Mg^{2+} ions which in turn encapsulates the F^- ion (Fig. 1c). In this way the polyanionic RNA avoids electrostatic repulsion with the bound F^-.

This motif composed of Mg^{2+}, F^-, phosphate, and water was computationally found to be stable largely based on the electrostatic interactions between F^- and Mg^{2+}.[18] A recent computational study proposes a stepwise assembly of the cluster with two Mg^{2+} bound in the *apo* form but the third Mg^{2+} only being incorporated together with F^-.[19] The riboswitch is strongly selective for F^- discriminating against other halides.[4]

Using solution nuclear magnetic resonance (NMR) spectroscopy, the fluoride binding riboswitch from *Bacillus cereus* was shown to adopt a highly similar fold.[10] Here, stem 1 and stem 2 form already in absence of Mg^{2+} and F^-,[10] whereas the formation of the pseudoknot requires presence of Mg^{2+}.[20,21] Further addition of F^- does not result in a significant change of the tertiary structure but suppresses dynamics involving a lowly populated excited *apo* state responsible for transcription termination.[10,22]

Pulsed electron-electron double resonance (PELDOR,[23,24] aka DEER for double electron-electron resonance[25]) spectroscopy permits measuring distances ranging from 20 Å to 100 Å and beyond[26–28] between paramagnetic centres such as spin labels,[26,29–31] paramagnetic metal ions,[32–37] amino acid radicals and radical cofactors.[38–40] This provides a unique opportunity for investigating conformational ensembles through their spin-spin distance distributions. Complementary to PELDOR, ^{19}F electron-nuclear double resonance (ENDOR) is an emerging technique to detect distances in a shorter range of about 5 - 20 Å between nitroxide spin labels and ^{19}F nuclei.[41–44] This has recently been expanded to distances between fluorine nuclei and other spin centres such as triarylmethyl (TAM) radicals,[45,46] tyrosyl radicals,[47] Cu^{II},[48] or Gd^{III}.[49–51]

In this work, we have investigated the *T. petrophila* riboswitch in solution by magnetic resonance methods. ^1H NMR spectroscopy allowed monitoring the formation of base pairs through signals in the imino region of the free RNA, the Mg^{2+}-bound *apo* riboswitch, and the F^--bound *holo* riboswitch. We have further studied the solution-state preorganisation of the tertiary fold in the free and *apo* riboswitches by PELDOR and we have employed ^{19}F ENDOR to investigate the structure of the binding pocket of the *holo* riboswitch.

Figure 1. (a) Graphical representation of the crystal structure of the sensing domain of the fluoride binding riboswitch from *T. petrophila* in the *holo* form (PDB: 4ENC)[4] with the F^- (green sphere) encapsulated by three Mg^{2+} (blue spheres); (b) schematic representation of the secondary structure of the 50 nucleotide construct of the fluoride binding riboswitch used here, with the two stem structures (green and blue), the pseudoknot structure (magenta), and the reversed Watson-Crick (A6·U38) and reversed Hoogsteen (A40·U48) base pairs (salmon). Residue numbers for spin-labelling sites are indicated. The residues whose O^- are involved in the coordination of the Mg^{2+} are highlighted with red asterisks; (c) Graphical representation of the binding pocket of the fluoride binding riboswitch (PDB: 4ENC)[4]

7

with the F- (green sphere) encapsulated by three Mg^{2+} (blue spheres). O- involved in the coordination are represented as red spheres. Additional ions not involved in the fluoride encapsulation and water molecules have been omitted for clarity throughout.
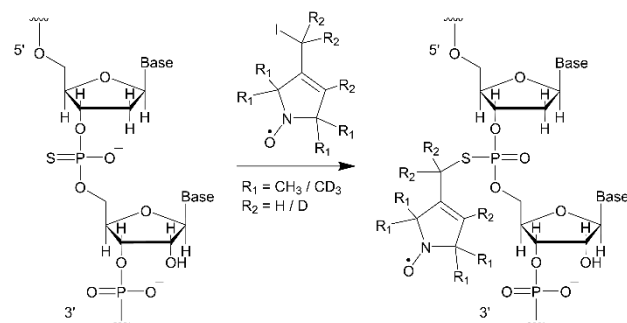
Results and Discussion

Design of spin labelled RNAs

PELDOR and ^{19}F ENDOR measurements require the presence of two or one paramagnetic spin labels, respectively. Several methods for site-directed spin labelling of nucleic acids at the base, ribose, or phosphate exist.[52,53] Custom synthesised RNA sequences with phosphorothioate modification are commercially available as is the spin label precursor for labelling at the phosphate backbone (Scheme 1) making this labelling approach readily accessible.[54,55] This procedure leads to the introduction of a diastereomeric pair. The spin labelled RNA is therefore expected to display both diastereomers of the phosphothiotriester.

All constructs were designed based on the crystal structure of the *T. petrophila* riboswitch.[4] The labelling sites were screened using *in silico* labelling with MtsslSuite[56] and MMM[57] to predict distances between pairs of spin labels or spin labels and F-. Furthermore, ensembles of spin label rotamers can also be extracted allowing analysis of orientation selection in the anisotropic EPR spectrum.[25] In absence of an experimental high-resolution structure a similar modelling can be applied using computational structure predictions that have become available also for nucleic acids with the introduction of AlphaFold3.[58]

To select the labelling sites, following general criteria were applied: (i) nucleotides involved in the formation of the cluster, i.e. residues coordinating Mg^{2+} were excluded (red asterisks in Fig. 1b); (ii) distances in the range between 18 – 45 Å were considered for PELDOR, and distances below 20 Å for ^{19}F ENDOR.

For PELDOR measurements four different constructs with two spin labels each were prepared: G29/G36 within stem 2 and the corresponding hairpin (blue), C14/C44 located in the region between the pseudoknot and stem 1 (green), A9/A49 across stem 1 and the 3' end of the RNA (magenta), and C4/C17 across the pseudoknot (red, Fig. S5 and Fig. S6).

Scheme 1. RNA labelling at the phosphorothioate modification of the backbone using either a protonated or a per-deuterated spin label. The sugar 5' to the labelling site is modified to a 2'-deoxyribose to avoid strand scission.[55]

For ^{19}F ENDOR, three constructs with one spin label each were chosen to result in linearly independent distance vectors between the respective label and the F-, allowing the trilateration of the F- position within the RNA structure (Fig. S7). G5 and G43 gave respective calculated mean distances (R_{model}) of 8.5 Å and 13.5 Å to the F-, well resolvable below the 20 Å threshold. For U46 however, *in silico* labelling displayed a bimodal distance distribution. The longer $R_{model,1}$ was predicted to be at about 21.5 Å, likely too long to be resolved by ^{19}F ENDOR. Nevertheless, the second, shorter $R_{model,2}$ of 15.5 Å was within the resolvable distance limit for ^{19}F ENDOR with nitroxide labels.[41]

To ensure reproducible folding of the riboswitch and infer any perturbation of folding by the spin label, we established a folding protocol (see SI section 1) using ^{1}H NMR spectroscopy. This allowed to identify the free aptamer in absence of Mg^{2+} and F-, the Mg^{2+}-bound *apo* riboswitch, and the F--bound *holo* riboswitch in analogy to results for the *B. cereus* fluoride riboswitch.[10]
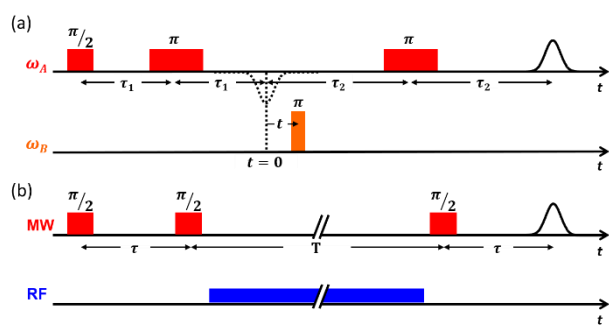
Comparison of ^{1}H NMR spectra in the imino regions of the free and *holo* aptamer for both, unlabelled and labelled RNA showed indistinguishable signals indicating appropriate folding after labelling (Fig. S10-16).

Alternatively, isothermal titration calorimetry (ITC) could be used to investigate labelled and mutated riboswitch constructs for retained aptamer functionality.

PELDOR measurements

PELDOR distance measurements using the 4-pulse DEER sequence (Scheme 2 (a)) are particularly well suited to study structural or conformational changes.[59–62] To map overall structure and preorganisation, PELDOR time traces were measured for the free, *apo* and *holo* aptamers for each of the four doubly spin labelled constructs. Sample preparation and experimental conditions for PELDOR are described in SI section 1. All spectra were recorded in frozen solution at 50 K. The time traces (Fig. 2) displayed differently pronounced dipolar modulations, which indicated different widths of the respective distance distributions. The time traces were analysed using Tikhonov regularisation implemented in the software DeerAnalysis (Fig. 2).[63] All obtained distance distributions, despite being broad, showed a distinct most probable distance, but did not show significant differences between the free, *apo* and *holo* aptamers for the C14/C44, G29/G36, and A9/A49 constructs. The C4/C17 construct displayed the most pronounced dipolar modulation resulting in the narrowest distance distribution with the lowest uncertainty. Here, the free and *apo* forms provided no difference within uncertainty but the *holo* form yielded a slightly reduced most probable distance. These results could be confirmed by an alternative processing with a different implementation of Tikhonov regularisation in the software DeerLab[64] and by neuronal network analysis with the software DeerNet[65] (Fig. S17). Interestingly, all distance distributions displayed significant distance probability up to 50 Å and beyond in all three processing approaches.
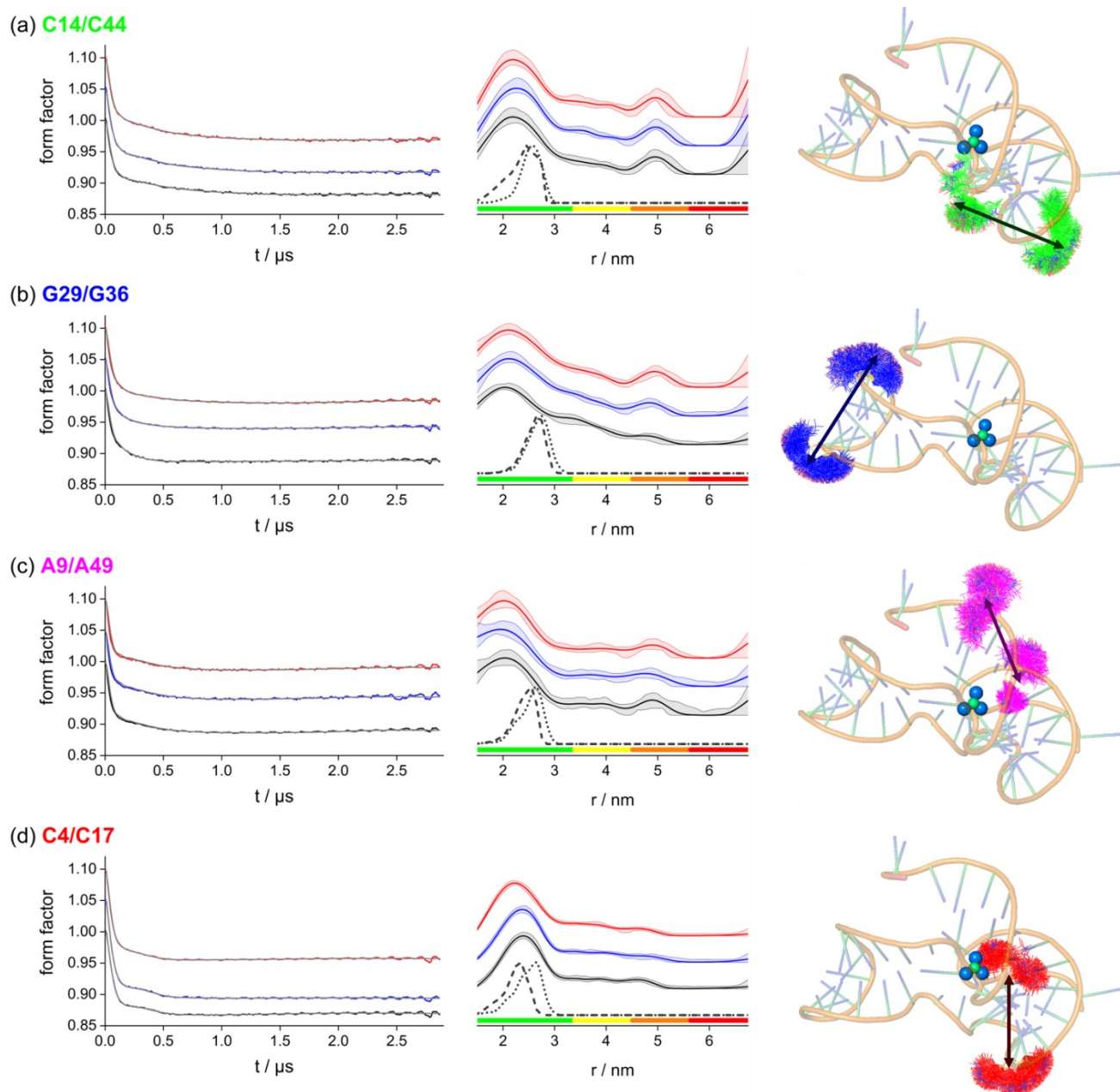
Scheme 2. Pulse sequences for 4-pulse DEER (a) and the Mims ENDOR (b) experiments with delays (t, τ), microwave (MW) pulses (π, π/2, red and orange), and a radiofrequency (RF) pulse (blue).

Comparing the experimental distance distributions with the modelling performed to select these constructs revealed substantially broader distributions in the former. Nevertheless, all modelled distance probability is fully covered by the experimental distributions. Particularly, the DeerLab analysis hinted at the presence of multiple distances corresponding to multiple conformations. This illustrated a conformational flexibility of the RNA backbone that was not reproduced by rotamer modelling on the single conformer in the crystal structure. Interestingly, for all cases but C4/C17 the most probable distances (i.e., maxima of the experimental distributions) were shorter than the maximum of the modelled distributions. This indicated the presence of more compact structural arrangements in solution than in the crystal structure. The similarity in the PELDOR data between the free, *apo*, and *holo* forms confirmed the preorganisation of stems 1 and 2 in the free RNA and their retention in the *apo* and *holo* aptamers. The shortening of the C4/C17 distance in the *holo* compared to the free and *apo* forms was consistent with an increased formation of the pseudoknot only upon addition of F$^-$ in contrast to the findings for the *B. cereus* fluoride riboswitch.[21] The long distances tailing up to 50 Å and beyond hint to an equilibrium involving a disordered or unfolded form of the aptamer. This was consistent with earlier fluoride binding ITC data that had been fitted to 0.75 and 0.87 binding sites[4] suggesting 13-25% of aptamers to binding F$^-$ and contributing very broad peaks to the distance distributions of all three forms.

https://doi.org/10.26434/chemrxiv-2024-f8czr-v2 **ORCID:** https://orcid.org/0000-0002-3384-271X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Figure 2. Background corrected PELDOR time traces (left) measured at Q-band MW frequency and 50 K and analysed using Tikhonov regularisation for the free (black), *apo* (blue), and *holo* (red) aptamer labelled at positions C14/C44 (a), G29/G36 (b), A9/A49 (c), and C4/C17 (d). Sample concentrations were adjusted to 10 µM RNA (a-c) and 20 µM RNA (d). The fit of the experimental time trace is given as grey line. Corresponding distance distributions (middle) with the respective 2σ confidence estimates. The colour bar is showing the reliability ranges (green: shape reliable, yellow: mean and width reliable, orange: mean reliable, red: no quantification possible) of the determined distributions. A vertical offset was introduced for both, the time traces and the distance distributions for clarity. The distance distributions obtained from *in silico* labelling with MtsslSuite (dashed line) and MMM (dotted line) are given for comparison. Graphical representation of RNA constructs (right, 4ENC[4]) spin labelled *in silico* with MtsslSuite and with arrows indicating the spin-spin distance assed by PELDOR measurements.

¹⁹F ENDOR spectroscopy

The fluoride binding riboswitch provides a unique opportunity to measure inter-spin distances to the active site in the *holo* aptamer, containing an endogenous F⁻, by ¹⁹F ENDOR spectroscopy. For this, ¹⁹F Mims ENDOR (pulse sequence in Scheme 2b) was performed with three constructs containing a single nitroxide label at positions G5, G43, or U46 (Fig. 1b). ¹⁹F ENDOR spectra of G5 and G43 were measured at W-band

(94 GHz) and four different excitation positions in the EPR spectrum (Fig. S19, left). The orientationally selected ENDOR spectra were weighted by the spectral intensity at the excitation position, added, and analysed (see section 1 in SI).

For construct U46, spectra were measured at Q-band (34 GHz), taking advantage of superior concentration sensitivity,[46] as the expected distance was longer (Fig. S20). At Q-band, ¹H ENDOR signals from the spin label methyl groups overlap with the ¹⁹F ENDOR signals (Fig. S21). This was circumvented by introducing a deuterated nitroxide spin label (see section 1 in SI).

10

Furthermore, this construct showed the slowest echo decay (i.e., slowest transverse dephasing, Fig. S22) contributing to improved sensitivity. At 34 GHz ^{19}F ENDOR spectra were measured only at two different spectral positions of the EPR line (Fig. S19, right), and no significant effect of orientation selection could be observed within the achievable resolution and signal-to-noise ratio.
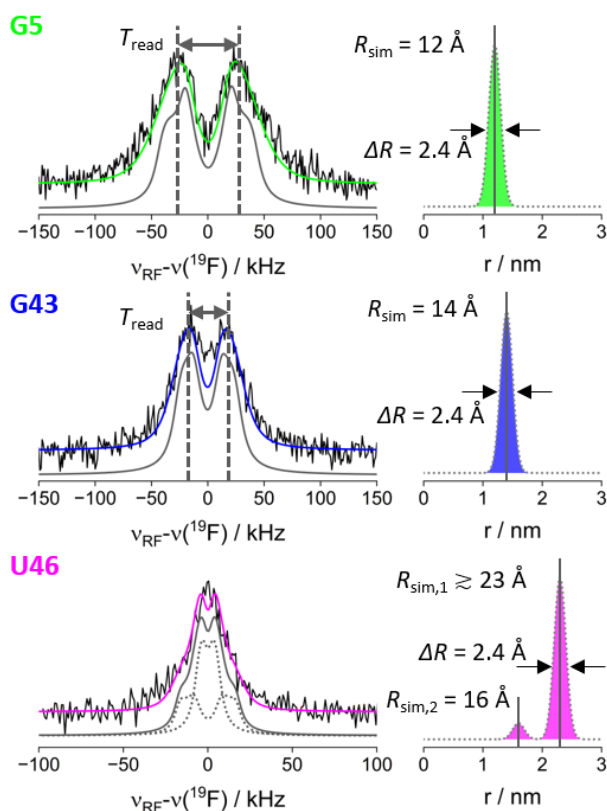
Figure 3. Mims ^{19}F ENDOR sum spectra (left column) of the *holo* aptamer and simulated spectra (coloured lines) using Gaussian distributions (right column) or based on T_{read} (grey lines, left). In case of U46 the simulated spectrum contains contributions of two distance components (dotted grey). Mean distance (R_{sim}) and FWH (ΔR) of the distance distributions used for simulations are given. Samples were composed of 100 μM, 150 μM, and 235 μM RNA for the RNA labelled in position G5, G43, and U46, respectively. Measurements were performed at W-band (G5 and G43) and Q-band (U46) and at 50 K. For W-band, a protonated spin label and for Q-band a deuterated one was used.

The three ^{19}F ENDOR spectra of the *holo* riboswitch are displayed in Fig. 3. For constructs G5 and G43, a well-resolved hyperfine splitting arising from the maxima of a Pake pattern[41] was observed. From this splitting, called T_{read}, the perpendicular component $T_\perp$ of the hyperfine tensor was estimated, giving first insight into the magnitude of the dipolar coupling and the corresponding inter-spin distance. Using the point-dipolar approximation the point-dipole distance (R_{read}) can be estimated as:[41]

$$T_{read} \approx T_\perp = \frac{\mu_0}{4\pi h}\left(\frac{g_e g_n \mu_B \mu_N}{R_{read}^3}\right) = \frac{C}{R^3} \qquad \text{Eq. 1}$$

with the vacuum permeability μ_0, the Planck constant h, the g-values of the electron and fluorine nucleus g_e and g_n,

respectively, the Bohr magneton μ_B and the nuclear magneton μ_N. With these constants C amounts to C = 74.52 MHz Å^3.

For G5, a coupling constant T_{read} = 50 ± 7 kHz was read off that corresponds (Eq. 1) to a point-dipole distance of R_{read} = 11.5 ± 0.5 Å. The error was estimated from the width of the maxima of the peaks. For G43, a smaller splitting from a coupling constant T_{read} = 30 ± 4 kHz was observed, corresponding to a point-dipole distance of 13.5 ± 0.6 Å. For U46 no splitting could be observed, however a clear peak was still detected. To ensure that this signal of U46 did not arise from free F$^-$ interacting with the nitroxide label, a control experiment was performed. ^{19}F ENDOR spectra of the *free* aptamer labelled in position G43 and U46 were measured in the presence of F$^-$ at W-band and Q-band, respectively. In absence of Mg^{2+} the F$^-$ binding site cannot form and indeed no ENDOR effect has been observed (Fig. S23). Thus, all ^{19}F ENDOR signals observed in the *holo* form arise from bound F$^-$.

To extract more information on conformational distributions from the ENDOR line shape we performed a spectral analysis in two steps. First, we simulated the spectra using our fast simulation routine SimSpec[43] (see section 1 in SI) without any prior model, just assuming a Gaussian distribution of distances centred at R_{read}. The centre and the width of the Gaussian distribution were manually adjusted to minimize the residual between experiment and simulated spectrum (Fig. S25 and S26). The simulations are displayed in Fig. 3 (left). The Gaussian distance distribution reproduced the line shapes of G5 and G43 very well using a broadening (full width at half maximum, FWHM) corresponding to 2.4 Å. The simulated T_{sim} component resulted in R_{sim} values as given in Fig. 3 (right) and in Table 1. For U46 a series of simulations was performed for Gaussian distributions of constant width 2.4 Å but varying R_{sim} between 18 and 28 Å. Examination of the residuals (Fig. S27 and Fig. S28) suggested that $R_{sim} \gtrsim 20$ Å can generally be considered a lower boundary for the distance, when no splitting occurs. We note that this is the first detection of a ^{19}F – nitroxide distance $\gtrsim 20$ Å, demonstrating an extension of the accessible distance range. Indeed, the resolution of the splitting depended on the choice of the line width parameter. Our choice of a Lorentzian line of 7 kHz (slightly larger than the power broadening of the RF pulse) is based on an ongoing investigation of intrinsic ENDOR line widths. This resulted in $R_{sim} \gtrsim 23$ Å. Using a larger line width of 12 kHz, resulted in a shorter limit of $R_{sim} \gtrsim 21$ Å (Fig. S28). Moreover, a single distance above 20 Å was not sufficient for reproducing the broad base of the ENDOR spectrum. This feature could be simulated by introducing an additional, shorter distance centred at about 16 Å and in a much smaller population (only 10% of the overall distance distribution). We note that the Mims ENDOR experiment enhances the contribution of shorter distances, making such a small contribution discernible in the spectrum. In order to examine to what extent conformational distributions affect the spectra, we also performed simulations of rigid Pake patterns based on a single R_{sim} value convoluted with the estimated intrinsic ENDOR line width. Comparison of spectral simulations (Fig. 3, Fig. S29) showed indeed additional broadening caused by a distance distribution between F$^-$ and spin label, which was most pronounced for the shortest distance in G5.

11

In the second step of analysis, we estimated the contribution of the spin label to the ENDOR line width by rotamer modelling of label orientations in MtsslSuite and MMM (Fig. S8), based on the available crystal structure (PDB 4ENC, see section 1 in SI). This modelling produces rotamer ensembles of spin label orientations as represented in Fig. 4b. For each rotamer, distance and orientation were considered. The distance was taken from the midpoint of the NO-bond to the F^- while the orientation was defined through the Euler angles α and β between the dipolar tensor and the spin label g-tensor, as illustrated in Fig. 4a. The Mims ENDOR spectrum for each of up to 400 rotamers was computed in SimSpec[43] (see section 1 in SI) and then added to generate the sum ENDOR spectrum (Fig. 4b). Both diastereomers were considered, however the modelling does not predict their relative weight. For G5, both diastereomers were considered in equal weight, while for G43 only one diastereomer was considered, as the second cannot be populated due to steric clashes. For U46, the weight of the diastereomer leading to the shorter distance was adjusted to reproduce the ENDOR spectrum.

Starting from the construct labelled in position G5 we obtained a substantial deviation of 3 Å between the predicted R_{model} and the observed R_{sim} distance. However, the distance distribution width with ΔR = 2.1 Å was very close to the simulated value of 2.4 Å (Fig. 4b). For construct G43 we found a good agreement between the Gaussian and the predicted distance distribution. For U46 a small deviation in the long distance ($R_{model,1}$) could be observed while the width of this distribution contribution was in good agreement with the Gaussian.

For all three constructs, we found that the width of the distance distributions predicted by the rotamer modelling, when keeping a rigid body of the RNA, was consistent with the width extracted from the model-free Gaussian distributions. This means that no substantial broadening arising from RNA backbone heterogeneity is observed in the ENDOR spectra. This is a significant result, which differed from the PELDOR findings, where the experimental distance distributions were considerably broader than the ones predicted from rotamer modelling. This is important, as it is known that rigid biomolecules (e.g., model protein GB1) allow reconciling PELDOR and RIDME data[66,67] and ENDOR data[50] all based on a single protein structure.
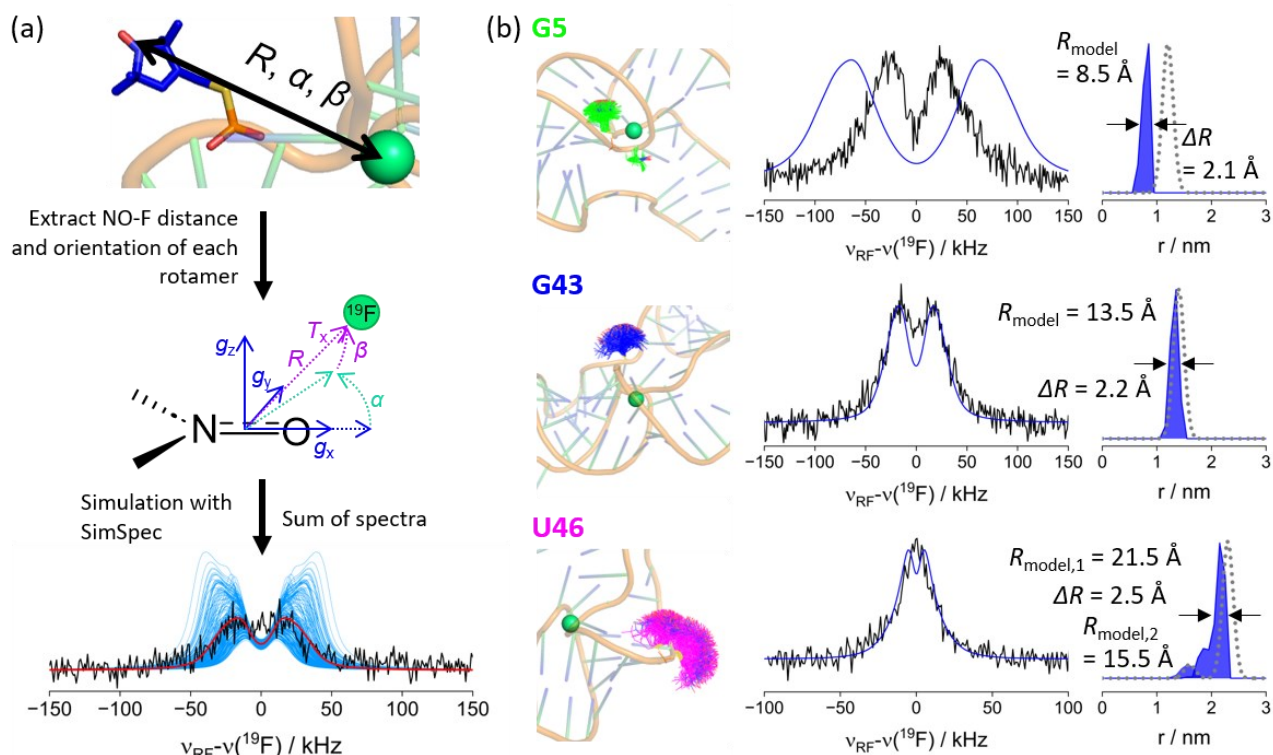
Furthermore, considering the mean distances, we observed good agreement only for G43 and U46. At G5, the significant deviation between the experimental and modelled distances can be rationalised by crystal contacts in the 5' region (Fig. S32) not being present in solution (*vide infra*). We conclude that the ENDOR data, although sensitive to small distance shifts of a few angstroms, do not report structural heterogeneity of the *holo* aptamer in the vicinity of the binding pocket. This indicates that the RNA structure around the active site is well-defined in the *holo* form.

The observed deviations between the modelled and experimental ^{19}F ENDOR distances around G5 prompted us to trilaterate the F^- position based on the experimental distances. This places the F^- in a position between the pseudoknot and the unpaired bases between the pseudoknot and stem 2 (Fig. S31a). Here, the F^--Mg^{2+} cluster could not be stabilised by coordination from the phosphate backbone (Fig. S31b), strongly indicating a different structure or conformation around G5 well beyond the error of the ^{19}F ENDOR measurement. The crystal structure shows significant intermolecular contacts (Fig. S32) for the first two bases from the 5' end, suggesting a structural variation in absence of these contacts in solution. The deviation in the distance between the nitroxide spin label and the F^- at G5 could be explained by a different orientation of the first five nucleotides at the 5' end. The distance between the G5 phosphate group and the F^- is 6.3 Å in the crystal structure. The spin label rotamers have distances of 7 - 7.5 Å between phosphate and NO group adding to a maximum NO-F distance of about 14 Å. Thus, the 12 Å distance seen in the ^{19}F ENDOR measurement (Fig. 3) could arise from a differently oriented phosphate group. Nevertheless, a high similarity between five crystal structures of the riboswitch in presence of various ions[4] as well as with the NMR derived structure of the *B. cereus* riboswitch[10] has been reported (Fig. S33).

Thus, the ENDOR and PELDOR results reveal different levels of structural heterogeneity – representing different functional dynamics - not directly obvious from crystal structures.

Table 1: Summary of the main contributions to the dipolar T, R, and ΔR parameters and uncertainties obtained for samples labelled at G5, G43 and U46 from the read-off spectra, the simulation with Gaussian distributions of distances, and simulations with the rotamer model as well as the line width (LW) used for simulations. Errors of T_{read} and R_{read} were estimated from the width of the maxima of the peaks in the spectra. Errors of T_{sim} and R_{sim} were estimated based on simulations with different T_{sim}/R_{sim}.

Sample	T_{read} / kHz	T_{sim} / kHz	R_{read} / Å	R_{sim} / Å	ΔR_{sim} / Å	R_{model} / Å	ΔR_{model} / Å	LW / kHz
G5	50 ± 7	43 ∈ [41, 53]	11.5 ∈ [10.9, 12.0]	12 ∈ [11.2, 12.2]	2.4	8.5	2.1	14
G43	30 ± 4	27 ∈ [25, 30]	13.5 ∈ [13.0, 14.2]	14 ∈ [13.6, 14.4]	2.4	13.5	2.2	14
U46	⩽20	⩽6	> 15	⩾23	2.4	21.5	2.4	7

Figure 4. (a) Schematic representation of the workflow for the simulation of ^{19}F ENDOR spectra based on the crystal structure spin labelled *in silico*. Distances (R) and orientations (α, β) of all rotamers were extracted, a ^{19}F ENDOR spectrum for each rotamer was simulated using SimSpec[43] and the sum of the simulated spectra was built; (b) Graphical representation of the fluoride riboswitch constructs labelled in position G5, G43, or U46 (left) with the respective Mims ^{19}F ENDOR sum spectra (centre, black) and simulated spectra (centre, blue) using the distance distributions from *in silico* spin labelling (right). Mean distance (R_{model}) and FWHM (ΔR) of the distance distributions from *in silico spin* labelling are given. The Gaussian distribution of distances is plotted as grey dotted line for comparison.

Conclusion

In this work we have shown that the free, *apo* and *holo* forms of the sensing domain of the fluoride riboswitch from *T. petrophila* could be reproducibly formed *in vitro*. However, a significant population of unfolded or disordered aptamer was present in all PELDOR samples manifesting as distance distributions tailing up to 50 Å and beyond. Nevertheless, PELDOR has revealed a preorganisation similar to the *B. cereus* fluoride riboswitch with the exception that also F$^-$ was required for pseudoknot formation.[10] Modelling the widths of distance populations based on the single backbone conformation of the crystal structure matched the experimental distributions for ^{19}F ENDOR but significantly underestimated the experimental PELDOR distance distributions. Thus, the combined PELDOR and ^{19}F ENDOR results indicated a preorganised riboswitch structure with a more rigid F$^-$-binding site and more conformations sampled at the periphery.

For further understanding of conformational flexibility and presence of unfolded forms, investigation of the fluoride riboswitch structure in solution using, e.g., NMR or fluorescence methods is needed. Molecular dynamic simulations are complicated by the unusual cluster of Mg^{2+} ions, but might lead to further insight.[18,19]

As a next step, extending the riboswitch from the sensing domain to the full-length expression platform will potentially allow forming the transcription terminator state in the absence of F$^-$ and observing the inherent conformational change by EPR methods, giving further insights into this intricate molecular mechanism of gene regulation.

Finally, we have shown that the use of a fluorine bearing ligand allows to selectively probe the ligand bound form by ^{19}F ENDOR in a complex mixture of bound and unbound molecules. In a more general context, similar strategies could be employed to study ligand binding of fluorinated pharmaceuticals to biological targets. ^{19}F ENDOR complements PELDOR with more precise distances on a shorter length scale. Here, we first observed a F$^-$-nitroxide distance in the range of 2 nm, which completely closes the gap in the distance range accessible from pulse dipolar spectroscopies. Therefore, the combination of the two EPR-based methods, for distance measurements in frozen solution under similar sample conditions, provides valuable information on structure and conformational distributions, constituting a powerful tool for validation of structural models of complex biological systems.

Supporting Information

The authors have cited additional references within the Supporting Information.[68-78]

Original spectroscopic data and codes will be made available after publication in the Göttinger Research Online Data Base.

13

Acknowledgements

Keywords: ENDOR • PELDOR • DEER • fluorine • RNA

[1] J. M. Berg, J. L. Tymoczko, G. J. Gatto jr., L. Stryer, *Stryer Biochemie*, Springer, **2018**.

[2] A. S. Mironov, I. Gusarov, R. Rafikov, L. E. Lopez, K. Shatalin, R. A. Kreneva, D. A. Perumov, E. Nudler, *Cell* **2002**, *111*, 747–756.

[3] W. Winkler, A. Nahvi, R. R. Breaker, *Nature* **2002**, *419*, 952–956.

[4] A. Ren, K. R. Rajashankar, D. J. Patel, *Nature* **2012**, *486*, 85–89.

[5] A. Serganov, E. Nudler, *Cell* **2013**, *152*, 17–24.

[6] A. V Sherwood, T. M. Henkin, *Annu. Rev. Microbiol.* **2016**, *70*, 361–374.

[7] S. S. Krajewski, F. Narberhaus, *Biochim. Biophys. Acta - Gene Regul. Mech.* **2014**, *1839*, 978–988.

[8] G. Nechooshtan, M. Elgrably-Weiss, A. Sheaffer, E. Westhof, S. Altuvia, *Genes Dev.* **2009**, *23*, 2650–2662.

[9] K. Kavita, R. R. Breaker, *Trends Biochem. Sci.* **2023**, *48*, 119–141.

[10] B. Zhao, S. L. Guffy, B. Williams, Q. Zhang, *Nat. Chem. Biol.* **2017**, *13*, 968–974.

[11] M. C. Speed, B. W. Burkhart, J. W. Picking, T. J. Santangelo, *Appl. Environ. Microbiol.* **2018**, *84*, e02306-17.

[12] S. Li, K. D. Smith, J. H. Davis, P. B. Gordon, R. R. Breaker, S. A. Strobel, *Proc. Natl. Acad. Sci.* **2013**, *110*, 19018–19023.

[13] N. R. Johnston, S. A. Strobel, *Arch. Toxicol.* **2020**, *94*, 1051–1069.

[14] J. L. Baker, N. Sudarsan, Z. Weinberg, A. Roth, R. B. Stockbridge, R. R. Breaker, *Science* **2012**, *335*, 233–235.

[15] S. Flisfisch, J. Meyer, J. H. Meurman, T. Waltimo, *Oral Dis.* **2008**, *14*, 296–301.

[16] M. Treshow, *Mycologia* **1965**, *57*, 216–221.

[17] L. H. Weinstein, A. Davison, *Fluorides in the Environment: Effects on Plants and Animals*, CABI Publishing PP - Cambridge, MA, **2004**.

[18] M. Chawla, R. Credendino, A. Poater, R. Oliva, L. Cavallo, *J. Am. Chem. Soc.* **2015**, *137*, 299–306.

[19] S. Kumar, G. Reddy, *J. Phys. Chem. B* **2023**, *127*, 9267–9281.

[20] K. E. Watters, E. J. Strobel, A. M. Yu, J. T. Lis, J. B. Lucks, *Nat. Struct. Mol. Biol.* **2016**, *23*, 1124–1131.

[21] J. Lee, S. E. Sung, J. Lee, J. Y. Kang, J. H. Lee, B. S. Choi, *Int. J. Mol. Sci.* **2021**, *22*, 1–12.

[22] R. Yadav, J. R. Widom, A. Chauvier, N. G. Walter, *Nat. Commun.* **2022**, *13*, 207.

[23] A. D. Milov, K. M. Salikhov, M. D. Shirov, *Fiz. Tverd. Tela* **1981**, *23*, 975–982.

[24] M. Pannier, S. Veit, A. Godt, G. Jeschke, H. W. Spiess, *J. Magn. Reson.* **2000**, *142*, 331–340.

[25] R. G. Larsen, D. J. Singel, *J. Chem. Phys.* **1993**, *98*, 5134–5146.

[26] G. Jeschke, *Annu. Rev. Phys. Chem.* **2012**, *63*, 419–446.

[27] B. Endeward, Y. Hu, G. Bai, G. Liu, T. F. Prisner, X. Fang, *Biophys. J.* **2022**, *121*, 37–43.

[28] O. Schiemann, C. A. Heubach, D. Abdullin, K. Ackermann, M. Azarkh, E. G. Bagryanskaya, M. Drescher, B. Endeward, J. H. Freed, L. Galazzo, D. Goldfarb, T. Hett, L. Esteban Hofer, L. Fábregas Ibáñez, E. J. Hustedt, S. Kucher, I. Kuprov, J. E. Lovett, A. Meyer, S. Ruthstein, S. Saxena, S. Stoll, C. R. Timmel, M. Di Valentin, H. S. Mchaourab, T. F. Prisner, B. E. Bode, E. Bordignon, M. Bennati, G. Jeschke, *J. Am. Chem. Soc.* **2021**, *143*, 17875–17890.

[29] J. L. Wort, K. Ackermann, D. G. Norman, B. E. Bode, *Phys. Chem. Chem. Phys.* **2021**, *23*, 3810–3819.

[30] T. E. Assafa, S. Nandi, D. Śmiłowicz, L. Galazzo, M. Teucher, C. Elsner, S. Pütz, S. Bleicken, A. Y. Robin, D. Westphal, I. Uson, R. Stoll, P. E. Czabotar, N. Metzler-Nolte, E. Bordignon, *Structure* **2021**, *29*, 114-124.e3.

[31] M. Rudolph, R. Tampé, B. Joseph, *Angew. Chemie Int. Ed.* **2023**, *62*, e202307091.

[32] K. Ackermann, S. Khazaipoul, J. L. Wort, A. I. S. Sobczak, H. El Mkami, A. J. Stewart, B. E. Bode, *J. Am. Chem. Soc.* **2023**, *145*, 8064–8072.

[33] K. Ackermann, C. A. Heubach, O. Schiemann, B. E. Bode, *J. Phys. Chem. Lett.* **2024**, *15*, 1455–1461.

[34] M. Oranges, J. L. Wort, M. Fukushima, E. Fusco, K. Ackermann, B. E. Bode, *J. Phys. Chem. Lett.* **2022**, *13*, 7847–7852.

[35] J. L. Wort, K. Ackermann, A. Giannoulis, A. J. Stewart, D. G. Norman, B. E. Bode, *Angew. Chemie Int. Ed.* **2019**, *58*, 11681–11685.

[36] T. F. Cunningham, M. R. Putterman, A. Desai, W. S. Horne, S. Saxena, *Angew. Chemie Int. Ed.* **2015**, *54*, 6330–6334.

[37] M. Oranges, A. Giannoulis, A. Vanyushkina, Y. F. Sirkis, A. Dalaloyan, T. Unger, X.-C. Su, M. Sharon, D. Goldfarb, *Biophys. J.* **2024**, *123*, 172–183.

[38] M. R. Seyedsayamdost, C. T. Y. Chan, V. Mugnaini, J. Stubbe, M. Bennati, *J. Am. Chem. Soc.* **2007**, *129*, 15748–15749.

[39] V. P. Denysenkov, D. Biglino, W. Lubitz, T. F. Prisner, M. Bennati, *Angew. Chemie Int. Ed.* **2008**, *47*, 1224–1227.

[40] M. Bennati, A. Weber, J. Antonic, D. L. Perlstein, J. Robblee, J. Stubbe, *J. Am. Chem. Soc.* **2003**, *125*, 14988–14989.

[41] A. Meyer, S. Dechert, S. Dey, C. Höbartner, M. Bennati, *Angew. Chemie Int. Ed.* **2020**, *59*, 373–379.

[42] A. Kehl, M. Hiller, F. Hecker, I. Tkach, S. Dechert, M. Bennati, A. Meyer, *J. Magn. Reson.* **2021**, *333*, 107091.

[43] H. Wiechers, A. Kehl, M. Hiller, B. Eltzner, S. F. Huckemann, A. Meyer, I. Tkach, M. Bennati, Y. Pokern, *J. Magn. Reson.* **2023**, *353*, 107491.

[44] M. Gauger, M. Heinz, A.-L. J. Halbritter, L. S. Stelzl, N. Erlenbach, G. Hummer, S. T. Sigurdsson, T. F. Prisner, *Angew. Chemie Int. Ed.* **2024**, *63*, e202402498.

[45] N. B. Asanbaeva, A. A. Sukhanov, A. A. Diveikina, O. Y. Rogozhnikova, D. V Trukhin, V. M. Tormyshev, A. S. Chubarov, A. G. Maryasov, A. M. Genaev, A. V Shernyukov, G. E. Salnikov, A. A. Lomzov, D. V Pyshnyi, E. G. Bagryanskaya, *Phys. Chem. Chem. Phys.* **2022**.

[46] N. B. Asanbaeva, D. S. Novopashina, O. Y. Rogozhnikova, V. M. Tormyshev, A. Kehl, A. A. Sukhanov, A. V Shernyukov, A. M. Genaev, A. A. Lomzov, M. Bennati, A. Meyer, E. G. Bagryanskaya, *Phys. Chem. Chem. Phys.* **2023**.

[47] A. Meyer, A. Kehl, C. Cui, F. A. K. Reichardt, F. Hecker, L.-M. Funk, M. K. Ghosh, K.-T. Pan, H. Urlaub, K. Tittmann, J. Stubbe, M. Bennati, *J. Am. Chem. Soc.* **2022**, *144*, 11270–11282.

[48] S. L. Schumann, S. Kotnig, Y. Kutin, M. Drosou, L. M. Stratmann, Y. Streltsova, A. Schnegg, D. A. Pantazis, G. H. Clever, M. Kasanmascheff, *Chem. – A Eur. J.* **2023**, *29*, e202302527.

[49] M. Judd, E. H. Abdelkader, M. Qi, J. R. Harmer, T. Huber, A. Godt, A. Savitsky, G. Otting, N. Cox, *Phys. Chem. Chem. Phys.* **2022**, *24*, 25214–25226.

[50] M. Seal, W. Zhu, A. Dalaloyan, A. Feintuch, A. Bogdanov, V. Frydman, X.-C. Su, A. M. Gronenborn, D. Goldfarb, *Angew. Chemie Int. Ed.* **2023**, *62*, e202218780.

[51] A. Bogdanov, V. Frydman, M. Seal, L. Rapatskiy, A. Schnegg, W. Zhu, M. Iron, A. M. Gronenborn, D. Goldfarb, *J. Am. Chem. Soc.* **2024**, *146*, 6157–6167.

[52] S. A. Shelke, S. T. Sigurdsson, in (Eds.: K. Nakatani, Y. Tor), Springer International Publishing, Cham, **2016**, pp. 159–187.

[53] O. Schiemann, N. Piton, J. Plackmeyer, B. E. Bode, T. F. Prisner, J. W. Engels, *Nat. Protoc.* **2007**, *2*, 904–923.

[54] Q. Cai, A. K. Kusnetzow, K. Hideg, E. A. Price, I. S. Haworth, P. Z. Qin, *Biophys. J.* **2007**, *93*, 2110–2117.

[55] P. Z. Qin, I. S. Haworth, Q. Cai, A. K. Kusnetzow, G. P. G. Grant, E. A. Price, G. Z. Sowa, A. Popova, B. Herreros, H. He, *Nat. Protoc.* **2007**, *2*, 2354–2365.

[56] G. Hagelueken, R. Ward, J. H. Naismith, O. Schiemann, *Appl. Magn. Reson.* **2012**, *42*, 377–391.

[57] Y. Polyhach, E. Bordignon, G. Jeschke, *Phys. Chem. Chem. Phys.* **2011**, *13*, 2356–2366.

[58] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J. M. Jumper, *Nature* **2024**, *630*, 493–500.

[59] T. Hett, T. Zbik, S. Mukherjee, H. Matsuoka, W. Bönigk, D. Klose, C. Rouillon, N. Brenner, S. Peuker, R. Klement, H.-J. Steinhoff, H. Grubmüller, R. Seifert, O. Schiemann, U. B. Kaupp, *J. Am. Chem. Soc.* **2021**, *143*, 6981–6989.

[60] C. M. Grytz, A. Marko, P. Cekan, S. T. Sigurdsson, T. F. Prisner, *Phys. Chem. Chem. Phys.* **2016**, *18*, 2993–3002.

[61] B. Joseph, A. Sikora, D. S. Cafiso, *J. Am. Chem. Soc.* **2016**, *138*, 1844–1847.

[62] S. McQuarrie, J. S. Athukoralage, S. A. McMahon, S. Graham, K. Ackermann, B. E. Bode, M. F. White, T. M. Gloster, *Nucleic Acids Res.* **2023**, *51*, 10590–10605.

[63] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger, H. Jung, *Appl. Magn. Reson.* **2006**, *30*, 473–498.

[64] L. Fábregas Ibáñez, G. Jeschke, S. Stoll, *Magn. Reson.* **2020**, *1*, 209–224.

[65] S. G. Worswick, J. A. Spencer, G. Jeschke, I. Kuprov, *Sci. Adv.* **2018**, *4*, eaat5218.

[66] K. Ackermann, A. Chapman, B. E. Bode, *Molecules* **2021**, *26*, 7534.

[67] V. Vitali, K. Ackermann, G. Hagelueken, B. E. Bode, *Appl. Magn. Reson.* **2024**, *55*, 187–205.

[68] G. Hagelueken, D. Abdullin, R. Ward, O. Schiemann, *Mol. Phys.* **2013**, *111*, 2757–2766.

[69] C. E. Tait, S. Stoll, *Phys. Chem. Chem. Phys.* **2016**, *18*, 18470–18485.

[70] K. Ackermann, C. Pliotas, S. Valera, J. H. Naismith, B. E. Bode, *Biophys. J.* **2017**, *113*, 1968–1978.

[71] G. Hagelueken, D. Abdullin, O. Schiemann, in *Methods Enzymol.* (Eds.: P.Z. Qin, K.B.T.-M. in E. Warncke), Academic Press, **2015**, pp. 595–622.

[72] G. Jeschke, *Protein Sci.* **2018**, *27*, 76–85.

[73] S. D. Pirrwitz J., *Verfahren Zur Herstellung von 2,2,6,6-Tetramethyl-4-Oxopiperidin*, **1984**, DD222017A (WP C 07 D/260 901 6).

[74] L. A. Shundrin, I. A. Kirilyuk, I. A. Grigor'ev, *Mendeleev Commun.* **2014**, *24*, 298–300.

[75] M. M. Haugland, A. H. El-Sagheer, R. J. Porter, J. Peña, T. Brown, E. A. Anderson, J. E. Lovett, *J. Am. Chem. Soc.* **2016**, *138*, 9069–9072.

[76] G. Úr, T. Kálai, K. Hideg, *Tetrahedron Lett.* **2016**, *57*, 778–780.

[77] P. L. Beaulieu, P. C. Anderson, D. R. Cameron, J. Croteau, V. Gorys, C. Grand-Maître, D. Lamarre, F. Liard, W. Paris, L. Plamondon, F. Soucy, D. Thibeault, D. Wernic, C. Yoakim, S. Pav, L. Tong, *J. Med. Chem.* **2000**, *43*, 1094–1108.

[78] T. L. Hwang, A. J. Shaka, *J. Magn. Reson. Ser. A* **1995**, *112*, 275–279.

Entry for the Table of Contents
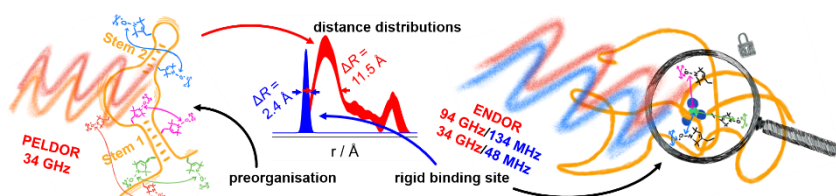
A fluoride sensing riboswitch was investigated by EPR-based distance measurements. Electron-electron double resonance revealed preorganisation of the stem structures and conformational flexibility while electron-nuclear double resonance demonstrates low structural heterogeneity around the fluoride site and distinct differences between the crystal and solution structures.

Institute and/or researcher Twitter usernames: @Bela_Bode; @StAndrewsChem