

Published in final edited form as:

*J Magn Reson.* 2021 May 01; 326: 106937. doi:10.1016/j.jmr.2021.106937.

## Optimal design of adaptively sampled NMR experiments for measurement of methyl group dynamics with application to a ribosome-nascent chain complex

Christopher A. Waudby,

Charles Burridge,

John Christodoulou

Department of Structural and Molecular Biology, UCL, Gower St, London WC1E 6BT, UK

### Abstract

NMR measurements of cross-correlated nuclear spin relaxation provide powerful probes of polypeptide dynamics and rotational diffusion, free from contributions due to chemical exchange or interactions with external spins. Here, we report on the development of a sensitivity-optimized pulse sequence for the analysis of the differential relaxation of transitions within isolated  $^{13}\text{CH}_3$  spin systems, in order to characterise rotational diffusion and side chain order through the product  $S^2\tau_c$ . We describe the application of optimal design theory to implement a real-time ‘on-the-fly’ adaptive sampling scheme that maximizes the accuracy of the measured parameters. The increase in sensitivity obtained using this approach enables quantitative measurements of rotational diffusion within folded states of translationally-arrested ribosome–nascent chain complexes of the FLN5 filamin domain, and can be used to place strong limits on interactions between the domain and the ribosome surface.

### Keywords

parametric estimation; sequential design; Cramér–Rao lower bound; co-translational folding; ribosome

### Introduction

Protein biosynthesis by the ribosome is a process of fundamental importance to the cell, ensuring that nascent polypeptides attain their functional native states while avoiding misfolding or aggregation [1]. Elaborate quality control networks have evolved to chaperone folding proteins, and to sense and eliminate misfolded states [2], but for many proteins a key additional facet of efficient biosynthesis is their progressive, ‘co-translational’ folding while still attached to the parent ribosome [3]. Co-translational folding can reduce the effective volume of conformational space accessible to a nascent polypeptide, facilitating the search for the native state, but it is increasingly apparent that this process can also be modulated or

---

Correspondence to: John Christodoulou.

c.waudby@ucl.ac.uk, j.christodoulou@ucl.ac.uk.

regulated by the ribosome itself [4]. In particular, a variety of polypeptides have now been observed or inferred to interact with the ribosome surface during translation. This has been associated with perturbations to the free-energy surfaces for folding, in the majority of cases leading to a destabilisation of the nascent polypeptide chain (NC) [5–10].

In our prior work, we have studied extensively the co-translational folding of the FLN5 immunoglobulin domain from the multidomain filamin protein of *Dictyostelium discoideum*, using a library of constructs comprising the FLN5 domain attached to the ribosome via varying lengths of the subsequent FLN6 domain and SecM arrest peptide (Fig. 1A,B) [9,11–14]. We have observed site-specific line broadening in unfolded NCs, which are likely to reflect interactions with the ribosome surface, and we have speculated that these interactions may partly underlie the observed delay in the initiation of co-translational folding several residues beyond the point at which the domain is fully emerged from the ribosome exit tunnel [9,14]. However, less is known about the behaviour of folded FLN5 NCs, interactions of which would perturb the other side of the folding equilibrium. We have found that such states are most effectively observed using selective [<sup>2</sup>H,<sup>13</sup>CH<sub>3</sub>] methyl labelling schemes (Fig. 1C,D) [9]. While we have not observed changes in chemical shifts between isolated FLN5 and FLN5 RNCs, indicating that the NCs are in their natively folded state, we are interested in understanding how the ribosome perturbs the tumbling of the domain, and how any such changes might be interpreted in terms of ribosome–NC interactions.

Methyl spin systems provide a rich array of transitions that may be interrogated to yield dynamic information (Fig. 2) [16,17]. Relaxation rates of these transitions are highly sensitive to restricted rotational diffusion or transient interactions with large species such as the ribosome surface (i.e. transferred relaxation), but also contain contributions from chemical exchange and interactions with external spins. These complicating effects can be eliminated by analysis of cross-correlated relaxation rates, i.e. spin state-dependent differences in relaxation rates, in order to provide ‘pure’ measurements of dynamics, encoded by the spectral densities  $J(\omega)$ . For methyl groups in slowly tumbling macromolecules, the spectral density function is dominated by the zero frequency term,  $J(0) = S^2\tau_c$ , where  $S^2$  is an order parameter describing the amplitude of motions of the three-fold symmetry axis, and  $\tau_c$  is the (effective) rotational correlation time [16]. This naturally provides a factorisation between the local dynamic properties of a particular methyl group, described by the  $S^2$  order parameter, and the global tumbling of the domain, described by the correlation time  $\tau_c$ . Provided that the structure of a domain does not change between the isolated protein and an NC (which would be evident from chemical shift perturbations), it is reasonable to assume that the  $S^2$  order parameter also will not change, and therefore that changes in  $J(0)$  between NCs and the isolated protein may be interpreted in terms of changes in the effective rotational correlation times.

A variety of experiments have been reported for the measurement of rotational diffusion via cross-correlated relaxation in <sup>13</sup>CH<sub>3</sub> spin systems [18–27]. The majority of these fall into two categories. First, measurements of HH/HH dipole-dipole/dipole-dipole (DD/DD) cross-correlated relaxation during the evolution of a <sup>1</sup>H single quantum coherence, which can be

detected either by direct measurement of the slow and fast  $^1\text{H}$  relaxation rates  $R_{2,H}^S$  and  $R_{2,H}^F$  (Fig. 2) [18], or via the build-up of double quantum [20] or, more sensitively, triple quantum  $^1\text{H}$  coherences [21] relative to the relaxation of the single quantum coherence. As  $^1\text{H}$  transitions are degenerate, multiple experiments must be acquired to isolate the various coherences being observed in these approaches. In a related approach, cross-correlated relaxation may be estimated as a function of  $S^2\tau_c$  through the relative resonance intensities in HSQC and HMQC experiments [27]. Second, measurements of differential relaxation in the  $^{13}\text{C}$  multiplet, which is dominated by CH/CH DD/DD cross-correlated relaxation, with additional contributions from CH/C DD/CSA (chemical shift anisotropy) cross-correlations. As  $^{13}\text{C}$  transitions are not degenerate, this differential relaxation can be read out using a constant-time  $^{13}\text{C}$  evolution period [19], most simply using an F1-coupled constant-time HSQC experiment [22,23] or, as applied to isolated lysine  $^{15}\text{NH}_3$  spin systems, using a standard F1-coupled HSQC [28].

Due to the low concentration of ribosome samples, together with the limited stability of attached nascent chains, optimisation of experimental sensitivity is a critical aspect of RNC NMR, and to this end we have previously explored the use of paramagnetic longitudinal relaxation enhancement agents [29] and non-uniform weighted sampling strategies [10,30]. In this regard, the potential to analyse cross-correlated relaxation rates from the  $^{13}\text{C}$  multiplet structure in a single, sensitive 2D spectrum is highly attractive. In particular, in this work we focus on measurements of isolated  $^{13}\text{CH}_3$  spin systems against a perdeuterated background, using a conventional (non-constant time) F1-coupled HSQC experiment in order to minimize unnecessary relaxation delays and so maximise sensitivity [28]. However, careful analysis of this sequence is required in order to account fully for CCR processes during all stages of the pulse sequence, some of which have been overlooked in previous works [19,22,23,28]. Therefore, here we present a complete expression for the evolution of magnetisation during this experiment. We further show that this can be measured with increased sensitivity by sampling the evolution in the time and phase domain non-uniformly, according to a mathematically optimal adaptive experimental design. Finally, we report applications of this method to characterise dynamics within folded FLN5 RNCs.

## Results

### Analysis of cross-correlated relaxation in methyl spin systems during an F1-coupled HSQC experiment

In this section we derive an expression for the observed signal in an F1-coupled HSQC experiment, accounting fully for cross-correlated relaxation processes throughout the sequence. We will show that this expression depends only on three parameters per methyl group – an overall amplitude, the  $^{13}\text{C}$   $R_2$ , and the  $S^2\tau_c$  – which can be estimated by least squares fitting in either the time or frequency domain. In a subsequent section we will then show how this expression provides a basis for optimal design of non-uniform sampling schemes in order to maximise the precision of the estimated parameters.

The relaxation of methyl spin systems has been treated and reviewed on previous occasions [16,19], but for completeness in this section we give a brief summary of relevant

aspects. Fig. 2 shows a diagram of the energy levels and single-quantum transitions in an isolated  $^{13}\text{C}_3$ -labelled methyl group. In the macromolecular limit and assuming rapid rotation about the three-fold symmetry axis, each transition relaxes in a single-exponential manner. However, the relaxation rates of individual  $^1\text{H}$  or  $^{13}\text{C}$  transitions vary due to cross-correlations between dipole-dipole (DD) interactions or between DD interactions and the  $^{13}\text{C}$  CSA [19].

There are ten  $^1\text{H}$  single quantum transitions present in the spin system (vertical transitions in Fig. 2). Due to cross-correlations between proton dipolar interactions ( $\eta_{HHH}$ ), these can be divided into the slowly relaxing inner lines  $|2\rangle\langle 3|$ ,  $|5\rangle\langle 6|$  and  $|7\rangle\langle 8|$ , which relax with rate  $R_{2,H}^S$ , and the more rapidly relaxing outer lines  $|1\rangle\langle 2|$  and  $|3\rangle\langle 4|$ . Depending on the  $^{13}\text{C}$  spin state, these outer lines are further split by cross-correlation between HH/HC dipolar interactions ( $\eta_{HHHC}$ ) into fast (+) and slow (-) relaxing transitions, which relax with rates  $R_{2,H}^{F+}$  and  $R_{2,H}^{F-}$  respectively (Fig. 2). Similarly, the eight  $^{13}\text{C}$  single quantum transitions (horizontal transitions in Fig. 2) relax with four different rates,  $R_{2,C}^{\alpha\alpha}$ ,  $R_{2,C}^{\alpha\beta}$ ,  $R_{2,C}^{\beta\beta}$  and,  $R_{2,C}^{\beta\beta\beta}$  depending on the spin states of the attached protons. In particular, cross-correlations between CH dipolar interactions ( $\eta_{CHCH}$ ) lead to slow relaxation of the two inner transitions and rapid relaxation of the two outer transitions, while cross-correlation between the CH dipolar interaction and the  $^{13}\text{C}$  CSA ( $\eta_{CHC}$ ) leads to slower relaxation where attached protons are in the  $\beta$  spin state (Fig. 2). Importantly, in the macromolecular limit all of these cross-correlations depend only on the spectral density at zero frequency,  $J(0) = S^2 \tau_C$ , and on the  $^{13}\text{C}$  CSA,  $\sigma_C$ , which depends primarily on the residue type. Values of  $18.2 \pm 1.5$  and  $25.8 \pm 5.6$  ppm were measured in protein L for isoleucine C $\delta$  and leucine/valine methyls respectively [22]. It will be demonstrated below that these mean values provide an acceptable degree of accuracy for our purposes, and therefore using this approximation all necessary CCR rates may be calculated as a function of  $S^2 \tau_C$  alone.

Fig. 3A shows the HSQC pulse sequence used in this work, in which the central  $^1\text{H}$  decoupling pulse is removed so that  $^{13}\text{C}$  multiplet structures can be resolved (Fig. 3B). The evolution of anti-phase  $^{13}\text{C}$  magnetisation during  $t_1$  would be expected to give rise to a 3:1:1:3 multiplet pattern, and deviations from this have previously been used to evaluate cross-correlated relaxation processes in lysine sidechain  $\text{NH}_3$  groups [28] and, using related constant-time evolution experiments, in methyl groups [19,22,23]. However, these studies have not examined the impact of cross-correlated relaxation during the remainder of the experiment, and we will show below that this can distort the expected 3:1:1:3 multiplet structure and lead to systematic errors. As noted in early studies of transverse relaxation in methyl groups via inverse detection [31], in order to describe the amplitudes of lines within  $^{13}\text{C}$  multiplets correctly it is therefore essential to calculate the evolution of magnetisation throughout the entire pulse sequence.

Following the initial  $90^\circ$   $\text{H}_x$  pulse, the density operator

$$\rho_a = -H_y = \frac{i}{2}(\sqrt{3}|1\rangle\langle 2| + 2|2\rangle\langle 3| + \sqrt{3}|3\rangle\langle 4| + |5\rangle\langle 6| + |7\rangle\langle 8|) + h.c. \text{ (where } h.c. \text{ indicates the Hermitian conjugate of the expression) evolves under the scalar coupling Hamiltonian } 2\pi J_{CH} \text{H}_z \text{C}_z \text{ for a total period } \tau = 1/2J_{CH}. \text{ The chemical shift is refocused during this}$$

interval, while relaxation occurs under the influence of cross-correlated HH/HH dipolar interactions and HH/HC dipolar interactions, following Fig. 2. Note that these cross-correlations are not affected by the central  $180^\circ$  ( $H_x + C_x$ ) pulse, and so at point b the density operator is

$$\rho_b = -C_z \left( \sqrt{3} e^{-R_2^F \tau} |1\rangle\langle 2| + 2e^{-R_2^S \tau} |2\rangle\langle 3| + \sqrt{3} e^{-R_2^F \tau} |3\rangle\langle 4| + e^{-R_2^S \tau} |4\rangle\langle 5| + e^{-R_2^S \tau} |5\rangle\langle 6| + e^{-R_2^F \tau} |6\rangle\langle 7| + e^{-R_2^S \tau} |7\rangle\langle 8| \right) + \text{h.c.}$$

Following the  $90^\circ$   $H_y$  pulse, which mixes together terms in the  $I = 3/2$  manifold, and the purge gradient  $g_2$ , the density operator at point c becomes:

$$\rho_c = C_z e^{-R_2^S \tau} \left[ \frac{3}{2}(\Delta + 1) |1\rangle\langle 1| + \frac{1}{2}(3\Delta - 1) |2\rangle\langle 2| - \frac{1}{2}(3\Delta - 1) |3\rangle\langle 3| - \frac{3}{2}(\Delta + 1) |4\rangle\langle 4| + |5\rangle\langle 5| - |6\rangle\langle 6| + |7\rangle\langle 7| - |8\rangle\langle 8| \right] \quad (1)$$

where the amplitude

$$\Delta = e^{-4\eta_{HHHH}\tau} \cosh 2\eta_{HHHC}\tau \quad (2)$$

accounts for the effects of cross-correlated relaxation during the INEPT transfer. Note that  $\Delta$ , like the rates  $\eta_{HHHH}$  and  $\eta_{HHHC}$ , is a function of  $\mathcal{J}(0)$  only, and is close to unity for small  $\mathcal{J}(0)$  (i.e. rapid tumbling), decreasing towards zero as  $\mathcal{J}(0)$  increases.

From point d to e,  $^{13}\text{C}$  coherences evolve under the influence of the  $^{13}\text{C}$  chemical shift and the heteronuclear scalar coupling, with spin-state dependent relaxation rates as indicated in Fig. 2. While  $^1\text{H}$  spin flips induced by interactions with external spins can couple adjacent transitions, because the  $^{13}\text{C}$  chemical shift and heteronuclear scalar coupling are not refocused in this period then provided that the rate of spin flips is much less than the frequency separation between adjacent lines of the multiplet,  $R_{1H,sel} \ll 2\pi J_{CH}$ , these coupling terms are non-secular and hence cross-relaxation is strongly suppressed.

The back-transfer from point f to g following the final  $90^\circ$   $H_x$  pulse can be analysed similarly to the initial INEPT transfer. At the point of acquisition, both fast and slow-relaxing  $^1\text{H}$  coherences are present and will give rise to degenerate resonances in the direct dimension, with amplitudes  $\text{Tr} \rho_g H_{\overline{S}F}$ , where  $H_{\overline{S}} = |3\rangle\langle 2| + \frac{1}{2}(|6\rangle\langle 5| + |8\rangle\langle 7|)$  and  $H_{\overline{F}} = \sqrt{3}(|2\rangle\langle 1| + |4\rangle\langle 3|)$ . Carrying out the calculation in this manner, we find that the observed magnetisation is:

$$\begin{aligned}
y_S(t_1) = & \frac{3}{16} e^{-2R_{2,H}^S t_1} \left( 3[\Delta + 1] e^{-R_{\alpha\alpha\alpha} t_1} \cos[(\omega_C + 3\pi J_{CH})t_1 + \phi] \right. \\
& + [3\Delta - 1] e^{-R_{\alpha\alpha\beta} t_1} \cos[(\omega_C + \pi J_{CH})t_1 + \phi] \\
& + [3\Delta - 1] e^{-R_{\alpha\beta\beta} t_1} \cos[(\omega_C - \pi J_{CH})t_1 + \phi] \\
& \left. + 3[\Delta + 1] e^{-R_{\beta\beta\beta} t_1} \cos[(\omega_C - 3\pi J_{CH})t_1 + \phi] \right)
\end{aligned} \tag{3}$$

$$\begin{aligned}
y_F(t_1) = & \frac{3}{16} e^{-2R_{2,H}^S t_1} \left( 3[\Delta + 1] e^{-R_{\alpha\alpha\alpha} t_1} \cos[(\omega_C + 3\pi J_{CH})t_1 + \phi] \right. \\
& + [3\Delta - 1] e^{-R_{\alpha\alpha\beta} t_1} \cos[(\omega_C + \pi J_{CH})t_1 + \phi] \\
& + [3\Delta - 1] e^{-R_{\alpha\beta\beta} t_1} \cos[(\omega_C - \pi J_{CH})t_1 + \phi] \\
& \left. + 3[\Delta + 1] e^{-R_{\beta\beta\beta} t_1} \cos[(\omega_C - 3\pi J_{CH})t_1 + \phi] \right)
\end{aligned} \tag{4}$$

where  $\phi$  indicates the phase of the final  $^{13}\text{C}$  90° pulse, and the relaxation rates  $R_{\alpha\alpha\alpha}$  etc. are defined in Fig. 2. As all signals are equally weighted by the factor  $e^{-2R_{2,H}^S t_1}$ , this will effectively be absorbed into an overall amplitude for each signal.

Following detection (with  $^{13}\text{C}$  decoupling) and Fourier transformation in  $t_2$ , each component above gives rise to a Lorentzian signal in the  $^1\text{H}$  dimension, with linewidths given by the slow and fast  $^1\text{H}$  relaxation rates respectively:

$$y(t_1, \omega_2) = \sum_{\text{spins}} y_S(t_1) \mathcal{L}(\omega_H, R_{2,H}^S) + y_F(t_1) \mathcal{L}(\omega_H, R_{2,H}^F) \tag{5}$$

The fast-relaxing component in this expression,  $y_F(t_1)$ , is suppressed by a factor relative to  $y_S(t_1)$ , which tends towards zero as  $S^2\tau_c$  increases (Eq. 2). Moreover, as the intensity of a cross-peak is inversely proportional to its linewidth, and given that  $R_{2,H}^F \gg R_{2,H}^S$ , to an excellent approximation the fast relaxing terms may be neglected in the above expression.

Examining Eq. 3 reveals the impact of both HH/HH and HH/HC cross-correlated relaxation during the initial INEPT transfer on the initial amplitudes of inner and outer lines in the  $^{13}\text{C}$  quartet. This builds on earlier analyses of the fully decoupled HSQC experiment, which derived the impact of HH/HH cross-correlations on the observed signal [17], as well as on the analysis of HH/HC and HH/HH cross-correlation relaxation for the measurement of side chain order [25]. For rapid tumbling,  $\sigma$  is close to one (Eq. 2) and so the multiplet exhibits the canonical 3:1:1:3 ratio expected from anti-phase magnetisation at point d. However, as  $J(0)$  increases  $\sigma$  tends towards zero, leading towards a ratio of 1:1:1:1 in the initial multiplet amplitudes (Fig. 3C). Indeed, this plot illustrates that the assumption of a 3:1:1:3 ratio is valid only for  $S^2\tau_c$  values below ca. 1 ns, and so even for methyls within small domains it is important to employ the full analysis given here to avoid systematic errors.

Eq. 3 may be used to fit data directly in the  $t_1$  time domain, by integration around the  $^1\text{H}$  chemical shift of an isolated resonance. Alternatively, the Fourier transform in  $t_1$  of the

expression is simply a sum of four Lorentzian signals corresponding to the four multiplet components, with amplitudes and linewidths as indicated (plus additional contributions depending on the particular window function applied), and this may be used to fit cross-sections from spectra in the frequency domain.

In Fig. 3D, we have estimated the relaxation rates of multiplet lines in the absence of interactions with other spins and additional effects such as exchange broadening. Using these rates, Fig. 3E shows the relative intensities of lines within an HSQC spectrum, processed with 20 Hz exponential line broadening, as a function of  $S^2\tau_c$ . Examination of this plot indicates that for small to moderate values of  $S^2\tau_c$  (up to ca. 50 ns), the greatest information content is contained in the relative intensities of inner vs outer lines, as these vary substantially over this range. However, for larger  $S^2\tau_c$  the intensities of the outer lines are reduced, and so the ratio of the  $\alpha\alpha\beta$  and  $\alpha\beta\beta$  lines, which depends on the  $^{13}\text{C}$  CSA, will become more informative. This indicates that estimates of  $S^2\tau_c$  for larger molecules may be more sensitive to site-specific variations in the CSA. Indeed, based on this principle a variant of an F1-coupled CT-HSQC was used previously to determine the average CSA values of ILV methyls employed in this work [22]. Alternatively, the product of intensities of upper and lower transitions in a CT-HSQC experiment has been shown to cancel out the dependence on the CSA entirely [23]. For completeness, we present below an expression for this quantity incorporating the corrections for cross-correlations determined earlier (where  $T$  is the length of the constant time period):

$$\frac{I_{\alpha\alpha\alpha}I_{\beta\beta\beta}}{I_{\alpha\alpha\beta}I_{\alpha\beta\beta}} = \frac{9(\Delta + 1)^2}{(3 - \Delta)^2} \exp(-\eta_{CHCH}T) \quad (6)$$

A final point remains to be considered regarding the effects of relaxation due to the interaction of spins with external protons, and the potential for the introduction of systematic errors. Relaxation during the evolution of  $^1\text{H}$  coherences (from point a to b and f to g in Fig. 3A) can be described in a Liouville subspace (for simplicity here showing only the  $|\alpha\rangle$   $^{13}\text{C}$  spin state):

$$\frac{d}{dt} \begin{pmatrix} |1\rangle\langle 2| \\ |2\rangle\langle 3| \\ |3\rangle\langle 4| \\ |5\rangle\langle 6| \\ |7\rangle\langle 8| \end{pmatrix} = - \begin{pmatrix} R_{2,H}^F + 9k_{HH} & -2\sqrt{3}k_{HH} & 0 & 0 & 0 \\ -2\sqrt{3}k_{HH} & R_{2,H}^S + 11k_{HH} & -2\sqrt{3}k_{HH} & 0 & 0 \\ 0 & -2\sqrt{3}k_{HH} & R_{2,H}^F + 9k_{HH} & 0 & 0 \\ 0 & 0 & 0 & R_{2,H}^S + \frac{5}{2}k_{HH} & 0 \\ 0 & 0 & 0 & 0 & R_{2,H}^S + \frac{5}{2}k_{HH} \end{pmatrix} \begin{pmatrix} |1\rangle\langle 2| \\ |2\rangle\langle 3| \\ |3\rangle\langle 4| \\ |5\rangle\langle 6| \\ |7\rangle\langle 8| \end{pmatrix} \quad (7)$$

where  $k_{HH} = \frac{1}{20} \left( \frac{\mu_0}{4\pi} \right)^2 \hbar h^2 \gamma_H^4 \left( \sum \frac{1}{r_{H,Ext}^6} \right) \tau_C$  and similar relations apply for the  $|\beta\rangle$   $^{13}\text{C}$  spin state and Hermitian conjugates [19]. In part, the additional diagonal terms can be absorbed into the definition of  $R_{2,H}^S$ , which in turn will affect only the overall amplitude of the signal. The significance of the remaining diagonal and off-diagonal terms depends primarily on the relative magnitudes of  $k_{HH}$  and  $R_{2,H}^F$ .  $R_{2,H}^F - R_{2,H}^S = 2\eta_{HHHH}$ . The sum over interproton distances may be approximated by the weighted average  $\langle r_{H,Ext} \rangle = \left( \sum \frac{1}{r_{H,Ext}^6} \right)^{-1/6}$ , and for [ $^2\text{H}$ ,  $^{13}\text{CH}_3$ -Ile] labelling in FLN5 this ranges between 4.5 and 12.3 Å depending on the methyl group. On the basis of this, the relative magnitude of external proton interactions are estimated to be an order of magnitude weaker than cross-correlated relaxation processes. However, for non-deuterated proteins, in which the weighted average ( $r_{H,Ext}$ ) is considerably smaller, the effects of external spins are likely to be larger and potentially non-negligible.

Interactions with external spins also introduce an additional perturbation to the evolution of the two inner lines during  $t_1$ , beyond the non-secular cross-relaxation terms already discussed. The proton density matrix elements relevant to  $aa\beta$  the line in  $\rho_C$  (Eq. 1) can be re-written as  $\rho_C \supset \frac{1}{2}(1 + \Delta)(|2\rangle\langle 2| + |5\rangle\langle 5| + |7\rangle\langle 7|) - \frac{1}{2}(1 - \Delta)(|2\rangle\langle 2| - |5\rangle\langle 5| - |7\rangle\langle 7|)$  (and a similar analysis follows for the  $aa\beta$  line). When  $\mathcal{K}(0)$  is small and little cross-correlated relaxation takes place during the initial INEPT transfer ( $\approx 1$ ), only the first term is present. This can be expanded as  $|2\rangle\langle 2| + |5\rangle\langle 5| + |7\rangle\langle 7| = |a\alpha\beta\rangle\langle a\alpha\beta| + |a\beta a\rangle\langle a\beta a| +$



$|\beta\alpha\alpha\rangle\langle\beta\alpha\alpha|$ , i.e. a simple population term, which relaxes with external protons with rate  $3k_{HH}$  [19]. However, when  $J(0)$  is larger and cross-correlated relaxation during the INEPT transfer cannot be neglected, the second term may become comparable to the first. This can be expanded as  $2|2\rangle\langle 2| - |5\rangle\langle 5| - |7\rangle\langle 7| = |a\alpha\beta\rangle\langle a\beta a| + |a\alpha\beta\rangle\langle\beta a a| + |a\alpha\rangle\langle a\alpha| + |a\beta\alpha\rangle\langle\beta a a| + |\beta\beta\alpha\rangle\langle a\alpha\beta| + |a\beta\alpha\rangle\langle a\beta a|$ , i.e. a symmetrised mixture of zero quantum  $^1\text{H}$  transitions, which relaxes more rapidly with external protons, with rate  $11k_{HH}$  [19]. The net effect of this is that the two inner transitions will be slightly broader than otherwise expected, leading to an underestimation of  $S^2\tau_c$ . As above, we estimate that the impact of this effect will be small for isolated  $^{13}\text{CH}_3$  groups in perdeuterated proteins, but may be more significant if applied to the analysis of non-deuterated molecules. Any application to such molecules should therefore be carried out with attention to these systematic effects.

## Experimental validation

Prior to application to RNCs, we carried out a series of measurements of methyl dynamics using isolated [ $^2\text{H}$ ,  $^{13}\text{CH}_3$ -Ile]-labelled FLN5 (Fig. 1B) to validate the accuracy of our analysis, against reference measurements acquired using the triple quantum build-up scheme [21]. Measurements were performed in a series of d8-glycerol concentrations at 298 K and 283 K, across which the viscosity is predicted to vary by an order of magnitude [32,33], in order to decrease the rate of tumbling and so probe rotational diffusion across a wide range of timescales. Fig. 4A–D shows coupled HSQC spectra acquired under these conditions, together with HMQC reference spectra for the determination of peak positions.

The spectrum of FLN5 in the absence of glycerol (298 K, Fig. 4A) shows strong outer lines, although it can be observed that these are broader than the less intense inner lines. As the glycerol concentration is increased, these outer lines become broader and less intense, and the asymmetry under exchange  $\alpha \leftrightarrow \beta$  introduced by cross-correlation with the  $^{13}\text{C}$  CSA starts to become discernible (Fig. 4B,C). At the highest glycerol concentration and lowest temperature acquired (Fig. 4D), for several residues the outer lines are barely observable, but the asymmetry between the inner  $a\alpha\beta$  and  $a\beta\beta$  lines nevertheless may be expected to provide a continued route to the quantification of rotational diffusion.

The extracted cross-sections were fitted to the Fourier transform of Eq. 3 and are plotted in Fig. 4A–D. Equivalent results could also be obtained by direct fitting of time-domain data (Fig. 4E–F). The  $^{13}\text{C}$  chemical shift was determined from the peak position in the accompanying HMQC spectrum, the  $^{13}\text{C}$  CSA was fixed to 18.2 ppm as previously reported for isoleucine  $C\delta$  spins [22], and the heteronuclear scalar coupling  $J_{HC}$  was fixed to 125 Hz. Therefore, each cross-section could be fitted as a function of only three parameters: an amplitude factor, the  $^{13}\text{C}$   $R_2$  relaxation rate, and  $S^2\tau_c$ , from which all necessary crosscorrelated relaxation rates could be calculated.  $S^2\tau_c$  values were also determined by measurement of the build-up of triple quantum magnetisation relative to the relaxation of single quantum magnetisation (Fig. 4G) [21]. A comparison of the two measurements shows excellent agreement ( $R^2 = 0.97$ ) across two orders of magnitude in  $S^2\tau_c$  (Fig. 4H), thus validating further applications of the new approach. Similar measurements of [ $^2\text{H}$ ,  $^{13}\text{CH}_3$ -ILV]-labelled FLN5 were also in good agreement ( $R^2 = 0.98$ , Fig. S1) indicating that less sparse labelling schemes are also amenable to analysis by this method.

## Optimal design of experimental measurements

In the previous sections, we have established and validated the analysis of cross-correlated relaxation in methyl groups by measurement of the differential relaxation of transitions in the  $^{13}\text{C}$  multiplet, and we have also shown that this analysis may be carried out in both the frequency (Fig. 4A–D) and time domains (Fig. 4E–F). From this point onwards, we will consider the analysis of time-domain data only, obtained following integration of nonoverlapping resonances in the  $^1\text{H}$  frequency dimension. As such data will be fitted directly, without Fourier transformation, there is no longer a requirement either to sample uniformly in the time domain or to acquire cosine- and sine-modulated pairs of points for quadrature detection. Instead, measurements may be acquired with arbitrary  $t_1$  evolution times and phases  $\phi$ , and so it is natural to ask whether a non-uniform sampling strategy might exist that provides greater sensitivity than conventional uniform sampling. Such a problem can be solved using the theory of optimal experimental design [34].

As we are in general considering the analysis of multiple resonances, we write our earlier expression for the time-evolution of observed signals (Eq. 3) as the vector-valued function  $y(\mathbf{x}; \boldsymbol{\theta})$ , termed the ‘response surface’, which describes the observed intensities of  $k$  spins as a function of the independent variables  $x = (t_1, \phi)$  and the  $m = 3k$  unknown parameters  $\boldsymbol{\theta} = (A^{(1)}, R_{2,C}^{(1)}, S_{\tau_c}^{2(1)}, \dots, S_{\tau_c}^{2(k)})$ . Therefore, our goal is to determine the optimal set of such experiments to perform, within the allowed space of experiments,  $\mathcal{X}$ , in order to estimate the parameters  $\boldsymbol{\theta}$  with minimal uncertainty.

We begin by defining an  $N$ -point experimental design  $\xi_N$  as a distribution of  $N$  measurements across  $n$  points  $x_j \in \mathcal{X}$ , with  $r_j$  repeated measurements at each support point  $i$  ( $\sum_i r_i = N$ ):

$$\xi_N = \begin{cases} x_1 & \cdots & x_n \\ r_1 & \cdots & r_n \end{cases} \quad (7)$$

The number of support points,  $N$ , required for an optimal design depends on the particular system under consideration. Clearly, the estimation of  $m = 3k$  parameters from a  $k$ -valued vector observation requires at least  $m/k = 3$  distinct measurements. It may also be proved as a consequence of Caratheodory’s theorem [34] that for single-valued observables ( $k = 1$ ) there exists an optimal design with at most  $\frac{1}{2}m(m+1) + 1 = 7$  support points. By extension, the analysis of  $k$  independent signals can therefore require no more than  $7k$  support points, corresponding to a measurement split between  $k$  sub-designs that are optimal for each particular signal. In reality however, most optimal designs require many fewer support points than the worst-case limit provided by Caratheodory’s theorem [35]. The number of support points required for vector-valued observations is also usually smaller, as measurements typically provide at least some useful information for all observed signals, thus reducing the total number of support points necessary.

Having made a set of observations  $y_i$  according to the experiment  $\xi_N$ , where the (uncorrelated) measurements have covariance matrix  $\sigma^2 \mathbf{I}_k$ , the data can be fitted to the

response surface in a least-squares sense to yield the maximum likelihood estimator of the parameters,  $\hat{\theta}$ . The Fisher information matrix

$$M_{\alpha\beta}(\xi, \hat{\theta}) = \sigma^{-2} \sum_{i=1}^n r_i \frac{\partial y^T(x_i, \hat{\theta})}{\partial \theta_\alpha} \cdot \frac{\partial y(x_i, \hat{\theta})}{\partial \theta_\beta} \quad (8)$$

describes the curvature of the log-likelihood function about  $\hat{\theta}$ , from which the parameter covariance matrix  $D = M^{-1}(\xi, \hat{\theta})$  may be defined. Note that the inverse of the Fisher information matrix is the Cramér-Rao lower bound for the variance of an unbiased estimator (understood as a comparison of positive definite matrices) and so indicates the maximum achievable precision for a given experimental design. This is also independent of the noise level, which simply scales the overall uncertainty.

To determine the *prediction* error for the response surface based on this estimate (required below), we consider a linearisation of the response surface about  $\hat{\theta}$ :  $y(x, \theta) \approx y(x, \hat{\theta}) + (\theta - \hat{\theta})^T F(x, \hat{\theta})$ , where  $F_{\alpha\beta} = \partial y / \partial \theta_\beta$  is the Jacobian of the response surface. As the errors in the parameter estimates are normally distributed with covariance  $D$ , from the transformation properties of multivariate normal distributions the predicted value of the response surface at point  $x$  is also normally distributed,  $\hat{y}(x, \hat{\theta}) \sim \mathcal{N}(y(x, \hat{\theta}), F^T D F)$ . We thus define the prediction uncertainty  $d(x, \xi, \hat{\theta}) = F^T(x, \hat{\theta}) M^{-1}(\xi, \hat{\theta}) F(x, \hat{\theta})$ .

To proceed further, we must now make our idea of the ‘optimality’ of an experiment mathematically precise. An intuitive and useful definition is to find the design  $\xi^*$  which minimises the volume of the parameter confidence ellipsoid,  $\det D$ :

$$\xi^* = \underset{\xi}{\operatorname{argmin}} \det D(\xi, \hat{\theta}) \quad (9)$$

This strategy is termed *D-optimal* design and is particularly attractive as it is invariant to linear changes of parameterisation and units. It can be proved (the General Equivalence Theorem [36]) that this problem is also equivalent to minimisation of the maximum prediction error (termed *G-optimality*):

$$\xi^* = \underset{\xi}{\operatorname{argmin}} \max_{x \in \mathcal{X}} \operatorname{Tr} d(x, \xi, \hat{\theta}) \quad (10)$$

This provides an intuitive design strategy: acquire measurements at points where the prediction uncertainty is greatest, and this in turn will minimise the uncertainty in the parameter estimates.

However, there is a difficulty in this approach that results from the dependence of  $M(\xi, \hat{\theta})$  on the unknown or uncertain parameter estimate  $\hat{\theta}$ , which introduces a circularity into the design. This can be resolved using an *adaptive* design (also referred to as a sequential design [34]) in which, following a set of seed measurements  $\xi_N$ , an updated parameter estimate is calculated and used to select the next measurement point:

$$\begin{aligned}
x_{N+1} &= \arg \max_{x \in \mathcal{X}} \text{Tr } d(x, \xi_N, \hat{\theta}_N) \\
\hat{\theta}_{N+1} &= \arg \min_{\theta \in \Theta} \sum_{i=1}^{N+1} \|y_i - y(x_i, \theta)\|_2 \\
M(\xi_{N+1}, \hat{\theta}_{N+1}) &= M(\xi_N, \hat{\theta}_{N+1}) + \sigma^{-2} F(x_{N+1}, \hat{\theta}_{N+1}) F^T(x_{N+1}, \hat{\theta}_{N+1})
\end{aligned} \tag{11}$$

This process is illustrated as a flowchart in Fig. 5A.

We initially tested the adaptive experimental design strategy using simulations to generate synthetic data. One such example is illustrated in Fig. 4B–D, using the uniformly sampled measurements of FLN5 relaxation rates obtained for I738, I674 and I695 in 50% d8-glycerol, 298 K (Fig. 3C) as the unknown ‘true’ values used to generate simulated samples. Six seed points were used to determine initial parameter estimates, with evolution times of 0.1, 1, 3, 10, 30 and 100 ms, and zero phase, following which 200 further samples were acquired according to the adaptive sampling scheme outlined above (Fig. 5A, B). The prediction error  $d(x, \xi_N, \hat{\theta}_N)$  was found to be a very rough surface, and so to ensure the global maximum was located a grid search was employed across  $t_1$  values from 0 to 100 ms and phases  $\phi \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ . For comparison, a uniformly sampled experiment was also simulated, with an identical number of total samples and a maximum  $t_1$  acquisition time of 100 ms (Fig. 5B). Measurements using the adaptive sampling scheme were focused at just six support points, which is towards the lower end of the theoretical range of 3 to  $7k = 21$  discussed earlier. We note that an advantage of the adaptive sampling approach is that the number of support points does not need to be predetermined but emerges naturally through the gradual accumulation of single optimised measurements.

The final set of simulated measurements and the fitted evolution curves are shown in Fig. 5C, while the evolution of the parameter estimates is plotted against the increasing number of measurements in Fig. 5D. From this, the progressive reduction in the parameter uncertainty may be observed, as well as the convergence of parameter estimates towards their ‘true’ values. For comparison, estimates determined by the equivalent uniformly sampled measurement are also plotted and may be seen to have greatly reduced precision.

To explore this in more detail, 500 independent simulations of the adaptive sampling scheme were carried out in order to obtain reliable estimates of the parameter uncertainty. A series of uniformly sampled measurements were then also simulated with acquisition times ranging from 10 to 200 ms (and a fixed total number of measurement points). The relative uncertainty in  $S^2\tau_c$  estimates is plotted for each residue as a function of the acquisition time in Fig. 5E. It may be observed from this that the performance of uniformly sampled measurements never exceeds that of the adaptive sampling scheme. Where the optimal acquisition time is employed, the gains in precision are modest, but in general this optimal time may be different for different residues – and is of course unlikely to be known prior to carrying out the measurement. Thus, the application of adaptive sampling schemes is likely to provide a modest but useful increase in experimental precision for a fixed total measurement time.

## Measurement of methyl dynamics in a FLN5+67 ribosome–nascent chain complex

Having established and validated the new experimental analysis and adaptive sampling scheme, we applied it to characterise the rotational diffusion of the folded FLN5 domain within the FLN5+67 RNC, relative to the isolated domain (Fig. 1A). In the case of FLN5 RNCs, four isoleucine resonances can be observed and distinguished from ribosome background signals in 2D NMR experiments, and three can still be resolved unambiguously in 1D spectra (Fig. 1C,D). While this may not be a large number of resonances, because folded domains tumble as rigid bodies even a small number of observations are sufficient to provide useful insight into their motion (in contrast to disordered states, the mobility of which may vary substantially along the sequence).

Adaptively sampled measurements of I738, I674 and I695 resonances within an FLN5+67 RNC sample were acquired using six seed points to provide initial parameter estimates, with  $t_1$  evolution times of 0.02, 1, 2, 5, 10 and 100 ms. Subsequent measurements were then interleaved with control measurements of translational diffusion, and together with western blotting analysis of a sample incubated in parallel with NMR measurements these were used to infer a sample lifetime of ca. 40 hours, corresponding to 120 adaptively-sampled measurements (Fig. 6A–C). Examination of estimated  $S^2\tau_c$  values as a function of the number of measurements indicates that the measured values do not depend strongly on the choice of this cut-off (Fig. 6D). In common with the simulated example in Fig. 5B–D, sampling was focused on just six support points, towards the lower end of the theoretical range of 3–21. Although the specific choice of points was not the same, there are nevertheless similarities in the overall distribution: beyond a first point at zero time, three points are clustered at short times (5–10 ms) followed by two at longer times (ca. 50 ms). We suggest this sampling pattern likely reflects the requirement to capture the rapid decay of outer lines together with the slower decay of the inner lines, while also distinguishing between the relaxation of upper and lower lines due to the  $^{13}\text{C}$  chemical shift anisotropy.

The  $S^2\tau_c$  values determined from this measurement were  $37.7 \pm 3.6$  ns (I738),  $29.1 \pm 3.0$  ns (I674) and  $27.3 \pm 2.6$  ns (I695). By comparison,  $S^2\tau_c$  values of  $6.2 \pm 0.2$  ns (I738),  $5.8 \pm 0.2$  ns (I674) and  $4.5 \pm 0.1$  ns (I695) were determined for the isolated domain under equivalent conditions. The values measured for the NC are instead approximately intermediate between measurements of isolated FLN5 in 40% and 50% glycerol. As outlined in the Introduction, assuming that  $S^2$  order parameters are the same within the NC and the isolated FLN5 domain – which is reasonable given that no chemical shift perturbations are observed, indicating that the NC has a similar folded structure to the isolated domain – then these differences may be interpreted in terms of relative changes in the rotational correlation time alone. On this basis, a regression analysis indicates that rotational diffusion of the NC is approximately 5.2 times slower than the isolated domain (Fig. 6E). While the correlation between RNC and isolated measurements is not perfect, given the experimental uncertainties that we have determined together with the small number of methyl groups observed, at this stage there does not appear to be strong evidence of site-specific effects that would indicate changes in the shape of the diffusion tensor as well as its magnitude.

## Discussion

In this manuscript, we have presented a careful analysis of a previously described approach for the analysis of rotational diffusion and side chain order in isolated methyl spin systems based on the fitting of multiplet lineshapes or intensities in coupled HSQC or constant-time HSQC experiments [19,22,23,28]. This has revealed the importance of analysing cross-correlated relaxation throughout the entirety of the sequence, which can result in initial multiplet amplitudes that deviate from a simple 3:1:1:3 ratio for slowly tumbling systems (Fig. 3C). By incorporating the effect of cross-correlations throughout the sequence (Eq. 3), we have shown that values of  $S^2\tau_c$  may be determined accurately over a wide range of correlation times, from 1 ns to over 100 ns (Fig. 4H, S1). We note that the expression we derive (Eq. 3) is equally applicable to the analysis of dynamics of  $^{15}\text{NH}_3$  groups in lysine side chains [28].

We have next developed and implemented an on-the-fly adaptive sampling scheme for this measurement based on the principles of optimal experimental design [34]. The optimal design of NMR experiments has been discussed in many previous works, generally in terms of the Cramér-Rao lower bound (the inverse of the Fisher information matrix  $\mathbf{M}(\xi)$ ), although the precise criteria employed to optimise this matrix quantity are not always clearly stated. An early application to protein NMR spectroscopy was the optimal selection of time points for relaxation measurements (in fact encouraged and supported by Chris Dobson) [37], while applications of the CRLB to parametric estimation of multidimensional NMR spectra, and the evaluation of non-uniform sampling schemes, was reported a few years later [38]. More recently, this approach has been applied to droplet size estimation using diffusion NMR experiments [39], and to the optimal parametric estimation of protein relaxation rates in a non-uniformly sampled accordion NMR experiment [40]. However, to the best of our knowledge the only previous report of an adaptive approach to design and acquire new NMR experiments ‘on-the-fly’ is ADAPT-NMR, used for automated protein resonance assignment [41]. The combination of the two approaches that we present here has demonstrated the increased precision or reduced acquisition time that can be provided by these methods (Fig. 5E), and points towards potential to applications of adaptive optimal designs to other areas of NMR such as relaxation measurements or non-uniform sampling.

We have applied the methods developed to the analysis of rotational diffusion within a ribosome–nascent chain complex. As noted in the introduction, understanding the dynamic behaviour of NCs as they emerge from the ribosome during biosynthesis is a key challenge in order to elucidate the mechanisms by which the ribosome may modulate or regulate the co-translational folding process [4]. The slower tumbling of the FLN5+67 NC that we have observed may reflect contributions from a number of factors. Firstly, the presence of the FLN6 linker may reduce the mobility of the attached domain. We have previously noted that amide resonances of this linker are strongly broadened in RNC spectra across many lengths [9], which may indicate the existence of interactions between the linker and the ribosome surface, or a conformational exchange process such as transient or partial folding. Such processes might in turn be expected to hinder the rotational diffusion of the attached domain. Moreover, if the FLN5 domain is unable to sample all orientations equally due to restrictions from the linker, then the  $J(0)$  spectral density will have an additional contribution

reflecting this restricted tumbling, proportional to the ca.  $2.5 \mu\text{s}$  rotational correlation time of the parent ribosome [42].

Lastly, the slower tumbling we observe may indicate transient interactions between FLN5 and the ribosome surface. Assuming rapid exchange, the observed rotational correlation times will be a population weighted average of those in the free and bound states, and from this we may infer that the maximum population of any bound state is approximately 1.5% (and additional contributions to the slower tumbling as discussed above would limit this population further). In this instance, such a weak interaction with the largely negatively-charged ribosome surface [43] is perhaps understandable given that the total charge on the FLN5 domain is  $-9$ . Indeed, using both NMR and fluorescence methods, ribosome-NC interactions have previously been identified to be driven, at least in part, by electrostatic interactions [7,10]. Nevertheless, a detailed understanding of such weak interactions, as well as other dynamic effects, may be expected to yield fundamental insights into the behaviour of folded NCs as they emerge from the ribosome, and the optimised experimental approaches that we have developed here will facilitate a new generation of sensitive and precise studies of these phenomena.

## In memory of Professor Sir Chris M. Dobson



Chris

Dobson (1949–2019) was a mentor to us both (CAW and JC) and we dedicate this article to him. Our period in his lab shaped our enthusiasm for NMR and our approaches to it. Chris had a unique versatility in his NMR exploration, and as such became an ambassador to the field. He is seen pictured here delivering the ICMRBS opening plenary in Florence in 2000, which is incidentally where I (JC) first met him. Impressed by his array of paradigm forming examples of protein folding and misfolding, I decided to join his lab. Chris' work was pioneering in an era where disorder and heterogeneity in biomolecules were shunned as peripheral and his early descriptions of the importance of such states in biology helped pave the way for what is now a flourishing field of study. It was clear to many that despite his enthusiasm for a wide variety of biophysical and computational methods to describe molecular mechanisms and landscapes, NMR was his 'go to' technique, one where he revelled in the 'truth' of residue-specific descriptions. Well over 200 of his manuscripts



(of >850 publications total with a h-index of 159) have *NMR* in their title. Chris was particularly excited by the possibility of looking at folding of nascent polypeptides on the ribosome. CryoEM and high-resolution X-ray crystallographic structures of ribosomes were beginning to emerge, but Chris was convinced that the invisible nascent chain in these structures could be visible to NMR. Achieving this required just ‘one or two technical issues’ to be overcome: developing the preparative biochemistry to enable stable selectively-labelled nascent polypeptides on their parent ribosomes to be produced *and* overcoming NMR sensitivity considerations associated with several factors, not least of which was the 2.4 MDa molecular weight of the complex together with the limited solution-state stability for NMR acquisition on 1–5  $\mu\text{M}$  samples! Together with some outstanding colleagues (in particular Profs Paola Fucini and Shang-Te Danny Hsu) we were able to show that this was not only possible, but that ultimately, we could use NMR to define a structural ensemble of an *in vivo* derived biosynthesis snapshot of a nascent chain undergoing its folding transition on the ribosome. Today, the work that we started in Chris’ lab, with his (and Carol Robinson’s) unerring support and instinctive understanding, continues to progress towards ever more detailed descriptions of the structure and dynamics of *de novo* folding processes. Going forward, it requires dedicated NMR advocates and exponents, working in a multidisciplinary team akin to Chris’ own group, to tease out the secrets of the nascent chain and honour Chris’ original vision.

JC was a Wellcome Trust Travelling Research fellow and CAW was a graduate student in Chris’ group.

## Materials and Methods

### Sample preparation

Isolated [ $^2\text{H}$ ,  $^{13}\text{CH}_3$ -Ile]-labelled FLN5 was expressed and purified using standard methods, as previously described [9]. The FLN5+67 RNC construct (Fig. 1A) has also been previously described, and was expressed using minor adaptations of previously reported methods [9,44]. *E. coli* BL21 (DE3) cells were progressively adapted to EM9 medium containing 100%  $\text{D}_2\text{O}$  and  $\text{d}_7$ -glucose, and then grown to an  $\text{OD}_{600}$  of 3. Selective isoleucine methyl labelling of the nascent chain was achieved by addition of 80 mg/L of the isoleucine precursor 2-ketobutyric-4- $^{13}\text{C}$ ,3,3- $\text{d}_2$  acid 30 minutes before induction with 1mM IPTG for 50 min at 310 K. RNCs were purified from harvested cells following our previously reported protocols, except for the substitution of a second sucrose cushion in place of the final sucrose gradient purification step [9]. Finally, RNCs were flash frozen and stored at  $-80^\circ\text{C}$  prior to NMR measurement.

### NMR measurements of isolated FLN5

Three samples of [ $^2\text{H}$ ,  $^{13}\text{CH}_3$ -Ile]-labelled FLN5 (ca. 100  $\mu\text{M}$ ) were prepared in deuterated Tico buffer (100%  $\text{D}_2\text{O}$  and  $\text{d}_8$ -HEPES, pH 7.5) containing 0, 40 and 52% (w/w)  $\text{d}_8$ -glycerol. NMR experiments were acquired at using a Bruker Avance III HD spectrometer operating at a  $^1\text{H}$  Larmor frequency of 700 MHz.  $^1\text{H}$ ,  $^{13}\text{C}$  HMQC and (uniformly sampled) F1-coupled  $^1\text{H}$ ,  $^{13}\text{C}$  HSQC experiments (Fig. 3A) were acquired with acquisition times of 107 ms in the  $^1\text{H}$  dimension and spectral widths of 20 ppm and 2.4 ppm in the  $^1\text{H}$  and  $^{13}\text{C}$

dimensions respectively. 384, 256, 256 and 64 complex points were acquired for 0%, 40%, 52% and 52% (283 K) samples respectively. Measurements of triple quantum build-up and single quantum relaxation were acquired for each sample and fitted as previously described [21]. All experiments were processed using nmrPipe [45] and analysed using MATLAB.

### Implementation of adaptive sampling

Adaptive sampling methods were implemented for Bruker spectrometers operating Topspin 3.5pl6. While a Python interface is provided within Topspin, this runs using the Java implementation, Jython, which does not support the required numerical libraries. Therefore, a separate instance of Python 3 is required, together with installations of NumPy, SciPy and Numba [46–48]. Simple routines have been developed to interface between Topspin and Python, using json formatted files to exchange data. All necessary code, scripts and pulse sequences are available online (<https://github.com/chriswaudby/adaptive-sampling>).

### NMR measurement of the FLN5+67 ribosome–nascent chain complex

Immediately prior to NMR measurement, RNC samples were buffer exchanged into deuterated Tico buffer (100% D<sub>2</sub>O and d<sub>8</sub>-HEPES, pH 7.5) supplemented with protease inhibitors (Roche), with a final sample concentration of 7.5 μM. Some of the sample was reserved in order to monitor the integrity of the NC over the course of the NMR measurements. This control sample was incubated at 298 K, and aliquots containing 2 pmol RNC were periodically flash frozen until NMR data collection was complete. These samples were analysed by anti-His western blotting, and the intensity of the band corresponding to tRNA-bound NC was determined relative to isolated standards using densitometry (with ImageJ software, <http://imagej.nih.gov/ij/>).

RNC NMR experiments were performed at 298 K using a Bruker Avance III HD spectrometer running Topspin 3.5pl6 and operating at a <sup>1</sup>H Larmor frequency of 950 MHz. F1-<sup>1</sup>H-coupled HSQC measurements were acquired with adaptive sampling as described above, with 2048 complex points in the direct dimension (107 ms acquisition time), 512 scans per measurement, and an inter-scan delay of 1.5 s (total measurement time 15 min). RNC integrity was monitored during data collection using <sup>13</sup>C-filtered <sup>1</sup>H STE diffusion measurements [49], which were interleaved between every six HSQC measurements, using a diffusion delay of 50 ms, bipolar smoothed-square gradient pulses with total length 4 ms and strengths of 0.024 and 0.457 T m<sup>-1</sup>, 512 scans and an inter-scan delay of 1.5 s (total measurement time ca. 30 min). The combined integrals of the isolated RNC methyls I738, I674 and I695 were analysed to provide a specific probe of the attachment of the NC.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgements

We acknowledge the use of the UCL Biomolecular NMR Centre and the staff for their support. This work was supported by the Francis Crick Institute through provision of access to the MRC Biomedical NMR Centre. The Francis Crick Institute receives its core funding from Cancer Research UK [FC001029], the UK Medical Research Council [FC001029], and the Wellcome Trust [FC001029]. This research was funded in whole, or in part, by

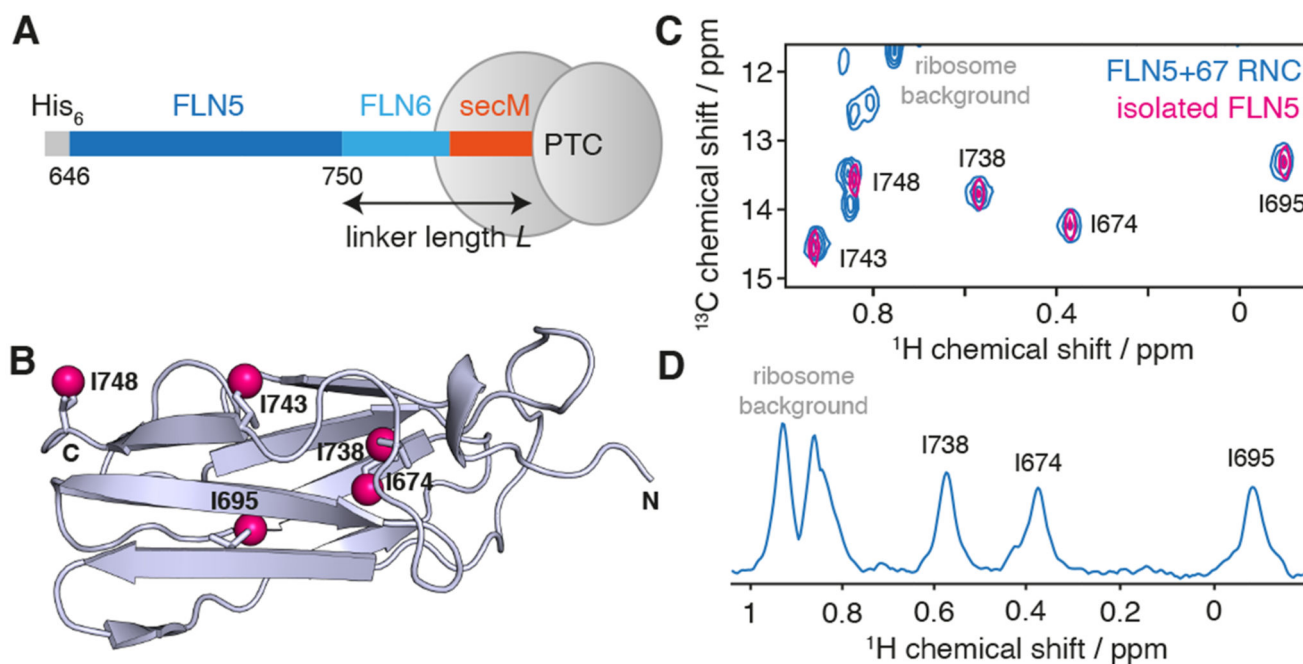
the Wellcome Trust [206409/Z/17/Z] and the BBSRC [BB/T002603/1]. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. This study made use of NMRbox: National Center for Biomolecular NMR Data Processing and Analysis, a Biomedical Technology Research Resource (BTRR), which is supported by NIH grant P41GM111135 (NIGMS) [50].

## References

- [1]. Chiti F, Dobson CM. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu Rev Biochem.* 2017; 86: 27–68. [PubMed: 28498720]
- [2]. Balchin D, Hayer-Hartl M, Hartl FU. In vivo aspects of protein folding and quality control. *Science.* 2016; 353 aac4354 [PubMed: 27365453]
- [3]. Cassaignau AME, Cabrita LD, Christodoulou J. How Does the Ribosome Fold the Proteome? *Annu Rev Biochem.* 2020; 89: 389–415. [PubMed: 32569518]
- [4]. Waudby CA, Dobson CM, Christodoulou J. Nature and Regulation of Protein Folding on the Ribosome. *Trends Biochem Sci.* 2019; 44: 914–926. [PubMed: 31301980]
- [5]. Kaiser CM, Goldman DH, Chodera JD, Tinoco I, Bustamante C. The Ribosome Modulates Nascent Protein Folding. *Science.* 2011; 334: 1723–1727. [PubMed: 22194581]
- [6]. Samelson AJ, Jensen MK, Soto RA, Cate JHD, Marqusee S. Quantitative determination of ribosome nascent chain stability. *Proc Natl Acad Sci U S A.* 2016; 113: 13402–13407. [PubMed: 27821780]
- [7]. Knight AM, Culviner PH, Kurt-Yilmaz N, Zou T, Ozkan SB, Cavagnero S. Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics. *ACS Chem Biol.* 2013; 8: 1195–1204. [PubMed: 23517476]
- [8]. Eichmann C, Preissler S, Riek R, Deuerling E. Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *Proc Natl Acad Sci U S A.* 2010; 107: 9111–9116. [PubMed: 20439768]
- [9]. Cabrita LD, Cassaignau AME, Launay HMM, Waudby CA, Wlodarski T, Camilloni C, Karyadi M-E, Robertson AL, Wang X, Wentink AS, Goodsell L, et al. A structural ensemble of a ribosome-nascent chain complex during cotranslational protein folding. *Nat Struct Mol Biol.* 2016; 23: 278–285. [PubMed: 26926436]
- [10]. Deckert A, Waudby CA, Wlodarski T, Wentink AS, Wang X, Kirkpatrick JP, Paton JFS, Camilloni C, Kukic P, Dobson CM, Vendruscolo M, et al. Structural characterization of the interaction of  $\alpha$ -synuclein nascent chains with the ribosomal surface and trigger factor. *Proc Natl Acad Sci U S A.* 2016; 113: 5012–5017. [PubMed: 27092002]
- [11]. Hsu S-TD, Fucini P, Cabrita LD, Launay H, Dobson CM, Christodoulou J. Structure and dynamics of a ribosome-bound nascent chain by NMR spectroscopy. *Proc Natl Acad Sci U S A.* 2007; 104: 16516–16521. [PubMed: 17940046]
- [12]. Hsu S-TD, Cabrita LD, Fucini P, Christodoulou J, Dobson CM. Probing side-chain dynamics of a ribosome-bound nascent chain using methyl NMR spectroscopy. *J Am Chem Soc.* 2009; 131: 8366–8367. [PubMed: 19492839]
- [13]. Cabrita LD, Hsu S-TD, Launay H, Dobson CM, Christodoulou J. Probing ribosome-nascent chain complexes produced in vivo by NMR spectroscopy. *Proc Natl Acad Sci U S A.* 2009; 106: 22239–22244. [PubMed: 20018739]
- [14]. Waudby CA, Wlodarski T, Karyadi M-E, Cassaignau AME, Chan SHS, Wentink AS, Schmidt-Engler JM, Camilloni C, Vendruscolo M, Cabrita LD, Christodoulou J. Systematic mapping of free energy landscapes of a growing filamin domain during biosynthesis. *Proc Natl Acad Sci U S A.* 2018; 115: 9744–9749. [PubMed: 30201720]
- [15]. McCoy AJ, Fucini P, Noegel AA, Stewart M. Structural basis for dimerization of the Dictyostelium gelation factor (ABP120) rod. *Nat Struct Biol.* 1999; 6: 836–841. [PubMed: 10467095]
- [16]. Ollerenshaw JE, Tugarinov V, Kay LE. Methyl TROSY: explanation and experimental verification. *Magn Reson Chem.* 2003; 41: 843–852.

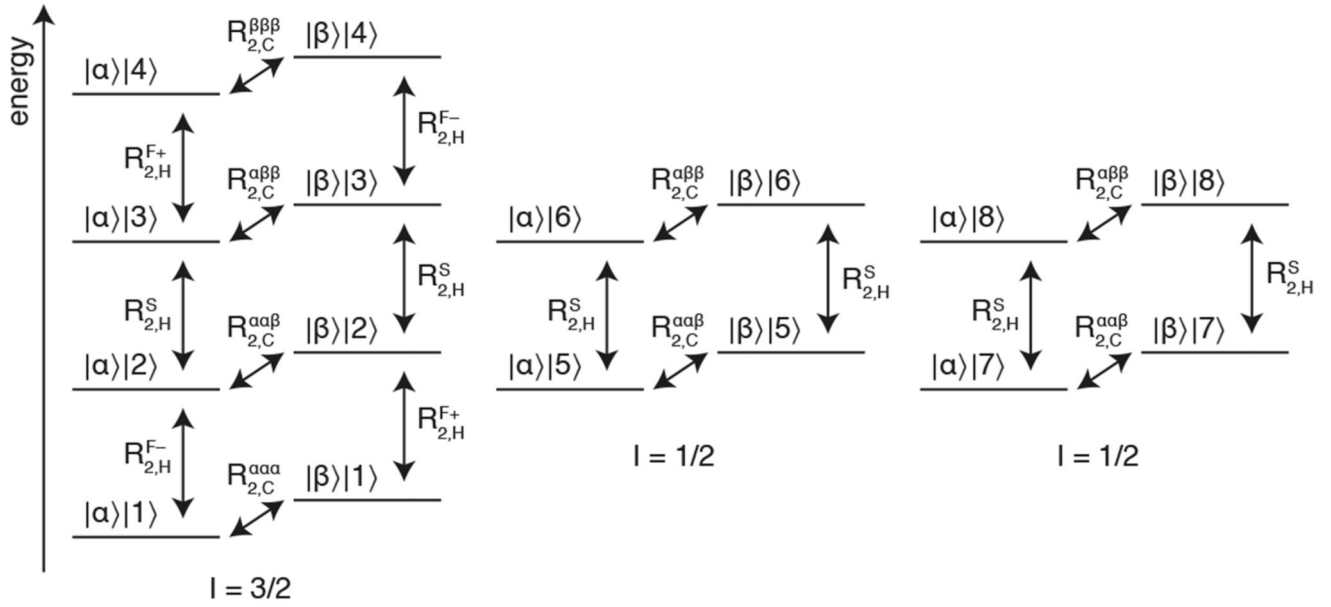
- [17]. Tugarinov V, Hwang PM, Ollerenshaw JE, Kay LE. Cross-correlated relaxation enhanced  $^1\text{H}$ [bond] $^{13}\text{C}$  NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc.* 2003; 125: 10420–10428. [PubMed: 12926967]
- [18]. Tugarinov V, Kay LE. Relaxation rates of degenerate  $^1\text{H}$  transitions in methyl groups of proteins as reporters of side-chain dynamics. *J Am Chem Soc.* 2006; 128: 7299–7308. [PubMed: 16734484]
- [19]. Liu W, Zheng Y, Cistola DP, Yang D. Measurement of methyl  $^{13}\text{C}$ - $^1\text{H}$  cross-correlation in uniformly  $^{13}\text{C}$ -,  $^{15}\text{N}$ -, labeled proteins. *J Biomol NMR.* 2003; 27: 351–364. [PubMed: 14512732]
- [20]. Tugarinov V, Sprangers R, Kay LE. Probing side-chain dynamics in the proteasome by relaxation violated coherence transfer NMR spectroscopy. *J Am Chem Soc.* 2007; 129: 1743–1750. [PubMed: 17249677]
- [21]. Sun H, Kay LE, Tugarinov V. An optimized relaxation-based coherence transfer NMR experiment for the measurement of side-chain order in methyl-protonated, highly deuterated proteins. *J Phys Chem B.* 2011; 115: 14878–14884. [PubMed: 22040035]
- [22]. Tugarinov V, Scheurer C, Brüschweiler R, Kay LE. Estimates of methyl  $^{13}\text{C}$  and  $^1\text{H}$  CSA values ( $\sigma$ ) in proteins from cross-correlated spin relaxation. *J Biomol NMR.* 2004; 30: 397–406. [PubMed: 15630560]
- [23]. Sabo TM, Michael Sabo T, Bakhtiari D, Walter KFA, McFeeters RL, Giller K, Becker S, Griesinger C, Lee D. Thermal coefficients of the methyl groups within ubiquitin. *Protein Science.* 2012; 21: 562–570. DOI: 10.1002/pro.2045 [PubMed: 22334336]
- [24]. Tugarinov V, Kay LE.  $^1\text{H}$ , $^{13}\text{C}$ - $^1\text{H}$ , $^1\text{H}$  dipolar cross-correlated spin relaxation in methyl groups. *J Biomol NMR.* 2004; 29: 369–376. [PubMed: 15213435]
- [25]. Sun H, Godoy-Ruiz R, Tugarinov V. Estimating side-chain order in methyl-protonated, perdeuterated proteins via multiple-quantum relaxation violated coherence transfer NMR spectroscopy. *Journal of Biomolecular NMR.* 2012; 52: 233–243. DOI: 10.1007/s10858-012-9604-y [PubMed: 22314703]
- [26]. Tugarinov V, Karamanos TK, Clore GM. Magic-Angle-Pulse Driven Separation of Degenerate  $^1\text{H}$  Transitions in Methyl Groups of Proteins: Application to Studies of Methyl Axis Dynamics. *Chemphyschem.* 2020; 21: 1087–1091. [PubMed: 32246547]
- [27]. Tugarinov V, Kay LE. Estimating Side-Chain Order in [ $^2\text{H}$ ; $^{13}\text{C}$ ]-Labeled High Molecular Weight Proteins from Analysis of HMQC/HSQC Spectra. *J Phys Chem B.* 2013; 117: 3571–3577. [PubMed: 23458382]
- [28]. Iwahara J, Jung Y-S, Clore GM. Heteronuclear NMR spectroscopy for lysine  $\text{NH}(3)$  groups in proteins: unique effect of water exchange on  $(^{15}\text{N})$  transverse relaxation. *J Am Chem Soc.* 2007; 129: 2971–2980. [PubMed: 17300195]
- [29]. Chan SHS, Waudby CA, Cassignau AME, Cabrita LD, Christodoulou J. Increasing the sensitivity of NMR diffusion measurements by paramagnetic longitudinal relaxation enhancement, with application to ribosome-nascent chain complexes. *J Biomol NMR.* 2015; 63: 151–163. [PubMed: 26253948]
- [30]. Waudby CA, Christodoulou J. An analysis of NMR sensitivity enhancements obtained using non-uniform weighted sampling, and the application to protein NMR. *J Magn Reson.* 2012; 219: 46–52. [PubMed: 22609525]
- [31]. Kay LE, Bull TE, Nicholson LK, Griesinger C, Schwalbe H, Bax A, Torchia DA. The measurement of heteronuclear transverse relaxation times in  $\text{ax}^3$  spin systems via polarization-transfer techniques. *J Magn Reson.* 1992; 100: 538–558.
- [32]. Cheng N-S. Formula for the Viscosity of a Glycerol-Water Mixture. *Industrial & Engineering Chemistry Research.* 2008; 47: 3285–3288. DOI: 10.1021/ie071349z
- [33]. Volk A, Kähler CJ. Density model for aqueous glycerol solutions. *Exp Fluids.* 2018; 59: 75.
- [34]. Fedorov, VV, Leonov, SL. *Optimal Design for Nonlinear Response Models.* CRC Press; 2013.
- [35]. Hyun SW, Yang M, Flournoy N. A procedure for finding an improved upper bound on the number of optimal design points. *Comput Stat Data Anal.* 2013; 58: 276–282.
- [36]. Kiefer J, Wolfowitz J. The Equivalence of Two Extremum Problems. *Canad J Math.* 1960; 12: 363–366.

- [37]. Jones JA. Optimal Sampling Strategies for the Measurement of Relaxation Times in Proteins. *J Magn Reson.* 1997; 126: 283–286.
- [38]. Ober RJ, Lin Z, Ye H, Ward ES. Achievable accuracy of parameter estimation for multidimensional NMR experiments. *J Magn Reson.* 2002; 157: 1–16. [PubMed: 12202128]
- [39]. Reci A, Sederman AJ, Gladden LF. Optimising magnetic resonance sampling patterns for parametric characterisation. *J Magn Reson.* 2018; 294: 35–43. [PubMed: 30005192]
- [40]. Carlström G, Elvander F, Swärd J, Jakobsson A, Akke M. Rapid NMR Relaxation Measurements Using Optimal Nonuniform Sampling of Multidimensional Accordion Data Analyzed by a Sparse Reconstruction Method. *J Phys Chem A.* 2019; 123: 5718–5723. [PubMed: 31194551]
- [41]. Bahrami A, Tonelli M, Sahu SC, Singarapu KK, Eghbalnia HR, Markley JL. Robust, integrated computational control of NMR experiments to achieve optimal assignment by ADAPT-NMR. *PLoS One.* 2012; 7 e33173 [PubMed: 22427982]
- [42]. Lavalette D, Amand B, Pochon F. Rotational relaxation of 70S ribosomes by a depolarization method using triplet probes. *Proc Natl Acad Sci U S A.* 1977; 74: 1407–1411. [PubMed: 323851]
- [43]. Wang J, Karki C, Xiao Y, Li L. Electrostatics of Prokaryotic Ribosome and Its Biological Implication. *Biophys J.* 2020; 118: 1205–1212. [PubMed: 32023436]
- [44]. Cassaignau AME, Launay HMM, Karyadi M-E, Wang X, Waudby CA, Deckert A, Robertson AL, Christodoulou J, Cabrita LD. A strategy for co-translational folding studies of ribosome-bound nascent chain complexes using NMR spectroscopy. *Nat Protoc.* 2016; 11: 1492–1507. [PubMed: 27466710]
- [45]. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR.* 1995; 6: 277–293. [PubMed: 8520220]
- [46]. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, et al. Array programming with NumPy. *Nature.* 2020; 585: 357–362. [PubMed: 32939066]
- [47]. SciPy 1.0 Contributors. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods.* 2020; 17: 261–272. [PubMed: 32015543]
- [48]. Lam, SK; Pitrou, A; Seibert, S. Numba; Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC - LLVM '15; New York, USA. 2015.
- [49]. Didenko T, Boelens R, Rudiger SGD. 3D DOSY-TROSY to determine the translational diffusion coefficient of large protein complexes. *Protein Engineering Design and Selection.* 2011; 24: 99–103. DOI: 10.1093/protein/gzq091
- [50]. Maciejewski MW, Schuyler AD, Gryk MR, Moraru II, Romero PR, Ulrich EL, Eghbalnia HR, Livny M, Delaglio F, Hoch JC. NMRbox: A Resource for Biomolecular NMR Computation. *Biophys J.* 2017; 112: 1529–1534. [PubMed: 28445744]



**Figure 1. Solution-state NMR spectroscopy of FLN5+6 RNCs.**

(A) Design of RNC constructs [9], showing the complete FLN5 filamin domain (646–750) flanked by an N-terminal His<sub>6</sub> affinity tag, and the subsequent FLN6 domain and 17 amino acid secM arrest motif, of total length  $L$ . (B) Crystal structure (1qfh [15]) of FLN5 showing the location of the five isoleucine residues, with C $\delta$  atoms highlighted as magenta spheres. (C) <sup>1</sup>H, <sup>13</sup>C HMQC correlation spectrum of a [<sup>2</sup>H, <sup>13</sup>CH<sub>3</sub>-Ile]-labelled FLN5+67 RNC (blue) and isolated [<sup>2</sup>H, <sup>13</sup>CH<sub>3</sub>-Ile]-labelled FLN5 (298 K, 950 MHz). (D) 1D spectrum showing the first increment of an <sup>1</sup>H, <sup>13</sup>C HMQC measurement of the FLN5+67 RNC.



$$\begin{aligned}
 |1\rangle &= |\alpha\alpha\alpha\rangle \\
 |2\rangle &= (|\alpha\alpha\beta\rangle + |\alpha\beta\alpha\rangle + |\beta\alpha\alpha\rangle) / \sqrt{3} \\
 |3\rangle &= (|\alpha\beta\beta\rangle + |\beta\alpha\beta\rangle + |\beta\beta\alpha\rangle) / \sqrt{3} \\
 |4\rangle &= |\beta\beta\beta\rangle \\
 |5\rangle &= (|\alpha\alpha\beta\rangle - |\alpha\beta\alpha\rangle) / \sqrt{2} \\
 |6\rangle &= (|\beta\alpha\beta\rangle - |\beta\beta\alpha\rangle) / \sqrt{2} \\
 |7\rangle &= (|\alpha\alpha\beta\rangle + |\alpha\beta\alpha\rangle - 2|\beta\alpha\alpha\rangle) / \sqrt{6} \\
 |8\rangle &= (2|\alpha\beta\beta\rangle - |\beta\beta\alpha\rangle - |\beta\alpha\beta\rangle) / \sqrt{6}
 \end{aligned}$$

$$\begin{aligned}
 \eta_{HHHH} &= \frac{9}{40} \left( \frac{\mu_0}{4\pi} \right)^2 \frac{\hbar^2 \gamma_H^4}{r_{HH}^6} S^2 \tau_c \\
 \eta_{HHHC} &= \frac{1}{5} \left( \frac{\mu_0}{4\pi} \right)^2 \frac{\hbar^2 \gamma_C \gamma_H^3}{r_{CH}^3 r_{HH}^3} S^2 \tau_c
 \end{aligned}$$

$$R_{2,H}^{F\pm} = R_{2,H}^S + 2\eta_{HHHH} \pm \eta_{HHHC}$$

$$R_{2,C}^{aaa} = R_{2,C} + 3\eta_{CHCH} + 3\eta_{CHC}$$

$$R_{2,C}^{aab} = R_{2,C} - \eta_{CHCH} + \eta_{CHC}$$

$$R_{2,C}^{abb} = R_{2,C} - \eta_{CHCH} - \eta_{CHC}$$

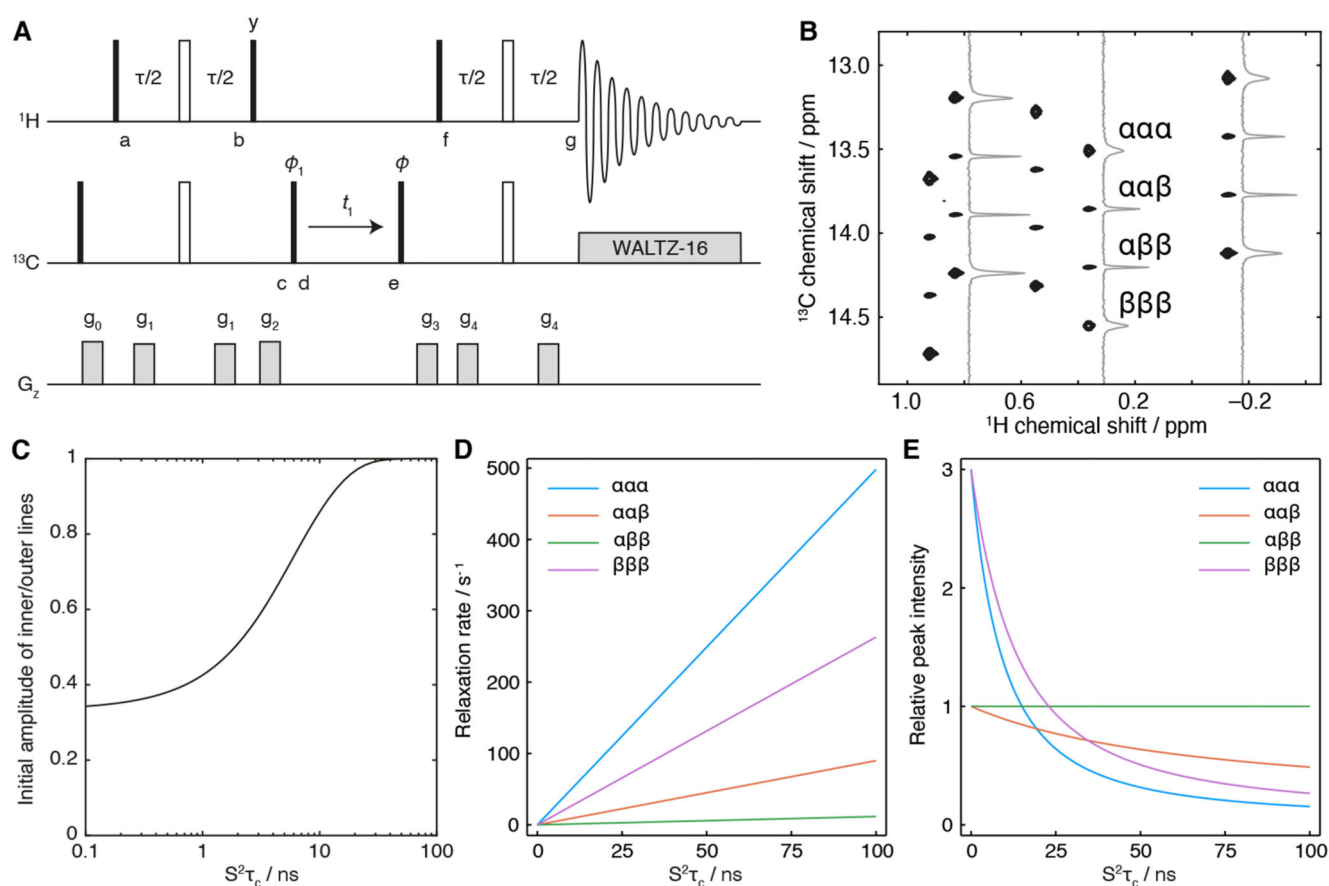
$$R_{2,C}^{bbb} = R_{2,C} + 3\eta_{CHCH} - 3\eta_{CHC}$$

$$\eta_{CHCH} = \frac{2}{45} \left( \frac{\mu_0}{4\pi} \right)^2 \frac{\hbar^2 \gamma_C^2 \gamma_H^2}{r_{CH}^6} S^2 \tau_c$$

$$\eta_{CHC} = \frac{4}{45} \left( \frac{\mu_0}{4\pi} \right) \frac{\hbar \gamma_C \gamma_H}{r_{CH}^3} \gamma_C B_0 \Delta\sigma_C S^2 \tau_c$$

**Figure 2.**

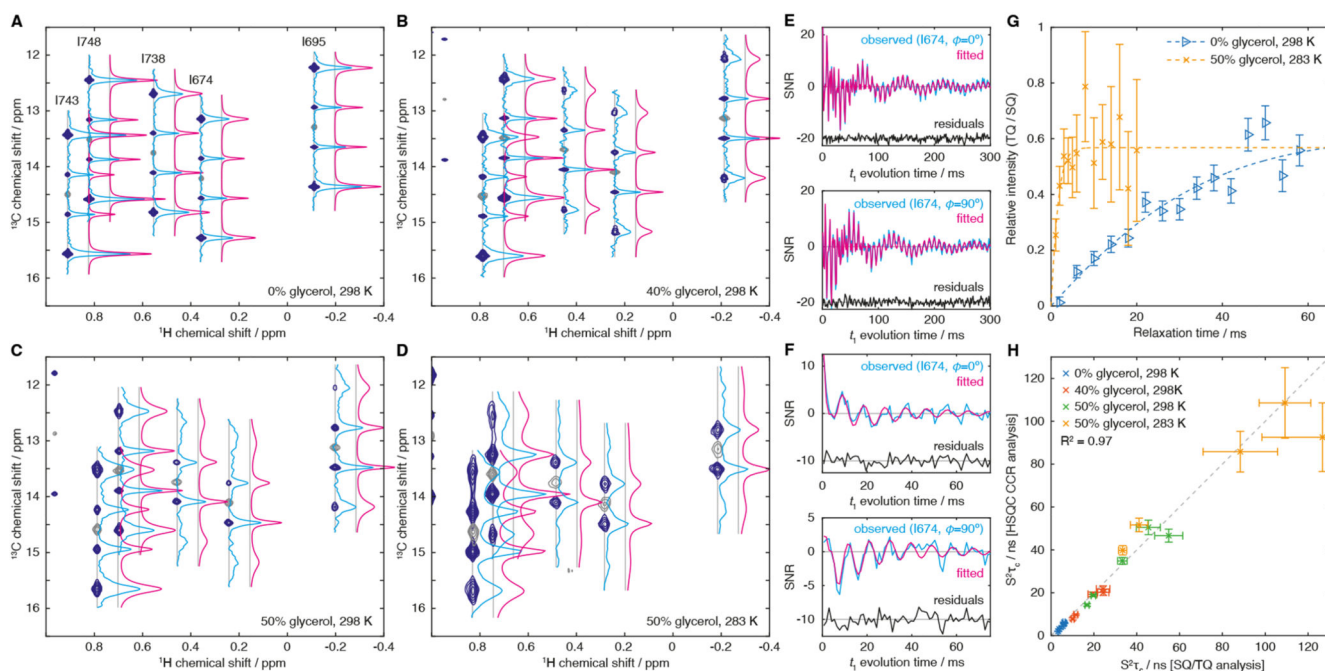
Energy level diagram for a  $^{13}\text{CH}_3$  spin system. Energy levels are labelled as the product of  $^{13}\text{C}$  and  $^1\text{H}$  spin states,  $|\text{C}\rangle|\text{H}\rangle$ , where the numbered  $^1\text{H}$  eigenstates are symmetrised linear combinations of the individual spin states as indicated [18].  $^1\text{H}$  eigenstates can be grouped into three independent manifolds and the spin quantum numbers,  $I$ , for these are shown below the diagram. Single quantum  $^1\text{H}$  and  $^{13}\text{C}$  transitions are indicated with arrows alongside their relaxation rates, expressions for which are provided below [19].



**Figure 3. Analysis of the F1-coupled HSQC experiment.**

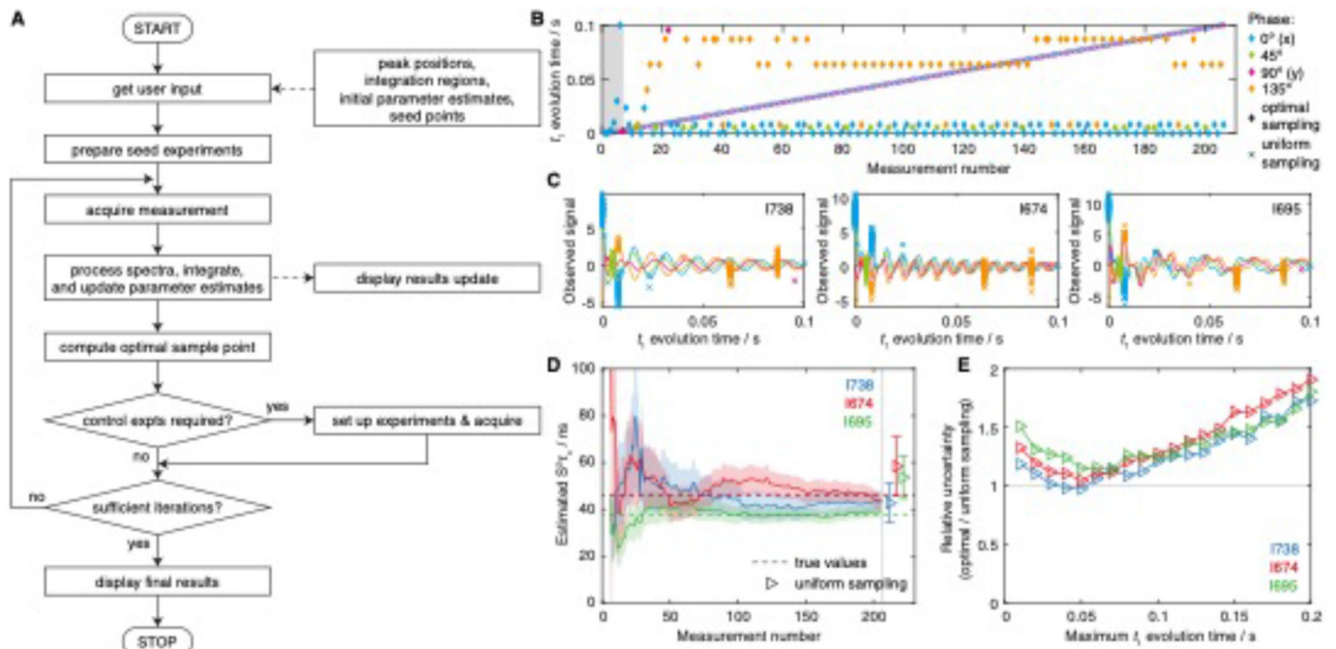
(A) Pulse sequence diagram of the F1-coupled  $^1\text{H}$ ,  $^{13}\text{C}$  HSQC experiment used in this work. The delay  $\tau$  is 4 ms ( $1/2J_{\text{ch}}$ ), phase  $\phi_1 = x, -x$ ; receiver phase  $\phi_{\text{rx}} = x, -x$ . States-TPPI quadrature detection is achieved by incrementing the phase of the pulse marked  $\phi$ . Alternatively, for measurements in the time domain, the phase  $\phi$  can be set to any value from 0 to  $180^\circ$ . Gradients are applied as 1 ms smoothed squares, with strengths  $g_0 = 0.0935 \text{ T m}^{-1}$ ,  $g_1 = 0.0385 \text{ T m}^{-1}$ ,  $g_2 = 0.275 \text{ T m}^{-1}$ ,  $g_3 = 0.19 \text{ T m}^{-1}$  and  $g_4 = 0.07 \text{ T m}^{-1}$ . (B) F1-coupled  $^1\text{H}$ ,  $^{13}\text{C}$  HSQC spectrum of [ $^2\text{H}$ ,  $^{13}\text{CH}_3$ -Ile]-labelled FLN5 (950 MHz  $^1\text{H}$  Larmor frequency) with cross-sections illustrating the differential relaxation of  $^1\text{H}$  spin states, as indicated, within the  $^{13}\text{C}$  multiplet. (C) Ratio of the initial amplitudes of inner and outer transitions,  $I_{\text{in}}/I_{\text{out}} = (3 - \dots)/(3 + \dots)$  (Eq. 3), plotted as a function of  $S^2\tau_c$  with calculated according to Eq. 2. (D) Variation in  $^{13}\text{C}$  single quantum relaxation rates as a function of  $S^2\tau_c$ , calculated at 800 MHz according to Fig. 2, with a  $^{13}\text{C}$  csa of 25 ppm. The  $^{13}\text{C}$   $R_2$  has been estimated in the absence of interactions with external protons. (E) Variation in the relative intensities of transitions in the  $^{13}\text{C}$  multiplet as a function of  $S^2\tau_c$ , calculated from the rates in panel D by Fourier transformation of Eq. 3, including 20 Hz exponential line broadening.





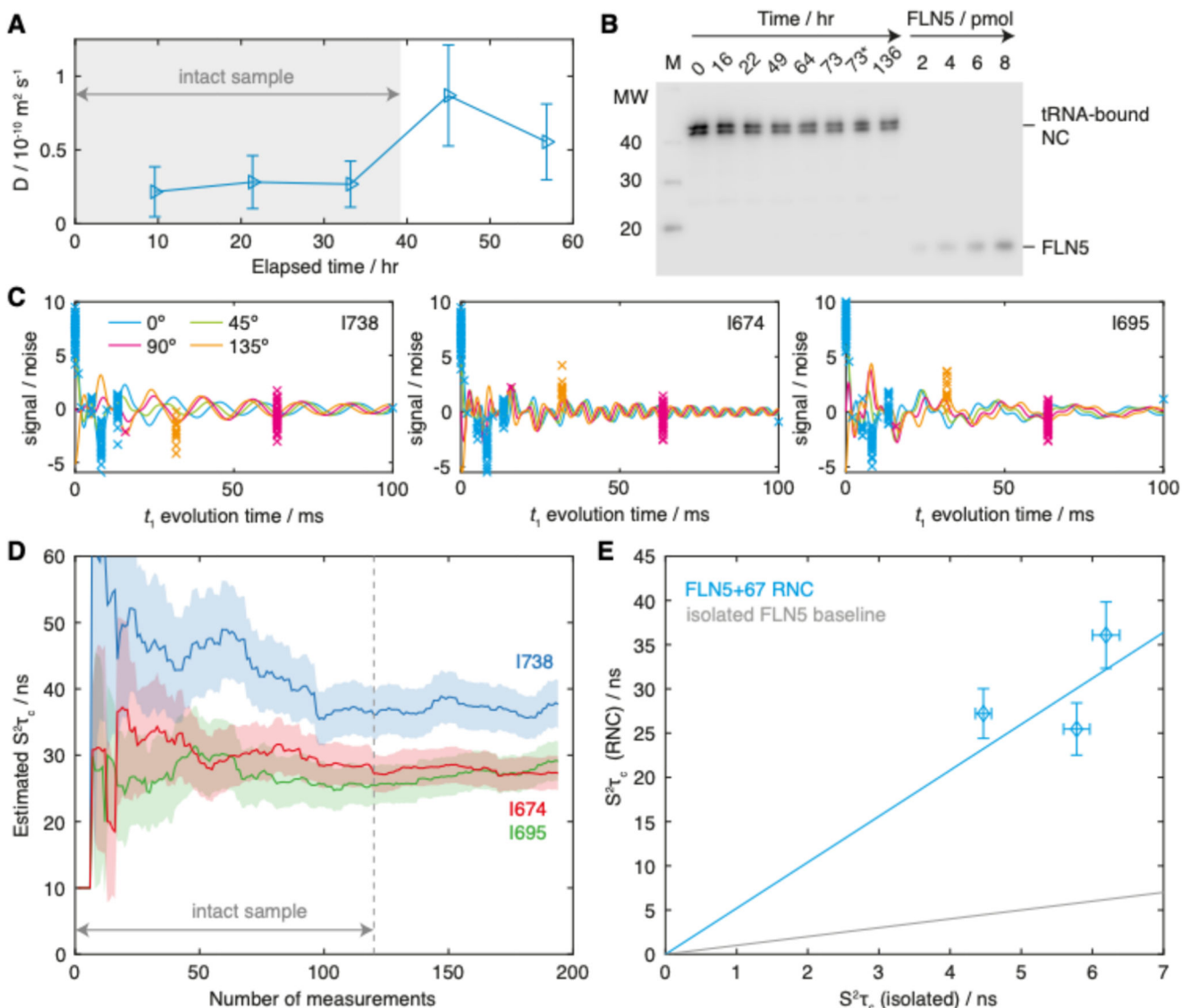
**Figure 4. Experimental measurements of methyl group dynamics in isolated FLN5.**

(A–D) F1-coupled  $^1\text{H}$ ,  $^{13}\text{C}$  HSQC spectra (Fig. 3A, dark blue) and reference  $^1\text{H}$ ,  $^{13}\text{C}$  HMQC (grey) acquired for isolated [ $^2\text{H}$ ,  $^{13}\text{CH}_3$ -Ile]-labelled FLN5 at varying temperatures and  $d_8$ -glycerol concentrations as indicated. Spectra were processed with exponential window functions in the indirect dimension (5 Hz (A), 5 Hz (B), 10 Hz (C) and 20 Hz (D)). Cross-sections through  $^{13}\text{C}$  multiplets are shown with blue lines, and fits to the Fourier transform of Eq. X, incorporating the effects of the window function, are shown in magenta (horizontally offset for clarity). (E–F) Time domain analysis of relaxation measurements, showing the observed  $t_1$  evolution of cosine ( $\phi=0^\circ$ ) and sine ( $\phi=90^\circ$ ) modulated components of the I674 resonance, in (E) 0% glycerol, 298 K, and (F) 50% glycerol, 283 K (blue lines). Cosine and sine modulated components were fitted simultaneously to Eq. 3 (magenta lines). (G) Reference measurements of  $S^2\tau_c$  in isolated FLN5 by measurement of the build-up of triple quantum magnetisation [21]. Data and fits are shown for the I674 resonance in 0% glycerol, 298 K, and 50% glycerol, 283 K, as indicated. (H) Comparison of  $S^2\tau_c$  measurements determined by both experimental approaches. Data are shown for all five isoleucine groups of FLN5, with solution conditions as indicated.



**Figure 5. Adaptive experimental design strategy and validation using synthetic data.**

(A) Flowchart illustrating the adaptive sampling acquisition process. (B) Measurement points acquired during a simulated experiment, following an optimal adaptive sampling scheme or using conventional uniform sampling. The acquisition of seed measurements is marked with grey shading, while the measurement phase,  $\phi$ , is shown with colour coding as indicated. (C) The final set of observed measurements (normalised by the noise level) are shown fitted to Eq. 3, with the measurement phase,  $\phi$ , coloured as in panel B. (D) Evolution of  $S^2\tau_c$  parameter estimates as a function of the number of measurements (solid lines). Shading is used to indicate the experimental uncertainty ( $\pm$  s.e.). The true values used in generating the synthetic data are indicated with dashed lines, while estimates obtained from a matched, uniformly sampled dataset are shown as points with error bars (horizontally offset for clarity). (E) Precision of  $S^2\tau_c$  parameter estimates obtained using a uniform sampling scheme, with maximum  $t_1$  evolution time as indicated, relative to those determined using an optimally designed adaptive sampling scheme. Uncertainties were determined as the mean standard error across 500 independent simulations. The total number of measurements was identical for all experiments.



**Figure 6. Measurement and analysis of methyl  $S^2\tau_c$  values in the FLN5+67 RNC.**

(A) Time course of NC diffusion coefficients measured for the isolated I738, I674 and I695 resonances. Grey shading indicates the total measurement time selected for further analysis. (B) Western blot (anti-histidine) time course of the FLN5+67 RNC, showing aliquots taken periodically from a sample of the RNC incubated in parallel with NMR measurements. The asterisk indicates an aliquot taken from the actual NMR sample. The locations of the higher molecular weight tRNA-bound NC, and of isolated FLN5, are indicated. (C) Adaptively sampled F1-coupled HSQC measurements for I738, I674 and I695 resonances of the FLN5+67 RNC (normalised by the noise level) are shown fitted to Eq. 3, with the measurement phase,  $\phi$ , coloured as indicated. (D) Evolution of  $S^2\tau_c$  parameter estimates as a function of the number of measurements (solid lines). Shading is used to indicate the experimental uncertainty ( $\pm$  s.e.). The dashed line indicates the cut-off for sample integrity, as determined by diffusion and western blot analysis (panels A and B). (E) Comparison of

isoleucine  $S^2\tau_c$  values measured for the FLN5+67 RNC and for isolated FLN5. A regression line through the origin is shown in blue (with a gradient of 5.2), and a reference line for the isolated domain is shown in grey.