Contents lists available at ScienceDirect

Progress in Nuclear Magnetic Resonance Spectroscopy

Protein folding on the ribosome studied using NMR spectroscopy

CrossMark

Christopher A. Waudby, Hélène Launay, Lisa D. Cabrita, John Christodoulou *

Institute of Structural and Molecular Biology, UCL and Birkbeck College, London WC1E 6BT, United Kingdom

Edited by David Neuhaus and David Gadian

A R T I C L E   I N F O

A B S T R A C T

NMR spectroscopy is a powerful tool for the investigation of protein folding and misfolding, providing a characterization of molecular structure, dynamics and exchange processes, across a very wide range of timescales and with near atomic resolution. In recent years NMR methods have also been developed to study protein folding as it might occur within the cell, in a *de novo* manner, by observing the folding of nascent polypeptides in the process of emerging from the ribosome during synthesis. Despite the 2.3 MDa molecular weight of the bacterial 70S ribosome, many nascent polypeptides, and some ribosomal proteins, have sufficient local flexibility that sharp resonances may be observed in solution-state NMR spectra. In providing information on dynamic regions of the structure, NMR spectroscopy is therefore highly complementary to alternative methods such as X-ray crystallography and cryo-electron microscopy, which have successfully characterized the rigid core of the ribosome particle. However, the low working concentrations and limited sample stability associated with ribosome–nascent chain complexes means that such studies still present significant technical challenges to the NMR spectroscopist. This review will discuss the progress that has been made in this area, surveying all NMR studies that have been published to date, and with a particular focus on strategies for improving experimental sensitivity.

© 2013 Published by Elsevier B.V.

Contents

---

* Corresponding author.
  E-mail address: j.christodoulou@ucl.ac.uk (J. Christodoulou).

1. Introduction

The process by which a nascent polypeptide acquires its native folded structure after emerging from the ribosomal exit tunnel has been the subject of intense research since the pioneering observations of Anfinsen and Levinthal [1–6]. The 'protein folding problem' is not one of purely scientific interest however, for the failure of proteins to reach a stable structure may result in the population of misfolded or aggregated species that are implicated in a very wide range of pathological disorders [7]. In some cases this may be the result of mutations that destabilize the native state, while other mutations may stabilize intermediates or transition states in the pathway to aggregated species such as amyloid fibrils. However, in many cases misfolding may occur in the absence of any apparent genetic or environmental factors, purely as a result of chance excursions into misfolded regions of the folding landscape. Such occurrences may give rise to late-onset disorders, such as Alzheimer's and Parkinson's diseases. Clearly, there is an evolutionary incentive to minimize such events, but only to the extent that they impinge upon the process of natural selection. With the large increase in human lifespan in the past several centuries, no selective pressures can yet have acted, and in this regard such diseases have been described as 'post-evolutionary' [8]. In all cases, however, by developing our understanding of folding landscapes, energetically and structurally, we may hope to rationally develop strategies to destabilize misfolded states and guide folding to the desired native state.

Protein folding is most commonly studied using isolated proteins in dilute aqueous solutions, either under equilibrium conditions (e.g. measurements of thermodynamic stability by reversible denaturation) or by studying relaxation kinetics following perturbations such as chemical denaturants, temperature or pressure. NMR spectroscopy may be used to follow folding in real-time, for example using stopped-flow methods to allow the characterization of intermediate states [9–11], but it has also proved to be a particularly powerful probe of protein folding at equilibrium, in the characterization of transient unfolding events through amide hydrogen exchange kinetics [12,13], or in the characterization of high-energy intermediate states through the analysis of relaxation dispersion phenomena [14]. Such measurements can characterize not only the thermodynamics of the excited state and the associated transition state, but also can determine structural parameters such as chemical shifts and residual dipolar couplings (RDCs) in the excited state. In combination with restrained molecular dynamics (MD) simulations, these may be used to develop detailed structural models of otherwise invisible intermediate states [15–18].

Within living cells, however, protein folding is not an equilibrium process, but may occur in a vectorial manner as the nascent polypeptide emerges from the ribosome, during translation, into the complex cellular environment [5,6,19]. The co-translational acquisition of folded structure has been demonstrated for a number of systems by a very wide range of biophysical, spectroscopic and imaging techniques, including functional assays [20–22], limited proteolysis [23], conformation-specific antibody recognition [24,25], assays of disulfide bond formation [26,27], FRET [28,29], time-resolved fluorescence anisotropy [30], cryoelectron microscopy [31,32], atomic force microscopy [33], single molecule force spectroscopy [34,35] and, of particular interest to the present review, NMR spectroscopy [36–40]. The tethering of the nascent chain to the ribosome during synthesis may affect the folding process in several ways. The simple presence of a nearby surface may result in entropic, excluded volume effects on the thermodynamics of the folding process, while interactions with the surface may preferentially stabilize folded or unfolded states, particularly given

the high local concentrations resulting from tethering [34]. Such effects may perturb the kinetics or thermodynamics of the folding reaction, potentially altering the folding pathway from that that would be observed in equilibrium studies of the isolated protein. However, the vectorial nature of protein synthesis will also give rise to an energy landscape that evolves as a function of chain length [41]. This may create 'shortcuts' through the energy landscape, allowing folding to proceed in a more efficient, directed manner and reducing the likelihood of populating misfolded states. This is a fundamentally non-equilibrium process, and one that is only accessible to exploration in co-translational systems [6,42].

The mechanism of co-translational folding is also of great interest in protein systems subject to topological constraints. A survey in 2006 determined that 0.8% of proteins in the Protein Data Bank contained a knot in their fold [43]. While undoubtedly rare, the existence of such knotted proteins presents interesting and extreme cases that challenge theoretical models of protein folding [44–47]. A prominent question is to understand the process of knot formation: can this occur co-translationally? Certainly, the presence of the ribosome places severe constraints on knot formation via the C-terminus. Again however, this information is not accessible through reversible folding studies, for in at least some cases knots have been shown to persist in highly denatured states [48]. Non-equilibrium kinetic studies of the first folding events following synthesis and release have begun to address this question [49], but beyond a small number of particular cases, much remains to be understood about the conditions under which knot formation may occur co-translationally.

What contributions, then, might NMR spectroscopy be able to make to the questions outlined above? The power of NMR lies in the ability to provide rich information on both protein structure and dynamics and, in contrast to most alternative methods, the ability to simultaneously resolve many spin systems using multidimensional spectroscopy supplies a very large number of probes, resulting in a description of the entire system with near atomic resolution. It is clear that solution-state NMR spectroscopy will not compete with X-ray crystallography and cryo-electron microscopy in the structural characterization of the rigid 2.3 MDa core of
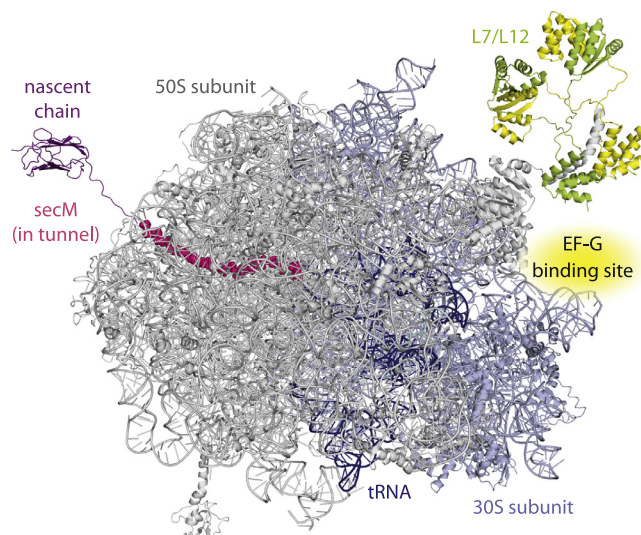
Fig. 1. Crystal structure of the *E. coli* 70S ribosome [53], showing complexed tRNA molecules aligned from a crystal structure of *Thermus thermophilus* [56], L7/L12 proteins modeled using NMR structures of the isolated proteins docked onto the L10 subunit [57–59], and a nascent polypeptide modeled upon a cryo-EM structure of a TnaC nascent chain [31].

the 70S ribosome (Fig. 1) [50–53], or in studies of structural preferences within the ribosomal exit tunnel [31,32]. On the surface of the ribosome however, and outside the tunnel, the flexibility of polypeptide chains gives rise to structural heterogeneity that has so far defied structural characterization by either X-ray crystallography or cryo-electron microscopy. It is here that NMR spectroscopy has proved to be a highly complementary technique, for the flexibility of such regions results in a reduction in their effective rotational correlation time, giving rise to sharp resonances despite the large size of the core ribosome particle. This observation led in 2004 to the high-resolution characterization of the structure and dynamics of the L7/L12 ribosome stalk region [54,55], which due to its flexibility was not observable in crystallographic studies of the ribosome. These initial findings stimulated the investigation of the structural and dynamical properties of ribosome-bound nascent chains during co-translational folding by NMR and, in 2007, we reported the first observation of a ribosome–nascent chain complex (RNC) by solution-state NMR spectroscopy [36].

However, while the potential of NMR to provide rich structural and dynamic information on co-translational protein folding is clear, it is still a subject that poses a large number of technical challenges. The preparation of RNCs, either through *in vitro* or *in vivo* methods, is a technically challenging process that requires the production of samples in which the ribosomes are programmed to uniformly arrest during translation a nascent chain of a given sequence. The production of RNCs for NMR studies must be carefully optimized in order to selectively isotopically label the nascent chain only, while delivering very high yields of stable and homogeneous complexes that are of a maximum nascent chain occupancy (the fraction of ribosomes carrying an attached nascent chain). Moreover, once produced, RNC samples are proving to be only moderately stable, with typical lifetimes of hours–days before the nascent chain is released from the complex or the ribosome particle itself starts to dissociate. Continual assessment of the sample integrity is therefore required throughout the data acquisition period to ensure that the measured signal is derived exclusively from ribosome-bound nascent chains. Thus, given the limited sample lifetimes, in combination with the maximum working concentrations possible for ribosomes (ca. 10 μM) and their large molecular weight (2.3 MDa), low sensitivity is the dominant characteristic of RNC NMR and experiments must be carefully designed to maximize the data obtained from each sample.

In this article, we will initially review the use of solution-state NMR spectroscopy for the study of the structure and dynamics of isolated ribosomes (Section 2). These observations were instrumental in demonstrating the feasibility of high-resolution studies of flexible domains within very large molecules, and led directly to studies of ribosome–nascent chain complexes by NMR. As outlined above however, there are many biochemical and spectroscopic challenges associated with the preparation and study of these large, unstable and low concentration samples by NMR, and these are discussed in detail in Section 3. We then review all NMR studies of RNCs that have been published to date (Section 4), considering both the biochemical and spectroscopic approaches taken in each case, and also the biological information that has been gleaned from these studies. Finally, we close by discussing the potential for the field in the future (Section 5).

2. Solution-state NMR studies of the ribosome

The first studies of ribosome structure by solution-state NMR spectroscopy were reported in 1980 and revealed that, despite the 2.3 MDa molecular weight of the ribosome (with a rotational correlation time of 2.5 μs at 298 K [60]), sharp resonances were apparent in the 1D ^{1}H spectrum of the 70S particle [61]. This indicates the existence of domains having significant mobility and effective correlation times one or two orders of magnitude reduced from that of the whole complex. Both 30S and 50S subunits were determined to contribute to the observed signal, although no resonances could be detected from purified rRNA components [61]. Similar resonances were later also observed from the large 60S subunit of the eukaryotic ribosome [62]. By selective exchange and removal of subunits, observed signals were traced to the S1 protein in the 30S subunit and the L7/L12 proteins in the 50S subunit [62,63], although the intensities of S1 resonances decrease significantly in the 70S particle. Moreover, addition of the elongation factor EF-G, which binds primarily to the 23S rRNA (Fig. 1), broadened the resonances of L7/L12 beyond detection, indicating that the dynamics of the stalk complex are quenched in the presence of the elongation factor [64]. NMR studies of the isolated L7/12 protein subsequently identified a conserved region on the surface of the C-terminal domain that interacts with EF-G, and also initiation factor 2 (IF2) and elongation factor Tu (EF-Tu), with millimolar affinity which, given the high local concentration of these translation factors when bound to the ribosome, results in large populations of the bound state on the ribosome [65].

These early results laid the path for later studies of the ribosome itself using high-resolution heteronuclear NMR methods. Well-resolved ^{1}H–^{15}N HSQC spectra can be obtained from isotopically labeled 70S ribosomes, and the observed resonances overlay closely with those of the individual 30S and 50S subunits (Fig. 2a and b) and, in particular, with the L7/L12 protein of the 50S subunit (Fig. 2c) [54,55]. Resonances are observed from the hinge and C-terminal domains (residues 42–120) of L7/L12 (Fig. 2d), and while there are small chemical shift perturbations in the hinge region, no significant chemical shift changes are observed in the C-terminal domain relative to the free protein, indicating that its structure is unchanged in the ribosome-associated form [55]. A key methodological development was the application of diffusion NMR measurements to verify the hydrodynamic radius of the observed species, providing direct confirmation that the spectra originate from ribosome-associated species rather than dissociated or degraded ribosomes (Fig. 2e) [54]. As discussed further below (Section 3.4), this is also an essential control measurement in the observation of RNC samples [36].

Analysis of the intensity of the L7/L12 resonances shows that two of the four copies present have sufficient mobility to be observed by solution-state NMR [54,55]. However, a small increase in the ^{15}N linewidth from an average of 6.3 ± 0.8 Hz to 9.3 ± 1.6 Hz was observed in the ribosome-bound form, indicating that rotational diffusion is still partially restricted [55]. ^{15}N nuclear spin relaxation measurements of the ribosome-bound form showed T_2 relaxation times of 30–40 ms in the structured core (residues 52–120), while those in the hinge region (residues 42–52) had somewhat slowed relaxation, with T_2 values of 70–90 ms [54]. An averaged apparent rotational correlation time of 13.6 ns was calculated, larger than that expected for the protein in the free state but certainly much less than that the value of 2.5 μs measured for the entire 70S particle [60], again indicating that tumbling of L7/L12 is largely decoupled from that of the ribosome. Finally, analysis of the relaxation data using a model of anisotropic rotational diffusion provided orientational restraints allowing a molecular model of this flexible region to be docked into a crystal structure of the core ribosome particle. Importantly, the resulting dynamic model is able to rationalize the interaction of L7/L12 with several other components of the ribosome including L10, L11 and EF-G, as observed in a series of cryoelectron and immunoelectron microscopic studies [66–69], providing an excellent example of the complementarity between NMR spectroscopy and other structural techniques. Moreover, these high-resolution studies of L7/L12
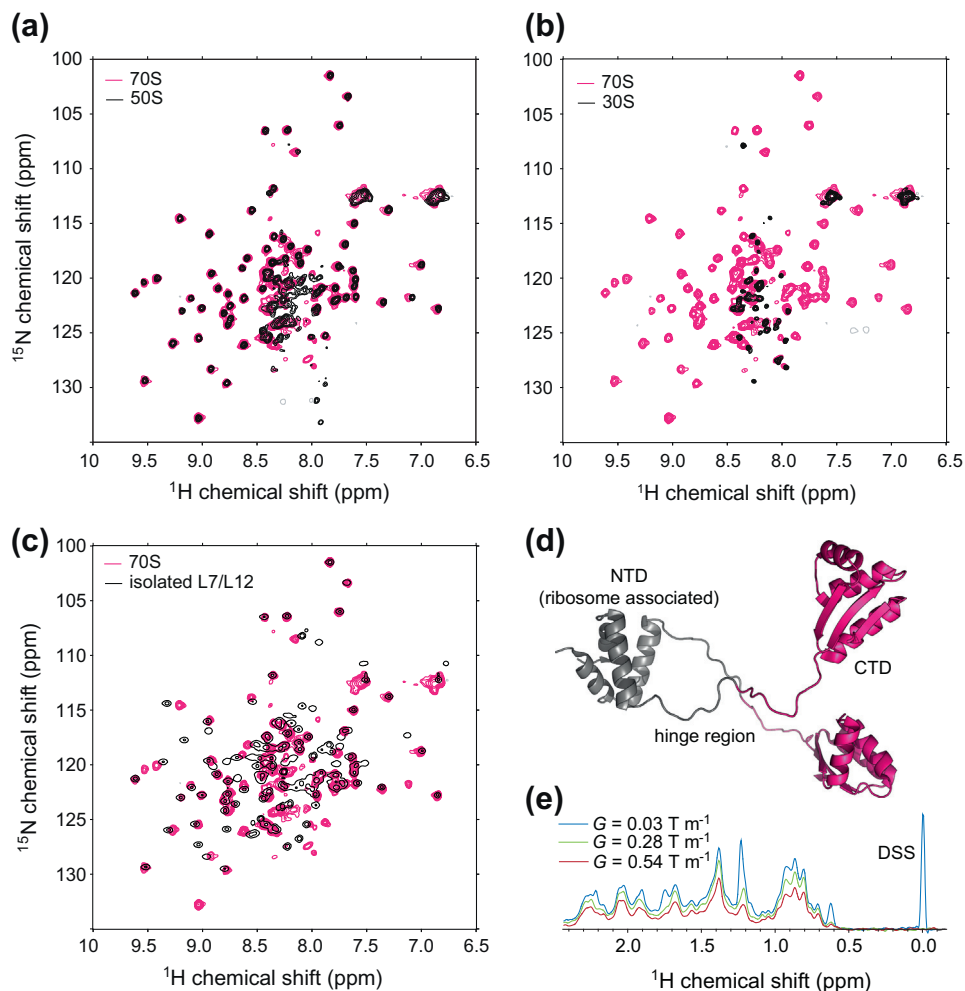
Fig. 2. ^{1}H,^{15}N HSQC spectrum (700 MHz, 298 K) of 70S ribosomes overlaid with spectra of the (a) 50S and (b) 30S subunits, and (c) with the isolated L7/L12 protein from the 50S subunit. (d) NMR structure of the isolated L7/L12 dimer [57], colored magenta to indicate residues observed in the ribosome-associated state. (e) ^{1}H STE PFG diffusion measurements of 70S ribosomes (700 MHz, 298 K, 100 ms diffusion delay, 4 ms bipolar gradients with strengths as indicated, data acquired at UCL Biomolecular NMR Facility) for determination of the hydrodynamic radius of observed species and assessment of ribosomal integrity.

were instrumental in demonstrating the feasibility of NMR studies of co-translational folding in ribosome–nascent chain complexes.

3. Technical aspects of the NMR of ribosome–nascent chain complexes

3.1. Choice of isotopic labeling scheme

Central to the study of RNCs by NMR is the ability to generate selectively isotopically-labeled and homogeneously stalled, ribosome-bound nascent chains. The selection of an isotopic labeling scheme for NMR studies of RNCs depends primarily on the folded or unfolded nature of the nascent chain species being targeted. The majority of RNC studies reported to date have used uniform ^{15}N (or ^{15}N/^{13}C) labeling, allowing the observation and characteriastion of both folded and disordered states in a single spectrum [36,38–40]. However, significantly greater sensitivity may be obtained using a ^{13}C labeling scheme to observe methyl groups, due to the greater proton concentration (three times that of amide groups) and to the favorable relaxation properties resulting from rapid rotation about the threefold symmetry axis [37].

In the future, RNC studies may be expected to benefit from the application of more advanced isotopic labeling schemes, which combine perdeuteration of samples with the selective introduction

of ^{13}CH$_3$ labels on sidechain methyl groups, most commonly on I, L and V residues, resulting in substantial reductions in transverse relaxation rates. Such labeling schemes, in combination with 'methyl-TROSY' ^{1}H–^{13}C HMQC experiments that avoid the mixing of slow and fast relaxing coherences [70], have enabled the high-resolution characterization of extremely large species such as the 20S proteasome (MW 670 kDa, τ_c approximately 180 ns at 65 °C) [71,72], and while these methods have not yet been applied to the study of RNCs, the potential gains in sensitivity and resolution appear to be extremely promising. However, a major limitation of ^{13}C-based experiments is the limited chemical shift dispersion. This is particularly problematic in the case of partially disordered states that are likely to be present in emergent nascent chains, and so when such species are being investigated it is preferable to take advantage of the large dispersion of the amide nitrogen chemical shift using a simple ^{15}N labeling scheme.

3.2. Preparation of ribosome–nascent chain complexes

The preparation of homogeneous and stable RNC samples in a reliable, cost-effective manner, and on the nmol scale required for NMR studies, is perhaps the most significant technical challenge associated with this field. Additionally, the production of homogeneous samples of programmed ribosomes requires that

translation must be arrested specifically at a given point in the protein sequence. Two main approaches have emerged to fulfill this need, each with different strengths: *in vitro* transcription–translation systems, in which cell extracts or reconstituted ribosomes are used to synthesize the nascent chain in a controlled, cell-free environment; and *in vivo* approaches, in which stalled RNCs are generated within living cells directly, before being harvested and purified.

The initial RNC samples studied by NMR were produced using an *Escherichia coli in vitro* coupled transcription–translation system, in combination with a linearized DNA template that lacks a stop codon to terminate translation [36,37]. In the absence of a stop codon, release factors cannot bind and the C-terminus of the nascent chain remains covalently attached to the P-site tRNA by the aminoacyl ester bond. With unlabeled ribosomes in the reaction mixture, the use of isotopically labeled amino acids allows for the translation of selectively labeled nascent chains, and together with a non-stop sequence, enables the nascent chains to remain bound to the ribosome. This strategy has the particular advantage that NMR-silent ribosomes are obtained, and therefore the resulting spectra are free from overlap with resonances from the L7/L12 stalk complex. The recently developed PURE (protein synthesis using recombinant elements) cell-free approach [73,74] also allows, in principle, control over all aspects of translation, including the presence or absence of auxiliary factors such as the ribosome-associated chaperone trigger factor (TF), release factors and other quality control mechanisms (e.g. tmRNA tagging in trans–translation) that release stalled ribosomes. However, *in vitro* approaches are typically extremely costly in terms of routine usage at the scales required for NMR spectroscopy, and together with the prospect of generating RNCs within their natural environment, this has stimulated the development of alternative *in vivo* approaches to sample production.

An *in vivo* approach for generating RNCs presents unique challenges compared to *in vitro* strategies, as there is less control of the ability to manipulate both translational arrest and the selective labeling of the nascent chain. To enable homogeneous translational arrest *in vivo*, protein motifs are introduced at the C-terminus of a DNA sequence that interact directly with the ribosome exit tunnel. Such sequences have been derived from proteins that modulate the translation of a downstream gene, providing an additional point of regulation for the control of gene expression [75]. For example, transient translational arrest of the secretion monitor protein in *E. coli* (SecM) regulates the translation of the downstream gene SecA, a protein involved in peptide secretion, according to the activity of the secretory system. Active translocation of SecM via an upstream signal peptide relieves translational arrest, whereupon the protein is exported to the periplasm and degraded. However, if secretion is impaired, stalling of the nascent SecM polypeptide induces a change in the secondary structure of the downstream mRNA to reveal a hidden initiation site for the SecA gene, and in this way the levels of SecA, and the activity of the secretory system, can be continually tuned to meet fluctuating demand [75,76]. The introduction of the 17 amino acid motif (F^{150}XXXXWIXXXXGIRAGP^{166}) derived from SecM to selectively stall nascent chains has proved to be instrumental in the development of RNC studies by NMR. Translational arrest is achieved from the presence of Pro166 within the P-site of the PTC, together with interactions made by a number of highly conserved residues within the SecM sequence with the ribosomal tunnel [76,77]. Selective stalling of nascent chains on *E. coli* ribosomes for structural studies using cryoEM has also been achieved using sequences derived from *E. coli* tryptophanase operon, TnaC [31], while sequences derived from the human cytomegalovirus and fungal arginine attenuator peptide have been used to stall eukaryotic ribosomes [78]. It is likely that these arrest sequences as well as others that have been

reported recently [79] may also prove to be useful for future NMR studies of RNCs.

To achieve selective isotopic labeling of the RNCs *in vivo*, the culturing conditions of the *E. coli* are manipulated, such that during growth a medium is used which permits the production of unlabeled ribosomes. The cells are then transferred into an isotopically rich medium in which translation takes place. Varying combinations of growth/expression media have been reported [39,40], though we find that high yields and reproducibility of RNC production required for NMR studies are achieved with the use of media compositions and growth conditions that can sustain high-cell density growth [38]. Following expression, the *in vivo* derived RNCs are subject to rigorous purification to recover homogeneous and highly programmed ribosomes, to ensure the strongest possible nascent chain NMR signal. Typical purification strategies involve a three-step strategy including sucrose cushions to isolate crude ribosomal material, affinity purification in which a tag (e.g. 6xHis or Strep) is used to isolate nascent chain containing ribosomes, and a sucrose density gradient to isolate intact monosomes. Finally, a combination of biochemical and NMR strategies are used to evaluate the integrity of the sample before and during acquisition. In particular, the ribosome concentration may be determined by UV absorption at 260 nm ($\varepsilon = 4.17 \times 10^7$ M^{-1} cm^{-1}) [54], while the attachment of the nascent chain may be verified by Western blotting directed against N-terminal purification tags [38,40].

3.3. NMR spectroscopy at very low concentrations

A defining feature of RNC NMR is the low sample concentration. In contrast to solution state NMR studies of isolated proteins or nucleic acids, which are typically carried out at concentrations of at least 100 μM and often several mM, ribosomes are limited to a concentration of 10 μM (we note however that given the large molecular weight of the ribosome, this still corresponds to a concentration of 23 mg/mL). The nascent chain concentration itself is also critical, and given the unstable and often complex behavior of RNC samples, the fraction of ribosomes harboring a nascent chain can change during the course of NMR data acquisition as the sample degrades. Clearly therefore, it is essential to optimize the acquisition process in order to maximize the signal to noise ratio (SNR) of the acquired data. Indeed, an entirely different mindset is required when compared to typical biomolecular NMR experiments, for given the limited lifetime of RNC samples it is simply not possible to compensate for inadequate signal by acquiring additional scans: it is essential that experiments are highly optimized before the experiments are begun. Therefore, the following four sections are dedicated to a review of methods for maximizing experimental sensitivity, through the spectrometer hardware and the design of pulse programs.

3.3.1. Hardware developments

At the level of the spectrometer hardware, the SNR in an experiment is described by:

$$\text{SNR} \propto \frac{n\gamma_e \sqrt{B_0^3 \gamma_d^3 t}}{\sqrt{R_S(T_S + T_A) + R_C(T_C + T_A)}} \tag{1}$$

where n is the quantity of spins in the detector, t is the total experimental time, B_0 is the magnetic field strength, γ_d (γ_e) is the gyromagnetic ratio of the nucleus being detected (excited), R_S and R_C are the resistances of the sample and the detector coil, and T_S, T_C and T_A are the temperatures of the sample, coil and preamplifier [80]. Given the low concentration and limited stability of RNC

samples, there is little opportunity to increase SNR through increasing n and t, and ^{1}H-based experiments already provide the greatest possible sensitivity with regard to the $\gamma^{5/2}$ scaling. However, hardware developments have had, and will continue to have, a very significant impact on the field. Given the scaling of SNR with $B_0^{3/2}$ for non-lossy samples, increases in spectrometer field strength provide large gains in sensitivity and, for example, the strongest field strength commercially available at the time of writing, 23.5 T (1000 MHz), would provide a 71% increase in sensitivity compared to the 16.4 T (700 MHz) spectrometers that have been used for several RNC studies to date [38,40]. However, for lossy samples the gain with field strength may be less.

The advent of cryogenic probe technology has revolutionized the capacity to study low concentrations of biomolecules: by reducing resistance and thermal noise in the coil and preamplifier (R_C, T_C and T_A in Eq. (1)), the SNR may be increased three or fourfold over a conventional probe [80]. At this point, sensitivity is limited primarily by rf power dissipation within the sample ($R_S T_S$ term in Eq. (1)), which has contributions from both ionic and dielectric conductivity. Where possible, this should be minimized by optimization of the buffer composition, such as reducing the ionic strength or using low-conductivity salts [81,82], but in the case of RNC samples the buffer components required to maintain the integrity of the ribosome particle places conflicting demands on the buffer composition. Dielectric power dissipation is not however uniform across the sample volume, but is dominated by electric-field 'hotspots' at the edges of the sample parallel to the B_1 field [83]. An alternative possibility is therefore to use (without concentration of the sample) NMR tubes with smaller diameters [84], or shaped tubes designed and oriented to minimize electric-field hotspots [1–6,83,85]. Despite the smaller total quantity of spins in such tubes, at moderate ionic strengths genuine improvements in SNR may be obtained.

3.3.2. Optimal sampling for two-dimensional NMR experiments

Two-dimensional (2D) NMR spectroscopy is the most basic component of biomolecular NMR capable of providing residue-specific probes of structure and dynamics. The indirect dimension of 2D experiments is encoded as a series of N points, with the chemical shift evolution period incremented by a dwell time, Δt, between each point, resulting in a total acquisition time $T_{aq} = N \cdot \Delta t$. The dwell time has a reciprocal relation to the spectral width, while the frequency resolution is related to the total acquisition time, $\Delta f = 1/T_{aq}$. Both the spectral width and the number of increments must be set prior to acquisition, and this choice has important consequences for the sensitivity of the experiment.
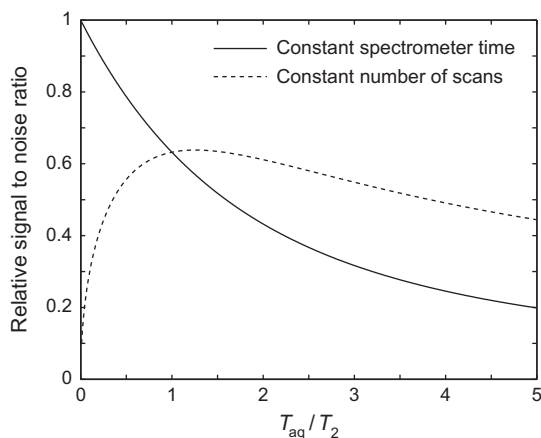

Fig. 3. Relative sensitivity of 2D NMR experiments as a function of the acquisition time, T_{aq}, relative to the tranverse relaxation time, T_2 (Eqs. (2) and (3)).

The signal intensity, $S(t)$, is related to the indirect evolution time, t, and the transverse relaxation time, T_2, of the spin system being observed: $S(t) \propto \exp(-t/T_2)$. In the common case where the sample is stable and a constant number of scans are acquired at each point, the total spectrometer time is directly proportional to the acquisition time, and may be increased without restriction. The measurement noise is therefore proportional to $\sqrt{T_{eq}}$, and so the net signal-to-noise ratio is described by Eq. (2) [7,86]. Plotted in Fig. 3 (dashed line), it may be observed that the SNR improves as the acquisition time is increased, up to a broad maximum at $T_{aq} \approx 1.26 T_2$, beyond which point SNR gradually decreases.

$$\text{SNR} \propto \frac{T_2}{\sqrt{T_{aq}}}[1 - \exp(-T_{aq}/T_2)] \qquad (2)$$

In contrast, if the sample is unstable then optimizing sensitivity becomes a question of most effectively partitioning a limited amount of spectrometer time between indirect time points. In this case, as the total spectrometer time is constant, the noise is also constant and so the experimental sensitivity (signal-to-noise ratio) is proportional to the acquired signal only. For a given acquisition time T_{aq}, this is given by:

$$\text{SNR} \propto \frac{T_2}{T_{aq}}[1 - \exp(-T_{aq}/T_2)] \qquad (3)$$

Eq. (3) is plotted in Fig. 3 (solid line), from which it is easily seen that SNR decreases monotonically with the acquisition time, and therefore that the sensitivity is also inversely related to the frequency resolution, Δf. Thus, for unstable samples where spectrometer time is limited, the acquisition of multidimensional NMR data presents a fundamental trade-off between sensitivity and resolution. Therefore, to maximize sensitivity, the minimum required spectral resolution for an RNC spectrum should be carefully considered prior to the acquisition of RNC data, and the acquisition time (i.e. number of increments) adjusted accordingly. As an aside, we also note from Eq. (3) that for a given acquisition time the SNR is independent of the spectral width and therefore there is no sensitivity advantage in folding spectra (minimizing both the spectral width and the number of increments by careful aliasing of resonances). This result, applicable to sensitivity-limited acquisition, is in contrast to the more common situation of sampling-limited acquisition, where a reduction in the number of increments can translate to substantial savings in spectrometer time [8,87].

Multidimensional NMR data is most commonly acquired with uniform sampling of the Nyquist grid, followed by apodization and processing with the discrete Fourier transform. The selection of an appropriate window function is important and has a large effect on the ultimate SNR of the spectrum. An alternative approach is the use of non-uniform sampling (NUS) methods, whereby the Nyquist grid is undersampled and advanced computational techniques are employed to reconstruct the frequency domain spectrum [9–11,88,89]. In the sampling-limited case, where the SNR of individual points is high, NUS approaches can provide order-of-magnitude savings in spectrometer time for the acquisition of multidimensional spectra. In the case of dilute samples such as RNCs, where many scans must be accumulated for every point in the indirect dimension, NUS methods have been reported to increase experimental sensitivity [12,13,90]. However, given the non-linear methods employed in the processing of NUS data, it is not straightforward to define the SNR of NUS spectra and alternative, indirect measures have been suggested, such as the probability of detecting weak peaks [14,91]. An alternative to NUS, applicable in the sensitivity-limited case where many scans are acquired for each point in the indirect dimension, is to sample every point in the Nyquist grid with a variable number of scans, according to some weighting function (Fig. 4a and b). This approach,
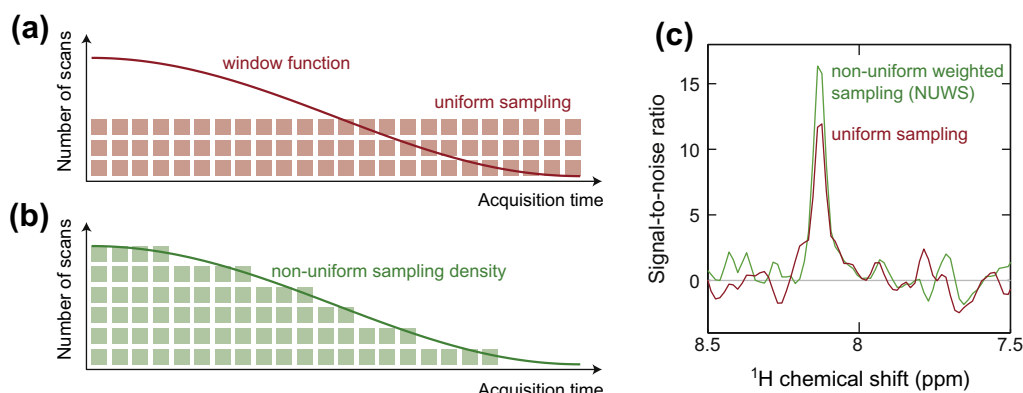
Fig. 4. (a) Uniform and (b) non-uniform weighted sampling schemes for the acquisition of the indirect dimension of multidimensional spectra, illustrating the relationship between the window function, applied to uniformly sampled data, and the sampling density function applied during NUWS acquisition [24,25,92]. (c) Comparison of the sensitivity of uniform and NUWS sampling schemes. One-dimensional cross-sections through ^{1}H,^{15}N-SOFAST-HMQC spectra of α-synuclein (5 μM, 277 K, 700 MHz), recorded with identical experiment durations and acquired/processed with cosine-squared sampling densities/window functions respectively. A 23% increase in sensitivity was obtained with the NUWS scheme. Adapted from [26,27,92].

termed non-uniform weighted sampling (NUWS), is an intermediate between uniform sampling and non-linear NUS methods, and data may be processed and their sensitivity analyzed in a conventional manner via the discrete Fourier transform [15–18,92]. However, while some reports have suggested sensitivity increases of 100–200% may be obtained from this approach [5,6,19,93,94], more recently we have compared the sensitivity of uniformed sampled and NUWS spectra on a like-for-like basis, accounting for an intimate connection between the acquisition sampling scheme and the apodization window function, and we find that the 'real world' sensitivity increase is only ca. 10–20% (Fig. 4c) [20–22,92]. Nevertheless, the NUWS approach may be readily implemented and accurately preserves peak intensities and lineshapes [23,92], and so it may yet provide a useful contribution to the study of RNC samples by NMR.

3.3.3. Optimization of transverse relaxation

Up to this point, the discussion of sensitivity has focused on the choice of acquisition time and the apodization and sampling strategies. However, losses due to transverse relaxation are also extremely significant, both during chemical shift encoding periods (Eq. (3)) and also during magnetization transfer delays. This is strongly affected by the choice of nucleus and labeling scheme, but also by the particular R_2 relaxation rates of the various coherences selected by the pulse program being executed. The choice of isotopic labeling schemes has already been discussed above (Section 3.1), and so the discussion here shall focus primarily on the optimal choice of pulse programs.

A large variety of pulse programs are available for the observation of ^{15}N-labeled samples. The majority of published RNC studies have employed the SOFAST HMQC experiment (a longitudinally-optimized variant of the HMQC, discussed further in the following section) [28,29,36,38,39,95]. In principle, the HMQC experiment might be expected to be less sensitive than a sensitivity-enhanced HSQC, due to the factor of $\sqrt{2}$ arising from sensitivity enhancement in the HSQC, and the effect of ^{1}H–^{1}H couplings active during the multiple-quantum evolution period in the HMQC experiment. However, an empirical comparison of their sensitivity, for the proteins ubiquitin (8.6 kDa, 298 K) and SiR-FP18 (18 kDa, 298 K), found that the SOFAST HMQC had a sensitivity approximately 50% greater when compared to longitudinally-optimized HSQC experiments, which was attributed to greater losses in HSQC experiments arising from pulse imperfections and B_1-field inhomogeneities [30,95].

An alternative route to maximizing the sensitivity of ^{15}N-based experiments is through experiments that take advantage of cross-correlated relaxation effects, both to minimize tranverse relaxation rates (the TROSY effect) and to increase the efficiency of magnetization transfer steps (CRINEPT). The TROSY effect, in a ^{1}H–^{15}N spin system, arises from destructive interference between cross-correlated dipolar and chemical shift anisotropy (CSA) relaxation mechanisms [31,32,96]. The sign of the interference effect (constructive/destructive) is dependent upon the ^{1}H and ^{15}N spin-state, but in fully decoupled experiments these spin states are rapidly exchanged, and so their associated relaxation rates are averaged. For small proteins, this loss in sensitivity due to faster transverse relaxation is more than compensated for by the gain in intensity that results from collapsing the multiplet structure. In contrast, for larger proteins it is preferable to avoid decoupling (in one or both dimensions) and instead record spin-state selective spectra, in which only the slowest relaxing coherence is observed. The use of such TROSY methods has permitted high-resolution studies of extremely large systems, such as the 800 kDa GroEL chaperone (τ_c approximately 310 ns at 308 K) [33,97]. Further sensitivity gains may be obtained through the use of CRINEPT polarization transfer elements, which exploit cross-correlation relaxation mechanisms to enhance the transfer of polarization from H_x to H_{x}-N_z. [34,35,98,99]. A number of 2D CRINEPT/TROSY experiments have been proposed, which vary in their combination of polarization transfer steps and spin-state selectivity, reflecting various trade-offs between the length and complexity of the pulse sequence, loss in signal due to spin-state selectivity, and loss in sensitivity and resolution due to non-optimal transverse relaxation [36–40,97,98]. The greatest sensitivity was found for the [^{15}N,^{1}H]-CRINEPT-HMQC-[^{1}H]-TROSY experiment [34,97], and this experiment has been successfully used to observe fast-relaxing folded resonances in an SH3 RNC system [40,41] (discussed further in Section 4). This experiment utilizes CRINEPT transfers but is ^{1}H-decoupled during the indirect evolution period and hence both slow and fast-relaxating spin states are preserved. In contrast, the direct dimension is acquired without decoupling resulting in a doublet structure, but for large molecules, only the slowly-relaxation TROSY coherence is observable. The lack of spin-state selectivity results in a high sensitivity for this experiment, but in more flexible regions of the RNC the anti-TROSY line may not be completely suppressed by its rapid relaxation, and this can result in increased overlap of peaks within crowded spectra [6,40,42].

In the case of ^{13}C-based observations of side-chain methyl groups (particularly favorable for folded nascent chains,

Section 3.1), HMQC sequences are the preferred class of experiment as the sequence is fully optimized for the methyl-TROSY effect. In analogy to dipolar/CSA interference in the ^{15}N TROSY effect, the transverse relaxation rate of some methyl coherences is also greatly decreased by the field-independent cancellation of intra-methyl dipolar fields [43,70]. To maximize the sensitivity gain, pulse sequences must minimize the mixing of these slow and fast relaxating coherences. The HMQC experiment is fully optimal in this regard, and therefore for slowly tumbling systems the sensitivity may be substantially greater than HSQC-based experiments [44–47,100]. Moreover, in contrast to HSQC experiments, Ernst angle excitation may be employed, and a fully longitudinal relaxation optimized experiment, the methyl SOFAST HMQC [48,95,101], has been successfully employed for the observation of RNC side-chain methyl groups [37,38,49].

Finally, although not so far employed in the study of RNCs, fractional deuteration or perdeuteration of samples would be expected to greatly improve the sensitivity in all cases, by elimination of unwanted dipolar relaxation pathways. Therefore, despite the greater expense of sample preparation, we expect that this may be an important avenue to explore in future studies.

3.3.4. Optimization of longitudinal relaxation

As discussed in the previous section, the sensitivity of macromolecular NMR is limited by the rapid tranverse relaxation of the observed signal. However, the slow recovery of longitudinal magnetization also has a significant effect on experimental sensitivity. Indeed, as longtitudinal relaxation is substantially slower than transverse relaxation, the bulk of spectrometer time is allocated to the inter-scan recovery delay, during which no signal is being actively acquired. Since the earliest days of pulsed NMR, the choice of this delay, T_{rec}, and the excitation flip angle, α, has been recognized to affect the sensitivity of the NMR experiment (signal-to-noise per unit time) [50–53,102]. For a fixed excitation angle of 90°, the optimum relaxation delay is $1.26T_1$. Ernst angle excitation, where the scan time is reduced and the excitation flip angle is set to an optimum value, $\cos\alpha = \exp(-T_{rec}/T_1)$, can provide a small additional sensitivity gain in compatible pulse programs, such as the HMQC. This may be of the order of 10% and increases as the scan time is decreased, reflecting a net gain in SNR from the accumulation of many individually noisy scans. Ultimately, the repetition rate is limited either by the combined length of the pulse sequence and the acquisition time, or by restrictions on the spectrometer duty cycle.

A more recent trend in pulse sequence design has been the acceleration of longitudinal relaxation through the use of selective pulses, wherein only the observed subset of spins are excited [31,32,95,103]. This preserves a bath of cold spins, the effect of which is to accelerate the longitudinal relaxation of the observed spins through cross-relaxation pathways. The return to equilibrium is therefore described by an effective relaxation time, T_{1eff}, which may be significantly shorter than the T_1. For example, in the case of ubiquitin a reduction in T_{1eff} from 0.9–1.4 s to 0.3–0.4 s was observed upon selective excitation of amide protons [54,55,95]. In larger, fractionally deuterated or perdeuterated molecules however, many fewer spins are present to act as a heat sink, reducing the effectiveness of band-selective methods, but this may be at least partially counterbalanced by the increased rate of cross-relaxation within slowly tumbling molecules, again resulting in useful sensitivity gains [36,101].

In the past several years, a variety of longitudinally-optimized experiments utilizing band-selective excitation have been described, including the LHSQC [53,103], SOFAST-HMQC [56,95,104], methyl-SOFAST-HMQC [57–59,101], the BEST family of 3D experiments [31,105,106] and BEST TROSY experiments [60,107,108]. The improvement in sensitivity in these experiments

is dependent upon the sample and experimental conditions, but gains of a factor of 1.5–2.5 are commonly obtained. In particular, the increased sensitivity of both ^{15}N and ^{13}C SOFAST-HMQC experiments has been instrumental in a number of RNC studies [36–39,61]. The more recently described BEST TROSY experiments [61,107] may offer advantages for future studies, however, due to the decreased transverse relaxation rate (as discussed above) but also to an additional strategy for the enhancement of longitudinal polarization [62,108]. TROSY experiments, due to their spin-state selectivity, are able to merge the initial polarization of both protons and nitrogens resulting in a small sensitivity gain of 10–20% [62,63,109]. Additionally, H_z magnetization arising from longitudinal relaxation during the chemical shift encoding delay is converted to $-N_z$ by the ST2-PT transfer element [64,109]. This can be inverted to N_z following the acquisition period, and in BEST TROSY experiments where the inter-scan delay is short this additional polarization may survive and contribute substantially to the observed signal [65,108]. However, as the recovered polarization arises from relaxation during the chemical shift encoding delay, the sensitivity gain will not be constant, but rather will be manifested indirectly as a reduced apparent R_2 relaxation rate in the indirect dimension. As discussed earlier, and according to Eq. (2), this will result in increased SNR in the resulting spectrum.

Finally, an alternative approach to longitudinal relaxation enhancement is the doping of the sample with a paramagnetic ion or complex. Gd^{3+} complexes were originally used to accelerate ^{13}C longitudinal relaxation in directly-detected experiments [54,55,110], and later to reduce the T_1 of solvent water protons, resulting in sensitivity gains for labile amides [55,111]. However, proton transverse relaxation rates are also greatly increased in the presence of Gd^{3+}, resulting in line broadening and decreased sensitivity. A solution to this problem was found in the use of paramagnetic species such as Ni^{2+} with very short electronic relaxation times, under which conditions the effect on R_1 and R_2 rates is identical. As $R_1 \ll R_2$ in the macromolecular limit, such paramagnetic species may have a substantial effect on longitudinal relaxation while affecting resonance linewidths only marginally [54,112]. Moreover, this approach may be combined with band selective excitation and rapid recycling experiments such as the SOFAST-HMQC, as discussed above, and in this way sensitivity improvements of a factor of ca. 1.5–2 have been reported [36,113]. While this approach has not yet applied to the study of RNCs (and indeed the compatibility of paramagnetic agents with the ribosome would need to be carefully assessed), the potential increase in sensitivity is highly attractive.

3.4. Application and optimization of NMR diffusion measurements

A critical measurement in the study of RNCs by NMR is the continual assessment of the diffusion coefficient, D, of the species being observed. NMR diffusion measurements were first employed in high-resolution NMR studies of the ribosomal L7/L12 stalk complex, to test and confirm that the observed species did in fact arise from ribosome-bound molecules rather than from released or degraded protein (Fig. 2e) [54,57], and the use of such measurements has remained essential in studies of RNCs [36,38,54,55]. At the heart of this approach is the difference in hydrodynamic radius, r_h, between small proteins in free solution, typically 2–3 nm, and the 2.3 MDa 70S ribosome, 12.6 nm [55,114]. Given the Stokes–Einstein relation, $r_h = k_BT/6\pi\eta D$, where k_B is the Boltzmann constant, T is the temperature, and η is the solution viscosity, this is equivalent to an approximately 4-fold difference in the translational diffusion coefficient, and this may be readily distinguished using pulsed-field gradient stimulated-echo (PFG-STE) NMR diffusion measurements [54,115].

The use of NMR for the study of translational diffusion is a rich field, with diverse biomolecular applications from the study of protein conformation [60,116] and protein–ligand interactions [66–69,117,118] to high-resolution in-cell NMR spectroscopy [36,38–40,119], and a number of excellent books are available describing the subject [37,120,121]. Here, we will focus on three aspects of NMR PFG diffusion measurements central to the study of RNCs: the ability to discriminate released material from ribosome associated nascent chains; selectivity for nascent chain resonances; and the sensitivity of the experiments.

The signal intensity in a PFG-STE experiment, $I(G)$, is described as a function of the gradient strength, G, by the Stejskal–Tanner equation [70,115]:

$$I(G) = I(0) \exp(-\gamma^2 G^2 \delta^2 \sigma^2 \Delta D) \qquad (4)$$

where γ is the gyromagnetic ratio of the nucleus on which position is encoded, δ is the length of the encoding and decoding gradient pulses, σ is the shape factor of the gradient pulses, and Δ is the length of the diffusion delay between encoding and decoding gradients. Corrections to Eq. (4) for finite gradient pulse lengths and bipolar gradients have been omitted, as these may vary with the pulse program being employed [71,72,115,122–124].

The lifetime of an RNC sample is limited both by the rate of nascent chain release from the ribosome, and by the degradation of the ribosome itself. In our experience, nascent chain release is often more rapid than ribosome degradation, and so while relatively strong resonances may be observed from the L7/L12 stalk complex, ribosome diffusion is not a sufficiently reliable probe to verify nascent chain attachment (although observation of ribosome breakdown is certainly indicative of nascent chain release). Therefore, to fully assess the sample integrity it is essential to measure the diffusion of the nascent chain specifically. This is greatly facilitated by the use of isotope-edited or heteronuclear diffusion experiments, which effectively eliminate unlabeled ribosomal resonances.

In the case of ^{15}N-labeled RNC samples, the current experiment of choice is the ^{15}N-XSTE experiment [36–38,122]. In this experiment, spatial encoding/decoding occurs on ^{1}H spins during 5.4 ms INEPT transfer periods ($1/2^1J_{NH}$), which sandwich a ^{15}N stimulated echo. Thus, the gradient encoding occurs on high-γ ^{1}H spins and long gradient pulse lengths δ may be employed during the INEPT delays which, according to Eq. (4), results in high sensitivity to small displacements, essential for the accurate characterization of slow moving ribosomal species. Storage as N_z in the stimulated echo clearly provides isotope editing, but also takes advantage of the longer ^{15}N T_1 to reduce longitudinal relaxation losses during the diffusion delay Δ, increasing the sensitivity of the experiment. An additional advantage, when compared with alternative ^{15}N-edited ^{1}H-stimulated echo experiments [73,74,125,126], is that the XSTE is not sensitive to amide hydrogen exchange during long diffusion delays. This effect is most significant for unfolded nascent chains, and may reduce the sensitivity or bias the measurement towards larger diffusion coefficients. Finally, although not so far employed in the study of RNCs, the use of longitudinal relaxation optimized experiments such as BEST-XSTE [76,77,123] or SORDID [31,124] may be of future interest in enhancing the sensitivity of XSTE experiments.

The diffusion of ^{13}C-labeled RNC samples may also be measured using a ^{13}C-XSTE experiment [78,122]. However, due to the larger $^1J_{CH}$ coupling constant the INEPT transfer delay is significantly shorter than in the ^{15}N experiment (3.5 ms compared to 5.4 ms), and this limits the maximum gradient pulse length, δ, to ca. 3 ms. Therefore, to maintain the ability to discriminate slowly diffusing species, large increases in the diffusion delay Δ may be required, with an associated loss in sensitivity due to increased

longitudinal relaxation. As an alternative, a ^{1}H-STE-^{1}H, ^{13}C-HMQC experiment may be acquired [79,127]. This experiment has a number of advantages over the ^{13}C-XSTE. Firstly, there is no restriction on the gradient pulse length, δ, as spatial encoding/decoding does not occur during INEPT transfers, resulting in greater sensitivity to small diffusion coefficients. In contrast to amides, methyl protons are non-exchangeable and therefore no loss of sensitivity or accuracy arises from the use of ^{1}H stimulated echoes. Finally, while losses to longitudinal relaxation are somewhat greater for the ^{1}H-STE compared to the ^{13}C-XSTE, transverse relaxation is significantly slower in the ^{1}H-STE-HMQC experiment, as it is optimized for the methyl-TROSY effect with minimal mixing of slow and fast relaxing coherences [39,40,70,127]. Therefore, particularly for the broad resonances associated with folded RNC species for which ^{13}C-based observation is most favorable, the ^{1}H-STE-HMQC experiment may be expected to have significantly greater sensitivity than the ^{13}C-XSTE.

4. Observations of co-translational folding using NMR spectroscopy

In this section, we present a systematic review of the five NMR investigations of RNCs in the literature at the time of writing. The first observation of a ribosome nascent-chain complex by NMR was reported in 2007 by Christodoulou and colleagues [36,38]. The gelation factor ddFLN (also known as ABP-120), from *Dictyostelium discoideum*, was selected as a model system [80,128]. This is a tandem repeat protein comprising an N-terminal actin-binding domain, followed by five highly conserved filamin domains having an immunoglobulin fold [38,40,129–131]. To model the progressive emergence and co-translational folding of the full-length protein from the ribosome, a fragment was designed comprising the two C-terminal domains, ddFLN5 and ddFLN6, minus the C-terminal beta-strand of ddFLN6. A cell-free transcription–translation system was employed using ^{13}C,^{15}N-labeled amino acids and a DNA construct lacking a stop codon to create a homogeneous population of stalled ribosomes, which were subsequently purified by sucrose gradient ultracentrifugation. NMR spectra were acquired using the ^{15}N-SOFAST-HMQC experiment to maximize the experimental sensitivity [80,95], and the ^{15}N-XSTE experiment [81,82,122] was used to measure the diffusion of the nascent chain and verify its attachment to the ribosome. Following acquisition, samples were incubated with 1 mM puromycin to induce release of stalled nascent chains, and the 62% increase in signal intensity that was subsequently observed provided further confirmation of the ribosome attachment. Finally, in addition to these spectroscopic measures, separate radiolabeled samples were prepared using [^{35}S]-methionine and incubated alongside NMR samples to provide an orthogonal biochemical probe of sample integrity during the measurement timecourse.

The 2D HMQC spectra obtained for these RNC samples were analyzed by comparison with spectra of the released two domain construct, and the single domain ddFLN5, in the presence and absence of the ribosome particle. The RNC spectrum contained a number of poorly dispersed and highly overlapped resonances with proton chemical shifts between 8.2 and 8.6 ppm, characteristic of disordered states, but also a number of very low intensity dispersed resonances with proton chemical shifts between 6.5 and 10 ppm. These resonances overlapped closely with the spectrum of isolated ddFLN5, with only very small chemical shift changes of ca. 0.18 ± 0.17 ppm. This important result provided, for the first time, direct evidence that folded structures can be populated co-translationally, and that the folded structure is not significantly perturbed by the presence of the ribosome. The poorly dispersed resonances also overlaid well with spectra of the full-length

construct in isolation, allowing them to be attributed to the ddFLN6 linker, which remained disordered both in isolation (due to the absence of the terminal beta-strand) and in the RNC, indicating that the presence of the ribosome did not strongly alter the folding equilibrium. An analysis of the ^1H linewidths of cross-peaks determined average linewidths of 59 ± 26 Hz in the RNC, compared with 41 ± 11 Hz for the isolated construct in the presence of ribosomes, and some evidence was obtained for site-selective broadening in the N-terminal hemisphere of ddFLN5. Together, these observations indicate that the tumbling of the nascent ddFLN5 domain is dynamically decoupled from the tumbling of the ribosome through the flexibility of the disordered ddFLN6 linker region, but that there may still be weak and transient interactions between the nascent chain and the ribosome surface.

A second study of the same nascent chain system investigated the use of methyl groups to probe the structure and dynamics of side-chains within the nascent chain [37,83]. As discussed earlier, methyl groups in general have intrinsically higher sensitivity when compared to amides, and in this case the SNR was approximately 2.5 times greater for methyl resonances in a $^1H,^{13}C$-SOFAST-HMQC relative to amide resonances in a $^1H,^{15}N$-SOFAST-HMQC acquired under the same conditions. For very dilute systems such as RNCs, this is a highly significant gain in sensitivity. The observed RNC spectrum contained several well-resolved, dispersed resonances which overlaid closely with those of isolated ddFLN5, with extremely small proton chemical shift changes of 0.003 ± 0.014 ppm, again indicating that the nascent chain populates a folded state

with a structure very similar to that observed for the free protein. As was the case in the $^1H,^{15}N$ spectrum, differential line broadening of methyl resonances was also observed, and following incubation with puromycin to induce release of the nascent chain, these peaks were found to increase in intensity. However, the most highly broadened methyl resonances were located in the core of the domain, in contrast to broadened amide resonances which were clustered in the N-terminal hemisphere of the domain. The reason for this discrepancy has not yet been fully rationalized, though further analysis with careful statistical consideration of measurements at very low SNR, will permit a more complete understanding of line broadening and dynamics within this RNC system.

The two studies discussed above used cell-free translation systems to produce the RNC samples. While this is a powerful and highly flexible approach allowing control over all aspects of isotopic labeling, the cost of regularly producing samples on the scale required for NMR is prohibitive. This stimulated the development of alternative in vivo approaches [38,39,84], using the 17 amino acid translational arrest motif derived from the E. coli secretion monitor protein, SecM (Section 3.2) [76,132,133], with subsequent purification by affinity chromatography and ultracentrifugation.

The first of the above studies [38] continued the characterization of the co-translational folding of the ddFLN protein. Two constructs were prepared comprising either ddFLN5 + 6 or the single ddFLN5 domain, fused to a C-terminal SecM motif and an N-terminal hexa-His tag. By passaging the cells selectively into a
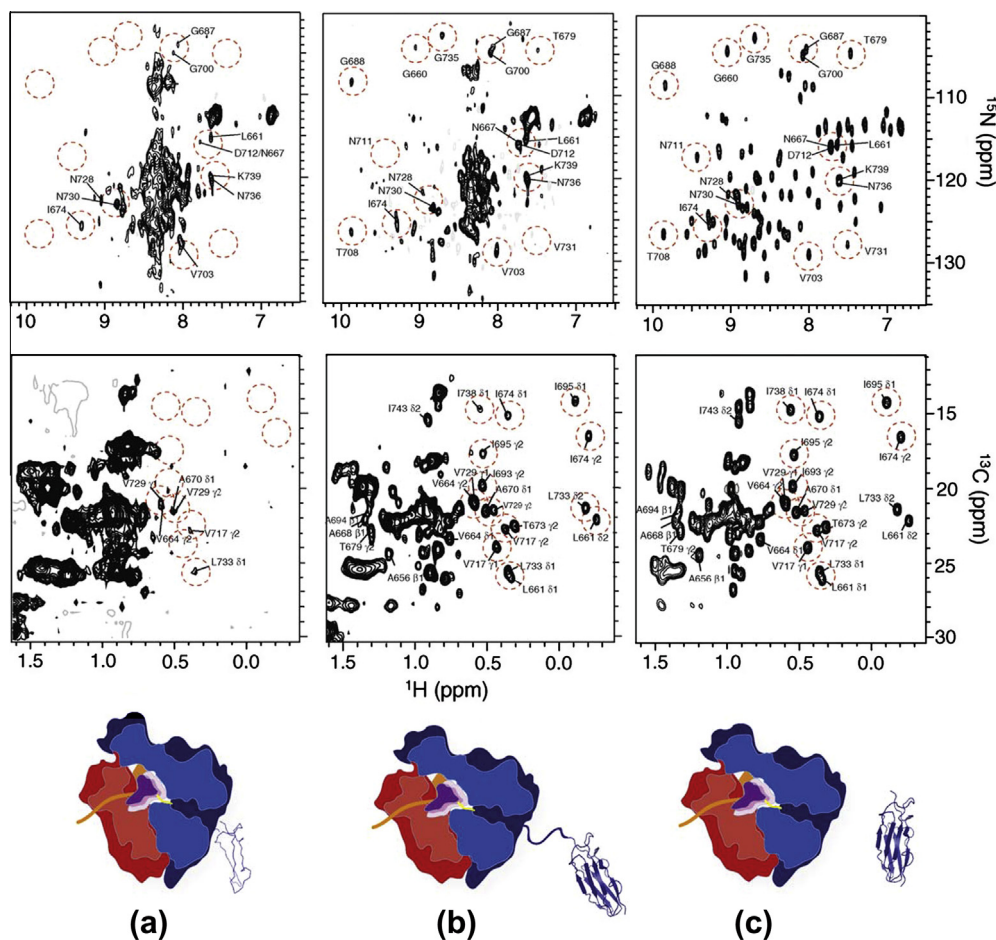
Fig. 5. Observing the progressive folding of a ddFLN5 nascent chain by NMR. $^1H–^{15}N$ SOFAST-HMQC (top) and $^1H–^{13}C$ HMQC (bottom) spectra of (a) a folding-incompetent ddFLN5 RNC, in which the C-terminus of ddFLN5 is 21 aa from the PTC; (b) a folding-competent ddFLN5 + 6 RNC, in which an 89 residue fragment of ddFLN6 results in a total linker length of 110 aa; and (c) isolated ddFLN5. Spectra were acquired at 700 MHz, 298 K. Reprinted from [6], data taken from [38].

rich isotopically labeled medium for protein expression, ^{13}C/^{15}N-labeled nascent chains could be produced specifically against a background of isotopically silent ribosomes. A two-stage purification was employed selecting on the basis of molecular weight (sucrose gradient ultracentrifugation) and for the presence of the N-terminal hexa-His tag (metal affinity chromatography), to obtain a highly homogeneous population of stalled ribosomes carrying selectively labeled nascent chains. As previously, both ^{1}H,^{13}C and ^{1}H,^{15}N-SOFAST-HMQC spectra were acquired (Fig. 5). ^{15}N-XSTE diffusion experiments were recorded to monitor the attachment of the nacent chain to the ribosome.

The comparison of the two RNC constructs provides a preliminary exploration of the dependence of co-translational protein folding on the length of the translated polypeptide, and in particular, the distance of the domain from the peptidyl-transferase center (PTC). In the shorter construct, the C-terminus of ddFLN5 is 21 residues from the PTC, while in the longer construct, the C-terminus of the domain is 110 residues from the PTC. By comparison, the ribosome exit tunnel encloses approximately 30 residues, and so in both constructs the terminal beta strand of ddFLN5 or ddFLN6 is not expected to be accessible for folding. In accordance with this, NMR spectra of the longer ddFLN5 + 110 RNC contained well-dispersed cross-peaks overlaying closely with the spectra of isolated ddFLN5 (Fig. 5b and c), in addition to a cluster of poorly resolved amide resonances with proton chemical shifts of 8–8.5 ppm, which were attributed to a disordered and flexible ddFLN6 linker. NMR spectra of the shorter ddFLN5 + 21 RNC were also typical of a disordered state (Fig. 5a), and cross-peaks in the ^{1}H,^{15}N HMQC spectrum overlaid well with a spectrum of isolated ddFLN5 denatured in 8 M urea.

In a parallel development and application of an *in vivo* sample preparation approach, the co-translational folding of the protein barnase was investigated by NMR [39]. The folding mechanism of barnase has been exceptionally well characterized in the isolated protein [134–136], and so NMR studies of co-translational folding in RNCs may be expected to yield a very interesting and informative comparison. Again, a SecM sequence was employed to create a stalled population of ribosomes, which were subsequently purified using a tobacco etch virus (TEV) protease-cleavable N-terminal Strep tag. A 50 residue sequence from the protein RpoB was incorporated between the barnase domain and SecM fragment to create a linker region, extending a total of 86 residues from the PTC, with sufficient flexibility to allow independent tumbling of the barnase domain and therefore observation by solution-state NMR. ^{1}H, ^{15}N-SOFAST-HMQC spectra of the RNC were acquired, with excellent SNR, and 100 dispersed cross-peaks could be identified. Of these, 72 could be assigned by comparison with chemical shifts deposited in the Biological Magnetic Resonance Data Bank (BMRB) [137], and only small chemical shift changes were observed in the RNC (^{1}H and ^{15}N standard deviations of 0.04 ppm and 0.4 ppm respectively) indicating that the nascent chain is able to fold co-translationally to a highly native-like structure.

The final study to be discussed here is an investigation of the length-dependent folding of the SH3 domain from α-spectrin [40]. This is a small and moderately stable domain (63 residues, ΔG_{F-U} = 3.9 kcal mol^{-1} [138]) comprising five β-strands arranged in an orthogonal β-sandwich. The domain was previously studied as an RNC using biochemical methods, where protection of the nascent chain from proteinase K digestion indicated the formation of stable tertiary structure [139]. For the NMR study, an *in vivo* approach based on a 27 residue SecM-derived stalling motif was used to produce ^{13}C/^{15}N isotopically labeled RNC samples, using a cleavable N-terminal 3xStrepII-Smt3 tag as the basis for purification by affinity chromatography. A TEV protease cleavage site was also inserted at the C-terminus of the SH3 domain, upstream of a short linker sequence and the SecM motif, so in total the SH3 domain

was 50 residues from the PTC. The integrity of the nascent chain was assessed by the progressive emergence of cross-peaks in the 2D spectra on timescales of >48 h, indicative of nascent chain release; and also by ^{14}N and ^{15}N-edited diffusion measurements, to characterize the hydrodynamic radii of both the nascent chain and the unlabeled ribosomes. Finally, an increase in the nascent chain diffusion coefficient was observed following TEV cleavage of the RNC, which provided an additional proof of ribosomal attachment. 2D NMR spectra were acquired using longitudinal relaxation optimized [^{1}H,^{15}N]-CRINEPT-HMQC-[^{1}H]-TROSY (discussed in Section 3.3) and ^{1}H,^{13}C-HMQC experiments. The transverse relaxation optimized CRINEPT methodology is likely to have provided a useful increase in sensitivity, but there is also an increase in spectral overlap arising from the anti-TROSY line, particularly for more slowly-relaxing disordered resonances. Interestingly, a comparison of the relative intensities of the TROSY and anti-TROSY lines may in the future provide a simple, semi-quantitative probe of local rotational dynamics in the nascent chain complex.

2D NMR spectra of the SH3 RNC showed 22 amide and 23 side-chain resonances, distributed throughout the domain sequence, with chemical shifts very similar to those of the isolated domain ($\Delta\delta_{HN}$ 0.040 ± 0.005 ppm, $\Delta\delta_{HC}$ 0.007 ± 0.002 ppm); other resonances could not be distinguished due to overlap with ribosomal background signals. As was found for ddFLN5 and barnase, this indicates that the domain is able to fold co-translationally to a structure closely resembling that found in free solution. An analysis of amide proton linewidths determined average linewidths of 36 ± 2 Hz for the RNC, compared to 28 ± 1 Hz upon release and 23 ± 1 Hz for the free protein. This ca. 10 Hz increase in linewidth was compared to linewidths of over 1000 Hz that are expected for the 70S ribosome, indicating that less than 1% of the nascent chain is bound to the ribosome surface at a given moment (assuming fast exchange between ribosome-surface-bound and non-ribosome-surface-bound forms) and therefore that the ribosome is unlikely to strongly modulate the folding landscape of this protein through interactions with its folded state.

To compare the effect of ribosomal attachment on the (unobserved) unfolded state of the SH3 RNC, a protein engineering strategy was employed, introducing an 'm10' variant containing two destabilizing mutations that result in complete unfolding of the domain [140]. [^{1}H,^{15}N]-CRINEPT-HMQC spectra of the m10 RNC showed limited proton chemical shift dispersion, characteristic of a disordered state, and which overlaid closely with the free m10 protein, with negligible chemical shift changes. This indicates that the conformational sampling of the unfolded polypeptide is not significantly perturbed by the presence of the ribosome. Thus, the ribosome appears to be essentially inert to both the folded and unfolded states of the protein.

Finally, to investigate the length dependence of folding, in analogy to the study of ddFLN [38] two 'folding snapshots' were created in which 5 and 22 residues were truncated from the C-terminus of the domain. This resulted in effective linker lengths of 26 and 9 residues respectively from the PTC (compared to a linker length of 50 residues in the folded wild-type and unfolded m10 RNCs), and given that ca. 30 residues within an extended linker are expected to be buried by the ribosomal exit tunnel, these lengths correspond to the emergence of two or four of the five β-strands comprising the SH3 domain. 2D NMR spectra of the RNCs were acquired, and in both cases the spectra exhibited very limited chemical shift dispersion, indicating that the nascent chains remained disordered with no formation of stable tertiary structure. From this it was concluded that complete emergence of the SH3 domain is required before co-translational folding may occur. However, it should perhaps be noted that due to the 31 residue SecM motif employed in this study, the C-terminal residues were completely absent from

the RNC construct, rather than being occluded by the exit tunnel, and that the truncated domains also failed to fold in free solution. Therefore, at this stage there is some difficulty in identifying what effect, if any, the presence of the ribosome has upon the folding of these truncations, but this will undoubtedly be an interesting question for future investigations to address.

5. Future prospects

5.1. Mapping the co-translational protein folding landscape

The NMR studies of RNCs discussed above have begun to address a very basic question: at what distance from the PTC can the formation of stable tertiary structure first occur? The answer to this question will depend on the structure adopted by the nascent chain within the ribosome exit tunnel [141,142]. The tunnel is 80–100 Å in length (depending on the definition of the exit point) and 10–20 Å wide [143], hence for a polypeptide in a fully extended conformation approximately 25 residues are expected to be occluded by the ribosome, while if a fully α-helical conformation were adopted as many as 50–60 residues might be occluded. We note however that while compact α-helical structure has been observed within the exit tunnel by cryoelectron microscopy, in general only short α-helical segments have been reported [31,77] and therefore in reality the figure of 50–60 residues is likely overestimated. In particular, the stalled SecM nascent chain has been observed to have an extended conformation, which is required in order to make a number of specific interactions with the inner surface of the exit tunnel [32].

Many experimental techniques have also been applied to characterize the length dependence of co-translational folding, such as the intrinsic fluorescence of the β-barrel protein eCFP, where a progressive increase in fluorescence was observed on increasing the linker length from 25 to 55 amino acids (measured from the PTC to the first structured residue of the β-barrel) [22]. Existing NMR studies have surveyed a wide range of linker lengths that separate folding competent domains from the PTC: 21 and 110 aa for ddFLN5 [38], 86 aa for barnase [39], and 9, 22 and 50 aa for SH3 [40]. While each protein system will clearly have its own peculiarities, it is nevertheless informative to compare the folding at the various lengths observed in these studies (Fig. 6). In no case is non-native structure observed (as probed by analysis of chemical shifts), and a folding transition is found to occur between 22 and 50 amino acids from the PTC, coincident with the emergence of the domains from the ribosome exit tunnel. At the resolution of current studies, however, it is not possible to determine the transition point more precisely, and therefore at present the NMR data are compatible with either extended or α-helical structure within the ribosome exit tunnel.

NMR is uniquely well-suited to characterizing the onset of folding, and to the broader enterprise of mapping the entire protein folding landscape as a function of chain length – the 'co-translational protein folding landscape' or folding funnel [41]. The measurement of chemical shift values gives a very direct probe of structure, allowing the identification of native structure, and the characterization of intermediate states. Moreover, as NMR is sensitive to folding kinetics through the effects of chemical exchange, such phenomena may in the future also allow the analysis of the kinetic barriers between folded, unfolded and intermediate states.

A complementary approach to the mapping of the co-translational folding landscape using stalled RNCs is through the analysis of isolated C-terminal truncations, mimicking the progressive emergence of domains from the ribosome exit tunnel. A number of systems have been studied in this way, using a variety of biophysical and spectroscopic methods including NMR together with circular dichroism (CD) spectroscopy, intrinsic fluorescence, and dye binding assay analyses [145–149]. Interestingly, a review of several proteins including CI2, barnase, staphylococcal nuclease and apomyoglobin observed that folding of fragments generally only occurred when almost the full chain length is present [150] (although the analysis of apomyoglobin fragments is complicated by their tendency to oligomerise [151]). For example, CI2, a small (64 amino acid) α + β domain that in isolation folds in a two-state manner, was found to populate a molten-globule-like state at intermediate lengths, and the final transition to its folded structure only occurred in the presence of the final two residues [146]. Changes in Hα chemical shift values were used to map the folding pathway on the structure according to the parameter $\Phi^{NMR} = (\delta_{observed} - \delta_{random\ coil})/(\delta_{native} - \delta_{random\ coil})$, where random coil chemical shift values were predicted from the local sequence [147]. This parameter, generally ranging between 0 and 1, is conceptually analogous to Φ-values used in the study of protein folding by protein engineering [152], and provides a simple measure of the structure of a residue in a given state relative to the native state. Its analysis in the case of CI2 determined that key stabilizing interactions present in the intermediate state were similar to those formed in the folding nucleus of the equilibrium folding pathway [147]. Given its simplicity, this parameter may in the future be a useful tool for the analysis of intermediate states observed within RNC samples.

The study of co-translational protein folding landscapes using series of isolated truncations provides an excellent starting point for systematic investigations of RNCs. Of particular interest will be the mechanisms by which the ribosome might perturb the landscape, as well as the magnitude of any effect (Section 5.2). Ultimately, we would like to understand folding as it occurs within the native environment of the cell [19], and the further perturbations to the folding landscape that might arise from macromolecular crowding effects and the presence of molecular chaperones (Section 5.3). The study of protein folding *in vivo* is being approached by many biophysical methods [21,153,154], but the emerging interest and development of methods for high-resolution NMR within living cells – 'in-cell NMR' – suggests that NMR spectroscopy also has the potential to make important contributions in this area [119,155–159].

5.2. Protein–ribosome interactions

Understanding the co-translational protein folding landscape, and in particular how it might be modulated by interactions with the ribosome itself, is crucial to understanding protein folding as it occurs within the cell. Preferential interactions with folded or unfolded states would provide a means by which the ribosome selectively stabilizes these states, in doing so affecting the position of the folding equilibrium, as well as the folding kinetics. Force microscopy studies of stalled ribosome bound nascent chain complexes have begun to investigate the effect of the ribosomal
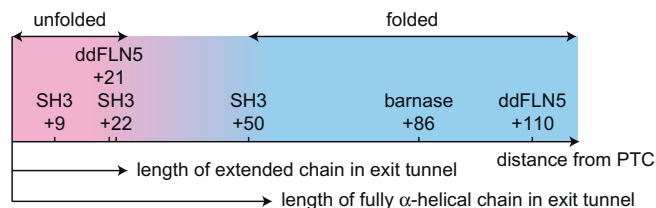
Fig. 6. Summary of linker lengths for folded and unfolded RNCs observed by NMR spectroscopy [38,40,144], compared with the estimated number of residues enclosed by the ribosome exit tunnel for maximally extended and α-helical conformations [141–143].

environment on the protein folding processes, and in the case of T4 lysozyme observed that the presence of the ribosome *decreased* the rate of folding [34]. This could be attributed to interactions of folded and unfolded states with the ribosome surface, at least part electrostatic in nature, acting to stabilize the states and so resulting in a larger free energy barrier between the states.

Solution-state NMR is exquisitely sensitive to interactions of proteins with the ribosome surface. Given the extremely slow tumbling of the ribosome, proton linewidths of the order of 1000–10,000 Hz could be expected for nascent chains interacting strongly with the ribosome, which is several orders of magnitude larger than observed for a small protein in isolation. This large
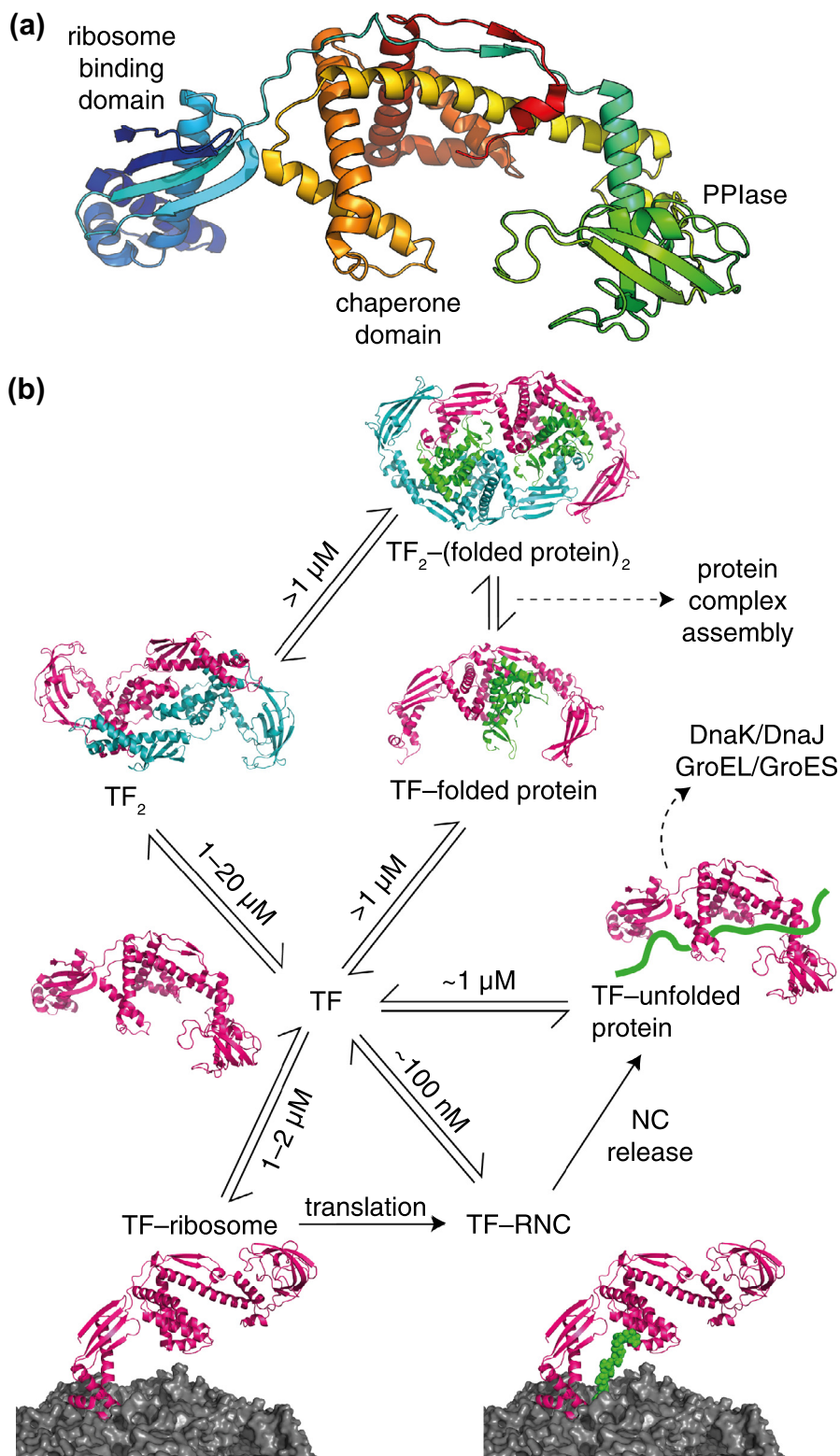
Fig. 7. (a) Crystal structure and domain structure of *E. coli* trigger factor monomer [164] shown in 'chainbow' coloring (blue at N-terminal to red at C-terminal). (b) Schematic illustration of trigger factor binding equilibria (color code: TF chain(s), magenta and cyan; translated protein, green; ribosome, gray).

difference in tranverse relaxation rates has been used to analyse the interaction of antibiotics with the ribosome [160]. In the context of RNCs, changes in resonance linewidths have been used to estimate the fraction of nascent chain interacting with the ribosome surface at any given instant, and in the case of the SH3 RNC discussed above, not more than 1% of folded or unfolded nascent chains were determined to be interacting with the ribosome [40]. Given the high sensitivity to low fractions of bound material, and as peak linewidths are readily measured even from noisy data, we expect that NMR will prove to be a useful tool for the further study of these interactions.

5.3. Interaction of nascent chains with trigger factor

Trigger factor (TF) is a 48 kDa bacterial protein (Fig. 7a) that is the only molecular chaperone known to interact directly with the 70S ribosome [161–165]. It is ATP-independent, and present in the cytosol at a concentration of ca. 50 µM, in significant molar excess to the ribosome itself, which has an intracellular concentration of ca. 20 µM [166]. As such, it is most probably the first chaperone encountered by a nascent chain on emerging from the ribosome exit tunnel. Deletion of TF can be tolerated, although levels of the complementary chaperone DnaK are upregulated and the fraction of nascent chains interacting with DnaK increases from 15% to 40%. However, the combined deletion of TF and DnaK is lethal above 30 °C [167–169]. TF is involved in several binding equilibria (Fig. 7b): the free protein exists as a homodimer, with reported K_d values between 1 and 20 µM, but interacts with the ribosome in its monomeric form with a K_d of 1–2 µM and a residence time of ca. 10 s [170–172]. The presence of a nascent chain increases the ribosome binding affinity by an order of magnitude to ca. 100 nM, modulating both the association and dissociation rates of the complex [144,172]. TF has highly efficient peptidyl prolyl isomerase (PPIase) activity [173], but the primary function is thought to be as a 'holdase', interacting strongly with disordered hydrophobic and/or basic sequences to inhibit premature folding into misfolded states [27,144,172]. Any physiological function of the TF dimer itself is unclear at the present date, although both monomeric and dimeric TF have been found to interact with a number of folded proteins, including the ribosomal S7 protein, suggesting that TF may have a secondary role in assisting the assembly of protein complexes [174].

The TF monomer comprises three independently folding domains (Fig. 7a). The structure of the molecule has been determined crystallographically [164], but some NMR characterizations of domains have also been reported [175–177]. Overall, the structure forms a cradle over the ribosome exit tunnel, binding via the N-terminal domain to the L23 and L29 proteins and 23S rRNA of the large subunit [164,165]. This domain is also required for dimerization, and a 26 kDa homodimer of the NTD (TF$_{1-117}$) has been studied by NMR and the backbone and side-chain chemical shifts fully assigned [178]. The central peptidyl prolyl isomerase (PPIase) domain (TF$_{151-251}$) has an FKBP fold that was first determined by NMR, and an analysis of ps–ns backbone dynamics found that many residues were in exchange on a µs–ms timescale as had previously been observed for FKBP itself, indicating that dynamics were conserved in addition to structure [176]. The final domain is the C-terminal chaperone domain, which is implicated in substrate binding, although protein substrates may in fact bind to a variety of sites across the entire inner surface of the protein [164,179]. The structure of this domain has been solved crystallographically [164,180], but an NMR characterization was undertaken to resolve differences in subdomain orientation between these crystal structures [177]. However, observed residual dipolar couplings were not consistent with either structure. An analysis of ^{15}N relaxation indicated the domain has substantial flexibility, sug-

gesting that the conformation observed by RDCs may be an average of multiple conformations, arising from hinge motions of subdomains. Certainly, however, this provides a clear example where crystal structures do not represent the major conformation found in solution [177].

This brief summary has only scratched the surface of a fascinating chaperone system. Clearly, a series of careful and elegant biochemical and kinetic studies have yielded great insights into the chaperone mechanism of TF. However, a structural understanding of the interaction between the chaperone and its protein substrates is currently absent, and NMR could be a powerful source of information in this regard. From an NMR perspective TF is a complex system to study given the large molecular weight of TF and the ribosome, and the limited concentration of RNCs, but at a basic level the analysis of chemical shift changes and line broadenings in the presence of TF appears to be a reasonably feasible goal. The complete analysis of such data, however, will require the unraveling of a challenging system of coupled equilibria – dimerization, ribosome binding, and substrate binding – but the eventual results will surely be well worth the effort.

5.4. Interfacing with molecular dynamics simulations

The direct simulation of proteins using molecular dynamics (MD) methods offers, with the appropriate analysis, validation and visualization, deep insights into protein dynamics, folding and function [181]. While MD has historically been applicable to very simple systems over short timescales, the exponential growth in computational power means that the simulation of complex systems over biologically relevant timescales is becoming an increasingly realistic prospect [182]. At the present time, it is feasible to run all-atom simulations of small protein domains (~10^4 atoms) on timescales of up to a few ms [183–185], while simulations of the 70S ribosome (3×10^6 atoms) have provided a glimpse of dynamics on timescales of 100 ns [186]. Indeed, the rate of improvement in biomolecular MD has outpaced Moore's Law [182], due in part to the continual improvement in algorithm design, but also to the development of specialized hardware dedicated to the evaluation of molecular potentials [183,187]. At the current rate of progress, by 2030 it may become feasible to conduct all-atom simulations of the ribosome on the biologically important millisecond timescale [182].

In the interim, a variety of simplified, coarse-grained simulation methods have been developed in order to model large and complex systems such as the ribosome and the process of co-translational protein folding [188]. Elcock describes a Gō-type model [189] of the coupled protein synthesis and folding of CI2, barnase and Semliki forest virus protein (SFVP) [190]. The energetics of the model were calibrated using the experimental studies of C-terminal truncations of CI2 and barnase discussed above (Section 5.1), and the simulation results closely matched both the experimentally observed length dependence of folding, and the increased co-operativity of thermal unfolding observed near the full chain length. This validated simulation method was then used to explore the coupling of synthesis to folding. The ribosome was not observed to promote folding relative to the isolated truncations, and the folding pathways were not affected by the presence of the ribosome. In the case of the two-domain protein SFVP however, significant differences were observed in the *de novo* folding of the domain compared to the reversible folding in free solution: in particular, the emerging C-terminal domain was found to accrete onto the folded N-terminal domain. A particular virtue of MD methods is that simulation parameters may be varied to non-physical values to explore the properties of the system in more detail. By artificially increasing the stability of the CI2 domain in this manner, partially folded intermediate states were observed to become trapped
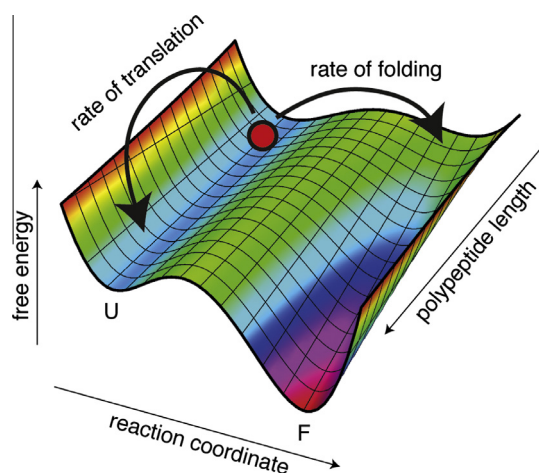
within the ribosome exit tunnel, suggesting that delaying the folding of domains until the C-terminus has emerged may confer evolutionary advantages [190].

Further molecular dynamics studies, also using coarse-grained Gō-type models, have continued to investigate the length dependence of co-translational folding. Based solely on steric constraints, it was observed that native and non-native elements of tertiary structure can readily form within the ribosome vestibule (the final 20 Å of the 100 Å long ribosome exit tunnel) [191]. By analysis of the temperature dependence of RNC folding, perturbations to the free energy landscape due to the presence of the ribosome could be dissected into their enthalpic and entropic contributions, while analysis of RNC folding kinetics following a temperature quench enabled an analysis of folding pathways [192]. In contrast to earlier studies of CI2 and barnase [190] here it was observed, for a variety of small protein domains, that ribosomal attachment did affect the folding pathway, decreasing the diversity of the transition state ensemble and favoring the initiation of folding via N-terminal residues [192].

The differences between these simulation results highlight the continued importance of experimental observations in calibrating and validating the predictions of computer models. An extremely fruitful avenue in this regard is the combination of MD simulations with experimentally-based restraints. Chemical shift-based structure determination is a prime example of this approach, either through the use of chemical shift data to steer molecular fragment replacement methods [193–195], or as energy terms to bias force-fields in MD simulations towards the native state [196]. A particular advantage of these methods for RNC studies is that they appear to be relatively robust to sparse or missing data, as the available chemical shift data are combined with general protein behavior encoded via the simulation forcefield or fragment database [196,197]. Of course, the large number of atoms in the ribosome–nascent chain complex will impose significant computational demands. Nevertheless, given the complexity of RNC NMR, extracting the maximum information from the available spectra is an imperative, and in this regard the application of restrained MD appears a promising direction for future research.

5.5. Dynamical coupling of translation and folding

We close this review by considering once again the co-translational protein folding landscape. A fundamental difference between co-translational folding as it occurs within the cell, and as studied by NMR, is that co-translational folding in the cell is a kinetic process, potentially occurring far from equilibrium. By contrast, stalled RNC samples for study by NMR spectroscopy will have reached equilibrium during the lengthy expression and purification process. Thus, the states that may be kinetically accessed co-translationally may only be a subset of those that are observed in stalled nascent chains, and NMR observations may therefore chart the landscape, without necessarily determining the route that the nascent chain takes across it. This may be inferred, however, from a careful consideration of the kinetics of barrier crossing, relative to the rate of peptide synthesis (Fig. 8).

In the last few years, experimental evidence has begun to accumulate demonstrating that, at least in some cases, the coupling of translation to folding may be of biological significance [198,199]. Translation does not occur at a uniform rate, but is a stochastic process chained to the search for the cognate tRNA, and rare codons in the mRNA transcript may therefore introduce delays in translation and folding processes [200]. In particular, rare codons have been found to cluster in multi-domain proteins approximately 100 base pairs downstream of domain boundaries, a length corresponding to the complete emergence of the domain from the ribosome exit tunnel [201]. This suggests that these translation delays may be significant in allowing additional time for folding to occur, prior to synthesis of the following domain, reducing the probability of misfolding events. In support of this hypothesis, substitution of rare codons with common variants was found to decrease the folding yield [201], while up-regulation of low-abundance tRNAs promoted the aggregation of several cellular proteins [202]. 'Silent' single nucleotide polymorphisms to synonymous rare codons have also been implicated in altered gene function in mammalian cells [203].

If we compare the typical translation rate, k_{trans}, to the rate of exchange between folded and unfolded states, k_{ex}, we can categorize folding into three kinetic regimes. When folding is slow ($k_{ex} \ll k_{trans}$) the full length of the protein is expected to be synthesized and released from the ribosome before folding occurs, and hence the folding process will be identical to that that would be observed in reversible folding/unfolding studies of the isolated protein. By contrast, when folding is fast ($k_{ex} \gg k_{trans}$) the entire surface of the co-translational folding landscape may be sampled prior to release. In this regime, the co-translational acquisition of structure will be observed, but it may be of little consequence to the ultimate folding yield because the repeated unfolding and refolding will erase the memory of any co-translational effects. Therefore, it is the intermediate exchange regime in which biologically interesting behavior may emerge, and the effective co-translational 'steering' of the folding reaction might occur.

In bacterial cells the translation rate k_{trans} is approximately 15–20 amino acids s^{-1}, while for eukaryotic cells it is approximately 1–5 amino acids s^{-1}, and therefore the intermediate folding regime is likely to correspond to the slow to intermediate regimes of NMR chemical exchange, when k_{ex} is comparable to the difference in frequency between states, $\Delta\omega$. While systems in this regime are spectroscopically highly complex, particularly given the low sensitivity of RNC NMR, nevertheless it is under these conditions that translation can most effectively couple to the folding process to mimimise the sampling of states prone to misfolding. Deciphering the length-dependent folding of such RNC systems in thermodynamic, kinetic and structural detail by NMR will surely be a key challenge for the community in the years ahead.

Fig. 8. Schematic co-translational folding landscape, illustrating the changing stability of the states as a function of the polypeptide length, and the significance of the relative rates of translation and barrier crossing in determining the pathway of co-translational folding.

Acknowledgements

References

[1] C. Levinthal, Are there pathways for protein folding, J. Chim. Phys. (1968).
[2] C.B. Anfinsen, Principles that govern the folding of protein chains, Science 181 (1973) 223–230.
[3] A.R. Dinner, A. Sali, L.J. Smith, C.M. Dobson, M. Karplus, Understanding protein folding via free-energy surfaces from theory and experiment, Trends Biochem. Sci. 25 (2000) 331–339.
[4] C.M. Dobson, Protein folding and misfolding, Nature 426 (2003) 884–890.
[5] D.V. Fedyukina, S. Cavagnero, Protein folding at the exit tunnel, Annu. Rev. Biophys. 40 (2011) 337–359.
[6] L.D. Cabrita, C.M. Dobson, J. Christodoulou, Protein folding on the ribosome, Curr. Opin. Struct. Biol. 20 (2010) 33–45.
[7] F. Chiti, C.M. Dobson, Protein misfolding, functional amyloid, and human disease, Annu. Rev. Biochem. 75 (2006) 333–366.
[8] C.M. Dobson, Getting out of shape, Nature 418 (2002) 729–730.
[9] J. Balbach, V. Forge, W.S. Lau, N.A. van Nuland, K. Brew, C.M. Dobson, Protein folding monitored at individual residues during a two-dimensional NMR experiment, Science 274 (1996) 1161–1163.
[10] A. Corazza, E. Rennella, P. Schanda, M.C. Mimmi, T. Cutuil, S. Raimondi, et al., Native-unlike long-lived intermediates along the folding pathway of the amyloidogenic protein beta2-microglobulin revealed by real-time two-dimensional NMR, J. Biol. Chem. 285 (2010) 5827–5835.
[11] E. Rennella, T. Cutuil, P. Schanda, I. Ayala, V. Forge, B. Brutscher, Real-time NMR characterization of structure and dynamics in a transiently populated protein folding intermediate, J. Am. Chem. Soc. 134 (2012) 8066–8069.
[12] S.W. Englander, Protein folding intermediates and pathways studied by hydrogen exchange, Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 213–238.
[13] C.E. Dempsey, Hydrogen exchange in peptides and proteins using NMR spectroscopy, Prog. Nucl. Magn. Reson. Spectrosc. 39 (2001) 135–170.
[14] P. Neudecker, P. Lundström, L.E. Kay, Relaxation dispersion NMR spectroscopy as a tool for detailed studies of protein folding, Biophys. J. 96 (2009) 2045–2054.
[15] D.M. Korzhnev, X. Salvatella, M. Vendruscolo, A.A. Di Nardo, A.R. Davidson, C.M. Dobson, et al., Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR, Nature 430 (2004) 586–590.
[16] P. Vallurupalli, D.F. Hansen, E. Stollar, E. Meirovitch, L.E. Kay, Measurement of bond vector orientations in invisible excited states of proteins, Proc. Natl. Acad. Sci. USA 104 (2007) 18473–18477.
[17] G. Bouvignies, P. Vallurupalli, D.F. Hansen, B.E. Correia, O. Lange, A. Bah, et al., Solution structure of a minor and transiently formed state of a T4 lysozyme mutant, Nature 477 (2011) 111–114.
[18] P. Neudecker, P. Robustelli, A. Cavalli, P. Walsh, P. Lundström, A. Zarrine-Afsar, et al., Structure of an intermediate state in protein folding and aggregation, Science 336 (2012) 362–366.
[19] A. Gershenson, L.M. Gierasch, Protein folding in the cell: challenges and progress, Curr. Opin. Struct. Biol. 21 (2011) 32–41.
[20] A.N. Fedorov, T.O. Baldwin, Contribution of cotranslational folding to the rate of formation of native protein structure, Proc. Natl. Acad. Sci. USA 92 (1995) 1227–1231.
[21] A.V. Nicola, W. Chen, A. Helenius, Co-translational folding of an alphavirus capsid protein in the cytosol of living cells, Nat. Cell Biol. 1 (1999) 341–345.
[22] D.A. Kelkar, A. Khushoo, Z. Yang, W.R. Skach, Kinetic analysis of ribosome-bound fluorescent proteins reveals an early, stable, cotranslational folding intermediate, J. Biol. Chem. 287 (2012) 2568–2578.
[23] B. Kleizen, T. van Vlijmen, H.R. de Jonge, I. Braakman, Folding of CFTR is predominantly cotranslational, Mol. Cell. 20 (2005) 277–287.
[24] T. Tsalkova, O.W. Odom, G. Kramer, B. Hardesty, Different conformations of nascent peptides on ribosomes, J. Mol. Biol. 278 (1998) 713–723.
[25] P.L. Clark, J. King, A newly synthesized, ribosome-bound polypeptide chain adopts conformations dissimilar from early in vitro refolding intermediates, J. Biol. Chem. 276 (2001) 25411–25420.
[26] A. Jansens, E. van Duijn, I. Braakman, Coordinated nonvectorial folding in a newly synthesized multidomain protein, Science 298 (2002) 2401–2403.
[27] A. Hoffmann, A.H. Becker, B. Zachmann-Brand, E. Deuerling, B. Bukau, G. Kramer, Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding, Mol. Cell. 48 (2012) 63–74.
[28] C.A. Woolhead, P.J. McCormick, A.E. Johnson, Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins, Cell 116 (2004) 725–736.
[29] A. Khushoo, Z. Yang, A.E. Johnson, W.R. Skach, Ligand-driven vectorial folding of ribosome-bound human CFTR NBD1, Mol. Cell. 41 (2011) 682–692.
[30] J.P. Ellis, P.H. Culviner, S. Cavagnero, Confined dynamics of a ribosome-bound nascent globin: cone angle analysis of fluorescence depolarization decays in the presence of two local motions, Protein Sci. 18 (2009) 2003–2015.
[31] B. Seidelt, C.A. Innis, D.N. Wilson, M. Gartmann, J.-P. Armache, E. Villa, et al., Structural insight into nascent polypeptide chain-mediated translational stalling, Science 326 (2009) 1412–1415.
[32] S. Bhushan, T. Hoffmann, B. Seidelt, J. Frauenfeld, T. Mielke, O. Berninghausen, et al., SecM-stalled ribosomes adopt an altered geometry at the peptidyl transferase center, PLoS Biol. 9 (2011) e1000581.
[33] A. Loksztejn, Z. Scholl, P.E. Marszalek, Atomic force microscopy captures folded ribosome bound nascent chains, Chem. Commun. (Camb.) 48 (2012) 11727–11729.
[34] C.M. Kaiser, D.H. Goldman, J.D. Chodera, I. Tinoco, C. Bustamante, The ribosome modulates nascent protein folding, Science 334 (2011) 1723–1727.
[35] A. Katranidis, W. Grange, R. Schlesinger, T. Choli-Papadopoulou, D. Brüggemann, M. Hegner, et al., Force measurements of the disruption of the nascent polypeptide chain from the ribosome by optical tweezers, FEBS Lett. 585 (2011) 1859–1863.
[36] S.-T.D. Hsu, P. Fucini, L.D. Cabrita, H. Launay, C.M. Dobson, J. Christodoulou, Structure and dynamics of a ribosome-bound nascent chain by NMR spectroscopy, Proc. Natl. Acad. Sci. USA 104 (2007) 16516–16521.
[37] S.-T.D. Hsu, L.D. Cabrita, P. Fucini, J. Christodoulou, C.M. Dobson, Probing side-chain dynamics of a ribosome-bound nascent chain using methyl NMR spectroscopy, J. Am. Chem. Soc. 131 (2009) 8366–8367.
[38] L.D. Cabrita, S.-T.D. Hsu, H. Launay, C.M. Dobson, J. Christodoulou, Probing ribosome–nascent chain complexes produced in vivo by NMR spectroscopy, Proc. Natl. Acad. Sci. USA 106 (2009) 22239–22244.
[39] A. Rutkowska, M. Beerbaum, N. Rajagopalan, J. Fiaux, P. Schmieder, G. Kramer, et al., Large-scale purification of ribosome–nascent chain complexes for biochemical and structural studies, FEBS Lett. 583 (2009) 2407–2413.
[40] C. Eichmann, S. Preissler, R. Riek, E. Deuerling, Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy, Proc. Natl. Acad. Sci. USA 107 (2010) 9111–9116.
[41] P.L. Clark, Protein folding in the cell: reshaping the folding funnel, Trends Biochem. Sci. 29 (2004) 527–534.
[42] A.N. Fedorov, T.O. Baldwin, Cotranslational protein folding, J. Biol. Chem. 272 (1997) 32715–32718.
[43] P. Virnau, L.A. Mirny, M. Kardar, Intricate knots in proteins: function and evolution, PLoS Comput. Biol. 2 (2006) e122.
[44] W. Li, T. Terakawa, W. Wang, S. Takada, Energy landscape and multiroute folding of topologically complex proteins adenylate kinase and 2ouf-knot, Proc. Natl. Acad. Sci. USA 109 (2012) 17789–17794.
[45] R. Potestio, C. Micheletti, H. Orland, Knotted vs. unknotted proteins: evidence of knot-promoting loops, PLoS Comput. Biol. 6 (2010) e1000864.
[46] J.I. Sułkowska, E.J. Rawdon, K.C. Millett, J.N. Onuchic, A. Stasiak, Conservation of complex knotting and slipknotting patterns in proteins, Proc. Natl. Acad. Sci. USA 109 (2012) E1715–E1723.
[47] M.C. Prentiss, D.J. Wales, P.G. Wolynes, The energy landscape, folding pathways and the kinetics of a knotted protein, PLoS Comput. Biol. 6 (2010) e1000835.
[48] A.L. Mallam, J.M. Rogers, S.E. Jackson, Experimental detection of knotted conformations in denatured proteins, Proc. Natl. Acad. Sci. USA 107 (2010) 8189–8194.
[49] A.L. Mallam, S.E. Jackson, Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins, Nat. Chem. Biol. 8 (2012) 147–153.
[50] I.S. Gabashvili, R.K. Agrawal, C.M. Spahn, R.A. Grassucci, D.I. Svergun, J. Frank, et al., Solution structure of the E. coli 70S ribosome at 11.5 A resolution, Cell 100 (2000) 537–549.
[51] N. Ban, P. Nissen, J. Hansen, P.B. Moore, T.A. Steitz, The complete atomic structure of the large ribosomal subunit at 2.4 A resolution, Science 289 (2000) 905–920.
[52] P. Nissen, J. Hansen, N. Ban, P.B. Moore, T.A. Steitz, The structural basis of ribosome activity in peptide bond synthesis, Science 289 (2000) 920–930.
[53] B.S. Schuwirth, M.A. Borovinskaya, C.W. Hau, W. Zhang, A. Vila-Sanjurjo, J.M. Holton, et al., Structures of the bacterial ribosome at 3.5 A resolution, Science 310 (2005) 827–834.
[54] J. Christodoulou, G. Larsson, P. Fucini, S.R. Connell, T.A. Pertinhez, C.L. Hanson, et al., Heteronuclear NMR investigations of dynamic regions of intact Escherichia coli ribosomes, Proc. Natl. Acad. Sci. USA 101 (2004) 10949–10954.
[55] F.A.A. Mulder, L. Bouakaz, A. Lundell, M. Venkataramana, A. Liljas, M. Akke, et al., Conformation and dynamics of ribosomal stalk protein L12 in solution and on the ribosome, Biochemistry 43 (2004) 5930–5936.
[56] M. Selmer, C.M. Dunham, F.V. Murphy, A. Weixlbaumer, S. Petry, A.C. Kelley, et al., Structure of the 70S ribosome complexed with mRNA and tRNA, Science 313 (2006) 1935–1942.
[57] E.V. Bocharov, A.G. Sobol, K.V. Pavlov, D.M. Korzhnev, V.A. Jaravine, A.T. Gudkov, et al., From structure and dynamics of protein L7/L12 to molecular switching in ribosome, J. Biol. Chem. 279 (2004) 17697–17706.
[58] M. Diaconu, U. Kothe, F. Schlünzen, N. Fischer, J.M. Harms, A.G. Tonevitsky, et al., Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation, Cell 121 (2005) 991–1004.
[59] T.M. Schmeing, K.S. Huang, S.A. Strobel, T.A. Steitz, An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA, Nature 438 (2005) 520–524.
[60] D. Lavalette, B. Amand, F. Pochon, Rotational relaxation of 70S ribosomes by a depolarization method using triplet probes, Proc. Natl. Acad. Sci. USA 74 (1977) 1407–1411.

[61] T.R. Tritton, Proton NMR observation of the *Escherichia coli* ribosome, FEBS Lett. 120 (1980) 141–144.

[62] C.A. Cowgill, B.G. Nichols, J.W. Kenny, P. Butler, E.M. Bradbury, R.R. Traut, Mobile domains in ribosomes revealed by proton nuclear magnetic resonance, J. Biol. Chem. 259 (1984) 15257–15263.

[63] A.T. Gudkov, G.M. Gongadze, V.N. Bushuev, M.S. Okon, Proton nuclear magnetic resonance study of the ribosomal protein L7/L12 in situ, FEBS Lett. 138 (1982) 229–232.

[64] V.N. Bushuev, A.T. Gudkov, Nuclear magnetic resonance techniques for studying structure and function of ribosomes, Meth. Enzymol. 164 (1988) 148–158.

[65] M. Helgstrand, C.S. Mandava, F.A.A. Mulder, A. Liljas, S. Sanyal, M. Akke, The ribosomal stalk binds to translation factors IF2, EF-Tu, EF-G and RF3 via a conserved region of the L12 C-terminal domain, J. Mol. Biol. 365 (2007) 468–479.

[66] H.M. Olson, A. Sommer, D.S. Tewari, R.R. Traut, D.G. Glitz, Localization of two epitopes of protein L7/L12 to both the body and stalk of the large ribosomal subunit. Immune electron microscopy using monoclonal antibodies, J. Biol. Chem. 261 (1986) 6924–6932.

[67] R.K. Agrawal, P. Penczek, R.A. Grassucci, J. Frank, Visualization of elongation factor G on the *Escherichia coli* 70S ribosome: the mechanism of translocation, Proc. Natl. Acad. Sci. USA 95 (1998) 6134–6138.

[68] R.K. Agrawal, A.B. Heagle, P. Penczek, R.A. Grassucci, J. Frank, EF-G-dependent GTP hydrolysis induces translocation accompanied by large conformational changes in the 70S ribosome, Nat. Struct. Biol. 6 (1999) 643–647.

[69] L. Montesano-Roditis, D.G. Glitz, R.R. Traut, P.L. Stewart, Cryo-electron microscopic localization of protein L7/L12 within the *Escherichia coli* 70 S ribosome by difference mapping and Nanogold labeling, J. Biol. Chem. 276 (2001) 14117–14123.

[70] V. Tugarinov, P.M. Hwang, J.E. Ollerenshaw, L.E. Kay, Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes, J. Am. Chem. Soc. 125 (2003) 10420–10428.

[71] R. Sprangers, L.E. Kay, Quantitative dynamics and binding studies of the 20S proteasome by NMR, Nature 445 (2007) 618–622.

[72] A.M. Ruschak, T.L. Religa, S. Breuer, S. Witt, L.E. Kay, The proteasome antechamber maintains substrates in an unfolded state, Nature 467 (2010) 868–871.

[73] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, et al., Cell-free translation reconstituted with purified components, Nat. Biotechnol. 19 (2001) 751–755.

[74] T. Matsuura, H. Yanagida, J. Ushioda, I. Urabe, T. Yomo, Nascent chain, mRNA, and ribosome complexes generated by a pure translation system, Biochem. Biophys. Res. Commun. 352 (2007) 372–377.

[75] T. Tenson, M. Ehrenberg, Regulatory nascent peptides in the ribosomal tunnel, Cell 108 (2002) 591–594.

[76] H. Nakatogawa, K. Ito, The ribosomal exit tunnel functions as a discriminating gate, Cell 108 (2002) 629–636.

[77] S. Bhushan, M. Gartmann, M. Halic, J.-P. Armache, A. Jarasch, T. Mielke, et al., Alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel, Nat. Struct. Mol. Biol. 17 (2010) 313–317.

[78] S. Bhushan, H. Meyer, A.L. Starosta, T. Becker, T. Mielke, O. Berninghausen, et al., Structural basis for translational stalling by human cytomegalovirus and fungal arginine attenuator peptide, Mol. Cell. 40 (2010) 138–146.

[79] K. Ito, S. Chiba, K. Pogliano, Divergent stalling sequences sense and control cellular physiology, Biochem. Biophys. Res. Commun. 393 (2010) 1–5.

[80] H. Kovacs, D. Moskau, M. Spraul, Cryogenically cooled probes—a leap in NMR technology, Prog. Nucl. Magn. Reson. Spectrosc. 46 (2005) 131–156.

[81] A.E. Kelly, H.D. Ou, R. Withers, V. Dötsch, Low-conductivity buffers for high-sensitivity NMR measurements, J. Am. Chem. Soc. 124 (2002) 12013–12019.

[82] G.M. Hautbergue, A.P. Golovanov, Increasing the sensitivity of cryoprobe protein NMR experiments by using the sole low-conductivity arginine glutamate salt, J. Magn. Reson. 191 (2008) 335–339.

[83] T.M. de Swiet, Optimal electric fields for different sample shapes in high resolution NMR spectroscopy, J. Magn. Reson. 174 (2005) 331–334.

[84] M.W. Voehler, G. Collier, J.K. Young, M.P. Stone, M.W. Germann, Performance of cryogenic probes as a function of ionic strength and sample tube geometry, J. Magn. Reson. 183 (2006) 102–109.

[85] L.C. Robosky, M.D. Reily, D. Avizonis, Improving NMR sensitivity by use of salt-tolerant cryogenically cooled probes, Anal. Bioanal. Chem. 387 (2006) 529–532.

[86] D. Rovnyak, J.C. Hoch, A.S. Stern, G. Wagner, Resolution and sensitivity of high field nuclear magnetic resonance spectroscopy, J. Biomol. NMR 30 (2004) 1–10.

[87] E. Lescop, P. Schanda, R. Rasia, B. Brutscher, Automated spectral compression for fast multidimensional NMR and increased time resolution in real-time NMR spectroscopy, J. Am. Chem. Soc. 129 (2007) 2756–2757.

[88] J.C. Hoch, Modern spectrum analysis in nuclear magnetic resonance: alternatives to the Fourier transform, Meth. Enzymol. 176 (1989) 216–241.

[89] S.G. Hyberts, H. Arthanari, G. Wagner, Applications of non-uniform sampling and processing, Top. Curr. Chem. 316 (2012) 125–148.

[90] S.G. Hyberts, K. Takeuchi, G. Wagner, Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data, J. Am. Chem. Soc. 132 (2010) 2145–2147.

[91] S.G. Hyberts, S.A. Robson, G. Wagner, Exploring signal-to-noise ratio and sensitivity in non-uniformly sampled multi-dimensional NMR spectra, J. Biomol. NMR (2012).

[92] C.A. Waudby, J. Christodoulou, An analysis of NMR sensitivity enhancements obtained using non-uniform weighted sampling, and the application to protein NMR, J. Magn. Reson. 219 (2012) 46–52.

[93] A. Kumar, S.C. Brown, M.E. Donlan, B.U. Meier, P.W. Jeffs, Optimization of two-dimensional NMR by matched accumulation, J. Magn. Reson. 95 (1991) (1969) 1–9.

[94] D. Rovnyak, M. Sarcone, Z. Jiang, Sensitivity enhancement for maximally resolved two-dimensional NMR by nonuniform sampling, Magn. Reson. Chem. (2011).

[95] P. Schanda, E. Kupce, B. Brutscher, SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds, J. Biomol. NMR 33 (2005) 199–211.

[96] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, Proc. Natl. Acad. Sci. USA 94 (1997) 12366–12371.

[97] R. Riek, J. Fiaux, E.B. Bertelsen, A.L. Horwich, K. Wüthrich, Solution NMR techniques for large molecular and supramolecular structures, J. Am. Chem. Soc. 124 (2002) 12144–12153.

[98] R. Riek, G. Wider, K. Pervushin, K. Wüthrich, Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules, Proc. Natl. Acad. Sci. USA 96 (1999) 4918–4923.

[99] R. Riek, K. Pervushin, K. Wüthrich, TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution, Trends Biochem. Sci. 25 (2000) 462–468.

[100] J.E. Ollerenshaw, V. Tugarinov, L.E. Kay, Methyl TROSY: explanation and experimental verification, Magn. Reson. Chem. 41 (2003) 843–852.

[101] C. Amero, P. Schanda, M.A. Durá, I. Ayala, D. Marion, B. Franzetti, et al., Fast two-dimensional NMR spectroscopy of high molecular weight protein assemblies, J. Am. Chem. Soc. 131 (2009) 3448–3449.

[102] R.R. Ernst, Application of Fourier transform spectroscopy to magnetic resonance, Rev. Sci. Instrum. 37 (1966) 93.

[103] K. Pervushin, B. Vögeli, A. Eletsky, Longitudinal (1)H relaxation optimization in TROSY NMR spectroscopy, J. Am. Chem. Soc. 124 (2002) 12898–12902.

[104] P. Schanda, B. Brutscher, Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds, J. Am. Chem. Soc. 127 (2005) 8014–8015.

[105] P. Schanda, H. Van Melckebeke, B. Brutscher, Speeding up three-dimensional protein NMR experiments to a few minutes, J. Am. Chem. Soc. 128 (2006) 9042–9043.

[106] E. Lescop, P. Schanda, B. Brutscher, A set of BEST triple-resonance experiments for time-optimized protein resonance assignment, J. Magn. Reson. 187 (2007) 163–169.

[107] J. Farjon, J. Boisbouvier, P. Schanda, A. Pardi, J.-P. Simorre, B. Brutscher, Longitudinal-relaxation-enhanced NMR experiments for the study of nucleic acids in solution, J. Am. Chem. Soc. 131 (2009) 8571–8577.

[108] A. Favier, B. Brutscher, Recovering lost magnetization: polarization enhancement in biomolecular NMR, J. Biomol. NMR 49 (2011) 9–15.

[109] K.V. Pervushin, G. Wider, K. Wüthrich, Single transition-to-single transition polarization transfer (ST2-PT) in [15N,1H]-TROSY, J. Biomol. NMR 12 (1998) 345–348.

[110] A. Eletsky, O. Moreira, H. Kovacs, K. Pervushin, A novel strategy for the assignment of side-chain resonances in completely deuterated large proteins using 13C spectroscopy, J. Biomol. NMR 26 (2003) 167–179.

[111] S. Hiller, G. Wider, T. Etezady-Esfarjani, R. Horst, K. Wüthrich, Managing the solvent water polarization to obtain improved NMR spectra of large molecular structures, J. Biomol. NMR 32 (2005) 61–70.

[112] S. Cai, C. Seu, Z. Kovacs, A.D. Sherry, Y. Chen, Sensitivity enhancement of multidimensional NMR experiments by paramagnetic relaxation effects, J. Am. Chem. Soc. 128 (2006) 13474–13478.

[113] F.-X. Theillet, A. Binolfi, S. Liokatis, S. Verzini, P. Selenko, Paramagnetic relaxation enhancement to improve sensitivity of fast NMR methods: application to intrinsically disordered proteins, J. Biomol. NMR 51 (2011) 487–495.

[114] D.E. Koppel, Study of *Escherichia coli* ribosomes by intensity fluctuation spectroscopy of scattered laser light, Biochemistry 13 (1974) 2712–2719.

[115] E.O. Stejskal, J.E. Tanner, Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient, J. Chem. Phys. 42 (1965) 288–292.

[116] D.K. Wilkins, S.B. Grimshaw, V. Receveur, C.M. Dobson, J.A. Jones, L.J. Smith, Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques, Biochemistry 38 (1999) 16424–16431.

[117] W.S. Price, Recent advances in NMR diffusion techniques for studying drug binding, Aust. J. Chem. 56 (2003) 855–860.

[118] L.H. Lucas, C.K. Larive, Measuring ligand–protein binding using NMR diffusion experiments, Concepts Magn. Reson. 20A (2004) 24–41.

[119] C.A. Waudby, M.D. Mantle, L.D. Cabrita, L.F. Gladden, C.M. Dobson, J. Christodoulou, Rapid distinction of intracellular and extracellular proteins using NMR diffusion measurements, J. Am. Chem. Soc. 134 (2012) 11312–11315.

[120] W.S. Price, NMR Studies of Translational Motion, Cambridge University Press, 2009.

[121] P.T. Callaghan, Translational Dynamics and Magnetic Resonance, OUP Oxford, 2011.

[122] F. Ferrage, M. Zoonens, D.E. Warschawski, J.-L. Popot, G. Bodenhausen, Slow diffusion of macromolecular assemblies by a new pulsed field gradient NMR method, J. Am. Chem. Soc. 125 (2003) 2541–2545.

[123] R. Augustyniak, F. Ferrage, R. Paquin, O. Lequin, G. Bodenhausen, Methods to determine slow diffusion coefficients of biomolecules: applications to Engrailed 2, a partially disordered protein, J. Biomol. NMR 50 (2011) 209–218.

[124] R. Augustyniak, F. Ferrage, C. Damblon, G. Bodenhausen, P. Pelupessy, Efficient determination of diffusion coefficients by monitoring transport during recovery delays in NMR, Chem. Commun. (Camb.) 48 (2012) 5307–5309.

[125] A.V. Buevich, J. Baum, Residue-specific real-time NMR diffusion experiments define the association states of proteins during folding, J. Am. Chem. Soc. 124 (2002) 7156–7162.

[126] S. Rajagopalan, C. Chow, V. Raghunathan, C.G. Fry, S. Cavagnero, NMR spectroscopic filtration of polypeptides and proteins in complex mixtures, J. Biomol. NMR 29 (2004) 505–516.

[127] T. Didenko, R. Boelens, S.G.D. Rüdiger, 3D DOSY-TROSY to determine the translational diffusion coefficient of large protein complexes, Protein Eng. Des. Sel. 24 (2011) 99–103.

[128] R.J.C. Gilbert, P. Fucini, S. Connell, S.D. Fuller, K.H. Nierhaus, C.V. Robinson, et al., Three-dimensional structures of translating ribosomes by Cryo-EM, Mol. Cell. 14 (2004) 57–66.

[129] P. Fucini, C. Renner, C. Herberhold, A.A. Noegel, T.A. Holak, The repeating segments of the F-actin cross-linking gelation factor (ABP-120) have an immunoglobulin-like fold, Nat. Struct. Biol. 4 (1997) 223–230.

[130] A.J. McCoy, P. Fucini, A.A. Noegel, M. Stewart, Structural basis for dimerization of the Dictyostelium gelation factor (ABP120) rod, Nat. Struct. Biol. 6 (1999) 836–841.

[131] P. Fucini, B. Köppel, M. Schleicher, A. Lustig, T.A. Holak, R. Müller, et al., Molecular architecture of the rod domain of the Dictyostelium gelation factor (ABP120), J. Mol. Biol. 291 (1999) 1017–1023.

[132] M.S. Evans, K.G. Ugrinov, M.-A. Frese, P.L. Clark, Homogeneous stalled ribosome nascent chain complexes produced in vivo or in vitro, Nat. Methods 2 (2005) 757–762.

[133] C. Schaffitzel, N. Ban, Generation of ribosome nascent chain complexes for structural and functional studies, J. Struct. Biol. 158 (2007) 463–471.

[134] A.R. Fersht, The sixth Datta Lecture. Protein folding and stability: the pathway of folding of barnase, FEBS Lett. 325 (1993) 5–16.

[135] A.R. Fersht, V. Daggett, Protein folding and unfolding at atomic resolution, Cell 108 (2002) 573–582.

[136] X. Salvatella, C.M. Dobson, A.R. Fersht, M. Vendruscolo, Determination of the folding transition states of barnase by using Phil-value-restrained simulations validated by double mutant PhilJ-values, Proc. Natl. Acad. Sci. USA 102 (2005) 12389–12394.

[137] E.L. Ulrich, H. Akutsu, J.F. Doreleijers, Y. Harano, Y.E. Ioannidis, J. Lin, et al., BioMagResBank, Nucl. Acids Res. 36 (2008) D402–8.

[138] J.C. Martínez, L. Serrano, The folding transition state between SH3 domains is conformationally restricted and evolutionarily conserved, Nat. Struct. Biol. 6 (1999) 1010–1016.

[139] A. Hoffmann, F. Merz, A. Rutkowska, B. Zachmann-Brand, E. Deuerling, B. Bukau, Trigger factor forms a protective shield for nascent polypeptides at the ribosome, J. Biol. Chem. 281 (2006) 6539–6545.

[140] F.J. Blanco, I. Angrand, L. Serrano, Exploring the conformational properties of the sequence space between two proteins with different folds: an experimental study, J. Mol. Biol. 285 (1999) 741–753.

[141] J. Lu, C. Deutsch, Folding zones inside the ribosomal exit tunnel, Nat. Struct. Mol. Biol. 12 (2005) 1123–1129.

[142] D.N. Wilson, R. Beckmann, The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling, Curr. Opin. Struct. Biol. 21 (2011) 274–282.

[143] N.R. Voss, M. Gerstein, T.A. Steitz, P.B. Moore, The geometry of the ribosomal polypeptide exit tunnel, J. Mol. Biol. 360 (2006) 893–906.

[144] A. Rutkowska, M.P. Mayer, A. Hoffmann, F. Merz, B. Zachmann-Brand, C. Schaffitzel, et al., Dynamics of trigger factor interaction with translating ribosomes, J. Biol. Chem. 283 (2008) 4124–4132.

[145] G. De Prat Gay, J. Ruiz-Sanz, J.L. Neira, L.S. Itzhaki, A.R. Fersht, Folding of a nascent polypeptide chain in vitro: cooperative formation of structure in a protein module, Proc. Natl. Acad. Sci. USA 92 (1995) 3683–3686.

[146] G. De Prat Gay, J. Ruiz-Sanz, J.L. Neira, F.J. Corrales, D.E. Otzen, A.G. Ladurner, et al., Conformational pathway of the polypeptide chain of chymotrypsin inhibitor-2 growing from its N terminus in vitro. Parallels with the protein folding pathway, J. Mol. Biol. 254 (1995) 968–979.

[147] J.L. Neira, L.S. Itzhaki, A.G. Ladurner, B. Davis, G. De Prat Gay, A.R. Fersht, Following co-operative formation of secondary and tertiary structure in a single protein module, J. Mol. Biol. 268 (1997) 185–197.

[148] J.L. Neira, A.R. Fersht, Exploring the folding funnel of a polypeptide chain by biophysical studies on protein fragments, J. Mol. Biol. 285 (1999) 1309–1333.

[149] C.C. Chow, C. Chow, V. Raghunathan, T.J. Huppert, E.B. Kimball, S. Cavagnero, Chain length dependence of apomyoglobin folding: structural evolution from misfolded sheets to native helices, Biochemistry 42 (2003) 7090–7099.

[150] S. Cavagnero, N. Kurt, Folding and misfolding as a function of polypeptide chain elongation, in: R.M. Murphy, A.M. Tsai (Eds.), Misbehaving Proteins: Protein (Mis)Folding, Aggregation, and Stability, Springer, 2006, pp. 217–245.

[151] S. Rajagopalan, N. Kurt, S. Cavagnero, High-resolution conformation and backbone dynamics of a soluble aggregate of apomyoglobin119, Biophys. J. 100 (2011) 747–755.

[152] A.R. Fersht, A. Matouschek, L. Serrano, The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding, J. Mol. Biol. 224 (1992) 771–782.

[153] Y. Han, A. David, B. Liu, J.G. Magadán, J.R. Bennink, J.W. Yewdell, et al., Monitoring cotranslational protein folding in mammalian cells at codon resolution, Proc. Natl. Acad. Sci. USA 109 (2012) 12467–12472.

[154] M. Guo, Y. Xu, M. Gruebele, Temperature dependence of protein folding kinetics in living cells, Proc. Natl. Acad. Sci. USA 109 (2012) 17863–17867.

[155] Z. Serber, A.T. Keatinge-Clay, R. Ledwidge, A.E. Kelly, S.M. Miller, V. Dötsch, High-resolution macromolecular NMR spectroscopy inside living cells, J. Am. Chem. Soc. 123 (2001) 2446–2447.

[156] P. Selenko, G. Wagner, Looking into live cells with in-cell NMR spectroscopy, J. Struct. Biol. 158 (2007) 244–253.

[157] A.Y. Maldonado, D.S. Burz, A. Shekhtman, In-cell NMR spectroscopy, Prog. Nucl. Magn. Reson. Spectrosc. 59 (2011) 197–212.

[158] C. Li, M. Liu, Protein dynamics in living cells studied by in-cell NMR spectroscopy, FEBS Lett. (2013).

[159] C.A. Waudby, C. Camilloni, A.W.P. Fitzpatrick, L.D. Cabrita, C.M. Dobson, M. Vendruscolo, et al., In-cell NMR characterization of the secondary structure populations of a disordered conformation of α-synuclein within E. coli cells, PLoS ONE, in press, http://dx.doi.org/10.1371/journal.pone.0072286.

[160] L. Verdier, J. Gharbi-Benarous, G. Bertho, N. Evrard-Todeschi, P. Mauvais, J.-P. Girault, Dissociation–equilibrium constant and bound conformation for weak antibiotic binding interaction with different bacterial ribosomes†, J. Chem. Soc. Perkin Trans. 2 (2000) 2363–2371.

[161] A. Hoffmann, B. Bukau, G. Kramer, Structure and function of the molecular chaperone Trigger Factor, Biochim. Biophys. Acta 2010 (1803) 650–661.

[162] R. Lill, E. Crooke, B. Guthrie, W. Wickner, The "trigger factor cycle" includes ribosomes, presecretory proteins, and the plasma membrane, Cell 54 (1988) 1013–1018.

[163] G. Kramer, T. Rauch, W. Rist, S. Vorderwülbecke, H. Patzelt, A. Schulze-Specking, et al., L23 protein functions as a chaperone docking site on the ribosome, Nature 419 (2002) 171–174.

[164] L. Ferbitz, T. Maier, H. Patzelt, B. Bukau, E. Deuerling, N. Ban, Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins, Nature 431 (2004) 590–596.

[165] D. Baram, E. Pyetan, A. Sittner, T. Auerbach-Nevo, A. Bashan, A. Yonath, Structure of trigger factor binding domain in biologically homologous complex with eubacterial ribosome reveals its chaperone action, Proc. Natl. Acad. Sci. USA 102 (2005) 12017–12022.

[166] E. Crooke, B. Guthrie, S. Lecker, R. Lill, W. Wickner, ProOmpA is stabilized for membrane translocation by either purified E. coli trigger factor or canine signal recognition particle, Cell 54 (1988) 1003–1011.

[167] E. Deuerling, A. Schulze-Specking, T. Tomoyasu, A. Mogk, B. Bukau, Trigger factor and DnaK cooperate in folding of newly synthesized proteins, Nature 400 (1999) 693–696.

[168] S.A. Teter, W.A. Houry, D. Ang, T. Tradler, D. Rockabrand, G. Fischer, et al., Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains, Cell 97 (1999) 755–765.

[169] P. Genevaux, F. Keppel, F. Schwager, P.S. Langendijk-Genevaux, F.U. Hartl, C. Georgopoulos, In vivo analysis of the overlapping functions of DnaK and trigger factor, EMBO Rep. 5 (2004) 195–200.

[170] H. Patzelt, G. Kramer, T. Rauch, H.-J. Schönfeld, B. Bukau, E. Deuerling, Three-state equilibrium of Escherichia coli trigger factor, Biol. Chem. 383 (2002) 1611–1619.

[171] R. Maier, B. Eckert, C. Scholz, H. Lilie, F.-X. Schmid, Interaction of trigger factor with the ribosome, J. Mol. Biol. 326 (2003) 585–592.

[172] C.M. Kaiser, H.-C. Chang, V.R. Agashe, S.K. Lakshmipathy, S.A. Etchells, M. Hayer-Hartl, et al., Real-time observation of trigger factor function on translating ribosomes, Nature 444 (2006) 455–460.

[173] C. Scholz, G. Stoller, T. Zarnt, G. Fischer, F.X. Schmid, Cooperation of enzymatic and chaperone functions of trigger factor in the catalysis of protein folding, Embo J. 16 (1997) 54–58.

[174] E. Martinez-Hackert, W.A. Hendrickson, Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone, Cell 138 (2009) 923–934.

[175] S.-T.D. Hsu, L.D. Cabrita, P. Fucini, C.M. Dobson, J. Christodoulou, Structure, dynamics and folding of an immunoglobulin domain of the gelation factor (ABP-120) from Dictyostelium discoideum, J. Mol. Biol. 388 (2009) 865–879.

[176] M. Vogtherr, D.M. Jacobs, T.N. Parac, M. Maurer, A. Pahl, K. Saxena, et al., NMR solution structure and dynamics of the peptidyl-prolyl cis-trans isomerase domain of the trigger factor from Mycoplasma genitalium compared to FK506-binding protein, J. Mol. Biol. 318 (2002) 1097–1115.

[177] Y. Yao, G. Bhabha, G. Kroon, M. Landes, H.J. Dyson, Structure discrimination for the C-terminal domain of Escherichia coli trigger factor in solution, J. Biomol. NMR 40 (2008) 23–30.

[178] S.-T.D. Hsu, C.M. Dobson, 1H, 15N and 13C assignments of the dimeric ribosome binding domain of trigger factor from Escherichia coli, Biomol. NMR Assign. 3 (2009) 17–20.

[179] S.K. Lakshmipathy, S. Tomic, C.M. Kaiser, H.-C. Chang, P. Genevaux, C. Georgopoulos, et al., Identification of nascent chain interaction sites on trigger factor, J. Biol. Chem. 282 (2007) 12186–12193.

[180] A.V. Ludlam, B.A. Moore, Z. Xu, The crystal structure of ribosomal chaperone trigger factor from *Vibrio cholerae*, Proc. Natl. Acad. Sci. USA 101 (2004) 13436–13441.

[181] R.O. Dror, R.M. Dirks, J.P. Grossman, H. Xu, D.E. Shaw, Biomolecular simulation: a computational microscope for molecular biology, Annu. Rev. Biophys. 41 (2012) 429–452.

[182] M. Vendruscolo, C.M. Dobson, Protein dynamics: Moore's law in molecular biology, Curr. Biol. 21 (2011) R68–70.

[183] D.E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R.O. Dror, M.P. Eastwood, et al., Atomic-level characterization of the structural dynamics of proteins, Science 330 (2010) 341–346.

[184] K. Lindorff-Larsen, S. Piana, R.O. Dror, D.E. Shaw, How fast-folding proteins fold, Science 334 (2011) 517–520.

[185] S. Piana, K. Lindorff-Larsen, D.E. Shaw, Protein folding kinetics and thermodynamics from atomistic simulation, Proc. Natl. Acad. Sci. USA 109 (2012) 17845–17850.

[186] P.C. Whitford, P. Geggier, R.B. Altman, S.C. Blanchard, J.N. Onuchic, K.Y. Sanbonmatsu, Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways, RNA 16 (2010) 1196–1204.

[187] J.L. Klepeis, K. Lindorff-Larsen, R.O. Dror, D.E. Shaw, Long-timescale molecular dynamics simulations of protein structure and function, Curr. Opin. Struct. Biol. 19 (2009) 120–127.

[188] C. Hyeon, D. Thirumalai, Capturing the essence of folding and functions of biomolecules using coarse-grained models, Nat. Commun. 2 (2011) 487.

[189] N. Go, Theoretical studies of protein folding, Annu. Rev. Biophys. Bioeng. 12 (1983) 183–210.

[190] A.H. Elcock, Molecular simulations of cotranslational protein folding: fragment stabilities, folding cooperativity, and trapping in the ribosome, PLoS Comput. Biol. 2 (2006) e98.

[191] E.P. O'Brien, S.-T.D. Hsu, J. Christodoulou, M. Vendruscolo, C.M. Dobson, Transient tertiary structure formation within the ribosome exit port, J. Am. Chem. Soc. 132 (2010) 16928–16937.

[192] E.P. O'Brien, J. Christodoulou, M. Vendruscolo, C.M. Dobson, New scenarios of protein folding can occur on the ribosome, J. Am. Chem. Soc. 133 (2011) 513–526.

[193] A. Cavalli, X. Salvatella, C.M. Dobson, M. Vendruscolo, Protein structure determination from NMR chemical shifts, Proc. Natl. Acad. Sci. USA 104 (2007) 9615–9620.

[194] Y. Shen, O. Lange, F. Delaglio, P. Rossi, J.M. Aramini, G. Liu, et al., Consistent blind protein structure generation from NMR chemical shift data, Proc. Natl. Acad. Sci. USA 105 (2008) 4685–4690.

[195] S. Raman, O.F. Lange, P. Rossi, M. Tyka, X. Wang, J. Aramini, et al., NMR structure determination for larger proteins using backbone-only data, Science 327 (2010) 1014–1018.

[196] P. Robustelli, K. Kohlhoff, A. Cavalli, M. Vendruscolo, Using NMR chemical shifts as structural restraints in molecular dynamics simulations of proteins, Structure 18 (2010) 923–933.

[197] Y. Shen, R. Vernon, D. Baker, A. Bax, De novo protein structure generation from incomplete chemical shift assignments, J. Biomol. NMR 43 (2009) 63–78.

[198] A.A. Komar, A pause for thought along the co-translational folding pathway, Trends Biochem. Sci. 34 (2009) 16–24.

[199] G. Zhang, Z. Ignatova, Folding at the birth of the nascent chain: coordinating translation with co-translational folding, Curr. Opin. Struct. Biol. 21 (2011) 25–31.

[200] E.P. O'Brien, M. Vendruscolo, C.M. Dobson, Prediction of variable translation rate effects on cotranslational protein folding, Nat. Commun. 3 (2012) 868.

[201] G. Zhang, M. Hubalewska, Z. Ignatova, Transient ribosomal attenuation coordinates protein synthesis and co-translational folding, Nat. Struct. Mol. Biol. 16 (2009) 274–280.

[202] I. Fedyunin, L. Lehnhardt, N. Böhmer, P. Kaufmann, G. Zhang, Z. Ignatova, TRNA concentration fine tunes protein solubility, FEBS Lett. 586 (2012) 3336–3340.

[203] C. Kimchi-Sarfaty, J.M. Oh, I.-W. Kim, Z.E. Sauna, A.M. Calcagno, S.V. Ambudkar, et al., A silent polymorphism in the MDR1 gene changes substrate specificity, Science 315 (2007) 525–528.

Glossary

BEST: band-selective excitation short-transient
CRINEPT: cross-correlated relaxation-enhanced polarization transfer
CTD: C-terminal domain
ddFLN: gelation factor (actin-binding protein 120) from *Dictyostelium discoideum*
EF-G: elongation factor G
EF-Tu: elongation factor Tu
FRET: Förster resonance energy transfer
HMQC: heteronuclear multiple quantum coherence
HSQC: heteronuclear single quantum coherence
IF2: initiation factor 2
INEPT: insensitive nuclei enhanced by polarization transfer
LHSQC: longitudinal relaxation optimized HSQC
MD: molecular dynamics
mRNA: messenger ribonucleic acid
NMR: nuclear magnetic resonance
NTD: N-terminal domain
NUS: non-uniform sampling
NUWS: non-uniform weighted sampling
PFG: pulsed field-gradient
PPIase: peptidyl prolyl isomerase
PTC: peptidyl transferase center
RDC: residual dipolar coupling
RNC: ribosome–nascent chain complex
rRNA: ribosomal ribonucleic acid
SecM: secretion monitor protein
SNR: signal to noise ratio
SOFAST: band-selective optimized-flip-angle short-transient
SORDID: signal optimization with recovery in diffusion delays
STE: stimulated echo
TEV: tobacco etch virus
TF: trigger factor
tmRNA: transfer-messenger ribonucleic acid
tRNA: transfer ribonucleic acid
TROSY: transverse relaxation optimized spectroscopy
XSTE: heteronuclear stimulated echo