**Structure and dynamics of a ribosome-bound nascent chain by NMR spectroscopy**

Shang-Te Danny Hsu*, Paola Fucini†‡§, Lisa D. Cabrita*, Hélène Launay*, Christopher M. Dobson*§,
and John Christodoulou*†¶

*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom; †AG-Ribosome, Max-Planck-Institute for Molecular Genetics, Ihnestrasse 73, D-14195 Berlin, Germany; and ‡Institut für Organische Chemie und Chemische Biologie, Johann Wolfgang Goethe-Universität Frankfurt am Main, Max-von-Laue-Strasse 7, D-60438 Frankfurt am Main, Germany

Edited by Adriaan Bax, National Institutes of Health, Bethesda, MD, and approved August 15, 2007 (received for review May 17, 2007)

Protein folding in living cells is inherently coupled to protein synthesis and chain elongation. There is considerable evidence that some nascent chains fold into their native structures in a cotranslational manner before release from the ribosome, but, despite its importance, a detailed description of such a process at the atomic level remains elusive. We show here at a residue-specific level that a nascent protein chain can reach its native tertiary structure on the ribosome. By generating translation-arrested ribosomes in which the newly synthesized polypeptide chain is selectively $^{13}$C/$^{15}$N-labeled, we observe, using ultrafast NMR techniques, a large number of resonances of a ribosome-bound nascent chain complex corresponding to a pair of C-terminally truncated immunoglobulin (Ig) domains. Analysis of these spectra reveals that the nascent chain adopts a structure in which a native-like N-terminal Ig domain is tethered to the ribosome by a largely unfolded and highly flexible C-terminal domain. Selective broadening of resonances for a group of residues that are colocalized in the structure demonstrates that there are specific but transient interactions between the ribosome and the N-terminal region of the folded Ig domain. These findings represent a step toward a detailed structural understanding of the cellular processes of cotranslational folding.

The three-dimensional structures of proteins are inherently determined by their amino acid sequences, and through the process of protein folding the hereditary information contained within the genetic sequence is converted into biological function (1). Our current mechanistic understanding of protein folding at the level of individual residues has come overwhelmingly from a combination of computer simulations and experimental studies of protein denaturation and renaturation in vitro, using biochemical and biophysical methods (2–6). In living cells, however, the folding processes are intricately linked to chain elongation on the ribosome, which occurs in a vectorial manner as the N-terminal part of the nascent chain emerges from the ribosome (7–10). Although recent studies have provided evidence for a degree of structural ordering of nascent chains (11–13), the molecular details of the contribution of the conformational and dynamical restraints imposed by the ribosome on nascent chain folding remain elusive, not the least because of the challenges that arise in applying to such systems the high-resolution physical techniques that can provide the level of structural information required.

Through a series of groundbreaking studies using x-ray crystallography and cryo-electron microscopy (cryo-EM), ribosome structures have been determined in a variety of functional states and have provided detailed insights into the protein translation machinery (for reviews, see refs. 14–17). In addition, by using cryo-EM methods, localized conformational changes have been observed in *Escherichia coli* ribosomes under translation arrest, with a range of different nascent chains threading through the ribosomal tunnel (12). The tRNAs to which the nascent chains are attached are clearly visible in these studies, but the nascent chains themselves have not been observed with any certainty, likely because of their inherent flexibility.

The technique that is most amenable to providing residue-specific structural information about dynamic systems is NMR spectroscopy (5, 18). The ribosome, however, has a mass of $\sim$2.3 MDa and contains $>7,500$ amino acid residues in the $>50$ proteins of which it is composed (19). This size suggests that both the resonance linewidths and the complexity of the spectra would be far too great for NMR to be used to study the properties of any nascent chain attached to such a complex. Nonetheless, we and others have shown (20, 21) that a number of well resolved resonances can be observed in NMR spectra of the ribosome itself as a result of the independent motion of localized regions of the structure. This has enabled us to define the structure and dynamics of the L7/L12 proteins in the mobile GTPase-associated region (GAR or stalk region) of the *E. coli* ribosome that is involved in the regulation of protein elongation (19). If at least part of a nascent chain were to have local dynamical properties similar to those observed for L7/L12, it might be feasible to observe resonances from individual residues of the newly synthesized polypeptide chain as it emerges from the ribosome.

**Results and Discussion**

To explore the possibility of using NMR to study nascent chains on ribosomes, we used a procedure similar to that used in our earlier cryo-EM study (12) to generate a ternary peptidyl-tRNA-ribosome species, i.e., a translation-arrested RNC. As in the cryo-EM study, we used a DNA template that encodes a tandem Ig domain repeat from the gelation factor ABP-120 of *Dictyostelium discoideum* (domains 5 and 6) (22), from which the stop codon was removed (see Materials and Methods) by restriction enzyme digestion (12). This template was designated as “Ig2.” We then selectively labeled the nascent chains of the RNCs by using a coupled transcription-translation *E. coli* cell-free system supplied with the Ig2 DNA template and $^{13}$C/$^{15}$N-labeled amino acids, so that the translation products, i.e., the nascent polypep-
such as free tRNAs and amino acids, any aggregated material, of large-scale reactions to generate appropriate quantities of the fluorescence spectroscopy (from single molecules to femtomoles) (13). SI Fig. 6). Schematic protocol of RNC preparation (for further details, see protocol for RNC preparation (for further details, see).

Fig. 1. Schematic protocol of RNC preparation for NMR studies. (A) Structural model of the ABP-120 domains 5 and 6 (Upper, residues 644–857 of Ig2; Protein Data Bank ID code 1QFH (22)) and the construct used for the Ig2 RNC (Lower). The N-terminal domain (NTD, domain 5) and the C-terminal domain (CTD, domain 6) are designated Ig2 NTD and Ig2 CTD, respectively. The DNA plasmid of the Ig2 was digested at the unique BstNI restriction site, leading to a truncated translation product ending at residue 838, the position at which the last β-strand of the Ig2 CTD begins (highlighted in black and also indicated in the structural model). (B) Schematic description of the experimental protocol for RNC preparation (for further details, see Materials and Methods and SI Fig. 6).

tide chains, are the only species present in the RNC with the potential to be detected by heteronuclear NMR spectroscopy. The RNC preparation protocol is depicted schematically in Fig. 1. A particular challenge is that the quantity of material required for NMR studies (10–100 pmol) is larger by a factor of >10⁴ than that needed for biophysical methods such as fluorescence spectroscopy (from single molecules to femtomoles) (13) or cryo-EM (~10 pmol) (12). We therefore carried out a series of large-scale reactions to generate appropriate quantities of the required RNCs. The resulting reaction mixture was subjected to sucrose gradient ultracentrifugation to separate the RNCs from any dissociated 50S and 30S ribosome subunits, small molecules such as free tRNAs and amino acids, any aggregated material, and, most importantly, any nascent chains free in solution [see supporting information (SI) Fig. 6].

To monitor the structural properties of the Ig2 nascent chain, we used recently developed SOFAST-HMQC spectroscopy (23) to record backbone fingerprint 1H–15N correlation spectra. This relaxation-optimized NMR technique makes use of selective pulses to achieve rapid repetition of data acquisition; in our case, this resulted in an ~4-fold reduction in data acquisition time relative to conventional techniques, an advance that proved to be critical at these low molar concentrations of the RNC samples. This approach enabled us to acquire NMR spectra with the spectral quality necessary to begin their detailed analysis within 6 h. The resulting 1H–15N correlation spectra of the Ig2 RNC (Fig. 2A) are of remarkable quality given the very large molecular mass of the complex and reveal that at least part of the nascent chain is dynamic. Further analysis of the spectrum shows that the average 1H linewidths (as detected in the 1H dimension of the SOFAST-HMQC experiment for 15N-labeled backbone amides) of the resonances observed in the Ig2 RNC spectrum, 59 ± 26 Hz, are in fact comparable to those of purified recombinant Ig2 protein in the presence of 70S ribosomes (designated Ig2/70S; Fig. 2B), which we measured to be 41 ± 11 Hz. By contrast, the values expected for the 70S ribosome particle are >1,000 Hz (24, 25). The addition of 70S ribosomes to purified Ig2 is done to account for the nonspecific increases in sample viscosity and crowding that are associated with the presence of ribosomes and is a similar control to that used in our previous NMR study of the E. coli ribosome (20).

The spectrum of the Ig2 RNC shows a large number of cross-peaks located in the region between 8.2 and 8.6 ppm along the 1H dimension (Fig. 2A), characteristic of resonances arising from amino acids within an unfolded polypeptide chain, but it includes in addition ~45 cross-peaks that are well dispersed, indicative of a highly folded protein. Overall, the spectrum shows...
a marked similarity to that of the Ig2/70S sample (Fig. 2B), and the majority of the well dispersed cross-peaks in the Ig2 RNC spectrum show only marginal chemical shift differences (\(\Delta \delta_{\text{HN}} = 0.18 \pm 0.17 \text{ ppm}\)) from those of the isolated Ig2 NTD. This result is significant because it shows the presence of a highly folded structure of the nascent Ig2 NTD even when attached to the ribosome (Fig. 2) (26, 27).

We also carried out \(^{15}\text{N}\)-filtered diffusion experiments [X-STE measurements (28)] to monitor the translational diffusion rate of the species giving rise to the observed resonances, an approach similar to that which we used previously to study the mobile region in the ribosome itself (20). The translational diffusion coefficient \(D_{\text{trans}}\) observed for the species labeled with the \(^{15}\text{N}\) isotope (i.e., the Ig2 nascent chain) is at least four times smaller than that of purified Ig2 in the presence of 70S ribosomes (Ig2/70S), thus setting a lower bound for the apparent molecular mass for the species that gives rise to the observed resonances of \(\approx 1\ \text{MDa}\) and thereby confirming that the nascent chain is attached to the ribosome (Fig. 3 and Table 1).

To further verify that the nascent chain is associated with the ribosome, nascent chains labeled with \(^{15}\text{S}\)Met were found to be efficiently released from the ribosome by the addition of puromycin, a phenylalanine-mimicking antibiotic that reacts specifically with the peptidyl-tRNA in the 70S ribosome (Fig. 3C). Addition of puromycin to the NMR sample after recording the \(^{1}\text{H}–^{15}\text{N}\) correlation spectra (Fig. 2C) resulted in an overall signal decrease of \(62 \pm 7\%\), as assessed by integrating the signal intensities over the region from 7.5–8.5 ppm along the \(^{1}\text{H}\) dimension (see Fig. 4A). The presence in a sample of isolated Ig2, of a concentration of 70S ribosomes similar to that present in the Ig2 RNC sample (12 \(\mu M\)) reduces the relative \(D_{\text{trans}}\) of a result of the increased sample viscosity. Under the assumption that the change in sample viscosity does not affect the monomer–dimer equilibrium of Ig2, and that the sample viscosity is dominated by the concentration of 70S ribosomes, the apparent MM of Ig2 derived from aqueous solution (32 kDa) was then used to determine the apparent MM of the Ig2 RNC. When human lysozyme is used as a reference without correcting for sample viscosity, the apparent MM of Ig2 RNC is found to be 2.4 MDa.

### Table 1. Relative translational diffusion coefficients, \(D_{\text{trans}}\), determined by \(^{15}\text{N}\)-filtered X-STE NMR diffusion experiments

<table>
<thead>
<tr>
<th>Protein</th>
<th>Relative (D_{\text{trans}}) %</th>
<th>Molecular mass, kDa</th>
<th>Apparent molecular mass, kDa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human lysozyme</td>
<td>100 (1)</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Ig2 NTD</td>
<td>107 (3)</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Ig2</td>
<td>76 (4)</td>
<td>21</td>
<td>32</td>
</tr>
<tr>
<td>Ig2/70S</td>
<td>63 (2)</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>Ig2 RNC</td>
<td>18 (&lt;10)</td>
<td>&gt;1,300</td>
<td></td>
</tr>
</tbody>
</table>

\*Relative \(D_{\text{trans}}\) was defined relative to a \(^{1}\text{H}\)-labeled human lysozyme sample (error estimates are indicated in parentheses), as described elsewhere (28). The relative \(D_{\text{trans}}\) of the Ig2 RNC is estimated from the intensity ratio of the X-STE spectra with 10% and 70% \(G_{\text{max}}\) integrated over the region from 7.6 to 9.3 ppm in the \(^{1}\text{H}\) dimension (see Fig. 4A). Only two gradient points were employed for the Ig2 RNC because of the limited sample concentrations.

\(\Delta \delta_{\text{HN}} = 0.10 \pm 0.07 \text{ ppm}\) (Fig. 2C).

To obtain insight into the dynamical restraints on the nascent chain resulting from its attachment to the ribosome, we examined the linewidths of the resolved resonances of the Ig2 NTD in the spectrum of the Ig2 RNC (Fig. 2A). We found that a number of cross-peaks of the Ig2 NTD exhibit differential line broadening (Fig. 4); in particular, the cross-peaks of 10 residues in the Ig2 RNC spectrum are broadened beyond detection (highlighted in Figs. 2 and 4A). Remarkably, mapping of the residues whose resonances exhibit significantly broader \(^{1}\text{H}\) linewidths on the structure of the Ig2 NTD reveals that all of these, apart from V729, are located in loop regions of the native structure (filled orange circles in Fig. 5). The localized line broadening of the resonances of G697 and N677, as well as the...
missing cross-peak of I748, can be attributed to their spatial proximity to the flexible and structurally heterogeneous Ig2 CTD that tethers the nascent chain to the peptidyl transferase center through the ribosomal tunnel. The much greater line broadening observed in the N-terminal hemisphere of the Ig2 NTD (Fig. 5), however, suggests that there are preferential, albeit transient, intermolecular interactions between this region of the nascent chain and the ribosome, which are lost upon its release.

Conclusions

We have demonstrated that NMR spectra of remarkable quality can be obtained from a nascent polypeptide chain attached to a ribosome, through a combination of selective isotope labeling using cell-free expression, ultrafast NMR spectroscopy, NMR diffusion measurements, and biochemical assays. Analysis of the spectra of the Ig2 construct studied here has shown not only that the C-terminal Ig domain is unfolded when attached to the ribosome, but also that it is sufficiently mobile to enable us to show that the remainder of the nascent chain, the Ig2 NTD, has folded to its fully native state despite retaining some interactions with the ribosomal surface.

The ability to record high-resolution NMR spectra of nascent chains provides an avenue of opportunity for obtaining a detailed description of the steps involved in the process of cotranslational folding at a residue-specific level. This understanding could be spurred by recent advances in our ability to define protein structures and dynamics (5, 29, 30) by using a variety of NMR parameters that can be collected even from large biomolecular assemblies of great biological interest (31, 32). Together, these developments serve as a vantage point for probing (in synergy with the information obtainable from x-ray crystallographic and cryo-EM studies) the manner in which auxiliary factors such as molecular chaperones help an elongating polypeptide chain to fold into a specific structure and to avoid misfolding and its potential pathogenic consequences (6, 33).

Materials and Methods

Sample Constructs. The DNA sequence of domains 5 and 6 of the gelation factor ABP-120 from D. discoideum (residues 644–838) was cloned and digested by BstNI, as described previously (12). A stop codon was reintroduced in the truncation site by PCR, and the resulting DNA template is designated Ig2. Similarly, a stop codon was introduced at the boundary between the domains to obtain an isolated domain 5 construct (residues 644–750), i.e., the NTD of the Ig2, designated Ig2 NTD. Both DNA constructs were transformed into an E. coli BL21 strain for protein overexpression.

Preparation of Ribosome-Nascent Chain Complexes. We used a coupled transcription–translation E. coli cell-free system

**Fig. 4.** Differential line broadening in the Ig2 NTD of the Ig2 RNC. (A) Representative regions of the SOFAST-HMQC spectra of the Ig2 RNC (Left) and Ig2/70S (Right), as shown in Fig. 2 A and B, respectively. Cross-peaks that show significantly broader linewidths in the Ig2 RNC spectrum are indicated with dashed circles. The cross-peak of residue I748, which is adjacent to the domain boundary between the Ig2 NTD and CTD (Fig. 1A) cannot be identified in the Ig2 RNC spectrum (Left). (B) ’H linewidth profiles of the Ig2 NTD in the Ig2 RNC (black open circles) and Ig2/70S (green line) obtained from the corresponding ’H–’N SOFAST-HMQC spectra (Fig. 2A and B). The residues in the Ig2 NTD in the Ig2 RNC whose cross-peaks exhibit severe line broadening are indicated by filled red circles, with those that exhibit linewidths >1 standard deviation above the average (dashed horizontal line) indicated by filled orange circles. Residues that are part of the β-strands of the Ig2 NTD are indicated by black arrows.

**Fig. 5.** Structural mapping of the locations of the Ig2 RNC that experience NMR line broadening. (A) Schematic representation of Ig2 RNC showing relative sizes (approximate) of the 70S ribosome and the Ig2 NTD (boxed), with the loop regions in the N-terminal hemisphere that appear to interact with the ribosome indicated in red. (B) Mapping of the residues whose resonances are broadened on the structure of Ig2 NTD, depicted in ribbon representation [Protein Data Bank ID code 1QFH (22)]. The heavy atoms of those residues whose resonances are broadened by >1 standard deviation from the average (filled orange circles in Fig. 2B) and those that are broadened beyond detection in the Ig2 RNC spectra (filled red circles in Fig. 2B) are shown as orange and red spheres, respectively, and labeled with their residue identities. Residue I748, which lies at the domain boundary and whose cross-peak is absent in the Ig2 RNC spectrum (Fig. 4A), is shown by green spheres. The figure was generated by using PyMol (www.pymol.org).
transcript. To monitor the reaction product, [35S]methionine was incorporated to translation arrest when the ribosome reaches the end of this codon. Reproducible NMR samples after being subjected to the purification process were pooled and pelleted by ultracentrifugation at 4°C for 22 h. Sucrose gradients were fractionated and monitored by optical absorbance at 260 nm. Fractions corresponding to 70S ribosomes were pooled and pelleted by ultracentrifugation at 4°C for 22 h in an SW42T rotor (rotor speeds were verified by Beckman Coulter, Fullerton, CA) for NMR samples [maximum relative centrifugal force (max. ref) 68,000 × g] and for 18 h in a SW40 rotor for [35S]-containing samples (max. ref 71,000 × g). These sucrose gradients were fractionated and monitored by optical absorbance at 254 nm. Fractions corresponding to 70S ribosomes were pooled and pelleted by ultracentrifugation at 4°C for 22 h in a 45Ti rotor (max. ref 126,000 × g) and a 50Ti rotor (max. ref 98,500 × g) for NMR and radioactively labeled samples, respectively. The pellets were then resuspended with binding buffer to 10 mM Hepes, pH 7.6/140 mM NH4Cl/6 mM MgCl2/0.005 mM spermine/2 mM spermidine/2 mM DTT) by shaking at 4°C for 1 h. The concentration of 70S ribosomes was determined by the optical absorbance at 260 nm (one A260 unit = 24 pmol 70S). The resuspended RNC samples were divided into aliquots, flash frozen in liquid nitrogen, and stored at −80°C. Typically, the cell-reaction size is 10 ml, which proved sufficient to give two reproducible NMR samples after being subjected to the purification procedure described above.

Preparation of Isolated Ig2 and E. coli 70S Ribosomes. Uniformly [13C,15N]-labeled Ig2 NTD and Ig2 polypeptides were overexpressed in minimal medium and purified as described previously (22). The purified proteins were then concentrated to 300–500 μM in buffer (20 mM KPi, pH 7.0/2 mM EDTA/2 mM 2-mercaptoethanol). E. coli 70S ribosomes were purified by zonal sucrose-gradient centrifugation and dialyzed against Tico buffer (10 mM Hepes, pH 7.6/30 mM NH4Cl/6 mM MgCl2/2 mM 2-mercaptoethanol). Purified 70S ribosomes were divided into aliquots, flash frozen in liquid nitrogen, and stored at −80°C. Uniformly [15N]-labeled Ig2, which has an identical sequence to that of the Ig2 nascent chain prepared from the cell-free reactions, was mixed with purified 70S ribosomes solution (Ig2 and 70S ribosome concentrations were 100 μM and 12 μM, respectively, in this Ig2/70S sample). The Ig2/70S sample was then exchanged into the same binding buffer used for RNC samples, for spectral comparison.

NMR Spectroscopy. All RNC samples (70S ribosome concentrations of 12 μM) were prepared in binding buffer. NMR data were recorded by using Bruker 700-MHz and 900-MHz spectrometers (Bruker BioSpin, Karlsruhe, Germany), both equipped with triple resonance cryogenic probeheads. As a lock solvent, 10% D2O was present in all NMR samples. [1-15N]-SOFAST-HMQC spectra (23) were recorded by using a 300-ms recycling delay with 1.024 × 64 complex points collected in the 1H and 15N dimensions, respectively, and 1,024 transients per increment. [15N]-edited X-STE diffusion measurements (28) with varied diffusion coding gradient strengths (3 ms total duration) were recorded by using a diffusion delay of 200 ms with 2,048 transients. These two experiments, SOFAST-HMQC and X-STE diffusion measurements, were carried out in an interleaved manner, allowing the integrity of the RNC to be ascertained. These experiments were also recorded with the reconstituted Ig2/70S sample for comparison. 1H PFG-LED diffusion measurements (34) were also recorded for RNC samples, to monitor the translational diffusion of the 70S ribosome by using the observable proton resonances of the stalk L7/L12 proteins (20). HNCO experiments (1H−15N or 1H−13C correlation projections) were recorded and revealed that the observed 1H−15N correlations in the SOFAST-HMQC spectra could not arise from free amino acids because the presence of a peptide bond is the prerequisite for detection of HN,CO−, correlations (35). For backbone assignment of the isolated Ig2 NTD spectra, the standard triple resonance experiments [HNCA, HN(CO)CA, CBCA(CO)NH, HNCCB, HNCO, HN(CA)CO, 15N-TOCSY-HSQC spectra (35)] were recorded at 25°C at 700 MHz in the absence of the ribosome. The chemical shifts of individual spin systems (HN, NCA, CB, CO, and HA) were collected manually for automated assignment using the program MARS (36). High-resolution 1H−15N HSQC spectra (2,048 × 200 complex points) of the Ig2 NTD and Ig2 were recorded at several temperatures ranging from 10°C to 25°C to facilitate assignment of the resolved cross-peaks in Ig2 RNC SOFAST-HMQC spectra by following the minimum chemical shift displacement criterion. All NMR data were processed and analyzed by XWIN-NMR (Bruker BioSpin), NMRPipe (37), and Sparky (38) software packages. The linewidth analyses were carried out by using the two-dimensional cross-peak fitting routine in Sparky, assuming Gaussian line shapes with visual inspection of the fitting quality in each case. Twenty-nine and 59 cross-peaks in the Ig2 RNC and Ig2/70S, respectively, were analyzed to obtain the mean values of linewidths (Fig. 4). Thirty-two nonoverlapping cross-peaks that are commonly present in the spectrum of Ig2 RNC before and after the addition of puromycin, and also in the spectrum of Ig2/70S, were used to obtain the 1H- and 15N-weighted chemical shift displacement given by $\Delta\delta = [(\delta_{\text{Ig2}} - \delta_{\text{RNC}})^2 + (0.65\Delta\delta_{\text{Ig2}})^2]^{1/2}$.

We thank the staff and acknowledge the use of the Biomolecular NMR Facility, Department of Chemistry, University of Cambridge. S.-T.D.H. is a recipient of a Netherlands Ramsay Fellowship and a Human Frontier Long-Term Fellowship. L.D.C. is a National Health and Medical Research Council J. Martin Fellow, P.F. and J.C. are recipients of a Human Frontier Young Investigators Award (RGY67/2007). C.M.D. and J.C. acknowledge funding from The Wellcome Trust and The Leverhulme Trust. The 900-MHz spectra were recorded at the SON NMR Large-Scale Facility, Utrecht, The Netherlands, which is funded by the European Union project “EU-NMR-European Network of Research Infrastructures for Providing Access and Technological Advancement in Bio-NMR.”