Check for updates

OPEN

Interactions between nascent proteins and the ribosome surface inhibit co-translational folding

Anaïs M. E. Cassaignau[1], Tomasz Włodarski[1], Sammy H. S. Chan [1], Lauren F. Woodburn[1], Ivana V. Bukvin [1], Julian O. Streit [1], Lisa D. Cabrita[1], Christopher A. Waudby [1✉] and John Christodoulou [1,2✉]

Most proteins begin to fold during biosynthesis on the ribosome. It has been suggested that interactions between the emerging polypeptide and the ribosome surface might allow the ribosome itself to modulate co-translational folding. Here we combine protein engineering and NMR spectroscopy to characterize a series of interactions between the ribosome surface and unfolded nascent chains of the immunoglobulin-like FLN5 filamin domain. The strongest interactions are found for a C-terminal segment that is essential for folding, and we demonstrate quantitative agreement between the strength of this interaction and the energetics of the co-translational folding process itself. Mutations in this region that reduce the extent of binding result in a shift in the co-translational folding equilibrium towards the native state. Our results therefore demonstrate that a competition between folding and binding provides a simple, dynamic mechanism for the modulation of co-translational folding by the ribosome.

I n the cell, most nascent polypeptide chains begin to fold during biosynthesis[1–3]. In many cases, co-translational folding increases the ability of a protein to efficiently attain its native structure[4–7]. This may in part be due to the ribosome modulating the conformational ensembles sampled by nascent chains[8–11]. The ribosome constrains disordered chains close to its charged surface[12], and can promote the early formation of compact states during co-translational folding[13–16]. In general, nascent chains have been found to be destabilized when they are bound to the ribosome[5,17–19]. This could be linked to interactions between the disordered nascent chain and the ribosome surface[20], some of which bear a partial electrostatic character[21,22]. Ribosome–nascent chain interactions have also been suggested to attenuate co-translational folding rates[23], and to compete with co-translational assembly between a nascent chain and its binding partner[24]. Collectively, these studies point clearly towards a role for the ribosome in shaping the onset of co-translational folding[3]. However, due in large part to the technical difficulty of measuring the intramolecular equilibria associated with ribosome–nascent chain interactions, a link between the energetics of ribosome interactions and co-translational folding outcomes has not yet been established in quantitative terms.

In this article we study the co-translational folding of FLN5, a 105 residue immunoglobulin-like domain from the tandem repeat protein filamin (Fig. 1a), using SecM-arrested ribosome–nascent chain complexes (RNCs) in which FLN5 is tethered to the ribosome via variable lengths of the following domain, FLN6 (Fig. 1b)[20,25]. Measurements of the accessibility of a C-terminal cysteine to covalent modification by PEG-maleimide (PEGylation) showed that the entire FLN5 domain emerges beyond the ribosome exit tunnel for linkers comprising at least 31 residues[20] (FLN5+31). However, NMR observations demonstrated that FLN5 remains partially unfolded until the linker extends beyond 42 residues. This offset between the emergence of FLN5 and its folding suggests that the ribosome has a destabilizing effect on co-translational folding, which we speculated

could be due to interactions between the unfolded nascent chain and the ribosome surface[20]. We use NMR spectroscopy together with protein engineering, molecular dynamics simulations and PEGylation measurements of nascent chain stability to identify a series of interaction sites on the nascent chain of varying affinities, and their impact on co-translational folding.

Results

Identification of regions of FLN5 RNCs interacting with the ribosome surface. Since the FLN5+31 RNC represents the first point during protein biosynthesis when the entire sequence is available for folding, we have used this biosynthetic snapshot as a starting point to examine how the ribosome might modulate the dynamic properties of a nascent chain. Previously, a comparison of the 2D ^{1}H,^{15}N NMR correlation spectrum of a FLN5+31 RNC against an isolated unfolded variant, Y719E, revealed site-selective line broadening that we interpreted as evidence of such interactions (Fig. 1c,d). From this, we have identified three main regions for investigation here. First, two clusters of aromatic residues, located within the core of the folded domain (Fig. 1a), were identified as displaying substantial broadening; these are referred to here as 'F_3' (residues F665–F675) and 'Y_3' (residues Y715–Y727) sites (Fig. 1d, green). Second, strong line broadenings were observed in the mildly basic C-terminal region of FLN5 (residues N728–C747, the 'C-terminal segment' (Fig. 1d, red). These observations correlate well with measurements of the interaction between individual amino acids and purified 70S ribosomes, which are strongest for aromatic and basic side chains (Fig. 1e and Extended Data Fig. 1). Third, resonances of the FLN6 tether are broadened to varying extents across FLN5 RNCs of increasing lengths, which, in part, relates to its occlusion within the ribosomal exit tunnel[20] (Fig. 1d, cyan).

Interactions of F_3 and Y_3 aromatic clusters with the ribosome surface. We prepared a series of constructs, termed F_3A_3, A_3Y_3

[1]Institute of Structural and Molecular Biology, University College London, London, UK. [2]Institute of Structural and Molecular Biology, Birkbeck College, University of London, London, UK. ✉e-mail: c.waudby@ucl.ac.uk; j.christodoulou@ucl.ac.uk
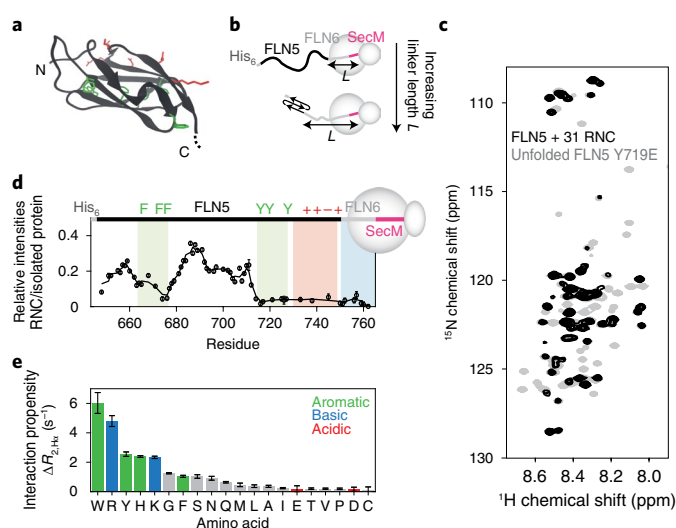
Fig. 1 | Identifying interactions of FLN5 RNCs with the ribosome surface. **a**, Cartoon representation of the FLN5 structure (PDB: 1QFH[48]). Aromatic (green) and surface-exposed (red) residues that have been mutated in this study are shown using stick representations. **b**, Design of FLN5+L RNC constructs, where L indicates the linker length, comprising a fragment of the following FLN6 domain and the 17-amino-acid SecM arrest peptide[20]. **c**, ^{1}H,^{15}N SOFAST-HMQC spectrum of a FLN5+31 RNC (blue) overlaid with that of the isolated, unfolded FLN5 Y719E variant (grey) (283 K, 950 MHz). Contour levels of the RNC spectrum are 3.8-fold lower than for isolated Y719E. **d**, Intensities of FLN5+31 RNC resonances relative to the isolated, unfolded FLN5 Y719E variant. A three-point moving average is shown. Shading indicates the approximate location of ribosome interaction sites characterized in this study. **e**, Transferred transverse ^{1}H NMR relaxation measurements of the interaction of free amino acids with purified 70S ribosomes (283 K, 700 MHz). All error bars indicate standard errors derived from the spectral noise.

and A$_3$A$_3$, in which the clusters of aromatic residues were mutated to alanine (Fig. 2a and Extended Data Fig. 2). The close overlay between ^{1}H,^{15}N NMR spectra of the isolated variants and disordered FLN5 indicates they are natively unfolded, as expected given large changes to the hydrophobic core (Extended Data Fig. 2c–f). On the ribosome, the ^{1}H,^{15}N correlation spectra of the corresponding FLN5+31 RNC variants showed substantial increases in cross-peak intensities relative to wild-type (wt) FLN5+31 RNC, broadly localized to the alanine mutation sites (Fig. 2b and Extended Data Fig. 3). Residues between A665 and G700 within the FLN5+31 A$_3$A$_3$ RNC variant reached a relative intensity close to 1, indicating that the mobility of this segment of the nascent chain is comparable to that of the isolated unfolded protein. These data demonstrate that the aromatic residues mediate a large proportion of the interactions at these sites. By contrast, in all variants the C-terminal segment of FLN5 (N728–C747) remained broadened beyond detection. This suggests additional interactions within this nascent chain segment (Fig. 2b).

To measure the strength of these aromatic interactions, we developed and acquired sensitivity-optimized measurements of the transverse cross-correlated relaxation (CCR) rate (Extended Data Fig. 4). Interactions of nascent chain segments with the ribosome surface will result in transferred relaxation, and thus an increase in the CCR rate, $\Delta\eta_{xy}$, relative to the isolated unfolded state. The increase, $\Delta\eta_{xy} = p_B\eta_{xy}^{\text{bound}}$, is proportional to the bound fraction p_B, and to the CCR rate of the bound state, η_{xy}^{bound}. Based on the known rotational correlation time (τ_C) of the ribosome[26], η_{xy}^{bound} is estimated to be ~7,000 s^{-1}.

CCR measurements were acquired for FLN5+31 wt, F$_3$A$_3$, A$_3$Y$_3$ and A$_3$A$_3$ RNCs, and the corresponding isolated proteins (Fig. 2c and Extended Data Fig. 4). In general, we observed that increased CCR rates were associated with reduced resonance intensities. In the wt RNCs, increases in η_{xy} of ~15 s^{-1} were measured around the aromatic F$_3$ cluster. Within the Y$_3$ site, residues beyond T714 were broadened beyond detection in wt and A$_3$Y$_3$ RNCs, but examination of flanking residues indicates that η_{xy} probably increases substantially beyond 15 s^{-1}. Increased relaxation was also observed for residues at the N terminus, which we ascribe to an interaction of the 6xHis tag as previously observed in α-synuclein RNCs[21].

We observe some evidence of cooperativity between the F$_3$ and Y$_3$ clusters, both by comparison of resonance intensities and of CCR rates: the elimination of one cluster leads to a small reduction in the interaction of the neighbouring cluster (Fig. 2b,d). However, in quantitative terms the extent of interaction is weak in all cases. Within the F$_3$ cluster, observed increases in CCR rates correspond to ribosome-bound nascent chain populations of ~0.1% (Fig. 2d). While resonances for the Y$_3$ cluster were strongly broadened in wt and A$_3$Y$_3$, only allowing for a partial quantitation of CCR rates, the ribosome-bound population between residues 708 and 716 also neared ~0.1%. Although these estimates assume a rigid bound state, even substantial flexibility that results in an order of magnitude decrease in η_{xy}^{bound} would indicate a bound population of only a few percent. Such weak interactions within these regions are not sufficient to perturb the co-translational folding process.

Interaction of the C-terminal segment with the ribosome surface. We next sought to investigate the interaction of the C-terminal segment (N728–C747) (Fig. 1d). The elimination of the Y$_3$ cluster in the F$_3$A$_3$ and A$_3$A$_3$ variants resulted in newly observable resonances flanking this segment. Careful inspection revealed that small chemical shift perturbations (CSPs) were observed for these resonances (V717–G725, encompassing part of the Y$_3$ region), relative to the isolated protein, and that their magnitude increased towards the (unobserved) ribosome-binding segment (Fig. 3a,b and Extended Data Fig. 5). Similarly, CSPs were observed at the C-terminal end of the binding segment between I748 and A751. Focusing on the A$_3$A$_3$ variant, we found that these CSPs were substantially reduced at high ionic strength (Fig. 3c), suggesting that they are associated with an interaction mediated at least in part by an electrostatic contribution.

We then explored the effect of the RNC linker length on the observed CSPs. We hypothesized that a shorter RNC would experience a higher effective ribosome concentration[20] and therefore modulate the extent of binding. Indeed, as the linker length is increased from 26 to 110 amino acids, CSPs were observed to decrease (Fig. 3d) while resonance intensities increased (Extended Data Fig. 5). The CSPs at different lengths were collinear, indicating that changes in the ^{1}H and ^{15}N chemical shifts were strongly correlated (Fig. 3d). These collinear, correlated CSPs are an unambiguous indication that the C-terminal segment is rapidly exchanging between a ribosome-bound and free state, such that the observed chemical shift reflects a population-weighted average of unbound and bound states[27]. Resonances of the C-terminal residues I749–A751 showed deviations from collinearity at short linker lengths, which we attribute to proximity to the exit tunnel. For this reason, they have been excluded from further analysis. Together, these observations provide compelling evidence for a strong ribosome interaction involving nascent chain residues between Y727 and C747.

Alanine mutations within the F$_3$ and Y$_3$ clusters did not substantially perturb the observed C-terminal CSPs, indicating that there is no detectable cooperativity between interactions of these aromatic clusters and the C-terminal segment (Extended Data Fig. 5). Furthermore, no modulations in chemical shifts or intensities were observed upon varying magnetic field strength (Extended Data Fig. 5). Given the fast chemical exchange behaviour observed, we can infer
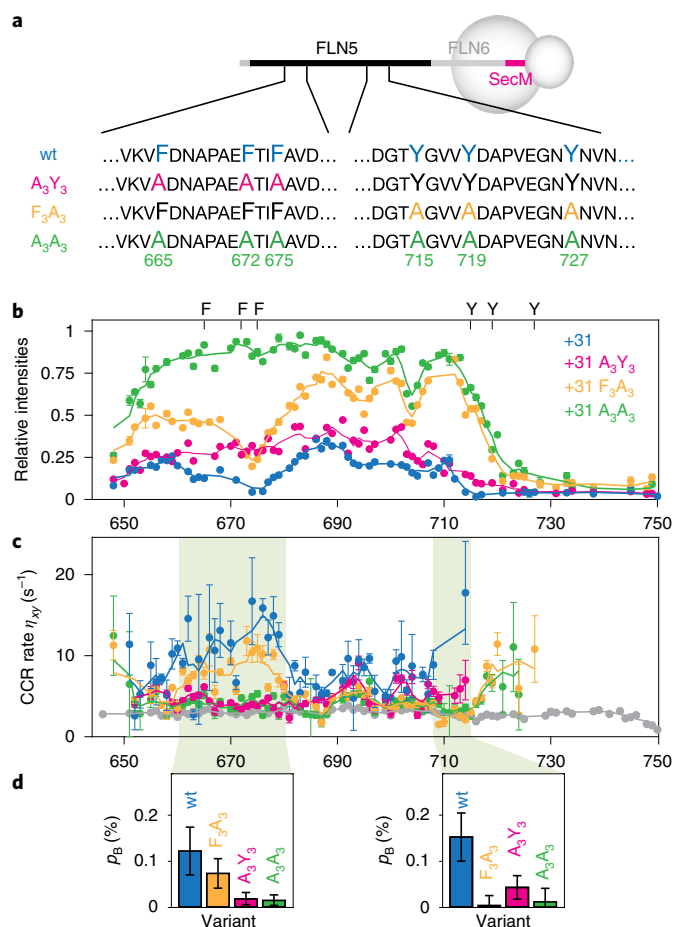
Fig. 2 | Quantification of interactions of aromatic clusters with the ribosome surface. a, Design of aromatic to alanine FLN5 variants. **b**, ^{1}H,^{15}N SOFAST-HMQC resonance intensities (950 MHz, 283 K) of FLN5+31 RNC variants relative to the corresponding isolated unfolded state. **c**, Transverse cross-correlated relaxation rates (950 MHz, 283 K) of FLN5+31 RNC variants (colours as in **b**), and isolated FLN5 A_3A_3 (grey). Errors were derived from the spectral noise for F_3A_3, A_3Y_3 and A_3A_3 variants, while the mean and standard error from three biological repeats are shown for the wt. Three-point moving averages are shown. **d**, Bound state populations, p_B, of the highlighted nascent chain segments (residues 660–680 and residues 708–716) determined by analysis of transferred cross-correlated relaxation. Error bars represent the standard deviation of residues within these segments.

from the largest frequency difference, $\Delta\nu$, of 30 Hz (for the ^{15}N resonance of E724 at 22.3 T) that the dissociation of the bound state is rapid, with a lifetime much less than 3 ms ($\tau \ll 1/4\pi\Delta\nu$, that is, $k_{off} \approx \tau^{-1} \gg 360\,\text{s}^{-1}$).

Impact of the linker sequence composition on FLN5 interactions. We investigated the influence of the FLN6 linker on FLN5–ribosome interactions. We previously reported that replacing all FLN6 residues with a poly(GS) sequence (Fig. 3e) did not perturb co-translational folding equilibria[20]. Here we find that substituting poly(GS) linkers within A_3A_3 RNCs led to only small shifts in the CSPs and resonance intensities in the C-terminal segment, indicating that the extent of binding is not greatly perturbed (Fig. 3f and Extended Data Fig. 6), and suggesting little impact of the nature of the tether on the co-translational folding of FLN5.

Quantification and molecular modelling of the C-terminal segment interaction. Having identified a ribosome interaction site

in the C-terminal segment, we next examined its molecular basis. Although the C-terminal segment does not contain aromatic residues, it contains three basic side chains (R734, K739, K746) that we hypothesized would contribute towards the electrostatic interactions (Fig. 1e), given the reduced CSPs observed at increased ionic strength (Fig. 3c). To explore this, we designed a FLN5 construct, $A_3A_3E_6$, in which six residues (surface-exposed within the folded state) were replaced with acidic glutamate residues, thus reversing the net charge within this segment from +1 to −6 (Figs. 1a and 4a). The E_6 mutations were found to greatly reduce the magnitude of CSPs in the C-terminal segment (Fig. 4b), indicating that these mutations reduce the affinity of the segment for the ribosome surface. However, some CSPs and line broadening persist (Extended Data Fig. 7), suggesting that the binding interaction of this region is not abrogated completely. Similar reductions in CSPs were also observed between FLN5+31 wt and E6 RNCs, although due to the increased line broadening fewer resonances adjacent to the interacting segment could be resolved (Extended Data Fig. 7).

To quantify the effect of the E_6 mutations on the binding interaction, we have analysed further the observed CSPs, which reflect a population-weighted average of unbound and bound states. These residues provide a convenient 'ruler' to compare the interactions of different nascent chains with the ribosome. However, to determine the absolute amount of binding, the chemical shift of the fully bound state must be determined. To achieve this, we carried out a global analysis of CSPs observed in the A_3A_3 and $A_3A_3E_6$ RNC variants across multiple linker lengths (Fig. 4c). Two assumptions were required for this analysis. First, while the strength of interactions in these variants clearly varies as a function of linker length, we assume that the difference in free energy of binding between variants, $\Delta\Delta G^{A3A3-A3A3E6}_{U_{free}-U_{bound}}$ (where U_{free} and U_{bound} represent unfolded states with the C-terminal segment unbound and ribosome-bound, respectively), is independent of RNC length. This is equivalent to the assumption that C-terminal segments in both variants experience the same effective ribosome concentration at a given linker length, which is supported by the similar structural and dynamic properties of the A_3A_3 and $A_3A_3E_6$ variants: no chemical shift perturbations or differences in ^{15}N R_2 relaxation rates are observed beyond the immediate vicinity of the E_6 mutations (Extended Data Fig. 8). Second, while unbound resonance positions vary between A_3A_3 and $A_3A_3E_6$ RNCs due to local sequence effects, we assume that the chemical shift change upon binding, $\Delta\delta_{max}$, is the same for both variants. Given this, the observed chemical shift perturbations of four well-resolved resonances in A_3A_3 and $A_3A_3E_6$ RNCs were fitted to determine the chemical shift differences between free and bound states, and the difference in affinities of the two variants, $\Delta\Delta G^{A3A3-A3A3E6}_{U_{free}-U_{bound}} = 1.9 \pm 0.1\,\text{kcal mol}^{-1}$ (Fig. 4d and

Extended Data Fig. 8). These results indicate that at a short RNC length of 26 amino acids, 90% and 22% of the A_3A_3 and $A_3A_3E_6$ RNCs are bound to the ribosome, respectively, then at a longer RNC length of 42 amino acids these values decreased to 60% and 7%, respectively, and by 110 amino acids, the interaction essentially disappears in the $A_3A_3E_6$ variant (Fig. 4e).

Next, we used coarse-grained (CG) molecular dynamics simulations of the A_3A_3 and $A_3A_3E_6$ RNCs to probe this interaction further, and to explore the location of nascent chain interaction sites on the ribosome surface (Fig. 4g). The strength of electrostatic interactions between the nascent chain and the ribosome surface in the CG model was calibrated using simulations of the A_3A_3 FLN5+42 RNC, in order to achieve ~60% binding of the C-terminal segment (730–746), as determined from our NMR observations (Fig. 4e). Simulations of other RNC lengths (+31, +67 and +110), and of the +42 E_6 variant, were then carried out with no further adjustment of parameters. We found that this CG model accurately identified
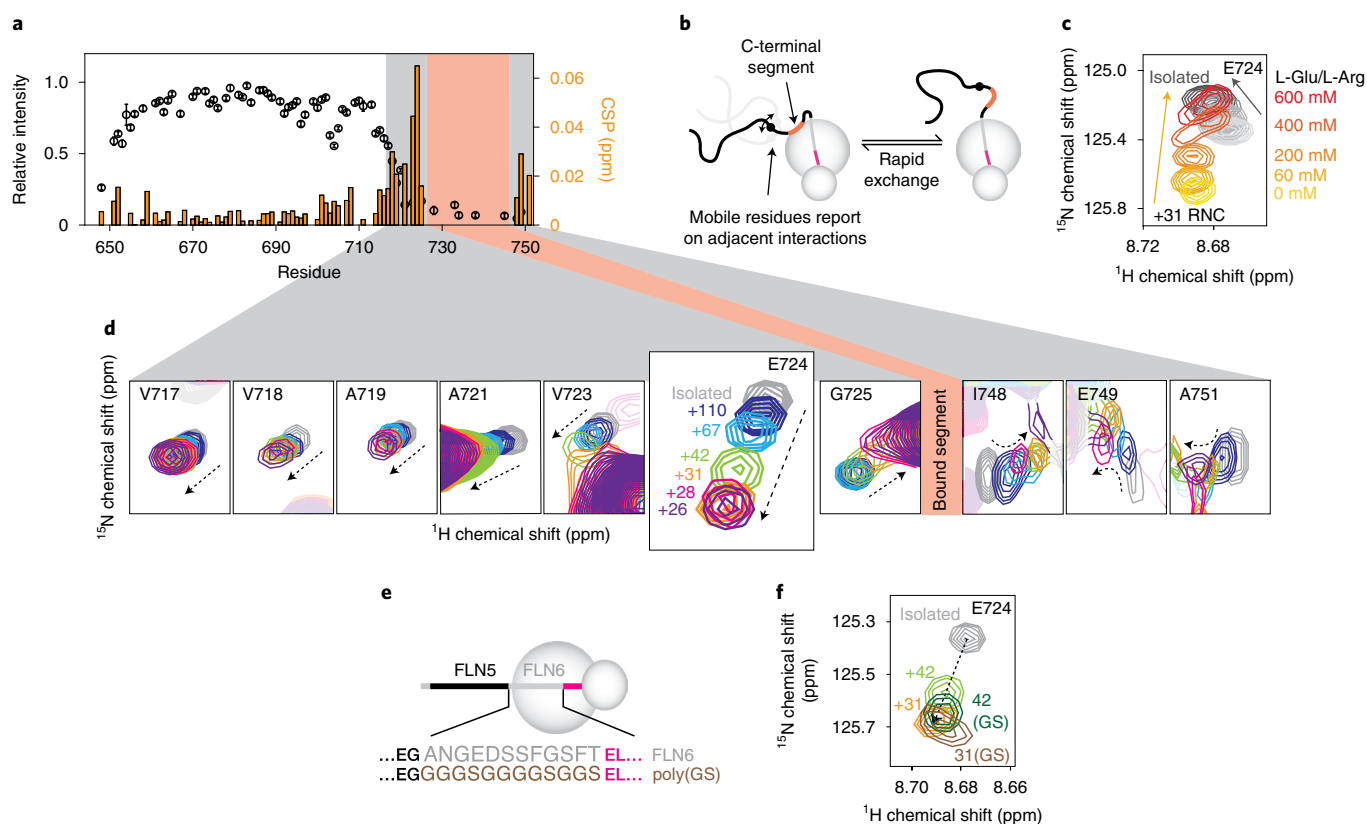
Fig. 3 | A ribosome-binding site in a C-terminal segment of FLN5. a, Resonance intensities (circles) and CSPs (bars) for the FLN5+31 A_3A_3 RNC relative to isolated A_3A_3. **b**, A model for the C-terminal segment in an equilibrium between ribosome-bound and unbound states. **c**, Magnified view of the E724 resonance in ^{1}H,^{15}N SOFAST-HMQC spectra of an FLN5+31 A_3A_3 RNC (yellow–red) and isolated FLN5 A_3A_3 (grey), shown at increasing ionic strengths (0–600 mM Glu/Arg). **d**, ^{1}H,^{15}N SOFAST-HMQC spectra showing isolated FLN5 A_3A_3 and FLN5 A_3A_3 RNC resonances, with linker lengths as indicated for the E724 resonance. Intensities have been rescaled for clarity. The location of the bound nascent chain segment is indicated with red shading. **e**, Design of poly(GS) linker RNCs. **f**, ^{1}H,^{15}N SOFAST-HMQC spectra of FLN5 A_3A_3 RNCs, centred on the E724 resonance and containing the native FLN6 or poly(GS) linker sequences as indicated. All data were acquired at 950 MHz, 283 K.

electrostatics-based interactions of the C-terminal segment of the nascent chain that were consistent with the extent of binding measured from our CSP analysis across all lengths (Fig. 4f). The C-terminal binding segment (N728–C747) in the A_3A_3 FLN5+31 RNC contacted primarily the 23S rRNA region located adjacent to the exit tunnel (mainly helices H24 and H50), and also a loop of the nearby ribosomal protein, uL24 (Fig. 4h). Moreover, for many individual RNA bases in this region, in particular those within helices H24, H47, H50 and H59 (Extended Data Fig. 9), the contact probability correlated strongly with the extent of binding across nascent chain lengths as determined experimentally from our CSP analysis (Fig. 4i). These observations point to a nascent chain interaction site at the ribosome exit vestibule. We note that a similar interaction site is also observed for folded FLN5+47 RNC by cryo electron microscopy[28], and that helices H24 and H50 are highly structurally conserved across Bacteria and Eukaryotes. It may be that this region on the ribosomal surface has a role in modulating the co-translational folding of FLN5 and perhaps generally for nascent chains.

Modulation of co-translational folding by C-terminal interactions. In the previous sections, we have developed a detailed description of the interactions formed between unfolded FLN5 nascent chains and the ribosome surface. By stabilizing the unfolded state, such interactions will in general lead to an inhibition of folding (Fig. 5a), and we set out in this section to test this linkage in quantitative terms. Since interactions of the F_3 and Y_3 aromatic clusters are weak, we have focused our analysis on the strong interactions of the

C-terminal segment, which we have previously shown is essential for folding of the domain[29]. A simple calculation (detailed in the Supplementary Information) relates the amount of binding within the unfolded state (p_B) to its change in stability:

$$\Delta G_{\text{Ufree}-\text{U}} = RT\ln(1 - p_B) \qquad (1)$$

We have used this expression, together with our measurements of interactions in unfolded RNCs of various lengths, to predict the effect on the co-translational folding equilibria of folding-competent FLN5 RNCs in the absence of the destabilizing A_3A_3 mutations (Fig. 4e, right-hand axis). While this calculation assumes that the A_3A_3 mutations themselves do not perturb the interaction, we have previously noted that the impact of E_6 mutations on interactions of wt and A_3A_3 RNCs was similar (Extended Data Fig. 7), and that there was no detectable cooperativity between the aromatic clusters and interactions of the C-terminal segment (Extended Data Fig. 5). On this basis, we predict a destabilization of wt FLN5 RNCs by over 1 kcal mol^{-1} at short linker lengths (Fig. 4e, right-hand axis). In the absence of other changes in thermodynamic stability, we also predict that the E_6 variant should shift co-translational folding equilibria towards the folded state (Fig. 5a and Fig. 4e, right-hand axis).

To test the predicted effect of interactions on folding experimentally, we employed a cysteine mass-tagging approach, PEGylation[30], as a reporter of folding and co-translational folding equilibria. 2D NMR observations indicated that wt and E_6 FLN5 have similar structures both in isolation and as RNCs (Extended Data Fig. 10),
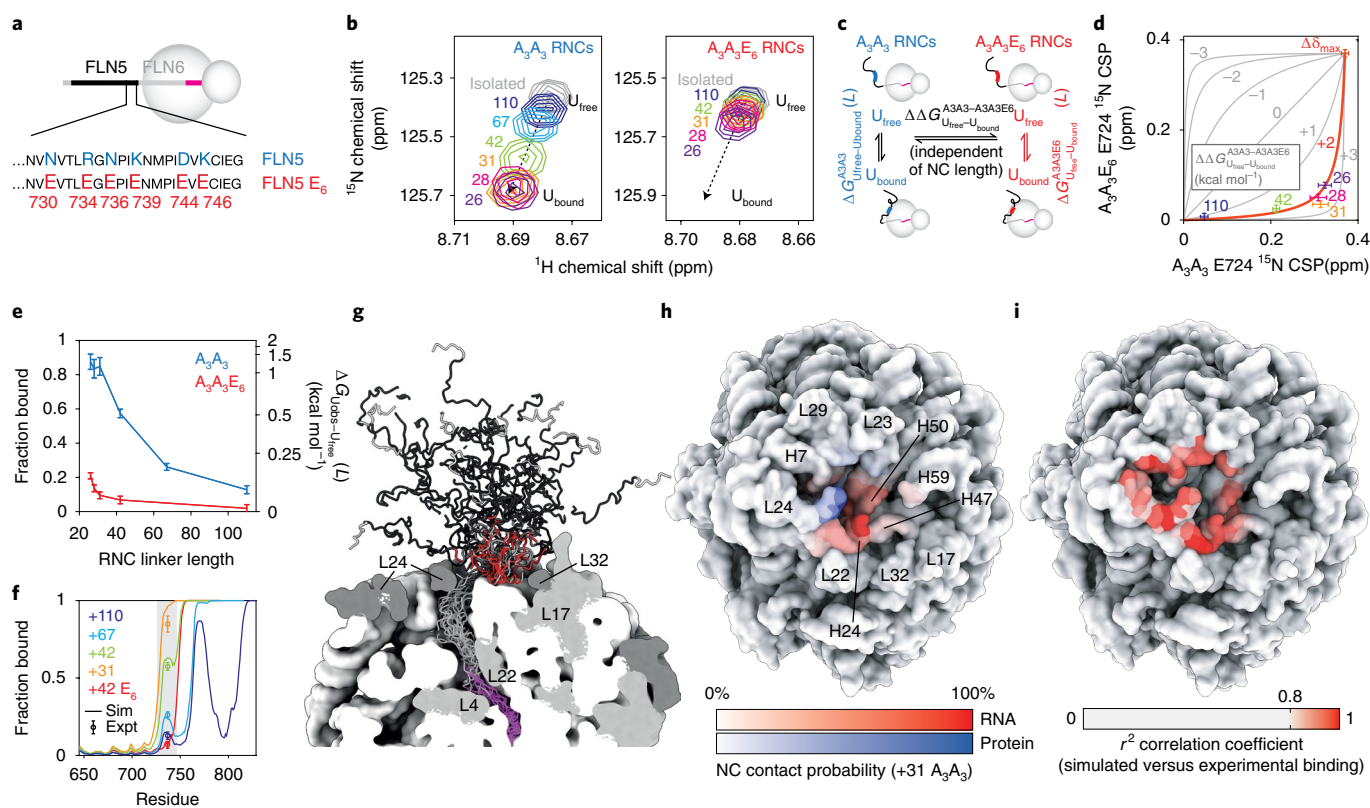
Fig. 4 | Quantification and molecular modelling of interactions between the FLN5 C-terminal region and the ribosome surface. a, Design of the FLN5 E_6 variant. **b**, $^1H,^{15}N$ SOFAST-HMQC spectra (283 K, 950 MHz) showing the E724 resonance within FLN5 A_3A_3 RNCs and FLN5 $A_3A_3E_6$ RNCs (linker lengths as indicated). Intensities have been rescaled for clarity. **c**, Equilibria underlying the joint analysis of A_3A_3 and $A_3A_3E_6$ nascent chain–ribosome interactions. **d**, Correlation plot of E724 ^{15}N chemical shift perturbations in A_3A_3 versus $A_3A_3E_6$ RNCs, with linker lengths as indicated. A global fit across multiple resonances (Extended Data Fig. 7) is shown (red line) from which $\Delta\Delta G^{A3A3-A3A3E6}_{U_{free}-U_{bound}}$ is derived. **e**, Binding of the FLN5 C-terminal region as a function of linker length. The right-hand axis shows the predicted perturbation to the observed co-translational folding equilibrium in folding-capable RNCs (equation (2)). **f**, Interaction propensities of unfolded FLN5 nascent chain residues with the ribosome surface based on coarse-grained molecular dynamics simulations. Markers with error bars indicate bound fractions of the C-terminal region (shaded) determined experimentally (Fig. 4e). **g**, A structural ensemble of the FLN5+67 nascent chain determined through molecular dynamics simulations. The nascent chain is coloured as in **a**, with the interacting segment highlighted in red. Ribosomal RNA and proteins are coloured white and grey, respectively. **h**, Contact probabilities between ribosome protein or RNA residues and the C-terminal region of the FLN5+31 A_3A_3 nascent chain, determined through coarse-grained molecular dynamics simulations. **i**, Correlation across multiple nascent chain lengths between nascent chain–ribosome contact probabilities and the experimentally determined binding between the ribosome and the C-terminal region of FLN5 nascent chains.

and so the mutations C747V and A721C were introduced to replace the native, C-terminal cysteine with one that is unambiguously emerged beyond the exit tunnel at all linker lengths examined here, and that is protected from modification by PEG-maleimide unless the domain is unfolded (Extended Data Fig. 10). The isolated variants, FLN5 C721 V747 (the 'pseudo-wild-type') and FLN5 C721 V747 E_6, were reacted with PEG-maleimide in increasing concentrations of urea and the extent of protection was determined by electrophoresis. The resulting measurements fitted closely to a two-state unfolding model and indicated that the E_6 variant was destabilized relative to the wt by $\Delta\Delta G^{E6-wt}_{N-U} = 4.0 \pm 0.2$ kcal mol^{-1} (where N and U represent native and unfolded states), in excellent agreement with equivalent CD measurements (Fig. 5b and Extended Data Fig. 10).

We next subjected FLN5 and FLN5 E_6 RNCs to PEGylation following the same protocol (Fig. 5c and Extended Data Fig. 10). As ribosomes dissociate at high concentrations of urea[17], we restricted our analysis to measurements below 3 M urea, with the exception of FLN5+110 RNC for which data up to 3.75 M were used (Fig. 5c, dashed lines). The stabilities of FLN5+110 wt and E_6 RNCs are indistinguishable from the corresponding isolated proteins,

indicating that the difference in thermodynamic stability between isolated FLN5 and FLN5 E_6 is conserved on the ribosome at long linker lengths ($\Delta\Delta G^{E6-wt}_{N-U} = 4.0 \pm 0.3$ kcal mol^{-1}). However, at short linker lengths, the difference in stability between FLN5 and FLN5 E_6 was decreased ($\Delta\Delta G^{E6-wt}_{N-U} = 2.3 \pm 0.2$ kcal mol^{-1} for FLN5+31) (Fig. 5c,d). Consistent with our hypothesis, wt nascent chains therefore become destabilized relative to E_6 under strongly interacting conditions. Based on our earlier measurements of bound populations (Fig. 4e), we can predict the change in stability relative to the isolated domains (Fig. 5d), and these predictions closely match direct observations using PEGylation (Fig. 5d). Therefore, we conclude that the C-terminal segment can indeed modulate co-translational FLN5 folding through the competition between binding and folding.

Conclusion

For many domains, co-translational folding has been reported to be destabilized on the ribosome[5,17–19]. This effect is generally inferred to arise through interactions between emerging nascent chains and the ribosome, which have been observed for a range of different nascent chain sequences[12,21–24]. However, a direct link between
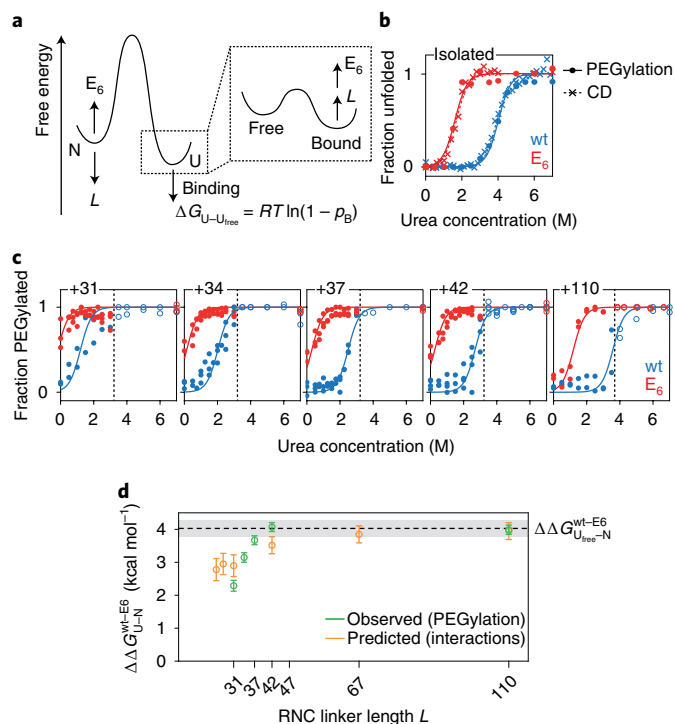
Fig. 5 | Analysis of folding and co-translational folding in FLN5 variants. **a**, Model for the free energy landscape associated with FLN5 RNCs, and how it is modulated by linker length L, and E_6 mutations. **b**, Urea denaturation measurements of the stability of isolated wt and E_6 FLN5 C721 V747, via the extent of PEGylation, and by CD spectroscopy. Solid lines show fits to a two-state unfolding model with a shared m value $\langle m_{D-N} \rangle = 1.67 \pm 0.06$ kcal mol^{-1}M^{-1}. **c**, Extent of PEGylation measured as a function of urea concentration for wt and E_6 FLN5 RNCs. Measurements at low urea concentrations (filled circles) were fitted to a two-state unfolding curve using the m value determined from the isolated proteins. **d**, Difference in free energy of co-translational folding between wt and E_6 RNCs, as measured by PEGylation (green) and predicted from CSP analysis (orange). Errors bars represent the standard error derived from fitting.

the energetics of interactions and of co-translational folding has not previously been established. In this article we have therefore systematically examined a series of interactions between unfolded FLN5 nascent chains and the ribosome surface, in order to determine their effect on co-translational folding. While some of these interactions, between aromatic clusters and the ribosome surface, are too weak to perturb the energetics of folding substantially, we have identified a strongly interacting C-terminal segment, with over 90% bound at short linker lengths. The length of the sequestered segment (22 amino acids) is longer than the C-terminal truncation that can be tolerated by isolated FLN5 before it unfolds (12 amino acids)[29], supporting the crucial role of this segment to enable native structure formation. Importantly, our analysis establishes quantitative agreement between the strength of the observed interactions and the energetics of co-translational folding itself, providing a residue-specific demonstration of the ability of the ribosome surface to directly modulate co-translational folding, effectively acting as a holdase.

A notable consequence of our analysis is that strong interactions are required to appreciably perturb the co-translational folding landscape, for example, a destabilization of 1 kcal mol^{-1} requires over 80% binding (equation (2)). Such interactions were indeed observed for the C-terminal segment at short linker lengths, but the

bound population decreases sharply with increasing linker length, to below 50% at linker lengths beyond 47 amino acids (Fig. 4e). This rapid, short-range effect may provide a mechanism by which the engagement of molecular chaperones with the emerging NC is regulated.

We observe that the molecular determinants of ribosome interactions, that is, positively charged and aromatic residues, are similar to those recognized by other molecule chaperones and processing complexes[31]. These include signal recognition particle[32], the ribosome-associated chaperone trigger factor[33–35], and SecB[36], both of which function as holdases for nascent polypeptide chains, and other downstream chaperones such as DnaK[37]. Indeed, over the past few years NMR studies have been instrumental in revealing with exquisite detail the mechanisms underlying substrate recognition and chaperone action by these molecules[35,36,38–42].

During biosynthesis, substrates emerge from the ribosome in their high-energy unfolded states, and so in contrast to chaperones that act post-translationally, interactions with holdases carry no energetic cost. This is reflected in the ATP independence of both trigger factor and SecB holdases—as well as in interactions with the ribosome surface itself. We speculate that this short-range holdase activity could have a number of functional roles: sequestration of hydrophobic segments until later residues have been synthesized; to reduce the risk of misfolding, particularly in tandem repeat proteins such as FLN[29,43]; or to delay folding ahead of co-translational assembly[44–46] or the engagement of downstream chaperones such as TF and SecB, which may be involved in secretory pathways[35,36]. This suggests that the ribosome is more than an inert hub that orchestrates interactions of auxiliary factors and chaperones[47], and in fact is itself an active participant in the co-translational folding process. In conclusion, we demonstrate the holdase effect of the ribosome as a bespoke form of regulation over co-translational folding.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41557-021-00796-x.

References

1. Balchin, D., Hayer-Hartl, M. & Hartl, F. U. In vivo aspects of protein folding and quality control. *Science* **353**, aac4354–aac4354 (2016).
2. Hartl, F. U. & Hayer-Hartl, M. Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**, 1852–1858 (2002).
3. Cassaignau, A. M. E., Cabrita, L. D. & Christodoulou, J. How does the ribosome fold the proteome? *Annu. Rev. Biochem.* **89**, 389–415 (2020).
4. Kim, S. J. et al. Protein folding. Translational tuning optimizes nascent protein folding in cells. *Science* **348**, 444–448 (2015).
5. Liu, K., Rehfus, J. E., Mattson, E. & Kaiser, C. M. The ribosome destabilizes native and non-native structures in a nascent multidomain protein. *Protein Sci.* **26**, 1439–1451 (2017).
6. Alexander, L. M., Goldman, D. H., Wee, L. M. & Bustamante, C. Non-equilibrium dynamics of a nascent polypeptide during translation suppress its misfolding. *Nat. Commun.* **10**, 2709 (2019).
7. Zhang, G., Hubalewska, M. & Ignatova, Z. Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat. Struct. Mol. Biol.* **16**, 274–280 (2009).
8. Waudby, C. A., Dobson, C. M. & Christodoulou, J. Nature and regulation of protein folding on the ribosome. *Trends Biochem. Sci* **44**, 914–926 (2019).
9. O'Brien, E. P., Christodoulou, J., Vendruscolo, M. & Dobson, C. M. New scenarios of protein folding can occur on the ribosome. *J. Am. Chem. Soc.* **133**, 513–526 (2011).
10. Frydman, J., Erdjument-Bromage, H., Tempst, P. & Hartl, F. U. Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase. *Nat. Struct. Biol.* **6**, 697–705 (1999).

11. Sander, I. M., Chaney, J. L. & Clark, P. L. Expanding Anfinsen's principle: contributions of synonymous codon selection to rational protein design. *J. Am. Chem. Soc.* **136**, 858–861 (2014).

12. Ellis, J. P., Bakke, C. K., Kirchdoerfer, R. N., Jungbauer, L. M. & Cavagnero, S. Chain dynamics of nascent polypeptides emerging from the ribosome. *ACS Chem. Biol.* **3**, 555–566 (2008).

13. Holtkamp, W. et al. Cotranslational protein folding on the ribosome monitored in real time. *Science* **350**, 1104–1107 (2015).

14. Notari, L., Martínez-Carranza, M., Farías-Rico, J. A., Stenmark, P. & von Heijne, G. Cotranslational folding of a pentarepeat β-helix protein. *J. Mol. Biol.* **430**, 5196–5206 (2018).

15. Kemp, G., Kudva, R., de la Rosa, A. & von Heijne, G. Force–profile analysis of the cotranslational folding of HemK and filamin domains: comparison of biochemical and biophysical folding assays. *J. Mol. Biol.* **431**, 1308–1314 (2019).

16. Liutkute, M., Maiti, M., Samatova, E., Enderlein, J. & Rodnina, M. V. Gradual compaction of the nascent peptide during cotranslational folding on the ribosome. *eLife* **9**, e60895 (2020).

17. Samelson, A. J., Jensen, M. K., Soto, R. A., Cate, J. H. D. & Marqusee, S. Quantitative determination of ribosome nascent chain stability. *Proc. Natl Acad. Ssci. USA* **113**, 13402–13407 (2016).

18. Liu, K., Maciuba, K. & Kaiser, C. M. The ribosome cooperates with a chaperone to guide multi-domain protein folding. *Mol. Cell* **74**, 310–319.e7 (2019).

19. Jensen, M. K., Samelson, A. J., Steward, A., Clarke, J. & Marqusee, S. The folding and unfolding behavior of ribonuclease H on the ribosome. *J. Biol. Chem.* **295**, 11410–11417 (2020).

20. Cabrita, L. D. et al. A structural ensemble of a ribosome-nascent chain complex during cotranslational protein folding. *Nat. Struct. Mol. Biol.* **23**, 278–285 (2016).

21. Deckert, A. et al. Structural characterization of the interaction of α-synuclein nascent chains with the ribosomal surface and trigger factor. *Proc. Natl Acad. Sci. USA* **113**, 5012–5017 (2016).

22. Knight, A. M. et al. Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics. *ACS Chem. Biol.* **8**, 1195–1204 (2013).

23. Kaiser, C. M., Goldman, D. H., Chodera, J. D., Tinoco, I. & Bustamante, C. The ribosome modulates nascent protein folding. *Science* **334**, 1723–1727 (2011).

24. Marino, J., Buholzer, K. J., Zosel, F., Nettels, D. & Schuler, B. Charge interactions can dominate coupled folding and binding on the ribosome. *Biophys. J.* **115**, 996–1006 (2018).

25. Cassaignau, A. M. E. et al. A strategy for co-translational folding studies of ribosome-bound nascent chain complexes using NMR spectroscopy. *Nat. Protoc.* **11**, 1492–1507 (2016).

26. Lavalette, D., Amand, B. & Pochon, F. Rotational relaxation of 70S ribosomes by a depolarization method using triplet probes. *Proc. Natl Acad. Sci. USA* **74**, 1407–1411 (1977).

27. Jagger, A. M., Waudby, C. A., Irving, J. A., Christodoulou, J. & Lomas, D. A. High-resolution ex vivo NMR spectroscopy of human Z α$_1$-antitrypsin. *Nat. Commun.* **11**, 1–13 (2020).

28. Javed, A. et al. Visualising nascent chain dynamics at the ribosome exit tunnel by cryo-electron microscopy. Preprint at https://www.biorxiv.org/content/10.1101/722611v3.full (2019).

29. Waudby, C. A. et al. Systematic mapping of free energy landscapes of a growing filamin domain during biosynthesis. *Proc. Natl Acad. Sci. USA* **115**, 9744–9749 (2018).

30. Lu, J. & Deutsch, C. Pegylation: a method for assessing topological accessibilities in Kv1.3. *Biochemistry* **40**, 13288–13301 (2001).

31. Bose, D. & Chakrabarti, A. Substrate specificity in the context of molecular chaperones. *IUBMB Life* **69**, 647–659 (2017).

32. Schibich, D. et al. Global profiling of SRP interaction with nascent polypeptides. *Nature* **536**, 219–223 (2016).

33. Patzelt, H. et al. Binding specificity of *Escherichia coli* trigger factor. *Proc. Natl Acad. Sci. USA* **98**, 14244–14249 (2001).

34. Tomic, S., Johnson, A. E., Hartl, F. U. & Etchells, S. A. Exploring the capacity of trigger factor to function as a shield for ribosome bound polypeptide chains. *FEBS Lett.* **580**, 72–76 (2005).

35. Saio, T., Guan, X., Rossi, P., Economou, A. & Kalodimos, C. G. Structural basis for protein antiaggregation activity of the trigger factor chaperone. *Science* **344**, 1250494–1250494 (2014).

36. Huang, C., Rossi, P., Saio, T. & Kalodimos, C. G. Structural basis for the antifolding activity of a molecular chaperone. *Nature* **537**, 202–206 (2016).

37. Rüdiger, S., Germeroth, L., Schneider-Mergener, J. & Bukau, B. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* **16**, 1501–1507 (1997).

38. Sekhar, A., Rosenzweig, R., Bouvignies, G. & Kay, L. E. Hsp70 biases the folding pathways of client proteins. *Proc. Natl Acad. Sci. USA* **113**, E2794–E2801 (2016).

39. He, L., Sharpe, T., Mazur, A. & Hiller, S. A molecular mechanism of chaperone-client recognition. *Sci. Adv.* **2**, e1601625 (2016).

40. Libich, D. S., Tugarinov, V. & Clore, G. M. Intrinsic unfoldase/foldase activity of the chaperonin GroEL directly demonstrated using multinuclear relaxation-based NMR. *Proc. Natl Acad. Sci. USA* **112**, 8817–8823 (2015).

41. Jiang, Y., Rossi, P. & Kalodimos, C. G. Structural basis for client recognition and activity of Hsp40 chaperones. *Science* **365**, 1313–1319 (2019).

42. Faust, O. et al. HSP40 proteins use class-specific regulation to drive HSP70 functional diversity. *Nature* **587**, 489–494 (2020).

43. Borgia, M. B. et al. Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins. *Nature* **474**, 662–665 (2011).

44. Shieh, Y.-W. et al. Operon structure and cotranslational subunit association direct protein assembly in bacteria. *Science* **350**, 678–680 (2015).

45. Shiber, A. et al. Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature* **561**, 268–272 (2018).

46. Bertolini, M. et al. Interactions between nascent proteins translated by adjacent ribosomes drive homomer assembly. *Science* **371**, 57–64 (2020).

47. Pechmann, S., Willmund, F. & Frydman, J. The ribosome as a hub for protein quality control. *Mol. Cell* **49**, 411–421 (2013).

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

Methods

Sample preparation and quality control. DNA constructs of tandem immunoglobulin domains FLN5 and FLN6, and of poly(GS) glycine serine repeat sequence variants and Y719E variants, were described previously[20]. Aromatic and glutamate mutants were generated using site-directed mutagenesis. A stronger SecM variant was used for preparation of [^{2}H,^{13}CH$_3$-Ile, Leu and Val (ILV)]-labelled RNCs for PEGylation experiments, with sequence 'FSTPVWIWWWPRIRGPPPPWT'. ^{15}N and [^{2}H,^{13}CH$_3$-ILV]-labelled proteins were grown and purified as described previously[49], except in the case of disordered FLN5 variants where the final chromatography step was performed using a HiLoad 16/600 Superdex 200 pg (GE Healthcare) column in the presence of 6 M urea, followed by buffer exchange into Tico buffer[25]. ^{15}N and [^{2}H,^{13}CH$_3$-ILV]-labelled RNCs were generated in BL21(DE3) *Escherichia coli* as described previously[25], where the only modification was the replacement of the final step of purification (sucrose gradient chromatography) with an additional sucrose cushion step followed by hydrophobic affinity chromatography using a HiTrap FF butyl column (GE Healthcare). The occupancy and integrity of RNC samples were monitored as described previously[25].

NMR spectroscopy. NMR data were acquired on 500 and 700 MHz Bruker Avance III, and 800 and 950 MHz Bruker Avance III HD spectrometers, all equipped with TCI cryoprobes. Data were processed and analysed with nmrPipe, Sparky, CCPN Analysis and MATLAB[50,51]. All samples were prepared at nascent chain concentrations of between 7 and 13 μM in Tico buffer, pH 7.5, containing 10% D$_2$O and 0.001% (w/v) DSS[25]. For ^{15}N-labelled samples, ^{1}H,^{15}N SOFAST-HMQC experiments[52] were recorded with typical acquisition times of 50.4 ms and 29.5 ms in the direct and indirect dimensions, respectively, and a 100 ms recycle delay. Transverse cross-correlated relaxation was measured using ^{1}H,^{15}N BEST-TROSY-CCR experiments (Extended Data Fig. 4), acquired at 950 MHz with a recycle delay of 200 ms and an acquisition time of 49 ms in the direct dimension. For measurements of isolated proteins, a relaxation delay of 111 ms was used, while shorter relaxation delays were used for measurements of RNCs (22 ms for FLN5+31; 44 ms for FLN5+31 F$_3$A$_3$ and FLN5+31 A$_3$A$_3$). An acquisition time of 43 ms was used in the indirect dimension for all measurements except for FLN5+31, for which a 21 ms acquisition time was used. Samples were doped with 15 mM NiDO2A (nickel(II) 1,4,7,10-tetraazacyclododecane-1,7-bis(acetic acid)) to enhance sensitivity (Extended Data Fig. 4)[53]. All experiments were interleaved with SORDID diffusion measurements[54] with a diffusion delay of 190 ms, an acquisition time of 49 ms, and 4 ms trapezoidal gradient pulses with strengths of 0.058 and 0.387 T m^{-1}. Ribosome background labelling was assessed using ^{15}N filtered/edited difference spectroscopy[25]. Chemical shift assignments for FLN5 variants were verified using ^{1}H,^{15}N NOESY-HSQC and TOCSY-HSQC experiments, with 200 ms and 70 ms mixing times, respectively.

For [^{2}H,^{13}CH$_3$-ILV]-labelled samples, ^{1}H,^{13}C HMQC spectra were acquired at 800 MHz with acquisition times of 100 ms and 7 ms in the direct and indirect dimensions, respectively. Sample integrity was assessed through ^{13}C-edited ^{1}H STE diffusion measurements[25] and ^{1}H R_2 relaxation measurements (^{13}C-edited and incorporating a filter to select slow relaxing coherences[55]), in which signals detected after a 100 ms relaxation delay were interpreted as indicating nascent chain release.

^{1}H R_2 rates of isolated amino acids (200 μM in D$_2$O) were measured in the presence and absence of 1 μM 70S ribosomes, using a 500 Hz PROJECT pulse train (700 MHz, 283 K)[56].

CD spectroscopy. CD spectra were acquired using a Chirascan-plus CD spectrometer (Applied Photophysics). All samples (20 μM) were incubated for ≥3 h at 283 K prior to measurement in Tico buffer (12 mM Hepes, 30 mM NH$_4$Cl, 6 mM MgCl$_2$, pH 7.5). CD signals at 211 nm were fitted globally using MatLab to a two-state unfolding model with a common m value, $\langle m_{\mathrm{D-N}} \rangle$ (ref. [57]):

$$y = \alpha_N + \alpha_D \frac{e^{\frac{\langle m_{\mathrm{D-N}} \rangle ([\mathrm{D}] - [\mathrm{D}]_{50\%})}{RT}}}{1 + e^{\frac{\langle m_{\mathrm{D-N}} \rangle ([\mathrm{D}] - [\mathrm{D}]_{50\%})}{RT}}} \quad (2)$$

where [D] is the denaturant concentration, [D]$_{50\%}$ is the midpoint of folding, and α terms represent baselines for native (N) and denatured states. Stabilities were then calculated as $\Delta G_{\mathrm{D-N}} = m_{\mathrm{D-N}} [\mathrm{D}]_{50\%}$.

Molecular dynamics. Simulations of FLN5+31, +42, +67 and +110 RNCs were run in Gromacs 4.5.7[58] using Cα structure-based models generated with SMOG 2.0[59,60], extended to represent RNA using three beads located at P, C4' and N3 atoms. A rigid ribosome model was created based on the structure 4ybb[61] and including only nascent chain-accessible residues, that is, the exit tunnel and the surrounding surface. The A$_3$A$_3$ variant was modelled by removing all contacts involving the mutated sites, and the simulation temperature was then tuned so that isolated wt and A$_3$A$_3$ variants were folded and unfolded, respectively. Electrostatic interactions were introduced using Debye–Hückel theory[62], with parameters chosen to reproduce the experimentally observed bound population of the FLN5+42 RNC and then applied without change to the remaining lengths. Three independent simulations were carried out for each system, with ~2 × 10^8 steps per run.

PEGylation. Samples (100–200 pmol isolated proteins or 8 pmol of RNC) were incubated in increasing urea concentrations for at least 2 h at 283 K prior to measurements. Methoxypolyethylene glycol maleimide (PEG-Mal) was added to a final concentration of 2 mM and incubated for 10 min before quenching the reaction with 0.23 M DTT. Samples were separated on denaturing 12% (w/v) polyacrylamide Bis-Tris gels (pH 5.7) which were subsequently Coomassie stained (for isolated proteins) or analysed by Western blot (for RNCs). Quantitation was performed using ImageStudio (Licor). As a control, Western blots of dilutions of a FLN5+47 RNC variant showed excellent linearity ($R^2 = 0.98$) over the concentration range used in analyses (Supplementary Fig. 14). Populations of unfolded (PEGylated) nascent chains were determined relative to the total nascent chain concentration (PEGylated and unPEGylated) and fitted to equation (2) to determine the thermodynamic stability as described above.

Data availability

Data supporting the findings of this study are included in the article, source data and Supplementary Information files. Assignments have been deposited in the BMRB under accession codes 51023 and 51028. Source data are provided with this paper.

Code availability

NMR pulse sequences are available on https://github.com/chriswaudby/pp and also as Supplementary Information. Source data are provided with this paper.

References

49. Hsu, S.-T. D. et al. Structure and dynamics of a ribosome-bound nascent chain by NMR spectroscopy. *Proc. Natl Acad. Sci. USA* **104**, 16516–16521 (2007).
50. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
51. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
52. Schanda, P. & Brutscher, B. 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**, 8014–8015 (2005).
53. Chan, S. H. S., Waudby, C. A., Cassaignau, A. M. E., Cabrita, L. D. & Christodoulou, J. Increasing the sensitivity of NMR diffusion measurements by paramagnetic longitudinal relaxation enhancement, with application to ribosome-nascent chain complexes. *J. Biomol. NMR* **63**, 151–163 (2015).
54. Augustyniak, R., Ferrage, F., Damblon, C., Bodenhausen, G. & Pelupessy, P. Efficient determination of diffusion coefficients by monitoring transport during recovery delays in NMR. *Chem. Commun.* **48**, 5307–5309 (2012).
55. Tugarinov, V., Ollerenshaw, J. E. & Kay, L. E. Dipolar dynamic frequency shifts in multiple-quantum spectra of methyl groups in proteins: correlation with side-chain motion. *Magn. Reson. Chem.* **44**, S122–S129 (2006).
56. Aguilar, J. A., Nilsson, M., Bodenhausen, G. & Morris, G. A. Spin echo NMR spectra without J modulation. *Chem. Commun.* **48**, 811–813 (2012).
57. Pace, C. N. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* **131**, 266–280 (1986).
58. Pronk, S. et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).
59. JK, N. et al. SMOG 2: A versatile software package for generating structure-based models. *PLoS Comput. Biol.* **12**, e1004794 (2016).
60. Clementi, C., Nymeyer, H. & Onuchic, J. N. Topological and energetic factors: what determines the structural details of the transition state ensemble and 'en-route' intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* **298**, 937–953 (2000).
61. J, N. et al. High-resolution structure of the *Escherichia coli* ribosome. *Nat. Struct. Mol. Biol.* **22**, 336–341 (2015).
62. Noel, J. K. & Whitford, P. C. How EF-Tu can contribute to efficient proofreading of aa-tRNA by the ribosome. *Nat. Commun.* **7**, 13314 (2016).

Author contributions

Competing interests

The authors declare no competing interests.

Additional information

Extended data are available for this paper at https://doi.org/10.1038/s41557-021-00796-x.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41557-021-00796-x.

Correspondence and requests for materials should be addressed to Christopher A. Waudby or John Christodoulou.

Peer review information *Nature Chemistry* thanks Bjoern Burmann and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.
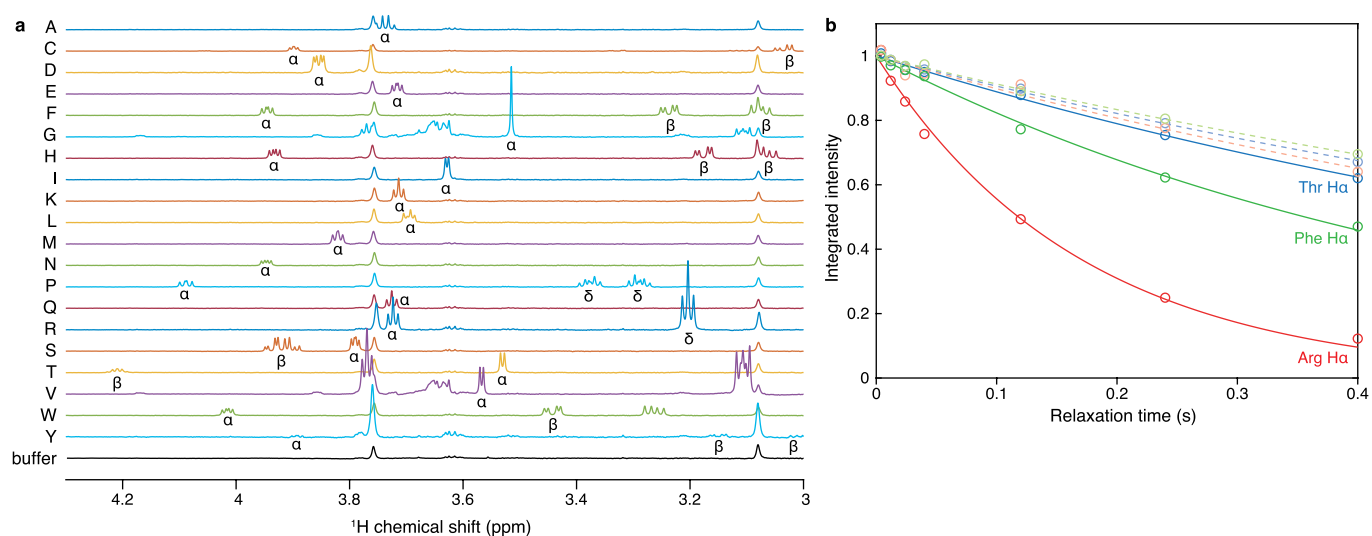
Extended Data Fig. 1 | Interaction of free amino acid with empty 70S ribosomes. (**a**) Hα region of 1D ^{1}H NMR spectra of isolated amino acids (700 MHz, 283 K), showing atom assignments. (**b**) Representative ^{1}H R_2 relaxation measurements for the indicated amino acid resonances in the presence (solid lines) and absence (dashed lines) of 1 μM 70S ribosomes.
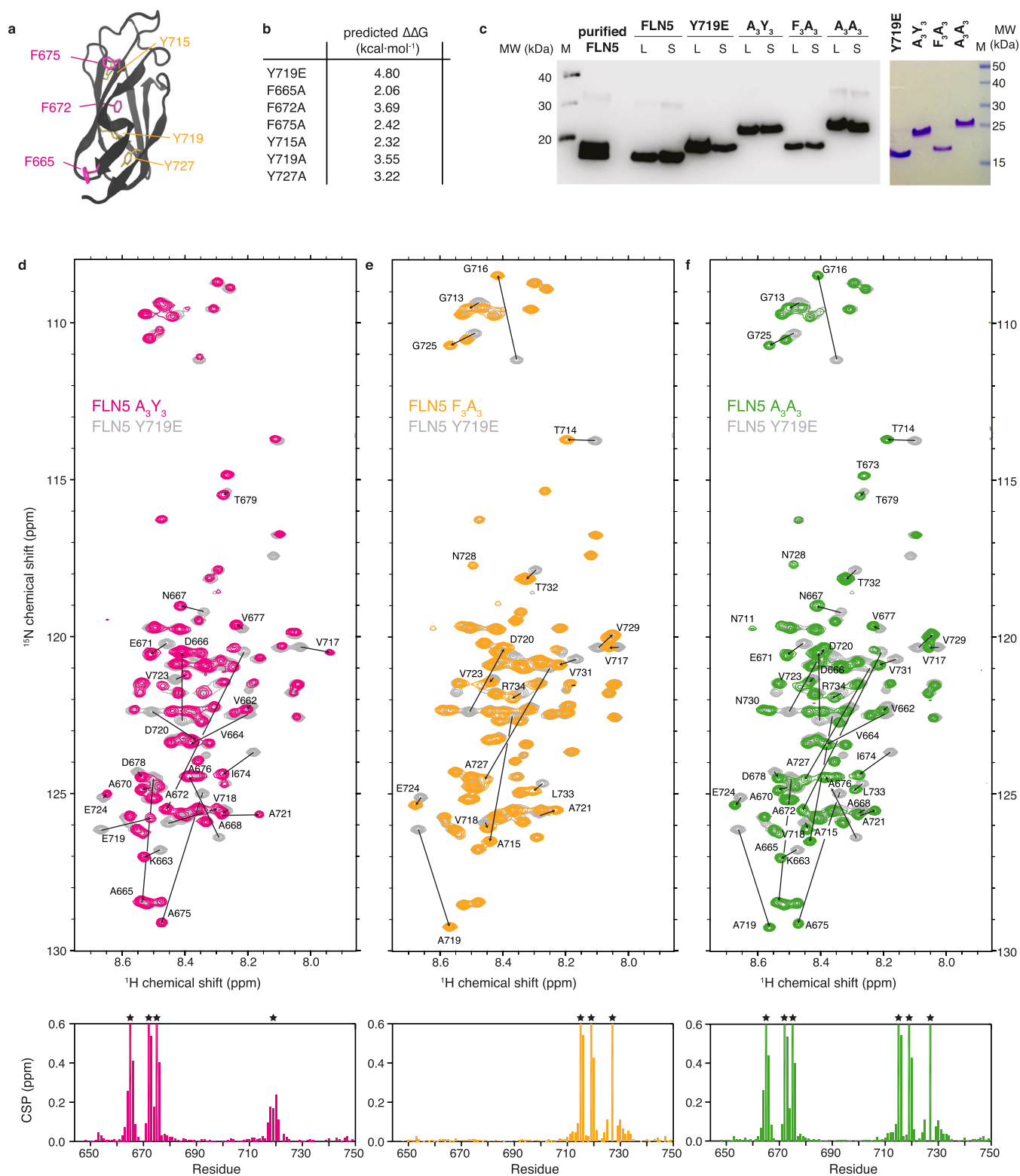
**Extended Data Fig. 2 | Biochemical and NMR analysis of isolated aromatic variants.** (**a**) Aromatic residues that have been mutated in this study mapped onto the PDB structure of FLN5 (PDB 1QFH). (**b**) POPMUSIC stability predictions of the FLN5 aromatic mutations[65]. (**c**) Western blot (anti-His tag) of an expression text of FLN5, FLN5 Y719E, FLN5 A_3Y_3, FLN5 F_3A_3 and FLN5 A_3A_3 proteins (L: total lysate, S: soluble lysate) and Coomassie stained SDS PAGE gel of purified FLN5 Y719E, FLN5 A_3Y_3, FLN5 F_3A_3 and FLN5 A_3A_3 proteins. (**d**–**f**) ^{1}H,^{15}N correlation spectrum (283 K) of FLN5 Y719E overlaid with that of (**d**) FLN5 A_3Y_3, (**e**) FLN5 F_3A_3 and (**f**) FLN5 A_3A_3, with assignments of the shifted resonances indicated. Combined amide chemical shift perturbations, $\Delta\delta = \sqrt{\Delta\delta_H^2 + (\Delta\delta_N/5)^2}$, are plotted below. The mutation sites relative to FLN5 Y719E have been indicated with star symbols on the CSP plot.
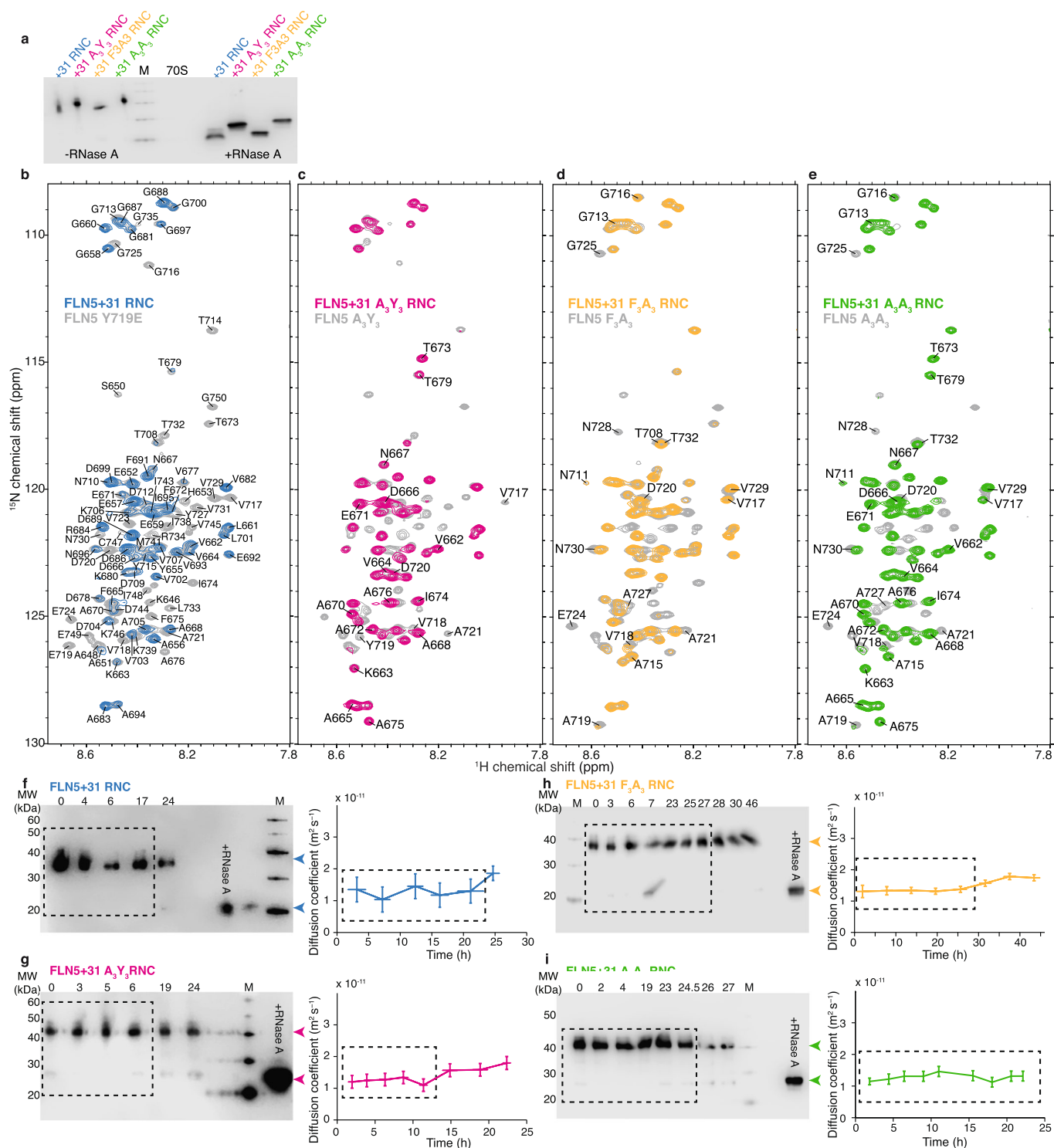
Extended Data Fig. 3 | Biochemical and NMR analysis of aromatic RNC variants. (**a**) Anti-His tag western blot of FLN5+31 wt, A_3Y_3, F_3A_3 and A_3A_3 RNCs (10 pmol each) purified for NMR studies, in their tRNA-bound and released forms. The migration patterns on SDS page reflect those of the isolated mutant proteins (cf. Extended Data Fig. 2). (**b–e**) ^{1}H,^{15}N SOFAST-HMQC spectra (283 K, 950 MHz) of (b) FLN5+31, (c) FLN5+31 A_3Y_3, (d) FLN5+31 F_3A_3 and (e) FLN5+31 A_3A_3 RNCs against the corresponding isolated, unfolded FLN5 variants, with assignments. (**f–i**) Assessment of the integrity of ^{15}N-labelled RNCs[25]: Anti-His western blots and translational diffusion coefficients measured using ^{15}N SORDID experiments are shown. The timeframes during which nascent chains have been assessed as being intact and used for analysis are indicated (dashed boxes). Vertical error bars represent standard errors derived from the spectral noise; horizontal bars (where applicable) indicate the acquisition periods of diffusion measurements.
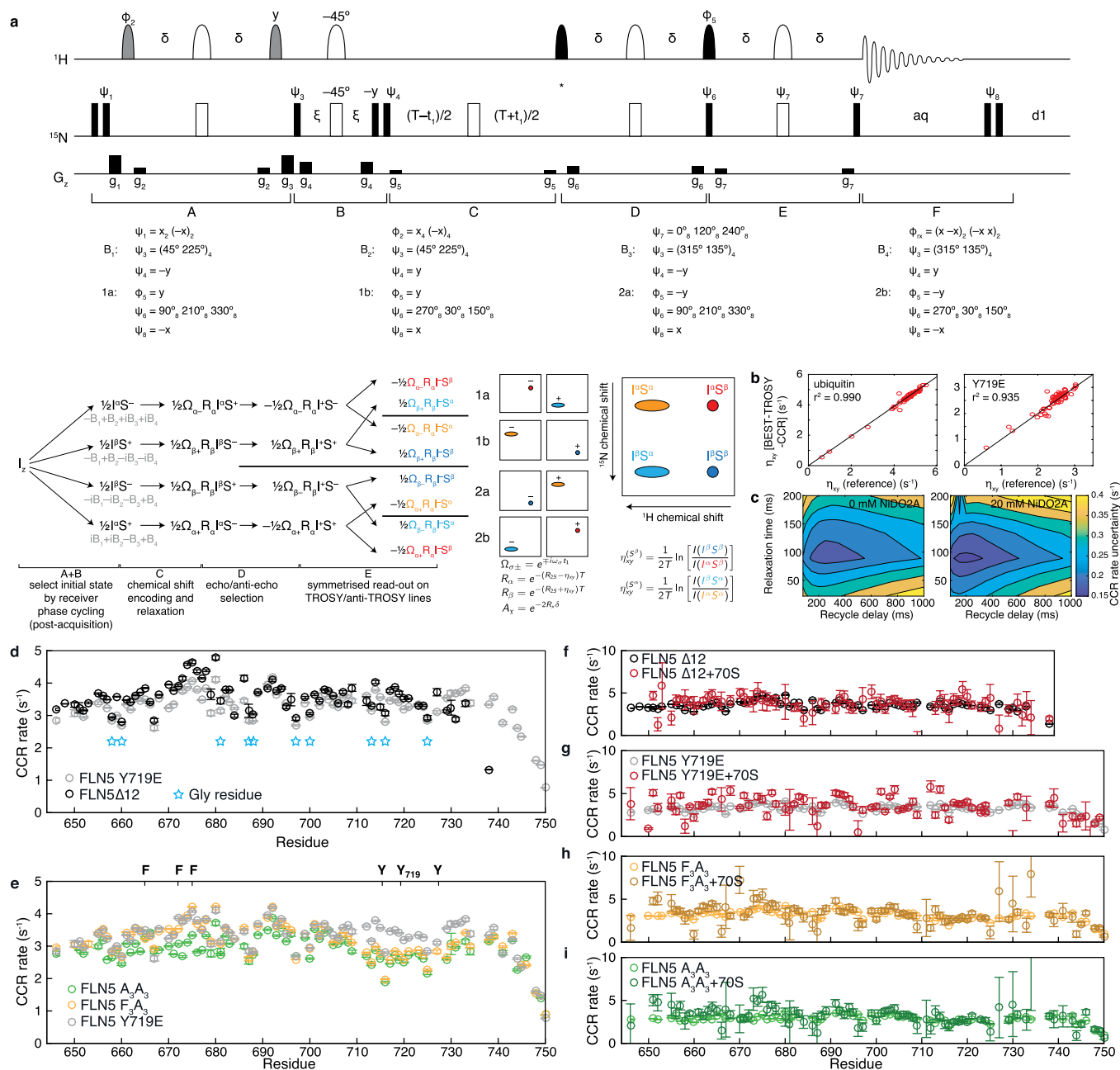
Extended Data Fig. 4 | The BEST-TROSY-CCR experiment and application to isolated FLN5 variants. (**a**) Pulse sequence and processing scheme for the BEST-TROSY-CCR experiment. Solid coloring indicates 90° pulses; empty shapes indicate 180° pulses. ^{1}H pulse lengths are calculated for application at an offset of 8.2 ppm (950 MHz). Gray ^{1}H 90° pulses are Pc9_4_90 pulses (1958 µs), solid ^{1}H 90° shapes indicate Eburp2 pulses (1251 µs, * indicates time reversed pulse), and ^{1}H 180° shapes indicate Reburp pulses (1432 µs). ^{15}N pulses are applied as high-power rectangular pulses. 16 subspectra are acquired, corresponding to the possible combinations of the phase programs B_1–B_4 and the phase programs 1a,1b,2a and 2b as indicated. All pulses are applied with a phase of x (0°) unless otherwise indicated. ^{15}N phases are shown as spin dynamical phases and must be modified for application on Bruker spectrometers[66]. Gradients are applied as SMSQ shapes: g_1 (31.4%, 1ms), g_2 (23%, 300 µs), g_3 (21%, 1ms), g_4 (31%, 300 µs), g_5 (11%, 1ms), g_6 (16.7%, 300 µs), g_7 (45%, 300 µs). (**b**) Comparison of CCR rates measured for ubiquitin and FLN5 Y719E using our pulse sequence (111 ms relaxation delay, 283 K) against a reference measurement using symmetric recombination[67]. (**c**) Sensitivity enhancement on the CCR measurements of FLN5 Y719E (283 K, 950 MHz) obtained using the PLRE agent NiDO2A, an inert Ni^{II} chelate. (**d**) CCR rates of FLN5 Y719E and FLN5Δ12. Blue stars indicate Gly residues, which on average have lower CCR rates. (**e**) CCR rates of FLN5 Y719E, FLN5 F$_3$A$_3$ and FLN5 A$_3$A$_3$. (**f**–**i**) Comparison of CCR rates of FLN5 variants at 10 µM, in presence and absence of an equimolar concentration of 70 S ribosomes: (f) FLN5Δ12, (g) FLN5 Y719E, (h) FLN5 F$_3$A$_3$ and (i) FLN5 A$_3$A$_3$. All errors were derived from the spectral noise.
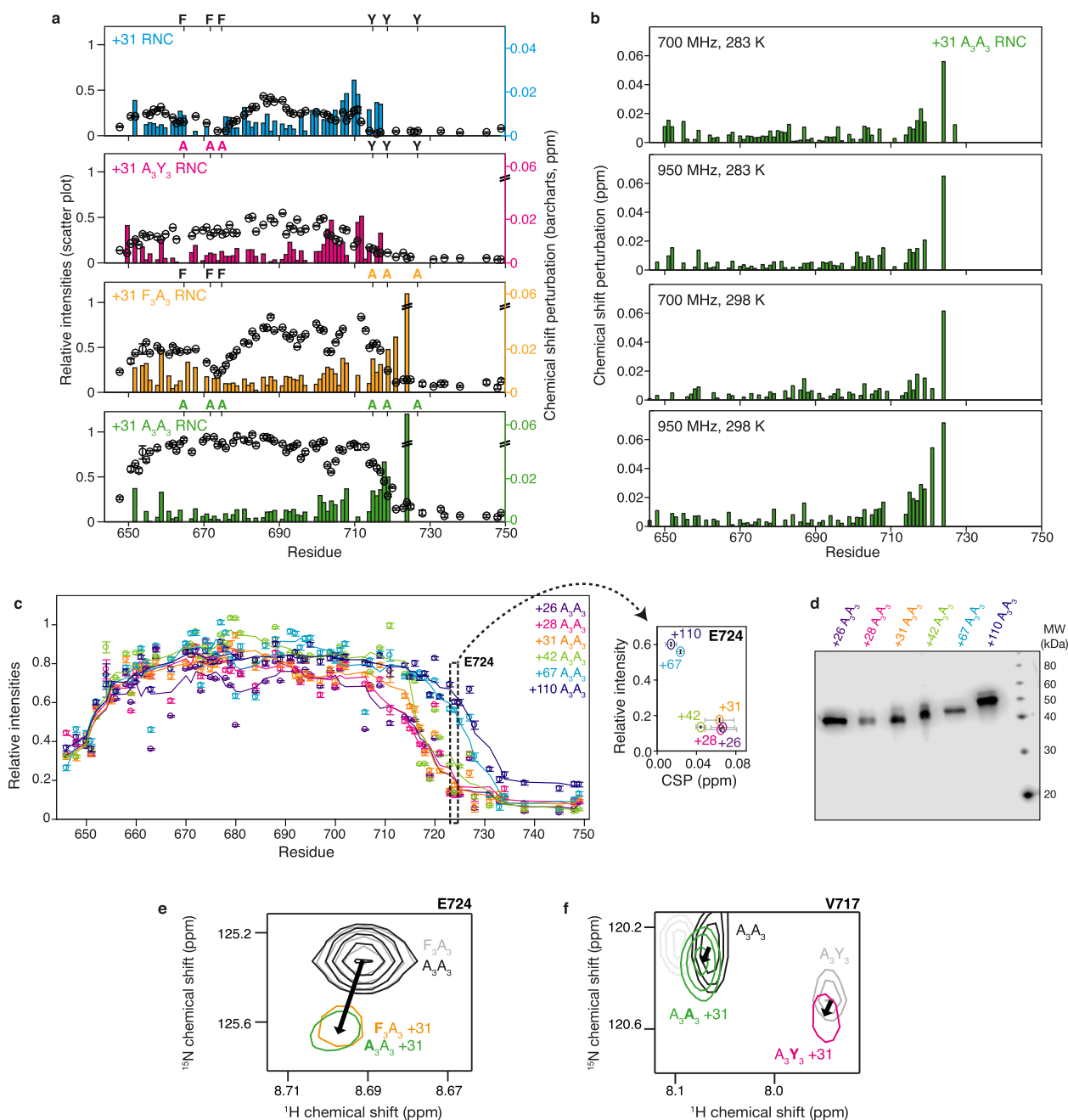
Extended Data Fig. 5 | CSPs of FLN5 RNCs against isolated FLN5 variants, and comparison of chemical shift perturbations in the C-terminal region of FLN5 RNCs between aromatic cluster variants. (**a**) Relative intensities (LH axis) of FLN5+31, FLN5+31 A_3Y_3, FLN5+31 F_3A_3 and FLN5+31 A_3A_3 RNCs (283 K) and chemical shift perturbations (RH axis, $\Delta\delta = \sqrt{\Delta\delta_H^2 + (\Delta\delta_N/5)^2}$), relative to their corresponding isolated, unfolded FLN5 variants. (**b**) Chemical shift perturbations of FLN5+31 A_3A_3 RNC relative to isolated FLN5 A_3A_3, at fields and temperatures as indicated. (**c**) Relative intensities of FLN5+26, 28, 31, 42, 67, 110 A_3A_3 RNCs relative to isolated FLN5 A_3A_3 (283 K, 950 MHz). inset: Intensities and combined amide chemical shift perturbations ($\Delta\delta = \left[\Delta\delta_H^2 + (\Delta\delta_N/5)^2\right]^{1/2}$) of the E724 amide resonance observed relative to isolated FLN5 A_3A_3 in ^{1}H,^{15}N SOFAST-HMQC spectra of FLN5 A_3A_3 RNCs, with varying linker lengths as indicated. (**d**) Anti-His tag western blot of FLN5+26, 28, 31, 42, 67, 110 A_3A_3 RNCs purified for NMR studies, in their tRNA-bound form. (**e**) Comparison of C-terminal chemical shift perturbations between A_3A_3 and F_3A_3 variants. The difference in perturbations between A_3A_3 and F_3A_3 variants was 0.009 ± 0.050 ppm (^{15}N) and 0.002 ± 0.006 ppm (^{1}H). (**f**) Comparison of C-terminal chemical shift perturbations between A_3A_3 and A_3Y_3 variants. Due to chemical shift perturbations between the A_3A_3 and A_3Y_3 variants, and the reduced intensity of C-terminal residues in the A_3Y_3 variant (Fig. 2b), E724 cannot be resolved in the A_3Y_3 spectrum and residue V717 is analyzed instead. The difference in perturbations between A_3A_3 and A_3Y_3 variants was 0.060 ± 0.044 ppm (^{15}N) and 0.001 ± 0.006 ppm (^{1}H). Based on these negligible differences, mutations in the F_3 and Y_3 clusters do not seem to affect the CSP which report on the ribosome interaction of the C-terminal segment (residues N728 to C747).
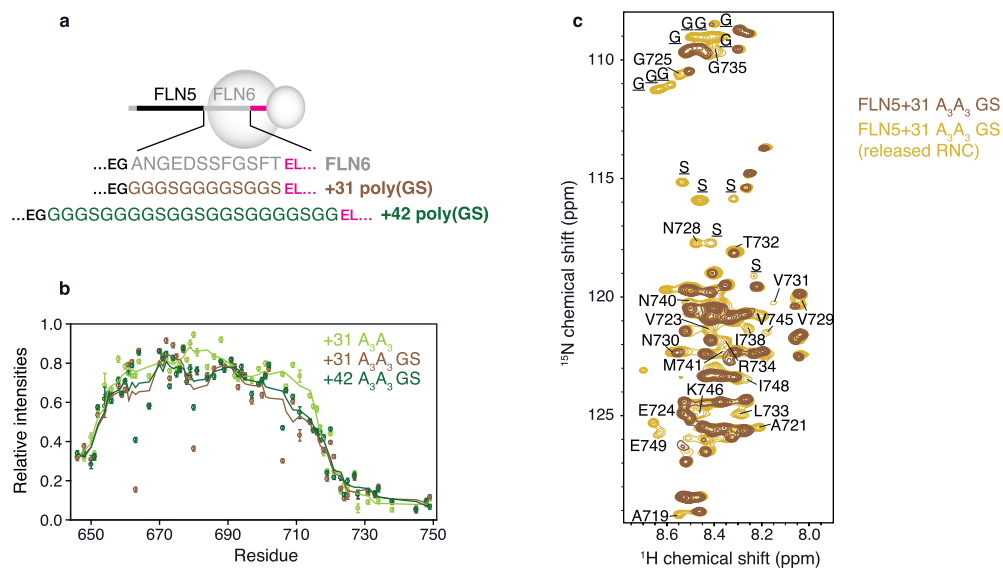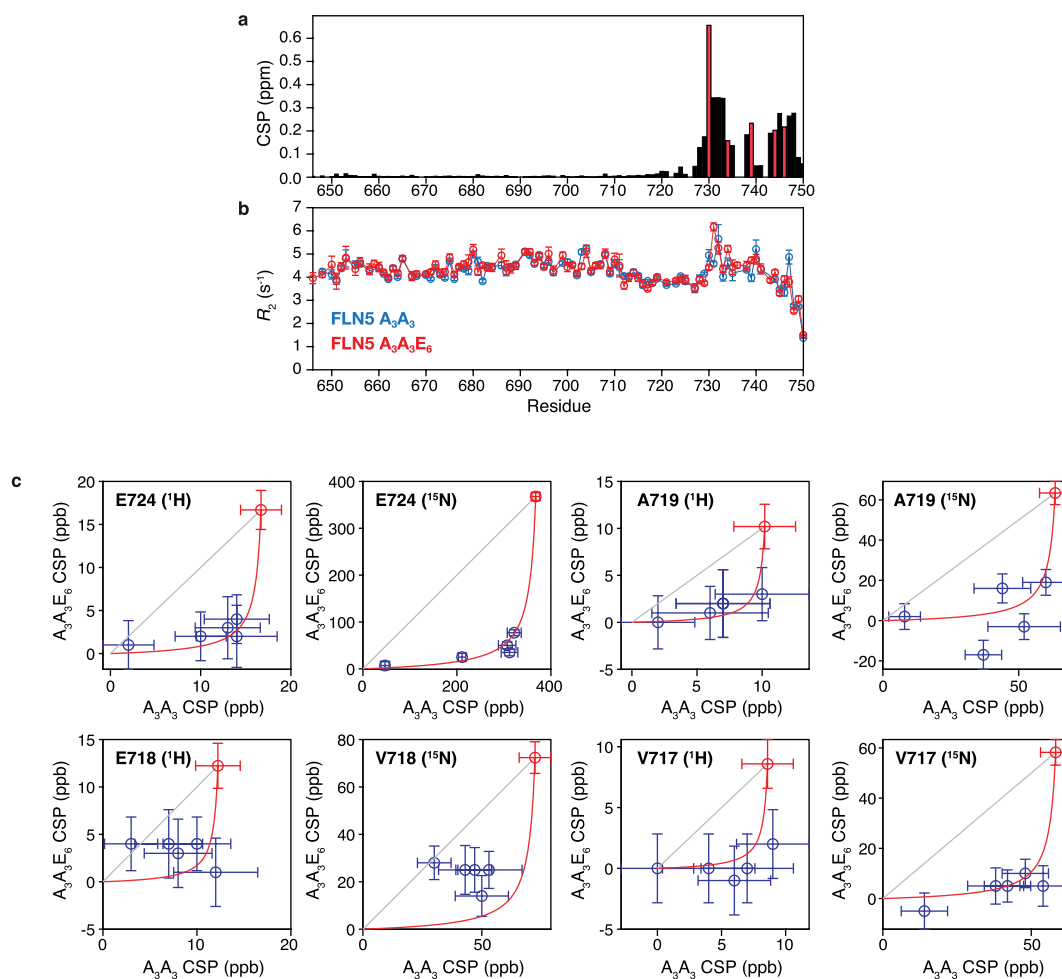
Extended Data Fig. 6 | Analysis of FLN5-poly(GS) RNCs. (**a**) Design of GS linker RNCs. (**b**) ^{1}H,^{15}N SOFAST-HMQC resonances intensities for FLN5+31 A_3A_3, +31 A_3A_3 GS and +42 A_3A_3 GS RNCs relative to isolated FLN5 A_3A_3. (**c**) Overlay of a ^{1}H,^{15}N spectrum of FLN5+31 A_3A_3 GS RNC and a ^{1}H,^{15}N spectrum of the same RNC after treatment with 20 mM EDTA to induce nascent chain release. The residues only observable in EDTA-treated FLN5+31 A_3A_3 GS RNC are assigned, including resonances marked 'G' and 'S' which we attribute to the poly(GS) linker.
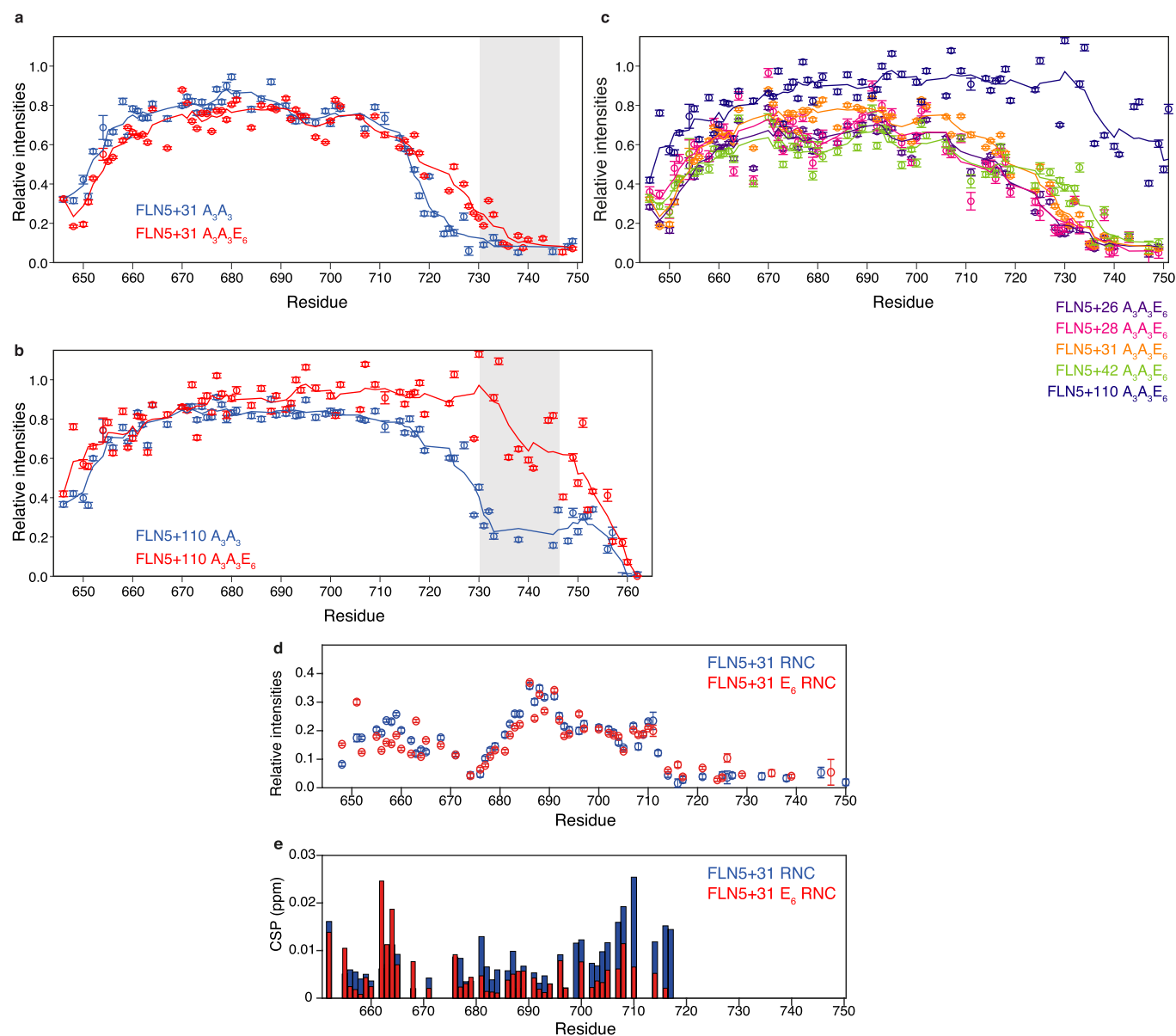
Extended Data Fig. 7 | Characterization of FLN5 E6 variants. (**a**) ^{1}H,^{15}N SOFAST-HMQC resonance intensities of FLN5+31 A_3A_3 RNC relative to isolated FLN5 A_3A_3, and of FLN5+31 $A_3A_3E_6$ RNC relative to isolated FLN5 $A_3A_3E_6$ (283 K, 950 MHz). The region between residues 730–746 is highlighted in light grey. (**b**) Same as in (a), except for FLN5+110 A_3A_3 and $A_3A_3E_6$ RNCs. (**c**) ^{1}H,^{15}N SOFAST-HMQC resonance intensities of FLN5+26, 28, 31, 42, 110 A_3A_3 E_6 RNCs relative to isolated FLN5 $A_3A_3E_6$ (283 K, 950 MHz). All errors were derived from the spectral noise. (**d**) Relative intensities of FLN5+31 and FLN5+31 E_6 RNCs (283 K) relative to their corresponding isolated, unfolded FLN5 variants. (**e**) Chemical shift perturbations ($\Delta\delta = \sqrt{\Delta\delta_H^2 + (\Delta\delta_N/5)^2}$) of FLN5+31 and FLN5+31 E_6 RNCs (283 K) relative to their corresponding isolated, unfolded FLN5 variants.

Extended Data Fig. 8 | Global analysis of FLN5 A3A3 and A3A3E6 chemical shift perturbations. (**a**) Chemical shift perturbations of FLN5 A_3A_3 vs. FLN5 $A_3A_3E_6$ ($\Delta\delta = \sqrt{\Delta\delta_H^2 + (\Delta\delta_N/5)^2}$). E_6 mutation sides are indicated using red bar charts. (**b**) T_2 relaxation measurements of isolated FLN5 A_3A_3 and FLN5 $A_3A_3E_6$ (**c**) Correlation plots of ^1H and ^{15}N chemical shift perturbations in +26, +28, +31, +42 and +110 $A_3A_3/A_3A_3E_6$ RNCs for residues as indicated (283 K, 950 MHz). A global fit of $\Delta\Delta G_{Ubound-Ufree}^{A3A3-A3A3E6}$ is shown (red line) and red markers indicate the fitted values of $\Delta\delta_{max}$ ($\chi^2 = 70.17$, dof = 66, $\chi^2/\nu = 1.06$).
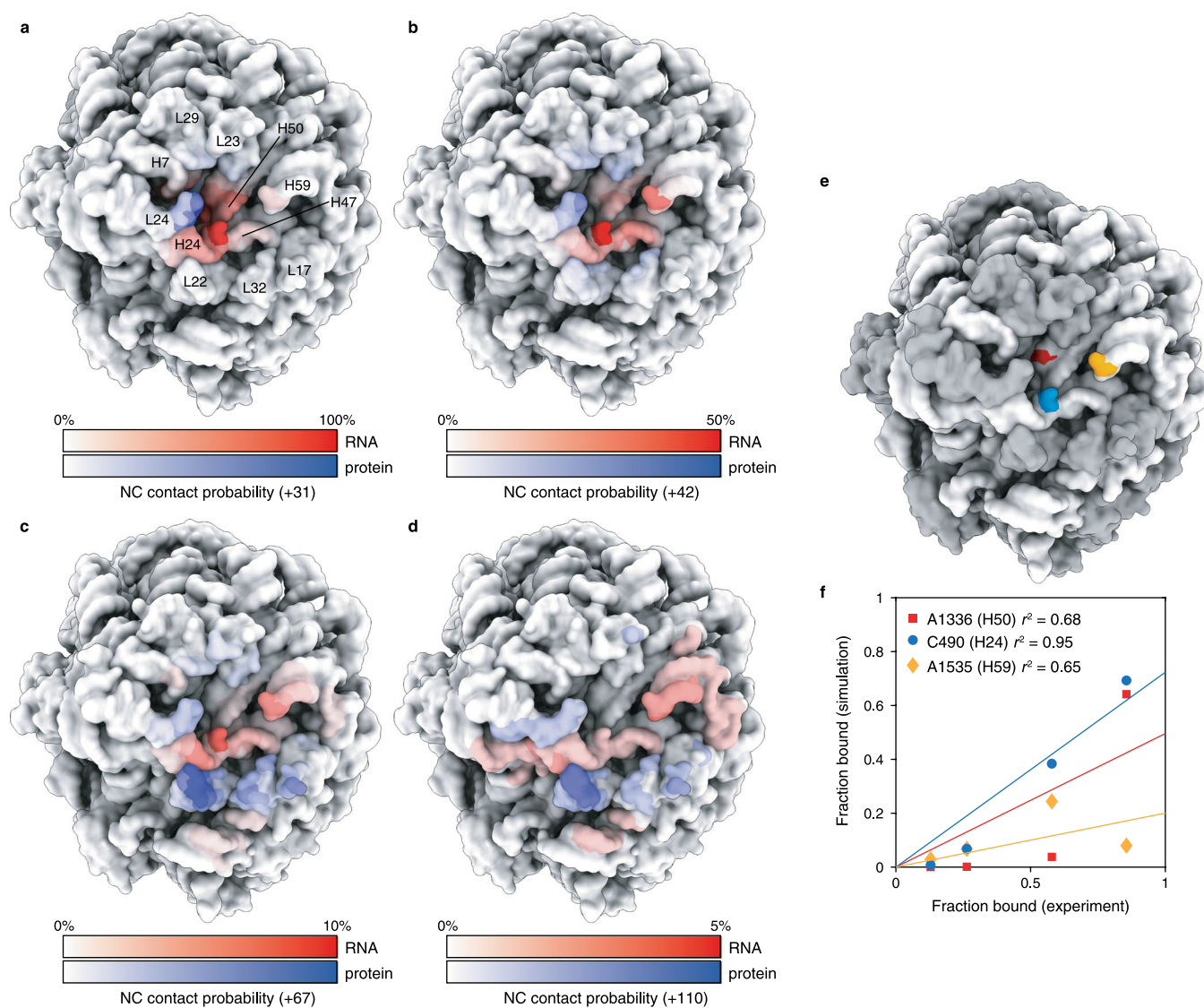
Extended Data Fig. 9 | Molecular modelling of interactions between the FLN5 C-terminal region and the ribosome surface. (**a**–**d**) Contact probabilities between ribosome protein and RNA residues and the C-terminal region of FLN5 nascent chains, determined through coarse grained molecular dynamics simulations, for linker lengths (a) +31, (b) +42, (c) +67, and (d) +110. (**e**) Coarse-grained ribosome structure highlighting the location of RNA residues A1336, C490 and A1535, and (**f**) correlation plots between simulated and experimentally determined nascent chain–ribosome interactions for these residues. Data are shown for linker lengths +31, +42, +67 and +110, and best fit lines through the origin are plotted with correlation coefficients as indicated.
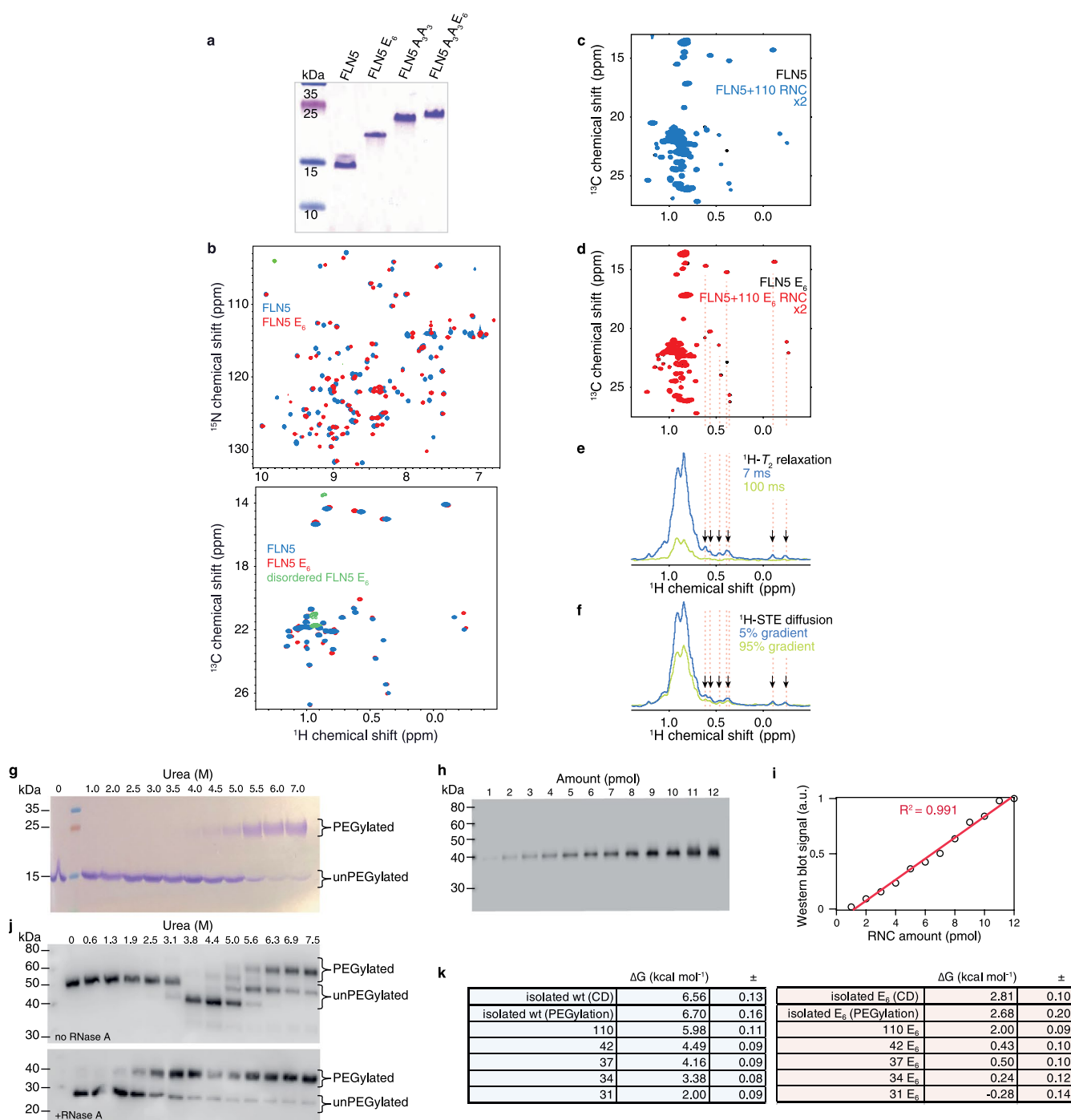
Extended Data Fig. 10 | Folding of FLN5 and FLN5 E_6 on and off the ribosome. (**a**) Coomassie-stained SDS PAGE gel of purified FLN5 protein variants. The net charge of FLN5 varies from -8.7 to -16.7 at pH 7.5 through the introduction of Glu residues, likely explaining the altered migration pattern on PAGE. (**b**) ^{1}H,^{15}N correlation spectra of ^{15}N-labelled proteins (backbone), and ^{1}H,^{13}C correlation spectra of [^{2}H,^{13}CH$_3$-ILV]-labelled proteins (sidechains). (**c**) ^{1}H,^{13}C HMQC spectrum of [^{2}H,^{13}CH$_3$-ILV]-labelled FLN5+110 RNC. (**d**) Same as c, but for FLN5+110 E_6 RNC. (**e**) Representative quality control measurements for nascent chain attachment by analysis of ^{1}H transverse relaxation. Dispersed methyl resonances (indicated by arrows) detected after a 100 ms relaxation delay were taken as an indication of nascent chain release. (**f**) Representative quality control measurement by measurement of translational diffusion of dispersed nascent chain resonances (indicated by arrows) using ^{1}H STE measurements with relative gradient strengths as indicated. (**g**) As a representative example, isolated FLN5 C721 V747 is shown on a Coomassie-stained 12% Bis-tris SDS-PAGE, after PEGylation measurements in urea. (**h**) As a representative example, FLN5+110 C721 V747 RNC is shown on a 12% Bis-tris SDS-PAGE, detected via anti-His tag western blot, after PEGylation measurements in urea. Intermediate concentrations of urea cause the tRNA-bound form of the nascent chain to migrate faster on Bis-Tris gels, as shown by the uniform migration pattern when the tRNA is degraded using RNase A, however this does not interfere with quantitation. (**i**) Western blot of 1-12 pmol of a FLN5+47 RNC. (**j**) Correlation plot of the western blot signal vs. concentration, from the blot shown in i. (**k**) Tabulated results from the fits of the data shown in Fig. 5b. The analysis was restricted to datapoints up to 3 M urea, and 3.75 M urea for FLN5+110 RNC[17] (cf. solid markers in Fig. 5b).