Ribosomal synthesis and folding of peptide-helical aromatic foldamer hybrids

Joseph M. Rogers, Sunbum Kwon, Simon J. Dawson, Pradeep K. Mandal and Ivan Huc

Translation, the mRNA-templated synthesis of peptides by the ribosome, can be manipulated to incorporate variants of the 20 cognate amino acids. Such approaches for expanding the range of chemical entities that can be produced by the ribosome may accelerate the discovery of molecules that can perform functions for which poorly folded, short peptidic sequences are ill suited. Here, we show that the ribosome tolerates some artificial helical aromatic oligomers, so-called foldamers. Using a flexible tRNA-acylation ribozyme—flexizyme—foldamers were attached to tRNA, and the resulting acylated tRNAs were delivered to the ribosome to initiate the synthesis of non-cyclic and cyclic foldomer–peptide hybrid molecules. Passing through the ribosome exit tunnel requires the foldamers to unfold. Yet foldamers encode sufficient folding information to influence the peptide structure once translation is completed. We also show that in cyclic hybrids, the foldomer portion can fold into a helix and force the peptide segment to adopt a constrained and stretched conformation.

In peptide biosynthesis, the ribosome, a large ribonucleoprotein complex, progresses along an mRNA template and brings together a succession of adapter tRNAs whose 3′-end hydroxyl groups are acylated with cognate amino acids. Three nucleotides of each tRNA (anticodon) match a given set of three nucleotides on the mRNA (codon), and one amino acid is delivered to the growing peptide chain. The nascent peptide emerges out of the so-called ribosome exit tunnel. Specific aminocyclization of tRNAs, catalysed by the corresponding aminocyl–tRNA synthetases, dictates codons’ assignment and constitutes the genetic code. Several methods have been developed to artificially expand or reprogram the genetic code by attaching non-proteinogenic amino acids to tRNAs. For example, versatile tRNA-acylation ribozymes, known as flexizymes, conveniently catalyse such attachments. Non-proteinogenic aminocyclization–tRNAs prepared by flexizymes can be added to solutions containing the translation-associating proteins, cofactors and ribosomes required for in vitro translation, allowing for the incorporation of a variety of synthetic building blocks and expanding the chemical register of mRNA-encoded peptides.

Genetic code reprogramming has also revealed the tolerance of the ribosomal machinery to non-proteinogenic amino acids: α-amino acids bearing non-natural side chains, many D-amino acids, and even some β-amino acids have been incorporated within peptidic sequences. Furthermore, chemistry that permits cyclization, as well as ester and N-alkylated amide functions, have been shown to be accepted by the ribosome. In addition to single amino acids, preformed di-, tri- or penta-peptides and other organic building blocks have also been attached directly to tRNA\textsubscript{\text{DCAU}}, the tRNA that initiates peptide synthesis by matching the start AUG codon of mRNA, and have been incorporated as small appendages at the N terminus of the translated peptide. These developments open up the prospect of hijacking the ribosome machinery to produce mRNA-encoded chemical entities even more remote from α-peptides, with the notable advantages that (1) mRNA encoding gives access to sequence randomization and selection from libraries containing trillions of unique sequences, and (2) enriching chemical diversity increases the range of accessible functions, including resistance to proteolytic degradation, cell penetration capabilities or protein recognition through the folding of short sequences. The stakes are thus high. However, beyond the key steps that have already been made, it is not yet known where definite limitations lie in this endeavour.

In the following, we present our discovery that the ribosome accepts objects far larger, and more distinct from peptides, than those previously considered. Specifically, we show that helical aromatic foldamers (helically folded oligomeric backbones with aryl rings in their main chain) larger than 900 Da can be incorporated at the N terminus of a translated peptide or protein. In the resulting foldamer–peptide hybrids, the peptide component may represent less than half of the whole entity. The process was found to require foldamer unfolding to pass the exit tunnel. Yet, the folding propensity of these foldamers largely exceeds that of short peptides, and their incorporation amounts to encoding additional folding information that can effectively bias the peptide structure. For example, we show that macrocycles can be translated by the ribosome, whereby a foldamer segment adopts a helical conformation and forces a peptide segment to remain extended. Reciprocally, the peptide segment controls the handedness of the foldamer helix. These results expand the scope of ribosomal expression of mRNA-encoded non-natural sequences, and also introduce the benefits of using the stable conformations of foldamers to control peptide conformation.

Results and discussion

Our initial motivation stemmed from the consideration that short peptides are inherently flexible and not prone to folding, whereas some aromatic foldamers feature outstanding conformational stability and might thus serve as structural templates to assist peptide folding. Several approaches already exist to stabilize peptide conformation, including, but not limited to, the use of non-natural building blocks such as β-amino acids or turn mimics, as well as
macrocyclization (for example through the stapling of side chains in α-helices\textsuperscript{23} or by joining the ends of a sequence\textsuperscript{24}) and conformational templates that initiate α-helicity\textsuperscript{23,24}. Non-natural building blocks and cyclization have also been combined\textsuperscript{25}. These conformationally restrained peptides generally perform better in the tight and selective recognition of biomolecules, as opposed to flexible sequences that must pay a higher entropic cost upon binding.

In contrast, some foldamers have folding properties that are superior to those of peptides, even peptides containing unnatural amino acids\textsuperscript{26,27}. Aromatic oligoamide foldamers are in this class. Oligomers derived from 8-amino-2-quinolinecarboxylic acid\textsuperscript{19} (henceforth referred to as Q, Fig. 1a) predictably form helical structures with remarkable stability, including in water\textsuperscript{28,29}. For example, an octamer shows no denaturation at 120 °C in dimethylsulfoxide (DMSO)\textsuperscript{19}. Q monomers are comparable to dipeptides in size (Fig. 1a), but, with their rigid aromatic structure, fewer rotatable bonds and a different pattern of hydrogen-bond donors and acceptors, their sequences form distinctly non-proteinogenic helices with a diameter about twice as large as that of an α-helix (Fig. 1b). Like peptides, Q oligomers can be decorated with different side chains. However, quite unlike peptides, folding is a property of the backbone and is independent of the side chain. As a hint that robust folding of Q might be conferred to peptides, they have been shown to be stable enough to accommodate individual α-amino acids, the conformation of which is then dictated by the aromatic backbone\textsuperscript{30,31}. We thus engaged in assessing the extent to which peptide–foldamer hybrid structures could be produced by ribosomal translation of mRNA and possibly exhibit synergistic folding properties.

**Acylation of tRNA with foldamer substrates.** Foldamer substrates were designed for attachment to tRNA, catalysed by the flexizyme eFx. To ensure good yields, all substrates in this study included a C-terminal glycine spacer followed by a phenylalanine cyanomethylester (Gly-Phe-CME, Fig. 1c), which is an appropriate

---

**Fig. 1 | Aromatic oligoamide foldamers and their initiation of in vitro translation.** a. Quinoline-based (Q) and pyridine-based (P) monomers with backbone (shown in bold) comparable to that of a dipeptide (Gly–Gly). Q can be functionalized with a variety of side chains (R). b. Intramolecular hydrogen bonds (dashed lines) within Q oligomers, together with aromatic stacking, allow for the formation of stable helices, as shown by the crystal structure of Q\textsubscript{4} (right, side chains are omitted for clarity)\textsuperscript{19}. c. Schematic representation of the acylation of tRNAs with N-acetylated foldamers, having a C-terminal glycyI-phenylalanyl cyanomethyl ester, by a flexizyme. Foldamer–tRNA adducts can be used to initiate in vitro translation. d. Foldamer–peptide hybrids (tRNA, yellow; ribosome 50S, light blue; ribosome 30S, dark blue; amino acids, white circles). e. Foldamer substrates 1–11 used in this study (CME, cyanomethyl ester). f. Structure of a foldamer–peptide (GFKKKFDYKDDDDK) hybrid. The C-terminal peptide segment DYKDDDDK is a flag-tag sequence. Unless stated otherwise, DNA/mRNA encoding for this peptide sequence was used.
Fig. 2 | In vitro translation of oligomeric foldamer-peptide/protein hybrids. a. Formulae of the expected targets (using substrates 1 and 4-10). ‘Peptide’ refers to GFKKFKDYKKDDDK (shown in Fig. 1f). Dashed arrows represent the iterative exploration of foldamer sequences in this study. Red crosses indicate that the subsequent sequence was not translated (that is, from 5 and 9). Thus, increasing the number of Q from 2 to 3 resulted in unsuccessful translation. Elongating (PQ) by one Q also prevented translation. However, increasing the P content (using substrate 10) allowed a pentameric foldamer–protein hybrid to be formed. b, MALDI–TOF–MS spectra of foldamer–peptide products successfully formed using in vitro translation. From top to bottom: peptide initiated with 1 (red); initiation using 4 (red) and 6 (blue); initiation using 7 (blue); initiation using 8 (blue); and initiation using 10 (purple). No product could be observed for substrates 5, 7 and c. 6.1 (blue); and initiation using 10 (purple). No product could be observed for substrates 5 and 9. c, MALDI–TOF–MS spectra (left) of a foldamer–protein hybrid formed using a DNA template for protein G (white, right). The major product corresponds to protein synthesis starting from the second (non-AUG) codon. C, calculated for [M+H]+; O, observed mass.

Peptide donor for tRNA acylation mediated by eFx. Three substrates were synthesized containing Q monomers with a range of side chains: aspartate-like (1), ornithine-like (2) and tetraethylene glycol (3) (Fig. 1a). Aminoacylation of these substrates was tested by charging onto an RNA microhelix, using an appropriate DMSO-containing reaction buffer (Supplementary Fig. 1). Note that this RNA microhelix was used as a tRNA mimic for reasons of detection simplicity, because its acylation can be readily quantified using band intensity on acid polyacrylamide gel electrophoresis (PAGE) and because it has been established that the yield of microhelix charging correlates with activity using tRNA. The longer foldamers 4–11, comprising several Q units with aspartate-like side chains and P units (6-aminomethyl-2-pyridinecarboxylic acid, Fig. 1a), were also synthesized and could be charged onto the RNA microhelix (Supplementary Figs. 2–4), showing the high substrate tolerance of eFx. Because peptide synthesis in Escherichia coli is initiated with N-formyl methionine-charged tRNA<sub>Met</sub><sub>CAU</sub>, we carried out the acylation of in vitro transcript tRNA<sub>Met</sub><sub>CAU</sub> by foldamer substrates 1–11 and examined their use in translation.

Scope of in vitro translation initiation with appended Q<sub>m</sub> oligomers. Initial experiments showed that foldamer–tRNA<sub>Met</sub><sub>CAU</sub> from substrates 1–3, containing Q monomers with variable side chains, could all initiate the synthesis of dodecapeptide KKKFDYKDDDDK (Fig. 1f). The expected foldamer–peptide hybrids were detectable by matrix-assisted laser desorption/ionization–time of flight mass...
Inclusion of pyridine-based monomers in Qₜ sequences. We sought to reduce the conformational stability of the foldamer helix to test whether it would permit translation of larger initiation units. This can be achieved by replacing Q with the more flexible P monomer (Fig. 1a). P has the same backbone atoms as Q, and sequences combining P and Q fold into similarly shaped helices. However, P possesses a smaller surface for aromatic stacking and an additional rotatable bond at the methylene group that interrupts conjugation; thus, both factors should contribute to lowering the overwhelming stability of the Q helices. For instance, P, oligomers do not fold. In sequences 6–9, P and Q units alternate (Fig. 1c). Using tRNA^{Stem}_{CAU} acylated with these substrates, translation could be initiated with 6, 7 and 8 (albeit with lower efficiency than with 1–3) but not with the longest sequence 9 (Fig. 2a,b). As for 5, increasing the concentration of IF2 from 0.4 to 8 μM brought no improvement. To find a foldamer with five aryl units suitable for translation, we produced a variant of 9 with an N-terminal P instead of Q (substrate 10). The foldamer–tRNA^{Stem}_{CAU} charged with 10 was accepted by the ribosome and could be used in translation initiation.

Substrate 10 is long enough to span two full aromatic helix turns and is equivalent to a dodecameric peptide in length. It constitutes the largest appendage to tRNA^{fMet} and is equivalent to a dodecameric peptide in length. It constitutes the sequence of IF2 from 0.4 to 8 μM with lower efficiency than with 1. Furthermore, initiation appears to work regardless of the side chain, and even a large tetraethylene glycol was accepted. Foldamer–tRNA^{Stem}_{CAU} initiation appears to work regardless of the side chain, and even aromatic foldamers happen to be a favourable case. Noteworthy is the concomitant presence of translated products lacking the foldamer initiators (Supplementary Fig. 7). It is possible that, after assembly of the ribosome–mRNA complex, dissociation of the large foldamer–tRNA^{Stem}_{CAU} occurs, followed by the restart of translation at the second codon.

Solution studies of foldamer helices. To underpin our hypothesis on unfolding of foldamers in the ribosome exit tunnel, we investigated their conformation and stability in an aqueous environment. Foldamers 12–15 were synthesized, corresponding to initiators 7, 5, 8 and 9 lacking the C-terminal Phe-CME, respectively. Multidimensional NMR allowed for the complete assignment of their ¹H NMR spectra (Supplementary Figs. 10–13). Rotating-frame Overhauser spectroscopy (ROESY) experiments showed correlations indicating that all are helically folded (Supplementary Figs. 14–17). Although the patterns and intensities of the nuclear Overhauser effect (NOE) signals were not entirely identical, correlations between amide protons of adjacent residues were identified in all cases (Fig. 3a), consistent with the canonical quinoline helix in which amide protons reside at the inner rim (Fig. 3b). Other medium-range NOE signals between non-adjacent residues (i to i+2 or i to i+3) represent typical through-space correlations in quinoline helices where side chain, aromatic and Gly protons locate on the outer rim (Fig. 3b). The ¹H NMR spectra of 12–15 all featured at least one strongly upfield-shifted singlet near 6.4 ppm (Fig. 3c), corresponding to the H3 proton of the first aromatic ring starting from the C terminus (always a Q) that is shielded by the ring current effect of a stacked quinoline (the second quinoline ring from the C terminus) in the helical conformation.

Having established folding, we next investigated helix stability. NH/ND exchange rates allow for the assessment of exposure to the surrounding water and in part reflect conformational dynamics (Supplementary Figs. 18–21). The pseudo-first-order rate constants of amide NH/ND exchange in 12–15 could be calculated and were shown to be significantly lower for 13 and 15. Within a given molecule, NH/ND exchange rates vary according to chemical nature and position in the sequence (Fig. 3d). When examining comparable functions, for example the terminal NH(Gly) of different oligomers, exchange was found to be faster for 12 and 14 than for 13 (Fig. 3c). The range of rates is considerable: the amide protons of 12, except its NH(Gly), are quantitatively exchanged within 100 s, whereas the NH of the central Q of pentamer 15 exchanges with a half-life of 66 min (Supplementary Figs. 18 and 21). Overall, stability increases both with increasing length and on increasing the proportion of Q. Conformational stability may thus well be the factor that prevents translation initiation with 5 and 9.

Conformational stability was also reflected in the ¹H NMR spectra. Helical conformations are inherently chiral, even if the molecules possess no stereogenic centres. As a result, CH₃ protons in Q (side chain) and P and Gly (main chain) are diastereotopic, and their signals may be anisochronous when exchange between left- and right-handed helices is slow on the NMR timescale, that is, when the helices have sufficient kinetic stability. Consistent with the higher stability of 13 and 15 revealed by NH/ND exchange rates, anisochronous CH₃ signals are obvious in the spectra of these
sequences, where multiple AB systems are present (Fig. 3c) that coalesced at elevated temperature (45°C, Supplementary Figs. 23 and 25). In contrast, the diastereotopic protons appear as sharp averaged singlets in the spectrum of 12 as left- and right-handed helices exchange rapidly (Supplementary Fig. 22). The spectrum of 14 shows an intermediate situation with some broadening (Supplementary Fig. 24), suggesting that its dynamics, albeit faster than for 13 and 15, are slower than for α-peptidic chains for which fast exchange on the NMR timescale is observed. In summary, these results suggest that translation initiation occurs only when appended foldamer helices have a sufficient ability to transiently unfold to pass through the ribosome exit tunnel, and that this is allowed for helices that are nevertheless more stable than a typical, isolated peptide helix.

Ribosomal synthesis of macrocyclic foldamer–peptide hybrids. To further expand the chemical diversity produced by the ribosome and to exploit ribosomal uptake of helical foldamers with better defined conformations than short peptides, we endeavoured to express foldamer–peptide cyclic hybrids. In a macrocycle, the peptide and foldamer are attached at two points, which might result in a coupling of their respective conformations. Genetic code reprogramming with an N-terminal chloroacetyl (ClAc) amino acid allows for efficient macrocyclization through thioether formation

---

**Fig. 3 | NMR studies of aromatic foldamer segments.** a. Chemical formulae of aromatic foldamer segments used for solution-state conformation studies with identified NOEs in aqueous solution (9:1 H2O/D2O, 50 mM NaHCO3). Red arrows indicate correlations between backbone amide protons. Blue arrows indicate correlations between sequentially non-adjacent residues. Solid and dashed arrows indicate strong and medium/weak interactions, respectively. b. Molecular model of 12 with observed NOEs. c, d. Pseudo-first-order rate plots of NH/ND exchange and half-life for NH(Gly) in D2O at 25 °C of 12, 13 and 14 (c) and amide protons of 13 (d). [A] and [A]0 indicate the integrated peak area of amide protons at a given time and at t = 0, respectively. Monitoring of NH(Gly) of 15 was not possible due to signal overlap. e. Parts of 1H NMR spectra of 12, 13, 14 and 15 in D2O with 50 mM NaHCO3 at 25 °C. Peaks in blue indicate H3 protons of quinoline rings. Peaks in orange indicate CH2 protons of Q (side chain), P (main chain) and Gly (main chain) residues.
with the thiol of a cysteine residue further along the sequence\(^1\). Precursor 11 (Fig. 1e) was prepared to include a terminal ClAc to form foldamer-containing macromolecules (Fig. 4a), and this substrate could initiate translation (Supplementary Fig. 4). Additional mRNA templates were designed to encode a cysteine either immediately after the foldamer initiator (template C2) or further along the peptide (templates C3, C4 and C5), and these were used for in vitro translation, initiated using 11. Within 30 min of translation, cyclized species were detected as the major product for templates C2 and C3 proceeded at a slower, similar rate (\(k_{\text{cycl}} = 3 \times 10^{-4} \text{s}^{-1}\)), Supplementary Figs. 8 and 9).

The effect of peptide stapling by a foldamer on macrocycle conformation was then investigated in solution. Macrocycles 16 and 17 were synthesized as analogues of cyclized translation products from C4 and C5 templates (Fig. 4b). Their circular dichroism (CD) spectra showed positive bands at 280 and 340 nm, indicating bias toward a right-handed foldamer helix\(^4\) as a consequence of the \(l\)-chirality of the peptide. Heating had little effect on the CD spectra (Fig. 4c and Supplementary Figs. 26 and 27), hinting at helix stability even under harsh conditions such as 7:3 H\(_2\)O/DMSO (vol/vol) at 65 °C. 1H NMR spectra of 16 and 17 showed poorly resolved signals at
room temperature (Supplementary Figs. 28 and 29), presumably due to nonspecific aggregation promoted by their amphipathic and zwitterionic nature. However, sharp spectra at 65 °C (Fig. 4d) allowed for the observation of quinoline H3 proton singlets near 6.4 ppm, indicating shielding by aromatic stacking. In addition, NH-CH$_3$, J'-couplings of P units showed obvious diastereotopic patterns (Fig. 4e and Supplementary Figs. 30 and 31). Thus, the PQPQ segment in both 16 and 17 is helical, chirally biased, and stabilized with respect to non-cyclic sequence 14. Helix conformation control by stereocentres can be remote$^{41}$. Here, the nearest aromatic centres are five atoms away from the foldamer helix at both C and N termini, which contrasts greatly with earlier examples of absolute handedness control$^{44}$.

Crystals of 17 could be obtained but diffracted only at 2.5 Å. The n-enantiomer of 17 was then also synthesized and mixed with l-17 to produce a raceme. Crystals of racemic 17 diffracted at 0.9 Å, and the structure in the solid state could be resolved in centrosymmetrical space group $I$ (Supplementary Figs. 32 and 33). Two independent molecules in the asymmetric unit show essentially superimposable main-chain conformation, suggesting that this conformation is robust (Supplementary Fig. 34). The structure proved to be consistent with solution-phase observations (Fig. 4f,g and Supplementary Movie 1). The aromatic segment PQPQ of l-17 is folded into a compact right-handed helix. Conversely, d-17 is left-handed. The helical fold is similar to that of canonical Q, helices. In contrast, the six-amino-acid peptide segment adopts a loop-like extended conformation. Two hydrogen bonds are established between two peptide NH protons belonging to Lys2 and Cys5, and the two CO of the carbonyl-amino-methyl groups in position 6 of the two P units. These inter-segment interactions presumably mediate helix handedness control while restraining the peptide conformation. The peptide side chains are not involved, suggesting that the conformation may be sequence independent. The reciprocal conformational effects between these chemically distinct biotic and abiotic folded entities—a peptide and an aromatic oligoamide—are quite unique. The structure mirrors stapled helical peptides$^{39}$ in which the peptide is $\alpha$-helical and the hydrocarbon staple is extended. Similar structures have also been observed in thioether-bridged macrocyclic peptides$^{45,46}$. Such macrocycles were synthesized by the ribosome and were selected from large libraries to bind to a particular partner protein. When bound, these macrocycles use a helix to present a peptide loop for binding, with the helix involved to different degrees in the binding interface (Supplementary Fig. 35). This raises an attractive possibility of using the structures described here, helical foldamers presenting a peptide loop, for molecular recognition.

Conclusions

We have demonstrated that aromatic foldamer-appended peptides could be synthesized by the ribosome using genetic code reprogramming. Foldamer–tRNA prepared by flexizyme can be delivered as a single unit to the ribosome and successfully initiate peptide synthesis. The foldamers represent not only the largest single building blocks to be accepted by the ribosome, but also chemical entities very distinct from the peptide sequences that the ribosome has been optimized to accept by evolution. We find that translation requires a certain degree of conformational flexibility, possibly to allow for unfolding of the foldamer segments to pass through the ribosome exit tunnel. Yet, foldamers with folding propensities stronger than those of peptides of comparable sizes—a non-cyclic decapetide rarely folds—are accepted for translation. It can be conceived that fine-tuning of folding–unfolding behaviour might allow translation of even larger appendages. In addition, macrorcyclization chemistry was successfully applied to these foldamer–peptide hybrids. Helically folded aromatic foldamers in cyclic peptides act as a novel class of peptide linker that can serve as a structural scaffold to present the peptide in three-dimensional space. Solution and solid-state studies reveal a reciprocal influence of the peptide and foldamer helix conformations, presumably by mutual reduction of the entropy penalty inherent to folding. For this type of hybrid molecule, one may expect functional improvements for the peptide regions in terms of selectivity and binding affinity to target biomolecules. Conversely, the foldamer may be highly water-soluble and possess cell-penetrating properties$^{47}$. An obvious subsequent step is to investigate the amenability of these constructs to sequence randomization and selection of functional species from libraries.

Data availability. The crystallographic data and experimental details of the structural refinement for the X-ray crystal structure of 17 have been deposited at the Cambridge Crystallographic Data Centre (CCDC) under deposition no. CCDC 1554263. The data can be obtained free of charge from CCDC (http://www.ccdc.cam.ac.uk/data_request/cif). Other data that support the findings of this study (spectroscopic or mass spectrometric data) are available from the corresponding author upon reasonable request.

Received: 1 July 2017; Accepted: 24 November 2017; Published online: 19 March 2018

References