FEBS openbio

FEBS PRESS science publishing by scientists

Co-translational folding of α-helical proteins: structural studies of intermediate-length variants of the λ repressor

Yuya Hanazono†, Kazuki Takeda and Kunio Miki

Department of Chemistry, Graduate School of Science, Kyoto University, Japan

Keywords
cotranslational folding; crystal structure; λ repressor

Correspondence
K. Miki, Department of Chemistry, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan
Fax: +81 75 753 4032
Tel: +81 75 753 4029
E-mail: miki@kuchem.kyoto-u.ac.jp

Present address
†Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai 980-8579, Japan

(Received 26 February 2018, revised 17 May 2018, accepted 14 June 2018)

doi:10.1002/2211-5463.12480

Nascent polypeptide chains fold cotranslationally, but the atomic-level details of this process remain unknown. Here, we report crystallographic, *de novo* modeling, and spectroscopic studies of intermediate-length variants of the λ repressor N-terminal domain. Although the ranges of helical regions of the half-length variant were almost identical to those of the full-length protein, the relative orientations of these helices in the intermediate-length variants differed. Our results suggest that cotranslational folding of the λ repressor initially forms a helical structure with a transient conformation, as in the case of a molten globule state. This conformation subsequently matures during the course of protein synthesis.

Database
Structural data are available in the PDB under the accession numbers 5ZCA and 3WOA.

Proteins are synthesized on the ribosomes and fold into thermodynamically stable structures. The newly synthesized polypeptides can fold cotranslationally because the folding speed and folding pathways are limited by the rate of translation [1–4]. The codon translation rate is 20–30 amino acids per second (in *Escherichia coli*) or 2–4 amino acids per second (in a eukaryote) [5]. Consequently, various intermediate states of nascent proteins can exist for a long period because the timescale of the cotranslational folding (on the order of seconds to minutes) is much longer than that of the full-length protein folding (on the order of microseconds) [6]. Therefore, the characterizations of the intermediate states of nascent proteins are important to understand the process of the cotranslational folding. Moreover, nascent proteins, which control their own translation and quality, are involved in the regulation of the life process. Therefore, the elucidation of the mechanism of the nascent protein folding is of great importance for understanding the biological systems [7,8]. However, it is difficult to elucidate the dynamics of the cotranslational folding. The folding of nascent proteins is assisted by molecular chaperones [9] and the ribosomal surface [10–12]. In addition, the nascent proteins are affected by many weak protein–protein interactions because the *in vivo* conditions are highly crowded [13]. A small number of reports have analyzed cotranslational folding by means of nuclear magnetic resonance [14–16], fluorescence resonance energy transfer [16–18], and computational methods [19–21].

Recently, we reported the structures of a series of WW domain N-terminal fragments with increasing numbers of amino acids to reveal the atomic-level

Abbreviations
CD, circular dichroism; MBP, maltose-binding protein; OPEP, optimized potential for efficient structure prediction; TFE, 2,2,2-trifluoroethanol.

FEBS Open Bio **8** (2018) 1312–1321

details of cotranslational folding [22]. Unexpectedly, the intermediate-length fragments formed helical structures even though the full-length protein has no helical regions. This suggests a structural change from a structure in which short-range interactions are decisive to one in which long-range interactions of a particular peptide length are decisive. Therefore, the nascent proteins eventually reach the native structures by adopting stable transient conformations.

Next, to reveal the atomic-level details of the short-range interactions of alpha-helical proteins in nascent protein folding, we focused on the N-terminal domain of the λ repressor. This domain has a five-helix bundle, and the folding mechanisms of its wild-type and numerous variants have already been investigated using various methods [23–31]. These studies revealed that the N-terminal domain of the λ repressor can fold in diverse ways, including by two-state folding, downhill folding, and helical-intermediate folding, depending on changes in the sequence, temperature, and solvent. The full-length folding of the λ repressor N-terminal domain is driven by the formation of a hydrophobic core with helices. However, the folding pathway of the λ repressor cannot form such a hydrophobic core in the early stage of the peptide extension. Here, we report the results of our structural studies of two intermediate-length fragments of the λ repressor N-terminal domain (residues 1–20: λ_{1-20}; 1–45: λ_{1-45}). Intermediate-length fragments of the λ repressor adopt a helical structure in the same way as the full-length λ repressor (λ_{1-92}). However, the relative orientation of the two helices in λ_{1-45} is not identical to that of the full-length λ repressor.

Materials and methods

Preparation of proteins

The genes for expression of the intermediate-length λ repressor N-terminal domain (λ_{1-20} or λ_{1-45}) fused with MBP at its C terminus were inserted into a pET22b vector using NdeI/HindIII sites. The linker sequences, which were Gly-Ser-Gly for λ_{1-20} and Gly-Ser-Gly-Met for λ_{1-45}, were inserted between the λ repressor fragment and MBP. The fragments of the MBP and λ repressor were amplified from the pKM596 vector (Addgene plasmid 8837) [32] and artificial gene synthesis (Hokkaido System Science, Sapporo, Japan), respectively. These constructs were transformed into Rosetta2(DE3)pLysS and grown at 37 °C in LB medium containing 100 $\mu g \cdot mL^{-1}$ ampicillin and 34 $\mu g \cdot mL^{-1}$ chloramphenicol. The protein expression was induced when the OD_{600} reached 0.6 by the addition of 1 mM IPTG at 37 °C for 3 h. After cells were harvested, the pellet was resuspended in 50 mM Tris/HCl pH 7.5 and 150 mM NaCl (Buffer A) and disrupted by sonication. The suspension of disrupted cells was centrifuged at 40 000 ***g*** for 30 min at 4 °C. The supernatant was applied to an MBPTrap column (GE Healthcare, Little Chalfont, UK) equilibrated with Buffer A. The bounded protein was eluted with Buffer A containing 10 mM maltose. Then, the pooled sample was applied to a HiLoad 16/60 Superdex 200 column (GE Healthcare) equilibrated with Buffer A.

The peptide of λ_{1-20} was synthesized by the Fmoc solid-phase method and purified to > 95% by GL Biochem Ltd (Shanghai, China). The peptides of λ_{1-45} and λ_{1-92} were subcloned into the pET22 vector using NdeI/HindIII sites. These were fused with MBP at the N-terminal and linked with a Gly-Ser-Gly-Ile-Glu-Gly-Arg linker, which contained a factor Xa recognition sequence. These constructs were expressed and purified as described above. After the gel filtration, these samples were cleaved with factor Xa (Novagen, Madison, WI, USA) in a solution containing 50 mM Tris/HCl pH 8.0, 100 mM NaCl, and 5 mM $CaCl_2$ for 16 h at 20 °C. The cleaved fragments were separated by a Superdex 75 10/300 column (GE Healthcare) equilibrated with Buffer A.

Crystallographic analysis

Two intermediate-length λ repressor N-terminal domains (λ_{1-20}, λ_{1-45}) fused with MBP (λ_{1-20}-MBP, λ_{1-45}-MBP) were concentrated to 20 $mg \cdot mL^{-1}$ in 10 mM Tris/HCl pH 7.5, 150 mM NaCl, and 10 mM maltose. The crystals of λ_{1-20}-MBP and λ_{1-45}-MBP were grown in a solution made up of a 1 : 1 mixture of the protein solution and reservoir solution. The reservoir conditions differed for the individual variants as follows: 1.6 M triammonium citrate was used for λ_{1-20}-MBP, and 1.6 M DL-malic acid (pH 7.0) was used for λ_{1-45}-MBP. The X-ray diffraction intensities were collected at BL41XU of SPring-8 (Harima, Japan). Diffraction data sets were processed and scaled using the HKL2000 software package [33]. The structure was solved by the molecular replacement method with the program MOLREP [34] in the CCP4 software suite [35]. MBP (PDBID: 1ANF) [36] was used as a search model. The program phenix.autobuild [37] was employed for autotracing. The output structure was manually improved with the program COOT [38]. The structure was refined using the program phenix.refine [37]. The refined structure was validated with the program MOLPROBITY [39]. The crystallographic and refinement statistics are listed in Table 1. The superimpositions were performed with the program LSQKAB [40]. All figures for the molecular models were prepared using the program PYMOL [41]. Coordinates and structure factors of λ_{1-20}-MBP and λ_{1-45}-MBP have been deposited in the Protein Data Bank under the accession numbers 5ZCA and 3WOA, respectively.

Table 1. Crystallographic and refinement statistics.

	λ_{1-20}-MBP	λ_{1-45}-MBP
Crystal data		
Space group	$P2_12_12_1$	$P2_12_12$
Cell parameters		
a (Å)	49.0	150.9
b (Å)	58.3	52.6
c (Å)	124.6	58.3
Resolution range (Å)	50–1.80 (1.83–1.80)	50–2.00 (2.03–2.00)
Reflections (total/unique)	214 992/33 855	218 770/32 189
Redundancy	6.4 (6.2)	6.8 (6.2)
Completeness (%)	100 (100)	99.9 (99.6)
$I/\sigma(I)$	17.5 (1.6)	15.1 (11.3)
R_{sym}[a] (%)	8.3 (81.7)	12.6 (17.9)
Refinement		
No. of atoms	3467	3727
Protein atoms	3071	3205
Ligand/Ion	36	23
Water molecules	360	499
Average B-factor (Å^2)	29.6	18.2
Protein atoms	28.7	17.0
λ repressor/ linker/MBP	83.3/61.7/26.5	27.1/53.7/15.5
Maltose	22.5	10.7
Water molecules	37.3	26.3
R_{work}[b]/R_{free}[c] (%)	17.3/21.5	15.8/19.4
RMSD bonds (Å)/angle (°)	0.006/1.0	0.009/1.1
Ramachandran plot (%)[d]	97.7/2.3/0	98.8/1.2/0

Values in parentheses refer to the highest resolution shell.
[a] $R_{sym} = \Sigma_{hkl}\Sigma_i | I_{hkl,i} - \langle I_{hkl} \rangle | / \Sigma_{hkl}\Sigma_i I_{hkl,i}$.
[b] $R_{work} = \Sigma_{hkl} || F_{obs}| - |F_{calc} || / \Sigma_{hkl} |F_{obs}|$.
[c] R_{free} was calculated using the 5% of the reflections that were not included in the refinement as a test set.
[d] Favored/allowed/outliers.

De novo modeling

Conformational modeling of the intermediate-length λ repressor N-terminal domain was performed by the program PEP-FOLD (web server http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/) [42]. Energy evaluation by PEP-FOLD depends on the optimized potential for efficient structure prediction (OPEP) coarse-grained force field.

Circular dichroism spectroscopy

Two intermediate-length proteins (λ_{1-20} and λ_{1-45}) and a full-length (λ_{1-92}) protein were separately dissolved in 5 mM potassium phosphate buffer (pH 7.5) in the presence of 0–50% 2,2,2-trifluoroethanol (TFE). All samples were measured using a J-805 CD spectropolarimeter (Jasco, Tokyo, Japan) in a range from 190 to 250 nm with a 1-mm quartz cuvette. The secondary structure content was analyzed with the program JWSSE-408 (Jasco) using a reference data set [43]. The thermal transition curves of the half-length proteins and a full-length protein were measured with an ellipticity at 222 nm in a range from 5 °C to 85 °C.

Results

Crystal structure of the intermediate-length λ repressor N-terminal domain

We determined the two crystal structures of the intermediate-length λ repressor N-terminal domain (λ_{1-20}-MBP and λ_{1-45}-MBP) at 1.8 Å and 2.0 Å, respectively (Fig. 1 and Table 1). To fix the C terminus so that it corresponded with that in the protein synthesis on the ribosome, MBP was fused just behind the N-terminal fragments. As in the case of the full-length protein, the intermediate-length variants were made of alpha-helices. There was little interaction between the regions of the λ repressor and MBP in the crystal (Fig. 1A,D). In addition, MBP was not an obstacle for the helix–helix interaction in the crystal of λ_{1-45}-MBP. The helical region of λ_{1-20}, which contains a portion of helix 1 of the full-length λ repressor, was almost the same as that of the full-length protein. λ_{1-20} and the full-length proteins (PDBID: 1LMB) could be superimposed on each other with a root-mean-square deviation (RMSD) of 0.77 Å. λ_{1-45} also formed a helical conformation, whose regions are almost identical to those of the full-length protein. However, helix 2 of λ_{1-45} was slightly shorter than helix 2 of the full-length protein (Fig. 2A). In addition, the region of residues 40–45, which are a part of helix 3, did not form any helical structures. The λ_{1-45} and full-length proteins could be superimposed on each other with an RMSD of 3.59 Å for the region of residues 10–39 between helices 1 and 2 (Fig. 2B). On the other hand, the RMSD for the region of residues 40–45 was 9.81 Å. Moreover, the interaction between helices 1 and 2 of λ_{1-45} was different from that of the full-length protein. The helix–helix interaction in λ_{1-45} consisted mainly of a hydrophobic interaction between Y22, L29, L31, and V36 (Fig. 2C). In the case of the full-length protein, the relative disposition of helices 1 and 2 was dictated by the hydrophobic core, which was constructed of Y22, L29, L31, V36, F51, L65, and L69 (Fig. 2D). Although the interaction between helices 1 and 2 of the half-length variant was similar to that of the full-length protein, the interactions between the side chains of the half-length and full-length proteins, which contained Y22, L29, L31, and V36, were different. It is thus possible that the relative orientations between helices change during synthesis of the nascent polypeptide on the ribosome.

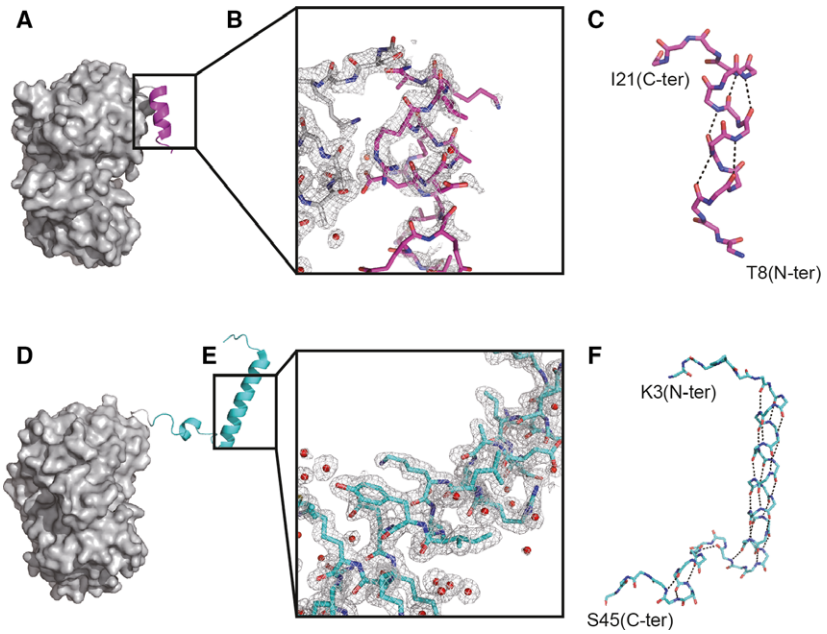

Fig. 1. Crystal structure of the intermediate-length λ repressor N-terminal domain fused with MBP. (A) Overall structure of the λ_{1-20}-MBP. The regions of the λ repressor and linker are shown as a ribbon model, and the MBP is shown as a surface model. The λ_{1-20} and linker region are colored magenta and gray, respectively. (B) A 2*Fo*–*Fc* map for the λ_{1-20}-MBP is represented as a gray mesh contoured at the 0.8σ level. Water molecules are shown as red spheres. (C) Backbone diagram of the λ_{1-20} shown as a stick model. Hydrogen bonds are shown as black dashed lines. (D) Overall structure of the λ_{1-45}-MBP. The λ_{1-45} and linker region are colored cyan and gray, respectively. (E) A 2*Fo*-*Fc* map of the λ repressor regions for the λ_{1-45}-MBP contoured at the 1.0σ level. (F) Backbone diagram of the λ_{1-45} shown as a stick model.

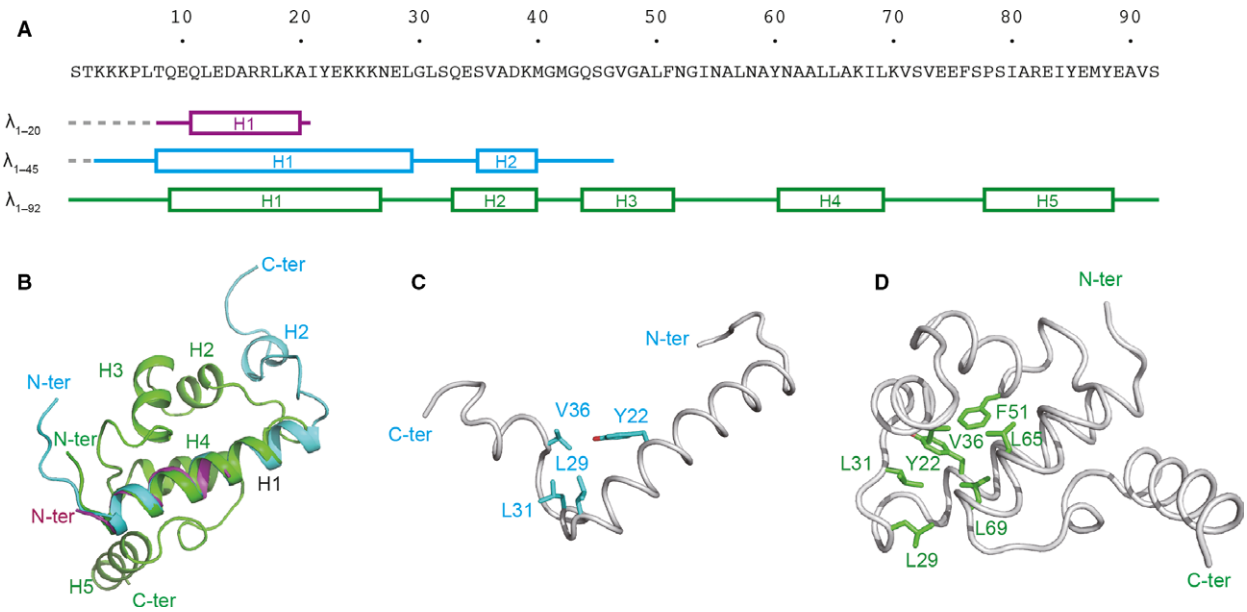

Fig. 2. Structural difference between the intermediate-length and full-length λ repressor. (A) Regions of the secondary structure of the λ_{1-20}, λ_{1-45}, and λ_{1-92}. The cylinders indicate alpha-helices. (B) Superimposition of the λ_{1-20} (magenta), λ_{1-45} (cyan), and λ_{1-92} (green). These proteins are superimposed in the region of helix 1 (residues 9–26). (C) Hydrophobic interaction between helices 1 and 2 of the λ_{1-45}. (D) Hydrophobic core of the λ_{1-92}.

De novo modeling of the intermediate-length λ repressor N-terminal domain

To reveal the structural change of the λ repressor fragments as their amino acid length increased, we carried out conformational modeling with the OPEP coarse-grained force field. First, we performed the conformational modeling of $\lambda_{1–20}$ and $\lambda_{1–45}$ (Figs 3A and S1). The structures of $\lambda_{1–20}$ and $\lambda_{1–45}$ harbored a helical conformation. In addition, the regions of helical conformation were approximately the same as those in the crystal structures (Fig. 3B). The conformational modeling yielded results similar to those of the X-ray crystallography. Therefore, a *de novo* approach by this method can provide a reliable structure. Subsequently, we performed the conformational modeling of $\lambda_{1–15}$, $\lambda_{1–25}$, $\lambda_{1–30}$, $\lambda_{1–35}$, $\lambda_{1–40}$, and $\lambda_{1–50}$. All of the predicted structures of the intermediate-length λ repressor harbored helical conformations in the same range as $\lambda_{1–20}$ or $\lambda_{1–45}$. This result suggested that the intermediate-length fragments of the λ repressor were able to form a stable helical conformation. The relative orientation of the conformational model between the helices 1 and 2 was different from the orientation of the crystal structures (Fig. 3C). This orientation of the predicted structure of $\lambda_{1–45}$ with the best score was mainly dictated by the hydrophobic interaction between I21, Y22, L29, V31, V36, M40, and M42 (Fig. 3D). The interaction among Y22, V36, and M40 was well observed in the predicted models (Fig. S2), whereas Y22 interacted with V36 and without M40 in the crystal structure (Fig. 2C). In the case of $\lambda_{1–92}$, Y22 interacted with V36 and F51 (Fig. 2D). This result suggests that the orientation of $\lambda_{1–45}$ between the helices 1 and 2 was produced by the weak and transient interaction via the native conformation.

CD spectroscopy analysis

The circular dichroism (CD) spectra for $\lambda_{1–20}$, $\lambda_{1–45}$, and $\lambda_{1–92}$ were independent of the protein concentration between 0.01 and 0.3 mg·mL^{-1} (Fig. 4A). This indicates that the concentration of the proteins had little effect on the secondary structures in solution. The negative ellipticities of $\lambda_{1–20}$ and $\lambda_{1–45}$ at 222 nm were smaller than that of $\lambda_{1–92}$ (Fig. 4B). We next analyzed the secondary structure of $\lambda_{1–20}$, $\lambda_{1–45}$, and $\lambda_{1–92}$ from the CD spectrum. $\lambda_{1–20}$, $\lambda_{1–45}$, and $\lambda_{1–92}$ contained 16%, 43%, and 73% helical structures in aqueous solution (Fig. 4C). The TFE titration results

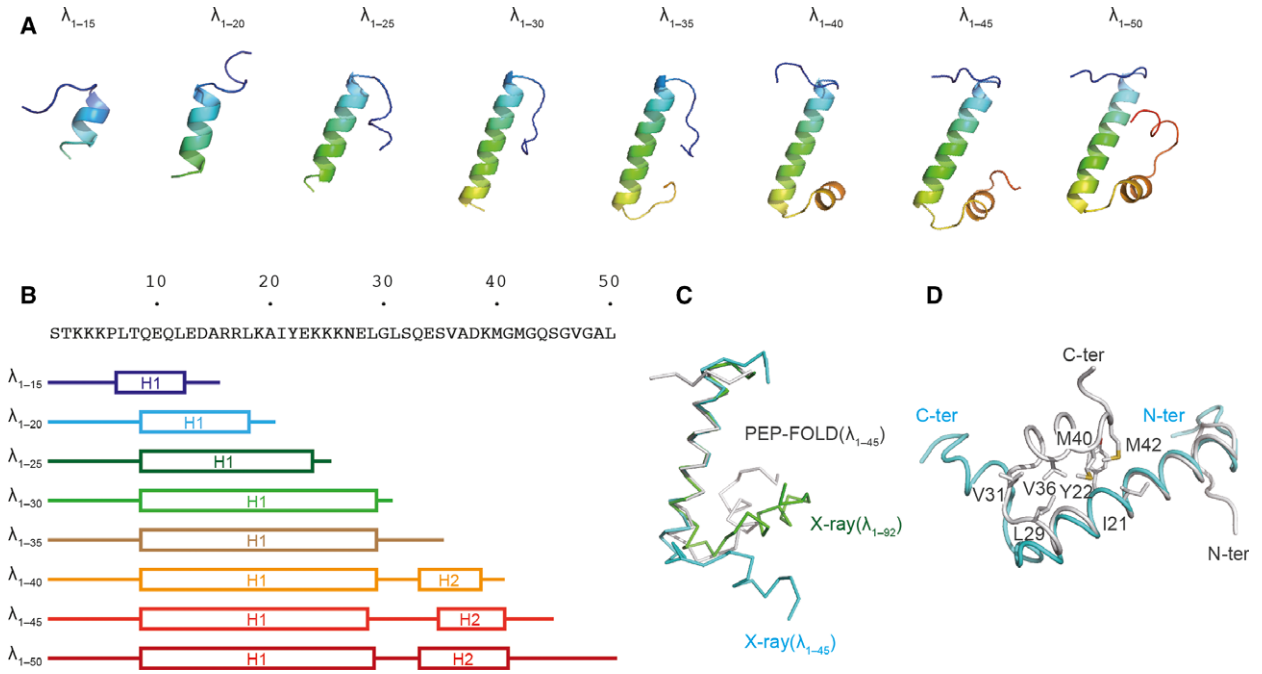

Fig. 3. Conformational modeling by the PEP-FOLD program. (A) The most representative of the best models for each of the intermediate-length λ repressors. (B) Secondary structure of the best models. (C) Comparison between the conformational model (gray) and the crystal structures ($\lambda_{1–45}$: cyan; and $\lambda_{1–92}$: green). (D) Hydrophobic interaction between helices 1 and 2 of the conformational model of $\lambda_{1–45}$ with the best score. The conformational model and the crystal structures are shown in gray and cyan, respectively.

indicate that all of the λ repressor fragments showed a similar propensity to form a helical conformation, regardless of the concentration of TFE. In the presence of a 30% TFE concentration whose dielectric constant is close to the *in vivo* condition [44,45], λ_{1-20}, λ_{1-45}, and λ_{1-92} contained 36%, 64%, and 73% helical structures, respectively (Fig. 4C). These values were in good accord with the crystallographic results, which gave helical contents of 20%, 52%, and 67%, respectively. Thermal stability was assessed by following changes in the spectrum at 222 nm with increasing temperature. The decreases in negative ellipticity of the intermediate-length variants were approximately linear, whereas that of λ_{1-92} was sigmoidal (Fig. 4D). The melting points of the intermediate-length variants were obscure, whereas the melting point of λ_{1-92} was clearly defined at 55 °C. It has been reported that the thermal denaturation transitions of the short peptides that have no hydrophobic core tend to be more linear than those of the globular proteins [46].

Discussion

Our crystallographic results showed that the λ repressor of N-terminal fragments could fold into helical structures with lengths of 20 and 45 residues. The helical regions of the fragments were in accord with the full-length λ repressor. A *de novo* approach produced similar results. CD spectrum analysis indicated that the major part of the intermediate-length λ repressor retained a substantial portion of helices. A previous report established that the nascent proteins exist under highly crowded conditions when the polypeptides are synthesized [47]. Further, it is known that dielectric constant is significantly reduced under such crowded conditions [48]. The results of our TFE titration using the CD spectrum showed that the intermediate-length fragments folded into a helical structure even in the absence of TFE. These results support the notion that the crystal structures of the intermediate-length fragments were well representative of diverse conditions. The various relative orientations between helices 1 and

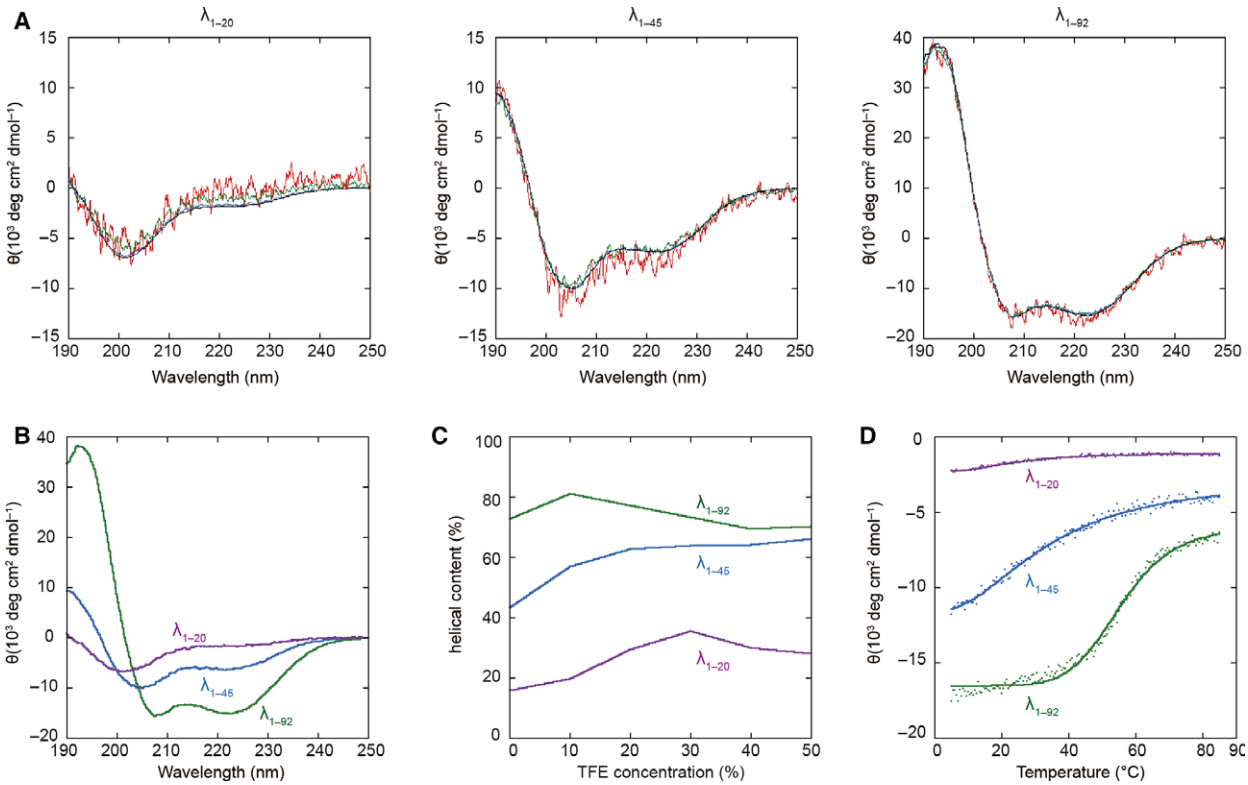

Fig. 4. Circular dichroism spectroscopic studies comparing the intermediate-length and full-length λ repressor. (A) CD spectrum of the λ repressor at different concentrations. (A) CD spectrum of the λ_{1-20}, λ_{1-45}, and λ_{1-92} of 0.01 mg·mL^{-1} (red), 0.03 mg·mL^{-1} (green), 0.1 mg·mL^{-1} (blue), and 0.3 mg·mL^{-1} (black). (B) Comparison of the CD spectra of the λ_{1-20} (magenta), λ_{1-45} (cyan), and λ_{1-92} (green) in 0.1 mg·mL^{-1}. (C) Change of the helical content as a function of TFE concentration. The λ_{1-20}, λ_{1-45}, and λ_{1-92} are colored magenta, cyan, and green, respectively. (D) Thermal denaturation of the λ_{1-20} (magenta), λ_{1-45} (cyan), and λ_{1-92} (green) obtained from the ellipticity at 222 nm. These data are fitted by sigmoidal curves.

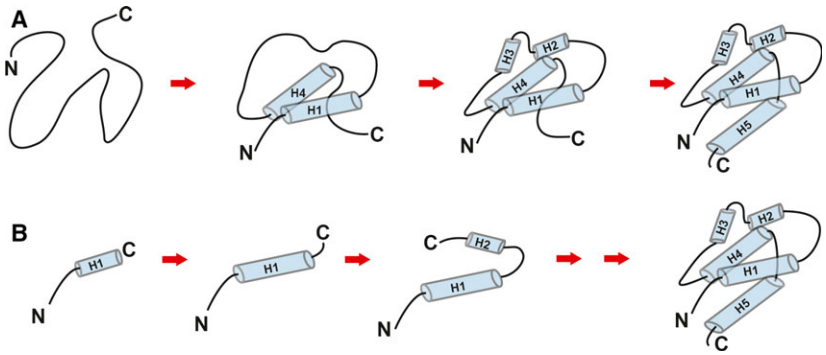

Fig. 5. Schematic model of the folding pathway of the λ repressor. (A) The folding pathway of the full-length protein. Cylinders indicate alpha-helices. (B) The cotranslational folding pathway.

2 were observed in the crystal and predicted structures. This indicates that the relative orientation is dictated by accidental weak hydrophobic interaction. The conformation of λ_{1-45} in crystal was shown to be a suitable model of the transient structures in peptide extension.

Based on the crystallographic, computational, and spectroscopic studies, the helical conformation is formed at the early stage of protein synthesis on the ribosomes. A theoretical analysis reported that the newly synthesized polypeptides are prone to forming a helical conformation in the ribosomal tunnel [49]. In fact, alpha-helical or helical-like structures have been observed within the tunnel by cryo-electron microscopy experiments [50,51]. On the other hand, the relative orientation between the helices is different from the native structure. In general, globular proteins have tightly packed hydrophobic cores and the hydrophobic effect is the major driving force behind their folding [52,53]. The λ repressor N-terminal domain also has hydrophobic cores, and the cores are important for protein stability [54]. However, the hydrophobic core matures in a later phase of the cotranslational folding. Therefore, the most stable conformation for the length is dominant in an early stage of the cotranslational folding. The transient conformation in full-length folding is known as a molten globule state, which has a native-like secondary structure but without a tightly packed conformation [55]. Based on the results of our CD spectroscopic analysis, the fragments of the intermediate-length λ repressor form a helical conformation but not a stable tertiary structure.

The folding path of the full-length λ repressor N-terminal domain suggests that helices 1 and 4 mainly come into contact in the earlier stages based on molecular dynamics [27,56,57] and mutational analysis [58]. Later, helices 2 and 3 are organized to form a helical conformation, and helix 5 finally folds (Fig. 5A). However, this folding path is only possible in the case of the full-length protein. The nascent protein could be predisposed to fold hierarchically because its folding depends on the rate of ribosome biosynthesis in the living cells, and thus, the nascent protein takes much more time to finish its folding than the full-length protein folding. In cotranslational folding, the region of helix 1 may adopt a helical conformation and helix 2 may subsequently adopt an orientation with the most stable conformation for each length (Fig. 5B).

In this study, we revealed that the intermediate-length λ repressor N-terminal domain forms an alpha-helical structure, the secondary structures of which are almost the same as the full-length structures. However, the interaction between helices 1 and 2 of the intermediate-length variant is different from the same interhelix interaction in the full-length protein, because the half-length variant has no hydrophobic cores. Formation of the λ repressor initially takes place via the local interaction between helices 1 and 2. A complete picture of this phenomenon cannot be derived merely from a folding investigation of the full-length protein. Therefore, the present results will contribute to elucidation of the process of cotranslational folding.

Acknowledgements

We would like to express our gratitude to the beam line staff at BL41XU of SPring-8 (Proposal Number: 2015A1077).

Author contributions

YH performed the biochemical and crystallographic experiments. YH, KT, and KM discussed the results and wrote the manuscript.

References

1 Hingorani KS and Gierasch LM (2014) Comparing protein folding in vitro and in vivo: foldability meets the fitness challenge. *Curr Opin Struct Biol* **24**, 81–90.

2 Kubelka J, Hofrichter J and Eaton WA (2004) The protein folding 'speed limit'. *Curr Opin Struct Biol* **14**, 76–88.

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

4 Thommen M, Holtkamp W and Rodnina MV (2017) Co-translational protein folding: progress and methods. *Curr Opin Struct Biol* **42**, 83–89.

5 Netzer WJ and Hartl FU (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* **388**, 343–349.

6 Jacobson GN and Clark PL (2016) Quality over quantity: optimizing co-translational protein folding with non-'optimal' synonymous codons. *Curr Opin Struct Biol* **38**, 102–110.

7 Ito K and Chiba S (2013) Arrest peptides: cis-acting modulators of translation. *Annu Rev Biochem* **82**, 171–202.

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

9 Mashaghi A, Kramer G, Bechtluft P, Zachmann-Brand B, Driessen AJ, Bukau B and Tans SJ (2013) Reshaping of the conformational search of a protein by the chaperone trigger factor. *Nature* **500**, 98–101.

10 Evans MS, Sander IM and Clark PL (2008) Cotranslational folding promotes beta-helix formation and avoids aggregation in vivo. *J Mol Biol* **383**, 683–692.

11 Kaiser CM, Goldman DH, Chodera JD, Tinoco I Jr and Bustamante C (2011) The ribosome modulates nascent protein folding. *Science* **334**, 1723–1727.

12 Knight AM, Culviner PH, Kurt-Yilmaz N, Zou T, Ozkan SB and Cavagnero S (2013) Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics. *ACS Chem Biol* **8**, 1195–1204.

13 Elcock AH (2010) Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Curr Opin Struct Biol* **20**, 196–206.

14 Cabrita LD, Cassaignau AME, Launay HMM, Waudby CA, Wlodarski T, Camilloni C, Karyadi ME, Robertson AL, Wang X, Wentink AS *et al.* (2016) A structural ensemble of a ribosome-nascent chain complex during cotranslational protein folding. *Nat Struct Mol Biol* **23**, 278–285.

15 Cabrita LD, Hsu ST, Launay H, Dobson CM and Christodoulou J (2009) Probing ribosome-nascent chain complexes produced in vivo by NMR spectroscopy. *Proc Natl Acad Sci USA* **106**, 22239–22244.

16 Buhr F, Jha S, Thommen M, Mittelstaet J, Kutz F, Schwalbe H, Rodnina MV and Komar AA (2016) Synonymous codons direct cotranslational folding toward different protein conformations. *Mol Cell* **61**, 341–351.

17 Holtkamp W, Kokic G, Jager M, Mittelstaet J, Komar AA and Rodnina MV (2015) Cotranslational protein folding on the ribosome monitored in real time. *Science* **350**, 1104–1107.

18 Kim SJ, Yoon JS, Shishido H, Yang Z, Rooney LA, Barral JM and Skach WR (2015) Protein folding. Translational tuning optimizes nascent protein folding in cells. *Science* **348**, 444–448.

19 O'Brien EP, Christodoulou J, Vendruscolo M and Dobson CM (2011) New scenarios of protein folding can occur on the ribosome. *J Am Chem Soc* **133**, 513–526.

20 O'Brien EP, Vendruscolo M and Dobson CM (2012) Prediction of variable translation rate effects on cotranslational protein folding. *Nat Commun* **3**, 868.

21 Trovato F and O'Brien EP (2017) Fast protein translation can promote co- and posttranslational folding of misfolding-prone proteins. *Biophys J* **112**, 1807–1819.

22 Hanazono Y, Takeda K and Miki K (2016) Structural studies of the N-terminal fragments of the WW domain: Insights into co-translational folding of a beta-sheet protein. *Sci Rep* **6**, 34654.

23 Burton RE, Huang GS, Daugherty MA, Calderone TL and Oas TG (1997) The energy landscape of a fast-folding protein mapped by Ala–>Gly substitutions. *Nat Struct Biol* **4**, 305–310.

24 Huang GS and Oas TG (1995) Structure and stability of monomeric lambda repressor: NMR evidence for two-state folding. *Biochemistry* **34**, 3884–3892.

25 Huang GS and Oas TG (1995) Submillisecond folding of monomeric lambda repressor. *Proc Natl Acad Sci USA* **92**, 6878–6882.

26 Kim SJ, Matsumura Y, Dumont C, Kihara H and Gruebele M (2009) Slowing down downhill folding: a three-probe study. *Biophys J* **97**, 295–302.

27 Larios E, Pitera JW, Swope WC and Gruebele M (2006) Correlation of early orientational ordering of engineered λ6–85 structure with kinetics and thermodynamics. *Chem Phys* **323**, 45–53.

28 Liu F, Du D, Fuller AA, Davoren JE, Wipf P, Kelly JW and Gruebele M (2008) An experimental survey of the transition between two-state and downhill protein folding scenarios. *Proc Natl Acad Sci USA* **105**, 2369–2374.

29 Liu F, Gao YG and Gruebele M (2010) A survey of lambda repressor fragments from two-state to downhill folding. *J Mol Biol* **397**, 789–798.

30 Liu F and Gruebele M (2007) Tuning lambda6-85 towards downhill folding at its melting temperature. *J Mol Biol* **370**, 574–584.
31 Liu Y, Strumpfer J, Freddolino PL, Gruebele M and Schulten K (2012) Structural characterization of lambda-repressor folding from all-atom molecular dynamics simulations. *J Phys Chem Lett* **3**, 1117–1123.
32 Fox JD, Routzahn KM, Bucher MH and Waugh DS (2003) Maltodextrin-binding proteins from diverse bacteria and archaea are potent solubility enhancers. *FEBS Lett* **537**, 53–57.
33 Otwinowski Z and Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326.
34 Vagin A and Teplyakov A (1997) MOLREP: an Automated program for molecular replacement. *J Appl Crystallogr* **30**, 1022–1025.
35 Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760–763.
36 Quiocho FA, Spurlino JC and Rodseth LE (1997) Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure* **5**, 997–1015.
37 Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221.
38 Emsley P, Lohkamp B, Scott WG and Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.
39 Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS and Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12–21.
40 Kabsch W (1976) A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr Section A* **32**, 922–923.
41 Schrödinger LLC (2010) *The PyMOL Molecular Graphics System, Version 1.4*. Schrödinger LLC, New York, NY.
42 Shen Y, Maupetit J, Derreumaux P and Tuffery P (2014) Improved PEP-FOLD Approach for Peptide and Miniprotein Structure Prediction. *J Chem Theory Comput* **10**, 4745–4758.
43 Reed J and Reed TA (1997) A set of constructed type spectra for the practical estimation of peptide secondary structure from circular dichroism. *Anal Biochem* **254**, 36–40.
44 Kleizen B and Braakman I (2004) Protein folding and quality control in the endoplasmic reticulum. *Curr Opin Cell Biol* **16**, 343–349.
45 Harada R, Sugita Y and Feig M (2012) Protein crowding affects hydration structure and dynamics. *J Am Chem Soc* **134**, 4842–4849.
46 Tanizaki S, Clifford J, Connelly BD and Feig M (2008) Conformational sampling of peptides in cellular environments. *Biophys J* **94**, 747–759.
47 Fort AG and Spray DC (2009) Trifluoroethanol reveals helical propensity at analogous positions in cytoplasmic domains of three connexins. *Biopolymers* **92**, 173–182.
48 Chin DH, Woody RW, Rohl CA and Baldwin RL (2002) Circular dichroism spectra of short, fixed-nucleus alanine helices. *Proc Natl Acad Sci USA* **99**, 15416–15421.
49 Ziv G, Haran G and Thirumalai D (2005) Ribosome exit tunnel can entropically stabilize alpha-helices. *Proc Natl Acad Sci USA* **102**, 18956–18961.
50 Bhushan S, Gartmann M, Halic M, Armache JP, Jarasch A, Mielke T, Berninghausen O, Wilson DN and Beckmann R (2010) alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nat Struct Mol Biol* **17**, 313–317.
51 Matheisl S, Berninghausen O, Becker T and Beckmann R (2015) Structure of a human translation termination complex. *Nucleic Acids Res* **43**, 8615–8626.
52 Finney JL, Bowron DT, Daniel RM, Timmins PA and Roberts MA (2003) Molecular and mesoscale structures in hydrophobically driven aqueous solutions. *Biophys Chem* **105**, 391–409.
53 Klein-Seetharaman J, Oikawa M, Grimshaw SB, Wirmer J, Duchardt E, Ueda T, Imoto T, Smith LJ, Dobson CM and Schwalbe H (2002) Long-range interactions within a nonnative protein. *Science* **295**, 1719–1722.
54 Burton RE, Myers JK and Oas TG (1998) Protein folding dynamics: quantitative comparison between theory and experiment. *Biochemistry* **37**, 5337–5343.
55 Baldwin RL and Rose GD (2013) Molten globules, entropy-driven conformational change and protein folding. *Curr Opin Struct Biol* **23**, 4–10.
56 Portman JJ, Takada S and Wolynes PG (2001) Microscopic theory of protein folding rates. I. Fine structure of the free energy profile and folding routes from a variational approach. *J Chem Phys* **114**, 5069–5081.
57 Portman JJ, Takada S and Wolynes PG (2001) Microscopic theory of protein folding rates. II. Local reaction coordinates and chain dynamics. *J Chem Phys* **114**, 5082–5096.
58 Prigozhin MB, Sarkar K, Law D, Swope WC, Gruebele M and Pitera J (2011) Reducing lambda repressor to the core. *J Phys Chem B* **115**, 2090–2096.

Supporting information

Additional Supporting Information may be found online in the supporting information section at the end of the article:

Fig. S1. The most representative of the five best models for each of the intermediate-length λ repressors.
Fig. S2. Hydrophobic interaction between helices 1 and 2 of the conformational models of λ_{1-45}.