Biochimica et Biophysica Acta 1864 (2016) 1281–1291

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Structure and dynamics of the retro-form of the bacteriophage T5 endolysin

Victor P. Kutyshenko [a,*], Galina V. Mikoulinskaia [b], Nikolai V. Molochkov [a], Dmitry A. Prokhorov [a], Sergei A. Taran [b], Vladimir N. Uversky [c,d,**]

[a] Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia
[b] Branch of Shemyakin & Ovchinnikov's Institute of Bioorganic Chemistry, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia
[c] Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL, USA
[d] Laboratory of Structural Dynamics, Stability and Folding of Proteins, Institute of Cytology, Russian Academy of Sciences, Tikhoretsky Ave. 4, St. Petersburg 194064, Russia

ARTICLE INFO

Article history:
Received 11 April 2016
Received in revised form 29 June 2016
Accepted 30 June 2016
Available online 1 July 2016

Keywords:
Endolysin
Retro-protein
Protein folding
Protein dynamics
High-resolution NMR
Flexibility
Intrinsic disorder
Intrinsically disordered protein

ABSTRACT

Using high-resolution NMR spectroscopy we conducted a comparative analysis of the structural and dynamic properties of the bacteriophage T5 endolysin (EndoT5) and its retro-form; i.e., a protein with the reversed direction of the polypeptide chain (R-EndoT5). We show that structurally, retro-form can be described as the molten globule-like polypeptide that is easily able to form large oligomers and aggregates. To avoid complications associated with this high aggregation propensity of the retro protein, we compared EndoT5 and R-EndoT5 in the presence of strong denaturants. This analysis revealed that these two proteins possess different internal dynamics in solutions containing 8 M urea, with the retro-form being characterized by larger dimensions and slower internal dynamics. We also show that in the absence of denaturant, both forms of the bacteriophage T5 endolysin are able to interact with micelles formed by the zwitterionic detergent dodecylphosphocholine (DPC), and that the formation of the protein-micelle complexes leads to the significant structural rearrangement of polypeptide chain and to the formation of stable hydrophobic core in the R-Endo T5.

1. Introduction

The role of the directionality of a polypeptide chain in protein folding and function remains an open question. In early studies, it has been hypothesized that the generation of the retro-forms of protein, in which the peptide bond direction is reversed without reversal of the amino acid chirality, should expand the conformational space accessible to the normally-read sequence [1], potentially producing the mirroring protein structure [1–4]. It was also hypothesized that some 'macromolecular chirality' might be present, since such retro-proteins were expected to be structured, possessing structures mirroring those of the normally-read sequences and were expected to behave as the inverso-forms of the natural proteins; i.e., proteins built from the all-D amino acids [1]. However, several computational and experimental studies conducted on retro-proteins produced rather contradictory data and did not provide clear evidence of the general validity of these hypotheses.

In fact, as it is briefly outlined below, the existing literature data support the notion that retro-proteins and retro-peptides are able to gain secondary structure. However, the existence of well-defined native-like 3D structure in retro-proteins under any conditions was not reported so far. Although many proteins in the Protein Data Bank (PDB) have inverse sequence similarity (ISS) to each other [5] a large-scale analysis of such ISS proteins, where the corresponding C_α atoms of forwardly and inversely aligned proteins were structurally aligned, revealed that the retro-proteins did not have folds similar to the folds of the original sequences [3,5], suggesting that the inversion of sequences dramatically affects the property of the protein to fold [5]. In agreement with these statistical analysis, the full-atom simulations and experimental characterization of the retro-forms of the domain B of staphylococcal protein A and the SH3 domain and B1 domain of Streptococcal protein G revealed that these retro-proteins were essentially unfolded [4]. Similarly, the retro-form of rubredoxin was mostly unstructured [6]. Also, using circular dichroism and high–resolution

Abbreviations: DPC, dodecylphosphocholine; DSS, sodium 4,4-dimethyl-4-silapentane-sulfate; EndoT5, bacteriophage T5 endolysin; IDP, intrinsically disordered protein; R-EndoT5, retro form of bacteriophage T5 endolysin; SD, spin diffusion; SSD, spectra of the spin diffusion.

* Corresponding author.
** Correspondence to: V.N. Uversky, Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL, USA.
E-mail addresses: kutyshenko@rambler.ru (V.P. Kutyshenko), vuversky@health.usf.edu (V.N. Uversky).

http://dx.doi.org/10.1016/j.bbapap.2016.06.018
1570-9639/

NMR it was shown that the retro-form of the chimeric protein SHH from the "SH3-Bergerac" family (where the β-turn N47D48 of the spectrin SH3-domain is substituted for the *KITVNGKTYE* sequence) does not contain rigid 3D structure, but is characterized by the presence of residual secondary structure [7]. On the other hand, several retro-proteins were able to fold into conformations totally different from the conformations attainable by their original sequences [8,9]. Other retro-proteins possessed well-developed secondary structures similar to those of their natural counterparts [10,11]. Furthermore, some retro-proteins were involved in the formation of extensive intermolecular contacts [10–12] and were even able to form amyloid-like fibrils [13]. Finally, several experimental examples showed that although retro-sequences may not only have the stable 3D structure [14], they still possess functions similar to the original sequences [12,15–17].

Emerging recent publications on the retro-proteins and retro-peptides have more applied nature showing what the synthesizable retro-polypeptides can be used for [18–20]. However, the question of why the retro-proteins typically fail to fold or at least to undergo the hydrophobic collapse remains open. The answer to this question is a matter of principle, since it is believed that, for foldable proteins, hydrophobic collapse serves as the primary step of the protein self-organization into unique globular structure [21]. Although retro-proteins contain sufficient number of hydrophobic residues, they are commonly not folded and exist in extended conformations, suggesting that the overall hydrophobicity of a polypeptide chain is not sufficient for hydrophobic collapse and some nucleation sites, for which the order of amino acid residues is important, must be present [22]. In this work, we conducted a comparative analysis of the structural and dynamic properties of the bacteriophage T5 L-alanyl-D-glutamate peptidase, EndoT5 (PDB (2MXZ) and BMRB (25437)), and its retro-form, R-EndoT5.

2. Materials and methods

2.1. Cloning of the T5-retrolysin gene and transformation of the Escherichia coli BL21 (DE3)

The T5-retrolysin gene (416 bp) was synthesized based on the codon usage preference for the expression in *E. coli* (taken from the Codon Usage Database at http://www.kazusa.or.jp/codon/) and cloned by blunt ends (*EcoRV* restriction enzyme site) in the pUC57plasmid (Fermentas, Lithuania). The correctness of the gene design and assembly was confirmed by sequencing of both DNA strands. DNA sequencing was performed using the ABI PRISM® BigDye™ Terminator v. 3.1 kit followed by the analysis of reaction products on Applied Biosystems 3730 DNA Analyzer.

The synthetic gene without the stop codon was subcloned into the expression vector pET28a (+) (Novagen) by *NcoI* and *XhoI* restriction sites. The resulting expression construct pT5R contained T5-retrolysin gene with the hexahistidine-tag encoding sequence at the 3′-end. The nucleotide sequence of the T5-retrolysin gene and the corresponding amino sequence of the resulting R-EndoT5 are shown in Fig. S1. Plasmid pT5R was used to transform the electrocompetent cells of the *E. coli* strain BL21 (DE3). Positive clones were selected on LB-agar containing 50 μg/ml kanamycin.

2.2. Isolation and purification of the R-EndoT5

Cells of the *E. coli* strain BL21 (DE3) were transformed with the pT5R plasmid. Transformants were grown in the 200 ml liquid LB-medium containing 50 μg/ml kanamycin. Synthesis of the R-EndoT5 was induced with 1 mM IPTG when the culture density reached A_{550} of 1.0, and 3 h after the beginning of induction, cells were collected by centrifugation at 6000 g for 10 min. Harvested *E. coli* BL21 (DE3) cells (1.3 g) bearing plasmid pT5R were suspended in 10 ml of buffer A containing 25 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM EDTA and 1 mM PMSF. The cell suspension was sonicated at 75 W for 1 min (2 sonication cycles for 30 s). The suspension was centrifuged at 10,000 g for 30 min. The pellet containing inclusion bodies with the target protein, was resuspended in buffer A and centrifuged under the same conditions. The pellet was then dissolved in 10 ml of buffer B (pH 8.0), containing 10 mM Tris, 100 mM NaH_2PO_4 and 8 M urea and applied to a 4 ml Ni Sepharose column equilibrated with buffer B. The column was washed first with buffer B and then with buffer C (pH 6.3), containing 10 mM Tris, 100 mM NaH_2PO_4 and 8 M urea. Protein was eluted with Buffer D (pH 5.9), containing 10 mM Tris, 100 mM NaH_2PO_4 and 8 M urea. The combined eluate was dialyzed against acetate buffer. In order to control the homogeneity of the purified R-EndoT5, protein was analyzed by the SDS PAGE according to Laemmli [23] (see Fig. S2) using standard marker proteins, such as β-galactosidase (116 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), lactate dehydrogenase (35 kDa), REase Bsp981 (25 kDa), β-lactoglobulin (18.4 kDa), and hen egg white lysozyme (14.4 kDa). Preparation of the native EndoT5 and its ^{13}C- and ^{15}N-enriched analogue is described in [24]. Analogous procedure was used for preparation of the ^{13}C- and ^{15}N-enriched analogue of the R-EndoT5.

2.3. Native PAGE electrophoresis

Electrophoresis in a 15% polyacrylamide gel was carried out in a continuous buffer system (0.2 M Tris-acetate buffer, pH 4.5) at room temperature for one day (current 6 A). Since R-EndoT5 is prone to aggregate under native conditions, electrophoresis was conducted in two parallel gels: one contained 2.5 M urea for dissolving protein aggregates, the other did not. Gels were fixed in 10% trichloroacetic acid, stained with Coomassie Brilliant Blue G-250, and washed with distilled water. In the absence of urea, R-EndoT5 did not enter into the separating gel, whereas in the presence of urea moved according to its isoelectric point.

2.4. Preparation of protein samples for structural analysis

NMR samples (90% H_2O + 10% D_2O) were prepared using heavy water from company CIL (99.9%), dodecylphosphocholine-d38 (98%, CIL), deuterated acetate (99%, CIL), and 0.03% sodium azide. Protein concentrations were about 1 mM in 10 mM Tris-HCl buffer (pH 7) or 10 mM acetate buffer (pH 4). Changes in pH were measured using the Digital pH-meter HANNA (Germany) equipped with a combined microelectrode. To remove imidazole, dialysis against 10 mM Tris-HCl buffer (pH 7.8) with 0.5 mM EDTA was used. Aliquots of the dodecylphosphocholine-d38 detergent were added to some samples. The molar protein/detergent ratio in all samples was 1/60.

Samples for CD spectra analysis were prepared by dilution of the NMR samples to have the final protein concentration in the range of 0.5 mM. Solutions of 8 M urea were prepared by dissolving the required amounts of urea in 10 mM acetate buffer (pH 4), then protein was added. Concentration was monitored by NMR or measured using the UV-2401 PC spectrophotometer, Shimadzu (Japan). The spectrophotometric measurements were performed at 25 °C in a quartz cuvette with the path length of 1 mm to 10 mm. The molar extinction coefficients for the proteins were calculated from amino acid sequence using the program available at the ExPASy website: http://www.iut-arles.up.univ-mrs.fr/w3bb/d_abim/compo-p.html.

2.5. NMR spectroscopy

NMR measurements were conducted on the AVANCE 600 III spectrometer, with an operating frequency of 600 MHz, at 298 K, a spectral width of 24 ppm and the 90-degree pulse of 11 μs. To achieve a good signal/noise ratio typically 128 repetitions were sufficient, which were accumulated in the 32 K memory cells of the control computer. In some cases (NOE), greater number of accumulation (1000–20,000)