Biochimica et Biophysica Acta 1843 (2014) 866–874

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Selection against toxic aggregation-prone protein sequences in bacteria

Susanna Navarro, Anna Villar-Piqué, Salvador Ventura *

Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain
Departament de Bioquímica i Biologia Molecular, Facultat de Biociències, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain

ARTICLE INFO

Article history:
Received 14 December 2013
Received in revised form 16 January 2014
Accepted 21 January 2014
Available online 26 January 2014

Keywords:
Protein aggregation
Protein folding
Protein evolution
Amyloid peptide
Bacteria

ABSTRACT

Despite genetic variation has the potential to arise new protein functions, spontaneous mutations usually destabilize the native fold. Misfolded proteins tend to form cytotoxic intracellular aggregates, decreasing cell fitness and leading to degenerative disorders in humans. Therefore, it is thought that selection against protein misfolding and aggregation constrains the evolution of protein sequences. However, obtaining experimental data to validate this hypothesis has been traditionally difficult. Here we exploit bacteria as a model organism to address this question. Using variants of the Alzheimer's related Aβ42 peptide designed to exhibit different in vivo aggregation propensities we show here that, in cell competition experiments, the most aggregation-prone variants are always purged out from the growing population. Flow cytometry analysis of cellular metabolism and viability demonstrates that this purifying effect responds to a clear correlation between physiological burden and intrinsic aggregation propensity. Interestingly, the fitness cost of aggregation appears to be associated with aggregation rates rather than with overall protein solubility. Accordingly, we show that, by reducing in vivo aggregation rates, the model osmolyte proline is able to buffer the metabolic impact of protein aggregation. Overall, our data provide experimental support for the role of toxic protein aggregation on the cell fitness landscape and the evolution of natural protein sequences.

1. Introduction

Most genetic mutations decrease the probability of protein-coding sequences to pack into compact and stable conformations [1]. Protein misfolding usually results in the formation of aggregates, which harms cells and reduces fitness [2]. In this way, protein aggregation is linked to the onset of an increasing number of human degenerative disorders, including Alzheimer's or Parkinson's diseases [3,4]. Protein aggregation is a generic structural property of polypeptide chains that directly competes with native protein folding [5,6]. Therefore, it might constrain the evolution of proteins, in such a way that protein sequences would be shaped by natural selection to encode polypeptides that would not aggregate at the cellular concentrations and locations in which they should function [7,8], the so-called "Life on the Edge" hypothesis [9]. This purifying effect of protein misfolding might explain why in all organisms, from bacteria to humans, the specific rate of synonymous and non-synonymous substitutions is slower in highly expressed proteins [10,11], since their mutation will generate relatively more misfolded species.

Despite the deleterious effect of protein aggregation on cell physiology is well documented [12,13], the magnitude of this effect has not been sufficiently characterized experimentally. Quantification of the influence of genetic mutations on aggregation and fitness is essential to validate the purifying role of aggregation during the evolution of protein sequences and might also provide molecular mechanistic insights on the basis of human conformational disorders [14]. However, quantifying how sequential traits and the associated aggregation propensity impact cellular metabolism and the cell fitness landscape is a challenging task because, *in vivo*, the intrinsic properties of proteins are modulated by an array of different factors [15–17]. Bacterial systems provide simple models to tackle this question under near-physiological conditions [18]. In bacteria, protein aggregation has been shown to impair cell division, causing loss of cell reproductive ability or aging [19,20]. Importantly, the aggregates formed by amyloidogenic proteins in bacteria share many conformational properties with those causing disease [21–23]. In particular, we have shown previously that a fusion of the Alzheimer's related Aβ42 peptide with GFP forms amyloid-like insoluble deposits when expressed in bacteria and impairs cell division [20,24]. Here we used this model system to address the relationship between protein aggregation propensity and protein evolution by analyzing the precise impact of three designed Aβ42 variants differing in their *in vivo* aggregation properties on cell metabolism and viability using flow cytometry and measuring how these toxic effects correlate with the fitness cost they impose to the cell population.

2. Materials and methods

2.1. Bacterial strains and cell growth

Strains *Escherichia coli* BL21 DE3 transformed with pET-28a vector (Novagen, INC., Madison, WI) encoding for Aβ42wt-GFP (Green

Abbreviations: IPTG, isopropyl β-D-1-thiogalactopyranoside; CTC, 5-cyano-2,3-ditolyl tetrazolium chloride; TMAO, trimethylamine n-oxide; GFP, green fluorescent protein; wt, wild type; PDB, protein data bank; IP, propidium iodide

* Corresponding author at: Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, E-08193 Bellaterra Spain. Tel.: +34 93 586 8956.
E-mail address: salvador.ventura@uab.es (S. Ventura).

0167-4889/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbamcr.2014.01.020

fluorescence protein), F19D and F19D/L34P variants were grown aerobically in liquid Luria–Bertani (LB) medium containing appropriate antibiotics in a rotary shaker at 37 ºC and 250 rpm. Overnight cultures were diluted 100-fold in LB and allowed to grow to an optical density at 600 nm (OD_{600}) of 0.5. At the indicated OD, protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 0.5 M of osmolytes proline and trimethylamine n-oxide (TMAO) was added, when necessary. Cultures were let to grow for 24 h.

2.2. Prediction of aggregation propensities

Theoretical aggregation-prone regions were predicted with TANGO using the default settings [25].

2.3. Fluorescence measurements

After indicated incubation times, bacteria cells were harvested by centrifugation and washed with phosphate-buffered saline (PBS buffer) at pH 7.0. GFP fluorescence in intact cells expressing the Aβ42wt-GFP, F19D and F19D/L34P variants was measured on a Jasco FP-8200 fluorescence spectrophotometer (Jasco Corporation, Japan) in the 500–600 nm range using an excitation wavelength of 470 nm.

2.4. Optical fluorescence microscopy

10 μL of cultured cells was centrifuged and washed in PBS buffer, and deposited on top of glass slides. Images were obtained under UV light using a filter for GFP excitation (450–500 nm) and an emission filter (515–560 nm) in a Leica fluorescence microscope (Leica Microsystems, Germany).

2.5. Electrophoresis and immunoblots

E. coli cells expressing Aβ42wt-GFP, F19D and F19D/L34P variants were lysed chemically. The cell lysate was centrifuged for 20 min at 12,000 *g*. The supernatant was collected separately and the pellet was diluted in the same volume as supernantant. Samples of 10 μL of lysate were separated on a 10% SDS-polyacrylamide gel electrophoresis containing 0.1% SDS. Proteins were transferred to a polyvinylidene difluoride membrane (Immobilon, Millipore) with a Mini Trans-Blot Cell (Bio-Rad). Membranes were blocked in 0.1 M Tris–HCl, pH 7.5, 0.5% Tween 20, 1% Triton X-100, 5% bovine serum albumin o/n and probed for GFP expression using an anti-GFP antibody (Roche) diluted 1:1000 o/n in blocking solution. The secondary peroxidase-conjugated antibody was goat anti-mouse IgG (BioRad) diluted 1:3000 in blocking solution. Detection was carried out using an enhanced chemiluminescence (ECL) SuperSignal kit (Pierce).

2.6. Bacterial growth and preparation for flow cytometry analysis

Overnight cultures of *E. coli* cells expressing Aβ42wt-GFP, F19D and F19D/L34P mutants were diluted 100-fold in LB and allowed to grow to an optical density at 600 nm (OD_{600}) of 0.5. Then cultures were induced with 1 mM IPTG and samples were taken from exponential and stationary growth phases after 4, 8 and 24 h, respectively. Protein expression levels were comparable among Aβ42-GFP variants in these conditions. For analysis of osmolyte effects, proline and TMAO were added 30 min before protein induction. Each sample was vortexed and diluted to an OD of 0.2. Then diluted 10-fold in 0.22-μm filtered PBS. Subsequently, the cell suspension was centrifuged at 3500 *g* for 10 min, and the cells were resuspended in 1 mL of PBS. All experiments were performed using three biological replicates. Non-induced cultures of the three variants were incubated and analyzed in the same manner to serve as negative controls.

2.7. IP staining

To analyze cell cultures viability, samples were stained with propidium iodide (PI) dye (BD Biosciences) to a final concentration of 50 μM in PBS. To prepare killed bacteria as a positive control, 1 mL of sample was heated at 90 ºC for 10 min prior to PI staining. The analyzed data correspond to the bacteria population gated for SSC vs FITC-A so that cell debris and non-intact cells were excluded from the analysis.

2.8. CTC staining

To evaluate the respiratory activity of *E. coli* cultures expressing Aβ42wt-GFP, F19D and F19D/L34P mutants, cells were stained with 5-cyano-2,3-ditolyl tetrazolium chloride (CTC) (Invitrogen). 1 mL of diluted cells in PBS was treated with 100 μL of the 50 mM CTC working solution, vortexed gently to mix and incubated for 30 min at 37 ºC, protected from light.

2.9. Flow cytometry

Flow cytometry was performed using a BD FACSCanto flow cytometer (BD Biosciences) equipped with 488 nm and 635 nm lasers. Cells were first gated (P1) by forward scatter (FSC) and side scatter (SCC) both set on logarithmic amplification. A threshold was set on FSC to reduce electronic background noise. Cells in P1 were then re-gated (P2) for green fluorescence emission (GFP) to separate bacterial cells from cellular debris. GFP fluorescence emission was measured on a 530/30 nm band pass filter and CTC fluorescence emission on a 670LP filter. All signals were amplified. We acquired a total of 20,000 events logarithmically (five decades). Data were acquired with the FACSDiva Software (BD Biosciences). Data analysis was performed with the FlowJo software.

2.10. Competitive growth selection

Aβ42wt-GFP, F19D and F19D/L34P constructs were grown aerobically in LB medium containing 50 μg/mL kanamicin at 37 ºC and 250 rpm. The overnight cultures were diluted 100-fold in a 50 mL falcon tube containing LB medium, 50 μg/mL kanamicin and 1 mM IPTG, and allowed to grow at 37 ºC. At 8, 24 and 48 h harvest time, the DNA was purified using the Plamid Miniprep kit (Thermo Scientific) following the manufacturer instructions and sent for sequencing. Chromatograms of the samples were analyzed with the software Sequence Scanner (Applied Biosciences) to assign the dominant construct.

3. Results

3.1. Design of Aβ42 peptide variants displaying different aggregation propensity

We have shown previously that the bacterial IBs formed by the Alzheimer's disease-related Aβ42 peptide display an amyloid-like structure when this peptide is expressed alone [22] or as a fusion to different fluorescent proteins [24]. We have constructed a library of 20 different Aβ42-GFP mutants differing only in the residue in position 19 of the Aβ42 peptide, which is a Phe residue in the wild type sequence [26]. All these fusions accumulate essentially in the insoluble cellular fraction. However, the presence of active GFP in such aggregates differs significantly, displaying a high correlation with the aggregation propensity of the Aβ42 mutants, due to the existence of an in vivo kinetic competition between the folding of the GFP domain and the aggregation of the fusion protein, directed by the Aβ moiety [27]. The Asp mutant (F19D) exhibited the slowest aggregation rate [28]. Here we wanted to further increase this dynamic range and generate an Aβ42-GFP variant that not only would aggregate slower than the F19D mutant but would also shift its solubility at equilibrium from being mainly insoluble to mostly soluble, while keeping changes to a minimum. An analysis of the Aβ42