Chinese Chemical Letters xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Chinese Chemical Letters

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

Original article

Self-assembly and morphological characterization of two-component functional amyloid proteins

Qi Qi[a,b,c], Tian-Xin Zhao[a], Bo-Lin An[a], Xuan-Yong Liu[b], Chao Zhong[a,*]

[a] School of Physical Science and Technology, ShanghaiTech University, Shanghai 200120, China
[b] Shanghai Institute of Ceramics, Chinese Academy of Sciences, Shanghai 200050, China
[c] University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:
Received 8 October 2016
Received in revised form 10 November 2016
Accepted 29 November 2016
Available online xxx

Keywords:
Amyloid
Biofilms
Atomic force microscopy
Self-assembly
Nanomaterials
Supramolecular structure

ABSTRACT

Functional amyloid has been increasingly applied as self-assembling nanostructures to construct multi-functional biomaterials. However, little has been known how different side domains, varied fusion positions and subunits affect self-assembly and morphologies of amyloid fibrils. Here, we constructed three groups of two-component amyloid proteins based on CsgA, the major protein components of *Escherichia coli* biofilms, to bridge these gaps. We showed that all fusion proteins have amyloid features, as indicated by Congo red assay. Atomic force microscopy (AFM) indeed reveals that these fusion proteins are able to self-assemble into fibrils, with an average diameter of 0.5–2 nm and length of hundreds of nanometers to several micrometers. The diameter of fibrils increases with the increase of the molecular weight of fusion domains, while the dynamic assembly of recombinant proteins was delayed as a result of the introduction of fusion domains. Moreover, fusion of the same functional domains but at intermediate position seems to cause the most interference on fibril assembly compared with those fused at C or N-terminus, as mainly short and irregular fibrils were detected. This phenomenon appears more pronounced for randomly coiled mussel foot proteins (Mfps) than for rigid chitin-binding domain (CBD). Finally, increase of the molecular weight of tandem repeats in protein monomer seemed to increase the fibril diameter of the resultant fibrils, but either reduction of the tandem repeats of CsgA to one single belta-sheet loop or increase in the number of tandem repeats of CsgAs from one to four produced shorter and intermittent fibrils compared with CsgA control protein. These studies therefore provide insights into self-assembly of two-component amyloid proteins and lay the foundation for rational design of multi-functional molecular biomaterials.

1. Introduction

Amyloid has long been thought to be intimately related with neurodegenerative diseases like Alzheimer's disease and Parkinson's disease [1,2]. However, in recent decades, an increasing number of studies suggest that there exists a special type of amyloid structures in nature, usually referred to as "functional amyloids", that can play normal biological functions, including but not limited to adhesive biofilms of bacteria, catalytic scaffolds and hormones reservoirs [3–5]. Amyloid has thus been extensively utilized as nanomaterials to construct functional materials [6–8]. In particular, design of functional amyloid-based biomaterials with well-defined sequence-functionality relationship *via* genetic engineering has attracted great interest. Such modular genetic strategy provides a new avenue to rationally design multi-functional molecular biomaterials with several features including self-assembling fibril structures, outstanding stability and excellent mechanical properties. In addition, these amyloid structures can be further elaborated by introducing multiple functional domains into amyloid backbones to achieve multi-functionalities [9–11]. Despite great advances, no systematic studies have been done regarding how varied functional domains, positions and subunits affect fibril self-assembly and how they will affect the final morphology of assembled structures.

Herein, we used a functional amyloid system based upon CsgA, which is the subunit of culi—the major protein components of *Escherichia coli* biofilm [3,12], to study the impacts of varied fusion domains, fusion positions as well as different subunits on amyloid fibril assembly and morphological changes.

* Corresponding author.
E-mail address: zhongchao@shanghaitech.edu.cn (C. Zhong).

http://dx.doi.org/10.1016/j.cclet.2016.12.008

2. Results and discussion

Curli is the major extracellular proteinaceous fibers produced by *E. coli* and *Salmonella enterica* that is related to biofilm formation, immune activation, host colonization and cell invasion [13] (Fig. 1a). Curli fibril typically generates characteristic X-ray fiber diffraction pattern of amyloid proteins [14,15], in which the meridional and equatorial reflection is 4.6 Å and 8.7 Å, respectively (Fig. 1b). Purified CsgA can self-assemble into single amyloid fibrils at early stage with diameter of 1–2 nm (Fig. 1c) and eventually aggregate to form multi-strand fibrils or fiber bundles with diameter from several nanometers to tens of nanometers (Fig. 1d).

As the major subunit for curli fibers, CsgA is composed of five imperfect repeating loops [13]. Due to its remarkable properties, CsgA has been applied as the scaffold to construct a variety of multi-functional biomaterials [9–11]. In this study, we constructed several series of multi-component proteins based on CsgA. In the first group, SpyTag [16] (*Streptococcus pyogenes* peptide tag), CBD [17] (chitin binding domain) and SUMO [18] (small ubiquitin-related modifier) were respectively added at the C terminus of CsgA, intending to study the influence of side domains on fibril assembly. In addition, we are interested in probing whether morphology of these two-component amyloid fibrils would change by fusing the same domains, including CBD, Mfp3 (mussel foot protein 3) and Mfp5 (mussel foot protein 5), at different positions. Finally, we selected diverse subunits—single loop of CsgA, CsgA, 2 tandem repeats of CsgA ($2\times$CsgA) and 4 tandem repeats of CsgA ($4\times$CsgA) and determined how they vary in morphology (Fig. 2a–c).

We started with genetic construction of several recombinant plasmids including pET22b/CsgA, pET22b/CsgA-SpyTag, pET22b/CsgA-SUMO, pET22b/CsgA-CBD, pET22b/NCsgA-CBD-CCsgA, pET22b/CBD-CsgA, pET22b/CsgA-Mfp3, pET22b/NCsgA-Mfp3-CCsgA, pET22b/Mfp3-CsgA, pET22b/CsgA-Mfp5, pET22b/NCsgA-Mfp5-CCsgA, pET22b/Mfp5-CsgA, pET22b/$2\times$CsgA and pET22b/$4\times$CsgA through Gibson assembly (Fig. S1–S12 in Supporting information). Plasmids were then transformed into the DE3 (BL21) competent strains. Target proteins were induced for 2 h by addition of 0.5 mmol/L IPTG and then were purified through cobalt resins.

All the protein bands were identified with expected molecular weights, as show in SDS-PAGE, and were further approved by western blotting (Fig. 2d). All freshly purified proteins tend to aggregate and form suspension after solution incubation for several days. In order to confirm if these aggregates are amyloid fibrils, we carried out Congo red staining, a dye assay that is often used to specifically recognize amyloid feature of proteins [12]. The results suggest that all protein aggregates could be stained by congo red, implying that these protein aggregates are indeed composed of amyloid fibrils (Fig. 2e).

To study how varied fusion domains affect fibril assembly and fibril morphologies, we constructed CsgA, CsgA-SpyTag, CsgA-CBD and CsgA-SUMO, respectively, containing fusion domains with increased molecular weight (Fig. 2a). The AFM images for amyloid fibrils incubated on mica for 18 h were presented in Fig. 3a–d. All proteins are able to self-assemble into fibril structures. While CsgA-SpyTag fibril morphologically resembles CsgA fibril with continuous and non-branched feature, CsgA-CBD and CsgA-SUMO fibrils are either branched or much shorter compared with CsgA fibrils. Among the three side fusion domains, spytag has the lowest molecular weight (1.47 kDa) and seem to have the least influence on fibril assembly, possibly due to less spatial occupation and thereby less interference on the amyloid cores. We also conducted statistical analysis to compare the average diameter of these amyloid fibrils. The distribution of diameter of amyloid fibrils is shown in Fig. 3e. After fitting with Gaussian function, the average diameter is 0.87 ± 0.09 nm for CsgA, 1.12 ± 0.04 nm for CsgA-SpyTag, 1.52 ± 0.05 nm for CsgA-CBD and 1.76 ± 0.08 nm for CsgA-SUMO (Fig. 3e). Therefore, the average diameter of fibrils is in direct proportion to the molecular weight of the fused domains. When fused onto CsgA amyloid backbones, the additional functional domains tend not to participate in the core construction, but are exposed and randomly distributed around the core structures [19]. These domains occupy some extra spaces and thus contribute to the increased diameter of fibrils compared with CsgA. To assess if side domains will affect the dynamic self-assembly of amyloid fibrils, we carried out THT kinetic assembly assay for CsgA, CsgA-CBD and CsgA-SUMO proteins. The results proved that CsgA-fusion proteins have elongated lag phase after introducing either CBD or SUMO domain compared with CsgA, while the lag phases between the two fusion proteins did not show much difference (Fig. 3f). We therefore confirmed that introduction of functional domains will retard fibril growth of fusion proteins.

Collectively, the above data suggest that fusion of functional domains does not interrupt fibril assembly, but retards fibril growth to certain extent. Introduction of functional side domain onto CsgA also tends to increase the average diameter of resultant fibrils, but often leads to shorter and branched fibrils possibly due to steric hinderance caused by fusion domains.

To assess how varied fusion position will affect fibril assembly, we constructed three similar groups of recombinant proteins, in which the same side domain (CBD, Mfp3 or Mfp5 domain) was fused at N terminus, intermediate position and C terminus of CsgA, respectively (Fig. 2b). Chitin binding domains (CBD) are C-terminus domain of Bacillus circulans chitinase that can bind specifically to insoluble chitin, which was reported to have rigid and compact twisted β-sandwich structures [17]. Mussel foot proteins are interfacial adhesion proteins of mussels and usually possess unstructured random coil structures [20].

The AFM images of these series of fibrils were present in Fig. 4. The results indicate that all proteins are able to form self-assembled nanofibrils, though morphologies of fibrils in each

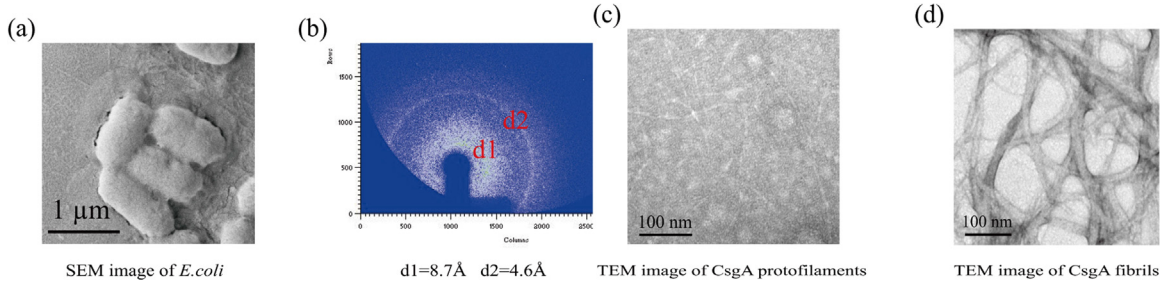

Fig. 1. Characterization of curli fibrils produced by *E. coli* (a) Scanning electron microscope (SEM) image of *E. Coli biofilms*, in which curli fibrils form around the bacteria. (b) X-ray fiber diffraction of CsgA fibrils. d1 and d2 stands for the equatorial reflection and meridional reflection, respectively. (c) Transmission electron microscope (TEM) image of CsgA fibrils collected at early stage of self-assembly. (d) TEM image of CsgA fibrils collected at later stage of self-assembly.

Please cite this article in press as: Q. Qi, et al., Self-assembly and morphological characterization of two-component functional amyloid proteins, Chin. Chem. Lett. (2016), http://dx.doi.org/10.1016/j.cclet.2016.12.008