Biochimica et Biophysica Acta xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Review

Folding mechanisms of periplasmic proteins ☆

Camille Goemans, Katleen Denoncin, Jean-François Collet *

de Duve Institute, Université Catholique de Louvain, 75-39 Avenue Hippocrate, B-1200 Brussels, Belgium
WELBIO, 75-39 Avenue Hippocrate, B-1200 Brussels, Belgium

ARTICLE INFO

Article history:
Received 1 August 2013
Received in revised form 11 October 2013
Accepted 16 October 2013
Available online xxxx

Keywords:
Periplasm
Folding
Chaperone
DegP
HdeA
Spy

ABSTRACT

More than one fifth of the proteins encoded by the genome of *Escherichia coli* are destined to the bacterial cell envelope. Over the past 20 years, the mechanisms by which envelope proteins reach their three-dimensional structure have been intensively studied, leading to the discovery of an intricate network of periplasmic folding helpers whose members have distinct but complementary roles. For instance, the correct assembly of ß-barrel proteins containing disulfide bonds depends both on chaperones like SurA and Skp for transport across the periplasm and on protein folding catalysts like DsbA and DsbC for disulfide bond formation. In this review, we provide an overview of the current knowledge about the complex network of protein folding helpers present in the periplasm of *E. coli* and highlight the questions that remain unsolved. This article is part of a Special Issue entitled: Protein trafficking.

1. Introduction

Unraveling the mechanisms by which proteins fold into their correct three-dimensional structure is a fundamental but complex question in basic biology. Although all the information necessary for a protein to attain its native structure is contained in its amino acid sequence, efficient protein folding *in vivo* requires the participation of various factors, including molecular chaperones, folding catalysts and proteases.

In Gram-negative bacteria such as *Escherichia coli*, a relatively well understood quality control machinery is present in the cytoplasm to ensure the proper folding of newly-synthesized polypeptide chains as they emerge from the ribosome. Indeed, successful folding of nascent proteins is essential for bacterial viability. However, although protein synthesis takes place in the cytoplasm, more than 20% of the proteins encoded by the *E. coli* genome are destined to the bacterial cell envelope. In this article, we will review the mechanisms of protein folding in this extracytoplasmic compartment.

The envelope of Gram-negative bacteria is composed of two membranes: the inner membrane (IM), which is in direct contact with the cytoplasm, and the outer membrane (OM), which constitutes the interface between the cell and the external environment [1] (Fig. 1). The IM and the OM have different structures and composition [1]. The IM is a classical phospholipid bilayer and IM proteins often are integral proteins crossing the membrane with one or more hydrophobic α-helices. A few lipoproteins are also anchored to the outer leaflet of the IM via a lipid moiety [2] (Fig. 1). Unlike the IM, the OM is an asymmetric bilayer composed of phospholipids and lipopolysaccharides (LPS) in the inner and outer leaflet, respectively [3]. OM proteins can be of two types: lipoproteins, most of which are anchored by a lipid moiety in the inner leaflet of the OM and face the periplasm [4], and integral membrane proteins, known as OMPs. These latter generally adopt a β-barrel conformation and serve as channels or «porins» that enable free diffusion of ions and hydrophilic molecules across the membrane [5,6] (Fig. 1).

The IM and the OM are separated by the periplasm, a viscous and oxidizing compartment that contains a thin layer of peptidoglycan and represents 10 to 20% of the total cell volume [7]. More than 300 proteins are present in the periplasm [8] where they perform a large variety of physiological functions, such as protein folding, uptake and transport of nutrients and detoxification of harmful substances.

Secreted proteins are synthesized in the cytoplasm as pre-proteins that are translocated across the IM by different secretory machineries, depending on the signal sequence they carry. The majority of secreted proteins carry a signal sequence recognized by the Sec apparatus [9–11], which transports them through the IM in an unfolded conformation. Although some proteins are secreted co-translationally by the Sec machinery, most are targeted post-translationally to the envelope [12,13]. In this latter case, pre-proteins first bind to the chaperone SecB whose role is to maintain them in their fully unfolded state until they reach the translocase [12,13]. In the co-translational targeting mechanism, the signal sequence of the protein is recognized by the signal recognition particle (SRP) while it emerges from the ribosome and the entire SRP–ribosome–nascent protein complex then binds to the Sec translocase [12]. A small subset of approximately 30 proteins

☆ This article is part of a Special Issue entitled: Protein trafficking.

* Corresponding author at: Tel.: +32 2 764 75 62; fax: +32 2 764 75 98.
E-mail address: jfcollet@uclouvain.be (J.-F. Collet).

0167-4889/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbamcr.2013.10.014

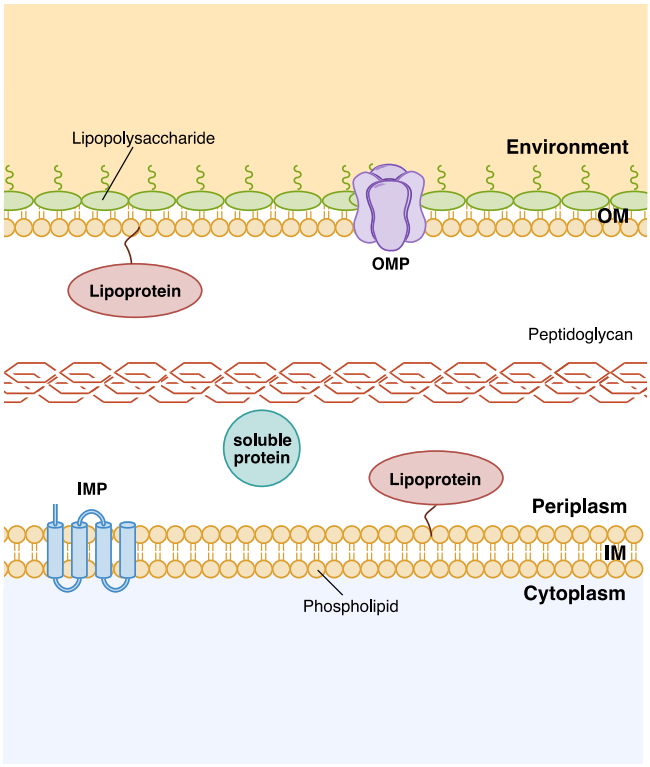

Fig. 1. General structure of the *E. coli* envelope. The envelope of *E. coli* is composed of the inner membrane (IM), the periplasm and the outer membrane (OM). The IM is a symmetric bilayer containing phospholipids and integral membrane proteins (IMP) with α-helical transmembrane domains. The OM is an asymmetric lipid bilayer with phospholipids in the inner leaflet and lipopolysaccharides in the outer leaflet. The OM also contains integral proteins (OMP), known as β-barrels. Both membranes comprise lipoproteins, anchored in the membrane by a lipid moiety and facing the periplasm. The periplasm is a viscous compartment comprised between the two membranes and which contains a thin layer of peptidoglycan.

is transported across the IM by the Twin Arginine transport (Tat) pathway. In this case, the signal peptide presents a characteristic twin-arginine motif [14]. In contrast to the Sec machinery, Tat-translocation substrates fold in the cytoplasm before crossing the membrane and are consequently functional directly after translocation [15]. Noteworthy, Gram-negative bacteria have evolved additional machineries to mediate protein secretion, such as the systems used by pathogenic bacteria to infect host cells. We refer the reader to the chapters of this special issue on protein translocation for more details on these systems.

In this chapter, we will focus on the fate of the polypeptides as they exit the Sec translocon and enter the periplasm, using *E. coli* as a model. The unraveling of the protein quality control mechanisms of the *E. coli* periplasm started in the early 90s with the discovery of catalysts of disulfide bond formation and of peptidyl-prolyl *cis-trans* isomerization [16–19]. The next step was the identification of periplasmic chaperones as a result of the independent work of several groups. Indeed, in 1996, Missiakas *et al.* [20] identified SurA, FkpA and Skp as potential periplasmic chaperones by searching for periplasmic factors decreasing the stress response induced by the accumulation of misfolded proteins in the periplasm [20]. That same year, SurA and Skp were shown to be involved in the folding of OMPs, confirming their identification as periplasmic chaperones [21,22]. Within the next four years, LolA was identified as a general chaperone for most lipoproteins [23] while DegP, which had been known as a protease for a long time, was identified as a protein whose chaperone activity dominates at low temperatures [24]. Other periplasmic chaperones have been identified more recently, including HdeA, a protein that plays an important role in acid survival [25,26], and Spy [27]. We will summarize here these 20 years of multidisciplinary research that led to the discovery of an intricate protein network controlling the folding and integrity of envelope proteins.

2. Periplasmic molecular chaperones

Proper folding of periplasmic proteins requires the assistance of molecular chaperones that are thought to differentiate properly folded proteins from their non-native conformations by recognizing the surface-exposed hydrophobic areas displayed by these latter. A remarkable feature of periplasmic chaperones is that, in contrast to their cytoplasmic counterparts, they assist protein folding without the need of ATP for their activity. So far, periplasmic chaperones have been shown to be involved in two major processes, the maturation of proteins located in the OM (OMPs and lipoproteins) and the protection of periplasmic proteins under stress conditions. In the following section, we will first describe the chaperones involved in OMPs assembly. Then, we will focus on LolA, a chaperone dedicated to lipoprotein transport before describing the stress-induced chaperones.

2.1. Chaperones involved in the biogenesis of OMPs

Newly synthesized OMPs that cross the IM through the Sec machinery need to be escorted by chaperones as they travel through the periplasm to reach the OM. Indeed, since these proteins penetrate the periplasm in an unfolded conformation, they are prone to aggregation in this aqueous compartment. Two parallel folding pathways, which prevent OMPs aggregation during their periplasmic transit, have been described in *E. coli*. The major chaperone pathway involves SurA, while the secondary pathway consists of two proteins, Skp and DegP.

2.1.1. SurA

Originally isolated as a protein essential for survival in stationary phase [28], SurA was later described both as a chaperone that assists the folding of OMPs [20,21,29] and as a peptidyl-prolyl *cis-trans* isomerase (PPIase) [30]. Enzymes with PPIase activity catalyze the *cis-trans* isomerization of peptide bonds involving a proline residue (see Section 3.1).

SurA is composed of four distinct regions: a large N-terminal domain, two PPIase domains of the parvulin family and a short C-terminal helix (Fig. 2). The structure of SurA reveals that the N- and C-terminal domains

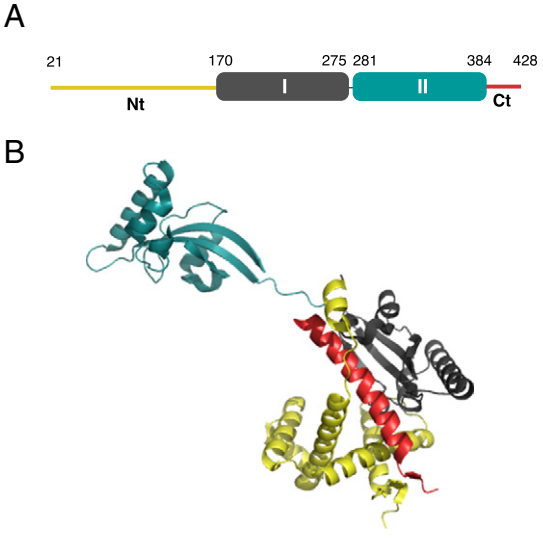

Fig. 2. Structure of *E. coli* SurA. (A) Schematic diagram of the domains of the mature SurA (no signal sequence). The numbers refer to amino acid position. (B) Ribbon representation of SurA (PDB entry code 1M5Y) [31]. The N-terminal domain (yellow) is followed by the two PPIase domains belonging to the parvulin family. The first PPIase domain (grey) has no PPIase activity, whereas the second one (turquoise blue) is active. The C-terminal tail is shown in red. Since polypeptide linkers between domains were poorly ordered, some could not be traced in the structure.

Please cite this article as: C. Goemans, et al., Folding mechanisms of periplasmic proteins, Biochim. Biophys. Acta (2013), http://dx.doi.org/10.1016/j.bbamcr.2013.10.014