Journal of Structural Biology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Review

Ribosomal proteins as documents of the transition from unstructured (poly)peptides to folded proteins

Andrei N. Lupas *, Vikram Alva

Department of Protein Evolution, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany

ARTICLE INFO

Article history:
Received 2 February 2017
Received in revised form 23 April 2017
Accepted 24 April 2017
Available online xxxx

Keywords:
Ancient peptides
Protein evolution
Protein folding
Ribosome
Ribosomal protein
RNA
Secondary structure

ABSTRACT

For the most part, contemporary proteins can be traced back to a basic set of a few thousand domain prototypes, many of which were already established in the Last Universal Common Ancestor of life on Earth, around 3.5 billion years ago. The origin of these domain prototypes, however, remains poorly understood. One hypothesis posits that they arose from an ancestral set of peptides, which acted as cofactors of RNA-mediated catalysis and replication. Initially, these peptides were entirely dependent on the RNA scaffold for their structure, but as their complexity increased, they became able to form structures by excluding water through hydrophobic contacts, making them independent of the RNA scaffold. Their ability to fold was thus an emergent property of peptide-RNA coevolution. The ribosome is the main survivor of this primordial RNA world and offers an excellent model system for retracing the steps that led to the folded proteins of today, due to its very slow rate of change. Close to the peptidyl transferase center, which is the oldest part of the ribosome, proteins are extended and largely devoid of secondary structure; further from the center, their secondary structure content increases and supersecondary topologies become common, although the proteins still largely lack a hydrophobic core; at the ribosomal periphery, supersecondary structures coalesce around hydrophobic cores, forming folds that resemble those seen in proteins of the cytosol. Collectively, ribosomal proteins thus offer a window onto the time when proteins were acquiring the ability to fold.

© 2017 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Life today results from the information storage provided by nucleic acids (mainly DNA) and the catalytic activity of polypeptides. Since the time of the Last Common Ancestor of all living beings on Earth (LUCA), these macromolecules have been following a tripartite division of labor, in which the information stored in DNA is converted to proteins in a process substantially dependent on RNA; this unidirectional flow of information from nucleic acids to proteins is considered the central dogma of molecular biology (Crick, 1970).

It seems impossible that this elaborate interplay of complex macromolecules could have emerged *de novo* from abiotic processes and it is generally accepted that life must have started in a simpler form, which has been the subject of much theorizing and some experimentation. Among the possibilities considered, by far the most popular and best supported has been that of RNA forming the first systems capable of autocatalytic replication, acting as both the information bearer and the agent of catalysis (e.g. Gesteland et al., 2006; Higgs and Lehman, 2015; Jeffares et al., 1998; Joyce, 2002; Lazcano et al., 1988; but for an alternative view see for example Kurland, 2010). This hypothesis, first formulated by Alexander Rich (Rich, 1962) and given the name of 'RNA world' by Walter Gilbert (Gilbert, 1986), rests substantially on the observation that even today, RNA still acts both as information carrier and catalyst in the biosynthesis of proteins, accepting information from DNA in a transcription step and transferring it to a ribozyme (the ribosome) for translation to a polypeptide sequence. While many obstacles remain to be overcome on the path from inorganic compounds to the first RNA polymers (Bernhardt, 2012; Shapiro, 2007) and a number of simpler, pre-RNA molecules have been discussed as the first information-bearing, autocatalytic entities (e.g. Engelhart and Hud, 2010; Lazcano and Miller, 1996; Orgel, 2000; Trevino et al., 2011), the RNA world is now well established and widely considered to have been the direct precursor to the DNA-protein world of today.

Abbreviations: Last Universal Common Ancestor, LUCA; peptidyl transferase center, PTC.

* Corresponding author at: Department of Protein Evolution, Max Planck Institute for Developmental Biology, Spemannstr. 35, 72076 Tübingen, Germany.
E-mail address: andrei.lupas@tuebingen.mpg.de (A.N. Lupas).

http://dx.doi.org/10.1016/j.jsb.2017.04.007
1047-8477/© 2017 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Please cite this article in press as: Lupas, A.N., Alva, V. Ribosomal proteins as documents of the transition from unstructured (poly)peptides to folded proteins. J. Struct. Biol. (2017), http://dx.doi.org/10.1016/j.jsb.2017.04.007

The involvement of polypeptides in the RNA world, initially in the form of short peptides, most likely occurred very early. Indeed, recent evidence suggests that RNA and peptides co-evolved from the beginning, their building blocks originating from the same chemical reaction networks (Patel et al., 2015; Ritson and Sutherland, 2013). RNA faces a number of limitations in stability and catalytic repertoire, particularly in its inability to mediate redox reactions with free radicals (Bernhardt, 2012; Joyce, 2002); peptides would have offered it many benefits, such as coordinating metals and small molecules, mounting iron-sulfur clusters for redox catalysis, promoting the stability and structural specificity of RNA folding by binding into its grooves, mediating complex formation, and functionalizing the first membranes. While the first peptides to join nucleic acid-based autocatalytic replicators were probably of abiotic origin, natural selection would soon have favored forms encoded and synthesized by nucleic acids. For one, abiotic peptide formation is highly inefficient (e.g. Cleaves et al., 2009; Schreiner et al., 2011), making availability a limiting factor of autocatalytic growth from the start. Also, most peptides, even those composed of the 20 proteinogenic amino acids, are of no structural and functional use to RNA, placing a premium on synthesizing only useful forms and passing the information on to the next generation. Given the broad spectrum of steps needed to fulfill even the basic requirements of an information-bearing chemical system capable of autocatalytic replication, it seems clear that the RNA-peptide world must have achieved considerable complexity well before its transition to the DNA-protein world we observe today. In making this transition, the RNA-peptide world faced a considerable challenge: whereas the chemistry of the RNA-to-DNA transition seems unproblematic (Ritson and Sutherland, 2014), there is a major obstacle on the path from peptides to proteins, known as the protein folding problem.

2. The protein folding problem from an evolutionary perspective

Both nucleic acids and proteins must assume defined three-dimensional structures for their biological activity, but their ability to do so is starkly different. Nucleic acids fold spontaneously and robustly, based primarily on a number of simple base-pairing rules, and can in general be denatured and renatured reversibly by chemical agents or temperature without substantial loss of material (witness for example the polymerase chain reaction). Protein structure, in contrast, is an altogether more complex property and the process by which proteins reach their structure (folding) is easily disrupted and readily undone by even minor changes in temperature or the chemical environment. Once denatured, proteins tend to aggregate and can either not be renatured, or only with large loss of material, making denaturation a substantially irreversible process. The easy loss of structure in most proteins is due to the low free energy of folding (often equivalent to just a few hydrogen bonds), which places them energetically close to the unfolded state. Their tendency to aggregate upon denaturation is due to the dominant role of the hydrophobic effect in folding, which leads folded proteins to mainly segregate hydrophobic residues to the protein core and hydrophilic residues to the surface. When the hydrophobic residues of the core become exposed in the denatured state, they tend to coalesce into heterogeneous tangles, which are generally impossible to resolve and must be degraded.

The closeness of the structured and unstructured states in most proteins and the many problems arising to living beings from this are documented in the elaborate protein quality control and degradation systems that are universal to life (e.g. Bukau et al., 2006; Gottesman et al., 1997; McClellan and Frydman, 2001). Even in healthy organisms not exposed to stressful conditions, protein misfolding represents an important challenge, as seen for example for the cystic fibrosis transmembrane conductance regulator, of which, in healthy humans, only about a third of the synthesized copies reach the membrane in a folded state (Ward et al., 1995). In old age and disease these problems become potentiated, leading for example in humans to a host of degenerative diseases (Gregersen et al., 2006; Voisine et al., 2010), such as cystic fibrosis, Alzheimer's, Parkinson's, and Huntington's diseases. Given these considerations, it may come as a surprise that natural proteins nevertheless represent a best-case set, because in their overwhelming majority polypeptides do not appear to have a folded structure at all. It is very difficult to estimate the actual proportion of folding polypeptides with any degree of accuracy, since the protein folding problem is still substantially unsolved and the number of sequence possibilities for a polypeptide chain exceeds the number of particles in the known universe already at a chain length of around 60 residues. Nevertheless, a rough estimate is given by screens of polypeptide libraries, which have produced a success rate of less than one in a billion, even when these libraries were biased for specific patterns of hydrophobicity or derived from a random fragmentation of genomic DNA (Keefe and Szostak, 2001; Matsuura et al., 2002; Riechmann and Winter, 2000; Wei et al., 2003).

Given the difficulty polypeptides encounter to reach and maintain a folded state, and the exceedingly low likelihood of newly emerged polypeptides to even have such a state, it is entirely non-trivial to explain how life came to rely so extensively on folded proteins. Looking at proteins today it is clear that nature is bypassing the protein folding problem by generating new proteins through the amplification, differentiation, and recombination of a basic set of autonomously folding prototypes (domains). Through their similarity in sequence and structure, these domains can be classified into a hierarchy of families, superfamilies, and folds (Andreeva et al., 2015; Dawson et al., 2017; Schaeffer et al., 2017), showing that, though seemingly boundless, the diversity of natural proteins is actually rather narrowly circumscribed (see e.g. Koonin et al., 2002). In total, these classifications, as well as large-scale surveys, suggest that there are no more than some 10^4 domain families, prototypes for many of which were already present at the time of LUCA, around 3.5 billion years ago (Koonin, 2003; Kyrpides et al., 1999; Ranea et al., 2006). Domain classifications have been a very powerful tool in retracing the evolution of the protein world that underpins life today, but the origins of domain prototypes themselves have long remained unclear and only started to emerge in recent years.

3. Proteins from peptides

As outlined above, the staggering size of protein sequence space and the low incidence of folded exemplars within it essentially preclude an origin of folded domains by random concatenation of amino acids. An alternative scenario proposes that the first folded domains did not arise from random processes, but from the increased complexity of the peptides that had evolved in the RNA world (Lupas et al., 2001; Soding and Lupas, 2003). In this scenario, the evolutionary pressures operating on peptides within their replicative systems led to the selection of biophysical properties that eventually yielded protein folding as an emergent property.

This scenario proceeds from the assumption that one of the properties under selection from the start must have been the ability of peptides and RNA to interact specifically, an evolutionary pressure resulting as much from a competition of primordial RNAs for a limited pool of peptides as from the greater functional effec-

Please cite this article in press as: Lupas, A.N., Alva, V. Ribosomal proteins as documents of the transition from unstructured (poly)peptides to folded proteins. J. Struct. Biol. (2017), http://dx.doi.org/10.1016/j.jsb.2017.04.007

Download English Version:

https://daneshyari.com/en/article/5591563

Download Persian Version:

https://daneshyari.com/article/5591563

Daneshyari.com