FEBS Letters 584 (2010) 3842–3849

journal homepage: www.FEBSLetters.org

Review

Building arks for tRNA: Structure and function of the Arc1p family of non-catalytic tRNA-binding proteins

Eleftherios Karanasios [a,1], George Simos [a,b,*]

[a] Laboratory of Biochemistry, School of Medicine, University of Thessaly, BIOPOLIS, Larissa 41110, Greece
[b] Institute of Biomedical Research and Technology (BIOMED), 51 Papanastasiou Str., Larissa 41222, Greece

ARTICLE INFO

Article history:
Received 23 July 2010
Revised 14 August 2010
Accepted 16 August 2010
Available online 20 August 2010

Edited by Michael Ibba

Keywords:
tRNA
Arc1p
Trbp111
p43
p38
p18

ABSTRACT

Following the intricate architecture of the eukaryotic cell, protein synthesis involves formation of many macromolecular assemblies, some of which are composed by tRNA-aminoacylation enzymes. Protein–protein and protein–tRNA interactions in these complexes can be facilitated by non-catalytic tRNA-binding proteins. This review focuses on the dissection of the molecular, structural and functional properties of a particular family of such proteins: yeast Arc1p and its homologues in prokaryotes and higher eukaryotes. They represent paradigms of the strategies employed for the organization of sophisticated and dynamic nanostructures supporting spatio-temporal cellular organization.

1. Introduction

The central dogma of molecular biology states that the genetic information is encoded in DNA, transcribed to mRNA and finally translated to protein [1]. While complementarity is enough to drive transcription, charged tRNAs are the key molecules bridging the worlds of nucleic acids and proteins. After years of research, the biochemistry of protein synthesis is becoming definite and defined in terms of the factors involved, but the cell biology of tRNA synthesis and function still adds new pieces into the jigsaw: in the traditional view, pre-tRNA transcripts originating in the nucleus are subsequently processed at their 5′ and 3′ termini, decorated with several nucleotide modifications and, for intron-containing tRNA species, spliced before the mature molecules are exported to the cytoplasm, where they are charged with the appropriate amino acid and participate in protein synthesis [2]. However according to recent advances, tRNA aminoacylation occurs also in the nucleus, splicing does not occur exclusively in the nucleus and tRNA transport is not unidirectional but can include nuclear export, retrograde nuclear import and nuclear re-export (see recent review [3]).

Along its travels tRNA encounters and interacts with processing, modification and aminoacylation enzymes as well as transport and translation factors that guide it through the cellular compartments and onto the ribosome. Apart of these proteins, tRNAs are often escorted by tRNA-binding proteins, which perform no apparent catalytic function but can be involved in the formation of multi-enzyme complexes [4]. A well-characterized prototype of these proteins, yeast Arc1p, is the subject of this review. It will be examined together with its homologues in prokaryotes, such as Trbp111, or higher eukaryotes, such as p43 (pro-EMAPII), p38 and p18. Non-catalytic appended domains of aminoacyl-tRNA synthetases (aaRSs), that are evolutionary and functionally linked to Arc1p and its orthologues, have been the main subject of recent reviews [5,6] and will be, therefore, only briefly mentioned when necessary.

2. Arc1p mediates the formation of a multi-synthetase complex in lower eukaryotes

Arc1p (its name being derived from **a**a**R**S **c**ofactor **1**) was originally identified in 1996 as the product of a yeast gene, mutations in which caused synthetic lethality when combined with

Abbreviations: aaRS, aminoacyl-tRNA synthetase; NLS, nuclear localization signal; NES, nuclear export signal

* Corresponding author at: Laboratory of Biochemistry, School of Medicine, University of Thessaly, BIOPOLIS, Larissa 41110, Greece. Fax: +30 2410 685545.
E-mail addresses: ek340@cam.ac.uk (E. Karanasios), simos@med.uth.gr (G. Simos).
[1] Present address: Cambridge Institute for Medical Research, Wellcome Trust/MRC Building, Addenbrooke's Hospital, Hills Road, Cambridge CB2 0XY, UK.

0014-5793/$36.00
doi:10.1016/j.febslet.2010.08.023

mutations in the LOS1 gene [7]. Los1p was known at the time to be a nuclear-pore associated protein involved in tRNA biogenesis [8] but was subsequently shown to be the yeast tRNA exportin mediating transport of tRNAs from the nucleus to the cytoplasm [9]. Affinity purification of Arc1p revealed its tight association with two proteins identified as yeast cytosolic MetRS and GluRS [7].

This was the first demonstration of a multi-aaRS complex in unicellular eukaryotes. Higher multi-cellular eukaryotes were already known to contain a much larger multi-enzyme complex comprising nine aaRSs (including MetRS and GluRS) and three non-catalytic proteins, p18, p38 and p43 [10]. The formation of this complex was thought to be mediated by the non-catalytic appended domains of the aaRSs, which were characteristic of the eukaryotic enzymes and absent from their prokaryotic homologues. Affinity purifications as well as in vitro binding studies demonstrated that the yeast aaRS complex could only be formed in the presence of Arc1p, which could associate simultaneously with both MetRS and GluRS, while the enzymes could not directly interact with each other [7,11].

3. The primary structure organization and homologies of Arc1p

Primary structure analysis of Arc1p revealed a basic polypeptide sequence of 376 residues, comprising three fairly distinguishable domains named N, M and C (Fig. 1) [7]. The N-terminal (N) domain (residues 1–131; Arc1-N) displayed at first little homology to other known proteins [7]. However, it was later shown to contain a conserved sequence motif, originally identified in the bacterial glutathione *S*-transferase (GST) super-family but, also present in many other eukaryotic components of the translation machinery including the N-terminal appended domains of many eukaryotic GluRSs, mammalian MetRS, ValRS and CysRS, eukaryotic elongation factors eEF-1β and eEF-1γ and the two non-catalytic components of the mammalian multi-synthetase complex, p18 an p38 [5,12,13]. As described in detail below, this GST-like N-domain of Arc1p was subsequently shown to be responsible for its interactions with the corresponding domains in GluRS and MetRS [11,12,14,15], suggesting a prototype mode of association that was also later demonstrated for most of the other aaRS-containing complexes [5].

The middle (M) domain of Arc1p stretches between residues 132 and 200 (Arc1-M), is very basic, rich in alanines and lysines and shows sequence similarity to H1 histones [7]. As also discussed below, it binds RNA non-specifically and can, thus, contribute to the affinity of Arc1p for tRNA [7,14]. Finally, the C-terminal (C) domain (residues 201–376; Arc1-C) is neutral and exhibits significant homology (54% identity) to the human protein EMAP II (endothelial-monocyte-activating polypeptide II) [16], which was subsequently shown to represent a C-terminal proteolytic fragment of the third non-catalytic component of the mammalian multi-synthetase complex, p43 [17]. Similar homology was later also shown to the C-terminal appendix of *Caenorhabditis elegans* and plant MetRS [14,18] and human TyrRS [5,14].

The first half of the C-domain of Arc1p (residues 202–306, termed Arc1-C_1 [12]) contains a distinct structural motif called the OB-fold and is also homologous to the C-terminal appendix of *Escherichia coli* MetRS, the N-terminal part of the β subunit of *E. coli* PheRS [14] and, most importantly, to the core sequence of the bacterial structure-specific tRNA-binding protein Trbp111 [19]. The OB fold is also part of the anticodon binding domain of three unrelated aaRSs (AsnRS, AspRS, and LysRS). In contrast, the second half of the C-domain (Arc1-C_2, residues 307–376) has homologies only among eukaryotic proteins. As described below, the M and C parts of Arc1p constitute its tRNA-binding domain (TRBD). Overall, Arc1p shares homologies with all the three non-catalytic components of the mammalian multi-synthetase complex (p18, p38 and p43), probably reflecting a functional conservation between the yeast and the mammalian complexes.

4. The tRNA binding properties of Arc1p

The original identification of Arc1p as a protein genetically linked to the tRNA transport pathway and its physical association with two tRNA-aminoacylation enzymes probed experiments to test its ability to interact with tRNA. Using electrophoretic mobility shift assays, it was shown that recombinant Arc1p could bind to in vitro transcribed yeast $tRNA^{Met}$. This binding was competed by other tRNAs while competition by either unstructured or highly structured RNA molecules other than tRNA (such as 5S rRNA) failed to saturate the tRNA binding site of Arc1p suggesting specific interaction with the tRNA tertiary structure [7]. Furthermore, an acceptor minihelix (comprising TψC arm-acceptor stem), but not an anticodon hairpin minihelix, partially challenged tRNA binding and indicated that the conserved TψC loop should contribute to the interaction [7]. Performing similar experiments with truncation mutants of Arc1p, made clear that the N-domain of Arc1p was dispensable for tRNA binding. In contrast, both the M-domain and the C-domain could individually interact with tRNA, albeit with low affinity. However, when the two domains were combined (as in an Arc1-ΔN mutant) binding affinity for tRNA was similar to that of the full-length protein, which was calculated to correspond to a K_d of approximately 10 nM [14]. Additional cross-linking and gel retardation experiments [7,14] indeed suggested that a high affinity and specific tRNA-binding domain (TRBD) is formed in Arc1p by its M-part (which provides non-specific RNA–protein affinity) and C-part (which provides tRNA-structure specificity), in accordance also with the homologies described above. Overall, these results predicted that the TRBD recognized the elbow of the general L form of tRNA, formed by the D- and TψC-loops, where conserved invariant bases are clustered, and left the anticodon arm and the acceptor helix free for interactions with aminoacylation enzymes.

When Arc1p was challenged with total yeast tRNA, it preferentially bound to a small group of tRNAs, which included the "cognate" tRNAs for the Arc1p-MetRS-GluRS complex ($tRNA^{Glu}$ and $tRNA^{Met}$), suggesting that Arc1p, in addition to its tRNA-structure specificity, may also exhibit limited tRNA-species specificity [11] and hinting that selective pressure rather than chance forced the incorporation of MetRS and GluRS in the yeast and probably also the mammalian complex. Analysis of the sequence of these tRNAs revealed the presence of certain unique elements to this group, including the C2–G71 and G51–C63 base pairs in the acceptor arm, the G10–C25–G45 triad in the base of the anticodon stem and C32 and G57 in the anticodon and TψC loops, respectively. These nucleotides may act as positive determinants for the interaction with Arc1p, a notion which has, however, never been confirmed experimentally (e.g., by foot-printing). The preference of Arc1p for G–C elements may, nevertheless, explain its previous identification as a protein (termed G4p1) with affinity for quadruplex (G4) nucleic acids, structures formed in vitro by nucleic acids that contain guanine tracts [20].

5. The tRNA binding properties of Trbp111 and p43

5.1. Trbp111

The demonstration of the tRNA-binding function of Arc1p triggered analogous studies in its TRBD-containing homologues in bacteria (Trbp111) and mammals (p43), for which also crystal structure information became soon available. Trbp111 (tRNA