



HAL
open science

Protein Fold Usages in Ribosomes: Another Glance to the Past

Inzhu Tanoz, Youri Timsit

► **To cite this version:**

Inzhu Tanoz, Youri Timsit. Protein Fold Usages in Ribosomes: Another Glance to the Past. International Journal of Molecular Sciences, 2024, 25, 10.3390/ijms25168806 . hal-04733039

HAL Id: hal-04733039

<https://hal.science/hal-04733039v1>

Submitted on 11 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Article

Protein Fold Usages in Ribosomes: Another Glance to the Past

Inzhu Tanoz¹ and Youri Timsit^{1,2,*}

¹ Aix-Marseille Université, Université de Toulon, IRD, CNRS, Mediterranean Institute of Oceanography (MIO), U.M. 110, 13288 Marseille, France; inzhu.tanoz@etu.univ-amu.fr

² Research Federation for the Study of Global Ocean Systems Ecology and Evolution, FR2022/Tara GOSEE, 3 Rue Michel-Ange, 75016 Paris, France

* Correspondence: youri.timsit@mio.osupytheas.fr

Abstract: The analysis of protein fold usage, similar to codon usage, offers profound insights into the evolution of biological systems and the origins of modern proteomes. While previous studies have examined fold distribution in modern genomes, our study focuses on the comparative distribution and usage of protein folds in ribosomes across bacteria, archaea, and eukaryotes. We identify the prevalence of certain ‘super-ribosome folds,’ such as the OB fold in bacteria and the SH3 domain in archaea and eukaryotes. The observed protein fold distribution in the ribosomes announces the future power-law distribution where only a few folds are highly prevalent, and most are rare. Additionally, we highlight the presence of three copies of proto-Rossmann folds in ribosomes across all kingdoms, showing its ancient and fundamental role in ribosomal structure and function. Our study also explores early mechanisms of molecular convergence, where different protein folds bind equivalent ribosomal RNA structures in ribosomes across different kingdoms. This comparative analysis enhances our understanding of ribosomal evolution, particularly the distinct evolutionary paths of the large and small subunits, and underscores the complex interplay between RNA and protein components in the transition from the RNA world to modern cellular life. Transcending the concept of folds also makes it possible to group a large number of ribosomal proteins into five categories of urfolds or metafolds, which could attest to their ancestral character and common origins. This work also demonstrates that the gradual acquisition of extensions by simple but ordered folds constitutes an inexorable evolutionary mechanism. This observation supports the idea that simple but structured ribosomal proteins preceded the development of their disordered extensions.



Citation: Tanoz, I.; Timsit, Y. Protein Fold Usages in Ribosomes: Another Glance to the Past. *Int. J. Mol. Sci.* **2024**, *25*, 8806. <https://doi.org/10.3390/ijms25168806>

Academic Editor: Oxana V. Galzitskaya

Received: 19 July 2024

Revised: 7 August 2024

Accepted: 8 August 2024

Published: 13 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: protein fold; ribosome evolution; super-ribosome folds; molecular convergence; power-law distribution; RNA-protein co-evolution

1. Introduction

Ce texte cite une «certaine encyclopédie chinoise » où il est écrit que «Les animaux se divisent en: (a) appartenant à l'Empereur, (b) embaumés, (c) apprivoisés, (d) cochons de lait, (e) sirènes, (f) fabuleux, (g) chiens en libertés, (h) inclus dans la présente classification, (i) qui s'agitent comme des fous, (j) innombrables, (k) dessinés avec un pinceau très fin en poil de chameau, (l) et caetera, (m) qui viennent de casser la cruche, (n) qui de loin semblent des mouches» (Borges). Préface de «les mots et les choses» Michel Foucault.

1.1. About Usages in Complex Systems

The usage or preferred usage of the elements that compose a complex system can reveal valuable information about its nature, the laws that govern it, its evolution, or the respective functions of each of its components. Certain systems that are the fruit of different evolutionary processes, such as the universe, life, human languages, or scientific literature, present analogies in the occurrence of their constituents. For example, far from being homogeneous, the use of words in a language, references in scientific papers, or codons or folds in genomes are biased and often form part of what is known as a power law or Zipf's

law [1–7]. However, these analogies may derive from different evolutionary mechanisms and organisational principles of their own. In linguistics, the use of characters, words, or ideograms can, for example, reveal the grammatical principles governing the language and its evolution. The preferential use of synonyms can also provide interesting information about the semantic features, the author style or the text purpose's [8–11]. Biological systems are also characterised by biases in the use of certain elements, which not only reveal their properties but can also provide clues to their history. For example, in modern bacteria, the bias of riboswitches towards some ligands may be evidence of systems that existed in the RNA world [12]. Another phenomenon—codon usage bias—is the preferential use of certain synonymous codons in genomes, genes, or even different gene regions [13,14]. Growing evidence shows that codon usage may modulate the speed of protein biosynthesis in the ribosome. The resulting non-uniform ribosome local kinetics—denoted as translational rhythm—modulate the translation rates and co-translational protein folding of the ribosome [15,16]. Fold usage in structural databases and extant genomes revealed several important features about the laws that govern the protein universe [3,17]. These studies revealed that the distribution of folds also follows a power law, and that there is, therefore, a minority of folds—the superfolds—that are extremely frequent in databases and genomes, such as P-loop NTPases, Rossmann-folds, ferredoxin-like domains, and TIM-barrels [4,18–25]. However, many factors affect the distribution of folds, which, for example, differs according to kingdoms, species, or an organisms' lifestyles. For example, the vibrational energy exchanges among protein residues [26] and allostery [27] have been shown to correlate with protein topology. Protein dynamics vary according to their fold [28], so that each has its own dynamic personality [29]. On the other hand, alpha helices are more robust to mutations than beta strands, [30] and some folds are more likely to evolve and develop new functions than others [31]. Fold usage can therefore provide functional and evolutionary insights as well as information on a fold's ancestral character, bearing in mind that superfolds are considered to be very ancient folds, which would have predated the radiation of the three kingdoms [32–36].

The analysis of fold usage has not yet, to our knowledge, been extensively applied to the functional and evolutionary study of ribosomes, the organelles responsible for protein biosynthesis. The ribosomes are ribonucleoprotein particles (roughly one-third protein and two-thirds RNA for cytoplasmic ribosomes that are composed of a large (LSU) and a small (SSU) subunit). Ribosomes perform translation—mRNA coded protein synthesis—in the three domains of life [37–43]. Since the 2000s, Solving the ribosome structures from the three kingdoms structures from the three kingdoms, with or without their various substrates and factors, has revealed not only the breathtaking details of the mechanisms of each stage of translation (initiation, elongation, termination, and recycling), but also how their complexity has increased over the course of evolution. Thus, the ribosome is an ideal system for studying evolutionary transitions because it constitutes, in particular its peptidyl-transfer center, a 'window to the past' [44–47]. Thought to be an autonomous system before the radiation of the three kingdoms, it may represent a precious missing link in evolution [48–52]. According to various hypotheses, its structure could be used to infer the "first steps" in the evolution of a peptide/RNA world. First of all, ribosomes are a model of choice for understanding the mechanisms and evolution of the concerted assembly of nucleoprotein complexes [53,54]. From a functional point of view, a ribosome is also an extremely complex system that carries out protein biosynthesis. Over the last 20 years, structural studies have made it possible to follow the various stages and dynamics of translation at the angstrom scale [55,56]. These studies have revealed the details of an extremely complex dynamic that underpins the fidelity of protein biosynthesis. These studies also highlighted the synchronization of translation players and communication between the functional sites. Ribosomes are also considered relics of the transition from the RNA world to the first peptide/RNA life forms [57–59], reflecting the co-evolution of translation players and the genetic code [60–62]. Consisting of an interesting repertoire of RNA modules [63,64] and ancestral proteins, they could reflect

the first forms of organization and the primordial mutualism of these two categories of polymers [65]. As such, they would be at the crossroads of the evolutionary processes that probably preceded the last universal common ancestor (LUCA) [66] at the end of the so-called “RNA world”. Numerous studies converge on the idea that ribosomes evolved by accretion and that their constituents have their own evolutionary scenarios [52]. Ribosomal RNA (rRNA) and the ribosomal proteins that make up the two subunits would thus have had different origins and would then have co-evolved to form ‘modern’ ribosomes [67–69]. An important component of ribosomes that play an essential role in their assembly and translation are ribosomal proteins [70–73].

Ribosomal proteins are unique in that they are composed of a globular domain, usually anchored to the surface, and long, disordered, filamentous extensions that thread their way inside the crevices formed by the rRNA [74]. The role of these extensions is still enigmatic, but numerous studies converge on the idea that they play multiple roles. For example, they may participate in the assembly of the large [75] or small subunit [76]. However, some experimental studies have shown that they can both play a role in assembly [77] and participate in very specific tasks during translation [78], such as allosteric coordination between functional sites [79]. The evolution and phylogeny of ribosomal proteins [80,81] has been documented by numerous studies, showing that they have followed complex evolutionary scenarios in different kingdoms. In addition, based on studies of the gradual evolution of protein motifs [82,83], it has been proposed that ribosomal proteins evolved from their disordered extensions because they contact deep, ancient regions of the rRNA [84]. Other studies have shown that ribosomal protein folds subsequently acquired these extensions, which weave increasingly complex networks as they pass from prokaryotes to eukaryotes [85–87]. These networks are thought to be involved in communication between distant functional parts of the ribosome and to synchronize increasingly complex tasks in the course of evolution [88].

To complement existing studies on the evolution and functions of ribosomal proteins, our short paper examines the distribution and usage of protein folds in ribosomes across different kingdoms. While it has been noted that ribosomes contain proteins that adopt very ancient folds, a detailed comparative analysis of protein fold distribution in the ribosomes of the three kingdoms has not, to our knowledge, been carried out. In the light of studies showing how, for example, Small Beta Barrel domains tolerate insertions, decorations, and other variations [89], and knowing that segment insertion and deletion (INDELS) plays an important role in protein evolution [90–94], we also examine how the different ribosomal protein folds acquire extensions. Are there different categories of extensions associated with specific folds? Finally, comparing the use of folds in ribosomes and in extant genomes provides interesting information for assessing differences in selective pressures at different stages of evolution.

1.2. “To Be or Not to Be” a Protein Fold

While we still wonder about the meaning of Shakespeare’s famous “to be or not to be” [95], the biophysics of protein folds [96] could well be engaged in such metaphysical reflection today, seeking their place and meaning in the complex landscape of evolution [97]. The protein fold has been a very useful conceptual unit that has enlightened both structural biology and evolution. However, while ‘structure–function relationships’ have shaped the landscape of molecular biology for almost 60 years, a number of studies have gradually shaken the immutability of the notion of protein’s structure and topology on one side and the concept of function on the other side. “To be or not to be” a protein fold is an important question to examine before analysing “fold usage” in genomes and in ribosome.

A fold is a biological entity that is the product of a spontaneous process, the hierarchical folding of a polypeptide chain into secondary and tertiary structures. Several decades of experimental and theoretical works have converged on the idea that proteins cooperatively fold to their native conformation, driven by negative ΔG , on funnel-shaped energy landscapes in the microsecond or millisecond time range [98–101]. The rate-limiting

event in the folding reaction is the formation of a transition-state whose topology is determined by a set of interactions involving a small number of key residues and is close to that of the native state. Thus, protein sequences encode the information needed in both the hydrophobic folding nucleus and surface to guide the protein-folding reaction and generate the native fold [102–106]. However, a recent work propose an alternative view in which protein folding may follow a wide variety of rugged and dynamic landscapes, resulting in a broad range of thermodynamic and kinetic stability [107]. In addition, while some proteins are capable of folding on their own, others rely on molecular chaperones which are crucial for the folding and assembly of many multi-domain and multi-subunit proteins [108–110].

The origins of the protein fold classification seem to go back to the 1970s and 80s, when the first attempts emerged to categorise protein domains by structural similarity [111,112]. Hierarchical classification systems then progressively counted more than a thousand folds classified according to their contents in secondary structures and their topologies [113–115]. However, as seen in the famous “encyclopédie chinoise” presented by Michel Foucault (see above), the desire to classify living organisms very quickly came up against serious conceptual obstacles. Although very fruitful in the development of structural biology, fold categorizations have gradually found it difficult to adapt to the changing nature and complexity of macromolecules. Protein fold classifications soon came up against numerous pitfalls. One of them is, for example, the discovery that many proteins or protein regions do not adopt a structure and that some of them can only fold in the presence of partners [116]. These discoveries have led to a gradual recognition of the role of disorder in protein functionality, and intrinsically unstructured proteins have become one of the new paradigms of molecular biology in recent decades [117]. Another surprise in the same vein was the discovery by molecular dynamics studies that a significant proportion of folds or domains do not behave as autonomous folding units [118]. So, paradoxically, “not to be” a domain is therefore a relatively common property of protein folds. To complicate matters a little further, it was realised quite early on that the same sequence could alternatively adopt a beta-sheet or alpha-helix structures [119,120]. These switch phenomena, firstly localised in small protein sequences, were subsequently found to be much more frequent and able to extend to large regions or even entire proteins [121–124]. “Structural disorder” has been, indeed, found to be the physiological and functional state for most IDPs/IDRs. To be disordered or not to be disordered: is that still a question for proteins in the cell? [125]. Interestingly, the disorder content seems to increase during evolution, reaching up to 33% in eukaryotic proteomes. Thus, an increase in disordered regions in proteins seems to be associated with new signaling functions present in eukaryotes [126–130] and taking into account the cellular context [131].

1.3. Folding or Misfolding: That’s a Question of Ribosome?

An important step was then taken with the discovery of two different folding states of the same ribosomal protein, bL20, trapped within a single crystal [132]. This study has demonstrated experimentally how the interplay of structural context and electrostatics could influence protein folding. Figure 1a,b also show how the two forms, the partially unfolded form and the native form, stabilize each other in the crystal packing, in the manner of chaperones. Later, numerous studies showed that many “metamorphic” or “transformer” proteins can reversibly adopt distinct folds [133–140]. Among these chameleon-like sequences [141], one category—prions—has attracted particular attention because the conformational conversion of their cellular form (PrP^c) into the pathologic isoform (PrP^{Sc}) has been found to cause fatal neurodegenerative diseases such as Bovine spongiform encephalopathy, scrapie of sheep, and Creutzfeldt–Jakob disease. Prions were the first example of transmissible particles composed exclusively of a modified protein (PrP^{Sc}) that are devoid of nucleic acid [142,143]. Prions have been found in the neurons and glial cells of mammals and birds and also have homologs in yeast, as well as containing a disordered N-terminal region and a helical N-terminal domain [144,145]. The

conversion of the α -helices of the normal functional (PrP^c) protein into β -sheets initiates a self-perpetuating aggregation process in prion diseases [146–152]. Interestingly, the ribosome and RNA play a central role in misfolding processes and the onset of these diseases [153–157].

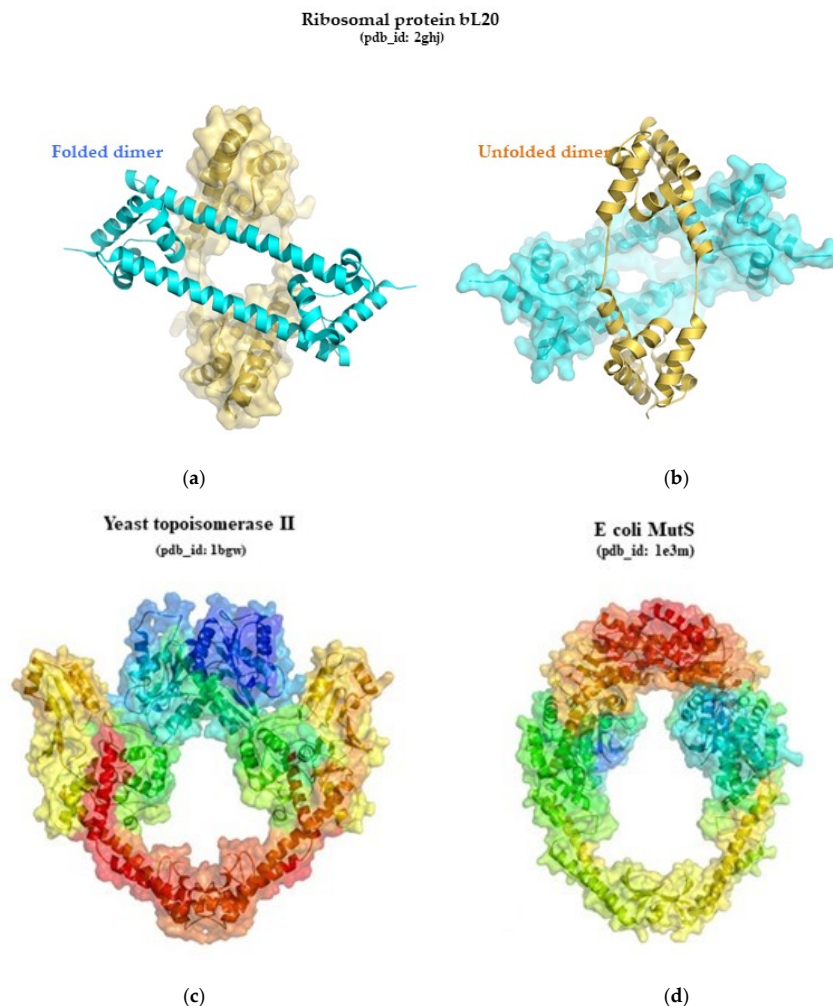


Figure 1. (a,b): the two folding states of the ribosomal protein bL20 (pdb_id: 2ghj); (c,d): similar rings formed by topoisomerase II (pdb_id: 1bgw) and MutS (pdb_id: 1e3m) with opposite chain polarity (colored with rainbow color code in Pymol (from blue, N-ter to red C-ter)).

Beyond the metaphysical aspect of oscillating between “being or not being” a protein fold, folds have been also “transcended” into *urfolds* [158], *metafolds* [159], or “concepts”, [160] entities that considerably broaden the conceptualization of the protein structure universe. Many structures, such as, for example, small β -barrels (SBB), display striking analogies in their properties that “transcend” the concept of folds, allowing to group slightly different folds into a single category. Small beta barrels which include OB-folds and SH3 domains [89] have remarkably similar properties, probably reflecting a common and ancient origin [161]. Along with β -propeller blades [162] and large families, which includes, for example, P-loop NTPases and Rossmann-folds [33,163,164], protein structures can be viewed as entities that transcend the current topological classification [165].

In addition, the discovery of enzyme promiscuity [166], that a folds such as the TIM-barrel can be associated with a large number of functions [167], and, conversely, that different, non-homologous folds have evolved convergently to catalyse similar biochemical reactions [168] defies the classical paradigm of “structure–function relationships”. Convergent evolution at the molecular level is also illustrated by the independent evolution

of propulsion systems for cell motility in the three kingdoms [169]. The reception and emission of light in bioluminescent organisms is also the result of fascinating convergent evolutionary mechanisms [170]. After the “not to be”, the long journey of fold metaphysics also takes us into the realms of “Je est un autre” (“I is another”) (A. Rimbaud), where molecular mimicry [171–174] makes us realize that evolution has played “imitation games”, such as protein folds being interchangeable. Convergent evolution can achieve feats of imitation that make a protein and a nucleic acid resemble each other, even to the point of luring in molecular recognition processes [175,176]. A spectacular illustration of this phenomenon of molecular mimicry is the virtually equivalent ring shape formed by the protein chains of opposite polarity of MutS and topoisomerase II [177] (Figure 1c,d). Interestingly, most of these mechanisms were already present before the radiation of the three kingdoms, and have helped shape the evolution of the ribosomal system.

2. Results

2.1. Ribosomal Protein Fold Distribution

2.1.1. Overview

In the ribosomes of bacteria, archaea, and eukarya, the most prevalent protein fold class is $\alpha\beta$ (approximately 50%), followed by mainly β (approximately 25%) and then mainly α (approximately 17%) (Figure 2a,b). A similar protein fold class distribution with mixed folds belonging to the $\alpha\beta$ class, being the most prevalent followed by mainly β and then mainly α , was observed in modern genomes [4]. Remarkably, this reveals continuity between ribosomes and modern genomes, as the proportions of the secondary structure content have remained consistent. The prevalence of $\alpha\beta$ -class folds aligns with the findings of Winstanley et al. [36], who assert that folds in the $\alpha\beta$ class are more ancient. Additionally, the most common architecture in the ribosomes of the three kingdoms is the two-layer sandwich. In terms of protein fold occurrence, ribosomes exhibit a pattern where only a few topologies are prevalent, while the majority occurs just once (Figure 2c,d). The prevalent folds that occur in the ribosomes two and more times are super ribosome folds (SRF).

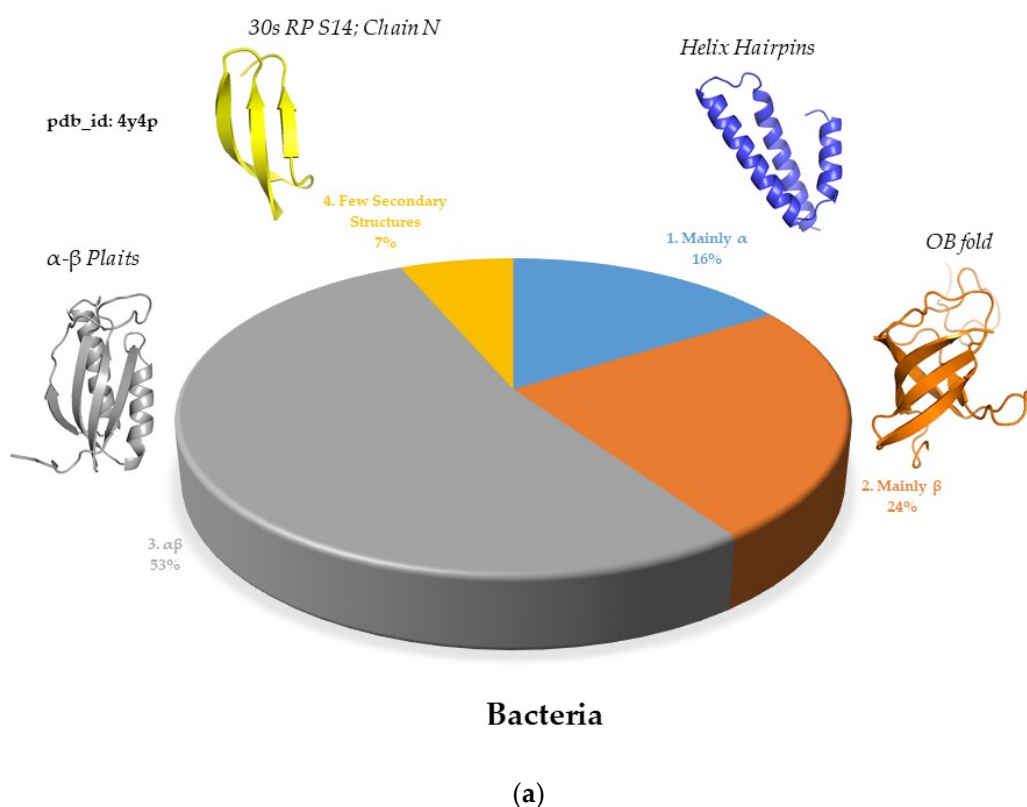
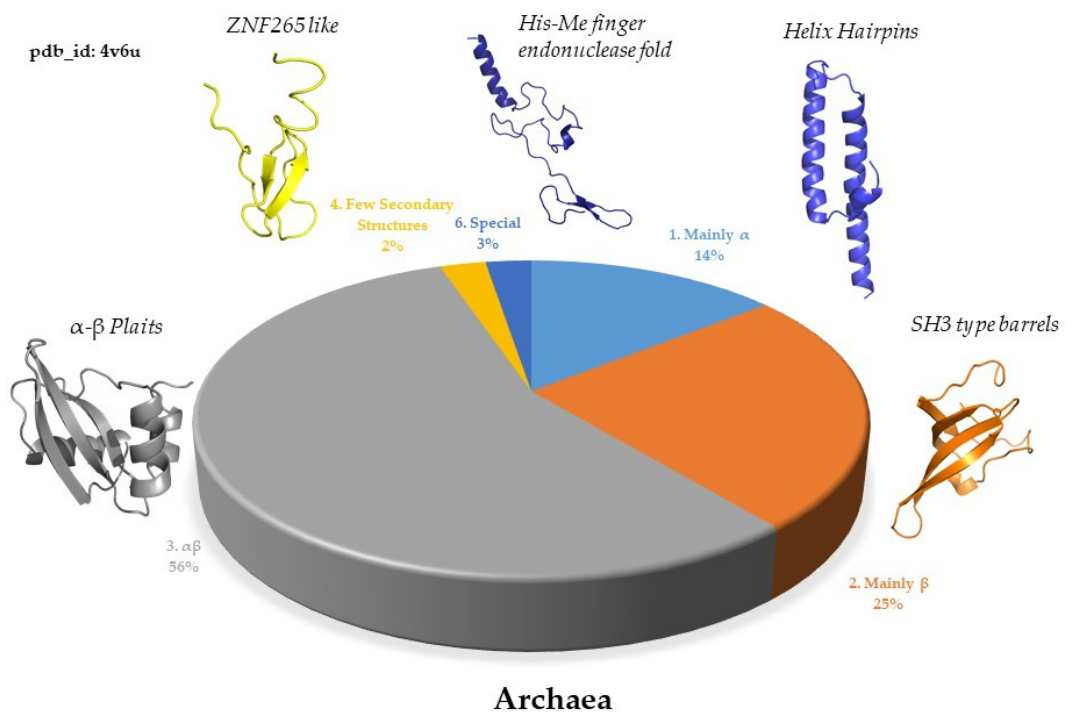
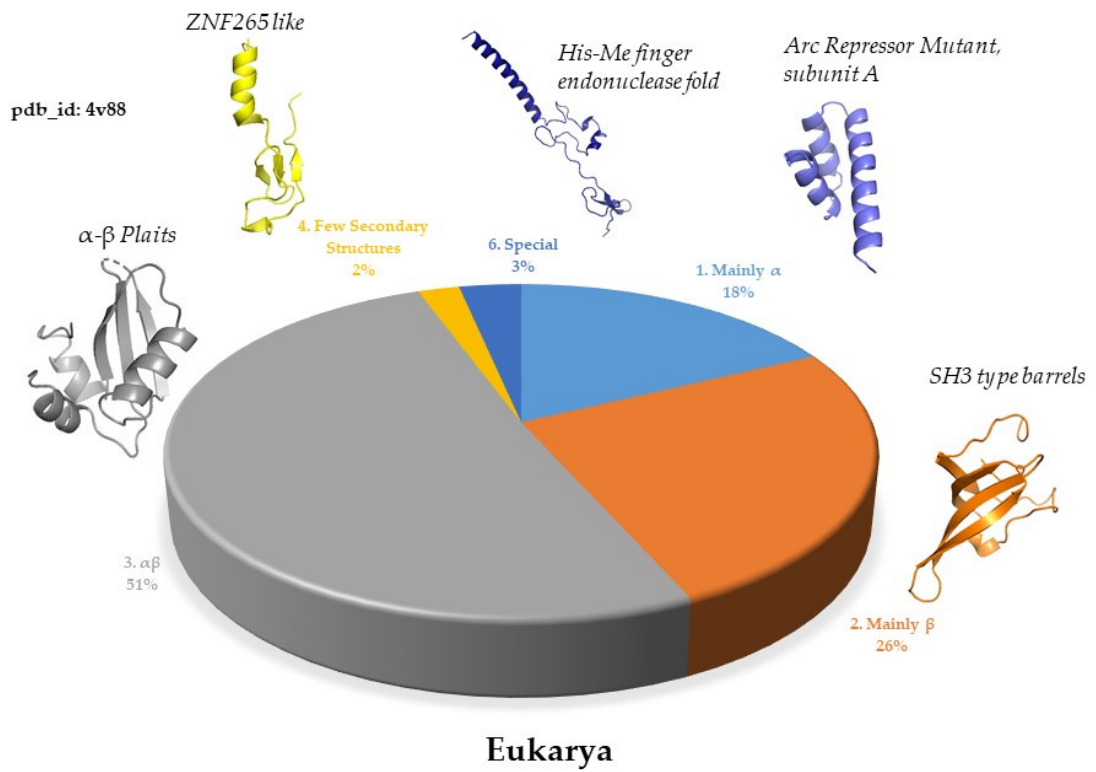


Figure 2. Cont.



(b)



(c)

Figure 2. Cont.

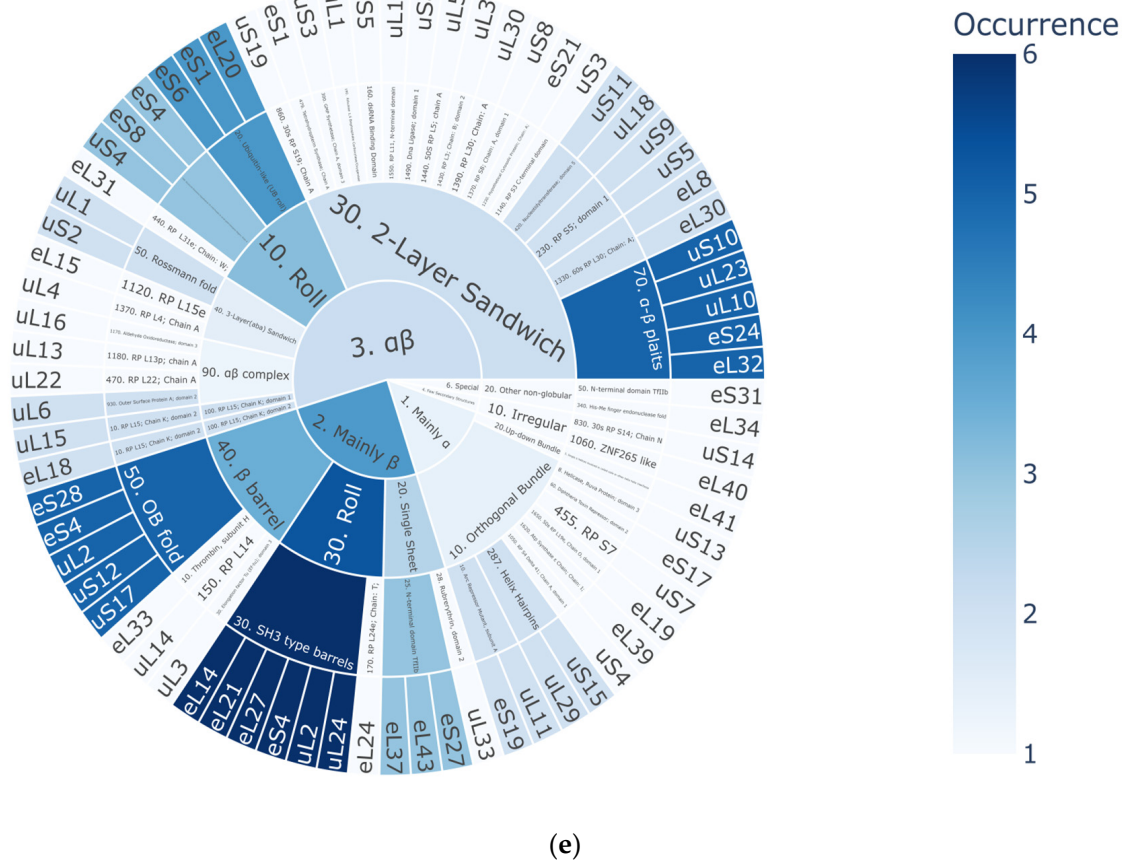
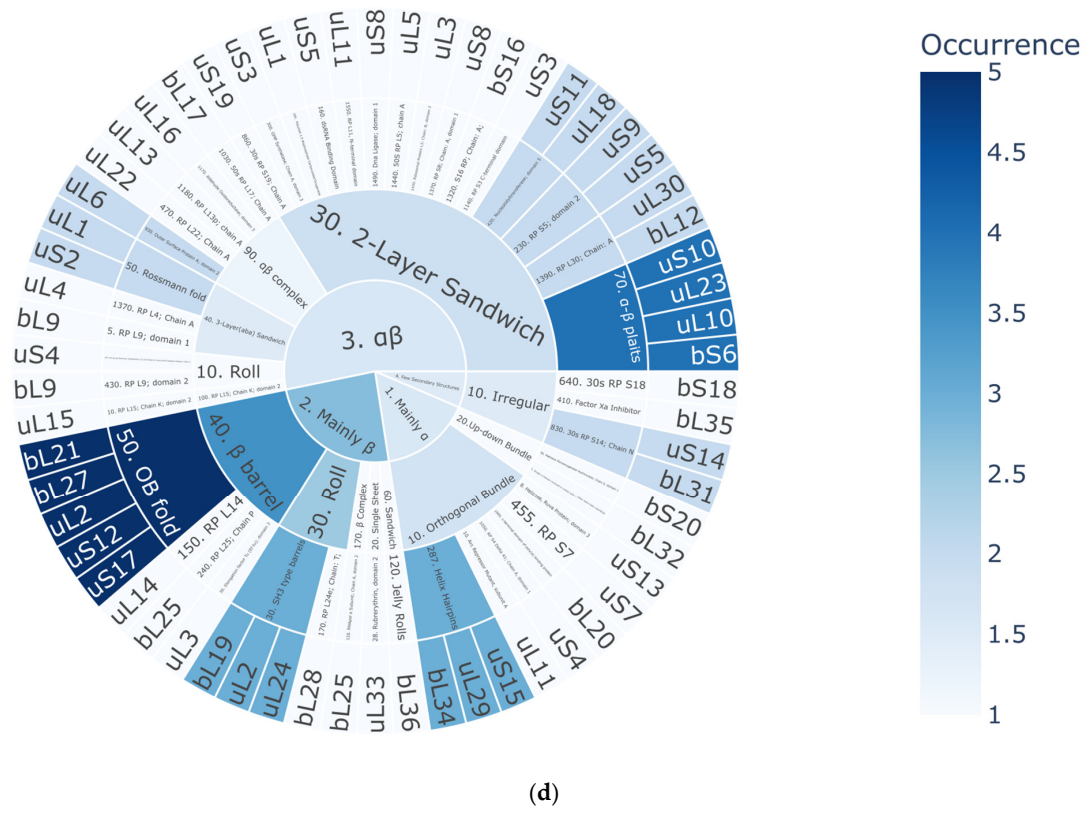


Figure 2. Cont.

being the most frequent, occurring seven and six times, respectively. However, we should note that, according to the recent finding of “ribosome heterogeneity”, slight variations in protein and fold content may also occur among species, among different strains of the same bacterial species, and even between different tissues within the same organism in multicellular organisms [179–183]. The Arc Repressor Mutant, subunit A, which is not a SRF in bacterial and archaeal ribosomes, becomes a SRF in eukaryotic ribosomes by being present in a third of the eukaryote-specific ribosomal proteins, and it remains a superfold in modern genomes. In contrast, other SRFs are more or less shared across the three kingdoms. Interestingly, the bacterial SRFs such as OB folds, α - β Plaits, Helix Hairpins, and SH3 type barrels have remained highly occurring folds and their percentages in modern genomes are likewise high [19]. In contrast, some archaeal and eukaryotic SRFs, such as the N-terminal domain of TflIb and the UB roll, are not superfolds in current genomes. Conversely, some superfolds present in modern genomes, like the Rossmann fold, are not highly represented in ribosomes.

Table 1. Comparison of numbers of proteins, folds, and nucleotides in the three kingdoms’ ribosomes and their LSU and SSU.

Kingdom		Proteins	Folds	Protein/Fold Ratio	Nucleotides
Bacteria	LSU	33	32	1.03	3036 [†]
	SSU	19	20	0.95	1521 [†]
	Total	52	46*	1.13	4557 [†]
Archaea	LSU	39	33	1.18	3175 [†]
	SSU	26	25	1.04	1495 [†]
	Total	65	49*	1.32	4670 [†]
Eukarya	LSU	44	36	1.22	3675 [†]
	SSU	33	29	1.14	1800 [†]
	Total	77	54*	1.42	5475 [†]

* The total number of folds in the ribosome was calculated excluding repetitions. [†] These are the numbers provided by PDB [184] for 4y4p (bacteria), 4v6u (archaea), and 4v88 (eukarya). The number of nucleotides varies among species.

2.1.2. OB and SH3 Domains—The Small Barrel SRFs

Among the most prevalent folds across the three kingdoms are the OB fold and SH3-type barrels, both of which are small β -barrels. These compact, evolutionarily ancient, and functionally versatile protein domains are characterized by their small size (around 100 residues over five β -strands) and highly conserved structural framework [89]. The prevalence of an OB fold and SH3 domain in bacterial and eukaryotic ribosomes, respectively, believed to share a common ancestry [161], aligns with their crucial roles in fundamental biological processes such as transcription, translation, replication, and genome stability [185,186]. The SH3 and OB domains share significant structural similarities, with most residues in four β -strands occupying equivalent positions. Their common elements include the shared strand, the entire CS-sheet, the central strand of the NC-sheet, and a short loop of the NC-sheet, forming the SH3/OB common core. Variations in strand β 5 include its absence in some OB domains, displacement in others, and distortion by additional β -strands in some SH3 domains. Other variable regions generally cause minor perturbations, except for notable cases such as uL24 (SH3) that contains a β hairpin, slightly deforming the CS-sheet [161].

It is interesting to note that the OB fold is overrepresented already in ribosomes due to its unique and versatile properties of RNA binding, and it is also prominent in modern genomes, serving various functions. The widespread presence of the OB fold in modern genomes highlights its two key characteristics: fold-related binding sites that evolution can adapt for diverse functions and the ability to bind a wide range of oligonucleotide sequences. The latter enables rapid and extensive diversification [187]. In contrast, proteins with the SH3 domain in extant organisms can bind to a variety of proline-rich motifs,

participating in signal transduction and the regulation of enzymatic activity [188]. This suggests that ribosomal proteins containing the SH3 domain may share similar properties, supporting the idea of allosteric networks among ribosomal proteins [85] and increased connectivity among eukaryotic ribosomal proteins [87].

2.1.3. The Premise of the Power-Law Behavior

In modern genomes, the distribution of protein folds follows a power-law pattern, with most folds being rare and only a few being highly occurrent [18]. While the number of genes increases exponentially, the diversity of protein folds does not expand at the same rate. When the number of folds in terms of percentages is plotted against the fold occurrences on a log–log graph, a straight line can be observed which indicates the presence of a power-law relationship. This power-law distribution of protein folds illustrates the evolutionary principle that a few highly efficient and adaptable folds can fulfill a wide array of functional roles, while numerous other folds may be more specialized or less common.

Despite the R^2 values on the log–log graphs (Figure 3a–c) not being very close to one, there are clear trends indicating the premises of a power-law relationship. We observe this power-law behavior across different ribosomes, and it becomes more pronounced when we examine the protein fold distributions from bacteria to eukarya (Figure 3d–f). Eukaryotic ribosomes contain more SRFs and they take up a bigger proportion compared to their bacterial and archaeal counterparts, with the prevalence of these SRFs progressively increasing from bacteria to eukarya. The protein-to-fold ratio also increases (Table 1), indicating that, with the emergence of archaea- and eukarya-specific proteins, the number of protein folds does not increase as significantly (Figure 4). This supports the idea that ribosomes might represent a crucial link between prebiotic chemistries and the emergence of the first genomes and cells. Thus, ribosomes, along with their translation factors and aminoacyl-tRNA synthetases, can be viewed as a primitive autonomous biological system [49].

Importantly, as mentioned earlier, the prevalent folds in ribosomes differ from those found in extant genomes. For instance, the Rossmann fold, which is one of the most prevalent superfolds in extant genomes due to its ability to accommodate a wide variety of catalytic functions, appears in the three kingdoms' ribosomes only three times. More precisely, it is present in the form of minimal Rossmann-like motif (RLM). Both uS2, one of the domains of uL1 [189], and uL4 (see below in the section urfolds) contain $\beta 1$, $\alpha 1$, and $\beta 2$ elements followed by an additional helix and an additional strand ending with $\alpha 2$ and $\beta 3$ elements [190].

2.1.4. Different Fold Usages in the LSU and SSU

In the ribosomes of all three kingdoms, the general compositions of the large (LSU) and small (SSU) ribosomal subunits differ significantly due to the abundance of topologies specific to each ribosomal protein. This fits well with evolutionary studies suggesting that the two subunits evolved independently [52,67,68]. Although the exact timeline of their relative ages remains debated, it is believed that each subunit has distinct molecular origins. The evidence indicates that both subunits independently developed their RNA folding capabilities and ancestral functions before co-evolving to form the modern ribosome.

We noticed a bigger degree of differentiation between the LSU and SSU in archaeal and eukaryotic ribosomes in comparison to those in bacterial ones. Specifically, when comparing the LSU and SSU of bacterial ribosomes, it is evident that, while each ribosomal protein typically has distinct topologies, they share the same prevalent folds. Even though SH3-type barrels appear three times exclusively in the LSU and RP S5 domain 2 appears twice only in the SSU; some of the other most common folds, such as the OB fold and α - β Plaits, have similar occurrences in both subunits (Figure 4a). In contrast, in archaeal and eukaryotic ribosomes, the differences in fold distribution between the subunits are more pronounced (Figure 4b). For example, the SH3-type barrel appears six times in the LSU but only once in the SSU. Conversely, in the SSU, the OB fold is the most prevalent, occurring

four times, while it appears only once in the LSU. This pattern supports their independent evolutionary paths and specialized functions within the ribosome.

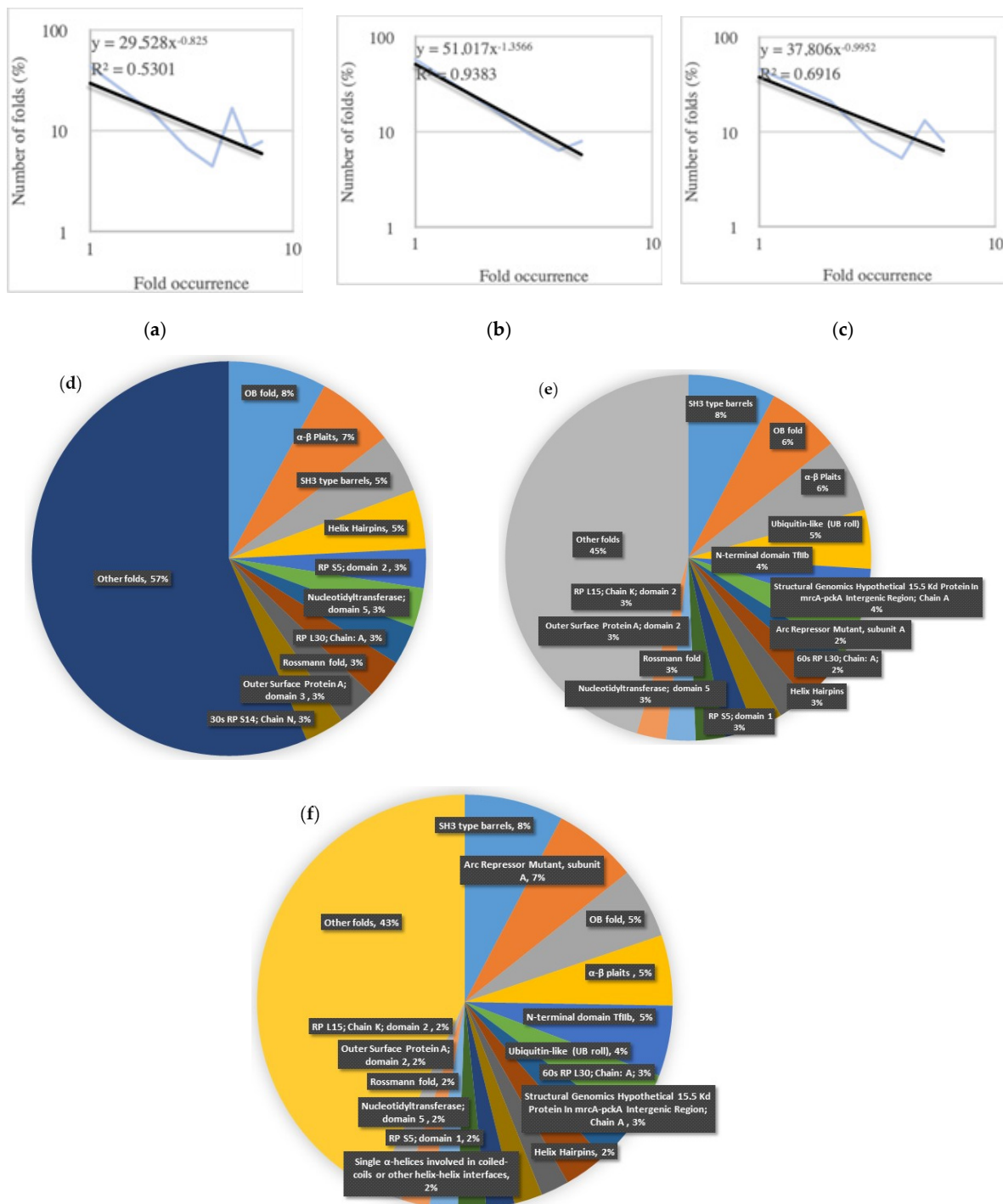


Figure 3. Ribosomal protein fold content represented by counting the occurrences of different protein folds and then grouping those with similar frequencies to display the premises of the power-law distribution, where the number of folds (F) with a specific genomic occurrence (V) follows the equation $F = aV^{-b}$. The vertical axis represents the normalized number of folds relative to the total fold types in the genome on a logarithmic scale, plotted against the occurrence of each fold also on a logarithmic scale in (a) bacterial ribosomes; (b) archaeal ribosomes; and (c) eukaryotic ribosomes [18]. The lines in black are the trend lines and lines in light blue represent the real data. The “super-ribosome folds” of (d) bacterial ribosomes; (e) of archaeal ribosomes; and (f) in eukaryotic ribosomes.

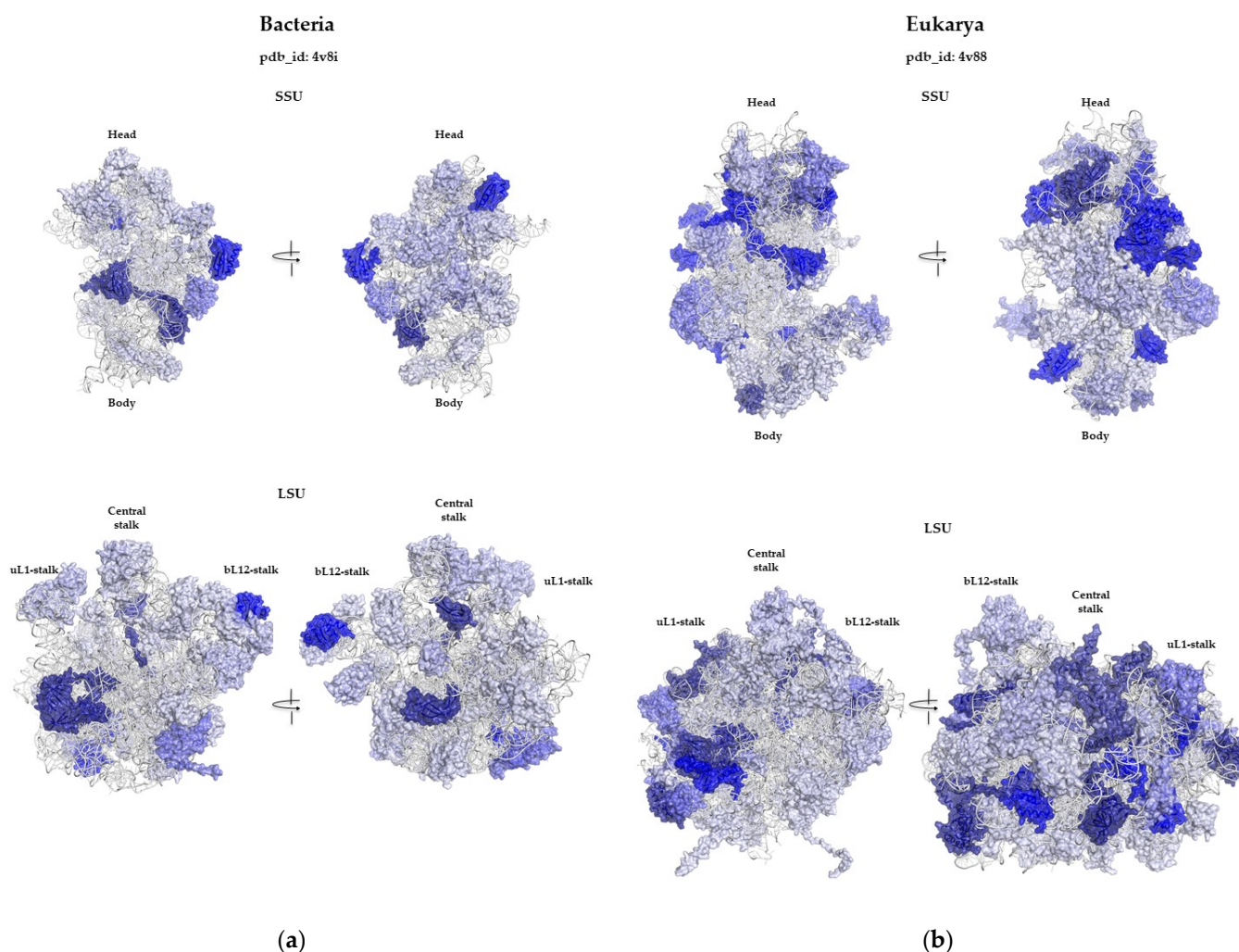


Figure 4. The most prevalent folds colored by their occurrences in the LSU and SSU of (a) bacterial pdb_id: 4v8i and (b) eukaryotic pdb_id: 4v88 ribosomes. The colors correspond to the colors used in Figure 2d–f.

2.2. Ribosomal Protein Folds Specific to Each Kingdom

2.2.1. Minimal Fold Content

We also examined the folds present in the ribosomes of the three kingdoms to determine the minimal fold content (Supplementary Table S1). While most ancestral folds generally correspond to universal ribosomal proteins, certain folds that are superfolds in extant genomes such as single α -helices involved in coiled-coils or other helix-helix interfaces (bL32, eL41, and eS30) and Ribosomal protein L24e; Chain: T (bL28 and eL24) are found in kingdom specific ribosomal proteins across the three kingdoms. However, almost all ribosomal protein folds that later evolved into superfolds in modern genomes are shared and thus part of the minimal fold content of the three kingdoms' ribosomes. Conversely, protein folds specific to each kingdom do not become superfolds in contemporary genomes.

2.2.2. Different Behavior of Kingdom Specific Folds in Bacteria and Archaea/Eukarya

It is important to note that, since the ribosomes of archaea and eukarya have the same ribosomal proteins, with eukaryotic ribosomes containing 12 additional ones, all the ribosomal protein folds that are shared between bacterial and archaeal ribosomes are also shared between bacterial and eukaryotic ribosomes [178].

When examining the protein folds specific to ribosomes in each kingdom, several patterns emerge. In bacterial ribosomes, there are 45 distinct topologies, of which 11 are

unique to bacteria (Figure 5a). In contrast, archaea and eukaryotes share 14 specific folds that are not found in bacterial ribosomes. Additionally, eukaryotic ribosomes possess five protein folds that are unique to them. This analysis, therefore, reveals a new finding: in the three kingdoms, only approximately half of the kingdom-specific ribosomal proteins have kingdom-specific protein folds (Table 2). For example, among the 18 bacteria-specific ribosomal proteins (Figure 5b), only 9 (bL9, bL17, bL20, bL25, bL35, bL36, bS16, bS18, bS20) exhibit bacteria-specific ribosomal protein folds, while the others (bL12, bL19, bL21, bL27, bL28, bL31, bL32, bL34, bS6) possess folds that are already present in universal ribosomal proteins, although there are two in archaeal and eukaryotic ribosomes. Similarly, out of the 31 archaea- and eukarya-specific ribosomal proteins, only 18 exhibit the archaea- and eukaryotic-specific ribosomal protein folds, and out of the 12 eukarya-specific ribosomal proteins, only 5 have folds that are not present in the ribosomes of the other kingdoms.

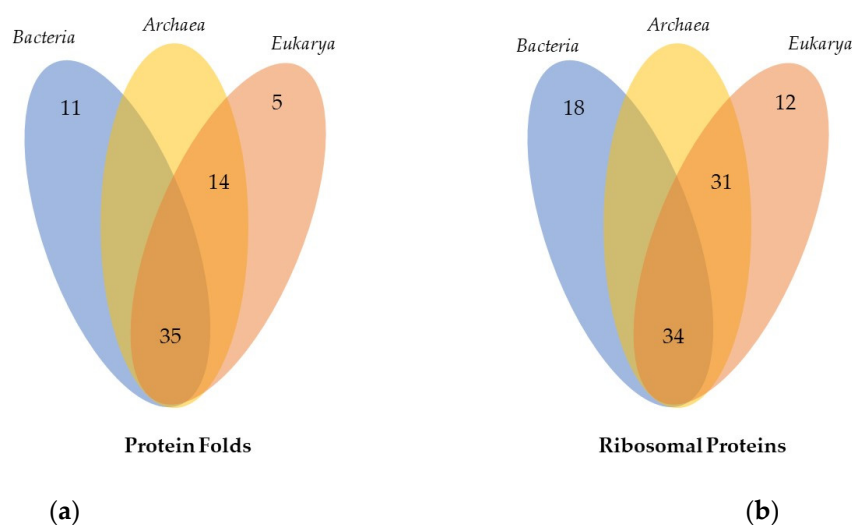


Figure 5. Venn diagrams representing: (a) common and specific ribosomal protein folds for each kingdom; (b) universal and specific ribosomal proteins for each kingdom.

Table 2. Kingdom-specific proteins with kingdom-specific and universal ribosomal protein folds.

	Specific Ribosomal Protein	Specific Ribosomal Protein Fold	Specific Ribosomal Protein	Universal Ribosomal Protein Fold
Bacteria	bL9	3.40.5 RP L9; domain 1 & 3.10.430 RP L9; domain 2	bL12	3.30.1390 RP L30; Chain: A
	bL17	3.90.1030 50s RP L17; Chain A	bL19	2.30.30 SH3 type barrels
	bL20	1.10.1900 c-terminal domain of poly(a) binding domain	bL21	2.40.50 OB fold
	bL25	2.40.240 RP L25; Chain P & 2.170.120 RNAPol α Subunit; Chain A, domain 2	bL27	2.40.50 OB fold
	bL35	4.10.410 Factor Xa Inhibitor	bL28	2.30.170 RP L24e; Chain: T
	bL36	2.60.120 Jelly Rolls	bL31	4.10.830 30s RP S14; Chain N
	bS16	3.30.1320 S16 RP; Chain: A	bL32	1.20.5 Single α -helices involved in coiled-coils or other helix-helix interfaces
	bS18	4.10.640 30s RP S18	bL34	1.10.287 Helix Hairpins
	bS20	1.20.58 Methane Monooxygenase Hydroxylase; Chain G, domain 1	bS6	3.30.70 α - β Plaits
Archaea and Eukarya	eL8	3.30.1330 60s RP L30; Chain: A	eL14	2.30.30 SH3 type barrels
	eL15	3.40.1120 RP L15e	eL18	3.100.10 RP L15; Chain K; domain 2
	eL19	1.10.1650 50s RP L19e, Chain O, domain 1	eL21	2.30.30 SH3 type barrels

Table 2. Cont.

Specific Ribosomal Protein	Specific Ribosomal Protein Fold	Specific Ribosomal Protein	Universal Ribosomal Protein Fold
eL20	3.10.20 UB roll	eL24	2.30.170 RP L24e; Chain: T
eL30	3.30.1330 60s RP L30; Chain: A	eL27	2.30.30 SH3 type barrels
eL31	3.10.440 RP L31e; Chain: W	eL32	3.30.70 α - β Plaits
eL33	2.40.10 Thrombin, subunit H	eL41	1.20.5 Single α -helices involved in coiled-coils or other helix-helix interfaces
eL34	6.20.340 His-Me finger endonuclease fold	eS4 (3 folds)	2.30.30 SH3 type barrels & 2.40.50 OB fold & 3.10.290 Structural Genomics Hypothetical 15.5 Kd Protein In mrcA-pckA Intergenic Region; Chain A
eL37	2.20.25 N-terminal domain TfiIb	eS8	3.10.290 Structural Genomics Hypothetical 15.5 Kd Protein In mrcA-pckA Intergenic Region; Chain A
eL39	1.10.1620 Atp Synthase ϵ Chain; Chain: I	eS19	1.10.10 Arc Repressor Mutant, subunit A
eL40	4.10.1060 ZNF265 like	eS24	3.30.70 α - β Plaits
eL43	2.20.25 N-terminal domain TfiIb	eS28	2.40.50 OB fold
eS1	3.10.20 UB roll & 3.30.479 Tetrahydropterin Synthase; Chain A		
eS6	3.10.20 UB roll		
eS21	3.30.1230 Hypothetical Cytosolic Protein; Chain A		
eS27	2.20.25 N-terminal domain TfiIb		
eS31	6.20.50 N-terminal domain TfiIb		
Eukarya			
eL22	3.30.1360 Gyrase A; domain 2	eL6	2.30.30 SH3 type barrels
eL29	6.10.140 Helix Hairpins	eL13	1.10.10 Arc Repressor Mutant, subunit A
eL38	3.30.720 Signal recognition particle alu RNA binding heterodimer, srp9/1	eL36	1.10.10 Arc Repressor Mutant, subunit A
RACK1	2.130.10 Methylamine Dehydrogenase; Chain H	eS10	1.10.10 Arc Repressor Mutant, subunit A
eS26	3.30.1740 First Zn finger domain of poly(adp-ribose) polymerase-1	eS25	1.10.10 Arc Repressor Mutant, subunit A
		eS30	1.20.5 Single α -helices involved in coiled-coils or other helix-helix interfaces
		Eukarya specific ribosomal protein	Archaea and Eukarya specific ribosomal protein fold
		eS7	2.20.25 N-terminal domain TfiIb
		eS12	3.30.1330 60s RP L30; Chain: A

Interestingly, the topologies unique to bacterial ribosomes each occur only once within their ribosomal structure, indicating a limited reuse of these specific folds. However, in archaeal and eukaryotic ribosomes, some of the unique protein folds are not only repeated multiple times but are also among the most prevalent. For instance, protein folds such as the N-terminal domain of TfiIb, the UB roll, and the 60S ribosomal protein L30; Chain A are frequently observed in both archaeal and eukaryotic ribosomes.

This difference in the occurrence of specific folds highlights the evolutionary divergence between these kingdoms. In bacterial ribosomes, the uniqueness of certain folds (which is also illustrated by the low bacterial protein/fold ratio—see Table 1) may reflect more specialized functions or bacterial-specific evolutionary constraints. Conversely, the repeated use of specific folds in archaeal and eukaryotic ribosomes suggests a greater degree of functional redundancy and versatility. These reused folds are likely essential for the complex regulatory and structural requirements of archaeal and eukaryotic ribosomes, contributing to their evolutionary success and adaptability.

2.3. Imitation Game: Fold Mimicry across Kingdoms

Another interesting point is that, in many situations, folds can interchange to bind to equivalent areas of bacterial and eukaryotic ribosomes. We have systematically selected and compared equivalent and conserved rRNA regions that are bound by different folds in bacterial ribosomes on the one hand, and archaeal or eukaryotic ribosomes on the other. Figure 6 and Table 3 show that, according to the current CATH classification, folds of other categories, in the LSU and SSU, replace six folds and four folds, respectively. This analysis shows, first, that there is a wide variety of ways in which different protein folds can bind to an equivalent rRNA structure. Figure 4 shows, for example, that, in the eS8/bS20 pair, beta-sheets and alpha-helices can be easily interchanged for binding the same rRNA motif and even penetrating deeply into very narrow rRNA crevices. A more detailed analysis shows that the 10 cases listed in Figure 6 differ in the way the folds replace each other to bind to the rRNA.

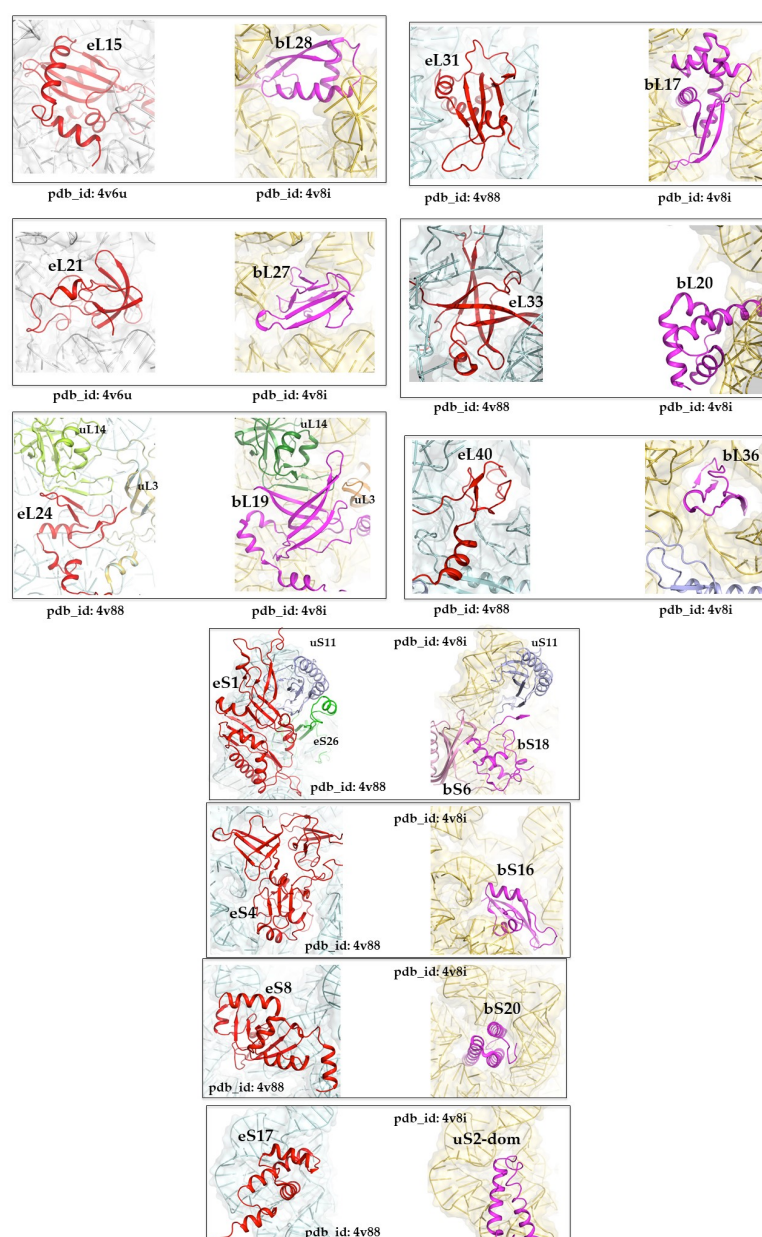


Figure 6. Pairs of different folds that bind equivalent structures in bacterial (pdb_id: 4v8i) and eukaryotic ribosomes (pdb_id: 4v6u and 4v88).

Table 3. Pairs of different proteins and their folds that bind equivalent structures in bacterial (pdb_id: 4v8i) and eukaryotic ribosomes (pdb_id: 4v6u and 4v88).

	Eukarya		Bacteria	
	Protein	CATH	Protein	CATH
LSU	eL15	3.40.1120 RP L15e	bL28	2.30.170 TP L24e; Chain T
	eL21	2.30.30 SH3 type barrels	bL27	2.40.50 OB fold
	eL24	2.30.170 RP L24e; Chain T	bL19	2.30.30 SH3 type barrels
	eL31	3.10.440 RP L31e; Chain W	bL17	3.90.1030 50s RP L17; Chain A
	eL33	2.40.10 Thrombin, subunit H	bL20	1.10.1900 C-terminal domain of poly(a) binding protein
	eL40	4.10.1060 ZNF265 like	bL36	2.60.120 Jelly Rolls
SSU	eS1	3.10.20 UB roll/3.30.479 Tetrahydropterin Synthase; Chain A	bS6/bS18	3.30.70 α - β Plaits/4.10.640 30s RP S18
	eS4	2.30.30 SH3 type barrels	bS16	3.30.1320 S16 RP; Chain A
	eS8	3.10.290 Structural Genomics Hypothetical 15.5Kd Protein In mrcA-pckA Intergenic Region; Chain A	bS20	1.20.58 Methane Monooxygenase Hydroxylase; Chain G, domain 1
	eS17	1.10.60 Diphtheria Toxin Repressor; domain 2	uS2-dom	3.40.50 Rossmann fold

The eL15/bL28 pair is representative of an extreme situation in which two different folds occupy the same RNA cavity (bounded by the H10, H16, H52, and H79 RNA helices), binding in completely different ways. The interaction surfaces and RNA regions contacted by each of the two proteins are totally different. However, the two folds share three equivalent anchor points: (i) Arg 50 (bL28) and Arg 32 (eL15) point to C2200 (H79), (ii) Lys 3 (bL28) and Lys 103 (eL15) point to G1369 (H52), and (iii) Arg 40 (bL28) and Lys 89 (eL15) are oriented towards U2232 (H75).

In the eS8/bS20 pair, the situation also shows two proteins of very different folds and sizes binding to a similar region. While eS8 is much larger than bS20 and interacts with a larger surface area of RNA, both proteins enter the same narrow cavity via two different secondary structure elements, helices.

In the following cases, the different folds seem to precisely target a common RNA interaction surface. eS1, which appears to derive from the duplication of a motif based on three small antiparallel beta-sheets, binds to the same region as the bS6/bS18 pair in the common space they occupy. So, although they have different folds, these proteins will target the same RNA interaction surface in proximity to uS11. eL33 and bL20, which have completely opposite folds in secondary structure content (beta barrel for eL33 and helix bundle for bL20), also concentrate their RNA binding on the same region of helix H41. bS16 occupies the same region as the N-terminal domain of eS4, which appears to have triplicated a similar domain containing a repeat motif of three small anti-parallel beta-sheets. Although presenting different folds, equivalent regions of both proteins interact with the same RNA surface. The pair eL31/bL17 occupies a roughly similar volume and penetrates the same crevasse bordered by helices H47, H61, H 100, and H10, establishing similar interactions despite their different folds. bL17 extends further, however, interacting with helix H57.

In the following three cases, our analysis leads us to believe that pairs whose folds have not been classified as equivalent nevertheless come from a common ancestor. Thus, it seems likely that the pairs eL21/bL27, eL24/bL19, and eL40/bL36 each have a related fold from a common ancestor. Interestingly, the alpha-helix-rich eS17 occupies the same region

as the additional coiled-coil domain of the uS2 protein, and shows surprising convergence by superimposing an almost similar pattern opposite the body/head junction of the SSU. Note, however, that in the eL21/bL27 pair, both of which appear to adopt a small beta barrel, both feature a different-sized N-terminal extension (shorter in eL21) that dips into a common area between LSU helices H38, H81, H85, H86, and H87. The globular domain of eL21, which appears to have rotated relative to bL27, interacts with more or less the same surface of rRNA.

eL24 and bL19 are both located at the interface of the two subunits, and each interacts with the universal protein uL14. Both occupy a remarkably similar space and interact similarly with neighboring rRNA helices H96, H101, and uL14. However, whereas eL24 forms a beta sheet parallel to uL14, bL19 forms an antiparallel sheet. This appears to be due to the relative rotation of the two proteins in the cavity, bringing non-homologous beta-sheets into contact with uL14.

eL40 and bL36 have similar topologies, containing three small antiparallel beta-sheets. They make very similar contacts in a very narrow cavity bounded by LSU helices H42, H89, H91, and H97. We believe they both originate from a common ancestor and have rotated relative to each other through bacterial/archaeal radiation.

Thus, the permutability or mimicry of folds for the recognition of equivalent rRNA regions, and probably similar functions, shows that processes of molecular convergence have already played an important role, right from the radiation of the three kingdoms.

2.4. Fold Evolution/Convergence

Some ribosomal protein folds may have undergone point mutations, insertions, and duplications, leading to the emergence of entirely new folds. Various experimental studies demonstrate that, in certain instances, the genetic routes between completely distinct protein folds can be unexpectedly short [191]. Many proteins exhibit intricate structures composed of multiple layers of secondary elements tightly packed together. These intricate architectures often evolve from simpler, more compact folds. This process involves the elongation of the primary sequence and the subsequent packing of the extended region against the exterior of the ancestral fold [192]. This computational analysis revealed that several ribosomal proteins resemble the minimal motifs of larger protein folds. For example, the protein bL36 exhibits the topology of a Greek key, which is a common elementary motif for different protein folds such as jelly roll and immunoglobulin folds. The repetitive nature of the TIM barrel architecture suggests its evolution from the recombination or duplication of a smaller structure [193]. Analysis revealed that the globular part of eL32 closely matches the minimal structural unit of the TIM barrel, suggesting that eL32 could be its ancestor that underwent recombination or duplication.

As mentioned earlier, the proteins uL1 and uS2 exhibit minimal RLM [190]. The majority of the big ribosomal proteins consist of two or three domains, with only exceptions being uL4 and uL5. In the CATH system [194], both uL4 and uL5 have their own protein folds, RP 14; Chain A and 50s RP L5; Chain A, respectively. However, a closer look at these protein structures reveals some similarities with already existing protein folds. It was already noted that uL4 is homologous to protein tyrosine phosphatase (PTPase) [195]. Remarkably, its hydrophobic core has four parallel beta-strands which show a strong analogy with minimal RLMs, thus grouping it with uL1 and uS2 (Figure 7). In addition to this, the fact that its beta sheets are parallel suggests some similarity with the RLM. This implies that it can be categorized together with uS2 and the small domain of uL1. Similarly, uL5 exhibits a $\beta\alpha\beta\beta\alpha\beta$ motif, with the four β -strands forming an antiparallel β -sheet and the α -helices packed on one side, and can be grouped together with uS6, the first domain of uL1, uL6, L7/L12, bL9, uL22, and uL30 [196,197]. This joins the idea of the existence of unfolds and meta-folds [89,158,159].

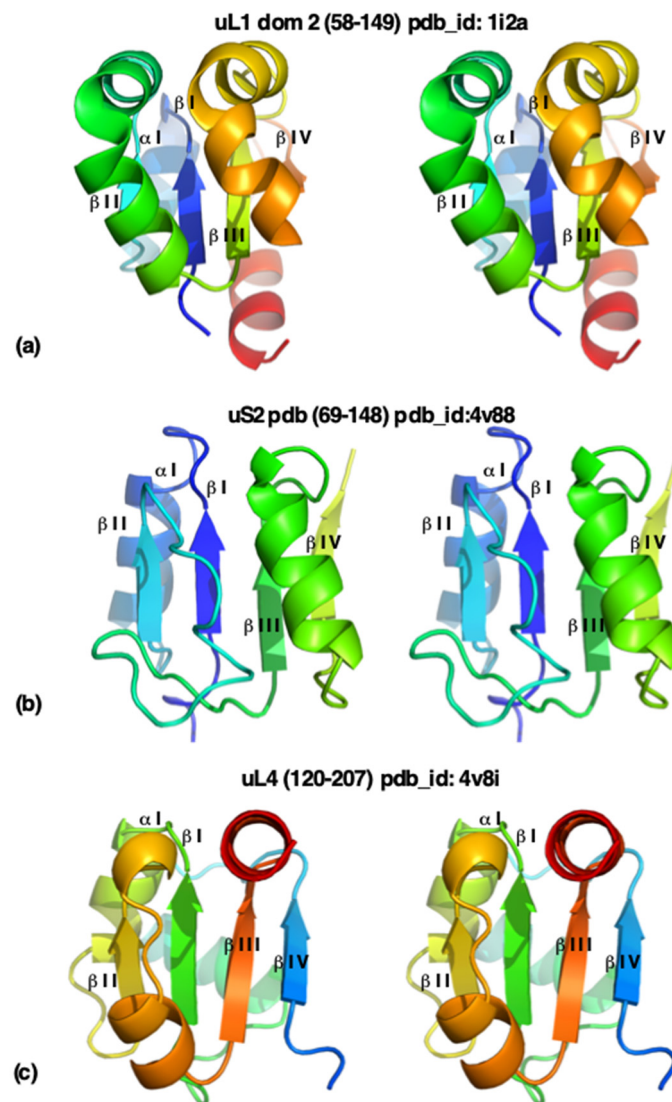


Figure 7. Stereo views of the core of uL1-domain 2 (a), of uS2 (b), and of uL4 (c) colored with the rainbow code in pymol. Similar beta-strands and alpha-helices are indicated.

2.5. The Inexorable Expansion of Fold Extensions

How do ribosomal protein folds acquire extensions? First, when examining the structured extensions of ribosomal proteins, it has been observed that these extensions typically arise by elongating pre-existing secondary structures within the fold. For instance, ribosomal proteins with β -barrel folds, such as the OB fold, often extend two β -sheets to form a β -hairpin. Similarly, ribosomal proteins composed mainly of α helices tend to lengthen the helices at their N- and C-termini to create extensions. Second, flexible loops can form in various locations: between two sheets, between two helices, or between a sheet and a helix. Notably, these extensions develop without disturbing the overall protein fold.

Analyzing the extensions of universal ribosomal proteins across the three kingdoms revealed that the universal extensions are conserved (Figure 8). There are also certain extensions that are absent in bacteria and are present in both archaea and eukarya and vice versa (Supplementary Table S3a–d). For example, the proteins uL3, uL4, uL13, uL18, uL22, uL23, and uL33 in LSU and uS17 in SSU are the most “evolvable” because they have acquired extensions at each evolutionary transition (Figure 9). In contrast, certain proteins such as uL1, uL6, uL10, uL11, and uL15 in the LSU and uS7, uS9, uS10, uS12, uS13, and uS19 do not appear to develop extensions over the course of evolution.

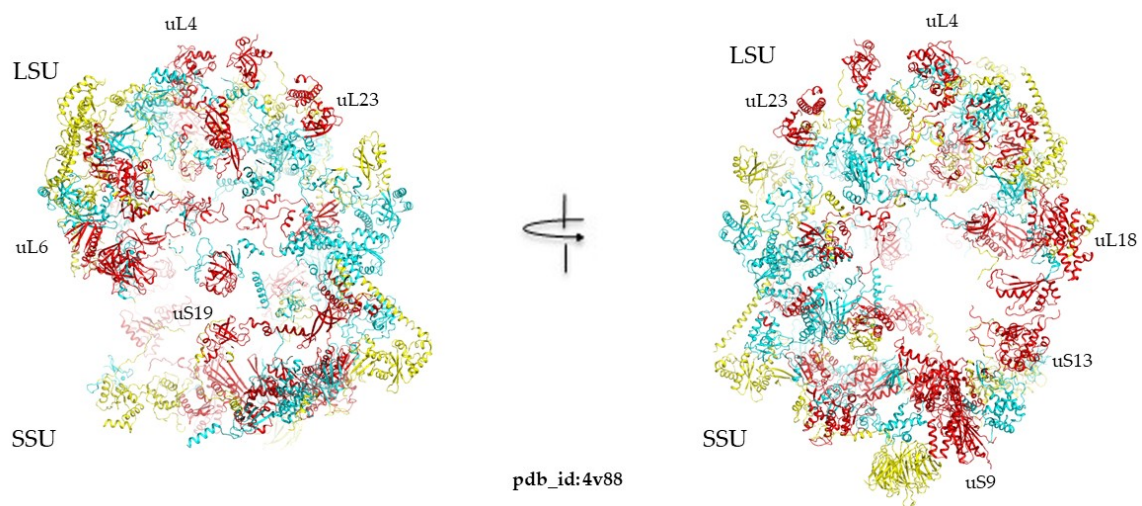


Figure 8. Emergence of extensions: in red: universal proteins and extensions, in blue: proteins and extensions acquired in archaeal and eukaryotic ribosomes, in yellow: proteins and extensions specific to eukaryotic ribosomes p_{db}_id: 4v88. The rRNA is not represented to highlight the proteins.

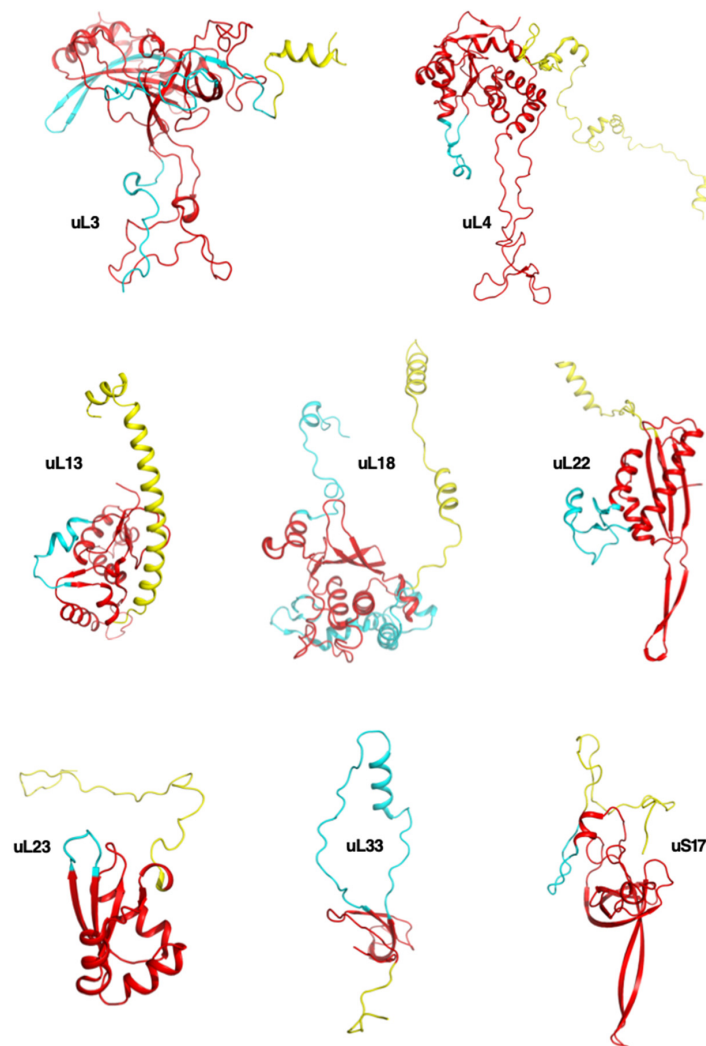


Figure 9. The most evolvable universal ribosomal proteins and their extensions. In red: universal proteins and their universal extensions. In blue: extensions acquired in archaeal ribosomes. In yellow: extensions specific to eukaryotic ribosomes p_{db}_id: 4v88.

Furthermore, a comparison of archaeal and eukaryotic ribosomal proteins showed that eukaryotic ribosomes have acquired 53 new extensions [87]. The acquisition of extensions can also occur within a single kingdom. For instance, ribosomal protein bS16 typically lacks extensions, but in the *Mycobacterium* ribosomes, it features a polypeptide chain ending with a helix at the interaction site. Additionally, structural analysis of ribosomal protein folds in modern genomes identified proteins with the same fold as a lot of the ribosomal proteins, such as uL22 presenting the RP L22; Chain A fold and the ones containing Rossmann fold, Elongation factor Tu, OB, and SH3 domains but without these extensions. More generally, many ribosomal proteins are small β -barrels that tend to be decorated with extensions [89]. These findings do not support the idea that extensions preceded folded proteins [84]. Instead, our analysis indicates that folded proteins emerged first and were later decorated with extensions. This aligns with the broader trend of increasing structural disorder with organism complexity, as the percentage of intrinsically disordered proteins rises from bacteria to archaea to eukaryotes [198,199].

3. Discussion

3.1. What's New from the Past?

This study delves into the intriguing usage of protein folds within ribosomes across different kingdoms—bacteria, archaea, and eukaryotes. This analysis shows that fold usages in the ribosomes are similar to those of modern genomes in that they are also part of what might be called the infancy of a power law, and are strongly biased in favor of superfolds. Thus, similar evolutionary principles and mechanisms shape the organization of ancestral nucleo-protein systems and modern proteomes. One of the key findings is the identification of SRFs, which are highly prevalent within ribosomes and appear to have significant functional roles. The OB fold and SH3-type barrels are notable examples of these SRFs. Interestingly, this study finds that some bacterial SRFs have retained their prominence in modern genomes, suggesting their ancient and essential roles. However, despite similarities in the distribution laws, a notable difference was observed. Whereas the superfolds of modern genomes essentially correspond to folds that can accommodate a large number of catalytic functions such as the TIM-barrel or the Rossmann fold, the SRFs are essentially associated with allosteric functions or RNA recognition. In this respect, it is particularly interesting to note that the proto Rossmann folds already present in the ribosome in the uL1, uS2, and uL4 proteins are not yet catalytic and have not yet achieved the success of their descendants. This fold, which will undergo a vertiginous expansion in modern genomes due to its huge diversity of catalytic functions, probably began its history in the ribosome due to its ability to recognize RNA and its allosteric functions.

3.2. Ancestral Evolutionary Mechanisms with Different Selective Pressure

These differences, therefore, reflect distinct selective pressures on cellular systems and on the ancestral forms of nucleo-protein complexes. Interestingly, a distinct pattern of fold usage is observed in the three kingdom ribosomes, highlighting the evolutionary paths and functions of ribosomal proteins. Figure 8 provides a comprehensive visual summary of the observed patterns, illustrating the distribution of protein folds in the ribosomes of the three kingdoms. Bacterial and eukaryotic ribosomes are clearly distinct in terms of fold usage. The protein-to-fold ratio, which increases from bacteria to eukarya, reflects the evolutionary trend towards more complex and specialized protein functions without a proportional increase in fold diversity. Furthermore, this study highlights the less-differentiated nature of bacterial large subunits (LSU) and small subunits (SSU) compared to those of archaea and eukaryotes. This differentiation in fold content aligns with the broader evolutionary narratives of these kingdoms, where archaea and eukaryotes have developed more specialized and complex ribosomal structures. A notable observation is the preferential use of SH3 domains rather than OB folds in eukaryotic ribosomes. This shift towards a SRF more conducive to allosteric functions probably reflects increased signaling in eukaryotic ribosomes with more complex functions. Transcending the concept

of folds also makes it possible to group a large number of ribosomal proteins into only five categories of urfolds or metafolds. This creates continuity in the fold space of the ribosomal protein, which could attest to their ancestral character and common origins (Table 4). For example, grouping into broader classes such as urfolds on the one hand, and categorizing ribosomal proteins by their folds (Figure 5 and Table 2) on the other, provides a fresh look at specific kingdom proteins. A much smaller number of protein families make up the ribosome, and almost half of all specific-kingdom proteins have common origins.

Table 4. Categories of identified ribosomal urfolds/metafolds and proteins that exhibit them.

Urfolds/Metafolds	Universal ¹	B ²	AE ³	E ⁴
SBBs	uL2, uL3, uL24, uS12, uS17	bL19, bL21, bL27	eL14, eL21, eL24, eL27, eL33, eS4, eS28	eL6
$\beta\alpha\beta\beta\alpha\beta$	uL1 (domain1), uL5, uL6, L7/L12, uL10, uL13, uL16, uL23, uS3, uS5 (C terminal domain), uS8, uS9, uS10	bL9 (C terminal domain), bS6	eL30, eL32, eS24	eL22
RLM	uL1 (domain 2), uL4, uS2			
Perpendicularly stacked α helices	uS13, uS2 (coiled coil domain)	bL20	eS17	
3 parallel α helices	uL29, uS15	bL34, bS20	eL19	

¹ Universal: universal ribosomal proteins. ² B: bacteria-specific ribosomal proteins. ³ AE: archaea- and eukarya-specific ribosomal proteins. ⁴ E: eukarya-specific ribosomal proteins.

3.3. Imitation Games at the Ribosome Age

The observation of the interchangeability of a large number of folds for the recognition of the same ribosomal RNA structures shows that, from the earliest evolutionary stages, evolution already had a large semantic repertoire for performing the same functions. The convergence or mimicry of folds is once again an evolutionary trait found at different evolutionary periods. This shows that evolutionary processes have been surprisingly similar since the first forms of life.

3.4. Order First, Disorder Next?

The demonstration that a large proportion of the ribosomal protein's ancestral folds gradually acquired various categories of extensions during their transitions to the three kingdoms also reveals similar evolutionary mechanisms in ancient nucleo-protein complexes and modern proteomes that are prone to insertion/deletion events. The inexorable expansion of extensions observed throughout ribosome evolution suggests that this trend probably preceded the radiation of the three kingdoms, and thus contributed to the acquisition of universal protein extensions already present in ancestral ribosomes. For example, it is likely that, already in the first evolutionary steps that formed the protoribosomes, the ancestral folds of the uL2, uL3, and uL4 proteins acquired extensions that reach the peptidyl-center or peptide tunnel to enable allostery and signaling between functional centres. These findings support an "order first, disorder next" evolutionary model in which simple but ordered ribosomal proteins preceded the appearance of disordered extensions. This model is also in good agreement with all the data showing that protein disorder increases over the course of evolution and reaches an apex in eukaryotes (Figure 10).

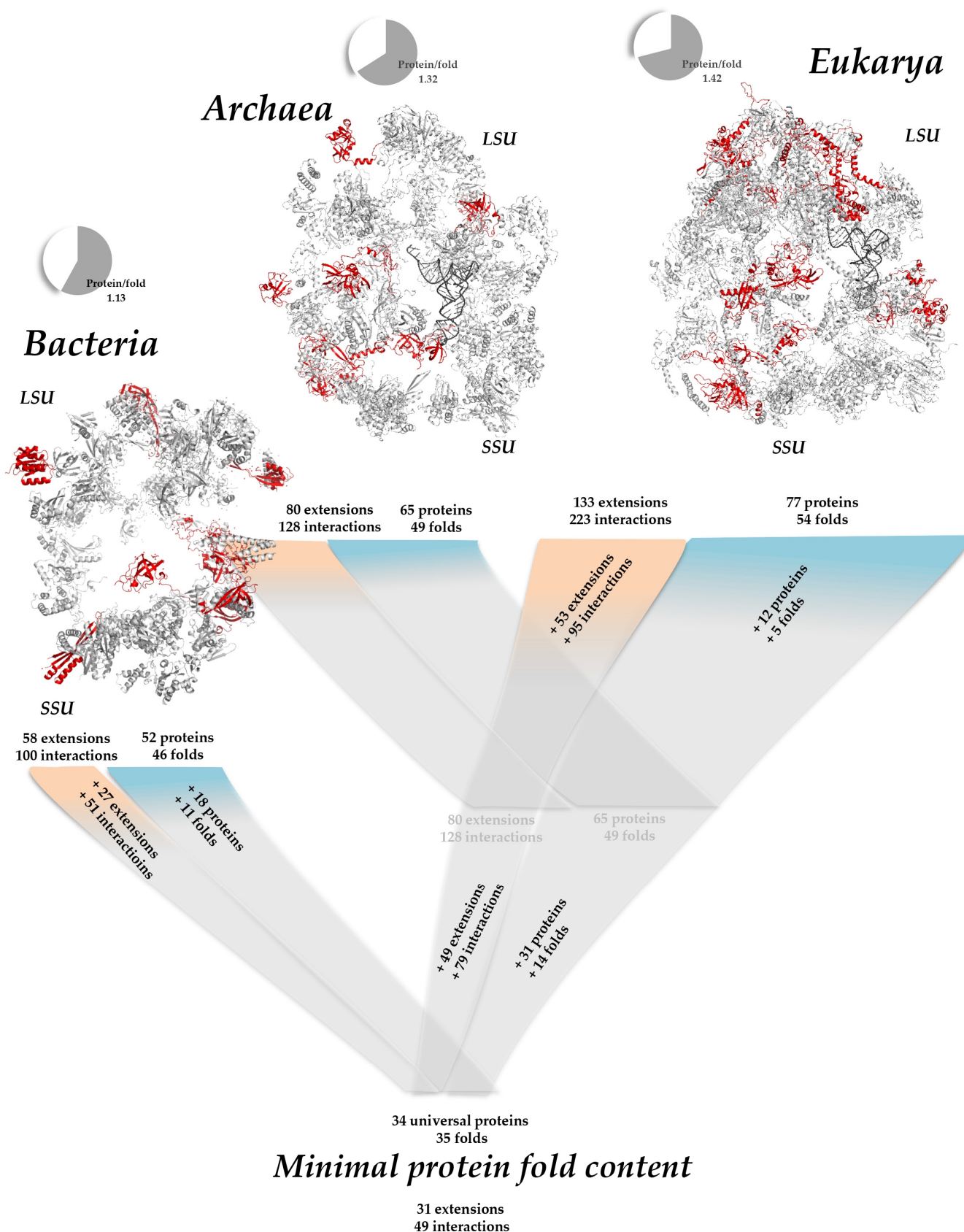


Figure 10. Distribution and evolution of protein folds and their extensions in ribosomes across the three kingdoms [200]. pdb_id: for bacteria 4y4p, for archaea 4v6u, for eukarya 4ug0.

4. Materials and Methods

The PDB [184] served as the primary resource for accessing high-resolution crystallographic and cryo-EM ribosome structures representative of bacteria (4Y4P, 4YBB, 4V9H, 7P7Q, 1DD3), archaea (4V6U), and eukarya (4V88, 3J7R, 4UG0). The ribosomal proteins were considered in the context of the whole ribosomes, and some proteins were isolated when needed. Additionally, it provided structures of proteins possessing known folds, facilitating comparative analyses with ribosomal proteins. To categorize protein folds, the CATH database [194], containing hierarchic classification of protein domain structures [115], was utilized. The classification system comprises four main levels: protein class (C), architecture (A), topology (T), and homologous superfamily (H). We used the fold identification search, by sequence and by structure, of the CATH database by isolating the ribosomal proteins from the overall structure of the ribosomes. For identifying proteins with analogous structures to ribosomal proteins whose folds were not directly identified by CATH, Foldseek [200] and Distance matrix alignment (Dali) [201] servers were employed. When a protein with similar structure was identified, its topology was attributed to the ribosomal protein in question. Visualization and structural alignments were performed using PyMOL 2.5.2 [202].

5. Conclusions

In summary, this comparative analysis of ribosomal protein folds across different kingdoms provides valuable insights into the evolutionary mechanisms and functional adaptations that have shaped modern proteomes. Surprisingly, this study suggests that the evolutionary mechanisms and principles observed in modern genomes, such as the preferential use of folds, convergence, and the acquisition of insertions, were already at work before the radiation of the three kingdoms.

This short paper provides a new look at ribosome evolution by focusing on the protein fold contents and occurrences. The identification of SRFs and their roles underscores the intricate interplay between protein structure and function, offering a deeper understanding of the evolutionary dynamics of ribosomes and their components. These SRFs announce the future power-law distribution observed in extant genomes. OB folds and SH3-type barrels are found to be the most common folds in bacterial and archaeal/eukaryotic ribosomes, respectively. Notably, the Rossmann fold, a prevalent superfold in contemporary proteomes, is not yet dominant in ribosomes. This suggests that different evolutionary constraints other than enzymatic capacities shape the ribosomal protein fold landscape. The observation that different folds can recognize equivalent rRNA structures illustrates an early mechanism of molecular convergence. This comparative analysis deepens our understanding of ribosomal evolution, especially the unique evolutionary trajectories of the large and small subunits, and highlights the intricate interactions between RNA and protein components during the shift from the RNA world to contemporary cellular life. This and our previous works also demonstrate that the gradual acquisition of extensions by simple but ordered folds constitutes an inexorable evolutionary mechanism, which finds its apex in the transition from archaeal ribosomes to eukaryotic ribosomes. This observation supports the idea that simple but structured ribosomal proteins preceded the development of their disordered extensions. Our work also reveals an interesting phenomenon of congruence: the increasing acquisition of extensions involved in the interconnection of functional centres, on the one hand, and the preferential use of SH3 domains conducive to allostery, on the other, show that evolution has made every effort to develop signaling in eukaryotic ribosomes.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25168806/s1>.

Author Contributions: Conceptualization, methodology, formal analysis, investigation, data curation, writing—original draft preparation, writing—review and editing, visualization, I.T. and Y.T.; resources, supervision, project administration, funding acquisition, Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: The project leading to this publication has received funding from the Excellence Initiative of Aix-Marseille Université-A*Midex, a French “Investissements d’Avenir programme” AMX-21-IET-018.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Acknowledgments: We thank Magali Lescot and Maximilian P. Kohl for constant support and helpful discussions. We are also grateful to Christopher Mazzerbo for the technical support.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Newman, M. Power Laws, Pareto Distributions and Zipf’s Law. *Contemp. Phys.* **2005**, *46*, 323–351. [[CrossRef](#)]
2. Caetano-Anollés, G. The Compressed Vocabulary of Microbial Life. *Front. Microbiol.* **2021**, *12*, 655990. [[CrossRef](#)] [[PubMed](#)]
3. Gerstein, M. Patterns of Protein-Fold Usage in Eight Microbial Genomes: A Comprehensive Structural Census. *Proteins* **1998**, *33*, 518–534. [[CrossRef](#)]
4. Abeln, S.; Deane, C.M. Fold Usage on Genomes and Protein Fold Evolution. *Proteins Struct. Funct. Bioinform.* **2005**, *60*, 690–700. [[CrossRef](#)] [[PubMed](#)]
5. Semple, S.; Ferrer-i-Cancho, R.; Gustison, M.L. Linguistic Laws in Biology. *Trends Ecol. Evol.* **2022**, *37*, 53–66. [[CrossRef](#)] [[PubMed](#)]
6. Koonin, E.V. Are There Laws of Genome Evolution? *PLoS Comput. Biol.* **2011**, *7*, e1002173. [[CrossRef](#)] [[PubMed](#)]
7. Sikorav, J. The Utility of Scientific Papers. *Scientometrics* **2005**, *21*, 49–68. [[CrossRef](#)]
8. Ohta, M.; Yamamoto, K. Power-Law Relation and Complexity in the Shape of Chinese Character (Kanji). *J. Phys. Soc. Jpn.* **2019**, *88*, 064803. [[CrossRef](#)]
9. Chierichetti, F.; Kumar, R.; Pang, B. On the Power Laws of Language: Word Frequency Distributions. In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, Tokyo, Japan, 7–11 August 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 385–394.
10. Seržant, I.A.; Moroz, G. Universal Attractors in Language Evolution Provide Evidence for the Kinds of Efficiency Pressures Involved. *Humanit. Soc. Sci. Commun.* **2022**, *9*, 58. [[CrossRef](#)]
11. Piantadosi, S.T. Zipf’s Word Frequency Law in Natural Language: A Critical Review and Future Directions. *Psychon. Bull. Rev.* **2014**, *21*, 1112–1130. [[CrossRef](#)]
12. Breaker, R.R. The Biochemical Landscape of Riboswitch Ligands. *Biochemistry* **2022**, *61*, 137–149. [[CrossRef](#)] [[PubMed](#)]
13. Ikemura, T. Correlation between the Abundance of Escherichia Coli Transfer RNAs and the Occurrence of the Respective Codons in Its Protein Genes: A Proposal for a Synonymous Codon Choice That Is Optimal for the *E. coli* Translational System. *J. Mol. Biol.* **1981**, *151*, 389–409. [[CrossRef](#)] [[PubMed](#)]
14. Plotkin, J.B.; Kudla, G. Synonymous but Not the Same: The Causes and Consequences of Codon Bias. *Nat. Rev. Genet.* **2011**, *12*, 32–42. [[CrossRef](#)]
15. Komar, A.A.; Samatova, E.; Rodnina, M.V. Translation Rates and Protein Folding. *J. Mol. Biol.* **2023**, *436*, 168384. [[CrossRef](#)] [[PubMed](#)]
16. Liu, Y. A Code within the Genetic Code: Codon Usage Regulates Co-Translational Protein Folding. *Cell Commun. Signal.* **2020**, *18*, 145. [[CrossRef](#)] [[PubMed](#)]
17. Orengo, C.A.; Jones, D.T.; Thornton, J.M. Protein Superfamilies and Domain Superfolds. *Nature* **1994**, *372*, 631–634. [[CrossRef](#)] [[PubMed](#)]
18. Qian, J.; Luscombe, N.M.; Gerstein, M. Protein Family and Fold Occurrence in Genomes: Power-Law Behaviour and Evolutionary Model. *J. Mol. Biol.* **2001**, *313*, 673–681. [[CrossRef](#)] [[PubMed](#)]
19. Buchan, D.W.A.; Shepherd, A.J.; Lee, D.; Pearl, F.M.G.; Rison, S.C.G.; Thornton, J.M.; Orengo, C.A. Gene3D: Structural Assignment for Whole Genes and Genomes Using the CATH Domain Structure Database. *Genome Res.* **2002**, *12*, 503–514. [[CrossRef](#)] [[PubMed](#)]
20. Wolf, Y.I.; Brenner, S.E.; Bash, P.A.; Koonin, E.V. Distribution of Protein Folds in the Three Superkingdoms of Life. *Genome Res.* **1999**, *9*, 17–26. [[CrossRef](#)]
21. Hegyi, H.; Lin, J.; Greenbaum, D.; Gerstein, M. Structural Genomics Analysis: Characteristics of Atypical, Common, and Horizontally Transferred Folds. *Proteins* **2002**, *47*, 126–141. [[CrossRef](#)]
22. Romei, M.; Carpentier, M.; Chomilier, J.; Lecointre, G. Origins and Functional Significance of Eukaryotic Protein Folds. *J. Mol. Evol.* **2023**, *91*, 854–864. [[CrossRef](#)]
23. Romei, M.; Sapriel, G.; Imbert, P.; Jamay, T.; Chomilier, J.; Lecointre, G.; Carpentier, M. Protein Folds as Synapomorphies of the Tree of Life. *Evol. Int. J. Org. Evol.* **2022**, *76*, 1706–1719. [[CrossRef](#)] [[PubMed](#)]
24. Kauko, A.; Lehto, K. Eukaryote Specific Folds: Part of the Whole. *Proteins* **2018**, *86*, 868–881. [[CrossRef](#)]
25. Ponting, C.P.; Russell, R.R. The Natural History of Protein Domains. *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *31*, 45–71. [[CrossRef](#)]
26. Maggi, L.; Carloni, P.; Rossetti, G. Vibrational Energy in Proteins Correlates with Topology. *J. Phys. Chem. Lett.* **2018**, *9*, 6393–6398. [[CrossRef](#)]

27. Xie, J.; Lai, L. Protein Topology and Allostery. *Curr. Opin. Struct. Biol.* **2020**, *62*, 158–165. [[CrossRef](#)] [[PubMed](#)]
28. Schaeffer, R.D.; Daggett, V. Protein Folds and Protein Folding. *Protein Eng. Des. Sel.* **2011**, *24*, 11–19. [[CrossRef](#)]
29. Henzler-Wildman, K.; Kern, D. Dynamic Personalities of Proteins. *Nature* **2007**, *450*, 964–972. [[CrossRef](#)] [[PubMed](#)]
30. Abrusán, G.; Marsh, J.A. Alpha Helices Are More Robust to Mutations than Beta Strands. *PLOS Comput. Biol.* **2016**, *12*, e1005242. [[CrossRef](#)]
31. Tóth-Petróczy, Á.; Tawfik, D.S. The Robustness and Innovability of Protein Folds. *Curr. Opin. Struct. Biol.* **2014**, *26*, 131–138. [[CrossRef](#)]
32. Bukhari, S.A.; Caetano-Anollés, G. Origin and Evolution of Protein Fold Designs Inferred from Phylogenomic Analysis of CATH Domain Structures in Proteomes. *PLoS Comput. Biol.* **2013**, *9*, e1003009. [[CrossRef](#)] [[PubMed](#)]
33. Laurino, P.; Tóth-Petróczy, Á.; Meana-Pañeda, R.; Lin, W.; Truhlar, D.G.; Tawfik, D.S. An Ancient Fingerprint Indicates the Common Ancestry of Rossmann-Fold Enzymes Utilizing Different Ribose-Based Cofactors. *PLOS Biol.* **2016**, *14*, e1002396. [[CrossRef](#)] [[PubMed](#)]
34. Kim, K.M.; Caetano-Anollés, G. The Evolutionary History of Protein Fold Families and Proteomes Confirms That the Archaeal Ancestor Is More Ancient than the Ancestors of Other Superkingdoms. *BMC Evol. Biol.* **2012**, *12*, 13. [[CrossRef](#)] [[PubMed](#)]
35. Ranea, J.A.G.; Buchan, D.W.A.; Thornton, J.M.; Orengo, C.A. Evolution of Protein Superfamilies and Bacterial Genome Size. *J. Mol. Biol.* **2004**, *336*, 871–887. [[CrossRef](#)] [[PubMed](#)]
36. Winstanley, H.F.; Abeln, S.; Deane, C.M. How Old Is Your Fold? *Bioinformatics* **2005**, *21*, i449–i458. [[CrossRef](#)] [[PubMed](#)]
37. Ramakrishnan, V. Ribosome Structure and the Mechanism of Translation. *Cell* **2002**, *108*, 557–572. [[CrossRef](#)] [[PubMed](#)]
38. Melnikov, S.; Ben-Shem, A.; Garreau De Loubresse, N.; Jenner, L.; Yusupova, G.; Yusupov, M. One Core, Two Shells: Bacterial and Eukaryotic Ribosomes. *Nat. Struct. Mol. Biol.* **2012**, *19*, 560–567. [[CrossRef](#)] [[PubMed](#)]
39. Bashan, A.; Agmon, I.; Zarivach, R.; Schluenzen, F.; Harms, J.; Berisio, R.; Bartels, H.; Franceschi, F.; Auerbach, T.; Hansen, H.A.S.; et al. Structural Basis of the Ribosomal Machinery for Peptide Bond Formation, Translocation, and Nascent Chain Progression. *Mol. Cell* **2003**, *11*, 91–102. [[CrossRef](#)]
40. Klinge, S.; Voigts-Hoffmann, F.; Leibundgut, M.; Ban, N. Atomic Structures of the Eukaryotic Ribosome. *Trends Biochem. Sci.* **2012**, *37*, 189–198. [[CrossRef](#)]
41. Wilson, D.N.; Doudna Cate, J.H. The Structure and Function of the Eukaryotic Ribosome. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, a011536. [[CrossRef](#)]
42. Armache, J.-P.; Anger, A.M.; Márquez, V.; Franckenberg, S.; Fröhlich, T.; Villa, E.; Berninghausen, O.; Thomm, M.; Arnold, G.J.; Beckmann, R.; et al. Promiscuous Behaviour of Archaeal Ribosomal Proteins: Implications for Eukaryotic Ribosome Evolution. *Nucleic Acids Res.* **2013**, *41*, 1284–1293. [[CrossRef](#)] [[PubMed](#)]
43. Ban, N.; Nissen, P.; Hansen, J.; Moore, P.B.; Steitz, T.A. The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution. *Science* **2000**, *289*, 905–920. [[CrossRef](#)] [[PubMed](#)]
44. Tirumalai, M.R.; Rivas, M.; Tran, Q.; Fox, G.E. The Peptidyl Transferase Center: A Window to the Past. *Microbiol. Mol. Biol. Rev.* **2021**, *85*, e00104-21. [[CrossRef](#)]
45. Agmon, I. The Dimeric Proto-Ribosome: Structural Details and Possible Implications on the Origin of Life. *Int. J. Mol. Sci.* **2009**, *10*, 2921–2934. [[CrossRef](#)]
46. Agmon, I. Three Biopolymers and Origin of Life Scenarios. *Life* **2024**, *14*, 277. [[CrossRef](#)]
47. Belousoff, M.J.; Davidovich, C.; Zimmerman, E.; Caspi, Y.; Wekselman, I.; Rozenszajn, L.; Shapira, T.; Sade-Falk, O.; Taha, L.; Bashan, A.; et al. Ancient Machinery Embedded in the Contemporary Ribosome. *Biochem. Soc. Trans.* **2010**, *38*, 422–427. [[CrossRef](#)] [[PubMed](#)]
48. Opron, K.; Burton, Z.F. Ribosome Structure, Function, and Early Evolution. *Int. J. Mol. Sci.* **2018**, *20*, 40. [[CrossRef](#)] [[PubMed](#)]
49. Root-Bernstein, M.; Root-Bernstein, R. The Ribosome as a Missing Link in the Evolution of Life. *J. Theor. Biol.* **2015**, *367*, 130–158. [[CrossRef](#)]
50. Root-Bernstein, R.; Root-Bernstein, M. The Ribosome as a Missing Link in Prebiotic Evolution II: Ribosomes Encode Ribosomal Proteins That Bind to Common Regions of Their Own mRNAs and rRNAs. *J. Theor. Biol.* **2016**, *397*, 115–127. [[CrossRef](#)]
51. Root-Bernstein, R.; Root-Bernstein, M. The Ribosome as a Missing Link in Prebiotic Evolution III: Over-Representation of tRNA- and rRNA-like Sequences and Plieofunctionality of Ribosome-Related Molecules Argues for the Evolution of Primitive Genomes from Ribosomal RNA Modules. *Int. J. Mol. Sci.* **2019**, *20*, 140. [[CrossRef](#)]
52. Fox, G.E. Origin and Evolution of the Ribosome. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a003483. [[CrossRef](#)]
53. Rodgers, M.L.; Woodson, S.A. A Roadmap for rRNA Folding and Assembly during Transcription. *Trends Biochem. Sci.* **2021**, *46*, 889–901. [[CrossRef](#)] [[PubMed](#)]
54. Davis, J.H.; Williamson, J.R. Structure and Dynamics of Bacterial Ribosome Biogenesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2017**, *372*, 20160181. [[CrossRef](#)]
55. Steitz, T.A. A Structural Understanding of the Dynamic Ribosome Machine. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 242–253. [[CrossRef](#)] [[PubMed](#)]
56. Noller, H.F.; Lancaster, L.; Mohan, S.; Zhou, J. Ribosome Structural Dynamics in Translocation: Yet Another Functional Role for Ribosomal RNA. *Q. Rev. Biophys.* **2017**, *50*, e12. [[CrossRef](#)]
57. Noller, H.F. Evolution of Protein Synthesis from an RNA World. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, a003681. [[CrossRef](#)]

58. Bose, T.; Fridkin, G.; Davidovich, C.; Krupkin, M.; Dinger, N.; Falkovich, A.H.; Peleg, Y.; Agmon, I.; Bashan, A.; Yonath, A. Origin of Life: Protoribosome Forms Peptide Bonds and Links RNA and Protein Dominated Worlds. *Nucleic Acids Res.* **2022**, *50*, 1815–1828. [[CrossRef](#)] [[PubMed](#)]
59. Wolf, Y.I.; Koonin, E.V. On the Origin of the Translation System and the Genetic Code in the RNA World by Means of Natural Selection, Exaptation, and Subfunctionalization. *Biol. Direct* **2007**, *2*, 14. [[CrossRef](#)]
60. Grosjean, H.; Westhof, E. An Integrated, Structure- and Energy-Based View of the Genetic Code. *Nucleic Acids Res.* **2016**, *44*, 8020–8040. [[CrossRef](#)]
61. Hartman, H.; Smith, T. The Evolution of the Ribosome and the Genetic Code. *Life* **2014**, *4*, 227–249. [[CrossRef](#)]
62. Koonin, E.V.; Novozhilov, A.S. Origin and Evolution of the Universal Genetic Code. *Annu. Rev. Genet.* **2017**, *51*, 45–62. [[CrossRef](#)] [[PubMed](#)]
63. Lescoute, A.; Westhof, E. Topology of Three-Way Junctions in Folded RNAs. *RNA* **2006**, *12*, 83–93. [[CrossRef](#)]
64. Cruz, J.A.; Westhof, E. Sequence-Based Identification of 3D Structural Modules in RNA with RMDetect. *Nat. Methods* **2011**, *8*, 513–521. [[CrossRef](#)] [[PubMed](#)]
65. Vitas, M.; Dobovišek, A. In the Beginning Was a Mutualism—On the Origin of Translation. *Orig. Life Evol. Biosph.* **2018**, *48*, 223–243. [[CrossRef](#)] [[PubMed](#)]
66. Forterre, P. The Universal Tree of Life: An Update. *Front. Microbiol.* **2015**, *6*, 717. [[CrossRef](#)] [[PubMed](#)]
67. Petrov, A.S.; Gulen, B.; Norris, A.M.; Kovacs, N.A.; Bernier, C.R.; Lanier, K.A.; Fox, G.E.; Harvey, S.C.; Wartell, R.M.; Hud, N.V.; et al. History of the Ribosome and the Origin of Translation. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 15396–15401. [[CrossRef](#)] [[PubMed](#)]
68. Caetano-Anollés, G.; Caetano-Anollés, D. Computing the Origin and Evolution of the Ribosome from Its Structure—Uncovering Processes of Macromolecular Accretion Benefiting Synthetic Biology. *Comput. Struct. Biotechnol. J.* **2015**, *13*, 427–447. [[CrossRef](#)] [[PubMed](#)]
69. Harish, A.; Caetano-Anollés, G. Ribosomal History Reveals Origins of Modern Protein Synthesis. *PLoS ONE* **2012**, *7*, e32776. [[CrossRef](#)] [[PubMed](#)]
70. Wilson, D.N.; Nierhaus, K.H. Ribosomal Proteins in the Spotlight. *Crit. Rev. Biochem. Mol. Biol.* **2005**, *40*, 243–267. [[CrossRef](#)]
71. Graifer, D.; Karpova, G. Roles of Ribosomal Proteins in the Functioning of Translational Machinery of Eukaryotes. *Biochimie* **2015**, *109*, 1–17. [[CrossRef](#)]
72. Brodersen, D.E.; Clemons, W.M.; Carter, A.P.; Wimberly, B.T.; Ramakrishnan, V. Crystal Structure of the 30 s Ribosomal Subunit from *Thermus thermophilus*: Structure of the Proteins and Their Interactions with 16 s RNA1. *J. Mol. Biol.* **2002**, *316*, 725–768. [[CrossRef](#)] [[PubMed](#)]
73. Klein, D.J.; Moore, P.B.; Steitz, T.A. The Roles of Ribosomal Proteins in the Structure Assembly, and Evolution of the Large Ribosomal Subunit. *J. Mol. Biol.* **2004**, *340*, 141–177. [[CrossRef](#)] [[PubMed](#)]
74. Peng, Z.; Oldfield, C.J.; Xue, B.; Mizianty, M.J.; Dunker, A.K.; Kurgan, L.; Uversky, V.N. A Creature with a Hundred Waggly Tails: Intrinsically Disordered Proteins in the Ribosome. *Cell. Mol. Life Sci.* **2014**, *71*, 1477–1504. [[CrossRef](#)]
75. Timsit, Y.; Acosta, Z.; Allemand, F.; Chiaruttini, C.; Springer, M. The Role of Disordered Ribosomal Protein Extensions in the Early Steps of Eubacterial 50 S Ribosomal Subunit Assembly. *Int. J. Mol. Sci.* **2009**, *10*, 817–834. [[CrossRef](#)]
76. Calidas, D.; Lyon, H.; Culver, G.M. The N-Terminal Extension of S12 Influences Small Ribosomal Subunit Assembly in *Escherichia coli*. *RNA* **2014**, *20*, 321–330. [[CrossRef](#)] [[PubMed](#)]
77. Tutuncuoglu, B.; Jakovljevic, J.; Wu, S.; Gao, N.; Woolford, J.L. The N-Terminal Extension of Yeast Ribosomal Protein L8 Is Involved in Two Major Remodeling Events during Late Nuclear Stages of 60S Ribosomal Subunit Assembly. *RNA* **2016**, *22*, 1386–1399. [[CrossRef](#)] [[PubMed](#)]
78. Lawrence, M.G.; Shamsuzzaman, M.; Kondopaka, M.; Pascual, C.; Zengel, J.M.; Lindahl, L. The Extended Loops of Ribosomal Proteins uL4 and uL22 of *Escherichia coli* Contribute to Ribosome Assembly and Protein Translation. *Nucleic Acids Res.* **2016**, *44*, 5798–5810. [[CrossRef](#)]
79. Rhodin, M.H.J.; Dinman, J.D. A Flexible Loop in Yeast Ribosomal Protein L11 Coordinates P-Site tRNA Binding. *Nucleic Acids Res.* **2010**, *38*, 8377–8389. [[CrossRef](#)] [[PubMed](#)]
80. Yutin, N.; Puigbò, P.; Koonin, E.V.; Wolf, Y.I. Phylogenomics of Prokaryotic Ribosomal Proteins. *PLoS ONE* **2012**, *7*, e36972. [[CrossRef](#)]
81. Melnikov, S.; Manakongtreecheep, K.; Söll, D. Revising the Structural Diversity of Ribosomal Proteins across the Three Domains of Life. *Mol. Biol. Evol.* **2018**, *35*, 1588–1598. [[CrossRef](#)]
82. Alva, V.; Söding, J.; Lupas, A.N. A Vocabulary of Ancient Peptides at the Origin of Folded Proteins. *eLife* **2015**, *4*, e09410. [[CrossRef](#)] [[PubMed](#)]
83. Lupas, A.N.; Alva, V. Ribosomal Proteins as Documents of the Transition from Unstructured (Poly)Peptides to Folded Proteins. *J. Struct. Biol.* **2017**, *198*, 74–81. [[CrossRef](#)] [[PubMed](#)]
84. Kovacs, N.A.; Petrov, A.S.; Lanier, K.A.; Williams, L.D. Frozen in Time: The History of Proteins. *Mol. Biol. Evol.* **2017**, *34*, 1252–1260. [[CrossRef](#)] [[PubMed](#)]
85. Poirot, O.; Timsit, Y. Neuron-like Networks between Ribosomal Proteins within the Ribosome. *Sci. Rep.* **2016**, *6*, 26485. [[CrossRef](#)] [[PubMed](#)]

86. Timsit, Y.; Bennequin, D. Nervous-like Circuits in the Ribosome Facts, Hypotheses and Perspectives. *Int. J. Mol. Sci.* **2019**, *20*, 2911. [[CrossRef](#)] [[PubMed](#)]
87. Timsit, Y.; Sergeant-Perthuis, G.; Bennequin, D. Evolution of Ribosomal Protein Network Architectures. *Sci. Rep.* **2021**, *11*, 625. [[CrossRef](#)] [[PubMed](#)]
88. Timsit, Y.; Grégoire, S.-P. Towards the Idea of Molecular Brains. *Int. J. Mol. Sci.* **2021**, *22*, 11868. [[CrossRef](#)] [[PubMed](#)]
89. Youkharibache, P.; Veretnik, S.; Li, Q.; Stanek, K.A.; Mura, C.; Bourne, P.E. The Small β -Barrel Domain: A Survey-Based Structural Analysis. *Structure* **2019**, *27*, 6–26. [[CrossRef](#)] [[PubMed](#)]
90. Ajawatanawong, P.; Baldauf, S.L. Evolution of Protein Indels in Plants, Animals and Fungi. *BMC Evol. Biol.* **2013**, *13*, 140. [[CrossRef](#)]
91. Wolf, Y.; Madej, T.; Babenko, V.; Shoemaker, B.; Panchenko, A.R. Long-Term Trends in Evolution of Indels in Protein Sequences. *BMC Evol. Biol.* **2007**, *7*, 19. [[CrossRef](#)]
92. Deryusheva, E.I.; Selivanova, O.M.; Serdyuk, I.N. Loops and Repeats in Proteins as Footprints of Molecular Evolution. *Biochemistry* **2012**, *77*, 1487–1499. [[CrossRef](#)] [[PubMed](#)]
93. Tóth-Petróczy, Á.; Tawfik, D.S. Protein Insertions and Deletions Enabled by Neutral Roaming in Sequence Space. *Mol. Biol. Evol.* **2013**, *30*, 761–771. [[CrossRef](#)] [[PubMed](#)]
94. Light, S.; Sagit, R.; Sachenkova, O.; Ekman, D.; Elofsson, A. Protein Expansion Is Primarily Due to Indels in Intrinsically Disordered Regions. *Mol. Biol. Evol.* **2013**, *30*, 2645–2653. [[CrossRef](#)] [[PubMed](#)]
95. Wilson, J.R. “To Be, or Not to Be”: Shakespeare against Philosophy. *Shakespeare* **2018**, *14*, 341–359. [[CrossRef](#)]
96. Sikosek, T.; Chan, H.S. Biophysics of Protein Evolution and Evolutionary Protein Biophysics. *J. R. Soc. Interface* **2014**, *11*, 20140419. [[CrossRef](#)] [[PubMed](#)]
97. Valas, R.E.; Yang, S.; Bourne, P.E. Nothing about Protein Structure Classification Makes Sense except in the Light of Evolution. *Curr. Opin. Struct. Biol.* **2009**, *19*, 329–334. [[CrossRef](#)] [[PubMed](#)]
98. Nassar, R.; Dignon, G.L.; Razban, R.M.; Dill, K.A. The Protein Folding Problem: The Role of Theory. *J. Mol. Biol.* **2021**, *433*, 167126. [[CrossRef](#)] [[PubMed](#)]
99. Onuchic, J.N.; Wolynes, P.G. Theory of Protein Folding. *Curr. Opin. Struct. Biol.* **2004**, *14*, 70–75. [[CrossRef](#)] [[PubMed](#)]
100. Mirny, L.; Shakhnovich, E. Protein Folding Theory: From Lattice to All-Atom Models. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 361–396. [[CrossRef](#)]
101. Rossmann, M.G.; Argos, P. Protein Folding. *Annu. Rev. Biochem.* **1981**, *50*, 497–532. [[CrossRef](#)]
102. Lindorff-Larsen, K.; Røgen, P.; Paci, E.; Vendruscolo, M.; Dobson, C.M. Protein Folding and the Organization of the Protein Topology Universe. *Trends Biochem. Sci.* **2005**, *30*, 13–19. [[CrossRef](#)] [[PubMed](#)]
103. Dobson, C.M. Protein Folding and Misfolding. *Nature* **2003**, *426*, 884–890. [[CrossRef](#)] [[PubMed](#)]
104. Fersht, A.R.; Daggett, V. Protein Folding and Unfolding at Atomic Resolution. *Cell* **2002**, *108*, 573–582. [[CrossRef](#)] [[PubMed](#)]
105. Mirny, L.A.; Shakhnovich, E.I. Universally Conserved Positions in Protein Folds: Reading Evolutionary Signals about Stability, Folding Kinetics and Function. *J. Mol. Biol.* **1999**, *291*, 177–196. [[CrossRef](#)] [[PubMed](#)]
106. Gruebele, M. Protein Folding: The Free Energy Surface. *Curr. Opin. Struct. Biol.* **2002**, *12*, 161–168. [[CrossRef](#)] [[PubMed](#)]
107. Sorokina, I.; Mushegian, A.R.; Koonin, E.V. Is Protein Folding a Thermodynamically Unfavorable, Active, Energy-Dependent Process? *Int. J. Mol. Sci.* **2022**, *23*, 521. [[CrossRef](#)]
108. Agard, D.A. To Fold or Not to Fold. *Science* **1993**, *260*, 1903–1904. [[CrossRef](#)] [[PubMed](#)]
109. Hartl, F.U.; Bracher, A.; Hayer-Hartl, M. Molecular Chaperones in Protein Folding and Proteostasis. *Nature* **2011**, *475*, 324–332. [[CrossRef](#)] [[PubMed](#)]
110. Rothman, J.E.; Schekman, R. Molecular Mechanism of Protein Folding in the Cell. *Cell* **2011**, *146*, 851–854. [[CrossRef](#)]
111. Levitt, M.; Chothia, C. Structural Patterns in Globular Proteins. *Nature* **1976**, *261*, 552–558. [[CrossRef](#)]
112. Chothia, C. Principles That Determine the Structure of Proteins. *Annu. Rev. Biochem.* **1984**, *53*, 537–572. [[CrossRef](#)] [[PubMed](#)]
113. Chothia, C.; Finkelstein, A.V. The Classification and Origins of Protein Folding Patterns. *Annu. Rev. Biochem.* **1990**, *59*, 1007–1035. [[CrossRef](#)] [[PubMed](#)]
114. Murzin, A.G.; Brenner, S.E.; Hubbard, T.; Chothia, C. SCOP: A Structural Classification of Proteins Database for the Investigation of Sequences and Structures. *J. Mol. Biol.* **1995**, *247*, 536–540. [[CrossRef](#)]
115. Orengo, C.; Michie, A.; Jones, S.; Jones, D.; Swindells, M.; Thornton, J. CATH—A Hierarchic Classification of Protein Domain Structures. *Structure* **1997**, *5*, 1093–1109. [[CrossRef](#)] [[PubMed](#)]
116. Uversky, V.N. Dancing Protein Clouds: The Strange Biology and Chaotic Physics of Intrinsically Disordered Proteins. *J. Biol. Chem.* **2016**, *291*, 6681–6688. [[CrossRef](#)] [[PubMed](#)]
117. Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V.N. Introducing Protein Intrinsic Disorder. *Chem. Rev.* **2014**, *114*, 6561–6588. [[CrossRef](#)]
118. Towse, C.-L.; Daggett, V. When a Domain Is Not a Domain, and Why It Is Important to Properly Filter Proteins in Databases: Conflicting Definitions and Fold Classification Systems for Structural Domains Make Filtering of Such Databases Imperative. *BioEssays* **2012**, *34*, 1060–1069. [[CrossRef](#)] [[PubMed](#)]
119. Abel, K.; Yoder, M.D.; Hilgenfeld, R.; Jurnak, F. An Alpha to Beta Conformational Switch in EF-Tu. *Structure* **1996**, *4*, 1153–1159. [[CrossRef](#)]
120. Polekhina, G.; Thirup, S.; Kjeldgaard, M.; Nissen, P.; Lippmann, C.; Nyborg, J. Helix Unwinding in the Effector Region of Elongation Factor EF-Tu-GDP. *Structure* **1996**, *4*, 1141–1151. [[CrossRef](#)]

121. Bhattacharjee, N.; Biswas, P. Statistical Analysis and Molecular Dynamics Simulations of Ambivalent α -Helices. *BMC Bioinform.* **2010**, *11*, 519. [[CrossRef](#)]
122. Guo, J.-T.; Jaromczyk, J.W.; Xu, Y. Analysis of Chameleon Sequences and Their Implications in Biological Processes. *Proteins* **2007**, *67*, 548–558. [[CrossRef](#)] [[PubMed](#)]
123. Kuznetsov, I.B.; Rackovsky, S. On the Properties and Sequence Context of Structurally Ambivalent Fragments in Proteins. *Protein Sci. Publ. Protein Soc.* **2003**, *12*, 2420–2433. [[CrossRef](#)] [[PubMed](#)]
124. Young, M.; Kirshenbaum, K.; Dill, K.A.; Highsmith, S. Predicting Conformational Switches in Proteins. *Protein Sci. Publ. Protein Soc.* **1999**, *8*, 1752–1764. [[CrossRef](#)]
125. Pauwels, K.; Lebrun, P.; Tompa, P. To Be Disordered or Not to Be Disordered: Is That Still a Question for Proteins in the Cell? *Cell. Mol. Life Sci.* **2017**, *74*, 3185–3204. [[CrossRef](#)] [[PubMed](#)]
126. Ward, J.J.; Sodhi, J.S.; McGuffin, L.J.; Buxton, B.F.; Jones, D.T. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *J. Mol. Biol.* **2004**, *337*, 635–645. [[CrossRef](#)] [[PubMed](#)]
127. Basile, W.; Salvatore, M.; Bassot, C.; Elofsson, A. Why Do Eukaryotic Proteins Contain More Intrinsically Disordered Regions? *PLOS Comput. Biol.* **2019**, *15*, e1007186. [[CrossRef](#)] [[PubMed](#)]
128. Lobanov, M.Y.; Galzitskaya, O.V. Occurrence of Disordered Patterns and Homorepeats in Eukaryotic and Bacterial Proteomes. *Mol. BioSyst.* **2012**, *8*, 327–337. [[CrossRef](#)] [[PubMed](#)]
129. Ahrens, J.B.; Nunez-Castilla, J.; Siltberg-Liberles, J. Evolution of Intrinsic Disorder in Eukaryotic Proteins. *Cell. Mol. Life Sci.* **2017**, *74*, 3163–3174. [[CrossRef](#)] [[PubMed](#)]
130. Lermyte, F. Roles, Characteristics, and Analysis of Intrinsically Disordered Proteins: A Minireview. *Life* **2020**, *10*, 320. [[CrossRef](#)]
131. Fuxreiter, M. Electrostatics Tunes Protein Interactions to Context. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2209201119. [[CrossRef](#)]
132. Timsit, Y.; Allemand, F.; Chiaruttini, C.; Springer, M. Coexistence of Two Protein Folding States in the Crystal Structure of Ribosomal Protein L20. *EMBO Rep.* **2006**, *7*, 1013–1018. [[CrossRef](#)] [[PubMed](#)]
133. Dishman, A.F.; Tyler, R.C.; Fox, J.C.; Kleist, A.B.; Prehoda, K.E.; Babu, M.M.; Peterson, F.C.; Volkman, B.F. Evolution of Fold Switching in a Metamorphic Protein. *Science* **2021**, *371*, 86–90. [[CrossRef](#)] [[PubMed](#)]
134. Dishman, A.F.; Volkman, B.F. Metamorphic Protein Folding as Evolutionary Adaptation. *Trends Biochem. Sci.* **2023**, *48*, 665–672. [[CrossRef](#)]
135. Porter, L.L.; Artsimovitch, I.; Ramírez-Sarmiento, C.A. Metamorphic Proteins and How to Find Them. *Curr. Opin. Struct. Biol.* **2024**, *86*, 102807. [[CrossRef](#)]
136. Porter, L.L. Fluid Protein Fold Space and Its Implications. *BioEssays* **2023**, *45*, e2300057. [[CrossRef](#)]
137. Schafer, J.W.; Porter, L.L. Evolutionary Selection of Proteins with Two Folds. *Nat. Commun.* **2023**, *14*, 5478. [[CrossRef](#)]
138. Knauer, S.H.; Artsimovitch, I.; Rösch, P. Transformer Proteins. *Cell Cycle* **2012**, *11*, 4289–4290. [[CrossRef](#)]
139. Knauer, S.H.; Rösch, P.; Artsimovitch, I. Transformation: The next Level of Regulation. *RNA Biol.* **2012**, *9*, 1418–1423. [[CrossRef](#)] [[PubMed](#)]
140. Das, M.; Chen, N.; LiWang, A.; Wang, L.-P. Identification and Characterization of Metamorphic Proteins: Current and Future Perspectives. *Biopolymers* **2021**, *112*, e23473. [[CrossRef](#)]
141. Bahramali, G.; Goliaei, B.; Minucheher, Z.; Salari, A. Chameleon Sequences in Neurodegenerative Diseases. *Biochem. Biophys. Res. Commun.* **2016**, *472*, 209–216. [[CrossRef](#)]
142. Prusiner, S.B. Prions. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 13363–13383. [[CrossRef](#)] [[PubMed](#)]
143. Collinge, J. Mammalian Prions and Their Wider Relevance in Neurodegenerative Diseases. *Nature* **2016**, *539*, 217–226. [[CrossRef](#)] [[PubMed](#)]
144. Riek, R.; Hornemann, S.; Wider, G.; Billeter, M.; Glockshuber, R.; Wüthrich, K. NMR Structure of the Mouse Prion Protein Domain PrP(121–231). *Nature* **1996**, *382*, 180–182. [[CrossRef](#)] [[PubMed](#)]
145. Zahn, R.; Liu, A.; Lühns, T.; Riek, R.; Von Schroetter, C.; López García, F.; Billeter, M.; Calzolari, L.; Wider, G.; Wüthrich, K. NMR Solution Structure of the Human Prion Protein. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 145–150. [[CrossRef](#)] [[PubMed](#)]
146. Pan, K.M.; Baldwin, M.; Nguyen, J.; Gasset, M.; Serban, A.; Groth, D.; Mehlhorn, I.; Huang, Z.; Fletterick, R.J.; Cohen, F.E. Conversion of Alpha-Helices into Beta-Sheets Features in the Formation of the Scrapie Prion Proteins. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10962–10966. [[CrossRef](#)]
147. Stubbs, G.; Stöhr, J. Structural Biology of PrP Prions. *Cold Spring Harb. Perspect. Med.* **2017**, *7*, a024455. [[CrossRef](#)] [[PubMed](#)]
148. Nafe, R.; Hattingen, E. The Spectrum of Molecular Pathways in Gliomas—An Up-to-Date Review. *Biomedicines* **2023**, *11*, 2281. [[CrossRef](#)] [[PubMed](#)]
149. Louros, N.; Schymkowitz, J.; Rousseau, F. Mechanisms and Pathology of Protein Misfolding and Aggregation. *Nat. Rev. Mol. Cell Biol.* **2023**, *24*, 912–933. [[CrossRef](#)] [[PubMed](#)]
150. Wickner, R.B.; Edsles, H.K.; Wu, S.; Gregg, K. Prions Are the Greatest Protein Misfolding Problem, and Yeast Has Several Solutions. *PLOS Pathog.* **2023**, *19*, e1011333. [[CrossRef](#)]
151. Spagnoli, G.; Rigoli, M.; Orioli, S.; Sevillano, A.M.; Faccioli, P.; Wille, H.; Biasini, E.; Requena, J.R. Full Atomistic Model of Prion Structure and Conversion. *PLOS Pathog.* **2019**, *15*, e1007864. [[CrossRef](#)]
152. Kovač, V.; Čurin Šerbec, V. Prion Protein: The Molecule of Many Forms and Faces. *Int. J. Mol. Sci.* **2022**, *23*, 1232. [[CrossRef](#)] [[PubMed](#)]

153. King, O.D.; Gitler, A.D.; Shorter, J. The Tip of the Iceberg: RNA-Binding Proteins with Prion-like Domains in Neurodegenerative Disease. *Brain Res.* **2012**, *1462*, 61–80. [[CrossRef](#)] [[PubMed](#)]
154. Scheckel, C.; Imeri, M.; Schwarz, P.; Aguzzi, A. Ribosomal Profiling during Prion Disease Uncovers Progressive Translational Derangement in Glia but Not in Neurons. *eLife* **2020**, *9*, e62911. [[CrossRef](#)]
155. Louka, A.; Zacco, E.; Temussi, P.A.; Tartaglia, G.G.; Pastore, A. RNA as the Stone Guest of Protein Aggregation. *Nucleic Acids Res.* **2020**, *48*, 11880–11889. [[CrossRef](#)]
156. Iben, S. To Aggregate or Not to Aggregate—Is It a Matter of the Ribosome? *BioEssays* **2023**, *45*, 2200230. [[CrossRef](#)]
157. Wickner, R.B.; Edsles, H.K.; Shewmaker, F.P.; Kryndushkin, D.; Nemecek, J.; McGlinchey, R.; Bateman, D. The Relationship of Prions and Translation. *WIREs RNA* **2010**, *1*, 81–89. [[CrossRef](#)] [[PubMed](#)]
158. Mura, C.; Veretnik, S.; Bourne, P.E. The Urfold: Structural Similarity Just above the Superfold Level? *Protein Sci. Publ. Protein Soc.* **2019**, *28*, 2119–2126. [[CrossRef](#)] [[PubMed](#)]
159. Alva, V.; Koretke, K.K.; Coles, M.; Lupas, A.N. Cradle-Loop Barrels and the Concept of Metafolds in Protein Classification by Natural Descent. *Curr. Opin. Struct. Biol.* **2008**, *18*, 358–365. [[CrossRef](#)] [[PubMed](#)]
160. Konagurthu, A.S.; Subramanian, R.; Allison, L.; Abramson, D.; Stuckey, P.J.; Garcia De La Banda, M.; Lesk, A.M. Universal Architectural Concepts Underlying Protein Folding Patterns. *Front. Mol. Biosci.* **2021**, *7*, 612920. [[CrossRef](#)]
161. Alvarez-Carreño, C.; Penev, P.I.; Petrov, A.S.; Williams, L.D. Fold Evolution before LUCA: Common Ancestry of SH3 Domains and OB Domains. *Mol. Biol. Evol.* **2021**, *38*, 5134–5143. [[CrossRef](#)]
162. Kopeck, K.O.; Lupas, A.N. β -Propeller Blades as Ancestral Peptides in Protein Evolution. *PLoS ONE* **2013**, *8*, e77074. [[CrossRef](#)]
163. Longo, L.M.; Jabłońska, J.; Vyas, P.; Kanade, M.; Kolodny, R.; Ben-Tal, N.; Tawfik, D.S. On the Emergence of P-Loop NTPase and Rossmann Enzymes from a Beta-Alpha-Beta Ancestral Fragment. *eLife* **2020**, *9*, e64415. [[CrossRef](#)] [[PubMed](#)]
164. Medvedev, K.E.; Kinch, L.N.; Dustin Schaeffer, R.; Pei, J.; Grishin, N.V. A Fifth of the Protein World: Rossmann-like Proteins as an Evolutionarily Successful Structural Unit. *J. Mol. Biol.* **2021**, *433*, 166788. [[CrossRef](#)] [[PubMed](#)]
165. Alva, V.; Remmert, M.; Biegert, A.; Lupas, A.N.; Söding, J. A Galaxy of Folds. *Protein Sci. Publ. Protein Soc.* **2010**, *19*, 124–130. [[CrossRef](#)] [[PubMed](#)]
166. Khersonsky, O.; Tawfik, D.S. Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. *Annu. Rev. Biochem.* **2010**, *79*, 471–505. [[CrossRef](#)] [[PubMed](#)]
167. Nagano, N.; Orengo, C.A.; Thornton, J.M. One Fold with Many Functions: The Evolutionary Relationships between TIM Barrel Families Based on Their Sequences, Structures and Functions. *J. Mol. Biol.* **2002**, *321*, 741–765. [[CrossRef](#)] [[PubMed](#)]
168. Galperin, M.Y.; Koonin, E.V. Divergence and Convergence in Enzyme Evolution. *J. Biol. Chem.* **2012**, *287*, 21–28. [[CrossRef](#)] [[PubMed](#)]
169. Beeby, M.; Ferreira, J.L.; Tripp, P.; Albers, S.-V.; Mitchell, D.R. Propulsive Nanomachines: The Convergent Evolution of Archaeella, Flagella and Cilia. *FEMS Microbiol. Rev.* **2020**, *44*, 253–304. [[CrossRef](#)] [[PubMed](#)]
170. Timsit, Y.; Lescot, M.; Valiadi, M.; Not, F. Bioluminescence and Photoreception in Unicellular Organisms: Light-Signalling in a Bio-Communication Perspective. *Int. J. Mol. Sci.* **2021**, *22*, 11311. [[CrossRef](#)]
171. Hall, R. Molecular Mimicry. *Adv. Parasitol.* **1994**, *34*, 81–132. [[CrossRef](#)]
172. Tsonis, P.A.; Dwivedi, B. Molecular Mimicry: Structural Camouflage of Proteins and Nucleic Acids. *Biochim. Biophys. Acta* **2008**, *1783*, 177–187. [[CrossRef](#)] [[PubMed](#)]
173. Nissen, P.; Kjeldgaard, M.; Nyborg, J. Macromolecular Mimicry. *EMBO J.* **2000**, *19*, 489–495. [[CrossRef](#)] [[PubMed](#)]
174. Chemes, L.B.; de Prat-Gay, G.; Sánchez, I.E. Convergent Evolution and Mimicry of Protein Linear Motifs in Host-Pathogen Interactions. *Curr. Opin. Struct. Biol.* **2015**, *32*, 91–101. [[CrossRef](#)]
175. Beckett, D. Functional Switches in Transcription Regulation; Molecular Mimicry and Plasticity in Protein-Protein Interactions. *Biochemistry* **2004**, *43*, 7983–7991. [[CrossRef](#)]
176. Putnam, C.D.; Tainer, J.A. Protein Mimicry of DNA and Pathway Regulation. *DNA Repair.* **2005**, *4*, 1410–1420. [[CrossRef](#)]
177. Timsit, Y. Convergent Evolution of MutS and Topoisomerase II for Clamping DNA Crossovers and Stacked Holliday Junctions. *J. Biomol. Struct. Dyn.* **2001**, *19*, 215–218. [[CrossRef](#)] [[PubMed](#)]
178. Ban, N.; Beckmann, R.; Cate, J.H.; Dinman, J.D.; Dragon, F.; Ellis, S.R.; Lafontaine, D.L.; Lindahl, L.; Liljas, A.; Lipton, J.M.; et al. A New System for Naming Ribosomal Proteins. *Curr. Opin. Struct. Biol.* **2014**, *24*, 165–169. [[CrossRef](#)] [[PubMed](#)]
179. Lecompte, O. Comparative Analysis of Ribosomal Proteins in Complete Genomes: An Example of Reductive Evolution at the Domain Scale. *Nucleic Acids Res.* **2002**, *30*, 5382–5390. [[CrossRef](#)]
180. Sulima, S.O.; Dinman, J.D. The Expanding Riboverse. *Cells* **2019**, *8*, 1205. [[CrossRef](#)]
181. Dinman, J.D. Pathways to Specialized Ribosomes: The Brussels Lecture. *J. Mol. Biol.* **2016**, *428*, 2186–2194. [[CrossRef](#)]
182. Chen, Y.-X.; Xu, Z.; Ge, X.; Hong, J.-Y.; Sanyal, S.; Lu, Z.J.; Javid, B. Selective Translation by Alternative Bacterial Ribosomes. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 19487–19496. [[CrossRef](#)]
183. Martinez-Seidel, F.; Beine-Golovchuk, O.; Hsieh, Y.-C.; Eshraky, K.E.; Gorka, M.; Cheong, B.-E.; Jimenez-Posada, E.V.; Walther, D.; Skirycz, A.; Roessner, U.; et al. Spatially Enriched Paralog Rearrangements Argue Functionally Diverse Ribosomes Arise during Cold Acclimation in Arabidopsis. *Int. J. Mol. Sci.* **2021**, *22*, 6160. [[CrossRef](#)] [[PubMed](#)]
184. Berman, H.G.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [[CrossRef](#)] [[PubMed](#)]

185. Achsel, T.; Stark, H.; Lührmann, R. The Sm Domain Is an Ancient RNA-Binding Motif with Oligo(U) Specificity. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3685–3689. [[CrossRef](#)]
186. Theobald, D.L.; Mitton-Fry, R.M.; Wuttke, D.S. Nucleic Acid Recognition by OB-Fold Proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 115–133. [[CrossRef](#)] [[PubMed](#)]
187. Murzin, A.G. OB(Oligonucleotide/Oligosaccharide Binding)-Fold: Common Structural and Functional Solution for Non-Homologous Sequences. *EMBO J.* **1993**, *12*, 861–867. [[CrossRef](#)] [[PubMed](#)]
188. Dalgarno, D.C.; Botfield, M.C.; Rickles, R.J. SH3 Domains and Drug Design: Ligands, Structure, and Biological Function. *Biopolymers* **1997**, *43*, 383–400. [[CrossRef](#)]
189. Nikonov, S.; Nevskaya, N.; Eliseikina, I.; Fomenkova, N.; Nikulin, A.; Ossina, N.; Garber, M.; Jonsson, B.H.; Briand, C.; Al-Karadaghi, S.; et al. Crystal Structure of the RNA Binding Ribosomal Protein L1 from *Thermus Thermophilus*. *EMBO J.* **1996**, *15*, 1350–1359. [[CrossRef](#)] [[PubMed](#)]
190. Medvedev, K.E.; Kinch, L.N.; Schaeffer, R.D.; Grishin, N.V. Functional Analysis of Rossmann-like Domains Reveals Convergent Evolution of Topology and Reaction Pathways. *PLOS Comput. Biol.* **2019**, *15*, e1007569. [[CrossRef](#)]
191. Pillai, A.S.; Hochberg, G.K.A.; Thornton, J.W. Simple Mechanisms for the Evolution of Protein Complexity. *Protein Sci.* **2022**, *31*, e4449. [[CrossRef](#)]
192. Jiang, H.; Blouin, C. Insertions and the Emergence of Novel Protein Structure: A Structure-Based Phylogenetic Study of Insertions. *BMC Bioinform.* **2007**, *8*, 444. [[CrossRef](#)]
193. Goldman, A.D.; Beatty, J.T.; Landweber, L.F. The TIM Barrel Architecture Facilitated the Early Evolution of Protein-Mediated Metabolism. *J. Mol. Evol.* **2016**, *82*, 17–26. [[CrossRef](#)] [[PubMed](#)]
194. CATH: Protein Structure Classification Database; University College London: London, UK, 2024.
195. Worbs, M.; Huber, R.; Wahl, M.C. Crystal Structure of Ribosomal Protein L4 Shows RNA-Binding Sites for Ribosome Incorporation and Feedback Control of the S10 Operon. *EMBO J.* **2000**, *19*, 807–818. [[CrossRef](#)]
196. Nakashima, T.; Yao, M.; Kawamura, S.; Iwasaki, K.; Kimura, M.; Tanaka, I. Ribosomal Protein L5 Has a Highly Twisted Concave Surface and Flexible Arms Responsible for rRNA Binding. *RNA* **2001**, *7*, 692–701. [[CrossRef](#)] [[PubMed](#)]
197. Ramakrishnan, V.; White, S.W. Ribosomal Protein Structures: Insights into the Architecture, Machinery and Evolution of the Ribosome. *Trends Biochem. Sci.* **1998**, *23*, 208–212. [[CrossRef](#)] [[PubMed](#)]
198. Dunker, A.K.; Romero, P.; Obradovic, Z.; Garner, E.C.; Brown, C.J. Intrinsic Protein Disorder in Complete Genomes. *Genome Inform.* **2000**, *11*, 161–171.
199. Oldfield, C.J.; Cheng, Y.; Cortese, M.S.; Brown, C.J.; Uversky, V.N.; Dunker, A.K. Comparing and Combining Predictors of Mostly Disordered Proteins. *Biochemistry* **2005**, *44*, 1989–2000. [[CrossRef](#)]
200. van Kempen, M.; Kim, S.S.; Tumescheit, C.; Mirdita, M.; Lee, J.; Gilchrist, C.L.; Söding, J.; Steinegger, M. Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **2024**, *42*, 243–246. [[CrossRef](#)] [[PubMed](#)]
201. Holm, L.; Laiho, A.; Toronen, P.; Salgado, M. DALI shines a light on remote homologs: One hundred discoveries. *Protein Sci.* **2023**, *23*, e4519. [[CrossRef](#)]
202. PyMOL Molecular Graphics System, Version 2.2.2; Schrodinger, LLC: New York, NY, USA, 2020.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.