Article

Diversification of Ferredoxins across Living Organisms

Nomfundo Nzuza [1,†], Tiara Padayachee [1,†], Wanping Chen [2], Dominik Gront [3], David R. Nelson [4] and Khajamohiddin Syed [1,*]

1   Department of Biochemistry and Microbiology, Faculty of Science and Agriculture, University of Zululand, KwaDlangezwa 3886, South Africa; nomfundonzuza11@gmail.com (N.N.); teez07padayachee@gmail.com (T.P.)
2   Department of Molecular Microbiology and Genetics, University of Göttingen, 37077 Göttingen, Germany; chenwanping1@foxmail.com
3   Faculty of Chemistry, Biological and Chemical Research Center, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland; dgront@gmail.com
4   Department of Microbiology, Immunology and Biochemistry, University of Tennessee Health Science Center, Memphis, TN 38163, USA; drnelson1@gmail.com
*   Correspondence: khajamohiddinsyed@gmail.com; Tel.: +27-(035)-902-6857
†   These authors contributed equally to the work.

Abstract: Ferredoxins, iron-sulfur (Fe-S) cluster proteins, play a key role in oxidoreduction reactions. To date, evolutionary analysis of these proteins across the domains of life have been confined to observing the abundance of Fe-S cluster types (2Fe-2S, 3Fe-4S, 4Fe-4S, 7Fe-8S (3Fe-4s and 4Fe-4S) and 2[4Fe-4S]) and the diversity of ferredoxins within these cluster types was not studied. To address this research gap, here we propose a subtype classification and nomenclature for ferredoxins based on the characteristic spacing between the cysteine amino acids of the Fe-S binding motif as a subtype signature to assess the diversity of ferredoxins across the living organisms. To test this hypothesis, comparative analysis of ferredoxins between bacterial groups, *Alphaproteobacteria* and *Firmicutes* and ferredoxins collected from species of different domains of life that are reported in the literature has been carried out. Ferredoxins were found to be highly diverse within their types. Large numbers of alphaproteobacterial species ferredoxin subtypes were found in *Firmicutes* species and the same ferredoxin subtypes across the species of Bacteria, Archaea, and Eukarya, suggesting shared common ancestral origin of ferredoxins between Archaea and Bacteria and lateral gene transfer of ferredoxins from prokaryotes (Archaea/Bacteria) to eukaryotes. This study opened new vistas for further analysis of diversity of ferredoxins in living organisms.

Keywords: Archaea; Bacteria; domains of life; Eukarya; evolution; ferredoxins; lateral gene transfer; iron-sulfur proteins

1. Introduction

Ferredoxins, iron-sulfur (Fe-S) cluster proteins, are ubiquitously present in all domains of life due to their involvement in fundamental metabolic processes such as photosynthesis, nitrogen fixation, and assimilation of hydrogen, nitrogen, and sulfur [1]. These proteins are primarily involved in the transfer of electrons in oxidation-reduction reactions [1]. Fe-S clusters are the simplest electron-transfer groups found in biological systems and are believed to have evolved early during chemical evolution [2,3]. Due to their rudimentary function and internal sequence symmetry, ferredoxins are considered to be a living protein fossil [4,5]. Based on the analysis of ferredoxin sequences, it has been proposed that all proteins evolved through tandem duplications of shorter proteins, which themselves may have emerged through the duplication of even shorter and simpler ancestral peptides [4–6].

The electron-transfer ability of ferredoxins to many proteins, including cytochrome P450 monooxygenases (CYPs/P450s) has been well established [7–10]. P450s are heme-thiolate proteins ubiquitously present in organisms across biological kingdoms [11]. With respect to ferredoxins binding partners, functional diversification and redundancy was

observed. The ferredoxins from the hyperthermophilic Archaeon, *Thermococcus kodakarensis* [12], and the three model photosynthetic organisms: *Synechocystic* sp. PCC6803 [13], *Zea mays* [14], and *Chlamydomonas reinhardtii* [15] were found to have diverse binding partners despite being paralogs indicating their functional diversification. Interestingly, despite belongings to different Fe-S cluster types, ferredoxins in myxobacteria, *Sorangium cellulosum* So ce56 [16] and actinomycobacteria, *Mycobacterium tuberculosis* H37Rv is reported [17] to have complementary functions indicating their functional redundancy. Apart from their primary electron transfer function, ferredoxins are also reported to have regulatory functions [18,19].

Ever since the inception of the name "ferredoxin" in 1962 following the isolation of ferredoxin from *Clostridium pasteurianum* [20], to date, a large number of ferredoxin are reported in all domains of life [21]. The well-known and well-studied ferredoxins include human ferredoxin known as adrenodoxin [22,23], ferredoxin from *Pseudomonas putida* known as putidaredoxin [24] and ferredoxin 1 (FDX1) from plants [25]. Ferredoxins belonging to different Fe-S cluster types have been crystallized, and some progress in understanding their interactions with P450s has been reported [7–9,26].

Ferredoxins like other Fe-S cluster proteins are classified into different types based on the number of Fe-atoms in their cluster. To date, 2Fe-2S, 3Fe-4S, 4Fe-4S, 7Fe-8S (3Fe-4S and 4Fe-4S) and 2[4Fe-4S] cluster types are reported for ferredoxins [7]. Each of the Fe-S cluster types have their own characteristic Fe-S binding motif where Fe-atom binding cysteine amino acids reside. The characteristic Fe-S binding motif includes four cysteines for 2Fe-2S, three cysteines and a proline that follows after the third cysteine for 3Fe-4S, four cysteines and a proline that follows after the fourth cysteine for 4Fe-4S and the 7Fe-8S ferredoxins having both 3Fe-4S and 4Fe-4S clusters features. The 2[4Fe-4S] ferredoxins have two 4Fe-4S motifs but the spacing between Fe-atom binding cysteines are different compared to 4Fe-4S cluster type motif. The 2[4Fe-4S] proteins are divided into two subfamilies with small proteins (approximately 55 amino acids) having isopotential Fe-S clusters and larger proteins, known as Alvin (Alv) ferredoxins, having clusters with different potentials [27]. Sequence analysis revealed the presence of an extra cysteine in 2[4Fe-4S]Alv ferredoxins exactly after three amino acids following the last cysteine of the second 4Fe-4S binding cluster [27]. Based on the arrangement of cysteines in the Fe-S binding motif, 2Fe-2S cluster proteins are also classified as plant-type, mitochondrial-type, bacterial-type and thioredoxin-type [28,29].

It is believed that among Fe-S cluster types, 4Fe-4S clusters are the first to have evolved as the conditions that mimic the primitive earth resulted in the formation of 4Fe-4S clusters [2,30,31]. Studies indicated that the 4Fe-4S cluster is sensitive to oxygen compared to the 2Fe-2S cluster [32–34] and thus, it is predicted that after the Great Oxidation Event, organisms preferred 2F-2S clusters due their oxygen tolerance. This phenomenon was also observed in a study where comparative analysis of 2Fe-2S and 4Fe-4S cluster proteins and flavodoxin proteins across the domains of life was carried out, and it has been found that 4Fe-4S cluster type proteins are abundantly present in anaerobic organisms whereas 2Fe-2S cluster type proteins are abundant in aerobic organisms [21]. A study dating back to 1966 on observing the symmetrical origin of 2[4F-4S] ferredoxin from *C. pasteurianum* by Eck and Dayhoff led to the proposal that all proteins emerged through the duplication of even shorter and simpler ancestral peptides [5]. Recently, it has been shown that indeed 2[4Fe-4S] ferredoxins have drifted from their symmetric roots via gene duplication followed by mutations [35]. Numerous studies reporting the gene duplication events as the source to the growth and diversification of Fe-S proteins has been listed in a recent article [21].

Studies on the evolution of ferredoxins are rare, and from the available data, one can understand that ferredoxins originated from a common ancestor, but divergent evolution played a critical role in their diversity [3,36–38]. In addition to this, ferredoxins are believed to be evolved from different ancestral genes that are diversified from a common ancestor [3,6,36–38]. Only a handful of studies reported lateral (or horizontal) gene transfer (LGT) of ferredoxins [18,39,40]. The ferredoxin domain of cyanobacterial origin was

found to be acquired by photosynthetic eukaryotes thorough HGT in chloroplast DnaJ-like proteins [18]. Eukaryotic protists such as *Giardia lamblia* and *Entamoeba histolytica* have ferredoxins predicted to be acquired from anaerobic bacteria by HGT [39]. Based on the percentage similarity, *Halobacterium salinarium* ferredoxin is predicted to be acquired by LGT from cyanobacterial species [40].

Apart from the above handful of studies that are confined to a small number of ferredoxins and a few species indicating LGT of ferredoxins, to date, diversity of ferredoxins within the cluster types was not studied. The current genome sequencing era has resulted in the availability of a large number of organisms' genomes that have been sequenced. This gives us an opportunity to look into ferredoxin protein sequences and better understand their diversity across the domains of life. Thus, this study aimed to address this research gap by performing genome data mining, annotation and comparative analysis of ferredoxins between the ancient bacterial group *Alphaproteobacteria* and *Firmicutes* and ferredoxins from the species belonging to different domains of life that are reported in the literature. In order to understand the diversity of ferredoxins within Fe-S cluster types in organisms, we here propose a subtype classification of the Fe-S cluster types based on the characteristic spacing between the cysteine amino acids of Fe-S binding motif as a subtype signature, considering the fact that this motif is conserved in ferredoxins per se in Fe-S cluster proteins. Furthermore, we also propose a nomenclature system for easy identification of cluster type and subtype for a ferredoxin.

2. Methodology

2.1. Species and Database

In this study, 241 alphaproteobacterial species and 227 *Firmicutes* species genomes that are available for public use at the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [41] were used (Table S1). These species are known to have P450s in their genome [42,43] and thus were selected for ferredoxins analysis considering P450s need ferredoxins for their function and so, as one can expect, there is the presence of a higher number of ferredoxins in these species. The list of species used in the study along with their species codes are presented in Table S1.

2.2. Genome Data Mining and Annotation of Ferredoxins

Ferredoxins belonging to different Fe-S cluster types reported in the literature (Table 1) were used as reference proteins. Protein BLAST [44] was performed using these reference proteins against each of the bacterial species genomes at KEGG [41]. The hit proteins were manually checked for the presence of ferredoxin Fe-S cluster type canonical motif and the proteins that have the Fe-S binding motif were selected. The selected proteins were then subjected to protein BLAST [44] at the National Center for Biotechnology and Information (NCBI) [45] against the Protein Data Bank (PDB) database [46] and analyzed for the presence of characteristic motif of ferredoxins at Pfam database [47], InterPro database [48] and NCBI Conserved Domains Database (CDD) [49]. Proteins that had a hit against ferredoxins at PDB database and have ferredoxin motifs as indicated by different databases were selected for further analysis. The rationale for using the PDB database as the first priority to sort hit proteins is that the BLAST against this database not only helped in identifying ferredoxin proteins but also helped in accurately identifying Fe-atom binding cysteine amino acids based on the alignment of the hit protein sequence with the crystallized ferredoxin sequence. In addition to this, in some cases, BLAST against PDB database also helped to sort proteins accurately whether they were 3Fe-4S or 4Fe-4S cluster proteins and 7Fe-8S or 2[4Fe-4S] cluster proteins. Hit proteins that were found to be dehydrogenases, oxidases or reductases with Fe-S binding motif were not considered as ferredoxins as the Fe-S clusters in these proteins exceeded more than two and also, they belong to a different protein family. Furthermore, sequences that had hits to ferredoxins but not in full-length were also not considered for further analysis. We also noted that manual sorting of Fe-S cluster proteins into different types by consulting as many as possible

databases such as PDB database [46], Pfam database [47], InterPro database [48] and NCBI CDD [49] is important for accurate assigning of Fe-S proteins to different types and also accurately identifying the Fe-atom binding cysteine amino acids. In this study, great care was taken to select only ferredoxins for further analysis.

Table 1. Information on ferredoxins that are used as reference proteins for datamining of ferredoxins in the bacterial species of *Alphaproteobacteria* and *Firmicutes*.

2Fe-2S.		
GenBank Accession Number or PDB Code	**Species Name**	**Reference**
NP_004100.1 (Adrenodoxin)	*Homo sapiens*	[22]
1PDX (Putidaredoxin)	*Pseudomonas putida*	[24]
ABB56370.1	*Synechococcus elongatus* PCC 7942 = FACHB-805	[50]
ABB56928.1	*Synechococcus elongatus* PCC 7942 = FACHB-805	[50]
WP_013424358.1 (FraEuI1c_3227)	*Frankia* sp. EuI1c (*Frankia inefficax* sp.)	[51]
WP_012394830.1 (mmi:MMAR_3155)	*Mycobacterium marinum*	[52]
3Fe-4S/4Fe-4S		
CAB59502.1	*Streptomyces coelicolor* A3(2)	[53]
WP_013425251.1 (FraEuI1c_4132)	*Frankia* sp. EuI1c (*Frankia inefficax* sp.)	[51]
WP_013426476.1 (FraEuI1c_5370)	*Frankia* sp. EuI1c (*Frankia inefficax* sp.)	[51]
WP_011740769.1 (Mmar_2879)	*Mycobacterium marinum*	[52]
WP_012395565.1 (Mmar_3973)	*Mycobacterium marinum*	[52]
WP_012396301.1 (Mmar_4763)	*Mycobacterium marinum*	[52]
NP_215277.1 (FdX-Rv0763c)	*Mycobacterium tuberculosis* H37Rv	[17,54]
NP_216302.1 (FdxE-Rv1786)	*Mycobacterium tuberculosis* H37Rv	[17,54]
ABB57779.1	*Synechococcus elongatus* PCC 7942 = FACHB-805	[50]
7Fe-8S		
NP_215693 (FdxC-Rv1177),	*Mycobacterium tuberculosis* H37Rv	[17]
NP_216523.1 (FdxA-Rv2007c)	*Mycobacterium tuberculosis* H37Rv	[17]
2VKR	*Acidianus ambivalens*	[55]
1H98	*Thermus thermophilus*	[56]
ABB56846.1	*Synechococcus elongatus* PCC 7942 = FACHB-805	[50]
2[4Fe-4S]		
2ZVS	*Escherichia coli* K-12	[27]
2FDN	*Clostridium acidurici*	[57]
WP_013068980.1 (FDI)	*Rhodobacter capsulatus*	[58]
2[4Fe-4S]Alv		
1BLU	*Allochromatium vinosum*	[59]
2FGO	*Pseudomonas aeruginosa*	[57]
WP_023923722.1 (FDIII)	*Rhodobacter capsulatus*	[58]
1RGV	*Thauera aromatica* K172	[60]

Note: For easy identification purpose for some ferredoxins their most popular literary names were included in parenthesis right after their GenBank accession number or PDB code.

2.3. Ferredoxins Subtype Classification and Nomenclature

Ferredoxins selected under different Fe-S cluster types were then subjected to multiple sequence alignment (MSA) at Clustal Omega database [61]. Based on the MSA patterns proteins were separated into different groups to such an extent that there were no amino acid gaps between the cysteine amino acids of the Fe-S cluster binding motif. Ferredoxins that were not aligned with other proteins were then individually subjected to analysis of spacing between cysteine amino acids of the Fe-S cluster binding motif. Then the spacing patterns between cysteine amino acids of the Fe-S cluster binding motif were presented as the characteristic signature of a subtype. Analysis of ferredoxins revealed that proline amino acid is not always conserved in 2[4Fe-4S] ferredoxins and thus is not included in the Fe-S binding motif signature for this type ferredoxins. Although proline was included for 7Fe-8S cluster proteins, only cysteine residues and the amino acid spacing between these residues can be taken as a signature as non-conservation of proline will be expected when more ferredoxins are analyzed. In order to represent different subtypes in a type, a nomenclature system was developed such that each protein starts with its Fe-S cluster type followed by the abbreviation "ST" for "subtype" and then a numeral indicating its subtype number within a type (Figure 1). Subtype numbers were assigned following the occurrence of new ferredoxins in the study. Ferredoxins identified in the bacterial species of *Alphaproteobacteria* and *Firmicutes* and retrieved from the literature are presented along with their nomenclature in Table S2.
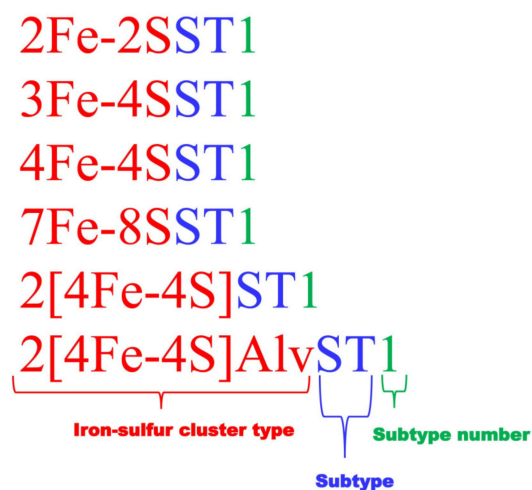

Figure 1. Ferredoxin nomenclature based on the spacing between the cysteine amino acids of the Fe-S cluster binding motif. Ferredoxins start with their Fe-S cluster type, followed by their ST indicating subtype and then the numeral indicating its subtype number in that type. Proteins grouped into different subtypes have the same characteristic spacing between the cysteine amino acids of the Fe-S cluster binding motif.

2.4. Assigning Fe-S Cluster Subtypes to the Ferredoxins Retrieved from the Literature

Ferredoxins that are reported in the literature were collected by going through the published articles and also data mining at PDB database [46]. Most of the ferredoxins belonging to the species of Archaea and some Eukarya were retrieved from an article published by Campbell and co-workers [21]. A search of the PDB database was carried out using Fe-S cluster type name, and the ferredoxin sequences were collected after manually looking at the Fe-S clusters of the ferredoxin. These ferredoxins were then subjected to subtype nomenclature as described in the Section 2.3. Ferredoxins were retrieved from literature and the data mined from the PDB database were presented along with their nomenclature in Table S3.

2.5. Phylogenetic Analysis of Ferredoxins

Phylogenetic analysis of ferredoxins was carried out following the procedure described recently by our laboratory [43,62]. The phylogenetic tree of ferredoxins was constructed using protein sequences. Firstly, the MAFFT v6.864 [63] was used to align the protein sequences that are part of the Trex web server [64]. The alignments were then used to interpret the best tree by the Trex web server [64]. Lastly, a web-based tool, VisuaLife, was used to create, visualize, and color the tree [65].

2.6. Generation of Ferredoxin Subtype Profile Heat Maps

Ferredoxin subtype profile heat maps were generated following the procedure described recently by our laboratory [43,62]. The heat map was generated using the ferredoxin subtype data to show the presence or absence of ferredoxin subtypes in three domains of life, Archaea, Bacteria, and Eukarya. The data were represented as (−3) for subtype absence (green) and (3) for subtype presence (red). A tab-delimited file was imported into Mev (Multi-experiment viewer) [66]. Hierarchical clustering using a Euclidean distance metric was used to cluster the data. Ferredoxin subtypes formed the vertical axis and three domains of life formed the horizontal axis.

3. Results and Discussion

3.1. Alphaproteobacterial and Firmicutes Species Have Different Fe-S Cluster Type Ferredoxins in Their Genomes

Genome data mining of 241 alphaproteobacterial species and 227 *Firmicutes* species revealed presence of 1307 and 281 ferredoxins in their genomes (Figure 2 and Table S1). This suggests that alphaproteobacterial species have four and a half times more ferredoxins in their genomes compared to *Firmicutes* species (Figure 2). Among alphaproteobacterial species *Sphingomonas wittichii* have the highest number of twelve ferredoxins (Table S1) and in *Firmicutes* species *Kyrpidia spormannii* have the highest number of six ferredoxins (Table S1). Six different Fe-S cluster type ferredoxins were found in alphaproteobacterial species whereas only four Fe-S cluster type ferredoxins were found in *Firmicutes* species (Figure 2). The 3Fe-4S and 2[4Fe-4S]Alv cluster type ferredoxins were not identified in the *Firmicutes* species analyzed in this study.
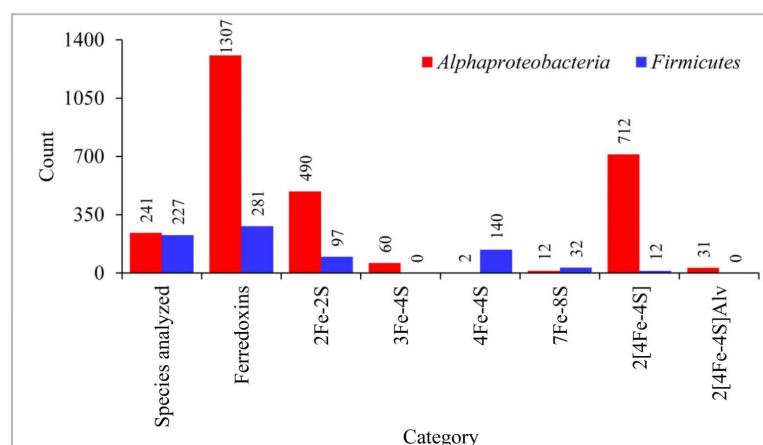

Figure 2. Comparative analysis of ferredoxins iron-sulfur (Fe-S) cluster features between alphaproteobacterial species and *Firmicutes* species. The number next to the bar indicates the count for that category. Detailed information on species and their ferredoxins is presented in Table S1.

Analysis of ferredoxin Fe-S cluster types in alphaproteobacterial species revealed that 2[4Fe-4S] cluster type ferredoxins was found to be most abundant with 712 ferredoxins followed by 2Fe-2S with 490 ferredoxins, 3Fe-4S with 60 ferredoxins, 2[4Fe-4S]Alv with 31 ferredoxins), 7Fe-8S with 12 ferredoxins and 4Fe-4S cluster with two ferredoxins

(Figure 2). Of the four Fe-S cluster types found in *Firmicutes* species, the 4Fe-4S was the most abundant with 140 ferredoxins followed by 2Fe-2S with 97 ferredoxins, 7Fe-8S with 32 ferredoxins and 2[4Fe-4S] with 12 ferredoxins (Figure 2). Comparison of ferredoxin Fe-S cluster types between these two bacterial groups revealed that alphaproteobacterial species have more ferredoxins belonging to the clusters 2[4Fe-4S] and 2Fe-2S whereas *Firmicutes* species have more ferredoxins belonging to the clusters 4Fe-4S and 7Fe-8S (Figure 2). This suggests that these two bacterial groups have different preferences with respect to Fe-S cluster type. A point to be noted is that the results of this study regarding the presence of more 4Fe-4S cluster type ferredoxins in *Firmicutes* species corroborate with a previous study where only 2Fe-2S and 4Fe-4S cluster types were analyzed as part of global protein electron carrier proteins analysis across the domain of life [21].

3.2. Highly Diverse and Common Ancestral Origin of Ferredoxins between Alphaproteobacteria and Firmicutes

In this study, we tested the amino acid spacing pattern between the cysteine amino acids of Fe-S cluster binding motif as a signature for ferredoxins subtype classification, which was therefore developed to understand diversity of ferredoxins using alphaproteobacterial species and *Firmicutes* species as model organisms. This took into consideration the fact that alphaproteobacterial species are ancient species compared to *Firmicutes* species [67] and thus they serve as the best models for understanding the diversity of ferredoxins within Fe-S cluster types. Furthermore, ferredoxins that are common to these species will be a good indication of their common ancestral origin. Analysis of ferredoxin Fe-S cluster subtypes revealed that alphaproteobacterial species and *Firmicutes* species have highly diverse ferredoxins in their genomes and some ferredoxin were found to share a common ancestor as ferredoxin belonging to the same subtype were found in both groups. This suggests that the subtype classification is helpful in understanding the diversity of ferredoxins within a cluster type. A detailed analysis on ferredoxin Fe-S cluster subtypes between these two bacterial groups is presented below:

3.2.1. 2Fe-2S

Based on the amino acid spacing pattern between the cysteine amino acids of the Fe-S cluster binding motif, 490 2Fe-2S ferredoxins of alphaproteobacterial species can be grouped into 29 subtypes, whereas 97 2Fe-2S ferredoxins of *Firmicutes* species can be grouped into 11 subtypes (Table 2). Among alphaproteobacterial 2Fe-2S ferredoxin subtypes, subtype 1 has the most ferredoxins followed by subtype 2 (Table 2), indicating subtype 1 ferredoxins are highly preferred by these species. *Firmicutes* species have the highest number of subtype 20 ferredoxins in their genomes followed by subtype 18 (Table 2). Comparative analysis of subtypes revealed that eight subtypes were found to be common between the alphaproteobacterial species and the *Firmicutes* species (Table 2), suggesting the common ancestral origin of these eight subtype ferredoxins.

3.2.2. 3Fe-4S

Sixty 3Fe-4S ferredoxins found in alphaproteobacterial species can be grouped into five different subtypes (Table 2). Among the five subtypes, subtype 1 ferredoxins present in higher numbers followed by subtypes 2 and 3 (Table 2), indicating ferredoxins with subtype 1 are preferred by alphaproteobacterial species. Interestingly, 3Fe-4S ferredoxin were not found in *Firmicutes* species analysed in this study.

3.2.3. 4Fe-4S

The two 4Fe-4S ferredoxins found in alphaproteobacterial species can be grouped into the same subtype 1 and 140 4Fe-4S ferredoxins found in *Firmicutes* species can be grouped into five different subtypes (Table 2). Of the five subtypes, subtype 2 ferredoxins were found in higher numbers in *Firmicutes* species indicating their preference for these species (Table 2). Contrary to the 2Fe-2S ferredoxin subtypes, no common 4Fe-4S subtypes were

found between alphaproteobacterial species and *Firmicutes* species (Table 2), indicating 4Fe-4S ferredoxins are highly diverse between these two bacterial groups.

Table 2. Subtype classification of ferredoxin across the domains of life. Ferredoxins collected from species belonging to different biological kingdoms that are reported in the literature and data mined at Protein Data Bank (PDB) also presented in the table as per their subtypes under the domains: Archaea, Bacteria, and Eukarya. Ferredoxin protein sequences as per their nomenclature are presented in Tables S2 and S3.

Subtypes	Cysteine Spacing Signature	No. AA	No of Ferredoxins				
			Alphaproteobacteria	*Firmicutes*	From Literature		
					Archaea	Bacteria	Eukarya
2Fe-2S							
Subtype 1	$CX_5CX_2CX_{36}C$	47	296			2	64
Subtype 2	$CX_5CX_2CX_{37}C$	48	81	1		4	22
Subtype 3	$CX_4CX_2CX_{29}C$	39	11	1	7	11	55
Subtype 4	$CX_5CX_2CX_{35}C$	46	12	3		2	
Subtype 5	$CX_5CX_2CX_{38}C$	49	12	1			
Subtype 6	$CX_4CX_2CX_{34}C$	44	11	11	7	1	
Subtype 7	$CX_4CX_2CX_{51}C$	61	8			1	
Subtype 8	$CX_4CX_2CX_{31}C$	41	7	6	1	1	
Subtype 9	$CX_4CX_2CX_{33}C$	43	2	3	19		2
Subtype 10	$CX_5CX_2CX_{39}C$	50	2			1	
Subtype 11	$CX_4CX_2CX_{35}C$	46	1				
Subtype 12	$CX_4CX_2CX_{25}C$	35	1				
Subtype 13	$CX_7CX_{34}CX_3C$	48	1				
Subtype 14	$CX_4CX_{29}C$	36	1				
Subtype 15	$CX_7CX_{29}C$	39	1				
Subtype 16	$CX_7CX_{35}C$	45	1				
Subtype 17	$CX_5CX_2CX_{33}C$	44	1				
Subtype 18	$CX_5CX_2CX_{34}C$	45	2	20	2	1	1
Subtype 19	$CX_5CX_2CX_{35}C$	46	1				
Subtype 20	$CX_5CX_2CX_{32}C$	43		44		1	5
Subtype 21	$CX_4CX_{31}CX_3C$	42		5			
Subtype 22	$CX_2CX_{41}CX_3C$	50		2			
Subtype 23	$CX_7CX_{38}CX_3C$	52				1	
Subtype 24	$CX_4CX_2CX_{30}C$	40			70		1
Subtype 25	$CX_5CX_2CX_{30}C$	41					1
Subtype 26	$CX_{12}CX_{30}CX_3C$	49	21				
Subtype 27	$CX_{12}CX_{31}CX_3C$	50	4				
Subtype 28	$CX_8CX_{44}CX_3C$	59	3				
Subtype 29	$CX_8CX_{33}CX_3C$	48	3				
Subtype 30	$CX_8CX_{32}CX_3C$	47	2				
Subtype 31	$CX_4CX_{36}CX_3C$	47	1		1		
Subtype 32	$CX_8CX_{38}CX_3C$	53	1				
Subtype 33	$CX_8CX_{39}CX_3C$	54	1				
Subtype 34	$CX_9CX_{33}CX_3C$	49	1				
Subtype 35	$CX_{12}CX_{33}CX_3C$	52	1				
Subtype 36	$CX_3CX_1CX_{38}C$	46				1	
Subtype 37	$CX_{12}CX_{32}CX_3C$	51				1	
Subtype 38	$CX_4CX_2CX_{28}C$	38			40		3
Subtype 39	$CX_4CX_2CX_{46}C$	56			2		
Subtype 40	$CX_4CX_2CX_{49}C$	59			3		
Subtype 41	$CX_4CX_2CX_{45}C$	55			2		
Subtype 42	$CX_4CX_2CX_{65}C$	75			1		

Table 2. *Cont.*

Subtypes	Cysteine Spacing Signature	No. AA	No of Ferredoxins				
			Alphaproteobacteria	*Firmicutes*	From Literature		
					Archaea	Bacteria	Eukarya
Subtype 43	$CX_4CX_2CX_{50}C$	60			2		
Subtype 44	$CX_4CX_2CX_{47}C$	57			2		
Subtype 45	$CX_4CX_2CX_{48}C$	58			1		
Subtype 46	$CX_5CX_2CX_{52}C$	63			2		
Subtype 47	$CX_5CX_2CX_{31}C$	42			1		
Subtype 48	$CX_5CX_2CX_{28}C$	39			2		
Subtype 49	$CX_5CX_2CX_{27}C$	38			10		
Subtype 50	$CX_5CX_2CX_{82}C$	93			2		
Subtype 51	$CX_5CX_2CX_{29}C$	40			1		
Subtype 52	$CX_4CX_2CX_{32}C$	42			1		
Subtype 53	$CX_5CX_2CX_{42}C$	53					1
Subtype 54	$CX_4CX_2CX_{22}C$	32					2
Subtype 55	$CX_4CX_2CX_{29}C$	39			2		
3Fe-4S							
Subtype 1	$CX_5CX_{38}CP$	47	26		2	7	
Subtype 2	$CX_5CX_{37}CP$	46	16			13	
Subtype 3	$CX_5CX_{36}CP$	45	14			2	
Subtype 4	$CX_5CX_{40}CP$	49	3				
Subtype 5	$CX_5CX_{36}CP$	45	1				
Subtype 6	$CX_5CX_{35}CP$	44				5	
Subtype 7	$CX_5CX_{49}CP$	58				2	
4Fe-4S							
Subtype 1	$CX_5CX_3CX_{33}CP$	46	2				
Subtype 2	$CX_2CX_2CX_{43}CP$	52		107			
Subtype 3	$CX_2CX_2CX_{45}CP$	54		24		1	
Subtype 4	$CX_2CX_2CX_{37}CP$	46		6			
Subtype 5	$CX_2CX_2CX_{44}CP$	53		2			
Subtype 6	$CX_2CX_2CX_{39}CP$	48		1			
Subtype 7	$CX_2CX_2CX_{36}CP$	45				2	
Subtype 8	$CX_2CX_2CX_{34}CP$	43				1	
Subtype 9	$CX_2CX_2CX_{38}CP$	47			1		1
Subtype 10	$CX_5CX_3CX_{32}CP$	45			1		
Subtype 11	$CX_5CX_3CX_{30}CP$	43			12		
Subtype 12	$CX_5CX_3CX_{31}CP$	44			2		
7Fe-8S							
Subtype 1	$CX_7CX_3CPX_{17}CX_2CX_2CX_3CP$ [*]	43	6	32		13	
Subtype 2	$CX_5CX_3CPX_{40}CX_2CX_2CX_3CP$	64	4				
Subtype 3	$CX_5CX_3CPX_{20}CX_2CX_2CX_3CP$	44	1				
Subtype 4	$CX_{10}CX_3CPX_{22}CX_2CX_2CX_3CP$	51	1				
Subtype 5	$CX_5CX_3CPX_{26}CX_2CX_2CX_3CP$	50			1		
Subtype 6	$CX_5CX_3CPX_{24}CX_2CX_2CX_3CP$	48				1	1
Subtype 7	$CX_{10}CX_3CPX_{17}CX_2CX_2CX_3CP$	46				1	
Subtype 8	$CX_5CX_3CPX_{22}CX_2CX_2CX_3CP$	46			9		
Subtype 9	$CX_5CX_3CPX_{18}CX_2CX_2CX_3C$	42			1		
Subtype 10	$CX_3CX_3CPX_{22}CX_2CX_2CX_3CP$	44			2		
2[4Fe-4S]							
Subtype 1	$CX_2CX_4CX_3CX_{18}CX_2CX_2CX_3C$	42	267				
Subtype 2	$CX_2CX_2CX_3CX_{18}CX_2CX_8CX_3C$	46	90			4	
Subtype 3	$CX_2CX_2CX_3CX_{20}CX_2CX_2CX_3C$	42	33	2	52	1	

Table 2. *Cont.*

Subtypes	Cysteine Spacing Signature	No. AA	No of Ferredoxins				
			Alphaproteobacteria	*Firmicutes*	From Literature		
					Archaea	Bacteria	Eukarya
Subtype 4	$CX_7CX_2CX_3CX_{23}CX_2CX_2CX_3C$	50	5				
Subtype 5	$CX_2CX_2CX_3CX_{42}CX_2CX_2CX_3C$	64	3				
Subtype 6	$CX_2CX_2CX_3CX_{18}CX_2CX_7CX_3C$	45	2				
Subtype 7	$CX_2CX_2CX_3CX_{18}CX_2CX_6CX_3C$	44	2				
Subtype 8	$CX_2CX_2CX_3CX_{24}CX_2CX_2CX_3C$	46	2		2		
Subtype 9	$CX_2CX_2CX_3CX_{18}CX_2CX_2CX_3C$	40	1	6	78	3	1
Subtype 10	$CX_2CX_2CX_3CX_{21}CX_2CX_2CX_3C$	43	6		2		
Subtype 11	$CX_2CX_2CX_3CX_{18}CX_3CX_2CX_3C$	42			2		
Subtype 12	$CX_2CX_2CX_3CX_{28}CX_2CX_2CX_3C$	50	132		1		2
Subtype 13	$CX_2CX_2CX_3CX_{27}CX_2CX_2CX_3C$	49	2				
Subtype 14	$CX_5CX_2CX_3CX_{20}CX_2CX_2CX_3C$	45	21				
Subtype 15	$CX_2CX_2CX_3CX_{19}CX_2CX_2CX_3C$	41	1	4	58		
Subtype 16	$CX_2CX_2CX_3CX_{40}CX_2CX_2CX_3C$	50	1				
Subtype 17	$CX_2CX_2CX_3CX_{29}CX_2CX_2CX_3C$	51	144				2
Subtype18	$CX_4CX_2CX_3CX_{18}CX_2CX_2CX_3C$	42			1	1	
Subtype 19	$CX_3CX_2CX_3CX_{20}CX_2CX_2CX_3C$	43				2	
Subtype 20	$CX_2CX_2CX_3CX_{17}CX_2CX_2CX_3C$	39			24	1	
Subtype 21	$CX_3CX_3CX_3CX_{37}CX_1CX_3CX_3C$	61				1	
Subtype 22	$CX_2CX_2CX_3CX_{26}CX_2CX_2CX_3C$	48					6
Subtype 23	$CX_2CX_2CX_3CX_{30}CX_2CX_2CX_3C$	52					1
Subtype 24	$CX_2CX_2CX_3CX_{33}CX_2CX_2CX_3C$	55			22		
Subtype 25	$CX_2CX_2CX_3CX_{32}CX_2CX_2CX_3C$	54			23		
Subtype 26	$CX_2CX_2CX_3CX_{23}CX_2CX_2CX_3C$	45			2		
Subtype 27	$CX_2CX_2CX_3CX_{34}CX_2CX_2CX_3C$	56			1		
Subtype 28	$CX_2CX_2CX_3CX_{14}CX_2CX_2CX_3C$	36			2		
Subtype 29	$CX_2CX_2CX_3CX_{22}CX_2CX_2CX_3C$	44			2		
Subtype 30	$CX_2CX_2CX_2CX_{38}CX_2CX_2CX_3C$	59			1		
Subtype 31	$CX_4CX_2CX_3CX_{19}CX_2CX_2CX_3C$	43			24		
Subtype 32	$CX_5CX_2CX_3CX_{19}CX_2CX_2CX_3C$	44			3		
Subtype 33	$CX_2CX_2CX_3CX_{16}CX_2CX_2CX_3C$	38			16		
	2[4Fe-4S]Alv						
Subtype 1	$CX_2CX_2CX_3CX_{18}CX_2CX_8CX_3CX_3C$	50	10			5	
Subtype 2	$CX_2CX_2CX_3CX_{39}CX_2CX_2CX_3CX_3C$	65	9				
Subtype 3	$CX_2CX_2CX_3CX_{43}CX_2CX_2CX_3CX_3C$	69	5			1	
Subtype 4	$CX_2CX_2CX_3CX_{42}CX_2CX_2CX_3CX_3C$	68	1				
Subtype 5	$CX_2CX_2CX_3CX_{40}CX_2CX_2CX_3CX_3C$	66	2				
Subtype 6	$CX_2CX_2CX_3CX_{38}CX_2CX_2CX_3CX_3C$	64	1				
Subtype 7	$CX_2CX_2CX_3CX_{46}CX_2CX_2CX_3CX_3C$	72	2				
Subtype 8	$CX_2CX_2CX_3CX_{44}CX_2CX_2CX_3CX_3C$	70	1				
Subtype 9	$CX_2CX_2CX_3CX_{30}CX_2CX_2CX_3CX_3C$	56			2		
Subtype 10	$CX_2CX_2CX_3CX_{19}CX_2CX_2CX_3CX_3C$	45			8		

Note: *, Only 7Fe-8S ferredoxins from *M. tuberculosis* H37Rv (Rv2007c) were found to have "arginine (R)" instead of "proline (P)". Proline is not conserved in 2[4Fe-4S] cluster ferredoxins and thus not included in the signature. Although proline was included for 7Fe-8S cluster ferredoxins, only cysteine residues and the amino acid spacing between these residues can be taken as a signature. No AA indicates number of amino acids in cysteine spacing signature motif.

3.2.4. 7Fe-8S

Twelve 7Fe-8S ferredoxins found in alphaproteobacterial species grouped into four different subtypes, whereas all the 32 *Firmicutes* species ferredoxins belongs to the subtype 1 (Table 2). Subtype 1 ferredoxins were found to be common between these two bacterial

groups (Table 2) indicating not only a common ancestral origin of these ferredoxin subtypes but also, they are wide-spread in *Firmicutes* species.

3.2.5. 2[4Fe-4S]

Seven hundred and eleven 2[4Fe-4S] ferredoxins found in alphaproteobacterial species can be grouped into 16 subtypes whereas the 12 2[4Fe-4S] ferredoxins of *Firmicutes* species can be grouped into three different subtypes (Table 2). Subtype 1 ferredoxins are more abundant in alphaproteobacterial species followed by subtypes 2 and 3 (Table 2). All the three ferredoxin subtypes of *Firmicutes* species can be found in alphaproteobacterial species (Table 2) indicating a common ancestral origin of these ferredoxin subtypes from alphaproteobacterial species to *Firmicutes* species.

3.2.6. 2[4Fe-4S]Alv

2[4Fe-4S]Alv ferredoxin was identified only in alphaproteobacterial species. The 31 2[4Fe-4S]Alv ferredoxins of alphaproteobacterial species were grouped into eight different subtypes (Table 2). Of the eight subtypes, subtypes 1 and 2 have more ferredoxins than the rest of the subtypes (Table 2).

3.3. Ferredoxin Fe-S Cluster Types Canonical Motifs

Based on the ferredoxins analyzed in the study and their Fe-S binding motif patterns, three different canonical motifs can be deduced for 2Fe-2S ferredoxins, $CX_{3-5}CX_{1-2}CX_{22-82}C$, $CX_{2-12}CX_{30-44}CX_3C$ and $CX_{4-7}CX_{29-35}C$. Among the three canonical motifs, the majority of the ferredoxin subtypes fall under the canonical motif $CX_{3-5}CX_{1-2}CX_{22-82}C$ followed by $CX_{2-12}CX_{30-44}CX_3C$ and $CX_{4-7}CX_{29-35}C$ (Table 2). This suggests that $CX_{3-5}CX_{1-2}CX_{22-82}C$ canonical motif for 2Fe-2S ferredoxins is highly prevalent in organisms. Contrary to 2Fe-2S ferredoxins, only one Fe-S binding canonical motif was observed for the rest of the ferredoxin Fe-S cluster types. $CX_5CX_{35-49}CP$ for 3Fe-4S; $CX_{2-5}CX_{2-3}CX_{30-45}CP$ for 4Fe-4S; $CX_{3-10}CX_3CPX_{17-40}CX_2CX_2CX_3CP$ for 7Fe-8S; $CX_{2-7}CX_{2-4}CX_{2-3}CX_{14-42}CX_{1-2}CX_{2-8}CX_3C$ for 2[4Fe-4S] and $CX_2CX_2CX_3CX_{18-46}CX_2CX_{2-8}CX_3CX_3C$ for 2[4Fe-4S]Alv. Analysis of Fe-S cluster binding motif amino acid patterns revealed that in all ferredoxins subtypes, the cysteine amino acid is invariantly conserved and the proline amino acid is not always conserved in 2[4Fe-4S] ferredoxins (Table 2). However, proline is invariantly conserved in 3Fe-4S and 4Fe-4S ferredoxins following the final cysteine amino acid of the Fe-S cluster binding motif, and this conservation makes these ferredoxins distinct compared to 2Fe-2S ferredoxins (Table 2). In addition to this, the amino acid spacing between the first and second cysteine amino acid is five in the Fe-S cluster binding motif in 3Fe-4S ferredoxins compared to 2Fe-2S ferredoxins (Table 2). With the exception of *Mycobacterium tuberculosis* H37Rv ferredoxin (Rv2007c), all the 7Fe-8S ferredoxins analyzed in this study have proline after the final cysteine amino acid in their Fe-S cluster binding motif. An interesting observation is that some subtypes have the same number of amino acids in the Fe-S cluster binding motif but the amino acid patterns especially the positional arrangement of cysteine amino acids-are different, indicating the amino acid patterns of Fe-S cluster binding motif is indeed a characteristic signature for a ferredoxin subtype (Table 2).

3.4. Evolutionary Linkage of Ferredoxins Subtype Classification

The subtype classification of ferredoxins proposed in this study is solely based on the arrangement of cysteine amino acids in the Fe-S cluster binding motif and to use this subtype classification as a criterion for assessing the diversity of ferredoxins one should evaluate the evolutionary linkage of subtypes, if any. Thus, here we performed evolutionary analysis of ferredoxins (Figure 3). As shown in Figure 3, with the exception of some ferredoxins, most of the ferredoxin were grouped as per their types and also as per their subtypes suggesting our subtype classification criteria certainly followed the evolutionary trend. This indicates that the amino acid spacing pattern between the cysteine amino acids of Fe-S cluster binding motifs is evolutionarily conserved and passed across

the species. One interesting aspect that can be drawn from the evolutionary analysis is that although 4Fe-4S ferredoxin aligned independently on the tree, their placement with 2[4Fe-4S] suggests the domains of 4Fe-4S ferredoxins certainly passed to 2[4Fe-4S] ferredoxins. This observation also strongly supports the hypothesis proposed by Eck and Dayhoff that all proteins emerged through the duplication of even shorter and simpler ancestral peptides [5]. Some of the ferredoxin Fe-S cluster types branched on the tree and aligned with other Fe-S cluster types, indicating they are evolutionarily linked to each other by sharing high sequence similarity in certain parts of the Fe-S cluster domain.
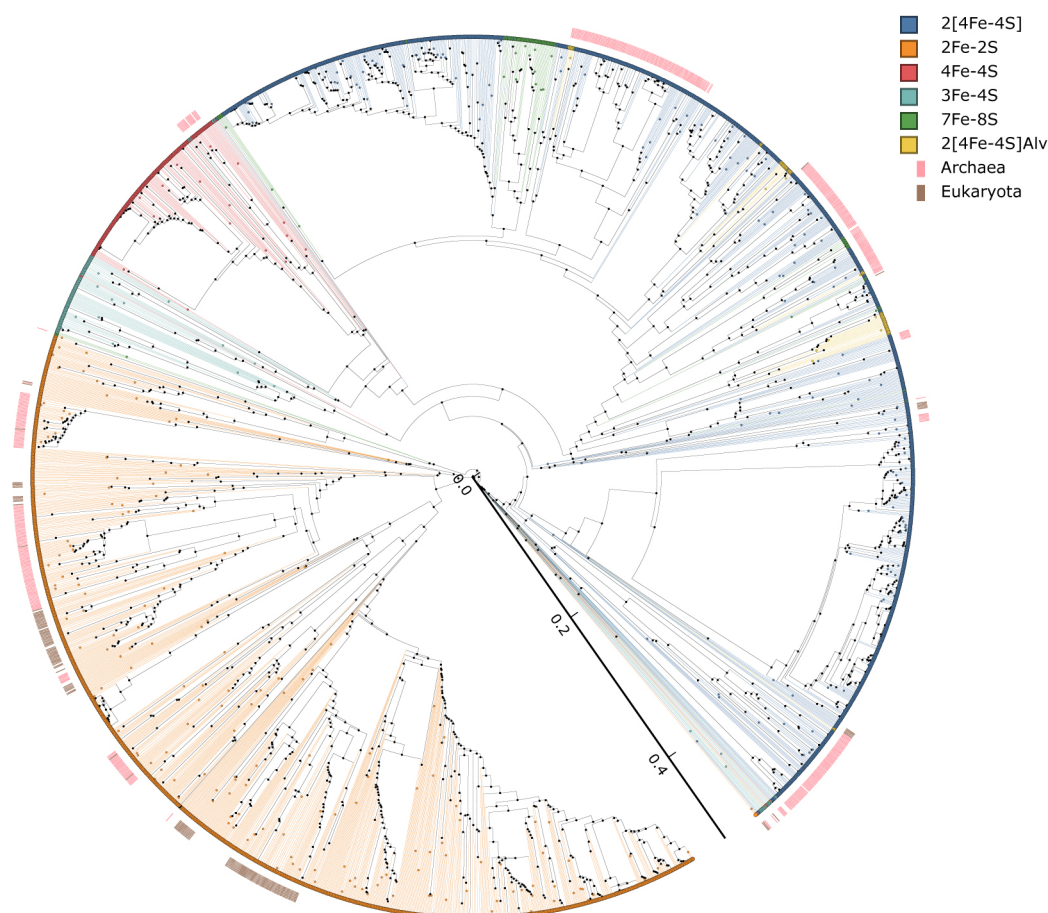
Figure 3. Phylogenetic analysis of ferredoxins. Ferredoxins belonging to different Fe-S cluster types were highlighted in different colors and indicated in the figure. Archaea and Eukaryota ferredoxin sequences were marked with pink and brown stripe, respectively. All other sequences originate from Bacteria. Ferredoxin protein sequences used to construct the phylogenetic tree are presented in Tables S2 and S3.

3.5. LGT of Ferredoxins

The current genomic era resulted in genome sequencing of large numbers of organisms and thus the availability of ferredoxin sequences from all domains of life. This has given us an opportunity to look into the LGT of ferredoxins. The subtype classification of ferredoxins as proposed in this study will certainly help in easily understanding the LGT of ferredoxins in domains of life as one can assume the ferredoxins belonging to the same subtype are indeed evolutionarily linked (as discussed in the above section). In order to understand LGT of ferredoxins across the living organisms, we have collected and annotated 538, 95 and 171 ferredoxins from Archaea, Bacteria (excluding ferredoxins from *Alphaproteobacteria/Firmicutes*) and Eukarya, respectively (Table 2 and Table S3).

Comparative analysis of ferredoxin subtypes revealed LGT of ferredoxins across the domains of life (Figure 4). Subtypes 3, 9 and 18 in 2Fe-2S and subtypes 9 and 12 in 2[4Fe-4S]

were found to be present across Archaea, Bacteria and Eukarya, indicating their LGT from Archaea/Bacteria to Eukarya (Figure 4). LGT of 2Fe-2S subtypes 24 and 38 between Archaea and Eukarya was observed (Figure 4). LGT of 4Fe-4S subtype 9 between Archaea and Eukarya, 7Fe-8S subtype 6 between Bacteria and Eukarya was observed (Figure 4). LGT of 2[4Fe-FS] subtype 17 between Bacteria and Eukarya was also observed (Figure 4). The above ferredoxin subtypes were found to be aligned together on phylogenetic tree (Figure 3) strongly indicating the LGT of these ferredoxins from prokaryotes (Archaea/Bacteria) to eukaryotes.
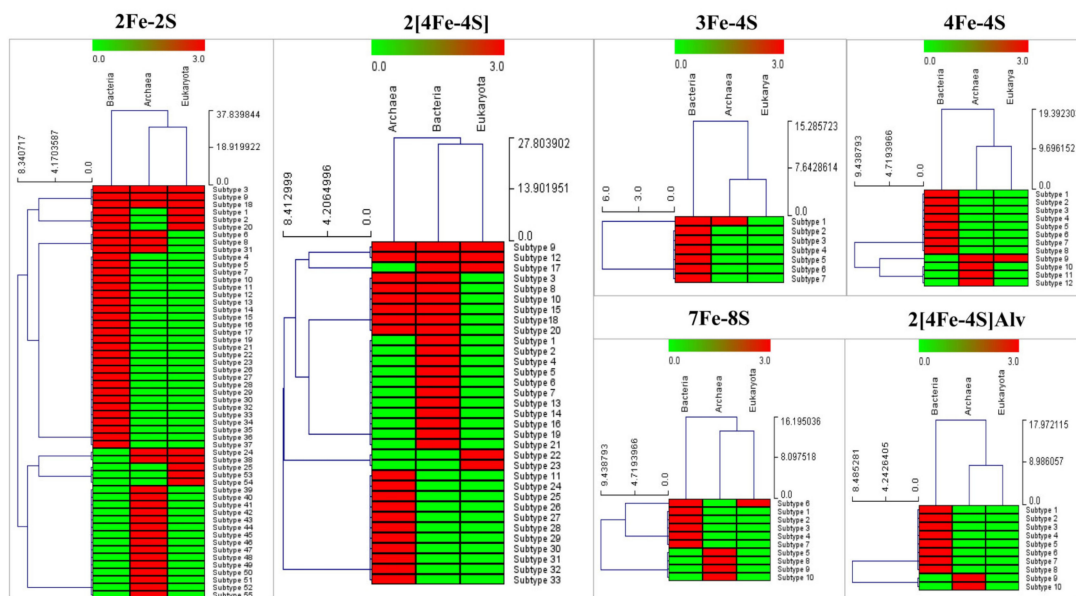
Figure 4. Heat map figure representing the presence (red) and absence (green) of ferredoxin subtypes in Archaea, Bacteria and Eukarya. Ferredoxin subtypes form the vertical axis and domains of life forms the horizontal axis.

This strongly suggests that ferredoxins from prokaryotes certainly passed to eukaryotes and subtype classification based on cysteine spacing signature motif amino acid patterns can be used to understand the LGT and the diversity of ferredoxins within a Fe-S cluster type.

4. Conclusions

Fe-S cluster proteins such as ferredoxins are the simplest known electron transfer proteins in biology. They evolved in such a way that they are capable of tuning both reduction potential and partner binding ability [21,35,68,69]. Due to the diversity in the primary amino acid sequences, to date, classification of ferredoxins was limited to the Fe-S cluster types. This led to the observation of different ferredoxin cluster types across the living organisms, but how these proteins evolved within these clusters and their distribution patterns across the living organisms were rarely reported. In this study we proposed ferredoxins subtype classification based on the amino acid spacing pattern between the cysteine amino acids of the Fe-S cluster binding motif as a subtype signature to understand the diversity and lateral (or horizontal) gene transfer of ferredoxins across the domains of life. We tested this hypothesis and found that subtype classification indeed serves as an effective tool to understand the diversity (based on the presence of number of subtypes within a Fe-S cluster type) and evolution of ferredoxins as the presence of the same ferredoxin subtypes across the domains of life was identified. For easy identification of ferredoxins belonging to different subtypes, a nomenclature system was also developed. Work is in progress on comprehensive analysis of ferredoxins from other bacterial, archaeal and eukaryotic species for better understanding the evolution of ferredoxins across the

domains of life and also to deduce the definitive characteristic signatures for ferredoxin subtypes.

References

1. Hall, D.; Cammack, R.; Rao, K. Role for ferredoxins in the origin of life and biological evolution. *Nature* **1971**, *233*, 136–138. [CrossRef] [PubMed]
2. Venkateswara Rao, P.; Holm, R. Synthetic analogues of the active sites of iron—Sulfur proteins. *Chem. Rev.* **2004**, *104*, 527–560. [CrossRef] [PubMed]
3. Cammack, R. Evolution and diversity in the iron-sulphur proteins. *Chem. Scr.* **1983**, *21*, 87–95.
4. Romero Romero, M.L.; Rabin, A.; Tawfik, D.S. Functional proteins from short peptides: Dayhoff's hypothesis turns 50. *Angew. Chem. Int. Ed.* **2016**, *55*, 15966–15971. [CrossRef]
5. Eck, R.V.; Dayhoff, M.O. Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences. *Science* **1966**, *152*, 363–366. [CrossRef]
6. Meyer, J. The evolution of ferredoxins. *Trends Ecol. Evol.* **1988**, *3*, 222–226. [CrossRef]
7. Hannemann, F.; Bichet, A.; Ewen, K.M.; Bernhardt, R. Cytochrome P450 systems—Biological variations of electron transport chains. *Biochim. Biophys. Acta* **2007**, *1770*, 330–344. [CrossRef]
8. Chiliza, Z.E.; Martínez-Oyanedel, J.; Syed, K. An overview of the factors playing a role in cytochrome P450 monooxygenase and ferredoxin interactions. *Biophys. Rev.* **2020**, *12*, 1217–1222. [CrossRef]

9. Li, S.; Du, L.; Bernhardt, R. Redox Partners: Function Modulators of Bacterial P450 Enzymes. *Trends Microbiol.* **2020**, *28*, 10. [CrossRef]

10. Kelly, S.L.; Kelly, D.E. Microbial cytochromes P450: Biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us? *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **2013**, *368*, 20120476. [CrossRef]

11. Nelson, D.R. Cytochrome P450 diversity in the tree of life. *Biochim. Biophys. Acta Proteins Proteom.* **2018**, *1866*, 141–154. [CrossRef]

12. Burkhart, B.W.; Febvre, H.P.; Santangelo, T.J. Distinct physiological roles of the three ferredoxins encoded in the hyperthermophilic archaeon *Thermococcus kodakarensis*. *mBio* **2019**, *10*, e02807–e02818. [CrossRef]

13. Cassier-Chauvat, C.; Chauvat, F. Function and regulation of ferredoxins in the cyanobacterium, *Synechocystis* PCC6803: Recent advances. *Life* **2014**, *4*, 666–680. [CrossRef]

14. Onda, Y.; Matsumura, T.; Kimata-Ariga, Y.; Sakakibara, H.; Sugiyama, T.; Hase, T. Differential interaction of maize root ferredoxin: NADP+ oxidoreductase with photosynthetic and non-photosynthetic ferredoxin isoproteins. *Plant Physiol.* **2000**, *123*, 1037–1046. [CrossRef]

15. Terauchi, A.M.; Lu, S.-F.; Zaffagnini, M.; Tappa, S.; Hirasawa, M.; Tripathy, J.N.; Knaff, D.B.; Farmer, P.J.; Lemaire, S.D.; Hase, T. Pattern of expression and substrate specificity of chloroplast ferredoxins from *Chlamydomonas reinhardtii. J. Biol. Chem.* **2009**, *284*, 25867–25878. [CrossRef]

16. Ewen, K.M.; Hannemann, F.; Khatri, Y.; Perlova, O.; Kappl, R.; Krug, D.; Hüttermann, J.; Müller, R.; Bernhardt, R. Genome mining in *Sorangium cellulosum* So ce56: Identification and characterization of the homologous electron transfer proteins of a myxobacterial cytochrome P450. *J. Biol. Chem.* **2009**, *284*, 28590–28598. [CrossRef]

17. Ortega Ugalde, S.; de Koning, C.P.; Wallraven, K.; Bruyneel, B.; Vermeulen, N.P.E.; Grossmann, T.N.; Bitter, W.; Commandeur, J.N.M.; Vos, J.C. Linking cytochrome P450 enzymes from *Mycobacterium tuberculosis* to their cognate ferredoxin partners. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 9231–9242. [CrossRef]

18. Petitjean, C.; Moreira, D.; López-García, P.; Brochier-Armanet, C. Horizontal gene transfer of a chloroplast DnaJ-Fer protein to Thaumarchaeota and the evolutionary history of the DnaK chaperone system in Archaea. *BMC Evol. Biol.* **2012**, *12*, 1–14. [CrossRef]

19. Mustila, H.; Allahverdiyeva, Y.; Isojärvi, J.; Aro, E.; Eisenhut, M. The bacterial-type [4Fe–4S] ferredoxin 7 has a regulatory function under photooxidative stress conditions in the cyanobacterium *Synechocystis* sp. PCC 6803. *Biochim. Biophys. Acta BBA-Bioenerg.* **2014**, *1837*, 1293–1304. [CrossRef]

20. Mortenson, L.; Valentine, R.; Carnahan, J. An electron transport factor from *Clostridiumpasteurianum. Biochem. Biophys. Res. Commun.* **1962**, *7*, 448–452. [CrossRef]

21. Campbell, I.J.; Bennett, G.N.; Silberg, J.J. Evolutionary relationships between low potential ferredoxin and flavodoxin electron carriers. *Front. Energy Res.* **2019**, *7*, 79. [CrossRef]

22. Cai, K.; Tonelli, M.; Frederick, R.O.; Markley, J.L. Human mitochondrial ferredoxin 1 (FDX1) and ferredoxin 2 (FDX2) both bind cysteine desulfurase and donate electrons for iron–sulfur cluster biosynthesis. *Biochemistry* **2017**, *56*, 487–499. [CrossRef]

23. Ewen, K.M.; Kleser, M.; Bernhardt, R. Adrenodoxin: The archetype of vertebrate-type [2Fe–2S] cluster ferredoxins. *Biochim. Biophys. Acta BBA-Proteins Proteom.* **2011**, *1814*, 111–125. [CrossRef]

24. Pochapsky, T.C.; Jain, N.U.; Kuti, M.; Lyons, T.A.; Heymont, J. A refined model for the solution structure of oxidized putidaredoxin. *Biochemistry* **1999**, *38*, 4681–4690. [CrossRef]

25. Sawyer, A.; Winkler, M. Evolution of *Chlamydomonas reinhardtii* ferredoxins and their interactions with [FeFe]-hydrogenases. *Photosynth. Res.* **2017**, *134*, 307–316. [CrossRef]

26. Zhang, W.; Du, L.; Li, F.; Zhang, X.; Qu, Z.; Han, L.; Li, Z.; Sun, J.; Qi, F.; Yao, Q. Mechanistic Insights into Interactions between Bacterial Class I P450 Enzymes and Redox Partners. *ACS Catal.* **2018**, *8*, 9992–10003. [CrossRef]

27. Saridakis, E.; Giastas, P.; Efthymiou, G.; Thoma, V.; Moulis, J.-M.; Kyritsis, P.; Mavridis, I.M. Insight into the protein and solvent contributions to the reduction potentials of [4Fe–4S] 2+/+ clusters: Crystal structures of the *Allochromatium vinosum* ferredoxin variants C57A and V13G and the homologous Escherichia coli ferredoxin. *JBIC J. Biol. Inorg. Chem.* **2009**, *14*, 783–799. [CrossRef]

28. Zanetti, G.; Pandini, V. Ferredoxin. In *Encyclopedia of Biological Chemistry*, 2nd ed.; Lennarz, W.J., Lane, M.D., Eds.; Academic Press: Waltham, MA, USA, 2013; pp. 296–298.

29. Bertini, I.; Luchinat, C.; Provenzani, A.; Rosato, A.; Vasos, P.R. Browsing gene banks for Fe2S2 ferredoxins and structural modeling of 88 plant-type sequences: An analysis of fold and function. *Proteins Struct. Funct. Bioinform.* **2002**, *46*, 110–127. [CrossRef]

30. Beinert, H.; Holm, R.H.; Münck, E. Iron-sulfur clusters: Nature's modular, multipurpose structures. *Science* **1997**, *277*, 653–659. [CrossRef]

31. Meyer, J. Iron–sulfur protein folds, iron–sulfur chemistry, and evolution. *JBIC J. Biol. Inorg. Chem.* **2008**, *13*, 157–170. [CrossRef]

32. Imlay, J.A. Iron-sulphur clusters and the problem with oxygen. *Mol. Microbiol.* **2006**, *59*, 1073–1082. [CrossRef] [PubMed]

33. Jagannathan, B.; Golbeck, J.H. Understanding of the binding interface between PsaC and the PsaA/PsaB heterodimer in photosystem I. *Biochemistry* **2009**, *48*, 5405–5416. [CrossRef] [PubMed]

34. Hsueh, K.-L.; Yu, L.-K.; Chen, Y.-H.; Cheng, Y.-H.; Hsieh, Y.-C.; Ke, S.-C.; Hung, K.-W.; Chen, C.-J.; Huang, T.-H. FeoC from *Klebsiella pneumoniae* contains a [4Fe-4S] cluster. *J. Bacteriol.* **2013**, *195*, 4726–4734. [CrossRef] [PubMed]

35. Mutter, A.C.; Tyryshkin, A.M.; Campbell, I.J.; Poudel, S.; Bennett, G.N.; Silberg, J.J.; Nanda, V.; Falkowski, P.G. De novo design of symmetric ferredoxins that shuttle electrons in vivo. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 14557–14562. [CrossRef]

36. Yasunobu, K.T.; Tanaka, M. The evolution of iron-sulfur protein containing organisms. *Syst. Zool.* **1973**, *22*, 570–589. [CrossRef]

37. Harayama, S.; Polissi, A.; Rekik, M. Divergent evolution of chloroplast-type ferredoxins. *FEBS Lett.* **1991**, *285*, 85–88. [CrossRef]
38. Bruschi, M.; Guerlesquin, F. Structure, function and evolution of bacterial ferredoxins. *FEMS Microbiol. Lett.* **1988**, *54*, 155–175. [CrossRef]
39. Nixon, J.E.; Wang, A.; Field, J.; Morrison, H.G.; McArthur, A.G.; Sogin, M.L.; Loftus, B.J.; Samuelson, J. Evidence for lateral transfer of genes encoding ferredoxins, nitroreductases, NADH oxidase, and alcohol dehydrogenase 3 from anaerobic prokaryotes to *Giardia lamblia* and *Entamoeba histolytica*. *Eukaryot. Cell* **2002**, *1*, 181–190. [CrossRef]
40. Pfeifer, F.; Griffig, J.; Oesterhelt, D. The fdx gene encoding the [2Fe-2S] ferredoxin of *Halobacterium salinarium* (*H. halobium*). *Mol. Gen. Genet. MGG* **1993**, *239*, 66–71. [CrossRef]
41. Kanehisa, M.; Sato, Y.; Furumichi, M.; Morishima, K.; Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **2019**, *47*, D590–D595. [CrossRef]
42. Nzuza, N.; Padayachee, T.; Syed, P.R.; Kryś, J.D.; Chen, W.; Gront, D.; Nelson, D.R.; Syed, K. Ancient Bacterial Class *Alphaproteobacteria* Cytochrome P450 Monooxygenases Can Be Found in Other Bacterial Species. *Int. J. Mol. Sci.* **2021**, *22*, 5542. [CrossRef]
43. Padayachee, T.; Nzuza, N.; Chen, W.; Nelson, D.R.; Syed, K. Impact of lifestyle on cytochrome P450 monooxygenase repertoire is clearly evident in the bacterial phylum *Firmicutes*. *Sci. Rep.* **2020**, *10*, 1–12. [CrossRef]
44. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
45. Sayers, E.W.; Beck, J.; Bolton, E.E.; Bourexis, D.; Brister, J.R.; Canese, K.; Comeau, D.C.; Funk, K.; Kim, S.; Klimke, W. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **2021**, *49*, D10. [CrossRef]
46. wwPDB Consortium. Protein Data Bank: The single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* **2019**, *47*, D520–D528. [CrossRef]
47. El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S.R.; Luciani, A.; Potter, S.C.; Qureshi, M.; Richardson, L.J.; Salazar, G.A.; Smart, A. The Pfam protein families database in 2019. *Nucleic Acids Res.* **2019**, *47*, D427–D432. [CrossRef]
48. Mitchell, A.L.; Attwood, T.K.; Babbitt, P.C.; Blum, M.; Bork, P.; Bridge, A.; Brown, S.D.; Chang, H.-Y.; El-Gebali, S.; Fraser, M.I. InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **2019**, *47*, D351–D360. [CrossRef]
49. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **2020**, *48*, D265–D268. [CrossRef]
50. Koksharova, O.A.; Klint, J.; Rasmussen, U. The first protein map of *Synechococcus* sp. strain PCC 7942. *Microbiology* **2006**, *75*, 664–672. [CrossRef]
51. Lau, I.C.; Feyereisen, R.; Nelson, D.R.; Bell, S.G. Analysis and preliminary characterisation of the cytochrome P450 monooxygenases from *Frankia* sp. EuI1c (*Frankia inefficax* sp.). *Arch. Biochem. Biophys.* **2019**, *669*, 11–21. [CrossRef]
52. Child, S.A.; Bradley, J.M.; Pukala, T.L.; Svistunenko, D.A.; Le Brun, N.E.; Bell, S.G. Electron transfer ferredoxins with unusual cluster binding motifs support secondary metabolism in many bacteria. *Chem. Sci.* **2018**, *9*, 7948–7957. [CrossRef]
53. Bentley, S.D.; Chater, K.F.; Cerdeño-Tárraga, A.-M.; Challis, G.L.; Thomson, N.; James, K.D.; Harris, D.E.; Quail, M.A.; Kieser, H.; Harper, D. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2). *Nature* **2002**, *417*, 141. [CrossRef]
54. McLean, K.J.; Warman, A.J.; Seward, H.E.; Marshall, K.R.; Girvan, H.M.; Cheesman, M.R.; Waterman, M.R.; Munro, A.W. Biophysical characterization of the sterol demethylase P450 from *Mycobacterium tuberculosis*, its cognate ferredoxin, and their interactions. *Biochemistry* **2006**, *45*, 8427–8443. [CrossRef]
55. Frazão, C.; Aragão, D.; Coelho, R.; Leal, S.S.; Gomes, C.M.; Teixeira, M.; Carrondo, M.A. Crystallographic analysis of the intact metal centres [3Fe–4S] 1+/0 and [4Fe–4S] 2+/1+ in a Zn2+-containing ferredoxin. *FEBS Lett.* **2008**, *582*, 763–767. [CrossRef]
56. Macedo-Ribeiro, S.; Martins, B.M.; Pereira, P.B.; Buse, G.; Huber, R.; Soulimane, T. New insights into the thermostability of bacterial ferredoxins: High-resolution crystal structure of the seven-iron ferredoxin from *Thermus thermophilus*. *JBIC J. Biol. Inorg. Chem.* **2001**, *6*, 663–674. [CrossRef]
57. Dauter, Z.; Wilson, K.S.; Sieker, L.C.; Meyer, J.; Moulis, J.-M. Atomic resolution (0.94 Å) structure of *Clostridium acidurici* ferredoxin. Detailed geometry of [4Fe-4S] clusters in a protein. *Biochemistry* **1997**, *36*, 16065–16073. [CrossRef]
58. Saeki, K.; Suetsugu, Y.; Tokuda, K.-I.; Miyatake, Y.; Young, D.; Marrs, B.; Matsubara, H. Genetic analysis of functional differences among distinct ferredoxins in *Rhodobacter capsulatus*. *J. Biol. Chem.* **1991**, *266*, 12889–12895. [CrossRef]
59. Moulis, J.M.; Sieker, L.C.; Wilson, K.S.; Dauter, Z. Crystal structure of the 2 [4Fe-4S] ferredoxin from *Chromatium vinosum*: Evolutionary and mechanistic inferences for [3/4Fe-4S] ferredoxins. *Protein Sci.* **1996**, *5*, 1765–1775. [CrossRef]
60. Unciuleac, M.; Boll, M.; Warkentin, E.; Ermler, U. Crystallization of 4-hydroxybenzoyl-CoA reductase and the structure of its electron donor ferredoxin. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 388–391. [CrossRef]
61. Madeira, F.; Park, Y.M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A.R.; Potter, S.C.; Finn, R.D. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47*, W636–W641. [CrossRef]
62. Khumalo, M.J.; Nzuza, N.; Padayachee, T.; Chen, W.; Yu, J.H.; Nelson, D.R.; Syed, K. Comprehensive Analyses of Cytochrome P450 Monooxygenases and Secondary Metabolite Biosynthetic Gene Clusters in *Cyanobacteria*. *Int. J. Mol. Sci.* **2020**, *21*, 656. [CrossRef] [PubMed]
63. Katoh, K.; Kuma, K.; Toh, H.; Miyata, T. MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **2005**, *33*, 511–518. [CrossRef] [PubMed]

64. Boc, A.; Diallo, A.B.; Makarenkov, V. T-REX: A web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic Acids Res.* **2012**, *40*, W573–W579. [CrossRef] [PubMed]

65. Kryś, J.D.; Gront, D. VisuaLife: Library for interactive visualization in rich web applications. *Bioinformatics* **2021**. [CrossRef] [PubMed]

66. Howe, E.A.; Sinha, R.; Schlauch, D.; Quackenbush, J. RNA-Seq analysis in MeV. *Bioinformatics* **2011**, *27*, 3209–3210. [CrossRef]

67. Battistuzzi, F.U.; Feijao, A.; Hedges, S.B. A genomic timescale of prokaryote evolution: Insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* **2004**, *4*, 44. [CrossRef]

68. Hosseinzadeh, P.; Lu, Y. Design and fine-tuning redox potentials of metalloproteins involved in electron transfer in bioenergetics. *Biochim Biophys. Acta BBA-Bioenerg.* **2016**, *1857*, 557–581. [CrossRef]

69. Kim, J.Y.; Nakayama, M.; Toyota, H.; Kurisu, G.; Hase, T. Structural and mutational studies of an electron transfer complex of maize sulfite reductase and ferredoxin. *J. Biochem.* **2016**, *160*, 101–109. [CrossRef]