



Comparative Analysis of Some Algal Ferritins Belonging To Chlorophyta and Rhodophyta

¹Sana Raoof, ²Bhavani.V, ³Nagalakshmi.T, ⁴Ashwini.M, ⁵G. Uma Sree,* ⁶Janardhan.k

^{1, 2, 3, 4, 5, 6}Department of Biotechnology, AV College PG Centre, Hyderabad

*E-mail- janardhank_2001@rediffmail.com

Received: 23rd Feb 2015, Accepted: 6th April 2015, Published: 1st May 2015

Abstract

Ferritin is the most important protein involved in the Iron metabolism and storage. Iron is an extremely important element whose function is involved in the oxygen transport and it participates in the electron transfer reactions and various redox potential reactions. Ferritin, the globular cytoplasmic protein is highly conserved and is involved in the regulation of the flow of iron into and out of the cell. The structural integrity of the protein makes it convenient to store 45,000 molecules of iron in it. In the current paper we have worked on the analysis of ferritin from different species of Algae using bioinformatics tools. Algal ferritins studies are undertaken with reference to some of the members of Chlorophyta and Rhodophyta. *Chlamydomonas reinhardtii* is a model member of Chlorophyta involving two unlinked genes FER1 and FER2 encoding ferritin sub-units. Comparative studies are performed in different classes of algae i.e. Chlorophyta and Rhodophyta to understand variations in ferritin sequences.

Key words- Ferritin, electron transfer, Chlorophyta, Rhodophyta, FER1 and FER2.

Introduction

Metal interactions with organic compounds are numerous in all living organisms. They are essential for many biochemical processes occurring within cells, and concern not only metabolism, but also some regulatory mechanisms of gene expression. Among metals, iron is of special interest because it is required in most of the cellular redox reactions and it is one of the major metals involved in electron transfer chains. However, its strong reactivity with oxygen makes it a difficult element to handle by aerobic organisms. Indeed, both its insolubility in the form of ferric hydroxides, or its toxicity through the Fenton reaction producing hydroxyl radicals (which are among the most chemically reactive species), have introduced evolutionary constraints in order to enable this metal to be safely utilized by living organisms [1]. The narrow efficient iron concentration required for cellular needs is strictly controlled by biological processes acting both at the transport and the storage levels. In multicellular organisms, transport mechanisms regulate iron traffic from

uptake to long distance tissular distribution, and ultimately to subcellular allocation [2]. Although the structure of ferritins is highly conserved between plants and animals, their cellular localization differs. Furthermore, regulation of ferritin gene expression in response to iron excess occurs at the transcriptional level in plants, in contrast to animals which regulate ferritin expression at the translational level.

Iron is toxic in uncontained situations because it catalyzes the production of free radical. Thus iron in the cell is stored in “FERRITIN” [10].

In algae, ferritins are plastid-located proteins able to form a holosphere, which can contain up to 4500 Fe atoms. This process allows us to buffer free iron, making it available under a safe form, and reveals a key role for ferritin in iron homeostasis and protection against iron-mediated oxidative stress [4]. Pre-ferritin, with a plastid targeting sequence, is encoded by a multigene family in plants. Increased ferritin production in high-iron-supplied cells is accomplished, at least in part, by transcriptional regulation of one or more ferritin genes [5]. This pattern of expression is consistent with a role for ferritin as an iron storage molecule under conditions of iron overload [6]. Despite extensive studies conducted during the last decade, many factors regulating the expression of ferritin genes in plants remain unknown (Marek Figlerowicz *et.al.*, 2009).

Ferritins are a broad superfamily of iron storage proteins, found in all the living kingdom, except in yeast (Andrews *et al.*, 2003; Briat *et al.*, 2006; Arosio *et al.*, 2008).

Three subclasses of these proteins can be defined:

- (1) haem-free ferritins present both in pro- and eukaryotes;
- (2) haem-containing bacterioferritins, found only in bacteria; and
- (3) DNA binding proteins from starved cells (Dps), called miniferritins,

Present in prokaryotes (Smith, 2004). Ferritins and bacterioferritins are composed of 24 subunits whereas only 12 identical subunits form the Dps proteins. These subunits assemble in a spherical protein shell defining a central cavity able to

accommodate between 2000 and 4000 ferric iron atoms for ferritins and 500 atoms for miniferritins [3].

***Chlamydomonas* model.** Significant insights into plant metal metabolism can be gained from studying other model photosynthetic organisms such as the unicellular green alga *Chlamydomonas reinhardtii*. With its simple growth requirements *C. reinhardtii* is a valuable experimental model for the study of metalloprotein biosynthesis and metal-responsive gene regulation in photosynthetic organisms. While efficient transport mechanisms for iron uptake are an essential element in all pro- and eukaryotic cells, its intracellular availability and storage have to be tightly regulated, not only to buffer supply and demand, but also to prevent cell damage from undesirable reactions of free radicals, formed catalytically by free Fe ions. Ferritin represents the most common form of iron storage in all domains of life [2].

Materials and Methods

Comparative studies are done in different classes of algae i.e. Chlorophyta and Rhodophyta to understand variations in ferritin sequences. In the present investigation we have worked on comparing the ferritin protein sequences among the five different species of Algae. They are:

1. Ferritin 2 of *Chlamydomonas reinhardtii*, with the accession no. ABW87266.1.
2. Ferritin of *Ulva fasciata*, with the accession no. ABR23157.1.
3. Ferritin of *Coccomyxa subellipsoidea* C-169, with the accession no. XP_005644133.1.
4. Ferritin (ISS) *Ostreococcus tauri*, with the accession no. XP_003075182.1.
5. Ferritin of *Pyropia yezoensis*, with the accession no. AFR78246.1.
6. Ferritin chloroplast precursor *Chondrus crispus*, with the accession no. CDF36738.1.

The comparative analysis was done by the use of the following methods discussed hereunder:

1. ProtParam:

In the present investigation, primary structure analysis of proteins was done with the help of ProtParam. ProtParam computes various physicochemical properties that can be deduced from a protein sequence. No additional information is required about the protein under consideration (Walker, 2005). The properties include theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient, instability index, aliphatic index and grand average hydropathy (GRAVY) were computed using the Expasy's ProtParam server (Neelima,

2009). ProtParam can be accessed at <http://www.expasy.org/tools/protparam.html>.

2. GOR(Garnier-Osguthorpe-Robson)

The secondary structure analysis was done using the tool GOR. The GOR method is based on information theory and was developed by J.Garnier, D.Osguthorpe and B.Robson (J.Mol.Biol.120, 97, 1978). The OR method uses both information theory and Bayesian statistics for predicting the secondary structure of proteins (Garnier *et al.*, 1978). It is based on probability parameters derived from empirical studies of known protein tertiary structures solved by X-ray crystallography. The GOR method takes into account not only the propensities of individual amino acids to form particular secondary structures, but also the conditional probability of the amino acid to form a secondary structure given that its immediate neighbors have already formed that structure. The method is therefore essentially Bayesian in its analysis. The predicted secondary structure is the one of highest probability compatible with a predicted helix segment of at least four residues and a predicted extended segment of at least two residues^[7]. GOR can be accessed at <http://gor.bb.iastate.edu/>.

3. Prosite:

In the present analysis, PROSITE was used for the identification of domains. The PROSITE database consists of a large collection of biologically meaningful signatures that are described as patterns or profiles. Each signature is linked to documentation that provides useful biological information on the protein family, domain or functional site identified by the signature. Each PROSITE signature is linked to an annotation document where the user can find information on the protein family or domain detected by the signature, origin of its name, taxonomic occurrence, domain architecture, function, 3D structure, main characteristics of the sequence and some references. Recently, for families or domains whose structure is known, a direct link to a representative PDB entry is provided in the documentation, in order to make the description of the 3D structure more comprehensible. A popular way to identify similarity between proteins is to perform a pairwise alignment. When the identity is >40% this method gives good results. A multiple sequence alignment (MSA) gives a more general view of a conserved region by providing a better picture of the most conserved residues, which are usually essential for the protein function. The various amino acids can then be weighed according to their degree of conservation. PROSITE currently contains patterns and profiles specific for more than a

thousand protein families or domains [8]. The database is accessible at <http://www.expasy.org/prosite/>.

4. Nuclear export signals:

Nuclear export signals (NES) are extremely important regulators of the sub cellular location of proteins. This regulation has an impact on transcription and other nuclear processes, which are fundamental to the viability of the cells. The NES are calculated by considering the Hidden Markov Model and Neural Networks for each residue to the given sequence. NetNES is accessible at <http://www.cbs.dtu.dk/services/NetNES-1.1/>

5. Clustal W:

Progressive alignment is a widely used heuristic method to compute multiple sequence alignments (MSAs). Progressive alignment basically consists of three steps. First, a distance value between each pair of input sequences is computed (pairalign). Second, a phylogenetic tree is calculated based on the distance matrix (guided tree). Finally, pairwise alignment of various profiles is carried out following the branching order in the phylogenetic tree to form the final MSA (malign). A popular tool based on this method is ClustalW [12]. The program can be run on-line from the EBI web server: <http://www.ebi.ac.uk/tools/clustalw2>

Results And Discussion

The results obtained from protparam, among the 6 ferritin sequences belonging to different members of Chlorophyta and Rhodophyta it was found that the amino acid 'Alanine' is present in highest frequency when compared to the other amino acids. The *Chondrus crispus* has low frequency of alanine when compared to other members. Theoretical PI was found to be high in *Chlamydomonas reinhardtii* and *Ostreococcus tauri* (Table-1).

After analyzing with GOR all the amino acids that participate in the formation of Alpha Helix were identified. From the results of Protparam and GOR it was revealed that the species, *Chlamydomonas reinhardtii* and *Ostreococcus tauri* which belongs to phylum Chlorophyta are closely related than the *Chondrus crispus* and *Pyropia yezoensis* which belongs to phylum Rhodophyta (Table-1).

With the use of the Prosite tool it was found that the initial pattern of the domain region in all the 6 sequences was similar and it was found to be of same length and almost all the domains showed GLUTAMIC ACID as the amino acid which was involved in the functional activity of the protein, except for *Chlamydomonas reinhardtii* in which it was found that HISTIDINE is involved in the functional activity of the protein (Fig-1).

Table1: The results obtained from PROTPARAM & GOR

	Accession number	Total no. of Amino acids	Helices	Extended strand	Random coil	High % of A.A	Theoretical PI	No. of -vely charged AA (Asp + Glu)	No. of +vely charged AA (Arg + Lys)	Instability index
1	ABW87266.1 <i>Chlamydomonas reinhardtii</i>	298	48.32%	9.73%	41.95%	Ala (A) 39 13.1% Leu (L) 32 10.7%	5.59	39	30	42.92 unstable protein
2	ABR23157.1 <i>Ulva fasciata</i>	237	41.30%	13.50%	45.15%	Ala (A) 29 12.2% Glu (E) 24 10.1%	5	36	24	46.73 unstable protein

3	XP_005644133.1 <i>Coccomyxa subellipsoidea</i> C-169	185	55.68%	16.22%	28.11%	Ala (A) 23 12.4% Glu (E) 17 9.2% Leu (L) 17 9.2%	5.14	28	19	42.94 unstable protein
4	XP_003075182.1 <i>Ostreococcus tauri</i>	242	56.20%	6.61%	37.19%	Ala (A) 34 14.0% Glu (E) 24 9.9%	5.43	37	28	45.25 unstable protein
5	AFR78246.1 <i>Pyropia yezoensis</i>	264	37.50%	13.26%	49.24%	Ala (A) 42 15.9% Val (V) 24 9.1%	4.52	33	17	35.54 unstable protein
6	CDF36738.1 <i>Chondrus crispus</i>	249	41.77%	11.65%	46.59%	Ala (A) 27 10.8%	5.34	34	24	55.65 unstable protein

Fig.1. Results obtained from Prosite

ABW87266.1| [*Chlamydomonas reinhardtii*]

MLLSTSCAVSAVSSRLDSGRAGRWPRAAAAAPSATGVRRGVLQTRRALSDSEGREPAREDPVIARAREQVQSGLVFQP
MGEVQPLVAAMDQQLMDPKAEPGLAATYSLARSNYSPDLESIGINEQINIELNMSYVYTSMYNFFARDDVGLPGFAAY
FRHNSDDEREHAHLLMNYQTQRGGRVRLALAPPETFWHAEKGDALHATELALSLEKLNFKLRDLHTVAQTVGDA
DATHFIEDYLLHEQSKDVKEAAVLVSQVRRAGRGHGVFHVDTLLAQEYGDKGAGGLDGNAGAGPLAG

ABR23157.1| [*Ulva fasciata*]

MASRCTVKPMRTARAVPAVAAQRRVRKAAAEVVTGMVFQPFSEVQSELSTVNDAPVTQSYARVDYHPSCEAAAMN
EQINIEYITISYVYHALHSYFARDNVGLPGFAKFFSEASEEERGHQAQLMDYQVKRGGRVELKPLSAPEMEFANDDKGE
ALYAMELALSLEKLNFKLQALHAIADENQDPALCDFIEGELLQEQVDSVKQHAIEYVSQQLRRVGKGVGVYIFDRELDE
GEGQGA

XP_005644133.1| [*Coccomyxa subellipsoidea* C-169]

MNSPDVDDSFARLNYHQDSEIGVNEQINHEYSMSYQYHAMSNYFNRDNVGLPGFAAFFRASSLEERNHAQQLMDFQA
TRGGRVKLAALAAPPDYNHEQKGDALYAMELALALEKLNFKMLFELHETAIEKRKDNMTDFVEEMLSEQAQGVKE
VSEYVAQLRRVGKGLGVYEFDAKMAQRATTLA

XP_003075182.1| [*Ostreococcus tauri*]

MRVARERRVARERAGDGKRTRSSAQSSARWRARVEASERPNVDADASAAIEATSTHPVSQTTSYSFARSGFSAECERA
LNAQINVEYNVSYIYHAMWAFFERDNVALKGFADHFKAEALEERSHAEQLMEYVNLRGGRVELQQLPPQTEYDHPE
KGCALYALELSLEKLNNDKLCELHRVAEDAGDAHMCDFLEGAMMDPQVQSVREVAEMVATLLRMGPPGDGMAA
WHFDQYLQNKA

AFR78246.1| [*Pyropia yezoensis*]

MTMAAFASVAPVGVSTFAPGASLSTGRPGAGAVAAPSTSRSAARMTFSSGSPSGGETIDFSDVDVTDAGAQFSGMVFTP
DTADAPLSRANVGFSQACQDAVNNQIQVEYTASYAYHAMFAYFNRTVALPGFAKYFEEQSLEERTHADEFMRMYN
KRGGQVVLKPLAVPSMSFNNTDGTSDAVYAMDHLQLEKFVWAKLEEVAAAANADNDLSLADLIDDYVQEQQVQAVK
KAADMVAQLKRVGTPHGVWHFDQEVLLGGEDAQA

CDF36738.1| [*Chondrus crispus*]

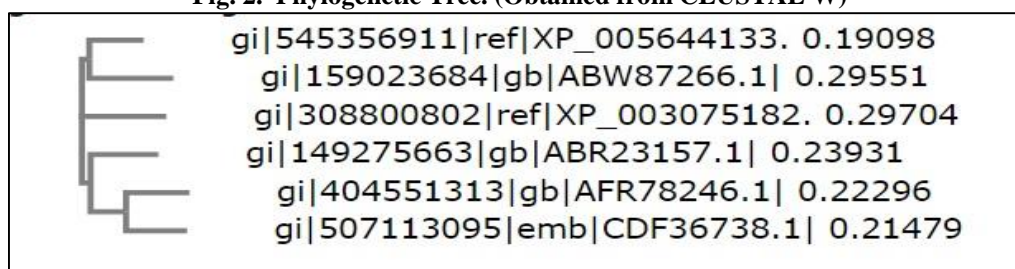
MSPQAFVSPVPLSSLRPSLSTLTQRRPSATPSRTATRAMMADDRSSIDFGKVDIKGAGENLSGLIFTPESESPLKTRCDV
GYTEQQEAAVNRHINVETASYAYHAIWAYFDKDTVLPFGFAKYFAEQSEEERTHAAEFMKYQNKRGGSVELQPIAV
PEMTFTQEDGTSDAVYAMDHLQLEKFVFKLRALHKAQDAGDEQMDMIDDYLTHQIDAIKTAADYVAQLKRVG
TPHGIYHFDLNLQGC

Table.2. Results obtained from NES

Sl.No	Organism	Seq-Pos-Residue	ANN	HMM	NES	Prediction
1	ABR23157.1 <i>Ulva fasciata</i>	Sequence-160 -L Sequence-161-A Sequence-162-L Sequence-164-L	0.527 0.101 0.268 0.528	0.904 0.897 0.957 0.931	1.132 0.717 0.813 1.271	yes yes yes yes
2	XP_005644133.1 <i>Coccomyxa subellipsoidea</i> C-169	Sequence-109-L Sequence-110-A Sequence-111-L Sequence-112-A Sequence-113-L Sequence-116-L	0.806 0.098 0.175 0.090 0.455 0.423	0.897 0.863 0.951 0.917 0.945 0.322	1.133 0.732 0.837 0.802 1.302 0.606	yes yes yes yes yes yes
3	ABW87266.1 <i>Chlamydomonas reinhardtii</i>	Sequence-209-L Sequence-210-S Sequence-211-L	0.330 0.093 0.352	0.857 0.839 0.942	0.665 0.638 0.797	yes yes yes
4	XP_003075182.1 <i>Ostreococcus tauri</i>	Sequence-164-L Sequence-166-L Sequence-168-L Sequence-170-L	0.139 0.629 0.279 0.469	0.931 0.951 0.959 0.828	0.678 1.222 0.854 1.174	yes yes yes yes
5	AFR78246.1 <i>Pyropia yezoensis</i>	Sequence-188-L Sequence-190-L	0.182 0.327	0.937 0.936	0.655 0.703	yes yes
6	CDF36738.1 <i>Chondrus crispus</i>	Sequence-178-L Sequence-180-L Sequence-182-L	0.146 0.365 0.613	0.903 0.925 0.934	0.620 0.701 1.225	yes yes yes

From the NES results it was found that the amino acids Alanine and Leucine will participate in the protein export

Fig. 2. Phylogenetic Tree. (Obtained from CLUSTAL W)



The above results are obtained from ClustalW, by performing multiple sequence alignment and constructing a phylogenetic tree to assess the pattern of evolution based on their genetic characteristics. The dendrogram was divided into three separate clusters I, II, III. The cluster I showed *Coccomyxa subellipsoidea* and *Chlamydomonas Reinhardtii* to be genetically related and the cluster II has only *Ostreococcus tauri* which indicates that it is more diverse from the other members. In the cluster III, *Ulva fasciata*, *Pyropia yezoensis* and *Chondrus crispus* falls together in the same cluster showing their genetic relatedness.

Conclusion

Ferritin has long been known as an iron storage protein. In the present investigation using different bioinformatics tools (ProtParam, PROSITE, GOR, NetNes, ClustalW) we have tried to analyze which amino acid is present in abundance, which amino acid is playing role in the protein export and functional activity and also the genetic relatedness among these selected organisms. From the results obtained it is concluded that among the 6 different members of Chlorophyta and Rhodophyta selected for the comparative analysis of ferritin, the amino acid Alanine was found to be present in highest frequency and Alanine and Leucine together might be performing the protein export function. Glutamic acid

is the amino acid found to be playing role in the functional activity of the protein in most of them. The dendrogram revealed that *Ostreococcus tauri* is genetically diverse from the other organisms selected for the analysis.

Acknowledgement

We express our thanks to Mrs. G. Ramya, IIT for her critical remarks.

References

- [1] Jean-Francois Briat*, Karl Ravet, Nicolas Arnaud, Céline Duc, Jossia Boucherez, Brigitte Touraine, Françoise Cellier and Frederic Gaymard (2009). "New insights into ferritin synthesis and function highlight a link between iron homeostasis and oxidative stress in plants". *Annals of Botany*. doi:10.1093/aob/mcp. Pg 128.
- [2] Sharon La Fontaine,1,2 Jeanette M. Quinn, Stacie S. Nakamoto, M. Dudley, Vera Go'hre, Jeffrey L. Moseley, Janette Kropat, and Sabeeha Merchant1, (2002). "Copper-Dependent Iron Assimilation Pathway in the Model Photosynthetic Eukaryote *Chlamydomonas reinhardtii*". *Eukaryote Cell*. DOI: 10.1128/EC.1.5.Pg 736–757.
- [3] Briat, J. F., Duc, C., Ravet, K., and Gaymard, F. (2010). "Ferritins and iron storage in plants". *Biochim. Biophys. Acta*. 1800.
- [4] Ravet, K., Touraine, B., Boucherez, J., Briat, J. F., Gaymard, F., and Cellier, F. (2009). "Ferritins control interaction between iron homeostasis and oxidative stress in *Arabidopsis*". *Plant J*. 57, 400–41206–814.
- [5] Briat, J. F., S. Lobreaux, N. Grignon, and G. Vansuyt. (1999). "Regulation of plant ferritin

synthesis: how and why". *Cell. Mol. Life Sci*. 56: Pg 155–166.

[6] Crichton, R. R., and J.-L. Pierre. (2001). "Old iron, young copper: from Mars to Venus". *Biometals*. 14:99 Pg 112.

[7] Taner Z. Sen,1,2 Robert L. Jernigan,1,2 Jean Garnier3 and Andrzej Kloczkowski1,* (2005). "GOR V server for protein secondary structure prediction". *Structural Bioinformatics*. Vol. 21 no. 11 2005, pg 2787–2788.

[8] Lorenza Bordoli1, Alexandre Gattiker, Edouard De Castro, Philipp Bucher2 and Amos Bairoch., (2004). "Recent improvements to the PROSITE database". *Nucleic Acids Research*. Vol. 32.

[9] Pradeep N V1, Anupama2*, Ankitha K3, Pooja J4., (2012). "Categorizing Phenomenal features of α -Amylase (*Bacillus species*) using Bioinformatic Tools". *Advances in Life Science and Technology*. ISSN Pg 2224-7181.

[10] Swathi. K , 2Hymavathi CH , 3Janardhan .K., (2013). "Comparative Analysis of Some Fungal Ferritins Belonging to Microsporidia & Ascomycota". *Helix*. Vol.2: Pg 305-309.

[11] Harrison P.M. & P. Arosio., (1996). "The ferritins: molecular properties, iron storage function and cellular regulation". *Biochim Biophys Acta* 1275. Pg 161-203.

[12] M.A. Larkin., G. Blackshields., N.P. Brown., R. Chenna., P.A. McGettigan., H. McWilliam., F. Valentin., I.M. Wallace., A. Wilm., R. Lopez., J.D. Thompson., T.J. Gibson and D.G. Higgins., (2007) "Clustal W and Clustal X version 2.0". *Bioinformatics*. Vol. 23, Pg 2947–2948.