



COMPARATIVE ANALYSIS ON LARGE HYDROPHOBIC RESIDUES AND SMALL HYDROPHOBIC RESIDUES IN DIFFERENT ORGANISMS

VINOBHA C.S.¹ AND RAJASEKARAN E.^{2*}

¹Department of Bioinformatics, School of Bioengineering, Faculty of Engineering & Technology, SRM University, Kattankulathur, 603203, Kancheepuram Dt., Tamil Nadu, India.

²Department of Bioinformatics, School of Biotechnology and Health Sciences, Karunya University, Karunya Nagar, Coimbatore, 641114, Tamil Nadu, India

*Corresponding author. E-mail: sekaran@karunya.edu

Received: December 06, 2010; Accepted: December 20, 2010

Abstract- One of the major tasks carried by biologist today is to understand the nature of proteins. How this large protein molecule folds themselves into some form and carryout the prescribed biochemical reactions. Hydrophobic interaction is the dominant force towards this task. To understand this interaction, a simple statistical analysis on the contribution of hydrophobic residues was carried out. Large Hydrophobic Residues (LHR) such as Phenylalanine (F), Isoleucine (I), Leucine (L), Methionine (M) and Valine (V) – (FILMV) as well as small hydrophobic residues (SHR) Glycine (G), Alanine (A), Proline (P), Cysteine (C) and Tryptophan (W) - (GAPCW) were studied in all proteins of given organisms. The organisms include *Homo sapiens*, *Macaca Mullatta*, *Pan troglodytes*, *Canis familiaris*, *Gallus gallus*, *Mus musculus*, *Rattus norvegicus*, *Bos taurus*, *Drosophila melonogaster*, *Monodelphis domestica*, *Danio rerio*, *Stronglycentrolus purpuratus*, *Anopheles gambiae*, *Apis mellifera*, *Arabidopsis thaliana*, *Tribolium castaneum*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and *Caenorhabditis elegans*. It is observed that the protein prefers to have 27% large hydrophobic residues to maintain the required hydrophobicity. In animal, particularly in human, it is observed less. It is interesting to note that small hydrophobic residues balance this lack in number by a factor of 1:3. So is the reason why the length of the animal proteins increases. This new finding on the contribution of hydrophobic residues in protein stability will be discussed in detail.

Keywords: proteomics; hydrophobicity analysis; large hydrophobic residues; small hydrophobic residues; sequence analysis; bioinformatics;

Introduction

One of the major tasks carried by biologist today is to understand the nature of proteins. How this large protein molecule folds themselves into some form and carryout the prescribed biochemical reactions? Hydrophobic interaction is the dominant force towards this task [1-5]. To understand this interaction, several analyses on the contribution of hydrophobicity were carried out earlier. Large Hydrophobic Residues (LHR) such as Phenylalanine (F), Isoleucine (I), Leucine (L), Methionine (M) and Valine (V) are the major contributors towards this hydrophobic force. Sometimes the Small Hydrophobic Residues (SHR) such as Glycine (G), Alanine (A), Proline (P), Cysteine (C) and Tryptophan (W) (GAPCW) are even balancing these LHR wherever required. This was the main focus of this work. Particularly the LHR and GAPCW were studied statistically in different organisms including *Homo sapiens*,

Macaca mullatta, *Pan troglodytes*, *Canis familiaris*, *Gallus gallus*, *Mus musculus*, *Rattus norvegicus*, *Bos taurus*, *Drosophila melonogaster*, *Monodelphis domestica*, *Danio rerio*, *Stronglycentrolus purpuratus*, *Anopheles gambiae*, *Apis mellifera*, *Arabidopsis thaliana*, *Tribolium castaneum*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and *Caenorhabditis elegans*.

Methodology

The complete set of protein sequences of the above said organisms were downloaded from NCBI database. A set of statistical analysis has been carried on the amount of LHR and GAPCW. Overall percentage of LHR and GAPCW were computed using program written in 'C'. The results were tabulated and compared. A graph has been created for better comparison and discussed.

Results and Discussion

The percentage of Large Hydrophobic Residues (LHR) and Small Hydrophobic Residues (GAPCW) were calculated individually for all the organisms and given in the Table 1. Large hydrophobic residues meant that it contributed more to the hydrophobicity and vice versa. *Homo sapiens* have low percentage of LHR and high percentage of GAPCW, but the percentage value was similar. In *Ceanorhabditis elegans*, the percentage of LHR was high and the percentage of GAPCW was low. In all other organisms, when the fraction of LHR increased, the fraction of GAPCW decreased and vice versa.

Figure 1 shows the comparison between the large hydrophobic and small hydrophobic residues. The percentage of LHR when compared to percentage of GAPCW is 1:3. To balance the hydrophobicity in the protein, the percentage of LHR increased when the percentage of GAPCW decreased and vice versa. *Homo sapiens*' protein is lengthier than the other organisms and also than the plant and yeast proteins. Increases of LHR were only in animals and not in yeast or plant. The order of LHR increased from animal, plant to yeast whereas the order of GAPCW decreased from yeast, plant to animal.

The fraction of hydrophobicity of *Homo sapiens*, *Macaca mulatto* and *Pan troglodytes* were closely related. Next to that the fraction of hydrophobicity of *Canis familiaris*, *Gallus gallus*, *Mus musculus*, *Rattus norvegicus*, *Bos taurus* were closely related. Then *Drosophila melanogaster*, *Monodelphis domestica*, *Danio rerio* were also closely related. Similarly, *Strongylocentrolus purpuratus*, *Anopheles gambiae* and *Apis mellifera* were closely related followed by *Arabidopsis thaliana* and *Tribolium castaneum*. Even *Saccharomyces cerevisiae*, *Schizosaccharomyce pombe* and *Ceanorhabditis elegans* were also closely related. Zebra fish and *Bos taurus* had better protein than *Homo sapiens*. *Arabidopsis thaliana* has the perfect protein when compared to other organisms' protein. The

percentages of LHR and GAPCW in animals was lesser and greater than that of the plant respectively. *Homo sapiens* had low percentage of LHR and high percentage of GAPCW, but the percentage value is similar. *Ceanorhabditis elegans* had high percentage of LHR and low percentage of GAPCW.

Conclusion

The percentage of LHR was decreased in proteins during evolution. To balance this loss in hydrophobic component, other hydrophobic residues, GAPCW was added to the protein. It seems that *Arabidopsis thaliana* had proper amount of these hydrophobic residues compared to other organisms. *Homo sapiens* had low percentage of LHR and high percentage of GAPCW, but maintaining the hydrophobicity. The percentage of LHR in animals was generally lesser than that of the plant. In other terms the percentage of GAPCW in animals was greater than that of the plant. Generally the loss of one LHR is replaced with three GAPCW.

References

- [1] Rajasekaran E., Rajadurai M., Vinobha C.S., Senthil R. (2008) *J. Comput. Intelli. in Bioinfo.* 1(2-3):115-118.
- [2] Rajasekaran E., Vinobha C.S., Vijayasarathy M., Senthil R., Sankarganesh P. (2009) *IEEE, IACSIT-SC.* 452-453.
- [3] Vinobha C.S., Rajadurai M., Rajasekaran E. (2009) *J. Natural Sci. and Tech.* 1:40-41.
- [4] Vinobha C.S., Rajasekaran E. and Rajadurai M. (2010) *Int. J. Bioinfo. Res.* 2(2):7-11.
- [5] Vinobha C.S., Rajasekaran E. and Rajadurai M. (2010) *J. Comput. Intelli. in Bioinfo.* 3(1):71-77.

Table 1- Percentage of large hydrophobic residues (FILMV) and small hydrophobic residues (GAPCW) in different organisms

ORGANISMS	LHR	GAPCW
<i>Homo sapiens</i>	25.4	25.45
<i>Macaca mullatta</i>	26.11	24.15
<i>Pan troglodytes</i>	26.13	23.98
<i>Canis familiaris</i>	26.24	23.74
<i>Gallus gallus</i>	26.28	24.16
<i>Mus musculus</i>	26.45	23.42
<i>Rattus norvegicus</i>	26.53	23.36
<i>Bos Taurus</i>	26.65	23.75
<i>Drosophila melanogaster</i>	26.85	20.64
<i>Monodelphis domestica</i>	26.86	23.81
<i>Danio rerio</i>	26.89	21.33
<i>Stronglycentrolus purpuratus</i>	27.01	21.68
<i>Anopheles gambiae</i>	27.06	22.49
<i>Apis mellifera</i>	27.9	19.69
<i>Arabidopsis thaliana</i>	28.16	20.94
<i>Tribolium castaneum</i>	28.24	20.36
<i>Saccharomyces cerevisiae</i>	28.25	17.7
<i>Schizosaccharomyces pombe</i>	28.67	18.75
<i>Caenorhabditis elegans</i>	28.84	19.99

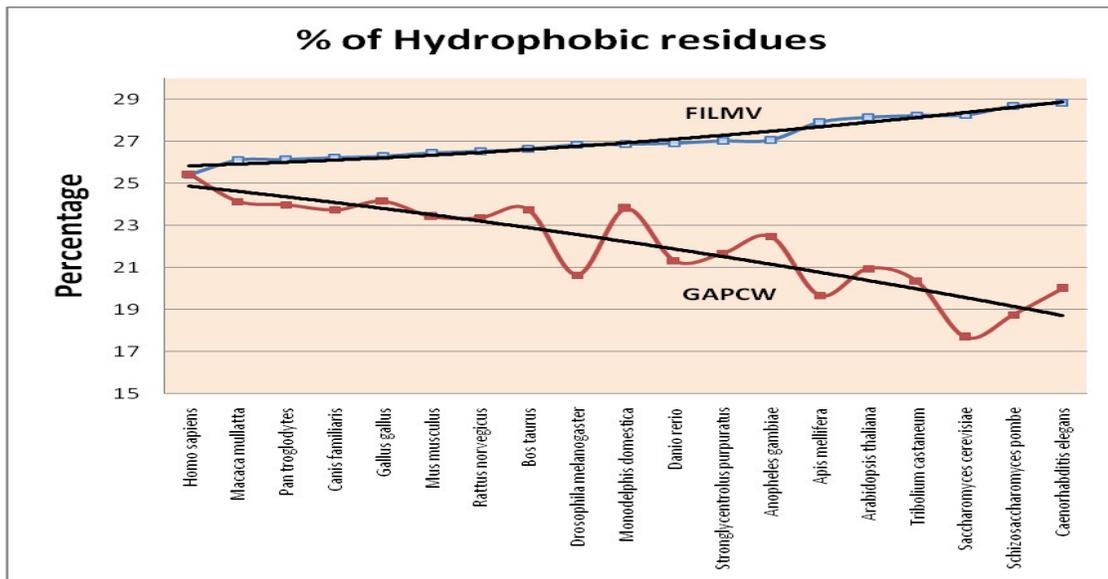


Fig. 1-Percentage of large hydrophobic residues (FILMV) and small hydrophobic residues (GAPCW) in different organisms