Journal of Computational Intelligence in Bioinformatics.
ISSN: 0973-385X Volume 1 Number 2-3 (2008) pp. 115–118

http://www.ripublication.com/jcib.htm

Are the proteins being hydrated during evolution?

**Ekambaram Rajasekaran[1], Maruthamuthu Rajadurai[2],
Chinnaiah Swaminathan Vinobha [3*] and R. Senthil[4]**

*[1]Department of Biotechnology, Periyar Maniammai University,
Thanjavur – 613403, Tamil Nadu, India.
[2]PG Department of Biotechnology and Bioinformatics, Bishop Heber College,
Tiruchirappalli-17, Tamilnadu, India
[3]Department of Bioinformatics, Ponnaiyah Ramajayam College,
Thanjavur – 614904, Tamil Nadu, India.
[4]Department of Bioinformatics, Marudu Pandiyar College,
Thanjavur – 620024, Tamil Nadu, India.
Corresponding author Email id: csvinobha@gmail.com

Abstract

Proteins prefer to have 31.44 % of carbon and 27% of large hydrophobic residues in its structure for stability. This paper analyzes the variation of these large hydrophobic residues in different species and its relevance to protein hydration, stability and disease. Protein sequences of 20 different living systems, including human, chimpanzee, dog, cow, mouse, rat, chicken, fruitfly, zebrafish, C.elegan, mosquito, honeybee, beetle, yeast, *S.purpuratus, K.lactis, S.pombe and A.thaliana* have been analyzed. For comparison the Influenza virus and *P.falciparum* are studied. The results show that there is reduction of large hydrophobic residues during evolution that causes hydration in proteins. The length of protein sequences increases accordingly. The reduction is high in heterosexual animals than that of plant and fungi. The reduction is observed based on the food habits, sexual behavior and environmental factors of the living system. To conclude many proteins are being hydrated during evolution that cause disease. This is very much in human.

Keywords: Protein Length; Protein hydration; Protein Evolution; Protein Analysis; disease protein.

Introduction

Proteins are the working force in all living systems**[1].** These proteins evolved to have a defined structure and specified function **[2-4].** These proteins are translated from mRNA. There are attempts to understand these biological macromolecules [5, 6]. Recently, it is reported that the protein coding frames of mRNA sequences prefers to have definite amount of XTX (X= A, T, G or C) that corresponds to number of large hydrophobic residues **[7].** That is proteins prefer to have 27% of large hydrophobic residues in its structure for stability[7]. These large hydrophobic residues include Phenylalanine (F), Isoleucine (I), Leucine (L), Methionine (M) and Valine (V) - (FILMV). Hydrophobic interactions are the dominant force in all biochemical reactions. Carbon is the only element contributes towards interaction that is dominated by FILMV. This paper analyzes the variation of FILMV in different species and its relevance to protein hydration, stability and disease.

Materials and Methods

The complete sets of protein sequences of human, chimpanzee, dog, cow, mouse, rat, chicken, fruitfly, zebrafish, C.elegan, mosquito, honeybee, beetle, yeast, *S.purpuratus, K.lactis, S.pombe, A.thaliana,* Influenza virus and *P.falciparum* are taken from ftp://ftp.ncbi.nlm.nih.gov/genomes/. The fraction of all amino acid in each species is calculated. The fraction of large hydrophobic residues, F, I, L, M and V are grouped separately. This number is obtained for all species and plotted fraction of FILMV versus species as shown in figure 1.

Results and Discussion

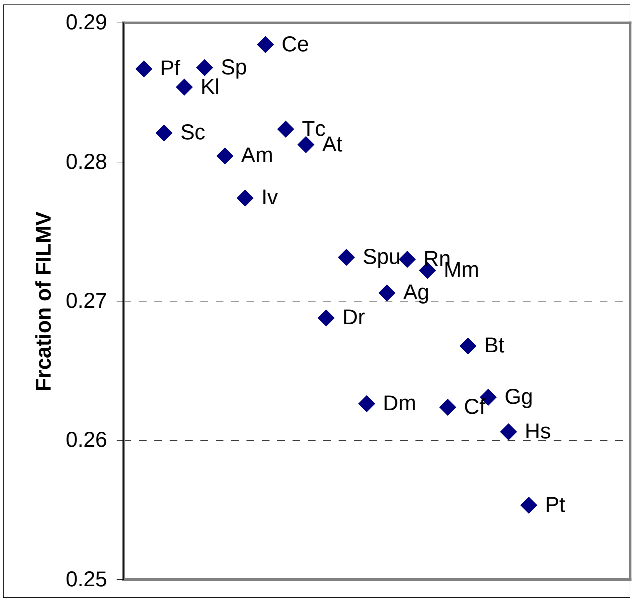

Figure 1: Reduction in amount of FILMV in heterosexual species.

The average fraction of FILMV in different species is shown in the figure. The species include human (Hs), chimpanzee (Pt), dog (Cf), cow (Bt), mouse (Mm), rat (Rn), chicken (Gg), fruitfly (Dm), zebrafish (Dr), C.elegan (Ce), mosquito (Ag), honeybee (Am), beetle (Tc), yeast (Sc), *S.purpuratus* (Spu), *K.lactis* (Kl), *S.pombe* (Sp), *A.thaliana* (At), Influenza virus (Iv) and *P.falciparum* (Pf). Plant and fungi show a higher amount of FLMV than the expected value of 27%. Most of the animals lack these large hydrophobic residues. Particularly human and chimpanzee show much lower amount of large hydrophobic residues. This mean that more number of proteins being hydrated during evolution. This is alarmingly important in human. This is reason why several proteins cause diseases.

During the evolution the length of the proteins increases in all species. It is phenomenal in heterosexuals because of mixing of DNA take place during reproduction. Apart from the heterosexual reproduction, the food habits and environmental factors also contribute towards the alteration in protein length. Small hydrophobic residues such as glycine, alanine, proline, cytocine and tryptophan compensate some of the loss in FILMV to some extent. A study is carried out on the extent of small hydrophobic residues contribution. It is about 3:1 ratio. Three small hydrophobic residues compensate one large hydrophobic residue. But it could have been 6:1 ratio and requires more detailed analysis.

This compensation by small hydrophobic residues increases the proteins length. Another study on the protein length has been carried out on species. It reveals the protein length increases considerably during evolution. Lengthier proteins found more in heterosexuals than that of plant or fungi. The number of protein sequences with length greater than 2000 amino acids are calculated. The species. *S.cerevisiae*, *K.lactis*, *S.pombe* and *A.thaliana* are having least number which posses a higher degree of order ness in the protein sequences. At the same time the animals such as human, chimpanzee, cow and dog show a higher number of lengthier proteins with lower degree of order. These changes also related to different food habit and environmental factors. The number of proteins with length greater than 5000 are counted for each species. There are no sequences having length greater than 5000 in *S.cerevisiae*, *K.lactis* and *S.pombe* and only two sequences in *A.thaliana* while dog found with 33 sequences. The rat and mouse are having 29 and 18 sequences respectively. It confirms the presence of lengthier proteins in heterosexuals compared to plant and fungi. That is the length of the proteins increases considerably in heterosexuals.

Conclusion

The conclusion is that there is a reduction in amount of large hydrophobic residues observed in animals. That is during evolution the proteins are being hydrated that cause genetic diseases. The length of the proteins also observed to be increases considerably during evolution. Lengthier proteins found more in heterosexuals than in plant or fungi. It is because of mixing of DNA take place during reproduction. In short many proteins are being hydrated during evolution. This is reason why several proteins cause disease. This is very much in human.

References

[1] Ponnuswamy, P.K., “Hydrophobic characteristics of folded proteins”, Prog Biophys Mol Biol., 59 (1993), 57-103.

[2] Venkatachalam, C.M., “Stereo chemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide unit”, Biopolymers, 6 (1968), 1425–1436.

[3] Galzitskaya, O.V., Ivankov, D.N. and Finkelstein, A.V. “Folding nuclei in proteins”. FEBS Letters, 489 (2001), 113–118.

[4] Frishman, D. and Argos, P. “Knowledge based protein secondary structure assignment”, Proteins 23 (1995), 566-579.

[5] Querol, E., Perez-Pons, J.A. and Mozo-Villarias, A., “Analysis of protein Conformational characteristics related to thermo stability”, Protein Eng., 9 (1996), 265-271.

[6] Huang, L.T., Saraboji, K., Ho, S. Y., Hwang, S.F., Ponnuswamy, M. N. and Gromiha, M. M. “Prediction of protein mutant stability using classification and regression tool”, Biophys Chem. 125 (2006), 462-170.

[7] Jayaraj, V., Suhanya, R., Vijayasarathy, M., Anandagopu, P. and Rajasekaran, E. “Role of Large Hydrophobic Residues in Proteins”, *Bioinformation* (under review).