Study on Carbon Distribution at Protein Regions of Disorder

K. Akila, N. Sneha, and E. Rajasekaran

***Abstract*—Hydrophobic interaction is the dominant force in proteins that arises due to carbon. The content and distribution of this carbon make a protein to fold in some form to have a specific function. Diseased proteins are expectedly differing from this carbon distribution pattern. This is taken as focus of this work by taking neurodegenerative protein, the SOD for study. The carbon distribution analysis program is able to identify the disordered region clearly. A long stretch of hydrophilic/hydrophobic regions with different combination of amino acid is considered as disordered regions. This disorder is due to unfolding or misfolding. The program developd for this purpose can further analysis the proteins for identification of mutational sites for stabilisation.**

***Index Terms*—Neurodegenerative disorder; SOD; card analysis; hydrophobicity; mutation; mutational site.**

I. Introduction

Hydrophobic interaction is the dominant force arises due to carbon in proteins. The content and distribution of this carbon make a protein to fold in some fashion to have a specific function. Disorder proteins are expectedly differing from this carbon distribution pattern. The disorders are mainly due to evolutionary concept. In particular the reduction of thymine in mRNAs causes reduction of large hydrophobic residues that alters number of large hydrophobic residues in proteins [1-2]. Due to this reduction stretch of sequences synthesized with less carbon content. These portions are hydrophilic in nature and disorder. To identify these regions, carbon is reference atom.

Recently it is reported that globular proteins prefer to have 31.45% of carbon for its stability and can be used as standard for carbon distribution analysis [3].

Allotment of carbon is responsible for disorders in proteins [4]. The carbon distribution analysis in disordered proteins is the focus of this work by taking neurodegenerative protein. There are many neurodegenerative disease targets like super oxide dismutase (SOD), tau protein, huntingenton and prion.The SOD is taken as case study as it is a simple and well studied protein. It is a metal binding protein must have adequate carbon to float.

Manuscript received December 28, 2011; revised February 13, 2012.

E. Rajasekaran and N. Sneha are with the Department of Bioinformatics, School of Biotechnology and Health Sciences, Karunya University, Karunya Nagar, Coimbatore-641114, Tamil Nadu, India (Tel.:+91-9524825876; e-mail:ersekaran@gmail.edu).

K. Akila is with the Bishop Heber College, Trichy-620 017.

II. Methodology

The amino acid sequence of SOD (P00441) is downloaded from the Uniprot database. The carbon distribution along the sequence was analysed using CARBANA program available online [3]. A outer length of 700 atoms are chosen here to have broader view on carbon distribution. The individual portion of the same sequence was subjected to Carbon distribution (CARd) analysis. This was about 10 amino acid length (~150 atoms). An inner length of 35 atoms is adapted. The CARd analysis was conducted at every 10 amino acid length with an interval of 5 amino acids (~75 atoms). This is to identify the inner length which has hydrophilic character or disorder based on carbon distribution.

III. Results and Discussion

Carbon distribution along the SOD sequence is analysed unsing CARBANA program. Average carbon content along the sequence is computed. A graph of carbon % at amino acid positions are shown in Fig. 1.

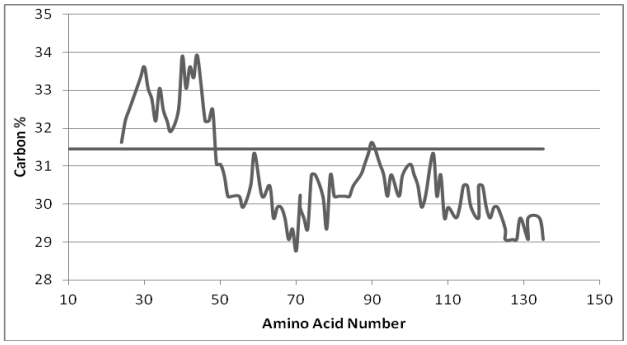
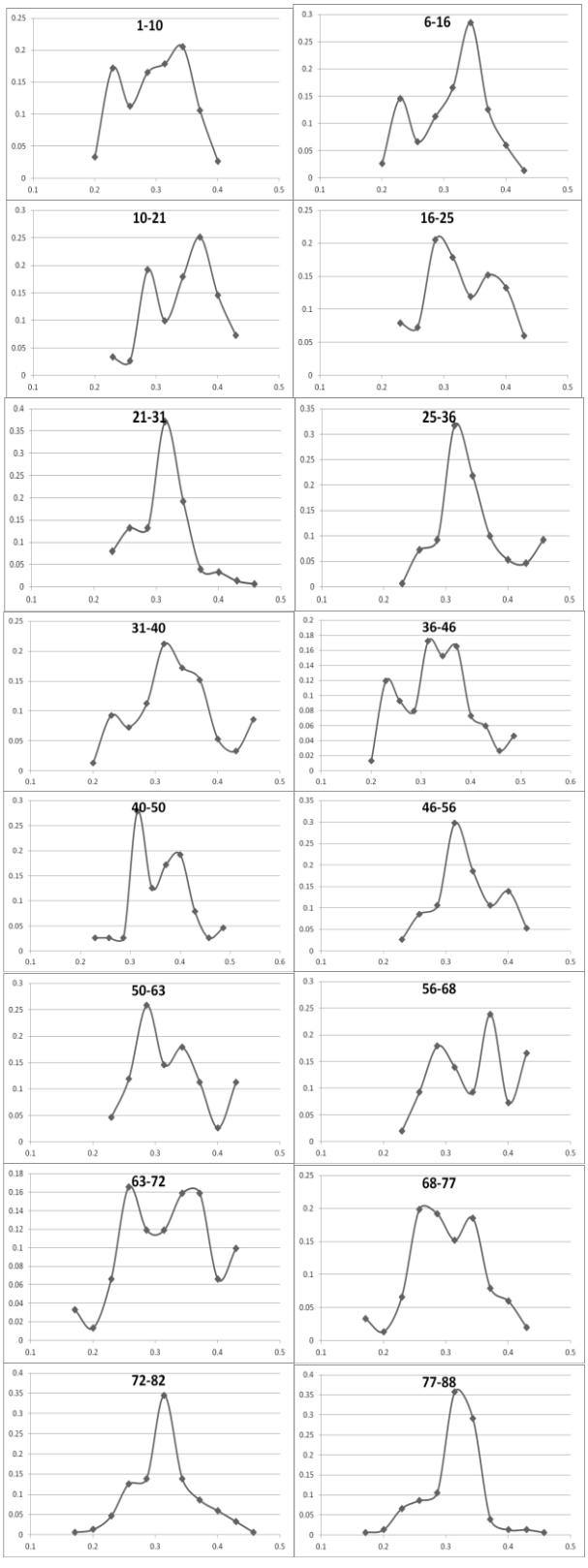


Fig. 1. Carbon content along the sequence of SOD protein. Points below the line indicate hydrophilic regions. Here from 149 to 154 are in hydrophilic region. This region is disorder and causes neurogenerative disease.

This result shows that there is a long hydrophilic region (149-254) due to which the protein unable to float to carry out the specified function. As SOD is a metal binding protein and must have adequate carbon to float which is not happening. Addition of further carbon in these region might improve a proteins activity. To do that a further CARd anlysis is carried out here by taking 10 amino acid (150 atoms of outer length) at a time and find out the inner length (35 atoms) that are having specified carbon content. The results are shown in fig 2.

Outer lengths 72-82, 77-88, 82-93, 88-99, 114-123, 119-129 and 141-152 are having normal distribution and

considered as stable regions. In fact the outer lengths 88-99, 114-123 and 129-141 are in perfect distribution. Amino acids in these lengths are arranged with perfect carbon distribution. These portion of sequences can be collected as carbon based patterns that can be utilised for protein stabilisation. That is mutation can be suggested based on these patterns. Outer lengths 93-104 and 99-110 are in order but hydrophobic in nature. Unless it active site, the carbon can be reduced for stability.



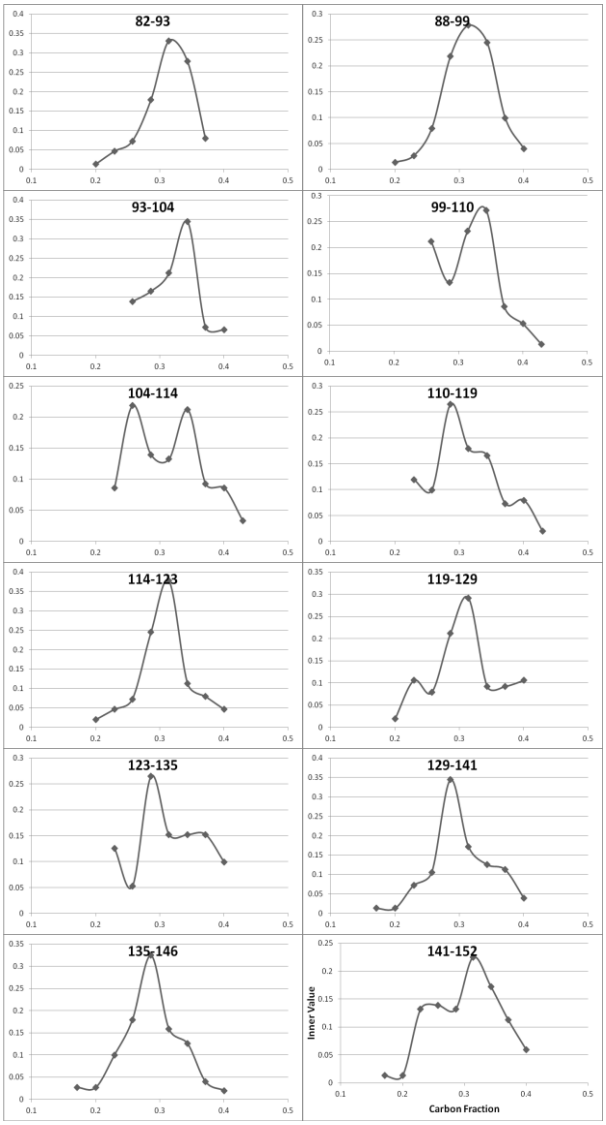


Fig. 2. CARd analysis on SOD protein with 10 amino acid (150 atoms) of outer length and 35 atoms of inner length.

Outer lengths 50-63, 110-119, 123-135, 129-141 and 135-146 are having normal carbon distribution but hydrophilic. These portions might unfold and cause disorder. Addition of carbon in these sites are suggested for better activity. Similarly the outer lengths 56-68, 63-72, 68-77 and 104-114 are hydrophilic plus disorder in carbon distribution. These portions needs to be given more attension for mutation. These sites of disorder are suggested for mutation for removing hydrophilicity and/or disorder of carbon distribution. Infact 56-68, 63-72 and 68-77 are metal binding sites. A careful mutation is required in such way that the activity is not lost. There are metal binding sites such as 46-56, 77-88, 82-93 and 114-123 are having normal carbon distribution. This can be adopted in the other sites as well.

IV. Conclusion

Carbana program is able to identify disorders in proteins. A long stretch of hydrophilic/hydrophobic regions considered as disordered regions. This disorder is due to unfolding or misfolding due to reduction of carbon or carbon

rich stretch. Carbon analysis on SOD identifies a long disordered hydrophilic region. A possible sites in these regions are suggeted for mutation using CARd analysis program.

REFERENCES

[1] E. Rajasekaran and P. Anandagopu, "Reduction of thymine in mRNA sequence of tumor protein," *J. Adv. Biotech.,* vol. 9, no. 7, pp. 9-10, 2010.

[2] E. Rajasekaran, M. Rajadurai, C. S. Vinobha, and R. Senthil, "Are the proteins being hydrated during evolution?" *J. Comput. Intelli. Bioinfo*. vol.1, no.2-3, pp.115-119, 2008.

[3] E. Rajasekaran, C. S. Vinobha, M. Vijayasarathy, R. Senthil, and P. Shankarganesh, "The nature of proteins," *IACSIT-SC, IEEE*, pp.452-453, 2009.

[4] E. Rajasekaran, K. Akila, and M. Vijayasarathy, "Allotment of carbon is responsible for disorders in proteins," *Bioinformation*, vol. 6, no. 8, pp. 291-292, 2011.