

## CARBON DISTRIBUTION IN PROTEIN LOCAL STRUCTURE DIRECT SUPEROXIDE DISMUTASE TO DISEASE WAY

**E. Rajasekaran\*, Sneha Nirmala John and J. Jannet Vennila**

*School of Biotechnology and Health Sciences, Karunya University, Coimbatore 641114, Tamil Nadu, India.*

**Abstract:** During the reaction involving superoxide dismutase (SOD), the oxidized Cu(II) form converts one superoxide molecule into oxygen while itself getting reduced to Cu(I) form. This reduced form of Cu(I) again converts another superoxide molecule into hydrogen peroxide and returns to oxidized Cu(II) form. Amyotrophic sclerosis is associated to instability in SOD and oxidative damage. Amyotrophic sclerosis mutants of human SOD form aggregates due to destabilization. Mostly these aggregates are due to hydrophobic interaction which comes from carbon present in the SOD. Some of these mutations found in amyotrophic sclerosis are studied here. Results reveal that most of these mutations alter local hydrophobic structures that are responsible for disorder. Carbon distribution analysis program is effective in identification and modification of mutational sites in biological sequences. This can be better exploited for gene therapy.

**Keywords:** Superoxide dismutase; carbon distribution study; CARd; mutation; metalloprotein; amyotrophic lateral sclerosis; FALS.

### Introduction

The copper containing superoxide dismutase (SOD) catalyses the reaction that convert superoxide into oxygen and hydrogen peroxide and prevent oxidative damage to cell. During the reaction, oxidized Cu(II) form converts one superoxide molecule into oxygen and becomes reduced to Cu(I) form in the enzyme. This reduced form of Cu(I) again converts another superoxide molecule into hydrogen peroxide and becomes oxidized to Cu(II) form back again (Rotilio *et al.*, 1972; Fielden *et al.*, 1974). Changes in Cu binding, instability in SOD and oxidative damage are associated with familial amyotrophic lateral sclerosis (FALS) disease (Furukawa and O'Halloran *et al.*, 2006). Demetallated SOD is pathogenic (Rakhit *et al.*, 2004). Why does it cause disease? Many point mutations in SOD reportedly cause amyotrophic sclerosis (Deng *et al.*, 1993).

These mutations understandably change the binding capability of Cu by altering the SOD character. ALS mutants of human superoxide dismutase form fibrous aggregates via framework destabilization (DiDonato *et al.*, 2003). Mostly these aggregates are due to hydrophobic interaction. Hydrophobic interaction is a force that comes from carbon atoms present in the SOD. Any point mutations in SOD, certainly varies the number of carbon. How significant are these changes in the overall function of the SOD? Carbon distribution analysis on several numbers of proteins show that single mutation can alter the protein function based on carbon distribution profile (Zelko *et al.*, 2002). Is it the same with point mutation reported for SOD? To address these issues, some of the SOD mutations found in amyotrophic sclerosis were studied here. Point mutations A4V, G37R, H43R, G85R, G93A and I113T in human SOD are analysed for carbon distribution using CARd program.

Corresponding Author: **E. Rajasekaran**

E-mail: [ersekaran@gmail.com](mailto:ersekaran@gmail.com)

Received: November 21, 2012

Accepted: November 25, 2012

Published: November 30, 2012

### Methodology

The protein structure and sequences were downloaded from PDB [[www.rcsb.org](http://www.rcsb.org)]. The PDB

identifier is 1SPD. The carbon distribution investigations were carried out using CARD program (Rajasekaran, 2012). First the entire sequence was scanned for carbon density using 35 inner and 700 outer lengths. The outer length (700) corresponds to 45 amino acid long stretch. So the carbon density is computed by these many amino acids at a time. This gives the hydrophobic and hydrophilic regions along the sequence in a broader term. Then the mutational sites are analysed for local carbon density. For this, different outer lengths were used. Lengths of 120, 210 and 255 were taken. These lengths correspond to 8, 14 and 17 amino acid long stretch.

### Results and Discussion

Carbon content along the SOD protein sequence was computed as shown in Figure 1. It is a statistical mean carbon content at different amino acid location. The outer window length of 700 atoms are chosen with mid point of reference amino acid. Carbon content in inner lengths are averaged statistically. The least amino acid number that can appear in the graph is 22 since it is the mid point of 45 amino acid long stretch taken for statistics. So the disease mutation A4V will not be visible with large outer length. Other mutational sites are visible and labelled in the graph. There appears hydrophilic region after 88<sup>th</sup> amino acid. Mostly these mutations occur at edges of carbon rich or carbon less stretch, similar to edges of secondary structures. Most of the time any mutations that occur at the edge of the secondary structure unit that alter function and leads to disease. These edge mutations totally change the adjoining secondary structure that leads to different structure and function. From the graph it is noted that the FALS mutations occur both in hydrophilic or hydrophobic regions. The SOD protein show a long stretch (88-154) of hydrophilic region. This region require an additional amount of carbon content. This need to be computed carefully before incorporation of amino acid in these region that will remain stable in the mRNA. If someone looks for binding sites for lead compounds (Antonyuk *et al.*, 2010), amino acid around 40 is ideal, as there is carbon rich site. From the crystal structure, this site appears to be on the surface and available for

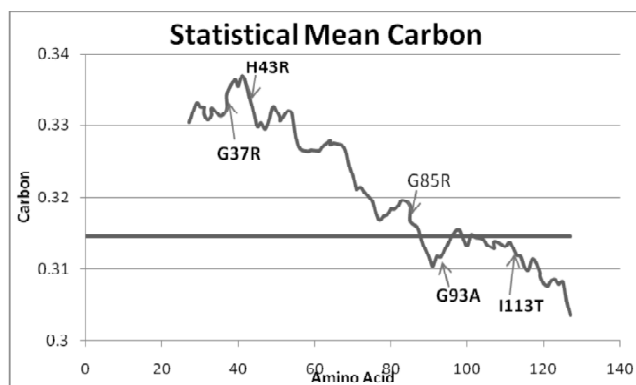


Figure 1: Carbon content along the native superoxide dismutase protein. The diseased mutational sites are labelled.

binding of aromatic compounds. Now lets look at the individual mutational site.

The carbon distribution profile for native and disease mutant (A4V) is given Figure 2. This site is in interface of protein-protein interaction. Any increase in hydrophobicity (Alanine by Valine) strengthens the dimerisation. Carbon distribution profile is calculated with outer lengths 120 and 135. Beyond this, the outer length is not possible to compute for higher lengths as the residue appear at 4<sup>th</sup> position. Looking at the distribution graph, the native protein shows a less normal distribution compared to mutant. Mutant seems stable compared to native. Cardoso 2002, argue that a instability in protein due to A4V mutation. But here we observe that stability on mutation. In this mutation, the cytosine in the second place of codon is substituted with thymine. Thymine content and distribution is important to produce adequate carbon in the protein. Amino acid change alone need not be the reason for disease. The mRNA expression and stability are also factors to be considered for disease.

Carbon distribution profile for native and mutant G37R is given Figure 3. The profiles are computed for outer lengths of 120, 210 and 255 which correspond to 8, 14 and 17 amino acid. At length 8, both native and mutant are away from normal distribution. Instability is observed locally. But at higher lengths, these profiles are normal. At these lengths the carbon distribution is well maintained. That is the amino acids are organized in such way that it give stability from carbon distribution point of view. In fact the mutant shows normal distribution over the

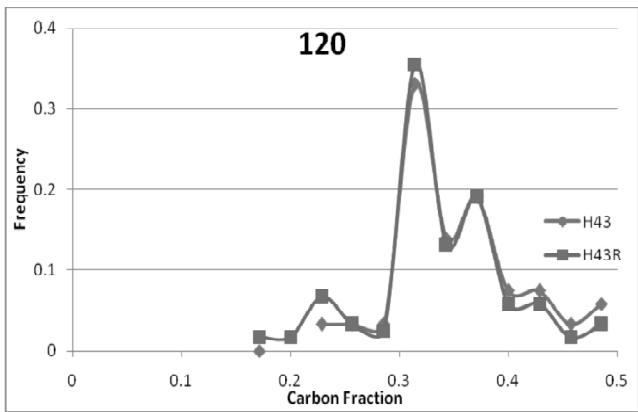
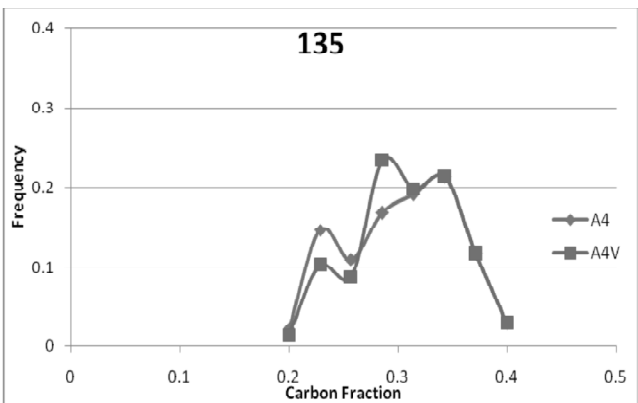
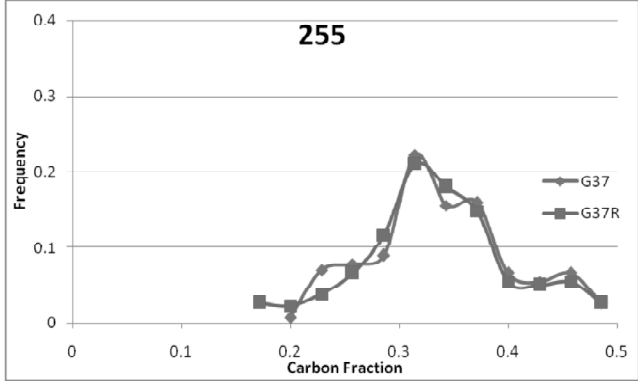
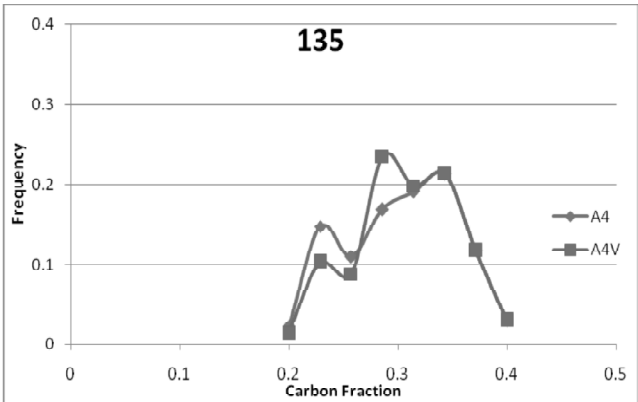
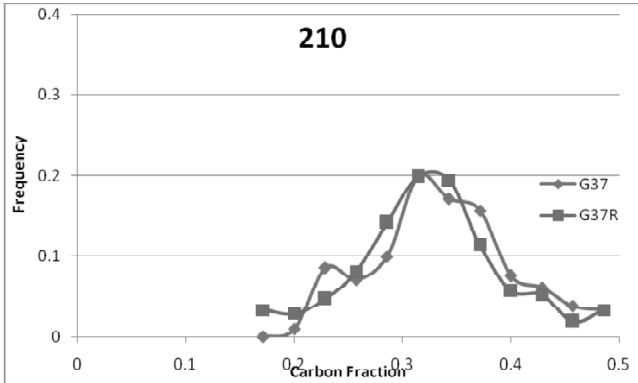
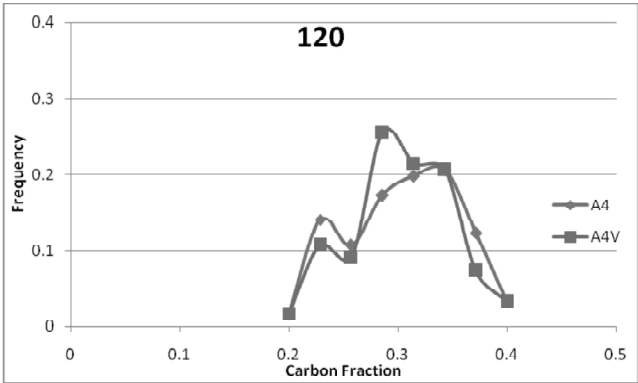


Figure 2: Carbon distribution profile for native and disease mutant (A4V) in different outer lengths.

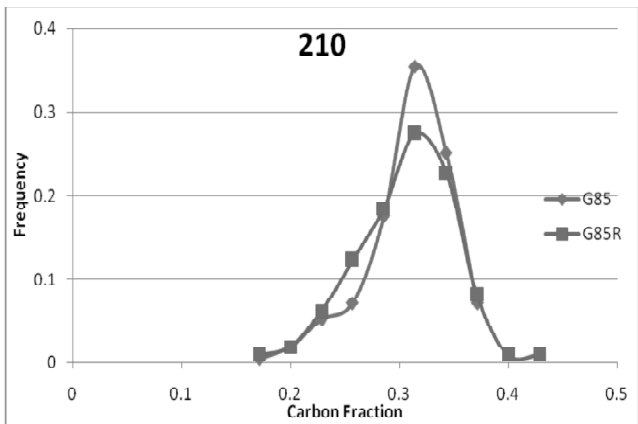
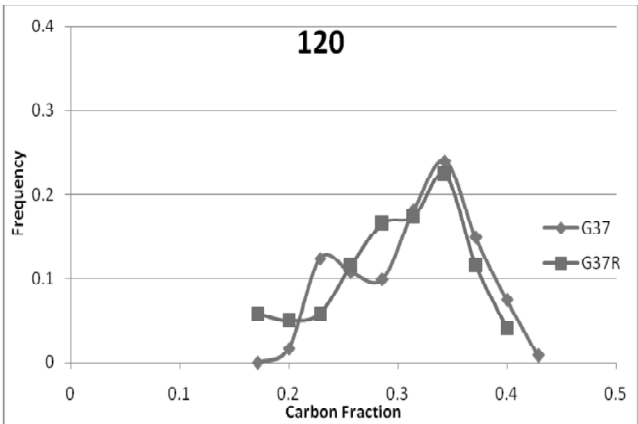


Figure 3: Carbon distribution profile for native and disease mutant (G37R) in different outer lengths.

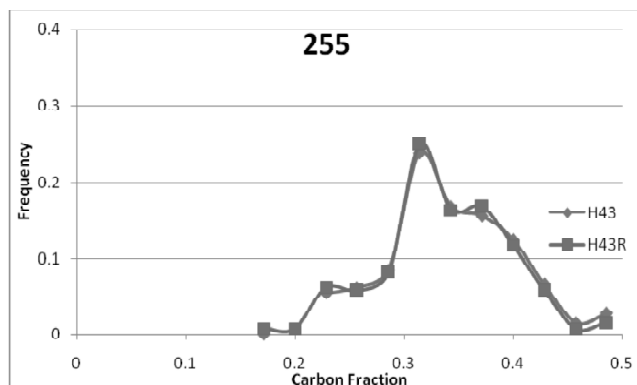


Figure 4: Carbon distribution profile for native and disease mutant (H43R) in different outer lengths.

native. The mutant increases the stability. Of course the charge needs to be taken into account here as neutral glycine was substituted with positively charged arginine. Since it is a metal binding protein, the charge is a matter, though it can be neutralized with ions and polar solvent.

Figure 4 shows the profile for native and disease mutant H43R. Surprisingly, there is no change in the carbon distribution profile in native and mutant and in all three lengths. Hydrophobic interaction is not perturbed. Considering the charge, both native and mutant residues are positively charged. So charge is not going to affect. Then what causes disease? One possible explanation may be the adenine (second nucleotide in codon) in native mRNA is substituted with guanine in the mutant. Competition between thymine less and rich is a matter. Finding cause at mRNA level may be appropriate.

The profile for native and mutant G85R (Cao *et al.*, 2008) is given in Figure 5. There is considerable change in carbon distribution between native and mutant in different lengths of calculations. Though normal distribution is observed in both forms, the native shows a narrow tall distribution curve compared mutant. The native shows stable over the mutant. This instability in hydrophobic interaction seemly causes disease.

Disease causing mutation G93A is given in figure 6. No change in carbon distribution profile except minor difference at lower length (8 amino acid long). At lower length, the mutant show more hydrophobic over the native one. From

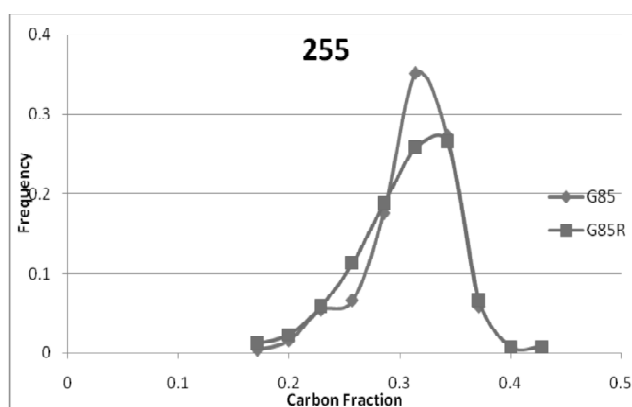
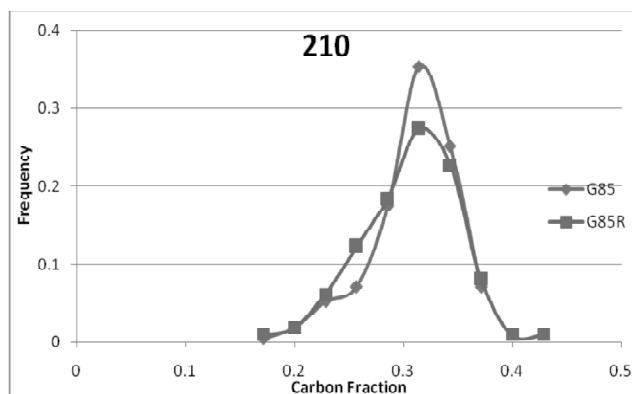
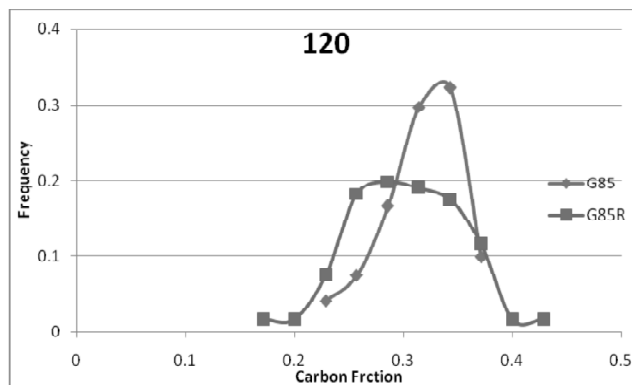
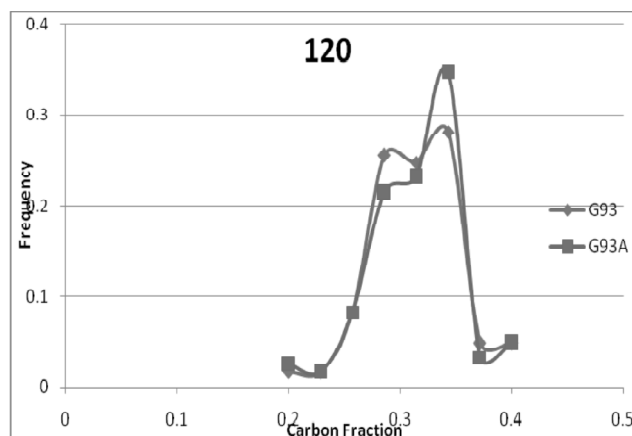


Figure 5: Carbon distribution profile for native and disease mutant (G85R) in different outer lengths.



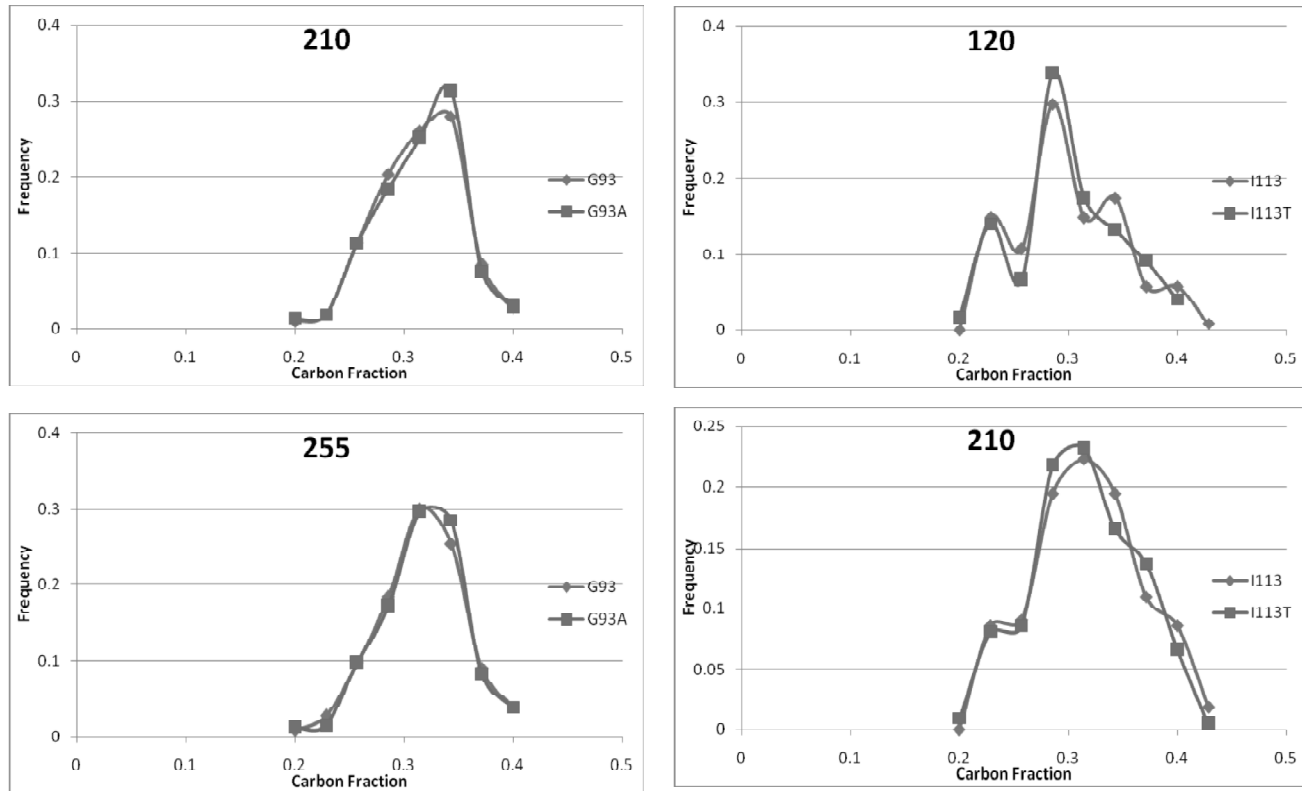


Figure 6: Carbon distribution profile for native and disease mutant (G93A) in different outer lengths.

Figure 1, one can observe that G93A occurs in hydrophilic region. Possibly the mutation can take it to hydrophobic stretch that can aggregate the protein causing unusual function.

Disease causing mutation I113T is given in Figure 7. Again there is no significant change in carbon distribution profile, except at lower length. At this lower length (8 amino acid long stretch) the mutant show less normal distribution curve compared to the one for native protein. So change in the local structure could affect the stability. This could cause disease. Also the mutational site (see Figure 1) is in hydrophilic region. Substitution of hydrophobic residue (I) by hydrophilic one (T) reduces the hydrophobic character further. This again could be reason for FALS. This result is in agreement with report by Hasnain and coworkers (Hough *et al.*, 2004).

Thymine distribution in different frames is required to produce stable and fully functional proteins. Superoxide dismutase is one of the finest examples of protein instability that comes from back strand frame 4 content.

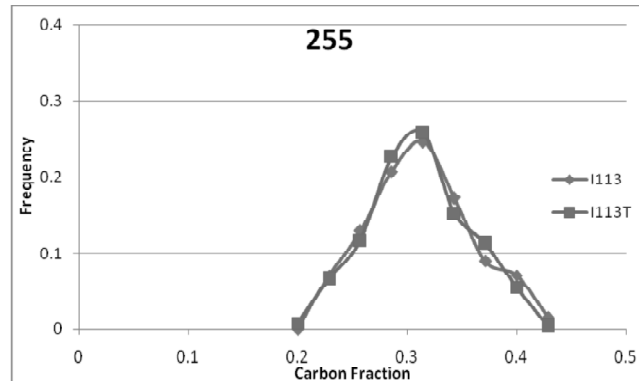


Figure 7: Carbon distribution profile for native and disease mutant (I113T) in different outer lengths.

## Conclusion

Carbon distribution analysis has been carried out on FALS mutants of SOD. Though there are several point mutations reported, mostly the disease is due to change in hydrophobic nature of local structure. Often the gene expression and the associated mRNA stabilities are factor causing FALS disease. One has to carry out a serious thymine distribution analysis to understand the disease. Though there is change in charge on disease causing mutation, it can be neutralized by ions and polar solvent present in the biological condition. Some of the disease mutation stabilizes

the proteins while some do not. From carbon distribution point of view, destabilization occurs at local structures. Carbon distribution analysis program can be better exploited for this kind of analysis and modification of protein for gene therapy.

Many mutations in superoxide dismutase (SOD) cause FALS. It is mainly due to changes in the local architectural integrity. These changes promote the formation of filamentous aggregates. This study provides molecular level understanding for FALS disease. Further thymine distribution in SOD mRNAs may be another cause of disease which requires deeper understanding of frame shift and translation.

### Abbreviations

SOD, Superoxide Dismutase; CARd, Carbon Distribution Analysis Program; FALS, Familial Amyotrophic Lateral Sclerosis; PDB, Protein Data Bank.

### References

- Antonyuk, S.V., Strange, R.W., and Hasnain, S.S. (2010) Structural discovery of small molecule binding sites in Cu-Zn human superoxide dismutase familial amyotrophic lateral sclerosis mutants provides insights for lead optimization. *J. Med. Chem.* 53, 1402-1406.
- Cao, X., Antonyuk, S. V., Seetharaman, S. V., Whitson, L. J., Taylor, A. B., Holloway, S. P., Strange, R. W., Doucette, P. A., Valentine, J. S., Tiwari, A., Hayward, L. J., Padua, S., Cohlberg, J. A., Hasnain, S. S., and Hart, P. J. (2008) Structures of the G85R variant of SOD1 in familial amyotrophic lateral sclerosis. *J. Biol. Chem.* 283, 16169-16177.
- Didonato, M., Craig, L., Huff, M.E., Thayer, M.M., Cardoso, R.M., Kassmann, C.J., Lo, T.P., Bruns, C.K., Powers, E.T., Kelly, J.W., Getzoff, E.D., and Tainer, J.A. (2003) ALS mutants of human superoxide dismutase form fibrous aggregates via framework destabilization. *Mol. Biol.* 332, 601-615.
- Fielden, E.M., Roberts, P.B., Bray, R.C., Lowe, D.J., Mautner, G.N., Rotilio, G., and Calabrese, L. (1974) Mechanism of action of superoxide dismutase from pulse radiolysis and electron paramagnetic resonance. Evidence that only half the active sites function in catalysis. *Biochem. J.* 139, 49-60.
- Gagliardi, S., Cova, E., Davin, A., Guareschi, S., Abel, K., Alvisi, E., Laforenza, U., Ghidoni, R., Cashman, JR, Ceroni, M., and Cereda, C. (2010) SOD1 mRNA expression in sporadic amyotrophic lateral sclerosis. *Neurobiol. Dis.* 39, 198-203.
- Hough M.A., Grossmann J.G., Antonyuk S.V., Strange R.W., Doucette P.A., Rodriguez J.A., Whitson L.J., Hart P.J., Hayward L.J., Valentine J.S., and Hasnain S.S. (2004) Dimer destabilization in superoxide dismutase may result in disease-causing properties: Structures of motor neuron disease mutants, *Proc. Natl. Acad. Sci. USA*, 101, 5976-5981.
- Rajasekaran, E. (2012) CARd: Carbon distribution analysis program for protein sequences. *Bioinformatics*, 8, 508-512.
- Rakhit, R., Crow, J.P., Lepock, J.R., Kondejewski, L.H., Cashman, N.R., and Chakrabartty, A. (2004) Monomeric Cu,Zn-superoxide dismutase is a common misfolding intermediate in the oxidation models of sporadic and familial amyotrophic lateral sclerosis. *J. Biol. Chem.* 279, 15499-504.
- Rotilio, G., Morpurgo, L., Giovagnoli, C., Calabrese, L., and Mondovì, B. (1972) Studies of the metal sites of copper proteins. Symmetry of copper in bovine superoxide dismutase and its functional significance. *Biochemistry*, 11, 2187-2192.
- Zelko, I.N., Mariani, T.J., and Folz, R.J. (2002) Superoxide dismutase multigene family: a comparison of the CuZn-SOD (SOD1), Mn-SOD (SOD2), and EC-SOD (SOD3) gene structures, evolution and expression. *Free Radic. Biol. Med.* 33, 337-49.