Carbon distribution to toxic effect in toxin proteins

Kannaiyan Akila[2], Rajendran Kaliaperumal[2] & Ekambaram Rajasekaran[1]*

[1]Department of Bioinformatics, School of Biotechnology and Health Sciences, Karunya University, Karunya Nagar, Coimbatore – 641114, Tamil Nadu, India; [2]Department of Biotechnology, Periyar Maniammai University, Thanjavur – 613403, Tamil Nadu, India; Ekambaram Rajasekaran – Email: ersekaran@gmail.com; *Corresponding author

Abstract:
The role of hydrophobic force in biological function through the formation of several local macro-molecular structures is evident. Carbon is the element that contributes to biological function in living systems. We show that carbon distribution is related to protein activity using an example. The carbon distribution profile is foreseen to help undestand unfolded and misfolded regions of protein structures. The carbon distribution profile in a toxin protein that is found associated with the toxic shock syndrome is described in this study. The carbon profile provides insight to the association of specific residues responsible for toxicity.

Keywords: carbon distribution, mutational study, hydrophobic interation, toxic protein, mitogenicity, carbon profile

Background:
Microorganisms are the source of toxins having mutation capabilities. These toxins enter the human and animal body through food, water, air and by other physical contacts. The immune system is involved in neutralizing these toxins by its defense mechanism during infection or disease progression. Infectious organisms produce different types of toxins like exotoxins, endotoxins, interotoxins, neurotoxins and sometimes enzymes with toxicity [1]. These toxins are often proteins secreted by the bacterial cell (for example, the anthrax toxin is an exotoxin produced by *Bacillus anthracis*) which are regularly destroyed by heat. It is of interest to understand carbon distribution to its toxic function in these proteins. The toxic shock syndrome toxin (TSST) is a super antigen expressed by *Staphylococcus aureus* (UniProtKB number P06886 and PDB ID 1QIL). The disease linked to this protein is characterized by fever, hypotension, multiple-organ disorder, and desquamation (skin peeling). TSST is 25% hydrophobic and stable in aqueous and basic solutions. Investigations produce data on the terminus responsible for toxic activity. Super antigens bind non-specifically and activate populations of T cells by forming H-bonds. Studies on this super antigen indicate that the binding domain lies at the top of the back side of this toxin, though the complete interaction remains to be determined. There have also been indications that the binding site is mapped to the major groove of the central alpha helix or the short amino terminal

alpha helix. Residues in the beta motif of TSST are known to interact with the invariant region of the alpha chain of other proteins. The role of carbon in TSST for protein disorder and activity is of interest [2]. Here, we describe the carbon distribution in TSST. These studies will provide insights to the optimal design and synthesis proteins with proper distribution of amino acids. It should be noted that TSST is taken as an example for carbon content analysis in this study. It should be also noted that thymine distribution in mRNAs is available in related studies [3].

Methodology:
We downloaded the protein sequence for TSST from PDB (www.rcsb.org/pdb/) with the PDB ID 1QIL. This entry describes three monomer chains. The carbon distribution in this protein sequence is analyzed using the CARd program as described elsewhere [4]. The carbon content along the sequence is calculated with outer and inner length of 700 and 35 atoms, respectively. The mean and statistical mean along the sequence is subsequently computed. The number of carbon atoms divided by total atoms (700 in this case) is the mean carbon content for a stretch of 700 atoms in a row. The statistical mean of carbon content is computed using CARd as described elsewhere [4]. The mutational site (135) of the protein is also analyzed for carbon distribution using CARd with an outer length of 270 atoms (~17 aa) and an inner length of 35 atoms.
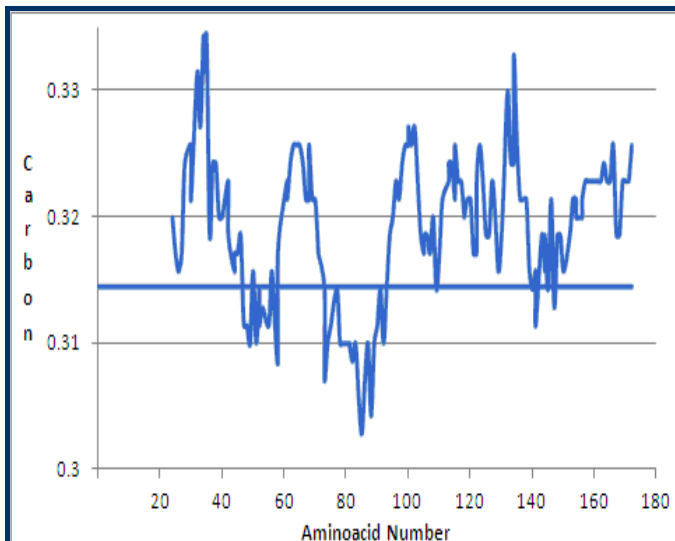
BIOINFORMATION

Figure 1: Hydrophobic index along TSST. The fraction of carbon is plotted against amino acid residue numbers in sequence. The region above the line corresponding to 0.3145 is considered as the carbon rich regions. It should be noted that the mutational site 135 is in the carbon rich region.

Discussion:

The TSST protein (UniProtKB number P06886) is 234 amino acid residues in length. The first 40 residues form the signal peptide. The remaining 194 amino acids residues in the sequence are subjected to carbon distribution profile analysis using an improved version of the CARd program. The carbon distribution in TSST is shown in **Figure 1**. The hydrophobic regions are 1-45, 58-73, 93-139 and 145-194 in the sequence. Each pocket is about 45 residues long. The TSST contains excess of carbon content as most of the regions are above the line value of 0.3145. The excess carbon indicates protein disorder. The mutational site H135 is also in the carbon rich region. It should also be observed that both the terminals are hydrophobic in nature. The secondary structure analysis reveals that most of them are beta-barrel. The mutational site (135) is present in the hydrophobic stretch. The mutation H135A results in the loss of T-cell mitogenicity and toxicity in experimental animals **[3].** The carbon distribution profile reveals the loss in biological function of the protein. This shows the relationship of carbon role in toxicity of toxic proteins. The average carbon content (0.322), mean carbon (0.338), median (0.343) and mode (0.314) of the distribution curve at site 135 in native types are reported. However, in the mutant protein, a carbon content of 0.326 and mean carbon of 0.336 is observed. The median and mode remains same. The comparison **(Figure 2)** of carbon content and mean carbon reveals that the mean carbon is always higher than the average carbon in the native protein. This is reverse at site 135. The mean carbon is lower than the average carbon at position 134. This is not true in the mutant protein. This is also

true at residue position 35 as a potential site for protein stabilization.
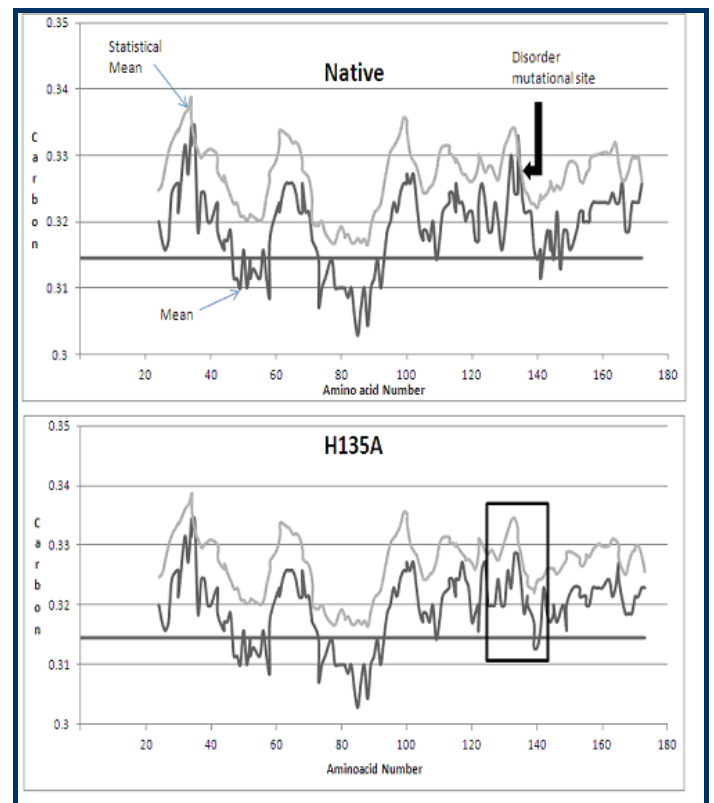

Figure 2: Carbon content (mean and statistical mean) in the toxin protein. The statistical mean is greater than the mean carbon in the distribution. It should be noted that at the mutational site (135) the trend is seen reverse while it is normal thereafter.

Conclusion:

The carbon distribution in a toxic protein is described in relation to its function. We show that carbon distributions in the local structure have relation to protein activity. The carbon content is higher at the active site which is destabilized by mutation causing protein disorder.

References:

[1] Papageorgiou AC *et al. Protein Sci.* 1996 **5**: 1737 [PMID: 8844860]

[2] Rajasekaran E, *Bioinformation*. 2012 **8**: 508 [PMID: 22829720]

[3] Rajasekaran E *et al. Bioinformation.* 2011 **6**: 291 [PMID: 21769187]

[4] Rajasekaran E & Akila K, *Int J Biosci Biochem Bioinfo.* 2011 **1**: 249