



Category: Bioinformatics

# Structural Analysis of Low Complexity Regions of Proteins

**Manish Kumar, Bandana Kumari and Ravindra Kumar**

Department of Biophysics, University of Delhi South Campus, Benito Juarez Road, New Delhi 110021, INDIA

Presenting author: [manish@south.du.ac.in](mailto:manish@south.du.ac.in)

## Abstract

Low complexity regions (LCR) in a protein sequence are regions of biased composition. Despite well-established importance and abundance, compositional and structural properties of LCRs are poorly understood and their structural status as ordered or disordered is at best ambiguous. Often, LCRs are considered as a part of disordered protein segments which most likely do not form any secondary structure but exist as solvent-exposed, disordered coil. We have analyzed the secondary structure content and surface accessibility of a non-redundant dataset of Protein Data Bank proteins and found that unlike popular belief, LCRs might have secondary structures (mostly helix) and they might not always exist as highly accessible disordered region. We also observed that in a LCR, all constituting amino acids might have same secondary structure or there may be combination of different secondary structures. We also observed that proteins whose structures were determined by X-ray crystallography were found to possess ordered LCRs while those whose structure was determined by NMR possessed disordered LCRs. Consensus disorder prediction by DISOPRED, IUPred, and IsUnstruct also supported our inference. Trans-membrane (TM) region of proteins are highly dominated by  $\alpha$ -helices and relatively very few have  $\beta$ -sheets which may be a possible reason for predominant occurrence of  $\alpha$ -helices. But a very small fraction (<5%) of helices suggests that our observation was not biased due to presence of TM helices. Comparison of enrichment/depletion profile of disorder promoting amino acids for LCRs in ordered and disordered regions of proteins revealed that they have different enrichment and depletion patterns of amino acids. In order to see whether the proteins constituting whose structure is solved by NMR are those in which folding upon binding phenomena is reported; we did GO term enrichment analysis using DAVID (release 6.7) at default threshold. A major fraction was found to be involved in regulatory processes suggesting the possibility for presence of disordered stretches, which might acquire structure after binding with the appropriate ligand.

Our analysis suggests that the structural state of LCR depends on the overall environment of the protein in which it is present. It is not the exclusive property of the LCRs.

**Citation:** Kumar, M., Kumari, B. and Kumar, R. Structural Analysis of Low Complexity Regions of Proteins [Abstract]. In: Abstracts of the NGBT conference; Oct 02-04, 2017; Bhubaneswar, Odisha, India: *Can J Biotech*, Volume 1, Special Issue (Supplement), Page 219.  
<https://doi.org/10.24870/cjb.2017-a204>