Research Article **Bioinformatics**

International Journal of Pharma and Bio Sciences **ISSN 0975-6299**

PREDICTION OF ACTIVE SITES IN SOME CANCER RELATED PROTEINS BY CARD ANALYSIS

K.AKILA[1*], S.RAJASREE[2] AND R.SHANTHAKUMAR[2]

[1]Department of Biotechnology, Periyar Maniammai University, Vallam, Thanjavur, Tamil Nadu,
[2]CoRx Lifesciences and Pharmaceutical (CLAP) Private Limited, Trichy,Tamil Nadu.

ABSTRACT

Carbon is the element that contributes towards the hydrophobic effect and biological function of living systems. It is reported that globular proteins prefer to have 31.45% carbon for stability, which can be used as the standard for carbon measurement and comparison. Carbon distribution analysis has been carried out using the tool CARd, which can clearly identify the hydrophobic and hydrophilic regions along the sequence and can also pinpoint an amino acid which is causing instability. Cancer is one of the most feared diseases of the 20th century and spreading further with continuance and growing incidence in 21st century. The role of carbon along the protein sequence and amino acid composition of hydrophobic regions in selected cancer related proteins (BAK, PTEN, E6, p53, Retinoblastoma) were focused in this study. Carbon content and distribution were computed and the active sites identified were reported here. This study which dealt with the active site prediction through the analysis of carbon distribution in cancer related proteins promises significant applications in biological and pharmaceutical sciences such as drug design and genetic engineering.

KEYWORDS: Carbon, Cancer, BAK, PTEN, E6, p53, Retinoblastoma

K. AKILA
Department of Biotechnology, Periyar maniammai University,
Vallam, Thanjavur, Tamil Nadu.

***Corresponding author**

INTRODUCTION

Proteins are fundamental intracellular macromolecules and are the chief sources of physiological functions in all biological systems. Protein sequence analysis at residual level facilitates the finding of sequence similarities, hydropathy profiles, motifs, protein families and specific domains. By moving one step further down at atomic level, much elaborate analysis of the peptide sequences can be performed. The key elements in all amino acids are hydrogen, carbon, oxygen, nitrogen and sulphur. These elements are responsible for endowing the amino acids, the properties of hydrophobicity/hydrophilicity, which play an important role in protein interactions. The hydrophobic amino acids characteristically own greater number of carbon atoms as carbon is the main element, which contributes to hydrophobic interactions in proteins. Cancer is one of the most deadly diseases of the past century and diffuses further with persistence and mounting incidence in 21st century. There are more than 100 types of cancer, including breast cancer, skin cancer, lung cancer, colon cancer, prostate cancer, and lymphoma. Defective apoptosis (planned cell death) signifies the foremost causative factor in the development and progression of cancer. The ability of tumor cells to evade engagement of apoptosis can play a noteworthy role in their resistance to conventional therapeutic regimens. The current study focuses on proteins that are involved in cancer and intends to study the relationship of carbon distribution with protein structure and activity. It is reported that globular proteins prefer to have 31.45% of carbon for their stabilities and function[4]. This fraction of carbon content can be used as the thumb standard for measurements and comparisons of hydrophobic effects in different globular proteins. Using this kind of carbon scaling, one can pinpoint amino acids which are related to causing disease which in turn requires atomic level representation to acquire better results. The carbon distribution analysis of cancer proteins including BAK, PTEN, p53, e6 and retinoblastoma protein are the focus of this work. The BAK protein is a pro-apoptotic member of the Bcl-2 gene family, which is implicated in initiating apoptosis. Deregulation of the BAK gene is of concernin human gastrointestinal cancers demonstrating that the gene plays a part in the pathogenesis of some cancers[3]. Phosphatase and tensin homolog (PTEN) is a protein that, in humans, is encoded by the PTEN gene[7]. PTEN is one of the most commonly missing tumor suppressors in human cancer; in fact, up to 70% of men with prostate cancer are likely to have lost a copy of the PTEN gene at the time of diagnosis. During tumor development, mutations and deletions of PTEN take place that inactivate its enzymatic activity leading to augmented cell proliferation and decreased cell death. Recurrent genetic inactivation of PTEN occurs in glioblastoma, endometrial cancer, and prostate cancer; and reduced expression is found in many other tumor types such as lung and breast cancer. Moreover, PTEN mutation also leads to several inherited predispositions to cancer[1]. p53, the cellular tumor antigen p53 or phosphoprotein p53 or tumor suppressor p53 is a protein that in humans is encoded by the *TP53* gene. If the *TP53* gene is damaged, tumor suppression is drastically decreased. Certain pathogens can also influence the p53 protein that the *TP53* gene expresses. One such example, human papillomavirus (HPV), encodes a protein, E6, which binds to the p53 protein and inactivates it. The high-risk human papillomavirus (HPV) E6 proteins are constantly expressed in HPV-associated lesions and cancers. High-risk HPV E6 proteins form a complex with the E3 ubiquitin ligase UBE3A (E6AP) and the p53 tumor suppressor, targeting p53 proteasomal degradation[6]. The Retinoblastoma protein is a protein product of the RB1 gene. The retinoblastoma protein (abbreviated pRb, *RB* or *RB1*) is a tumor suppressor protein that is dysfunctional in several major cancers. Mutations in RB1 are considered to contribute to the development of bladder cancer, and these genetic changes may aid to predict whether tumors will grow rapidly and spread to other tissues[2].

METHODOLOGY

The five different cancer protein sequences were collected from National Centre for Biotechnological Information (NCBI) and Kyoto Encyclopedia of Genes and Genomes (KEGG). The carbon distribution in these proteins was computed using CARd program[4], which uses the principle of 31.45% of carbon[4]. The CARd program can be accessed online (www.rajasekaran.net.in/tools/carbana.html).

DATASETS

The sequence data for following cancer related proteins were retrieved from NCBI.

Table 1
List of proteins taken for CARd analysis

Name of the Protein	Accession Number	Source	Sequence Length
BAK	AAA74466	Homo sapiens	211
PTEN	AAD13528	Homo sapiens	403
E6	ACR25130	Human Papilloma Virus	158
p53	BAC16799	Homo sapiens	393
Retinoblastoma associated protein	NP_000312	Homo sapiens	928

RESULTS AND DISCUSSION

Carbon is the only element that contributes towards the dominant force, hydrophobic interaction in proteins. Proteins evolve based on carbon content and may influence the coding of genes. It is reported that proteins prefer to have a 31.45% of carbon for their stability[5]. The results of carbon distribution in the five different cancer proteins have been shown in figures 1-5. The BAK protein is a pro-apoptotic member of the Bcl-2 gene family which is involved in initiating apoptosis. BAK protein has higher carbon contents at region 75 to 165, which corresponds to misfolding region. Phosphatase and tensin homolog (PTEN) is a protein that, in humans, is encoded by the PTEN gene. A mutation of this gene is a step in the development of many cancers. PTEN protein has the higher carbon content at the (misfolding) regions 70-120 and 170-270. The high-risk human papillomavirus (HPV) E6 proteins are consistently expressed in HPV-associated lesions and cancers. E6 protein has a hydrophobic pattern with length more than 35-95 amino acids and a hydrophilic pattern with a length of 95-155 amino acids. p53 also known as cellular tumor antigen p53 or phosphoprotein p53 or tumor suppressor p53 is a protein that in humans is encoded by the *TP53* gene. Hydrophobic patterns at 20-120 and 220-370 have been observed in p53. The Retinoblastoma protein is a protein product of the RB1 gene. The retinoblastoma protein is a tumor suppressor protein that is dysfunctional in several major cancers. On contrary, retinoblastoma aasociated protein has abnormal distribution of carbon throughout the sequence.

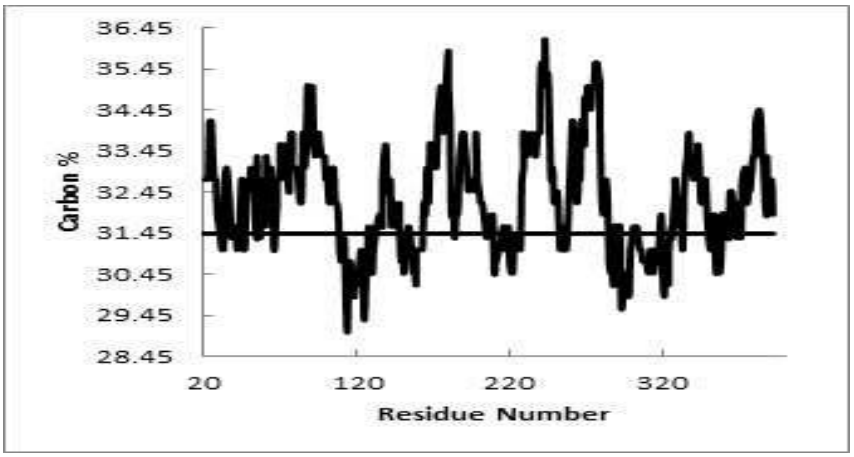

Figure 1
Carbon distribution of BAK protein

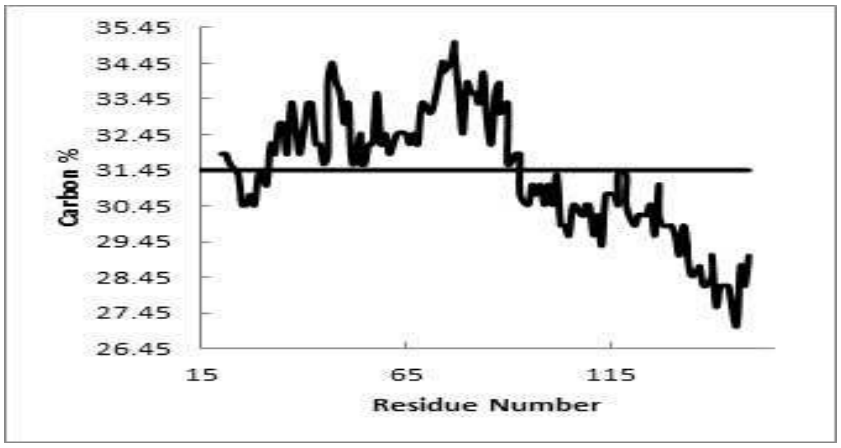

Figure 2
Carbon distribution of PTEN protein

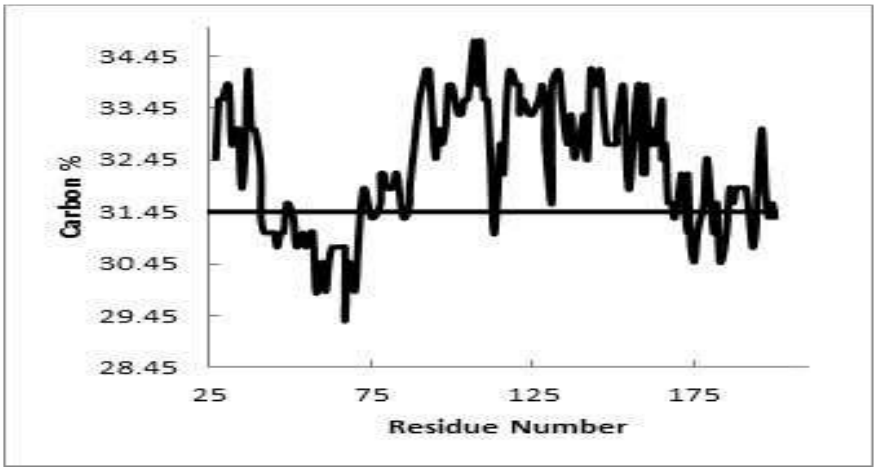

Figure 3
Carbon distribution of e6 protein of HPV16

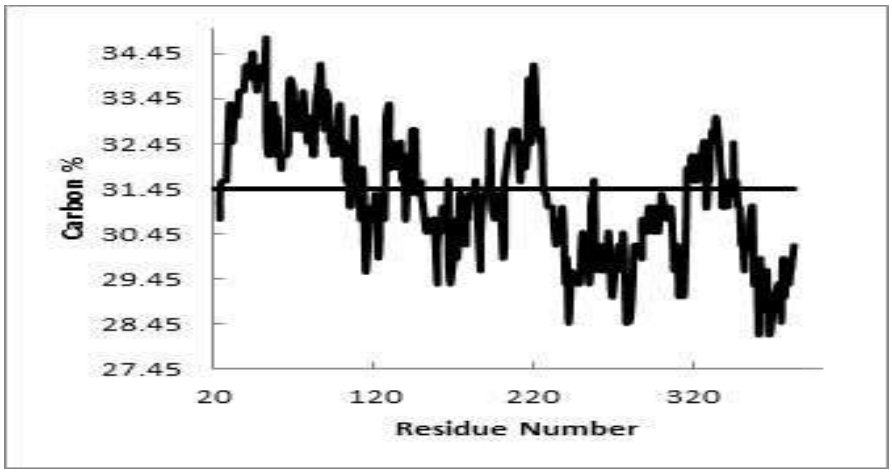

Figure 4
Carbon distribution of p53

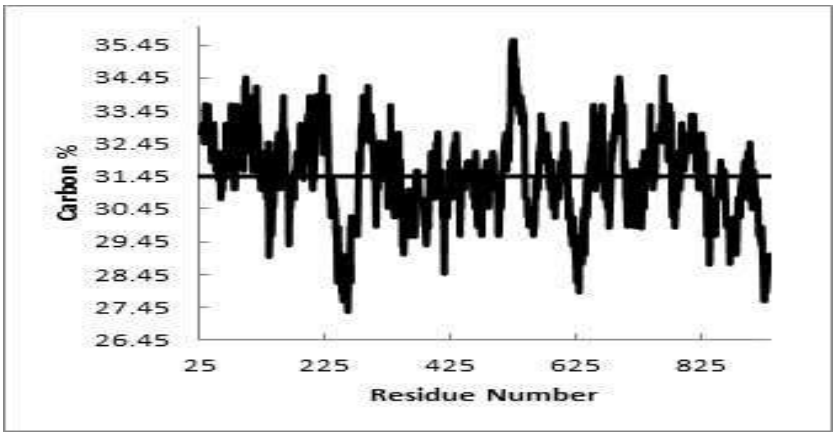

Figure 5
Carbon distribution of retinoblastoma-associated protein

The residues which show high carbon content are considered as active sites of the proteins. The active site of the protein has several significant applications in biological and pharmaceutical sciences such as drug design, genetic engineering, and diagnosis[8]. Hence, this study which has focused on the active site prediction through the analysis of carbon distribution in cancer related proteins promises new discoveries in pharmaceutical and clinical aspects.

CONCLUSION

Computational methods prove to be of great use in understanding the functionally important sites of proteins. The active sites in selected cancer related proteins were identified and reported here. The principle behind this calculation was that proteins prefer to have 31.45% carbon for stability. The carbon content analysis has thus paved way to arrive at the functional sites of cancer proteins studied. Thus, the carbon distribution study along the protein chain is validated as the most significant step towards understanding the biological features such as active or functional site prediction, which can provide possible approaches for the design of new drugs to overcome cancer.

REFERENCES

1. Chen Z, Trotman LC, Shaffer D, Lin HK, Dotan ZA, Niki M, Koutcher JA, Scher HI, Ludwig T, Gerald W. Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. Nat, 11:725–730,(2005).
2. Murphree AL, Benedict WF. Retinoblastoma: clues to human oncogenesis. Science,223:1028–1033, (1994).
3. Qiang-Song Tong, Li-Duan Zheng, Liang Wang, Jun Liu, and Wei Qian. BAK overexpression mediates p53-independent apoptosis inducing effects on human gastric cancer cells. BMC Cancer, 4: 33, (2004).
4. Rajasekaran E CARd: Carbon distribution analysis program for proteins sequences. Bioinformation, 8(11): 508-512, (2012).
5. Rajasekaran E. Vinobha CS. Vijayasarathy M. Senthil R. Shankarganesh P. The Nature of Proteins. IACSIT-SC IEEE, 452-453, (2009).
6. Scheffner M, Huibregtse JM, Vierstra RD, Howley PM. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. Cell, 75:495–505, (1993).
7. Steck PA, Pershouse MA, Jasser SA, Yung WK, Lin H, Ligon AH, Langford LA, Baumgard ML, Hattier T, Davis T, Frye C, Hu R, Swedlund B, Teng DH and Tavtigian SV. Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. Nat. Genet., 16: 64-67, (1997).
8. Senthil R, Sathish S, Jannet Vennila J and Rajasekaran E. Prediction of ligand binding site in globular proteins. Journal of Advanced Bioinformatics Applications and Research, 2;98-99, (2011).