J Biochem Tech (2012) 3(4): 306-308
ISSN: 0974-2328

Sequence analysis of the tumour suppressor protein p53 and its implications

Sultan Mohideen A K, Asrar Sheriff M

Received: 02 March 2011 / Received in revised form: 12 March 2011, Accepted: 12 March 2011, Published online: 20 February 2012,

Abstract

The p53 is a transcription factor encoded by the tumour suppressor gene TP53, involved in regulating the cell responses to DNA damage to conserve genomic stability. p53 protein sequence analysis was carried out to unravel the structural details of normal and mutant forms. The tumour suppressor protein p53 is mutated in more than 50% of invasive cancers and 30% of these mutations are found in six major hot spot codons located in its DNA binding core domain. This study was conducted to gain insights into normal and mutant variants of p53 and to understand their clinical implications in the etiology of cancer. pBLAST analysis was performed between normal and two mutant p53 protein sequences and various types of mutations like substitutions of amino acids were identified. Importance of insilico study on p53 as a tool for prediction and diagnosis of mutations in human cancers and its medical significance are discussed.

Key words: Sequence analysis, p53, mutation, substitution, human cancer.

Introduction

p53 has been described as the "Guardian of the Genome" referring to its role in preventing genome mutation (Read, 1999). TP53 gene acts as a tumour suppressor and when it is mutated, it is involved in a wide variety of human cancers including breast, lung, liver and colorectal (Hollstein et al. 1991). p53 is 393 amino acids long and has seven domains that are characteristic of transcriptional activators including an activation region at its terminus, a sequence -specific DNA binding region within its central protein and an oligomerization domain within its C– terminal region (Harms and Chen 2005). In addition to genes regulating cell cycle and cell death, TP53 can stimulate the transcription of genes encoding products that can affect both DNA synthesis and repair. One of the most well studied TP53 target genes is P21/CIP1/WAF1, which can inhibit cyclin dependent kinases (Deiry et al. 1993).TP53 can bind and inactivate PCNA, which plays an essential role in DNA replication (Waga et al.1994). It also activates RRM2B/P53R2 and DDB2 (damage – specific DNA binding protein 2) genes that have distinct roles in DNA repair (Tanaka et al. 2000). When the TP53 gene is altered due to various factors, tumour suppression is reduced. Individuals who inherit only one functional copy of TP53 gene have the likelihood of developing Li Fraumeni's syndrome (Hollstein et al. 1991). More than 50 proteins of human tumours contain a mutation or deletion of TP53 gene (Tyner et al. 2002). Sequence analysis was attempted to identify the structural details of normal p53 and its mutant variants using bioinformatics tools. This would help in predicting the mutation sites in the p53 amino acid residues and their by help in understanding p53 dysfunction as a result of oncogenic mutations.

Sultan Mohideen A K

PG & Research Department of Zoology, The New College, Chennai, Tamilnadu , India

*E-mail: smdbio@yahoo.co.in

Asrar Sheriff M

PG & Research Department of Zoology, The New College, Chennai, Tamilnadu , India

Materials and Methods

Data retrieval

The protein sequences of normal and two mutant forms of p53 were retrieved from NCBI using the accession number NM_000546, PDB 3D05 and PDB 2WGX respectively. The sequence was obtained in the FASTA format for further analysis.

Data analysis

Amino acid data analyses were carried out using EMBOSS. Analysis of the the amino acids in the normal p53 sequence was estimated through PepStats which detects the various properties of the protein sequence. The potential antigenic sites of the protein sequence were predicted using the ANTIGENIC-P. The functionally and structurally important residues in the protein sequence were identified using ConSeq (http://www.conseq.bioinfo.tau.ac.il/). The amino acid sequences of the normal and mutant forms of p53 were subjected to pBLAST for performing pairwise sequence analysis.

Results

p53 is derived from the TP53 gene (Fig-1) and has 393 amino acids with a molecular weight of 43653.06.

```
>tumor suppressor protein p53 [Homo sapiens]
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIE
QWFTEDPGPDEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPP
GTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVE
YLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTII
TLEDSSGNLLGRNSFEVRVCACPGRDRRTEEENLRKKGEPHHELPPGSTKR
ALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGK
EPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD
```

Figure 1. Amino acid sequence of normal p53

The average weight of each residue was estimated to be 111.076. The amino acid composition of p53 ranged from 1.0% to 11.5%. The amino acid proline was found to be 11.5% and that of tryptophan 1%. The total number of negatively charged residues (Aspartic acid and Glutamic acid) was found to be 50% when compared to the positively charged residues (Arginine and Lysine) with 46%. The instability index (II) was computed to be 73.59. The Aliphatic index shows 59.08 mole%. The Hydrophobic aliphatic amino acids were identified as isoleucine, leucine and valine. The concentration of negatively charged residues Glutamate and Aspartate was found to be 7.6% and 5.1% , whereas that of positively charged residues Lysine and Arginine was 5.1% and 6.6% respectively. The hydrophobic amino acids Isoleucine, Leucine and Valine contribute about 14.7% . The protein motif outfile shows 15 hits in the sequence. The amino acids ranging from 213 to 237 contain 25 residues. Valine was found to be dominant in the sequence with 16%. Similarly the residues from 70-101, 131-150, 169-182 271-279 show the distribution of Cysteine, Lysine, Cysteine and Valine residues. The amino acids position ranging from 1-17, 751-154, 164-180, 196-227, 239-291 and 381-397 shows functional, highly conserved and exposed residues. Residues from 55-108, 282-326, 351-380 and 384-393 represents exposed amino acids which is indicative of hydrophobic nature . Pair wise sequence analysis of normal p53 and mutant forms of p53 (Fig 2) revealed amino acid substitution of arginine by serine at the 249 amino acid residue (PDB 3D05).

```
>lcl|60867 3D05:A|PDBID|CHAIN|SEQUENCE
Length=200

Score =  424 bits (1089),  Expect = 2e-123, Method:
Compositional matrix adjust.
 Identities = 199/200 (99%), Positives = 199/200
(99%), Gaps = 0/200 (0%)
Query                                                94
SSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQ
LWVDSTPPP  153

SSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQ
LWVDSTPPP
Sbjct                                                1
SSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQ
LWVDSTPPP  60
Query                                               154
GTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVE
YLDDRNTFR  213

GTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVE
YLDDRNTFR
Sbjct                                                61
GTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVE
YLDDRNTFR  120
```

```
Query                                                214
HSVVVPYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNL
LGRNSFEVR  273
HSVVVPYEPPEVGSDCTTIHYNYMCNSSCMGGMNR
TIITLEDSSGNLLGRNSFEVR
Sbjct                                                121
HSVVVPYEPPEVGSDCTTIHYNYMCNSSCMGGMNRSPILTIITLEDSSGNL
LGRNSFEVR  180
Query  274  VCACPGRDRRTEEENLRKKG  293
                    VCACPGRDRRTEEENLRKKG
   Sbjct  181  VCACPGRDRRTEEENLRKKG  200
```

Figure 2. Pairwise sequence analysis of normal and mutant p53 (3DO5)

In mutant p53 (PDB 2WGX) many amino acid substitutional mutations are evident like substitution of Methionine by leucine at 133 amino acid position, valine by alanine at 203, tyrosine by phenylalanine at 236, asparagine by tyrosine at 239, threonine by isoleucine at 253 and asparagine by aspartic acid at 268 amino acid position. These mutations occur in the central DNA binding domain of normal p53 corresponding to amino acid sequence ranging between 102 and 292 residues (Fig 3).

```
>lcl|30595 2WGX:A|PDBID|CHAIN|SEQUENCE
Length=219

Score =  452 bits (1164),  Expect = 3e-132, Method:
Compositional matrix adjust.
 Identities = 213/219 (98%), Positives = 216/219
(99%), Gaps = 0/219 (0%)

Query                                                 94
SSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQ
LWVDSTPPP  153

SSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNK+FCQLAKTCPVQ
LWVDSTPPP
Sbjct                                                  1
SSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNKLFCQLAKTCPVQ
LWVDSTPPP  60

Query                                                154
GTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVE
YLDDRNTFR  213
GTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLR
EYLDD       RNTFR
Sbjct                                                 61
GTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRAE
YLDDRNTFR  120

Query                                                214
HSVVVPYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNL
LGRNSFEVR  273
HSVVVPYEPPEVGSDCTTIHYN+ MC   SSCMGGMNRRPIL
IITLEDSSGNLLGR+SF EVR
Sbjct                                                121
HSVVVPYEPPEVGSDCTTIHYNFMCYSSCMGGMNRRPILIIITLEDSSGNL
LGRDSFEVR  180

Query  274  VCACPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNT
312
           VCACPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNT
Sbjct  181  VCACPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNT
219
```

Figure 3. Pairwise sequence analysis of normal and mutant p53(2WGX)

Discussion

Mutations in the evolutionarily conserved codons of the p53 tumour suppressor gene are common in diverse types of human cancers like colon, lung, oesophagus, breast, liver, brain etc., (Hollstein et al. 1991) p53 is a transcription factor that guards the genome stability

and normal cell growth. Stresses like DNA damage, oncogenic assault will turn on p53 function which leads to cell cycle arrest for DNA repair, senescence for permanent growth arrest or apoptosis for programmed cell death (Chen et al. 2010). Alteration of this gene occurs in both somatic and germline mutations in cancer-prone families with Li-Fraumeni syndrome. More than 50% of human tumours contain a mutation or deletion of TP53 gene. The gene sequence can also be damaged in cells by mutagens, increasing the likelihood that the cells will begin uncontrolled cell division (Tyner et al. 2002). According to Prives (1999) alteration of the TP53 gene is the most frequent genetic alteration in human cancer which leads to the accumulation of mutant p53 in the nucleus of tumour cells.

The p53 comprised of 393 amino acids which are the building blocks of a protein which determines the secondary and tertiary structures. The negatively charged residues Glutamate and Aspartate and the positively charged residues Lysine and Arginine can create strong and electrostatic repulsion between the residual groups of the same sign. The hydrophobic amino acids Isoleucine, Leucine and Valine are usually found in the inner face of alpha-helices (Mount 2001) but when they occur on the surface of a protein are more likely to be a part of antigenic sites (Kolaskar and Tongaonkar 1990). Hydrophobicity is an essential driving force in protein folding. The hydrophobic character of amino acid residues is complex, and no single hydrophobic parameter can represent the complete range of amino acid behaviours (Attwood 2001).

In the present study, pair wise sequence analysis revealed that in the two mutant forms of p53, there were many amino acid substitutional mutations. Mutation or deletion of p53 can lead to more rapid proliferation and reduced apoptosis thereby triggering the development of tumours (Osada and Takahashi, 2002). Alteration in the central DNA binding domain of p53 could be responsible for sporadic cancers and somatic mutations as reported by Hainaut and Hollstein (2000). Luca et al. (1998) and Rutherford et al. (2002) reported on germline mutations in the human p53 gene. Hollstein et al. (1991) observed that analysis of mutations in p53 could provide clues to the etiology of diverse tumours such as the occurrence of transitions in colon, brain and lymphoid malignancies and transversions in the cancers of lung and liver. According to Bai and Zhu (2006) structural studies of p53 have revealed that the majority of p53 mutations found in cancers are missense mutations that are mostly located in the central DNA-binding domain. p53 is essential for preventing inappropriate cell proliferation and maintaining genome integrity following genotoxic stress (Vousden and Lu 2002). Various intracellular and extracellular stimuli, such as DNA damage through ionizing and UV radiation, application of cytotoxic drugs or chemotherapeutic agents, and infectious virus, heat shock, hypoxia, and oncogene over expression can result in p53 activation which triggers diverse biological responses, both at the level of a single cell as well as in the whole organism (Vogelstein et al. 2000).

Conclusion

Further insilico studies pertaining to secondary and tertiary structures of p53 could lead to applications for drug designing, target novel anti cancer drugs to bind and stabilize p53 protein and help in unraveling the significant insights on the role of p53 in carcinogenesis and mutagenesis. The present study could provide new targets and insights for diagnostic and therapeutic interventions for human cancers and help in unraveling avenues for further research in tumour biology.

Reference

Attwood TK and Smith DJP (2001) Introduction to Bioinformatics. Pearson Education, USA

Deiry WS et al (1993) WAF 1 a potential mediator of p53 tumour suppression cell. Cell 75:817-825

Chen F et al (2010) Current strategies to target p53 in cancer. Biochem Pharmacol. 80(5):724-30

Hainaut P and Hollstein M (2000) p53 and human cancers: The first ten thousand mutations. Adv Cancer Res 77:81-137

Harms KL and Chen X (2005) The C terminus of p53 family proteins in a cell fat determinant. Mol.Cell.Biol. 25(5):2014-2030.

Hollstein M et al (1991) p53 mutations in human cancers. Science 253(5015):49–53

Kolaskar AS and Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. Febs Lett 276*:*172-174

Bai L, Wei-Guo Zhul (2006) p53: Structure, Function and Therapeutic Applications. J Cancer Mol 2(4):141-153

Luca JW, Strong LC, Hansen MF (1998) A germline missense mutation R337C in exon 10 of the human p53 gene. Hum Mutat 1:S58-S61

Mount DW (2001) Bioinformatic sequence and genome analysis. Cold Spring Harbor Laboratory Press, NY.

Osada H, Takahashi T (2002) Genetic alterations of multiple tumour suppressors and oncogenes in the carcinogenesis and progression of lung cancer. Oncogene 21:7421-34

Prives C, Hall PA (1999) The p53 Pathway. J Pathol 187:112-127

Suad O, Rozenberg H et al (2009) Structural basis of restoring sequence-specific DNA binding and transactivation to mutant p53 by suppressor mutations. J Mol Biol 385(1):249-65

Khoo KH, Joerger AC, Freund SM, Fersht AR (2009) Stabilising the DNA-binding domain of p53 by rational design of its hydrophobic core. Protein Eng Des Sel 22(7):421-30

Read et al (1999) Chapter 18: Cancer Genetics", Human molecular genetics 2. Wiley, New York

Rutherford J et al. (2002) Investigations on a clinically and functionally unusual and novel germline p53 mutation. Br J Cancer 86(10):1592-1606

Tanaka H et al (2000) A ribonucleotide reductase gene involved in a p53 dependent cell cycle checkpoint for DNA damage. Nature 404:42-49

Tyner SD (2002) p53 mutant mice that display early ageing-associated phenotypes. Nature 415(6878):45-53

Vogelstein B et al (2000) Surfing the p53 network. Nature 408:307-310

Vousden KH and Lu X (2002) Live or let die: the cell's response to p53. Nat Rev Cancer 2:594-604

Waga S et al (1994) The p21 inhibitor of cyclin dependent kinase controls DNA replication by interaction with PCNA. Nature 369:574-578