INTERACTION OF HSP 90 WITH P53 AND ITS MUTATED FORM AND THEIR COMPARISON

A DISSERTATION
SUBMITTED IN PARTIAL FULFILLMENT OF

BACHELOR OF TECHNOLOGY

UNDER THE ACADEMIC AUTONOMY
NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA

BY RAVI BHUSHAN
ROLL NO.107BT11

UNDER THE GUIDANCE OF

PROFESSOR AND HEAD SUBHANKAR PAUL
DEPARTMENT OF BIOTECHNOLOGY AND MEDICAL ENGG.
NATIONAL INSTITUTE OF TECHNOLOGY
ROURKELA -769008, ODISHA, INDIA, 2011

NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA

CERTIFICATE

This is to certify that the project entitled "Interaction of Hsp 90 with p53 and its mutated form and their comparison" submitted by Ravi Bhushan of Biotechnology branch, National Institute of Technology, Rourkela, for the degree of bachelor of technology is a record based on the result obtained in the bonafide research work carried out by him under my guidance and supervision.

To the best of my knowledge, the matter embodied in the project has not been submitted to any other university/institute for the award of any degree or diploma

Date :

Dr. Subhankar Paul
Professor andHead
Department of Biotechnology and Medical Engineering
NIT Rourkela

ACKNOWLEDGEMENT

This project is the account of one year of devoted work at national institute of technology, Rourkela, India, which would not have been possible without the help and support of many.

With deep regards and profound respect, I avail the opportunity to express my gratitude to my supervisor prof. Subhankar Paul, Department of Biotechnology and Medical Engineering, NIT Rourkela, for his support, guidance and advice.

I acknowledge my sincere regards to all the faculty members, Department of Biotechnology and Medical Engineering, NIT Rourkela for their enthusiasm in promoting the research, their support and dedication to students.

I extend my severe thanks to Mustafa, my project assistant who worked with me since one year

Finally I am indebted to my parents for their patience, endurance, ceaseless and moral support which made the completion of course possible.

Date :

Ravi Bhushan
107BT11
Department of Biotechnology and Medical Engineering
NIT Rourkela
Odisha, India.

CONTENTS

10. MATERIALS AND METHODS

11. RESULTS AND DISCUSSION

For human Hsp 90-alpha isoform

For wild type human p53

ABBREVIATIONS

Hsp	Heat Shock Protein
GRAVY	Grand average of hydropathy
SAPS	Statistical Analysis of Protein Sequences
MEME	Multiple Em for Motif Elicitation
I-TASSER	**I**terative **T**hreading **ASSE**mbly **R**efinement
BLAST	Basic Local Alignment Search Tool
SAGE	Serial analysis of gene expression
SOPMA	self –optimized prediction method

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The binding of one signaling protein to another serve to recruit a signaling protein to a location where it is activated and where it is needed to carry out its function, it induce conformational changes that affect activity or accessibility of additional binding domains, permitting additional protein interactions. If in a cell interaction disappearance will make cell deaf and blind, paralytic and finally will disintegrate.Hsp90 are molecular chaperones involved in stress response and normal biosynthetic and homeostatic control mechanisms of the cell .It interact with different clients such as protein (kinases, transcription factor, and structurally unrelated protein), activate and direct it to proteasomal degradation. Disrupting the interaction between Hsp 90 and its client protein controls its biogenesis, stability and activity. The question arises is how the Hsp 90 recognizes p53 protein and what region of Hsp 90 are interacting with it. In this study the basis of interaction of Hsp 90 with p53 is predicted by identifying unveiled hydrophobicity patches,location of binding site ,charge distribution and docking. Hydrophobic patches with hydropathy index similarity between Hsp 90 and p53 is predicted.The charge distribution and overall hydrophobicity was also compared . Protein –protein docking is studied using HEX which predicts that human Hsp 90 and wild type p53 form stable complex ,this is compared with mutant type p53.

INTRODUCTION

Bioinformatics is the field of science in which biology, computer science and information technology merge into a single discipline.it is conceptualizing biology in terms of molecules and applying informatics techniques to understand and organize the information associated with these molecules. Using bioinformatics drug can be designed with minimal work at wet lab. With the help of gene sequence protein sequence is determined, from which determination of structure can be done using structure prediction technique. With further resolving protein surface force surrounding molecule can be determined. Docking allows prediction of ligand which will bind to the protein surface. The influxes of biological data and advancement of computer technology have broadened the scope of biology. Need for a large database is felt to store, view and deconstruct information.

Protein structure analysis is an important part of bioinformatics which approximates the structure of a molecule in very less time, which helps the biochemists, who take years and centuries to study and predict the structure of larger molecules.

Heat shock protein belongs to the family of protein whose expression increases in response to variety of metabolic insults. It promotes synthesis and folding of protein throughout the cell and supports protein assembly, secretion, trafficking, protein degradation and regulation of transcription factor and protein kinases. Transcription factor p53 activation occurs due to cellular

stress and several other pathways of its activation. In normal cellular condition it remains in "standby mode", once activated it activates number of genes and take part in DNA repair. In most cases of human cancer cell it remains in mutated form which prevents transcriptional activation of genes.

P53 is a tumour suppressing protein encoded by tp53 gene.p53 protein regulates cell cycle and functions as tumour suppressor protein that is involved in preventing cancer. It remain in cell in mutated or deactivated form as it is degraded rapidly via proteolysis and ubiquitination.Due to this only low concentration of this normal protein is present in the cell.

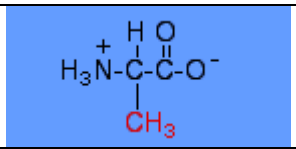
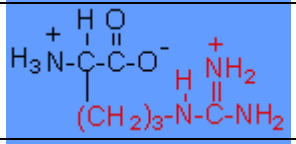
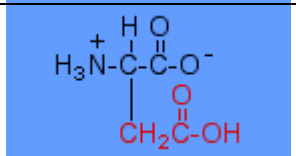
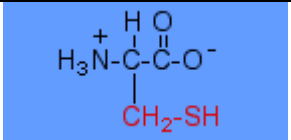
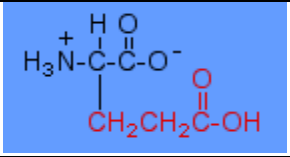
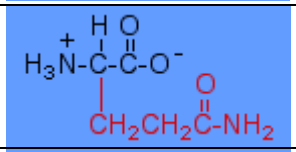
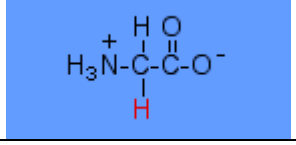
OBJECTIVE

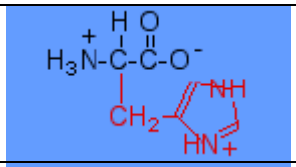
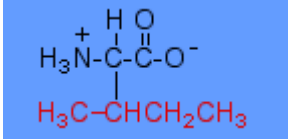
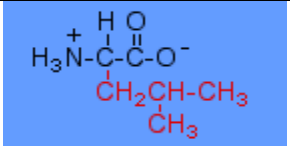
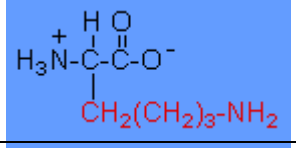
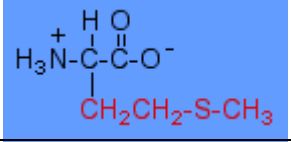
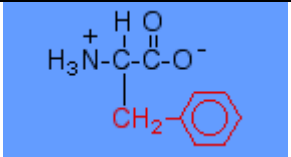
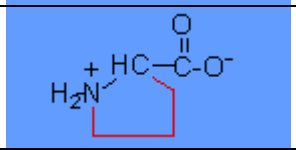
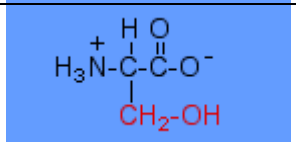
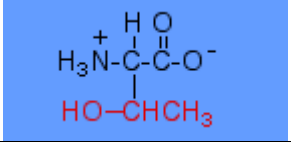
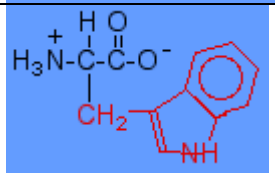
TO STUDY INTERACTION OF HSP90 WITH P53 AND ITS MUTATED FORM AND THEIR COMPARISON

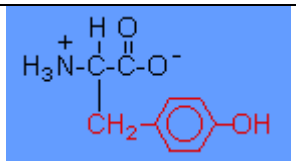
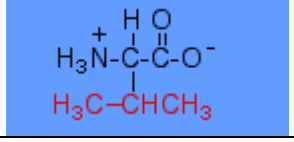
LITERATURE REVIEW

Table (1): Structures of Amino Acids [22]

The structure of amino acids differs by the type of R group present in the basic amino acid structure. there name are abbreviated with single alphabet.

Amino Acid Name	abrev	Structure of R group (red)	Comments
Alanine	A		Neutral None-polar
Arginine	R		Basic Polar
Asparagine	N		Neutral Polar
Aspartic Acid	D		Acidic Polar
Cysteine	C		Neutral Slightly Polar
Glutamic Acid	E		Acidic Polar
Glutamine	Q		Neutral Polar
Glycine	G		Neutral Non-polar

Histidine	H		Basic Polar
Isoleucine	I		Neutral Non-polar
Leucine	L		Neutral Non-polar
Lysine	K		Basic Polar
Methionine	M		Neutral Non-polar
Phenyl-alanine	F		Neutral Non-polar
Proline	P		Neutral Non-polar
Serine	S		Neutral Polar
Threonine	T		Neutral Polar
Trypto-phan	W		Neutral Slightly polar

Tyrosine	Y		Neutral Polar
Valine	V		Neutral Non-polar

Structure of Hsp90:

It consists of 732 amino acid sequence .the numbering indicates the amino acid sequence which define function domains. This is basically present as homo dimer in cytoplasm. The homo dimer consists of N-terminal ATP-binding domain, a middle region, and C-terminal homo dimerization domain which contain tetratricopeptide repeat protein (TRP) binding motif. CR region act as flexible linker between N-terminal and middle domains. The N-terminal domain contains peculiar adenine nucleotide-binding pocket called as Bergerat fold .Hydrolysis of ADP to ATP in the bergerat fold shows structure alteration which is essentially important for chaperoning activity of Hsp 90 dimer. The structures of middle region contain highly charged and flexible linker sequence that connects to the middle region of Hsp 90.It interacts with gamma phosphate of ATP molecule. Middle region plays a key role in binding of many client protein to Hsp 90.C-terminal has an important role in cooperative intermolecular and intermolecular interaction that regulate use of ATP by chaperone.Hsp90 interaction with client protein is regulated by many co-chaperone. [15]

Interaction cycle:

In open state of Hsp 90 ATP binding induces major conformational change in the dimer which forms a ring like closed structure. The N-terminal association is the rate limiting step for overall ATPase cycle. [15]

Interaction study of human hsp90-p53:

P53 a transcription factor is degraded rapidly through proteolysis and the ubiquitin proteasome pathway, low concentration of p53 reside in cell due to high turnover ,but mutated p53 is not strongly affected by proteolysis and it accumulates in the cell. Hsp 90 bind tightly to p53 and inhibits its function leading to degradation of mutant p53.Hsp 90 binds to wild type p53 which is more transient than mutant protein. It is held in dynamic equilibrium by transient association with Hsp 90 which contain complex of p23 and CYP40 which help in maintaining conformation for DNA binding. Most mutant p53 are not able to achieve the conformation required for DNA binding.[14][15]

BIOINFORMATICS TOOL

Swiss prot database:

It is a protein sequence database with annotated data; data includes function of protein, domain structure, post translational modifications. Amino acid sequence of p53 and hsp90 is obtained from Swiss prot[11][12]

SOPMA:

Self-Optimized Prediction Method is used to predict secondary structure of protein that uses multiple alignments. A protein query sequence is retrieved from NCBI in FASTA format and pasted on the SOPMA submission form. The SOPMA was run and the result was saved. It predict number of alpha helix and beta bridges and other detail of protein structure[9][8]

Surface Hydrophobicity Plot:

Hydrophobicity plot is the graphical representation of the distribution of polar and apolar residues along a protein sequence. It predict hydrophobic and hydrophilic region. This works under two types of scale kyte-doolittle used for hydrophobic character and hopp woods scale used for antigenic region. Region above value 0 are hydrophobic and window size 5-7 is good for finding surface expose region in kyte-dolittle scale.[3][4]

Table 2: Hydrophobicity Scale of Amino acid[10]

Amino Acid	Kyte-Doolittle scale	Hopp-Woods scale
Alanine	1.8	-0.5
Arginine	-4.5	3.0
Asparagine	-3.5	0.2
Aspartic acid	-3.5	3.0
Cysteine	2.5	-1.0
Glutamine	-3.5	0.2
Glutamic acid	-0.4	0.0
Glycine	-0.4	0.0
Histidine	-3.2	-0.5
Isoleucine	4.5	-1.8
Leucine	3.8	-1.8
Lysine	-3.9	3.0
Methionine	1.9	-1.3
Phenylalanine	2.8	-2.5
Proline	-1.6	0.0
Serine	-0.8	0.3
Threonine	-0.7	-0.4
Tryptophan	-0.9	-3.4
Tyrosine	-1.3	-2.3
Valine	4.2	-1.5

Protparam tool:

This is a tool which calculate various physical and chemical parameters of a given protein in Swiss Prot or TrEMBL .Parameters are molecular weight ,no.of amino acid, theoretical pI, atomic composition, extinction coefficient ,estimated half life, instability index and grand average of hydrophobicity (GRAVY)[7]

SAPS

Statistical analysis of protein sequence (SAPS) evaluates protein sequence properties by statistical criteria. Properties contain compositional biases, sample and runs of charge and other amino acid, different kind of extent and repetitive structures, local motifs, and anomalous spacing between identical residue types. Sequence with at least 200 residue is reliable and should be of the following format; raw, plain, EMBL, Swiss Prot, Genbank,PIR,Fasta,NBRF, and GCG.[9]

Ident and sim tool:

This is a tool which calculates percentage similarity of hydrophobic patch between Hsp90 and p53.This also calculates identical charged-charged interaction.[4]

I-TASSER:

Iterative threading assembly refinement program is an online 3dimensional structure prediction server for protein .It predicts structure based on confidence score and tm score. The output

contains full length secondary and tertiary structure prediction, and functional annotation on legend binding site.[6]

BLAST:

Basic local alignment search tool is a tool for comparing sequence among all known sequence in database .It predicts alignment based on confidence score. The outputs contain many sequences with most similar alignment and property. We take database with high e-value. BLASTS also help in finding PDB id with high confidence score. That ID use in protein docking with help of HEX.6. [5]

HEX:

This is used to study protein- protein docking in which one of the proteins is docked as ligand and other as receptor. This removes all water molecules and other hetero atoms from input file. Hex rotates each protein about its coordinate origin and varies the separation between two origins. Highest scoring orientation are saved and returned to user.[2]

ESyPred3D:

This is an automated homology program used to predict three –dimensional structure of proteins. Alignments are obtained by combining, weighing and screening the results of several multiple alignment programs. Steps involved are search for a template, aligning query and template sequence, building 3D model using last alignment and the structure of the template and assessing the final 3D model.[9]

MATERIALS AND METHODS

Amino acid sequence:

The three protein involved hsp90, wild type p53, and mutated p53, their amino acid sequence was obtained from Swiss Prot database.

3-d protein structure:

3 dimensional protein structures were obtained from PDB (protein data bank). PDB id and chain id was obtained by BLAST tool of FSTA sequence. This ID was used to obtain PDB structure and their 3-D modal for good prediction of docking. The protein includes human wild type p53, mutant p53and Hsp90.

Hydrophobicity patch:

Hydrophobicity patches in Hsp90 and p53 (client protein) and mutant p53 was identified using surface hydrophobicity plot scale and window size being Kyte-doolittle and 7 respectively.[10]

Hydropathy index:

Hydropathy index of Hsp 90 and p53 was estimated by grand average of hydropathy (GRAVY) value which is the sum of the hydropathy value for all amino acid in a protein sequence divided by no. of residue, using Protparam tool.[11][12]

Overall percentage hydrophobicity:

Overall percentage of hydrophobicity of Hsp 90 and p53 was calculated as ratio of no. of amino acid having high hydrophobicity especially (L, V, I, F, M) and total no. of amino acid sequence.

Similarity of hydrophobicity between hsp90 and client protein:

Percentage similarity of hydrophobic patches between Hsp90 and p53 was calculated using Ident and sim tool. [9]

Prediction of 3D protein structure by I-TASSER:

Prediction of 3- dimensional structure of protein was done based on confidence score and tm score. It was fully computerized algorithm to obtained 3-D structure that was obtained by online I-TASSER server.[6]

Protein- protein Docking :

Protein- protein docking was done using HEX tool in which human Hsp 90 was docked with wild type and mutant type p53. Result was obtained with help of HEX6 tool shown energy data for docking that predicts which type of docking is more stable.[2]

Protein interaction site:

Protein –protein interaction predicted using SHARP2 tool. PDB id for atomic coordinates of the crystal structure of a p53 and Hsp 90 and chain id for promoter of choice is entered. Type of protein is selected and parameters are set and are submitted to get the location of the interaction site.[1]

RESULT AND DISCUSSION

Sequence of Human Hsp 90

>gi|153792590|ref|NP_001017963.2| heat shock protein HSP 90-alpha isoform 1 [Homo sapiens]

MPPCSGGDGSTPPGPSLRDRDCPAQSAEYPRDRLDPRPGSPSEASSPPFLRSRAPVNWYQEKAQVFLWHL
MVSGSTTLLCLWKQPFHVSAFPVTASLAFRQSQGAGQHLYKDLQPFILLRLLMPEETQTQDQPMEEEEVE
TFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIRYESLTDPSKLDSGKELHINLIPNKQDRT
LTIVDTGIGMTKADLINNLGTIAKSGTKAFMEALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDE
QYAWESSAGGSFTVRTDTGEPMGRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDK
EVSDDEAEEKEDKEEEKEKEEKESEDKPEIEDVGSDEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWT
RNPDDITNEEYGEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKNNIKLYVRRV
FIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFTELAEDKENYKKFY
EQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKDYCTRMKENQKHIYYITGETKDQVANSAFV
ERLRKHGLEVIYMIEPIDEYCVQQLKEFEGKTLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDI
LEKKVEKVVVSNRLVTSPCCIVTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQ
KAEADKNDKSVKDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPP
LEGDDDTSRMEEVD

Physio-chemical parameters of Hsp90 alpha isoform computed by Protparam tool (expasy.org)

Number of amino acids: 854

Molecular weight: 98161.1

Theoretical pI: 5.07

Amino acid composition:

Ala (A)	43	5.0%
Arg (R)	39	4.6%
Asn (N)	32	3.7%
Asp (D)	60	7.0%
Cys (C)	10	1.2%
Gln (Q)	33	3.9%
Glu (E)	100	11.7%
Gly (G)	40	4.7%
His (H)	15	1.8%
Ile (I)	51	6.0%
Leu (L)	77	9.0%
Lys (K)	83	9.7%
Met (M)	22	2.6%
Phe (F)	31	3.6%
Pro (P)	37	4.3%
Ser (S)	56	6.6%
Thr (T)	47	5.5%
Trp (W)	7	0.8%
Tyr (Y)	28	3.3%
Val (V)	43	5.0%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged amino acid (Asp + Glu): 160

Total number of positively charged amino acid (Arg + Lys): 122

Atomic composition:

Carbon	C	4336
Hydrogen	H	6874
Nitrogen	N	1156
Oxygen	O	1371
Sulfur	S	32

Formula: $C_{4336}H_{6874}N_{1156}O_{1371}S_{32}$

Total number of atoms: 13769

Extinction coefficients:

Unit of extinction coefficients M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 80845

Abs 0.1% (=1 g/l) 0.824, assuming all pairs of Cys residues form cystines

Ext. coefficient 80220

Abs 0.1% (=1 g/l) 0.817, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the Hsp90 sequence considered is M (Met).

The estimated half-life is: 30 hours.

>20 hours (yeast, in vivo).

>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is calculated to be 47.71

This categories the protein as unstable.

Aliphatic index: 78.09
Grand average of hydropathicity (GRAVY): -0.699

Secondary structure prediction of Hsp90-alpha isoform using SOPMA tool

```
        10        20        30        40        50        60        70
         |         |         |         |         |         |         |
MPPCSGGDGSTPPGPSLRDRDCPAQSAEYPRDRLDPRPGSPSEASSPPFLRSRAPVNWYQEKAQVFLWHL
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccchhhhhhheeeeee
MVSGSTTLLCLWKQPFHVSAFPVTASLAFRQSQGAGQHLYKDLQPFILLRLLMPEETQTQDQPMEEEEVE
ehttcceeeeetcccceeecccccchhhhhhccttttthhhhhtcchhhhhhecccccccccccccchhhhh
TFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIRYESLTDPSKLDSGKELHINLIPNKQDRT
hhhhhhhhhhhhhhhhhhcccthhhhhhhhhhccchhhhhhhhhhhcccccccctteeeeeeccctttce
LTIVDTGIGMTKADLINNLGTIAKSGTKAFMEALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDE
eeeeetccccchhhhhhhhhhhhhttchhhhhhhhttcceeeeeccccchhhhhhhteeeeeeccccccc
QYAWESSAGGSFTVRTDTGEPMGRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDK
ceeeeccttceeeeeecccccccttceeeeeectthhhhhhhhhhhhhhhhhcctteetcceeeeeehcccc
EVSDDEAEEKEDKEEEKEKEEKESEDKPEIEDVGSDEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWT
cccchhhhhhhhhhhhhhhhhhhhtcccchhhhcccchhhhhhttchhhhhhhhhhhccchhhccccceeee
RNPDDITNEEYGEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKNNIKLYVRRV
ccccccchhhhhhhhhhhhhtcccchhheeeecttcccceeeeeecccccccchhhhhhhhccheeeeeee
FIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFTELAEDKENYKKFY
eecccchhhhhhhhhhhhhhheecttcccccccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh
EQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKDYCTRMKENQKHIYYITGETKDQVANSAFV
hhhhhheeeeeecccccchhhhhhhhhhccccccccchhhhhhhhhhhcttteeeeeetccccchhcccchhh
ERLRKHGLEVIYMIEPIDEYCVQQLKEFEGKTLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDI
hhhhhttceeeeeeccchhhhhhhhhhhcttccceeehhhhheeecccchhhhhhhhhhhhhhhhhhhhhh
LEKKVEKVVVSNRLVTSPCCIVTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQ
hhhhhhheeehhhccccchheeeecccccchhhhhhhhhhhhhccccchhhhhhhcceeecctteehhhhhhh
KAEADKNDKSVKDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPP
hhhccccchhhhhhhhhhhhhhhhhhttcccccchhhhhhhhhhhhhhhccccccccccchhcchhhhhhcch
LEGDDDTSRMEEVD
Hhhhhhhhhhhhhh
```

This result shows the secondary structure of Hsp90 amino acid sequence.

Sequence length: 854

SOPMA:

Alpha helix (Hh): 407 is 47.66%

3_{10} helix (Gg) : is 0.00%

Pi helix (Ii) : is 0.00%

Beta bridge (Bb) : is 0.00%

Extended strand (Ee) : 123 is 14.40%

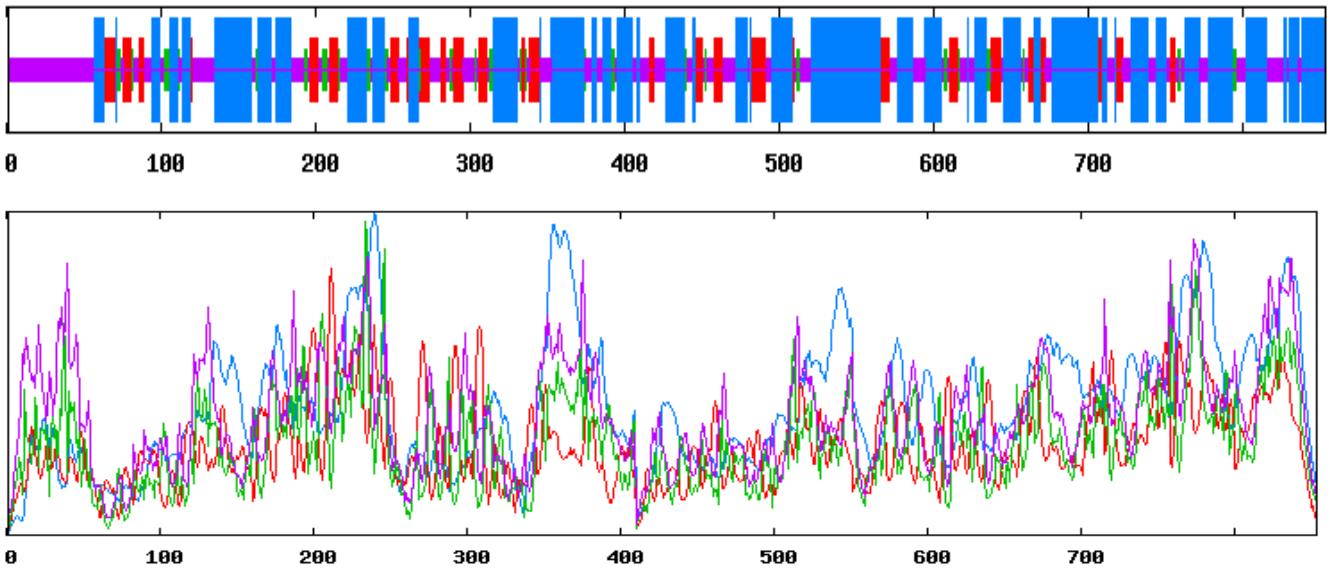
Beta turn (Tt) : 48 is 5.62%

Bend region (Ss) : is 0.00%

Random coil (Cc) : 276 is 32.32%

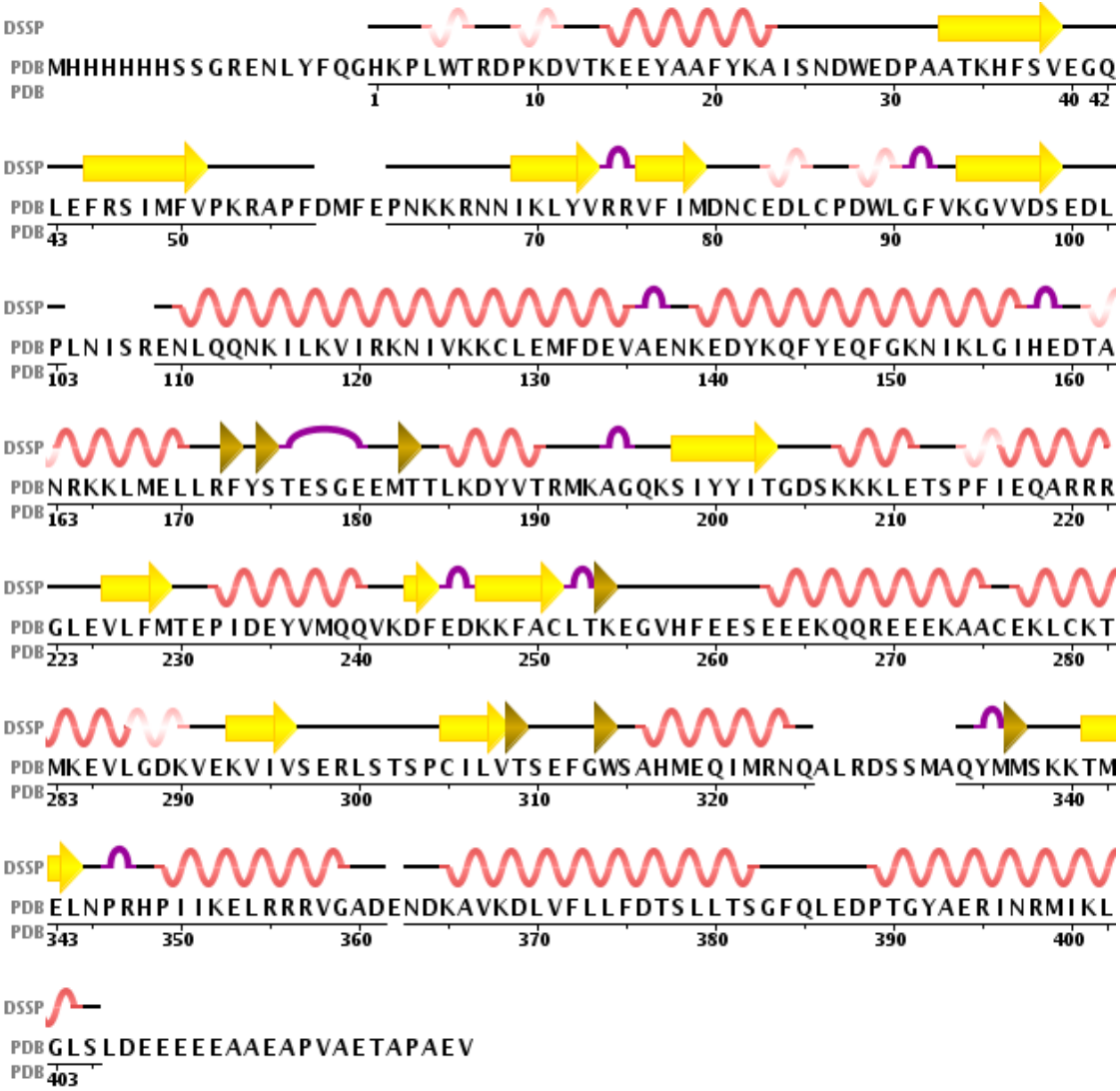
Ambiguous states (?) : is 0.00%

Other states : is 0:00%



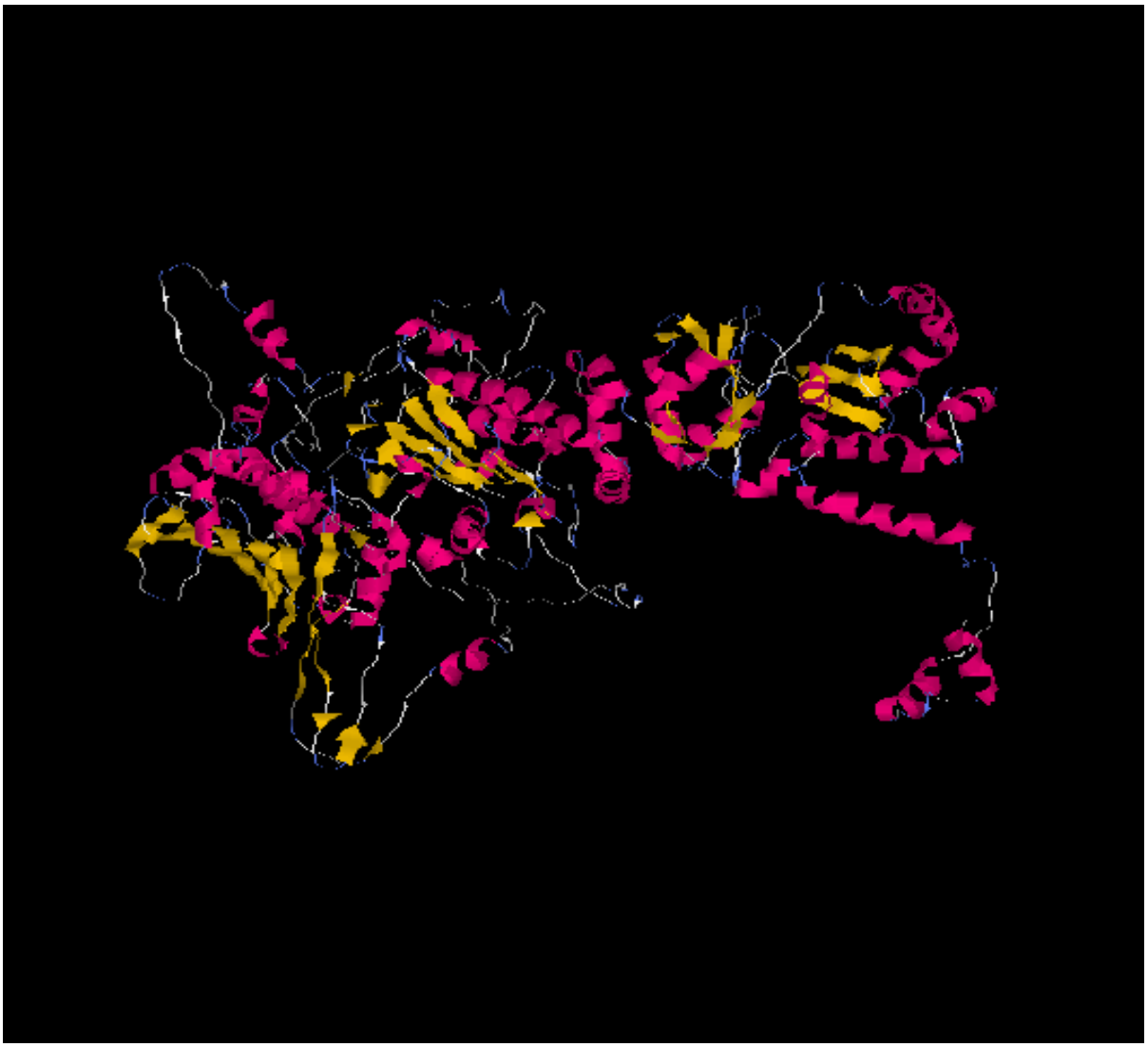
(fig-**1)**

PDB secondary structure of Hsp90



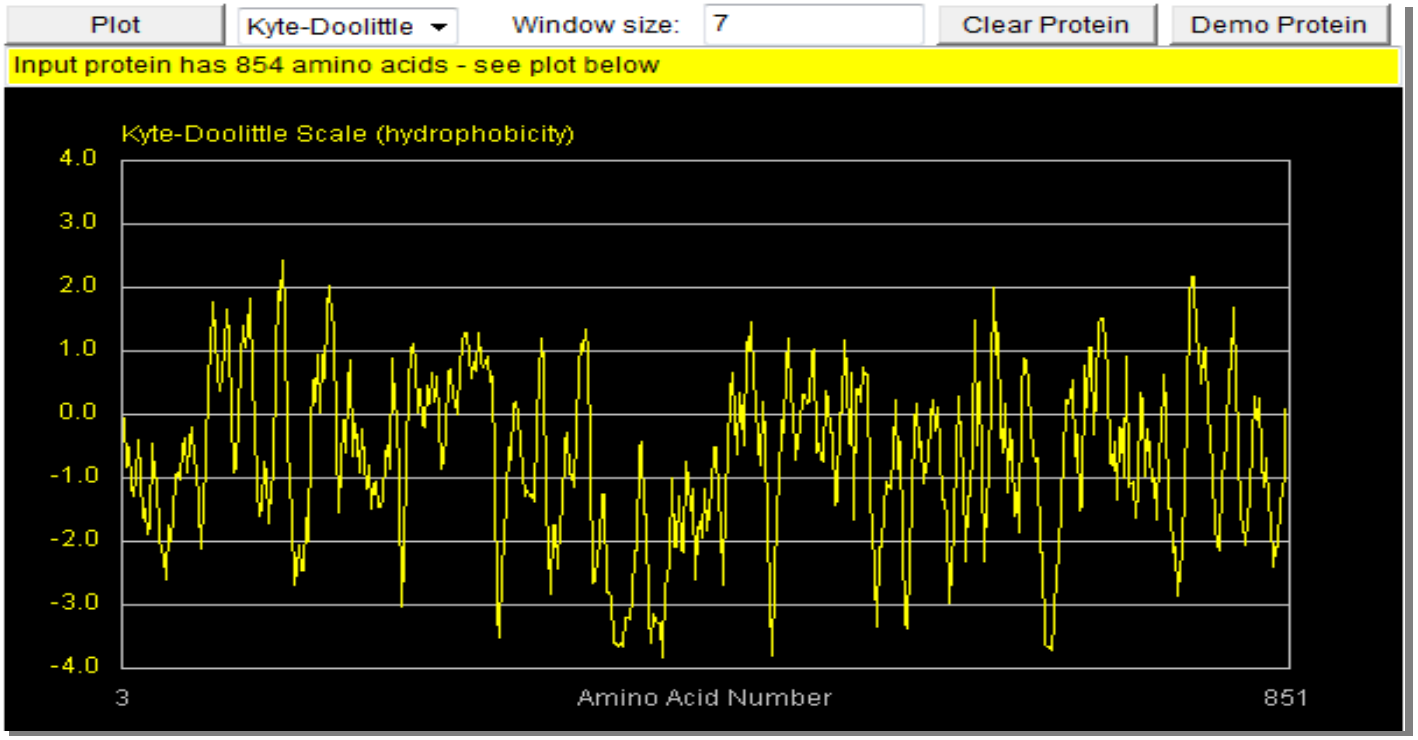
(Fig-2)

I-TASSER:- online 3dimensional structure prediction of Hsp90-alpha isoform



(fig-3)

Hydrophobicity plot of hsp 90-alpha isoform



(Fig-4)

Hydrophobicity plot was predicted using surface hydrophobicity plot using FASTA sequence of Hsp90 and window size 7 which is suitable for studying protein –protein interaction .From the hydrophobic plot peak was obtained at a specific location of amino acid .Including the peak location 7 amino acid sequence was selected which were hydrophobic. This forms hydrophobic patches .The patches obtained is detailed in table below.

Table (3): Predicted Hydrophobic Patch in Hsp90 alpha and in Human

protein	%Overall Hydrophobicity	Hydropathy Index of the predicted patch
Hsp90 alpha isoform	26.6%	**2.316(ILLRLLM),**2.029(**LMSLIIN**), 2.100(**LEVIYMI**), 2.186 (**LVILLYE**)

Sequence of Human P53

GenBank: BAC16799.1

GenPeptGraphics

>gi|23491729|dbj|BAC16799.1| P53 [Homo sapiens]
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGPDEAPRMPEAA
PRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKT
CPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRN
TFRHSVVVPYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVHVCACPGR
DRRTEEENLRKKGEPHHELPPGSTKRALSNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALEL
KDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD

Physio chemical property of p53

Number of amino acids: 393

Molecular weight: 43683.1

Theoretical pI: 6.37

Amino acid composition:

Ala (A)	24	6.1%
Arg (R)	26	6.6%
Asn (N)	14	3.6%
Asp (D)	20	5.1%
Cys (C)	10	2.5%
Gln (Q)	15	3.8%
Glu (E)	30	7.6%

Gly (G)	23	5.9%
His (H)	13	3.3%
Ile (I)	8	2.0%
Leu (L)	32	8.1%
Lys (K)	20	5.1%
Met (M)	12	3.1%
Phe (F)	11	2.8%
Pro (P)	43	10.9%
Ser (S)	39	9.9%
Thr (T)	22	5.6%
Trp (W)	4	1.0%
Tyr (Y)	9	2.3%
Val (V)	18	4.6%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 50
Total number of positively charged residues (Arg + Lys): 46

Atomic composition:

Carbon	C	1897
Hydrogen	H	2978
Nitrogen	N	550
Oxygen	O	593
Sulfur	S	22

Formula: $C_{1897}H_{2978}N_{550}O_{593}S_{22}$

Total number of atoms: 6040

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 36035
Abs 0.1% (=1 g/l) 0.825, assuming all pairs of Cys residues form cystines

Ext. coefficient 35410
Abs 0.1% (=1 g/l) 0.811, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 71.93
This classifies the protein as unstable.

Aliphatic index: 59.08

Grand average of hydropathicity (GRAVY): -0.758

Secondary structure of wild type p53

```
        10        20        30        40        50        60        70
         |         |         |         |         |         |         |
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGPDEAPRMPEAA
ccccccccccccccchhhhhhhhhhccttheeccccccccchhhccccccchhccccccccchhhhccccc
PRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKT
ccccccccccccccccccccccccccccccccccccccceeeeeeccccccccceeecchhhhhhhhhhhtc
CPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRN
cceeeeecccccccceeehhhhhcccchhhhhhhhccccccccccccccccccceeeeecccccheeeccccc
TFRHSVVVPYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVHVCACPGR
cccceeeeecccccccccchheeeehhhhcccccccccccceeeeeeecctttcceetcceeeeeeecctttc
DRRTEEENLRKKGEPHHELPPGSTKRALSNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALEL
ccccchhhhhhccccccccccccccccccccccccccccccccccccceeeeeeecccchhhhhhhhhhhhhh
KDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD
Hhhhhhccccccccchhhccccccccctttcceeeeeccccccc
```

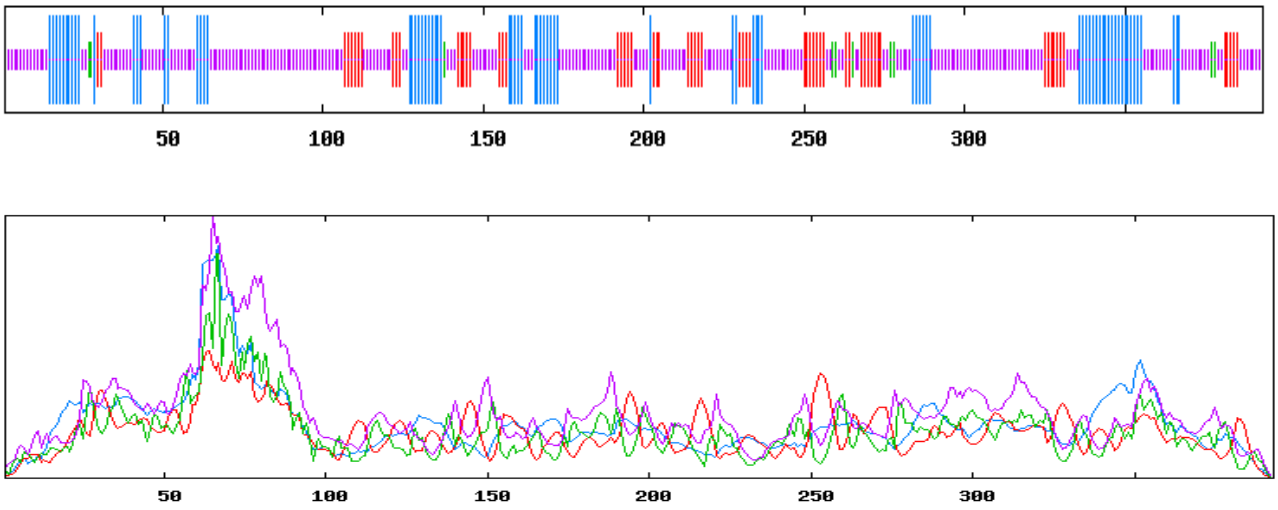
Sequence length : 393

SOPMA tool result

Alpha helix (Hh) : 80 is 20.36%
3_{10} helix (Gg) : 0 is 0.00%
Pi helix (Ii) : 0 is 0.00%
Beta bridge (Bb) : 0 is 0.00%
Extended strand (Ee) : 64 is 16.28%
Beta turn (Tt) : 10 is 2.54%
Bend region (Ss) : 0 is 0.00%
Random coil (Cc) : 239 is 60.81%
Ambigous states (?) : 0 is 0.00%
Other states : 0 is 0.00

Parameters :

Window width : 17
Similarity threshold : 8
Number of states : 4



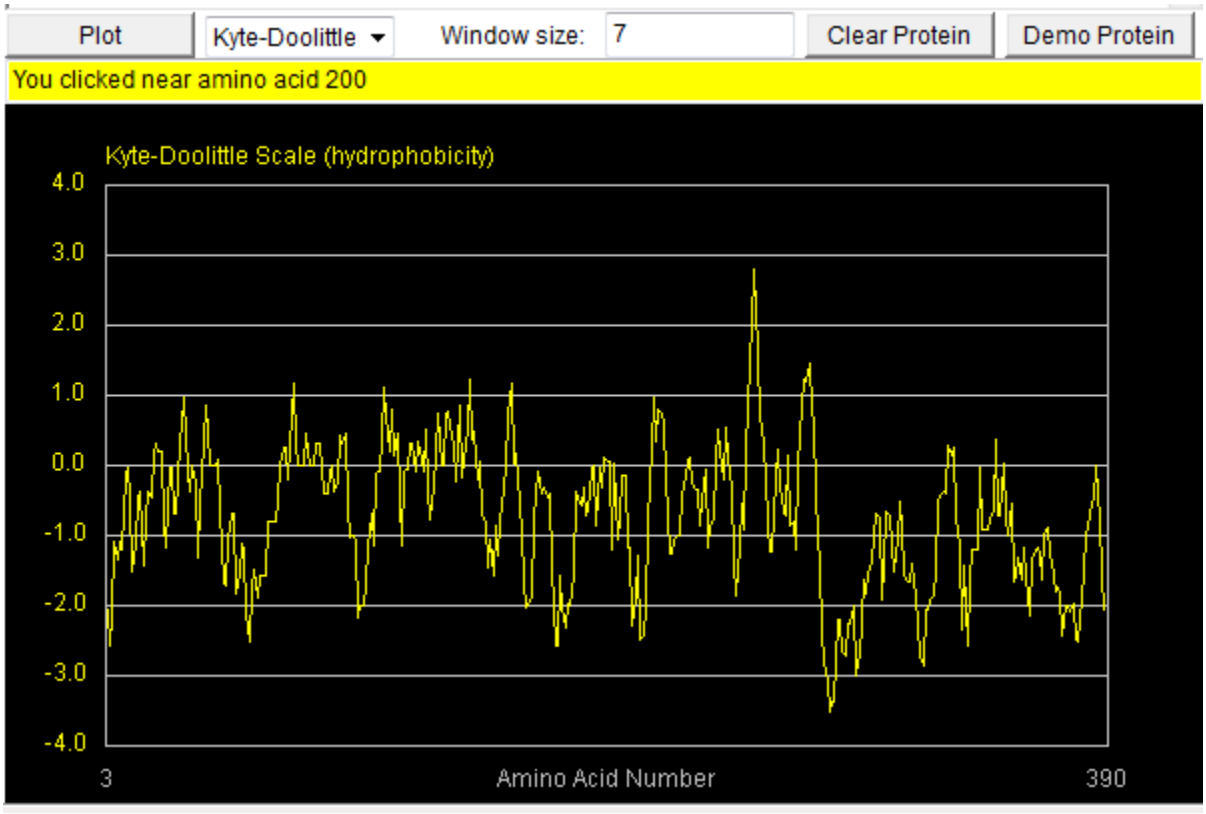
(Fig-5)

I-TASSER:- online 3dimensional structure prediction of p53

(fig-6)

This predicts three dimensional structure of protein online automatically using FASTA sequence of amino acid. This is used to compare three dimensional structure obtained from predicted PDB id.

Hydrophobicity plot of wild type p53



(fig-7)

Hydrophobic patch of p53 was obtained as it was obtained for Hsp90.The patch obtained are listed in table-2 below.

Table 4: Predicted Hydrophobic Patch in wild type p53 that came from fig . (Transcription Factor) in Human

Transcription Factor	Predicted Hydrophobic Patch	Location of Patch
p53	GMNRRP**ILTIITL**EDSSGNLLG	Middle domain

Mutant p53 [Homo sapiens]

GenBank: ACI25593.1

GenPept Graphics

>gi|208342286|gb|ACI25593.1| mutant p53 [Homo sapiens]

MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGPDEAPRMPEAA
PRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKT
CPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEVNLRVEYLDDRN
TFRHSVVVPYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGR
DRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALEL
KDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD

Physio-chemical properties of mutant p53

Number of amino acids: 393

Molecular weight: 43754.3

Theoretical pI: 6.48

Amino acid composition:

Ala (A)	24	6.1%
Arg (R)	27	6.9%
Asn (N)	14	3.6%
Asp (D)	20	5.1%
Cys (C)	10	2.5%
Gln (Q)	15	3.8%
Glu (E)	30	7.6%
Gly (G)	22	5.6%
His (H)	12	3.1%
Ile (I)	8	2.0%
Leu (L)	32	8.1%
Lys (K)	20	5.1%

Met (M)	12	3.1%
Phe (F)	11	2.8%
Pro (P)	44	11.2%
Ser (S)	38	9.7%
Thr (T)	22	5.6%
Trp (W)	4	1.0%
Tyr (Y)	9	2.3%
Val (V)	19	4.8%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 50
Total number of positively charged residues (Arg + Lys): 47

Atomic composition:

Carbon	C	1902
Hydrogen	H	2991
Nitrogen	N	551
Oxygen	O	592
Sulfur	S	22

Formula: $C_{1902}H_{2991}N_{551}O_{592}S_{22}$
Total number of atoms: 6058

Extinction coefficients:
Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 36035
Abs 0.1% (=1 g/l) 0.824, assuming all pairs of Cys residues form cystines

Ext. coefficient 35410
Abs 0.1% (=1 g/l) 0.809, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 72.64

This classifies the protein as unstable.

Aliphatic index: 59.82**Grand average of hydropathicity (GRAVY):** -0.752

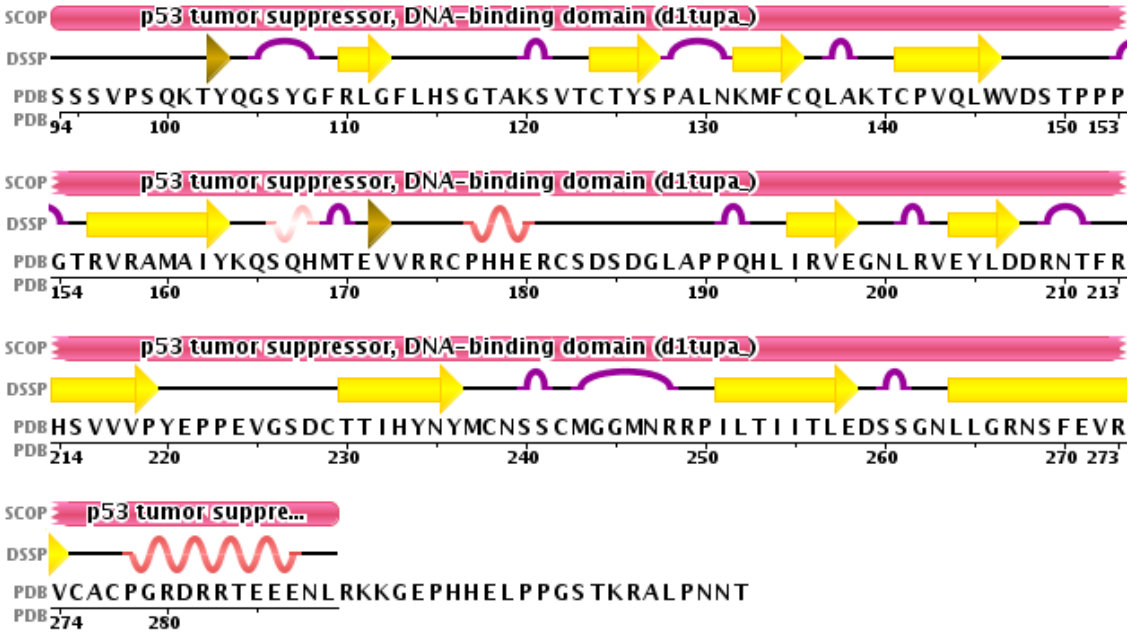
Secondary structure of mutant p53

```
        10         20         30         40         50         60         70
         |          |          |          |          |          |          |
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGPDEAPRMPEAA
ccccccccccccccchhhhhhhhhhccccceecccccchhhhhhecccccccccccccccccccccccc
PRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKT
ccccccccccccccccccccccccccccccccccccccceeeeeecccccccceeecchhhhhhhhhhttc
CPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEVNLRVEYLDDRN
cceeeeeccccttceeeeeeeecchhhhhhhhccccccccccccccccccheeeetccccheeeccccc
TFRHSVVVPYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGR
cccceeeeecccccccchhheeeehhhccccccccccceeeeeeecctteeeccceeeeeeecctte
DRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALEL
ccchhhhhccccccccccccccccccccccccccccccccccccccceeeeeecccchhhhhhhhhhhhhh
KDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD
Hhhhccccccccccchheeccccccccttcceeeeeccccccc
```

Sequence length : 393

This result shows the secondary structure of amino acids according to their folding.

PDB secondary structure of mutant p53

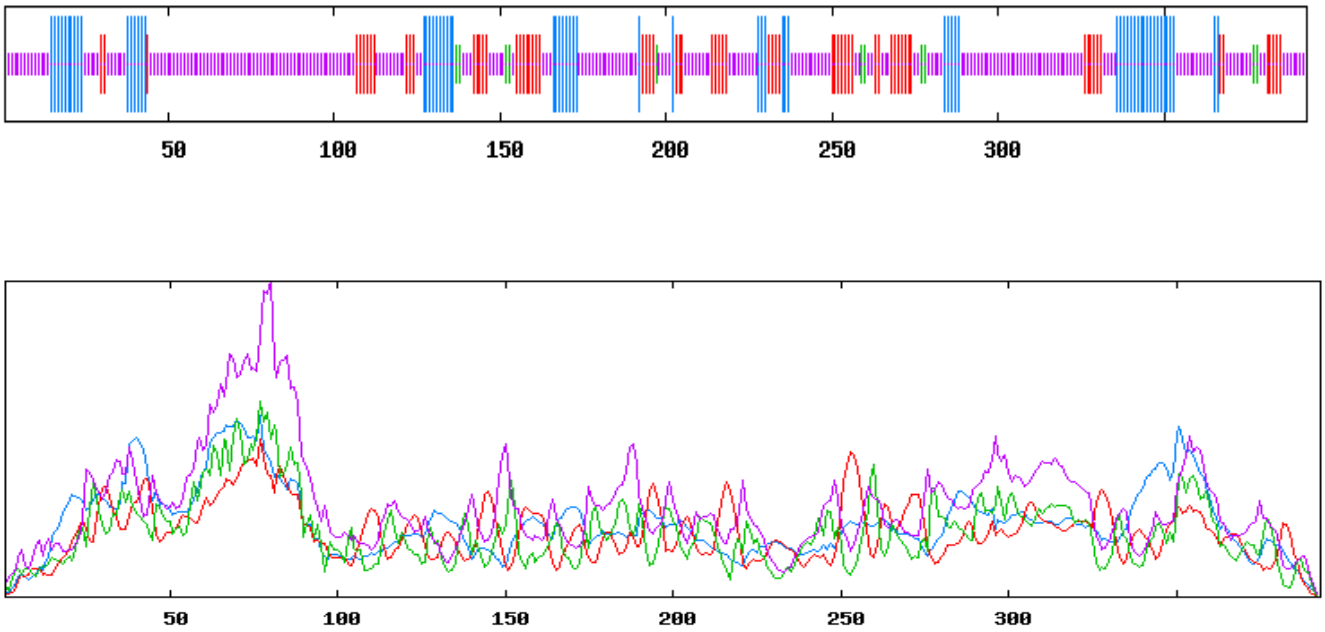


(fig-8)

Alpha helix (Hh) : 67 is 17.05%
3_{10} helix (Gg) : 0 is 0.00%
Pi helix (Ii) : 0 is 0.00%
Beta bridge (Bb) : 0 is 0.00%
Extended strand (Ee) : 70 is 17.81%
Beta turn (Tt) : 11 is 2.80%
Bend region (Ss) : 0 is 0.00%
Random coil (Cc) : 245 is 62.34%
Ambiguous states (?) : 0 is 0.00%
Other states : 0 is 0.00%

Parameters :

Window width : 17
Similarity threshold : 8
Number of states : 4

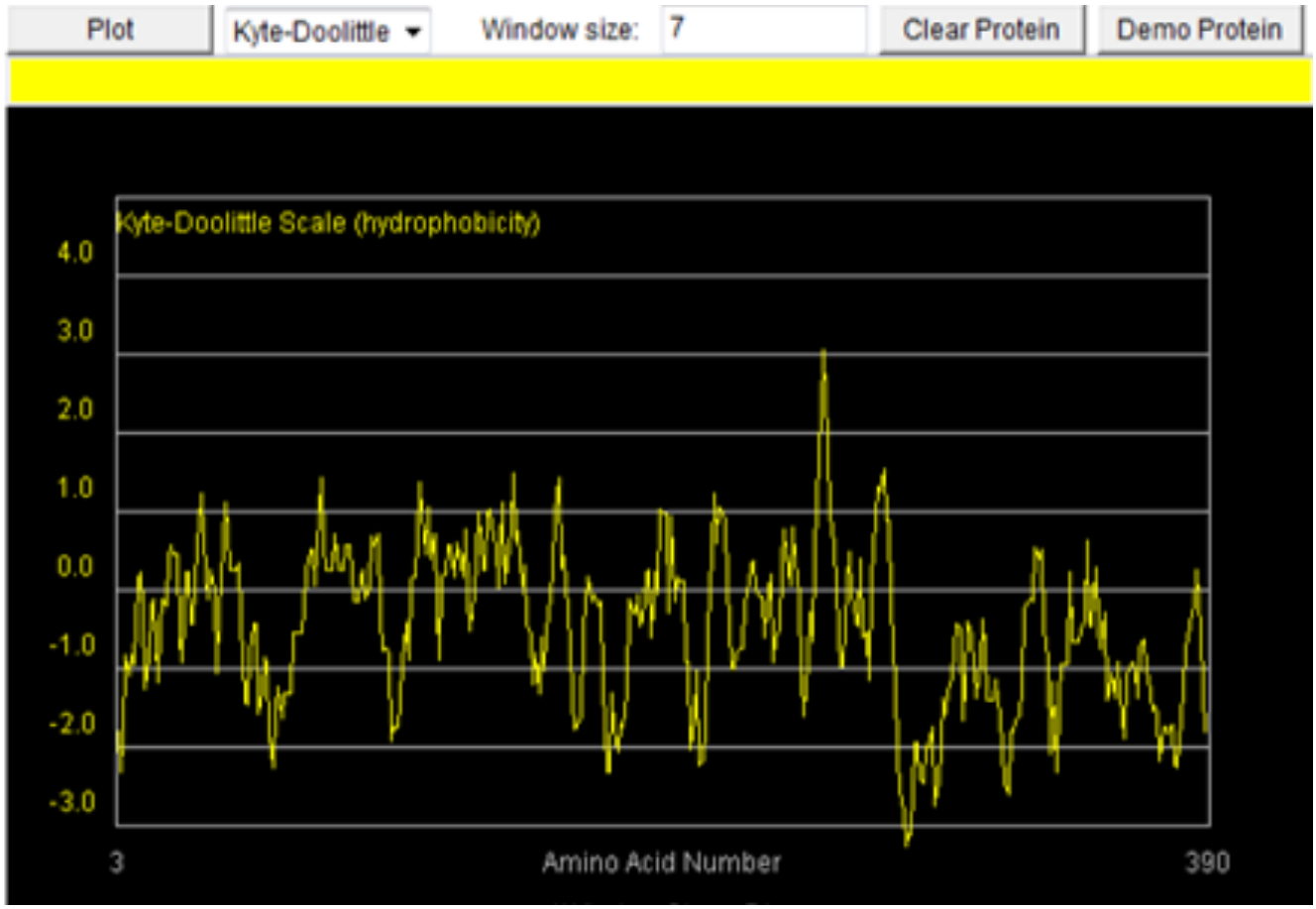


I-TASSER:- online 3dimensional structure prediction of (fig-9)

(fig-10)

The sequence of mutant p53 obtained from PDBid and i-tasser was compared same as wild type p53.

Hydrophobicity plot of mutant p53



(fig-11)

Single hydrophobic patch was seen which was similar to wild type p53, but with greater hydrophocity.The patch is predicted as shown in table-

Table (5): predicted hydrophobic patch of mutant p53

Transcription Factor	Predicted Hydrophobic Patch	Location of Patch
Mutant p53	GMNRRP**ILTIITL**EDSSGNLLG	Middle domain

Alignment of p53 with mutant p53

99.2% identity in 393 residues overlap; Score: 2097.0;Gap frequency: 0.0%

```
1   MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP
1   MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP
    ************************************************************

61  DEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAK
61  DEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAK
    ************************************************************

121 SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHE
121 SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHE
    ************************************************************

181 RCSDSDGLAPPQHLIRVEVNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNYMCNS
181 RCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNYMCNS
    ****************** ****************************************
241 SCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGRDRRTEEENLRKKGEPHHELP
241 SCMGGMNRRPILTIITLEDSSGNLLGRNSFEVHVCACPGRDRRTEEENLRKKGEPHHELP
    ******************************** ****************************

301 PGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG
301 PGSTKRALSNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG
    ******** ****************************************************

361 GSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD
361 GSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD
    *********************************
```

These 3 amino acid pair that was not alike shown single base substitution mutation in human p53.This mutation was responsible for change in their interaction property with Hsp90.

Table (6) mutation point in amino acid sequence

Wild type p53	Mutant p53	**Position of amino acid**
G - Glycine	V- Valine	199
H –Histidine	R –Arginine	273
P- Proline	S –Serine	309

CHARGE DISTRIBUTIONAL ANALYSIS

With the help of **SAPS** tool, charge distribution of amino acid inside human Hsp90 was obtained. The distribution of charges in the protein sequence was evaluated in terms of clusters, hydrophobic patches and high scoring segments of human p53.

Table (7) Correlation of hydrophobic patches with score value:

Hydrophobic patches	Score
(LVIFM)	2.00 maximum score of hydrophobicity in amino acid
(AGYCW)	1.00 (50%)hydrophobicity
(BZX)	0.00 no hydrophobic patches
(PH)	-2.00
(STNQ)	-4.00 hydrophilic
(KEDR)	-8.00 maximum hydrophilic

Correlation between hydropbobic patches and charge distribution of Hsp90 amino acid sequence

Shown in [table 1&3] human p53& Hsp 90 both had opposite charge distribution due to their amino acid sequence. These apposite charges on each other attract to reduce their electrostatic energy. These charge distribution was also responsible for making extra and intra H-bond with each other and water molecules. In protein folding and docking with other protein H-bolding is important characteristic. Hydrophobic property of protein that developed due to hydrophobic patches involved main role in interaction of p53 with Hsp90.

Table(8)

positive (KR)	KR:122 (14.3%)	Shows 14.3% of positive charge
negative (ED)charge	ED:160 (18.7%)	Shows 18.7% of negative charge this represented more negative score of this protein
total charge (KRED)	KRED:282 (33.0%)	
net charge (KR-ED)	KR-ED:-38(4.4%)	Hsp90 protein negatively charged
Major Hydrophobic (LVIFM)	LVIFM:224 (26.2%)	Hydrophobic patches confine to LVIFM.

Table (9) Correlation between hydrophobic patch and charge distribution of p53 amino acid sequence:

positive (KR)	KR:46 (11.7%)	Shows 11% of positive charge
negative (ED)charge	ED:50 (12.7%)	Shows 12% of negative charge this represented more negative score of this protein
total charge (KRED)	KRED:96(24.4%)	Total charge due to positively & negatively charge amino acid in p53
net charge (KR-ED)	KR-ED :-4(1.0%)	Negative charge score of p53 protein
Major Hydrophobic (LVIFM)	LVIFM:81(20.6%)	Hydrophobic patches confine to LVIFM.

positive (KR)	KR:47 (12.0%)	Shows 12% of positive charge
negative (ED)charge	ED:50 (12.7%)	Shows 12.7% of negative charge this represented more negative score of this protein
total charge (KRED)	KRED:97 (24.70%)	
net charge (KR-ED)	KR-ED : -3(-0.8%)	Mutant p53 protein negatively charged
Major Hydrophobic (LVIFM)	LVIFM:82 (20.9%)	Hydrophobic patches confine to LVIFM.

Table (10) Correlation between hydrophobic patch and charge distribution of mutant p53 amino acid sequence

The result of charge distribution and hydrophobic patch of wild type p53 and mutant p53 when compared showed that mutant p53 was more positive by one positive charge and slightly more hydrophobic than wild type 53, overall % hydrobhobicity being more for mutant type p53.This was used to predict that mutant p53 can interact better with Hsp90 than wild type p53.

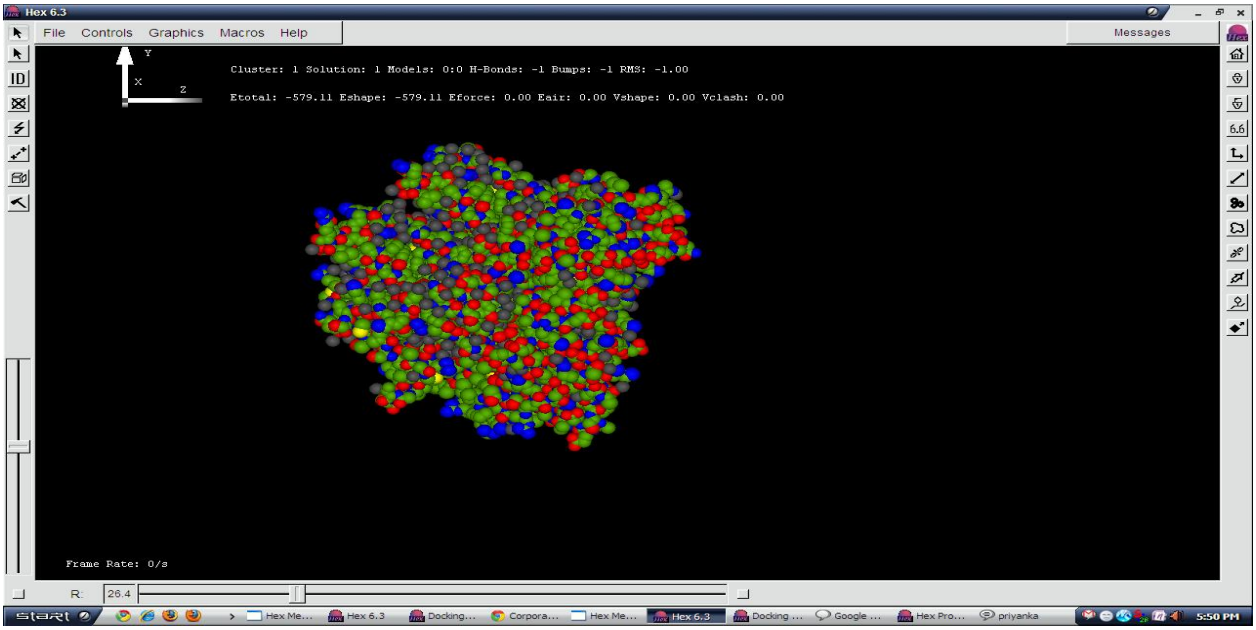
Interaction between hsp90 and p53 and Interaction between human Hsp 90 and mutant p53 complex by hex (protein -protein docking)

Hsp90 co-chaperone	% Overall hydrophobicity	Predicted Hydrophobic patch	Hydropathy index of the predicted patch	% Similarity between predicted hydrophobic patch and **LEVIYMI** on Middle domain of Hsp90 beta	% Similarity between predicted hydrophobic patch and **LVILLYE** on C Terminal domain of Hsp90 beta	% Similarity between predicted hydrophobic patch and **LMSLIIN** on N Terminal domain of Hsp90 beta
Wild type p53	20.6%	**TIITLED** (Middle domain)	1.783	42.86%	42.86%	42.86%

Table 11: Interaction between Hsp90 and wild type p53

For docking process BLAST,a similarity based search tool was used which compare query sequence with the sequence stored in server and the sequence with highest score was considered for the PDB id .In case of PDB id of wild type and mutant p53 .sequence are much similar therefore they give similar sequence PDBid,but from the list of sequences first five was considered and PDB id which is not similar and with highest score was taken. PDB id is required for hex docking.Hex.6 is a software which help in docking of protein which predicts most stable interaction with specific orientation on the basis of total free energy, and orientation with less E-value give more stable protein-protein interaction. On the basis of total free energy it was found that interaction of hsp90 with wild type p53 was more stable than mutated p53.

Figure (12): Hsp90–mutant p53 (PDB id-1TSR) complex having docking energy -579.67 .



Figure(13): Hsp90–wild p53 type (PDB id 2BIM) complex having docking energy -1061.67

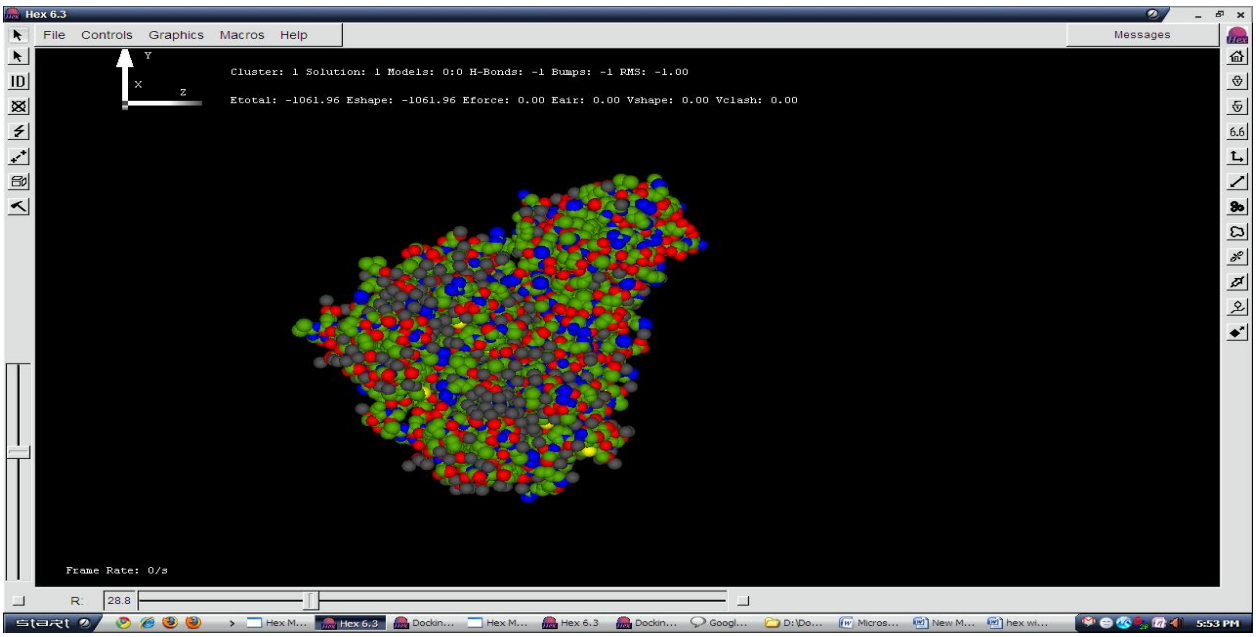


Figure (14):Hsp90-wild type p53 complex having docking energy -516.16

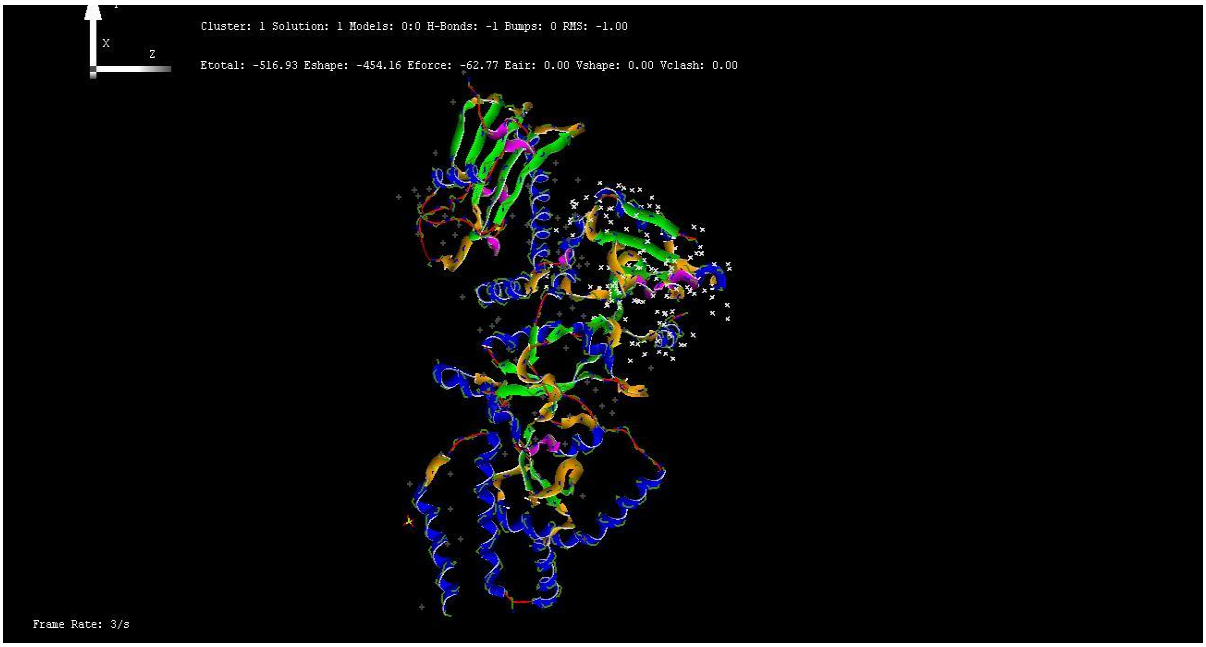
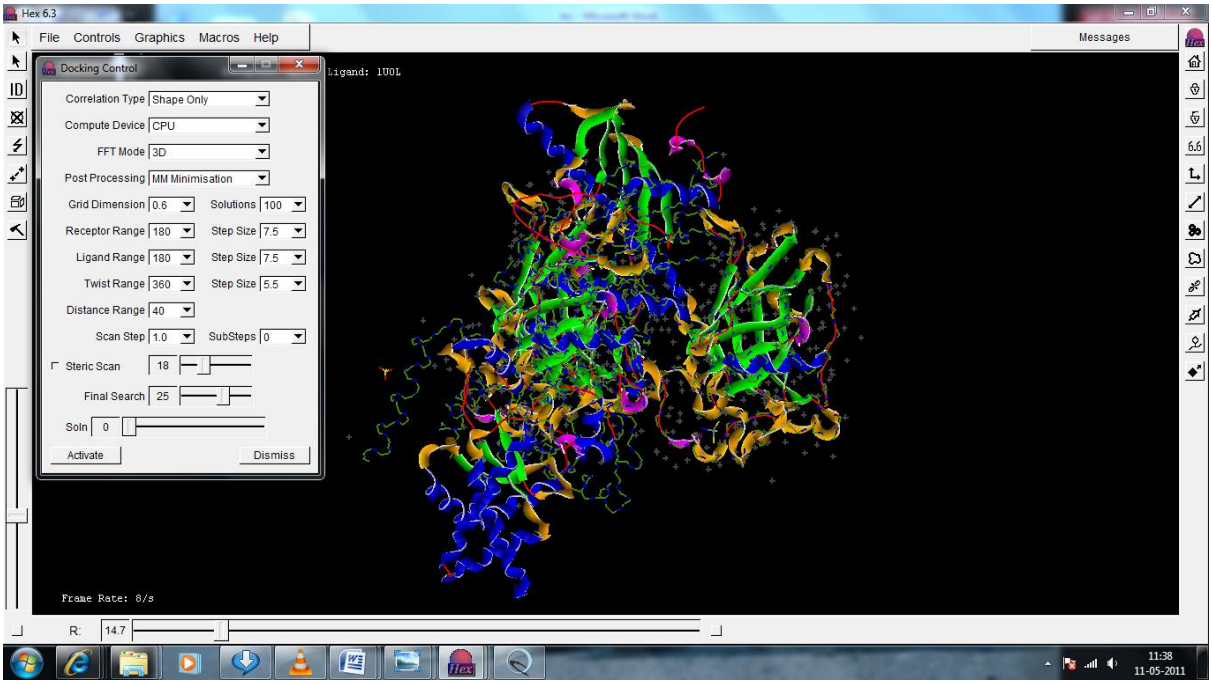


Figure (15):Hsp90-Mutant p53 complex having docking energy 510.1.



Observation of inter molecular and intra molecular H-bonding in both figure

Figure (16-A) Hsp90-wild type p53

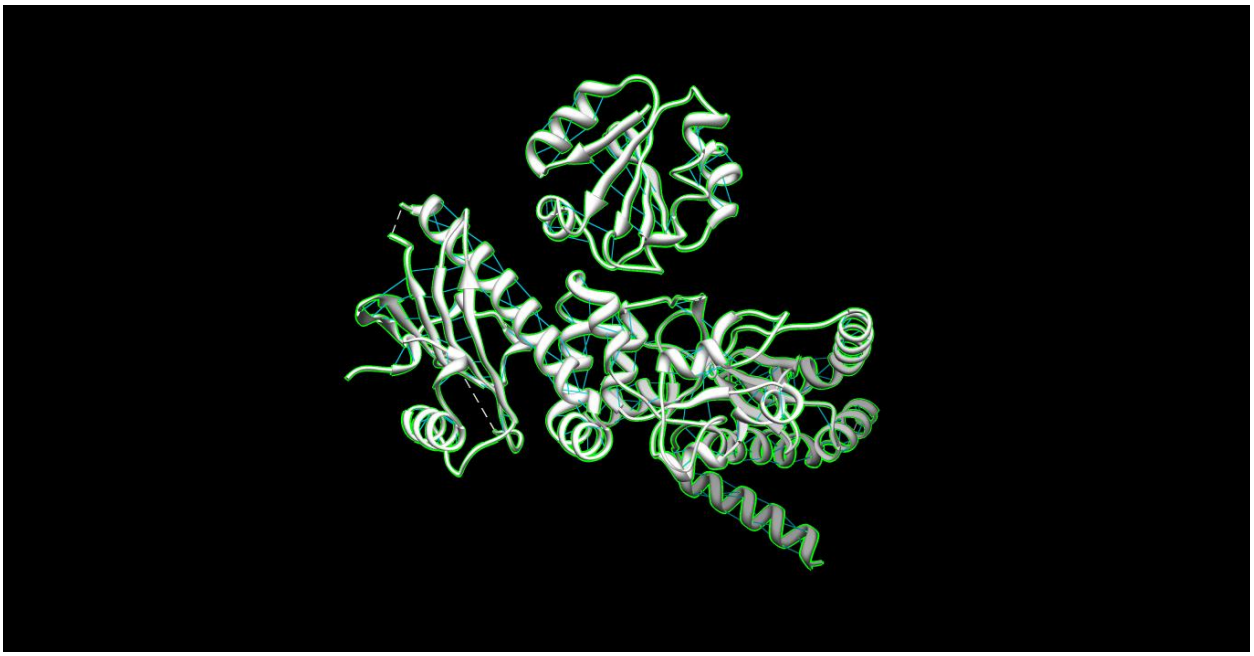
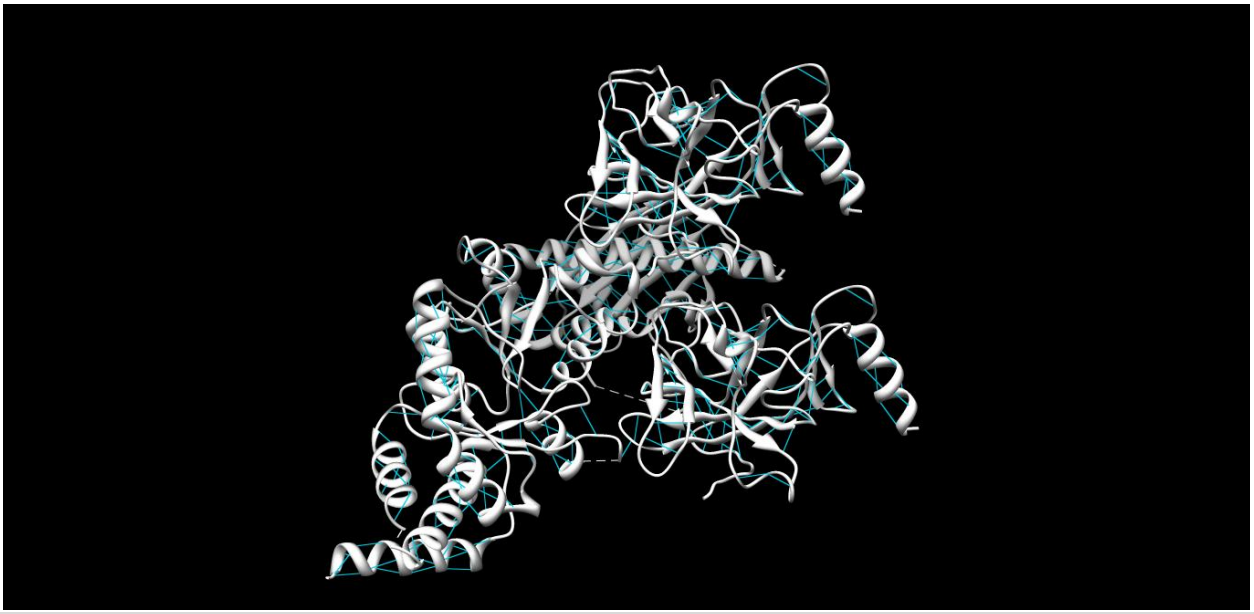


Figure (16-B) Hsp90-mutant p53



CONCLUSION

The results from surface hydrophobicity plot, docking and charge distribution analysis of Hsp90, wild type p53 and mutant p53, each parameter were compared .The result from hydrophobicity plot and charge distribution analysis showed that mutant should interact with hsp90 better than wild type. The docking result showed that binding efficiency of mutant p53 is less with respect to wild type p53.This anomalous observation can be explained as it might be that due to more hydrophobic patches it sometimes hide in the core of protein or the charge repulsion might have occurred, or different type of folding in mutant p53 may be a reason for less interaction .Due to less interaction potential of mutant p53 ,it cannot properly control the cell cycle regulation mechanism and initiation of apoptotic cell death the way wild type p53 does which may lead to the growth of tumour or cancer. Wild type p53 is a liable protein with high turnover rate; it is degraded via proteolysis and the ubiquitin-proteasome pathway. Therefore only low concentration of p53 resides in cell. The mutated form is not affected by proteolysis or ubiquitination and accumulates in the cell causing tumour growth.

REFERENCES

1. Murakami,Y. & Jones, S. (2005) SHARP^2 ; Protein-Protein Predictions using Patch Analysis. Bioinformatics, submitted.
2. Murakami,Y. (2005) Implementation of a protein-protein interaction prediction method as a bioinformatics tool on the Internet. MSc Bioinformatics thesis. Department of Biochemistry, University of Sussex, Falmer, BN1 9QG, United Kingdom.
3. http://www.bioinformatics.sussex.ac.uk/SHARP2/send_result.html
4. http://www.bioinformatics.sussex.ac.uk/SHARP2/help_parameters.html
5. http://www.ncbi.nlm.nih.gov/pubmed/7462208
6. http://zhanglab.ccmb.med.umich.edu/I-TASSER/
7. http://expasy.org/tools/protparam.html
8. http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html
9. http://expasy.org/tools/
10. http://www.vivo.colostate.edu/molkit/hydropathy/scales.html
11. http://www.ncbi.nlm.nih.gov/protein/BAC16799.1
12. http://www.ncbi.nlm.nih.gov/protein?term=hsp90%20homo%20sepiance
13. http://www.bioinformatics.sussex.ac.uk/SHARP2/help_parameters.html
14. Walerych, D., Kudla, G., Muller, L., King, F. W., Helwak, A., Boros, J., Zylicz, A., Zylicz, M**.** (2004) **Hsp90 Chaperones Wild-type p53 Tumor Suppressor Protein.** *J. Biol. Chem.* 279: 48836-48845
15. Buchner, J., Scheibel, T. (1998) **Two chaperone sites in Hsp90 differing in substrate specificity and ATP dependence.** Proc. Natl. Acad. Sci. USA 95:1495-1499.
16. Jennifer, R. S., Paul, W. (2007) **Targeting the cancer chaperone Hsp90** Drug Discovery Today: Therapeutic Strategies 19, 207–236.

17. Kyte, J., Doolittle, R. E. (1982). J Mol Biol 157:105-132.
18. Mayer, M. P., Prodromou, C., Frydman, J. (2009) **The Hsp90 mosaic: a picture emerges** Nature, Pages 106-119.
19. Muller, P., Ceskova, P., Vojtesek, B. (2005) **Hsp90 Is Essential for Restoring Cellular Functions of Temperature-sensitive p53 Mutant Protein but Not for Stabilization and Activation of Wild-type p53: implications for cancer therapy.** *J. Biol. Chem.* 280: 6682-6691.
20. Zhang, Y. (2008) **I-TASSER server for protein 3D structure prediction** BMC Bioinformatics
21. Tomala, K., Korona, R. (2008) **Molecular chaperones and selection against mutations** Biology Direct, 3:5.
22. http://www.elmhurst.edu/~chm/vchembook/561aminostructure.html