ORIGINAL ARTICLE Open Access

ISSN 2278-1404

International Journal of Fundamental & Applied Sciences

An *in silico* based characterization and analysis of Human matrix metalloproteinases (MMPs)

Pritam Kumar Panda*, Selvaa Kumar C*

School of Biotechnology and Bioinformatics, D.Y. Patil University, CBD Belapur, Navi Mumbai-400614, India

Manuscript received 02nd Feb, 2015, revised 20th March, 2015, accepted 26th March, 2015

Abstract

Background and aims: Potentially involved proteins which are implicated as a specific target for any diseased condition may implicate certain unusual features in several pathological conditions. Human Matrix metalloproteinase (MMP) family of endopeptidases is one such family responsible for many beneficial as well as several pathological critical diseases. With the advent of field of bioinformatics and computational efforts can aid researchers to comprehend their system of work. **Methodology:** An insilico characterization of the MMP family has been carried out to analyze their primary, secondary, structural and functional perspective . The research has been focused on specific MMPs in which the further study was based on Mutational analysis confirming the pathogenicity of MMPs in cancer metastasis. The basic approach was to screen large protein families which plays dual role during normal and diseased conditions. **Results:** Thus it is hypothesized that cysteine rich and highly thermostable MMPs might be key players in diseased conditions. **Conclusion:** It can also be concluded that the disease responsive MMPs might be considered as promising targets for treatment of cancer.

Keywords: MMP; Physio-chemical analysis; MSA; SOPMA; I-Mutant; Polyphen

1. Introduction

MMPs are a family of zinc containing endopeptidases, which is a subset of the metzincin superfamily of metalloproteinases. These regulatory proteases are the extracellular matrix (ECM) remodelers characterized by their substrate specificity to degrade ECM proteins[1]. Based on this, they have been classified as collagenases, gelatinases, stromelysins, matrilysins, membrane type MMPs (MT-MMPs) and other unclassified MMPs which are responsible for the tissue remodeling and degradation of the extracellular matrix (ECM), including collagens, elastins, gelatin, matrix glycoproteins, and proteoglycan[2]. They are regulated by hormones, growth factors, and cytokines, and are involved n ovarian functions. Structurally, MMPs consist of four domains: an amino terminal hydrophobic pro- domain, a Zn2+ containing catalytic domain, a flexible hinge region and a carboxy terminal hemopexin-like domain responsible for their substrate specific nature . Out of the 26 MMPs reported till date, 23 have been identified in humans. Our study reports an *in silico* comparative characterization and analysis of human MMPs using various bio-computational tools, pertaining to their physico-chemical, secondary structural and functional features[3]. Any typical but significant feature may have various connotations with respect to the role of MMPs in pathological conditions. The aim here is to identify potential disease responsive MMPs that might possibly be implicated for their role in diseases. Moreover, such an in depth knowledge of all human MMPs would greatly aid researchers to identify the MMPs of interest relevant to their respective working systems. This would further set a precedent for similar comparative characterization studies for other large protein families, using the numerous resources from the field of computational biology..The general applicability of the "cysteine-switch"[4] activation mechanism to the members of the matrix metalloproteinase (MMP) gene family share the characteristic that they are synthesized in a latent, inactive, form. Recent evidence suggests that this latency in human fibroblast collagenase (HFC) is the result of formation of an intramolecular complex between the single cysteine residue in its propeptide domain and the essential zinc atom in the catalytic domain, a complex that blocks the active site. This is referred to as the "cysteine-switch" mechanism of activation. The propeptide domain that contains the critical cysteine residue and the catalytic domain that contains the zinc-binding site are the only two domains common to all of the MMPs. The amino acid sequences

*Corresponding author
Full Address :
School of Biotechnology and Bioinformatics, D.Y. Patil University, CBD Belapur, Navi Mumbai-400614, India
Phone no. +91-22-27563600
E-mail: pritampkp15@gmail.com

surrounding both the critical cysteine residue and a region of the protein chains containing two of the putative histidine zinc- binding ligands are highly conserved in all of the MMPs.

2. Materials And Methods

2.1 Protein sequence retrieval

UniProtKB/Swiss-Prot[5], a high quality manually annotated and non-redundant protein sequence database, was used to retrieve the complete sequences of the 23human MMP's. These sequences were used for further analysis using various online bio-computational tools.

2.2 Phyologenetic Analysis

Based on Multiple Sequence alignment the phylogenetic unrooted tree has been deduced using Phylip DRAWGRAM of SDSC Biology workbench using MSA of MMPs and Cladogram of Clustal W[6] MSA Tool (**Suppl. Figure I**).

2.3 Physico-chemical analysis

The computation of various physical and chemical parameters, such as amino acid composition, molecular weight, isoelectric point (pI), total number of negative and positive charged residues, extinction coefficient, instability index, aliphatic index and Grand Average of Hydropathy (GRAVY), was done using ExPASy's ProtParam tool[7] (http://us.expasy.org/ tools/ protparam.html).

2.4 Secondary structural analysis

SOPMA tool (Self-Optimized Prediction Method with Alignment)[8] of NPS@ (Network Protein Sequence Analysis) server was used to characterize the secondary structural features of the proteins such as, alpha helix, 310 helix, Pi helix, beta bridge, extended strand, beta turn, bend region, random coil, ambiguous states and other states.

3. Results And Discussion

3.1 Primary Structure Analysis

MMPs are secreted in latent form as pro-MMPs and these zymogens are required to be cleaved for activation. They are found to exhibit pro and active forms, characterized by a difference in molecular weights. Out of all the 23 characterised MMPs 13 MMPs were stable and 10 MMPs are unstable according to Primary structure analysis (**TableI**). This indicates that although all the MMPs are considered to be involved in several chronic diseases, the unstable MMPs might involved in cancer related diseases as such for instance MMP9 (**Figure I**) which can easily activate all other MMPs and vice versa.

TableI: Protparam Analysis/Physiochemical Properties

MMP	Number of amino acids:	Molecular weight	Theoretical pI	instability index	Aliphatic index	hydropathicity (GRAVY)	Formula
2	660	73882.3	5.26	27.46 Stable	61.09	-0.446	$C_{3343}H_{4973}N_{861}O_{979}S_{31}$
1	469	54006.9	6.47	35.46 Stable	65.27	-0.572	$C_{2458}H_{3666}N_{656}O_{699}S_{13}$
9	707	78458.2	5.69	41.10 Unstable	65.13	-0.394	$C_{3517}H_{5298}N_{958}O_{1035}S_{28}$
8	467	53412.1	6.38	29.47 Stable	69.16	-0.451	$C_{2437}H_{3628}N_{640}O_{698}S_{11}$
7	267	29676.8	7.73	32.39 Stable	76.37	-0.369	$C_{1330}H_{2053}N_{361}O_{387}S_{12}$
3	477	53977.3	5.77	27.90 Stable	76.65	-0.386	$C_{2461}H_{3740}N_{638}O_{708}S_{12}$
10	476	54152.2	5.49	34.39 Stable	71.74	-0.370	$C_{2476}H_{3690}N_{632}O_{710}S_{15}$
11	488	54589.9	6.38	51.62 Unstable	74.08	-0.332	$C_{2503}H_{3736}N_{686}O_{679}S_{9}$
12	470	54001.5	8.75	25.91 Stable	73.23	-0.383	$C_{2478}H_{3716}N_{644}O_{687}S_{15}$
13	471	53819.8	5.32	20.36 Stable	73.10	-0.435	$C_{2456}H_{3657}N_{629}O_{707}S_{16}$
14	582	65839.9	7.63	47.29 Unstable	67.73	-0.465	$C_{2992}H_{4517}N_{805}O_{841}S_{21}$
15	669	75806.9	7.03	44.36 Unstable	67.34	-0.567	$C_{3429}H_{5167}N_{961}O_{953}S_{22}$
16	607	69521.4	8.72	36.13 Stable	69.51	-0.484	$C_{3179}H_{4786}N_{842}O_{874}S_{23}$
17	603	66652.8	6.08	46.90 Unstable	68.69	-0.454	$C_{2998}H_{4510}N_{842}O_{862}S_{16}$
19	508	57356.9	7.22	35.90 Stable	71.44	-0472	$C_{2597}H_{3936}N_{704}O_{741}S_{15}$
20	483	54386.8	8.92	26.10 Stable	68.67	-0.407	$C_{2469}H_{3747}N_{653}O_{704}S_{17}$
21	569	65043.5	9.19	45.06 Unstable	67.47	-0.543	$C_{2961}H_{4430}N_{824}O_{818}S_{12}$
23	390	43934.8	9.94	55.08 Unstable	83.59	-0.183	$C_{1990}H_{3072}N_{582}O_{517}S_{16}$
28	520	58939.0	9.70	49.47 Unstable	75.50	-0.444	$C_{2676}H_{4064}N_{772}O_{723}S_{10}$
24	645	73231.4	9.30	50.21 Unstable	68.53	-0.562	$C_{3328}H_{5052}N_{928}O_{912}S_{18}$
25	562	62554.0	8.76	51.82 Unstable	69.48	-0.449	$C_{2826}H_{4305}N_{795}O_{791}S_{15}$
26	261	29708.4	5.96	29.25 Stable	75.44	-0.385	$C_{1351}H_{1993}N_{361}O_{383}S_{9}$
27	513	59025.8	8.83	25.03 Stable	77.95	-0.278	$C_{2712}H_{4110}N_{704}O_{740}S_{19}$

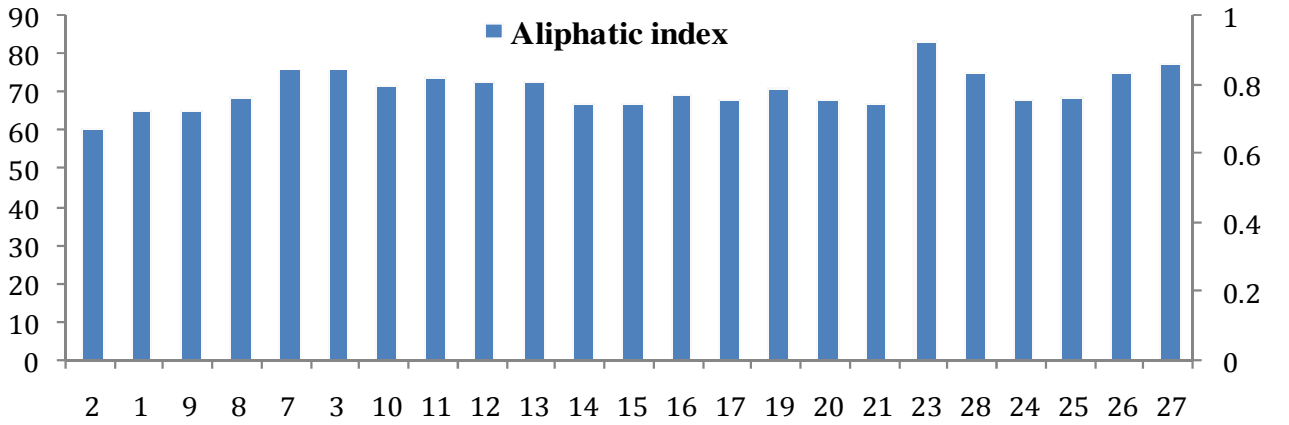


Figure I: Computation of aliphatic index by ExPASy's ProtParam tool.

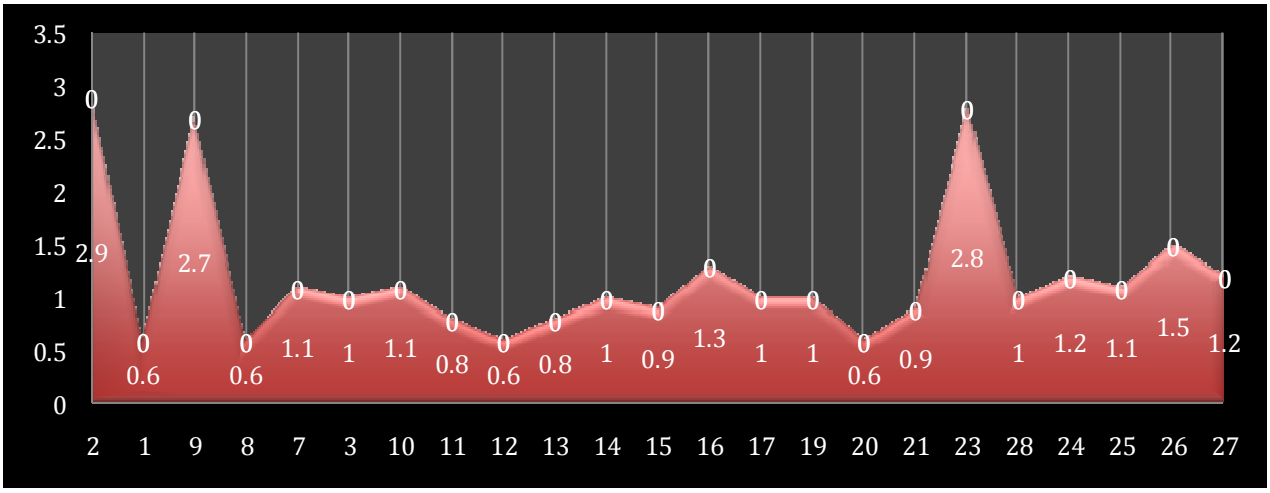


Figure II: Percentage of cysteine residues in human MMPs computed by ExPASy's ProtParam tool.

3.2 Analysis of Cysteine Residues in All MMPs

Analysis of amino acid composition indicates that while the percentage of cysteine residues in majority of MMPs lies in the range of 0.6-1.3%, MMP-2, 9 and 23 show a significant rise with values 2.9, 2.7 and 2.8 percent, respectively (**Figure II**). High percentage of cysteine residues in MMP-2 and 9 might be correlated with presence of cysteine switch motif and role of these MMPs in pathological conditions. These gelatinases have been previously implicated in carcinomas and cardio-vascular disorders. High cysteine content of the unclassified MMP-23 might be attributed to the presence of cysteine array in its structure. Highly significant presence of cysteine suggests its role as a critical residue for MMP activity and thus these MMPs may be investigated for possible role in diseased conditions. Further analysis of the amino acid composition can help to locate amino acid presence at an unusual level and be correlated with specific pathological conditions.

3.3 Secondary Structure Analysis

The secondary structure analysis was done of all 23 MMPs using SOPMA to determine the composition of alpha helices, beta strands and coils which inturn determine the stability of the structure of MMPs derived from their sequence perspective (**Figure III, Figure IV**) which is represented graphically to identify the highest residual property containing secondary structural elements.

3.4 Classification of MMPs

The protein structure of the MMPs: Matrix metalloproteinases (MMPs) can be divided into eight distinct structural groups, five of which are secreted and three of which are membrane-type MMPs (MT-MMPs). Secreted MMPs: The minimal-domain MMPs contain an amino-terminal signal sequence (Pre) that directs them to the endoplasmic reticulum, a propeptide (Pro) with a zinc -interacting thiol (SH) group that maintains them as inactive zymogens and a catalytic domain with a zinc-binding site (Zn).

3.5 Blast analysis

From BLAST[9] Analysis it has been deduced that the Matrix Metallo protein families belongs to Humans as well as other eukaryotic organisms which are homologous to Homosapiens in terms of functions (**Table II**). The sequences of MMP have been retrieved using Swissprot /Uniprot of reviewed one of different organisms and alignment was done to retrieve phylogeny lineage among them. Since from the BLAST analysis it

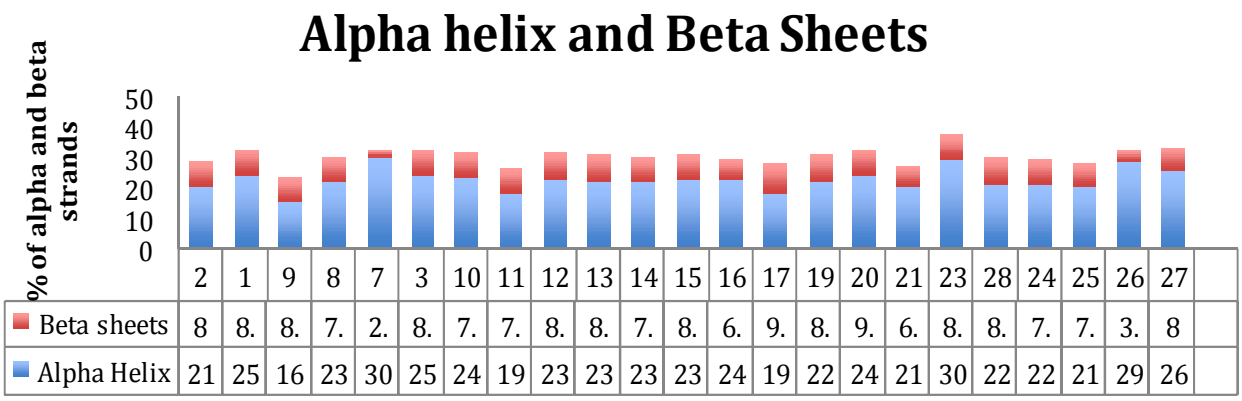


Figure III :Graphical representation of Alpha helix and beta sheets composition in MMPs

Figure IV :Graphical representation of Coils and extended strans composition in MMPs.

Table II:Compositional Matrix Adjustments in BLASTp

Sl. no.	Organism	Identity	Query Coverage
1	*Rattus norvegicus*	95	100
2	*Mus musculus*	95	100
3	*Bos taurus*	96	100
4	*Gallus gallus*	83	95
5	*Canis lupus familiaris*	45	83
6	*Equus caballus*	42	93
7	*Xenopus laevis*	43	68
8	*Sus scrofa*	86	85
9	*Fetis catus*	42	63
10	*Sachaaromyces Cerevicae*	39	68

Table III :Residue substitution in aminoacid composition of MMPs and its effect

MMP	Natural Variant/ Mutagenesis	Residues	Effect
MMP9	279-279	Q→R	Common polymorphism; may be associated with susceptibility to IDD.
	402-402	E → Q	Loss of activity.
MMP2	101-101	R→H	Multicentric osteolysis and arthritis syndrome
	400-400	Missing Residue	Winchester syndrome
	404-404	E→K	Winchester syndrome

has been observed that the residues present in Human MMPs are also being observed in several organisms listed below and they are highly significant in respect to their identity and similarity based on compositional matrix.

3.6 Phylogenetic Analysis

From the compositional matrix derived from Blastp, the organisms which are orthologous to human MMPs are considered to identify the clustal distance (**Suppl. Figure I**) and their classification group in which they are categorized for further analysis of clinical trials to identify the specific role of MMPs in cancer metastasis.

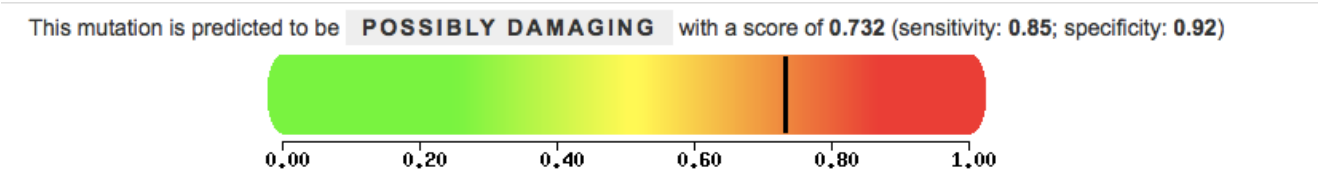
3.7 Specific MMPs role in cancer

MMPs play a salient role in cancer. MMP-2 and MMP-9

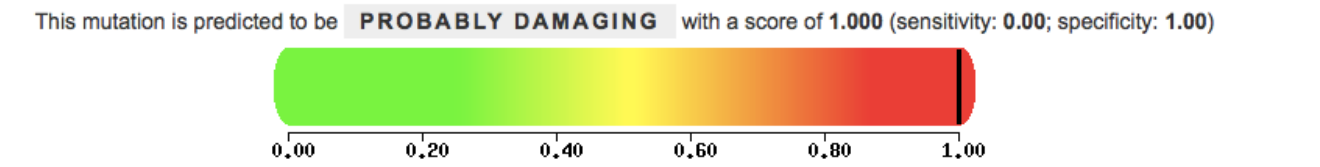
Table IV: Imutant 2.0 mutational analysis/stability effect upon mutation.

MMP	Position	WT	NEW	Stability	RI	pH	Temp
2	101	R	H	Decrease	8	7	25
	400	V	All resi-dues	Decrease	Varies	7	25
	401	E	K	Decrease	2	7	25
9	279	Q	R	Decrease	5	7	25
	402	E	Q	Decrease	0	7	25

R→H



E → Q



E→K

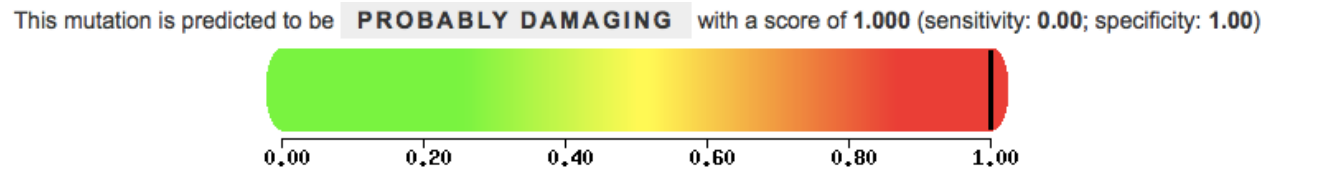


Figure V: Polyphen Analysis:Prediction of mutation specificity and sensitivity

(72kDa gelatinase and 92kDa gelatinase) are the prominent MMPs responsible for basement membrane ECM protein degradation that facilitates the migration of tumor cells to blood vessels. In this regard, specifically designed synthetic MMP inhibitors are not likely to prove efficacious as a cancer therapy if they interfere with anti-angiogenesis pathways or immune-mediated tumor killing. So MMP9 And MMP2 has been chosen for the recent study purpose to understand the structural and sequential modifications occurs during cancer metastasis.

3.8 Mutational Analysis

The mutational aspect or the residual substitution derived from uniprot which is annotated (**Table IV**) was derived and mutational analysis was done to cross reference the results. The mutational analysis was carried out by I-Mutant2.0[10] and polyphen to confirm the mutagenesis, to check the stability of proteins (**Table III**).T he polyphen[11] mutational tool has been performed to find the mutational probability of its sensitivity and specificity (**Figure V**) towards the damaging property to the protein.

4. Conclusion

The dual role of MMPs in normal and diseased state yielded a new insight and perspectives which can be used to identify the main cause of cancer which can be emerged as a therapeutic target by using computational biology. The comparative analysis and characterization of all 23 MMPs has been carried out in context of sequential analysis. This aid the researchers to experimentally determine the state of disease caused and to know the actual mechanism of MMPs alteration *in vitro*. Since the characterization shows that the MMP2,MMP9 might be a key player in pathological conditions and stability analysis was done which aid the researchers to target specific MMPs. The mutational analysis of MMP2 and MMP9 has been done using computational tools to study the residual substitution effect during pathological conditions and it has been proved right hypothetically. Further studies with the help of experimental research and testing need to be carried out to validate this proposal. Additionally, this study may be taken as a prototype for similar experimental investigational studies with regard to several proteins

involved in cancer metastasis, wherein such characterization might aid in giving a direction to further research in the cure of cancer.

Conflict of interest

The author' s declares none.

References

1. Jaiswal A, Chhabra C, Malhotra, U, Kohli S, and Rani V.Comparative analysis of human matrix metalloproteinases: Emerging therapeutic targets in diseases. *Bioinformation*. **6(1)** (2011) 23–30.
2. Bäck M, Ketelhuth DF, Agewall S. Matrix metalloproteinases in atherothrombosis, *Progress in Cardiovascular Diseases.* **52(5)** (2010) 410-28.
3. Verma RP, Hansch C., Matrix metalloproteinases (MMPs): chemical-biological functions and (Q) SARs. *Bioorganic and Medicinal Chemistry*. **15(6)** (2007) 2223-68.
4. Van Wart HE, Birkedal-Hansen H. The cysteine switch: a principle of regulation of metalloproteinase activity with potential applicability to the entire matrix metalloproteinase gene family. *Proceedings of National Academy of Sciences USA*. **87(14)** (1990) 5578-82.
5. The UniProt Consortium. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **42** (2014) D191-D198 .
6. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ and Higgins DG. ClustalW and ClustalX version 2 *Bioinformatics* **23(21)** (2007) 2947-2948.
7. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server;(In) John M. Walker (ed): *The Proteomics Protocols Handbook*, Humana Press (2005);pp. 571-607
8. C. Geourjon and G. Deléage1,SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer Applications in the Biosciences*. **11(6)** (1995) 681-684.
9. Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ. Basic local alignment search tool."*Journal of Molecular. Biology*. **215** (1990) 403-410.
10. Capriotti, E, Fariselli P and Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Research.* **33** (suppl 2) (2005) W306-W310.
12. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR.,A method and server for predicting damaging missense mutations. *Nature Methods*. **7(4)** (2010) 248 -249