IOSR Journal of Pharmacy and Biological Sciences (IOSR-JPBS)
e-ISSN: 2278-3008, p-ISSN:2319-7676. Volume 10, Issue 4 Ver. II (Jul - Aug. 2015), PP 41-51
www.iosrjournals.org

Phylogenetic and in-silico structural and functional analysis of RIP's as immunolesioning agents

Sowmya Dhawan[1] and Usha chouhan[2]
[1]*Department of Bioinformatics, Maulana Azad National Institute of Technology, Bhopal (M.P) India*
[2]*Department o Mathematics, Maulana Azad National Institute of Technology, Bhopal (M.P), India*

Abstract: *Creating models that mimics the major neurodegenerative diseases like Parkinson's is of great concern in the developing world. In order to study the lesions created in the brain by the Parkinson's disease and for the screening of many drug molecules, there is a need of effective models and effective agents to create such models. In the present study, the Ribosome inactivating protein of Saponaria officinalis, Saporin 6 is used which shows immunolesioning property. In this paper an in-silico approach is described for the search of the closely related RIP for saporin 6 and to study their similarities and differences. Also a method is developed for the simultaneous superimposition of protein molecules and for the calculation of the RMSD values in order to obtain most similar structure of the phytoprotiens showing immunolesioning property. Consequently, we obtained Dianthin 30 a phytoprotein, which shows very similar structural and physico-chemical properties with Saporin 6, although showing differential catalytic activity. By functional characterization, we also have investigated the contribution of amino acid changes in the active sites of the proteins leading towards their differential catalytic activity.*

Key words: *Parkinson's disease, Immunolesioning agents, Ribosome-inactivating protein (RIP), Saporin, Dianthin, Catalytic activity.*

I. Introduction

Parkinson's disease is the second most prevalent neurodegenerative disorder after the Alzheimer's disease [1]. It is a chronic, progressive neurological disorder in which is characterized by the damage of an area of the brain called substantial nigra. This area influences all involuntary movements. The disorder is idiopathic consequently no cure as such exists for it, but the symptoms can be controlled by using a combination of drugs, therapies and surgery as the last resort [2]. It affects more than 1% of the population near the age of 65 years and more than 4% of populations over the age of 85 are affected [3, 4].

The disease remains unrecognized at an early stage, since at that point symptoms are often non-specific and can include weakness, tiredness and fatigue. The primary symptoms include: muscular rigidity, resting tremor, difficulty with movement initiation (Bradykinesia), slowed voluntary movement, difficulty with balance and difficulty with walking. A person may also start the secondary symptoms which include: depression, senility, postural deformity and difficulty in speaking. Regardless of these primary and secondary symptoms, an impairment of hand functions during daily activities is also found [5]. Many movement disorders such as P.D have been attributed to disturbances of basal ganglia, but the precise neuroanatomy and function of these neuronal is not yet fully understood [6]. Thus in order to study the disease, creating animal models that mimicking Parkinson's disease is relevant as well as for the drug screening. Thus compounds called immunolesioning agents that induce similar conditions came in to existence as a tool in molecular neurosurgery.

This strategy involves producing highly selective neural lesions by targeting an immunotoxin into specific neurons based on their binding to surface membrane targets, such as neurotransmitter or growth factor receptors and so on. The goal is to produce lesions of unprecedented selectivity to match the staggering complexity of the organization of the nervous system [7].These lesions have been used extensively to investigate the role of these dopamine neurons with respect to the behavior, to examine the brains capacity to recover from a compensate for specific neurochemical depletions and to investigate the promotive effects of experimental and clinical approaches which are relevant for the treatment of Parkinson's disease [8].

Ribosomal inactivating proteins (RIP's) have been used to prepare immunotoxins or other conjugates, either by chemical linkage or as recombinant fusion proteins mostly with monoclonal antibodies but also with other suitable carriers, e.g. hormones (hormonotoxins), cytokines, neuropeptides [9]. Ribosome inactivating proteins (RIP's) are the most widely studied phytoprotein which can inhibit protein synthesis or translation in a cell by depurinating (N-glycoside activity) the 28sr RNA are known to serve as apoptosis inducers [10]. They act by catalytically depurinating an adenine residue present in the universal stem loop region of 23/26/28s ribosomal RNA (also known as alpha- sarcin loop), causing an irreversible arrest in protein synthesis leading to cell death of mammalian target cells [11,12]. The plant based neurotoxins are the most preferred suicide transport agents. Since animal neurotoxins act rapidly by blocking synaptic neurotransmitter exocytosis,

resulting in selective interferences with ionic channels and receptors located on the neuronal cell surface. In contrast plant neurotoxins act inside the cells [6]. Other than the immunolesioning property which is comparatively less reported, the RIP's show several pharmacological properties mainly include anti-HIV, antitumor, antibacterial, antifungal, antiparasitic, insecticidal and abortifacient properties [13]. Ribosome inactivating protein induces apoptosis by decreasing the action of anti apoptotic factors .They have proved to be very effective drug against AIDS by acting directly on HIV Infected cells by depurinating the RNA [14]. The RIPs are the better cure for certain allergies but are also having allergenic properties as they are raw eaten in the form of vegetables [15]. Since the drugs of natural origin are more efficient and cost effective over the synthetic drugs, therefore new drug alternatives from plant should be identified and designed in order to obtain drugs with negligible side effects. RIPs also served as molecular tools to selectively kill specific cell types in order to study their physiological or behavioral relevance.

In this paper RIP- Phytoprotein mainly those inhibiting immunolesioning properties is taken for analysis. Saporin a ribosome – inactivating protein is an agent of choice for making anti neuronal immunotoxins and neuropeptide-toxin conjugate [6]. RIPs of Saponaria officinalis is named Saporin, and the most abundant form is Saporin-6. Saporin –S6 belongs to a multigene family of proteins that includes more than nine different isoforms isolated from various plant tissues, such as leaf, root and seed. All isoforms differ from each other in both their physico-chemical and biological properties. Saporin–6 is the most representative of the seed isoforms, accounting for approximately 7% of the total seed protein content. Notably, this one is the most studied among the type 1 RIPs because of its strong activity both in cell- free systems and in cell lines [16, 17]. Type 1 is single chain proteins with molecular weight of approximately 30 kDa. The Type 2 RIPs are potent toxins, with an active A chain linked with a B chain with lectin properties where as Type 1 RIPs being devoid of the B chain, cannot bind to cells in which they enter with difficulty and consequently their toxicity is much lower than that of ricin and related toxins. But these toxins may escape the endosome or may traffic other intracellular compartments and since many of the Type 2 toxins are of less specificity and gets spreaded to other tissues. Thus Type 1 is more used [18].

In general anti –neuronal immunotoxins consist of monoclonal antibodies directed against a specific neuronal antigen and coupled to a RIP such as saporin. Saporin since the lack of the B chain, it cannot penetrate in to a neuron unless coupled to a delivery vector such as a monoclonal antibody that will mediate the endocytosis of the intact conjugate. Because of its wide neuronal targeting, it is been used as a general suicide transport agent for the study of neurodegenerative diseases [6]. Thus saporin is a valuable neurological tool for studying structure and function of adrenergic transmission. They can target midbrain dopaminergic neurons and may be useful in producing a lesion very similar to the naturally occurring neural degeneration seen in Parkinson's disease [19]. This immunotoxin provides researchers with a powerful lesioning tool, more specific and effective than chemical, surgical or electrolytic lesioning. This intern lead to the discovery of many animal models for drug screening.

Neurological research has been relied upon analysis of the effects of the lesions on the nervous systems structure and function. Molecular neurosurgery using chimeric toxins which seek out and destroy specific neurons that express a selected cell surface target molecule is a recent advance in lesioning approach. Although biochemical properties of RIPs have been extensively studied, the enzymatic mechanism of RIPs is still elusive.

The phylogenetic analysis was done and several homologous protein sequences were identified. Since the sequence similarity shows resemblance to functional similarity, and amino acid sequences determines the protein three dimensional structures and structural similarity between the proteins is very good predictor of functional similarity. Therefore on Structural analysis the protein structure identified showed immunolesioning property similar to Saporin6.This protein structure is superimposed and the sequence alignments, structure alignments and Protein Data Bank coordinates and RMSD statistics were generated [20]. As a result of our analysis Dianthus Caryophyllus RIP was identified which may also show immunolesioning property. The antiviral property of this protein has been studied and reported earlier [21]. On studying molecular insights of the three dimensional structures of proteins can help to study molecular mechanisms – such as site directed mutagenesis ,mapping of disease – related mutations, and structure – based design of inhibitors and in other molecular surgery tools.

The physicochemical and the structural properties of the proteins are well understood with the use of computational tools. The statistics about a protein sequence such as number of amino acids, sequence length, and the physicochemical properties of a protein such as molecular weight, atomic composition, extinction coefficient, GRAVY, aliphatic index, instability index, etc. can be computed by computational tools for the prediction and characterization of protein structure. The amino acid sequence provides most of the information required for determining and characterizing the molecule's function and physical and chemical properties.

It was reported earlier that Dianthin 30 is less potent, than saporin isoforms, which may affect their cytotoxic activity [22].The present study has also investigated the amino acid changes in the active site leading towards the differential catalytic activity of the proteins.

II. Materials And Methods

The present analysis involves ribosome inactivating protein namely Saporin from the plant Saponaria officinalis, commonly called as Soap Wort. There are total 9 RIPs isolated from soap wort as tabulated below.

Table 1: Saporin Isoforms

Phyto- Protien/Gene Name	Protein Name	No. of amino acid residues	Sequence Status	PDB ID/ Uniport Acession no
Saporin 1	r Rna N-glycosidase	292	Complete	A6H5D1
SAP 1	RIP Saporin 1	40	Fragment	P98185
SAP2	RIP Saporin2	292	Complete	P27559
SAP3	RIP Saporin3	236	Fragment	P27560
SAP4	RIP Saporin4	157	Fragment	P27561
SAP5	RIP Saporin5	253	Complete	Q41389
SAP6	RIP Saporin6	299	Complete	1QI7/P20656
SAP7	RIP Saporin7	253	Complete	Q41391
SAP9	RIP Saporin9	253	Complete	Q7M1Z2

Only full length mature sequences are used in bioinformatics analysis. However the best characterized and most widely utilized type 1 RIPs are the Saporin 6 .It is because SAP 6 is the most abundant form and are extremely stable. Thus because of the low toxicity , easy availability, safety in handling and their extreme stability they are used as therapeutic or research tool when conjugated to other biological molecules that target specific cell types [17].

The protein sequences were retrieved from Uniport database and the 3-D structures were taken from RCSB Protein Data Bank. Further the Insilico analysis was done using the following bioinformatics software tools: BLASTp, Clustal-W, Geneious, Swiss model, Superpose etc. Only full length sequences were considered for Insilico studies. The methodology involves the following steps.

Step 2.1: Sequence search across the Uniport Knowledge Base Database: The full length sequences of ribosome inactivating protein Saporin 6 of Saponaria officinalis were retrieved in the FASTA Format from Uniport knowledge base. (uniportKB)

Step 2.2: The similarity search done across Ribosome Inactivating Protein (RIP) sequences of Saponaria officianalis: The similarity search was done with the help of protein BLAST (Basic Local Alignment Search Tool) Software available at NCBI (National Centre of Biotechnological Information) Blastp which involves the pair wise alignment algorithm [23]. The sequences showing more than 75% similarity were taken for further analysis...

Step 2.3: Performing multiple sequence alignment for the similar RIP Sequences: Clustal-Omega performs the global multiple sequence alignment and the similar sequences were aligned [24].

Step 2.4: Generating phylogenitic tree by using UPGMA algorithm: In this step after aligning the entire protein sequences, the phylogentic tree was generated for the Ribosome Inactivating Protein (RIP's) Saporin 6.The distance matrix was calculated and the orthologs and paralogs were identified respectively.

Step 2.5: Retrieval of protein 3-D Structures from Protein Data Bank (PDB): The three dimensional protein structures were retrieved from the PDB.

Step 2.6: Secondary Structure Prediction: GOR IV was used to predict the secondary structure.

Step 2.7: Structure alignment of the closely related sequence: To have structural similarity it has been found for protein structure after superimposition should have the RMSD values less than or equal to 2Å.Using SUPERPOSE web server the proteins 3-D structures were superposed against Saporin and the proteins with RMSD 2 Å or less than that were predicted as the proteins homologous to the Saporin 6 Showing the immunolesioning property.

Step 2.8: Analysis of Molecular structures and physicochemical properties: Structures and properties of nucleotide and corresponding amino acid sequences of antiviral and RIP gene from Dianthus Caryophyllus and Saponaria officinalis were obtained respectively by using ProtParam tool, which is available through EXPASY server[25].

Step 2.9: Identification of domains: Both the protein sequence was analyzed using SMART (Simple Modular Architecture Research Tool) and Pfam. SMART is an online resource for identification and annotation of protein domains and analysis of protein domain architecture [26]. Pfam is widely used database of protein families and domains. Pfam is used to organize sequences, to find the origin and evolution of proteins and for identifying interesting new targets [27].

Step 2.10: Prediction of function: The biological function was studied and compared using ESG, a sequence similarity based protein function prediction server [28].

Step 2.11: Antigenicity Prediction: The server VaxiJen was used to analyze the antigenicity of the query protein. VaxiJen was developed to allow antigen classification based on physicochemical properties of proteins. Protein sequences can be submitted as single protein or as a multiple sequence file in fasta format [29].

Step 2.12: Binding Site Prediction: RaptorX- Binding, a web server that predicts the binding sites of a protein sequence, based up on the 3D model, is used for the prediction of the binding sites [30].

III. Results And Discussion

As a result of similarity 11 RIP sequences were found to similar with the RIP Saporin6**.**But there was only 1 sequence of different genus which showed more than 75 % similarity with SAP 6 other than the isoforms of Saponaria officinalis. The 11 RIPs obtained as a result of similarity search are tabulated under table 3.

3.1 Sequence search across the Uniport Knowledge Base Database

The sequence selected and details are given in the table 2.

Table 2: Ribosome Inactivating Protein taken for the study

Phyto-Protein/ Gene Name	Protein Name	No. of amino acid residues	Sequence Status	PDB ID/ Uniport Accession Number
SAP6	RIP Saporin6	299	Complete	1QI7/P20656

STEP 3.2 Similarities search was performed using protein Blast

Table 3: Blast results

ENTRY	ENTRY NAME	PROTEIN NAME	ORGANIM	LENGTH	IDENTITY	GENE NAME	E- VALUE
P20656	RIP6_SAP0F	Ribosome inactivating protein saporin -6	Saponaria officinalis	299	100%	SAP6	0.00
Q2QEH3	Q2QEH3_SAPOF	rRNA N-glycosidase	Saponaria officinalis	292	99%		0.00
A6H5D1	A6H5D1_SAOPOF	rRNA N-glycosidase	Saponaria officinalis	292	96%	SAPORIN-1	0.00
Q41391	RIP7_SAPOF	Ribosome – Inactivating Protein Saporin -7	Saponaria officinalis	253	98%	SAP7	2.0×10^{-172}
P27561	RIP4_SAPOF	Ribosome – Inactivating Protein Saporin -4	Saponaria officinalis	157	98%	SAP4	1.0×10^{-100}
Q05148	Q05148_SAPOF	rRNA N-glycosidase	Saponaria officinalis	103	97%		1.0×10^{-60}
Q41389	RIP5_SAPOF	Ribosome – Inactivating Protein Saporin -5	Saponaria officinalis	253	96%	SAP5	2.0×10^{-168}
P27559	RIP2_SAPOF	Ribosome – Inactivating Protein Saporin -2	Saponaria officinalis	292	95%	SAP2	0.00
Q71MIZ2	RIP9_SAPOF	Ribosome – Inactivating Protein Saporin -9	Saponaria officinalis	253	92%	SAP9	5.0×10^{-162}
P27560	RIP3_SAPOF	Ribosome – Inactivating Protein Saporin -3	Saponaria officinalis	236	89%	SAP3	4.0×10^{-136}
P24476	RIP0_DIACA	Antiviral protein DAP -30	Dianthus caryophyllus	293	79%	DAP 30	8.0×10^{-154}

STEP 3.3

Multiple sequence alignment (MSA)

Generally the alignment of three or more biological sequence of similar length has been performed. From that output homology and the evolutionary relationship can be inferred between the sequences studied. Among them sap 2, Sap9, saporin 1, Dap 30 showed 78% sequence similarity. Then with sap 6, sap 5 and sap 7 it showed 79% similarity.

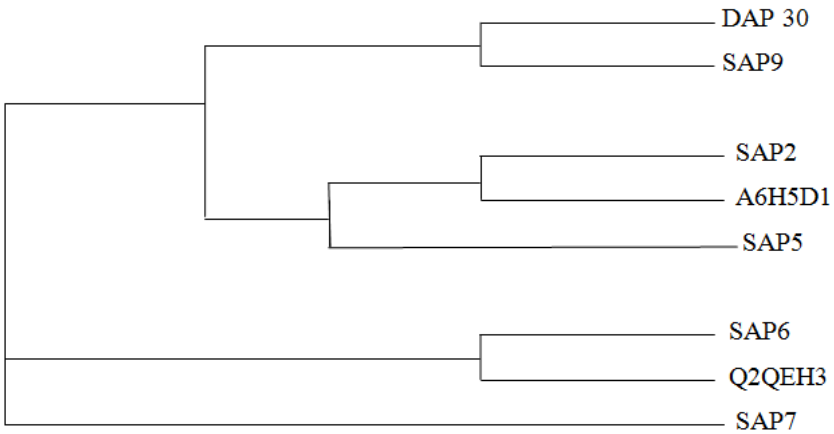
MSA Using Clustal Omega for SAP-6
CLUSTAL O(1.2.0) multiple sequence alignment

```
sp|P24476|RIP0_DIACA                                                    MKIYLVAAIAWILFQSSSWTT-
DAATAYTLNLANPSASQYSSFLDQIRNNVRDTSLIYGG
sp|Q7M1Z2|RIP9_SAPOF       ----------VTSITLDLVNPTAGQYSSFVDKIRNNVKDPNLKYGG
sp|P27559|RIP2_SAPOF
MKIYVVATIAWILLQFSAWTTTDAVTSITLDLVNPTAGQYSSFVDKIRNNVKDPNLKYGG
tr|A6H5D1|A6H5D1_SAPOF
MKIYVVATIAWILLQFSAWTTTDAVTSITLDLVNPTAGQYSSFVDKIRNNVKDPNLKYGG
sp|Q41389|RIP5_SAPOF       ----------VTSITLDLVNPTAGQYSSFVDKIRNNVKDPNLKYGG
sp|P20656|RIP6_SAPOF
MKIYVVATIAWILLQFSAWTTTDAVTSITLDLVNPTAGQYSSFVDKIRNNVKDPNLKYGG
tr|Q2QEH3|Q2QEH3_SAPOF
MKIYVVATIAWILLQFSAWTTTDAVTSITLDLVNPTAGQYSSFVDKIRNNVKDPNLKYGG
sp|Q41391|RIP7_SAPOF       ----------VTSITLDLVNPTAGQYSSFVDKIRNNVKDPNLKYGG
                             .*: **:*.**:*.*****:*:*****:* .* ***
sp|P24476|RIP0_DIACA
TDVAVIGAPSTTDKFLRLNFQGPRGTVSLGLRRENLYVVAYLAMDNANVNRAYYFKNQIT
sp|Q7M1Z2|RIP9_SAPOF                                                    TDIAVIGPPS-
KDKFLRINFQSSRGTVSLGLKRDNLYVVAYLAMDNTNVNRAYYFRSEIT
sp|P27559|RIP2_SAPOF                                                    TDIAVIGPPS-
KDKFLRINFQSSRGTVSLGLKRDNLYVVAYLAMDNTNVNRAYYFKSEIT
tr|A6H5D1|A6H5D1_SAPOF                                                  TDIAVIGPPS-
KEKFLRINFQSSRGTVSLGLKRDNLYVVAYLAMDNTNVNRAYYFRSEIT
sp|Q41389|RIP5_SAPOF                                                    TDIAVIGPPS-
KEKFLRINFQSSRGTVSLGLKRDNLYVVAYLAMDNTNVNRAYYFRSEIT
sp|P20656|RIP6_SAPOF                                                    TDIAVIGPPS-
KEKFLRINFQSSRGTVSLGLKRDNLYVVAYLAMDNTNVNRAYYFRSEIT
tr|Q2QEH3|Q2QEH3_SAPOF                                                  TDIAVIGPPS-
KEKFLRINFQSSRGTVSLGLKRDNLYVVAYLAMDNTNVNRAYYFRSEIT
sp|Q41391|RIP7_SAPOF                                                    TDIAVIGPPS-
KEKFLRINFQSSRGTVSLGLKRDNLYVVAYLAMDNTNVNRAYYFRSEIT
                  **:**** ** .:****:***. *********:*:**********:*********:.:**
sp|P24476|RIP0_DIACA
SAELTALFPEVVVANQKQLEYGEDYQAIEKNAKITTGDQSRKELGLGINLLITMIDGVNK
sp|Q7M1Z2|RIP9_SAPOF
SAELTALFPEATAANHKALEYTEDYHSIEKNAQITEGDKSRKELGLGINLLSSTMDTVNK
sp|P27559|RIP2_SAPOF
SAELTALFPEATTANQKALEYTEDYQSIEKNAQITQGDKSRKELGLGIDLLLTFMEAVNK
tr|A6H5D1|A6H5D1_SAPOF
SAELTALFPEATTANQKALEYTEDYQSIEKNAQITQGDKSRKELGLGIDLLLTSMEAVNK
sp|Q41389|RIP5_SAPOF
SAELTALFPEATTANQKALEYTEDYQSIEKNAQITQGDKSRKELGLGIDLLLTSMEAVNK
sp|P20656|RIP6_SAPOF
SAESTALFPEATTANQKALEYTEDYQSIEKNAQITQGDQSRKELGLGIDLLSTSMEAVNK
tr|Q2QEH3|Q2QEH3_SAPOF
SAELTALFPEATTANQKALEYTEDYQSIEKNAQITQGDQSRKELGLGIDLLSTSMEAVNK
sp|Q41391|RIP7_SAPOF
SAELTALFPEATTANQKALEYTEDYQSIEKNAQITQGDKSRKELGLGIDLLLTSMEAVNK
                  *** *******...**:* *** ***.:******:** **:**********:** : :: ***
```

```
sp|P24476|RIP0_DIACA
KVRVVKDEARFLLIAIQMTAEAARFRYIQNLVTKNFPNKFDSENKVIQFQVSWSKISTAI
sp|Q7M1Z2|RIP9_SAPOF
KVRVVKNEARFLLIAIQMTAEAVRFRYIQNLVTKNFPNKFNSENKVIKFEVNWKKISTAI
sp|P27559|RIP2_SAPOF
KARVVKNEARFLLIAIQMTAEVARFRYIQNLVTKNFPNKFDSDNKVIQFEVSWRKISTAI
tr|A6H5D1|A6H5D1_SAPOF
KARVVKNEARFLLIAIQMTAEVARFRYIQNLVTKNFPNKFDSDNKVIQFEVSWRKISTAI
sp|Q41389|RIP5_SAPOF
KARVVKNEARFLLIAIQMTAEVARFRYIQNLVTKNFPNKFDSDNKVIQFEVSWRKISTAI
sp|P20656|RIP6_SAPOF
KARVVKDEARFLLIAIQMTAEAARFRYIQNLVIKNFPNKFNSENKVIQFEVNWKKISTAI
tr|Q2QEH3|Q2QEH3_SAPOF
KARVVKDEARFLLIAIQMTAEAARFRYIQNLVIKNFPNKFNSENKVIQFEVNWKKISTAI
sp|Q41391|RIP7_SAPOF
KARVVKNEARFLLIAIQMTAEAARFRYIQNLVIKNFPNKFNSENKVIQFEVNWKKISTAI
                 *.****:***************..********* *******:*:****:*:*:* ******
sp|P24476|RIP0_DIACA
FGDCKNGVFNKDYDFGFGKVRQAKDLQMGLLKYLGRPKSSSIEANSTDDTADVL------
sp|Q7M1Z2|RIP9_SAPOF      HGDAKNGVFNKDYDFGFGKVRLVKDLQMGLLMHLGKPK----------
-----sp|P27559|RIP2_SAPOF
YGDAKNGVFNKDYDFGFGKVRQVKDLQMGLLMYLGKPKSSNEANSTAYATTVL-------
tr|A6H5D1|A6H5D1_SAPOF
YGDAKNGVFNKDYDFGFGKVRQVKDLQMGLLMYLGKPKSSNEANSTAYATTVL-------
sp|Q41389|RIP5_SAPOF      YGDAKNGVFNKDYDFGFGKVRQVKDLQMGLLMYLGKPK----------
---
sp|P20656|RIP6_SAPOF
YGDAKNGVFNKDYDFGFGKVRQVKDLQMGLLMYLGKPKSSNEANSTVRHYGPLKPTLLIT
tr|Q2QEH3|Q2QEH3_SAPOF
YGDAKNGVFNKDYDFGFGKVRQVKDLQMGLLMYLGKPKSSNEANSTAYATTVL-------
sp|Q41391|RIP7_SAPOF      YGDAKNGVFNKDYDFGFGKVRQVKDLQMGLLMYLGKPK----------
---
                 .**.****************** .******** :**:**
```

STEP 3.4

PHYLOGRAM



While analyzing the phylogram, it has been found that the Dap 30 is orthologs to sap6 since it is distantly placed in the phylogentic tree. The orthologs are genes in different species that evolved from a common ancestral gene by specification. orthologos retain the same function in the course of evolution. Dap 30 or Dianthin antiviral protein from the leaves of carnation Dianthus caryophyllus.

STEP 3.5

The three dimensional protein structures of mainly Saporin 6 and Dianthin 30 were retrieved from the PDB and details are given in table 4.

Table 4

PHYTOPROTIEN	NO.OF AMINOACID RESIDUE	PDB ID
Saporin 6 (SAP 6)	253	1QI7
Dianthin 30(DAP 30)	255	1LP8

Step 3.6: Secondary Structure Prediction: The secondary structure prediction defined each residue into alpha helix, extended strand or random coil secondary structures. GOR IV analysis revealed that alpha helix is more than extended strand and random coil in both the proteins as shown in the table 5.

Table 5: GOR IV analysis results

SECONDARY STRUCTURE	PHYTOPROTEIN			
	SAP -6(1Q17)		DAP- 30(1RLO)	
	No of residues	Percentage	No of residues	Percentage
Alpha helix	111	43.87%	110	43.14%
Extended Strand	47	18.58%	41	16.08%
Random coil	95	37.55%	104	40.78%

Step 3.7: Structure alignment of the closely related sequence: Comparison between structures is done by Superpose web server, which used the sequence alignment to guide the superimposition. The back bone superimposing between Saporin6 and Dianthin 30 which gave a root mean square deviation of 0.68 Å. After superimposing the structures we got the protein having the structure homologous to the ribosome inactivating protein of Saponaria officinalis. It has been predicted that only Dianthin 30 is orthologous for Saporin 6 protein, i.e. Both of them retain same function in the course of evolution and are more closely related protein structures and are homologous in their structure and function.

As a result of superimposition, shown in figure 1 it has been found that the protein structures with RMSD less than 2 Å are found to be structurally similar to the RIP Saporin 6.Since Dianthin 30 is similar to Saporin 6 with RMSD 0.68 Å , therefore can show the immunolesioning property like Saporin 6.

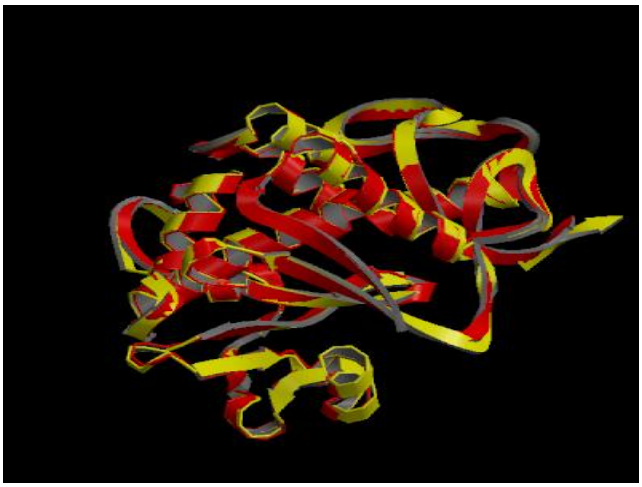


Figure: 1: Superimposition of Saporin 6 and Dianthin 30

Step 3.8 Analysis of molecular structures and physicochemical properties: The physicochemical properties shown by the ProtParam is indiced in the table 6 may be pertinent to determine that they are a group of gene with significant functional and close genetic relation. The mature form of Saporin –S6 IS 253 amino acids long and Dianthin 30 is 255 amino acids long. The average molecular weights of RIPs of TYPE 1 are calculated to be below 30 k Da. The isoelectric point is the PH at which the surface of the protein is covered with a charge but the net charge of the protein is zero.

The computed theoretical isoeletric point values were 9.45 and 9.48 shows that this protein belonged to the same basic protein. This high P^I value is because of the presence of high lysine residues in the total amino acid content. The $P^I > 7$ indicates that these are basic in character. The computed isoelectric point P^I will be useful for developing buffer systems for purification by isoelectric focusing method. The number of positively charged residues is higher than the negatively charged.

Although the Expasy's PortParam computes the extension coefficient for a range of (276,278,279,280 and 282 nm) wavelength, 280 nm is favored because proteins absorb strongly there while other substances commonly in the protein solutions do not. The computed protein concentration and extension coefficient help in the protein-protein and protein –ligand interactions in the solution.

On the basis of instability index the Expasy's PortParam classifies both the protein as stable (Instability Index< 40).The aliphatic index(AI) which is defined as the relative volume of the protein occupied by the aliphatic side chains (A ,V,I and L) is regarded as a positive factor for the increase of thermal stability of globular proteins. A similar and relatively higher aliphatic index of both the proteins infers that they may be stable for a wide range of temperature, to denaturation by urea or guanidine and to the attack by proteolytic enzymes. This is also very stable in response to chemical modifications such as those necessary for derivatization and conjugation procedures and it is resistant to many freeze thaw cycles. Altogether these characteristics render these proteins as interesting candidates for the construction of immunoconjugates .The low GRAVY Index of both the proteins infers that they could result in better interaction with water.

The only difference between the proteins is in their estimated half life , where Sap 6 showed much higher half life of 100 hours than Dianthin 30 with half life of only 4.4 hours in mammalian reticulate , but similar in yeast and E.coli. Since the N- terminal of DAP 30 is alanine where as SAP 6 has valine in the N-terminal.

Table 6: Parameters computed using Expasy's ProtParam tool.

Phytoproteins		**SAP 6**	**DAP 30**
Accession Number		1QI7	1RL0
Sequence length		253	255
Molecular Weight		28588.7	28599.6
P^I- Isoelectirc Point		9.45	9.48
-R- Number of negative residues		27	24
+R- Number of positive residues		36	33
EC- Extinction coefficient at 280nm		23380	23380
Estimated Half Life (in hours)	**Mammalian reticulate**	100	4.4
	Yeast	20	20
	E.coli	10	10
II-Instability Index		24.68 –STABLE	24.51-STABLE
AI-Aliphatic Index		87.87	89.88
GRAVY-Grand Average hydropathy		-0.338	-0.264

Step 3.9: Identification of Domains: Smart and Pfam identifies the specific region that encodes the domain as shown in the table 7 and 8 respectively for SAP and DAP.

Identification of Domains for SAP: Smart revealed the domain is present between amino acid 29 and 242. Pfam classified the domain in the Pfam A family group with id PF00161.

Identification of Domains for DAP: Smart revealed the domain is present between amino acid 28 and 242. Pfam classified the domain in the Pfam A family group with id PF00161.

Table 7: Region identified by SMART that encodes the domain

Phytoprotein	**Name**	**Begin**	**End**	**E-value**
SAP 6	Pfam RIP	28	242	6.9e- 52
DAP 30	Pfam RIP	28	242	8.8e- 50

Table 8: Region identified by PFAM that encodes the domain

Phytoprotein	**Source**	**Domain**	**Start**	**End**
SAP6	PfamA	RIP	29	242
DAP30	PfamA	RIP	28	242

Step 3.10: Prediction of Function: ESG reveal the molecular function, biological function and the percentage distribution of the Protein in different cellular components as shown in Table9, 10, 11 and 12 respectively.

Results for protein 1QI7

Table 9: Molecular function terms revealed by ESG

SL.NO	**Probability**	**Term**	**Description**
1	100%	GO:0016787	Hydrolase activity
2	100%	GO:0030598	rRNA N- glycosylase activity
3	5%	GO:0005529	Sugar binding

Table 10: Biological Process terms revealed by ESG

SL.NO	Probability	Term	Description
1	100%	GO:0006952	Defense response
2	100%	GO:0009405	Pathogenesis
3	100%	GO:0017148	Negative regulation of translation
4	21%	GO:0009615	Response to virus
5	8%	GO:0050688	Regulation of defense response to virus

It is predicted that 100% molecular function is on both hydrolyses and r RNA N-glycosylase activity. The main biological function of the protein is pathogenesis and negative regulation of translation.

Results for protein 1RLO

Table11: Molecular function terms revealed by ESG

SL.NO	Probability	Term	Description
1	99%	GO:0016787	Hydrolase activity
2	99%	GO:0030598	rRNA N- glycosylase activity
3	8%	GO:0005529	Sugar binding

Table12: Biological Process terms revealed by ESG

SL.NO	Probability	Term	Description
1	99%	GO:0006952	Defense response
2	99%	GO:0009405	Pathogenesis
3	99%	GO:0017148	Negative regulation of translation
4	21%	GO:0009615	Response to virus
5	8%	GO:0050688	Regulation of defense response to virus

It is predicted that 99% molecular function is on both hydrolyses and rRNA N-glycosylase activity. The main biological function of the protein is pathogenesis and negative regulation of translation.

Step 3.11: Antigenicity Prediction: VaxiJen classified both the phytoproteins as protective antigen, because it has a value of 0.5653 and 0.5844 respectively which is above the normal threshold value of 0.5.Thus both showed similar antigenic determinants as per the results obtained.

Step 3.12: Binding Site Prediction: RaptorX predicted pockets, ligands and binding residues, as tabulated below in 13 and 14.

Table 13: Prediction for 1Q17

POCKET	MULTIPLICITY	LIGAND	BINDING RESIDUES
1	80	ADE	L71,Y72,V73,E118,D119,Y120,I171,A175,E176, R179,E205,V206
2	26	SO_4	Y116,T117,D119,S122
3	18	NAG	K226,D227

Table14: Prediction for 1RL0

POCKET	MULTIPLICITY	LIGAND	BINDING RESIDUES
1	75	ADE	L72,Y73,V74,E119,D120,Y121,I172,A176,E177 ,R180,Q206,W209
2	28	SO_4	G118,E119,D120,A123
3	22	NAG	K227,D228

For the binding site prediction, an important measure is the pocket multiplicity, which is used to judge the quality of a predicted pocket. It represents the frequency with which the selected pocket was found in a set of Ligand –binding protein structure. The higher the better, when it is above 40, there is a good chance that the predicted pocket is true.

According to RaptorX results, both the proteins show a very high level of pocket multiplicity in case of Ligand ADE, i.e. 80 and 75 respectively for SAP and DAP which is highly above the threshold of 40.From this it is proved that RIPs can remove some adenine bases from double stranded DNA /RNA molecules there by arresting the protein synthesis.

There by it can also be inferred that the difference in the catalytic activity showed in the earlier experimental assays of the two proteins must be due to the change in the amino acid residues in the binding site.

IV. Conclusion

This study contribute to the understanding of the agents which help in specific elimination of cells that express dopamine, which is very much useful in studying the role of dopaminergic neurons in the devasting disease of old age Parkinson's. There are several approaches introduced for the proteome analysis, here we have used the protein sequences in order to get the alternatives for existing immunolesioning agents. In the present study Saponaria Officinalis, Ribosome Inactivating Proteins sequences especially Sap6 was taken. Earlier studies have reported its immunolesioning property clearly. Here we have come up with a most closely related structural homologue for this protein .As a result of our analysis this phytoprotein can show the immunolesioning property similar in case RIPs of Saponaria officinalis.The structural homologue obtained, Dianthin-30 also shows similar physicochemical properties, Biological functions, and antigenicity like Sap 6 .This gives add on to the structural analysis, that it can also be used as neural lesioning agent or neurotoxins to create lesions very similar to the naturally occurring neural degeneration seen in Parkinson's disease when conjugated with a Dopamine transporter, other than its antiviral property which has been reported earlier. In the present study we have also investigated the amino acid changes in the active site leading towards the differential catalytic activity of the proteins. It was reported earlier that Dianthin 30 is less potent, than saporin isoforms, which may affect their cytotoxic activity.

Because of the wide availability, the biochemical features of only one group of closely related saporin isoforms, collectively named as sap6 the basic protein of our study, is most studied and characterized. Similarly since dianthin also is widely available phytoprotein, further in vivo research is needed for effective recognition of properties. Dianthin 30 was found in larger amounts than Dianthin 32 which is found only in the leaves, where as Dianthin 30 is found throughout the plant.

The studies will help in the immunological characterization of these proteins to define their potential for immunological diagnosis. Because of low toxicity or safety in handling and their extreme stability, they are used as a therapeutic or research tool when conjugated to other biological molecules that target specific cell types. This can be used as molecular tools to selectively kill specific cell types in order to study their physiological or behavioral relevance. Thus by destruction of midbrain dopaminergic neurons using these immunotoxins, will lead to the discovery of animal model mimicking Parkinson's disease for the drug screening.

In conclusion, the present study will encourage the extension of molecular level investigation on these RIPs to elucidate important differences and simulation that can be linked to their biological activities in various substrates and its use in molecular surgery in devasting conditions like neurodegeneration, thus ultimately this is of great importance for their pharmacological and medical applications. There are many potential reagents waiting to be used to make targeted toxins, and with time, many new ones will emerge with promising properties. The future is bright and filled with opportunity for expanded use of targeted toxins in molecular neurosurgery, to characterize the complexities of nervous system organization and potentially treats neurological disorders.

Acknowledgement

The authors are grateful to Department of Biotechnology, New Delhi, India for providing Bioinformatics Infrastructure facility to carry out this work.

References

[1]. Claire Gibrat., Martine Saint-Pierre., et al. (2009), Differences between subacute and chronic MPTP mice models: Investigation of dopaminergic neuronal degeneration and α-synuclein inclusions, Journal of Neurochemistry, 109(5), 1469-1482.
[2]. En.wikipedia.org/wiki/parkinsons disease
[3]. Uday Muthane. (2008), Movement disorders in India, Advances in clinical neuroscience and rehabilitation, 8(2), 22-23.
[4]. Kathrin Brockmann., Claudia Schult., et al.(2013), SNCA: Major genetic modifier of age at onset of Parkinson's Disease, Movement Disorders,28(9),1217-1221.
[5]. Albert,F., Diemayr,G., et al. (2010),Coordination of grasping and walking in Parkinon's Disease, Experimental Brain research,202(3), 709-721.
[6]. Hava Lorberboum-Galski.,& Philip Lazarovivi.(2003),Chimeric toxins: mechanisms of action and therapeutic applications
[7]. Ronald, G., Wiley,& Douglas A.Lappi.(2005), Molecular neurosurgery with targeted toxins,Humana Press Inc ,1-307
[8]. Schwarting,RK., & Huston, JP.(1996), The unilateral 6 –hydroxydopamine lesion model in behavioral brain research. Analysis of functional deficits, recovery and treatments, Progress in Neurobiology, 50 (2-3), 275-331.
[9]. Frankel, A.E., Kretman, R.J., et al.(2000) Targeted toxins, Clinical cancer Research, 6, 326-334.
[10]. Mrinal kumar Das., Radhey Shyam Sharma., and Vandana Mishra. (2012),Induction of apoptosis by ribosome inactivating proteins, Importance of N-glycosidase activity,Applied Biochem Biotechnology, 166, 1552- 1561.
[11]. Fiorenzo Stripe.,(2004),Ribosome –inactivating proteins, Toxicon,44(4),371-383.
[12]. Micheal lord, J., & Martin, R., .hartley. (2010), Toxic plants proteins,Springer.

[13]. Ranjitha kumar,B.D., & Pradeep Kumar,S. (2011),Ribosome inactivating proteins : An overview,Journal of Pharmacy Research, 4(3),769-772.
[14]. Edward Gnana Jothi, G., & Sahaya Jose Majilla,G., et al. (2012) Analysis of Ribosome Inactivating protein(RIP): A bioinformatics approach, International Journal of Advancements in Research and Technology,1(5).
[15]. Barbieri, L., Polito, L., et al.,(2006),Ribosome –inactivating proteins in edible plants and purification and characterization of a new ribosome –inactivating protein from Cucurbita moschata,Biochimica et Biophysica Acta, 17650(5),783-792.
[16]. Letizia Polito., Massimo Bortolotti., et al. (2013), Saporin-S6: A useful tool in cancer therapy, Toxins,5(10),1698-1722.
[17]. Mathew, J., walsh, Jennifer, E., Dodd, et al.(2013),Ribosome inactivating proteins, potent poisons and molecular tools, Virulence ,4(8), 774-784.
[18]. Kiresten Nielsen., & Rebecca, S., Boston. (2001), Ribosome –Inactivating Proteins: A plant perspective, Annual review plant physiology and plant molecular biology, 58, 785-816.
[19]. Wiley, RG.,Harrison, MB., et al..(2003), Destruction of midbrain dopaminergic neurons by using immunotoxin to dopamine transporter, Cellular and Molecular Neurobiology, 23(4-5), 839-850.
[20]. Maiti,R., Van Domselaar, GH ., et al. (2004), Superpose web server, Nucleic acid Research,32, 590-594.
[21]. Fiorenzo Stripe., David, G., Willams, et al.(1981), Dianthins,ribosome –damaging proteins and antiviral properties from Dianthus Caryophyllus L.(carnation),Biochem Journal ,195,399-405.
[22]. Serena Fabbrini., M., Emilia Rappocciolo, et al.(1997), Characterization of a Saporin isoform with lower ribosome inhibiting activity, Biochem J, 322, 719-727.
[23]. Altschul,S.f.,Gish,W., et al. (1990),Basic local alignment search tool ,Journal of molecular biology,215,403-410.
[24]. Sievers,F., Wilm,A., et al. (2011),Fast ,scalable generation of high –quality protein multiple sequence alignments using Clustal Omega, Molecular systems biology,7(1),539.
[25]. Gasteiger,E., Hoogland,C., et al.(2005),Protein identification and analysis tools on the EXPASy Server, The proteomics Protocols Handbook, 571-607.
[26]. Jorg Schultz., Frank Milpetz, et al. (1998), SMART, A simple modular architecture research tool: Identification of signaling domains, Proceedings of National Academy of Sciences, 95(11), 5857-5864.
[27]. Marco Punta., Penny,C.,Coggil, et al.(2012),The Pfam protein families database, Nucleic Acid Research,40(1),290-301.
[28]. Chitale,M., Hawkins,T., et al.(2009),ESG: Extended similarity Group Method for automated protein function prediction,Bioinformatics,25(14),1739-1745.
[29]. Irini, A., Doytchinova, Darren, R., Flower. (2007), VaxiJen :A server for Prediction of protective antigens, tumor antigens, and subunit vaccines,BMC Bioinformatics, 8(4).
[30]. Jian Peng.,& Jinbo xu., (2012),RaptorX: exploiting structure information for protein alignment by statistical inference, Proteins,79,161-171.