Indian Journal of Biotechnology
Vol 16, October 2017, pp 655-662

An *in silico* mutagenesis effect on mode of action of curcin, a toxic protein from *Jatropha curcas*

S Vimala Devi[1*], Chetna Tyagi[2], S K Dhyani[1], A K Mishra[2], A Kumar[2], A K Handa[1]

[1]Tree Improvement Lab, Central Agroforestry Research Institute, Jhansi 284 003, India

[2]Agricultural Knowledge Management Unit (AKMU), Indian Agricultural Research Institute, New Delhi 110 012, India

Received 8 May 2015; revised 17 February 2016; accepted 23 February 2016

Curcin is a toxic protein found in the seeds of *Jatropha curcas* and belongs to the family of ribosome inactivating proteins (RIPs), which bind to the rRNA of ribosomes to inactivate them. Curcin is a type 2 RIP protein consisting of a single chain. Its presence makes it difficult to utilize the *Jatropha* seed cakes for animal feed. Inhibition of its biological activity against ribosomes through mutagenesis of protein structure is a novel and valid approach to understand its mode of action. Such mutations were introduced into the protein sequence and the protein was modeled to see their effect on interaction with GAGA loop of mouse 28S rRNA. Interestingly, with the removal or mutation of critical residues of the active site, the binding of curcin and rRNA molecule increased thus indicating towards the steric hindrance provided by these residues. The removal of these residues seems to increase the cavity size, thus, making it easier for the approaching rRNA molecule to bind closely. However, as all RIPs interact with rRNA by N-glycosidation action, the presence of amino acid (aa) residues may be crucial for providing chemical toxicity. This study paves way for a step ahead to see the effect of their removal or mutation in terms of N-glycosidase activity for better understanding.

Keywords: Ribosomal inactivating protein, curcin, docking, mutagenesis

Introduction

Curcin is a ribosome inactivating protein (RIPs), which depurinate rRNA, thus inhibiting protein synthesis[1]. The RIPs are RNA N-glycosidases that inactivate ribosomes by site specifically cleaving the single N–C glycosidic bond between adenine and ribose at A4324 in the 28s rRNA[2]. The depurination of the specific adenine prevents the elongation factor 2 (EF-2) from binding to the 60S subunit, thus RIPs can inhibit protein synthesis. On the basis of structure of the genes and mature proteins[3], RIP superfamily of proteins is classified into three types. Curcin is classified as type 2 RTP, having two chains, Chain A and Chain B, linked by disulfide bridges. The A chain have the ribosome inactivating property; the chain B, contains a lectin domain which interacts with the cell surface galactosides and facilitates the entry of the A chain into the cytoplasm of the cell. Thus, some, but ~15 not all, type 2 RIPs have stronger toxicities than type 1 RIPs because type 1 RIPs can only enter into cells with difficulty, even though they are very active towards isolated ribosome. A comparison of the amino acid (aa) sequences of curcin with other RIP, e.g., ricin A-chain and trichosanthin revealed that there existed relatively high resemblance among them. The percentages of identity between curcin and ricin A ~14 chain, and between curcin and trichosanthin were found to be 54% (156/287) and 57% (138/241), respectively[4].

Curcin is a major constituent of *J. curcas* latex and seeds and is frequently classified as a lectin and comparable to ricin and arbrin proteins from castor bean and *Abrus precatorius*, respectively, with the implication that it has similar toxicity function like inhibiting protein synthesis in whole cells and cell-free systems[5]. *Jatropha curcas* is one of the important source of renewable biodiesel resource because of its low gestation period, ease in production and its oil which has high oleic and linolenic content. To increase its use as economically viable source, value addition of its seed cake as animal feed is good option. However, it is hard to utilize the *Jatropha* seed cakes for animal feed after extracting the oil, because of several anti-nutrients and various toxins, such as curcin, trypsin inhibitors, phorbolesters, esterases, lipase, phytates and saponins[6]. Of these,

*Author for correspondence:
Tel: 0510-2730214, 2730213; Fax: 0510-2730364
vimals123@gmail.com

toxins identified from *Jatropha*, curcin and phorbolester are known to have harmful effects on humans and farm animals. A range of methods like extraction with polar organic solvent and/or heat/$NaHCO_3$ treatment have been used to detoxify the defatted seed meal, however economically feasible method for large-scale detoxification is yet to develop[7,8], hence it is critical to eliminate the risk of toxicity. One effective means of doing so is to shut down and/or modify the biosynthesis of these toxins at molecular level. Though the first report of the cloned sequence of curcin protein is available since 2000, but so far only two reports are available for gene silencing. In the two reports made by Lin *et al*[9], and Patade *et al*[10], only few plants expressed reduced level of gene expression and still the *Jatropha* varieties without curcin in their seeds, has not been a reality till now. Studies regarding the structural and functional properties of curcin and their mechanism of action have been reported but the effect of mutagenesis on its amino acid (aa) residues is not clear. *In silico* analysis provides an insight into the mechanism of action before and after modifications, which will help in designing appropriate inhibitor targets. In order to assist in reducing the toxicity at molecular level and to widen the scope of research, an *in silico* study was performed in search for a curcin structure modification that renders it inactive with the focus on the mode of action and its interacting amino acid residues.

Materials & Methods

Data Retrieval and Protein Modeling

All sequences deposited in the public domain for curcin related proteins were retrieved from National Centre for Biotechnology Information (NCBI)[11] (Table 1) and aligned using MultAlin[12] (Fig. 1). The sequences were having 98% similarity, a full length protein sequence of curcin from *Jatropha curcas* was retrieved from NCBI Genpept database with the accession number ABZ04128.1 in the FASTA format and were used for further study. The sequence was submitted to different modeling servers for modeling of curcin protein using Modweb[13-16] and I-TASSER

Table 1 — Genbank accession ID for curcin protein & curcin precursor from NCBI Genpept database

Gen bank ID	Seq from *Jatropha curcas*	Protein size	Source lab	Year
ABZ04128.1	Full length Curcin	309 aa	Reliance Life Sciences, Navi Mumbai, India	2008
ACO53803.1	Curcin	293 aa	National Center for Genetic Engineering and Biotechnology, Pathumthani Thailand	2008
AEA72440.1	Curcin precursor	309 aa	Defence Institute of Bio-Energy Research, Haldwani, India	2014
AAL58089.1	Curcin precursor	293 aa	Sichuan University, Sichuan, China	2001
AAL86778.1	Curcin precursor	293 aa	Sichuan University, Sichuan, China	2002
ADN39429.1	Curcin1	309 aa	Temasek Life Science Laboratory, National University of Singapore, Singapore	2009
ADN39428.1	Curcin 2A	309 aa	Temasek Life Science Laboratory, National University of Singapore, Singapore	2009

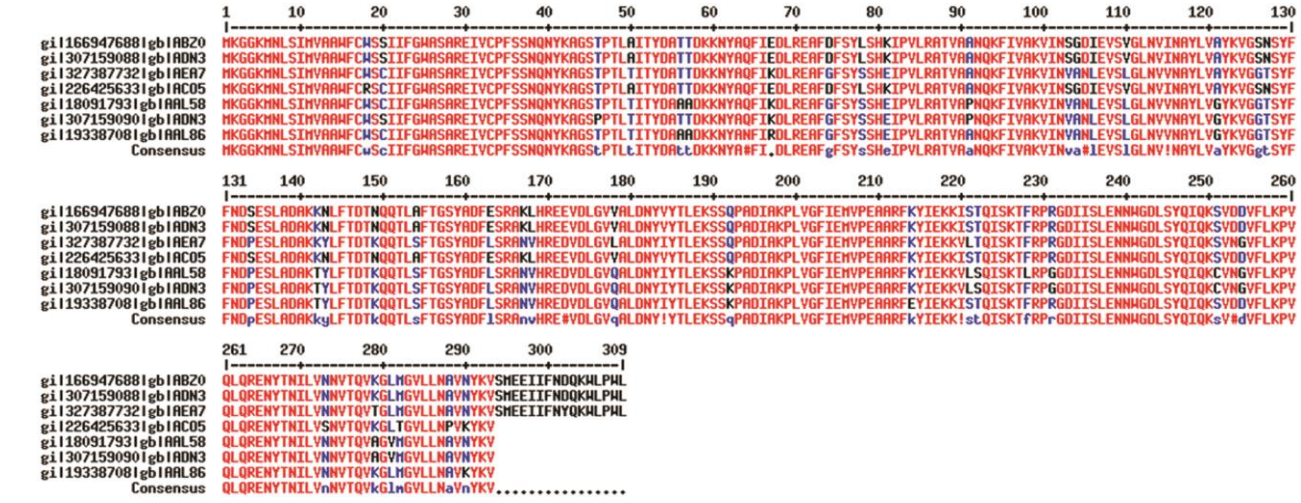

Fig. 1 — Multiple alignment of the curcin protein sequences retrieved from NCBI database.

servers[17,18]. While the Modweb gives the homology modeling and I-TASSER is an *ab initio* protein modeling server, which predicts protein structures based on a template as well as using *ab initio* methods. The active site predicted by I-TASSER based on protein templates were also taken into account for docking.

Protein-RNA Docking Using Haddock

The 28S rRNA structure to which the curcin, type 2 RIP is known to interact was retrieved from PDB with accession ID: 430D[19]. Further modification was made using Viewerlite[20] by deletion of Mg ions and water molecules to prevent hindrance to protein-RNA interaction. Such extraneous molecules must be removed from the downloaded structure before docking or molecular dynamics (MD) simulations as they can interfere with the mode of interaction. In order to target the Sarcin-Ricin loop or the GAGA loop, the bases G14, A15, G16 and A17 were chosen. The rRNA model was docked to curcin using HADDOCK[21] with default parameters. Haddock is a user-friendly interface with access to a dedicated cluster and e-NMR grid infrastructure and requires PDB structures of molecules to be docked and the interacting aa residues as input. The docking process can also be customized using more complex interfaces. Haddock stands for High Ambiguity Driven protein-protein DOCKing, though it also calculates docking interfaces for protein-nucleic acids complexes. The docking information is derived from identified protein interfaces rather than using an *ab initio* approach. As both ricin and curcin are RIPs, the mode of interaction of ricin becomes crucial to understand to find similarities or differences. Hence for comparison, ricin was also docked to rRNA to ascertain its binding site using Haddock with default parameters.

Protein Structure Modeling of Post-Modification Curcin and Docking with rRNA

After ascertaining the aa residues involved in the active site of curcin, these residues were removed from the primary sequence and modeled again. The critical residues found to exhibit interaction with rRNA; namely, Asn116, Tyr118, Asp133, and Glu135 were removed from the sequence and then modeled using Modweb. This model was again docked to rRNA to see the effect of modification on the interaction.

Mutational Analysis of Curcin and its Effect on rRNA Binding

The curcin model was also subjected to mutation in terms of aa residues involved in binding with rRNA molecule. Further to elucidate the role of crucial aa residues involved in interaction, these residues were mutated to different aa residues with different physiological properties. The mutations were carried out using mutate tool of Swiss-PDBViewer[22]. These mutated variants were again docked to rRNA molecule targeting the same binding site to find out their role in binding. The first variant constituted the mutation Tyr118 to Ala118, second with mutation Tyr118 to Gly118, and third with Tyr118 and Asp133 to Gly118 and Gly133, respectively.

Results and Discussion

The curcin genes cloned and sequenced from different laboratories across four countries and deposited in public domain (Table 1) were retrieved from the NCBI database. The multiple sequence alignment using MULTALIGN, produced sequence similarity of almost 98% (Fig. 1). The sequence id: ABZ04128.1 shared 100% homology with the sequence Id: ADN39429.1 and these were the longest full length sequence retrieved, so for all the modeling studies, the ABZ04128.1 sequences were used.

Protein Modeling of Curcin

The protein 3D structure of curcin was deduced using both homology-based approach and *ab initio* approach (Table 2 & 3). The structures obtained by ModWeb and I-TASSER were found to be structurally similar as shown in Fig. 2. The ModWeb structure was based on the template 2VLC[23] which corresponds to the crystal structure of type 2 ribosomal inactivating natural cinnamomin isoform. The protein segment of 2VLC from 4-299 became the template region.

I-TASSER models were based on the known protein structures retrieved from PDB. The top templates were 3KU0[24], 1CE7[25], 1BR6[26], 3KTZ[27], 3MVG[28] and 1ONK[29] (Table 4). Secondary structure of the modeled protein along with a confidence score of a particulars aa residue being a part of a secondary structure was predicted. Residues 12-23, 60-72, 159-167, 177-187, 195-222, 231-262, 275-279 and 296-298 are part of α-helices with a confidence score of 7 or more. Residues 49-52, 95-101, 107-113 and 118-123 are regions belonging to the β-sheet secondary structure with a confidence score of 7 or more. Rest all residues have either α or β secondary structure conformation with low confidence value or a high confidence value of belonging to loop/coil structure.

The algorithm also predicted the solvent accessibility with respect to each aa residue as a part of curcin. The values range from 0 for buried residues

Table 2 — ModWeb result for full length curcin

Database Annotation	Protein size (aa)	Modeled segment	Size (aa)	Seq id (%)	PDB code	PDB segment	PDB comment
recname: full = rrnan-glycosidase;ec = 3.2.2.22;	309	51-308	258	33.00	2vlcA	4-299	Crystal structure of natural cinnamomin (isoform iii)
recname: full = rrnan-glycosidase;ec = 3.2.2.22;	309	51-290	240	34.00	1mrjA	2-245	Studies on crystal structures active center geometry and depurine mechanism of two ribosome-inactivating proteins
recname: full = rrnan-glycosidase;ec = 3.2.2.22;	309	51-288	238	38.00	3bwhA	2-241	Atomic resolution structure of cucurmosin, a novel type 1 RIP from the sarcocarp of *Cucurbita moschata*
recname: full = rrnan-glycosidase;ec=3.2.2.22;	309	51-288	238	38.00	1cf5A	2-241	Beta-momorcharin structure at 2.55a

Table 3 — I-TASSER result: Secondary structure prediction and solvent accessible residue prediction for curcin sequence

Primary sequence of curcin	Secondary structure prediction	Prediction of solvent accessibility
MKGGKMNLSIMVAAWFCWSS	CCCCCSSSSHHHHHHHHHHHHHHHHC	865430100000111200000101232333
IIFGWASAREIVCPFSSNQN	CCCCCCCCCCCCCCCCCCCCSSSSSSSCCC	223233444344342220302231434473 0
YKAGSTPTLAITYDATTDKK	CHHHHHHHHHHHHHHHCCCCCCCCCC	340043017404673411000002462644 3
NYAQFIEDLREAFDFSYLSHKIPVLRATV	SSCCCCCCCCCSSSSSSSCCCCCSSSSSSCC	200102033474320000022310000000 1
AANQKFIVAKVINSGDIEVSVGL	CSSSSSSSCCCCSSSSCCCCCCCCCCCCCC	443211033244344344200441443404 2
NVINAYLVAYKVGSNSYFFNDSE	CCCSSSCCCCCCHHHHHHHHCCCCCCCC	423153027306351661411352034003 3
SLADAKKNLFTDTNQQTLAFTG	CCHHHHHHHHHHHHCCCCCHHHHHHH	024354443003000000000112204302
SYADFESRAKLHREEVDLGVVA	HHHHHHHHHHHHHHHHHHHHHHHHHC	430352134523143301411530340031 0
LDNYVYTLEKSSQPADIAKPLVG	CCCCCCHHHHHHHHHHHHHHHHHHH	251473413430403247443120320430 2
FIEMVPEAARFKYIEKKISTQISKT	CCCCCCCCSSSSCCCCCSSSSSSHHHHHHH	410000001243443343124324412336
FRPRGDIISLENNWGDLSYQIQK	HHHSSCCCCCCCCHHHHHHHCCCCCCCC	
SVDDVFLKPVQLQRENYTNILVN		
NVTQVKGLMGVLLNAVNYKVS		
MEEIIFNDQKWLPWL		

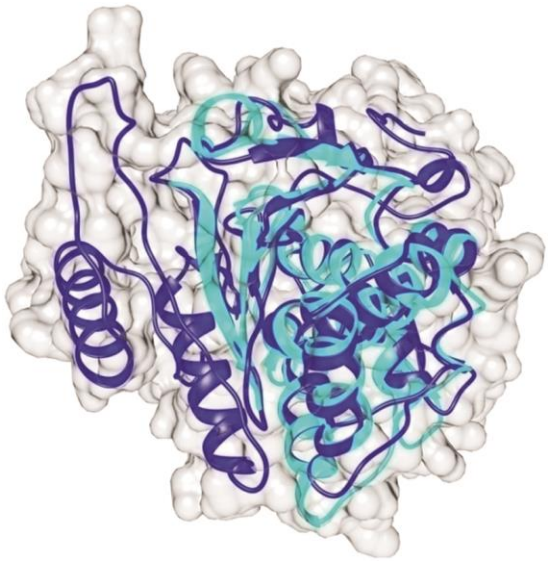

Fig. 2 — Diagrammatic representation of structural similarity between curcin models obtained from ModWeb (in Cyan) and from I-TASSER (in blue).

to 9 for highly exposed residues. Buried residues are generally not involved in any interaction with other biomolecules or ligands and are generally hydrophobic in nature. During protein folding, these get buried to avoid water solvent which makes up for a protein's environment. The exposed aa residues lie on top and form the surface of proteins and generally take part in all interactions. Curcin, being a RIP binds to 28S rRNA. The templates used were ranked on the basis of coverage, normalized Z score and identity. The identity 1 represents percentage sequence identity of these templates in the aligned region with curcin sequence. Identity 2 specifies the percentage sequence identity of the whole template chains with the curcin sequence. Coverage represents the number of aligned residues divided by the length of the query protein. The normalized Z score value of more than one for an alignment is taken to be good. The best model from Modweb was compared with the one obtained from I-TASSER. A good structural similarity was found between the two models. As the model obtained from

Table 4 — Top ten threading templates used by I-TASSER for modeling of curcin protein sequence

Rank	PDB Hit	Identity1	Identity2	Coverage	Normalized Z score	Threading algorithm
1	3KU0	0.33	0.28	0.79	2.34	MUSTER
2	1CE7	0.30	0.24	0.77	2.44	dPPAS
3	1BR6	0.34	0.29	0.83	3.20	Neff-PPAS
4	3KU0	0.32	0.28	0.79	3.78	PPAS
5	3KU0	0.32	0.28	0.79	2.57	wdPPAS
6	1BR6	0.33	0.29	0.83	2.43	SPARKS-X
7	3KTZ	0.33	0.28	0.79	4.40	SP3
8	3KTZ	0.33	0.28	0.79	2.61	HHSEARCH2
9	3MVG	0.30	0.25	0.78	3.86	PROSPECT2
10	1ONK	0.33	0.26	0.79	4.46	FFAS03

Table 5 — Prediction of active site based on template proteins with similar binding site

Rank	$Cscore^{LB}$	PDBhit	TM-score	RMSD	IDEN[a]	Cov.	BS-score	Lig. Name	Predicted binding site residues
1	0.67	1ifsA	0.755	1.56	0.336	0.790	1.59	ADE	118, 119, 157, 158, 159, 204, 212
2	0.65	1il5A	0.767	1.62	0.327	0.806	1.64	DDP	116, 118, 158, 159, 212
3	0.50	3rtjA	0.758	1.55	0.336	0.793	1.58	Nucleic Acid	117, 118, 119, 157, 158, 159, 160, 204, 209, 212, 238, 239
4	0.46	3o5wA	0.753	1.40	0.322	0.783	1.43	H35	118, 157, 159, 163, 170, 201, 204, 209, 212, 238
5	0.29	1tcsA	0.745	1.65	0.339	0.783	1.40	NDP	117, 118, 119, 157, 159, 204, 208, 209, 212, 238, 239, 240, 241, 242, 243, 246, 247, 258
6	0.07	3bwhA	0.744	1.46	0.367	0.777	1.40	Mul.Part	255, 264, 268, 269, 270, 274, 276, 277
7	0.07	1uq50	0.757	1.58	0.329	0.793	1.31	PEPTIDE	215, 233, 237, 252, 253, 254, 255, 263, 270, 271, 272, 273, 274, 276, 277, 281
8	0.06	2aaiA	0.766	1.72	0.324	0.806	1.26	Mul.Part	256, 259, 273
9	0.06	3n3xA	0.742	1.65	0.311	0.780	1.34	Peptide	215, 218, 219, 222, 223, 228, 229, 230, 231, 264, 277, 282

I-TASSER is the complete one, it was selected for performing interaction analysis with 28S rRNA. The structural similarity between the modeled proteins can be deduced from Figure 2. The N-terminal region of curcin was modeled by I-TASSER and shows the presence of a helix-loop-helix structure. Taking the homology modeling techniques to be much accurate than *ab initio* modeling, the Modweb model was selected for docking against 28S rRNA.

Active Site Prediction by I-TASSER Based on TM Score

The prediction of binding sites using TM-score of I-TASSER compared curcin protein model to that of known similar PDB templates and yielded 9 different binding sites. The site predicted by comparing to the PDB template 3RTJ[30] with a TM-Score of 0.758, root mean square deviation (RMSD) value of 1.55, Identity 1 of 0.336, coverage of 0.793 and BS-Score of 1.58 is the appropriate one as the ligand is a nucleic acid in this case (Table 5). The template is of ricin A-chain active site. Curcin sequence is analogous to ricin A-chain sequence and seems to share structural characteristics as well. The predicted binding site residues of this site are 117, 118, 119, 157, 158, 159, 160, 204, 209, 212, 238 and 239.

Protein-RNA Docking using Haddock

To ascertain the rRNA binding site of curcin, the 28S rRNA structure was retrieved from PDB (PDB ID: 430D). The structure corresponds to rat 29 nucleotides long 28S rRNA containing the domain essential for binding with elongation factors. This domain is known as the sarcin/ricin loop and includes four crucial nucleotides G14, A15, G16 and A17. Out of these, the 15^{th} adenine is targeted by toxins like ricin and curcin. The docking run with the prediction interface of Haddock yielded a score of -74.9 +/- 0.5. The experimental information is incorporated in the form of ambiguous interaction restraints (AIRs) which are defined based on the active and passive residues defined by the user. For every active residue, one AIR is defined between that residue and all active and passive residues. This AIR energy term is also introduced in the calculation.

Haddock modeled 184 interaction (docked) structures from a maximum limit of 200 structures. These 184 were grouped into 6 clusters which represent 92.0% of water-refined Haddock generated structures. The top cluster is the most reliable according to HADDOCK. Its Z-score indicates how many standard deviations this cluster is located from the average in terms of score (the more negative the better). The HADDOCK score is a weighted sum of Van der Waals, electrostatic, desolvation and restraint violation energies together with buried surface area. The top ranked cluster - 4 with 8 structures gave the highest scoring structure. Its RMSD from the overall lowest-energy structure was 7.0+/- 0.3. The RMSD value should be close to zero as it represents deviation from the average structure. This structure has very low RMSD value. Its Van der Waals energy was - 41.8 +/- 7.1, electrostatic energy was -339.4 +/- 27.6, desolvation energy was 31.2 +/- 5.3, restraints violation energy was 35.8 +/- 23.82, buried surface area was 1433.0 +/- 108.4, and a good Z score of -1.4.

Protein Modeling and Docking after Curcin Modification

Curcin was modified after deletion of 4 crucial residues found to aid in binding to rRNA. These residues were Asn116, Tyr118, Asp133, and Glu135. After deletion, curcin was modeled again using ModWeb. The results are tabulated in Table 6. The difference in three dimensional loop structures before and after modification can be seen in Figure 3. The conformational superimposition of curcin protein structure before and after modification (removal of 4 interacting aa residues) revealed a slight conformational change in the domains (shown in green). The loop consisting residues 116, 117, and 118 protrudes more towards outside in comparison to the pre-modification curcin structure. However, this structural change does not hamper the binding of rRNA to curcin rather increases it as can be deduced from the post-modification Haddock score of -92.0+/-5.4. This complex is also highly stable in terms of the Z score of -1.8. A possible explanation of this observation can be the removal of bulky side-chain residues and thus providing more space for the approaching rRNA molecule to bind. A similar observation is seen when these residues are mutated to residues with small or no side-chains as discussed in the next section.

Mutational Analysis of Curcin and its Effect on rRNA Binding

The mutational analysis carried by modifying crucial aa residues to physicochemically different ones has been discussed to elucidate the role of these residues. The HADDOCK scores corresponding to the mutations were -122.0 +/- 8.5 with Z score of -1.5 for the first mutant (Tyr118→Ala118), a score of -115.6 +/- 8.0 with Z score of -1.5 for the second mutant (Tyr118→Gly118), a score of -81.2 +/- 3.2 with Z score of -1.8 for the third mutant (Tyr118→Gly118 and Asp133→Gly133). These scores show increased binding affinity after modification than before and indicate towards the steric hindrance caused by the presence of these residues to the approaching rRNA molecule. In the absence of these bulky side-chain residues, probably the cavity size increases and provides space for the rRNA to bind well. The hindrance providing region of curcin to the approaching rRNA molecule has been depicted in Figure 4 in green. The removal or mutation of these residues renders accessibility for interaction and thus can be a probable explanation for increased binding affinity post modification. An important observation

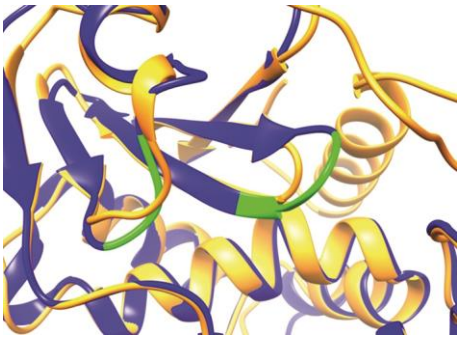

Fig 3 — Difference in loop structure pre- and post- modification.

Table 6 — ModWeb modeling results for curcin modification

Protein Size	Modeled Segment	Size	SeqId (%)	PDBcode	PDB Segment	PDB Comment
305 aa	51-284	234	38.00	3bwhA	2-241	Atomic resolution structure of cucurmosin, a novel type 1 RIP from the sarcocarp of *Cucurbita moschata*
305 aa	48-288	241	34.00	3ktzA	4-250	Structure of gap31
305 aa	57-283	227	32.00	1ggpA	12-245	Crystal structure of *Trichosanthes kirilowii* lectin-1 and its relation to the type 2 ribosome inactivating proteins
30 5aa	51-304	254	32.00	2vlcA	4-299	Crystal structure of natural cinnamomin (isoform III)

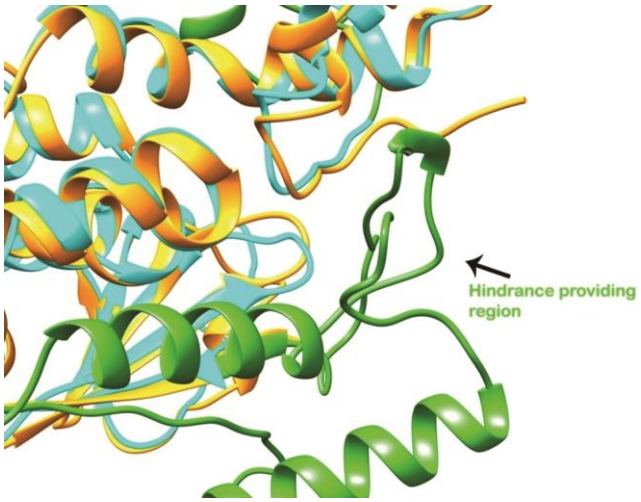

Fig. 4 — Hindrance providing region of the curcin model.

is that all RIPs render toxicity by N-glycosylation of a crucial adenine residue on the conserved GAGA loop of 28S rRNA molecule. This inhibits protein synthesis by interfering with the binding of rRNA with elongation factors. This gives a cue to another perspective of not just taking into account the physical mode of binding between rRNA and curcin but also the chemical mode of action. Probably, the crucial residues reported in this study, though provide hindrance, also cause N-glycosylation of the above said adenine residue (A15 of GAGA loop) and catalytically damaging ribosomes. Their removal or mutation may cause easy accessibility for the rRNA to bind but may also result in loss of N-glycosylation activity. This study, thus, paves way for a detailed experimental analysis on the chemical mode of action and gives site-specific clues as to where the rRNA molecule binds. The inhibitors designed to these regions might provide higher and reliable gene silencing.

Acknowledgement

Authors would like to thank AKMU, Indian Agricultural Research Institute and ICAR-Central Agroforestry Research Institutefor usage of all computational facilities.

References

1 Andrew J K, Wei H, Mark F, Ramiaramana & Ian A G, Potential of *Jatropha curcas* as a source of renewable oil and animal feed, *J Exp Bot*, 1093 (2009) 1-9.

2 Barbieri L, Battelli M G & Stirpe F, Ribosome-inactivating proteins from plants, *BBA-Rev Biomembranes*, 1154 (1993) 237-282.

3 Van Damme E J M, Hao Q, Chen Y, Barre A, Vandenbussche F *et al*, Ribosome-inactivating proteins: a family of plant proteins does more than inactivate ribosomes, *Crit Rev Plant Sci*, 20 (2001) 395-465.

4 Chow T, Feldman R A, Lovett M A & Piatak M, Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type ribosome-inactivating protein, *J Biol Chem* 265 (1990) 8670-8674.

5 Baloch Z, Cohen S & Coffman F D, Synergistic interactions between tumor necrosis factor and inhibitors of DNA topoisomerase I and II, *J Immunol*, 145 (1990) 2908–2913.

6 Staubmann R, Ncube I, Gubitz G M, Steiner W & Read J S, Esterase and lipase activity in *Jatropha curcas* L. seeds, *J Biotechnol*, 75 (1999) 117–126.

7 Martı′nez-Herrera J, Siddhuraju P, Francis G, Davila-Ortiz G & Becker K, Chemical composition, toxic/antimetabolic constituents, and effects of different treatments on their levels, in four provenances of *Jatropha curcas* L. from Mexico, *Food Chem*, 96 (2006) 80–89.

8 King A J, He W, Cuevas J A, Freudenberger M, Ramiaramanana D *et al*, Potential of *Jatropha curcas* as a source of renewable oil and animal feed, *J Exp Bot*, 60 (2009) 2897–2905.

9 Lin J, Li Y, Li Y, Xu Y, Yan F *et al*, Rapid analysis of *Jatropha curcas* gene functions by virus-induced gene silencing, *Plant Biotechnol J*, 7 (2009) 1–13.

10 Patade VY, Khatri D, Kumar K, Grover A, Kumari M, *et al*, RNAi mediated curcin precursor gene silencing in *Jatropha* (*Jatropha curcas* L.), *Mol Biol Rep*, 41 (2014) 4305-12.

11 www.ncbi.nlm.nih.gov/

12 Corpet F, Multiple sequence alignment with hierarchical clustering, *Nucleic Acids Res*, 16 (1988) 10881-10890.

13 Eswar N, John B, Mirkovic N, Fiser A, Ilyin V A *et al*, Tools for comparative protein structure modeling and analysis, *Nucleic Acids Res*, 31 (2003) 3375-3380.

14 Fiser A, Do R K G & Šali A, Modeling of loops in protein structures, *Protein Sci*, 9 (2000) 1753-1773.

15 Martí-Renom M A, Stuart A C, Fiser A, Sánchez R, Melo F *et al*, Comparative protein structure modeling of genes and genomes, *Annu Rev Bioph Biom*, 29 (2000) 291-325.

16 Šali A & Blundell T L, Comparative protein modelling by satisfaction of spatial restraints, *J Mol Biol*, 234 (1993) 779-815.

17 Zhang Y, I-TASSER server for protein 3D structure prediction, *BMC bioinformatics*, 9 (2008) 40.

18 Roy A, Kucukural A & Zhang Y, I-TASSER: a unified platform for automated protein structure and function prediction, *Nat Protoc*, 5 (2010) 725-738.

19 Correll C C, Munishkin A, Chan Y L, Ren Z, Wool I G *et al*, Crystal structure of the ribosomal RNA domain essential for binding elongation factors, *Proc Natl Acad Sci- Biol*, 95 (1998) 13436-13441.

20 Viewerlite, Version 5.0, Accelrys Inc.

21 de Vries S J, van Dijk M & Bonvin A M J J, The HADDOCK web server for data-driven biomolecular docking, *Nat Protoc*, *5* (2010) 883-897.

22 Guex N &Peitsch M C, SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling, *Electrophoresis*, 18 (1997) 2714-2723.

23 Azzi A, Wang T, Zhu D W, Zou Y S, Liu W Y *et al*, Crystal structure of native cinnamomin isoform III and its

comparison with other ribosome inactivating proteins, *Proteins*, 74 (2009) 250-255.

24 Li H G, Huang P L, Zhang D, Sun Y, Chen H C *et al*, A new activity of anti-HIV and anti-tumour protein GAP31: DNA adenosine glycosidase–structural and modeling insight into its functions, *Biochem Bioph Res Co*, 391 (2010) 340-345.

25 Krauspenhaar R, Rypniewski W, Kalkura N, Moore K, DeLucas L *et al*, Crystallisation under microgravity of mistletoe lectin I from *Viscum album* with adenine monophosphate and the crystal structure at 1.9 Å resolution, *Acta Crystallogr*, D 58 (2002) 1704-1707.

26 Yan X, Hollis T, Svinth M, Day P, Monzingo A F *et al*, Structure-based identification of a ricin inhibitor, *J Mol Biol*, 266 (1997) 1043-1049.

27 Paulsen R B, Seth P P, Swayze E E, Griffey R H, Skalicky J J *et al*, Inhibitor-induced structural change in the HCV IRES domain IIa RNA, *Proc Natl Acad Sci USA*, 107 (2010) 7263-7268.

28 Meyer A, Weber W, Singh T P & Betzel C, Native structure of IRIP, a type I ribosome inactivating protein from *Iris hollandica* var. at 1.25 A, doi : 10.2210/pdb3MVG/pdb.

29 Gabdoulkhakov A G, Savoshkina Y, Krauspenhaar R, Stoeva S, Konareva N *et al*, Mistletoe lectin I from *Viscum album*, doi : 10.2210/pfb10NK/pdb.

30 MonzingoA F & Robertus J D, X-ray analysis of substrate analogs in the ricin A-chain active site, *J Mol Biol*, 227 (1992) 1136-1145.