

Jatropha curcas hemagglutinin is similar to a 2S albumin allergen from the same source and has unique sugar affinities

Divya N. Nair · Vijay Singh · Yoshiki Yamaguchi ·
Desh Deepak Singh

Received: 28 April 2012 / Accepted: 25 June 2012
© Springer-Verlag 2012

Abstract We have previously reported the purification and preliminary X-ray characterization of a hemagglutinin from the seeds of *Jatropha curcas* and, with the detailed sequencing information available now, we find that it is similar to a 2S albumin allergen isolated from the same source. Through a search of *Jatropha* genome database (<http://www.kazusa.or.jp/jatropha/>), we map it to the sequence id JcCA0234191 (now referred to as Jcr4S00619.70 in the new version, release 4.5) which has a conserved alpha amylase inhibitor/seed storage protein domain found in the 2S albumin allergens. The putative sequence of the small and large chains of the protein is assigned and the total mass of the two subunits matches with the intact mass 10 kDa determined through MALDI. The protein retains hemagglutination activity between pH 6–9 and up to 60 °C on heat treatment and its hemagglutination activity is inhibited by sialic acid and fetuin. Bioinformatics studies show that the isolated protein sequence clusters in close association with a 2S albumin from *Ricinus communis* in phylogeny analysis and has a conservation of the characteristic four disulfide linkage pattern. Hemagglutinins and lectins are known to have allergenic effects through their interaction with immunoglobulin E and histamine release and earlier studies have shown that this interaction can be inhibited by lectin-specific sugars. We hope this report bridges the plant allergens and hemagglutinins further for

exploring possible mediation of allergenic activity through sialic acid and complex sugar interactions and generates further interest in the area.

Keywords Alpha amylase inhibitor · Plant food allergens · Hemagglutination · *Jatropha curcas*

Abbreviations

nsLTP	Non-specific lipid transfer protein
AAI	Alpha amylase inhibitor
NeuNAc	N-acetylneuraminic acid
MIC	Minimum inhibitory concentration
DTT	Dithiothreitol
BME	Beta mercaptoethanol
RIP	Ribosomal inactivating protein
JHg	<i>Jatropha</i> hemagglutinin
SS	Seed storage protein
RBC	Red blood cells

Introduction

Jatropha curcas L. is an interesting member of the *Euphorbiaceae* family important for its potential as a bio-fuel. The oil producing seeds from the plant are highly toxic owing to the presence of the diterpenoid phorbol esters which promote tumors and have severe pathological effects on kidneys and lungs (Li et al. 2010). A type-1 ribosomal inactivating protein (RIP), curcin isolated from *J. curcas* seeds, acts on ribosomes and inactivates protein synthesis (Stirpe et al. 1976) and the presence of these compounds make the seeds inedible. We have previously reported on the isolation and preliminary X-ray characterization of a hemagglutinin (referred to as JHg further) from this plant (Nair et al. 2011) and with the detailed sequencing result,

D. N. Nair · V. Singh · D. D. Singh (✉)
Department of Bioinformatics and Structural Biology,
Indian Institute of Advanced Research, Koba,
Gandhinagar 382007, Gujarat, India
e-mail: ddsingh@iiar.res.in

Y. Yamaguchi
Structural Glycobiology Team, Systems Glycobiology Research
Group, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

and analysis of the genome sequence of *J. curcas* (<http://www.kazusa.or.jp/jatropha/>) (Sato et al. 2011) we find that it is similar to the previously reported 2S albumin plant food allergen, Jat c1, from the same source (Maciel et al. 2009). Plant food allergens are either storage or defense-related proteins with four main families/superfamilies: Bet v1, profilin family, Cupin and Prolamin superfamily (Breiteneder and Clare Mills 2005) (<http://www.allergen.org>). The prolamin superfamily comprises the largest number of allergenic plant food proteins and consists of proline and glutamine-rich α -helical proteins with a conserved skeleton of eight cysteine residues that serve several biologic functions. They have three major groups: the seed storage 2S albumins found in tree nuts and seeds, the defense-related nsLTP found in soft fruits and vegetables, and the cereal α -amylase/trypsin inhibitors (Breiteneder and Ebner 2000; Martin et al. 2007).

In an effort to isolate lectins from *J. curcas* seeds, we identified JHg showing hemagglutination activity and on further analysis based on the detailed sequencing we characterized it further. In this communication, we show that JHg is similar to Jat c1, and its hemagglutination activity is inhibited by sialic acid and complex sugars. This hopefully gives additional interesting features for the family. The protein is further characterized biochemically and through bioinformatics analysis.

Materials and methods

Isolation and purification

The isolation and purification protocol for JHg has been reported in our earlier communication (Nair et al. 2011).

SDS PAGE and molecular weight determination

SDS PAGE was done using the method of Laemmli and tricine gel electrophoresis by the method of Schägger (2006), in which 20 % (w/v) acrylamide separating gel was used and the protein sample was denatured using 50 mM DTT. The separated protein bands were visualized by Coomassie Brilliant Blue stain. The exact molecular weight of JHg was determined with Ultraflex MALDI TOF/TOF (Bruker Daltonics) using 25KVA reflector mode method with an average of 500 shots and analyzing the data with Flex Analysis software.

Agglutination assay and inhibition by sugars

The agglutination assay was carried out as previously reported (Nair et al. 2011). Briefly, 50 μ l of JHg solution (0.8 mg/ml in PBS) serially diluted, 50 μ l of 1 % rabbit

erythrocytes (RBC) suspension in PBS and 50 μ l of PBS (pH 7.4) were added to a 96-well plate and button formation (sedimented cells) was observed visually, as well as after 1 h of incubation at room temperature, if there was agglutination or not. For inhibition assay, sugars in the concentration range of 3–100 mM were added to the wells in place of 50 μ l of buffer and the following sugars were tested: D-mannose (Hi-media), D-galactose (Himedia), lactose (SD finechem), N-acetyl D-glucosamine (Sigma), N-acetyl D-galactosamine (Himedia), N-acetylneuraminic acid (sialic acid) (Sigma), fucose (Sigma), fetuin (Sigma) and D-glucose (Sigma). The minimum concentration of sugar showing detectable button formation in the assay was taken as MIC.

Effect of temperature and pH

The effect of temperature on hemagglutination activity of JHg was determined by incubating different aliquots of 50 μ l of JHg (0.8 mg/ml in PBS) from 10–100 °C for 15 min and then cooling the tubes on ice for half an hour followed by the agglutination assay to check activity. The pH stability of JHg was determined by incubating different protein aliquots in buffers ranging from pH 2 to 10 for 3 h at room temperature followed by a stabilization at 4 °C for 2 h and neutralization of pH and finally checking the hemagglutination activity.

N-terminal sequencing

N-terminal sequencing was performed on Procise 494 cLC system (Applied Biosystems, USA) using 10 μ g of the purified protein for the 30 reaction cycles.

Bioinformatics analysis

NCBI BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) was used for sequence similarity search, multiple sequence alignment was done using clustal X (Jeanmougin et al. 1998) and the phylogenetic analysis of JcCA0234191.10 and 31 sequences of reviewed plant alpha amylase inhibitor (AAI) sequences downloaded from UniProt database (<http://www.uniprot.org/>) was done using neighbor joining method in MEGA program (Tamura et al. 2007). Domain search was done using the conserved domain database (CDD) (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) InterPro (<http://www.ebi.ac.uk/interpro>) and PFAM (<http://pfam.sanger.ac.uk/>). Molecular modeling was done using I-Tasser server (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>), structural homologs were found using DALI server (http://ekhidna.biocenter.helsinki.fi/dali_server/start) and the disulfide topology was found using DiANNA 1.1 Web server (<http://clavius.bc.edu/~clotelab/DiANNA/>).

Statistical analysis

Since the results for SDS PAGE and the sugar inhibition studies were qualitative, they were confirmed by repeating the results on more than three independent experiments.

Results

Molecular weight determination

JHg shows two bands between 5 and 10 kDa and another one between 3 and 5 kDa in tricine SDS PAGE (Fig. 1a) and MALDI TOF which shows a peak at 10.087 kDa (Fig. 1b).

Sugar binding specificity

The hemagglutination activity of the JHg was inhibited by sialic acid and fetuin with an MIC of 5 mM and 10 mM, respectively. None of the other sugars showed inhibition of hemagglutination activity in the tested concentration range.

Effect of pH and temperature on hemagglutination activity

JHg shows hemagglutination activity between pH 6 and 9 and loses activity rapidly when the pH is out of this range. The thermal denaturation result shows that it remains stable till 60 °C and above this temperature rapidly loses activity.

N-terminal sequencing

N-terminal sequencing of JHg shows the presence of the following residues: VRDIrKKEAERRDLSSKENYITQR RRr.

Bioinformatics analysis

The N-terminal amino acid sequence of JHg maps to the sequence id JcCA0234191.10 present in the *Jatropha* database (<http://www.kazusa.or.jp/jatropha/>) as the only significant hit, and BLAST output of JcCA0234191.10 as query shows a similarity with 2S albumin seed storage proteins from *Ricinus communis*, *Medicago truncatula*, *Hordeum vulgare* possessing alpha amylase inhibitors (AAIs) and seed storage protein domain. Phylogenetic analysis of 31 sequence of AAIs from plants downloaded from UniProt and JcCA0234191.10 shows four major clusters and JcCA0234191.10 is found in the clade with the closest member as 2SS-RICCO, a 2S albumin from *R. communis* in cluster 1 (Fig. 2).

The 3-D model of JcCA0234191.10 predicted by I-Tasser server shows four alpha-helices (H1, H2, H3, H4) (Fig. 3a) and it superposes very well (Fig. 3b) on the closest homolog, Ric c 3 protein (PDB code: 1PSY), a 2S albumin allergen whose 3-D structure has been solved (Pantoja-Uceda et al. 2003) with a root mean square deviation of 1.2 Å. The model of two chains of the mature protein obtained after truncation of the pre-protein (residues in the N-terminal and those between the two chains in JcCA0234191.10 sequence are removed) and corresponding sequences are mapped and shown (Fig. 3c).

Discussion

Jat c1 and JHg are similar

We have previously reported the purification and preliminary X-ray characterization of a hemagglutinin of about 10 kDa from the seeds of *J. curcas* (Nair et al. 2011). Literature search shows a report on isolation of a 2S

Fig. 1 **a** Tricine gel profile of JHg which shows two bands between 5–10 kDa and 3–5 kDa (lane 1). **b** MALDI TOF of JHg shows a peak at 10.087 kDa

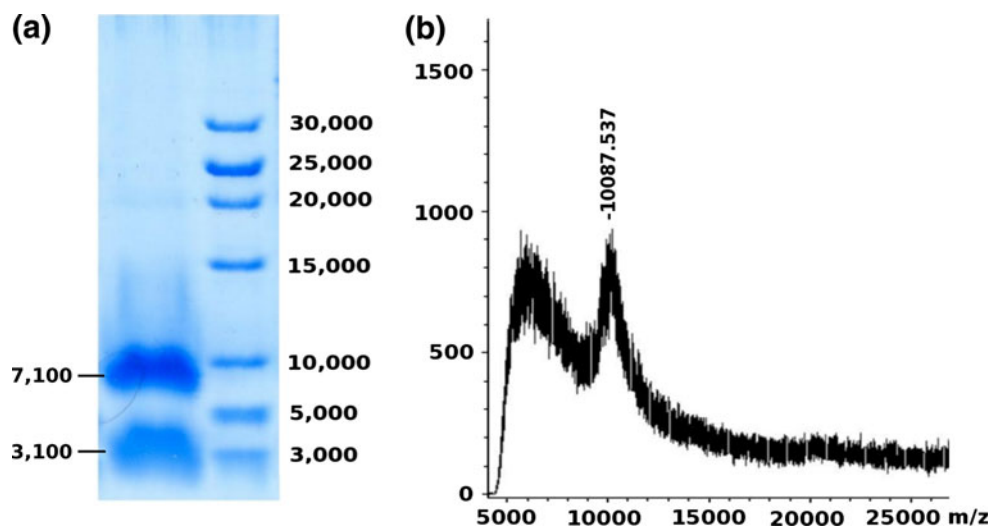
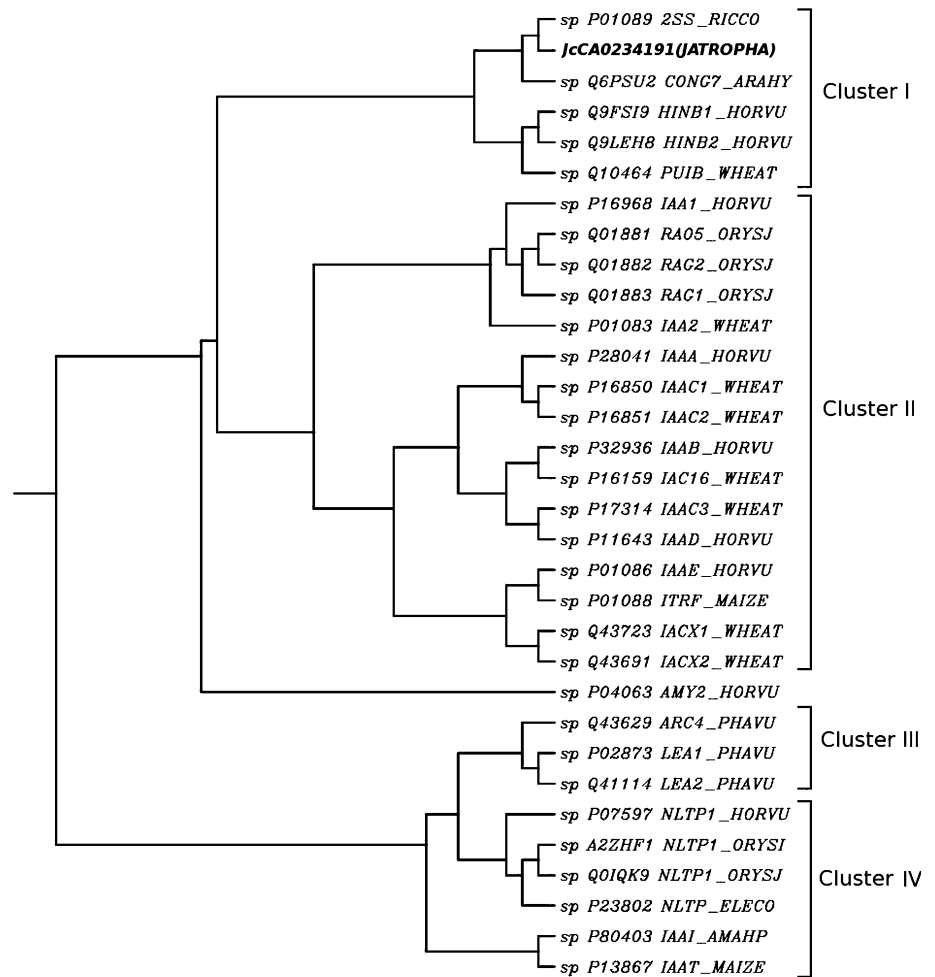


Fig. 2 The phylogenetic analysis of alpha amylase inhibitory (AAI) protein (31 sequences of plant AAI downloaded from UniProt database) shows four distinct clusters. Cluster I contains 2S albumin seed storage proteins and the JcCA0234191 marked as (Jat), cluster II belongs to cereal trypsin and alpha amylase inhibitor, cluster III contains lectin alpha amylase inhibitors and cluster IV contains nsLTP (non-specific lipid transfer protein)

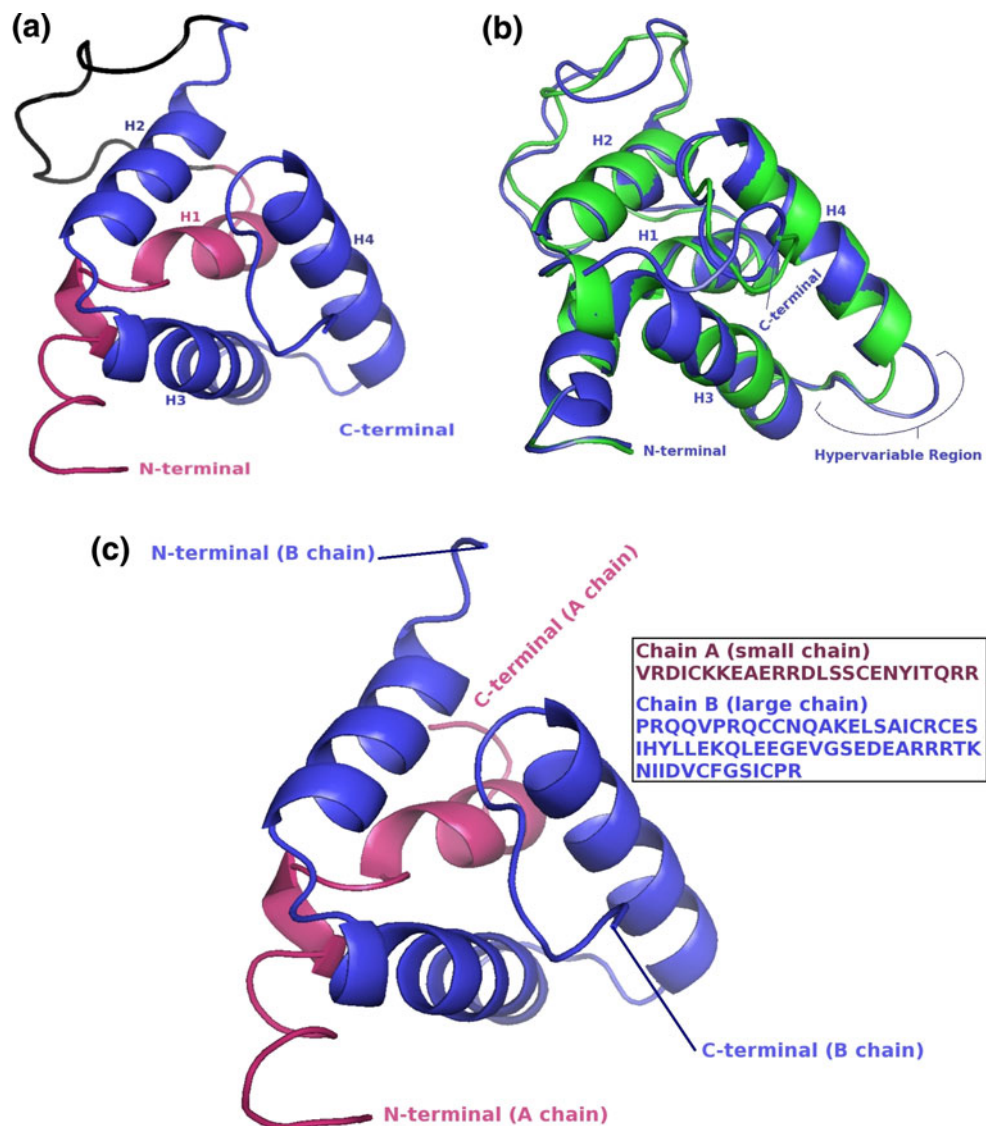


albumin allergen, Jat c1, of about 12 kDa from the same source (Maciel et al. 2009) which was shown to have two sub-units of 7.2 (large chain) and 5 kDa (small chain) with partial N-terminal sequencing result. The *J. curcas* genome has been sequenced recently (Sato et al. 2011) and a total of 40,929 complete and partial structures of protein coding genes have been deduced and deposited in the open source database (<http://www.kazusa.or.jp/jatroph/>). Blast analysis using the sequence obtained from N-terminal sequencing of JHg against the *Jatropha* database obtained JcCA0234191.10 (now referred to as Jcr4S00619.70 in the new version, release 4.5) as the hit. This protein has 2S albumin allergens as its close homolog and the AAI domain as found by conserved domain database (CDD) search. Since our earlier analysis showed the presence of a single band for JHg at 10 kDa in SDS PAGE and Jat c1 showed the presence of two bands, it was pertinent to resolve this ambiguity if the two proteins were similar. We denatured the protein sample with DTT instead of BME and ran it on tricine gel and this resulted in the presence of two bands (Fig. 1a). Also, it is reported that the full length 2S albumin proteins undergo a modification through

proteolytic cleavage with the loss of a linker peptide and short peptides from both the N- and C-terminal region (Crouch et al. 1983; Ericson et al. 1986), which explains the presence of two chains for JcCA0234191.10. The N-terminal sequencing result (VRDIrKKEAERRDLSSkE-NYITQRRr) obtained by us for JHg and the one reported for small chain of Jat c1 (VRDKCGEEAERTLXGCE-NYISQRR) matches significantly and this leads us to conclude that the small chain is of about 25 residues corresponding to a molecular weight of 3.1 kDa. Similarly, the partial sequence of large chain of Jat c1 (PRE-QVPRCCCNQALE) aligned with the initial 15 out of total 62 residues of JcCA0234191.10 toward the C-terminal, adding up to a molecular weight of 7.1 kDa. The two chains together give a mass of 10.2 kDa, close to the molecular ion peak obtained at 10.08 kDa, which we observe for the protein. Based on this analysis, the putative sequence for the small and large chain obtained from the precursor protein is shown for JHg (Fig. 3c).

Phylogenetic analysis of JcCA0234191 and the 31 AAI sequences downloaded from UniProt shows the presence of four clusters (Fig. 2). The first cluster contain 2S albumin

Fig. 3 **a** The modeled structure of the sequence of JcCA0234191. **b** The superposed structure of the modeled protein on its template (PDB: 1PSY). The hypervariable region is marked. **c** The 3-D model of the two chains of JcCA0234191, obtained after sequencing (see text) is shown and the sequences of the corresponding chains is shown in *inset box*



seed storage family proteins from *R. communis*, *Arachis hypogaea* and *H. vulgare* and the isolated hemagglutinin falls in this cluster in close association with 2SS-RICCO, a 2S albumin from *R. communis* (Pantoja-Uceda et al. 2003), thus confirming the family. The cereal α -amylase/trypsin inhibitors belong to the second cluster which consists of allergic proteins from *Triticum aestivum*, *Oryza sativa*, and *Zea mays* as its members, the third cluster consists of lectin AAI from *Phaseolus vulgaris*, whereas fourth cluster consists nsLTPs with members from *O. sativa* and *H. vulgare*.

The prolamin superfamily of plant food allergens is further divided into 2S albumin allergens, the defense-related nsLTP and cereal α -amylase/trypsin inhibitors. All the members consist of α -helical proteins with a conserved skeleton of at least eight cysteines in all of them and with the only difference in their disulfide linkage (Fig. 4b). Through a search of the Pfam and InterPro database, it is

found that the cysteine linkage is (C1–C5, C2–C3, C4–C7, C6–C8) for 2S albumins, (C1–C6, C2–C3, C4–C7, C5–C8) for nsLTP and (C1–C10, C2–C4, C3–C8, C5–C9, C6–C7) for cereal α -amylase/trypsin inhibitors which have additional two cysteine residues (<http://www.ebi.ac.uk/interpro/IEntry?ac=IPR006105>) (Oda et al. 1997). The multiple sequence alignment of some of these members along with the sequence of JcCA0234191.10 indicates that eight cysteine residues are conserved in all the sequences. The modeled structure of the isolated hemagglutinin shows a common fold of four helices (H1, H2, H3, H4), folded leaf, right-handed superhelix which has been stabilized by four disulfide bonds (Fig. 4a, b) as found in the 3-D structure of other members of 2S albumin allergens (Pantoja-Uceda et al. 2003). In between the third and fourth helices (H3 and H4), there is a short segment of “hypervariable region” (Fig. 3b), which has IgE binding epitopes essential for protein allergenicity (Moreno and Clemente 2008).

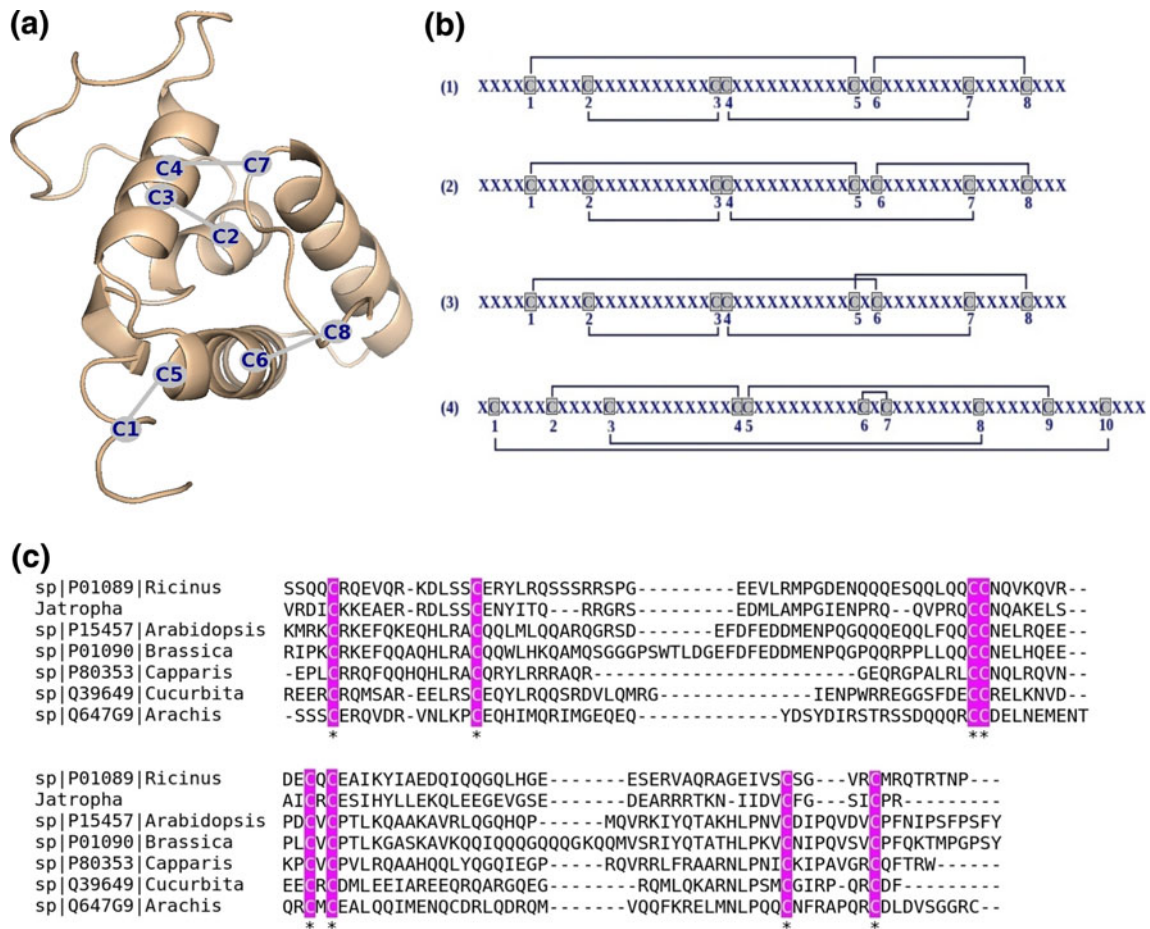


Fig. 4 **a** The disulfide linkage of cysteines present in the model of JcCA0234191. **b** The disulfide topology of the cysteine in different members of prolamin superfamily as found in the Pfam database: (1) JcCA0234191, (2) 2S albumin, (3) nsLTP and (4) cereal alpha

amylase inhibitor/trypsin inhibitor. **c** The eight conserved cysteine residues present in different 2S albumin protein sequences are colored and marked with *asterisk*

These features are conserved in the model generated for JHg by us and so we conclude that Jat c1, the name given to the 2S allergen reported earlier, and JHg are similar.

Unique sugar specificity of the isolated hemagglutinin

The hemagglutination activity of JHg is inhibited by sialic acid and fetuin, which is a glycoprotein, and none of the other simple sugars tested in the assay showed inhibitory activity. This is interesting because though lectins with sialic acid specificity are commonly found in animals, they are relatively rare in plants (Lehman et al. 2006). Some well-known examples of plant lectins with affinity for sialic acid are wheat germ agglutinin (WGA) (Peters et al. 1979; Wright 1990), *Morus alba* lectins (MLL 1 and 2) (Ratanapo et al. 1998), *Maackia* lectins (MAL, MAH) (Yamamoto et al. 1997; Imberty et al. 2000) and mistletoe lectin I (ML-I) (Müthing et al. 2004). Lectins from *Sambucus* spp. like *S. nigra* lectin (SNA), *S. canadensis* lectin (SCA) and

S. sieboldiana lectins (SSA) have specificity for Neu5A- α 2,6Gal and Neu5A α 2-3Gal β 1-4GlcNAc, but it is not clear yet whether the interaction is through the sialyl or galactosyl moiety (Shibuya et al. 1989; Kaku et al. 2007).

Hemagglutinins and lectins are known to have allergenic effects through their interaction with immunoglobulin E (IgE) and histamine release. In a study on 12 lectins isolated from plants, it was shown that IgE bound to many of them and this interaction was inhibited by glycoprotein and lectin-specific sugars. It was further shown that leukocytes isolated from allergic patients on stimulation with the tested lectins showed a significant release of histamine (Shibasaki et al. 1992). Since Jat c1 and JHg isolated by us are similar and hemagglutination activity of JHg is inhibited by sialic acid and complex sugars on fetuin, a glycoprotein, we hope future studies will enhance our understanding of allergenic responses of hemagglutinins from *Jatropha* and other closely related species and their overall biological context.

Acknowledgments The authors thank DST, Govt. of India and Puri Foundation for Education in India for financial support. We thank Professors M. Vijayan and A. Surolia for useful discussions and encouragement and Dr. Suguna and Dr. Alok Sharma from MBU, IISc, for conducting some experiments. The kind support of Dr. C.G. Suresh from NCL Pune in giving critical inputs is very gratefully acknowledged.

References

- Breiteneder H, Clare Mills EN (2005) Plant food allergens—structural and functional aspects of allergenicity. *Biotechnol Adv* 23:395–399
- Breiteneder H, Ebner C (2000) Molecular and biochemical classification of plant-derived food allergens. *J Allergy Clin Immunol* 106:27–36
- Crouch ML, Tenbarger KM, Simon AE, Ferl R (1983) cDNA clones of *Brassica napus* seed storage proteins: evidence from nucleotide sequence analysis that both subunits of napin are cleaved from a precursor polypeptide. *J Mol Appl Genet* 2:273–283
- Ericson ML, Rodin J, Lenman M, Glimellus K, Josefsson LG, Rask L (1986) Structure of the 1.7 S storage protein, napin and its precursor. *J Biol Chem* 261:14576–14581
- Imberty A, Gautier C, Lescar J, Pérez S, Wyns L, Loris R (2000) An unusual carbohydrate binding site revealed by the structures of two *Maackia amurensis* lectins complexed with sialic acid-containing oligosaccharides. *J Biol Chem* 275:17541–17548
- Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with Clustal X. *Trends Biochem Sci* 23:403–405
- Kaku H, Kaneko H, Minamihara N, Iwata K, Jordan ET, Rojo MA, Minami-Ishii N, Minami E, Hisajima S, Shibuya N (2007) Elderberry bark lectins evolved to recognize Neu5Ac alpha2,6-Gal/GalNAc sequence from a Gal/GalNAc binding lectin through the substitution of amino-acid residues critical for the binding to sialic acid. *J Biochem* 142:393–401
- Lehman F, Tiralongo E, Tiralongo J (2006) Sialic acid specific lectins: occurrence, specificity and function. *Cell Mol Life Sci* 63:1331–1354
- Li CY, Devappa RK, Liu JX, Lv JM, Makkar HPS, Becker K (2010) Toxicity of *Jatropha curcas* phorbol esters in mice. *Food Chem Toxicol* 48:620–625
- Maciel FM, Laberty MA, Oliveira ND, Felix SP, Soares AM, Vericimo MA, Machado OL (2009) A new 2S albumin from *Jatropha curcas* L. seeds and assessment of its allergenic properties. *Peptides* 30:2103–2107
- Martin DC, Anna P, Heimo B, Fatima F (2007) Nomenclature and structural biology of allergens. *J Allergy Clin Immunol* 119:414–420
- Moreno FJ, Clemente A (2008) 2S albumin storage proteins: what makes them food allergens? *Open Biochem J* 2:16–28
- Müthing J, Meisen I, Bulau P, Langer M, Witthohn K, Lentzen H, Neumann U, Peter-Katalinić J (2004) Mistletoe lectin I is a sialic acid-specific lectin with strict preference to gangliosides and glycoproteins with terminal Neu5Ac alpha 2–6Gal beta 1–4GlcNAc residues. *Biochemistry* 43:2996–3007
- Nair DN, Suresh CG, Singh DD (2011) Purification, crystallization and preliminary X-ray characterization of a hemagglutinin from the seeds of *Jatropha curcas*. *Acta Crystallogr, Sect F: Struct Biol Cryst Commun* 67:1534–1536
- Oda Y, Matsunga T, Fukuyama K, Miyazaki T, Morimoto T (1997) Tertiary and quaternary structures of 0.19 alpha-amylase inhibitor from wheat kernel determined by X-ray analysis at 2.06 Å resolution. *Biochemistry* 36:13503–13511
- Pantoja-Uceda D, Bruix M, Gimenez-Gallego G, Rico M, Santoro J (2003) Solution structure of Ric C3, a 2S albumin storage protein from *Ricinus communis*. *Biochemistry* 42:13839–13847
- Peters BP, Ebisu S, Goldstein JJ, Flashner M (1979) Interaction of wheat germ agglutinin with sialic acid. *Biochemistry* 18:5505–5511
- Ratanapo S, Ngamjunyaporn W, Chulavatnatol M (1998) Sialic acid binding lectins from leaf of mulberry (*Morus alba*). *Plant Sci* 139:141–148
- Sato S, Hirakawa H, Isobe S, Fukai E, Watanabe A et al (2011) Sequence analysis of the genome of an oil bearing tree *Jatropha curcas* L. *DNA Res* 18:65–76
- Schägger H (2006) Tricine-SDS-PAGE. *Nat Protoc* 1:16–22
- Shibasaki M, Sumazaki R, Isoyama S, Takita H (1992) Interaction of lectins with human IgE: IgE-binding property and histamine-releasing activity of twelve plant lectins. *Int Arch Allergy Appl Immunol* 61:32–39
- Shibuya N, Tazaki K, Song ZW, Tarr GE, Goldstein JJ, Peumans WJ (1989) A comparative study of bark lectins from three elderberry (*Sambucus*) species. *J Biochem* 106:1098–1103
- Stirpe F, Pession-Brizzi A, Lorenzoni E, Strocchi P, Montanaro L, Sperti S (1976) Studies on the proteins from the seeds of *Croton tiglium* and of *Jatropha curcas*. *Biochem J* 156:1–6
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Wright CS (1990) 2.2 Å resolution structure analysis of two refined N-acetylneuraminyl-lactose—wheat germ agglutinin isolectin complexes. *J Mol Biol* 215:635–651
- Yamamoto K, Konami Y, Irimura T (1997) Sialic acid-binding motif of *Maackia amurensis* lectins. *J Biochem* 121:756–761