

OPEN

Structural characterization and *in-silico* analysis of *Momordica charantia* 7S globulin for stability and ACE inhibition

Pooja Kesari^{1,3}, Shivendra Pratap^{1,3}, Poonam Dhankhar^{1,3}, Vikram Dalal¹, Manisha Mishra², Pradyumna Kumar Singh², Harsh Chauhan¹ & Pravindra Kumar^{1*}

Momordica charantia (*Mc*) seeds are widely used edible crop with high nutritional quality. The food and pharmaceutical industries use it as a natural anti-oxigenic agent. Herein, a ~52 kDa protein, which is a major part of seed proteome has been purified, biochemically characterized and structure has been determined. MALDI-ESI-MS identified peptide fragments and contig-deduced sequence suggested the protein to be homologous to 7S globulins. The crystal structure shows that protein has a bicupin fold similar to 7S globulins and the electron density for a copper and acetate ligand were observed in the C-terminal barrel domain. *In silico* study reveals that a tripeptide (VFK) from *Mc*7S possess a higher binding affinity for angiotensin converting enzyme (ACE) than already reported drug Lisinopril (LPR). The protein is a glycoprotein and highly stable under varying thermal and pH conditions due to its secondary structures. The DPPH (2,2-diphenyl-1-picryl-hydrazyl-hydrate) assay showed the protein to have an anti-oxigenic nature and can aid in scavenging free radical from sample. The protein can assist to enhance the nutritional and functional value of food by acting as a food antioxidant. Further, characterization of *Mc*7S required which might add in importance of *Mc*7S as antioxidant, anti-diabetic and anti-hypertensive.

Momordica charantia (*Mc*) commonly known as bitter melon or bitter gourd is a member of Cucurbitaceae family, which grows in tropical and sub-tropical areas. The plant is known for centuries in Ayurveda for its anti-bacterial, anti-fungal, anti-viral, anti-parasitic, hypoglycemic, anti-fertility, anti-tumorous, and anti-carcinogenic properties¹⁻⁵. It is preferred as a medication for a broad range of health applications, including the treatment of dysmenorrhea, eczema, emmenagogue, galactagogue, gout, jaundice, kidney (stone), leprosy, leucorrhoea, piles, pneumonia, psoriasis, rheumatism, scabies, T2DM (type 2 diabetes mellitus), obesity, hypertension, bacterial and viral infections, cancer, and even AIDS¹⁻⁵. It contains biologically active molecules including proteins, triterpenes, saponins, steroids, flavonoids, alkaloids, and acids⁶. Each part of plant i.e. seeds, roots, leaves and especially unripe fruits has its pharmacological properties⁴. The juice is used to cure a large number of conditions such as for articular pain relief and chronic fever in jaundice, liver disease, and digestive system diseases because of laxative, diuretic, and anti-helminthic effects. In the case of chronic skin diseases, it is applied locally to treat boils, burns, and rash. Moreover, components of *Mc* plant i.e. unripe fruit, seeds and aerial parts are largely known for its anti-diabetic properties due to the presence of insulin-mimetic in the seeds^{7,8}. The aqueous seed extract of *Mc* also scavenges the free radicals for protection against lipid peroxidation thereby reducing the risk of diabetic complications⁹. The strong anti-oxigenic activity may be due to the presence of phenolic compounds and saponins¹⁰.

The globular seed fraction of *Mc* majorly contains globulins, which are high-molecular weight proteins, initially thought to be catalytically inactive, however, with the passage of time; new activities have been associated with these proteins. Globulins belong to cupin superfamily and are categorized as 11–12S legumin and 7–8S vicilin based on its sedimentation coefficients. Vicilins are trimeric protein expressed by multiple structural genes and their extensive post-translational processing (glycosylation and protease action) results in a high degree of

¹Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, 247667, India. ²Plant Molecular Biology Division, Council of Scientific and Industrial Research (CSIR)-National Botanical Research Institute, Lucknow, India. ³These authors contributed equally: Pooja Kesari, Shivendra Pratap and Poonam Dhankhar. *email: pravinmcu@gmail.com

polymorphism¹¹. The expression of 7S globulins increases under dehydrating conditions and therefore they are thought to be involved in desiccation and oxidative stress as well¹².

In this study, we have structurally and functionally characterized a 52 kDa protein from the seeds of *Mc*. The MALDI-ESI-MS characterized fragments suggest the protein has homology with 7S globulin. In the absence of 7S globulin sequence, the contig sequences were aligned to deduce the protein sequence. The crystal structure of *Momordica charantia* 7S (*Mc7S*) globulin showed maximum homology to 7S vicilins. The structural analysis led to the identification of a bound copper and acetate ion in the protein. *In silico* study indicates that *Mc7S* release the peptide (VFK) which can bind with angiotensin-converting enzyme (ACE). Further, molecular docking and simulation confirm that binding of VFK to ACE tends to form stable ACE-VFK complex. Protein is highly stable upto 85 °C, so this protein can be used in food industry. Independently, the protein also possesses antioxidant activity and is glycosylated. The protein plays a significant role in food system and human health system by reducing oxidative stress and free radicals.

Results

Purification of *Mc7S* globulin. *Mc7S* globulin was purified to homogeneity by a three-step procedure involving ammonium sulphate precipitation, affinity chromatography and size exclusion chromatography. In ammonium sulphate precipitation step, 7S globulin was fractionated in 60–80% fraction, as observed on SDS-PAGE. This fraction showed two major protein bands of ~52 kDa and ~33 kDa and further, taken for affinity purification using Affi-Gel Blue matrix and the desired protein was eluted with a stepwise gradient of NaCl. The ~52 kDa protein was eluted in-between the 0.2 M–0.5 M NaCl fractions. The eluted fractions were pooled and dialyzed; before applying onto HiLoad 16/60 Superdex 200 pg column. The purity of ~52 kDa protein fraction was confirmed by the presence of a single band on SDS-PAGE gel and the Gel filtration chromatogram showed peak at ~69.2 ml (Fig. S1). The molecular weight of protein was estimated by using HMW calibration curve, which suggests that the protein exists as a trimer of ~156 kDa.

Sequence analysis. To identify the ~52 kDa protein isolated from the seeds, the protein was analysed using MALDI-ESI-MS. Three peptide fragments FALEAR, PHFNSR and LVGFGINAQNLR were identified (Fig. S2). The derived peptide sequences were searched against the EST database and protein sequence database. The literature suggests that the protein could be a seed storage protein; however no hit was obtained against the protein sequence database, therefore the transcriptome data of *Mc* was analysed to identify cDNA population encoding 7S globulin. The previously reported literature on the sequencing of bitter gourd results suggested that 42 contigs of bitter gourd share high sequence homology with cupin 2, PV100 from *Cucurbita maxima* (gi|3808062). These contigs were aligned using DNASTar software package. Each aligned fragment was searched against the nucleotide database and protein database to identify the nearest homologs. The contig-deduced nucleotide sequence was translated to the protein sequence. BLAST result against protein database showed that the derived protein sequence which shares the maximum homology with cupin 2, PV100 (gi|3808062) of cupin superfamily and shares 73% homology with vicilin from *Cucumis sativus* (XP_011650968.1). The alignment of contigs resulted in the identification of 419 amino acid residues having a molecular weight of ~47 kDa and theoretical pI of 7.22. Supplementary Data SD1 shows the alignment of few contig fragments. The composition of 20 amino acids of *Mc7S*, vicilin from *Solanum* and *Capsicum* is given in Table S1. The protein has a large number of charged residues and very few cysteine and tryptophan residues. It is rich in essential amino acids like phenylalanine, valine, threonine, leucine, isoleucine and lysine.

Crystal structure of *Mc7S*. Purified protein was concentrated to about 10 mg/ml and immediately used for crystallization. Diffraction-quality crystals of protein were obtained 0.1 M sodium acetate trihydrate pH 4.6 and 2 M sodium chloride. The crystal structure of protein was determined at a final resolution of 3.1 Å. It belongs to *P63* space group and has a single monomer in an asymmetric unit. The chain A of the crystal structure of vicilin from pecan (*Carya illinoensis*) (PDB ID: 5E1R) which share 52% sequence similarity with *Mc7S* was used as a template for model building. Since the template sequence has good query coverage the initial R_{factor} and R_{free} values were 0.31 and 0.41. The final R_{factor} and R_{free} values after iterative cycles of restrained refinement and model building were 0.19 and 0.27, respectively. The numbering of *Mc7S* residues is done as per canavalin. The Ramachandran plot generated using the MolProbity server for analysis of the main chain stereochemistry showed that 90.7% residues are in the favoured region and 9% residues in allowed region and 0.3% are in outlier of the plot. The final statistics of data collection and refinement parameters have been described in Table 1.

Overall structure of *Mc7S*. The monomer of *Mc7S* comprises of two cupin fold polypeptide chain separated by a pseudo-dyad axis of symmetry as observed in the other 7S family members (Fig. 1A). This pseudo-dyad axis divides monomer into two domains, namely N- and C-terminal domains. The core region of the N- and C-terminal domains are composed of β -sheets whereas the helices form the extended arms. The N-terminal barrel fold (residues 46–248) is composed of 11 β sheets, i.e. β 1 (Ser56–Ser58), β 2 (Gly65–Leu70), β 3 (Arg87–Ala93), β 4 (Thr97–Ile99), β 5 (Ala106–Val112), β 6 (Ala116–Val123), β 7 (Lys128–Ala136), β 8 (Gly139–Ile144), β 9 (Thr149–Val153), β 10 (Gln161–Pro168) and β 11 (Asp177–Leu179) surrounded by two extended α -helices α 2 (Asn194–Leu201), α 3 (Arg207–Val211). Similarly C-terminal barrel fold (residues 265–426) is composed of 12 β sheets β 12 (Leu250–Lys252), β 13 (Glu265–Ala270), β 14 (Ala284–Ile290), β 15 (Gly295–Asn301), β 16 (Thr305–Glu312), β 17 (Ser316–Cys321), β 18 (Glu337–His342), β 19 (Leu348–Val351), β 20 (Leu367–Met360), β 21 (Leu367–Ile374), β 22 (Arg381–Phe383) and β 23 (Phe422–Lys424) surrounded by two extended α -helices α 5 (Gln396–Thr402), α 6 (Lys408–Glu412). The sequence alignment of N- and C-terminal domain of *Mc7S* suggests that both domains share an evolutionary relationship (Fig. S3). There are around 44 and 49 residues on

PDB ID	6KM8
Wavelength (Å)	1.5418
Resolution range (Å)	50–3.1 (3.15–3.1)
Space group	$P6_3$
Cell dimensions	
a, b, c (Å), α, β, γ (°)	90.77, 90.77, 68.48, 90, 90, 120
Total no of Reflections	27725
Unique reflections	5314 (397)
Multiplicity	2.8 (2.1)
Completeness (%)	89.52 (66.44)
Mean $I/\sigma(I)$	11.17 (2.29)
R-sym	0.12 (0.407)
R-factor	0.19
R-free	0.27
RMSD bonds lengths (Å)	0.01
RMSD bond angles (°)	1.50
Number of non-hydrogen atoms	2753
Macromolecules	2741
Number of protein residues	345
Ramachandran favored (%)	90.7
Ramachandran Allowed (%)	9.0
Ramachandran outliers (%)	0.3
Wilson B factor (Å ²)	53.8
Average B-factor (Å ²)	53.0

Table 1. Data collection and refinement statistics of *Mc7S*. Statistics for the highest-resolution shell are shown in parentheses. $R_{\text{merge}} = \frac{\sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl,j}}$.

the interface between the N-terminal and C-terminal domains, respectively which form 18 hydrogen bonds and 8 salt bridges (Table S2).

As per gel filtration profile, the *Mc7S* globulin should exist as a trimer. The trimer was observed when the three monomer molecules were arranged in the form of an equilateral triangle where their extended α -helix regions interact with the other monomer in a head-to-tail fashion (Fig. 1B). For the formation of the trimer, the monomers are arranged in a three-fold axis of symmetry. Similar fold was also observed in reported 7S seed storage proteins, i.e. adzuki bean 7S globulin (PDB ID: 2EAA and 2EA7), jackbean canavalin (PDB ID: 2CAU), french bean phaseolin (PDB ID: 2PHL), 8S mungbean (PDB ID: 2CV6), Korean pine (PDB ID: 4LEJ), pecan (PDB ID: 5E1R), peanut AraH1 (PDB ID: 3SMH), soybean β -conglycinin (PDB ID: 1UIK), *Capsicum annum* Vic_CAPAN (PDB ID: 5YJS) and *Solanum melongena* SM80.1 (PDB ID: 5VF5). Thus, the trimer was generated by applying the cell symmetry in coot, showing three molecules from three asymmetric unit. The hydrophobicity of monomer of globulin is higher as compared to other seed storage protein¹³. However, these surface hydrophobic residues get buried inside during head-to-tail interaction of the tertiary structure. Also, charged residues from the extended helical arm of the monomer are interlocked in the tertiary structure through salt-bridges and hydrogen bonding. The interface of chain A-chain B forms 25 hydrogen bonds and 4 salt bridges whereas the interface of chain C-chain A form 25 hydrogen bonds and 4 salt bridges (Table S3).

In *Mc7S*, the density was observed for 342 residues of the monomer out of 383 residues. Four stretches (I: 125–127 residues, II: 215–220 residues, III: 225–247 residues, and IV: 325–334 residues) were omitted from the final model because no electron density at the main-chain level was available (Fig. 1C).

Sequence and structural comparison with other 7S seed storage globulin. The 7S globulins or vicilins share ~40% sequence similarity over the entire length of protein. The pBLAST search against 7S globulin sequence suggests that *Mc7S* share high similarity with vicilin from pecan (52%), Vic_CAPAN (44%) and SM80.1 (41%). Followed by 7S globulin-1 adzuki bean (36%), β -conglycinin (35%), 8S mungbean storage protein (35%), AraH1 (35%), Korean pine vicilin (33%), phaseolin (33%), and canavalin (33%). The Fig. S4 shows the multiple sequence alignment (MSA) of all these 7S globulins. The MSA showed that the loop I connecting $\beta 6$ – $\beta 7$ is longer in case of *Mc7S* as compared to other crystal structures. The loop region II and III involved in connecting the N-terminal domain to the C-terminal domain are disordered. The length of loop II is small in *Mc7S*, pecan, *Solanum*, french bean and jackbean. The loop II is longest in peanut, adzuki bean and 8S mungbean. The loop III has almost the same length in all 7S globulins. The loop region IV is near the copper-binding site and the length is longer in *Mc7S* followed by pecan, Korean pine and peanut as compared to other reported crystal structures.

Structural similarity search using Dali Lite indicated that *Mc7S* structure shares the highest similarity with SM80.1, Vic_CAPAN and pecan vicilin; followed by AraH1, Korean pine vicilin, β -conglycinin, 7S globulin-3, 8S mungbean. The least homology is observed with phaseolin and canavalin. All these structures have an overall C-alpha RMSD of less than 2 Å; the slight variation occurs around disorder loop III, loop IV and extended

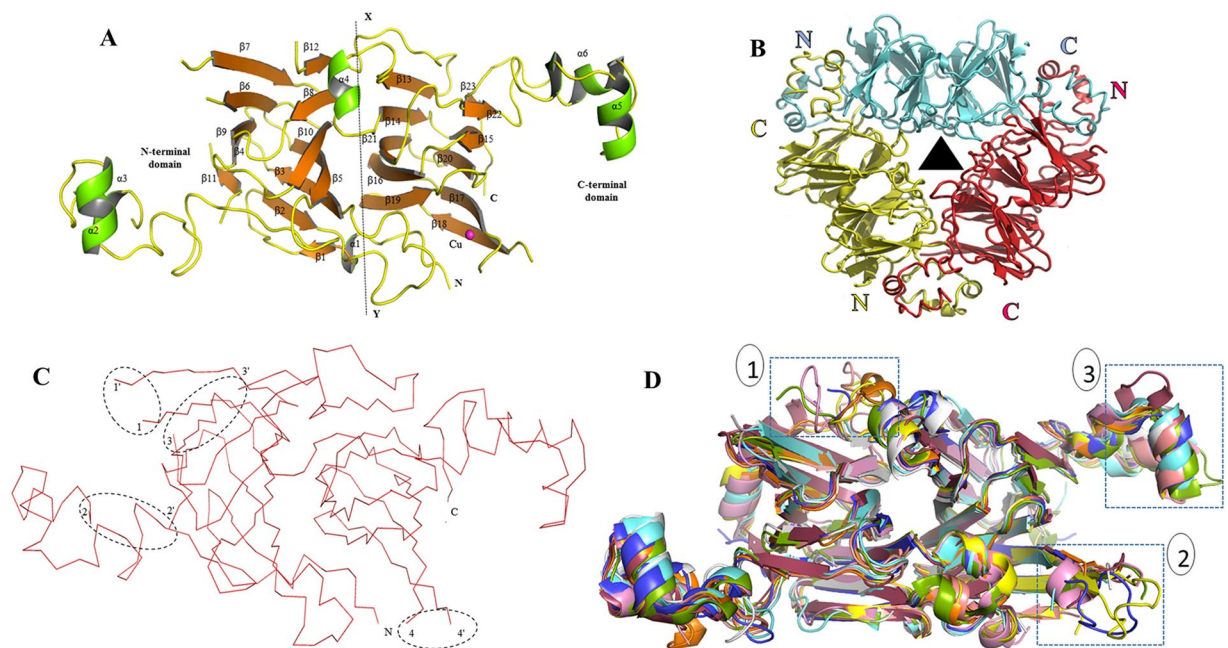


Figure 1. Crystal structure of *Mc7S* globulin. **(A)** The crystallographic asymmetric unit of *Mc7S* globulin contains a monomer which is represented by a cartoon form. In the monomer, the strands, helices, and random coils are shown in orange, green, and yellow, respectively. The pseudo-dyad axis of symmetry (X-Y axis) divides the monomer into N and C-terminal domain. Each domain is comprised of a cupin fold. The N and C-terminal ends have been marked as N and C. The β sheets and α -helices have been labelled as per occurrence. The magenta sphere is the Cu ion. **(B)** The biological *Mc7S* trimer assembly is formed by interaction of monomers in head-to-tail fashion. The 3 chains of *Mc7S* is represented in cartoon form in green, pink and purple, respectively. The hydrophobic surface residues are coloured in red. The triangle in the centre represents the axis perpendicular to the three-fold axis of symmetry. **(C)** Four disordered region of *Mc7S* (I: 125–127 residues, II: 215–220 residues, III: 225–247 residues, and IV: 325–334 residues) have been circled. The ending residues of the disordered loop is marked by apostrophe ('). **(D)** Structural superimposition of vicilin from *Solanum melongena* SM80.1 (5vf5-golden), pecan (cyan), AraH1 (magenta), Korean pine (green), β -conglycinin (orange), phaseolin (brown), 8S mungbean (purple), Aduzki bean (light red), Jackbean (blue), *Capsicum annum* Vic_CAPAN (pink) over *Mc7S* (white). The region of variation is marked by blue dotted box. Box 1 represent the region around disorder loop 3, the box 2 is the copper binding site whereas the box 3 is the C-terminal helical region.

C-terminal helical arm (Fig. 1D). The density for loop II was present in most of the structures except 8S mungbean; while the density of loop III was not observed in most of the 7S globulins. The loop region IV is near the copper-binding site and the density was not observed in any structure. The length of the loop region IV is longer in *Mc7S* as compared to other reported crystal structures. The maximum likelihood (ML) phylogenetic tree classified *Mc7S* with pecan, SM80.1 and Vic_CAPAN, followed by Korean pine vicilin, whereas the second subtree comprised of adzuki bean, canavalin, β -conglycinin, 8S mungbean, phaseolin and AraH1 (Fig. S5).

Metal binding site. The Fo-Fc map showed the clear electron density near the tyrosine (Tyr49) from the N-terminal domain and the cysteine (Cys321), and two histidines (His323 and His355) from the C-terminal domain. The site was conserved in the phylogenetic subtree members pecan, SM80.1, Vic_CAPAN and Korean pine vicilin. The crystal structure of 7S globulins from SM80.1, pecan, Vic_CAPAN and Korean pine showed electron density of copper at the site^{14–18}. The presence of copper ion in *Mc7S* globulin was further confirmed by inductively coupled plasma mass spectroscopy (ICP-MS) (Table S4). The metal possesses a trigonal planar geometry with copper in the centre as inferred from CheckMyMetal server (Fig. 2A). Structural superimposition suggests that copper-binding site is conserved in pecan (Cys652, His654 and His698), SM80.1 (Cys281, His283 and His325), Vic_CAPAN (Cys417, His419 and His464) and Korean pine vicilin (Cis338, His340 and His379) is shown in Fig. 2B. The copper ion is placed slightly away from the centre of the trigonal plane formed by SG of Cys321, ND1 of His323, and the NE2 of His374. The polder omit map (mFo-Fc) contoured at 2.5σ for copper ion and coordinating residues are shown in Fig. S6A.

In globulin-3 from adzuki bean, the electron density was observed for a calcium ion, at a different site. The site was not conserved in *Mc7S* globulin. Although, the site was conserved in members of other legume vicilins like canavalin, β -conglycinin, 8S mungbean and AraH1 from *Arachis hypogaea*; no calcium ion-electron density was observed. Thus, the ML phylogeny tree corresponds well with the metal ion binding tendency of vicilins.

Acetate binding site. The C-terminal β -barrel domain of *Mc7S* has an evident density for acetate moiety (Fig. 2C). The acetate interacts with NE2 atom of His299 and ND2 atom of Asn301 through its O atom. Superimposition

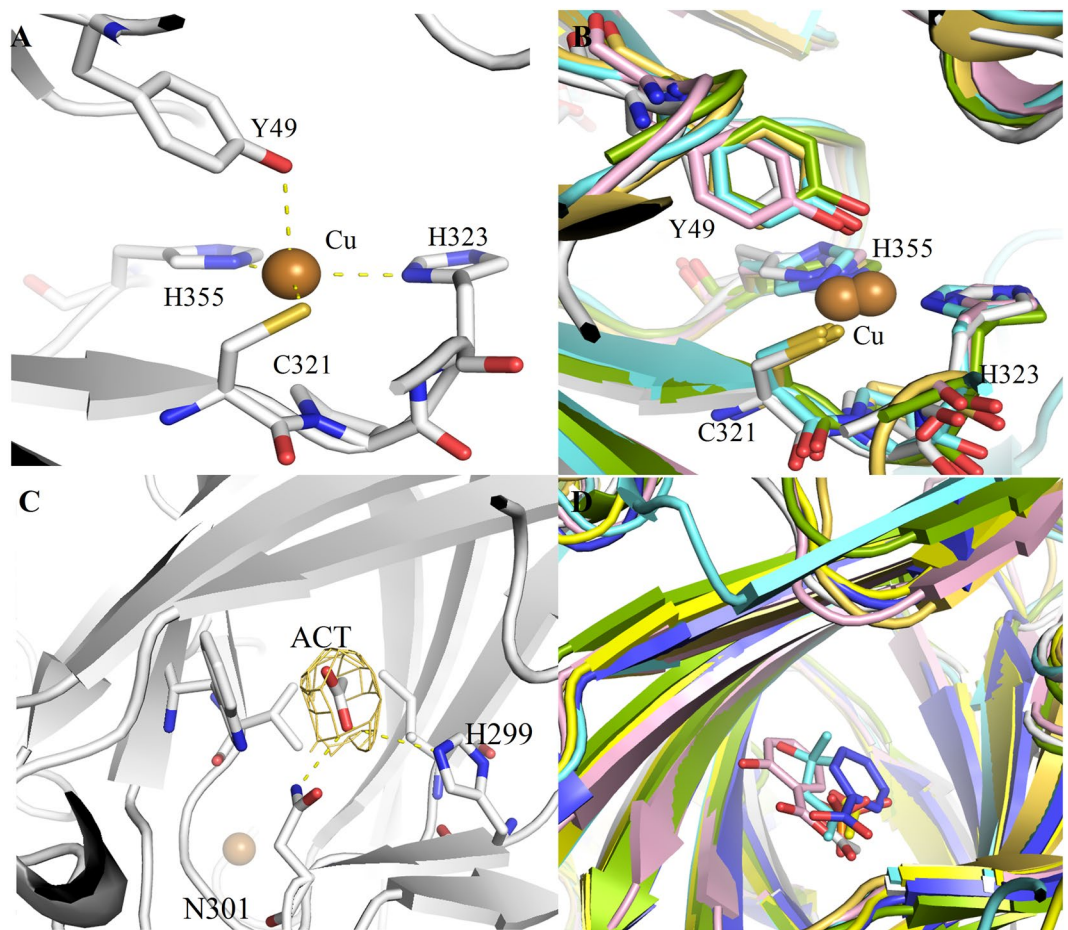


Figure 2. Ligand bound to *Mc7S*. (A) Copper (Cu) metal coordination site of *Mc7S* showing trigonal planar geometry. (B) Structural superimposition of the metal binding site of *Mc7S* (white), Korean pine vicilin (green), pecan (cyan), *Capsicum annum* Vic_CAPAN (pink) and *Solanum melongena* (golden). (C) Interaction of Acetate (ACT) with NE2 atom of His299 and ND2 atom of Asn301 through its O atom. (D) Superimposition of acetate (ACT) binding site of *Mc7S* (white) with *Solanum melongena* (golden), Adzuki bean (yellow), Jackbean (purple), Korean pine (green) and pecan (cyan). Residues are labeled as per *Mc7S*.

with other 7S globulin shows that the presence of acetate (in SM80.1 and adzuki bean), 2-Methyl-2, 4-pentenediol (MPD) (in pecan), phosphate (in Korean pine), salicylic acid (in Vic_CAPAN) and benzoic acid (canavalin) in the β -barrel of C-terminal domain (Fig. 2D). The acetate binding residues are conserved in most of the 7S globulin (Adzuki bean, β -conglycinin, and AraH1) with few exceptions. In SM80.1, the acetate is stabilized through four hydrogen bonds with Tyr259, Asn261, Arg266 and Lys351. In place of residues Tyr259, Arg266 and Lys351 of SM80.1, the *Mc7S* has His299, Arg381 and Trp306. In Vic_CAPAN, the salicylic acid forms interaction similar to acetate of SM80.1. The carboxy end of salicylic acid interacts with Tyr395, Asn397, Arg402 and Lys490. In canavalin, the carboxyl end of benzoic acid also forms similar interaction as acetate. It interacts with His297, Asn299, Arg376 and Asn284. The Asn284 of canavalin is replaced by Ala286 in *Mc7S*. The MPD also interacts in similar manner, with one end oxygen coordinated by conserved His630, Asn632 and Arg724; whereas the other end oxygen is coordinated by Asn617 (Asn284 of canavalin). In Korean pine, a phosphate ion is found at a similar site. The phosphate coordinates well with Glu383 and conserved Arg405. Therefore, there is a possibility that charge within the pocket can accommodate similar ligands.

In-silico ACE inhibitor prediction. BIOPEP server results showed that *Mc7S* globulin contained many ACE inhibiting peptide but these peptides are inactive within the protein¹⁹. The human body having many gastrointestinal enzymes and during enzymatic digestion, the protein was cleaved into many fragments¹⁹. Through *in silico* proteolysis, it was confirmed that 'VFK' (PEP) a tripeptide fragment was released on trypsin digestion which has ACE inhibitory activity. The polder omit map (mFo-Fc) contoured at 2.5σ for VFK residues in *Mc7S* are shown in Fig. S6B. ACE enzyme plays a significant role in blood pressure regulation and its inhibitors are widely used as a drug for the treatment of hypertension, myocardial infarction, heart failure, and diabetic nephropathy^{20,21}. The binding affinity of 'VFK' peptide was analyzed by docking into the active site of human ACE. The crystal structure of ACE complexed with Lisinopril (PDB ID 1O86) at 2.0 \AA was considered for docking of PEP (VFK-peptide) with ACE.

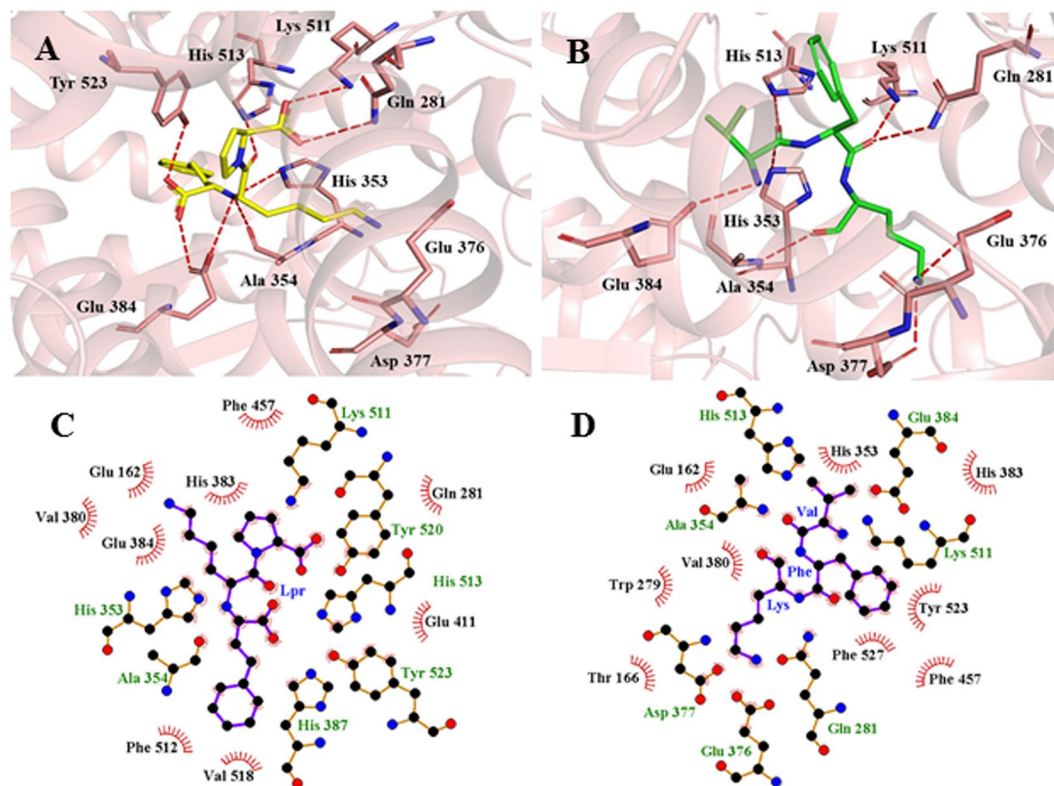


Figure 3. Molecular docking interaction analysis of human angiotensin converting enzyme (ACE) with LPR and PEP. The cartoon representation of ACE with LPR (A: yellow color) and PEP (green color). Interacting residues of ACE are shown in stick format in salmon color while red dotted line shows the intermolecular hydrogen bond interactions. 2D representation of interactions of ACE with LPR (C) and PEP (D) is generated by LIGPLOT.

Molecular docking of LPR and PEP along with ACE was done using HADDOCK webserver. LPR and PEP show hydrogen bonding interactions with His353, Glu384, Lys511, and His513 as shown in Fig. 3. PEP shows a higher binding affinity (-81.2 ± 3.8) as compared to LPR (-73.5 ± 0.5) as shown in Table S5. Density functional theory (DFT), a quantum mechanics approach was used to determine the molecular and electronic properties of LPR and PEP. High and low electron density regions in ligands were calculated by electrostatic potentials as shown in Fig. 4. Highest occupied and lowest unoccupied molecules orbitals (HOMO and LUMO) representation of LPR and PEP are shown in Fig. 4. The higher energy gap of PEP (4.48 eV) as compared to LPR (1.53 eV) indicates that PEP can bind efficiently with ACE as shown in Table S6. Overall molecular docking and quantum mechanics results reveal that PEP can bind more efficiently at the active site of ACE.

Molecular dynamics. Molecular dynamics of ACE-LPR and ACE-PEP complexes were performed using GROMACS 2019.2. Root mean square deviation (RMSD) along C α backbone atoms were calculated to study the dynamics flexibility of the C α atoms of ACE-LPR and ACE-PEP complexes. ACE-LPR and ACE-PEP complexes achieved convergence at 65 ns at 0.26 nm and stable during the molecular dynamics of 100 ns as shown in Fig. 5A. Both complexes exhibit an average RMSD of 0.25 nm as shown in Table S7. Radius of gyration (Rg) results indicate that ACE-PEP complex is as stable as ACE-LPR complex as shown in Fig. 5B. Solvent Accessible Surface Area (SASA) was generated to determine the volume of receptor surrounded by solvent molecules during the molecular simulation. SASA results depict that ACE-PEP complex is as compact as ACE-LPR complex as shown in Fig. 5C and Table S7. g_hbond tool of gromacs was used to generate the intra hydrogen bonds in protein-ligand complexes during the molecular dynamics of 100 ns. ACE-PEP complex showed the comparable number of intra hydrogen bonds to ACE-LPR complex as shown in Fig. 5D and Table S7. The distribution of hydrogen bond lengths signifies that ACE-PEP complex form hydrogen bond from high to low affinity which is comparable to ACE-LPR complex as shown in Fig. 5E. Overall, molecular dynamics results confirm that binding of PEP to ACE results in the formation of stable ACE-PEP complex.

Temperature and pH stability. Scan of far-UV (190–250 nm) CD was used to monitor the secondary structure as well as the conformational stability of Mc7S globulin under increasing temperatures (25 to 85 °C) and pH (3 to 11) range (Fig. S7). Spectra showed a small negative peak around 208 nm, a positive peak between 195 nm and 200 nm and a broad negative peak between 212–220 nm which are the characteristic features of protein having predominant β -sheet with little α -helix. It can be concluded that Mc7S globulin has predominant antiparallel β -sheet and little α -helix which is comparable with 7S seed storage protein (SSP)s such as phaseolin,

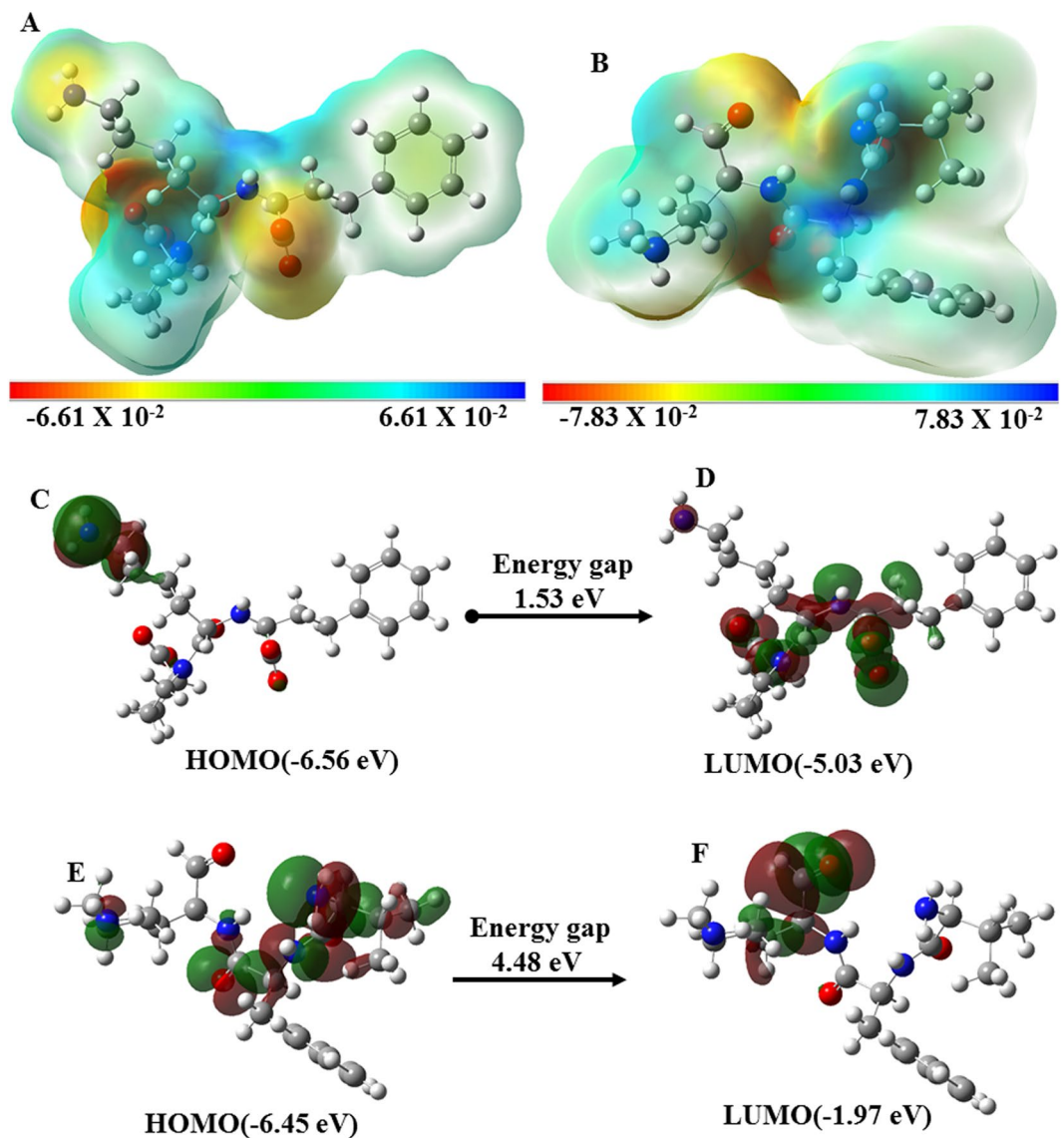


Figure 4. Molecular electrostatic potential (MEP), HOMO and LUMO plots of LPR and PEP generated by DFT/B3LYP method with 6–311G (d,p) basis set in Gaussian 16. MEP plots showing the red (electron rich) and blue (electron poor) regions for: (A) LPR and (B) PEP. HOMO and LUMO plots of LPR (C,D) and PEP (E,F) representing the positive (red) and negative (red) phase distribution in molecular orbital wave function.

canavalin, SM80.1, β -conglycinin^{15,22–24}. On changing temperature and pH, very minute change occurs in the range between 195–200 nm, in-conclusion only minor change in the helical region as compared to sheets region. The homo-trimeric state of 7S globulins provides thermo-stability to the protein. The helices from both terminals of each monomer protrude outward to interact with the neighbouring monomers, resulting in the formation of a trimer. It is observed that the electrostatic and hydrophobic interactions stabilize the quaternary structure.

Antioxidant activity. Antioxidants play a vital role in food processing and storage to reduce the oxidative stress as well as in human health to control the free radicals and oxidative stress-related diseases such as diabetes mellitus, cardiovascular diseases, and neurodegenerative diseases^{25,26}. The DPPH radical-scavenging activity was carried out to find the antioxidant activity of the *Mc7S* protein. DPPH is a synthetic, stable free-radical in aqueous or ethanol and has been widely used to assess the radical scavenging ability of various samples. By accepting an electron or hydrogen radical, DPPH becomes a stable diamagnetic molecule and change in the intensity of purple colour of free DPPH radical was measured at 517 nm. The DPPH scavenging activity increased reliably with the change in concentration of *Mc7S* globulin from 0.0–1.5 mg/ml, while the scavenging activity of control compound Ascorbic acid reached a maximum plateau from 0.5 to 1.0 mg/ml (Fig. S8). Thus, results showed that *Mc7S* has antioxidant activity. Additionally, the basic fold of seed storage 7S globulins has evolved from plant germin, an Mn-binding protein with oxalate oxidase and superoxide dismutase activities²⁷.

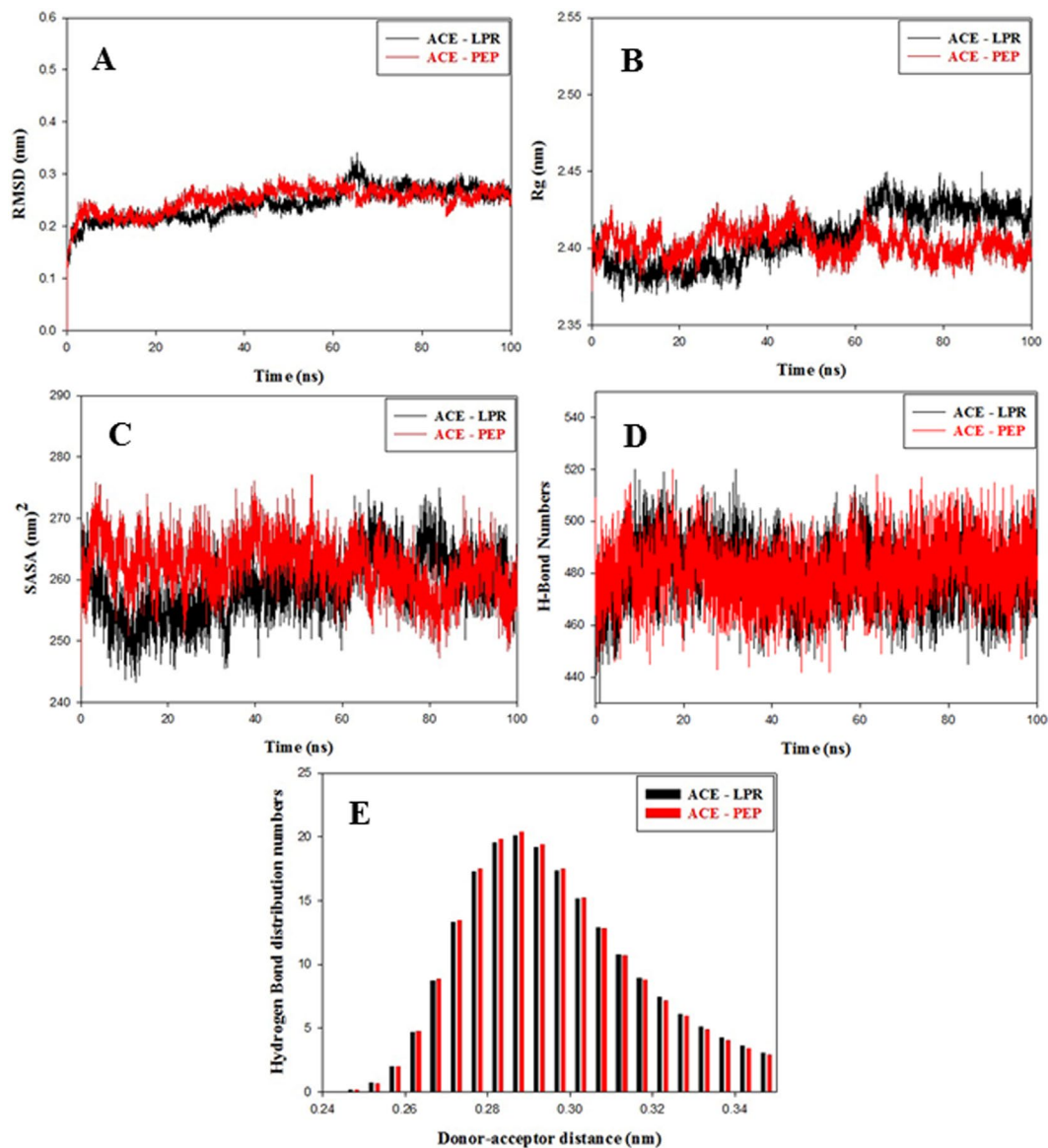


Figure 5. Molecular dynamics results of ACE - LPR (black color) and ACE - PEP (red color) complex during the molecular simulation of 100 ns. (A) Root Mean Square Deviation (RMSD), (B) Radius of gyration (Rg), (C) Solvent Accessible Surface Area (SASA), (D) Intra-hydrogen bond numbers and (E) Intra hydrogen bond distribution numbers.

Glycosylation assay. Glycosylation governs the stability of the protein and provide resistance against denaturation. It plays an important role in protein folding, stability, and interaction as well as in protein trafficking. The glycosylation changes the net charge on the protein surface and largely affects the emulsifying ability and emulsion stability. The 7S globulin pre-pro-proteins are synthesized in the endoplasmic reticulum, and glycosylated in-order to be transported to the vacuoles for storage. Molisch's test, a sensitive chemical test for detection of carbohydrate moiety was utilized to detect the glycosylation property of *Mc7S*. By the addition of *Mc7S* globulin, a purple coloured ring appeared at the junction of two liquids, which indicates that the purified protein has a carbohydrate moiety (Fig. S9). The N-glycosylated in native β -conglycinin and phaseolin is at Asn328 and Asn262, respectively. Sequence and structural superimposition show that Asn residue is absent from both the position in *Mc7S*, which suggests that N-glycosylation site is different from phaseolin and β -conglycinin. The N-glycosylation occurs on Asparagine residue occurring in the Asn-X-Ser/Thr stretch where X can be any amino acid. This stretch was observed at residue's position Asn255-Gln256-Thr257 in *Mc7S*. Therefore, the probable N-glycosylation site in *Mc7S* can be Asn255.

Discussion

The *Mc* plant has several nutritional and nutraceutical values. The SDS-PAGE profile of the seed extract suggested the presence of a ~52 kDa in major amount in *Mc*. The MALDI-ESI-MS identified fragments suggested the protein to be homologous to 7S globulin. The protein sequence was derived by aligning contig sequences deposited in the Short Read Archive (SRA) database. Three-dimensional structure of *Mc*7S globulin revealed that it has predominantly antiparallel β -sheets as compare to alpha-helix, similar to that of other 7S globulins. The trimer formation around the three-fold crystallographic axis of *Mc*7S resembles that of the other 7S globulins. Literature suggests the trimeric arrangement limits the proteolysis and provides thermal stability^{28,29}. Although the sequence homology of Korean pine with *Mc*7S is less, the ML phylogeny tree classified them together. Moreover, a similar copper-binding site was observed in Korean pine along with pecan, SM80.1 and Vic_CAPAN. As there was no copper ion added during the purification or crystallization trials, the presence of copper ion in *Mc*7S globulin was the property of the protein.

To determine the active biopeptide *Mc*7S, *in silico* study was performed and VFK (PEP) a tripeptide fragment was released by trypsin digestion. Molecular docking shows that PEP possesses a higher binding affinity for ACE as compared to LPR. Density functional theory analysis reveals that PEP is more reactive than LPR. Molecular dynamics simulations results confirm that binding of PEP to ACE results in formations of stable ACE-PEP complex. Several molecular docking and simulation studies were used to predict the efficiency of binding of the ligand with macromolecules^{30–35}. The ACE inhibitors such as captopril, lisinopril, enalapril, ramipril and perindopril have shown common side effects such as headache, dizziness, cough, low blood pressure, hypotension, nausea, and angioedema^{36–38}. Since *Momordica charantia* is considered as a novel plant with diverse applications in food industry as well as in the therapy of many diseases such as diabetes and atherosclerosis. Identification of ACE inhibiting peptide fragments from this plant will show better activity with lower side effects³⁹. Variations in secondary structures of *Mc*7S at different pH and temperatures were conducted using Circular dichroism. The *Mc*7S globulin is very stable at different pH and temperature; and probably has a high denaturing temperature similar to 7S globulin from soy protein⁴⁰. Therefore, it is suitable for an application in the food industry where products are needed to maintain under higher temperatures during processing.

The *Mc*7S has radical scavenging ability, as suggested by DPPH assay. The superoxide dismutase radical scavenging ability was observed in Vic_CAPAN, the 7S globulin from *Capsicum annum*¹⁷. The antioxidants play an important role, including defending against oxidative damage in the major signaling pathways of cells in the biological system. Although several commercial available synthetic antioxidants are available, their usage is restricted due to their toxic side effects⁴¹. Thus, efforts need to be made to identify and development of alternative natural antioxidants. The common oxidation products are formed from sulphur-containing (cysteine and methionine), aromatic (tryptophan, tyrosine, and phenylalanine) and imidazole-containing (histidine) residues compared to aliphatic residues. The amino acid composition of these residues in *Mc*7S is similar to SM80.1 and Vic_CAPAN; therefore, they may share functional properties too. Other factors that are responsible for the antioxidant ability of the protein are its amino acid composition, pH, its source, glycosylation, digestion, and absorption⁴². The 7S globulin from *Momordica charantia* is a glycoprotein similar to 7S globulin from *Pisum sativum*⁴³, β -conglycinin from *Glycine max*⁴⁴, and phaseolin from *Phaseolus vulgaris*⁴⁵. The glycosylation increases the stability and solubility of the protein.

In conclusion, the *Mc*7S protein was crystallized and characterized by using computational and *in-vivo* study. Crystal structure of *Mc*7S reveals that it is trimer and consists of antiparallel β -sheets. ICP-MS confirmed the presence of copper ions coordinated with Cys321, His323, and His374 of *Mc*7S. *In silico* analysis indicate the release of tripeptide (VFK) from protein during digestion by digestive enzymes in humans. Further, molecular docking, quantum mechanics and molecular dynamics confirm that PEP binds with ACE with higher binding affinity than LPR and tends to form stable ACE-PEP complex. Already reported ACE inhibitors such as captopril, lisinopril, enalapril, ramipril and perindopril have several side effects, so usage of PEP might result in lesser side effects. Secondary structure content of *Mc*7S sustains upto 85 °C which indicates to use it at high temperatures in food industry. Moreover, *Mc*7S globulin has antioxidant property which plays a significant role in the treatment of different diseases. Glycosylation property of *Ms*7S globulin plays role in protein stability and in antioxidant activity. This study suggesting that *Mc*7S protein and its peptide could be used as a natural source of antioxidant and as a raw material, respectively in food industry and human health system to reduce the oxidative stress and free radicals.

Methods

Purification of *Mc*7S globulin. The seeds of *Mc* were purchased from local market. The seed coats were removed and decorticated seeds were soaked overnight in buffer A (50 mM Tris Buffer pH 7.4). Soaked seeds were flash-frozen with liquid nitrogen and homogenized by mortar and pestle in buffer A. The crude extract was subjected to centrifugation at 12000 rpm, 4 °C for 2 h to obtain a clear supernatant. The supernatant was retrieved after careful removal of lipid layer. The clear supernatant was subjected to ammonium sulphate precipitation step in which different percentage of ammonium sulphate (0–20%, 20–40%, 40–60%, 60–80% and 80–100%) was slowly added while stirring at 4 °C. Eluted fractions were analysed on 12% SDS-PAGE and fractions containing 7S globulin were dialyzed thrice in 1 L of buffer A for 12 hrs at 4 °C. The dialyzed sample was applied onto Affi-Gel Blue matrix (Bio-Rad Laboratories, Hercules, California, USA) pre-equilibrated with buffer A. The column was washed with buffer A and bound proteins were eluted using a gradient of NaCl (0.1, 0.2, 0.3, 0.4, 0.5 and 1.0 M) in buffer A. Eluted fractions were examined on 12% SDS-PAGE and the fractions with 7S globulin were pooled and concentrated to 10 mg/ml using Amicon Ultra-15 (10 000 Da MWCO, Millipore). Further, size exclusion chromatography was performed for final step purification. The concentrated protein fraction was loaded on HiLoad 16/60 Superdex 200 pg column pre-equilibrated with buffer A and 150 mM NaCl, at a flow rate of 0.5 ml/min on the ÄKTA purifier (GE Healthcare). By measuring the absorbance at 280 nm, the protein elution profile

was monitored and analysed on 12% SDS-PAGE under reducing condition. Peak containing 7S globulin eluted at 69.2 ml, which correspond nearly to the predicted molecular weight of 7S globulin trimers. The result was validated by calibrating the column with high molecular weight standards.

Partial internal sequencing and gene sequence deduction from contigs. The purified protein was electrophoresed on 12% SDS-PAGE and the protein bands were excised and partially digested by Trypsin. The partially digested fragments were utilized for Matrix assisted laser desorption ionization electrospray ionization mass spectrometry (MALDI-ESI-MS) followed by reverse phase separation. The data obtained were analysed by the MASCOT search engine. The experiment was carried out at protein sequencing facility at National Botanical Research Institute, Lucknow, India. The obtained peptide fragments were searched in ESTs database, however, no hits were obtained. Therefore, the contig sequence data related to bitter melon (*Momordica charantia*) that has been deposited in the GenBank Short Read Archive (SRA) with the accession number SRP004091 was downloaded. The accession numbers for the individual experiments of normalized sequence data is SRX030203⁴⁶. The composition of cDNA transcripts analysed from this experiment was found to encode cupin 2, PV100 a seed storage protein. The sequence read archive nucleotide BLAST was utilized to identify these contig sequences using cupin 2, PV100 (gi|3808062) as query. DNASTAR Lasergene Molecular Biology Suite was used for assembling contigs. The gene sequence was converted to protein sequence using ExpAsy translate tool (<https://web.expasy.org/translate/>).

Crystallization, data collection, data processing, structure determination and refinement. The purified *Mc7S* globulin was crystallized by the sitting drop vapour-diffusion method in 96-well crystallization plates (Hampton Research) at 293 K. The 10 mg/ml protein concentration was used for setting crystal trays. Initial hits were obtained in crystal screen (Hampton Research, USA). The diffraction quality hexagonal plate-like crystals were obtained in 0.1 M sodium acetate trihydrate pH 4.6 and 2 M sodium chloride within 2–3 days. Crystal was soaked in a cryoprotectant solution containing reservoir solution along with 30% glycerol followed by flash freezing in liquid nitrogen prior to diffraction. Diffraction data were collected using a Bruker Microstar rotating anode X-ray generator (CuK α wavelength = 1.54 Å) and MAR345dtb image plate detector, positioned at 270 mm distance from crystal at 100 K temperature. X-ray exposure of 10 min and 1° oscillation per frame was used for collection of 100 diffraction images. The crystal diffracted to a resolution of 3.1 Å at the home source. HKL2000 was used for indexing, integrating and scaling the data⁴⁷. The structure of *Mc7S* globulin was solved by molecular replacement method using vicilin from pecan (*Carya illinoensis*) (PDB ID: 5E1R) as template⁴⁸. The transformed coordinates were refined with phenix.refine from Phenix⁴⁹. The manual model building was carried out with graphics programs COOT⁵⁰. The models were evaluated using the program Molprobity⁵¹. Metal geometry was checked using CheckMyMetal web server⁵². The figures were generated using the program PyMOL and ESPript program^{53,54}.

Metal analysis using inductively coupled plasma mass spectroscopy (ICP-MS). ICP-MS of the *Mc7S* protein sample was performed to detect the presence of copper metal ion content in the purified protein. The sample was prepared using a previously described method⁵⁵. The protein was concentrated using Amicon Centricon-10 concentrator up to ~10 mg/ml, further, the acid mineralization was done to solubilize the organic content which may otherwise precipitate during experimentation and lead to choking of the tubing system in the instrument. For this, an equal volume of an ultrapure analytical grade of HCl was mixed with ultrapure analytical grade HNO₃ in a ratio 1:1. This mixture was added to an equal volume of protein sample and heated in a dry bath for 2 h at 100 °C. The sample volume was brought to 5 ml using doubly deionized water. Liquid calibration standards (10 ppb, 100 ppb and 1000 ppb) were prepared for constructing a calibration curve. Signal intensities of the unknown samples were compared to the calibration curve and the intensity of the analyte was determined.

In silico ACE inhibitor prediction. The *Mc7S* protein sequence was evaluated for the presence of peptide fragments using BIOPEP server (<http://www.uwm.edu.pl/biochemia/index.php/en/biopep>). Further, Expasy peptide cutter server was utilized to check the digestion site in *Mc7S*. ‘Valine Phenylalanine Lysine-VFK’ a tri-peptide fragment confirmed by BIOPEP server was considered for molecular docking using HADDOCK (High Ambiguity Driven protein-protein DOcking)^{56,57}. Best generated conformations were analysed and visualized in PyMOL⁵³. Interaction figures were made in Pymol and LIGPLOT^{53,58}. Density functional theory (DFT) studies were performed using DFT/B3LYP method with 6–311G (d,p) basis set in Gaussian 16^{59–62}. The details of molecular docking and density functional theory is mentioned in Supplementary information.

Molecular dynamics. Molecular dynamics was done to determine the dynamic stability of ACE-LPR and ACE-PEP complexes²⁰. Crystal structure of *Mc7S* and human angiotensin converting enzyme in complex with Lisinopril (PDB ID:1O86) were considered as a starting point for the molecular simulation. Molecular dynamics was performed using GROMOS 54a7 force field in GROMACS 2019.2 suite on an Ubuntu-based workstation^{63,64}. LPR topology files were produced by PRODRG and partial atomic charges were calculated using DFT/B3LYP method with 6–311G (d,p) basis set in Gaussian 16^{59–62,65}. Long-range electrostatics were measured using Particle Mesh Ewald (PME) and coulomb interactions were determined within a cut off the radius of 12 Å⁶⁶. The final molecular dynamics was run for 100 ns. The molecular dynamics simulation study described in detail in Supplementary information.

Circular dichroism studies. Secondary structure and conformational stability of protein were analysed by using Jasco J-1500 CD spectrometer equipped with a Peltier thermostat cell holder (PTC-510) temperature controller. Far-UV (190–250 nm) CD spectra were recorded at a protein concentration of 0.2 mg/ml in sodium phosphate buffer (20 mM, pH 7.4) in 1 mm quartz cell at 25 °C. The conformational stability of protein was examined

by recording the spectra at different temperature (25–85 °C) and at different pH (3–11) in triplicates, and then averaged the data for analysis. The CD spectra were analyzed and intensities were expressed as mdeg.

DPPH radical-scavenging activity. The percent of DPPH radical scavenging activity was determined according to the Zheng *et al.* with some modification⁶⁷. The reaction mixture contains 1 ml of 0.1 mM DPPH solution in ethanol with 1 ml of protein sample in 50 mM Tris buffer pH 7.4 at different concentrations (62.5–1500 µg/ml) and incubated in the dark for 1 hr at room temperature. Using UV-Vis spectrophotometer, the color shifts were read at 517 nm. Ascorbic acid was used as positive control with different concentrations. The scavenging activity was expressed as follows:

$$\% \text{ scavenging activity} = \frac{Ac - As}{Ac} * 100$$

where *As* was the absorbance of mixture contained protein samples and *Ac* was the absorbance of control (only DPPH solution).

Glycosylation assay. The glycosylated nature of the purified protein was confirmed by Molisch Test⁶⁸. Briefly, 1 ml of protein sample (0.1 mg/ml) in buffer A (50 mM Tris Buffer pH 7.4) was placed in a test tube; 1–2 drops of Molisch's reagent (10% α-naphthol in ethanol) were added and mixed gently. Along the side of the inclined tube, 1–2 ml of concentrated sulphuric acid was slowly added and the same was done with the negative control (only buffer instead of protein) as well as with positive control (glucose solution).

Protein data bank submission. The atomic coordinates and structure factors of *Momordica charantia* 7S globulin has been deposited with accession code 6KM8.

Received: 13 August 2019; Accepted: 31 December 2019;

Published online: 24 January 2020

References

1. Beloin, N. *et al.* Ethnomedicinal uses of Momordicacharantia (Cucurbitaceae) in Togo and relation to its phytochemistry and biological activity. *Journal of Ethnopharmacology* **96**, 49–55 (2005).
2. Grover, J. & Yadav, S. Pharmacological actions and potential uses of Momordica charantia: a review. *Journal of ethnopharmacology* **93**, 123–132 (2004).
3. Ng, T., Chan, W. & Yeung, H. Proteins with abortifacient, ribosome inactivating, immunomodulatory, antitumor and anti-AIDS activities from Cucurbitaceae plants. *General Pharmacology: The Vascular System* **23**, 575–590 (1992).
4. Scartezzini, P. & Speroni, E. Review on some plants of Indian traditional medicine with antioxidant activity. *Journal of ethnopharmacology* **71**, 23–43 (2000).
5. Zafar, R. Momordica charantia-a review. *Hamdard medicus* **34**, 49–61 (1991).
6. Raman, A. & Lau, C. Anti-diabetic properties and phytochemistry of Momordica charantia L.(Cucurbitaceae). *Phytomedicine* **2**, 349–362 (1996).
7. Horax, R., Hettiarachchy, N., Kannan, A. & Chen, P. Protein extraction optimisation, characterisation, and functionalities of protein isolate from bitter melon (Momordica charantia) seed. *Food chemistry* **124**, 545–550 (2011).
8. Ng, T. B., Wong, C. M., Li, W. W. & Yeung, H. W. Insulin-like molecules in Momordica charantia seeds. *Journal of ethnopharmacology* **15**, 107–117 (1986).
9. Sathishsekar, D. & Subramanian, S. Antioxidant properties of Momordica Charantia (bitter gourd) seeds on Streptozotocin induced diabetic rats. *Asia Pacific journal of clinical nutrition* **14**, 153 (2005).
10. Padmashree, A., Sharma, G. K., Semwal, A. D. & Bawa, A. S. Studies on the antioxygenic activity of bitter gourd (Momordica charantia) and its fractions using various *in vitro* models. *Journal of the Science of Food and Agriculture* **91**, 776–782 (2011).
11. Autran, J.-C., Halford, N. G. & Shewry, P. R. In *Plant Nitrogen* 295–341 (Springer, 2001).
12. Wang, W.-Q., Möller, I. M. & Song, S.-Q. Proteomic analysis of embryonic axis of Pisum sativum seeds during germination and identification of proteins associated with loss of desiccation tolerance. *Journal of proteomics* **77**, 68–86 (2012).
13. Horax, R., Hettiarachchy, N., Over, K., Chen, P. & Gbur, E. Extraction, fractionation and characterization of bitter melon seed proteins. *J. Agric. Food Chem.* **58**, 1892–1897 (2010).
14. Jin, T. *et al.* Crystal structure of Korean pine (Pinus koraiensis) 7S seed storage protein with copper ligands. *Journal of agricultural and food chemistry* **62**, 222–228 (2013).
15. Jain, A., Kumar, A. & Salunke, D. M. Crystal structure of the vicilin from Solanum melongena reveals existence of different anionic ligands in structurally similar pockets. *Sci Rep.* **6**, 23600, <https://doi.org/10.1038/srep23600> (2016).
16. Zhang, Y. *et al.* Identification and characterization of a new pecan [Carya illinoensis (Wangenh.) K. Koch allergen, Car i 2. *J. Agric. Food Chem.* **64**, 4146–4151 (2016).
17. Shikhi, M., Nair, D. T. & Salunke, D. M. Structure-guided identification of function: role of Capsicum annum vicilin during oxidative stress. *Biochemical Journal* **475**, 3057–3071 (2018).
18. Wlodawer, A. *et al.* Detect, correct, retract: How to manage incorrect structural models. *The FEBS journal* **285**, 444–466 (2018).
19. Montoya-Rodríguez, A., Gómez-Favela, M. A., Reyes-Moreno, C., Milán-Carrillo, J. & González de Mejía, E. Identification of bioactive peptide sequences from amaranth (Amaranthus hypochondriacus) seed proteins and their potential role in the prevention of chronic diseases. *Comprehensive Reviews in Food Science and Food Safety* **14**, 139–158 (2015).
20. Natesh, R., Schwager, S. L., Sturrock, E. D. & Acharya, K. R. Crystal structure of the human angiotensin-converting enzyme–lisinopril complex. *Nature* **421**, 551 (2003).
21. Sukalo, A. *et al.* Treatment of hypertension with combination of lisinopril/hydrochlorothiazide. *Medical Archives* **70**, 299 (2016).
22. Ko, T. P., Day, J. & McPherson, A. The refined structure of canavalin from jack bean in two crystal forms at 2.1 and 2.0 Å resolution. *Acta Crystallographica Section D: Biological Crystallography* **56**, 411–420 (2000).
23. Lawrence, M. C., Izard, T., Beuchat, M., Blagrove, R. J. & Colman, P. M. Structure of Phaseolin at 2.2 Å Resolution: Implications for a Common Vicilin/Legumin Structure and the Genetic Engineering of Seed Storage Proteins. *Journal of molecular biology* **238**, 748–776 (1994).
24. Maruyama, N. *et al.* Crystal structures of recombinant and native soybean β-conglycinin homotrimers. *The FEBS Journal* **268**, 3595–3604 (2001).
25. Li, Y., Zheng, Y., Zhang, Y., Xu, J. & Gao, G. Antioxidant Activity of Coconut (Cocos nucifera L.) Protein Fractions. *Molecules* **23**, 707 (2018).

26. Castel, V., Andrich, O., Netto, F. M., Santiago, L. G. & Carrara, C. R. Total phenolic content and antioxidant activity of different streams resulting from pilot-plant processes to obtain *Amaranthus mantegazzianus* protein concentrates. *Journal of food engineering* **122**, 62–67 (2014).
27. Woo, E.-J., Dunwell, J. M., Goodenough, P. W., Marvier, A. C. & Pickersgill, R. W. Germin is a manganese containing homohexamer with oxalate oxidase and superoxide dismutase activities. *Nature Structural and Molecular Biology* **7**, 1036 (2000).
28. Deshpande, S. & DAMODARAN, S. Structure-digestibility relationship of legume 7S proteins. *Journal of Food Science* **54**, 108–113 (1989).
29. Maleki, S. J. *et al.* Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation. *The journal of Immunology* **164**, 5844–5849 (2000).
30. Dalal, V. *et al.* Repurposing an Ancient Protein Core Structure: Structural Studies on FmtA, a Novel Esterase of *Staphylococcus aureus*. *Journal of molecular biology* **431**(17), 3107–3123 (2019).
31. Malik, A., Dalal, V., Ankri, S. & Tomar, S. Structural insights into *Entamoeba histolytica* arginase and structure-based identification of novel non-amino acid based inhibitors as potential antiamebic molecules. *The FEBS journal* **286**, 4135–4155 (2019).
32. Singh, N., Dalal, V. & Kumar, P. Structure based mimicking of Phthalic acid esters (PAEs) and inhibition of hACMSD, an important enzyme of the tryptophan kynurenine metabolism pathway. *International journal of biological macromolecules* **108**, 214–224 (2018).
33. Singh, N., Dalal, V., Mahto, J. K. & Kumar, P. Biodegradation of phthalic acid esters (PAEs) and in silico structural characterization of mono-2-ethylhexyl phthalate (MEHP) hydrolase on the basis of close structural homolog. *Journal of hazardous materials* **338**, 11–22 (2017).
34. Singh, N., Dalal, V. & Kumar, P. Molecular docking and simulation analysis for elucidation of toxic effects of dicyclohexyl phthalate (DCHP) in glucocorticoid receptor-mediated adipogenesis. *Molecular Simulation*, 1–13 (2019).
35. Saini, G. *et al.* Molecular docking and dynamic approach to virtual screen inhibitors against Esbp of *Candidatus Liberibacter asiaticus*. *Journal of Molecular Graphics and Modelling* **92**, 329–340 (2019).
36. Sangole, N. V. & Dadkar, V. N. Adverse drug reaction monitoring with angiotensin converting enzyme inhibitors: A prospective, randomized, open-label, comparative study. *Indian journal of pharmacology* **42**, 27 (2010).
37. Cozier, G. E. *et al.* Crystal structures of sampatrilat and sampatrilat-Asp in complex with human ACE—a molecular basis for domain selectivity. *The FEBS journal* **285**, 1477–1490 (2018).
38. Masuyer, G., Schwager, S. L., Sturrock, E. D., Isaac, R. E. & Acharya, K. R. Molecular recognition and regulation of human angiotensin-I converting enzyme (ACE) activity by natural inhibitory peptides. *Scientific reports* **2**, 717 (2012).
39. Yoshime, L. T., de Melo, I. L. P., Sattler, J. A. G., de Carvalho, E. B. T. & Mancini-Filho, J. Bitter melon (*Momordica charantia* L.) seed oil as a naturally rich source of bioactive compounds for nutraceutical purposes. *Nutrire* **41**, 12 (2016).
40. Jiang, J., Xiong, Y. L. & Chen, J. Role of β -Conglycinin and Glycinin Subunits in the pH-shifting-Induced Structural and Physicochemical Changes of Soy Protein Isolate. *Journal of food science* **76** (2011).
41. Meenakshi, S., Umayaparthi, S., Arumugam, M. & Balasubramanian, T. *In vitro* antioxidant properties and FTIR analysis of two seaweeds of Gulf of Mannar. *Asian Pacific Journal of Tropical Biomedicine* **1**, S66–S70 (2011).
42. Shahidi, F. & Zhong, Y. Novel antioxidants in food quality preservation and health promotion. *European Journal of Lipid Science and Technology* **112**, 930–940 (2010).
43. Badenoch-Jones, J., Spencer, D., Higgins, T. J. V. & Millerd, A. The role of glycosylation in storage-protein synthesis in developing pea seeds. *Planta* **153**, 201–209 (1981).
44. Kitamura, K., Okubo, K. & Shibasaki, K. The purification of soybean 11S globulin with ConA-Sepharose 4B and Sepharose 6B. *Agricultural and Biological Chemistry* **38**, 1083–1085 (1974).
45. Carbonaro, M. 7S globulins from *Phaseolus vulgaris* L.: Impact of structural aspects on the nutritional quality. *Bioscience, biotechnology, and biochemistry* **70**, 2620–2626 (2006).
46. Yang, P., Li, X., Shipp, M. J., Shockey, J. M. & Cahoon, E. B. Mining the bitter melon (*Momordica charantia* L.) seed transcriptome by 454 analysis of non-normalized and normalized cDNA populations for conjugated fatty acid metabolism-related genes. *BMC plant biology* **10**, 250 (2010).
47. Otwinowski, Z. & Minor, W. [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods in enzymology* **276**, 307–326 (1997).
48. Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *Journal of applied crystallography* **30**, 1022–1025 (1997).
49. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallographica Section D: Biological Crystallography* **68**, 352–367 (2012).
50. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography* **60**, 2126–2132 (2004).
51. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic acids research* **35**, W375–W383 (2007).
52. Zheng, H. *et al.* Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. *Nature protocols* **9**, 156 (2014).
53. DeLano, W. The PyMOL Molecular Graphics System on World Wide Web, <http://www.pymol.org> (2002).
54. Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics (Oxford, England)* **15**, 305–308 (1999).
55. Figueroa, J. A. L., Stiner, C. A., Radzyukevich, T. L. & Heiny, J. A. Metal ion transport quantified by ICP-MS in intact cells. *Scientific Reports* **6**, 20551 (2016).
56. Kurkcuoglu, Z. *et al.* Performance of HADDOCK and a simple contact-based protein–ligand binding affinity predictor in the D3R Grand Challenge 2. *Journal of computer-aided molecular design* **32**, 175–185 (2018).
57. Van Zundert, G. *et al.* The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *Journal of molecular biology* **428**, 720–725 (2016).
58. Laskowski, R. A. & Swindells, M. B. 2778–2786 (ACS Publications, 2011).
59. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of chemical physics* **98**, 5648–5652 (1993).
60. Frisch, M. *et al.* Gaussian 16, Revision A. 03, Gaussian, Inc., Wallingford CT (2016).
61. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical review B* **37**, 785 (1988).
62. Schlegel, H. B. Optimization of equilibrium geometries and transition structures. *Journal of Computational Chemistry* **3**, 214–218 (1982).
63. Van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *Journal of computational chemistry* **26**, 1701–1718 (2005).
64. van Gunsteren, W. F. *et al.* Biomolecular simulation: the {GROMOS96} manual and user guide. (1996).
65. Schüttelkopf, A. W. & Van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallographica Section D: Biological Crystallography* **60**, 1355–1363 (2004).
66. Abraham, M. J. & Gready, J. E. Optimization of parameters for molecular dynamics simulation using smooth particle-mesh Ewald in GROMACS 4.5. *Journal of computational chemistry* **32**, 2031–2040 (2011).

67. Zheng, Y. *et al.* Fractionation, physicochemical properties, nutritional value, antioxidant activity and ACE inhibition of palm kernel expeller protein. *RSC Advances* **5**, 12613–12623 (2015).
68. Dreywood, R. Qualitative test for carbohydrate material. *Industrial & Engineering Chemistry Analytical Edition* **18**, 499–499 (1946).

Acknowledgements

CD studies and size-exclusion chromatography were performed at Department of Biotechnology and Macromolecular Crystallographic Unit at Institute Instrumentation Centre (IIC), IIT Roorkee, respectively. Author would like to thank partial internal sequencing facility at National Botanical Research Institute, Lucknow, India.

Author contributions

P.K.² guided the project scope; S.P. performed protein purification, crystallization and data collection; P.K.¹ solved and analysed the crystal structure; P.D. performed the purification, biophysical characterization and enzyme assays; V.D. did molecular docking, dynamics and quantum mechanics; M.M. and P.K.S. performed the MALDI experiment; H.C. guide for contig sequence analysis; P.K.¹, P.D. and V.D. generated figures, organized and drafted the manuscript. P.K.² supervised the project and helped in drafting the manuscript. All authors discussed the work and contributed to the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-58138-9>.

Correspondence and requests for materials should be addressed to P.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020