

Article

CRYSTAL STRUCTURE OF COCOSIN, A POTENTIAL FOOD ALLERGEN FROM COCONUT (*Cocos nucifera*)

tengchuan jin, Cheng Wang, Caiying Zhang, Yang Wang, Yu-Wei Chen, Feng Guo,
Andrew Howard, Min-Jie Cao, Tong-Jen Fu, Tara H. McHugh, and Yuzhu Zhang

J. Agric. Food Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jafc.7b02252 • Publication Date (Web): 17 Jul 2017

Downloaded from <http://pubs.acs.org> on July 31, 2017

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



ACS Publications

**CRYSTAL STRUCTURE OF COCOSIN, A POTENTIAL FOOD ALLERGEN FROM
COCONUT (*Cocos nucifera*)**

Tengchuan Jin,^{†¶||} * Cheng Wang,^{†||} Caiying Zhang,[†] Yang Wang,^{¶±} Yu-Wei Chen,^{¶⊥} Feng
Guo,^{¶√} Andrew Howard,[¶] Min-jie Cao,^{||} Tong-Jen Fu,[§] Tara H. McHugh,[‡] Yuzhu Zhang^{¶‡*}

[†] Laboratory of Structural Immunology, CAS Key Laboratory of Innate Immunity and Chronic
Diseases, CAS Center for Excellence in Molecular Cell Sciences, School of Life Sciences and
Medical Center, University of Science and Technology of China, Hefei 230027 China

[¶] Department of Biology, Illinois Institute of Technology, 3101 South Dearborn Street, Chicago,
IL 60616, USA

^{||} College of Food and Biological Engineering, Jimei University, Xiamen, Fujian, 361021, China

[§] U.S. Food and Drug Administration, Division of Food Processing Science and Technology,
6502 South Archer Road, Bedford Park, IL 60501, USA

[‡] Healthy Processed Foods Research Unit, USDA-ARS, Western Regional Research Center, 800
Buchanan Street, Albany, CA 94710, USA

*Correspondence:

Tengchuan Jin

Tel: +86-551-63600720

E-mail: jint@ustc.edu.cn

Yuzhu Zhang

Tel: +1-510-559-5981

23 Fax: +1-510-559-5818

24 E-mail: yuzhu.zhang@ars.usda.gov

25 E-mail addresses of all co-authors:

26 Cheng Wang: kingsar@163.com

27 Caiying Zhang: caiyingz@mail.ustc.edu.cn

28 Yang Wang: wangyxianyanta@sina.com

29 Yu-Wei Chen: ychen112@iit.edu

30 Feng Guo: guofeng@iit.edu

31 Andrew Howard: howard@iit.edu

32 Min-Jie Cao: mjcao@jmu.edu.cn

33 Tong-Jen Fu: tongjen.fu@fda.hhs.gov

34 Tara H. McHugh: tara.mchugh@ars.usda.gov

35 Present address:

36 [‡]Y.W.: Xi'an Innovation College of Yan'an University, China

37 [‡]Y.-W.C.: Merial Ltd., Duluth, GA 30601, USA

38 [√] F. G.: DuPont Industrial Biosciences, Genencor Technology Center, 925 Page Mill Road, Palo
39 Alto, CA 94304, USA

40

41

42

43 Running Title: X-ray structure of cocosin

44

45 Manuscript information: 24 pages, 1 table, 6 figures, and 6775 words.

ABSTRACT

Coconut (*Cocos nucifera*) is an important palm tree. Coconut fruit is widely consumed. The most abundant storage protein in coconut fruit is cocosin (a likely food allergen), which belongs to the 11S globulin family. Cocosin was crystallized near a century ago, but its structure remains unknown. By optimizing crystallization conditions and cryoprotectant solutions, we were able to obtain cocosin crystals that diffracted to 1.85 Å. The cocosin gene was cloned from genomic DNA isolated from dry coconut tissue. The protein sequence deduced from the predicted cocosin coding sequence was used to guide model building and structure refinement. The structure of cocosin was determined, for the first time, and it revealed a typical 11S globulin feature of a double layer doughnut-shaped hexamer.

Keywords:

Cocosin, 11S globulin, food allergen, hexameric structure, coconut

Introduction

Coconut Palm (*Cocos nucifera*) is the only species of the genus *Cocos* from the *Arecaceae* family. Coconut is a very important specialty crop in tropical regions and virtually every part of the palm tree has economic values.¹ The endosperm of coconut fruit is rich in protein and fat. Coconut extract or the coconut milk is an important part of daily diets in many regions of the world. Severe cases of allergy to coconut cream and coconut oil have been reported,² Coconut is

on the U.S. Food and Drug Administration's list of tree nuts and is considered a major food allergen for labeling purposes in U.S. food products.³

In a study of two patients with severe tree nut allergy, serum IgE (the E type immunoglobulin) from both patients reacted with two coconut proteins with 35 and 36.5 kDa in size. The 35 kDa coconut protein was shown to be immunologically similar to soy glycinin (a 11S legumin of seed storage proteins).² In another case report of a 28-year-old man who experienced anaphylaxis following ingestion of coconut ice cream, sera IgE of the patient recognized a 78 kDa protein.⁴ In a study of a patient with anaphylaxis to coconut and oral symptoms to tree nuts, cross-reactivity between coconut and hazelnut proteins was demonstrated. Serum IgE of the patient was reactive to the 35 and 50 kDa proteins in both coconut and hazelnut extracts.⁵ In another coconut allergy case report, a strong IgE binding protein of about 18 kDa and two weaker IgE binding proteins of about 25 and 75 kDa were identified by the serum of a 3-year-old boy who had symptoms of abdominal pain, vomiting, oral allergy syndrome, and edema of the eyelids immediately after oral contact with a coconut sweet.⁶ In a report of two coconut allergy cases, the sera reacted strongly to a protein of approximately 29 kDa in the coconut extract. Follow-up mass spectrum analysis revealed that this protein belongs to a 7S globulin subunit.⁷ In a recent report of a severe coconut allergy case, two IgE binding protein units of 27 kDa and 16 kDa were found, and the authors believed that these proteins were from the 7S globulin family.⁸ In summary, several coconut allergy cases have been reported in the medical literature. Most of the IgE reactive proteins fall into four size groups, i.e., 15-22 kDa, 25-36.5 kDa, 45-50 kDa and 75-78 kDa. Seed storage proteins are known to undergo proteolytic processing. The variations in the observed molecular size could arise from different proteolysis stages of coconut samples, in addition to the possible errors in size estimation. In IgE mediated

food allergies, allergens are molecules that elicit strong host immune responses through their association with the E type immunoglobulins. Studies of the identities, sequences, biochemical properties, and structures of allergens, and their interactions with the host immune system are critical for understanding the cause of anaphylaxis, and for the development of allergy prevention and treatment strategies. In contrast to the wealth of information available on allergens of other tree nuts and peanuts, there are only scarce studies on coconut allergens.

The 11S legumin seed storage proteins are abundant in many plant species and are extremely stable. They supply nutrition for seed germination. It is known that in mature seeds of a number of plant species, the 11S storage proteins exist as hexamers and are post-translationally cleaved by an endopeptidase, leaving a disulfide bond as the only covalent bond between the N-terminal acidic unit of ~ 40 kDa and the C-terminal basic unit of ~ 20 kDa.^{9, 10} The orthologues of the 11S seed storage proteins in many commonly consumed foods have been recognized as food allergens, including the peanut major allergen Ara h 3,¹¹ almond allergen Pru du 6,¹²⁻¹⁴ Pecan allergen Car i 4,¹⁵ hazelnut nut allergen Cor a 9,¹⁶ Brazil nut allergen Ber e 2,¹⁷ cashew allergen Ana o 2,¹⁸ pistachio allergen Pis v 2/5,¹⁹ and walnut allergen Jug n/r 4.^{20, 21}

To date, two allergens from coconut have been identified and studied, including a 7S globulin^{7, 22} and an 11S globulin.²² Members of these protein families are often allergens in peanut and other tree nuts. In a recent proteomics study of coconut pollen allergens, 12 proteins were identified including its 11S globulin.²³ In the reported coconut allergy cases, sera from many patients contained IgE that are reactive to the most abundant 11S legumin storage protein, cocosin, suggesting cocosin as a likely food allergen from coconut.^{2, 22, 23} Cocosin, named near a century ago, is an 11S legumin seed globulin found in the coconut endosperm.²⁴ It has a subunit

with a MW of approximately 54 kDa, which is composed of a 32 kDa acidic domain and a 22 kDa basic domain.²⁵ Cocosin is one of those first crystallized proteins in the history of macromolecular crystallography. There are three recently published papers describing the purification or crystallization of this protein.²⁵⁻²⁷ Because the crystals were too small for an X-ray diffraction study,²⁵ the resolution was not high enough and there was no sequence information available.²⁷ The detailed structure of cocosin has not been reported except that a backbone structure at 2.61 Å resolution was deposited in the Protein Data Bank (PDB code 1XGF). Due to the lack of complete genomic information of coconut, the sequence of cocosin is unavailable, which hindered its structural determination and 3-dimensional epitope allocation.

In this study, we purified cocosin protein from coconut endosperm extracts and obtained high quality crystals, which were used to acquire a high-resolution X-ray diffraction data set. We further cloned the cocosin gene to facilitate structure refinement. Here we report a high-resolution structure of cocosin, achieving the decades-long goal of elucidating the crystallographic structure of this protein.

Materials and Methods

Protein Extraction

Fifty gram of dry coconut endosperm were ground and extracted in 200 ml of extraction buffer (1 M NaCl, 20 mM Tris-HCl, pH 7.9), incubated in a 60 °C water bath for 30 minutes and subjected to centrifugation at 20,000g for 10 min at room temperature. The supernatant was collected, and dialyzed twice against distilled water, each for 12 hours. The precipitated microcrystals was collected by centrifugation and were defatted three times, each with 50 ml hexane for two hours with stirring. The defatted pellet was dissolved in the extraction buffer and

was dialyzed against distilled water for a second time. The final crystalline pellet was collected by centrifugation and was stored at -80°C for later use.

Purification

After the crystalline cocosin pellet was dissolved in extraction buffer, ammonium sulfate powder was added to the sample to a final concentration of 1.5 M. The sample was then filtered through $0.45\ \mu\text{m}$ pore size syringe filters prior to loading to a 10 ml phenyl Sepharose hydrophobic interaction column (GE Healthcare, Piscataway, NJ) pre-equilibrated with buffer A (1.5 M ammonium sulfate, 50 mM Tris-HCl, pH 7.9). The bound proteins were eluted with a 95 ml linear gradient of ammonium sulfate by mixing buffer A with 0 - 100 % buffer B (10 mM Tris-HCl, pH 7.9). Cocosin was eluted as a major peak and the fractions of the peak were pooled and loaded onto a 320 ml Superdex 200 column (GE Healthcare, Piscataway, NJ) pre-equilibrated with buffer C (200 mM NaCl, 20 mM Tris-HCl, pH 7.9). The protein was then eluted with buffer C. The gel filtration column was calibrated with thyroglobulin (669 kDa), ferritin (440 kDa), catalase (242 kDa), aldolase (158 kDa) and serum albumin (67 kDa) as molecular weight standards (GE Healthcare, Piscataway, NJ).

SDS-PAGE analysis

SDS-PAGE was performed with a standard protocol,²⁸ except that due to the low solubility of cocosin at the low salt sample buffer, cocosin samples were mixed with an equal volume of 8M urea first, then were mixed 3:1 with a 4X reducing sample buffer and boiled for 5 minutes.

Crystallization

FPLC purified cocosin was further purified by three rounds of crystallization. That is, cocosin protein was concentrated in 2 M NaCl to a concentration of greater than 100 mg/ml and was dialyzed against pure water at 4°C. Numerous small crystals readily formed. The crystals were collected by centrifugation at 4°C. The final purified sample was redissolved in 0.7 M NaCl solution and was further concentrated to about 30 mg/ml with Ultracel-30k filter devices (Millipore, Bedford, MA) at room temperature. Two strategies were adopted in the crystallization screen. To take advantage of the inverse salting-in effect, one μ l of the concentrated protein solution was sealed against one ml of different NaCl solutions at concentrations between zero and 2 M, using the hanging drop vapor diffusion method. The second strategy combined traditional crystallization methods with inverse salting-in effects. The drops were setup by mixing one μ l of the protein solution and one μ l of crystallization solutions from a Crystal Screen™ kit (Hampton Research, Aliso Viejo, CA). Instead of equilibrating the drops against the corresponding crystallization solutions as commonly practiced in typical crystallization screen, the drops were hung over one ml of water or lower ionic strength NaCl solutions. Commonly used additives including sucrose, glycerol, PEG400 and ethylene glycol were tested as cryoprotectants during flash cooling of cocosin crystals in liquid nitrogen.

X-ray diffraction and structure determination

X-ray data collection was performed using a MAR-225 CCD detector at the SER-CAT 22BM beamline at the Advanced Photon Source (APS), Argonne National Laboratories (Lemont, IL). The diffraction data were processed using the HKL2000 suite of programs²⁹ and XDS.³⁰ A structural model was derived by molecular replacement calculations using a protomer of the peanut allergen Ara h 3 (pdb 3C3V)³¹ as a search model with the program PHASER.^{32, 33}

The structure obtained was refined alternately in Coot,³⁴ refmac³⁵ in CCP4 package suite and phenix.refine in the Phenix package.³⁶ TLS parameters were generated by the TLSMD server³⁷ and Phenix,³⁶ and applied throughout the refinement. The crystal structures were validated by Molprobity server³⁸ and RCSB ADIT validation server.³⁹ Solvent accessible surface area was calculated with areaimol from the CCP4 suite.⁴⁰ Electrostatics surfaces were calculated by program APBS⁴¹ with PDB2PQR⁴² using AMBER force field and displayed with Pymol (Delano Scientific LLC, San Carlos, CA).

Cloning of cocosin

To facilitate structure refinement, we sought to clone the cocosin gene from the same coconut powder that we used to purify the cocosin protein. To isolate genomic DNA, ~200 mg desiccated coconut powder was grounded in liquid nitrogen and homogenized in one ml of QG buffer (QIAGEN, Valencia, CA). After filtration by filter paper, 200 µl of isopropanol was added to the filtrate and loaded onto a Qiagen minispin column. After washing with PE buffer, the genomic DNAs were eluted from the column with 100 µl of EB buffer. Hot-start Phusion DNA polymerase (NEB, Ipswich, MA) was used to amplify the cocosin gene. A pair of forward (aagcagcagccttcagcgtctctc) and reverse (aggaaacatccatgtcctagctcataggcgg) oligo primers were designed based on available and the most genetically related African oil palm (*Elaeis guineensis*) 11S glutelin mRNA sequence (NCBI accession code: AF193433).⁴³ Blunt end PCR product was directly ligated into a pBlue-script II vector (Agilent Technologies, Santa Clara, CA) digested with *EcoR* V. Positive clones were screened on X-gal plates, white colonies were sequenced in both directions. Putative introns were identified after sequence alignment with the *Elaeis guineensis* glutelin mRNA, and the peptide sequence deduced from the putative ORF was

matched to the partially refined structure. The DNA sequence of the cocosin gene was deposited into NCBI with accession number KY242371.

Phylogenetic analysis and structural comparisons

Translated cocosin protein sequence was aligned with 11S globulins from other species by *MUSCLE* with manual adjustments of gaps. A bootstrap consensus tree was built with sequence alignment data in MEGA7⁴⁴ with the maximum likelihood (ML) methods. Structure based sequence alignment of selected 11S globulins was done by ESPrnt 3.0.⁴⁵ The cocosin structure was aligned with selected 11S globulins from various sources including *Amaranthus hypochondriacus* (pdb: 3QAC),⁴⁶ pumpkin (pdb: 2EVX), almond nut (pdb: 3FZ3),⁴⁷ Rapeseed (pdb: 3KGL),⁴⁸ peanut (pdb: 3C3V),³¹ soy (pdb: 1FXZ).⁴⁹ Protomer structures of 11S globulins were superimposed in Coot, and the RMSDs of aligned atoms were calculated.

Protein Data Bank Accession Codes

The coordinates and structural factors of cocosin have been deposited in the Protein Data Bank with accession code 5WPW.

Results and Discussion

Purification of cocosin

We found that the solubility of cocosin as a function of the salt concentration of the buffer was very similar to that of other seed 11S globulins, such as edestin (the 11S globulin from hemp seeds), amandin (the 11S globulin from almond seeds, also known as prunin-1 and Pru du 6) and excelsin (the 11S globulin from Brazil-nut, also known as Ber e 2).^{10, 17} The solubility of cocosin

increases dramatically with the increase in salt concentration. Cocosin was extracted and purified by changing the ionic strength of buffers as described in the method section. Hydrophobic interaction chromatography (Figure 1A) and size exclusion chromatography (Figure 1B) were used to further purify cocosin. In the size exclusion chromatography step of the purification, cocosin was eluted at 166 ml (Figure 1B) as a single peak. The molecular mass of the eluted protein was estimated to be ~345 kDa based on the calibration of the column with molecular weight standards, indicating that cocosin exists as a monodispersed hexamer in solution. In addition, cocosin migrated at a nearly identical position on the gel filtration column as did peanut major allergen Ara h 3 (Figure 1C). When analyzed with reducing SDS-PAGE (Figure 1D, lane R), cocosin separated into two major bands with sizes of 32 kDa and 22 kDa, respectively, similar to the reported results for peanut allergen Ara h 3.⁵⁰ These data agreed with previous reports that cocosin underwent post-translational proteolytic modification during maturation.^{24, 27, 51}

Optimization of cocosin crystallization conditions

Previously reported cocosin crystals grew under conditions where PEGs or MPD were used as primary precipitants.²⁷ In a crystallization screen carried out in our laboratory at room temperature, a number of crystals appeared in the drops mixed with PEGs or MPD hanging over water or the kit solutions containing PEGs or MPD within a week. For those grew in the drops with PEGs that were equilibrated against the kit solutions, heavy precipitation was observed in the drops, while no significant precipitation was observed in identical drops equilibrated against water. Under the latter condition, water molecules were absorbed into the drop, which gradually decreased the concentration of PEGs. Due to the dilution effect and a large number of

nucleations, crystals did not grow to usable size. On the other hand, by utilizing inverse salting-in principle, which has been successfully used in the crystallization of several 11S seed globulins, including Ber e 2 (excelsin) from Brazil nut,¹⁷ edestin from hamper seed, and cucumbin from cucumber, we were able to grow sizable cocosin crystals by equilibrating sample drops against water without the use of any precipitant in the drops.

The ionic strength or the vapor pressure of the reservoir determines the volume expansion rate of the drop, which affects the nucleation numbers. In optimizing the crystallization conditions, a range of salt concentration from 0 to 2 M NaCl in the reservoir was systematically screened in the second round. Large crystals were obtained when the salt concentration was less than 1.2 M, but most of the crystals in the drop appeared cracked after growing for several days (Figure 2A). Furthermore, large crystals grew under the above conditions without any cryoprotectant were severely cracked after soaking into the cryoprotectants tested, including sucrose, glycerol, PEG400 and ethylene glycol. Another round of screening was designed to include the above cryoprotectants in the initial protein drops, and to equilibrate the drops against different concentrations of salt solution. In this screening, sucrose was found to significantly retard the nucleation and crystal growth. Small crystals were able to form in one week, but they never grew to usable size under the protein concentration tested. Proteins precipitated heavily upon mixing with PEG400 or ethylene glycol, which led to the formation of many small crystals. With the presence of 30% glycerol in the initial drops, perfect looking crystals were obtained when the drops were equilibrated against 0.5 - 1 M NaCl solutions. Best crystals were grown to the size of 0.5×0.5×0.5 mm within two weeks with 0.7 M NaCl in the reservoir (Figure 2B). Single crystals were picked by nylon loops and quickly soaked into a 25% glycerol cryoprotectant solution, and flash cooled in liquid nitrogen for data collection.

275

276 *Molecular cloning of cocosin*

277 There was no protein sequence available for any region of cocosin at the time we obtained
278 crystal diffraction data, which hindered its structural refinement. The 11S seed storage protein
279 (known as glutelin) of the African oil palm (*Elaeis guineensis*), which also belongs to the
280 *Arecaceae* family under the order of *Arecales* was available at NCBI database at the start of the
281 structure refinement. Anticipating that the protein sequence of cocosin and its coding sequence
282 would have a high percentage of identity with those of African oil palm glutelin, we designed
283 DNA oligos based on the African oil palm (*Elaeis guineensis*) 11S glutelin mRNA sequence
284 (AF193433)⁴³ and used them as primers in PCR experiments with coconut genomic DNA as
285 template. After testing different combinations of primer pairs, the gene sequence of cocosin was
286 successfully cloned. As shown in Figure 3A, a ~1600 base-pair PCR product was cloned and
287 sequenced. Guided by the mRNA sequence from African oil palm 11S glutelin, three regions of
288 putative introns were identified. The ORF search using the program ORF Finder at NCBI also
289 identified the same open read frame with 1401 base pair from which the 466 amino-acid
290 sequence of cocosin was deduced. The resulting protein product has a predicted molecular
291 weight of 52.66 kDa. The protein sequence was used in the subsequent model building and
292 refinement of cocosin structure. Blast search showed that cocosin shared 67% identity with the
293 African oil palm 11S glutelin (XP_010935904) across the full-length protein. The sequence
294 identities to other common tree nut 11S allergens such as Jun r 4, Cor a 9, Pis v 2, and Pru du 6
295 were 51%, 48%, 48% and 46%, respectively.

296 Phylogenetic analysis of the 11S storage proteins from selected trees showed that palm trees
297 including coconut and oil palm form a clade with another monocot magnolia, while setting them

apart from other dicot trees (Figure 3B). The evolutionary relationships based on the 11S globulins are in agreement with a large-scale genome wide phylogenetic analysis of angiosperms,⁵² suggesting that the 11S globulin genes are conserved genes.

Data collection and structure determination

During data collection, noticeable radiation damage of the cocosin crystals was observed, but we were able to collect a complete 1.85 Å data set. Data processing using the HKL2000 suite of programs²⁹ and XDS revealed a trigonal space group H3 with unit cell parameters $a = b = 92.2$ Å and $c = 212.9$ Å (Table 1). The data was scaled to 1.85 Å with a redundancy of 3.4. Assuming two cocosin molecules in an asymmetric unit and an average partial specific volume of $0.74 \text{ cm}^3 \text{ g}^{-1}$ for proteins, the Matthew's coefficient was $1.60 \text{ Å}^3/\text{Da}$, with a corresponding solvent content of the crystal $\sim 26.2 \%$.

The structure was readily solved by a molecular replacement calculation using the peanut major allergen Ara h 3 as a search model. Guided by the cocosin sequence deduced from its coding DNA sequence, most of the residues were assigned with confidence. The refinement was terminated when there were no significant changes in Rwork and Rfree and inspection of the difference density map suggested that no further corrections or addition were justified.

There are two copies of cocosin molecules in each asymmetric unit. The final model includes ~ 400 protein residues in each chain. Molecule A consists of residues 42-129, 143-206, 216-266, and 282-458 and molecule B consists of residues 44-129, 143-206, 217-267, and 282-458. Three regions of poor density are located at the surface exposed loops that have a high degree of conformational freedom. These regions correspond to those commonly found to be disordered in other published 11S globulins.^{46, 48} In addition, 214 water (HOH) molecules are

included in the final model. All of the data to 1.85 Å resolution are used in the final refinement, with $R_{\text{work}} = 22.9\%$ and $R_{\text{free}} = 29.2\%$. Molprobity validation showed 97.2% of all protein residues are in favorite regions and there are no outliers in the Ramachandran plot. There are no C-beta outliers. The average B-factor is 38.6 Å² for protein, and 50.0 Å² for solvent atoms. The geometry of the model is acceptable (Table 1) with rmsd bond of 0.008 Å and rmsd angle of 0.94°.

Overall structure of cocosin

Together with the 7S vicilins and convicilins, 11S seed storage proteins belong to the bi-cupin protein family. Each cupin module is composed of a conserved barrel domain. A protomer of cocosin has two jellyroll β-barrels in the center and two extended helical domains composed of three helices (Figure 4A). Two pairs of intramolecular disulfide bonds are observed in each protomer. SS1 formed between cys45 and cys78, and SS2 formed between cys121 and cys288. A close inspection showed that the SS2 is formed between the acidic and basic subunits (Figure 4B), the reduction of which is responsible for the separation of the 32 kDa and 22 kDa bands on the SDS-PAGE gel.

As for all of the 11S globulins, cocosin is a homo-hexamer formed by dimer of trimers. Trimers are formed through the swapping of their helical regions in a head-to-tail orientation. Such arrangement is commonly seen in 7S vicilin-type bi-cupins (Figure 4C). The trimer-trimer interface is stabilized by six pairs of parallel beta-strands (Figure 4D).

Electrostatic nature of packing interface

The cocosin hexamer surface is highly charged (Figures 6 A-C). Careful inspection of the cocosin crystal-packing interface identified its highly electrostatic nature. Notably, cocosin hexamers form a two dimensional layer of molecules at a-b plane through only one type of interface formed by two protomers (Figure 6D). The interface is formed between a long-edge with a short-edge of cocosin hexagon as shown in Figures 6B-C. The high symmetry nature of this molecular packing interface leads to a high tendency to expand crystal lattice under suitable crystallization conditions. The ionic interactions can also explain the salt sensitive nature of the protein's solubility and crystallizability.

Structure comparison with other 11S globulins

Structural alignment of cocosin to some available 11S globulins showed that they are highly conserved. The rmsd of aligned atoms falls into the range of 0.66 - 0.74 Å (Figure 5A). The two jellyroll β -barrels and two extended helical domains are nearly superimposable, while the variations exist at the terminals of broken loops due to their conformational flexibility. Furthermore, the hexameric structures among the 11S globulins are also almost identical. The structure of cocosin can be superimposed perfectly onto that of peanut allergen Ara h 3 (Figure 5B). Whether the structural conservation of cocosin to other allergenic 11S globulins leads to a similarity in allergenicity among these proteins needs further investigation.

Many 11S globulins have been characterized as food allergens. Coconut is also an allergenic food source. The characterization of the 11S storage protein in coconut is critical for the understanding of coconut induced allergy. Allergy epitope search with the Allergen Database for Food Safety (ADFS) (<http://allergen.nihs.go.jp/ADFS>) identified at least three hotspots of

potential allergenic epitopes which were labeled in Figure 5C. These predicted epitopes share high sequence similarity with reported epitopes in 11S globulin-type tree nut allergens Car i 4 (pecan allergen), Jug r 4 (walnut allergen), Ana o 2 (cashew allergen), and Cor a 9 (hazelnut allergen). Interestingly, these three predicted epitope regions are all part of the core β -barrels structures composed of β -strands and connecting loops. Judging from crystal structure, these regions are likely to be inaccessible for direct IgE binding on intact protein (data not shown), suggesting that partially digested or unfolded cocosin protein might be more allergenic compared with intact protein. The phenomenon of inaccessibility of the IgE antibody binding epitopes in correctly folded food allergen has been reported previously^{53, 54}.

In addition, cocosin is one of the early proteins to be purified by crystallization. It is well documented that cocosin readily crystallized during dialysis against low salt solutions because of the reverse salt-in effect. As a result, cocosin is used in many training courses for next generation crystallographers. However, its complete structure is unavailable until now. In this study, we were able to obtain high quality crystals that diffracted to around 1.85 Å, and with degenerative primers, we successfully cloned the cocosin gene from dry coconut tissue. With the available gene sequence, the cocosin structure was finally completed. Built on the result of continuous efforts from several generations of crystallographers, we achieved a significant milestone in determining the structure of cocosin, and demonstrated that cocosin shares the di-cupin structural features commonly observed in other allergenic 11S globulins.

Notes

The authors declare no competing financial interest. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

Acknowledgements

X-ray diffraction data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-BM beamline at the Advanced Photon Source (APS), Argonne National Laboratory.

Funding

Use of the APS was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38. This work was partially supported by a fund from the Illinois Institute of Technology and by Cooperative Agreement 5U01FD003801 between the U.S. Food and Drug Administration and Institute for Food Safety and Health, Illinois Institute of Technology. T.J. is supported by the Fundamental Research Funds for the Central Universities and the 100 Talents Program of CAS.

Table and Figures

Table 1. X-ray data collection and refinement table.

Cocosin	
Data Collection	
Space group	H3

Unit cell (a, b, c) (Å)	92.4, 92.4, 212.9
(α,β,γ) (°)	90, 90, 120
Resolution (Å) (Last shell)	50-1.85 (1.96-1.85)
No of reflections (total/unique)	199624/55347
Redundancy (last shell)	3.4 (3.4) *
Completeness (%)	95.1 (92.7) *
I/ σ (I) (last shell)	12.2 (2.1) *
R _{meas} (last shell) (%) [¶]	6.5 (78.4) *
Refinement	
Resolution(Å)	50-1.85
Number of protein atoms	5967
No. of solvent/hetero-atoms	214
Rmsd bond lengths (Å)	0.008
Rmsd bond angles (°)	0.94
R _{work} (%) [†]	22.9
R _{free} (%) [‡]	29.2
Ramachandran plot	
favored/disallowed**	97.2/0.0
PDB code	5WPW

409

410 * Asterisked numbers correspond to the last resolution shell.

411 [¶] $R_{\text{meas}} = \sum_h (n/n-1)^{1/2} \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h) \rangle$ are the i th and mean

412 measurement of the intensity of reflection h .

$R_{\text{work}} = \sum_h |F_{\text{obs}}(h) - F_{\text{calc}}(h)| / \sum_h |F_{\text{obs}}(h)|$, where $F_{\text{obs}}(h)$ and $F_{\text{calc}}(h)$ are the observed and calculated structure factors, respectively. No I/σ cutoff was applied.

R_{free} is the R value obtained for a test set of reflections consisting of a randomly selected 5% subset of the data set excluded from refinement.

**Values from Molprobiy server (<http://molprobiy.biochem.duke.edu/>).

Figure 1. Purification of cocosin. A. Hydrophobic interaction chromatography purification of cocosin. The protein was eluted as a major peak from the column with a 2 M to 0 M ammonium sulfate linear gradient. B. XK26/60 superdex-200 gel filtration of cocosin. C. Cocosin has a similar molecular size as that of Ara h 3. The superdex-200 column elution profiles of cocosin and Ara h 3 are superimposed. D. SDS-PAGE analysis of purified cocosin. Cocosin is cleaved into two major peptides linked by disulfide bonds. Lane R, reduced.

Figure 2. Crystallization of cocosin. A. Cocosin crystals obtained in the initial screen. Cocosin protein without glycerol was sealed against 1 M NaCl. Most of crystals were not of diffraction quality. B. Single crystals of cocosin after optimization. One μl of protein sample in 30% glycerol was sealed against 0.7 M NaCl. Crystals reached the size of 0.4 – 0.6 mm in each dimension in two weeks at 16 °C.

Figure 3. Genomic cloning of cocosin. A. The gene sequence of cocosin. The identified intron regions were shaded in grey. The deduced protein sequence is shown underneath the coding sequence. The ATG start codon is highlighted in green, and the TGA stop codon is highlighted in red. The two primers used to amplify the genomic sequence are underlined. B. Molecular

435 phylogeny of 11S globulins from common food producing trees. Sequences are aligned by
436 *MUSCLE*, and the phylogenetic tree was built in MEGA7 with a ML method.

437 **Figure 4. Overall structure of cocosin.** A. Structure of cocosin protomer. Cocosin protomer is
438 shown in ribbon, with its helices colored in red and strands colored in yellow. Two inter-subunit
439 disulfide bonds are colored in gold. B. Electron density map of SS2 formed between C121 and
440 C288. 2Fo-Fc electron density map was calculated at 2σ . C. Hexameric cocosin. Cocosin
441 hexamer is generated by crystal symmetry and six chains are colored differently.

442 **Figure 5. Structural comparison with other 11S globulins.** Available 11S globulin structures
443 are selected for analysis. A. Structural superposition of cocosin (5WPW) to available 11S
444 globulin protomers. The rmsds between aligned atoms in 5WPW (green) with the following
445 structures are 0.76Å (3QAC, cyan), 0.76 Å (2EVX, magenta), 0.66 Å (3FZ3, wheat), 94 Å
446 (3KGL, slate), 0.67 Å (3C3V, orange) and 0.74 Å (1FXZ, deepteal), respectively. B.
447 Superposition of cocosin and peanut allergen Ara h 3 hexamers. Cocosin is colored in green and
448 Ara h 3 is colored in orange. C. Structural based alignment of 11S globulins. The secondary
449 structural elements of cocosin are shown on the top of aligned sequences. Identical regions are
450 highlighted in red, and highly conserved regions are highlighted in box. Three predicted epitope
451 regions were designated by blue lines on the top of the cocosin sequence.

452 **Figure 6. Surface charge and packing interface of Cocosin.** A. Surface charge of cocosin
453 hexamer in the same orientation as in Fig 5B. B. Surface charge of the long-edge of cocosin
454 hexamer, same orientation as in Fig 4D. C. Surface charge of the short-edge of cocosin hexamer.
455 D. Cartoon representation of cocosin hexamer packing mode in the a-b plane. Each hexagon
456 represents one cocosin hexamer with three long-edges and three short-edges. The molecular
457 packing in the a-b plane is mediated by a long-edge to short-edge interaction.

458

459 **References**

- 460 (1) Foale, M., *The coconut odyssey: the bounteous possibilities of the tree of life*. Melbourne,
461 2003.
- 462 (2) Teuber, S. S.; Peterson, W. R., Systemic allergic reaction to coconut (*Cocos nucifera*) in 2
463 subjects with hypersensitivity to tree nut and demonstration of cross-reactivity to legumin-like
464 seed storage proteins: new coconut and walnut food allergens. *J. Allergy Clin. Immunol.* **1999**,
465 *103*, 1180-1185.
- 466 (3) Polk, B. I.; Dinakarbandian, D.; Nanda, M.; Barnes, C.; Dinakar, C., Association of tree nut
467 and coconut sensitization. *Ann. Allergy Asthma Immunol.* **2016**, *117*, 412-416.
- 468 (4) Rosado, A.; Fernandez-Rivas, M.; Gonzalez-Mancebo, E.; Leon, F.; Campos, C.; Tejedor,
469 M. A., Anaphylaxis to coconut. *Allergy*. **2002**, *57*, 182-183.
- 470 (5) Nguyen, S. A.; More, D. R.; Whisman, B. A.; Hagan, L. L., Cross-reactivity between
471 coconut and hazelnut proteins in a patient with coconut anaphylaxis. *Ann. Allergy Asthma*
472 *Immunol.* **2004**, *92*, 281-284.
- 473 (6) Tella, R.; Gaig, P.; Lombardero, M.; Paniagua, M. J.; Garcia-Ortega, P.; Richart, C., A case
474 of coconut allergy. *Allergy*. **2003**, *58*, 825-826.
- 475 (7) Benito, C.; Gonzalez-Mancebo, E.; de Durana, M. D.; Tolon, R. M.; Fernandez-Rivas, M.,
476 Identification of a 7S globulin as a novel coconut allergen. *Ann. Allergy Asthma Immunol.* **2007**,
477 *98*, 580-584.
- 478 (8) Michavila Gomez, A.; Amat Bou, M.; Gonzalez Cortes, M. V.; Segura Navas, L.; Moreno
479 Palanques, M. A.; Bartolome, B., Coconut anaphylaxis: Case report and review. *Allergol*
480 *Immunopathol. (Madr)* **2015**, *43*, 219-220.
- 481 (9) Badley, R. A.; Atkinson, D.; Hauser, H.; Oldani, D.; Green, J. P.; Stubb, J. M., The
482 structure, physical and chemical properties of the soy bean protein glycinin. *Biochim. Biophys.*
483 *Acta.* **1975**, *412*, 214-228.
- 484 (10) Albillos, S. M.; Jin, T.; Howard, A.; Zhang, Y.; Kothary, M. H.; Fu, T. J., Purification,
485 crystallization and preliminary X-ray characterization of prunin-1, a major component of the
486 almond (*Prunus dulcis*) allergen amandin. *J. Agric. Food Chem.* **2008**, *56*, 5352-5358.
- 487 (11) Mueller, G. A.; Maleki, S. J.; Pedersen, L. C., The molecular basis of peanut allergy. *Curr.*
488 *Allergy Asthma Rep.* **2014**, *14*, 429.
- 489 (12) Costa, J.; Mafra, I.; Carrapatoso, I.; Oliveira, M. B., Almond allergens: molecular
490 characterization, detection, and clinical relevance. *J. Agric. Food Chem.* **2012**, *60*, 1337-1349.
- 491 (13) Willison, L. N.; Tripathi, P.; Sharma, G.; Teuber, S. S.; Sathe, S. K.; Roux, K. H., Cloning,
492 expression and patient IgE reactivity of recombinant Pru du 6, an 11S globulin from almond. *Int.*
493 *Arch. Allergy Immunol.* **2011**, *156*, 267-281.
- 494 (14) Willison, L. N.; Zhang, Q.; Su, M.; Teuber, S. S.; Sathe, S. K.; Roux, K. H.,
495 Conformational epitope mapping of Pru du 6, a major allergen from almond nut. *Mol. Immunol.*
496 **2013**, *55*, 253-263.
- 497 (15) Sharma, G. M.; Irsigler, A.; Dhanarajan, P.; Ayuso, R.; Bardina, L.; Sampson, H. A.;
498 Roux, K. H.; Sathe, S. K., Cloning and characterization of an 11S legumin, Car i 4, a major
499 allergen in pecan. *J. Agric. Food Chem.* **2011**, *59*, 9542-9552.
- 500 (16) Blanc, F.; Bernard, H.; Ah-Leung, S.; Przybylski-Nicaise, L.; Skov, P. S.; Purohit, A.; de
501 Blay, F.; Ballmer-Weber, B.; Fritsche, P.; Rivas, M. F.; Reig, I.; Sinaniotis, A.; Vassilopoulou,

- E.; Hoffmann-Sommergruber, K.; Vieths, S.; Rigby, N.; Mills, C.; Adel-Patient, K., Further studies on the biological activity of hazelnut allergens. *Clin. Transl. Allergy*. **2015**, *5*, 26.
- (17) Guo, F.; Jin, T.; Howard, A.; Zhang, Y. Z., Purification, crystallization and initial crystallographic characterization of brazil-nut allergen Ber e 2. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2007**, *63*, 976-979.
- (18) Robotham, J. M.; Xia, L.; Willison, L. N.; Teuber, S. S.; Sathe, S. K.; Roux, K. H., Characterization of a cashew allergen, 11S globulin (Ana o 2), conformational epitope. *Mol. Immunol.* **2010**, *47*, 1830-1838.
- (19) Ahn, K.; Bardina, L.; Grishina, G.; Beyer, K.; Sampson, H. A., Identification of two pistachio allergens, Pis v 1 and Pis v 2, belonging to the 2S albumin and 11S globulin family. *Clin. Exp. Allergy* **2009**, *39*, 926-934.
- (20) Zhang, Y. Z.; Du, W. X.; Fan, Y.; Yi, J.; Lyu, S. C.; Nadeau, K. C.; Thomas, A. L.; McHugh, T., Purification and Characterization of a Black Walnut (*Juglans nigra*) Allergen, Jug n 4. *J. Agric. Food Chem.* **2017**, *65*, 454-462.
- (21) Wallowitz, M.; Peterson, W. R.; Uratsu, S.; Comstock, S. S.; Dandekar, A. M.; Teuber, S. S., Jug r 4, a legumin group food allergen from walnut (*Juglans regia* Cv. Chandler). *J. Agric. Food Chem.* **2006**, *54*, 8369-8375.
- (22) Manso, L.; Pastor, C.; Perez-Gordo, M.; Cases, B.; Sastre, J.; Cuesta-Herranz, J., Cross-reactivity between coconut and lentil related to a 7S globulin and an 11S globulin. *Allergy*. **2010**, *65*, 1487-1488.
- (23) Saha, B.; Sircar, G.; Pandey, N.; Gupta Bhattacharya, S., Mining Novel Allergens from Coconut Pollen Employing Manual De Novo Sequencing and Homology-Driven Proteomics. *J. Proteome. Res.* **2015**, *14*, 4823-4833.
- (24) Sjogren, B.; Spychalski, R., The molecular weight of cocosin. *J. Am. Chem. Soc.* **1930**, *52*, 4400-4404.
- (25) H.J. Carr, G. W. P., M. L. Parker and N. Lambert, Characterisation and Crystallisation of an 11S Seed Storage Globulin from Coconut (*Cocos nucifera*). *Food Chem.* **1990**, *38*, 11-20.
- (26) Garcia, R. N.; Arocena, R. V.; Laurena, A. C.; Tecson-Mendoza, E. M., 11S and 7S globulins of coconut (*Cocos nucifera* L.): purification and characterization. *J. Agric. Food Chem.* **2005**, *53*, 1734-1739.
- (27) Balasundaresan, D.; Sugadev, R.; Ponnuswamy, M. N., Purification and crystallization of coconut globulin cocosin from *Cocos nucifera*. *Biochim. Biophys. Acta.* **2002**, *1601*, 121-122.
- (28) Jin, T.; Albillos, S. M.; Chen, Y. W.; Kothary, M. H.; Fu, T. J.; Zhang, Y. Z., Purification and characterization of the 7S vicilin from Korean pine (*Pinus koraiensis*). *J. Agric. Food Chem.* **2008**, *56*, 8159-8165.
- (29) Otwinowski, Z.; Minor, W., Processing of X-ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymology*, Carter, C. W.; Sweet, R. M., Eds. Academic Press: New York, 1997; Vol. 276, pp 307-326.
- (30) Kabsch, W., Xds. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125-132.
- (31) Jin, T.; Guo, F.; Chen, Y. W.; Howard, A.; Zhang, Y. Z., Crystal structure of Ara h 3, a major allergen in peanut. *Mol. Immunol.* **2009**, *46*, 1796-1804.
- (32) McCoy, A. J.; Grosse-Kunstleve, R. W.; Storoni, L. C.; Read, R. J., Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallogr.* **2005**, *61*, 458-464.
- (33) Storoni, L. C.; McCoy, A. J.; Read, R. J., Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 432-438.

- (34) Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126-2132.
- (35) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **1997**, *53*, 240-255.
- (36) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213-221.
- (37) Painter, J.; Merritt, E. A., Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 439-450.
- (38) Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 12-21.
- (39) Yang, H.; Guranovic, V.; Dutta, S.; Feng, Z.; Berman, H. M.; Westbrook, J. D., Automated and accurate deposition of structures solved by X-ray diffraction to the Protein Data Bank. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 1833-1839.
- (40) Lee, B.; Richards, F. M., The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **1971**, *55*, 379-400.
- (41) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A., Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10037-10041.
- (42) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A., PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids. Res.* **2004**, *32*, W665-667.
- (43) Cha, T. S.; Habib Shah, F., Kernel-specific cDNA clones encoding three different isoforms of seed storage protein glutelin from oil palm *Elaeis guineensis*. *Plant Sci.* **2001**, *160*, 913-923.
- (44) Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S., MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731-2739.
- (45) Robert, X.; Gouet, P., Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, W320-324.
- (46) Tandang-Silvas, M. R.; Cabanos, C. S.; Carrasco Pena, L. D.; De la Rosa, A. P.; Osuna-Castro, J. A.; Utsumi, S.; Mikami, B.; Maruyama, N., Crystal structure of a major seed storage protein, 11S proglutinin, from *Amaranthus hypochondriacus*: insight into its physico-chemical properties. *Food Chem.* **2012**, *135*, 819-826.
- (47) Jin, T.; Albillos, S. M.; Guo, F.; Howard, A.; Fu, T. J.; Kothary, M. H.; Zhang, Y. Z., Crystal structure of prunin-1, a major component of the almond (*Prunus dulcis*) allergen amandin. *J. Agric. Food Chem.* **2009**, *57*, 8643-8651.
- (48) Tandang-Silvas, M. R.; Fukuda, T.; Fukuda, C.; Prak, K.; Cabanos, C.; Kimura, A.; Itoh, T.; Mikami, B.; Utsumi, S.; Maruyama, N., Conservation and divergence on plant seed 11S globulins based on crystal structures. *Biochim. Biophys. Acta.* **2010**, *1804*, 1432-1442.
- (49) Adachi, M.; Takenaka, Y.; Gidamis, A. B.; Mikami, B.; Utsumi, S., Crystal structure of soybean proglycinin A1aB1b homotrimer. *J. Mol. Biol.* **2001**, *305*, 291-305.

- (50) Koppelman, S. J.; Knol, E. F.; Vlooswijk, R. A.; Wensing, M.; Knulst, A. C.; Hefle, S. L.; Gruppen, H.; Piersma, S., Peanut allergen Ara h 3: isolation from peanuts and biochemical characterization. *Allergy*. **2003**, *58*, 1144-1151.
- (51) Carr, H. J.; Plumb, G. W.; Parker, M. L.; Lambert, N., Characterisation and Crystallisation of an 11S Seed Storage Globulin from Coconut (*Cocos nucifera*). *Food Chem.* **1990**, *38*, 11-20.
- (52) Bremer, B.; Bremer, K.; Chase, M. W.; Fay, M. F.; Reveal, J. L.; Soltis, D. E.; Soltis, P. S.; Stevens, P. F.; Anderberg, A. A.; Moore, M. J.; Olmstead, R. G.; Rudall, P. J.; Sytsma, K. J.; Tank, D. C.; Wurdack, K.; Xiang, J. Q. Y.; Zmarzty, S.; Grp, A. P., An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* **2009**, *161*, 105-121.
- (53) Pomes, A., Relevant B cell epitopes in allergic disease. *Int. Arch. Allergy Immunol.* **2010**, *152*, 1-11.
- (54) Zhou, Y.; Wang, J. S.; Yang, X. J.; Lin, D. H.; Gao, Y. F.; Su, Y. J.; Yang, S.; Zhang, Y. J.; Zheng, J. J., Peanut Allergy, Allergen Composition, and Methods of Reducing Allergenicity: A Review. *Int. J. Food Sci.* **2013**, *2013*, 909140.

Figure 1. purification cocosin

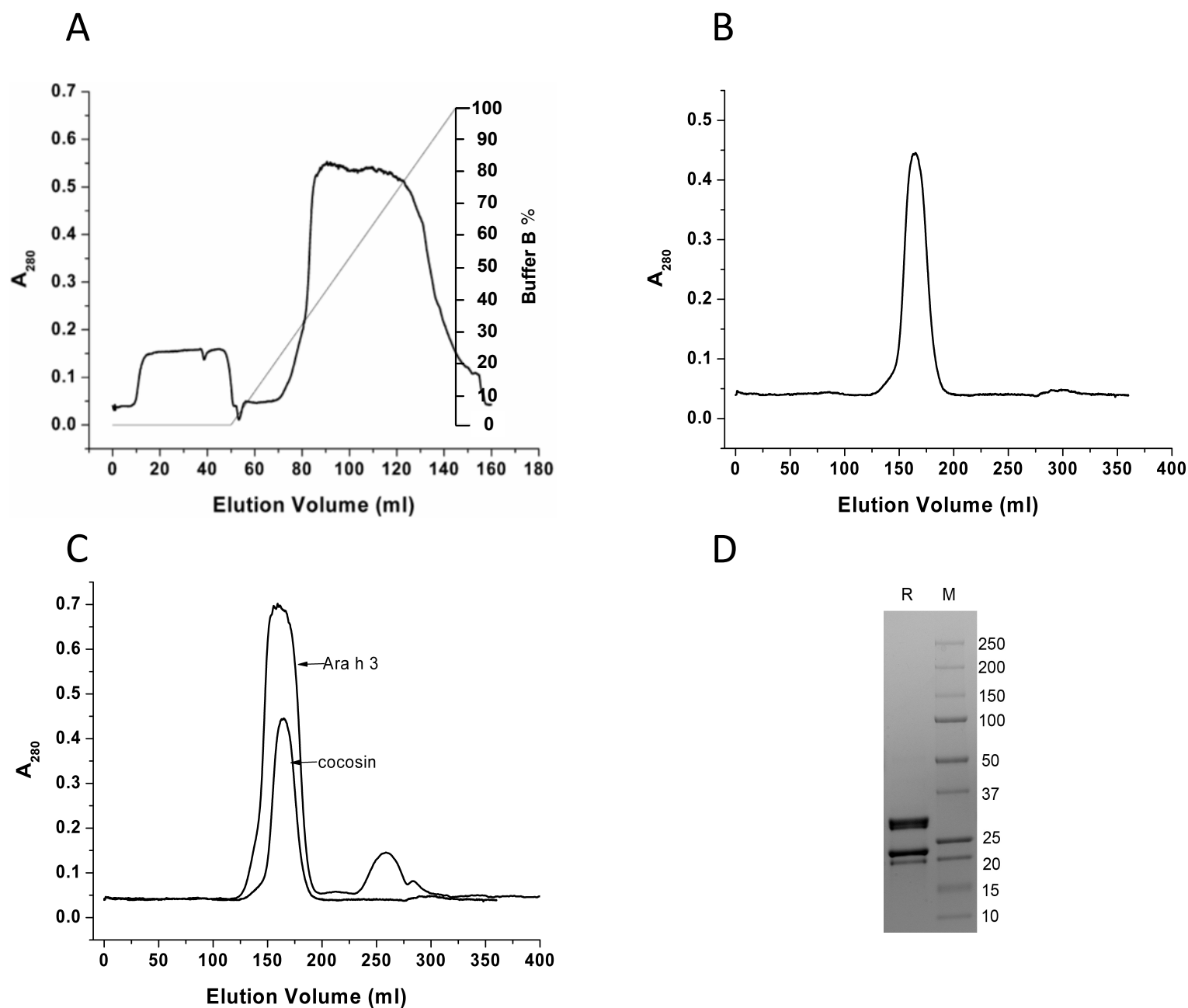
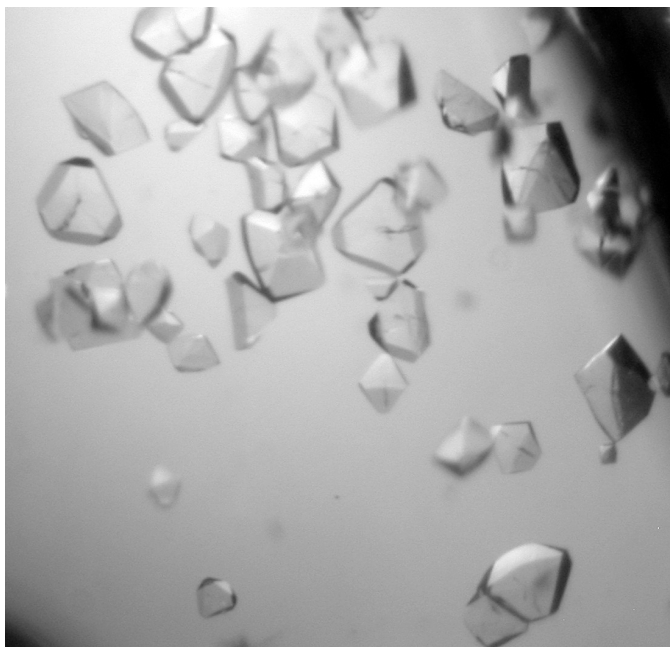


Fig1

Figure 2. crystallization and structural determination of cocosin

A



B

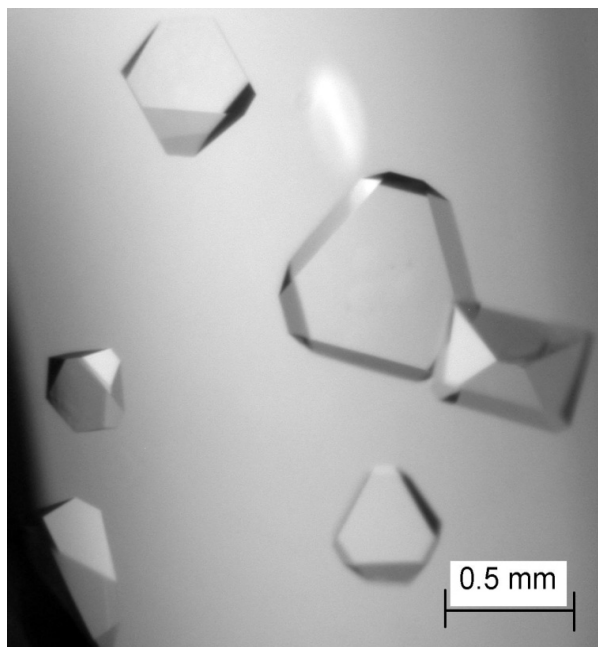


Fig2

Figure 4 Crystal structure cocosin

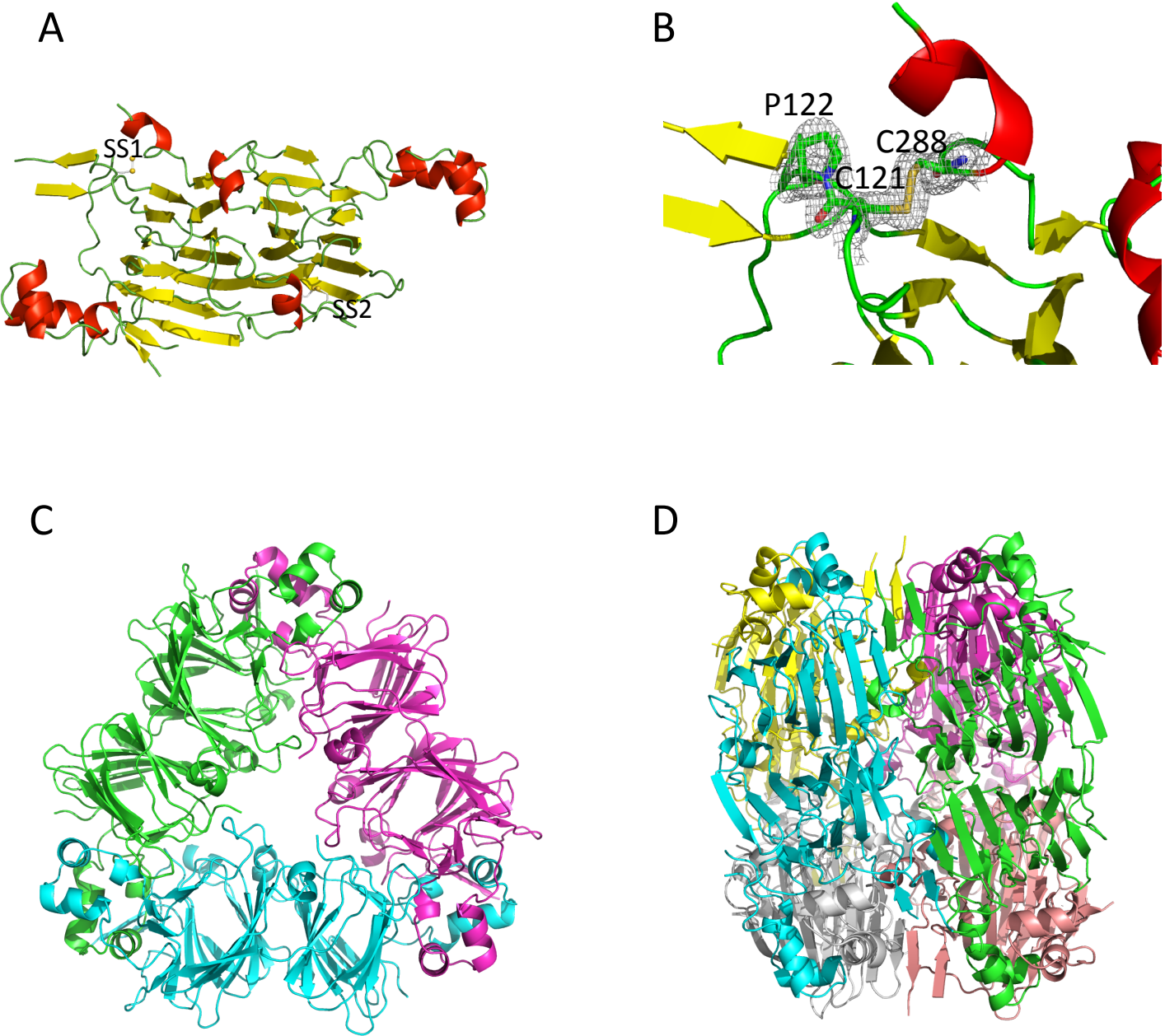


Fig4

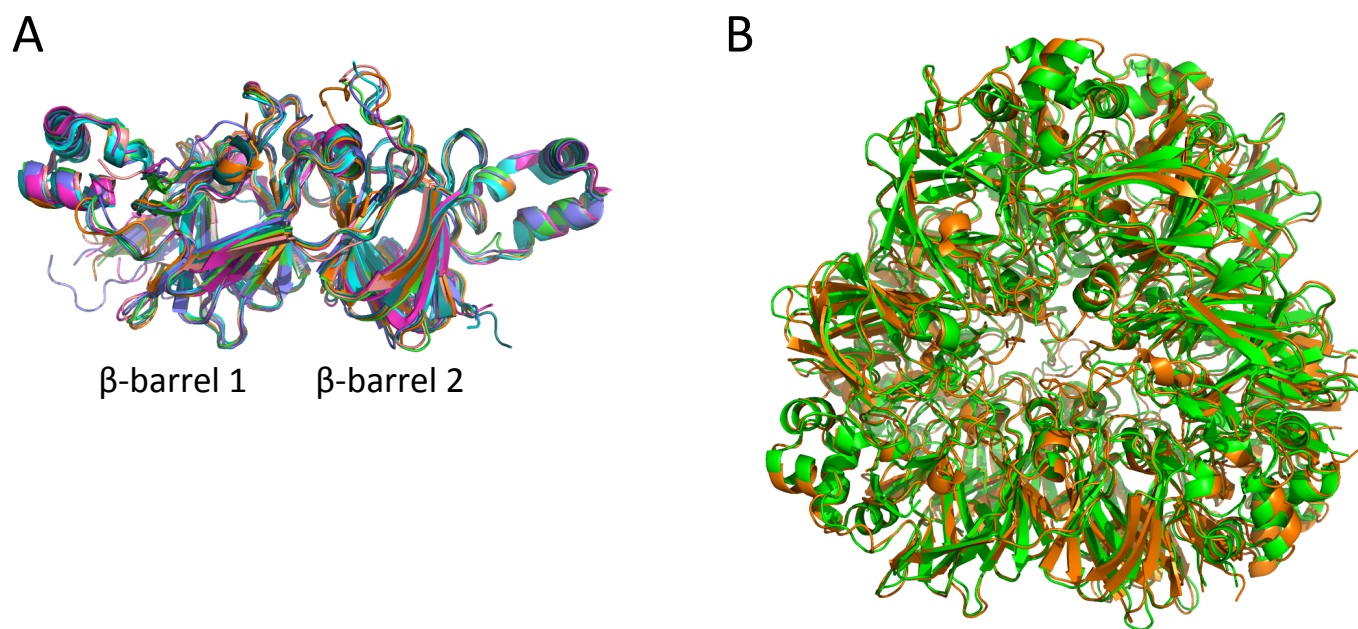


Fig5a

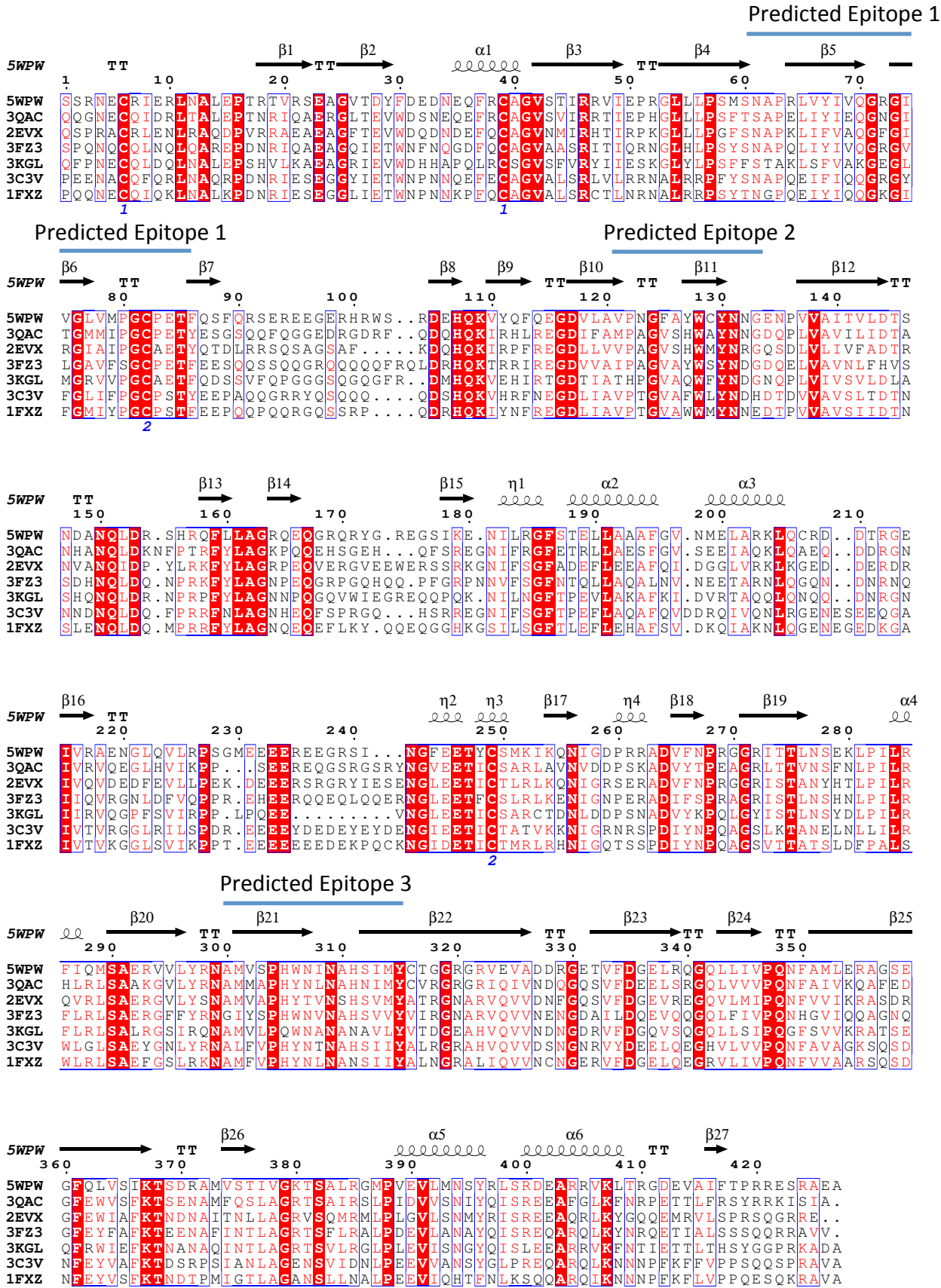


Fig5c

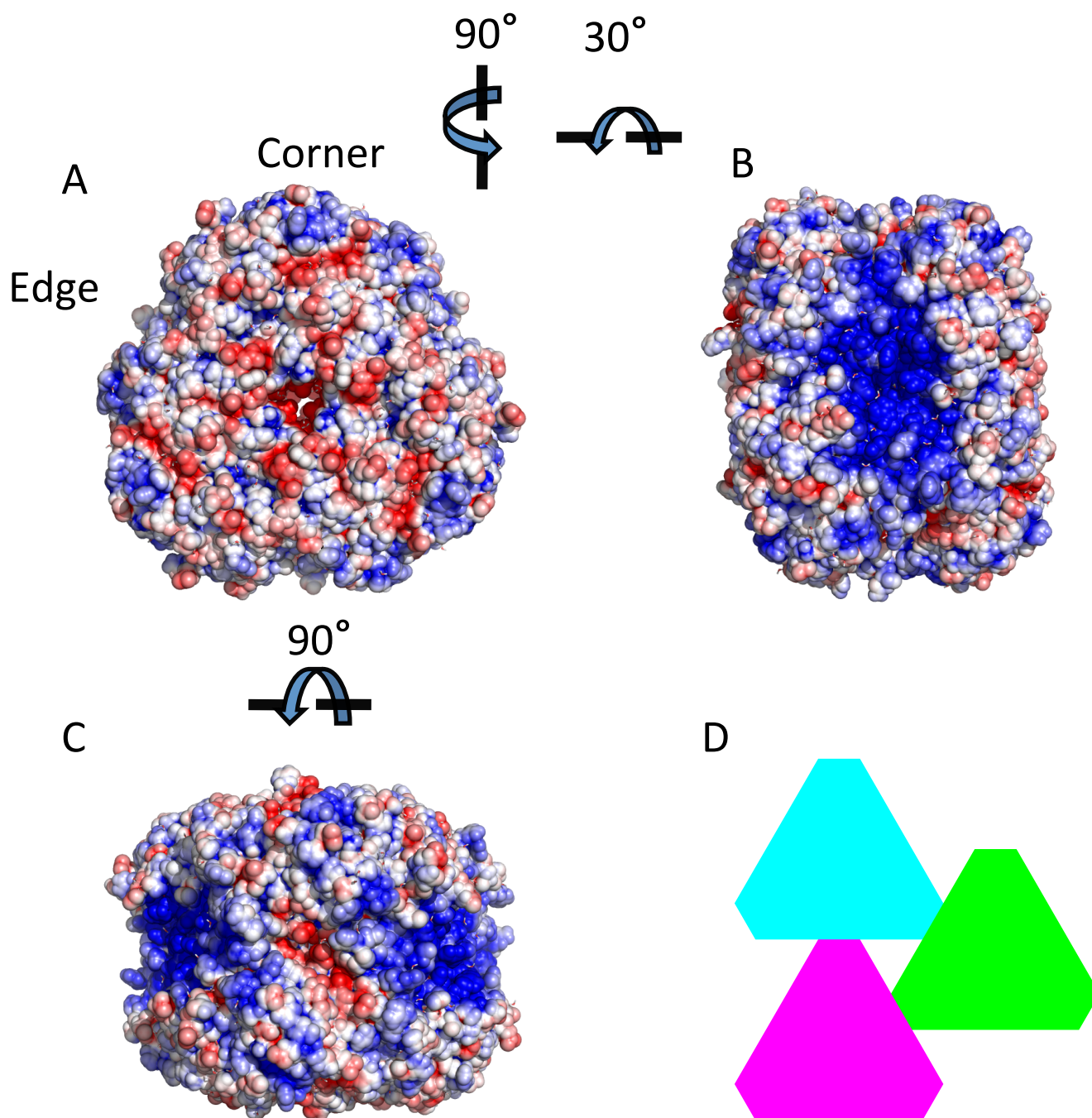


Fig6

TOC

