UNIVERSIDADE FEDERAL DO CEARÁ

CENTRO DE CIÊNCIAS

DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA

ANTÔNIO JOSÉ ROCHA

CLONAGEM DE cDNAS CODIFICANDO GLOBULINAS 7S (VICILINAS) DE DOIS GENÓTIPOS DE CAUPI [*Vigna unguiculata* (L.) Walp] COM RESPOSTA CONTRASTANTE AO CARUNCHO (*Callosobruchus maculatus*): SIMULAÇÕES COMPUTACIONAIS REVELAM COMO AS VICILINAS DO CAUPI INTERAGEM COM QUITINA.

FORTALEZA

2018

ANTÔNIO JOSÉ ROCHA

CLONAGEM DE cDNAS CODIFICANDO GLOBULINAS 7S (VICILINAS) DE DOIS GENÓTIPOS DE CAUPI [*Vigna unguiculata* (L.) Walp] COM RESPOSTA CONTRASTANTE AO CARUNCHO (*Callosobruchus maculatus*): SIMULAÇÕES COMPUTACIONAIS REVELAM COMO AS VICILINAS DO CAUPI INTERAGEM COM QUITINA.

Tese de doutorado apresentado à coordenação do Programa de Pós-Graduação em Bioquímica como requisito obrigatório para a obtenção do título de Doutor em Bioquímica pela Universidade Federal do Ceará-UFC.

Orientador: Prof. Dr. Thalles Barbosa Grangeiro

FORTALEZA
2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R571c Rocha, Antônio José.
Clonagem de cDNAs codificando globulinas 7S (vicilinas) de dois genótipos de caupi [Vigna unguiculata (L.) Walp] com resposta contrastante ao caruncho (Callosobruchus maculatus): simulações computacionais revelam como as vicilinas do caupi interagem com quitina. / Antônio José Rocha. – 2018.
112 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Bioquímica , Fortaleza, 2018.
Orientação: Prof. Dr. Thalles Barbosa Grangeiro.

1. Proteínas de armazenamento de sementes. 2. Leguminosae. 3. Callosobruchus maculatus. 4. local de ligação de quitina. 5. defesa vegetal. I. Título.

CDD 572

ANTÔNIO JOSÉ ROCHA

CLONAGEM DE cDNAS CODIFICANDO GLOBULINAS 7S (VICILINAS) DE DOIS GENÓTIPOS DE CAUPI [*Vigna unguiculata* (L.) Walp] COM RESPOSTA CONTRASTANTE AO CARUNCHO (*Callosobruchus maculatus*): SIMULAÇÕES COMPUTACIONAIS REVELAM COMO AS VICILINAS DO CAUPI INTERAGEM COM QUITINA.

Tese de doutorado apresentado à coordenação do Programa de Pós-Graduação em Bioquímica como requisito obrigatório para a obtenção do título de Doutor em Bioquímica pela Universidade Federal do Ceará-UFC.

Aprovada em: / /

BANCA EXAMINADORA

_____

Prof. Dr. Thalles Barbosa Grangueiro (Orientador)
Universidade Federal do Ceará-UFC

_____

Prof. Dr. Celso S. Nagano
Universidade Federal do Ceará-UFC

_____

Prof.Dr. José Hélio Costa
Universidade Federal do Ceará-UFC

_____

Profa. Dra. Vânia M. Ceccatto
Universidade Estadual do Ceará-UECE

_____

Prof. Dr. Daniel F. Feijó
Enstituto de Ensino Aprender Formação profissional

Com amor, dedico à minha mãe Raimunda Aquiles Rocha

AGRADECIMENTOS

Ao Professor Dr. Thalles Barbosa Grangeiro Agradeço por ser um orientador presente, um cientista admirável e por ter me possibilitado trabalhar em um projeto que me deu bastante satisfação. Professor, obrigado pelas oportunidades.

Aos professores Drs. Celso N. Nagano, Bruno A. Mathias Rocha, José Tadeus A. Oliveira e Valder Nogueira por suas contribuições essenciais para conclusão do manuscrito, a quem os admiro não somente por serem profissionais exemplares, mas também por serem pessoas altamente capacitados. Obrigado pela disponibilidade de contribuição para o trabalho e pela oportunidade de ter feito a parceria. Foi uma honra tê-los como colaboradores.

Aos prof. Drs. Bruno Lopes, Ito L. Barroso-Neto**,** Matheus Girão e ao Dr. Rômulo Farias pela contribuição no trabalho

Aos amigos do Laboratório de Genética e Biologia Molecular-LABGEM: Dr.Ednésio, Dr. Juscelino, Simone, Suelen, Jéssica Moura, e principalmente ao Dr. Edvar Monteiro -Júnior que contribui bastante para que esse trabalho fosse concluído. Agradeço a todos diretamente e indiretamente àqueles que contribuíram para esse trabalho e também pelos momentos de descontração, desabafos e pelo bom convívio.

Aos amigos do departamento: prof. Dr. Márcio Viana Ramos, Dr. Rafael de Sousa Miranda, Dra. Kátia Daniella da Cruz, saraiva, Ms. Mauro, Wallace, Débora, Eilton, Júnior, Sandro, etcc..

Às minhas amigas zeladoras do departamento de Biologia: Dna. Terezinha e Eliene, e da Bioquímica: Nega e Teca polos cafezinhos e brincadeiras!

É bom que posso sempre contar com vocês todos!

A todos os professores e os integrantes do Departamento de Bioquímica e Biologia Molecular da Universidade Federal do Ceará.

Aos meus pais Raimunda Aquiles Rocha e José Eudes Rocha (*In memoriam*) pelo apoio, dedicação e amor incondicional, vocês são o meu melhor presente de Deus e os responsáveis por todas as minhas conquistas.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP), Conselho

Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Departamento de Bioquímica e Biologia Molecular da Universidade Federal do Ceará, Departamento de Biologia, departamento de Física, Laboratório de Genética e Biologia Molecular-LABGEM, departamento de Engenharia de pesca e Centro Nacional de Processamento de Alto Desempenho-Universidade Federal do Ceará (CENAPAD-UFC**)** em cujos laboratórios esta pesquisa foi realizada.

A mente que se abre a uma nova idéia, jamais voltará ao seu tamanho original

Albert Einstein

RESUMO

O feijão caupi (*Vigna unguiculata*) é uma leguminosa que tem grande importância socioeconômica no Nordeste brasileiro. Entretanto, na fase de pós-colheita, um dos problemas enfrentados pelos agricultores é a infestação pelo *Calosobruchus maculatus* (Coleoptera: Bruchidae). Essa praga pode causar danos significativos às sementes, diminuindo seu teor nutricional e valor econômico. Diversos trabalhos correlacionam a resistência de alguns cultivares de caupi a uma ou mais variantes de vicilinas, globulinas 7S que são uma das principais proteínas de reserva dessas sementes. Essas variantes se ligam fortemente à quitina da membrana peritrófica de larvas de *C. maculatus*, prejudicando a absorção de nutrientes e, por consequência, causando mortalidade das mesmas. Por outro lado, variantes de vicilinas de cultivares susceptíveis ao *C. maculatus* se ligam fracamente à quitina. Mas até o momento, as diferenças estruturais que estão na origem da interação diferenciada entre as variantes de vicilinas e a quitina ainda são desconhecidas. Por essa razão, o objetivo principal do presente estudo foi investigar a base estrutural desse comportamento diferenciado das variantes de vicilinas 7S que poderiam explicar, pelo menos em parte, a resistência de sementes de algumas cultivares de caupi ao *C. maculatus*. Neste trabalho, as sequências parciais de cDNA que codificam vicilinas de caupi foram obtidas a partir de sementes em desenvolvimento dos genótipos EPACE-10 (suscetível ao *C. maculatus*) e IT81D-1053 (resistente ao bruquídeo), mediante amplificação por PCR, clonagem e sequenciamento. As sequências de aminoácidos deduzidas das sequências dos cDNAs foram então comparadas para se verificar as possíveis diferenças que poderiam existir entre as variantes de vicilinas 7S. Modelos moleculares tridimensionais das vicilinas de caupi, obtidos por modelagem por homologia, mostraram as características típicas do domínio dos membros da superfamília das cupinas. Simulações computacionais usando técnicas de *docking* e dinâmica molecular revelaram que cada trímero de vicilina continha três sítios de ligação à quitina, cada um localizado no vértice do oligômero em forma de triângulo, que é característico das globulinas 7S. Os modelos de interação previram que as moléculas de carboidrato, encaixadas nos sítios identificados, foram estabilizadas principalmente por pontes de hidrogênio, muitas delas mediadas por moléculas de água, uma característica estrutural comum observada em uma grande variedade de proteínas que se ligam a carboidratos. Além disso, muitos dos resíduos envolvidos nos locais de ligação à quitina de vicilinas de caupi são conservados em outras globulinas 7S. Utilizando um método baseado em mecânica quântica, as energias livres de ligação de moléculas de quito-oligossacarídeos aos oligómeros de β-vignina foram calculadas, variando de -690,6 kcal / mol (oligômeros de

GlcNAc ligados ao trímero de β-vignina R2) a -531,3 kcal / mol (fragmentos de quitina ligados ao oligômero β-vignina R3). Estes resultados concordam com trabalhos anteriores, que demonstraram que as vicilinas de caupi purificadas a partir de sementes de genótipos suscetíveis ou resistentes a *C. maculatus* têm a capacidade de se ligar *in vitro* à quitina e *in vivo* a estruturas quitinosas do intestino médio das larvas. Além disso, a conservação em outras globulinas 7S de muitos dos resíduos que constituem o sítio de ligação a quitina de β-vignina, como previsto neste trabalho, também suporta as hipóteses anteriores que mostraram que as vicilinas de outras espécies de leguminosas também podem se ligar à quitina. Os resultados do presente trabalho fornecem a primeira descrição do mecanismo molecular envolvido na interação entre vicilinas de caupi e quitina.

Palavras-chaves: Proteínas de armazenamento de sementes. Leguminosae. *Callosobruchus maculatus*. Local de ligação de quitina. Defesa vegetal

ABSTRACT

Cowpea (Vigna unguiculata) is a legume that has great socioeconomic importance in the Brazilian Northeast. However, in the post-harvest phase, one of the problems faced by farmers is the infestation by Calosobruchus maculatus (Coleoptera: Bruchidae). This pest can cause significant damage to the seeds, reducing their nutritional content and economic value. Several studies correlate the resistance of some cowpea cultivars to one or more variants of vicillins, 7S globulins that are one of the main reserve proteins of these seeds. These variants bind strongly to the chitin of the peritrophic membrane of C. maculatus larvae, impairing the absorption of nutrients and, consequently, causing their mortality. On the other hand, varicose veins of cultivars susceptible to C. maculatus bind poorly to chitin. But so far, the structural differences that are at the origin of the differentiated interaction between the variants of vicillin and chitin are still unknown. For this reason, the main objective of the present study was to investigate the structural basis of this differentiated behavior of variants of 7S vicillins that could explain, at least in part, the resistance of seeds of some cowpea cultivars to C. maculatus. In this work, partial cDNA sequences coding for cowpea vicillin were obtained from developing seeds of the EPACE-10 genotypes (susceptible to C. maculatus) and IT81D-1053 (brucellosis resistant) genotypes by PCR amplification, cloning and sequencing. The deduced amino acid sequences of the cDNA sequences were then compared to verify possible differences that could exist between the variants of 7S vicilins. Three-dimensional molecular models of cowpea vicillin, obtained by homology modeling, showed the typical characteristics of the domain of the members of the superfamily of the termites. Computational simulations using docking techniques and molecular dynamics revealed that each trimer of vicillin contained three chitin binding sites, each located at the apex of the triangle-shaped oligomer, which is characteristic of 7S globulins. The interaction models predicted that the carbohydrate molecules, embedded in the identified sites, were stabilized mainly by hydrogen bonds, many of them mediated by water molecules, a common structural feature observed in a large variety of proteins that bind to carbohydrates. In addition, many of the residues involved in the chitin binding sites of cowpea vicillin are conserved in other 7S globulins. Using a quantum mechanical-based method, the free energies of binding of chito-oligosaccharide molecules to the β-vignin oligomers were calculated, ranging from -690.6 kcal / mol (GlcNAc oligomers bound to β-vignin trimer R2) to -531.3 kcal / mol (chitin fragments bound to the β-vignin oligomer R3). These results agree with previous studies that demonstrated that cowpea vicillins purified from seeds of C. maculatus susceptible or resistant genotypes have the ability to bind in vitro to chitin

and in vivo to chitin structures of the midgut of larvae. In addition, the conservation in other 7S globulins of many of the residues constituting the β-vignin chitin binding site as predicted in this paper also supports previous hypotheses that have shown that the vicillins of other legume species can also bind to chitin. The results of the present work provide the first description of the molecular mechanism involved in the interaction between cowpea and chitin vicillins.

Keywords: Seed storage proteins. Leguminosae. *Callosobruchus maculatus*. Chitin-binding-site. Plant defense

LISTA DE ILUSTRAÇÕES

are shown. Water molecules (OH2) are represented as red spheres. Similar diagrams depicting the interaction models obtained for the other complexes are shown in Figs. S25-S27. The diagrams were generated using the program LigPlot+ (Laskowski and Swindells, 2011).

Figure 5 (S1)-1% agarose gel electrophoresis of cDNA fragments encoding cowpea vicilins amplified by RT-PCR from total mRNA, which was purified from developing seeds from genotypes IT81D-1053 and EPACE-10. Lane 1: DNA size markers; lane 2: PCR products from genotype IT81D-1053; lane 3: PCR products from genotype EPACE-10. The major amplified band from each sample, which was cloned into the pGEM-T Easy vector, is indicated by an arrow.

Figure 6 (S2)-Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype EPACE-10 isoform S1 (GenBank accession number: MG973241). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

Figure 7 (S3)-Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype EPACE-10 isoform S2 (GenBank accession number: MG973242). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

Figure 8 (S4)-Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype EPACE-10 isoform S3 (GenBank accession number: MG973243). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively

Figure 8 (S5)-Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype IT81D-1053 isoform R1 (GenBank accession number: MG973244). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence

determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively

Figure 9 (S6)-Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype IT81D-1053 isoform R2 (GenBank accession number: MG973245). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

Figure 10 (S7)-Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype IT81D-1053 isoform R3 (GenBank accession number: MG973246). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

Figure 11 (S8)-Alignment of partial cDNA sequences encoding cowpea vicilins, which were cloned from genotypes EPACE-10 (sequences S1, S2 and S3) and IT81D-1053 (sequences R1, R2 and R3).

Figure 12 (S9)-Alignment of the amino acid sequences of cowpea vicilins, which were deduced from partial cDNA sequences obtained from genotypes EPACE-10 and IT81D-1053. Sites containing identical residues are shaded in black, whereas positions with chemically similar residues are shaded in gray. The alignment was rendered using BOXSHADE v3.21 through the web server embnet.vital-it.ch/software/BOX_form.html.

Figure 13 (S10)-SDS-polyacrylamide gel electrophoresis of seed vicilins from cowpea (*V. unguiculata*). SDS-PAGE was performed as described by Laemmli (Laemmli, 1970) using 15% slab gels. Samples (20 µg) were prepared in 0.0625 M Tris-HCl (pH 6.8) containing 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 10% (v/v) glycerol and 0.001% (w/v) bromophenol blue. Protein bands were stained with 0.2% (w/v) Coomassie Brilliant Blue R250 in 50% methanol/10% acetic acid for 16 h. Destaining was carried out with 12.5% isopropanol/10% acetic acid. Lane 1: molecular weight markers. Lane 2: total vicilins extracted from genotype IT81D-1053. Lane 3: total vicilins extracted from genotype EPACE-10. The apparent molecular masses of the major protein bands identified in each fraction (labeled 1 to 11) are shown in parentheses.

Figure 14 (S11)-Size exclusion chromatography of cowpea vicilins on a column of Sephacryl S-200 HR coupled to an ÄKTA prime plus chromatography system. Vicilins from genotypes IT81D-1053 (**A**) and EPACE-10 (**B**) were resuspended in 50 mM Tris-HCl buffer, pH 8.0 (5 mg/mL), and 10 mg of each sample were loaded onto the column. The column was equilibrated and eluted with 50 mM Tris-HCl buffer, pH 8.0, at a constant flow rate (0.5 mL/min) and 3 mL fractions were collected.

Figure 15 (S12)-Ion exchange chromatography of cowpea vicilins on DEAE-Sepharose. Vicilins from cowpea genotypes IT81D-1053 (**A**) and EPACE-10 (**B**) were loaded onto a DEAE Sepharose Fast Flow (GE Healthcare) column (6 mL) equilibrated with 50 mM Tris-HCl buffer pH 8.0. The chromatography was performed at a constant flow rate (2 mL/min) and 3 mL fractions were collected.

Figure 16 (S13)-SDS-polyacrylamide gel electrophoresis of seed vicilins from cowpea (*V. unguiculata*) purified by ion exchange chromatography. Peaks I (lane 1; 20 µg) and II (lane 2; 20 µg) from the ion exchange chromatography of seed vicilins from genotypes IT81D-1053 (**A**) and EPACE-10 (**B**) were subjected to SDS-PAGE (15% polyacrylamide) and protein bands were stained with 0.2% (w/v) Coomassie Brilliant Blue R250 in 50% methanol/10% acetic acid for 16 h. Destaining was carried out with 12.5% isopropanol/10% acetic acid. Lane M: molecular weight markers.

Figure 17 (S14). MALDI-TOF mass spectra of tryptic digest of cowpea seed vicilins from genotype IT81D-1053. Spectra are from tryptic digestion of proteins from peak I (**A**) and peak II (**B**), as obtained by ion exchange chromatography. Assignment of masses to appropriate amino acid sequences is shown in Table S3.

Figure 17 (S15**)**-MALDI-TOF mass spectra of tryptic digest of cowpea seed vicilins from genotype EPACE-10. Spectra are from tryptic digestion of proteins from peak I (**A**) and peak II (**B**), as obtained by ion exchange chromatography. Assignment of masses to appropriate amino acid sequences is shown in Table S4.

Figure 18 (S16)-MS/MS spectra of the ion at m/z 1391. Amino acid sequence specific y-ions are indicated.

Figure 19 (S17)-MS/MS spectra of the ion at m/z 1960. Amino acid sequence specific y-ions are indicated.

Figure 20 (S18**)**-MS/MS spectra of the ion at m/z 1177. Amino acid sequence specific y-ions are indicated

Figure 21(S19)-Alignment of the amino acid sequence of cowpea vicilin from genotype

EPACE-10 (sequence S2) with the primary structure of the 7S globulin-1 (PDB ID: 2EA7) from adzuki bean (*V. angularis*). Identical and chemically-similar residues are shaded in black and gray, respectively. The alignment was rendered using BOXSHADE v3.21 through the web server embnet.vital-it.ch/software/BOX_form.html.

Figure 22 (S20)-Alignment of the amino acid sequence of cowpea vicilin from genotype EPACE-10 (sequence S3) with the primary structure of the 7S globulin-1 (PDB ID: 2EA7) from adzuki bean (*V. angularis*). Identical and chemically-similar residues are shaded in black and gray, respectively. The alignment was rendered using BOXSHADE v3.21 through the web server embnet.vital-it.ch/software/BOX_form.html.

Figure 23 (S21)-Alignment of the amino acid sequence of cowpea vicilin from genotype IT81D-1053 (sequence R2) with the primary structure of the 7S globulin-1 (PDB ID: 2EA7) from adzuki bean (*V. angularis*). Identical and chemically-similar residues are shaded in black and gray, respectively. The alignment was rendered using BOXSHADE v3.21 through the web server embnet.vital-it.ch/software/BOX_form.html.

Figure 24 (S22)-Alignment of the amino acid sequence of cowpea vicilin from genotype IT81D-1053 (sequence R3) with the primary structure of the 7S globulin-1 (PDB ID: 2EA7) from adzuki bean (*V. angularis*). Identical and chemically-similar residues are shaded in black and gray, respectively. The alignment was rendered using BOXSHADE v3.21 through the web server embnet.vital-it.ch/software/BOX_form.html.

Figure 25 (S23)-Superposition of the three-dimensional molecular models of β-vignin. The superposed models, shown as ribbon diagrams, are colored pink (model S2), yellow (model S3), green (model R2) and cyan (model R3). The modeled structures were superposed by overlapping 354 Cα-atoms of structurally equivalent residues with an RMSD of 1.479 Å.

Figure 26 (S24)-Root-mean-square deviations (RMSD; Å) of the backbone atoms of a tetra-*N*-acetyl-chitotetraose molecule [$(GlcNAc)_4$] docked in the chitin-binding site of cowpea vicilins, during the 40 ns MD simulations. The graphics show the backbone RMSD plots of $(GlcNAc)_4$ docked in the three-dimensional molecular models generated from sequences S2 (**A**), S3 (**B**), R2 (**C**) and R3 (**D**).

Figure 27 (S25)-Two-dimensional diagram of the interaction model between a $(GlcNAc)_4$ molecule docked in the chitin-binding site of β-vignin. Residues from model R3 involved in hydrogen bonds (green dashed lines) or hydrophobic contacts (brick-red spoked arcs) are shown. Water molecules (OH2) are represented as red spheres.

Figure 28 (S26)-Two-dimensional diagram of the interaction model between a $(GlcNAc)_4$ molecule docked in the chitin-binding site of β-vignin. Residues from model S2 involved in hydrogen bonds (green dashed lines) or hydrophobic contacts (brick-red spoked arcs) are shown. Water molecules (OH2) are represented as red spheres.

LISTA DE TABELAS

Table S9- Summary statistics of the MolProbity analysis of the three-dimensional molecular model of cowpea vicilin from sequence S2. The analysis was performed by submitting the three-dimensional molecular model to the MolProbity web server (http://molprobity.biochem.duke.edu/)

Table S10-Summary statistics of the MolProbity analysis of the three-dimensional molecular model of cowpea vicilin from sequence S3. The analysis was performed by submitting the three-dimensional molecular model to the MolProbity web server (http://molprobity.biochem.duke.edu/)

Table S11-Summary statistics of the MolProbity analysis of the three-dimensional molecular model of cowpea vicilin from sequence R2. The analysis was performed by submitting the three-dimensional molecular model to the MolProbity web server (http://molprobity.biochem.duke.edu/)

Table S12-Summary statistics of the MolProbity analysis of the three-dimensional molecular model of cowpea vicilin from sequence R3. The analysis was performed by submitting the three-dimensional molecular model to the MolProbity web server (http://molprobity.biochem.duke.edu/)

Table S13-Calculated binding free energies (kcal/ml) of $(GlcNAc)_4$ molecules bound to β-vignin molecular models. The values were based on a quantum mechanical method (Maranhão et al., 2017).

LISTA DE ABREVIAÇÕES E SIGLAS

R-Resistent

S-Susceptível

cDNA- complementar ao DNA

GlcNAc - N-acetil-D-glucosamina

2- BLAST- Basic Local Alignment Search Tool

LC-ESI-MS/MS- Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometric

UPLC-ultra-high-performance liquid chromatography

MALDI-TOF/TOF- Matrix-assisted laser desorption/ionization time-of-flight tandem

PMF-peptide mass fingerprint

NCBI- National Center for Biotechnology Information

ESI-MS/MS- Electrospray Ionization Tandem Mass

RMSD- root-mean-square deviation

PDB-protein data Ban

SUMÁRIO

1. INTRODUÇÃO GERAL

O feijão caupi (*Vigna unguiculata*) possui um importante papel sócio-econômico em regiões tropicais e subtropicais da África, Ásia e Américas (SINGH et al., 2003). No Brasil, principalmente nas regiões Norte e Nordeste, as sementes de caupi são uma das principais fontes de proteínas na nutrição humana. Na fase de pós-colheita, um dos problemas enfrentados pelos produtores é a infestação das sementes pelo caruncho do caupi, *Callosobruchus maculatus* Fabricius, 1775 (Coleoptera: Chrysomelidae). Essa praga causa danos às sementes, reduzindo o valor comercial do produto destinado ao consumo humano.

Na década de 70, a coleção de germoplasma de caupi do Instituto Internacional de Agricultura Tropical (IITA, Nigéria) foi avaliada com o objetivo de se identificar cultivares naturalmente resistentes ao C. maculatus. De oito mil cultivares analisadas, três apresentaram resistência ao caruncho do caupi: TVu 2027, TVu 11952 e TVu 11953 (SINGH et al., 1985).

Gatehouse e colaboradores (1979) atribuíram a resistência das sementes da cultivar resistente TVu 2027 ao ataque de C. maculatus a níveis elevados de inibidores de proteases do tipo tripsina em relação a oito cultivares suscetíveis, que apresentaram baixos níveis desse inibidor. Entretanto, Xavier-Filho e colaboradores (1989), utilizaram cultivares suscetíveis e resistentes ao ataque do caruncho do feijão caupi e mostraram que não existia correlação significativa entre os níveis de inibidores de tripsina e a resistência/suscetibilidade das mesmas, contrariando as sugestões de Gatehouse e colaboradores (1979). De fato, larvas de *C. maculatus* não usam majoritariamente proteases do tipo tripsina para sua digestão proteica, mas sim proteinases aspárticas e principalmente proteinases cisteínicas (Campos et al., 1989).

Posteriormente, Macedo e colaboradores (1993) sugeriram que a resistência ao *C. maculatus* poderia ser atribuída a uma forma variante de vicilina 7S, que era deletéria ao desenvolvimento das larvas. Estudos posteriores revelaram que as vicilinas dos cultivares resistentes se ligavam fortemente à quitina (in vitro) e à camada de quitina que reveste a membrana peritrófica do trato intestinal (SALES et al., 1996; SALES et al., 2001). Sales e colaboradores (1996; 2001) sugeriram que essa interação limitava a assimilação dos nutrientes no intestino das larvas de *C. maculatus*, dificultando assim a sobrevivência e desenvolvimento do inseto. Mais recentemente, Aquino (2009) realizou um trabalho in silico e modelou a proteína vicilina 7S de sementes de feijão caupi usando uma vicilina 7S de sementes de *Glycine max* como modelo comparativo. Em seguida, através de ferramentas de docking molecular determinou a presença de interação com quitina mostrando três domínios de ligação à quitina: Haveína, R&R consenso e o domínio conglicinina (CD). Além disso, foi mostrado que o domínio CD é o único que participa diretamente da ligação à quitina.

Diante desses dados, algumas lacunas precisam ser elucidadas. Embora Aquino (2009) tenha identificado um domínio de ligação à quitina, este achado foi baseado apenas em dados in silico. Ademais, Silva (2014) evidenciou diferenças nas sequências de aminoácidos deduzidas a partir do sequenciamento parcial do cDNA que codifica vicilinas 7S de dois genótipos de feijão caupi contrastantes quanto à resistência ao *C. maculatus*: IT81D-1053 (resistente) e EPACE 10 (susceptível). As sequências parciais de aminoácidos das vicilinas 7S de ambas cultivares divergiram especificamente em regiões ricas em glutamina. Tal análise encontrou locais potencialmente ligantes a radicais fosfato e radicais eletronegativos como grupamentos amínicos e carboxílicos, sendo que a região variável de glutaminas foi indicada como potencial ligante a esses grupos eletronegativos. Apesar dos achados descritos estes dados foram obtidos com base nas sequências parciais das vicilinas 7S e não com base nas sequências completas. Em resumo, até o momento ainda não se sabe com clareza o mecanismo estrutural de interação das vicilinas 7S de feijão caupi com quitina.

Tendo em vista essas importantes lacunas à respeito das vicilina 7S do feijão caupi, além dos prejuízos pós-colheita ocasionados por C. maculatus, é de suma importância que um estudo que objetive elucidar o mecanismo de interação de vicilinas 7S de feijão caupi com a quitina de C. maculatus seja realizado, a fim de que os prejuízos ocasionados por esta praga sejam minimizados.

Vicilinas ou vigninas são globulinas 7S de armazenamento de sementes que contêm duas cópias do domínio da superfamília de cupins. O domínio de cupinas é caracterizado por uma dobra de β-hélice de cadeia dupla, que é encontrada em diversas enzimas, bem como em proteínas de armazenamento de sementes não catalíticas (Dunwell et al., 2004). No entanto, não existe uma semelhança estrutural entre o domínio da cupina e as proteínas de ligação à quitina conhecidas, como a heveína e as lectinas de ligação à quitina, por exemplo. Essas proteínas canônicas de ligação à quitina têm locais típicos de ligação a carboidratos que reconhecem e unem um ou mais resíduos de N-acetil-D-glucosamina, a unidade repetitiva de quitina. Portanto, o mecanismo molecular envolvido na interação entre vicilinas de caupi e oligômeros de quitina permanece mal compreendido. No presente trabalho, as sequências parciais de cDNA que codificam vicilinas de caupi foram obtidas a partir dos genótipos EPACE-10 e IT81D-1053. Os modelos moleculares tridimensionais foram gerados, validados e submetidos a cálculos de encaixe molecular e simulações de dinâmica molecular, com o objetivo de revelar os locais putativos de ligação à quitina nas estruturas modeladas de vicilinas.

2-REFERENCIAL TEÓRICO

2. 1. Feijão caupi (*Vigna unguiculata*)

O feijão caupi [*Vigna unguiculata* (L.) Walpers], é uma planta dicotiledônea, pertencente à ordem Fabales, família Fabaceae, subfamília Faboideae, tribo Phaseoleae e à subtribo Phaseolinae. O gênero Vigna apresenta cerca de 160 espécies (PADULOSI; NG., 1997).

Essa cultura é originária do oeste da África, sendo cultivada nas principais regiões tropicais em todo o mundo (SINGH *et al*, 1985). O maior centro de diversidade fica também no oeste da África, localizado no Instituto Internacional de Agricultura Tropical (IITA, Nigéria) (PANT et al., 1982; NG). Entretanto, a maior diversidade de genótipos selvagens ocorre, principalmente, na região sudeste da África (PADULOSI; NG., 1997).

No Brasil, existe uma grande diversidade de cultivares de feijão caupi, entretanto, essa cultura recebe vários nomes dependendo da região do Brasil, como feijão de corda, massacar, feijão de vagem, feijão de vara, feijão de moita. Na Bahia é conhecido como feijão catador; feijão trepa-pau no Maranhão; feijão gurutuba no norte de Minas Gerais. Já no Rio de Janeiro é conhecido como feijão fradinho e no Sul como feijão miúdo. Já no Norte é conhecido como feijão da praia. (CAJAZEIRAS, 2000). Diante de todas as diversidades desses nomes populares a Embrapa cientificamente a denominou de feijão caupi em referência ao nome em inglês Cowpea.

O grão seco é o principal produto do consumo humano, sendo este uma das principais fontes de proteínas, carboidratos e lipídios para a população brasileira, principalmente a região Nordeste, e também na África (SINGH *et al*, 1985).

O teor proteico das sementes varia de 19,5 a 26,1% (MAIA et al, 2000). As proteínas totais podem ser divididas em quatro grupos de acordo com sua solubilidade: albuminas, globulinas, glutelinas e prolaminas, sendo as globulinas, as principais proteínas de reversa (cerca de 70% do conteúdo total de nitrogênio (AREMU 1990; OSBORN, 1988). Todos os aminoácidos essenciais encontram-se nas sementes de feijão caupi, sendo bastante rico em lisina e pobre em aminoácidos sulfurados (AREMU, 1990).

Alguns dos principais problemas do cultivo do feijão caupi são advindos da ação de fungos, nematóides e bactérias, além do ataque de pestes, como o caruncho (RIOS, 1988). Dentre as principais doenças que acometem o feijão caupi podem ser citadas: podridão das raízes, murcha de fusarium, mancha bacteriana, antracnose, além de doenças ocasionadas pelo vírus do mosaico do caupi (CPMV), e pelo vírus do mosaico severo do caupi (CPSMV), entre

outros (OLIVEIRA, 1981; RIOS, 1988). Já no período pós-colheita a principal praga que ataca seu armazenamento é Callosobruchus maculatus, conhecido popularmente como caruncho ou gorgulho (CREDLAND, 1986).

2. 2. *Callosobruchus maculatus* e sua interação com o feijão caupi (*Vigna unguiculata*)

Dentre os insetos que atacam o feijão caupi no período de armazenamento, o *Callosobruchus maculatus* tem sido seu principal predador, entretanto, esse bruquídeo não ataca às sementes do feijão-comum (*Phaseolus vulgaris*), devido à presença de inibidores de α-amilases. O principal predador do feijão comum é o bruquídeo Zabrotes subfasciatus. O *C. maculatus* pertence à classe Insecta, ordem Coleóptera, subordem Poliphaga, subfamília Chrysomeloidea, família Chrysomelidae, gênero Calosobruchus e à espécie *C. maculatus*. Essa espécie é originária da África, no entanto, hoje é frequentemente encontrada nas regiões tropicais e subtropicais do mundo, uma vez que foi propagada através do comércio (SOUTHGATE, 1979).

As larvas desses insetos só se alimentam e crescem no interior das sementes de leguminosas (Fabaceae). Já os adultos desta espécie, não se alimentam ou necessitam de água, são chamados insetos afagos (OLIVEIRA, 2013). Além disso, o tempo de desenvolvimento é de aproximadamente duas a três semanas, passando boa parte desse tempo acasalando-se e realizando a postura de ovos nos feijões (OLIVEIRA, 2013). Esse inseto possui um desenvolvimento holometabólico descrito por BASTOS (1968), passando pelas fases de ovo, larva, pupa e adulto.

O ataque do *C. maculatus* ao feijão caupi se inicia quando as fêmeas depositam seus ovos nas sementes estocadas. As larvas eclodem, penetram nas sementes e passam a se alimentar das proteínas de reserva e carboidratos. Após certo período de desenvolvimento elas se transformam em pupas das quais emergem insetos adultos (CREDLAND, 1986). Machos e fêmeas adultos acasalam-se e cerca de duas horas depois, as fêmeas começam a ovopositar em novas sementes. (DICK; CREDLAND, 1986; EDVARDSSON, 2005).

2. 3. Cultivares de *Vigna unguiculata* resistentes ao *Callosobruchus maculatus*

A cultivar TVu 2027 identificada pelo IITA na Nigéria deu origem a várias outras, como IT81D-1032, IT81D-1064 e IT81D-1045, mostrando diferentes níveis de resistência ao caruncho do feijão caupi (SINGH et al, 1985, SINGH et al, 2003). A resistência conferida a estes feijões variantes foi atribuída a vários fatores, que na sua maioria, são de natureza química.

Inicialmente Gatehouse e colaboradores (1979) propuseram que a resistência observada estava associada a altos níveis de inibidores de tripsina nessas sementes. Entretanto, uma hipótese proposta por Xavier-Filho e colaboradores (1989) descartou a relação dos inibidores de proteases com a resistência ao caruncho proposta por Gatehouse e colaboradores (1979). Um experimento usando três cultivares suscetíveis (CE-11, CE-12 e CE-524) e três cultivares resistentes (ITB1D-1045, ITB1D-1064 e TVu 2027) ao caruncho do caupi, mostraram que não existia correlação significativa entre os níveis de inibidores de tripsina e a resistência/suscetibilidade das mesmas. Nesse experimento, as cultivares resistentes IT81D-1045 e IT81D-1064 mostraram baixas concentrações de inibidores de tripsina, refutando a hipótese de Gatehouse e colaboradores (1979), muito embora uma das cultivares resistentes, TVu 2027, tenha apresentado altos níveis de inibidor de tripsina. Entretanto, três cultivares suscetíveis (CE-11, CE-12 e CE-524) também apresentaram altos níveis desse inibidor, o que descartou totalmente a teoria de Gatehouse e colaboradores (1979).

Atrelado a isso, Campos e colaboradores (1989) demonstraram em seus experimentos que as larvas de C. maculatus usam proteinases aspárticas e principalmente proteinases cisteínicas para sua digestão proteica e não proteases do tipo tripsina.

Ademais, Sales e colaboradores (1992) submeteram amostras de vicilinas de cultivares suscetíveis [CE-11 e CE-31 (Pitiuba)] e resistentes (IT81D-1032 e IT81D-1045) ao inseto à digestão proteica com pepsina, papaína e extratos do intestino de larvas de C. maculatus, contendo as proteases nativas. Os autores observaram que as vicilinas das cultivares resistentes foram mais refratárias à ação das proteases do que as vicilinas das cultivares suscetíveis.

Mais tarde, Macedo e colaboradores (1993) realizaram um experimento com duas cultivares resistentes ao *C. maculatus*, IT81D-1032 e IT81D-1045 que os levaram a sugerir que a forma variante da vicilina (globulina 7S) foi deletéria ao desenvolvimento das larvas.

Os estudos químicos e imunológicos mostraram que a vicilina variante IT81D-1045 tinha um grande poder de se ligar a estruturas do intestino médio que contém quitina aderindo-

se nos ápices epiteliais das células intestinas das larvas (SALES et al., 1996; SALES et al., 2001).

Além disso, Sales e colaboradores (2001) mostraram também que as vicilinas das cultivares resistentes e susceptíveis se ligavam fortemente e fracamente, respectivamente, à quitina (in vitro) e à camada da membrana peritrófica do trato intestinal. Diante disso, os autores sugeriram que essa interação da vicilina 7S com a camada de quitina que compõe o intestino do C. maculatus não deixava que as larvas assimilassem os nutrientes necessários para o seu desenvolvimento, dificultando sua sobrevivência ou levando à morte do inseto.

Portanto, os autores postularam que a suscetibilidade diferenciada das vicilinas 7S às proteases do inseto constitui um aspecto do mecanismo de resistência das sementes ao ataque pelo *C. maculatus*.

2. 4. Proteínas de reserva de sementes

As proteínas de reserva foram primeiramente classificadas por Osborne (1988), com base na sua extração e solubilidade em água (albuminas), solução salina (globulinas), solução alcoólica (prolaminas) e solução ácida ou básica (glutelinas) (SHEWRY et al., 1995).

A maioria das proteínas em sementes de leguminosas consiste de globulinas, proteínas de reserva sintetizadas durante o desenvolvimento da semente. Elas estão presentes não somente em dicotiledôneas, mas também em monocotiledôneas, incluindo cereais e palmeiras e em esporos de samambaias (TEMPLEMAN et al., 1987).

Durante a germinação da semente as globulinas são estocadas em corpos proteicos, ou esqueletos de nitrogênio e carbono que fornecem energia para o desenvolvimento das plântulas e são hidrolisadas durante o processo germinativo. As demais proteínas são principalmente albuminas, que incluem proteínas de reserva e de manutenção, lectinas, lipoxigenases, entre outras (SHEWRY et al., 1995; DODEMAN et al., 1997; MARUYAMA et al., 2003; WANG et al., 2003).

As globulinas são geralmente classificadas como proteínas 11S/12S (leguminas) e 7S/8S (vicilinas) de acordo com seu coeficiente de sedimentação. Essas proteínas são oligoméricas, e quando submetidas à dissociação, originam polipeptídeos heterogêneos. Essa heterogeneidade é evidente pelo tamanho e pela carga dos polipeptídeos e resulta de uma combinação de dois fatores: origem multigênica de cada globulina e modificações pós-traducionais. As principais proteínas de reserva do feijão caupi são as globulinas do tipo vicilina 7S, que são responsáveis pelo valor nutricional da semente (PEDALINO et al.,1992; SHEWRY, et al., 1995; DURANTI e GIUS, 1997; PAES et al., 2008; OLIVEIRA, 2013).

2. 5. Vicilinas (Globulinas 7S)

As vicilinas (7S) representam aproximadamente 80% das proteínas totais de sementes maduras. As globulinas de sementes que são tipicamente classificadas como vicilinas consistem em uma combinação de subunidades com massa molecular que varia de 150 a 190 kDa. Nestas subunidades de vicilinas, não há presença de pontes dissulfeto e a proteína é estabilizada por forças não covalentes (PAES et al., 2008; OLIVEIRA, 2013).

As vicilinas 7S Possuem dois domínios cupin_1 (superfamília cupinas: bicupinas) e são constituídas majoritariamente por β-vigninas e Proteína trimérica e possui 55 e 60 kDa em suas subunidades de polipeptídeos (Salles et al,1992; 2001). Elas consistem em uma combinação de 3 subunidades com massa molecular que varia de 156 a 177 kDa formando um trímero em formato de triângulo e apresentam uma composição de aminoácidos com altas concentrações de ácido aspártico, ácido glutâmico, arginina, fenilalanina e leucina, mas têm, no entanto, concentrações mínimas de aminoácidos sulfurados como metionina e cisteína (MACEDO et al.,1995). A ausência de resíduos de cisteínas nas globulinas do tipo vicilina 7S impede que as mesmas possam realizar pontes dissulfeto. O tamanho final da proteína madura varia consideravelmente devido aos processamentos pós-traducionais, como proteólise e glicosilação (SHEWRY et al., 1995; SHUTOV et al., 1995).

Estas proteínas de reserva parecem ser multifuncionais, fornecendo aminoácidos durante a germinação da planta e participando dos mecanismos de defesa das sementes, sendo tóxicas a insetos (MACEDO et al., 1993; SHUTOV et al., 1995; Sales et al., 2000; Souza, 2009). Existe um grau variado de proteólise e de glicosilação (4-8%) nas β–vigninas (Macedo et al, 1993; Salles et al, 2001).

6-REFERÊNCIAS

1. Aremu, C. Y. Proximate and amino acid composition of cowpea (*vigna unguiculata*, walp) protein concentrate prepared by isoeletric point precipitation. Food chemistry, v. 8, p. 61-68, 1990.
2. Aquino, R.erminação de motivos de ligação à quitina em vicilinas de *canavalia ensiformis* e *vigna unguiculata* através de métodos in silico e relação com suas toxicidades para o bruquídeo *callosobruchus* (coleoptera: bruchidae). 2009. Dissertação (mestrado em bioquímica) -centro de biociências, universidade federal do rio grande do norte, natal, 2009.
3. Bastos, J. A. M. Principais pragas das culturas e seus controles. 3.ed. São paulo: nobel, 1985, 329p.
4. Cajazeiras, J. B. Identificação de genótipos de caupi [*vigna unguiculata* (l.) Walp.] Resistentes ao caruncho (*Callosobruchus maculatus* fabricius, 1792), 2000. Dissertação (mestrado em bioquímica) - centro de ciências, universidade federal do ceará, fortaleza, 2000.
5. Campos, F. A. P.; Xavier-Filho, J.; Silva, C. P.; Ary, M. B. Resolution and partial characterization of proteinases and *a*amylases from midguts of larvae of the bruchid beetle *callo*sobruchus maculatus (f.). Comp. Biochemistry physiology, v. 92, p. 51-57, 1989.
6. Oliveira, G. B. Comparação entre a absorção de vicilinas de sementes suscetíveis e resistentes do feijão-de-corda, *vigna unguiculata* pelo epitélio intestinal de larvas do caruncho *Callosobruchus maculatus* e o processamento de vicilinas no corpo gorduroso após absorção – 2013. Dissertação (mestrado em bioquímica) - centro de ciências biológicas, universidade federal de santa catarina, florianópolis, 2013.
7. Credland, P.F. Effect of host availability on reproductive perfomance in *Callosobruchus maculatus* (f.) (coleoptera: bruchidae). Journal of stored products research, v. 22, p. 49-54, 1986.
8. Souza, S. M. Absorção e compartimentalização de vicilinas de *vigna unguiculata* por *callosobruchus maculatus* e *zabrotes subfasciatus* (coleoptera: chrysomelidae: bruquinae). 2009. Dissertação (mestrado em biociências e biotecnologia) - centro de biociências e biotecnologia, universidade estadual do norte fluminense, campos dos goytacazes, 2009.
9. Dick, K. M.; Credland, P. F. Chances in response of *Callosobruchus maculatus* (f.) (coleoptera: bruchidae) to a resistant variety of cowpea.journal of stored products research, v. 22, p. 227-233, 1986.
10. Dodeman, V. L.; Ducreux, G.; Kreis, M. Zygotic embryogenesis *versus* somatic embryogenesis.journal of experimental botany, v. 48, p.1493-1509, 1997.
11. Duranti, M.; Gius, C. Legume seeds: protein content and nutritional value. Field crops research, v. 53, p.31-45, 1997.
12. Edvardsson, M., Tregenza, T. Why do male callosobruchus maculatus harm their mates. Behavioural ecology, v.16, p. 788-793, 2005.
13. Gatehouse, A. M. R.; Gatehouse, J. A.; Dobie, P.; kilminster, a. M.; boulter, d. Biochemical basis of insect resistance in *vigna unguiculata*. Journal science food agriculture, v. 30, p. 948-958, 1979.
14. Macedo, M. R. L.; Andrade, L. B. S.; Moraes, R. A.; Xavier-Filho, J. Vicilin variants and the resistance of cowpea (*vigna unguiculata*) seeds to the cowpea weevil (*callosobruchus maculatus*).comparative biochemistry and physiology, v.105, p. 89-94, 1993.
15. Maia, F. M. M.; Oliveira, J. T. A.; Matos, M. R. T.; Moreira, R. A.; vasconcelos, I. M. Proximate composition aminoacid content and haemagglutinating and trypsin-inhibiting

activities of some brazilian vigna unguiculata (l) walp cultivars. Journal of the science of food and agriculture, v. 80, p. 453-458, 2000.
16. Maruyama, N.; Fukuda, T.; Saka, S.; Inui, N.; kotoh, J.; Miyagawa, M.; Hayashi, M.; Sawda, M.; Moriyama, T.; Utsumi, S. Molecular and structural analysis of electrophoretic variants of soybean seed storage proteins. Phytochemistry, v. 64, p. 701-708, 2003.
17. Oliveira, M. Z. A. Fungos associados à semente de caupi: identificação, patogenicidade e controle. 1981. Dissertação (mestrado em biologia molecular). Centro de ciências biológicas, universidade de brasília, brasília, 1981.
18. Osborn, T. C.; Alexander, D. C.; Sun, S. S.; Cardona, C.; Bliss, F. A. Insecticidal activity and lectin homology of arcelin seed protein. Science, p. 240, v. 207-210, 1988.
19. Padulosi, S.; NG, Q. N. Cowpea gene pool distribution and crop improvement. 1988. In: (ed.) Ehlers Id, hall, A. E. Cowpea (*vigna unguiculata* l. Walp.).fields crop res, v. 53, p. 187-204, 1997.
20. Paes, E. V.; Uchoa, A. F.; pinto, M. S. T.; Silva, C. P.; Fernandes, K. V. S.; Oliveira, A. E. A.; Xavier-Filho, J. Binding of *vigna unguiculata* vicilins to the peritrophicmembrane of *tenebrio molitor* affects larval development. Entomologia experimentalis et applicata, v. 129, p.11-17, 2008.
21. Pant, K. C.; Chandel, K. P. S.; Joshi, B. S. Analysis of diversity in indian cowpea genetic resourch, v. 14, p. 103-111, 1982.
22. Pedalino, M.; paino-d'urzo, M.; Delle Done, G.; Grillo, S.; Rao, R. The structure of cowpea (*vigna unguiculata* l. Walp.) Seed storage proteins.seed science and technology, v. 20, p. 223-231, 1992.

23. Rios, G.P. doenças fúngicas e bacterianas do caupi. In: Araujo, J.P.P; Watt E.E. o caupino brasil. Brasília: iita/embrapa, v. 1, p. 547-589, 1988.
24. Sales, M. P.; Fernandes, K. V. S.; Gomes, V. M.; Xavier-Filho, J. Chitin-binding proteins from cowpea (*vigna unguiculata*) seeds.brazilian journal of medical and biological research, v. 29, p. 319-326, 1996.
25. Sales, M. P.; Macedo, M. R. L.; Xavier-Filho, J. Digestibility of cowpea (*vigna unguiculata*) vicilins by pepsin, papain and bruchid midgut proteinases. Comparative biochemistry and physiology, v.103, p. 945-950, 1992.
26. Sales, M. P.; Pimenta, P. P.; Paes, N. S.; Grossi-de-Sá, M. F.; Xavier-Filho, J. Vicilins (7S storage globulins) of cowpea (*vigna unguiculata*) seeds bind to chitinious structures of the midgut of *callosobruchus maculatus* (coleoptera: bruchidae) larvae. Brazilian journal medical biology research, v. 34, p. 27–34, 2001.
27. Shewry, P. R.; Napier, J. A.; Tatham, A. S. Seed storage proteins: structures andbiosynthesis. The plant cell, v. 7, p. 945-956, 1995.
28. Shutov, A. D., Karhovskaya, I. A., Braun, H., Baümlein, H., Müntz, K. Legumin like and vicilin-like seed storage proteins: evidence for a common single-domain ancestral gene. Journal of molecular evolution, v. 41, p. 1057-1069, 1995.
29. Singh, B. B.; Ajeigbe, H. A.; Tarawali, S. A.; Fernandez-Rivera, S.; Abubakar, M. Improving the production and utilization of cowpea as food and fodder. Field crops research, v. 84, p. 169-177, 2003.
30. Singh, B. B.; singh, S. R.; Adjadi, O. Bruchid resistance in cowpea. Crop science, v. 25, p. 736-739, 1985.
31. Southgate, B. J. Biology of the bruchidae. *Annual review of entomology*, v. 24, p. 449-473, 1979.
32. Templematn, S.; Demaggioa, E.; Stetlerd, A. Biochemistry of fern spore germination: globulin storage proteins in *matteuccia struthiopteris* l. Plant physiology, v. 85, p. 343-349, 1987.

33. Wang, T. L.; Domoney, C.; Hedley, C. L.; Casey, R.; Grusak, M. A. Can we improve the nutritional quality of legume seeds? Plant physiology, v. 131, p. 886-891, 2003.

3. HIPÓTESES

I. Vicilinas (β-vigninas) de caupi devem possuir um ou mais sítios na sua estrutura, sítio esse envolvido na interação entre a proteína e cadeias de quitina;

II. Vicilinas de cultivares resistentes possivelmente possuem diferenças na sua estrutura em relação às vicilinas de cultivares susceptíveis;

III. Essas diferenças estruturais podem resultar em diferenças na força de interação com a quitina, o que poderia explicar, ao menos em parte, os efeitos contrastantes que essas proteínas de resersa causam em larvas de *C. maculatus*.

4. OBJETIVOS

Obter cDNAs codificando vicilinas de caupi, a partir de dois genótipos, um resistente (IT81D-1053) e outro susceptível (EPACE-10) ao C. maculatus, gerar modelos tridimensionais das proteínas (a partir das sequências deduzidas dos cDNAs) e usar simulações computacionais (docking e dinâmica molecular) para elucidar a localização dos sítios de ligação a quitina nas estruturas modeladas, analisando os tipos de ligações predominantes e os resíduos de aminoácidos envolvidos na interação entre as β-vigninas e oligômeros de NAG.

5. Cloning of cDNA sequences encoding cowpea (*Vigna unguiculata*) vicilins: molecular modeling, docking calculations and molecular dynamics simulations suggest a binding mode of cowpea vicilins to chitin oligomers

Antônio J. Rocha[a], Bruno L. Sousa[b], Matheus S. Girão[a], Ito L. Barroso-Neto[c], José E. Monteiro-Júnior[d], José T. A. Oliveira[a], Celso S. Nagano[e], Rômulo F. Carneiro[e], Ana C. O. Monteiro-Moreira[f], Bruno A. M. Rocha[a], Valder N. Freire[g], Thalles B. Grangeiro[d,*]

[a] Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará (UFC), Fortaleza, Ceará, Brazil

[b] Faculdade de Filosofia Dom Aureliano Matos, Universidade Estadual do Ceará, Av. Dom Aureliano Matos, 2060, Limoeiro do Norte, CE, 62930-000, Brazil

[c] Departamento de Química Analítica e Físico-química, UFC, Fortaleza, Ceará, 60455-760, Brazil

[d] Departamento de Biologia, UFC, Fortaleza, Ceará, 60440-900, Brazil

[e] Departamento de Engenharia de Pesa, UFC, Fortaleza, Ceará, Brazil

[f] Núcleo de Biologia Experimental, Universidade de Fortaleza, Fortaleza, Ceará, Brazil

[g] Departamento de Física, UFC, Fortaleza, Ceará, 60440-760, Brazil

*To whom all correspondence should be addressed

Phone: +558533669809; Fax: +558533669806

E-mail: tbgrangeiro@gmail.com

Submitted for International Journal of Biological Macromolecules

6. Cloning of cDNA sequences encoding cowpea (*Vigna unguiculata*) vicilins: molecular modeling, docking calculations and molecular dynamics simulations suggest a binding mode of cowpea vicilins to chitin oligomers

Antônio J. Rocha[a], Bruno L. Sousa[b], Matheus S. Girão[a], Ito L. Barroso-Neto[c], José E. Monteiro-Júnior[d], José T. A. Oliveira[a], Celso S. Nagano[e], Rômulo F. Carneiro[e], Ana C. O. Monteiro-Moreira[f], Bruno A. M. Rocha[a], Valder N. Freire[g], Thalles B. Grangeiro[d,*]

[a] Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará (UFC), Fortaleza, Ceará, Brazil

[b] Faculdade de Filosofia Dom Aureliano Matos, Universidade Estadual do Ceará, Av. Dom Aureliano Matos, 2060, Limoeiro do Norte, CE, 62930-000, Brazil

[c] Departamento de Química Analítica e Físico-química, UFC, Fortaleza, Ceará, 60455-760, Brazil

[d] Departamento de Biologia, UFC, Fortaleza, Ceará, 60440-900, Brazil

[e] Departamento de Engenharia de Pesa, UFC, Fortaleza, Ceará, Brazil

[f] Núcleo de Biologia Experimental, Universidade de Fortaleza, Fortaleza, Ceará, Brazil

[g] Departamento de Física, UFC, Fortaleza, Ceará, 60440-760, Brazil

*To whom all correspondence should be addressed

Phone: +558533669809; Fax: +558533669806

E-mail: tbgrangeiro@gmail.com

Abstract

Vicilins are 7S globulins which constitute the major seed storage proteins in leguminous species. The ability of cowpea vicilins to bind chitin has been implicated in the resistance of some genotypes to the cowpea weevil (*Callosobruchus maculatus*). Vicilins belong to the cupin superfamily, a conserved β-barrel fold that is found in a wide variety of enzymes as well as in non-catalytic seed storage proteins. The cupin fold does not share similarity with any known chitin-biding domain, like those found in hevein-like proteins and chitin-binding lectins. Therefore, it is poorly understood how these storage proteins bind to chitin oligomers and hence affect the larval development of bruchid beetles. In this work, partial cDNA sequences encoding cowpea vicilins were obtained from developing seeds from genotypes EPACE-10 and IT81D-1053. Three-dimensional molecular models of cowpea vicilins, obtained by homology modeling, showed the characteristic cupin fold. Computational simulations revealed that each vicilin trimer contained 3 chitin-binding sites, each site located at the vertex of the triangle-shaped oligomer. Interaction models predicted that docked carbohydrate molecules were stabilized mainly by hydrogen bonds, a common structural feature observed in a wide variety of carbohydrate-binding proteins. Furthermore, many of the residues involved in the chitin-binding sites of cowpea vicilins are conserved in other 7S globulins. The results of the present work support previous experimental evidences on the ability of vicilin-like proteins from cowpea and other leguminous species to bind *in vitro* to chitin as well as *in vivo* to chitinous structures of *C. maculatus* larvae's midgut.

Keywords

Seed storage proteins; Leguminosae; *Callosobruchus maculatus*; chitin-binding; plant defense

1. Introduction

Cowpea [*Vigna unguiculata* (L.) Walp.] is an important food crop that is cultivated in arid and semi-arid regions of Africa, Asia and the Americas (Ehlers and Hall, 1997). This leguminous species is adapted to high temperatures, drought and other abiotic stresses, and their dry seeds constitute a valuable source of proteins and calories for human consumption, especially for smallholder farmers (Singh et al., 2003). One major constraint to cowpea production in developing countries is the attack by the bruchid beetle *Callosobruchus maculatus* (F.), commonly known as the cowpea weevil, which is a cosmopolitan pest of stored cowpeas. Fertilized females of *C. maculatus* lay their eggs on the surface of the seeds, usually a single egg per seed. Eggs hatch after 5-6 days of oviposition and the first-instar larvae burrow directly into the seed. Each developing larva feeds within a single seed, excavating a chamber as it grows. All larval stages and pupation occur within a single seed. At 25 °C and 70% relative humidity, adults emerge approximately 36 days after the eggs were laid (Howe and Currie, 1964). Adults live on average 12-14 days, and during this time, they mate and the fertilized females lay eggs on undamaged seeds. Infestation of this bruchid species starts in the field and continues in storage, causing sometimes the complete destruction of seeds within a period of 3-4 months through secondary infestation (Murdock et al., 2003). Damaged seeds are unsuitable for human or animal consumption, have a decreased germination potential, lower nutritional qualities and a reduction in commercial value (Ojimelukwe and Ogwumike, 1999) (Melo et al., 2010). Due to financial and technical limitations, insect-resistant varieties would be the most effective method of controlling the cowpea weevil in northeastern Brazil and other underdeveloped regions of the world.

In the 1970's, researchers at the International Institute of Tropical Agriculture (IITA) in Ibadan, Nigeria, initiated a systematic screening of over 8,000 germplasm accessions of cowpea, aiming to develop bruchid resistant varieties. These efforts allowed the identification of three cowpea accessions, TVu2027, TVu11952 and TVu11953, which showed moderately seed resistance to

C. maculatus (Singh et al., 1985). These accessions showed similar levels of resistance, in which the percentage of adult emergence was delayed, staggered and lower in comparison to the same parameters obtained for insects reared on susceptible seeds. Using these resistant genotypes, especially TVu 2027, several improved breeding lines combining resistance to *C. maculatus* with other desirable traits have been developed (Singh and Singh, 1990). Gatehouse et al. were the first to hypothesize that resistance to cowpea weevil in TVu 2027 was due to elevated levels of trypsin inhibitors (Gatehouse et al., 1979) (Gatehouse and Boulter, 1983). However, this claim was not supported by further experimental evidences, which have shown that the levels of proteinase inhibitors targeting different types of *C. maculatus* midgut proteases or even α-amylase inhibitors are not correlated with the resistance or susceptibility to bruchid infestation (Baker et al., 1989) (Xavier-Filho et al., 1989) (Fernandes et al., 1993) (Reis et al., 1997). On the other hand, some authors have shown that cowpea vicilins purified from resistant lines derived from TVu 2027 are detrimental to *C. maculatus* larvae, and these vicilin variants are more refractory to digestion by the bruchid's midgut proteinases in comparison to vicilins from susceptible lines (Sales et al., 1992) (Macedo et al., 1993). Furthermore, it has been shown that cowpea vicilins bind *in vitro* to chitin (Sales et al., 1996) and *in vivo* to chitinous structures of the midgut of *C. maculatus* larvae (Firmino et al., 1996) (Sales et al., 2001). Binding of cowpea vicilins to larvae's midgut epithelium cell surface leads to the absorption of intact molecules and their transport into the haemolymph, fat body cells and malpighian tubules (Uchôa et al., 2006). Vicilin molecules are detected in these organs in all larval stages and pupae, and vicilin-derived peptides are also found in fat bodies of male and female adults, even 10 days after emergence (Souza et al., 2010). Recent data have demonstrated that internalization of vicilin molecules in the enterocytes of *C. maculatus* larvae is accomplished predominantly through receptor-mediated, clathrin-independent endocytosis and vesicular trafficking in the cytoplasm (Oliveira et al., 2014) (Kunz et al., 2017). These findings have lead to the hypothesis that once internalized, variant vicilins disturb the normal physiology of *C. maculatus*

enterocytes, adversely affecting the larval development. In this proposed mechanism, binding to chitin-containing structures in the larvae midgut is a crucial step, that ultimately provokes the toxic effects of cowpea vicilins on developing insects.

Vicilins are 7S seed storage globulins that contain two copies of the cupin superfamily domain. The cupin domain is characterized by a double-stranded β-helix fold, which is found in diverse enzymes as well as in non-catalytic seed storage proteins (Dunwell et al., 2004). However, there is no structural similarity between the cupin domain and known chitin-binding proteins, such as hevein and chitin-binding lectins, for example. These canonical chitin-binding proteins have typical carbohydrate-binding sites that recognize and bind one or more residues of *N*-acetyl-D-glucosamine, the repeating unit of chitin. Therefore, the molecular mechanism involved in the interaction between cowpea vicilins and chitin oligomers remains poorly understood. In the present work, partial cDNA sequences encoding cowpea vicilins were obtained from genotypes EPACE-10 and IT81D-1053. Three-dimensional molecular models were generated, validated and subjected to molecular docking calculations and molecular dynamics simulations, aiming to reveal the putative chitin-binding sites in the modeled vicilin structures.

2. Materials and methods

2.1 Plant material

Cowpea seeds (genotypes IT81D-1053 and EPACE-10) were kindly provided by F. R. Freire-Filho (Embrapa Meio-Norte, Teresina-PI, Brazil) and E. M. Teofilo (Centro de Ciências Agrárias, UFC, Fortaleza-CE, Brazil), respectively.

2.2 Plasmid, bacterial strain and reagents

The plasmid pGEM-T Easy and cells of *Escherichia coli* strain TOP10F′ were purchased from Promega (Madison, WI, USA) and Invitrogen (Carlsbad, CA, USA), respectively. All other reagents were of analytical grade.

2.3 RNA purification, cDNA synthesis, amplification and cloning of PCR products

Total RNA was purified from developing seeds (harvested 12 days after pollination), using the method described by Chang et al. (Chang et al., 1993). Conversion of total RNA to DNA was performed as previously described (Maranhão et al., 2017). First-strand cDNA products were then amplified by polymerase chain reaction (PCR) using the following oligonucleotide primers: 5′-ATTGTACACCGGGAGCACCAAG-3′ (forward) and 5′-GTAGARASTGYCCAAAATWGAAGATAA-3′ (reverse). These primers were designed according to expressed sequence tags (ESTs) from cowpea, deposited in the NCBI's EST database (www.ncbi.nlm.nih.gov/dbEST/). The ESTs that encoded amino acid sequences matching the N- and C-terminal ends of characterized vicilins from related leguminous species were used to design the forward and reverse primers, respectively. Amplification reactions and cloning of PCR products into the pGEM-T Easy vector were performed as previously described (Maranhão et al., 2017).

2.4 DNA sequencing and sequence analysis

DNA sequencing was performed at the Macrogen Inc. (Seoul, South Korea) using the Sanger's dideoxy chain termination method. DNA sequencing and sequence assembly were done as described elsewhere (Maranhão et al., 2017). Nucleotide sequences were translated to amino acid sequences using the ExPASy Translate tool (web.expasy.org/translate/). Manipulation, edition and alignment of sequences were performed using the program BioEdit v. 7.2.5 (Hall, 1999). Searches for homologous sequences in public databases were performed using BLAST (Altschul et al., 1990).

2.5 Purification of cowpea vicilins

Cowpea vicilins were extracted from seed flour according to the protocol first described by Samour et al. (Sammour et al., 1984) and including the modifications described by Macedo et al. (Macedo et al., 1993). The preparations were further purified by size exclusion chromatography on Sephacryl S-200 and ion exchange chromatography on DEAE-Sepharose. Protein fractions were analyzed by SDS-polyacrylamide gel electrophoresis according to the method of Laemmli (Laemmli, 1970), using 15% slab gels. Staining and destaining of protein bands were performed as described by Lobo et al. (Lobo et al., 2013). N-terminal amino acid sequencing of protein bands resolved by SDS-PAGE was performed as described by Landim et al. (Landim et al., 2017).

2.6 Mass spectrometry analysis

Proteins were resolved by SDS-PAGE and electrophoretic bands were subjected to in-gel digestion with trypsin (Promega), according to the protocol previously described (Shevchenko et al., 2006). Identification of tryptic peptides by tandem mass spectrometry (LC-ESI-MS/MS) was performed using a Synapt G1 HDMS Q-ToF mass spectrometer (Waters Co., Milford, MA, USA) coupled to a Waters ultra-high-performance liquid chromatography (UPLC) unit. Data

processing, MS/MS ions search and peptide identification were done as described in detail by Freire et al. (Freire et al., 2015). Tryptic peptides were also subjected to MALDI-TOF/TOF MS analysis (AutoFlex III mass spectrometer, Bruker Daltonics, USA). The spectrometer was operated in the reflector mode for MALDI-TOF MS peptide mass fingerprint (PMF), at m/z range of 700 to 4000, and in the "LIFT™" mode for MALDI-TOF/TOF MS/MS fragmentation experiments (at m/z range of precursor). Spectra were processed with Flex Analysis 3.4 software. PMF spectra were compared with *in silico* digestions of vicilin sequences using PEPTIDEMASS through the ExPASy server (Gasteiger et al., 2005). MS/MS spectra were interpreted manually. The ambiguities for isobaric amino acids were approached by amino acid sequences deduced from cDNA. Sequenced peptides were searched online against NCBI and UniProt databases.

2.7 Molecular modeling

Three-dimensional molecular models were generated using MODELLER (Sali et al., 1995) through the MPI Bioinformatics Toolkit (Alva et al., 2016) web service (toolkit.tuebingen.mpg.de). The crystal structure of the 7S globulin-1 (PDB ID: 2EA7) from adzuki bean (*V. angularis*) was used as template (Fukuda et al., 2008). Models generated using MODELLER were refined using GalaxyRefine (Heo et al., 2013) through the program's web server. Refined models were validated concerning their stereo-chemical properties (Ramachandran plots, steric overlaps, Cβ deviation parameters, rotamers, and bond angle quality) using MolProbity (Chen et al., 2010) through the software's web server (molprobity.biochem.duke.edu/).

2.8 Molecular docking calculations

A standard docking procedure was performed with AutoDock Vina v1.1.2 (Trott and Olson, 2010). Initially, a blind docking strategy was performed using a search space defined by a 40 Å × 40 Å × 40 Å cube, which was applied to three different portions of each vicilin monomer,

thus covering the whole protein surface. The ligands were $(GlcNAc)_4$ and $(GlcNAc)_5$, which were built using the program SWEET (Bohne et al., 1999). All of the torsional bonds of the ligands were free to rotate while the protein atoms were held rigid. Polar hydrogen atoms were added using the AutoDock Tools version 1.5.6, and Kollman united atom partial charges were assigned (Morris et al., 2009). Afterwards, a refined search with the same ligands was performed using a search space defined by a 20 Å × 20 Å × 20 Å cube, based on the oligosaccharides length, centered on the top-ranked chitin binding site. Exhaustiveness was set to 15, and for all other parameters, default values were used. Selection of the best results was done as previously described (Maranhão et al., 2017).

2.9 Molecular dynamics (MD) simulations

MD simulations (30 ns) were performed using the program NAMD v2.10 (Nanoscale Molecular Dynamics) (Phillips et al., 2005), with the force fields CHARMM27 (Best et al., 2012) and CHARMM36 (Guvench et al., 2011). The software VMD (Visual Molecular Dynamics) was used for file preparation (Humphrey et al., 1996). All MD simulations were performed essentially as described in detail by Maranhão et al. (Maranhão et al., 2017).

2.10 Binding energy calculations

Binding energy calculations were done using a dispersion-corrected density functional theory (DFT-D) method (Delley, 2000), using the procedure described in detail by Maranhão et al. (Maranhão et al., 2017).

3. Results and discussion

3.1 Cloning of partial cDNA fragments encoding β-vignin, the major vicilin from cowpea

PCR products with approximately 1,200 bp that presumptively encoded vignin were amplified from total cDNA, which was obtained from developing seeds of cowpea genotypes EPACE-10 and IT81D-1053 (Fig S1). Several clones were sequenced and 3 unique partial cDNA sequences from each genotype were identified. These sequences are herein referred to as S1, S2 and S3 from EPACE-10, which is susceptible to *C. maculatus*, and R1, R2 and R3 from IT81D-1053, which is resistant to *C. maculatus*. All cDNA sequences were 1,296 nucleotides long, except sequence S3, which had a length of 1,287 nucleotides (Figs. S2-S7). Sequence identity ranged from 99.9% (1 different nucleotide), between S1 and R1, to 98.6% (17 distinct nucleotides), between R2 and R3 (Table S1). Sequence S3 encoded a polypeptide chain with 429 amino acid residues, whereas the other 5 sequences encoded proteins with 432 residues. The shorter sequence S3 was due to a deletion of 9 nucleotides (Fig. S8), encoding a stretch of 3 amino acid residues (^{193}QDE^{195} in the other 5 sequences) (Fig. S9). Pairwise comparisons revealed that the amino acid sequences of S1 and R1 were identical, whereas the other pairs of compared structures had differences, ranging from 3 to 11 residues (Table S2). Molecular masses calculated from these amino acid sequences varied from 49347.02 (S3) to 49827.45 Da (S1 and R1) (Table S3). BLASTp searches against the NCBI protein database revealed that the 6 amino acid sequences had highest similarity with 7S globulins from Leguminosae species, such as adzuki bean (*V. angularis*; ~86%), mung bean (*V. radiata*; ~86%), common bean (*Phaseolus vulgaris*; ~64%) and soybean (*Glycine max*; ~68%). Moreover, searches against the CDD showed that each protein contained two domains, both belonging to the cupin_1 family (SMART accession number: SM00835) of the cupin superfamily (Table S4). In the 6 cowpea proteins, the N- and C-terminal cupin_1 domains were 147 and 157 residues long, respectively. The cupin_1 family, which represents the conserved barrel domain of the cupin superfamily, contains 11S (legumins) and 7S (vicilins) seed storage globulins and germins (Dunwell, 1998).

Legumins and vicilins are two-domain proteins (bicupins), whereas germins are single-domain molecules (monocupins). This analysis showed that the cDNA sequences from cowpea encoded 6 proteins that have typical structural features of 7S seed storage globulins (Fig. 1).

To verify the relationship between the products encoded by the cDNA sequences and cowpea seed storage proteins, a fraction enriched in 7S globulins was obtained from mature seeds of both genotypes. When subjected to SDS-PAGE, the vicilin fractions from EPACE-10 and IT81D-1053 showed very similar profiles (Fig. S10). Each pattern was characterized by 11 protein bands, with apparent molecular masses ranging from 110 to 24 kDa, in which the bands with 60 and 54 kDa were the most abundant ones in each genotype. Almost all resolved bands (21 out of 22) were identified as 7-8S globulins by ESI-MS/MS (Tables S5 and S6). These results agree with previous works, which have demonstrated that under denaturing and reducing conditions, cowpea vicilins are a heterogeneous mixture of polypeptides of various sizes (Cerdeira et al., 1985) (Khan et al., 1980) (Murray et al., 1983). When these vicilin fractions were subjected to size exclusion chromatography, one main peak was obtained (PII; Fig. S11), which was further resolved in two peaks (PI and PII) by ion exchange chromatography (Fig. S12). Each major peak from the ion exchange chromatography showed a similar pattern of polypeptides, when analyzed by SDS-PAGE, and both were enriched in the 60 and 54 kDa bands, which are characteristic of β-vignin, the major component of cowpea globulins (Freitas et al., 2004). When these 2 major bands were subjected to Edman degradation, the same N-terminal amino acid sequence (26 residues) was obtained from both β-vignin bands of each genotype: IVHREHQESQEESEPRGQNNPFYFDS. This sequence matched exactly the first 26 amino acid residues deduced from the 6 partial cDNA fragments, which were obtained from developing seeds of cowpea. The average molecular mass calculated from these 6 sequences was, approximately, 49.7 kDa, which is closer to the faster migrating β-vignin subunit. The differences between the values calculated from the sequences and those determined by SDS-PAGE are probably due to glycosylation. Indeed, some authors have demonstrated that the 2

main β-vignin polypeptides are both glycosylated, the 60 kDa subunit having a greater amount of carbohydrate than the 54 kDa subunit (Freitas et al., 2004) (de Souza Ferreira et al., 2018). To further strength the relationship between the sequences deduced from the cDNA fragments and the seed proteins, several tryptic peptides were identified by MALDI-TOF MS analysis from the β-vignin peaks resolved by ion exchange chromatography (Figs. S14-S18). The sequences of these peptides matched specific segments of the primary structures deduced from the cDNA fragments obtained from IT81D-1053 (Table S7) and EPACE-10 (Table S8). The experimentally determined peptides covered approximately 23.4 and 17.6% of the amino acid sequences of the proteins from EPACE1-10 and IT81D-1053, respectively. Therefore, the cDNA fragments cloned from EPACE-10 and IT81D-1053 encode β-vignin, which is the main fraction of the 7S globulins stored in cowpea seeds.

3.2 Molecular models of β-vignin and their interaction with chitin oligomers

To investigate how cowpea vicilins bind chitin, three-dimensional molecular models were generated, refined, validated and subjected to computational simulations using chito-oligosaccharides as potential ligands. Vicilin models were built from 4 amino acid sequences (S2 and S3 from EPACE-10 and R2 and R3 from IT81D-1053), using the three-dimensional structure of the 7S globulin-1 (PDB ID: 2EA7) from adzuki bean (*V. angularis*) as template (Figs. S19-S22). When the modeled structures (designated as models S2, S3, R2 and R3) were aligned, an RMSD of 1.479 Å was obtained (Fig. S23), indicating that the models had very similar folds. Furthermore, superposition of the model and template structures gave the following TM-score values: 0.7224 (model S2 *vs* 2EA7), 0.5013 (model S3 *vs* 2EA7), 0.7245 (model R2 *vs* 2EA7) and 0.7199 (model R3 *vs* 2EA7). This analysis showed that the overall fold of the modeled structures was correct (Fig. 2), as indicated by the TM-score values greater than 0.5 (Zhang and Skolnick, 2004). In the three-dimensional structures of the β-vignin models, the N- and C-terminal halves of the polypeptide chain assumed the characteristic cupin

fold, the domains being symmetrically related by a pseudo-twofold axis, and each one containing a core β-barrel and an extended loop region with 2 helices (Fig. 2), as typically found in 7S globulins (Fukuda et al., 2008). After refinement, the percentage of Ramachandran outliers was either 0 or less than 1%. Clashscore, which is the number of clashes ≥0.4 Å per 1000 atoms, ranged from 6.65 (88th percentile) to 13.52 (56th percentile), whereas MolProbity score varied from 1.38 (97th percentile) to 1.71 (89th percentile) (Tables S9-S12). MolProbity score is a log-weighted combination of the clashscore, percentage of Ramachandran outliers and percentage of bad side-chain rotamers, providing one value that reflects the approximate crystallographic resolution at which that combination of numbers would be expected (Chen et al., 2010). Therefore, the molecular models of cowpea vicilins were judged as of good quality, according to MolProbity validation criteria.

In native 7S globulins, monomers assemble into trimers, in which the subunits are held by non-covalent forces. These trimers have a triangle-like shape, arising through head-to-tail association of individual monomers (Fukuda et al., 2008). Symmetry operations were used to generate β-vignin homotrimers, and these oligomers were subjected to computational simulations. Initial molecular docking calculations suggested that each trimer of β-vignin had 3 sites that could bind chito-oligosaccharides. Each binding site, located in the vertex of the triangle-shaped oligomer, was constituted by some residues from one subunit and other residues from its neighbor chain (Fig. 3). To verify the stability of a tetra-*N*-acetyl-chitotetraose molecule docked in a single binding site of cowpea β-vignins, molecular dynamics simulations were performed. RMSD plots showed that, after an initial accommodation that took place in the first nanoseconds, the interaction of the bound oligosaccharide was very stable along most simulations, especially in the complexes between $(GlcNAc)_4$ and the models S3 and R2 (Fig. S24). In the stable complexes obtained after MD simulations, the interaction between amino acid residues and ligand atoms was mediated mainly by hydrogen bonds and hydrophobic contacts (Fig. 4). In each complex, a network of hydrogen bonds was observed, involving

hydroxyl groups from each GlcNAc unit of the bound oligosaccharide and several O and N atoms from the side chains of some residues of the modeled protein structure. Many of these polar interactions between the chito-oligosaccharide and β-vignin were predicted to be mediated by water molecules (Fig. 5). Carbohydrate-binding proteins, including a large diversity of enzymes as well as non-enzymatic proteins, such as lectins, have wide differences in their three-dimensional structures and carbohydrate-binding topologies. However, a common feature of the interactions between proteins and their carbohydrate ligands is the prevalence of hydrogen bond interactions, usually involving water molecules (Quiocho, 1986). Most of the β-vignin amino acid residues, involved in the interactions with the bound oligosaccharide in each chitin binding-site (ChBS), were mapped to 4 main segments in the protein's primary structure (Fig. 1). One of these segments was mapped in the N-terminal domain, whereas the other 3 regions were constituted by residues found in the C-terminal domain. Some of the residues from these 4 regions that were involved in the interaction between β-vignin and chito-oligosaccharides are conserved in other 7S globulins. Using a quantum mechanics-based method, the binding free energies of chito-oligosaccharide molecules to β-vignin oligomers were calculated, ranging from −690.6 kcal/mol (GlcNAc oligomers bound to β-vignin R2 trimer) to −531.3 kcal/mol (chitin fragments bound to β-vignin R3 oligomer) (Table S13). These results agree with previous works, which have demonstrated that cowpea vicilins purified from seeds of genotypes that are susceptible or resistant to *C. maculatus* have both the ability to bind *in vitro* to chitin (Sales et al., 1996) and *in vivo* to chitinous structures of the larvae's midgut (Firmino et al., 1996) (Sales et al., 2001). Furthermore, the conservation in other 7S globulins of many of the residues that constitute the β-vignin's ChBS, as predicted in this work, also support earlier reports that have shown that vicilins from other leguminous species can also bind to chitin (Macedo et al., 2008).

4. Conclusions

Partial cDNA sequences encoding β-vignin, the main component of the 7S globulins of cowpea, were obtained. Computational simulations were then used to predict the probable chitin-binding site in the modeled structures. Interaction models suggested that chito-oligosaccharides bound to β-vignin are stabilized mainly by hydrogen bonds, a typical structural feature of many carbohydrate-binding proteins.

Acknowledgements

This work was supported by research grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP).

5-References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Alva, V., Nam, S.-Z., Söding, J., Lupas, A.N., 2016. The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. Nucleic Acids Res. 44, W410–W415. https://doi.org/10.1093/nar/gkw348

Baker, T.A., Suzanne, N., Shade, R.E., Singh, B.B., 1989. Physical and chemical attributes of cowpea lines resistant and susceptible to Callosobruchus maculatus (F.) (Coleoptera: Bruchidae). J. Stored Prod. Res. 25, 1–8. https://doi.org/10.1016/0022-474X(89)90002-7

Best, R.B., Zhu, X., Shim, J., Lopes, P.E.M., Mittal, J., Feig, M., Mackerell, A.D., 2012. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. J. Chem. Theory Comput. 8, 3257–3273. https://doi.org/10.1021/ct300400x

Bohne, A., Lang, E., von der Lieth, C.W., 1999. SWEET - WWW-based rapid 3D construction of oligo- and polysaccharides. Bioinformatics 15, 767–768.

Bond, C.S., Schüttelkopf, A.W., 2009. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. Acta Crystallogr. D Biol. Crystallogr. 65, 510–512. https://doi.org/10.1107/S0907444909007835

Cerdeira, A.L., Cole, A.W., Luthe, D.S., 1985. Cowpea (Vigna unguiculata) Seed Protein Response to Glyphosate. Weed Sci. 33, 1–6. https://doi.org/10.1017/S0043174500083843

Chang, S., Puryear, J., Cairney, J., 1993. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Report. 11, 113–116. https://doi.org/10.1007/BF02670468

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21. https://doi.org/10.1107/S0907444909042073

de Souza Ferreira, E., Capraro, J., Sessa, F., Magni, C., Demonte, A., Consonni, A., Augusto Neves, V., Maffud Cilli, E., Duranti, M., Scarafoni, A., 2018. New molecular features of cowpea bean (Vigna unguiculata, l. Walp) β-vignin. Biosci. Biotechnol. Biochem. 82, 285–291. https://doi.org/10.1080/09168451.2017.1419855

Delley, B., 2000. From molecules to solids with the DMol3 approach. J. Chem. Phys. 113, 7756–7764. https://doi.org/10.1063/1.1316015

Dunwell, J.M., 1998. Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. Biotechnol. Genet. Eng. Rev. 15, 1–32.

Dunwell, J.M., Purvis, A., Khuri, S., 2004. Cupins: the most functionally diverse protein superfamily? Phytochemistry 65, 7–17.

Ehlers, J.D., Hall, A.E., 1997. Cowpea (Vigna unguiculata L. Walp.). Field Crops Res. 53, 187–204. https://doi.org/10.1016/S0378-4290(97)00031-2

Fernandes, K.V., Sabelli, P.A., Barratt, D.H., Richardson, M., Xavier-Filho, J., Shewry, P.R., 1993. The resistance of cowpea seeds to bruchid beetles is not related to levels of cysteine proteinase inhibitors. Plant Mol. Biol. 23, 215–219.

Firmino, F., Fernandes, K.V.S., Sales, M.P., Gomes, V.M., Miranda, M.R.A., Domingues, S.J.S., Xavier, J., 1996. Cowpea (Vigna unguiculata) vicilins associate with putative chitinous structures in the midgut and feces of the bruchid beetles Callosobruchus maculatus and Zabrotes subfasciatus. Braz. J. Med. Biol. Res. 29, 749–756.

Freire, J.E.C., Vasconcelos, I.M., Moreno, F.B.M.B., Batista, A.B., Lobo, M.D.P., Pereira, M.L., Lima, J.P.M.S., Almeida, R.V.M., Sousa, A.J.S., Monteiro-Moreira, A.C.O., Oliveira, J.T.A., Grangeiro, T.B., 2015. Mo-CBP3, an Antifungal Chitin-Binding

Protein from Moringa oleifera Seeds, Is a Member of the 2S Albumin Family. PloS One 10, e0119871. https://doi.org/10.1371/journal.pone.0119871

Freitas, R.L., Teixeira, A.R., Ferreira, R.B., 2004. Characterization of the proteins from Vigna unguiculata seeds. J. Agric. Food Chem. 52, 1682–1687. https://doi.org/10.1021/jf0300588

Fukuda, T., Maruyama, N., Salleh, M.R.M., Mikami, B., Utsumi, S., 2008. Characterization and crystallography of recombinant 7S globulins of Adzuki bean and structure-function relationships with 7S globulins of various crops. J. Agric. Food Chem. 56, 4145–4153. https://doi.org/10.1021/jf072667b

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R., Bairoch, A., 2005. Protein Identification and Analysis Tools on the ExPASy Server, in: The Proteomics Protocols Handbook. Humana Press, Totowa, New Jersey, pp. 571–607.

Gatehouse, A.M.R., Boulter, D., 1983. Assessment of the antimetabolic effects of trypsin inhibitors from cowpea (Vigna unguiculata) and other legumes on development of the bruchid beetle Callosobruchus maculatus. J. Sci. Food Agric. 34, 345–350. https://doi.org/10.1002/jsfa.2740340405

Gatehouse, A.M.R., Gatehouse, J.A., Dobie, P., Kilminster, A.M., Boulter, D., 1979. Biochemical basis of insect resistance in Vigna unguiculata. J. Sci. Food Agric. 30, 948–958. https://doi.org/10.1002/jsfa.2740301003

Guvench, O., Mallajosyula, S.S., Raman, E.P., Hatcher, E., Vanommeslaeghe, K., Foster, T.J., Jamison, F.W., Mackerell, A.D., 2011. CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. J. Chem. Theory Comput. 7, 3162–3180. https://doi.org/10.1021/ct200328p

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Heo, L., Park, H., Seok, C., 2013. GalaxyRefine: Protein structure refinement driven by side-chain repacking. Nucleic Acids Res. 41, W384–W388. https://doi.org/10.1093/nar/gkt458

Howe, R.W., Currie, J.E., 1964. Some laboratory observations on the rates of development, mortality and oviposition of several species of Bruchidae breeding in stored pulses. Bull. Entomol. Res. 55, 437–477. https://doi.org/10.1017/S0007485300049580

Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: visual molecular dynamics. J. Mol. Graph. 14, 33–38, 27–28.

Khan, M.R.I., Gatehouse, J.A., Boulter, D., 1980. The Seed Proteins of Cowpea (Vigna unguiculata L. Walp.). J. Exp. Bot. 31, 1599–1611.

Kunz, D., Oliveira, G.B., Uchôa, A.F., Samuels, R.I., Macedo, M.L.R., Silva, C.P., 2017. Receptor mediated endocytosis of vicilin in Callosobruchus maculatus (Coleoptera: Chrysomelidae) larval midgut epithelial cells. Comp. Biochem. Physiol. - Part B Biochem. Mol. Biol. 210, 39–47. https://doi.org/10.1016/j.cbpb.2017.06.003

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685.

Landim, P.G.C., Correia, T.O., Silva, F.D.A., Nepomuceno, D.R., Costa, H.P.S., Pereira, H.M., Lobo, M.D.P., Moreno, F.B.M.B., Brandão-Neto, J., Medeiros, S.C., Vasconcelos, I.M., Oliveira, J.T.A., Sousa, B.L., Barroso-Neto, I.L., Freire, V.N., Carvalho, C.P.S., Monteiro-Moreira, A.C.O., Grangeiro, T.B., 2017. Production in Pichia pastoris, antifungal activity and crystal structure of a class I chitinase from cowpea (Vigna unguiculata): Insights into sugar binding mode and hydrolytic action. Biochimie 135, 89–103. https://doi.org/10.1016/j.biochi.2017.01.014

Laskowski, R.A., Swindells, M.B., 2011. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J. Chem. Inf. Model. 51, 2778–2786. https://doi.org/10.1021/ci200227u

Lobo, M.D., Silva, F.D., Landim, P.G., da Cruz, P.R., de Brito, T.L., de Medeiros, S.C., Oliveira, J.T., Vasconcelos, I.M., Pereira, H.D., Grangeiro, T.B., 2013. Expression and efficient secretion of a functional chitinase from Chromobacterium violaceum in Escherichia coli. BMC Biotechnol. 13, 46. https://doi.org/10.1186/1472-6750-13-46

Macedo, L.L.P., Amorim, T.M.L., Uchôa, A.F., Oliveira, A.S., Ribeiro, J.K.C., de Macedo, F.P., Santos, E.A., de Sales, M.P., 2008. Larvicidal effects of a chitin-binding vicilin from Erythrina velutina seeds on the mediterranean fruit fly Ceratitis capitata. J. Agric. Food Chem. 56, 802–808. https://doi.org/10.1021/jf072746n

Macedo, M.L.R., Da Andrade, S.L.B., Moraes, R.A., Xavier-Filho, J., 1993. Vicilin variants and the resistance of cowpea (Vigna unguiculata) seeds to the cowpea weevil (Callosobruchus maculatus). Comp. Biochem. Physiol. - Part C Toxicol. Pharmacol. 105, 89–94. https://doi.org/10.1016/0742-8413(93)90063-Q

Maranhão, P.A.C., Teixeira, C.S., Sousa, B.L., Barroso-Neto, I.L., Monteiro-Júnior, J.E., Fernandes, A.V., Ramos, M.V., Vasconcelos, I.M., Gonçalves, J.F.C., Rocha, B.A.M., Freire, V.N., Grangeiro, T.B., 2017. cDNA cloning, molecular modeling and docking calculations of L-type lectins from Swartzia simplex var. grandiflora (Leguminosae, Papilionoideae), a member of the tribe Swartzieae. Phytochemistry 139, 60–71. https://doi.org/10.1016/j.phytochem.2017.04.007

Melo, R.A., Forti, V.A., Cicero, S.M., Novembre, A.D.L.C., de, M., 2010. Use of X-ray to evaluate damage caused by weevils in cowpea seeds. Hortic. Bras. 28, 472–476. https://doi.org/10.1590/S0102-05362010000400016

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785–2791. https://doi.org/10.1002/jcc.21256

Murdock, L.L., Seck, D., Ntoukam, G., Kitch, L., Shade, R.E., 2003. Preservation of cowpea grain in sub-Saharan Africa - Bean/Cowpea CRSP contributions. Field Crops Res. 82, 169–178. https://doi.org/10.1016/S0378-4290(03)00036-4

Murray, D.R., Mackenzie, K.F., Vairinhos, F., Peoples, M.B., Atkins, C.A., Pate, J.S., 1983. Electrophoretic Studies of the Seed Proteins of Cowpea, Vigna unguiculata (L.) Walp. Z. Für Pflanzenphysiol. 109, 363–370. https://doi.org/10.1016/S0044-328X(83)80118-4

Ojimelukwe, P.C., Ogwumike, F.C., 1999. Effects of infestation by bruchid beetles (Callosobruchus macultus) on the nutritional quality and sensory properties of cowpeas (Vigna unguiculata). J. Food Biochem. 23, 637–645.

Oliveira, G.B., Kunz, D., Peres, T.V., Leal, R.B., Uchôa, A.F., Samuels, R.I., Macedo, M.L.R., Carlini, C.R., Ribeiro, A.F., Grangeiro, T.B., Terra, W.R., Xavier-Filho, J., Silva, C.P., 2014. Variant vicilins from a resistant Vigna unguiculata lineage (IT81D-1053) accumulate inside Callosobruchus maculatus larval midgut epithelium. Comp. Biochem. Physiol. - Part B Biochem. Mol. Biol. 168, 45–52. https://doi.org/10.1016/j.cbpb.2013.11.001

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., Schulten, K., 2005. Scalable molecular dynamics with NAMD. J. Comput. Chem. 26, 1781–1802. https://doi.org/10.1002/jcc.20289

Quiocho, F.A., 1986. Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. Annu. Rev. Biochem. 55, 287–315. https://doi.org/10.1146/annurev.bi.55.070186.001443

Reis, C.M., Calvet, M.M., Sales, M.P., Fernandes, K.V.S., Gomes, V.M., Xavier, J., 1997. alpha-amylase inhibitors of legume seeds and their involvement in the resistance to Bruchid beetles. Arq. Biol. E Tecnol. 40, 413–418.

Sales, M.P., Gomes, V.M., Fernandes, K.V.S., Xavier, J., 1996. Chitin-binding proteins from cowpea (Vigna unguiculata) seeds. Braz. J. Med. Biol. Res. 29, 319–326.

Sales, M.P., Macedo, M.L.R., Xavier-Filho, J., 1992. Digestibility of cowpea (Vigna unguiculata) vicilins by pepsin, papain and bruchid (insect) midgut proteinases. Comp. Biochem. Physiol. - Part B Biochem. Mol. Biol. 103, 945–950. https://doi.org/10.1016/0305-0491(92)90220-L

Sales, M.P., Pimenta, P.P., Paes, N.S., Grossi-de-Sa, M.F., Xavier, J., 2001. Vicilins (7S storage globulins) of cowpea (Vigna unguiculata) seeds bind to chitinous structures of the midgut of Callosobruchus maculatus (Coleoptera : Bruchidae) larvae. Braz. J. Med. Biol. Res. 34, 27–34. https://doi.org/10.1590/S0100-879X2001000100003

Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., Karplus, M., 1995. Evaluation of comparative protein modeling by MODELLER. Proteins 23, 318–326. https://doi.org/10.1002/prot.340230306

Sammour, R.H., Gatehouse, J.A., Gilroy, J., Boulter, D., 1984. The homology of the major storage protein of jack bean (Canavalia ensiformis) to pea vicilin and its separation from α-mannosidase. Planta 161, 61–70. https://doi.org/10.1007/BF00951461

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., Mann, M., 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. 1, 2856–2860. https://doi.org/10.1038/nprot.2006.468

Singh, B.B., Ajeigbe, H.A., Tarawali, S.A., Fernandez-Rivera, S., Abubakar, M., 2003. Improving the production and utilization of cowpea as food and fodder. Field Crops Res., Approaches to improve the utilization of food-feed crops 84, 169–177. https://doi.org/10.1016/S0378-4290(03)00148-5

Singh, B.B., Singh, S.R., 1990. Breeding for bruchid resistance in cowpea, in: Bruchids and Legumes: Economy, Ecology and Coevolution, Series Entomologica. Kluwer Academic Publishers, Dordrecht, pp. 219–228.

Singh, B.B., Singh, S.R., Adjadi, O., 1985. Bruchid resistance in cowpea. Crop Sci. 25, 736–739.

Souza, S.M., Uchôa, A.F., Silva, J.R., Samuels, R.I., Oliveira, A.E.A., Oliveira, E.M., Linhares, R.T., Alexandre, D., Silva, C.P., 2010. The fate of vicilins, 7S storage globulins, in larvae and adult Callosobruchus maculatus (Coleoptera: Chrysomelidae: Bruchinae). J. Insect Physiol. 56, 1130–1138. https://doi.org/10.1016/j.jinsphys.2010.03.009

Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461. https://doi.org/10.1002/jcc.21334

Uchôa, A.F., DaMatta, R.A., Retamal, C.A., Albuquerque-Cunha, J.M., Souza, S.M., Samuels, R.I., Silva, C.P., Xavier, J., 2006. Presence of the storage seed protein vicilin in internal organs of larval Callosobruchus maculatus (Coleoptera : Bruchidae). J. Insect Physiol. 52, 169–178. https://doi.org/10.1016/j.jinsphys.2005.10.002

Xavier-Filho, J., Campos, F.A.P., Ary, M.B., Silva, C.P., Carvalho, M.M.M., Macedo, M.L.R., Lemos, F.J.A., Grant, G., 1989. Poor correlation between the levels of proteinase inhibitors found in seeds of different cultivars of cowpea (Vigna unguiculata) and the resistance/susceptibility to predation by Callosobruchus maculatus. J. Agric. Food Chem. 37, 1139–1143. https://doi.org/10.1021/jf00088a071

Zhang, Y., Skolnick, J., 2004. Scoring function for automated assessment of protein structure template quality. Proteins 57, 702–710. https://doi.org/10.1002/prot.20264

6-Figure captions

Figure 1. Multiple sequence alignment of the amino acid sequences of β-vignin with the primary structures of representative vicilin-like 7S globulins. Amino acid sequences of β-vignin obtained from *V. unguiculata* genotypes EPACE-10 (sequence S2) and IT81D-1053 (sequence R2) were aligned with those of *V. angularis* (adzuki bean) 7S globulin-3 (Adzuki 7S3; UniProtKB accession number: A0A0S3SX36), *V. radiata* 8S globulin (UniProtKB accession number: Q198W3), β-conglycinin (from *Glycine max*; UniProtKB accession number: P25974) and canavalin (from *Canavalia ensiformis*; UniProtKB accession number: P50477). Segments in the primary structures of β-vignins that were shown to contribute to their chitin-binding site (ChBS), as evidenced by computational simulations, are indicated. The alignment was edited using the program ALINE (Bond and Schüttelkopf, 2009).

Figure 2. Three-dimensional molecular model of β-vignin. N- and C-terminal cupin_1 domains are colored red and blue, respectively. (A) Ribbon diagram of the β-vignin monomer structure, as obtained by homology modeling (the structure modeled from sequence R2 is shown). (B) Superposition of the β-vignin model (blue) over the x-ray crystallographic structure of the *V. angularis* 7S globulin-1 (orange), which was used as template (PDB ID: 2EA7).

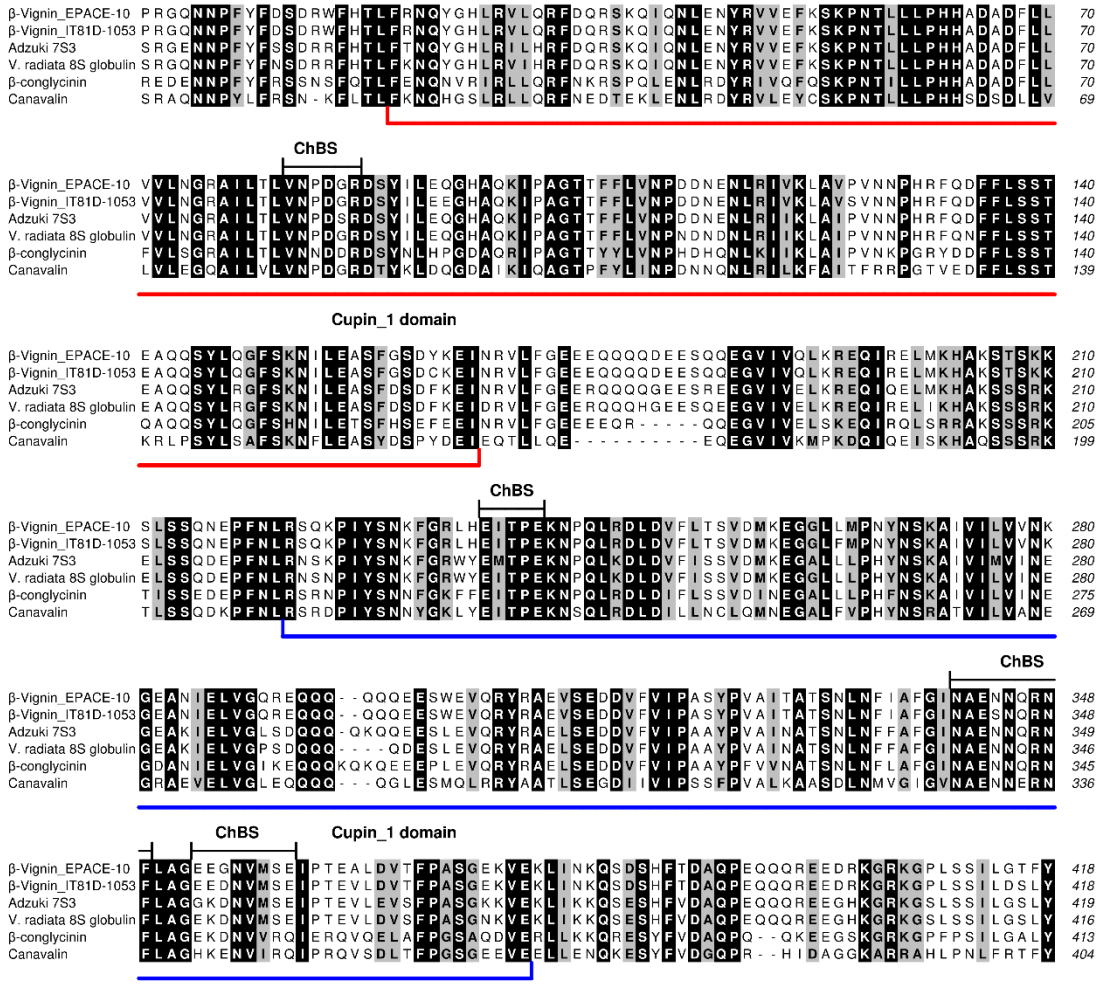
Figure 3. Ribbon diagrams of the β-vignin homotrimers. Homotrimers of the models S2 (A) and R2 (B) are shown. Subunits are colored pink, green and cyan. Chito-oligosaccharide molecules [$(GlcNAc)_4$] docked in the chitin-binding sites of each oligomer are also shown as stick models (carbon, nitrogen and oxygen atoms are colored yellow, blue and red, respectively).

Figure 4. Close-up view of a $(GlcNAc)_4$ molecule docked in the chitin-binding site of β-vignin. The panels show a ligand molecule docked in the ChBS of models S2 (A), S3 (B), R2 (C) and R3 (D). Water molecules are represented as red spheres, hydrogen bonds are shown as green

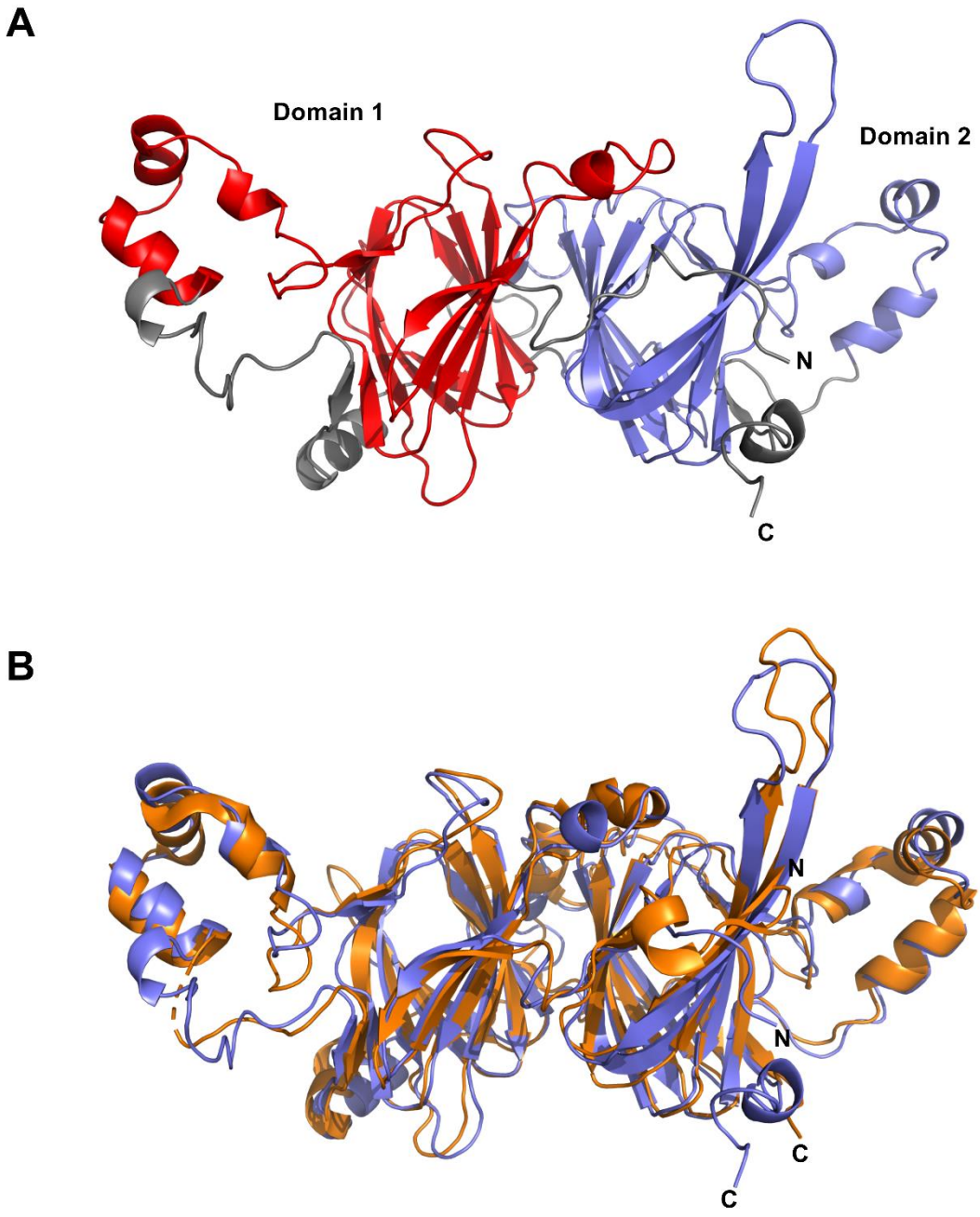
dotted lines (the distance cut-off is 3.5 Å) and the side chains of interacting residues are depicted as sticks. Nitrogen and oxygen atoms are colored blue and red, respectively.

Figure 5. Two-dimensional diagram of the interaction model between a $(GlcNAc)_4$ molecule docked in the chitin-binding site of β-vignin. Residues from model R2 involved in hydrogen bonds (green dashed lines) or hydrophobic contacts (brick-red spoked arcs) are shown. Water molecules (OH2) are represented as red spheres. Similar diagrams depicting the interaction models obtained for the other complexes are shown in Figs. S25-S27. The diagrams were generated using the program LigPlot+ (Laskowski and Swindells, 2011).

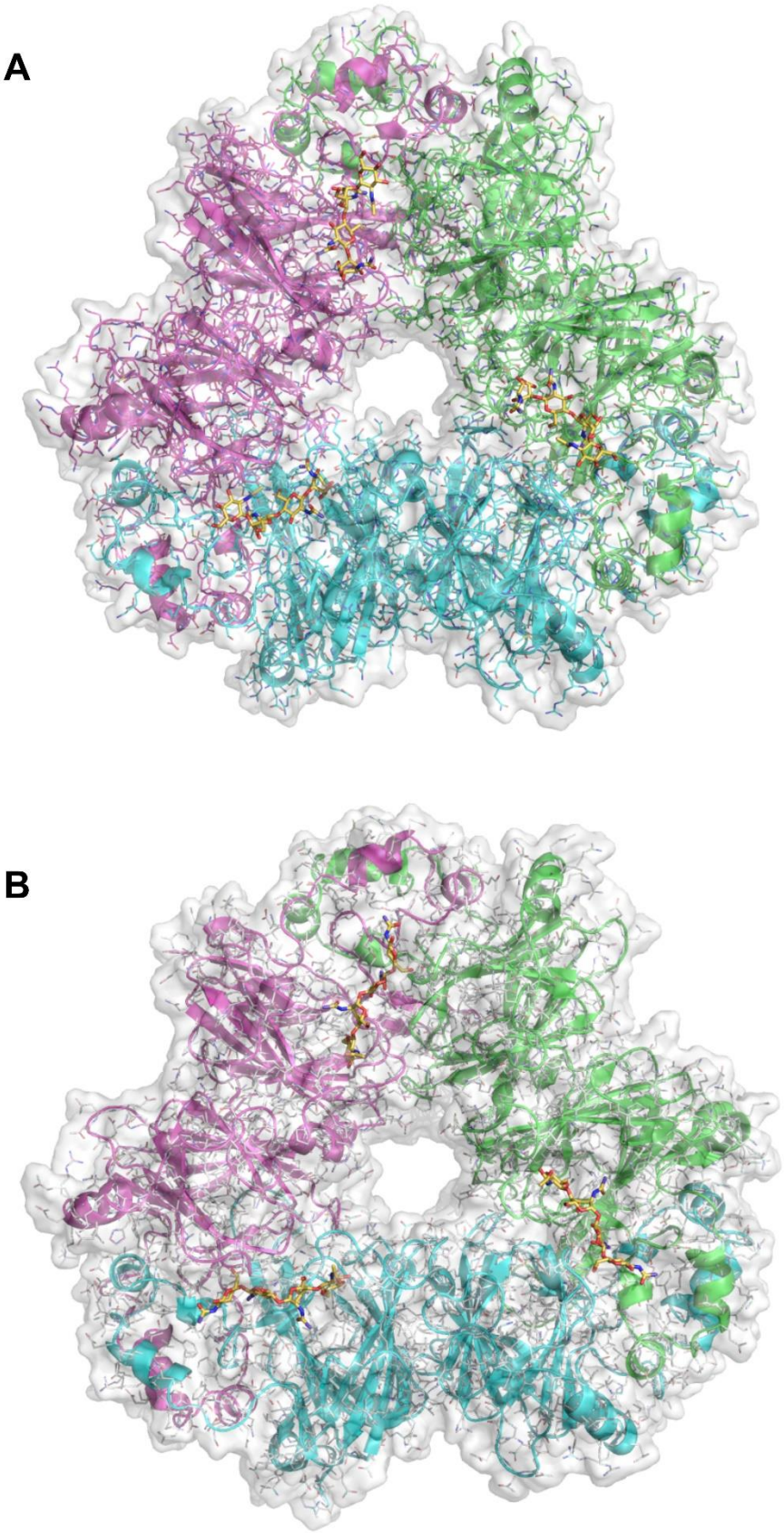
International Journal of Biological Macromolecules
Rocha et al. 2018
Figure 1



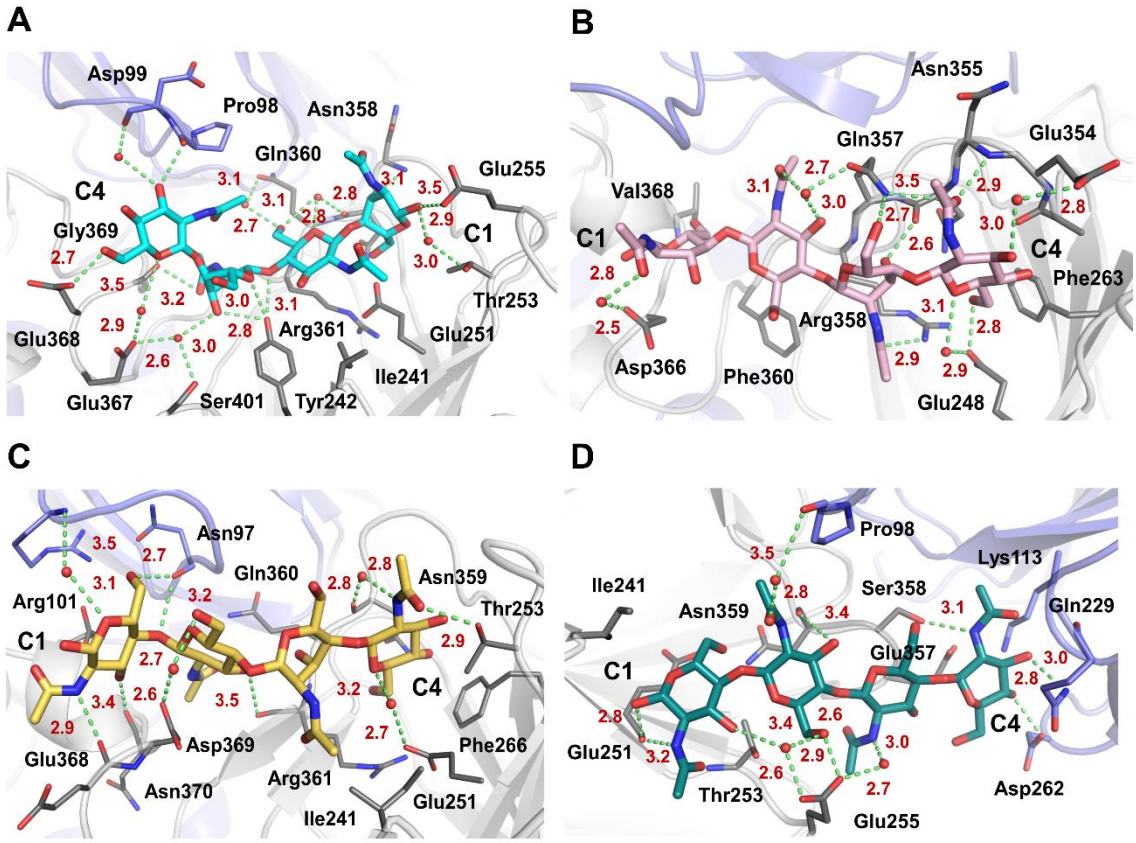
International Journal of Biological Macromolecules
Rocha et al. 2018
Figure 2



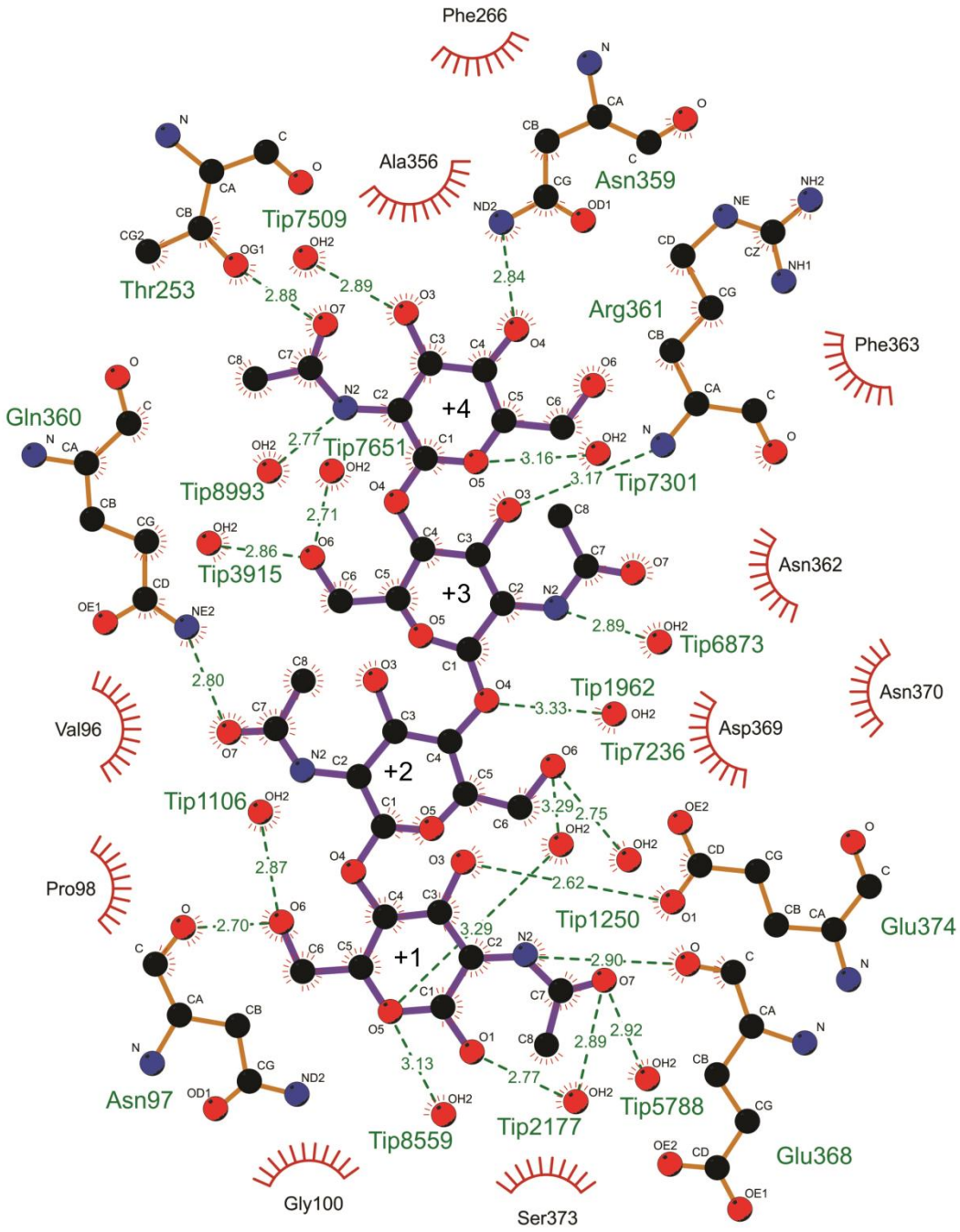
International Journal of Biological Macromolecules
Rocha et al.2018
Figure 3



International Journal of Biological Macromolecules
Rocha et al.2018
Figure 4



International Journal of Biological Macromolecules
Rocha et al.2018
Figure 5



7-Supplementary material

Cloning of cDNA sequences encoding cowpea (*Vigna unguiculata*) vicilins: molecular modeling, docking calculations and molecular dynamics simulations suggest a binding mode of cowpea vicilins to chitin oligomers

Antônio J. Rocha[a], Bruno L. Sousa[b], Matheus S. Girão[a], Ito L. Barroso-Neto[c], José E. Monteiro-Júnior[d], José T. A. Oliveira[a], Celso S. Nagano[e], Rômulo F. Carneiro[e], Ana C. O. Monteiro-Moreira[f], Bruno A. M. Rocha[a], Valder N. Freire[g], Thalles B. Grangeiro[d,*]

[a] Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará (UFC), Fortaleza, Ceará, 60440-900, Brazil

[b] Faculdade de Filosofia Dom Aureliano Matos, Universidade Estadual do Ceará, Av. Dom Aureliano Matos, 2060, Limoeiro do Norte, CE, 62930-000, Brazil

[c] Departamento de Química Analítica e Físico-química, UFC, Fortaleza, Ceará, 60455-760, Brazil

[d] Departamento de Biologia, UFC, Fortaleza, Ceará, 60440-900, Brazil

[e] Departamento de Engenharia de Pesa, UFC, Fortaleza, Ceará, brazil

[f] Núcleo de Biologia Experimental, Universidade de Fortaleza, Fortaleza, Ceará, Brazil

[g] Departamento de Física, UFC, Fortaleza, Ceará, 60440-760, Brazil

*To whom all correspondence should be addressed
Phone: +558533669809; Fax: +558533669806
E-mail: tbgrangeiro@gmail.com

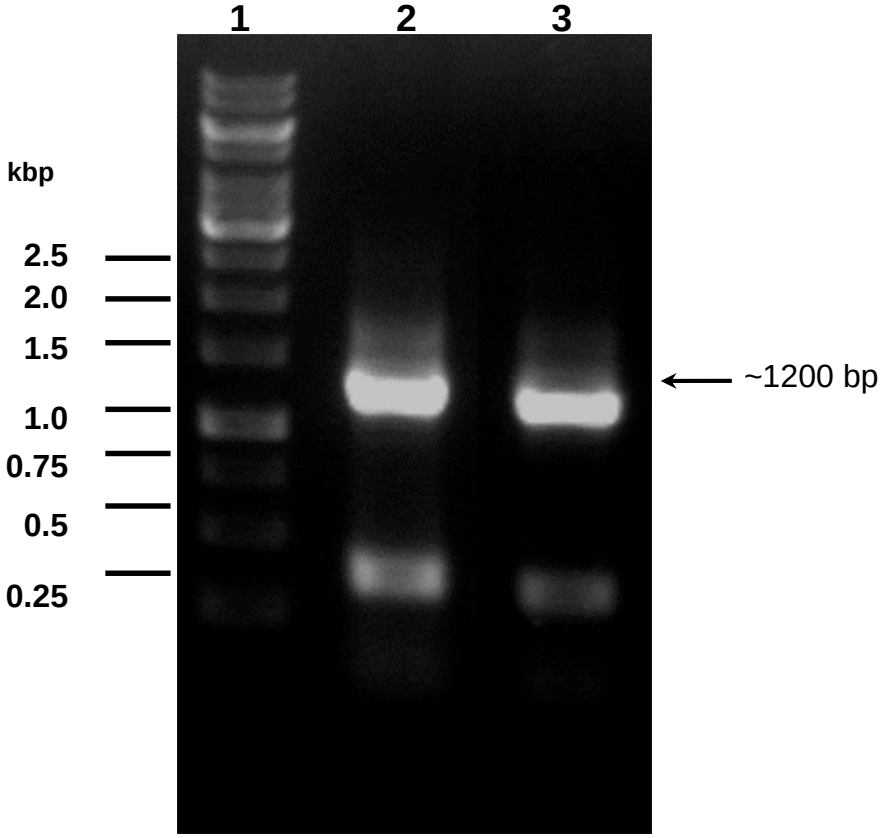


Figure S1. 1% agarose gel electrophoresis of cDNA fragments encoding cowpea vicilins amplified by RT-PCR from total mRNA, which was purified from developing seeds from genotypes IT81D-1053 and EPACE-10. Lane 1: DNA size markers; lane 2: PCR products from genotype IT81D-1053; lane 3: PCR products from genotype EPACE-10. The major amplified band from each sample, which was cloned into the pGEM-T Easy vector, is indicated by an arrow.

```
1     ATTGTACACCGGGAGCACCAAGAGAGCCAAGAAGAGTCGGAACCAAGAGGACAAAATAAC
        I  V  H  R  E  H  Q  E  S  Q  E  E  S  E  P  R  G  Q  N  N   20

61    CCCTTCTACTTCGACTCTGACAGGTGGTTCCACACTCTATTCAGAAACCAATATGGTCAC
        P  F  Y  F  D  S  D  R  W  F  H  T  L  F  R  N  Q  Y  G  H   40

121   CTTCGTGTCCTCCAGAGGTTCGACCAACGCTCCAAACAAATTCAGAATCTTGAAAACTAC
        L  R  V  L  Q  R  F  D  Q  R  S  K  Q  I  Q  N  L  E  N  Y   60

181   CGTGTTGTAGAGTTCAAGTCCAAACCCAACACCCTCCTTCTTCCTCACCATGCTGATGCC
        R  V  V  E  F  K  S  K  P  N  T  L  L  L  P  H  H  A  D  A   80

241   GATTTCCTCCTAGTTGTCCTTAATGGGAGAGCCATACTCACCTTGGTGAACCCTGACGGC
        D  F  L  L  V  V  L  N  G  R  A  I  L  T  L  V  N  P  D  G  100

301   AGAGACTCCTACATTCTTGAGCAAGGCCATGCTCAGAAGATCCCTGCAGGAACCACTTTC
        R  D  S  Y  I  L  E  Q  G  H  A  Q  K  I  P  A  G  T  T  F  120

361   TTTTTGGTTAACCCTGACGACAACGAGAATCTCAGAATAGTCAAACTCGCCGTACCCGTT
        F  L  V  N  P  D  D  N  E  N  L  R  I  V  K  L  A  V  P  V  140

421   AACAACCCTCACAGATTTCAAGACTTTTTCCTATCTAGCACAGAAGCCCAACAATCCTAC
        N  N  P  H  R  F  Q  D  F  F  L  S  S  T  E  A  Q  Q  S  Y  160

481   TTGCAAGGATTCAGCAAGAATATTCTAGAGGCCTCTTTCGGTAGTGACTACAAGGAGATA
        L  Q  G  F  S  K  N  I  L  E  A  S  F  G  S  D  Y  K  E  I  180

541   AACAGGGTTCTGTTTGGAGAGGAGGAACAGCAACAGCAAGACGAGGAGAGTCAGCAAGAA
        N  R  V  L  F  G  E  E  E  Q  Q  Q  Q  D  E  E  S  Q  Q  E  200

601   GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCT
        G  V  I  V  Q  L  K  R  E  Q  I  R  E  L  M  K  H  A  K  S  220

661   ACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCAACCTGAGAAGCCAAAAACCT
        T  S  K  K  S  L  S  S  Q  N  E  P  F  N  L  R  S  Q  K  P  240

721   ATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGAGAAAAACCCCCAGCTT
        I  Y  S  N  K  F  G  R  L  H  E  I  T  P  E  K  N  P  Q  L  260

781   CGAGACTTGGATGTGTTCCTCACTTCTGTGGATATGAAAGAGGGAGGTCTTCTTATGCCC
        R  D  L  D  V  F  L  T  S  V  D  M  K  E  G  G  L  L  M  P  280

841   AACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
        N  Y  N  S  K  A  I  V  I  L  V  V  N  K  G  E  A  N  I  E  300

901   CTTGTTGGGCAAAGAGAACAGCAACAGCAGCAGCAAGAGGAAAGCTGGGAAGTGCAGAGG
        L  V  G  Q  R  E  Q  Q  Q  Q  Q  Q  E  E  S  W  E  V  Q  R  320

961   TATAGAGCTGAGGTGTCTGAAGACGATGTATTTGTTATCCCAGCATCTTATCCAGTTGCC
        Y  R  A  E  V  S  E  D  D  V  F  V  I  P  A  S  Y  P  V  A  340

1021  ATCACCGCAACCTCCAATCTAAATTTCATTGCTTTCGGTATCAATGCTGAGAACAACCAG
        I  T  A  T  S  N  L  N  F  I  A  F  G  I  N  A  E  N  N  Q  360

1081  AGGAACTTCCTTGCAGGTGAGGAAGACAATGTGATGAGCGAGATACCTACAGAGGTGTTG
        R  N  F  L  A  G  E  E  D  N  V  M  S  E  I  P  T  E  V  L  380

1141  GATGTTACCTTCCCTGCATCTGGTGAGAAGGTTGAGAAGTTGATAAACAAGCAGAGTGAT
        D  V  T  F  P  A  S  G  E  K  V  E  K  L  I  N  K  Q  S  D  400

1201  TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGA
        S  H  F  T  D  A  Q  P  E  Q  Q  Q  R  E  E  D  R  K  G  R  420

1261  AAGGGTCCATTATCTTCAATTTTGGACACTTTCTAC
        K  G  P  L  S  S  I  L  D  T  F  Y  432
```

Figure S2. Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype EPACE-10 isoform S1 (GenBank accession number: MG973241). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

```
1     ATTGTACACCGGGAGCACCAAGAGAGCCAAGAAGAGTCGGAACCAAGAGGACAAAATAAC
        I  V  H  R  E  H  Q  E  S  Q  E  E  S  E  P  R  G  Q  N  N   20

61    CCCTTCTACTTCGACTCTGACAGGTGGTTCCACACTCTATTCAGAAACCAATATGGTCAC
        P  F  Y  F  D  S  D  R  W  F  H  T  L  F  R  N  Q  Y  G  H   40

121   CTTCGTGTCCTCCAGAGGTTCGACCAACGCTCCAAACAAATTCAGAATCTTGAAAACTAC
        L  R  V  L  Q  R  F  D  Q  R  S  K  Q  I  Q  N  L  E  N  Y   60

181   CGTGTTGTAGAGTTCAAGTCCAAACCCAACACCCTCCTTCTTCCTCACCATGCTGATGCC
        R  V  V  E  F  K  S  K  P  N  T  L  L  L  P  H  H  A  D  A   80

241   GATTTCCTCCTAGTTGTCCTTAATGGGAGAGCCATACTCACCTTGGTGAACCCTGACGGC
        D  F  L  L  V  V  L  N  G  R  A  I  L  T  L  V  N  P  D  G  100

301   AGAGACTCCTACATTCTTGAGCAAGGCCATGCTCAGAAGATCCCTGCAGGAACCACTTTC
        R  D  S  Y  I  L  E  Q  G  H  A  Q  K  I  P  A  G  T  T  F  120

361   TTTTTGGTTAACCCTGACGACAACGAGAATCTCAGAATAGTCAAACTCGCCGTACCCGTT
        F  L  V  N  P  D  D  N  E  N  L  R  I  V  K  L  A  V  P  V  140

421   AACAACCCTCACAGATTTCAAGACTTTTTCCTATCTAGCACAGAAGCCCAACAATCCTAC
        N  N  P  H  R  F  Q  D  F  F  L  S  S  T  E  A  Q  Q  S  Y  160

481   TTGCAAGGATTCAGCAAGAATATTCTAGAGGCCTCTTTCGGTAGTGACTACAAGGAGATA
        L  Q  G  F  S  K  N  I  L  E  A  S  F  G  S  D  Y  K  E  I  180

541   AACAGGGTTCTGTTTGGAGAGGAGGAACAGCAACAGCAAGACGAGGAGAGTCAGCAAGAA
        N  R  V  L  F  G  E  E  E  Q  Q  Q  Q  D  E  E  S  Q  Q  E  200

601   GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCT
        G  V  I  V  Q  L  K  R  E  Q  I  R  E  L  M  K  H  A  K  S  220

661   ACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCAACCTGAGAAGCCAAAAACCT
        T  S  K  K  S  L  S  S  Q  N  E  P  F  N  L  R  S  Q  K  P  240

721   ATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGAGAAAAACCCCCAGCTT
        I  Y  S  N  K  F  G  R  L  H  E  I  T  P  E  K  N  P  Q  L  260

781   CGAGACTTGGATGTGTTCCTCACTTCTGTGGATATGAAAGAGGGAGGTCTTCTTATGCCC
        R  D  L  D  V  F  L  T  S  V  D  M  K  E  G  G  L  L  M  P  280

841   AACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
        N  Y  N  S  K  A  I  V  I  L  V  V  N  K  G  E  A  N  I  E  300

901   CTTGTTGGGCAAAGAGAACAGCAACAGCAGCAGCAAGAGGAAAGCTGGGAAGTGCAGAGG
        L  V  G  Q  R  E  Q  Q  Q  Q  Q  Q  E  E  S  W  E  V  Q  R  320

961   TATAGAGCTGAGGTGTCTGAAGACGATGTATTTGTTATCCCAGCATCTTATCCAGTTGCC
        Y  R  A  E  V  S  E  D  D  V  F  V  I  P  A  S  Y  P  V  A  340

1021  ATCACCGCAACCTCCAATCTAAATTTCATTGCTTTCGGTATCAATGCTGAGAACAACCAG
        I  T  A  T  S  N  L  N  F  I  A  F  G  I  N  A  E  N  N  Q  360

1081  AGGAACTTCCTTGCAGGTGAGGAAGGCAATGTGATGAGCGAGATACCTACAGAGGCGTTG
        R  N  F  L  A  G  E  E  G  N  V  M  S  E  I  P  T  E  A  L  380

1141  GATGTTACCTTCCCTGCATCTGGTGAGAAGGTTGAGAAGTTGATAAACAAGCAGAGTGAT
        D  V  T  F  P  A  S  G  E  K  V  E  K  L  I  N  K  Q  S  D  400

1201  TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGA
        S  H  F  T  D  A  Q  P  E  Q  Q  Q  R  E  E  D  R  K  G  R  420

1261  AAGGGTCCATTATCTTCAATTTTGGGCACTTTCTAC
        K  G  P  L  S  S  I  L  G  T  F  Y  432
```

Figure S3. Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype EPACE-10 isoform S2 (GenBank accession number: MG973242). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

```
1     ATTGTACACCGGGAGCACCAAGAGAGCCAAGAAGAGTCGGAACCAAGAGGACAAAATAAC
        I  V  H  R  E  H  Q  E  S  Q  E  E  S  E  P  R  G  Q  N  N   20

61    CCCTTCTACTTCGACTCTGACAGGTGGTTCCACACTCTATTCAGAAACCAATATGGTCAC
        P  F  Y  F  D  S  D  R  W  F  H  T  L  F  R  N  Q  Y  G  H   40

121   CTTCGTGTCCTCCAGAGGTTCGACCAACGCTCCAAACAAATTCAGAATCTTGAAAACTAC
        L  R  V  L  Q  R  F  D  Q  R  S  K  Q  I  Q  N  L  E  N  Y   60

181   CGTGTTGTAGAGTTCAAGTCCAAACCCAACACCCTCCTTCTTCCTCACCATGCTGATGCC
        R  V  V  E  F  K  S  K  P  N  T  L  L  L  P  H  H  A  D  A   80

241   GATTTCCTCCTAGTTGTCCTTAATGGGAGAGCCATACTCACCTTGGTGAACCCTGACGGC
        D  F  L  L  V  V  L  N  G  R  A  I  L  T  L  V  N  P  D  G  100

301   AGAGACTCCTACATTCTTGAGCAAGGCCATGCTCAGAAGATCCCTGCAGGAACCACTTTC
        R  D  S  Y  I  L  E  Q  G  H  A  Q  K  I  P  A  G  T  T  F  120

361   TTTTTGGTTAACCCTGACGACAACGAGAATCTCAGAATAGTCAAACTCGCCGTACCCGTT
        F  L  V  N  P  D  D  N  E  N  L  R  I  V  K  L  A  V  P  V  140

421   AACAACCCTCACAGATTTCAAGACTTTTTCCTATCTAGCACAGAAGCCCAACAATCCTAC
        N  N  P  H  R  F  Q  D  F  F  L  S  S  T  E  A  Q  Q  S  Y  160

481   TTGCAAGGATTCAGCAAGAATATTCTAGAGGCCTCTTTCGGTAGTGACTACAAGGAGATA
        L  Q  G  F  S  K  N  I  L  E  A  S  F  G  S  D  Y  K  E  I  180

541   AACAGGGTTCTGTTTGGAGAGGAGGAACAGCAACAGGAGAGTCAGCAAGAAGGAGTGATT
        N  R  V  L  F  G  E  E  E  Q  Q  Q  E  S  Q  Q  E  G  V  I  200

601   GTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCTACTTCAAAG
        V  Q  L  K  R  E  Q  I  R  E  L  M  K  H  A  K  S  T  S  K  220

661   AAATCCCTTTCCTCCCAAAATGAACCATTCAACCTGAGAAGCCAAAAACCTATCTATTCC
        K  S  L  S  S  Q  N  E  P  F  N  L  R  S  Q  K  P  I  Y  S  240

721   AACAAGTTTGGAAGGTTGCATGAGATCACCCCGGAGAAAAACCCCCAGCTTCGAGACTTG
        N  K  F  G  R  L  H  E  I  T  P  E  K  N  P  Q  L  R  D  L  260

781   GATGTGTTCCTCACTTCTGTGGATATGAAAGAGGGAGGTCTTCTTATGCCCAACTACAAT
        D  V  F  L  T  S  V  D  M  K  E  G  G  L  L  M  P  N  Y  N  280

841   TCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAACCTGTTGGG
        S  K  A  I  V  I  L  V  V  N  K  G  E  A  N  I  E  P  V  G  300

901   CAAAGAGAACAGCAACAGCAGCAGCAAGAGGAAAGCTGGGAAGTGCAGAGGTATAGAGCT
        Q  R  E  Q  Q  Q  Q  Q  Q  E  E  S  W  E  V  Q  R  Y  R  A  320

961   GAGGTGTCTGAAGACGATGTATTTGTTATCCCAGCATCTTATCCAGTTGCCATCACCGCA
        E  V  S  E  D  D  V  F  V  I  P  A  S  Y  P  V  A  I  T  A  340

1021  ACCTCCAATCTAAATTTCATTGCTTTCGGTATCAATGCTGAGAACAACCAGAGGAACTTC
        T  S  N  L  N  F  I  A  F  G  I  N  A  E  N  N  Q  R  N  F  360

1081  CTTGCAGGTGAGGAAGACAATGTGATGAGCGAGATACCTACAGAGGTGTTGGATGTTACC
        L  A  G  E  E  D  N  V  M  S  E  I  P  T  E  V  L  D  V  T  380

1141  TTCCCTGCATCTGGTGAGAAGGTTGAGAAGTTGATAAACAAGCAGAGTGATTCCCACTTT
        F  P  A  S  G  E  K  V  E  K  L  I  N  K  Q  S  D  S  H  F  400

1201  ACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGAAAGGGTCCA
        T  D  A  Q  P  E  Q  Q  Q  R  E  E  D  R  K  G  R  K  G  P  420

1261  TTATCTTCTATTTTGGGCACTCTCTAC
        L  S  S  I  L  G  T  L  Y   429
```

Figure S4. Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype EPACE-10 isoform S3 (GenBank accession number: MG973243). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

```
1     ATTGTACACCGGGAGCACCAAGAGAGCCAAGAAGAGTCGGAACCAAGAGGACAAAATAAC
        I  V  H  R  E  H  Q  E  S  Q  E  E  S  E  P  R  G  Q  N  N   20

61    CCCTTCTACTTCGACTCTGACAGGTGGTTCCACACTCTATTCAGAAACCAATATGGTCAC
        P  F  Y  F  D  S  D  R  W  F  H  T  L  F  R  N  Q  Y  G  H   40

121   CTTCGTGTCCTCCAGAGGTTCGACCAACGCTCCAAACAAATTCAGAATCTTGAAAACTAC
        L  R  V  L  Q  R  F  D  Q  R  S  K  Q  I  Q  N  L  E  N  Y   60

181   CGTGTTGTAGAGTTCAAGTCCAAACCCAACACCCTCCTTCTTCCTCACCATGCTGATGCC
        R  V  V  E  F  K  S  K  P  N  T  L  L  L  P  H  H  A  D  A   80

241   GATTTCCTCCTAGTTGTCCTTAATGGGAGAGCCATACTCACCTTGGTGAACCCTGACGGC
        D  F  L  L  V  V  L  N  G  R  A  I  L  T  L  V  N  P  D  G  100

301   AGAGACTCCTACATTCTTGAGCAAGGCCATGCTCAGAAGATCCCTGCAGGAACCACTTTC
        R  D  S  Y  I  L  E  Q  G  H  A  Q  K  I  P  A  G  T  T  F  120

361   TTTTTGGTTAACCCTGACGACAACGAGAATCTCAGAATAGTCAAACTCGCCGTACCCGTT
        F  L  V  N  P  D  D  N  E  N  L  R  I  V  K  L  A  V  P  V  140

421   AACAACCCTCACAGATTTCAAGACTTTTTCCTATCTAGCACAGAAGCCCAACAATCCTAC
        N  N  P  H  R  F  Q  D  F  F  L  S  S  T  E  A  Q  Q  S  Y  160

481   TTGCAAGGATTCAGCAAGAATATTCTAGAGGCCTCCTTCGGTAGTGACTACAAGGAGATA
        L  Q  G  F  S  K  N  I  L  E  A  S  F  G  S  D  Y  K  E  I  180

541   AACAGGGTTCTGTTTGGAGAGGAGGAACAGCAACAGCAAGACGAGGAGAGTCAGCAAGAA
        N  R  V  L  F  G  E  E  E  Q  Q  Q  Q  D  E  E  S  Q  Q  E  200

601   GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCT
        G  V  I  V  Q  L  K  R  E  Q  I  R  E  L  M  K  H  A  K  S  220

661   ACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCAACCTGAGAAGCCAAAAACCT
        T  S  K  K  S  L  S  S  Q  N  E  P  F  N  L  R  S  Q  K  P  240

721   ATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGAGAAAAACCCCCAGCTT
        I  Y  S  N  K  F  G  R  L  H  E  I  T  P  E  K  N  P  Q  L  260

781   CGAGACTTGGATGTGTTCCTCACTTCTGTGGATATGAAAGAGGGAGGTCTTCTTATGCCC
        R  D  L  D  V  F  L  T  S  V  D  M  K  E  G  G  L  L  M  P  280

841   AACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
        N  Y  N  S  K  A  I  V  I  L  V  V  N  K  G  E  A  N  I  E  300

901   CTTGTTGGGCAAAGAGAACAGCAACAGCAGCAGCAAGAGGAAAGCTGGGAAGTGCAGAGG
        L  V  G  Q  R  E  Q  Q  Q  Q  Q  Q  E  E  S  W  E  V  Q  R  320

961   TATAGAGCTGAGGTGTCTGAAGACGATGTATTTGTTATCCCAGCATCTTATCCAGTTGCC
        Y  R  A  E  V  S  E  D  D  V  F  V  I  P  A  S  Y  P  V  A  340

1021  ATCACCGCAACCTCCAATCTAAATTTCATTGCTTTCGGTATCAATGCTGAGAACAACCAG
        I  T  A  T  S  N  L  N  F  I  A  F  G  I  N  A  E  N  N  Q  360

1081  AGGAACTTCCTTGCAGGTGAGGAAGACAATGTGATGAGCGAGATACCTACAGAGGTGTTG
        R  N  F  L  A  G  E  E  D  N  V  M  S  E  I  P  T  E  V  L  380

1141  GATGTTACCTTCCCTGCATCTGGTGAGAAGGTTGAGAAGTTGATAAACAAGCAGAGTGAT
        D  V  T  F  P  A  S  G  E  K  V  E  K  L  I  N  K  Q  S  D  400

1201  TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGA
        S  H  F  T  D  A  Q  P  E  Q  Q  Q  R  E  E  D  R  K  G  R  420

1261  AAGGGTCCATTATCTTCAATTTTGGACACTTTCTAC
        K  G  P  L  S  S  I  L  D  T  F  Y  432
```

Figure S5. Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype IT81D-1053 isoform R1 (GenBank accession number: MG973244). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

```
1      ATTGTACACCGGGAGCACCAAGAGAGCCAAGAAGAGTCGGAATCAAGAGGACAAAATAAC
         I  V  H  R  E  H  Q  E  S  Q  E  E  S  E  S  R  G  Q  N  N   20

61     CCCTTCTACTTCGACTCTGACAGGTGGTTCCATACTCTATTCAGAAACCAATATGGTCAC
         P  F  Y  F  D  S  D  R  W  F  H  T  L  F  R  N  Q  Y  G  H   40

121    CTTCGTGTCCTCCAGAGGTTCGACCAACGCTCCAAACAAATTCAGAATCTTGAAAACTAC
         L  R  V  L  Q  R  F  D  Q  R  S  K  Q  I  Q  N  L  E  N  Y   60

181    CGTGTTGTAGAGTTCAAGTCCAAACCCAACACCCTCCTTCTTCCTCACCATGCTGATGCC
         R  V  V  E  F  K  S  K  P  N  T  L  L  L  P  H  H  A  D  A   80

241    GATTTCCTCCTAGTTGTCCTTAATGGGAGAGCCATACTCACCTTGGTGAACCCTGACGGC
         D  F  L  L  V  V  L  N  G  R  A  I  L  T  L  V  N  P  D  G  100

301    AGAGACTCCTACATTCTTGAGCAAGGCCATGCTCAGAAGATCCCTGCAGGAACCACTTTC
         R  D  S  Y  I  L  E  Q  G  H  A  Q  K  I  P  A  G  T  T  F  120

361    TTTTTGGTTAACCCTGACGACAACGAGAATCTCAGAATAGTCAAACTCGCCGTACCCGTT
         F  L  V  N  P  D  D  N  E  N  L  R  I  V  K  L  A  V  P  V  140

421    AACAATCCTCACAGATTTCAAGACTTTTTCCTATCTAGCACAGAAGCCCAACAATCCTAC
         N  N  P  H  R  F  Q  D  F  F  L  S  S  T  E  A  Q  Q  S  Y  160

481    TTGCAAGGATTCAGCAAGAATATTCTAGAGGCCTCCTTCGATAGTGACTACAAGGAGATA
         L  Q  G  F  S  K  N  I  L  E  A  S  F  D  S  D  Y  K  E  I  180

541    AACAGGGTTCTGTTTGGAGAGGAGGAACAGAAACAGCAAGACGAGGAGAGTCAGCAAGAA
         N  R  V  L  F  G  E  E  E  Q  K  Q  Q  D  E  E  S  Q  Q  E  200

601    GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCT
         G  V  I  V  Q  L  K  R  E  Q  I  R  E  L  M  K  H  A  K  S  220

661    ACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCAACCTGAGAAGCCAAAAACCT
         T  S  K  K  S  L  S  S  Q  N  E  P  F  N  L  R  S  Q  K  P  240

721    ATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGGGAAAAACCCCCAGCTT
         I  Y  S  N  K  F  G  R  L  H  E  I  T  P  G  K  N  P  Q  L  260

781    CGAGACTTGGATGTGTTCCTCACTTCTGTGGATATGAAAGAGGGAGGTCTTCTTATGCCC
         R  D  L  D  V  F  L  T  S  V  D  M  K  E  G  G  L  L  M  P  280

841    AACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
         N  Y  N  S  K  A  I  V  I  L  V  V  N  K  G  E  A  N  I  E  300

901    CTTGTTGGGCAAAGAGAACAGCAACAGCAGCAGCAAGAGGAAAGCTGGGAAGTGCAGAGG
         L  V  G  Q  R  E  Q  Q  Q  Q  Q  Q  E  E  S  W  E  V  Q  R  320

961    TATAGAGCTGAGGTGTCTGAAGACGATGTATTTGTTATCCCAGCATCTTATCCAGTTGCC
         Y  R  A  E  V  S  E  D  D  V  F  V  I  P  A  S  Y  P  V  A  340

1021   ATCACCGCAACCTCCAATCTAAATTTCATTGCTTTCGGTATCAATGCTGAGAACAACCAG
         I  T  A  T  S  N  L  N  F  I  A  F  G  I  N  A  E  N  N  Q  360

1081   AGGAACTTCCTTGCAGGCGAGGAAGACAATGTGATGAGCGAGATACCTACAGAGGTGTTG
         R  N  F  L  A  G  E  E  D  N  V  M  S  E  I  P  T  E  V  L  380

1141   GATGTTACCTTCCCTGCATCTGGTGAGAAGGTTGAGAAGTTGATAAACAAGCAGAGTGAT
         D  V  T  F  P  A  S  G  E  K  V  E  K  L  I  N  K  Q  S  D  400

1201   TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGA
         S  H  F  T  D  A  Q  P  E  Q  Q  Q  R  E  E  D  R  K  G  R  420

1261   AAGGGTCCATTATCTTCTATTTTGGACACTTTCTAC
         K  G  P  L  S  S  I  L  D  T  F  Y  432
```

Figure S6. Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype IT81D-1053 isoform R2 (GenBank accession number: MG973245). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

```
1      ATTGTACACCGGGAGCACCAAGAGAGCCAAGAAGAGTCGGAACCAAGAGGACAAAATAAC
         I  V  H  R  E  H  Q  E  S  Q  E  E  S  E  P  R  G  Q  N  N   20

61     CCCTTCTACTTCGACTCTGACAGGTGGTTCCACACTCTATTCAGAAACCAATATGGTCAC
         P  F  Y  F  D  S  D  R  W  F  H  T  L  F  R  N  Q  Y  G  H   40

121    CTTCGTGTCCTCCAGAGGTTCGACCAACGCTCCAAACAAATTCAGAATCTTGAAAACTAC
         L  R  V  L  Q  R  F  D  Q  R  S  K  Q  I  Q  N  L  E  N  Y   60

181    CGTGTCGTAGAGTTCAAGTCCAAACCCAACACCCTCCTTCTTCCTCACCATGCTGATGCC
         R  V  V  E  F  K  S  K  P  N  T  L  L  L  P  H  H  A  D  A   80

241    GATTTCCTCCTAGTTGTCCTTAATGGGAGAGCCATACTCACCTTGGTGAACCCTGACGGC
         D  F  L  L  V  V  L  N  G  R  A  I  L  T  L  V  N  P  D  G  100

301    AGAGACTCCTACATTCTTGAGGAAGGCCATGCTCAGAAGATCCCTGCAGGAACCACTTTC
         R  D  S  Y  I  L  E  E  G  H  A  Q  K  I  P  A  G  T  T  F  120

361    TTTTTGGTTAACCCTGACGACAACGAGAATCTCAGAATAGTCAAACTCGCCGTATCCGTT
         F  L  V  N  P  D  D  N  E  N  L  R  I  V  K  L  A  V  S  V  140

421    AACAACCCTCACAGATTTCAAGACTTTTTCCTATCTAGCACAGAAGCCCAACAATCCTAC
         N  N  P  H  R  F  Q  D  F  F  L  S  S  T  E  A  Q  Q  S  Y  160

481    TTGCAAGGATTCAGCAAGAATATTCTAGAGGCCTCTTTCGGTAGTGACTGCAAGGAGATA
         L  Q  G  F  S  K  N  I  L  E  A  S  F  G  S  D  C  K  E  I  180

541    AACAGGGTTCTGTTTGGAGAGGAGGAACAGCAACAGCAAGACGAGGAGAGTCAGCAAGAA
         N  R  V  L  F  G  E  E  E  Q  Q  Q  Q  D  E  E  S  Q  Q  E  200

601    GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCT
         G  V  I  V  Q  L  K  R  E  Q  I  R  E  L  M  K  H  A  K  S  220

661    ACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCAACCTGAGAAGCCAAAAACCT
         T  S  K  K  S  L  S  S  Q  N  E  P  F  N  L  R  S  Q  K  P  240

721    ATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGAGAAAAACCCCCAGCTT
         I  Y  S  N  K  F  G  R  L  H  E  I  T  P  E  K  N  P  Q  L  260

781    CGAGACTTGGATGTGTTCCTCACTTCTGTGGATATGAAAGAGGGAGGTCTTTTTATGCCC
         R  D  L  D  V  F  L  T  S  V  D  M  K  E  G  G  L  F  M  P  280

841    AACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
         N  Y  N  S  K  A  I  V  I  L  V  V  N  K  G  E  A  N  I  E  300

901    CTTGTTGGGCAAAGAGAACAGCAACAGCAGCAGCAAGAGGAAAGCTGGGAAGTGCAGAGG
         L  V  G  Q  R  E  Q  Q  Q  Q  Q  Q  E  E  S  W  E  V  Q  R  320

961    TATAGAGCTGAGGTGTCTGAAGACGATGTATTTGTTATCCCAGCATCTTATCCAGTTGCC
         Y  R  A  E  V  S  E  D  D  V  F  V  I  P  A  S  Y  P  V  A  340

1021   ATCACCGCAACCTCCAATCTAAATTTCATTGCTTTCGGTATCAATGCTGAGAGCAACCAG
         I  T  A  T  S  N  L  N  F  I  A  F  G  I  N  A  E  S  N  Q  360

1081   AGGAACTTCCTTGCAGGTGAGGAAGACAATGTGATGAGCGAGATACCTACAGAGGTGTTG
         R  N  F  L  A  G  E  E  D  N  V  M  S  E  I  P  T  E  V  L  380

1141   GATGTTACCTTCCCTGCATCTGGTGAGAAGGTTGAGAAGTTGATAAACAAGCAGAGTGAT
         D  V  T  F  P  A  S  G  E  K  V  E  K  L  I  N  K  Q  S  D  400

1201   TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGA
         S  H  F  T  D  A  Q  P  E  Q  Q  Q  R  E  E  D  R  K  G  R  420

1261   AAGGGTCCATTATCTTCAATTTTGGACAGTCTCTAC
         K  G  P  L  S  S  I  L  D  S  L  Y  432
```

Figure S7. Partial cDNA sequence and deduced amino acid sequence of cowpea vicilin from genotype IT81D-1053 isoform R3 (GenBank accession number: MG973246). The sequences of the oligonucleotide primers, which were used to amplify the cDNA fragment, are underlined. The amino acids matching the N-terminal sequence determined from β-vignin subunits is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

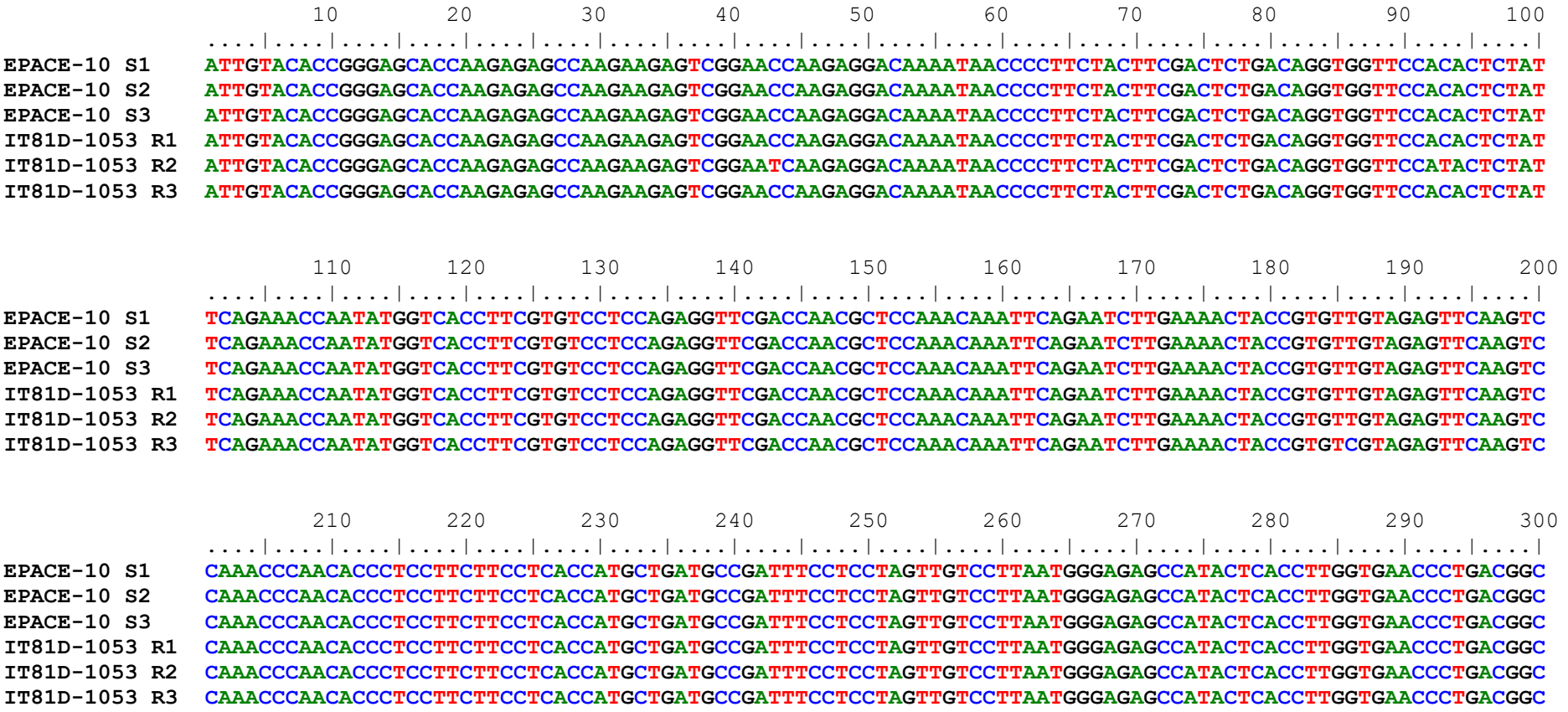


Figure S8. Alignment of partial cDNA sequences encoding cowpea vicilins, which were cloned from genotypes EPACE-10 (sequences S1, S2 and S3) and IT81D-1053 (sequences R1, R2 and R3).

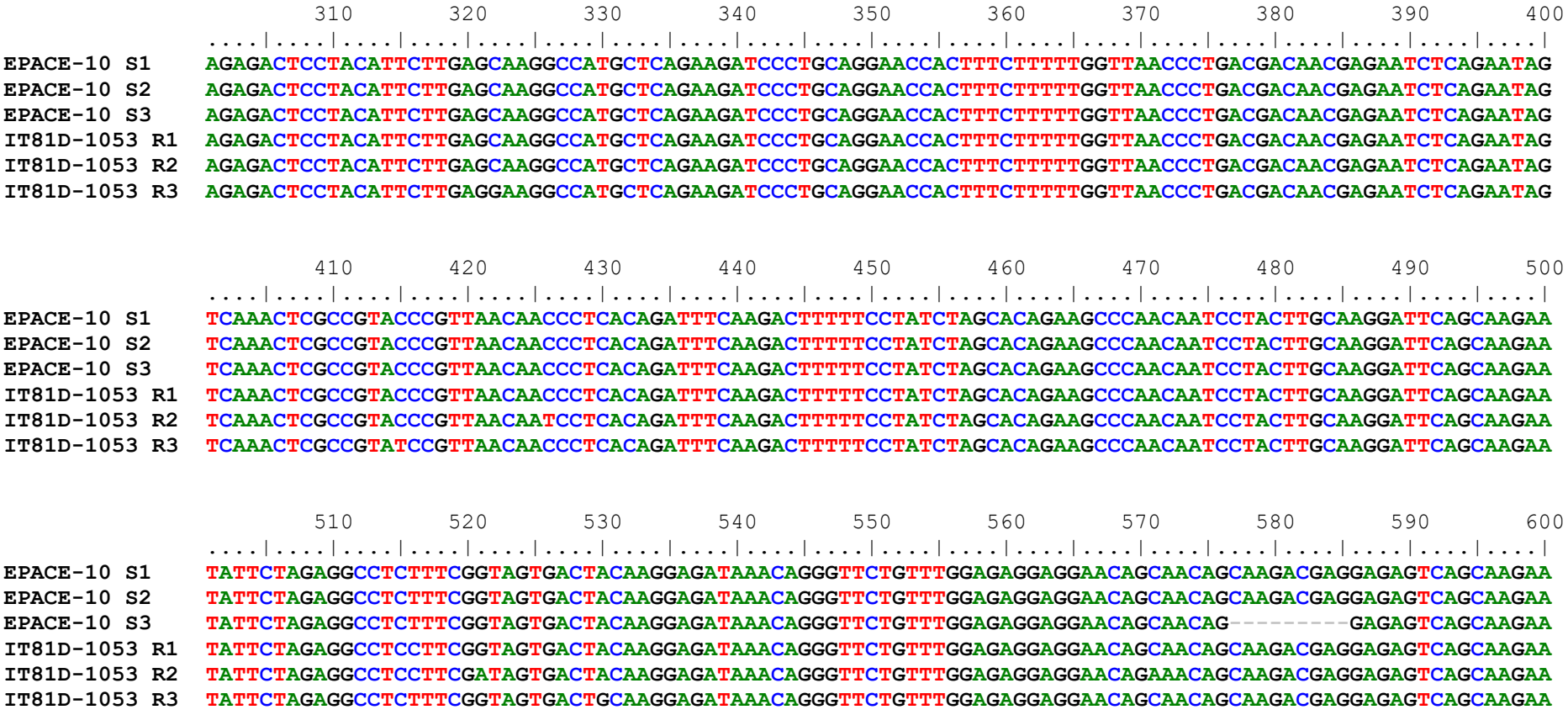


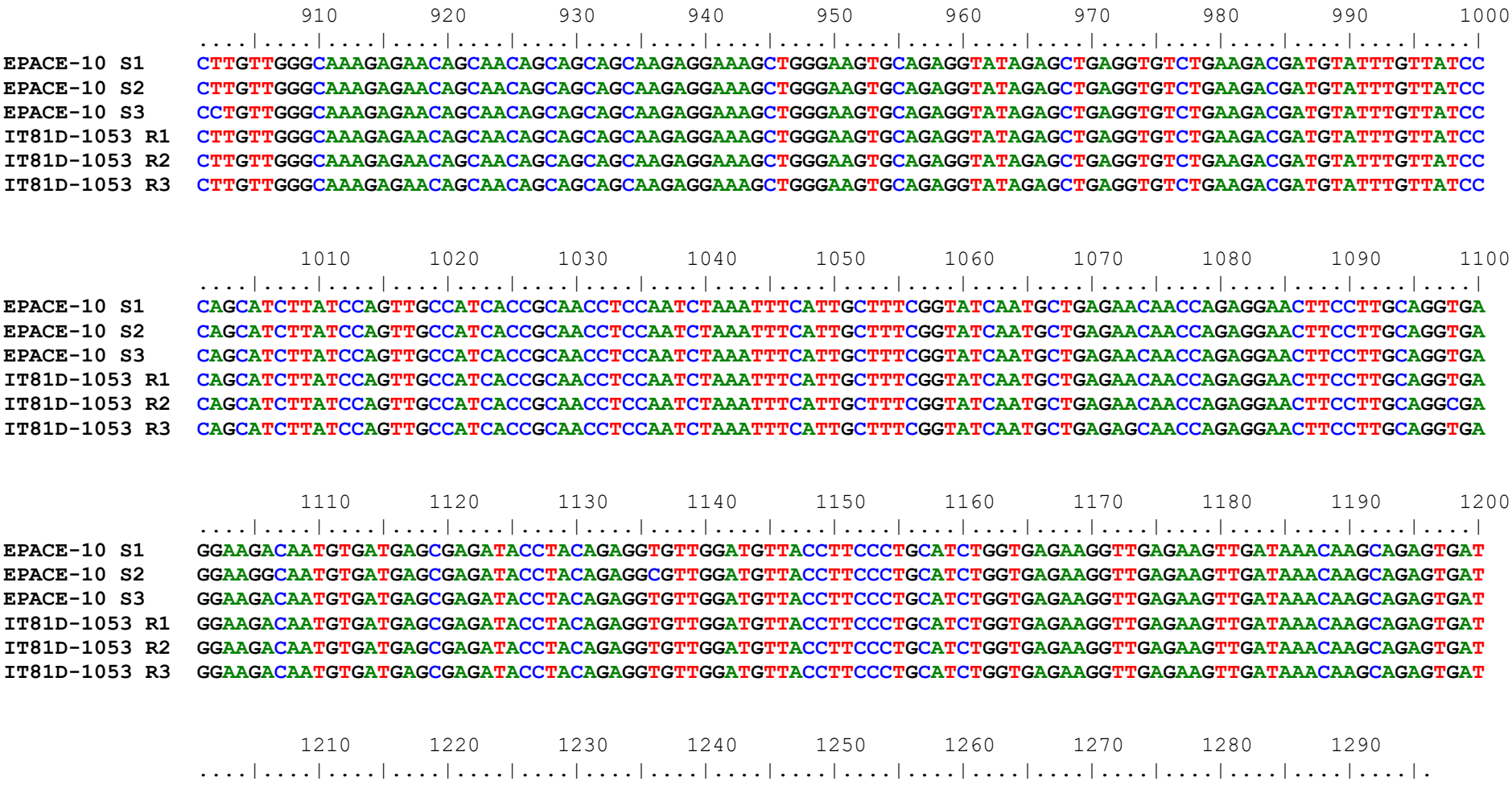
Figure S8. (Continued)

```
                  610       620       630       640       650       660       670       680       690       700
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EPACE-10 S1       GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCTACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCA
EPACE-10 S2       GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCTACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCA
EPACE-10 S3       GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCTACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCA
IT81D-1053 R1     GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCTACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCA
IT81D-1053 R2     GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCTACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCA
IT81D-1053 R3     GGAGTGATTGTGCAACTTAAAAGGGAACAGATTCGGGAACTGATGAAGCATGCTAAATCTACTTCAAAGAAATCCCTTTCCTCCCAAAATGAACCATTCA

                  710       720       730       740       750       760       770       780       790       800
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EPACE-10 S1       ACCTGAGAAGCCAAAAACCTATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGAGAAAAACCCCCAGCTTCGAGACTTGGATGTGTTCCT
EPACE-10 S2       ACCTGAGAAGCCAAAAACCTATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGAGAAAAACCCCCAGCTTCGAGACTTGGATGTGTTCCT
EPACE-10 S3       ACCTGAGAAGCCAAAAACCTATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGAGAAAAACCCCCAGCTTCGAGACTTGGATGTGTTCCT
IT81D-1053 R1     ACCTGAGAAGCCAAAAACCTATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGAGAAAAACCCCCAGCTTCGAGACTTGGATGTGTTCCT
IT81D-1053 R2     ACCTGAGAAGCCAAAAACCTATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGGGAAAAACCCCCAGCTTCGAGACTTGGATGTGTTCCT
IT81D-1053 R3     ACCTGAGAAGCCAAAAACCTATCTATTCCAACAAGTTTGGAAGGTTGCATGAGATCACCCCGGAGAAAAACCCCCAGCTTCGAGACTTGGATGTGTTCCT

                  810       820       830       840       850       860       870       880       890       900
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EPACE-10 S1       CACTTCTGTGGATATGAAAGAGGGAGGTCTTCTTATGCCCAACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
EPACE-10 S2       CACTTCTGTGGATATGAAAGAGGGAGGTCTTCTTATGCCCAACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
EPACE-10 S3       CACTTCTGTGGATATGAAAGAGGGAGGTCTTCTTATGCCCAACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
IT81D-1053 R1     CACTTCTGTGGATATGAAAGAGGGAGGTCTTCTTATGCCCAACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
IT81D-1053 R2     CACTTCTGTGGATATGAAAGAGGGAGGTCTTCTTATGCCCAACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
IT81D-1053 R3     CACTTCTGTGGATATGAAAGAGGGAGGTCTTTTTTATGCCCAACTACAATTCAAAGGCCATAGTGATACTAGTGGTTAATAAAGGAGAAGCAAATATTGAA
```

Figure S8. (Continued)



```
EPACE-10 S1     TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGAAAGGGTCCATTATCTTCAATTTTGGACACTTTCTAC
EPACE-10 S2     TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGAAAGGGTCCATTATCTTCAATTTTGGGCACTTTCTAC
EPACE-10 S3     TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGAAAGGGTCCATTATCTTCTATTTTGGGCACTCTCTAC
IT81D-1053 R1   TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGAAAGGGTCCATTATCTTCAATTTTGGACACTTTCTAC
IT81D-1053 R2   TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGAAAGGGTCCATTATCTTCTATTTTGGACACTTTCTAC
IT81D-1053 R3   TCCCACTTTACGGATGCACAGCCTGAGCAACAGCAAAGGGAGGAGGATCGTAAGGGAAGAAAGGGTCCATTATCTTCAATTTTGGACAGTCTCTAC
```

Figure S8. (Continued)

```
EPACE-10_S1      1
IVHREHQESQEESEPRGQNNPFYFDSDRWFHTLFRNQYGHLRVLQRFDQRSKQIQNLENY
EPACE-10_S2      1
IVHREHQESQEESEPRGQNNPFYFDSDRWFHTLFRNQYGHLRVLQRFDQRSKQIQNLENY
EPACE-10_S3      1
IVHREHQESQEESEPRGQNNPFYFDSDRWFHTLFRNQYGHLRVLQRFDQRSKQIQNLENY
IT81D-1053_R1    1
IVHREHQESQEESEPRGQNNPFYFDSDRWFHTLFRNQYGHLRVLQRFDQRSKQIQNLENY
IT81D-1053_R2    1
IVHREHQESQEESESRGQNNPFYFDSDRWFHTLFRNQYGHLRVLQRFDQRSKQIQNLENY
IT81D-1053_R3    1
IVHREHQESQEESEPRGQNNPFYFDSDRWFHTLFRNQYGHLRVLQRFDQRSKQIQNLENY
                 1 **************
*****************************************

EPACE-10_S      61
RVVEFKSKPNTLLLPHHADADFLLVVLNGRAILTLVNPDGRDSYILEQGHAQKIPAGTTF
EPACE-10_S      61
RVVEFKSKPNTLLLPHHADADFLLVVLNGRAILTLVNPDGRDSYILEQGHAQKIPAGTTF
EPACE-10_S      61
RVVEFKSKPNTLLLPHHADADFLLVVLNGRAILTLVNPDGRDSYILEQGHAQKIPAGTTF
IT81D-1053_R    61
RVVEFKSKPNTLLLPHHADADFLLVVLNGRAILTLVNPDGRDSYILEQGHAQKIPAGTTF
IT81D-1053_R    61
RVVEFKSKPNTLLLPHHADADFLLVVLNGRAILTLVNPDGRDSYILEQGHAQKIPAGTTF
IT81D-1053_R    61
RVVEFKSKPNTLLLPHHADADFLLVVLNGRAILTLVNPDGRDSYILEEGHAQKIPAGTTF
                61 **********************************************
************

EPACE-10_S1     121
FLVNPDDNENLRIVKLAVPVNNPHRFQDFFLSSTEAQQSYLQGFSKNILEASFGSDYKEI
EPACE-10_S2     121
FLVNPDDNENLRIVKLAVPVNNPHRFQDFFLSSTEAQQSYLQGFSKNILEASFGSDYKEI
EPACE-10_S3     121
FLVNPDDNENLRIVKLAVPVNNPHRFQDFFLSSTEAQQSYLQGFSKNILEASFGSDYKEI
IT81D-1053_R1   121
FLVNPDDNENLRIVKLAVPVNNPHRFQDFFLSSTEAQQSYLQGFSKNILEASFGSDYKEI
IT81D-1053_R2   121
FLVNPDDNENLRIVKLAVPVNNPHRFQDFFLSSTEAQQSYLQGFSKNILEASFDSDYKEI
IT81D-1053_R3   121
FLVNPDDNENLRIVKLAVSVNNPHRFQDFFLSSTEAQQSYLQGFSKNILEASFGSDCKEI
                121 ******************
********************************** ** ***

EPACE-10_S1     181
NRVLFGEEEQQQQDEESQQEGVIVQLKREQIRELMKHAKSTSKKSLSSQNEPFNLRSQKP
EPACE-10_S2     181
NRVLFGEEEQQQQDEESQQEGVIVQLKREQIRELMKHAKSTSKKSLSSQNEPFNLRSQKP
EPACE-10_S3     181 NRVLFGEEEQQQ---
ESQQEGVIVQLKREQIRELMKHAKSTSKKSLSSQNEPFNLRSQKP
IT81D-1053_R1   181
NRVLFGEEEQQQQDEESQQEGVIVQLKREQIRELMKHAKSTSKKSLSSQNEPFNLRSQKP
IT81D-1053_R2   181
```

```
NRVLFGEEEQKQQDEESQQEGVIVQLKREQIRELMKHAKSTSKKSLSSQNEPFNLRSQKP
IT81D-1053_R3  181
NRVLFGEEEQQQQDEESQQEGVIVQLKREQIRELMKHAKSTSKKSLSSQNEPFNLRSQKP
               181 ********** *
*****************************************

EPACE-10_S1    241
IYSNKFGRLHEITPEKNPQLRDLDVFLTSVDMKEGGLLMPNYNSKAIVILVVNKGEANIE
EPACE-10_S2    241
IYSNKFGRLHEITPEKNPQLRDLDVFLTSVDMKEGGLLMPNYNSKAIVILVVNKGEANIE
EPACE-10_S3    238
IYSNKFGRLHEITPEKNPQLRDLDVFLTSVDMKEGGLLMPNYNSKAIVILVVNKGEANIE
IT81D-1053_R1  241
IYSNKFGRLHEITPEKNPQLRDLDVFLTSVDMKEGGLLMPNYNSKAIVILVVNKGEANIE
IT81D-1053_R2  241
IYSNKFGRLHEITPGKNPQLRDLDVFLTSVDMKEGGLLMPNYNSKAIVILVVNKGEANIE
IT81D-1053_R3  241
IYSNKFGRLHEITPEKNPQLRDLDVFLTSVDMKEGGLFMPNYNSKAIVILVVNKGEANIE
               241 ************** **********************
**********************
```

Figure S9. Alignment of the amino acid sequences of cowpea vicilins, which were deduced from partial cDNA sequences obtained from genotypes EPACE-10 and IT81D-1053. Sites containing identical residues are shaded in black, whereas positions with chemically similar residues are shaded in gray. The alignment was rendered using BOXSHADE v3.21 through the web server embnet.vital-it.ch/software/BOX_form.html.

```
EPACE-10_S1    301
LVGQREQQQQQQEESWEVQRYRAEVSEDDVFVIPASYPVAITATSNLNFIAFGINAENNQ
EPACE-10_S2    301
LVGQREQQQQQQEESWEVQRYRAEVSEDDVFVIPASYPVAITATSNLNFIAFGINAENNQ
EPACE-10_S3    298
PVGQREQQQQQQEESWEVQRYRAEVSEDDVFVIPASYPVAITATSNLNFIAFGINAENNQ
IT81D-1053_R1  301
LVGQREQQQQQQEESWEVQRYRAEVSEDDVFVIPASYPVAITATSNLNFIAFGINAENNQ
IT81D-1053_R2  301
LVGQREQQQQQQEESWEVQRYRAEVSEDDVFVIPASYPVAITATSNLNFIAFGINAENNQ
IT81D-1053_R3  301
LVGQREQQQQQQEESWEVQRYRAEVSEDDVFVIPASYPVAITATSNLNFIAFGINAESNQ
               301
************************************************************ **

EPACE-10_S1    361
RNFLAGEEDNVMSEIPTEVLDVTFPASGEKVEKLINKQSDSHFTDAQPEQQQREEDRKGR
EPACE-10_S2    361
```

```
RNFLAGEEGNVMSEIPTEALDVTFPASGEKVEKLINKQSDSHFTDAQPEQQQREEDRKGR
EPACE-10_S3     358
RNFLAGEEDNVMSEIPTEVLDVTFPASGEKVEKLINKQSDSHFTDAQPEQQQREEDRKGR
IT81D-1053_R1   361
RNFLAGEEDNVMSEIPTEVLDVTFPASGEKVEKLINKQSDSHFTDAQPEQQQREEDRKGR
IT81D-1053_R2   361
RNFLAGEEDNVMSEIPTEVLDVTFPASGEKVEKLINKQSDSHFTDAQPEQQQREEDRKGR
IT81D-1053_R3   361
RNFLAGEEDNVMSEIPTEVLDVTFPASGEKVEKLINKQSDSHFTDAQPEQQQREEDRKGR
                361 ******** *********
****************************************


EPACE-10_S1     421 KGPLSSILDTFY
EPACE-10_S2     421 KGPLSSILGTFY
EPACE-10_S3     418 KGPLSSILGTLY
IT81D-1053_R1   421 KGPLSSILDTFY
IT81D-1053_R2   421 KGPLSSILDTFY
IT81D-1053_R3   421 KGPLSSILDSLY
                421 ******** . *
```

Figure S9. (Continued).

Table S1. Matrix of pairwise comparisons of cDNA sequences encoding cowpea vicilins, which were cloned from genotypes EPACE-10 (S1, S2 and S3) and IT81D-105 (R1, R2 and R3). For each pair of compared sequences, the percentage of sequence identity (above the diagonal) and the number of different nucleotides (below the diagonal) between them are shown. These numbers were calculated based on the multiple sequence alignment shown in Fig. S8.

	S1	S2	S3	R1	R2	R3
S1	−	99.7	99.6	99.9	99.3	99.3
S2	3	−	99.6	99.6	99.0	99.1
S3	4	5	−	99.6	99.1	99.2
R1	1	4	5	−	99.3	99.3
R2	9	12	11	8	−	98.6
R3	8	11	10	9	17	−

Table S2. Matrix of pairwise comparisons of vicilin amino acid sequences from cowpea genotypes EPACE-10 (S1, S2 and S3) and IT81D-105 (R1, R2 and R3). For each pair of compared sequences, the percentage of sequence identity (above the diagonal) and the number of different amino acid residues (below the diagonal) between them are shown. These numbers were calculated based on the multiple sequence alignment shown in Fig. S9.

	S1	S2	S3	R1	R2	R3
S1	−	99.3	99.3	100.0	99.0	98.3
S2	3	−	99.0	99.3	98.3	97.6
S3	3	4	−	99.3	98.3	98.1
R1	0	3	3	−	99.0	98.3
R2	4	7	7	4	−	97.4
R3	7	10	8	7	11	−

Table S3. Calculated molecular masses and theoretical isoelectric point (pI) values of cowpea β-vignins. The values were calculated by submitting the amino acid sequences to ExPASy's Compute pI/Mw tool (web.expasy.org/compute_pi/)

Amino acid sequence	Molecular mass (Da)	pI
S1	49827.45	5.22
S2	49683.32	5.33
S3	49347.02	5.38
R1	49827.45	5.22
R2	49803.43	5.27
R3	49717.30	5.17

Table S4. Results of the BLAST searches against the Conserved Domain Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) using as query the cowpea vicilin sequences from genotypes EPACE-10 (sequences S1, S2 and S3) and IT81D-1053 (R1, R2 and R3)

Sequence	Domain	Accession	Bit score	Interval (size)	E-value
S1	Cupin_1	smart00835	78.86	34-180 (147)	1.68e-17
	Cupin_1	smart00835	122.77	236-392 (157)	1.40e-33
S2	Cupin_1	smart00835	78.48	34-180 (147)	1.88e-17
	Cupin_1	smart00835	121.62	236-392 (157)	4.16e-33
S3	Cupin_1	smart00835	78.86	34-180 (147)	1.64e-17
	Cupin_1	smart00835	120.08	233-389 (157)	1.36e-32
R1	Cupin_1	smart00835	78.86	34-180 (147)	1.68e-17
	Cupin_1	smart00835	122.77	236-392 (157)	1.40e-33
R2	Cupin_1	smart00835	77.32	34-180 (147)	6.03e-17
	Cupin_1	smart00835	122.00	236-392 (157)	3.18e-33
R3	Cupin_1	smart00835	77.71	34-180	3.55e-17
	Cupin_1	smart00835	122.39	236-392	2.33e-33

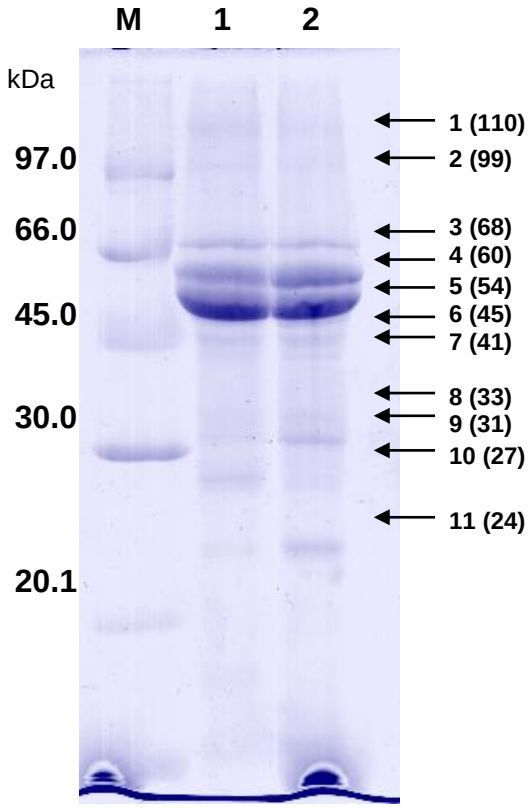


Figure S10. SDS-polyacrylamide gel electrophoresis of seed vicilins from cowpea (*V. unguiculata*). SDS-PAGE was performed as described by Laemmli (Laemmli, 1970) using 15% slab gels. Samples (20 µg) were prepared in 0.0625 M Tris-HCl (pH 6.8) containing 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 10% (v/v) glycerol and 0.001% (w/v) bromophenol blue. Protein bands were stained with 0.2% (w/v) Coomassie Brilliant Blue R250 in 50% methanol/10% acetic acid for 16 h. Destaining was carried out with 12.5% isopropanol/10% acetic acid. Lane 1: molecular weight markers. Lane 2: total vicilins extracted from genotype IT81D-1053. Lane 3: total vicilins extracted from genotype EPACE-10. The apparent molecular masses of the major protein bands identified in each fraction (labeled 1 to 11) are shown in parentheses.

Table S5. Identification of the protein bands from the vicilin fraction of cowpea (*V. unguiculata*) genotype IT81D-1053 by ESI-MS/MS. MS/MS ions searches against the NCBIprot database were performed using Mascot through the software's web server.

Protein band	Protein hit showing the most significant peptide matches	Mascot score[1]	Sequence coverage (%)	GenBank accession number
1	Vicilin protein (*V. unguiculata*)	241	7	CAP19902
2	Vicilin protein (*V. unguiculata*)	228	8	CAP19902
3	beta-conglycinin, beta chain-like (*Vigna radiata* var. radiata)	71	2	XP_014492536
4	beta-conglycinin, beta chain-like (*V. angularis*)	326	13	XP_017433627
5	Vicilin protein (*V. unguiculata*)	488	22	CAP19902
6	Adzuki bean (*V. angularis*) 7S globulin-1	100	4	2EA7
7	Vicilin protein (*V. unguiculata*)	161	6	CAP19902
8	Vicilin protein (*V. unguiculata*)	158	5	CAP19902
9	Adzuki bean (*V. angularis*) 7S globulin-1	91	4	2EA7
10	No significant hits	-	-	-
11	Vicilin protein (*V. unguiculata*)	163	5	CAP19902

Table S6. Identification of the protein bands from the vicilin fraction of cowpea (*V. unguiculata*) genotype EPACE-10 by ESI-MS/MS. MS/MS ions searches against the NCBIprot database were performed using Mascot through the software's web server.

Protein band	Protein hit showing the most significant peptide matches	Mascot score[1]	Sequence coverage (%)	GenBank accession number
1	Vicilin protein (*V. unguiculata*)	306	15	CAP19902
2	Vicilin protein (*V. unguiculata*)	235	7	CAP19902
3	Vicilin protein (*V. unguiculata*)	162	5	CAP19902
4	beta-conglycinin, beta chain-like (*V. angularis*)	375	11	XP_017433627
5	Vicilin protein (*V. unguiculata*)	495	19	CAP19902
6	Vicilin protein (*V. unguiculata*)	247	7	CAP19902
7	Vicilin protein (*V. unguiculata*)	231	8	CAP19902
8	Adzuki bean (*V. angularis*) 7S globulin-1	91	4	2EA7
9	Adzuki bean (*V. angularis*) 7S globulin-1	91	4	2EA7

10	Seed storage protein A (*V. luteola*)	83	4	AAZ06660
11	Vicilin protein (*V. unguiculata*)	227	6	CAP19902

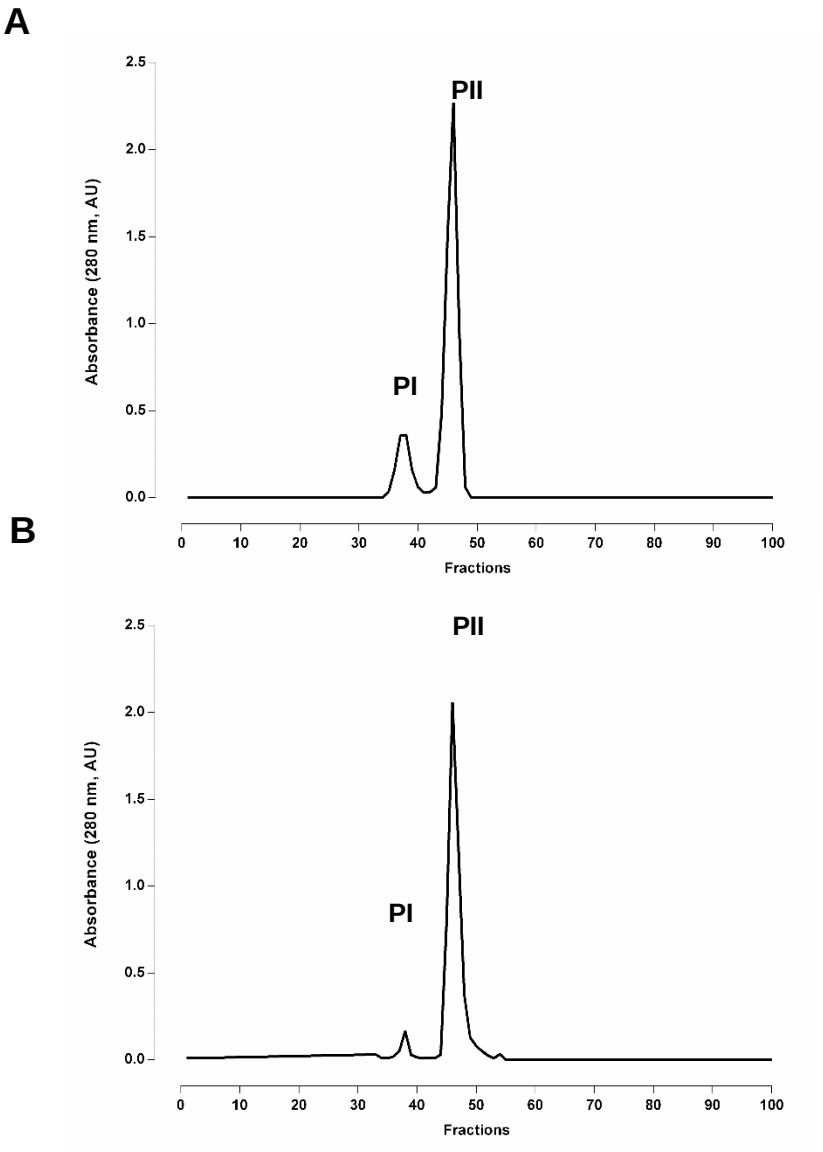


Figure S11. Size exclusion chromatography of cowpea vicilins on a column of Sephacryl S-200 HR coupled to an ÄKTA prime plus chromatography system. Vicilins from genotypes IT81D-1053 (**A**) and EPACE-10 (**B**) were resuspended in 50 mM Tris-HCl buffer, pH 8.0 (5 mg/mL), and 10 mg of each sample were loaded onto the column. The column was equilibrated and eluted with 50 mM Tris-HCl buffer, pH 8.0, at a constant flow rate (0.5 mL/min) and 3 mL fractions were collected.

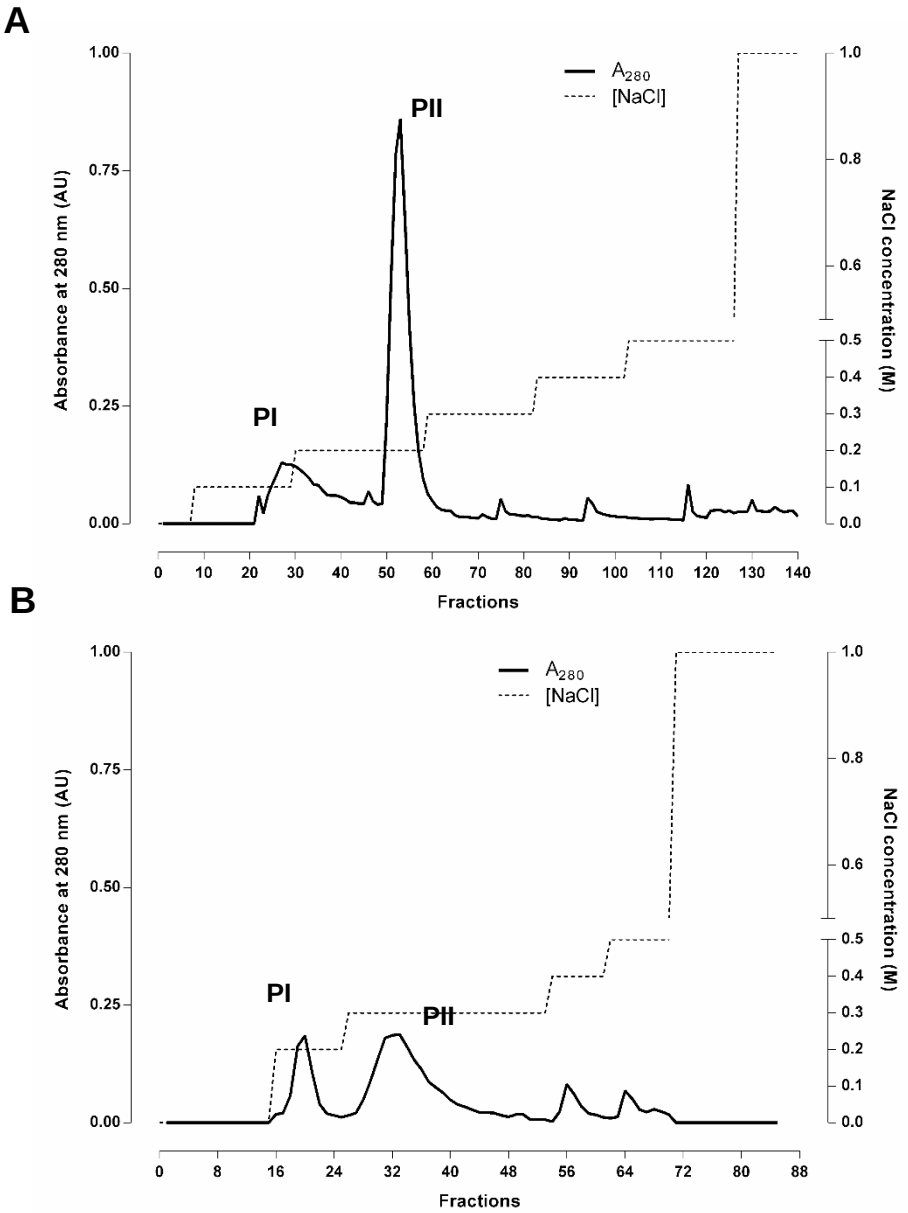


Figure S12. Ion exchange chromatography of cowpea vicilins on DEAE-Sepharose. Vicilins from cowpea genotypes IT81D-1053 (**A**) and EPACE-10 (**B**) were loaded onto a DEAE Sepharose Fast Flow (GE Healthcare) column (6 mL) equilibrated with 50 mM Tris-HCl buffer pH 8.0. The chromatography was performed at a constant flow rate (2 mL/min) and 3 mL fractions were collected.

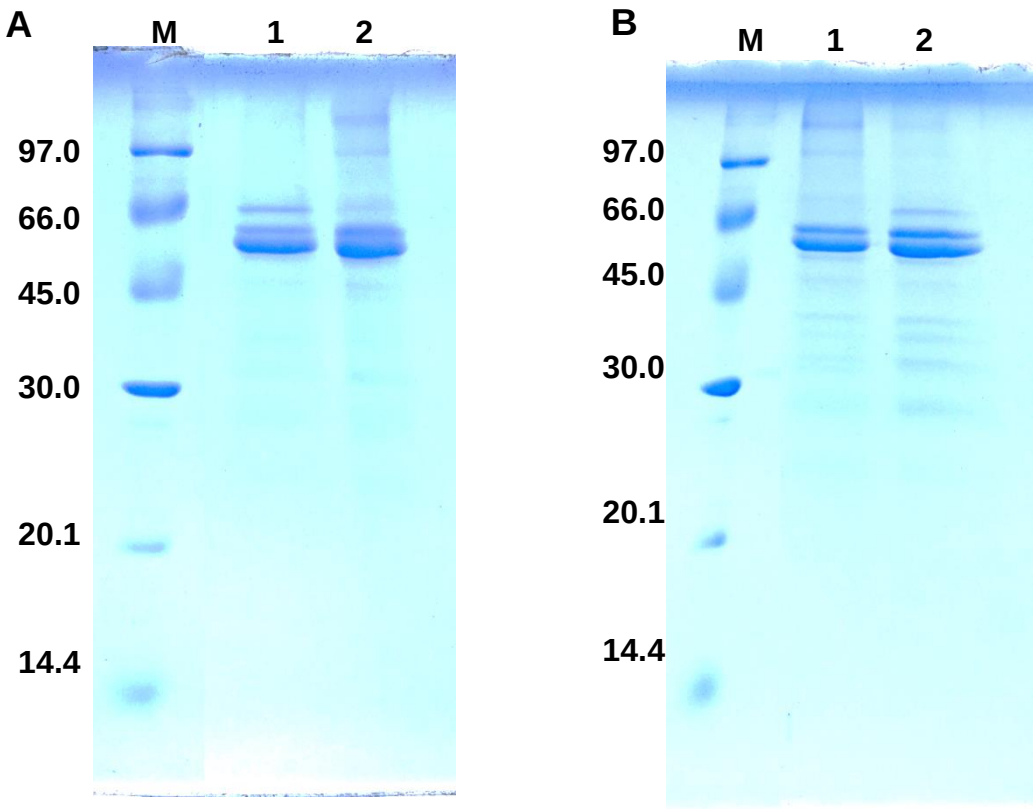


Figure S13. SDS-polyacrylamide gel electrophoresis of seed vicilins from cowpea (*V. unguiculata*) purified by ion exchange chromatography. Peaks I (lane 1; 20 µg) and II (lane 2; 20 µg) from the ion exchange chromatography of seed vicilins from genotypes IT81D-1053 (**A**) and EPACE-10 (**B**) were subjected to SDS-PAGE (15% polyacrylamide) and protein bands were stained with 0.2% (w/v) Coomassie Brilliant Blue R250 in 50% methanol/10% acetic acid for 16 h. Destaining was carried out with 12.5% isopropanol/10% acetic acid. Lane M: molecular weight markers.

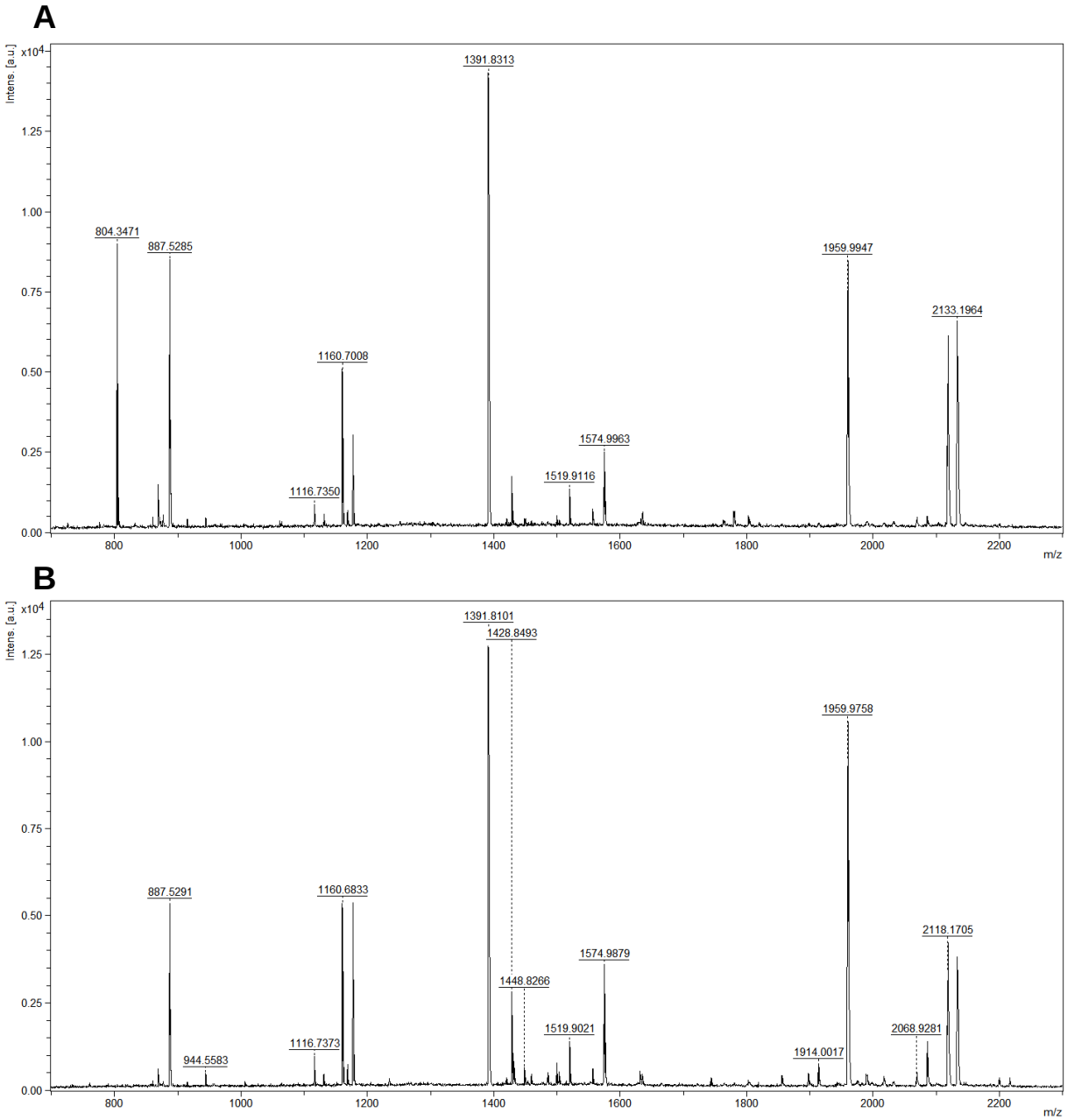


Figure S14. MALDI-TOF mass spectra of tryptic digest of cowpea seed vicilins from genotype IT81D-1053. Spectra are from tryptic digestion of proteins from peak I (**A**) and peak II (**B**), as obtained by ion exchange chromatography. Assignment of masses to appropriate amino acid sequences is shown in Table S3.

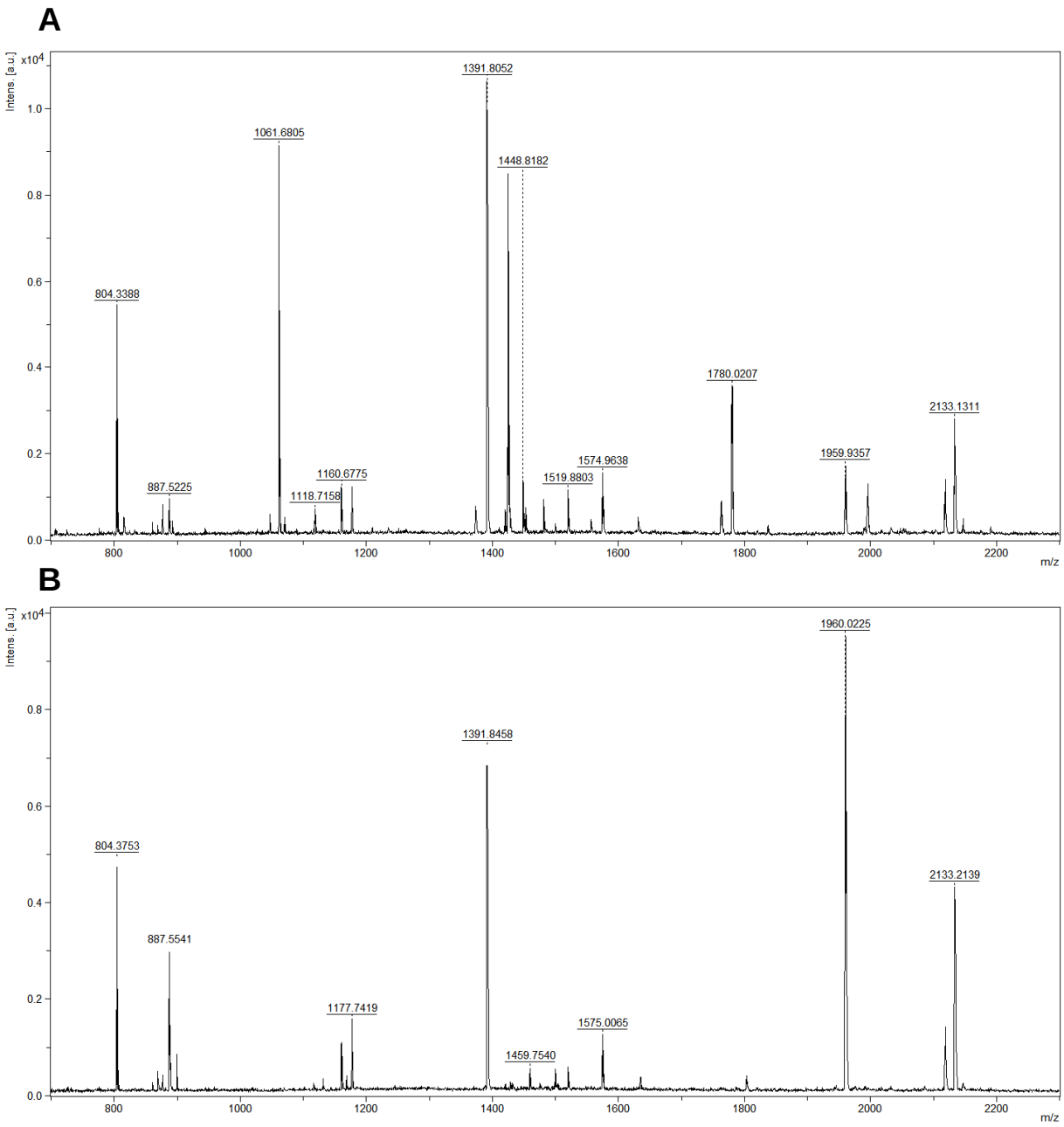


Figure S15. MALDI-TOF mass spectra of tryptic digest of cowpea seed vicilins from genotype EPACE-10. Spectra are from tryptic digestion of proteins from peak I (**A**) and peak II (**B**), as obtained by ion exchange chromatography. Assignment of masses to appropriate amino acid sequences is shown in Table S4.

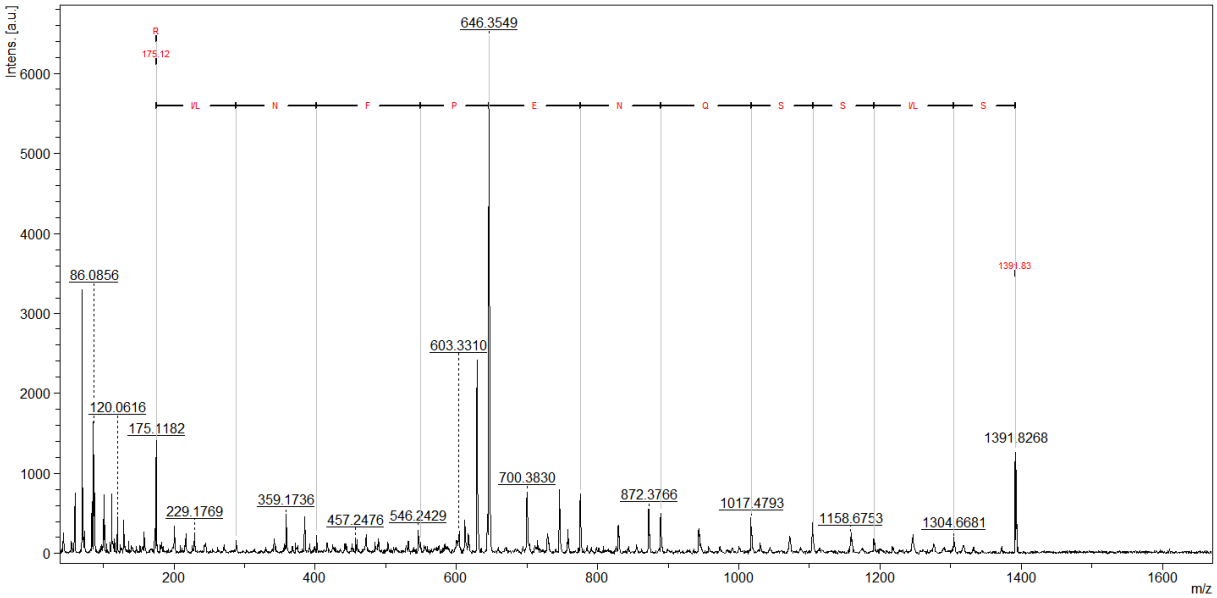


Figure S16. MS/MS spectra of the ion at m/z 1391. Amino acid sequence specific y-ions are indicated.

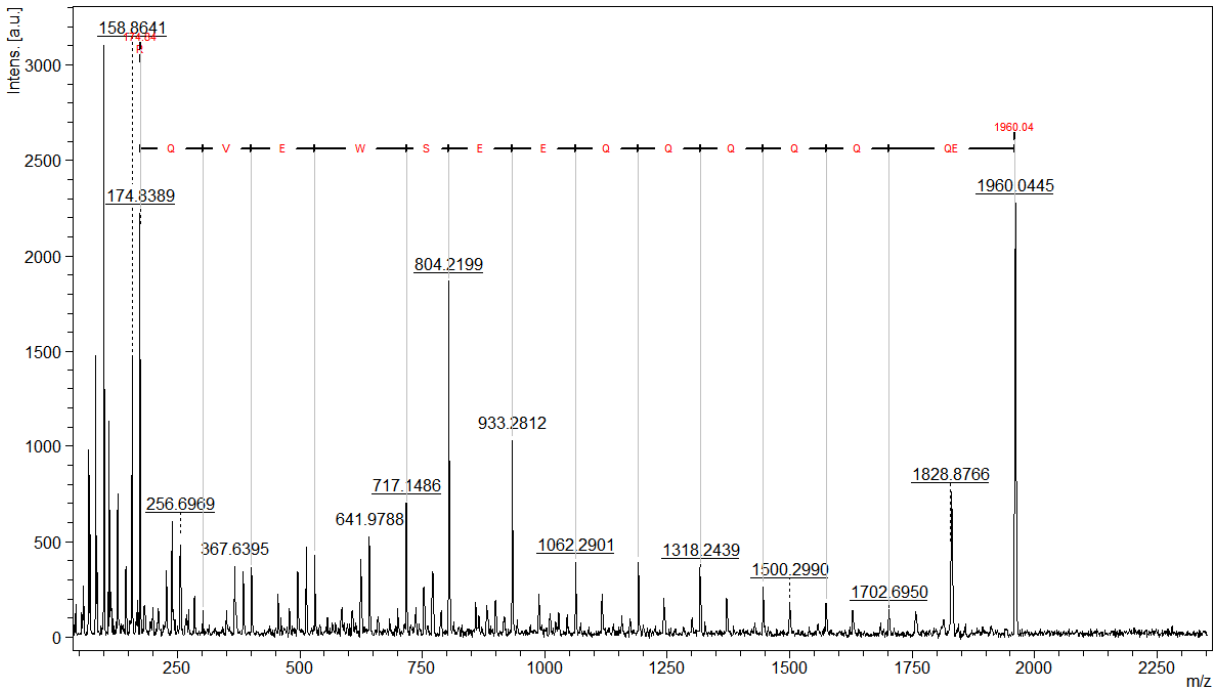


Figure S17. MS/MS spectra of the ion at m/z 1960. Amino acid sequence specific y-ions are indicated.

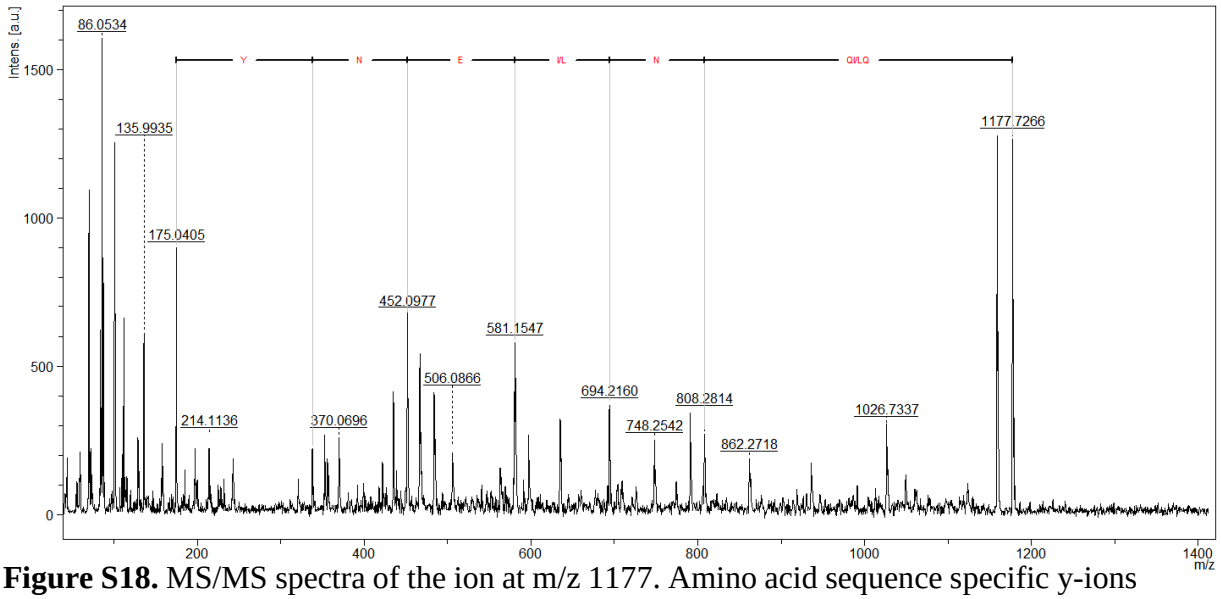


Figure S18. MS/MS spectra of the ion at m/z 1177. Amino acid sequence specific y-ions are indicated.

Table S7. Tryptic peptides from cowpea seed vicilins, purified from genotype IT81D-1053, as identified by MALDI-TOF mass spectrometry

Amino acid sequence	Sample	Mass			
		Identification method	Observed	Calculated	Δ (Da)
^{36}NQYGHLR^{42}	PI, PII	PMF	887.5275	887.4485	0.08
^{53}QIQNLENYR^{61}	PII	MS/MS	1177.7266	1177.5960	0.13
^{225}SLSSQNEPFNLR^{236}	PI	MS/MS	1391.8260	1391.6913	0.13
^{224}KSLSSQNEPFNLR^{236}	PI, PII	PMF	1519.9216	1519.7863	0.13
^{249}LHEITPEKNPQLR^{261}	PII	MS/MS	1574.9662	1574.8649	0.10
^{306}EQQQQQQEESWEVQR^{320}	PII	MS/MS	1960.0445	1959.8791	0.16
^{114}IPAGTTFFLVNPNDNDNLR^{132}	PI	MS/MS	2118.2031	2118.0614	0.14

PMF: peptide mass fingerprint; MS/MS: tandem mass spectrometry. Numbers in superscript represent the amino acid positions in the amino acid sequence deduced from cDNA sequences cloned from genotype IT81D-1053.

Table S8. Tryptic peptides from cowpea seed vicilins, purified from genotype EPACE-10, as identified by MALDI-TOF mass spectrometry

Amino acid sequence	Sample	Mass			
		Identification method	Observed	Calculated	Δ (Da)
^{36}NQYGHLR^{42}	PI, PII	PMF	887.5235	887.4485	0.07
^{43}VLQRFDQR^{50}	PI	MS/MS	1061.6843	1061.5850	0.09
^{225}SLSSQNEPFNLR^{236}	PI	MS/MS	1391.8268	1391.6913	0.12
^{237}SQKPIYSNKFGR^{248}	PI, PII	PMF	1424.8711	1424.7644	0.10
^{224}KSLSSQNEPFNLR^{236}	PI, PII	PMF	1519.8582	1519.7863	0.07
^{249}LHEITPEKNPQLR^{261}	PII	MS/MS	1574.9662	1574.8649	0.10
^{53}QIQNLENYRVVEFK^{66}	PI	MS/MS	1779,9870	1779.9388	0.04
^{306}EQQQQQQEESWEVQR^{320}	PII	MS/MS	1959.9955	1959.8791	0.11
^{114}IPAGTTFFLVNPDDNENLR^{132}	PII	MS/MS	2133.2380	2133.0611	0.18

PMF: peptide mass fingerprint; MS/MS: tandem mass spectrometry. Numbers in superscript represent the amino acid positions in the amino acid sequence deduced from cDNA sequences cloned from genotype EPACE-10.

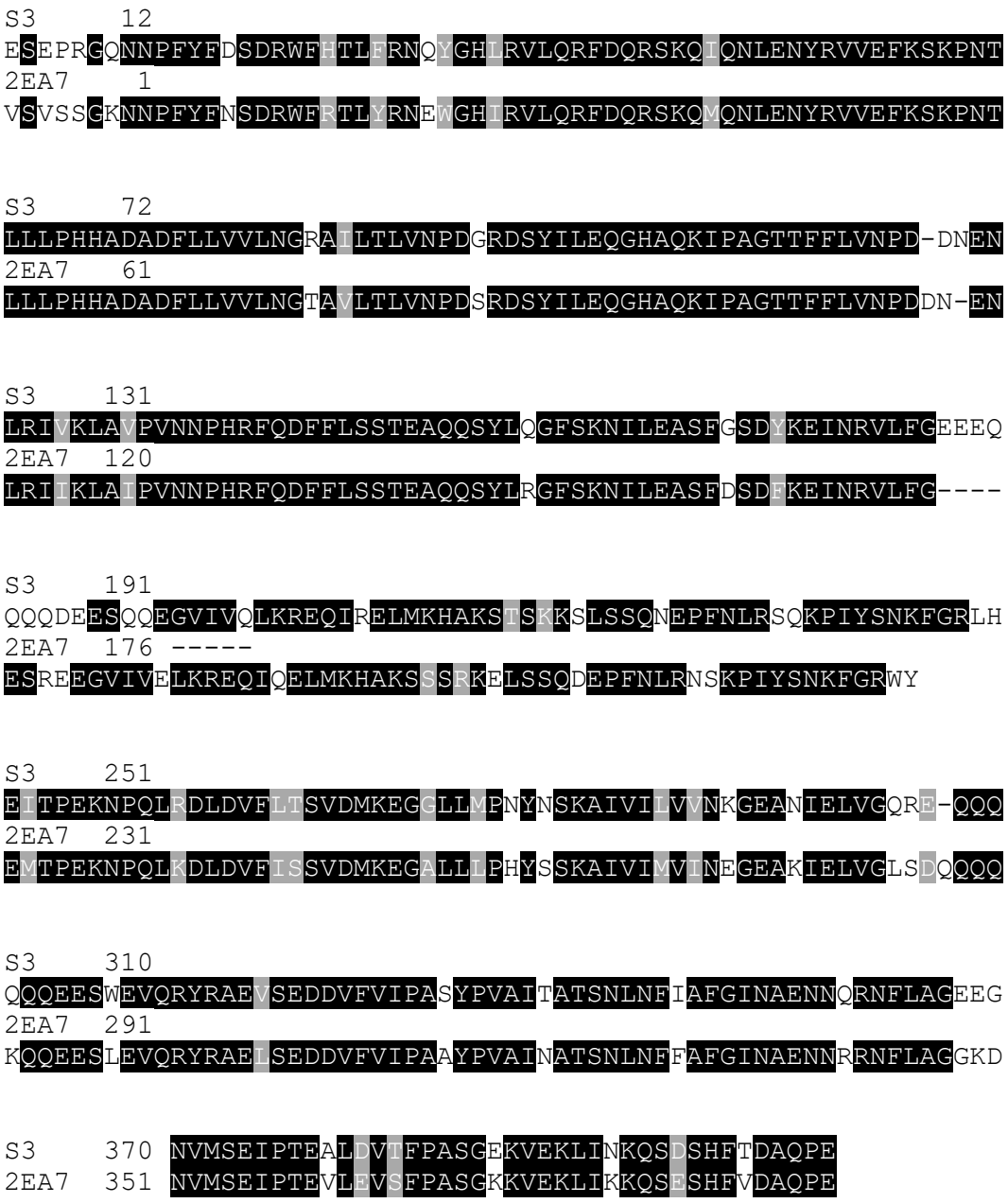


Figure S19. Alignment of the amino acid sequence of cowpea vicilin from genotype EPACE-10 (sequence S2) with the primary structure of the 7S globulin-1 (PDB ID: 2EA7) from adzuki bean (*V. angularis*). Identical and chemically-similar residues are shaded in black and gray, respectively. The alignment was rendered using BOXSHADE v3.21 through the web server embnet.vital-it.ch/software/BOX_form.html.

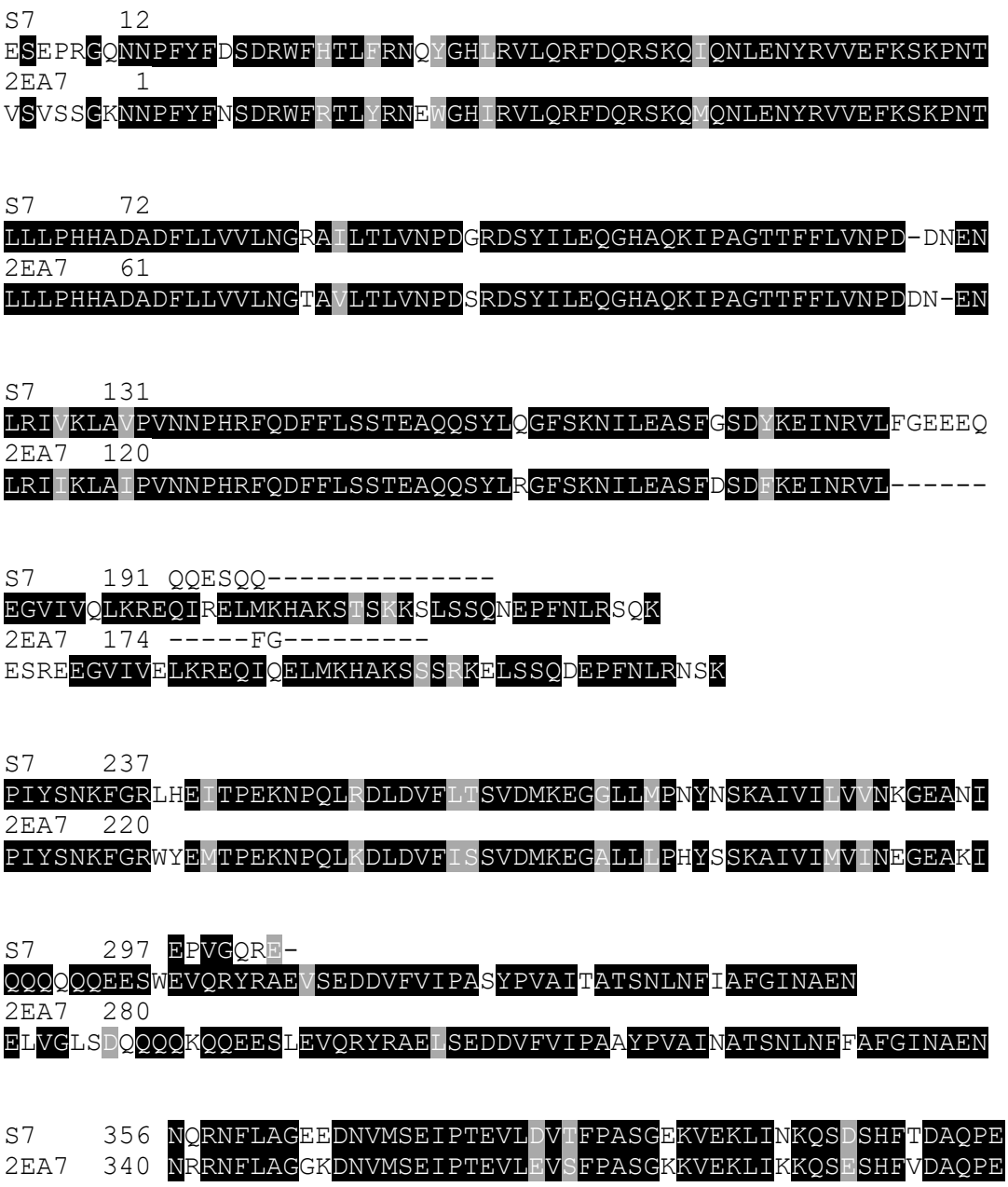


Figure S20. Alignment of the amino acid sequence of cowpea vicilin from genotype EPACE-10 (sequence S3) with the primary structure of the 7S globulin-1 (PDB ID: 2EA7) from adzuki bean (*V. angularis*). Identical and chemically-similar residues are shaded in black and gray, respectively. The alignment was rendered using BOXSHADE v3.21 through the web server embnet.vital-it.ch/software/BOX_form.html.

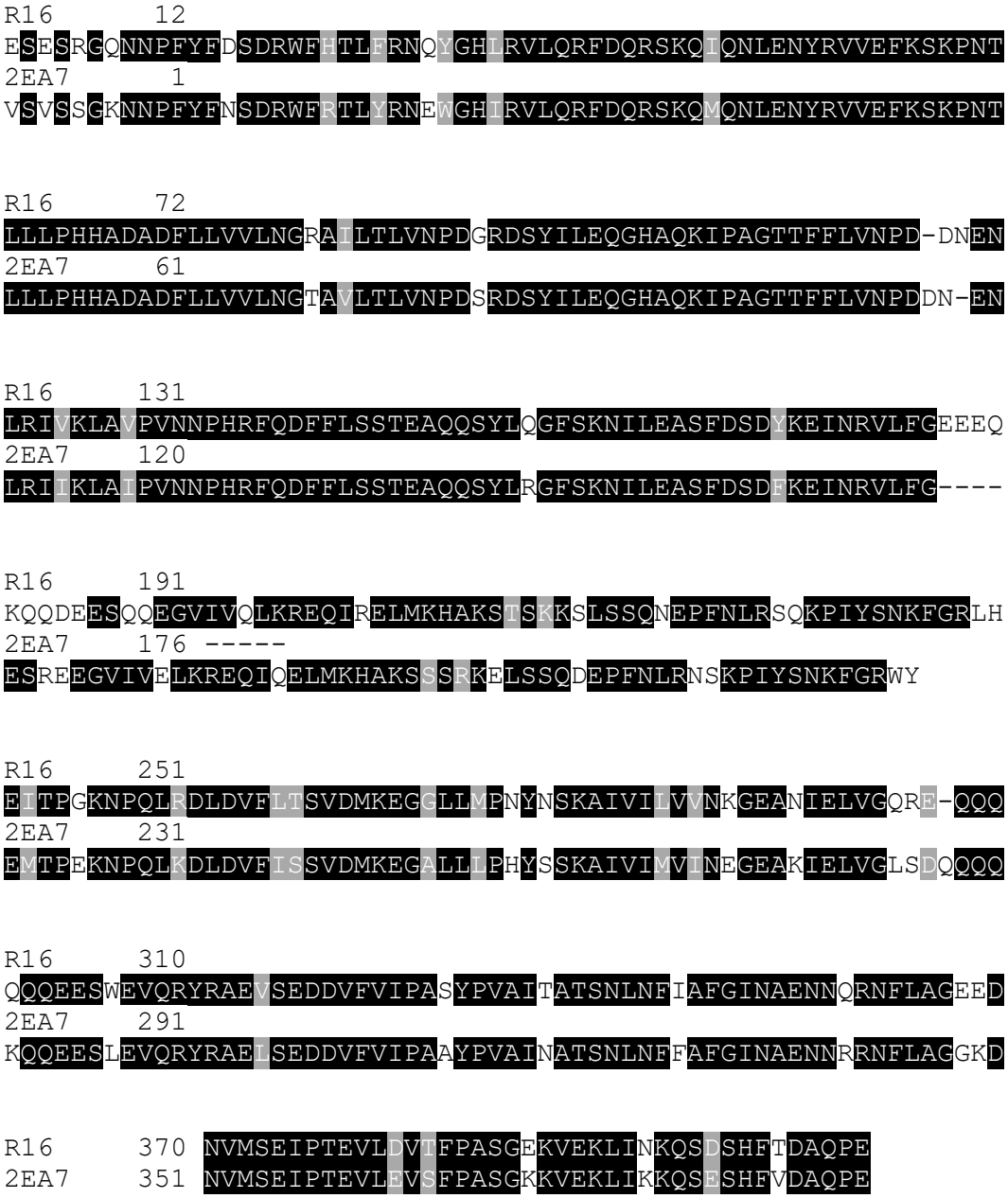


Figure S21. Alignment of the amino acid sequence of cowpea vicilin from genotype IT81D-1053 (sequence R2) with the primary structure of the 7S globulin-1 (PDB ID: 2EA7) from adzuki bean (*V. angularis*). Identical and chemically-similar residues are shaded in black and gray, respectively. The alignment was rendered using BOXSHADE v3.21 through the web server embnet.vital-it.ch/software/BOX_form.html.

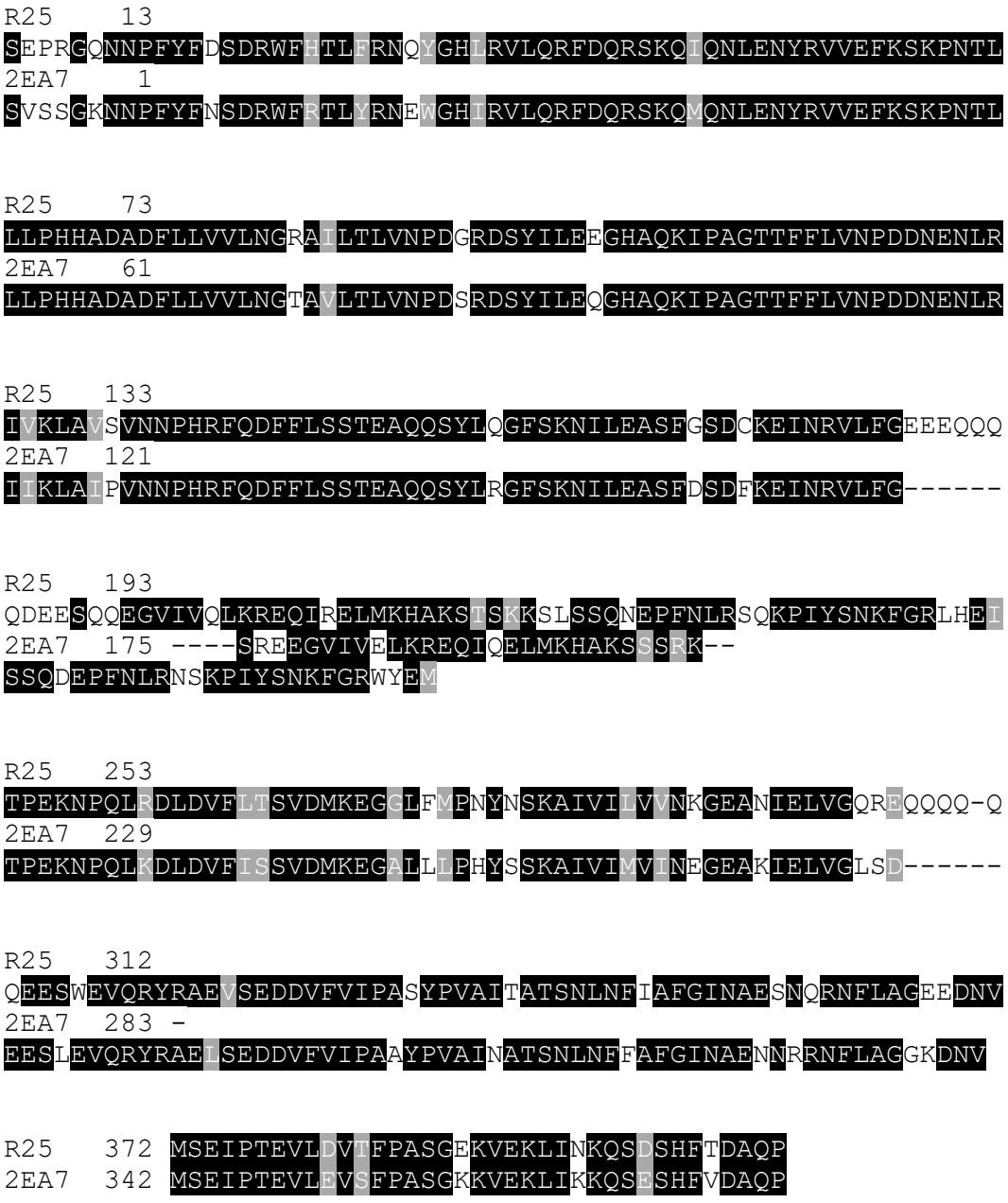


Figure S22. Alignment of the amino acid sequence of cowpea vicilin from genotype IT81D-1053 (sequence R3) with the primary structure of the 7S globulin-1 (PDB ID: 2EA7) from adzuki bean (*V. angularis*). Identical and chemically-similar residues are shaded in black and gray, respectively. The alignment was rendered using BOXSHADE v3.21 through the web server embnet.vital-it.ch/software/BOX_form.html.

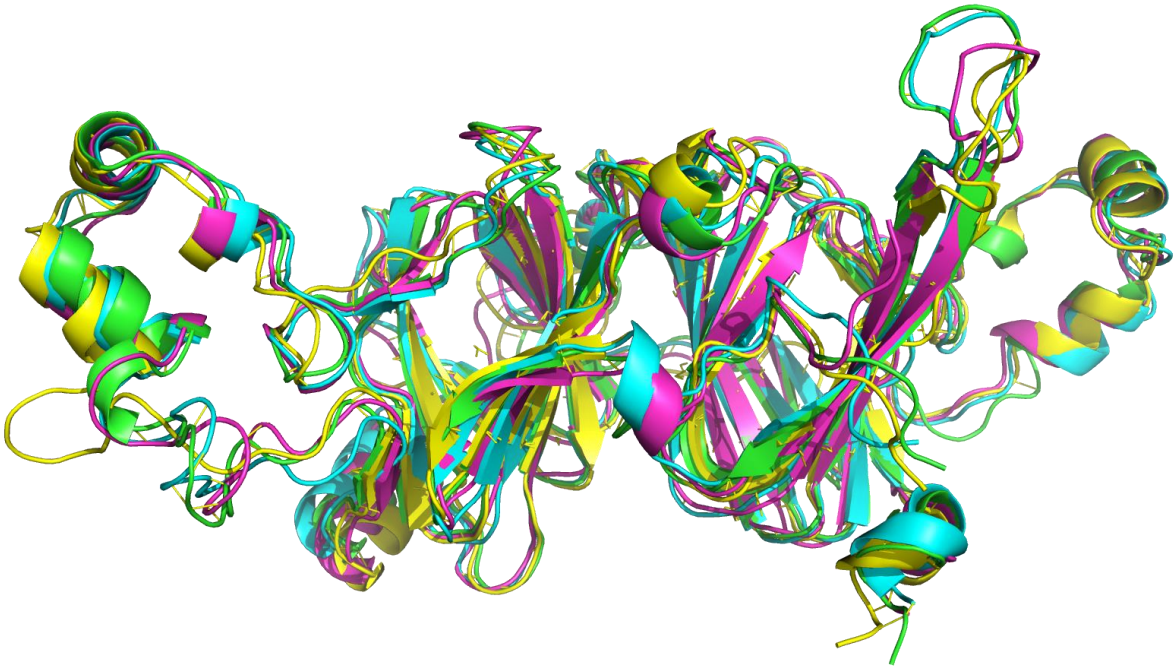


Figure S23. Superposition of the three-dimensional molecular models of β-vignin. The superposed models, shown as ribbon diagrams, are colored pink (model S2), yellow (model S3), green (model R2) and cyan (model R3). The modeled structures were superposed by overlapping 354 Cα-atoms of structurally equivalent residues with an RMSD of 1.479 Å.

Table S9. Summary statistics of the MolProbity analysis of the three-dimensional molecular model of cowpea vicilin from sequence S2. The analysis was performed by submitting the three-dimensional molecular model to the MolProbity web server (http://molprobity.biochem.duke.edu/)

	Raw model[1]	Refined model[2]
(a) All-atom contacts		
Clashscore, all atoms (N=1784, all resolutions)	59.86 (2nd percentile*)	6.65 (88th percentile*)
(b) Protein geometry		
Poor rotamers (Goal: <0.3%)	10 (2.57%)	2 (0.51%)
Favored rotamers (Goal: >98%)	350 (89.97%)	383 (98.46%)
Ramachandran outliers (Goal: <0.05%)	4 (0.93%)	0 (0.00%)
Ramachandran favored (Goal: >98%)	405 (94.19%)	421 (97.91%)
MolProbity score (N=27675, 0Å - 99Å)	2.95 (25th percentile*)	1.39 (97th percentile*)
Cβ deviations >0.25Å (Goal: 0%)	1 (0.24%)	18 (4.39%)
Bad bonds (Goal: 0%)	0 / 3581 (0.00%)	0 / 3581 (0.00%)
Bad angles (Goal: <0.1%)	64 / 4838 (1.32%)	63 / 4838 (1.30%)
(c) Peptide omegas		
Cis prolines	0 / 20 (0.00%)	0 / 20 (0.00%)

* 100th percentile is the best among structures of comparable resolution; 0th percentile is the worst. For clashscore the comparative set of structures was selected in 2004, for MolProbity score in 2006.

^ MolProbity score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution.

Table S10. Summary statistics of the MolProbity analysis of the three-dimensional molecular model of cowpea vicilin from sequence S3. The analysis was performed by submitting the three-dimensional molecular model to the MolProbity web server (http://molprobity.biochem.duke.edu/)

	Raw model[1]	Refined model[2]
(a) All-atom contacts		
Clashscore, all atoms (N=1784, all resolutions)	86.67 (0th percentile*)	13.52 (56th percentile*)
(b) Protein geometry		
Poor rotamers (Goal: <0.3%)	18 (4.64%)	1 (0.26%)
Favored rotamers (Goal: >98%)	350 (90.21%)	373 (96.13%)
Ramachandran outliers (Goal: <0.05%)	2 (0.47%)	3 (0.70%)
Ramachandran favored (Goal: >98%)	413 (96.72%)	417 (97.66%)
MolProbity score (N=27675, 0Å - 99Å)	3.12 (19th percentile*)	1.71 (89th percentile*)
Cβ deviations >0.25Å (Goal: 0%)	4 (0.98%)	23 (5.64%)
Bad bonds (Goal: 0%)	5 / 3557 (0.14%)	3 / 3557 (0.08%)
Bad angles (Goal: <0.1%)	82 / 4808 (1.71%)	79 / 4808 (1.64%)
(c) Peptide omegas		
Cis prolines	0 / 21 (0.00%)	0 / 21 (0.00%)

* 100th percentile is the best among structures of comparable resolution; 0th percentile is the worst. For clashscore the comparative set of structures was selected in 2004, for MolProbity score in 2006.

^ MolProbity score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution.

Table S11. Summary statistics of the MolProbity analysis of the three-dimensional molecular model of cowpea vicilin from sequence R2. The analysis was performed by submitting the three-dimensional molecular model to the MolProbity web server (http://molprobity.biochem.duke.edu/)

	Raw model[1]	Refined model[2]
(a) All-atom contacts		
Clashscore, all atoms (N=1784, all resolutions)	60.74 (2nd percentile*)	6.93 (87th percentile*)
(b) Protein geometry		
Poor rotamers (Goal: <0.3%)	11 (2.81%)	0 (0.00%)
Favored rotamers (Goal: >98%)	351 (89.54%)	384 (97.96%)
Ramachandran outliers (Goal: <0.05%)	2 (0.47%)	0 (0.00%)
Ramachandran favored (Goal: >98%)	415 (96.51%)	423 (98.37%)
MolProbity score ^ (N=27675, 0Å - 99Å)	2.82 (30th percentile*)	1.38 (97th percentile*)
Cβ deviations >0.25Å (Goal: 0%)	2 (0.49%)	19 (4.61%)
Bad bonds (Goal: 0%)	0 / 3588 (0.00%)	0 / 3588 (0.00%)
Bad angles (Goal: <0.1%)	66 / 4847 (1.36%)	66 / 4847 (1.36%)
(c) Peptide omegas		
Cis prolines (Expected: ≤1 per chain, or ≤5%)	0 / 19 (0.00%)	0 / 19 (0.00%)
Twisted peptides	1 / 412 (0.24%)	0/412 (0.00%)

* 100th percentile is the best among structures of comparable resolution; 0th percentile is the worst. For clashscore the comparative set of structures was selected in 2004, for MolProbity score in 2006.

^ MolProbity score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution.

Table S12. Summary statistics of the MolProbity analysis of the three-dimensional molecular model of cowpea vicilin from sequence R3. The analysis was performed by submitting the three-dimensional molecular model to the MolProbity web server (http://molprobity.biochem.duke.edu/)

	Raw model[1]	Refined model[2]
(a) All-atom contacts		
Clashscore, all atoms (N=1784, all resolutions)	51.8 (3rd percentile*)	8.1 (81st percentile*)
(b) Protein geometry		
Poor rotamers (Goal: <0.3%)	9 (2.30%)	1 (0.26%)
Favored rotamers (Goal: >98%)	358 (91.33%)	387 (98.72%)
Ramachandran outliers (Goal: <0.05%)	4 (0.93%)	2 (0.47%)
Ramachandran favored (Goal: >98%)	411 (95.58%)	419 (97.44%)
MolProbity score (N=27675, 0Å - 99Å)	2.77 (33rd percentile*)	1.55 (94th percentile*)
Cβ deviations >0.25Å (Goal: 0%)	2 (0.49%)	16 (3.88%)
Bad bonds (Goal: 0%)	0 / 3579 (0.00%)	0 / 3579 (0.00%)
Bad angles (Goal: <0.1%)	63 / 4834 (1.30%)	73 / 4834 (1.51%)
(c) Peptide omegas		
Cis prolines	0 / 19 (0.00%)	0 / 19 (0.00%)

* 100th percentile is the best among structures of comparable resolution; 0th percentile is the worst. For clashscore the comparative set of structures was selected in 2004, for MolProbity score in 2006.

^ MolProbity score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution.

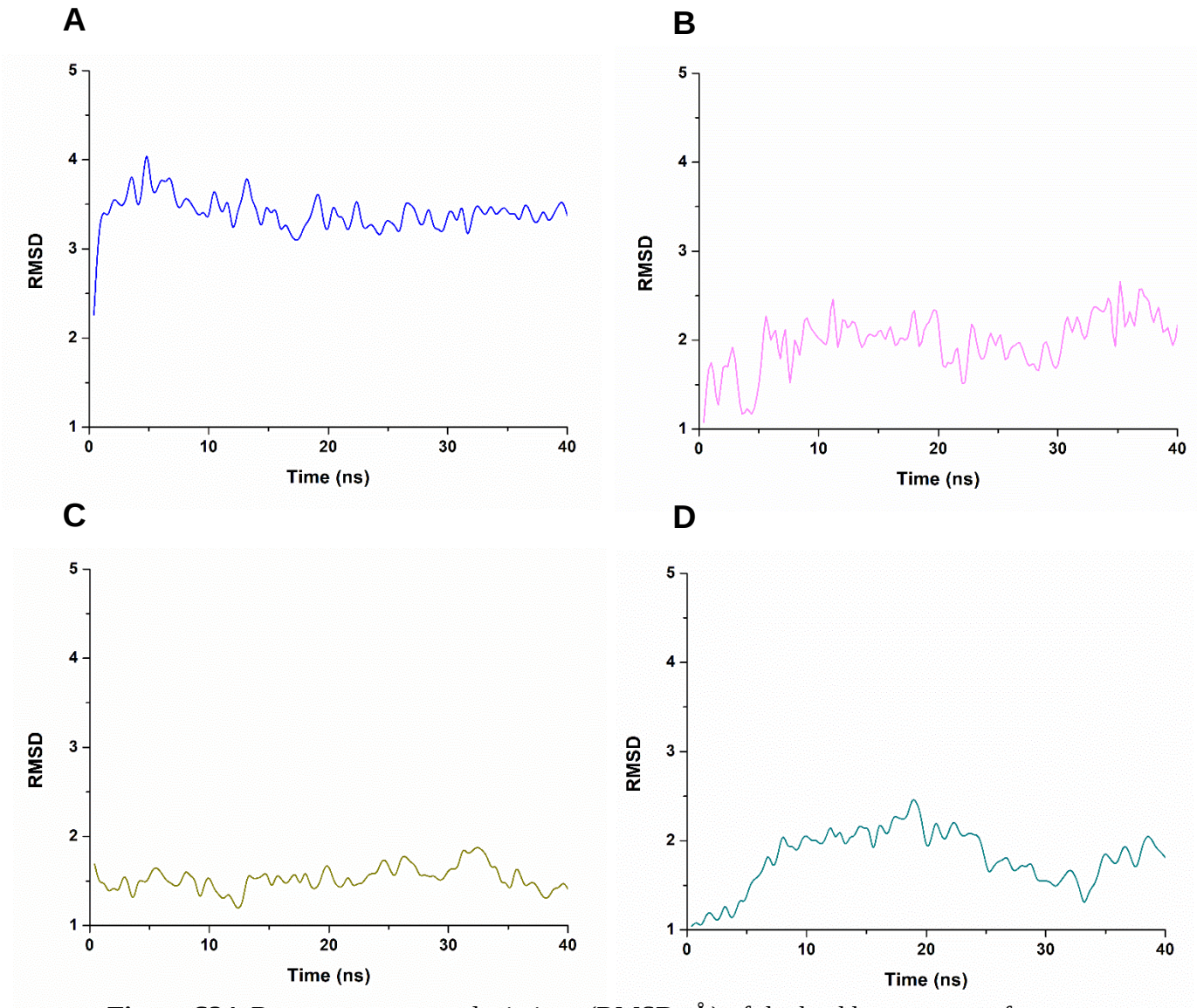


Figure S24. Root-mean-square deviations (RMSD; Å) of the backbone atoms of a tetra-*N*-acetyl-chitotetraose molecule [$(GlcNAc)_4$] docked in the chitin-binding site of cowpea vicilins, during the 40 ns MD simulations. The graphics show the backbone RMSD plots of $(GlcNAc)_4$ docked in the three-dimensional molecular models generated from sequences S2 (**A**), S3 (**B**), R2 (**C**) and R3 (**D**).

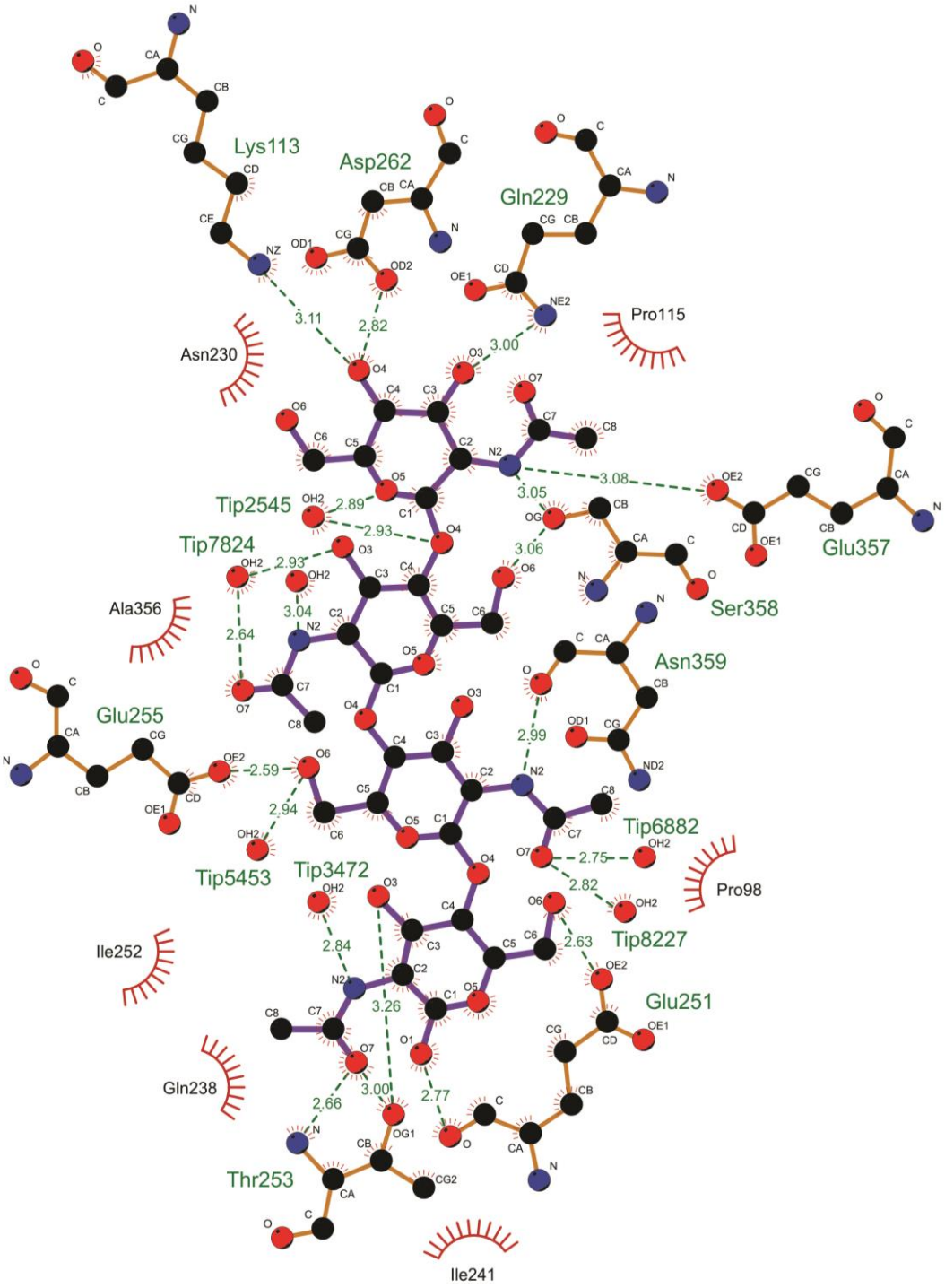


Figure S25. Two-dimensional diagram of the interaction model between a $(GlcNAc)_4$ molecule docked in the chitin-binding site of β-vignin. Residues from model R3 involved in hydrogen bonds (green dashed lines) or hydrophobic contacts (brick-red spoked arcs) are shown. Water molecules (OH2) are represented as red spheres.

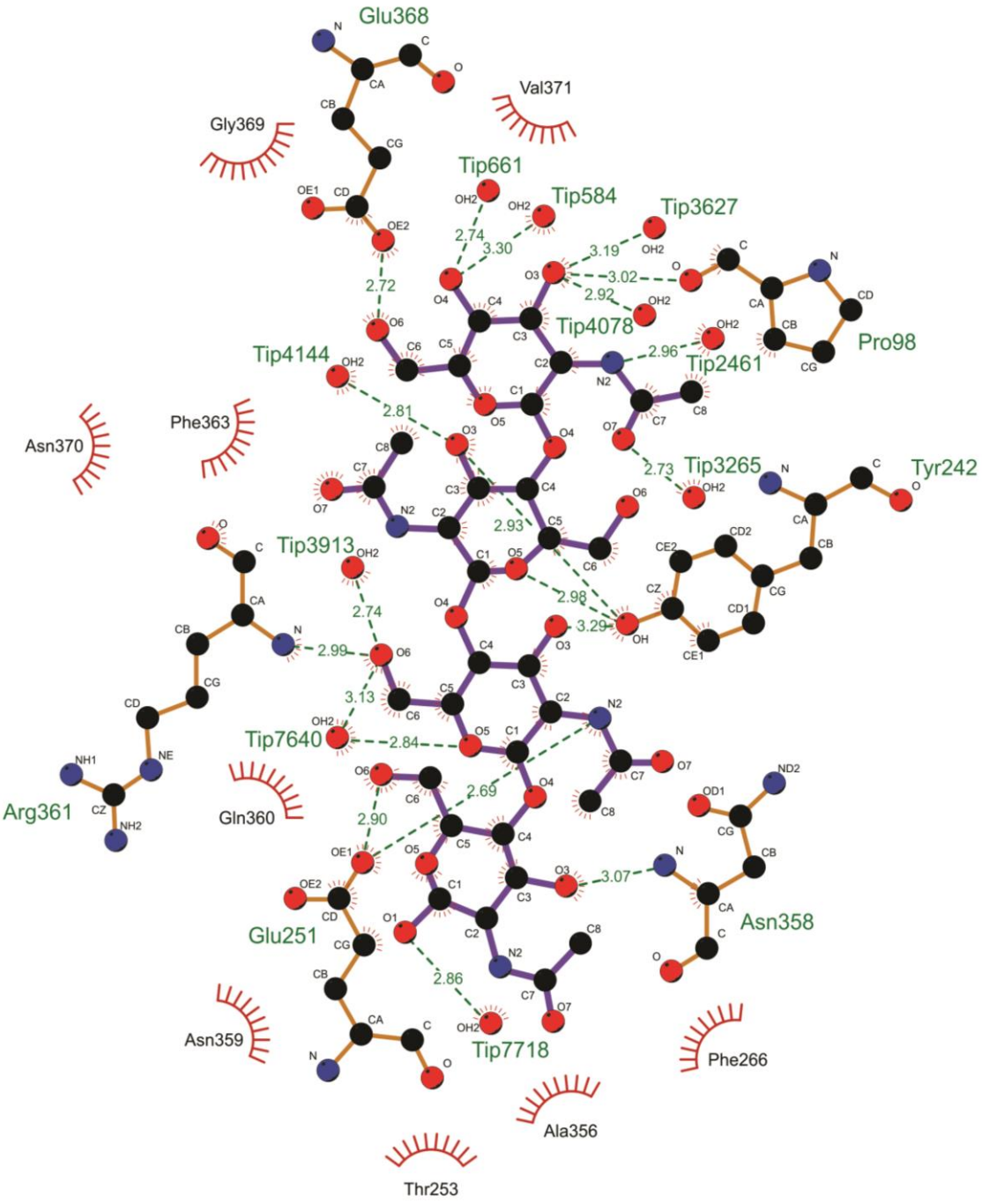


Figure S26. Two-dimensional diagram of the interaction model between a $(GlcNAc)_4$ molecule docked in the chitin-binding site of β-vignin. Residues from model S2 involved in hydrogen bonds (green dashed lines) or hydrophobic contacts (brick-red spoked arcs) are shown. Water molecules (OH2) are represented as red spheres.

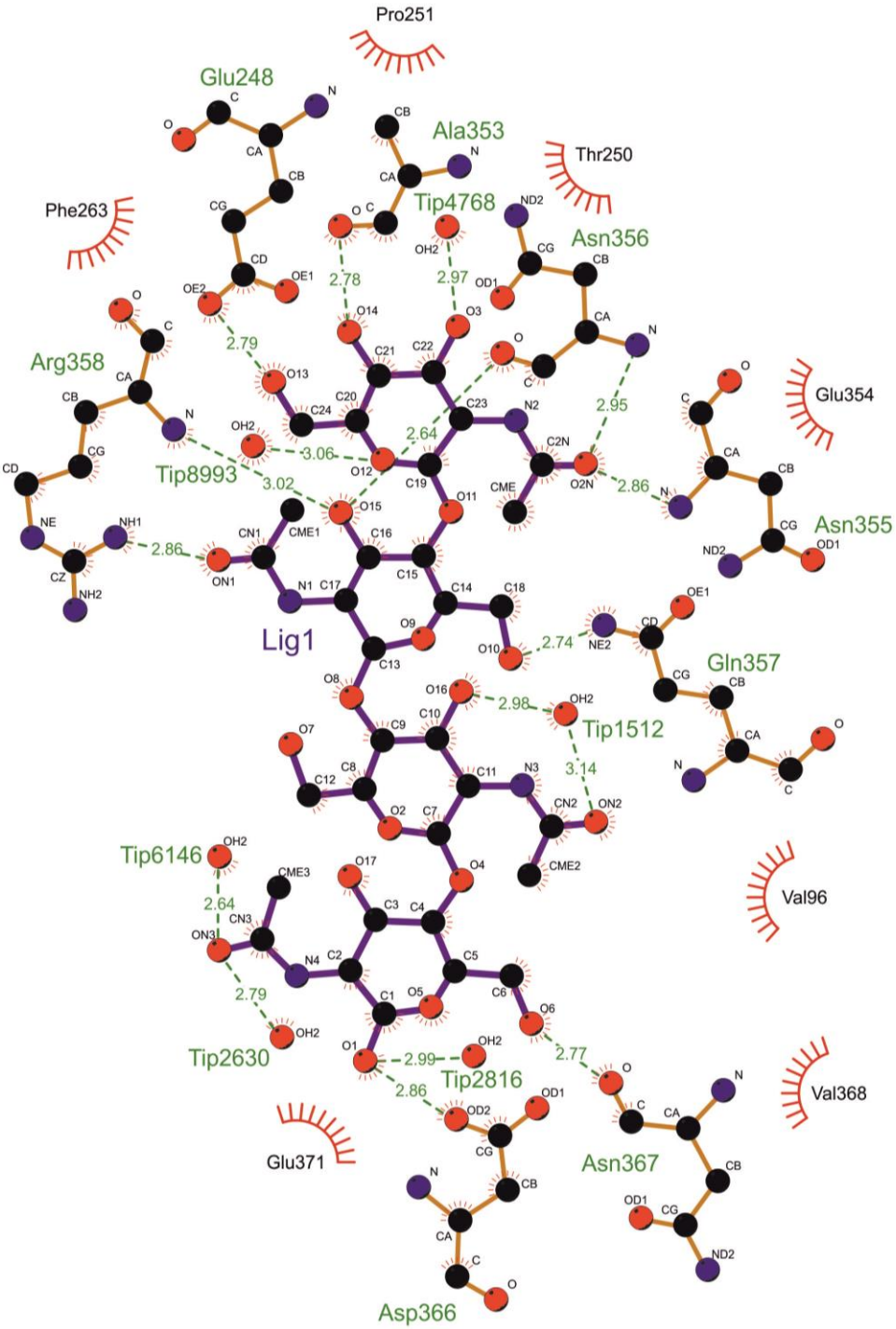


Figure S27. Two-dimensional diagram of the interaction model between a $(GlcNAc)_4$ molecule docked in the chitin-binding site of β-vignin. Residues from model S3 involved in hydrogen bonds (green dashed lines) or hydrophobic contacts (brick-red spoked arcs) are shown. Water molecules (OH2) are represented as red spheres.

Table S13. Calculated binding free energies (kcal/ml) of $(GlcNAc)_4$ molecules bound to β-vignin molecular models. The values were based on a quantum mechanical method (Maranhão et al., 2017).

Molecular model	One bound monosaccharide	Three bound monosaccharides
S2	−208.7	−626.1
S3	−180.7	−542.1
R2	−230.2	−690.6
R3	−177.1	−531.3

References

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685.

Maranhão, P.A.C., Teixeira, C.S., Sousa, B.L., Barroso-Neto, I.L., Monteiro-Júnior, J.E., Fernandes, A.V., Ramos, M.V., Vasconcelos, I.M., Gonçalves, J.F.C., Rocha, B.A.M., Freire, V.N., Grangeiro, T.B., 2017. cDNA cloning, molecular modeling and docking calculations of L-type lectins from Swartzia simplex var. grandiflora (Leguminosae, Papilionoideae), a member of the tribe Swartzieae. Phytochemistry 139, 60–71. https://doi.org/10.1016/j.phytochem.2017.04.007

7. CONSIDERAÇÕES FINAIS

Em suma, os genótipos de caupi contrastantes em relação à resposta ao C. maculatus contêm vicilinas (β-vignina) que possuem sítios de ligação a quitina. Além disso,os sítios de ligação a quitina, como identificados no presente trabalho, exibem um perfil de interação proteína-ligante semelhante ao encontrado em proteínas que classicamente interagem com carboidratos e que diferenças nas sequências de aminoácidos de β-vignina podem impactar nas energias de interação proteína-quitina.

REFERÊNCIAS

Aremu, C. Y. Proximate and amino acid composition of cowpea (*vigna unguiculata*, walp) protein concentrate prepared by isoeletric point precipitation. Food chemistry, v. 8, p. 61-68, 1990.

Aquino, R.erminação de motivos de ligação à quitina em vicilinas de *canavalia ensiformis* e *vigna unguiculata* através de métodos in silico e relação com suas toxicidades para o bruquídeo *callosobruchus* (coleoptera: bruchidae). 2009. Dissertação (mestrado em bioquímica) -centro de biociências, universidade federal do rio grande do norte, natal, 2009.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. Basic local alignment search tool. J. Mol. Biol. 215, 403–410, 1990.

Alva, V., Nam, S.-Z., Söding, J., Lupas, A.N. The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. Nucleic Acids Res. 44, W410–W415. https://doi.org/10.1093/nar/gkw348, 2016

Bastos, J. A. M. Principais pragas das culturas e seus controles. 3.ed. São paulo: nobel, 1985, 329p.

Baker, T.A., Suzanne, N., Shade, R.E., Singh, B.B. Physical and chemical attributes of cowpea lines resistant and susceptible to Callosobruchus maculatus (F.) (Coleoptera: Bruchidae). J. Stored Prod. Res. 25, 1–8. https://doi.org/10.1016/0022-474X(89)90002-7, 1989.

Best, R.B., Zhu, X., Shim, J., Lopes, P.E.M., Mittal, J., Feig, M., Mackerell, A.D. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. J. Chem. Theory Comput. 8, 3257–3273. https://doi.org/10.1021/ct300400x, 2012

Bohne, A., Lang, E., von der Lieth, C.W. SWEET - WWW-based rapid 3D construction of oligo- and polysaccharides. Bioinformatics 15, 767–768, 1999.

Bond, C.S., Schüttelkopf, A.W. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. Acta Crystallogr. D Biol. Crystallogr. 65, 510–512. https://doi.org/10.1107/S0907444909007835, 2009.

Cajazeiras, J. B. Identificação de genótipos de caupi [*vigna unguiculata* (l.) Walp.] Resistentes ao caruncho (*Callosobruchus maculatus* fabricius, 1792), 2000. Dissertação (mestrado em bioquímica) - centro de ciências, universidade federal do ceará, fortaleza, 2000.

Campos, F. A. P.; Xavier-Filho, J.; Silva, C. P.; Ary, M. B. Resolution and partial characterization of proteinases and *a*amylases from midguts of larvae of the bruchid beetle *callo*sobruchus maculatus (f.). Comp. Biochemistry physiology, v. 92, p. 51-57, 1989.

Cerdeira, A.L., Cole, A.W., Luthe, D.S. Cowpea (Vigna unguiculata) Seed Protein Response to Glyphosate. Weed Sci. 33, 1–6. https://doi.org/10.1017/S0043174500083843, 1985.

Chang, S., Puryear, J., Cairney, J. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Report. 11, 113–116. https://doi.org/10.1007/BF02670468, 1993.

Credland, P.F. Effect of host availability on reproductive perfomance in *Callosobruchus maculatus* (f.) (coleoptera: bruchidae). Journal of stored products research, v. 22, p. 49-54, 1986.

Chen, V.B., Arendall, W.B., 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. 66, 12–21. https://doi.org/10.1107/S0907444909042073, 2010.
Dick, K. M.; Credland, P. F. Chances in response of *Callosobruchus maculatus* (f.) (coleoptera: bruchidae) to a resistant variety of cowpea.journal of stored products research, v. 22, p. 227-233, 1986.
Dodeman, V. L.; Ducreux, G.; Kreis, M. Zygotic embryogenesis *versus* somatic embryogenesis.journal of experimental botany, v. 48, p.1493-1509, 1997.
Duranti, M.; Gius, C. Legume seeds: protein content and nutritional value. Field crops research, v. 53, p.31-45, 1997.
Edvardsson, M., Tregenza, T. Why do male callosobruchus maculatus harm their mates. Behavioural ecology, v.16, p. 788-793, 2005.
de Souza Ferreira, E., Capraro, J., Sessa, F., Magni, C., Demonte, A., Consonni, A., Augusto Neves, V., Maffud Cilli, E., Duranti, M., Scarafoni, A. New molecular features of cowpea bean (Vigna unguiculata, l. Walp) β-vignin. Biosci. Biotechnol. Biochem. 82, 285–291. https://doi.org/10.1080/09168451.2017.1419855, 2018.
Delley, B. From molecules to solids with the DMol3 approach. J. Chem. Phys. 113, 7756–7764. https://doi.org/10.1063/1.1316015, 2000.
Dunwell, J.M. Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. Biotechnol. Genet. Eng. Rev. 15, 1–32,1998.
Dunwell, J.M., Purvis, A., Khuri, S. Cupins: the most functionally diverse protein superfamily? Phytochemistry 65, 7–17, 2004.
Ehlers, J.D., Hall, A.E. Cowpea (*Vigna unguiculata* L. Walp.). Field Crops Res. 53, 187–204. https://doi.org/10.1016/S0378-4290(97)00031-2, 1997.
Fernandes, K.V., Sabelli, P.A., Barratt, D.H., Richardson, M., Xavier-Filho, J., Shewry, P.R. The resistance of cowpea seeds to bruchid beetles is not related to levels of cysteine proteinase inhibitors. Plant Mol. Biol. 23, 215–219, 1993.
Firmino, F., Fernandes, K.V.S., Sales, M.P., Gomes, V.M., Miranda, M.R.A., Domingues, S.J.S., Xavier, J. Cowpea (Vigna unguiculata) vicilins associate with putative chitinous structures in the midgut and feces of the bruchid beetles Callosobruchus maculatus and Zabrotes subfasciatus. Braz. J. Med. Biol. Res. 29, 749–756,1996.
Freire, J.E.C., Vasconcelos, I.M., Moreno, F.B.M.B., Batista, A.B., Lobo, M.D.P., Pereira, M.L., Lima, J.P.M.S., Almeida, R.V.M., Sousa, A.J.S., Monteiro-Moreira, A.C.O., Oliveira, J.T.A., Grangeiro, T.B. Mo-CBP3, an Antifungal Chitin-Binding Protein from Moringa oleifera Seeds, Is a Member of the 2S Albumin Family. PloS One 10, e0119871. https://doi.org/10.1371/journal.pone.0119871, 2015.
Freitas, R.L., Teixeira, A.R., Ferreira, R.B. Characterization of the proteins from Vigna unguiculata seeds. J. Agric. Food Chem. 52, 1682–1687. https://doi.org/10.1021/jf0300588, 2004.
Fukuda, T., Maruyama, N., Salleh, M.R.M., Mikami, B., Utsumi, S. Characterization and crystallography of recombinant 7S globulins of Adzuki bean and structure-function relationships with 7S globulins of various crops. J. Agric. Food Chem. 56, 4145–4153. https://doi.org/10.1021/jf072667b, 2008.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R., Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server, in: The Proteomics Protocols Handbook. Humana Press, Totowa, New Jersey, pp. 571–607, 2005.
Gatehouse, A.M.R., Gatehouse, J.A., Dobie, P., Kilminster, A.M., Boulter, D. Biochemical basis of insect resistance in Vigna unguiculata. J. Sci. Food Agric. 30, 948–958. https://doi.org/10.1002/jsfa.2740301003, 1979.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685
Macedo, L.L.P., Amorim, T.M.L., Uchôa, A.F., Oliveira, A.S., Ribeiro, J.K.C., de Macedo, F.P., Santos, E.A., de Sales, M.P. Larvicidal effects of a chitin-binding vicilin from Erythrina velutina seeds on the mediterranean fruit fly Ceratitis capitata. J. Agric. Food Che
m. 56, 802–808. https://doi.org/10.1021/jf072746n, 2008.
Macedo, M.L.R., Da Andrade, S.L.B., Moraes, R.A., Xavier-Filho, J. Vicilin variants and the resistance of cowpea (Vigna unguiculata) seeds to the cowpea weevil (Callosobruchus maculatus). Comp. Biochem. Physiol. - Part C Toxicol. Pharmacol. 105, 89–94. https://doi.org/10.1016/0742-8413(93)90063-Q. 1993.
Maranhão, P.A.C., Teixeira, C.S., Sousa, B.L., Barroso-Neto, I.L., Monteiro-Júnior, J.E., Fernandes, A.V., Ramos, M.V., Vasconcelos, I.M., Gonçalves, J.F.C., Rocha, B.A.M., Freire, V.N., Grangeiro, T.B. cDNA cloning, molecular modeling and docking calculations of L-type lectins from Swartzia simplex var. grandiflora (Leguminosae, Papilionoideae), a member of the tribe Swartzieae. Phytochemistry 139, 60–71. https://doi.org/10.1016/j.phytochem.2017.04.007, 2017.
Melo, R.A., Forti, V.A., Cicero, S.M., Novembre, A.D.L.C., de, M. Use of X-ray to evaluate damage caused by weevils in cowpea seeds. Hortic. Bras. 28, 472–476. https://doi.org/10.1590/S0102-05362010000400016, 2010.
Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785–2791. https://doi.org/10.1002/jcc.21256, 2009.
Murdock, L.L., Seck, D., Ntoukam, G., Kitch, L., Shade, R.E. Preservation of cowpea grain in sub-Saharan Africa - Bean/Cowpea CRSP contributions. Field Crops Res. 82, 169–178. https://doi.org/10.1016/S0378-4290(03)00036-4, 2003.
Murray, D.R., Mackenzie, K.F., Vairinhos, F., Peoples, M.B., Atkins, C.A., Pate, J.S. Electrophoretic Studies of the Seed Proteins of Cowpea, Vigna unguiculata (L.) Walp. Z. Für Pflanzenphysiol. 109, 363–370. https://doi.org/10.1016/S0044-328X(83)80118-4, 1983.
Maia, F. M. M.; Oliveira, J. T. A.; Matos, M. R. T.; Moreira, R. A.; vasconcelos, I. M. Proximate composition aminoacid content and haemagglutinating and trypsin-inhibiting activities of some brazilian vigna unguiculata (l) walp cultivars. Journal of the science of food and agriculture, v. 80, p. 453-458, 2000.
16. Maruyama, N.; Fukuda, T.; Saka, S.; Inui, N.; kotoh, J.; Miyagawa, M.; Hayashi, M.; Sawda, M.; Moriyama, T.; Utsumi, S. Molecular and structural analysis of electrophoretic variants of soybean seed storage proteins. Phytochemistry, v. 64, p. 701-708, 2003.
Khan, M.R.I., Gatehouse, J.A., Boulter, D. The Seed Proteins of Cowpea (Vigna unguiculata L. Walp.). J. Exp. Bot. 31, 1599–1611, 1980.
Kunz, D., Oliveira, G.B., Uchôa, A.F., Samuels, R.I., Macedo, M.L.R., Silva, C.P. Receptor mediated endocytosis of vicilin in Callosobruchus maculatus (Coleoptera: Chrysomelidae) larval midgut epithelial cells. Comp. Biochem. Physiol. - Part B Biochem. Mol. Biol. 210, 39–47. https://doi.org/10.1016/j.cbpb.2017.06.003, 2017.
Oliveira, M. Z. A. Fungos associados à semente de caupi: identificação, patogenicidade e controle. 1981. Dissertação (mestrado em biologia molecular). Centro de ciências biológicas, universidade de brasília, brasília, 1981.
Osborn, T. C.; Alexander, D. C.; Sun, S. S.; Cardona, C.; Bliss, F. A. Insecticidal activity and lectin homology of arcelin seed protein. Science, p. 240, v. 207-210, 1988.

19. Padulosi, S.; NG, Q. N. Cowpea gene pool distribution and crop improvement. 1988. In: (ed.) Ehlers Id, hall, A. E. Cowpea (*vigna unguiculata* l. Walp.).fields crop res, v. 53, p. 187-204, 1997.
Ojimelukwe, P.C., Ogwumike, F.C. Effects of infestation by bruchid beetles (Callosobruchus macultus) on the nutritional quality and sensory properties of cowpeas (Vigna unguiculata). J. Food Biochem. 23, 637–645, 1999.
Oliveira, G.B., Kunz, D., Peres, T.V., Leal, R.B., Uchôa, A.F., Samuels, R.I., Macedo, M.L.R., Carlini, C.R., Ribeiro, A.F., Grangeiro, T.B., Terra, W.R., Xavier-Filho, J., Silva, C.P. Variant vicilins from a resistant Vigna unguiculata lineage (IT81D-1053) accumulate inside Callosobruchus maculatus larval midgut epithelium. Comp. Biochem. Physiol. - Part B Biochem. Mol. Biol. 168, 45–52. https://doi.org/10.1016/j.cbpb.2013.11.001, 2014.
Oliveira, G. B. Comparação entre a absorção de vicilinas de sementes suscetíveis e resistentes do feijão-de-corda, vigna unguiculata pelo epitélio intestinal de larvas do caruncho Callosobruchus maculatus e o processamento de vicilinas no corpo gorduroso após absorção – 2013. Dissertação (mestrado em bioquímica) - centro de ciências biológicas, universidade federal de santa catarina, florianópolis, 2013
Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., Schulten, K. Scalable molecular dynamics with NAMD. J. Comput. Chem. 26, 1781–1802. https://doi.org/10.1002/jcc.20289, 2005.
Quiocho, F.A. Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. Annu. Rev. Biochem. 55, 287–315. https://doi.org/10.1146/annurev.bi.55.070186.001443, 1986.
Paes, E. V.; Uchoa, A. F.; pinto, M. S. T.; Silva, C. P.; Fernandes, K. V. S.; Oliveira, A. E. A.; Xavier-Filho, J. Binding of *vigna unguiculata* vicilins to the peritrophicmembrane of *tenebrio molitor* affects larval development. Entomologia experimentalis et applicata, v. 129, p.11-17, 2008.
Pant, K. C.; Chandel, K. P. S.; Joshi, B. S. Analysis of diversity in indian cowpea genetic resourch, v. 14, p. 103-111, 1982.
Pedalino, M.; paino-d'urzo, M.; Delle Done, G.; Grillo, S.; Rao, R. The structure of cowpea (*vigna unguiculata* l. Walp.) Seed storage proteins.seed science and technology, v. 20, p. 223-231, 1992.
Reis, C.M., Calvet, M.M., Sales, M.P., Fernandes, K.V.S., Gomes, V.M., Xavier, J. alpha-amylase inhibitors of legume seeds and their involvement in the resistance to Bruchid beetles. Arq. Biol. E Tecnol. 40, 413–418, 1997.
Rios, G.P. doenças fúngicas e bacterianas do caupi. In: Araujo, J.P.P; Watt E.E. o caupino brasil. Brasília: iita/embrapa, v. 1, p. 547-589, 1988.
Sales, M. P.; Fernandes, K. V. S.; Gomes, V. M.; Xavier-Filho, J. Chitin-binding proteins from cowpea (*vigna unguiculata*) seeds.brazilian journal of medical and biological research, v. 29, p. 319-326, 1996.
Sales, M. P.; Macedo, M. R. L.; Xavier-Filho, J. Digestibility of cowpea (*vigna unguiculata*) vicilins by pepsin, papain and bruchid midgut proteinases. Comparative biochemistry and physiology, v.103, p. 945-950, 1992.
Sales, M. P.; Pimenta, P. P.; Paes, N. S.; Grossi-de-Sá, M. F.; Xavier-Filho, J. Vicilins (7S storage globulins) of cowpea (*vigna unguiculata*) seeds bind to chitinious structures of the midgut of *callosobruchus maculatus* (coleoptera: bruchidae) larvae. Brazilian journal medical biology research, v. 34, p. 27–34, 2001.
Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., Karplus, M. Evaluation of comparative protein modeling by MODELLER. Proteins 23, 318–326. https://doi.org/10.1002/prot.340230306, 1995.

Sammour, R.H., Gatehouse, J.A., Gilroy, J., Boulter, D. The homology of the major storage protein of jack bean (Canavalia ensiformis) to pea vicilin and its separation from α-mannosidase. Planta 161, 61–70. https://doi.org/10.1007/BF00951461, 1984.
Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. 1, 2856–2860. https://doi.org/10.1038/nprot.2006.468, 2006.
Singh, B.B., Ajeigbe, H.A., Tarawali, S.A., Fernandez-Rivera, S., Abubakar, M. Improving the production and utilization of cowpea as food and fodder. Field Crops Res., Approaches to improve the utilization of food-feed crops 84, 169–177. https://doi.org/10.1016/S0378-4290(03)00148-5. 2003.
Singh, B.B., Singh, S.R. Breeding for bruchid resistance in cowpea, in: Bruchids and Legumes: Economy, Ecology and Coevolution, Series Entomologica. Kluwer Academic Publishers, Dordrecht, pp. 219–228, 1990.
Singh, B.B., Singh, S.R., Adjadi, O. Bruchid resistance in cowpea. Crop Sci. 25, 736–739, 1985.
Souza, S. M. Absorção e compartimentalização de vicilinas de *vigna unguiculata* por *callosobruchus maculatus* e *zabrotes subfasciatus* (coleoptera: chrysomelidae: bruquinae). 2009. Dissertação (mestrado em biociências e biotecnologia) - centro de biociências e biotecnologia, universidade estadual do norte fluminense, campos dos goytacazes, 2009.
Souza, S.M., Uchôa, A.F., Silva, J.R., Samuels, R.I., Oliveira, A.E.A., Oliveira, E.M., Linhares, R.T., Alexandre, D., Silva, C.P. The fate of vicilins, 7S storage globulins, in larvae and adult Callosobruchus maculatus (Coleoptera: Chrysomelidae: Bruchinae). J. Insect Physiol. 56, 1130–1138. https://doi.org/10.1016/j.jinsphys.2010.03.009, 2010.
Shewry, P. R.; Napier, J. A.; Tatham, A. S. Seed storage proteins: structures andbiosynthesis. The plant cell, v. 7, p. 945-956, 1995.
Shutov, A. D., Karhovskaya, I. A., Braun, H., Baümlein, H., Müntz, K. Legumin like and vicilin-like seed storage proteins: evidence for a common single-domain ancestral gene. Journal of molecular evolution, v. 41, p. 1057-1069, 1995.
Southgate, B. J. Biology of the bruchidae. *Annual review of entomology*, v. 24, p. 449-473, 1979.
Templematn, S.; Demaggioa, E.; Stetlerd, A. Biochemistry of fern spore germination: globulin storage proteins in *matteuccia struthiopteris* l. Plant physiology, v. 85, p. 343-349, 1987. Trott, O., Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461. https://doi.org/10.1002/jcc.21334, 2010.
Uchôa, A.F., DaMatta, R.A., Retamal, C.A., Albuquerque-Cunha, J.M., Souza, S.M.,Samuels, R.I., Silva, C.P., Xavier, J. Presence of the storage seed protein vicilin in internal organs of larval Callosobruchus maculatus (Coleoptera : Bruchidae). J. Insect Physiol. 52, 169–178. https://doi.org/10.1016/j.jinsphys.2005.10.002, 2006.
Xavier-Filho, J., Campos, F.A.P., Ary, M.B., Silva, C.P., Carvalho, M.M.M., Macedo, M.L.R., Lemos, F.J.A., Grant, G. Poor correlation between the levels of proteinase inhibitors found in seeds of different cultivars of cowpea (Vigna unguiculata) and the resistance/susceptibility to predation by Callosobruchus maculatus. J. Agric. Food Chem. 37, 1139–1143. https://doi.org/10.1021/jf00088a071, 1989
Wang, T. L.; Domoney, C.; Hedley, C. L.; Casey, R.; Grusak, M. A. Can we improve the nutritional quality of legume seeds? Plant physiology, v. 131, p. 886-891, 2003.
Zhang, Y., Skolnick, J. Scoring function for automated assessment of protein structure template quality. Proteins 57, 702–710. https://doi.org/10.1002/prot.2026, 2004.