

Article - Agriculture, Agribusiness and Biotechnology

In Silico Characterization of *Meloidogyne* Genus Nematode Cellulose Binding Proteins

Alana Manoela Fraga Menezes¹

<https://orcid.org/0000-0002-7057-8620>

Edilton de Albuquerque Cavalcanti Junior^{1,3}

Luiza Suely Semen Martins^{2,3}

Rômulo Maciel de Moraes Filho^{1,3}

<https://orcid.org/0000-0003-3988-4379>

¹Affiliation: Departamento de Agronomia, Universidade Federal Rural de Pernambuco

²Affiliation: Departamento de Biologia, Universidade Federal Rural de Pernambuco

³Affiliation: Programa de Pós-Graduação em Agronomia: Melhoramento Genético de Plantas, Universidade Federal Rural de Pernambuco;

Received: 2018.03.05; Accepted: 2019.03.19.

HIGHLIGHTS

- CBPs are remarkably conserved between *Meloidogyne* species
- CBPs exhibit two to three functional domains related to signal binding and cellulose hydrolase activities

Abstract: Root-knot nematodes are a group of endoparasites species that induce the formation of giant cells in the hosts, by which they guarantee their feeding and development. *Meloidogyne* species infect over 2000 plant species, and are highly destructive, causing damage to many crops around the world. *M. enterolobii* is considered the most aggressive species in tropical regions, such as Africa and South America. Phytonematodes are able to penetrate and migrate within plant tissues, establishing a sophisticated interaction with their hosts through parasitism factors, which include a series of cell wall degradation enzymes and plant cell modification. Among the parasitism factors documented in the *M. enterolobii* species, cellulose binding protein (CBP), a nematode excretion protein that appears to be associated with the breakdown of cellulose present in the plant cell wall. In silico analysis can be of great importance for the identification, structural and functional characterization of genomic sequences, besides making possible the prediction of structures and functions of proteins. The present work characterized 12 sequences of the CBP protein of nematodes of the genus *Meloidogyne* present in genomic databases. The results showed that all CBP sequences had signal peptide and that, after their removal, they had an isoelectric point that characterized them as unstable in an acid medium. The values of the average hydrophilicity demonstrated the hydrophilic character of the analyzed sequences. Phylogenetic analyzes were also consistent with the taxonomic classification of the nematode species of this study. Five motifs were identified, which are present in all sequences analyzed. These results may provide theoretical grounds for future studies of plant resistance to nematode infection.

Keywords: Bioinformatics; Phytonematodes; Parasitism Factors; Cell wall degrading enzymes; Homology modeling; Protein function prediction

INTRODUCTION

Root-knot nematodes are obligate endoparasites that induce the formation of giant cell on its host. It is estimated that more than 5,000 species of plants worldwide are infested by these organisms¹. Root-knot nematodes pierce the cell wall of the roots with a projectable stylet, by which they secrete substances that induce the differentiation of plant cells, forming multinucleated cells that undergo phenotypic, functional and metabolic alterations²⁻⁴. Galls are formed around the giant cells and the roots become distorted, compromising their functions and damaging the growth of the plant. It is through the giant cells that the phytonematodes guarantee their feeding, development and reproduction⁴. Phytonematodes of the genus *Meloidogyne* are highly destructive and are associated with losses of production in several crops worldwide. The presence of these parasites in different regions of the world has been increasingly reported, including in crops plants with resistance genes to other species of the genus⁵⁻⁷.

Among the species with the highest parasitic potential, *Meloidogyne enterolobii* is considered to be the most aggressive in comparison with other tropical species of root-knot nematodes^{8,9}. The species was reported in Africa, Central America, the United States, France and China¹⁰.

In Brazil, *M. enterolobii* was originally detected in guava orchards in 2001, in the States of Pernambuco and Bahia. Since then, this nematode has been a major concern in the country due to its rapid propagation and destructive potential, making it impossible to cultivate in areas where it is present¹¹.

The severity of its infestation is primarily due to the high capacity of this species to overcome resistance genes of their hosts⁹. This resistance-breaking ability is an important factor that gives this nematode species a high multiplication potential in relation to other species of this genus, in addition to inducing the formation of gall in the roots of plants more efficiently in tropical regions^{11,12}.

Through parasitism factors, phytonematodes are able to penetrate and migrate within plant tissues, establishing a sophisticated interaction with their hosts. Parasitism factors of these nematodes include a number of cell wall degradation enzymes (CWDEs) and secretory effector proteins (SEP) to suppress host defense responses¹³⁻¹⁹. The invasion to the roots of the plants requires the degradation of the cellular wall, constituted mainly by cellulose, hemicellulose and pectins, that form a physical barrier against the attack of pathogens²⁰. Because cellulose is an insoluble polymer and particularly difficult to be degraded, some microorganisms have developed strategies that involve the production and secretion of enzymes that act synergistically²¹. In these phytonematodes, several genes are found that are generally absent in animals, although they are similar to genes present in prokaryotic organisms, such as soil bacteria. Phylogenetic analyzes indicate that these nematodes have acquired genes from other organisms capable of giving them the ability to produce enzymes that degrade the cell wall of plants. This acquisition would be through the horizontal transfer of genes, which means transmission of genes between different organisms by mechanisms other than the vertical genetic inheritance from biological parents^{20,22}. Among the parasitism factors related to cell wall breakage documented in the *M. enterolobii* species, cellulose binding protein (CBP) had not its function clearly described yet. In phytonematodes, the first gene identified was the Mi CBP-1, in the species *M. incognita*⁴. The authors identified an N-terminal signal peptide in the protein sequence, indicating that Mi CBP-1 would be a secreted protein. Several lines studies^{4,21,23} suggest that CBP is secreted during parasitism. One of the indications would be that, at first, CBP genes are not found in any other parasitic nematode of animals or non-parasites, indicating that the mode of action of the protein would be on the plant's cellular tissue. In a study of cell wall modification during the parasitism of the phytonematode species *Heterodera*

schachtli, increasing levels of CBP mRNA were observed in stage J2, and a peak expressed in stage J3, suggesting an important role of the protein after penetration, already inside the plant root and particularly during the beginning of the formation and development of the syncytium (multinucleated mass of cytoplasm formed by the fusion of originally separated cells)²¹.

In the *M. javanica* species, the CBP gene encodes a protein that contains a signal peptide of secretion, which is expressed in eggs and in the pre-invasive stage J2, but not in adult females²⁴. *In silico* analysis can be of great importance for the identification, structural and functional characterization of genomic sequences, besides making possible the prediction of structures and functions of proteins. This type of experimentation represents a low cost alternative for studies on biological functions with very precise computational models when compared with laboratory conditions. Computational analyzes have been used in the characterization of several proteins and enzymes in both eukaryotes and prokaryotes, in search of genes linked to parasitism and resistance in host species²⁵⁻³¹. The identification of proteins involved in the relationship of parasitism of nematodes and plants can provide important tools for research that develop plant species with resistant genes and in the search for mechanisms that avoid the breakdown of this resistance by phytonematodes.

MATERIAL AND METHODS

Data search and sequence retrieval

The search for genomic and protein sequences was performed at National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>). Using *M. enterolobii* CBP sequence (ANH56394) as query, the BLASTp tool was used for localization of homologous and similar sequences in other nematode genera. Additionally this tool was also used to search for protein homologues from the plant parasitic nematodes genomes in the wormbase database (parasite.wormbase.org). This database contain three publically available genomes of completely sequenced nematodes from the *Meloidogyne* genus, *M. floridensis*, *M. incognita*, *M. hapla*.

Analysis of Protein Features

Physicochemical parameters of the identified protein sequences were estimated by the ProtParam software (<http://web.expasy.org/protparam>)³². The subcellular localization of the analyzed sequences is being predicted by the software CELLO2GO³³. Location of regions of peptide signals in the characterized proteins will be performed by the software TOPCONS (<http://topcons.cbr.su.se/>)³⁴.

Identification of conserved motifs and phylogenetic analysis

The presence of conserved motifs was analyzed by the MEME SUITE tool³⁵. Additionally, Prodom server³⁶ was used to search for conserved domains in the protein sequence. Sequence alignment was performed by the ClustalW algorithm implemented in the software MEGA6.06³⁷. For the phylogenetic analysis, a Neighbor Joining (NJ) tree was generated with 1000 bootstrap replications.

Tertiary structure prediction

Due to the non-existence of a three-dimensional model of CBPs, it was necessary to use the homology modeling methodology. This method predicts the three-dimensional structure of a protein using as a template similar sequences that already have their structure elucidated. 3D models of the proteins identified were generated by the Phyre2 server (<http://www.sbg.bio.ic.ac.uk/phyre2>)³⁸. Phyre2 server incorporates the Poing tool³⁹, which is an ab initio folding simulation to model regions of the proteins with no detectable homology to known structures. Poing tool also combines multiple templates to improve model accuracy. Ab initio prediction methods consist in modeling all the energetics involved in the process of folding, and then in finding the structure with lowest free energy. The accuracy of this methodology is limited to small proteins, with less than 100 residues and should be taken in consideration with caution^{38, 40}. The results were viewed by the UCSF Chimera

software⁴¹. Model quality was evaluated using the Molprobit server (<http://molprobit.biochem.duke.edu/>)⁴² by Ramachandran plot analysis. Z-score was calculated using interactive ProSA-web server (<https://prosa.services.came3.sbg.ac.at/prosa.php>) to recognize errors in 3-D structures⁴³.

RESULTS and DISCUSSION

Identification and characterization of CBPs

Several studies have demonstrated that CWDEs produced by phytonematodes have an important role in the establishment of the parasite-host relationship during the infection process^{4, 19, 21, 23, 24}. Following the first report of a CBP from a nematode in *M. incognita*⁴, several highly similar sequences were described in phytonematode species from the Meloidogynidae family^{4, 19, 21, 23, 24}, including two to three allelic variations of CBP from *M. incognita*, *M. javanica* and *M. arenaria*⁴

In this study, 11 CBP sequences from the nematodes of the Meloidogynidae family were found in the NCBI database, recovered in the FASTA format, and one sequence of the species *M. hapla* was retrieved from the Wormbase genomic database. The 12 sequences are listed in Table 1. All sequences were identified with presence of signal peptide by the TOPCONS server. This tool predicts the presence and location of signal peptide cleavage sites. TOPCONS shows all sequences analyzed with N-terminal peptide signal (Table 2). The presence of signal peptide is characteristic of excretory proteins. Physicochemical parameters of the CBP sequences were analyzed in order to investigate the structures and functionalities of the proteins. As shown in Table 2, after the removal of the signal peptide, the analyzed CBP sequences vary in size between 174 aa (*M. hapla*) to 187 aa (*M. enterolobii*) and have molecular weight of 18.3 kDa (*M. hapla*) and 19.9 kDa (*M. incognita*_CAM33386.1). The isoelectric point (pI) of the sequences was between 4.27 (*M. enterolobii*) and 4.93 (*M. incognita*_CAM33386.1). The isoelectric point is the pH at which the protein becomes insoluble and therefore unstable⁴⁴. The mean hydropathicity, identified in Table 2 as GRAVY, varied between -0.630 (*M. hapla*) and -0.761 (*M. incognita*_CAM33386.1). These results indicate that the analyzed CBPs are hydrophilic. Thus, the identified CBPs would have low activity at acidic pH, probably acting in the neutral or alkaline range. Due the acidic pH found in plant tissues, and the fact that *Meloidogyne* species and many other phytoparasitic organisms are often encountered in soil raises the possibility that the acidic CWDEs may have physiological functions that are important for survival outside the plant tissue⁴⁵. CELLO2GO server was used to analyze the subcellular location of the sequences. All the sequences were described as extracellular proteins, and may also occur in the cell interior, in the chloroplast (*M. arenaria*) and in the nucleus (*M. enterolobii*, *M. hapla*, *M. incognita*, *M. javanica*), which is in accordance with the prior identification of CBP, as a parasite-related excretion protein^{4, 21}.

Table 1 - Twelve cellulose binding protein (CBP) sequences and their species of origin, obtained from the NCBI database and the Wormbase genomic database.

Species	Access Number	Gen Bank Definition	Common name
<i>M. arenaria</i>	CAM33389	cellulose binding protein precursor	Root-knot nematode
<i>M. arenaria</i>	CAM33387	cellulose binding protein precursor,	Root-knot nematode
<i>M. arenaria</i>	CAM33384	cellulose binding protein precursor	Root-knot nematode
<i>M. enterolobii</i>	ANH56394	cellulose binding protein	Root-knot nematode
<i>M. incognita</i>	CAM33385	cellulose binding protein precursor,	Root-knot nematode
<i>M. incognita</i>	CAM33388	cellulose binding protein precursor,	Root-knot nematode
<i>M. incognita</i>	AAC05133	cellulose binding protein precursor cbp-1	Root-knot nematode
<i>M. incognita</i>	CAM33386	cellulose binding protein precursor,	Root-knot nematode
<i>M. javanica</i>	CAM33392	cellulose binding protein precursor,	Root-knot nematode
<i>M. javanica</i>	CAM33390	cellulose binding protein precursor,	Root-knot nematode
<i>M. javanica</i>	CAM33391	cellulose binding protein precursor,	Root-knot nematode
<i>M. hapla</i> *	Contig343.frz3.gene25	-	Northern root-knot nematode

* sequence retrieved from the Wormbase database

Table 2 - Analysis of primary structures, signal peptide and subcellular location of CBP, identified by the TOPCONS server.

Species	Protein Size	MW (kDa)	pI	GRAVY	SCPS	ScL
<i>M. arenaria</i> _CAM33389.1	203/182	21.9/19.6	4.41/4.46	-0.395/ 0.654	A22- A23	Ex; Ch
<i>M. arenaria</i> _CAM33387.1	203/182	21.9/19.6	4.37/4.41	-0.388/ 0.647	A22- A23	Ex; Ch
<i>M. arenaria</i> _CAM33384.1	203/182	21.9/19.5	4.43/4.47	-0.417/ 0.664	A22- A23	Ex; Ch
<i>M. enterolobii</i> _ANH56394.1	208/187	22.2/19.8	4.23/4.27	-0.495/ 0.734	A22- A23	Ex; Nu
<i>M. hapla</i> _Contig343.frz3.gene25	195/174	20.6/18.3	4.53/4.35	-0.369/ 0.630	A21- A22	Ex; Nu
<i>M. incognita</i> _CAM33385.1	203/182	21.8/19.5	4.41/4.46	-0.405/ 0.651	A22- A23	Ex
<i>M. incognita</i> _CAM33388.1	202/181	21.7/19.4	4.35/4.39	-0.397/ 0.657	A22- A23	Ex; Nu
<i>M. incognita</i> _AAC05133.1	203/182	21.9/19.6	4.38/4.42	-0.478/ 0.732	A22- A23	Ex; Nu

Cont.

Cont.

Species	Protein Size	MW (kDa)	pI	GRAVY	SCPS	ScL
<i>M. incognita</i> _CAM33386.1	204/183	22.3/19.9	4.84/4.93	-0.505/- 0.761	A22- A23	Ex; Nu
<i>M. javanica</i> _CAM33390.1	203/182	21.9/19.6	4.37/4.41	-0.395/- 0.654	A22- A23	Ex; Nu
<i>M. javanica</i> _CAM33391.1	203/182	21.9/19.6	4.41/4.46	-0.405/- 0.666	A22- A23	Ex; Nu
<i>M. javanica</i> _CAM33392.1	203/182	21.9/19.6	4.41/4.46	-0.383/- 0.641	A22- A23	Ex; Nu

* GRAVY, grand average of hydropathicity; MW, molecular weight; pI, isoelectric point are reported as precursor values/modeled protein values. ScL, subcellular locations are reported as Nc (nuclear), Ex (extracellular) and Ch (Chloroplast).

Identification of conserved motifs and phylogenetic analysis

Figure 1 shows the multiple alignment of the analyzed CBP sequences generated by the MEGA7 software using the ClustalW algorithm. The characterization of CBP in species of the genus *Meloidogyne* demonstrated the conservation of this protein in the organisms analyzed, suggesting its importance for the infection process of these parasites in their respective hosts. The conserved motifs of the CBP sequences were analyzed using the MEME Suite tool. Based on the results, five motifs were discovered. Three of the five motifs occur in all analyzed sequences (Figure 2). Motif 5 is absent in the CBP sequence in *M. hapla* species and motif 3 is not present in the sequence of *M. incognita*_CAM33386.1.

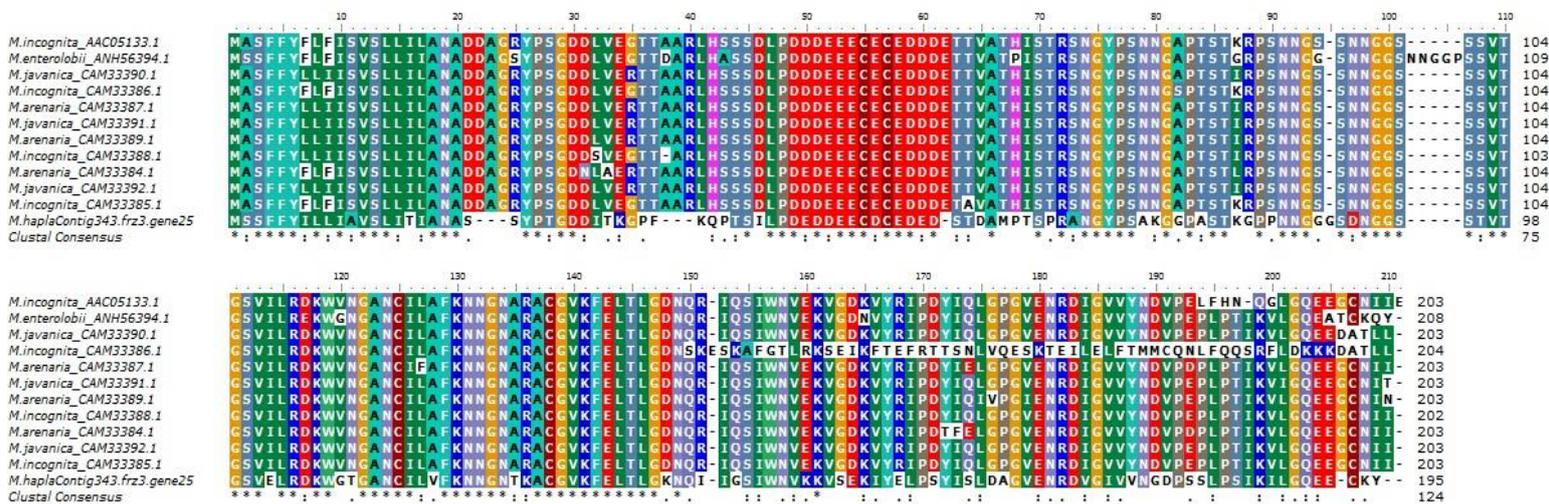


Figure 1. Multiple alignment of CBP sequences generated by the ClustalW algorithm.

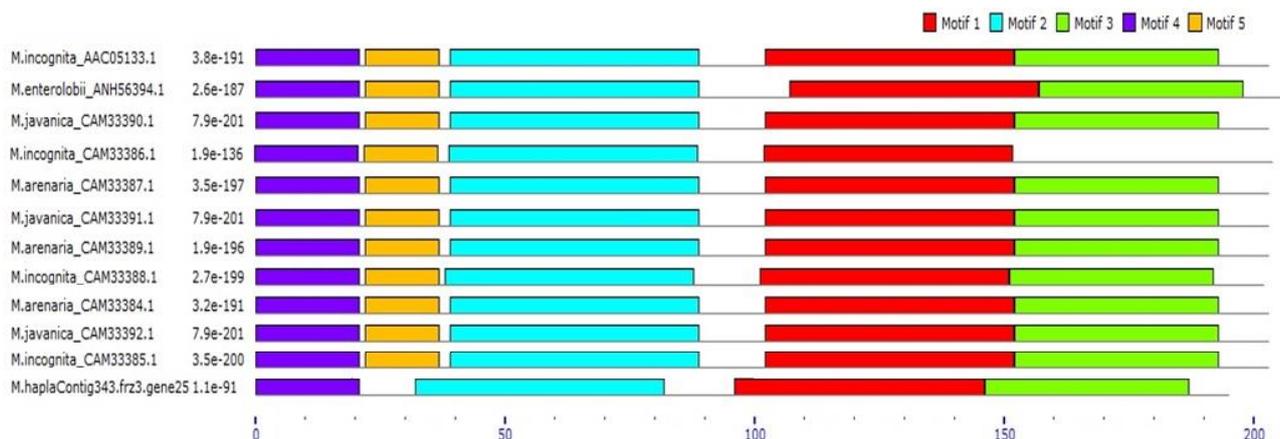


Figure 2 Conserved motifs identified in the evaluation of cellulose binding protein (CBP) using the MEME SUITE tool. The position of each block indicates where the subject was matched in the sequence. The width of the blocks indicates the width relative to the size of the sequence. The color and boundary of the blocks are used to identify the corresponding motif, as shown in the legend. The height of the blocks represents the significance of the correspondence, with higher blocks being more significant.

The results obtained by the ProDom server (Figure 3) demonstrate the presence of three functional domains, related to signal binding (PD069996 and PDD5D5D3) and cellulose hydrolase (PDC864N4). As observed in the Figure 3, Motif 3 is absent in *M. incognita*_CAM33386.1 sequence. This motif corresponds to the PDC864N4 domain, which is only partially present, and PDD5D5D3 domain, which is completely absent in *M. incognita*_CAM33386.1. Four allelic variations for *M. incognita* CBP were previously identified^{4, 24}, and *M. incognita*_CAM33386.1 was described as originated from a avirulent line of *M. incognita* from Lybia²⁴.

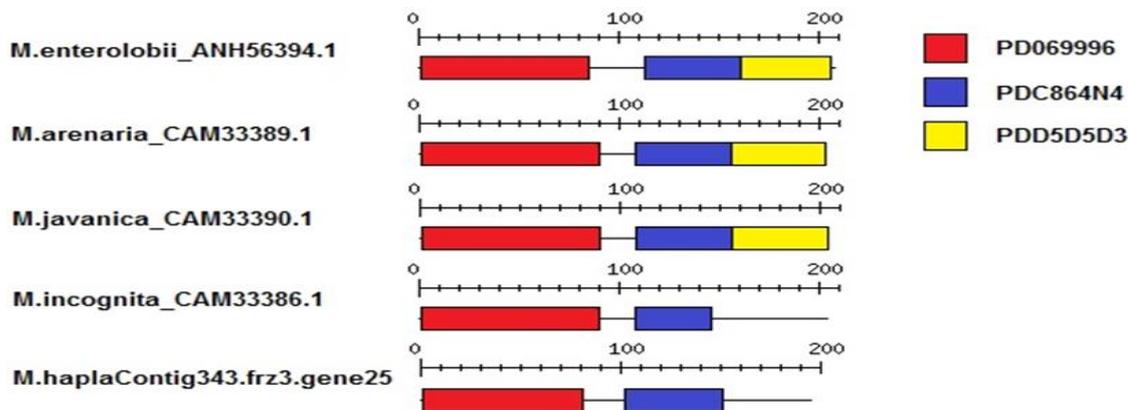


Figure 3 - Functional domains observed by the ProDom server. Blocks of different colors represent distinct domains found in the Pfam database.

The grouping found in the clustering tree was consistent with the taxonomic classification of the nematodes. The species *M. arenaria*, *M. enterolobii*, *M. incognita* and *M. javanica* reproduce by parthenogenesis and are characteristic of tropical regions, differing from *M. hapla*, which reproduces by meiotic parthenogenesis and present in temperate regions^{46, 47}. However, the variation observed in CBP within the G1 cluster, does not show a clear species specific relationship (Figure 4) which is consistent with the existence of many allelic variations for CBP in each species.

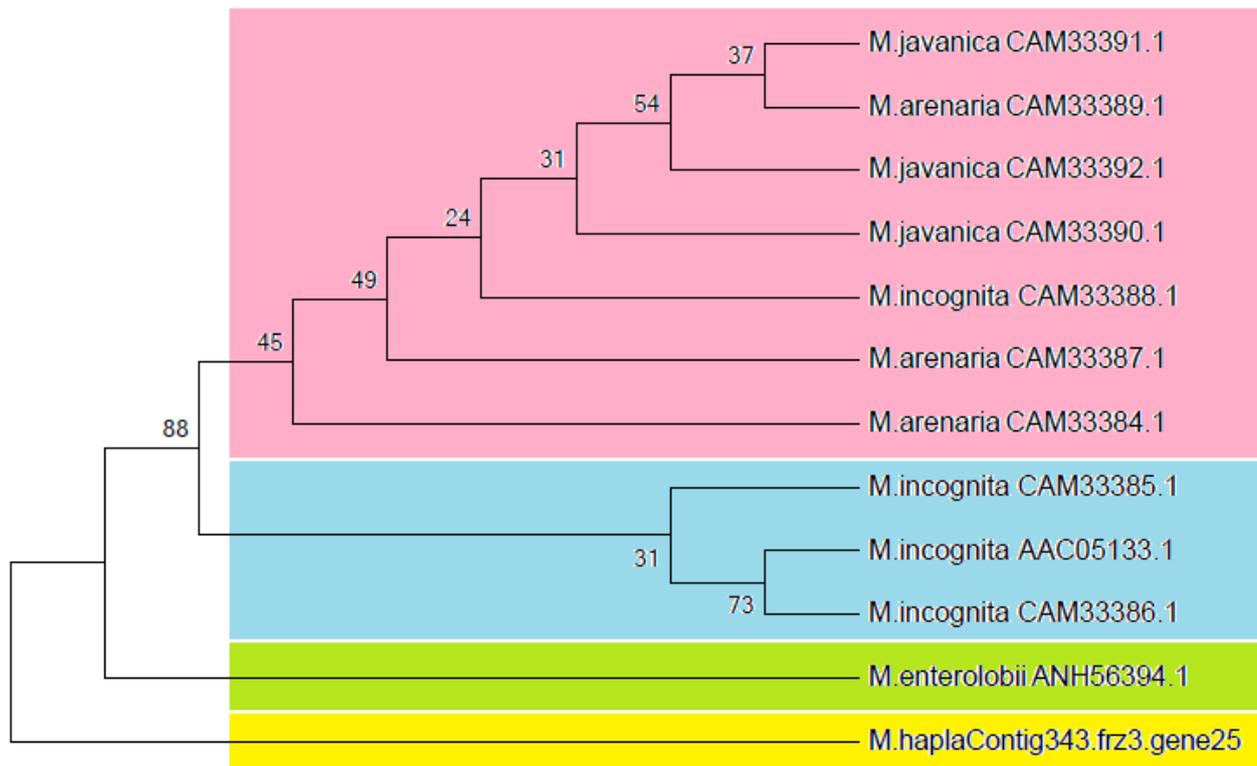


Figure 4 - Neighbor-joining (NJ) tree obtained from CBP sequences of species of the genus *Meloidogyne*. The bootstrap percentage (1,000) is indicated on each node of the tree.

Tertiary structure prediction

The knowledge of the 3D structure of proteins is of great importance for the understanding of their functions. Thus, tertiary structure prediction of *M. enterolobii* was performed using the Phyre2 server. This server uses the alignment of hidden Markov models via HHsearch⁴⁸ to improve the accuracy of the alignment and detection rate. Three templates were selected as models based on heuristics to maximize confidence, percentage identity and alignment coverage. The selected models were, *d1e5ba*, identified as Endo-1,4-beta xylanase D, *c3ndyG*, identified as Endoglucanase d, and *c2rttA*, identified as a chitin binding protein (*chi18ac*). Eighty-five (85) residues were modeled by *ab initio* methodologies due the lack of similarity to any available models. Figure 5a illustrates the model generated for the CBP of the *M. enterolobii* species. The hydrophobicity surface calculated by the UCSF Chimera software as can be seen in figure 5b, confirms the hydrophilic character of CBP. The electrostatic surface (fig 5c) shows a mostly negatively charged surface for Me-CBP. The three-dimensional model generated for the CBP of *M. enterolobii* reveals a protein with 40% of its structure consisting of Beta-strands, which is characteristic of carbohydrate binding modules⁴⁹. The enzymatic degradation of cell wall components, such as cellulose and xylans requires several types of enzyme such as endoglucanases, cellobiohydrolases (exoglucanases), or xylanases⁵⁰. Structurally, these enzymes generally consist of a catalytic domain and a carbohydrate-binding module (CBM2) which is a conserved region of approximately 100 amino acid residues, that can be found either at their N-terminal or C-terminal extremities⁵¹. Like other CBM domains, CBM2 is a beta-sheet domain⁵².

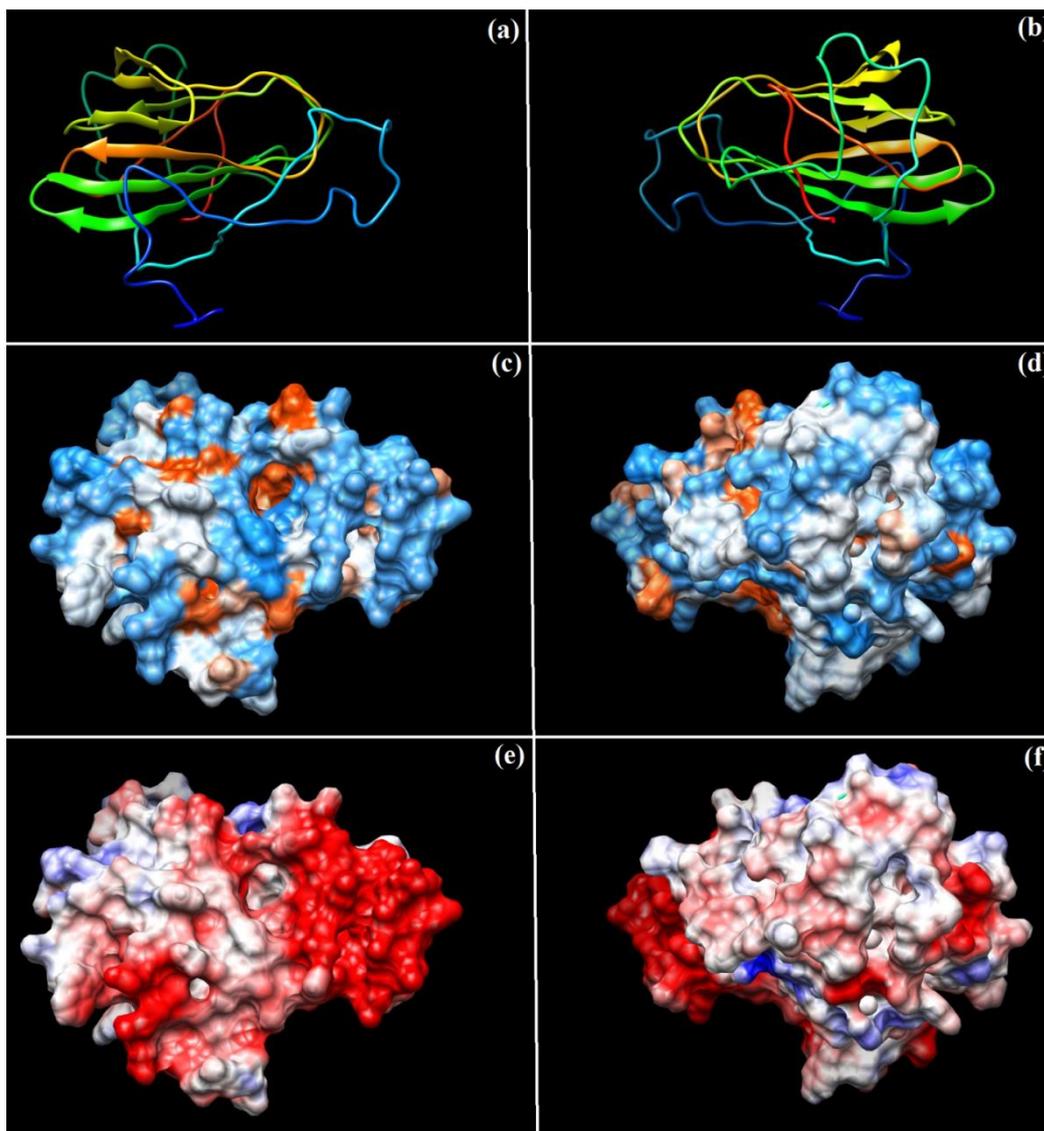


Figure 5 - Three-dimensional structure predicted by the Phyre2 server for the CBP of *M. enterolobii*. Front (a) and back (b) of the dark blue (N-terminal) region to the red (C-terminal) region. Front (c) and back (d) of the protein structure representing color gradient hydrophobicity, with blue being the most hydrophilic, red more hydrophobic region and neutral region of white color. Electrostatic surface represented in color gradient of the most negatively charged (red) to most positively charged (blue).

The generated model was submitted to Yasara energy minimization server⁵³, for molecular dynamics simulations and refinement through energy minimization. Verification of stereochemical quality of the model was carried out using Ramachandran plot analysis by the Molprobit server. Due to the presence of Ramachandran outliers, model refinement was also carried out with the KiNG software⁵⁴. The final models showed 90.81% of amino acid residues in favored regions and no outliers (Figure 6a). ProSA-web (Protein Structure Analysis web) was used to recognition errors in the tertiary structure prediction. The Z-score was used to measure the energy, as it indicated overall quality of the model. Positive Z-score values show that the structure is not stabilized while zero and negative scores represent energy stabilized structures. Me-CBP model showed a Z-score value of -3.2. The plot of residue scores shows local model quality by plotting energies as a function of amino acid sequence position (Figure 6b). Positive values correspond to problematic or erroneous parts of the structure. As was demonstrated in the graph of Figure 6c, Most of amino acid residues of Me-CBP are below zero on x-axis.

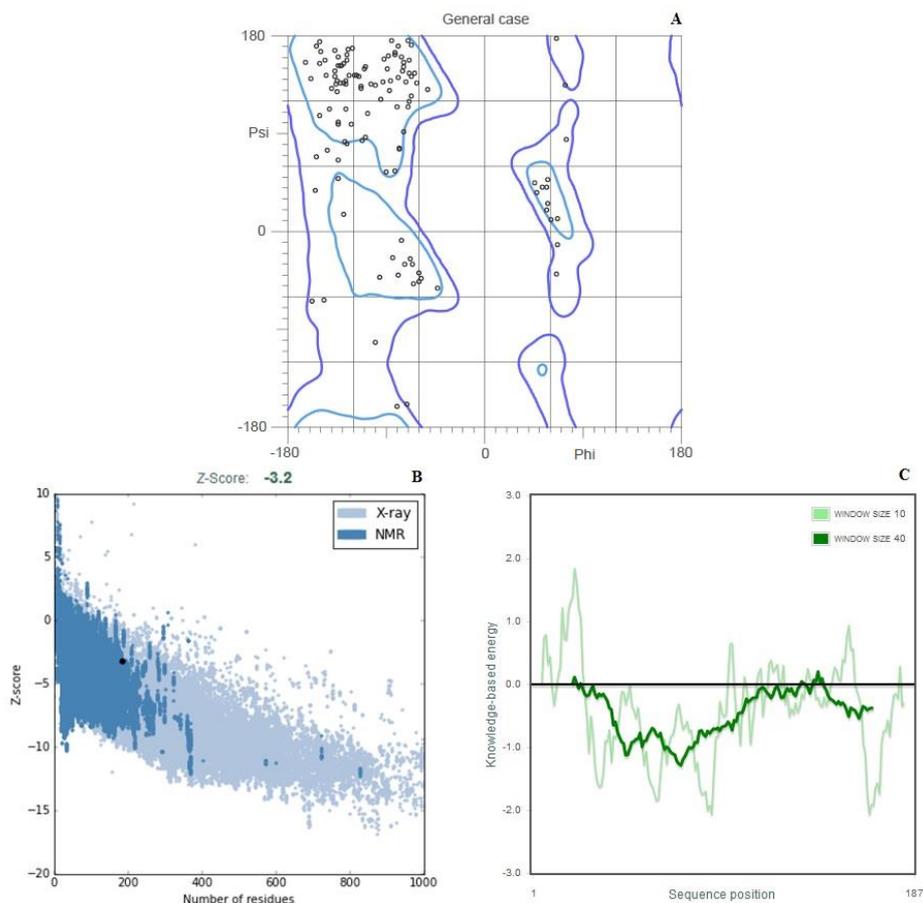


Figure 6 - Ramachandran plot of MeCBP generated by the MOLPROBITY server. (b) ProSA-web Z-score plot of MeCBP showing the Z value (black dot) and (c) ProSA-web plot of MeCBP showing the energy graph of residue scores of a native protein structure.

Prediction of the 3-D structure of Me-CBP would provide valuable insights into the molecular basis of these protein functions.

As suggested by previous studies ^{4, 21, 23, 24}, CBPs may have an important role in the establishment of the parasitism relationship between plant parasitic nematodes and their hosts, but CBP function as CWDE is not clearly comprehended, since this protein did not exhibit cellulase activity but bound to cellulose and plant cell walls ⁴. The development of a CBP three dimensional model could be of great importance for molecular docking studies and the understanding of its molecular interactions with cellulose molecules.

CONCLUSIONS

Bioinformatics analysis can play a vital role in the interpretation of proteomic data. These methodologies have been extensively used for predicting function and structure of proteins from its amino acid sequences ^{24, 27, 28, 52, 53, 55}. The identification of the characteristics of these proteins during parasitism can serve as an important tool in the construction of control strategies for root-knot nematodes and its interactions with their host plants. In this work we present the first tertiary model of CBP from *M. enterolobii*. These findings can provide useful information on the molecular basis of the functions of these proteins and the understanding of nematode infection processes.

Funding: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. Cavalcanti Junior EA and Moraes Filho RM were respectively supported by a masters and postdoctoral fellowship PNPD-CAPES.

Acknowledgments: We thank UFRPE for the institutional support to this research

Conflict of interest

"The authors declare no conflict of interest."

REFERENCES

1. Blok VC, Jones JT, Phillips MS, Trudgill DL. Parasitism genes and host range disparities in biotrophic nematodes: The conundrum of polyphagy versus specialization. *BioEssays*, 2008, 30:249-259.
2. Jones MGK, and Payne HL. Early stages of nematode-induced giant cell formation in roots of *Impatiens balsamina*. *Journal of Nematology*, 1978, 10:70-84.
3. Wiggers RJ, Starr JL, Price, HJ. DNA content and variation in chromosome number in plant cells affected by *Meloidogyne incognita* and *M. arenaria*. *Phytopathology*, 1990, 80:1391-1395.
4. Ding X, Shields J, Allen R, Hussey R.S A secretory cellulose-binding protein cDNA cloned from the root-knot nematode (*Meloidogyne incognita*). *MPMI*, 1998, 11:952-959.
5. Blok VC, Wishart J, Fargette M, Berthier K, Phillips MS. Mitochondrial DNA differences distinguishing *Meloidogyne mayaguensis* from the major species of tropical root-knot nematodes. *Nematology*, 2002, 4: 773-781.
6. Carneiro RMDG, Almeida MRA, Braga RS, Almeida C. A, Gioria, R. Primeiro registro de *Meloidogyne mayaguensis* parasitando plantas de tomate e pimentão resistentes à meloidoginose no Estado de São Paulo. *Nematol. Bra.*, 2006, 30:81-86.
7. Brito JA, Stanley JD, Kaur R, Cetintas R, Di Vito, M, Thies, JA, Dickson, DW. Effects of the Mi-1, N and Tabasco genes on infection and reproduction of *Meloidogyne mayaguensis* on tomato and pepper genotypes. *Journal of Nematology*, 2007, 39: 327-332.
8. Brito J, Powers, TO, Mullin PG, Inserra, RN, Dickson, DW. Morphological and molecular characterization of *Meloidogyne mayaguensis* isolates from Florida. *Journal of Nematology*, 2004, 36:232-240.
9. Kiewnick S, Dessimoz M, Franck L. Effects of the Mi-1 and the N root-knot nematode-resistance gene on infection and reproduction of *Meloidogyne enterolobii* on tomato and pepper cultivars. *Journal of Nematology* 2009, 41(2):134-139.
10. EPPO global database. An emerging root-knot nematode, *Meloidogyne enterolobii*: addition to the EPPO Alert List. EPPO Reporting Service, 2008, 5:9-10.
11. Carneiro RMDG, Moreira WA, Almeida MRA, Gomes ACMM. Primeiro registro de *Meloidogyne mayaguensis* em goiabeira no Brasil. *Nematol. Bra.*, 2001, 25:223-228.
12. Cetintas R, Kaur R, Brito JA, Mendes ML, Nyczepir AP, Dickson DW. Pathogenicity and reproductive potential of *Meloidogyne mayaguensis* and *M. floridensis* compared with three common *Meloidogyne* spp. *Nematropica*, 2007, 37:21-32.
13. Williamson VM, Gleason CA. Plant-nematode interactions. *Curr. Opin. Plant Biol.* 2003, 6: 327-333. [http://dx.doi.org/10.1016/S1369-5266\(03\)00059-1](http://dx.doi.org/10.1016/S1369-5266(03)00059-1)
14. Davis EL, Hussey RS, Baum TJ. Getting to the roots of parasitism by nematodes. *Trends Parasitol.* 2004, 20:134-141
15. Baum TJ, Hussey RS, Davis EL. Root-knot and cyst nematode parasitism genes: the molecular basis of plant parasitism. In: Genetic Engineering, (Setlow, J.K., ed), Springer Science+Business Media, LLC, New York, USA, 2007, Vol.28, pp. 17-34.
16. Davis EL, Hussey RS, Mitchum M.G, Baum, T.J. Parasitism proteins in nematode-plant interactions. *Curr. Opin. Plant Biol.* 2008, 11: 360-366. [http://dx.doi.org/10.1016/j.pbi.\(2008\).04.003](http://dx.doi.org/10.1016/j.pbi.(2008).04.003)
17. Perry RN, Moens N. Introduction to plant-parasitic nematodes: modes of parasitism. In: Genomics and molecular genetics of plant-nematode interactions (Jones J, Gheysen G and Fenoll C, eds.). Springer Science + Business Media, Dordrecht, 2011), pp 3-20.
18. Haegeman A, Mantelin S, Jones JT, Gheysen, G. Functional roles of effectors of plant-parasitic nematodes. *Gene*, 2012, 492, 19-31
19. Rai KM, Balasubramanian VK, Welker, CM, Pang M. et al. Genome wide comprehensive analysis and web resource development on cell wall degrading enzymes from phyto-parasitic nematodes. *BMC Plant Biol.* 2015, 15: 187. <http://dx.doi.org/10.1186/s12870-015-0576-4>.

20. Danchin EGJ, Rosso M, Vieria P, Almeida-Engler J, Coutinho PM, Henrissat B, Abad P. Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. *PNAS*, 2010, 107(41)17651–17656.
21. Hwezi T, Howe, P, Maier, TR, Hussey, RS, Mitchum, MG, Davis, EL, Baum, TJ. Cellulose binding protein from the parasitic nematode *Heterodera schachtii* interacts with Arabidopsis pectin methylesterase: cooperative cell wall modification during parasitism. *Plant Cell*, 2008,20:3080–3093.
22. Scholl EH, Thome, JL, Mccarter, JP, Bird, DM. Horizontally transferred genes in plant-parasitic nematodes: a high-throughput genomic approach. *Genome Biology*, 2003, 4(6)R39
23. Gao B, Allen R, Davis EL, Baum TJ, Hussey RS *Int J Parasitol*. 2004 Nov; 34(12):1377-83.
24. Adam MAM, Phillips MS, Jones JT, Block VC, Characterization of the cellulose-binding protein Mj-cbp-1 of the root knot nematode, *Meloidogyne javanica*. *Physiol Mol Plant Path*, 2008, 72(1-3): 21-28.
25. Darabi M, Seddigh, S. Bioinformatic characterization of aspartic protease (AP) enzyme in seed plants. *Plant Sys Evol*, 2015, 301(10), 2399-2417.
26. Feng BZ, Li PQ, Fu L, Yu XM. Exploring laccase genes from plant pathogen genomes: a bioinformatic approach. *Gen. Mol. Res.* 2015, 14 (4):14019–14036.
27. Han Y, Zheng, QS, Wei, YP, Chen, J, Liu, R, Wan, HJ. *In silico* identification and analysis of phytoene synthase genes in plants. *Gen. Mol. Res.* 2015, 14 (3): 9412-9422,.
28. Vatansever R, Filiz, E, Ozyigit, I.I. *In silico* identification and comparative analysis of molybdenum (Mo) transporter genes in plants. *Braz. J. Bot.* 2015, 39: 1-13.
29. Moraes Filho R, Martins, LSS. *In silico* comparative analysis of tylenchid nematodes pectate lyases. *Gen. Mol. Res.*, 2016, gmr.15038402.
30. Dhia Bouktila, Yosra Khalfallah, Yosra Habachi-Houimli, Maha Mezghani-Khemakhem, Mohamed Makni & Hanem Makni (2015). Full-Genome identification and characterization of NBS-encoding Disease Resistance Genes in wheat. *Molecular Genetics and Genomics*. 290(1):257-271. DOI: 10.1007/s00438-014-0909-2.
31. Dhia Bouktila, Yosra Habachi-Houimli, Yosra Khalfallah, Maha Mezghani-Khemakhem, Mohamed Makni & Hanem Makni (2014). Characterization of novel wheat NBS domain-containing sequences and their utilization, *in silico*, for genome-scale R-gene mining. *Molecular Genetics and Genomics*. 289: 599-613. doi: 10.1007/s00438-014-0834-4.
32. Gasteiger E, Hoogland, C, Gattiker, A, Duvaud, S. et al. Protein identification and analysis tools on the ExpASY server. In: *The proteomics protocols handbook* (Walker JM, ed.). Humana Press, New York, 2005, 571-607.
33. Yu CS, Cheng CW, Su WC, Chang KC, Huang SW, Hwang JK, Lu CH. CELLO2GO: a web server for protein subCELLular Localization prediction with functional gene ontology annotation. *PLoS One*, 2014. 9: e99368. <http://dx.doi.org/10.1371/journal.pone.0099368>
34. Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. The TOPCONS web server for combined membrane protein topology and signal peptide prediction. *Nucleic Acids Res.* 2015, 43: 401-407. <http://dx.doi.org/10.1093/nar/gkv485>
35. Bailey TL, Boden M, Buske, FA, Frith, M, Grant, CE, Clementi, L, Ren, J, Li, W.W, Noble, W.S. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research*, 2009, 37:202-208.
36. Servant F, Bru C, Carrère S, Courcelle E, Gouzy, J, Peyruc, D, Kahn, D. ProDom: Automated clustering of homologous domains. *Briefings in Bioinformatics*. 2002, 3(3):246-251
37. Tamura K, Peterson D, Peterson N, Stecher, G. et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 2011, 28: 2731-2739. <http://dx.doi.org/10.1093/molbev/msr121>
38. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 2015, 10: 845-858. [tp://dx.doi.org/10.1038/nprot.\(2015\).053](tp://dx.doi.org/10.1038/nprot.(2015).053)
39. Jefferys BR, Kelley LA, Sternberg, M.J. Protein folding requires crowd control in a simulated cell. *J. Mol. Biol.* 2010, 397: 1329-1338. <http://dx.doi.org/10.1016/j.jmb.2010.01.074>
40. Lee J, Wu, S, Zhang Y. Ab Initio Protein Structure Prediction. In: Rigden D.J. (eds) *From Protein Structure to Function with Bioinformatics*. Springer, Dordrecht, 2009.

41. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* 2004, 25: 1605-1612. [http://dx.doi.org/10.1002/jcc.\(2008\)4](http://dx.doi.org/10.1002/jcc.(2008)4)
42. Chen VB, Arendall WB, Headd JJ, Keedy DA, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 2010,66: 12-21. <http://dx.doi.org/10.1107/S0907444909042073>
43. Wiederstein M, Sippl M.J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 2007, 35: W407-W410. <http://dx.doi.org/10.1093/nar/gkm290>
44. Shaw KL, Grimsley GR, Yakovlev GI, Makarov AA, Pace CN. The effect of net charge on the solubility, activity, and stability of ribonuclease Sa. *Protein Sci.* 2001, 10: 1206-1215. <http://dx.doi.org/10.1110/ps.440101>
45. Favey SC, Bourson Y, Bertheau A, Kotoujansky et al. Purification of the acidic pectate lyase of *Erwinia chrysanthemi* 3937 and sequence analysis of the corresponding gene. *J Gen Microbiol*, 1992, 138:499-508. <http://dx.doi.org/10.1099/00221287-138-3-499>
46. Castagnone-Sereno P, Danchin EGJ, Perfus-Barbeoch Le Abad, P. Diversity and Evolution of Root-Knot Nematodes, Genus *Meloidogyne*: New Insights from the Genomic Era. *Annu Rev Phytopathol*, 2013, 51:203-220.
47. Castagnone-Sereno P, Danchin, EGJ. Parasitic success without sex: the nematode experience. *J. Evol. Biol.*, 2014, 27:1323-1333,.
48. Söding J. Protein homology detection by HMM-HMM comparison. *Bioinformatics.* 2005, 21: 951-960. <http://dx.doi.org/10.1093/bioinformatics/bti125>
49. Pires VM, Henshaw JL, Prates JA, Bolam DN, Ferreira LM, Fontes CM, Henrissat B, Planas A, Gilbert HJ, Czjzek M. The crystal structure of the family 6 carbohydrate binding module from *Cellulobacterium mixtus* endoglucanase 5a in complex with oligosaccharides reveals two distinct binding sites with different ligand specificities. *J Biol Chem.* 2004,14(20):21560-8. DOI:10.1074/jbc.M401599200
50. Gilkes N.R, Henrissat B, Kilburn D.G, Miller, R.C. Jr, Warren, R.A. Domains in microbial beta-1, 4-glycanases: sequence conservation, function, and enzyme families. *Microbiol Rev*, 1991, 55(2):303-315.
51. Meinke A, Gilkes, N.R, Kilburn, D.G, Miller R.C. Jr, Warren, R.A.. Bacterial cellulose-binding domain-like sequences in eucaryotic polypeptides. *Protein Seq Data Anal.* 1991, 4(6):349-53.
52. Xu GY, Ong E, Gilkes NR, Kilburn DG, Muhandiram DR, Harris-Brandts M, Carver JP, Kay LE, Harvey TS .Solution structure of a cellulose-binding domain from *Cellulomonas fimi* by nuclear magnetic resonance spectroscopy. *Biochemistry*, 1995, 34(21):6993-7009
53. Krieger E, Joo, K, Lee, J, Lee, J, Raman, S, Thompson, J, Tyka, M, Baker, D, Karplus, K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins.* 2009, 77(9):114-22.
54. Chen V.B, Davis I.W, Richardson, D.C. KING (Kinemage, Next Generation): a versatile interactive molecular and scientific visualization program. *Protein Sci.* 2009, 18: 2403-2409. <http://dx.doi.org/10.1002/pro.250>
55. Moraes Filho R, Menezes AF, Martins LSS. *In silico* modeling and characterization of phytoparasitic nematodes translationally-controlled tumor protein. *Gen. Mol. Res.* 2017, 16 (3): gmr16039800.
56. Moreira GMSG, Conceição FR, McBride AJA, Pinto LS. Structure Predictions of Two *Bauhinia variegata* Lectins Reveal Patterns of CTerminal Properties in Single Chain Legume Lectins. *PLoS ONE*, 2013, 8(11): e81338. doi:10.1371/journal.pone.0081338
57. Waghmare S, Buxi A, Nandurkar Y, Shelke A, Chavan, R. *In silico* sequence analysis, homology modeling and function annotation of leishmanolysin from *Leishmania donovani*. *J Parasit Dis.* 2016, 40(4):1266–1269 DOI 10.1007/s12639-015-0665-1