# Cellulolytic fungi from central Morocco: comparative analysis of enzyme activities, *in silico* prediction of physico-chemical properties and molecular docking

**Nait M'Barek Hasna* and Hajjaj Hassan**
Laboratory of Plant Biotechnology and Molecular Biology, Faculty of Sciences, Moulay Ismail University, BP 11201 Zitoune, Meknes city, MOROCCO
*h.naitmbarek@edu.umi.ac.ma; hasnae.nait.mbarek@gmail.com

***Abstract***
*Fungal cellulases are important biocatalysts implicated in the conversion of lignocellulose for the production of second-generation bioethanol. Predicting their catalytic properties and bonding ability to substrate using computational tools is a prominent way to better assign their industrial use. Cellulolytic potential of two wood decaying fungi from central Morocco is studied. Comparison is made between cellulolytic activities in the presence of microcrystalline cellulose as the sole carbon source, results of in silico computation of physico-chemical properties and molecular docking of enzymes to substrate.*

*The strain Penicillium brasilianum (Pb) shows a pertinent cellulolytic potential producing 1.1 IU/ml total cellulase activity, 3.2 IU/ml endoglucanase activity and up to 41.7 IU/ml Xylanase activity. It yields 63.6% hydrolysis. Trichoderma atroviride (Ta) isolate is less cellulolytic and hydrolyses the polysaccharide only to 1.26%. Computational analysis shows that Pb cellulases are particularly stable and dock well to microcrystalline cellulose with 2.65Å to 2.92Å hydrogen bond distances. The studied cellulase of Ta is also stable and has strong docking properties to substrate. Reporting these results to quantitative activities shows evidence of the proteome diversity among isolates of the same Ta species mainly controlled by the gene expression under the surrounding environmental conditions.*

**Keywords:** Lignocellulose, cellulase, *Penicillium brasilianum*, *Trichoderma atroviride*, *in silico*, molecular docking, Morocco.

## Introduction

The conversion of cellulose feedstocks is nowadays a pertinent alternative to produce sustainable bio-molecules and biofuels. Its valorization is a key-clue to the industrial sector, allowing the decrease of its environmental impact without creating much tension on the food-chain resources[1].

Already more than seven decades of scientific exploration of cellulose-degrading microorganisms have passed, starting from the discovery of the cellulolytic ascomycete *Trichoderma reesei* in the Solomon Islands and its use to develop an enzyme market with a panel of cocktails for very diversified applications[2]. Several strains were studied and characterized for their ability to degrade lignocellulosic biomass and the stability of their enzymes towards process extreme conditions.

Today, the most performant industrial cocktails are from fungal strains of the following genera: Trichoderma, Aspergillus, Neurospora, Penicillium, Paecilomyces, Fusarium, Humicola, Mucor, Chaetomium, Poria, Tyromyces, Phanerochaete, Trametes, Pleurotus and many others[2]. Several fungal cellulases are already in use as auxiliary technological agents in the food processing industry, in textile, biorefinery, animal feed and other sectors[3].

The diversity of enzymes involved in cellulose degradation by fungi demonstrates the ability of these microorganisms to adapt to different biotopes and stand environmental changes. The secret behind fungal performance is particularly due to the production of different classes of cellulases acting in a synergistic way. The most studied classes of these enzymes are the Endoglucanases (EC 3.2.14), Exoglucanases (EC 3.2.1.91), β-glucosidases (EC 3.2.1.21) and the recently discovered, the Lytic Polysaccharide Monooxygenases (LPMOs)[2,4].

One prominent way of predicting stability and functionality of cellulases is the application of *in silico* analysis to protein sequences. The development of computational tools has come as in great service to the scientific committee and enhanced the discovery of pertinent and very stable industrial proteins. Nowadays, the application of bioinformatics revolutionized the field of molecular biology and is a key to predict and characterize functional properties of cellulases before large-scale production[5].

This study enables the comparison of the measured cellulolytic potential of two wild-type fungi isolated in central Morocco with the predicted physico-chemical properties of their enzymes and their docking ability to microcrystalline cellulose using a computational structured methodology.

## Material and Methods

**Culture Medium:** Cellulose liquid medium: 10mL Czapek concentrate (30g $NaNO_3$; 5g KCl; 5g $MgSO_4.7H_2O$; 0,1g $FeSO_4.7H_2O$; in 100mL distilled water); 1mL metal trace solution (1g $ZnSO_4.7H_2O$; 0,5g $CuSO_4.5H_2O$; in 100mL

distilled water); 1g $K_2HPO_4$; 5g microcrystalline cellulose (Avicel, Sigma Aldrich); all in 1L distilled water; pH 6.7.

**Chemicals (all Sigma Aldrich):** Whatmann qualitative filter paper (grade 1); low viscosity carboxymethyl cellulose (CMC); xylan from Birchwood; 3,5-dinitrosalicylic acid; phenol; D-cellobiose; D-glucose; D-xylose; sodium citrate tribasic dihydrate; potassium sodium tartrate tetrahydrate and sodium sulfite ($Na_2SO_3$).

**Fungal strains:** Two fungi isolated from decaying wood in central Morocco are used in this study. They were identified as *Penicillium brasilianum* (Pb) and *Trichoderma atroviride* (Ta) using molecular identification and screened with qualitative assay to assess their positive cellulolytic potential.[6]

**Culture conditions:** Both strains are grown in 250mL shaking flasks each containing 100mL cellulose liquid medium. All samples are inoculated with $10^7$ spores, incubated at 25°C for two weeks and agitated at 120 rpm. Proteins, cellulase activities and cellulose hydrolysis yield are quantified over time and all samples are analyzed in duplicate.

**Protein measurement and (hemi-)cellulase activities:** Samples to be analyzed are centrifuged at 3000rpm for 15min and 2°C. Supernatants are then recovered and constitute the fraction containing our extracellular cellulases. Proteins are measured using Lowry method as described[7]. Cellulase activity assays in[7] are used to follow the kinetics over time: Filter Paper Assay (FPA) for total saccharifying cellulase, endo-β-1-4 glucanase (EGL), β-glucosidase (BGL) and endo-xylanase (EXyl)[7,8]. One activity unit is considered as the amount of enzyme capable of releasing 1µmol of glucose or xylose per minute.

**Cellulose hydrolysis yield:** Hydrolysis yield ($Y_H$) of cellulose Avicel is calculated for each cellulolytic fungi. Every two days aliquots of 50mL are recovered, boiled for 5min in a boiling water bath and reducing sugars are quantified using DNS method[9]. Results are expressed as quantities of glucose equivalent reducing sugars. The following formula is used to calculate the hydrolysis yield as described[10]:

$Y_H$ = [Reducing sugars (mg) * H / Polysaccharide content in substrate (mg)] * 100,

where H is the hydrolysis conversion factor enabling to calculate anhydrous form of sugar from the form released by hydrolysis. This conversion aims on subtracting proportional weight added to each molecule by the water of hydrolysis. H is equivalent to 0.9 for glucose conversion.

**Retrieving enzyme sequences and phylogeny construction**: Available cellulase sequences for both strains (Pb and Ta) are retrieved from the protein database of the National Center for Biotechnology Information (NCBI). Eighteen (18) cellulases are used in this study and presented in table 1. More details on enzyme linear sequences are provided in supplementary material (Table S1).

**Table 1**
**Cellulase enzyme sequences for *P. brasilianum* (Pb) and *T. atroviride* (Ta) retrieved from the protein database of the National Center for Biotechnology Information (NCBI)**

| Fungus | Accession | Sequence length (a.a) | Designation |
|---|---|---|---|
| *Pb* | ACB06750 | 421 | Endo-1,4-beta-D-glucanase [*Penicillium brasilianum*] |
| | CEJ55762 | 332 | Putative Endo-beta-1,4-glucanase B [*Penicillium brasilianum*] |
| | CEJ59332 | 474 | Putative Endoglucanase [*Penicillium brasilianum*] |
| | CEJ59869 | 661 | Putative Endoglucanase Cel5C [*Penicillium brasilianum*] |
| | CEJ60291 | 374 | Putative Extracellular cellulase CelA/allergen Asp F7-like [*Penicillium brasilianum*] |
| | CEJ61474 | 421 | Putative Endo-1,4-beta-D-glucanase [*Penicillium brasilianum*] |
| | CEO59140 | 406 | Putative Endoglucanase II [*Penicillium brasilianum*] |
| | OOQ83231 | 474 | Putative endo-beta-1,4-glucanase B [*Penicillium brasilianum*] |
| | OOQ83663 | 421 | Putative endo-beta-1,4-glucanase B [*Penicillium brasilianum*] |
| | OOQ86698 | 338 | Putative endo-beta-1,4-glucanase B [*Penicillium brasilianum*] |
| | OOQ87332 | 390 | Putative cellulase [*Penicillium brasilianum*] |
| | OOQ87745 | 405 | Endoglucanase EG-II [*Penicillium brasilianum*] |
| | OOQ88946 | 661 | Extracellular endoglucanase [*Penicillium brasilianum*] |
| | OOQ91129 | 374 | Putative extracellular cellulase CelA [*Penicillium brasilianum*] |
| *Ta* | AJP16798 | 418 | Endoglucanase II [*Trichoderma atroviride*] |
| | EHK45466 | 422 | Glycoside hydrolase family 5 protein [*Trichoderma atroviride* IMI 206040] |
| | EHK47103 | 447 | Glycoside hydrolase family 5 protein [*Trichoderma atroviride* IMI 206040] |
| | EHK50177 | 418 | Glycoside hydrolase family 5 protein [*Trichoderma atroviride* IMI 206040] |

Protein sequences are aligned in MEGA 6 software[11] using the CLUSTALW® algorithm[12]. Neighbor-Joining phylogenetic tree is constructed based on p-distance model and 1000 bootstrap replicates to evaluate the fidelity of branching nodes.

**Physico-chemical characterization of enzymes and prediction of secondary structure:** For each enzyme sequence, *in silico* characterization is done using ProtParam tool from the ExPASy server[13]. Computed physico-chemical parameters are: molecular weight (MW), theoretical Isoelectric point (pI), amino-acid composition, Extinction Coefficient (EC)[14], *In vivo* half-life[15–17], Instability index (Ii)[18], Aliphatic index (Ai)[15], Grand Average Hydropathicity (GRAVY)[19] and the number of negative (R-) and positive residues (R+).

As the extinction coefficient is an indication for the quantity of light absorbed by the protein at a designated wavelength, it is an important parameter to follow proteins in purification processes. *In vivo* half-life is an indication of the time taken by half of the amount of the produced protein in a cell to disappear after its synthesis. Secondary structural elements are assessed using Self Optimized Prediction from Multiple Alignment tool (SOPMA) (https://omictools.com/self-optimized-prediction-from-multiple-alignment-tool)[20]. This tool allows us to predict the secondary conformation of different parts of the protein sequence.

**Homology modeling of tridimensional structure and model validation:** Prediction of tridimensional structure of the studied enzymes is conducted using automated mode of the SWISS-MODEL server[21–23]. This allows the selection of a best-fit enzyme model representing the lowest energy interactions. Only sequences showing a good stability profile using SOPMA are selected for tridimensional modeling. Two distance-dependent interaction potentials of mean force, namely C-β atoms and all atoms, are calculated for all computed models. Local backbone geometry of 3D structure is estimated with the torsion factor which represents the torsion angle potential over three consecutive amino acids. The calculation of QMEAN score enables the assessment of overall quality of the model which is validated by submission to RAMPAGE, ERRAT[22] and ProSA[24,25] servers respectively.

**Molecular docking of enzyme and substrate:** After validation, models are subjected to docking with cellulose using PATCH-DOCK program, an algorithm based on shape complementarity principles[17,26]. Cellulose 3D structure is recovered from the PubChem database of the NCBI. OpenBabel 2.4.1 assures the conversion of files format from SDF to PDB and docked complexes are visualized using UCSF Chimera 1.11.2 software. Potential hydrogen bonds between cellulose and interacting amino acids of the active site are identified using the same program calculating the distances.

## Results and Discussion

**Saccharification potential of fungal strains:** Following up enzyme activities over time reveals that the strain *P. brasilianum* is a potential cellulolytic fungus. Extracellular supernatants recovered from cellulose media show total saccharifying cellulase activity of 1.1 IU/ml, endoglucanase activity of 0.16 IU/ml and xylanase activity of 37.4 IU/ml respectively. β-Glucosidase activity is not detected in all samples.

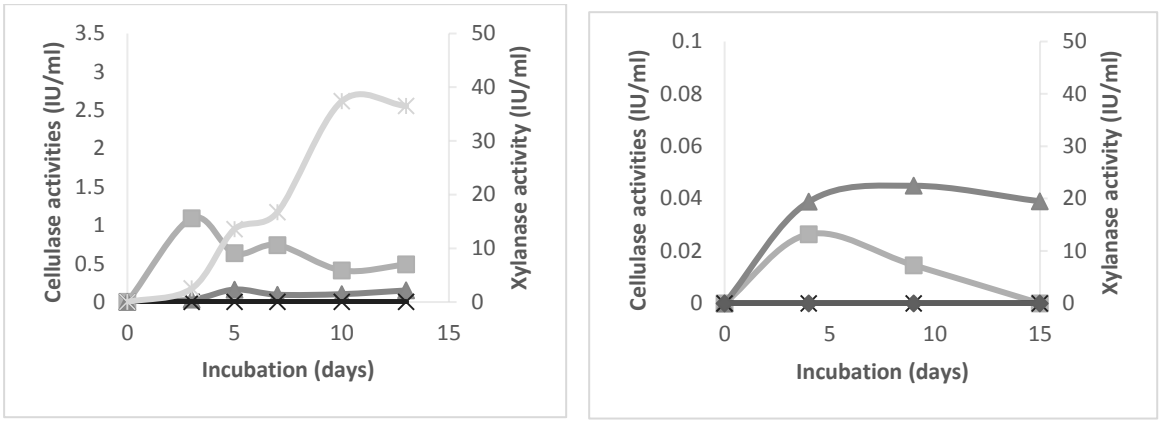
The kinetics of production of enzymes shows that total cellulase activity is maximally expressed on the third day of incubation while xylanase reaches its optimum in the tenth day. This illustrates the synergy and action complementarity of cellulolytic and xylanolytic enzymes in the hydrolysis process as stated by previous studies[27]. When glucose is added in low concentration to the medium in the second day of incubation, we observe a significant increase in total cellulase activity compared to standard conditions (data not shown). This is because the expression of cellulase enzymes by *P. brasilianum* is induced in the presence of glucose[28].

*T. atroviride* isolate has less cellulolytic performances with a maximum FPA activity of 0.026 IU/ml, 0.045 IU/ml endoglucanase activity and no detected β-Glucosidase and xylanase activities. Fig. 1 represents the results for kinetics of production of (hemi)-cellulolytic activities for both strains.

The maximum hydrolysis of cellulose for *P. brasilianum* is also obtained in the third day of incubation with a calculated yield of 63.6%. *T. atroviride* hydrolyses the polymer only to 1.26%, which is concordant with the measured activities.

**Multiple Sequence Alignment and phylogeny construction of cellulase enzymes:** Alignment of protein sequences reveals that cellulases from *P. brasilianum* and *T. atroviride* are very phylogenetically close with reliable nodes statistics (bootstrap percentage values ranging from 65% to 100%). Particularly, cellulases of *P. brasilianum* with accessions ACB06750, OOQ83663, CEJ61474, CEJ55762, OOQ86698, CEJ59332 and OOQ83231 are closely gathered from the evolutionary point of view. Great similarities are also observed among cellulases EHK47103, EHK45466, AJP16798, EHK50177 from *T. atroviride* and the enzyme clade gathering CEO59140 and OOQ87745 from *P. brasilianum* (Fig. 2).

**Physico-chemical parameters of enzymes:** As illustrated in table 2, retrieved sequences for both strains are 332 to 661 amino acids long and their molecular weight (MW) ranging from 36.2 to 70.4 kDa. All studied enzyme sequences have acidic nature. Instability index indicates that cellulase enzymes are very stable in test tube experiment and have *in vivo* half-life up to sixteen hours in standard conditions[17,26]. Cellulases from *P. brasilianum* with accessions CEJ55762 and OOQ86698 are particularly very stable proteins.




**(a)** **(b)**

**Fig. 1: Kinetics of production of cellulase and xylanase activities for a) *Penicillium brasilianum* and b) *Trichoderma atroviride* in the presence of microcrystalline cellulose as the sole carbon source. ■ Filter Paper Assay (total cellulase activity), ▲ Endo-β-1-4 glucanase activity, x β-Glucosidase activity and x Xylanase activity. IU/mL designs International Unit per Milliliter of supernatant**

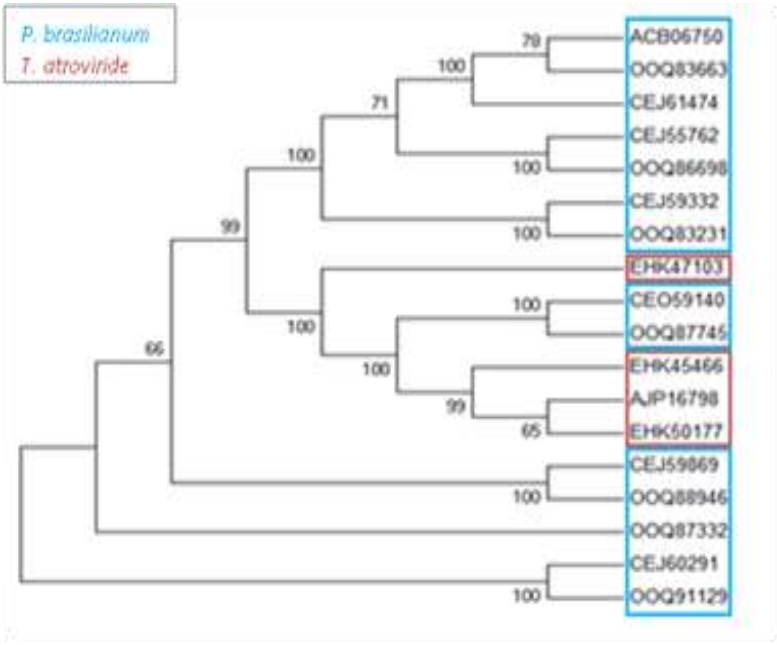



**Fig. 2: Neighbor-Joining tree with 1000 bootstraps replicates for cellulase sequences (1000 bootstraps replicates) retrieved from the protein database of the National Center for Biotechnology Information (NCBI)**

All studied enzymes present high aliphatic index and very low and negative GRAVY index reflecting their stability in a wide range of temperature with good interaction with water molecules (less hydrophobicity). We also observe that the studied proteins contain more negatively charged amino acids than the positively charged ones.

**Prediction of secondary structure and tridimensional conformation of cellulase enzymes:** Cellulase sequences of both strains present a secondary structure with good α-helix content (Table 2). We highlight that the good organization of different parts of proteins' sequences in random coil impacts their flexibility, giving them more conformational search during enzymatic turnover[17,29]. However, those parts of the protein sequence are most confronted to mutations and present evolutionary hotspots[29].

For the above-mentioned reasons, only enzyme sequences with stable physico-chemical parameters and bidimensional properties are selected to conduct further *in silico* analysis. We chose three cellulase sequences: CEJ55762 and OOQ86698 from *P. brasilianum* and EHK50177 from *T. atroviride* for having the lowest instability index and random coils and the biggest α-helix content.

Tridimensional modeling of the selected sequences using SWISS MODEL server reveals interesting results. Sequence identity of predicted models is higher than 76%, Global Model Quality Estimate (GMQE) is more than 0.7 and good QMEAN overall quality scores (Table 3). The best model fit (5i77.1.A) for cellulases CEJ55762 and OOQ86698 is obviously the same with insignificant difference in identity

and quality scores meaning that these two sequences have high similarity and approving phylogenetic tree results.

**Model validation and docking with microcrystalline cellulose:** Computing the tridimensional structures using RAMPAGE server leads to a global and final estimation of the quality of the best-fit 3D model by estimating the adequacy of each amino-acid position in the 3D structure. Scores of Residues in Favored Region (RFR) (>96%), Residues in Allowed Region (RAR) (2.6 - 2.8%) and Residues in Outlier Region (ROR) (<1%) reflect good quality of all templates (Table 4). In addition, ERRAT (Fig. S2, supplementary material) and ProSA plots approve the high quality of the tridimensional structures. Figure 3 represents graph results using RAMPAGE and ProSA servers.

**Table 2**
**ProtParam and Self Optimized Prediction from Multiple Alignment (SOPMA) computed parameters for cellulases of *Penicillium brasilianum* (Pb) and *Trichoderma atroviride* (Ta)**

| Fungus | Accession | MW[1] | pI[2] | EC[3] | *In vivo* HL[4] | Ii[5] | Ai[6] | GRAVY[7] | R-[8] | R+[9] | α-Helix % | Extended strand % | β-Turn % | Random coil % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Pb* | ACB06750 | 45063.61 | 4.77 | 107175 | >16h | 24.27 | 65.20 | -0.168 | 27 | 18 | 31.83 | 19.95 | 7.13 | 41.09 |
| | CEJ55762 | 36235.42 | 4.39 | 81485 | >16h | 9.41 | 75.84 | -0.129 | 35 | 17 | 36.45 | 22.59 | 11.14 | 29.82 |
| | CEJ59332 | 50638.85 | 4.84 | 114165 | >16h | 27.38 | 62.41 | -0.245 | 43 | 29 | 31.22 | 19.41 | 9.70 | 39.66 |
| | CEJ59869 | 70389.93 | 4.86 | 139730 | >16h | 31.94 | 74.84 | -0.015 | 47 | 29 | 22.39 | 29.50 | 8.62 | 39.49 |
| | CEJ60291 | 37967.30 | 4.40 | 41870 | >16h | 34.96 | 62.11 | -0.081 | 31 | 13 | 17.91 | 28.61 | 6.95 | 46.52 |
| | CEJ61474 | 45133.76 | 4.96 | 107175 | >16h | 24.48 | 64.04 | -0.186 | 26 | 19 | 30.64 | 21.38 | 7.13 | 40.86 |
| | CEO59140 | 43389.18 | 5.37 | 76485 | >16h | 32.76 | 73.23 | -0.174 | 27 | 19 | 25.62 | 24.38 | 12.56 | 37.44 |
| | OOQ83231 | 50655.81 | 4.75 | 114165 | >16h | 28.09 | 62.19 | -0.243 | 44 | 28 | 31.43 | 18.99 | 8.65 | 40.93 |
| | OOQ83663 | 45063.61 | 4.77 | 107175 | >16h | 23.71 | 65.20 | -0.166 | 27 | 18 | 31.12 | 20.67 | 7.13 | 41.09 |
| | OOQ86698 | 36905.20 | 4.49 | 81485 | >16h | 8.90 | 75.36 | -0.154 | 34 | 18 | 35.80 | 23.08 | 11.83 | 29.29 |
| | OOQ87332 | 42697.80 | 5.13 | 104655 | >16h | 31.18 | 80.38 | -0.034 | 32 | 21 | 34.36 | 24.87 | 9.23 | 31.54 |
| | OOQ87745 | 43290.05 | 5.37 | 76485 | >16h | 32.88 | 73.41 | -0.162 | 27 | 19 | 25.43 | 24.20 | 12.84 | 37.53 |
| | OOQ88946 | 70378.00 | 5.03 | 135720 | >16h | 33.64 | 75.58 | -0.011 | 45 | 29 | 22.39 | 28.74 | 8.62 | 40.24 |
| | OOQ91129 | 37831.03 | 4.40 | 40380 | >16h | 35.44 | 62.62 | -0.086 | 31 | 13 | 16.31 | 29.68 | 6.95 | 47.06 |
| *Ta* | AJP16798 | 44208.90 | 4.96 | 85090 | >16h | 30.96 | 70.24 | -0.194 | 25 | 17 | 28.95 | 21.53 | 9.57 | 39.95 |
| | EHK45466 | 44473.33 | 4.59 | 85550 | >16h | 29.67 | 73.60 | -0.127 | 27 | 16 | 23.70 | 22.75 | 9.72 | 43.84 |
| | EHK47103 | 47316.61 | 4.44 | 59860 | >16h | 29.47 | 72.10 | -0.191 | 49 | 27 | 27.07 | 21.70 | 9.40 | 41.83 |
| | EHK50177 | 44161.76 | 5.19 | 69620 | >16h | 28.15 | 66.34 | -0.238 | 23 | 16 | 29.19 | 21.05 | 10.05 | 39.71 |

[1]Molecular Weight in Dalton, [2]Isoelectric point, [3]EC Extinction Coefficient ($M^{-1}$ $cm^{-1}$) at $\lambda_{280}$ in water (assuming that all pairs of Cys residues form cysteines), [4]*In vivo* Half Life, [5]Instability index, [6]Aliphatic index, [7]Grand Average Hydropathicity, [8]Negative Residues, [9]Positive Residues

**Table 3**
**Tridimensional structure modeling of selected cellulolytic enzymes using SWISS MODEL server and quality verification**

| F. | Accession | Template | Sequence Id[1] | GMQE[2] | QMEAN[3] |
|---|---|---|---|---|---|
| *Pb* | CEJ55762 | 5i77.1.A | 76.67% | 0.82 | -1.32 |
| | OOQ86698 | 5i77.1.A | 76.67% | 0.80 | -1.55 |
| *Ta* | EHK50177 | 3qr3.1.A | 81.48% | 0.75 | -0.42 |

[1]Sequence Identity of the model, [2]Global Model Quality Estimate,
[3]Quality Model Estimate based on global and local absolute scoring.

**Table 4**
**Best model search and validation of tridimensional structure using RAMPAGE, ERRAT and ProSA servers**

| Fungus | Accession | Best-fit Template | RAMPAGE | | | ERRAT | ProSA |
|---|---|---|---|---|---|---|---|
| | | | RFR[1] | RAR[2] | ROR[3] | Quality factor | Z-score |
| *Pb* | CEJ55762 | 5i77.1.A | 97% | 2.7% | 0.3% | 97.26% | -8.56 |
| | OOQ86698 | 5i77.1.A | 96.7% | 2.6% | 0.7% | 97.635% | -8.64 |
| *Ta* | EHK50177 | 3qr3.1.A | 97.2% | 2.8% | 0% | 94.904% | -9.03 |

[1]Residues in Favored Region, [2]Residues in Allowed Region, [3]Residues in Outlier Region

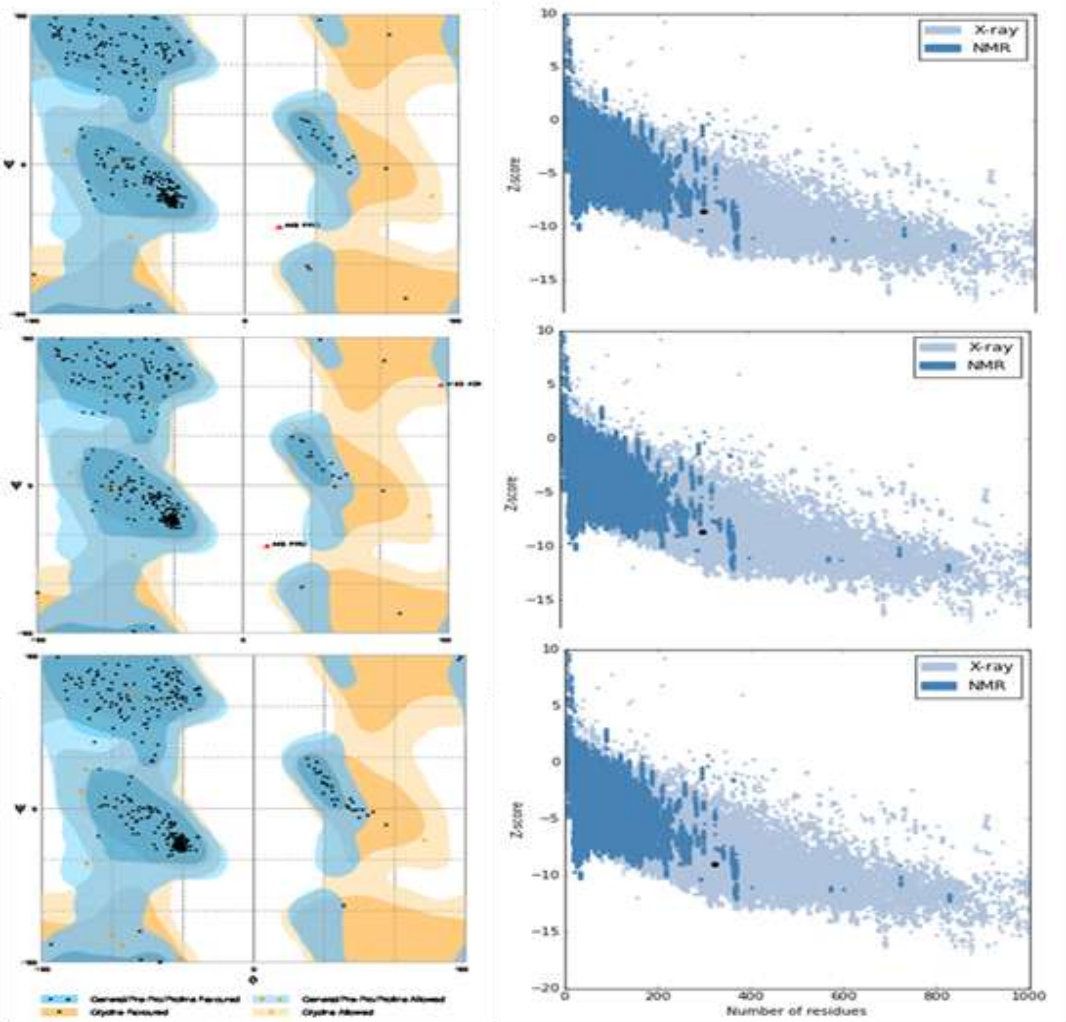



**Figure 3: RAMPAGE (left) and ProSA (right) plots for cellulases CEJ55762 (line 1), OOQ86698 (line 2) and EHK50177 (line 3) respectively. Black dots in RAMPAGE plots are amino acids of the Allowed and Favored Regions (blue zone) and the few others in Outlier Region (orange zone). Each black dot in ProSA graphs presents the position of cellulase compared to analyzed proteins of the ProSA database according to the number of residues and the Z-score of the protein**

Docking of cellulase 3D structures with cellulose is based on shape complementarity principles. Results for CEJ55762 from *P. brasilianum* show that GLU81 and GLY126 are the only amino acids from active site that are interacting with the substrate and forming 2.92Å and 2.65Å hydrogen bond distances, respectively (Fig. 4). For cellulase OOQ86698 of the same fungus, TRP206 bonds to cellulose with a 2.66Å distance. We then observe that even though models of both enzymes are closely the same, the difference in amino-acid composition plays a major role in docking properties. Cellulase EHK50177 from *T. atroviride* has more affinity to substrate with three amino acids implicated: THR108, TYR197 and TYR311 with 2.24Å, 3.32Å and 2.84Å hydrogen bond distances, respectively (Fig. 4).

However, measured cellulase activities for fungal isolates used in this study do not totally reflect these results. We think it is due to divergence in isoenzyme forms produced by *T. atroviride* isolate mainly controlled by gene expression under different environmental conditions.

## Conclusion

Isolation of new wild-type fungal species and the characterization of their cellulolytic enzymes open new insights in the conception of innovative bioprocesses. Fungi are good candidates known for the stability of their enzymes and their ability to stand process parameters. Computational characterization of these proteins is a basic step to assign the best utilization. With the use of *in silico* analysis, we are nowadays very close to predicting the functionality of cellulases and say in advance weather it is suitable or not for the industrial application. This study highlights the pertinence of new (hemi-) cellulolytic isolates from central Morocco and gives additional information on the relevance of xylanase activity produced by a special *Penicillium brasilianum* isolate for the breakdown of cellulosic biomass. Hence, the major perspective of our study is to characterize the enzymatic cocktail produced by this latter and to study the effect of substrate on the kinetics of production using a batch bioreactor system.

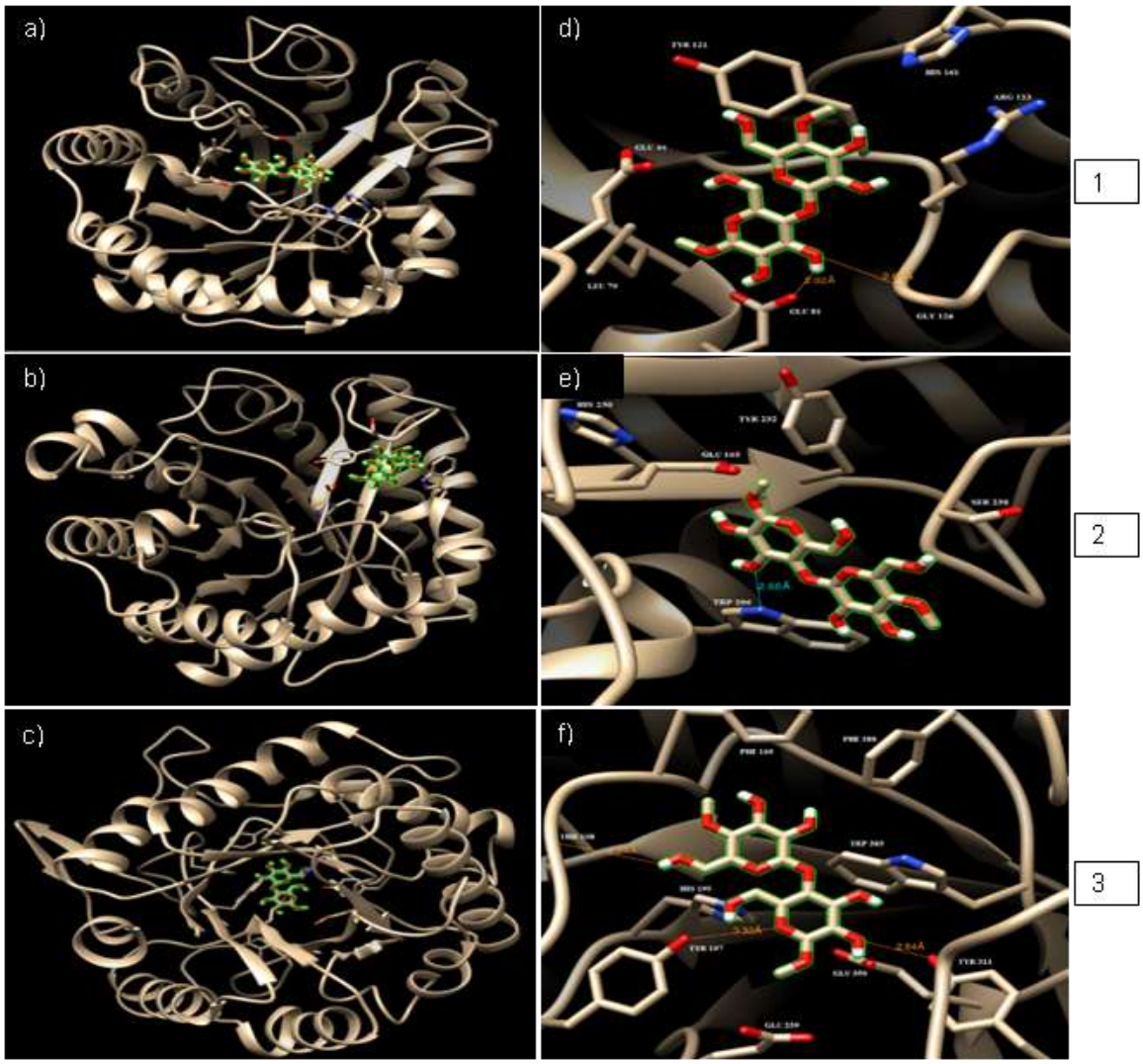



**Figure 4: Molecular docking of cellulases 1) CEJ55762, 2) OOQ86698 and 3) EHK50177 with microcrystalline cellulose visualized under UCSF Chimera 1.11.2 software. a), b) and c) are results of molecular docking. d), e) and f) represent active sites with hydrogen bonds distances**

**Supplementary material**
**Table S1**
**Linear sequences of cellulase from *Penicillium brasilianum* (*Pb*) and *Trichoderma atroviride* (*Ta*) retrieved from the National Center for Biotechnology Information (NCBI) and used in this study.**

| F. | Accession | Length (a.a.) | Protein linear sequence |
|---|---|---|---|
| *Pb* | ACB06750 | 421 | MKYPLLLATSAALALAGPQGFSKRASSFVWFGTSESGAEFGNQNIPGVLGTDYIWPDT SAIQTLRNAGMNIFRVAFLMERLVPTTLTSTPDSTYLQDLKSVLDYITSTGAYAIVDPHN FGRYYGNIINSTSDFAAFWTTVAKQFASNDKVIFDTNNEFNTEDQTLVLNLNQAAINAI RAAGATSQYIFVEGNSWSGAWTWTSVNTNLVSLTDPNNKIVYEMHQYLDSDGSGTSD TCVSSTIGQERVQSATEWLKSNGKLGFLGEFAGGANSVCQSAVTGMLDYMQANSDV WLGASWWAAGPWWGTYIYSIEPPSGTAYSYYLNILSAYFPSSSGSSTTTTSTTTRSTST STTVSTTKSTSTTTSATKSTSTTTSTTSTGSTATATASHWAQCGGIGWTGATTCASPYTC QVQNAYYSQCL |
| | CEJ55762 | 332 | MKLNTVILLAAAAGSAVAAPLKNAKRASSFEWFGTSESGAEFGEGNLPGVWGTDYIFP DTSTIQTLIDDGMNIFRVQFLMERLTPSGMTGSFDSDYLKNLTTVVNYITEAGAHAVID PHNYGRFNGAIITSTSDFQTFWTNVAGQFKTNSLVIFDTNNEYHDMDQTLVLNLNQAA INGIRAAGATSQYIFVEGNSYTGAWTWTTVNDNLKDLTDPQDKIVYEMHQYLDSDGS GTSETCVSITIGQERVTDATQWLIDNGKVGVLGEFAGGVNDQCKTAITGMLDYLAENS SVWKGAMWWAAGPWWGTYIYSMEPPSGVAYTGMLATLKSYFP |

| | | | |
|---|---|---|---|
| | CEJ59332 | 474 | MKFTNMVLAASAAGMAVAYPRGRDVVPEKRSVEKKRATGFTWVGVSESGAEFGSAI<br>PGTLGTDYTWPVTSQIQVLRDAGMNIFRVPFLMERLVPSSITGTLDATYLAALKSTVNY<br>ITESGAYAVLDPHNYGRYSGSIISSTDNFKAWWKTVATEFASNDKVIFDTNNEYHDMD<br>QTLVLNLNQAAIDGIRAAGATSQYIFVEGNAWTGAWSWTSNNDNMKDLTDTEDKIVY<br>EMHQYLDSDSSGTSETCVSSTIGKERLEAATTWLKDNAKKGFIGEFAGGVNSVCETAV<br>EGMLSYMSDNSDVWMGAEWWSAGPWWGSYMYSLEPTDGPAYSTYLSILEKYFVSGS<br>SSSSSSSSSTTTAKATTSTTTAAAQHTTTSTSTSTTAAVKVTTTPNNQVEVSSSASSTTTT<br>APAVVQTTLVSKPSTTKSASTTTTAASSSSTGGTVAHWYQCGGINWTGATACESGYTC<br>VVQNPYYSQCL |
| | CEJ59869 | 661 | MHSLCVIAIAALFGSAGAQVAGYGQCGGTGYSGSTTCTSGFYCTSQNPYYYQCIPGTA<br>TTSTTTTSKTSVTTTTSTTTTTKTSSTTSASTTLKTTTTQTATQTSSATGAATCTGAFSAI<br>SASDFVAKLHPGWNLGNTLDATPDEGSWNNAPVVASTFSTVQKAGFKSVRIPVTYAY<br>HFTGSSPDWTIDPTWLARVSEVVDMVTSLGLYAIVDAHHDSWIWADVTASGANLTMI<br>EEKFYRLWYQVGTTLACKSSLVAFEPINEPPCNTATDAAEINKLNQIFLQAINDAGGFN<br>PQRVVTLVGGGEDSVKTSEWFVAPTGFSNPYAIQFHYYSPYDFIFSAWGKTIWGSDTD<br>KTTVLTDFTLIRNNFTDVPILIGEYDASPTNTEPAARWKYVDYLIRSARSLNFACVLWD<br>NGLDHLDRNADTWRDPNSISIITKSTASSANSLPDSSEDSTATSQWSSAYIFHEYGTSVV<br>AQSLSFIFNGNTLTSISDSTGSTLTSGTDYSVSGANITFSASYLSKHITSTTSPGIIANLTLT<br>FSGGQSSPIVQLVQWKTPTLSSTSAVASSVSGSDLSIPITWGGLPEIAAVKALTTSGVYL<br>VDTWTVYLGPLQQARTTYSSQYNWDSSHVIITSSAISSVISAGVSTVFTFEFYPRVEGGV<br>NAVNFTLTV |
| | CEJ60291 | 374 | MKYQRLLSLGAVLFGATATVSAKHLGGRHEHNQCPKGYTVSVYTSYSTVYLTSTYAA<br>APEVTSTSSPVAVMETPAPVETTTAVGDEAASTSSSTTSSTPVPTSEAPVAMTTQSTVA<br>VVAELSTENAVAAVATSSTTSQETTSSASSTSTTEAAATATSSTSSASTSTKSSTSSSSGS<br>TSGEATFYGGNVAGGTCSYSGYTLPSGLFGTAFSGAAWDNAAECGACVAVTGPNGNT<br>IKAMIVDQCPECAASHLDLFQDAFAELSDISAGVIDITWEYTACDLSGPLKLKNKEGTS<br>QYWFSMQVVNANEPVTKLEVSTDGGSTWQSTTRTSYNYFEQSSGFGVDTVDVRVTGQ<br>SGTTVVVNSVGCSSGSEITASSNL |
| | CEJ61474 | 421 | MKYPLLLATSAALALAGPQGFSKRASSFVWFGTSESGAEFGNQNIPGVLGTDYIWPDT<br>SAIQTLRNAGMNIFRVPFMMERLVPTTLTSTPNSKYLQDLKSTVDYITSTGAYAIVDPH<br>NFGRYYGNIINSTSDFAAFWTTVAKQFASNDKVIFDTNNEFNTEDQTLVLNLNQAAINA<br>IRAAGATSQYIFVEGNSWSGAWTWTSVNTNLVSLTDPNNKIVYEMHQYLDSDGSGTS<br>DTCVSSTIGQERVQSATEWLKSNGKLGFLGEFAGGANSVCQSAVTGMLDYMQANSDV<br>WLGASWWAAGPWWGTYIYSIEPPSGTAYSYYLNILSAYFPSSSGSSTTTTTSTTTRSTST<br>STTVSTTKSISTTTSATKSTSTTTSTTSTGSTATATASHWAQCGGIGWTGATTCASPYTC<br>QVQNAYYSQCL |
| | CEO59140 | 406 | MKSQPGHSLLALLGIGGTVLGQQSAWGQCGGTGWTGQTTCVSGYHCSVQNNYYSQC<br>LPGTSSTSSGTSTTTTLVTKTSATTSTSSSVPTTKVKFAGVNIAGFDFGMVISGTQDQSQ<br>VFDESVDGVNQMKHFVNDDTFNIFRLPTGWQYIVANNLGGSLDSTNFGKYDKLVQGC<br>LSLGVYCIVDIHNYARWNGGIIGQGGPTDDQFVNLWTQLATHYKSQSKVIFGIMNEPH<br>DLNINTWATTVQKTVTAIRNAGATSQMILLPGTDYTSAANFIENGSGAALSAVVNPDG<br>STTNLIFDVHKYLDSDNSGTHAECATNNVDAFNTLAQWLRSGKRQALLSETGGGNVQ<br>SCATYMCQQLDALNANADVYLGWTSWSAGAFQTTYTLSEVPVNGVDQYLVQQCFVP<br>KWKS |
| | OOQ83231 | 474 | MKFTNMVLAASAAGMAVAYPRGRDVVPEKRSVEKKRATGFTWVGVSESGAEFGSAI<br>PGTLGTDYTWPVTSQIQVLRDAGMNIFRVPFLMERLVPSSITGTLDATYLAALKSTVNY<br>ITESGAYAVLDPHNYGRYSGSIISSTDNFKAWWKTVATEFASNDKVIFDTNNEYHDMD<br>QTLVLNLNQAAIDGIRAAGATSQYIFVEGNAWTGAWSWTSNNDNMKDLTDTEDKIVY<br>EMHQYLDSDSSGTSETCVSSTIGKERLEAATTWLKDNGKKGFIGEFAGGVNSVCETAV<br>EGMLSYMSDNSDVWMGAEWWSAGPWWGSYMYSLEPTDGPAYSTYLSILEKYFVSGS<br>SSSSSSSSSTTTAKATTSTTTAAAQHTTTSTSTSTTAAVEVTTTPNNQVEVSSSASSTTTT<br>APAVVQTTLVSRPSTCKSASTTTTAASSSSTGGTVAHWYQCGGINWTGATACESGYTC<br>VVQNPYYSQCL |
| | OOQ83663 | 421 | MKYPLLLATSAALALAGPQGFSKRASSFVWFGTSESGAEFGNQNIPGVLGTDYIWPDT<br>SAIQTLRNAGMNIFRVAFLMERLVPTTLTSTPDSTYLQDLKSTVDYITSTGAYAIVDPHN<br>FGRYYGNIINSTSDFAAFWTTVAKQFASNDKVIFDTNNEFNTEDQTLVLNLNQAAINAI<br>RAAGATSQYIFVEGNSWSGAWTWTSVNTNLVSLTDPNNKIVYEMHQYLDSDGSGTSD<br>TCVSSTIGQERVQSATEWLKSNGKLGFLGEFAGGANSVCQSAVTGMLDYMQANSDV<br>WLGASWWAAGPWWGTYIYSIEPPSGTAYSYYLNILSAYFPSSSGSSTTTTTSTTTRSTST<br>STIVSTTKSTSTTTSATKSTSTTTSTTSTGSTATATASHWAQCGGIGWTGATTCASPYTC<br>QVQNAYYSQCL |

| | | | |
|---|---|---|---|
| | OOQ86698 | 338 | MKLNTVILLAAAAGSAVAAPLKNAKRASSFEWFGTSESGAEFGEGNLPGVWPSTILKG<br>TDYIFPDTSTIQTLIDNGMNIFRVQFLMERLTPNGMTGSFDSDYLKNLTTVVNYITEAGA<br>HAVIDPHNYGRFNGAIITSTSDFQTFWTNVAGQFKTNSLVIFDTNNEYHDMDQTLVLN<br>LNQAAINGIRAAGATSQYIFVEGNSYTGAWTWTTVNDNLKDLTDPQDKIVYEMHQYL<br>DSDGSGTSETCVSTTIGQERVTDATQWLIDNGKVGVLGEFAGGVNDQCKTAITGMLD<br>YLAENSSVWKGAMWWAAGPWWGTYIYSMEPPSGVAYTGMLSTLKSYFP |
| | OOQ87332 | 390 | MILINKFFLYPLLVCLATARKVASTQYLDWKTFNAVGVNLGGWLEQESTIDTTWWSQ<br>HSGGAVDEWGLCAHLGGQCGPVLERRYATWITTADIDTLGAAGINVLRIPTTYAAWIE<br>VPGSQLYHGNQLSFLSSIANYAINKYRMHIIIDIHSLPGGVNGFPFGEAYGHYGWFNNA<br>TALKYSLEAVDAAIAFIQASKSPQSFTLEPINEPVDVQDLSLFGSPYCLTDSGAAYLASYI<br>HQVLSKVQAVNPRIPVMFQGSFKGEAYWSSNFSTGTNLVFDLHNYYFAGRSASSANIT<br>DYICEDAVTSAGDGKFPTFIGEWSIEAEIDNTFSSRQKSLETGLAAWKKYTQGSAYWTA<br>KFAGNTTVSGQGTQADYWNYEAFINLGYTGSTSSAVTC |
| | OOQ87745 | 405 | MKSQPGHSLLALLGIGGTVLGQQSAWGQCGGTGWTGQTTCVSGYHCSVQNSYYSQC<br>LPGTSSTSSGTSTTTLVTRTSATTSTSSSVPTTKVRFAGVNIAGFDFGMVISGTQDQSQV<br>FDESVDGVNQMKHFVNDDTFNIFRLPTGWQYIVANNLGGSLDSTNFGKYDKLVQGCL<br>SLGVYCIVDIHNYARWNGGIIGQGGPTDDQFVSLWTQLATHYKSQSKVIFGIMNEPHD<br>LNINTWATTVQKTVTAIRNAGATSQMILLPGTDYTSAANFIENGSGAALSAVVNPDGST<br>TNLIFDVHKYLDSDNSGTHAECATNNVDAFNTLAQWLRSGKRQALLSETGGGNVQSC<br>ATYMCQQLDALNANADVYLGWTSWSAGAFQTTYTLSEVPVNGVDQYLVQQCFVPK<br>WKS |
| | OOQ88946 | 661 | MHSLYVIAIAALFGSAHAQVAGYGQCGGTGYSGSTTCTSGFYCTSQNPYYYQCIPGTA<br>TTSTTTTSKTSVTTTTSTTTTTKTSSTTSASTTLKTTTTQAATQTSSSTGAATCTGTFSAIS<br>ASEFVAKLHPGWNLGNTLDATPDEGSWNNAPVVASTFSAVQKAGFKSVRIPVTYAYH<br>FTGSSPDWTIDPTWLARVSEVVDMVTSLGLYAIVDAHHDSWIWADVTASGANLTMIE<br>EKFYRLWYQVGTTLACKSSLVAFEPINEPPCNTATDAAEINKLNQIFLQAINDAGGFNP<br>QRVVTLVGGGEDSVKTSEWFVAPTGFSNPYAIQFHYYSPYDFIFSAWGKTIWGSDTDK<br>ATVLTDFTLIRNNFTDVPILIGEYDASPTNTEPAARWKYVDYLIRSARSLNFACVLWDN<br>GLDHLDRNAGTWRDPNSISIITKSTASSANSLPDSSEDSTATSQLSSAYIFHQYGTSVVA<br>QSLPFIFNGNTLTSISDSTGSTLTSGTDYSVSGANITFSASYLSKHITSTTSPGIIANLTLTF<br>SGGQSSPIVQLVQWKTPTLSSTSAVASSVSGSDLSIPITWGGLPEIAAVKALTTSGVYLV<br>DTWTVYLGPLQQARTTYSSQYNWDSSHVIITSSAISS<br>VISAGVSTVFTFEFYPRVEGGVNAVNFTLTV |
| | OOQ91129 | 374 | MKYQRLISLGAVLFGATATVSAKHLGGRHEHNQCPKGYTVSVYTSYSTVYLTSTSSAA<br>PEVTSTSSPVAVVETPAPVETTTAVGDEAASTSSSTTSSTPVPTSEAPVATTTQSTVAVV<br>AELSTENAVAAVATSSSTSQETTSSASSTSTTEAAATATSSTSSASTSTKSSTSSSSGSTS<br>GEATFYGGNVAGGTCSYSGYTLPSGLFGTAFSGAAWDNAAECGACVAVTGPNGNTIK<br>AMIVDQCPECAASHLDLFQDAFAELSDISAGVIDITWEYTACDLSGPLKLKNKEGTSQY<br>WFSMQVVNANEPVTKLEVSTDGGSTWQSTTRTSYNYFEQSSGFGVDTVDVRVTGQ<br>SGTTVVVNSVGCSSGSEITASSNL |
| *Ta* | AJP16798 | 418 | MNKSVAPLLLAASTLVGGVAAQQTVWGQCGGIGWSGPTNCAPGSACSTLNSYYAQCI<br>PGATSMTTSTRPPSGPTTTTRATSTTSSTPPTSSGDRFVGVNVAGFDFGCSTDGTCVTSK<br>VYPPLKNFTGSNNYPDGIGQMQHFVNDDGMTIFRLPVGWQYLVNNNLGGPLDPTSISK<br>YDQLVQGCLSLGVYCIVDIHNYARWNGGIIGQGGPTNAQFTSLWSQLASKYASQSRV<br>WFGIMNEPHDVNINTWAATVQEVVTAIRNAGATSQYISLPGNDWQSAAAFISDGSAAA<br>LSQVTNPDGSTTNLIFDVHKYLDSDNSGTHAECTTNNIDGAFAPLATWLRQNNRQAILT<br>ETGGGNVQSCIQYLCQQVQYLNQNSDVYLGYVGWGAGSFDNTYVLTETPTGSGNSW<br>TDTSLVSSCLARK |
| | EHK45466 | 422 | MINNKAALLFAAYAGVSGVAAQQQTTWGQCGGQGYSGPTSCVSGAACSTINPYYAQ<br>CIPATGIITSTTTRATSATSTLKSTTASASSTPPPSNGSGTQFAGINIAGFDFSCSTDGTCN<br>VSGAYPPLKNYDGANNYPDGVGQMQHFVKDDGFNIFRLPVGWQYLVNATLGATLNP<br>TNLGYYDQLVQGCLDTGAYCIIDIHNYARWNTGIIGQGGPTNAQFVNLWTQIATKYAS<br>EPKIWFGVMNEPHDVNITTWAATVQLVVTAIRNAGATSQYISLPGTDWQSAGSIITDGG<br>VAALGAITNPDGSKTNLIFDVHKYLDSDNSGTNSVCVTDNIDSAFAPLATWLRSNNRK<br>AILTETGGGNTPSCEQYLCQQIQYLNQNADVYMGYVGWAAGSFDPGYPLAETPVQNA<br>DGSWTDQPLVSLCLAR |
| | EHK47103 | 447 | MRPTSALAAALALASGALAGKIRYLGVAIPGIDFGCDIDGSCPTDTSSVPLLSYKGGDG<br>AGQMKHFAEDDGLNVFRISATWQFVTNNTVDGKLDELNWGSYNKVIDACLVTGAWC<br>MIDMHNFARYNGGIIGQGGVEDEVFVNLWVQIAKYYEKNDKIIFGLMNEPHDLDVNI<br>WAQTCQKVVTAIREAGATSQMILLPATNFASVGTYVDSGSAAALGAITNPDNSTDLLY<br>FDVHKYLDINNSGSHVECTTDNVQAFEDFATWLRDNKRQAIISETGASMDPSCMTDFC |

| | | | |
|---|---|---|---|
| | | | AQNKAISANSDVYIGFVGWGAGSFDTSYILTLTPLGEPGNYTDNKLMEQCILDQFTMD SKYAPTPTSISTATETGTSTPTPDPSSTSGNSPSSTTPSADKQTPSSSPKANPVNDPSSDTN NSSKDGNKDGKSAAPSGAKALSGSLLLTGAAFGYMMVAF |
| | EHK50177 | 418 | MNKPMGPLLLAATLMASGAIAQTQTVWGQCGGQGYSGPTNCASGSACSTLNPYYAQ CIPGATSFTTSTTSTKSPGSGSSTTSSASQPTGSGQTRFAGINIAGFDFGCTTDGTCVTSQI YPPLKNFGGTNNHPDGVGQMQHFVNDDKLNIFRLPVGWQYLVNNNLGGTLDSTAISN YDQLVQGCLATGAYCIVDIHNYARWNGAIIGQGGPTNAQFVSLWTQLATKYASQSKI WFGIMNEPHDVDINTWGTTVQAVVTAIRNAGATTQFISLPGTDYQSAGNFLTDGSSTA LSQVKNPDGSTTNLIFDLHKYLDSDNSGTHTECVTNNIATAFQPVATWLRQNKRQGILT ETGGGNTQSCIQDVCQQNQFLNQNSDVFLGYVGWGAGSFDSTYQLTLTPTQNGNTWT DTALAAACFSRA |

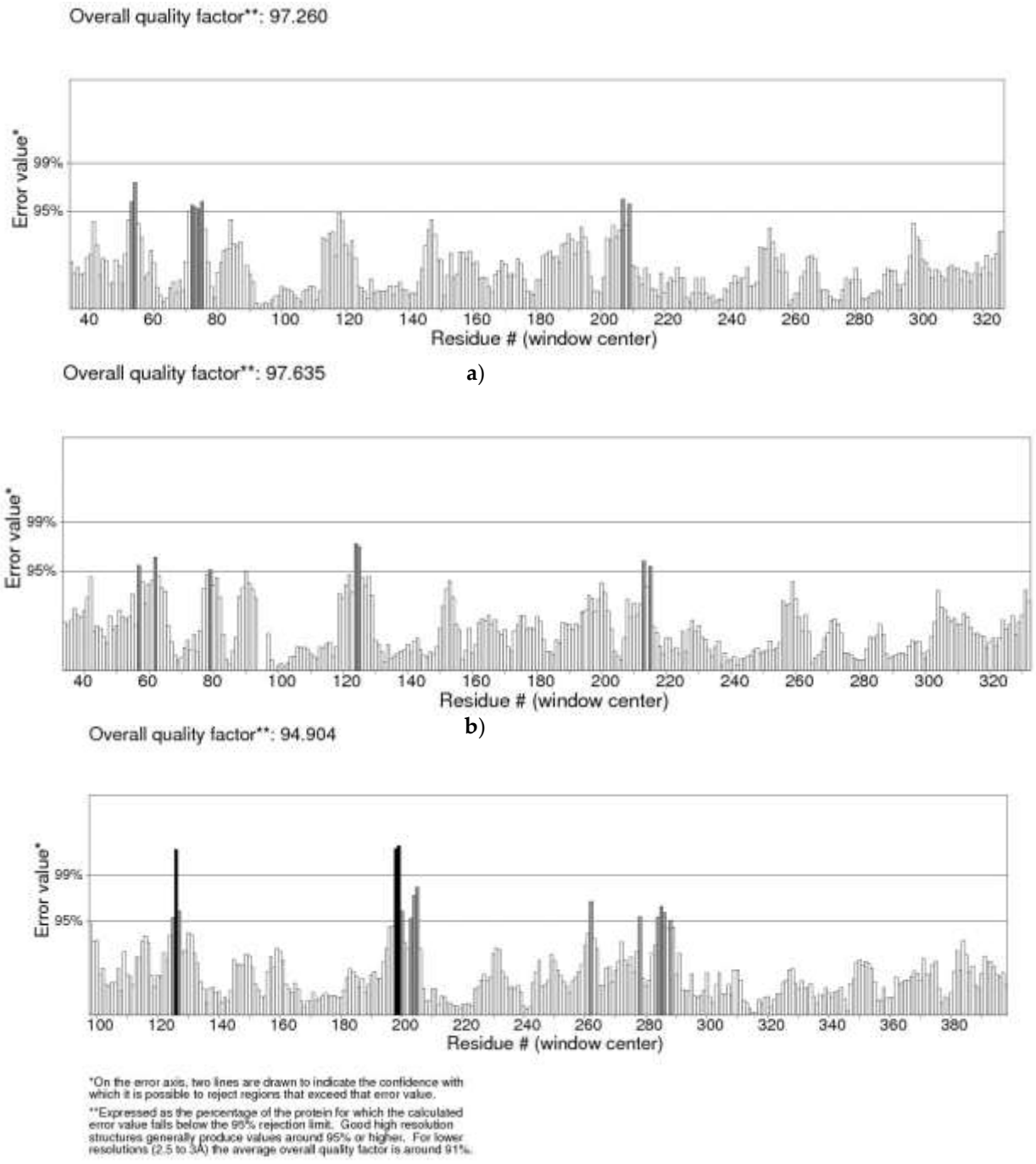



**Figure S1: ERRAT plots for cellulases a) CEJ55762, b) OOQ86698 (*Penicillium brasilianum*) and c) EHK50177 (*Trichoderma atroviride*) respectively.**

## References

1. Valentine J., Clifton-brown J., Hastings A., Robson P., Allison G. and Smith P., Food vs . fuel : the use of land for lignocellulosic 'next generation ' energy crops that minimize competition with primary food production, *Global Change Biology Bioenergy*, https://doi.org/10.1111/j.1757-1707.2011.01111.x, **4**, 1–19 **(2012)**

2. Bischof R.H., Ramoni J. and Seiboth B., Cellulases and beyond: The first 70 years of the enzyme producer Trichoderma reesei, *Microb. Cell Fact.*, https://doi.org/10.1186/s12934-016-0507-6 **(2016)**

3. Kuhad R.C., Gupta R. and Singh A., Microbial Cellulases and Their Industrial Applications, *Enzyme Res.*, https://doi.org/10. 4061/2011/280696, **2011**, 1–10 **(2011)**

4. Forsberg Z., Mackenzie A.K., Sorlie M., Rohr A.K., Helland R., Arvai A.S., Vaaje-Kolstad G. and Eijsink V.G.H., Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases, Proc. Natl. Acad. Sci., https://doi.org/10.1073/pnas.1402771111 **(2014)**

5. Suresh Chandra Kurup R., Snishamol C. and Nagendra Prabhu G., Cellulase Production by Native Bacteria Using Water Hyacinth as Substrate under Solid State Fermentation, *Malays. J. Microbiol.*, **1**, 25–29 **(2005)**

6. Pointing S.B., Qualitative methods for the determination of lignocellulolytic enzyme production by tropical fungi, *Fungal Divers.*, **2**, 17–33 **(1999)**

7. Ghose T.K., Measurement Of Cellulase Activities, *Pure Appl. Chem.*, **59**, 257–268 **(1987)**

8. Bailey M.J., Biely P. and Poutanen K., Interlaboratory testing of methods for assay of xylanase activity, *J. Biotechnol.*, https:// doi.org/10.1016/0168-1656(92)90074-J, **23**, 257–270 **(1992)**

9. Miller G.L., Blum R., Glennon W.E. and Bruton A.L., Measurement of Carboxymethylcellulase Activity, *Anal. Biochem.*, **2**, 127–132 **(1960)**

10. Resch M.G., Baker J.O. and Decker S.R., Low Solids Enzymatic Saccharification of Lignocellulosic Biomass **(2015)**

11. Tamura K., Dudley J., Nei M. and Kumar S., MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0, *Mol. Biol. Evol.*, **24**, 1596–1599 **(2007)**

12. Larkin M.A. et al, Clustal W and Clustal X version 2.0, *Bioinformatics*, **23**, 2947–2948 **(2007)**

13. Elisabeth G., Christine H., Alexandre G., Séverine D., Marc R.W., Ron D.A. and Amos B., Protein Identification and Analysis Tools on the ExPASy Server, In Walker J.M. eds., The Proteomics Protocols Handbook, Humana Press, https://doi.org/10.1385/15 92598900, 571–607 **(2005)**

14. Gill S.C. and Von Hippel P.H., Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem.*, **182**, 319–326 **(1989)**

15. Ikai A., Thermostability and aliphatic Index of Globular Proteins, *J. Biochem.*, **88**, 1895–1898 **(1980)**

16. Pradeep N.V., Anupama S., Vidyashree K.G. and Lakshmi P., In silico Characterization of Industrial Important Cellulases using Computational Tools, *Adv. Life Sci. Technol.*, **4**, 8–14 **(2012)**

17. Tamboli A.S., Rane N.R., Patil S.M., Biradar S.P., Pawar P.K. and Govindwar S.P., Physicochemical characterization, structural analysis and homology modeling of bacterial and fungal laccases using in silico methods, *Netw. Model Anal. Heal. Informatics Bioinforma.*, **4**, 1–12 **(2015)**

18. Guruprasad K., Reddy B.V.B. and Pandit M.W., Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence, *Protein Eng.*, **4**, 155–161 **(1990)**

19. Kyte J. and Doolittle R.F., A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.*, https://doi.org/ 10.1016/0022-2836(82)90515-0, **57**, 105–132 **(1982)**

20. Combet C., Blanchet C.G.C. and D.G., NPS@: Network Protein Sequence Analysis, *Trends Biochem. Sci.*, **25**, 147–150 **(2000)**

21. Guex N., Peitsch M.C. and Schwede T., Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective, *Electrophoresis*, https://doi.org/10.1002/elps.200900140, **30**, 162–173 **(2009)**

22. Colovos C. and Yeates T.O., Verification of protein structures: Patterns of nonbonded atomic interactions, *Protein Sci.*, https:// doi.org/10.1002/pro.5560020916, **2**, 1511–1519 **(1993)**

23. Sippl M.J., Recognition of Errors in Three-Dimensional Structures of Proteins, *Proteins Structure Funct. Genet.*, https:// doi.org/10.1002/prot.340170404, **17**, 355–362 **(1993)**

24. Schneidman-Duhovny D., Inbar Y., Nussinov R. and Wolfson H.J., PatchDock and SymmDock: Servers for rigid and symmetric docking, *Nucleic Acids Res.*, **33**, 363–367 **(2005)**

25. Schneidman-Duhovny D. et al, Taking geometry to its edge: Fast unbound rigid (and hinge-bent) docking, *Proteins Struct. Funct. Genet.*, **52**, 107–112 **(2003)**

26. Rogers S., Wells R. and Rechsteiner M., Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis, *Science*, **234**, 364–368 **(1986)**

27. Hu J., Arantes V. and Saddler J.N., The enhancement of enzymatic hydrolysis of lignocellulosic substrates by the addition of accessory enzymes such as xylanase: Is it an additive or synergistic effect?, *Biotechnol. Biofuels*, **4**, 1–13 **(2011)**

28. Gusakov A.V. and Sinitsyn A.P., Cellulases from Penicillium species for producing fuels from biomass, *Biofuels*, **3(4)**, 463–477 **(2012)**

29. Buxbaum E., Protein structure, In Fundamentals of Protein Structure and Function, Springer International Publishing Switzerland, https://doi.org/10.1007/978-0-387-68480-2, 378 **(2007)**.

(Received 15th April 2019, accepted 19th June 2019)