Biotechnology Reports 15 (2017) 55–62

Contents lists available at ScienceDirect

Biotechnology Reports

journal homepage: www.elsevier.com/locate/btre

# *In vitro* and *in silico* characterization of metagenomic soil-derived cellulases capable of hydrolyzing oil palm empty fruit bunch

Laura Marcela Palma Medina[a,1,2], Diana Catalina Ardila[a,1,3], María Mercedes Zambrano[b,d], Silvia Restrepo[c], Andrés Fernando González Barrios[a,*]

[a] Grupo de diseño de Productos y Procesos (GDPP), Department of Chemical Engineering, Universidad de los Andes, Carrera 1E N 18A-10, Bogotá, Colombia
[b] Molecular Genetics, Corporación Corpogen, Cra 5 N° 66A-34, Bogotá, Colombia
[c] Department of Biological Sciences, Universidad de los Andes, Carrera 1E N 18A-10, Bogotá, Colombia
[d] Gebix Center for Genomics and Bioinformatics of Extreme Environments, Cra 5 N° 66A-34, Bogotá, Colombia

ARTICLE INFO

*Keywords:*
Cellulases
Metagenomics
Oil palm empty fruit bunch
Fractal-like kinetic model
Molecular docking

ABSTRACT

Diversification of raw material for biofuel production is of interest to both academia and industry. One attractive substrate is a renewable lignocellulosic material such as oil palm (*Elaeis guineensis Jacq.*) empty fruit bunch (OPEFB), which is a byproduct of the palm oil industry. This study aimed to characterize cellulases active against this substrate. Cellulases with activity against OPEFB were identified from a metagenomic library obtained from DNA extracted from a high-Andean forest ecosystem. Our findings show that the highest cellulolytic activities were obtained at pH and temperature ranges of 4–10 and 30 °C–60 °C, respectively. Due to the heterogeneous character of the system, degradation profiles were fitted to a fractal-like kinetic model, evidencing transport mass transfer limitations. The sequence analysis of the metagenomic library inserts revealed three glycosyl hydrolase families. Finally, molecular docking simulations of the cellulases were carried out corroborating possible exoglucanase and β-glucosidase activity.

## 1. Introduction

In recent years there has been an increase in the use, trade, and production of biofuels, due to the need to replace fossil fuels with renewable energy sources. However, biofuel production has generated an ethical debate because the main raw materials can also be basic food crops (e.g., sugarcane, sugar beet and corn). In consequence, the search for alternative sources such as lignocellulosic material from industrial processes has increased [1–4]. One of the industries that currently generates higher amounts of suitable material for bioethanol production is the palm oil extraction. The palm oil industry generates lignocellulosic by-products from fresh fruit bunches, which in terms of mass contain 21% of palm oil, 27% of water and 52% of byproducts. A total of 44.2% of these byproducts constitutes empty fruit bunch (OPEFB) [5]. The utilization of OPEFB to obtain reducing sugars for bioethanol synthesis can, therefore, generate added value to the oil palm industry. The bioethanol produced from this alternative raw material is potentially competitive and has environmental advantages as it helps in reducing greenhouse gas emissions. Nevertheless, the industrialization of this process requires optimization at different stages, including the enzymatic hydrolysis by cellulases, which is a limiting step for the utilization of any cellulosic biomass [4].

The degradation of cellulose through enzymatic hydrolysis [6] is carried out by three kinds of cellulases: (1) endoglucanases, that can randomly hydrolyze internal glycoside linkages of the amorphous region of cellulose; (2) exoglucanases that progressively attack cellulose molecules at non-reducing ends of the chain, producing cellobiose molecules; and (3) β-glucosidases that hydrolyze cellobiose into glucose [7]. Although there are several sources of commercially available cellulases produced by microorganisms, their effectiveness depends on their affinity for the substrate [8]. It is expected that cellulases highly specific for OPEFB can produce fermentable sugars from this substrate in an optimized manner, making the process of saccharification easier and reducing costs. In consequence, the identification of cellulases capable of degrading OPEFB with high affinity and good reaction rates is a priority to optimize bioethanol production from this material.

* Corresponding author.
*E-mail address:* andgonza@uniandes.edu.co (A.F.G. Barrios).
[1] Both authors contributed equally to this work.
[2] Current Affiliation: Department of Medical Microbiology, University of Groningen. University Medical Center Groningen. Hanzeplein1, P.O. box 30001, 9700 RB Groningen, The Netherlands.
[3] Current Affiliation: Department of Bioengineering, The University of Pittsburgh, 302 Benedum Hall, Pittsburgh, PA 15219, USA.

http://dx.doi.org/10.1016/j.btre.2017.06.003
Received 2 August 2016; Received in revised form 17 May 2017; Accepted 6 June 2017
Available online 10 June 2017

Bioprospecting for microorganisms and their metabolic potential is increasingly used as a strategy to identify novel enzymes that may foster the biofuel industry [9,10]. However, the majority of microorganisms available in an environmental sample are unculturable, and therefore their study and assessment of their full biotechnological potential are difficult [11]. Metagenomics, a culture-independent strategy, has been used for discovering products from DNA isolated directly from the environment [11,12]. This approach has led to the discovery and characterization of a wide range of biocatalysts [11–13], which roused interest for the search of novel cellulases for biofuel production [14,15]. Previous metagenomics studies have reported as many as 105 new cellulases, 60 of which were obtained from soil samples [16]. This work reports the identification and characterization of cellulases from a metagenomic library of a high-Andean forest ecosystem, as part of the studies on metagenomics, microbial diversity and bioprospecting done by the GEBIX Center (*Colombian Center for Genomics and Bioinformatics of Extreme Environments*) in the Colombian National Natural Park "*Los Nevados*". The identified enzymes showed activity against OPEFB and are potentially useful in second-generation biofuel production.

## 2. Materials and methods

### 2.1. Metagenomic library

The metagenomic library, which consists of 18,432 clones, was constructed using DNA extracted from high-Andean forests soils from the National Natural Park "Los Nevados [17]. DNA was purified using the Ultra Clean Mega Soil DNA kit (MoBio) and fragments of approximately 30 kb were ligated to the pCC2FOS vector (Epicentre) and used to transform *Escherichia coli* EPI300™, following the manufacturer's indications (Epicentre). The identification of positive clones for cellulose degradation was done as published [17]. Briefly, clones capable of growing on minimum salt medium (MM; 0.2% $NaNO_3$, 0.1% $K_2HPO_4$, 0.05% $MgSO_4$, 0.05% KCL, 0.02% Peptone, 12.5 μg/ml chloramphenicol) with pretreated OPEFB as only source of carbon for 35 days at 30 °C and 200 rpm, were presumed to be cellulases carriers. Then, these colonies were enriched in Luria-Bertani (LB) medium for 2 days and transferred to a solid MM containing carboxymethyl cellulose. To detect cellulose hydrolysis, Congo red staining was used and the presence of a hydrolysis halo surrounding a colony was taken as a positive clone for cellulose degradation.

### 2.2. Oil palm empty fruit bunch pretreatment

OPEFB was shredded by grinding to a diameter of 5 mm in a hammer mill and pre-hydrolyzed by soaking in 1% (w/v) $H_2SO_4$ (100 mL for each 5 g of OPEFB) for 1 h, followed by autoclaving the acid-treated material at 121 °C for 15 min. The lignocellulosic material was washed with deionized water until the pH was close to 6.5, and then dried in an oven at 45 °C for at least 48 h [5]. Prior to each inoculation, the lignocellulosic material was autoclaved with the media.

### 2.3. Determination of cellulase activity

Cells harboring clones with cellulolytic activity [17], were grown in 2.5 L of MM [18], containing 12.5 μg/ml chloramphenicol, with pretreated OPEFB (1% w/v) as carbon source to the middle of the exponential growth phase ($OD_{600}$ of 0.4), collected by centrifugation for 30 min at 4500 rpm (2,000 × *g*), and resuspended in 5 mL of buffer (Tris-HCl 50 mM, NaCl 100 mM, EDTA 1 mM, 0.15% Triton X-100, pH 8). Cell membranes were disrupted using a Beadbeater (Biospec Products, Bartlesville, OK, USA). After centrifugation for 10 min at 13,000 rpm (17,000 × *g*), supernatants (crude extracts) were used for enzymatic assays.

Enzymatic reactions were performed in buffer (different buffers were used depending on the pH condition required for the reaction) containing OPEFB (2.5% w/v) and crude extract (1.25 mg/ml), and incubating at varying conditions for two hours with agitation (250 rpm). All assays were done in duplicate. After incubation, samples were centrifuged for 1 min at 13,000 rpm (17,000 × *g*) in order to eliminate OPEFB. The phenol-sulfuric acid assay [6] was used for quantification of sugars: 50 μL of sample were mixed with 30 μL of 5% phenol and 180 μL of 96% sulfuric acid, incubated for 5 min, and the absorbance was measured at 480 nm. The concentration of sugar was determined using a standard curve generated with eight glucose solutions of varying concentrations ranging between 10 and 400 μg/ml.

To determine the effect of metal ion addition, reactions were carried out in the presence of 10 mM each of $MgCl_2$, $CuSO_4$, $ZnSO_4$ and KCl. These reactions were performed in buffer McIlvaine [19] at pH 5.5 and 50 °C.

The pH effect on reaction efficiency was evaluated by performing hydrolysis experiments under different pH conditions using seven different buffers (KCl-HCl for pH 1 and 2.5, McIlvaine for pH 4, 5.5 and 7, Tris-HCl for pH 8.5 and Borax-NaOH for pH 10). The reactions were carried out at 50 °C in the presence or absence of metal ions. Temperature (10 °C to 70 °C) was analyzed at the pH that displayed the highest activity.

### 2.4. Cellulase kinetics evaluation

Each clone with evidence of cellulase activity was evaluated by obtaining a time profile of fermentable sugar concentration using different initial OPEFB substrate concentrations (5%, 6.25%, and 7.5%) under the specific reaction conditions of metal ions, temperature, and pH for each one. The crude extracts were added when reactions were started. Samples were collected every three minutes up to twenty-seven minutes. Reducing sugars were quantified using the phenol-sulfuric acid method [6]. The experiments were performed by quintuplicate. The results were adjusted to a semi-empirical fractal-like kinetic model (Eq. (1)) [20,21], where $[S]_0$ is the initial concentration of substrate, $k$ is a kinetic constant and $h$ is the fractal dimension. The kinetic constant represents the affinity and velocity of the reaction and the fractal dimension represents the influence of the transport phenomena on the reaction kinetics. The fractal dimension value represents the effect of the surrounding factors such as the average diffusion distance of the protein in the 3D space, the 2D diffusion over the cellulase surface and the adsorption rate. The parameter $k$ is mostly related to the efficiency and affinity of the proteins for the substrate and also the ability of the different kinds of cellulases to interact synergistically. To find the values of each constant, a regression was performed by the least squares method and corroborated by the open fitting curve toolbox (cftool) on Matlab™ (http://www.mathworks.com/products/curvefitting/).

$$P(t) = [S]_0(1 - \exp(-kt^{(1-h)})) \quad (1)$$

### 2.5. Fosmid DNA extraction and sequencing

Fosmid DNA was purified using PureLink™ Quick Plasmid Miniprep Kit (Invitrogen, Carlsbad), according to the manufacturer's protocol, and resuspended in a final volume of 75 μL. The DNA quality was verified by 2% agarose gel electrophoresis in 1× TAE buffer. The quantity and quality were analyzed by measuring absorbance 260 nm/280 nm in a Nanodrop™ ND-1000 (Thermo Scientific). The isolated DNA was sequenced by Ion Torrent (314 chip) (Life Technologies, Carlsbad) at the Huck Institutes of Life Sciences, Pennsylvania State University.

### 2.6. Determination of sequences coding for cellulases

The quality of the reads was checked with the FastQC package [22] and trimmed and filtered by quality using FASTX-ToolKit [23].