A novel high performance metagenome-derived alkali-thermostable endo-β-1,4-glucanase for lignocellulosic biomass hydrolysis in the harsh conditions

Shohreh Ariaenejad (shariaee@gmail.com)
Agricultural Biotechnology Research Institute of Iran
https://orcid.org/0000-0001-6385-5098

Atefeh Sheykhabdolahzadeh
Agricultural Biotechnology Research Institute of Iran

Morteza Maleki
Agricultural Biotechnology Research Institute of Iran

Kaveh Kavousi
University of Tehran

Mehdi Foroozandeh Shahraki
University of Tehran

Ghasem Hosseini Salekdeh
Agricultural Biotechnology Research Institute of Iran

Research

Keywords: endo-β-1,4-glucanase, metagenomics, enzymatic hydrolysis, rice straw, lignocellulosic biomass

Posted Date: February 18th, 2020

DOI: https://doi.org/10.21203/rs.2.23802/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published on October 19th, 2020. See the published version at https://doi.org/10.1186/s12896-020-00647-6.
Abstract

Background: Lignocellulosic biomass, is a great resource for the production of bio-energy and bio-based material since it is largely abundant, inexpensive and renewable. The requirement of new energy sources has led to a wide search for novel effective enzymes to improve the exploitation of lignocellulose, among which the importance of thermostable and halotolerant cellulase enzymes with high pH performance is significant.

Results: The primary aim of this study was to discover a novel alkali-thermostable endo-β-1,4-glucanase from the sheep rumen metagenome. Using a multi-step in-silico analysis, primary candidates with desired properties were found and subjected to cloning, expression, and purification followed by functional and structural characterization. The enzymes' kinetic parameters, including V max , Km, and specific activity, were calculated. The PersiCel4 demonstrated its optimum activity at pH 8.5 and a temperature of 85°C and was able to retain more than 70% of its activity after 150 hours of storage at 85°C. Furthermore, this enzyme was able to maintain its catalytic activity in the presence of different concentrations of NaCl, MgCl 2, CaCl 2, and MnCl 2. Our results showed that treatment with MnCl 2 could enhance the enzyme's activity by 89%. PersiCel4 was ultimately used for enzymatic hydrolysis of autoclave pretreated rice straw, the most abundant agricultural waste with rich cellulose content. In autoclave treated rice straw, enzymatic hydrolysis with the PersiCel4 increased the release of reducing sugar up to 260% after 72 hours in the harsh condition ( T= 85°C, pH = 8.5).

Conclusion: Considering the urgent demand for stable cellulases that are operational on extreme temperature and pH conditions and due to several proposed distinctive characteristics of PersiCel4, it can be used in the harsh condition for bioconversion of lignocellulosic biomass.

Background

Lignocellulose is the principal constituent of the biomass and is the most abundant bio-renewable organic resource on earth and therefore is of significant interest to governments, researchers, and industries. Lignocellulosic biomass is composed of three biopolymers, hemicellulose, cellulose and lignin which, depending on the type of the biomass, are intermeshed and organized into complex structures to varying degrees and different compositions[1].

Annually, processing of raw material gives rise to copious amounts of agricultural and industrial waste. From the quantity perspective, rice straw, wheat straw, corn straw, and sugarcane bagasse are the chief available agricultural wastes[2]. Among them, rice straw (RS) has the most abundance over the world, with approximately 600 to 900 million tons being produced yearly. However, only a small portion of this great resource is being used as animal feed, and the rest is squandered commonly by burning, which contaminates the air and causes various environmental perils.[3,4].

Naturally, complete hydrolysis of cellulose is achieved through the cooperative action of enzymes including cellobiohydrolase (EC 3.2.1.91), endo-β-1,4-glucanase (EC 3.2.1.4) and β-glucosidase (EC
3.2.1.21)[5]. Among the three types of cellulose-degrading enzymes, endo-β-1,4-glucanase is the crucial enzyme that first acts on the cellulose polymers’ amorphous sites and randomly cleaves the internal β-1,4-bonds and releases smaller fragments of varying random lengths[6]. Due to the sequence similarities of the catalytic domains of endo-β-1,4-glucanase enzymes were classified into glycoside hydrolase (GH) families of 5, 6, 7, 8, 9, 12, 44, 45, 48, 51 and 74[7,8].

Alongside with the vast involvement and importance in bio-ethanol production technologies, endo-β-1,4-glucanases are also applicable in various other industries such as pulp and paper, textile, food and feed processing, detergents[9].

The bioconversion process of lignocellulosic biomass consists of three steps, namely pretreatment of biomass, enzymatic hydrolysis of the polysaccharides, and fermentation of fermentable sugars. The densely packed structure of cellulose and hemicellulose together with lignin, which essentially serves as a protection for plants, highlights the importance of pretreatment strategies to solubilize and separate one or more of these components of biomass in order to make the remaining solid matter more easily accessible for the enzymatic hydrolysis[10,11]. Several pretreatment methods have been designed and used all of which attempt to accomplish main purposes including increasing the digestibility of solid biomass to augment the sugar yield during enzymatic hydrolysis step, evading the degradation of released sugars, and minimize the risk of emergence of inhibitors which could potentially interfere with subsequent steps[12]. It is reported that alkali pretreatment, results in more reducing sugar yield for agricultural waste in comparison to acid pretreatment[13].

Although, desired enzymatic attributes vastly differ for different applications, high catalytic activity, stability at extreme environmental conditions and tolerance to end-product inhibition and high concentrations of salts and ions are generally among the most favorable characteristics of industrial cellulases[14–16].

Generally, the enzymatic degradation of cellulose takes place at 40 °C to 50 °C at slow rates with incomplete hydrolysis, low sugar yield, and high probability of microbial contamination. Thermostable enzymes that are produced by thermophilic or thermotolerant microorganisms can address the limitations as mentioned earlier. Thus, thermostable cellulases with better catalytic activity at elevated temperatures offer several advantages including decreased hydrolysis time[17], reduced susceptibility to microbial contamination[18], easier recovery of volatile products such as ethanol[19], and a lower necessity for cooling after thermal pretreatment[20].

The pretreatment of biomass often takes place in the presence of acids or bases and at high temperatures. Nonetheless, the pH must be neutralized before the initiation of enzymatic hydrolysis, and the neutralization of these acids or bases generates salts[21]. The process of removing the formed salts before further downstream steps requires tons of water and energy; hence, halotolerant and halophile enzymes that are stable and operational in the presence of salts are in great demand[9,22].
The utilization of culture-independent approaches such as metagenomics permits access to almost all genetic information embedded within an environmental sample through DNA extraction followed by either cloning or sequencing[23,24]. Ample invaluable information about composition, function, organization, and hierarchy of microbial communities in various habitats has already been collected employing sequence-based metagenomics[25]. Any particular microbial habitat hosts a distinct microbial community with specific characteristics and capabilities aimed at accomplishing certain objectives within their environment. As an instance, the rumen microbiota possesses a vigorous hydrolyzing enzyme profile adapted to enhance the digestion and exploitation of lignocellulosic biomass, which dominates the ruminant diet[26,27]. Metagenomic screening provides the opportunity to explore the extensive biodiversity of nature and thus enabling the isolation of several novel enzymes from different environments.

Numerous studies introduced novel cellulases with unique attributes, making them suitable for particular applications[28]. Some of the endo-β-1,4-glucanases were directly isolated from microorganisms. For instance, a new thermostable and halotolerant endoglucanase was produced from Botrytis ricini URM 5627[29]. In another research, a thermo-halotolerant and alkali Stable GH6 endoglucanase from T. halotolerans were discovered that displays high catalytic activity in alkaline pH conditions and the presence of NaCl. Furthermore, pretreatment with high concentrations of NaCl improved the enzyme’s activity[30]. Many studies have focused on the isolation of novel endo-β-1,4-glucanases from metagenomic sources. For example, a novel halo-ionic liquids tolerant thermoacidophilic endo-β-1,4-glucanase was isolated from saline-alkaline lake soil microbial metagenomic DNA which could be applied in the hydrolysis of acid, and ionic liquid (IL) pretreated biomass[31]. Moreover, Song et al. isolated and characterized a new endo-β-1,4-glucanase from black-goat rumen metagenome[32] and Narra et al. mined a novel GH12 endo-glucanase which was applied in the enzymatic degradation of alkali-treated and delignified RS[33]. In contrast to alkaline or alkali-tolerant cellulases, which have found use in laundry detergents around the world, only a few halophilic or halotolerant cellulases have been reported[34–36].

In the current study, the sheep rumen metagenomic data was explored and screened, and several target candidates were found, one of which, PersiCel4, was cloned, expressed, purified, and characterized. PersiCel4, a novel halotolerant endo-β-1,4-glucanase, demonstrated high thermal and alkaline stability. This enzyme was then utilized to enhance the hydrolysis of pretreated RS and substantially increased the reducing sugar yield.

**Results And Discussion**

As the necessity for new alternative energy sources grows and given the amplitude of agricultural waste which is being produced and squandered every year, enzymatic hydrolysis of these feedstocks gains increasing attention and importance. Bio-conversion-based industries demand hydrolytic enzymes with high catalytic efficiency and tolerance to extreme conditions such as pH, temperature, and presence of different salts. Endo-β-1,4-glucanase is one of the major enzymes involved in the degradation process of
cellulose. Moreover, ruminants’ rumen enzymatic diversity is known to be rich and capable of degrading recalcitrant lignocellulosic biomass, the primary constituent of most ruminants’ diet. Hence, sheep rumen metagenome was screened with the focus to find halotolerant endo-β-1,4-glucanases which are particularly stable at alkaline pH and elevated temperatures. This enzyme could improve the sugar yield of the enzymatic hydrolysis phase.

**Prediction of thermostable endo-glucanase sequences**

One of the contigs assembled from the sheep rumen metagenome, named PersiCel4, passed all filters applied for identifying thermostable endo-glucanase. The PersiCel4 is an enzyme with amino acid length of 339. The blast results show that it is most similar to an endo-glucanase enzyme from GH5 family with Endohydrolysis of (1–4)-beta-D-glucosidic linkages in cellulose, lichenin and cereal beta-D-glucans catalytic activity and UniProt accession number B7UAM4 with E-Value of 3e-112 (Table 1).

Based on the CDD results, the PersiCel4 best matches the Pssm-ID 333879 with Bit Score 212.21 and E-value 2.57e-67 (Fig. 1A).

Phyre2 suggests that the most similar structure to the PersiCel4 with 100% confidence and 52% identity belongs to an endo-1,4-beta-glucanase from Bacillus subtilis 168 with 3PZV PDB code (Fig. 1B).

**Expression and purification for recombinant endo-β-1,4-glucanase**

From the metagenomic DNA of sheep rumen with degenerate primers and over-expressed under the control of the T7 promoter of pET-28a vector in E. coli BL21 (DE3), the sequence of PersiCel4 fragment was amplified. The purification step was checked by SDS-PAGE and using Ni-NTA Fast Start Kit for purified the N-terminus His-tagged protein. The single PersiCel4 band was visible with a molecular weight corresponding to the calculated 38 kDa (Fig. 1C).

**Properties of purified PersiCel4**

The kinetic parameters (Kₘ and Vₘₐₓ), of PersiCel4, was evaluated. The enzyme obeys based on a Lineweaver-Burk map in Michaelis-Menten kinetics concerning CMC as substrate, Kₘ and Vₘₐₓ values of PersiCel4 were 0.58 µmol ml⁻¹, 1.95 µmol min⁻¹mg⁻¹. The calculated specific activities for purified PersiCel4 was 22.425 Umg⁻¹.

Optimum pH was determined in room temperature. Figure 2A shows that the optimal pH value for PersiCel4 was 8.5, and it preserved more than 60% of its maximum activity in the pH range between 5 up to 10 (Fig. 2A). By fixing the pH for enzyme, the maximum activity was observed at 85 °C in the range of temperature between 30–90 °C, and the activity of the enzyme was decreased just 8% of its maximum value at 95 °C (Fig. 2B). However, PersiCel4 was also highly stable when stored for prolonged periods.
The PersiCel4 retained 70% of its relative activity after being stored for 168 hours at 85 °. The enzyme activity was diminished drastically after 288 hours of storage (Fig. 2C).

Because of slow pace and incomplete hydrolysis of cellulose in moderate temperatures, the susceptibility of the microbial contamination on one hand and the easier recovery of ethanol at high temperatures and decreased importance of cooling after thermal pretreatment, on the other hand, thermostable enzymes find great importance and value[17,20]. The competence of cellulases to operate in extreme pH conditions is another desirable attribute. Alkaline pretreatments on agricultural wastes such as RS have shown better sugar yield results compared to acid pretreatments[13] and alkali-stable cellulases have been reported to be suitable potential enzymes for lignocellulosic saccharification at such conditions[37]. As presented in Table 2, PersiCel4 depicted stability and functionality at more extreme pH and thermal conditions in comparison to some other known endo-\(\beta\)-1,4-glucanase representatives.
Table 2
The property comparison of PersiCel4 with some known endo-β-1,4-glucanase representatives.

| Name         | Source                                             | Optimum pH | Optimum temperature | Proposed Application                                      | References                  |
|--------------|----------------------------------------------------|------------|----------------------|----------------------------------------------------------|-----------------------------|
| nmGH45       | saline-alkaline lake soil microbial metagenomic DNA | 4.5        | 60–70                | lignocellulosic biomass hydrolysis                       | (Zhao et al., 2018)         |
| CelG5        | Phialophora sp. G5                                | 4.5        | 55–60                | brewing and feed industries                             | (Zhao et al., 2012)         |
| KG35         | metagenomic library of the black-goat rumen       | 6...7      | 30–50                | ND                                                       | (Song et al., 2017)         |
| Cel5G        | metagenomic library from soil                     | 4.8        | 50                   | industrial use                                          | (Liu et al., 2011)          |
| CS10         | metagenome gutmicroflora of Hermetia illucens     | 7          | 50                   | industrial use                                          | (Lee et al., 2014)          |
| cel5A        | mangrove soil                                     | 6.5–7.5    | 50                   | ND                                                       | (Gao et al., 2010)          |
| Cel5H        | Dictyoglomus thermophilum                         | 5          | 50–85                | ND                                                       | (Shi et al., 2013)          |
| CelRH5       | rhizosphere metagenomic library                   | 6.5        | 40                   | ND                                                       | (Wierzbicka-Woś et al., 2019)|
| Cel5R        | soil metagenome                                   | 5-6.5      | 58                   | ND                                                       | (Garg et al., 2016)         |
| cel7482, cel3623 and cel36 | Metagenomic sequences | 5.5        | 60–70                | ND                                                       | (Yang et al., 2016)         |
| Cel6H-f481   | compost metagenomic library                       | 5.5        | 50                   | ND                                                       | (Lee et al., 2018)          |
| PersiCel4    | Sheep rumen metagenome                            | 8.5        | 85                   | lignocellulosic biomass hydrolysis for bio-ethanol production | This study                  |

PersiCel4’s prolonged storage stability in addition to its optimal activity at 85 °C and pH 8.5, which are considerably higher than other endo-glucanase representatives, clearly distinguish this enzyme from many other reported endo-glucanases for industrial waste purposes.

Effects of salts on enzyme activity
As described before, since several pretreatment methods take place at either acidic or alkaline conditions and due to formation of various salts at the pH neutralization step, an important challenge is the removal of such salts that consumes an abundance of water and energy[21]. Given this fact, cellulolytic-enzymes that can tolerate the presence of different salts without reduction in their catalytic ability find attraction[9,22].

The PersiCel4 activity was investigated in the presence of different concentrations of NaCl (1 mM, 2.5 mM, 5 mM, 7.5 mM, 10 mM), MgCl$_2$, CaCl$_2$ and MnCl$_2$ at 1, 5 and 10 mM in comparison with control in standard conditions (without any salt). The effect of different NaCl concentrations on the activity of the PersiCel4 was measured and presented the enzymatic activity of 103.78, 99.66, 89.97, 82.36, and 82.77%, respectively for different NaCl concentration.

The enzyme activity demonstrated maintenance in different concentrations of NaCl. As it is shown in Table 3, besides the depicted tolerance to presence of NaCl, the presence of MgCl$_2$, CaCl$_2$, and MnCl$_2$ resulted in a significant increase in PersiCel4’s activity. This feature is yet to be added to other mentioned notable characteristics of PersiCel4 to nominate it for biofuel industries.
Table 3
PersiCel4’s relative activity (%) when treated with different concentrations of salts compared to control condition without the presence of salt.

| Concentration (mM) | Remaining enzyme activity % |
|--------------------|-----------------------------|
| control none       | 100                         |
| NaCl 1             | 103.78                      |
| 2.5                | 99.66                       |
| 5                  | 89.97                       |
| 7.5                | 82.36                       |
| 10                 | 82.77                       |
| MgCl2 1            | 122.63                      |
| 5                  | 91.59                       |
| 10                 | 107.08                      |
| CaCl2 1            | 122.68                      |
| 5                  | 111.68                      |
| 10                 | 103.54                      |
| MnCl2 1            | 167.37                      |
| 5                  | 176.89                      |
| 10                 | 189.58                      |

Effect of different pretreatment and enzymatic hydrolysis of rice straw

Lignocellulosic materials constitute economically renewable raw materials and are available in large quantities[38]. Several researchers have addressed the use of agricultural waste in the feasibility, sustainability, and economics of biofuels production[39–43].

One of the most abundant residues of lignocellulose in the world is RS[2,44,45]. Due to the complex structure of lignin and hemicelluloses, the conversion of RS to reducing sugar is a very complicated process[46]. Generally, by pretreatment of fibrous biomass through mechanical, physical, thermal, chemical or enzymatic pathways, it results in the conversion of complex cellulose and hemicellulose into reducing sugars. In the pretreatment process, cellulose and hemicellulose can be removed, and cellulose crystallization reduced, and porosity of the materials increased use recent reference. An effective
pretreatment method is the one which maximizes sugar productivity during the enzymatic hydrolysis step while preventing the degradation or loss of carbohydrates[12]. The effects of different preparations before enzymatic hydrolysis produced the highest sugar production.

Figure 3 shows the percentage of reduced sugar in each pretreatment obtained from untreated and different treated RS. The highest yield of reducing sugar observed when the RS (2%w/v) in sodium phosphate buffer (50 mM, pH 7) had done with an incubator at 100˚C, 30 min by thermal pretreatment.

In this study, it observes that autoclave pretreatment RS improves enzymatic hydrolysis resulting in a higher yield of reducing sugar than untreated RS. This may be due to the deficiency caused by the swelling of RS particles and an increase in the inner surface, which in turn has led to access of more enzymes to the internal structure, and thereby, increasing sugar yields[47,48].

In the next step, the enzymatic hydrolysis of thermal treated RS was evaluated using screening of the reducing sugar concentrations against time, at 85˚C (pH 8.5) (Fig. 4).

The results showed that PersiCel4 significantly enhanced the saccharification yield in 96 hours, specifically after 72 hours, which confirmed the boosting effect PersiCel4. In autoclave treated RS, the comparison between enzymatic hydrolysis with the PersiCel4 and without it, revealed a 260% increase in the released reducing sugar after 72 hours. In this study, it observes that the enzymatic hydrolysis of thermal pretreated RS was the most efficient method of converting the RS into reducing sugars. This study highlights the competence of the combined implementation of metagenomic screening and in-silico analysis in order to efficiently discover and isolate particular enzymes of interest. As described, PersiCel4 depicted various characteristics that are essential for enhanced hydrolysis of cellulose.

Conclusions

Waste management and energy crisis are among the most important issues that the world faces today. Such concerns have inspired the extensive search for capable novel enzymes to improve the recalcitrant lignocellulosic biomass conversion, such as agricultural wastes.

In this study, the sheep rumen metagenome was explored using in-silico analysis to find a novel alkali-thermostable halotolerant endo-β-1,4-glucanase. Ultimately, PersiCel4 was identified and isolated and characterized by multiple experiments, all of which verified the computationally predicted properties. PersiCel4 was successfully utilized for the hydrolysis of autoclave pretreated RS in extreme temperature and pH, which is the most plentiful agricultural waste. Regarding several proposed distinctive characteristics of PersiCel4, it can potentially find commercial applications the harsh conditions of the industrial bioconversion process.

Methods

Computational prediction of thermostable endo-β-1,4-glucanase enzyme sequences
The primary and preprocessing operations, including quality control and short read assembly, carried out on sheep rumen metagenomics data using FastQC and MEGAHIT assembler respectively. After that, the assembled contigs were explored for putative thermostable endo-glucanase sequences by the stages described below. In the first step, MetaGeneMark was employed for predicting potential microbial genes. The predicted endo-glucanase enzymes were separated from assembled contigs for next step analysis. Six thermostable endo-glucanase sequences from different GH families were identified from previous works and their sequences were downloaded from UniProt (Table 1). Using standalone NCBI BLAST, the mined enzymes from metagenome were aligned against the list of enzymes mined from literature. The most similar predicted metagenomics endo-glucanases to the known enzymes were determined based on pairwise blast with appropriate E-value.

Among predicted genes, those confirmed as endo-glucanases by NCBI Conserved Domains Database (CDD)[49] were selected for next step. The sequences of the filtered enzymes were assigned to the Phyre2 server[50] to predict their tertiary structures. Finally, one of the sequences that passed all filters, named PersiCel4, was selected for experimental assays.

### Table 1

| UniProt Accession | GH Family | EC     | Organism                          | length |
|-------------------|-----------|--------|-----------------------------------|--------|
| P0C2S4            | 9         | 3.2.1.4| Clostridium thermocellum (Ruminiclostridium thermocellum) | 625    |
| Q60033            | 12        | 3.2.1.4| Thermotoga maritima               | 274    |
| B7UAM4            | 5         | 3.2.1.4| Bacillus subtilis                 | 499    |
| (PersiCel4 e-value = 3e-112) | |       |                                   |        |
| P54583            | 5         | 3.2.1.4| Acidothermus cellulolyticus       | 562    |
| P96492            | 12        | 3.2.1.4| Thermotoga neapolitana            | 274    |
| G2QCS4.1          | 7         | 3.2.1.4| Thermotheleomyces thermophilus     | 464    |

**Cloning, expression, and purification of the endo-β-1,4-glucanase gene**

To acquire endo-β-1,4-glucanase gene, metagenome DNA template from sheep rumen. The endo-β-1,4-glucanase gene encoding sequence was amplified by polymerase chain reaction (PCR) using PersiCel4 F(5'- TAATAGGCTAGCATGAAAAGAATTTTGATTTTGGC – 3') and PersiCel4 R(5'- TGATAG GTCGAC TCACTCTATGTCTCCGCA – 3') primers which included NdeI and NotI restriction sites, respectively. PCR product was purified, digested with Nhel and Sall and ligated into pET-28a, resulting recombinant plasmid pet28[51,52].

Page 11/23
The resulting PCR products were detected on agarose gel 1.5% (w/v) and for purified using the gel extraction kit (BioRon, Germany). Purified DNA fragments were cloned and digested into the pET28a. The resulting plasmids were then transformed into the E. coli BL21 (DE3) and correct insertion was confirmed by sequencing. In the Luria-Bertoni (LB) medium, the recombinant strain pET28an was cultivated at the temperature of 37 °C. Adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM for 20 h at 20 °C, expression of enzymes were induced. By utilizing Ni-NTA Fast Start Kit (Qiagen, Hilden, Germany), N-terminal Histidine-tagged recombinant protein was purified and evaluated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

The candidate enzyme produced from successful cloning, expression, and purification were named PersiCel4. This endo-β-1,4-glucanase was subjected to further biophysical experiments. Nucleotide sequence of PersiCel4 was submitted to Genbank Database with accession number Mn016942. Protein concentrations were determined through the Bradford method using the bovine serum albumin as the standard. By measuring the optical density (OD) of chromatography eluent at 280 nm, the protein concentration was estimated.

**Spectroscopy studies**

Synergy™ HTX Multi-Mode Microplate Reader was used for UV–vis spectrophotometry. For matured protein concentration using the Bradford method, and bovine serum albumin (BSA) as the protein standard was used. The CMC substrate and reducing sugars amount was used for enzyme activity measurement by DNS method at 540 nm.

**Measurement of endo-β-1,4-glucanase activity and kinetic parameters**

The DNS method was used for determination of cellulase activity. After the occurrence of enzymatic hydrolysis of cellulose, the resulting reducing sugars can reduce the nitro of 3,5- dinitrosalicylic acid (DNS) into amino, forming a reddish-brown color due to the presence of amino compounds.[53] The intensity of brown color and the quantity of reducing sugar present a positive correlation in a particular concentration range which is diagnosable through the spectrophotometric method.

The cellulose substrate carboxymethyl cellulose (CMC) was used in the cellulose assay[54]. The intensity of brown color and the quantity of reducing sugar present a positive correlation in a particular concentration range which is diagnosable through the spectrophotometric method. 20 µL enzyme solution was added to 60 µL of 0.5% solubilized CMC in a test tube. Then 1.5 mL, 3,5-dinitrosalicylic acid (DNS) reagent was added and incubated in the water bath at 80˚C for 5 min. After that the optical absorbance was determined at 540 nm.

In the assay conditions described by Kamble RD, et al, the amount of enzyme that can 1 µmol of reduced sugars released per minute per millilitre is a unit of cellulose activity[55]. Ultimately, enzymes were incubated with CMC concentrations ranging from 0 to 10 mg/ml in 10 mM phosphate buffer at pH 8.
Using Michaelis-Menten plots at room temperature, $K_m$ and $V_{max}$ and Michaelis constant were determined.

**Determination of optimum pH and temperature and storage**

To determine the optimum pH of enzyme activity, 10 mM phosphate buffer was prepared at different pH (4 to 11), and after added enzyme was incubated with a substrate for 20 min at room temperature. The DNS was used to measure activity\[53\]. To determine the optimum temperature of enzyme activity, enzyme solution in 10 mM phosphate buffer (pH 8.5) with substrate was incubated in the different temperature (30–95 °C) for 20 min and DNS was used to measure activity.

For reporting PersiCel4 activity, relative activity was considered as the percentage of highest activity. The methods for determining storage stability was used by incubation of the purified enzymes for 12 days in kit elution buffer at 85 °C and the enzymes’ activities were examined in 24 h time intervals.

**Effects of salts on enzyme activity**

The PersiCel4 was pre-incubated at various concentrations of NaCl (1, 2.5, 5, 7.5, 10 mM) at 40°C for 30 min without the substrate and the activity was assayed according to the standard conditions. The activity of the enzyme not reacted with any NaCl concentrations was assessed as the control (%100).

To evaluate the effect of ionic liquid on enzyme activity the enzyme was pre-incubated with 1, 5 and 10 mM of MgCl$_2$, CaCl$_2$ and MnCl$_2$ at 40 °C for 30 min before adding the substrate. After pre-incubation, remaining activity was measured by the standard method described above and the effect of different ionic liquid concentration was tested. The enzyme activity without any chemicals was considered to be 100% activity and compared with the other values.

**Rice straw hydrolysate (RSH)**

Rice straw (RS) was collected from a local farm and washed with water and then dried at 70 °C in a hot air oven. By mixer grinder, size of dried RS has been reduced. The RS with suitable particle size (2–3 mm) pretreated by five different methods (alkaline, thermal, autoclave, microwave).

Alkaline pretreatments

In the RS (2%w/v), the pH was set at 12 by the addition of 2M NaOH, then mixed with magnetic stirrer for 30 min\[56\].

Autoclave pretreatments

The RS (2%w/v) in sodium phosphate buffer (50 mM, pH 7) have been done with autoclave at 121°C, 1.5 atm for 15 min\[57\].

The microwave pretreatments
About 0.2 g of RS was suspended in 10 mL of sodium phosphate buffer (50 mM, pH 7) at 30 min in microwave. For microwave pretreatment 700 W; 170 °C; 15 min were chosen[58].

**Thermal pretreatments**

For thermal pretreatment, the RS (2%w/v) in sodium phosphate buffer (50 mM, pH 7) have been done with an incubator at 100°C, 30 min[59].

The most effective pretreatment was found for this study by reducing sugar measurement. After pretreatment, the substrates are recovered by distillation and centrifugation several times. The pre-treated substrates were dried at 50 °C overnight and stored in a cool and dry place until further analyses.

The pre-treated RS was suspended in 50 mM sodium phosphate buffer (pH 8.5) and supplemented with enzymes. The substrates were saccharified in an incubator shaker at 120 rpm at 85 °C for 96 h and the released reducing sugars were determined using DNS method in the supernatant after the centrifugation.

**Abbreviations**

GH: glycoside hydrolase; IL: ionic liquid; CDD: conserved domains database; PCR: polymerase chain reaction; IPTG: isopropyl-β-D-thiogalactopyranoside; SDS-PAGE: dodecyl sulfate-polyacrylamide gel electrophoresis; BSA: bovine serum albumin; OD: optical density; DNS: nitro of 3,5- dinitrosalicylic acid; CMC: carboxymethyl cellulose; RS: rice straw; RSH: rice straw hydrolysate.

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Availability of data and materials**

We have read our policy on data and material release and give details of the accession numbers of any protein or nucleotide sequence data, below. We provide this information within the main text of our manuscript. Sequence data of this study have been deposited in the GenBank database with the accession codes “Submission No. Mn016942.

**Competing interests**

The authors declare that they have no competing interests.
Funding

This research was supported by grants from Agricultural Biotechnology Research Institute of Iran (ABRII).

Authors’ Contributions

SA performed the 3D-structure prediction, biochemical characterization, and also analyzed and interpreted the data. AS and MM performed the gene cloning, expression, and purification. KK and MFS contributed to the computational and bioinformatics data analysis. GHS supervised the study and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgement

We give thanks to the supports from members of the Institute of Biochemistry and Biophysics (IBB), University of Tehran who contributed to this study.

Author details

1 Department of Systems and Synthetic Biology, Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research Education and Extension Organization (AREO), Karaj, Iran 2 Laboratory of Complex Biological Systems and Bioinformatics (CBB), Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran 3 Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia.

References

1. Isikgor FH, Becer CR. Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. Polym Chem. 2015;
2. Kim S, Dale BE. Global potential bioethanol production from wasted crops and crop residues. Biomass and Bioenergy. 2004;
3. Saini JK, Saini R, Tewari L. Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. 3 Biotech. 2015.
4. Vlasenko EY, Ding H, Labavitch JM, Shoemaker SP. Enzymatic hydrolysis of pretreated rice straw. Bioresour Technol. 1997;
5. Kumar R, Singh S, Singh O V. Bioconversion of lignocellulosic biomass: Biochemical and molecular perspectives. J. Ind. Microbiol. Biotechnol. 2008.
6. Sweeney MD, Xu F. Biomass converting enzymes as industrial biocatalysts for fuels and chemicals: Recent developments. Catalysts. 2012.
7. Cantarel B, Coutinho P, Rancurel C, Bernard T, Lombard V, Henrissat B. Carbohydrate Active Enzymes database. Carbohydrate-Active Enzym. database an Expert Resour. Glycogenomics. Nucleic Acids
8. Henrissat B. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J. 1991;
9. Kuhad RC, Gupta R, Singh A. Microbial cellulases and their industrial applications. Enzyme Res. 2011.
10. Demirbaş A. Bioethanol from cellulosic materials: A renewable motor fuel from biomass. Energy Sources. 2005;
11. Leonowicz A, Matuszewska A, Luterek J, Ziegenhagen D, Wojtaś-Wasilewska M, Cho NS, et al. Biodegradation of lignin by white rot fungi. Fungal Genet. Biol. 1999.
12. Brodeur G, Yau E, Badal K, Collier J, Ramachandran KB, Ramakrishnan S. Chemical and physicochemical pretreatment of lignocellulosic biomass: A review. Enzyme Res. 2011.
13. Sukumaran RK, Singhania RR, Mathew GM, Pandey A. Cellulase production using biomass feed stock and its application in lignocellulose saccharification for bio-ethanol production. Renew Energy. 2009;
14. Singh A, Adsul M, Vaishnav N, Mathur A, Singhania RR. Improved cellulase production by Penicillium janthinellum mutant. Indian J Exp Biol. 2017;
15. Vaishnav N, Singh A, Adsul M, Dixit P, Sandhu SK, Mathur A, et al. Penicillium: The next emerging champion for cellulase production. Bioresour Technol Reports. 2018;
16. Di Donato P, Finore I, Poli A, Nicolaus B, Lama L. The production of second generation bioethanol: The biotechnology potential of thermophilic bacteria. J Clean Prod. 2019;
17. Viikari L, Terhi A. Thermostable Enzymes in Lignocellulose Hydrolysis. 2007;121–2.
18. Abdel-Banat BMA, Hoshida H, Ano A, Nonklang S, Akada R. High-temperature fermentation: How can processes for ethanol production at high temperatures become superior to the traditional process using mesophilic yeast? Appl. Microbiol. Biotechnol. 2010.
19. Taylor MP, Eley KL, Martin S, Tuffin MI, Burton SG, Cowan DA. Thermophilic ethanologenesis: future prospects for second-generation bioethanol production. Trends Biotechnol. 2009.
20. Turner P, Mamo G, Karlsson EN. Potential and utilization of thermophiles and thermostable enzymes in biorefining. Microb Cell Fact. 2007;
21. Klinke HB, Thomsen AB, Ahring BK. Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. Appl. Microbiol. Biotechnol. 2004.
22. Xing M-N, Zhang X-Z, Huang H. Application of metagenomic techniques in mining enzymes from microbial communities for biofuel synthesis. Biotechnol Adv. 2012;30:920–9.
23. Banik JJ, Brady SF. Recent application of metagenomic approaches toward the discovery of antimicrobials and other bioactive small molecules. Curr. Opin. Microbiol. 2010.
24. Handelsman J. Metagenomics: Application of Genomics to Uncultured Microorganisms. Microbiol Mol Biol Rev. 2005;
25. Edwards JL, Smith DL, Connolly J, McDonald JE, Cox MJ, Joint I, et al. Identification of carbohydrate metabolism genes in the metagenome of a marine biofilm community shown to be dominated by Gammaproteobacteria and Bacteroidetes. Genes (Basel). 2010;

26. Gharechahi J, Salekdeh GH. A metagenomic analysis of the camel rumen's microbiome identifies the major microbes responsible for lignocellulose degradation and fermentation. Biotechnol Biofuels. 2018;

27. Stewart RD, Auffret MD, Warr A, Walker AW, Roehe R, Watson M. Compendium of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and enzyme discovery. Nat Biotechnol. 2019;

28. Méndez-Líter JA, Gil-Muñoz J, Nieto-Domínguez M, Barriuso J, De Eugenio LI, Martínez MJ. A novel, highly efficient β-glucosidase with a cellulose-binding domain: Characterization and properties of native and recombinant proteins. Biotechnol Biofuels. 2017;10:1–15.

29. Silva TP, de Albuquerque FS, dos Santos CW V, Franco M, Caetano LC, Pereira HJV. Production, purification, characterization and application of a new halotolerant and thermostable endoglucanase of Botrytis ricini URM 5627. Bioresour Technol. 2018;

30. Yin YR, Zhang F, Hu QW, Xian WD, Hozzein WN, Zhou EM, et al. Heterologous expression and characterization of a novel halotolerant, thermostable, and alkali-stable GH6 endoglucanase from Thermobifida halotolerans. Biotechnol Lett. 2015;

31. Zhao J, Guo C, Zhang L, Tian C. Biochemical and functional characterization of a novel thermoacidophilic, heat and halo-ionic liquids tolerant endo-β-1,4-glucanase from saline-alkaline lake soil microbial metagenomic DNA. Int J Biol Macromol. 2018;

32. Song YH, Lee KT, Baek JY, Kim MJ, Kwon MR, Kim YJ, et al. Isolation and characterization of a novel endo-β-1,4-glucanase from a metagenomic library of the black-goat rumen. Brazilian J Microbiol. 2017;

33. Narra M, Dixit G, Divecha J, Kumar K, Madamwar D, Shah AR. Production, purification and characterization of a novel GH 12 family endoglucanase from aspergillus terreus and its application in enzymatic degradation of delignified rice straw. Int Biodeterior Biodegrad. 2014;

34. Gao Z, Ruan L, Chen X, Zhang Y, Xu X. A novel salt-tolerant endo-β-1,4-glucanase Cel5A in Vibrio sp. G21 isolated from mangrove soil. Appl Microbiol Biotechnol. 2010;

35. Hirasawa K, Uchimura K, Kashiwa M, Grant WD, Ito S, Kobayashi T, et al. Salt-activated endoglucanase of a strain of alkaliphilic Bacillus agaradhaerens. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol. 2006;

36. Voget S, Steele HL, Streit WR. Characterization of a metagenome-derived halotolerant cellulase. J Biotechnol. 2006;

37. Yu HY, Li X. Alkali-stable cellulase from a halophilic isolate, Gracilibacillus sp. SK1 and its application in lignocellulosic saccharification for ethanol production. Biomass and Bioenergy. 2015;

38. Millati R, Niklasson C, Taherzadeh MJ. Effect of pH, time and temperature of overliming on detoxification of dilute-acid hydrolyzates for fermentation by Saccharomyces cerevisiae. Process
39. Balat M, Balat H. Recent trends in global production and utilization of bio-ethanol fuel. Appl. Energy. 2009.

40. Demirbas MF, Balat M, Balat H. Potential contribution of biomass to the sustainable energy development. Energy Convers Manag. 2009;

41. Pimentel D, Patzek TW. Ethanol production: Energy and economic issues related to U.S. and Brazilian sugarcane. Biofuels, Sol Wind as Ren Energy Syst Benefits Risks. 2008.

42. Patzek TW, Croft GD. A global coal production forecast with multi-Hubbert cycle analysis. Energy. 2010;

43. Zhang LX, Ulgiati S, Yang ZF, Chen B. Energetics analysis of lignocellulosic materials and three wetland fish farming systems in Nansi Lake area, China. J Environ Manage. 2011;

44. De Faveri D, Torre P, Perego P, Converti A. Statistical investigation on the effects of starting xylose concentration and oxygen mass flow rate on xylitol production from rice straw hydrolyzate by response surface methodology. J Food Eng. 2004;

45. Karimi K, Emtiazi G, Taherzadeh MJ. Ethanol production from dilute-acid pretreated rice straw by simultaneous saccharification and fermentation with Mucor indicus, Rhizopus oryzae, and Saccharomyces cerevisiae. Enzyme Microb Technol. 2006;

46. Gong G, Liu D, Huang Y. Microwave-assisted organic acid pretreatment for enzymatic hydrolysis of rice straw. Biosyst Eng. 2010;

47. Agbor VB, Cicek N, Sparling R, Berlin A, Levin DB. Biomass pretreatment: Fundamentals toward application. Biotechnol. Adv. 2011.

48. Zhang QZ, Cai WM. Enzymatic hydrolysis of alkali-pretreated rice straw by Trichoderma reesei ZM4-F3. Biomass and Bioenergy. 2008;

49. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res. 2016;45:D200–3.

50. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc [Internet]. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2015;10:845. Available from: http://dx.doi.org/10.1038/nprot.2015.053

51. Ariaeenejad S, Maleki M, Hosseini E, Kavousi K, Moosavi-Movahedi AA, Salekdeh GH. Mining of camel rumen metagenome to identify novel alkali-thermostable xylanase capable of enhancing the recalcitrant lignocellulosic biomass conversion. Bioresour Technol. 2019;

52. Ariaeenejad S, Hosseini E, Maleki M, Kavousi K, Moosavi-Movahedi AA, Salekdeh GH. Identification and characterization of a novel thermostable xylanase from camel rumen metagenome. Int J Biol Macromol. 2019;

53. Miller GL. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem. 1959;31:426–8.
54. Ghose TK. Measurement of cellulase activities. Pure Appl Chem. 1987;
55. Kamble RD, Jadhav AR. Isolation, purification, and characterization of xylanase produced by a new species of Bacillus in solid state fermentation. Int J Microbiol. 2012;2012.
56. Kim J, Park C, Kim TH, Lee M, Kim S, Kim SW, et al. Effects of various pretreatments for enhanced anaerobic digestion with waste activated sludge. J Biosci Bioeng. 2003;95:271–5.
57. Wang CC, Chang CW, Chu CP, Lee DJ, Chang B V, Liao CS. Producing hydrogen from wastewater sludge by Clostridium bifermentans. J Biotechnol. 2003;
58. Hermiati E, Mangunwidjaja D, Sunarti TC, Suparno O, Prasetya B. Application of Microwave Heating in Biomass Hydrolysis and Pretreatment for Ethanol Production. Ann Bogor. 2010;
59. Remli NAM, Md Shah UK, Mohamad R, Abd-Aziz S. Effects of chemical and thermal pretreatments on the enzymatic saccharification of rice straw for sugars production. BioResources. 2014;

**Figures**
Figure 1

A) 3D structure of PersiCel4 Predicted from Phyre2 and PROSITE prediction server. 

B) The PersiCel4 matched based on the CDD results. 

C) SDS-PAGE analysis of recombinant endo-β-1,4-glucanase. SDS-PAGE was performed using a 12% polyacrylamide gel and stained with Coomassie Brilliant Blue. Lane 1 shows supernatant of E.coli’s lysis containing recombinant endo-β-1,4-glucanase. Lane 2-3, purified recombinant endo-β-1,4-glucanase. Lane 6 molecular weight of marker standard.
Figure 2

Characterization of purified PersiCel4. A) Effect of pH on PersiCel4 activity. B) Effect of temperature on PersiCel4 activity. C) The storage stability of PersiCel4.
Figure 3

Comparative analysis of reducing sugar release of RS after pretreatment by four different methods.
Figure 4

Curves are indicating the amount of reducing sugar in enzymatic hydrolysis of RS with PersiCel4 after 72 hours.