ABSTRACT

Cellulose is the most abundant biopolymer consisting of β-linked glucose molecules that are organized into higher order fibrillar structures[1]. It is the major polysaccharide compound in plant biomass and the most abundant source of renewable energy[2]. Cellulosic hydrolysis studies have revealed many cellulolytic microorganisms and their complex cellulases[1, 3, 4]. Cellulases enzymes are synthesized by a large diversity of microorganisms including both fungi and bacteria during their growth on cellulosic materials[5, 6]. Among them, the genera of *Clostridium*, *Cellulomonas*, *Thermomonospora*, *Trichoderma*, and *Aspergillus* are the most extensively studied cellulase producers[7, 8, 9].

The cellulase enzymes are extremely valuable owing to their significance in hydrolyzing cellulosic biomass and for their application in plant cell wall deconstruction and generation of second generation biofuels that is critical to the energy supply, environment and economy of the modern world. But the effective conversion of biomass into biofuels is limited by the nearly impenetrable architecture of plant cell walls, chemical and physical changes to biomass during pre-treatment and structural features of the biomass-degrading enzymes. Moreover, there is no single process which is cost-effective in the degradation of the natural cellulosic biomass and thus the use of the current commercial preparations of cellulase for deconstruction of cellulosic biomass is not economically feasible. So there is a need for unlocking novel cellulases from natural resources that could degrade cellulosic biomass efficiently and cost-effectively.

Genomes of the uncultured microbes encode a largely untapped reservoir of novel enzymes and bioactive molecules. The rumen of ruminants is also a largely untapped natural ecosystem that encodes highly efficient fibrolytic enzymes needed for the conversion of plant biomass into energy-rich simple sugars and low molecular weight fatty acids. The rumen, owing to the presence of unique diverse microbial consortia and microbial products, is one of the best microbial habitats to explore the sources of biomass-degrading enzymes like cellulases that may find use in the biofuel industry[10].

Metagenomics is the study of collective microbial genomes isolated directly from environmental samples for understanding the diversity of microbes, their functions, interactions and evolution.

The present investigation was undertaken to explore the goat rumen digesta for cellulase enzyme production using metagenomic approach. Metagenomic DNA (mgDNA) was extracted from goat rumen digesta using five different methods, P1[11], P2[12], P3[13], P4[14] and P5. The method P5 was developed in the laboratory[15] yielding 240.6 ± 8.1 μg of mgDNA per gram of goat rumen digesta, the highest DNA yield obtained among the methods used. The method P5 took only 3.5 h and yielded quality mgDNA suitable for restriction digestion, PCR amplification, and cloning.

The key step in the protocol P5 is the removal of tannins with the use of soluble polyethylene glycol (PEG). Tannins have been shown to inhibit PCR[16]. PEG forms an aggregate with tannins and other phenolics, which bind to the protein and cell debris upon lysis, forming a complex. This lysate, when centrifuged in the presence of phenol, results in the accumulation of protein-tannin complex at the interface between the organic and aqueous phases. The supernatant, thus obtained, is largely free from tannins[13].

The mgDNA was partially digested with *Bam*HI and the fragments (0.5-2.0 kb) were ligated into *Bam*HI-digested pUC19 cloning vector. The recombinant plasmid DNA was transferred to *E. coli* DH5α competent cells. The recombinant clones were identified by blue-white screening. The mgDNA library was assessed by digesting the recombinant pUC19 plasmid DNA isolated from randomly chosen recombinant bacterial colonies with *Kpn*I restriction enzyme. The recombinant clones showed a distinct restriction pattern confirming the presence of a DNA insert, each with a different size.

The mgDNA library was screened for cellulase positive clones on carboxymethyl cellulose (CMC) agar medium. After screening 2,000 recombinant bacterial colonies, only one cellulolytic clone was identified with the formation of yellow halo around the colony after 24-36 h of incubation at 37 ºC. The presence of ~1.5 kb DNA insert was observed with the restriction digestion of the recombinant pUC19 plasmid DNA. The same was confirmed by DNA sequencing which revealed a DNA sequence of 1,501 bp. On subjecting the DNA insert

sequence to ORF finder of NCBI, the presence of an ORF of 1,122 bp encoding the cellulase protein was revealed which showed 96% sequence homology with the cellulase of an uncultured bacterium (accession no. AFJ05146). This cellulase encoding gene was designated as 'CelT3'. The phylogenetic tree of CelT3 was constructed using Neighbor-Joining method. The cellulase gene sequence was submitted to GenBank under the accession no. KX226390.

The CelT3 gene was cloned into pET28a(+) expression vector and transferred to *E. coli* BL21(DE3) competent cells. The over-expression of CelT3 protein was induced with 0.5 mM IPTG for 5 h at 37°C. The size of recombinant protein, fused with N- terminal His-tag, was calculated to be ~44.909 kDa including the insert sequence encoding 373 amino acids (42.727 kDa) and pET28a(+) vector backbone sequence contributing 20 amino acids (2181.34 Da). On running the CelT3 protein on SDS-PAGE, a band of ~45 kDa was observed which corresponds to His-tagged CelT3 with vector backbone.

The Ramachandran plot validated the modelled structure of the CelT3 protein. The favored, allowed and outlier regions contained 96.6%, 2.8% and 0.6% amino acid residues, respectively. The plot was significant as high percentage of residues (>96%) resided in the favored region postulating that the protein is stable. The multiple sequence alignment of the CelT3 cellulase showed that it contained a glycosyl hydrolase family 5 (GH5) catalytic domain.

Biochemical characterization of the CelT3 cellulase revealed the maximum activity in the presence of CMC (17.74±1.19 U mg^{-1}) followed by Avicel (0.025±0.01 U mg^{-1}), filter paper (0.021±0.02 U mg^{-1}) and laminarin (0.016±0.01 U mg^{-1}). The kinetic study of the CelT3 enzyme revealed V_{max} and K_m values of 222.22 U mg^{-1} and 0.35 mg ml^{-1}, respectively. The purified protein exhibited a specific activity of 17.74 U mg^{-1} with 4.42 purification fold. For maximum activity of the enzyme, a temperature of 45°C and pH 6.0, presence of the divalent cations Ca^{2+} and Fe^{2+} were essential whereas Mg^{2+}, Co^{2+}, Zn^{2+} and Cu^{2+} inhibited the enzymatic activity of CelT3.

The mgDNA isolated from the goat rumen digesta was used for elucidation of cellulase gene (CelT3) through primer based amplification and cloning. The CelT3 containing recombinant pET28a(+) plasmid (pET28CT3) was transferred into *E. coli* BL21(DE3) and over expression of the enzyme confirmed. The

recombinant cellulase enzyme produced through the genetically engineered *E. coli* BL21(DE3) can be used for degrading lignocellulosic biomass in the production of nutritionally improved animal feed and as a co-additive in pulp bleaching in pulp and paper industries.

References

1. Lynd, L.R., et al. Consolidated bioprocessing of cellulosic biomass: an update, *Curr. Opin. Biotechnol.* **16,** 577--583, 2005.
2. Liu, J., et al. Cloning and functional characterization of a novel endo-β-1,4-glucanase gene from a soil-derived metagenomic library, *Appl. Microbiol. Biotechnol.* **89**, 1083--1092, 2011.
3. Lowe, S.E., Theodorou, M.K., & Trinci, A.P. Cellulases and xylanase of an anaerobic rumen fungus grown on wheat straw, wheat straw holocellulose, cellulose, and xylan, *Appl. Environ. Microbiol.* **53**, 1216--1223, 1987.
4. Ohmiya, K., et al. Application of microbial genes to recalcitrant biomass utilization and environmental conservation, *J. Biosci. Bioeng.* **95**, 549--561, 2003.
5. Kubicek, C.P. From cellulose to cellulase inducers: facts and fiction, in *Proceedings of the 2nd Symposium Trichoderma Reesei Cellulases and Other Hydrolases (TRICEL '93),* P. Suominen & T. Reinikainen, eds., Foundation for Biotechnical and Industrial Fermentation Research, Espoo, Finland, 1993, 181—188.
6. Lee, S.M., & Koo, Y.M. Pilot-scale production of cellulase using *Trichoderma reesei* Rut C-30 in fed-batch mode, *J. Microb. Biotech.* **11**, 229--233, 2001.
7. Kuhad, R.C., Manchanda, M., & Singh, A. Hydrolytic potential of extracellular enzymes from a mutant strain of *Fusarium oxysporum, Bioprocess Engineering* **20**, 133--135, 1999.
8. Sun, Y., & Cheng, J. Hydrolysis of lignocellulosic materials for ethanol production: a review, *Bioresour. Technol.* **83**, 1--11, 2002.
9. Kuhad, R.C., Gupta, R., & Khasa, Y.P. Bioethanol production from lignocellulosic biomass: an overview, in *Wealth from Waste,* B. Lal & P.M. Sarma, eds., The Energy and Resources Institute, New Delhi, India, 2011, 53—99.

10. Hess, M., et al. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen, *Science* **331**, 463--467, 2011.
11. Sharma, R., et al. Extraction of PCR-quality plant and microbial DNA from total rumen contents, *Biotechniques* **34**, 92--97, 2003.
12. Yu, Z., & Morrison, M. Improved extraction of PCR-quality community DNA from digesta and fecal samples, *Biotechniques* **36**, 808--812, 2004.
13. Krause, D.O., Smith, W.J., & McSweeney, C.S. Extraction of microbial DNA from rumen contents containing plant tannins, *Biotechniques* **31**, 294--298, 2001.
14. Popova, M., Martin, C., & Morgavi, D.P. Improved protocol for high-quality co-extraction of DNA and RNA from rumen digesta, *Folia Microbiol.* **55**, 368--372, 2010.
15. Bashir, Y., Rather, I.A., & Konwar, B.K. Rapid and simple DNA extraction protocol from goat rumen digesta for metagenomic analysis, *Pak. J. Pharm. Sci*. **28**, 2305--2309, 2015.
16. Stefanis, C., et al. Principal methods for isolation and identification of soil microbial communities, *Folia Microbiol.* **58**, 61--68, 2013.