Bioprospecting of Hot Springs of India for Thermophilic Xylanases Using Metagenomics Approach

SUMMARY

SUBMITTED TO THE

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY

(A CENTRAL UNIVERSITY)

FOR THE AWARD OF THE DEGREE OF

Doctor of Philosophy

in

BIOTECHNOLOGY

Submitted by

GARIMA CHAUHAN

UGC Fellow

Enrolment No: 953/15

Under the supervision of

Dr. MONICA SHARMA

Assistant Professor

DEPARTMENT OF BIOTECHNOLOGY

SCHOOL FOR BIOSCIENCES & BIOTECHNOLOGY

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY

(A CENTRAL UNIVERSITY) (NAAC: A++ Accredited)

VIDYA VIHAR, RAEBARELI ROAD

LUCKNOW-226025, UTTAR PRADESH (INDIA)

2024

Summary

Xylans are the primary components of the hemicellulose polysaccharide found in the cell walls of terrestrial plants, ranking as the second most prevalent polysaccharide on the planet (Basit et al., 2018). They are composed of poly-xylose sugar backbone connected through β-1,4 glycosidic linkages, along with several substituent groups in their side chains such as acetyl, 4-O-methyl-D-glucuronosyl, D-glucuronic acid, and α-arabino-furanosyl etc, which makes it heteropolysaccharide (Hsieh and Harris, 2019). Xylanases belong to the glycosidic hydrolase (GH) family and play an essential role in xylan hydrolysis. They primarily target the xylan backbone and break down the β-1,4 glycosidic bonds between xylopyranosyl residues at internal sites, resulting in a wide range of hydrolytic products such as xylobiose, xylotriose, xylotetraose, xylopentose, and branched longer xylo-oligomers (XOS). These hydrolytic products can be used as prebiotics (Kayserilioğlu et al., 2003).

Thermo-alkali stable xylanases have many advantages in the industrial domain, such as reducing the viscosity of liquids, accelerating growth rates, enhancing solubility in oil and polymeric substrates, unique substrate specificity, resistance to chemical agents, extending shelf life, and reducing chances of contamination (Bhardwaj et al., 2019). Thermophilic xylanases are of significant interest to the researchers due to their diverse applications in various sectors such as the feed and food industry, bio-refinery industry, textile, detergent, and pulp-paper industries (Basit et al., 2021). Most of the xylanases are obtained from 0.1-1% cultivable microorganisms present in natural habitat. On the contrary, culture-independent metagenomic strategies facilitate the recovery of genes that encode beneficial enzymes from environmental samples, eliminating the necessity of cultivating microorganisms (Sleator et al., 2008).

The need for xylanases which withstand elevated temperature and alkaline environments, especially in the field of pulp bleaching has encouraged us to employ this innovative strategy to discover genes encoding xylanases with thermo-alkali stability. This study discovered a highly thermostable and alkali-stable xylanase-encoding gene (xyn GM) from the metagenomic library of a Tapovan Hot-spring soil sample using a metagenomics approach. This is an endoacting thermo-alkalistable enzyme which may find application in

the pharmaceutical industry to produce xylooligosaccharides and in the pulp-paper industry for bleaching pulp.

Keeping above in view, the current research work entitled **'Bioprospecting of Hot Springs of India for Thermophilic Xylanases Using Metagenomics Approach'** has been undertaken with the aim to isolate and characterize a robust xylanase enzyme from various Hot springs of India and used in various industrial sectors. The objectives of this research work were:

1. To isolate metagenome from an environmental sample and standardization of DNA extraction protocol (soil).
2. To construct a genomic library and its screening for thermophilic carbohydrate degrading enzymes.
3. To perform sequence metagenomics using homology-based screening of the metagenome for xylanase enzyme.
4. Sequencing and analysis of the gene/enzyme (xylanase) using bioinformatics tools.
5. Recombinant expression and characterization of recombinant protein.

The major findings of the current research work include:

1. **To isolate the metagenome from an environmental sample and standardization of DNA extraction protocol (soil)**

Soil and sediment samples were collected from different hot springs across India, with temperatures ranging from 50 °C to 95 °C and pH levels varying from 5.5 to 8.0. A new protocol was developed to isolate metagenomes from two soil and sediment samples, TP-B-88.8 °C and TP-D-55.5 °C, obtained from the Tapovan hot spring. These two samples were selected to standardize the protocol due to their varying microbial loads and diverse microbiota.

The developed protocol uses a combination of three chemicals (PAC, PVPP, and $CaCl_2$) to extract humic acid-free metagenomic DNA from extreme environmental soil samples in the lysis buffer. The importance of this protocol lies in the fact that it yields good DNA quantity and high purity, which is generally absent in most previously published protocols and commercial soil DNA kits. As the protocol relies on direct lysis of the samples, the extracted metagenomic DNA is a good representative of various microbial communities, which is highly desirable in a metagenome-based study to avoid sample biasing.

The protocol was successfully validated by digestion of the extracted metagenome through various restriction enzymes and PCR amplification of the 16S rRNA gene using universal primers, making it suitable for various downstream molecular processes such as shotgun sequencing and metagenomic library construction. The developed protocol was used to isolate metagenomes from five different soil/sediment samples of various hot springs, including Tapovan (Uttarakhand), Atri Hot Spring (Odisha), Chawalpani (Madhya Pradesh), Manikaran (Himachal Pradesh), and Surajkund (Jharkhand). All the extracted metagenomes were free from humic acid and suitable for PCR amplification, restriction digestion, and metagenomic library construction.

2. To construct a genomic library and its screening for thermophilic carbohydrate degrading enzymes

Metagenomic DNA libraries were created from various soil samples using the *EcoRI* restriction enzyme at a concentration of 0.5-2 U to prepare the fragments of metagenomic DNA within the 1-6 kb size range. These fragments were then cut and ligated into *EcoRI*-digested pUC19 vectors. The transformed host *E. coli* DH5α was used to prepare the metagenomic library. Recombinant clones obtained from different metagenomic DNA libraries were selected and preserved using replica plating methods. Five thousand recombinant clones with metagenomic DNA were analysed for carbohydrate-active enzymes (CAZy). Recombinant clones were randomly chosen for sequencing, three of which tested positive for CAZy enzymes such as amylase, amylopullulanase, and xylanase. These enzymes were amplified through their respective degenerate primers. The amylase gene of 1503 bp was cloned into the linearized T-vector, which exhibited a maximum identity of 98.95% with *Bacillus licheniformis* SRCM 103914, *B. licheniformis* strain CP6, and *B. licheniformis* 584, with 81% query coverage at the nucleotide level.

The amplicon of the xylanase gene was approximately 526 bp in size and encoded partial CDS for the xylanase protein. Therefore, to obtain a fully functional xylanase enzyme, metagenomes from the Tapovan hotspring were sequenced, and two xylanase-encoding genes and other CAZy enzymes were retrieved.

3. To perform sequence metagenomics using homology-based screening of the metagenome for xylanase enzyme.

Environmental DNA (eDNA) was extracted from a Tapovan soil sample and used as a template for PCR amplification. The xylanase gene was then amplified using degenerate primers through PCR, resulting in a 927 bp PCR amplicon. The amplified gene was inserted into a linearized T-vector and transformed into *E. coli* DH5α host cells, resulting in the observation of 1698 colonies. Recombinant vectors carrying the xylanase gene were subsequently isolated.

Confirmation of the insert's presence in recombinant vectors was determined by colony PCR, successfully amplifying the 927 bp xylanase gene. Additionally, restriction digestion with the *EcoRI* enzyme was used to validate the presence of the

insert by visualizing the 927 bp xylanase gene amplicon and the 3016 bp backbone of the linearized T-vector on an agarose gel. The 927 bp xylanase gene was sequenced using the Sanger sequencing method and analysed with NEB cutter 2.0, which revealed the absence of *BamHI* and *XhoI* restriction sites. These sites were incorporated upstream of the xylanase degenerate primers to facilitate further cloning.

In the subsequent step, the xylanase gene with incorporated restriction sites was subcloned into an expression vector called pET28a (+), and *E. coli* BL21(DE3) was used as the host for this expression vector. Approximately 40 recombinant clones were obtained on an LB-agar-kanamycin plate, and the presence of the xylanase gene in the recombinant vectors was confirmed through colony PCR and restriction digestion with *BamHI* and *XhoI* restriction enzymes. The successful clones were sequenced using the Sanger sequencing method with the T7 forward primer, providing a confirmed and sequenced xylanase gene in the pET28a (+) expression vector.

4. Sequencing and analysis of the gene/enzyme (xylanase) using bioinformatic tools.

The obtained sequence data of the xylanase gene (xyn GM) was analysed using various bioinformatics tools. Compared with the NCBI database using BLASTn, it exhibited a high similarity of 99.31% with the xylanase gene (XynB) of *Cohnella* sp. strain A01, showing 93% coverage. However, a significant mismatch was observed at the 5' end of the gene in the pairwise alignment. Further analysis with BLASTx revealed a 100% match of the insert sequence with multispecies of *Cohnella* endo-1,4-beta-xylanase, as well as matches with other xylanase genes, such as *Paenibacillus mucilaginosus* K02 xylanase B, *Cohnella thermotolerance* endo-1,4-beta-xylanase, and *Paenibacillus woosongensis* endo beta-1,4-xylanase.

The ORF finder tool predicted 19 ORFs, with ORF1 encoding the xylanase gene containing 308 amino acids located on the +1 strand of DNA. Multiple sequence alignment unveiled two signature sequences of glycoside hydrolase family 10 (DVVNE and GLPIYVSELDM). Phylogenetic analysis indicated significant homology with other 1,4-beta-xylanases of multispecies of *Cohnella*.

Signal P 6.0 predicted a signal peptide at the N-terminal of the xylanase protein, with a high probability of 0.98. Cleavage sites were present between the 18th and 19th

positions. Prosite software identified a glycoside hydrolase family (GH10) domain in the xylanase gene, containing conserved Glu residues at positions 133 and 240, acting as a proton donor and nucleophile during catalytic reactions. ProtParam analysis revealed that the xylanase protein, consisting of 308 amino acids, had 24 acidic and 21 basic amino acids. The theoretical molecular weight and isoelectric point of recombinant xylanase were approximately 34 kDa and 6.08, respectively. The value of the instability index of xylanase was 26.07, which suggests that the protein is stable. Homology modelling using the Phyre2 online program indicated a 50% identity with GH10 endo-β-1,4-xylanase of *Xanthomonas axonopodis*pv. citri. These analyses provided a detailed understanding of the xylanase gene's characteristics, structural features, and potential functional aspects.

5. Recombinant expression and characterization of recombinant protein.

Expression experiments at varying temperatures and IPTG concentrations revealed limited success, with no significant expression observed at 37 °C and 30 °C. Only at 25 °C was basal level protein expression observed due to the presence of signal peptide sequence in the recombinant xylanase protein, which transports the recombinant protein in the extracellular medium. Induction of *E. coli* BL21 (DE3) cells carrying recombinant vector with 0.5 mM IPTG and harvesting after 5 hours resulted in the acquisition of periplasmic protein, identified as the recombinant xylanase crude enzyme. The purified recombinant xylanase protein, obtained through Ni^{2+}-NTA affinity chromatography, showed a single band of approximately 32.5 kDa on a 12% SDS-PAGE gel, consistent with the predicted molecular mass. Zymography on SDS-PAGE containing xylan substrate exhibited a yellow zone of hydrolysis, confirming the enzymatic activity of the recombinant xylanase.

The characterization of the xylanase enzyme encompassed an assessment of its physicochemical parameters, substrate specificity, and the effects of various additives on the activity of the purified xylanase enzyme. The enzyme demonstrated maximum activity at 70 °C temperature and pH 9.0, with >65% residual activity at 50 °C and 60 °C after 2 hours and >75% residual activity found at pH 9.0 for 3 hours.

Substrate specificity revealed xylanase preference for beechwood xylan, and no activity was observed with carboxymethylcellulose (CMC) substrate, indicating that it did not

have cellulase activity and could be used in the pulping process in the paper industry. Metal ions like Mn^{2+}, Ca^{2+}, and Co^{2+} positively influenced the enzyme activity at 1 mM concentration, while higher concentrations of certain metal ions led to the reduction in xylanase activity. Solvents and detergents displayed varying inhibitory effects, with chloroform, glycerol, and methanol showing high inhibition at a concentration of 10%. Inhibitors like EDTA, DTT, and β-mercaptoethanol exhibited slight inhibitory effects due to a single cysteine amino acid residue in the xylanase protein.

The K_M and V_{Max} values were determined using beechwood xylan as a substrate and yielded values of 20.9 mg/mL and 156.25 µmol/mg/min, respectively. The enzyme displayed thermostability with a half-life of 3.0 hours at 50 °C and 2.3 hours at 60 °C. The purified xylanase enzyme's activation energy (EaD) was 38.6 kJ/mol.

The analysis of hydrolytic products of beechwood xylan using thin-layer chromatography (TLC) demonstrated the production of xylobiose, xylotriose, xylotetraose, and xylopentose after one hour of incubation at 70 °C temperature. The production of various xylooligosaccharides other than xylose confirmed the endoacting activity of the purified xylanase. It could be used in the pharmaceutical industry to produce xylooligosaccharides that can be used as prebiotics and help improve human gut health.

References

- Basit, A., Jiang, W., & Rahim, K. (2021). Xylanase and its industrial applications. *Intech Open*. https://doi.org/10.5772/intechopen.92156
- Basit, A., Liu, J., Rahim, K., Jiang, W., & Lou, H. (2018). Thermophilic xylanases: from bench to bottle. *Critical Reviews in Biotechnology, 38*(7), 989-1002. https://doi.org/10.1080/07388551.2018.1425662
- Bhardwaj, N. K., Kumar, B., & Verma, P. (2019). A detailed overview of xylanases: an emerging biomolecule for current and future prospective. *Bioresources and Bioprocessing*, *6*(40). https://doi.org/10.1186/s40643-019-0276-2

- Hsieh, Y. S. Y., & Harris, P. J. (2019). Xylans of Red and Green Algae: What Is Known about Their Structures and How They Are Synthesised? *Polymers*, *11*(2), 354. https://doi.org/10.3390/polym11020354
- Sleator, R. D., Shortall, C., & Hill, C. (2008). Metagenomics. *Letters in Applied Microbiology*, *47*(5), 361-366. https://doi.org/10.1111/j.1472-765x.2008.02444.x

Garima Chauhan
UGC Fellow (Enrolment No.- 953/15)
Department of Biotechnology
School for Biosciences & Biotechnology
Babasaheb Bhimrao Ambedkar University
Lucknow, Uttar Pradesh-226025

Supervisor
Dr. Monica Sharma
Department of Biotechnology
School for Biosciences & Biotechnology
Babasaheb Bhimrao Ambedkar University
Lucknow, Uttar Pradesh-226025

Head of the Department
Prof. D R Modi
Department of Biotechnology
School for Biosciences & Biotechnology
Babasaheb Bhimrao Ambedkar University
Lucknow, Uttar Pradesh-226025