MOLECULAR CLONING, SEQUENCE ANALYSIS AND ENGINEERING OF *Pichia etchellsii* β-GLUCOSIDASE FOR SYNTHESIS OF OLIGOSACCHARIDES

RICHA BARANWAL

DEPARTMENT OF BIOCHEMICAL ENGINEERING AND BIOTECHNOLOGY

INDIAN INSTITUTE OF TECHNOLOGY DELHI

MOLECULAR CLONING, SEQUENCE ANALYSIS AND ENGINEERING OF *Pichia etchellsii* β-GLUCOSIDASE FOR SYNTHESIS OF OLIGOSACCHARIDES

By

RICHA BARANWAL

Department of Biochemical Engineering and Biotechnology

Submitted in fulfillment of the requirement of the degree of

DOCTOR OF PHILOSOPHY

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

APRIL 2010

DEDICATION

TO MY DEAR PARENTS, PARENTS IN LAW,

LOVING HUSBAND AND MY LITTLE ANGEL

YUG

CERTIFICATE

This is to certify that the thesis entitled **"Molecular cloning, sequence analysis and engineering of *Pichia etchellsii* β-glucosidase for synthesis of oligosaccharides"** being submitted by **Ms. Richa Baranwal** to the **Indian Institute of Technology Delhi**, for the award of the degree of **'Doctor of Philosophy'**, is a record of the bonafide research work carried out by her, which has been prepared under my supervision in conformity with the rules and regulations of the 'Indian Institute of Technology Delhi'. The research reports and the results presented in this thesis have not been submitted for any degree or diploma in any other University or Institute.

Prof. Saroj Mishra

ACKNOWLEDGEMENT

First and foremost, I am extremely thankful to my respected supervisor, ***Prof. Saroj Mishra****, for providing the vision for the genesis of this research as well as her continual help, support and valuable guidance during the course for its execution. Despite her hectic schedules in teaching, research and administration, she has always been very generous with her time to share her knowledge, experience and research insight through thoughtful discussion and critical analysis. I am grateful to her and feel privileged to have been her student. Her perfection, punctuality and organization have helped to shape me a researcher with lot of scientific skills.*

I sincerely express my deep gratitude to the members of my 'Scientific Research Committee', ***Prof. V. S. Bisaria, Dr. Prashant Mishra*** *and* ***Dr. Nivedita Karmakar Gohil*** *(Centre for Biomedical Engineering, IIT Delhi) for guiding me through all these years and critically reviewing the progress of this research work to make necessary improvements.*

I am thankful to all my learned teachers and faculty members of the Department of Biochemical Engineering and Biotechnology, IIT Delhi for showing me the beautiful spectrum of biotechnology in the light of modern sciences. Special thanks to ***Mr. V. K. Ghosh*** *for teaching me the basic concepts of microbiology. I shall be grateful to him for his contribution in numerous ways and his care in giving me new ideas. I am also thankful to* ***Dr. Indrani Das Gupta*** *and his student* ***Basanta Bora*** *(University of Delhi, South Campus) and* ***Gaurav Anand*** *(ICGEB, New Delhi) for their kind help in electroporation experiments. I also express my gratitude to* ***Prof. S. H. Dey*** *(IIT Kharagpur) and* ***Prof. Vinay Sharma*** *(Banasthali Vidyapeeth) who have been my 'guru' of the subject before I was here.*

I am grateful to IIT Delhi for providing me the institute fellowship and teaching assistantship for carrying out my research work and providing an ideal environment to pursue the doctoral study at this prestigious institute.

Thanks are also due in many regards to ***Renu madam, Mukesh Anand, Sumit ji,*** *and other staff of the DBEB, IIT Delhi. I am grateful to* ***Kishan ji*** *and* ***Bhagwan Singh ji*** *for providing me the necessary chemicals for research at times and* ***Ram Gopal ji*** *and* ***Raj Kumar ji*** *for providing me 'ready to use' glassware and consumables required for carrying out research. I wish to express my gratitude to* ***Neera madam*** *for helping me with vast resources of books and journals. I am also thankful to the official staff (****Sunita, Pushplata*** *and* ***Meena madam****) for their efforts to facilitate management of official and institutional responsibilities.*

I am sincerely thankful to all my seniors particularly ***Saloni, Pranita, Usha Sri, Rumpa, Shilpi, Ashwini, Ranjeeta, Anand, Parul, Anjali, Ashish*** *for passing down their research skills. I am thankful to* ***Asif*** *for being a responsible senior and providing me the tips for FPLC and protein purification. I am grateful to my batch mates* ***Mukesh, Vinod, Brijendra*** *and* ***Badilu*** *for their help in study during course work. I also express thanks to* ***Meenu, Neha, Swati, Jyoti, Yunus, Saurab, Gayatri*** *and* ***Prabha*** *for being remarkable juniors both in intelligence and spirit. It was also a great pleasure to work with M.Tech. Project students* ***Sandeep, Vinayak and Saumya Jain.***

IIT Delhi has brought me in contact with delightful and wonderful individuals with sharp intellects. It fills me with great pride to say that I have found truly lifelong friends in them. Many- many thanks to ***Bhawna Madan*** *for being a wonderful senior and a friend with a helping hand in every possible way, to* ***Raju****, for being a great friend supportive in many ways and a source of required reasons and logics and to* ***Roohi*** *for being a true friend and a wonderful person for providing help and support at every step. Special thanks*

*to **Arun** and **Akhilesh** and **Rakhi** for giving their helping hand in many ways which are memorable throughout life, to **Bhawna Agrawal** for being a wonderful friend and batch-mate. I am thankful to my friends **Shweta** and **Kavita** for their love and support through out the period spent in IIT Delhi.*

*Finally I am coming to the most important people in my life, my family for their everlasting love, faith and encouragement. This research would simply have been impossible without them. I am grateful to my father, **Prof. Nagendra Kumar** for being an icon of moral and professional honesty, to my mother **Ms. Vimlesh** for her loving care and her blessings. To my loving husband, **Mr. Shakti Baranwal** for being my pillar of strength, my joy and my guiding star. His support and confidence in my abilities have helped me to overcome difficulties in my life. Special thanks for my in laws **Mr. M. L. Baranwal** and **Ms. Usha Baranwal** for the sincere and utmost care, concern and trust. I am obliged to my adoring family members (**Gaurav, Nalini, Shailendra, Nilima, Vipul, Babita, Shalini, Tripti, Vibhuti, Rita, Anjali, Nishant and Ravi**) for their love, care and support. I will remain forever indebted to them.*

*At the end, I sincerely bend my head down for almighty **GOD** who make things happen.*

Richa Baranwal

ABSTRACT

The thermo-tolerant yeast *Pichia etchellsii* produces two inducible, wall bound β-glucosidases viz. BGLI and BGLII. Both the enzymes show transglycosylation activities leading to synthesis of glycoconjugates like alkyl glucosides. In this study BGLI was selected as the model enzyme in order to evaluate the feasibility of this large yeast β-glucosidase for synthesis of sugars. The *BGL1* gene, encoding BGLI enzyme was cloned and functionally expressed in *Pichia pastoris* under the control of *AOX1* promoter. Site-directed mutagenesis studies were done to confirm the putative nucleophilic residue important in enzyme catalysis. The recombinant wild type and mutant enzymes were investigated for synthesis of oligosaccharides.

The internal peptide sequences of native BGLI protein were aligned with GHF3 β-glucosidases using Clustal W program of DNASTAR to design the primers to fish out *BGL1* gene from genomic DNA of *P. etchellsii*. Step down PCR, 3' RACE and Primer Walking methods were used to find out complete nucleotide sequence of *BGL1*. The complete sequence of *BGL1* was assembled by sequencing products obtained from three PCR methods. The intronless *BGL1* gene (GenBank Accession No. EU914813) ORF consisted of 2,544 bp nucleotides encoding a protein of 847 amino acids. The ORF predicted a protein of molecular mass of 93.4 kDa. The predicted isoelectric point (pI) of the protein was 5.2 and A+T % and G+C % were 58.06 % and 41.95 % respectively. The sequence homology search using BLAST revealed sequence identity with several members of GHF3 and the sequence was concluded to belong to GHF3. The multiple sequence alignment using ClustalV program of DNASTAR showed maximum 98.6 % identity with a hypothetical 765 aa protein of *Kluyveromyces lactis* and 73.4 % with 845 aa β-glucosidase protein of *Kluyveromyces fragilis*.

The *BGL1* gene, encoding BGLI enzyme was cloned in yeast shuttle expression vector, pPIC9 in-frame with the native *Saccharomyces cerevisiae* α-factor secretion signal sequence and functionally expressed in *P. pastoris* GS115 strain under the control of *AOX1* promoter. GS115pPIC9*BGL1*-5 transformant of *P. pastoris* showed strong fluorescence under UV light on MUG as a substrate when screened for secretory expression of rBGLI by agar plate assay, confirming successful extracellular expression of recombinant enzyme. GS115pPIC9*BGL1*-5 was found to be His^+ Mut^+ as transformants grew well on both MD and MM plates.

The two clones GS115pPIC9*BGL1*-5 and GS115pPIC9*BGL1*-41 were selected for expression studies in baffled flasks based on results of PCR, dot blot and agar plate assay. The recombinant enzyme activity reached a maximum of 560 IU/L on eighth day in case of GS115pPIC9*BGL1*-5 and 415 IU/L for GS115pPIC9*BGL1*-41. The effect of temperature during culture growth and induction was optimized which indicated a maximum activity of 974 IU/L on day 6 which was about 1.74 fold higher than in unoptimized culture. Thus, temperature of cultivation was found to affect expression. Northern analysis confirmed that the increase was at mRNA level.

Native PAGE showed fluorescent high molecular weight band under UV light in GS115pPIC9*BGL1*-5 medium supernatant which was not present in control GS115pPIC9. This further confirmed the secretory expression of protein in culture medium. The enzyme was purified partially by a combination of ammonium sulfate precipitation, Sephadex G 200 chromatography and ion exchange chromatography to a specific activity of 3.5 IU/mg of protein with 15 % yield. *P. pastoris* secreted rBGLI had an apparent molecular weight of 97.3 kDa as shown on SDS-PAGE, similar to the native BGLI in size (97.7 kDa).

The rBGLI showed maximal enzyme activity at pH 6.0 and more than 80% stability between pH 3.5–9. The enzyme had temperature optimum of 50 °C under optimal pH with

*p*NPG as substrate. Half life values of enzyme at 30 °C, 35 °C and 37 °C were found to be 16.5 h, 14.4 h and 12.8 h. However at 40 °C, 45 °C and 50 °C $t_{1/2}$ values of 6.0 h, 3.4 h and 0.5 h were obtained. A decline in enzyme activity was observed with Cu^{2+} and Fe^{2+} (~38 and 40% respectively) ions and only 27.34 % of activity was left in presence of 1 mM Hg^{2+} ions. The properties of the enzyme produced in *P. pastoris* were found to be similar to the native BGLI from *P. etchellsii*.

The putative nucleophilic residue important in enzyme catalysis was identified using molecular modeling and site-directed mutagenesis. The program 3Djigsaw (available online at http://www.bmm.icnet.uk/servers/3djigsaw/) provided the modeled structure of *P. etchellsii* BGLI protein using *Hordeum vulgare* EXO I as a template. Modeled structure showed the Tim Barrel domain and the (α/β) 6 domain. Unlike the *H. vulgare* EXO I, the latter domain was larger than the Tim Barrel domain and could not be well-defined. The two catalytic amino acids, D 227 and E 590 were placed in a position similar to the ones found in the known structure of *H. vulgare*. Since both these residues also occupied similar positions in modeled structure, these were used as lead sites for mutagenesis. The aspartate residue (D-227) present in SDW repeat of *P. etchellsii* BGLI was selected for mutagenesis and changed to Glycine and Valine using QuikChange site-directed mutagenesis protocol (Stratagene). In zymogram analysis with culture supernatant of mutants (D227G, D227V), no high molecular weight band corresponding to the position of rBGLI showing fluorescence on MUG was obtained. This confirmed the importance of this residue in catalysis and it was proposed that it functioned as a nucleophilic residue.

The recombinant wild type and mutant enzymes were investigated for synthesis of oligosaccharides. Using rBGLI with glucose as the substrate (167 mM), disaccharides were observed at 8 h of reaction and reached a maximum of 13.14 mmol/L at 20 h. The accumulation of trisaccharides started later and reached a maximum of 4.29 mmol/L. A

total conversion of ~11 % was calculated. The formation of di- and tri- saccharides was confirmed by ESI-MS. The sugars were detected as Na^+ adducts (m/z = 203, 365 and 527).

The possibility of using the nucleophilic mutant, D227G as glycosynthase for enzymatic synthesis of oligosaccharides was also investigated. The principle was based on findings of Withers *et al.*, (1998). Glycosynthases are specifically mutated glycosidases that efficiently synthesize oligosaccharides but don't hydrolyze them. The synthesis of oligosaccharides by transglucosylation approach using α-D glucopyranosyl fluoride (10 mM) as a donor and *p*NPG, *p*NPC and cellobiose as acceptors was investigated. Higher oligosaccharides were detected on TLC when cellobiose was used as acceptor and not with *p*NPG and *p*NPC. The percent conversion of cellobiose to sugars was about 45%. The formation of oligosaccharides by mutant enzyme (D227G) in reaction vials containing cellobiose was also confirmed by ESI-MS. The sugars were detected as Na^+ adducts and identified as trisaccharides (m/z=527) and tetrasaccharide (m/z=689). The peaks corresponding to higher sugars were not observed in mass spectrum when α-D glucopyranosyl fluoride (10 mM) was used as a donor and *p*NPG and *p*NPC were used as acceptor.

CONTENTS