Strain improvement of *Bacillus amyloliquefaciens* SS35 for enhanced endoglucanase catalytic efficiency and identification of mutation causing the structure changes by cloning, expression and purification of glycoside hydrolase family 5 endoglucanase and its application in saccharification of *Sorghum durra*

PhD Thesis

by

Shweta Singh

January 2020

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI – 781039, ASSAM, INDIA

TH-2393_146106034

TH-2393_146106034

Strain improvement of *Bacillus amyloliquefaciens* SS35 for enhanced endoglucanase catalytic efficiency and identification of mutation causing the structure changes by cloning, expression and purification of glycoside hydrolase family 5 endoglucanase and its application in saccharification of *Sorghum durra*

A Thesis

Submitted in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy

by

Shweta Singh

Under supervision of

Professor Arun Goyal

January 2020

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI – 781039, ASSAM, INDIA

TH-2393_146106034

TH-2393_146106034

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

DEPARTMENT OF BIOSCIENCES & BIOENGINEERING

STATEMENT

I do hereby declare that the content embodied in this thesis entitled as **"Strain improvement of *Bacillus amyloliquefaciens* SS35 for enhanced endoglucanase catalytic efficiency and identification of mutation causing the structure changes by cloning, expression and purification of glycoside hydrolase family 5 endoglucanase and its application in saccharification of *Sorghum durra*"** is the result of investigations carried out by me in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati, India under the guidance of Professor Arun Goyal.

In keeping with the general practice of reporting scientific observations, due acknowledgements have been made wherever the work described is based on the findings of other investigators.

January, 2020 *Shweta Singh*

(146106034)

TH-2393_146106034

TH-2393_146106034

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

DEPARTMENT OF BIOSCIENCES & BIOENGINEERING

CERTIFICATE

It is certified that the work described in this thesis entitled **"Strain improvement of *Bacillus amyloliquefaciens* SS35 for enhanced endoglucanase catalytic efficiency and identification of mutation causing the structure changes by cloning, expression and purification of glycoside hydrolase family 5 endoglucanase and its application in saccharification of *Sorghum durra*"** by **Shweta Singh (Roll No. 146106034)** for the award of degree of Doctor of Philosophy is an authentic record of the results obtained from the research work carried out under my supervision at the Department of Biosciences & Bioengineering, Indian Institute of Technology Guwahati, Guwahati, India and this work has not been submitted elsewhere for a degree.

Dr. Arun Goyal *(MTech, PhD)*
(FAMI, FBRS, FABAP, FNABS, FNAAS, FIFIB)
Professor
(Thesis Supervisor)
Department of Biosciences & Bioengineering
Indian Institute of Technology Guwahati
Guwahati, 781 039, India

TH-2393_146106034

TH-2393_146106034

ACKNOWLEDGEMENTS

First of all, my eternal bow in reverence and gratitude is to "Almighty God" whose gracious blessings enabled me to complete my work on time. Words cannot express the depth of my gratitude for all those who directly or indirectly helped me in my endeavor.

I convey my heartfelt and esteemed sense of gratitude to my Ph.D. supervisor, Professor Arun Goyal, Department of Biosciences & Bioengineering, Indian Institute of Technology Guwahati for their guidance, moral support, encouragement and providing me with the necessary instructions and research facilities for my research work.

I would also like to express my sincere thanks to all my doctoral committee members Dr. Manish Kumar, Dr. Sachin Kumar and Prof. Ranjan Tamuli for their valuable suggestions and guidance that has led to the successful completion of my research work.

I am very much thankful to Department of Biosciences & Bioengineering and Central Instrumentation Facility (CIF), Indian Institute of Technology Guwahati for providing me instruments for my research work.

It is my extreme happiness to thank and acknowledge the present and former heads of the Department of Biosciences & Bioengineering, Indian Institute of Technology Guwahati, Prof. Latha Rangan, Prof. K. Pakshirajan, and Prof. Venkata. V. Dasu for providing me with the necessary facilities.

I am immensely thankful to my seniors Dr. Shuchi Singh, Dr. Arun Dhillon, Dr. Aruna Rani, and Dr. Rwivoo Baruah for their help and suggestions. I am greatly thankful to all research group members for their cooperation and support. I am also thankful to all the people with whom I have worked in the lab at the Department of Biosciences and Bioengineering for their kind support.

TH-2393_146106034

I extend my deep sense of appreciation to my loving friends Vivek, Smita, Krishan, Priyanka and Ruchika for their generous help and support during my stay in Guwahati.

I want to express my heartfelt thanks to my grandmother Late Maharani Devi, my parents Gyan Singh and Sheela Singh, my sisters Chitra Singh and Dr. Ambika Singh, my brothers Chetan Raj Singh and Jitendra Singh and my husband Dr. Gaurav Singh for their immense love, support and care.

I want to thank all teaching and non-teaching staff members of the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati for their support.

I wish to acknowledge continuous financial support from MHRD, Govt. of India for providing fellowship and also thankful to Department of Biotechnology, Govt. of India, New Delhi for providing financial assistance through its sponsored project.

Shweta Singh

January, 2020

TH-2393_146106034

Abstract

In bioethanol production, the enzymatic saccharification of lignocellulosic biomass for release of reducing sugars is the cost limiting step. Therefore, for reducing the production cost of bioethanol, strain improvement of cellulase producing microorganisms is important. Random mutagenesis by UV in microorganism resembles the natural evolution process. Therefore, it can be used for improvement of biochemical properties in industrial enzymes (cellulase). Wild-type strain of *Bacillus amyloliquefaciens* SS35 was exposed to UV irradiations to develop the UV2 mutant strain with improved endoglucanase catalytic efficiency and wide range pH stability. The gene encoding endoglucanase, *Ba*GH5-WT and *Ba*GH5-UV2 were amplified from wild-type and UV2 mutant of *Bacillus amyloliquefaciens* SS35, respectively, using degenerate primers for family 5 glycoside hydrolase (GH5) and were cloned in pET-28a(+) vector and expressed in *E. coli* BL21(DE3) pLysS cells. The recombinant mutant *Ba*GH5-UV2 showed 22-fold higher catalytic efficiency and wider range pH stability than recombinant wild-type *Ba*GH5-WT. The mutant enzyme, *Ba*GH5-UV2 showed substitution mutation of residue, Asp256 to Gly256. This mutation was in loop connecting the β_6 to α_6 of $(\beta/\alpha)_8$ TIM-barrel fold. Molecular dynamics simulation studies showed more stable 3-D structure for *Ba*GH5-UV2 than *Ba*GH5-WT. Molecular docking results showed that *Ba*GH5-UV2 gave maximum increase in Gibb's free energy ($\Delta G°$) against cellotetraose. Application of *Ba*GH5-UV2 in saccharification of acid or base pretreated *Sorghum durra* stalk in cocktail with *Ct*CBH5A and *Ct*GH1 was explored for pretreatment specific customization of enzymes in mixture design. This report provides the information for protein engineering in GH5 endoglucanases for improving their biochemical properties and pretreatment specific optimization of enzymes in mixture for enzymatic saccharification.

TH-2393_146106034

TH-2393_146106034

TH-2393_146106034

SYNOPSIS

Introduction

In the present scenario the human population is heavily dependent on oil, mainly the transportation sector with a daily requirement of 84 million barrels of fossil fuels (Daniel *et al.*, 2010). However, the increasing price of oil, its unstable supply and environmental related issues are the forces that lead to the production of renewable fuels. Lignocellulosic biomass is the most potential renewable feedstock for the sustainable production of bioethanol because of its whole year decentralized availability, low market value, less prone to microbial contamination and high cellulose content (Cardona *et al.*, 2007). It contains approximately, 40-60% cellulose (Putro *et al.*, 2016). The bioconversion of lignocellulosic biomass to fermentable sugars is difficult because of its physio-chemical structural and compositional factors, which hinders the enzymatic digestibility of cellulose (Van Zyl *et al.*, 2007). Only, the cellulase production accounts for approximately, 40% of the total cost of the process (Muthuvelayudham *et al.*, 2004). Despite the efforts from several laboratories across the world, no commercially efficient enzyme complex has been produced. Therefore, significant efforts are required to overcome the challenges of high production cost of cellulase enzyme, low enzyme yield and low enzyme activities (Kim *et al.*, 1987). Random mutagenesis of bacterial cellulase producing strains provides a better route for the improvement of biochemical properties of these enzymes (Chand *et al.*, 2005). This can be achieved by UV irradiation and chemical agents providing the tools to achieve the genetic and functional modifications in microorganism (Singh *et al.*, 2017). This can

TH-2393_146106034

be applied in a laboratory evolution of proteins and generates random genetic diversity of a gene with improved cellulase activities which has no prior structural knowledge (Leemhuis *et al.,* 2009; Packer *et al.,* 2015). For enhancing the industrial enzymes market, recombinant DNA technology was combined with directed evolution/strain improvement to suit the requirements of the users (Hodgson, 1994). Production of natural proteins from microorganism was enhanced by recombinant DNA technology to several folds. The desired gene from particular microbial strain was amplified by designing the degenerate primers from the most similar phylogenetic sp. on the basis of 16S rDNA sequence (Lorenz *et al.,* 2002). Cellulases are classified as glycoside hydrolase enzymes belonging to families 5-9, 12, 26, 44, 45, 48, 60 and 61 in CAZY database (www.cazy.org) (Henrissat 1993 and 1997). Glycoside hydrolase family 5 includes cellulases, exo-β-glucanases, mannanases, and endoceramidases etc. (Davies *et al.,* 1998). The homology modeled structure of Cel5A (endoglucanse) and its mutants predict structure of GH5A at N-terminus as catalytic module and CBM3 module at C-terminus folded in a typical $(\alpha/\beta)_8$ TIM-like barrel motif (Lin *et al.,* 2009). Molecular dynamic simulation studies predict the relationship of protein flexibility with its activity and stability. The flexibility of a protein plays an important role in their folded state to accomplish their function such as activity modulation, ligand binding and macromolecular interactions (Yu *et al.,* 2015). Molecular docking analysis predicts the ligand-receptor complex structure in the protein tertiary structure via computation methods (Meng *et al.,* 2011). *Sorghum durra* is a widely grown crop in India along with other major cash crops. Its availability throughout the year makes the *Sorghum durra* stalk as a potential feedstock for bioethanol production. Therefore, pretreated *Sorghum*

TH-2393_146106034

durra stalk was explored as a substrate for saccharification with cellulosic enzymes for bioethanol production (Jamaldheen *et al.*, 2018).

Present investigation is carried out on Strain improvement of *Bacillus amyloliquefaciens* SS35 for enhancing endoglucanase activity. The mutant enzyme (CMCase-UV2) from developed mutant strain was biochemically characterized and compared with the wild-type enzyme (CMCase-WT). Gene encoding mutant and wild-type enzyme was cloned in expression vector using recombinant DNA technology. Recombinant mutant enzyme (*Ba*GH5-UV2) was biochemically and structurally characterized and compared and wild-type enzyme (*Ba*GH5-WT). Application of *Ba*GH5-UV2 enzyme in saccharification of acid and base pretreated *Sorghum durra* stalk biomass in cocktail with *Ct*CBH5A and *Ct*GH1 using the Stat-Ease Design Expert software was explored which will be used in bioethanol production. This thesis work comprises 6 chapters

Chapter 1 is the "General Introduction" which embodies the brief review of literature on the forces responsible for the requirement of clean fuel such as bioethanol. It describes the advantages of bioethanol over gasoline i.e. it has high octane number, high heat of vaporization, oxygenated fuel and contains negligible amount of sulfur (Lynd *et al.*, 1996; Kar *et al.*, 2006). This chapter elaborately reviewed the major bottleneck in the production of second generation bioethanol from lignocellulosic biomass and their possible solutions to reduce the cost of enzyme production. It also described the carbohydrate-active enzymes, especially GH5 along with their structure that are important plant cell wall degrading enzymes. It also demonstrated the involvement of recombinant bacterial hydrolytic enzymes in saccharification of

TH-2393_146106034

lignocellulosic biomass. This chapter defines the significance of investigation undertaken and then lists the specific objectives.

Chapter 2 describes the UV/EMS directed evolution of *Bacillus amyloliquefaciens* SS35 enhanced the CMCase production in the cell free supernatant. Wild-type strain of *Bacillus amyloliquefaciens* SS35 was mutated in two stages first by UV irradiation and second EMS. In the first stage of UV mutagenesis, 180 min of UV exposure for wild-type strain (*Bacillus amyloliquefaciens* SS35) was sub-lethal resulting 1% of survivors. UV mutant colonies were screened on the basis of plate staining method by 0.3% Congo red and CMCase production. Out of 14 UV mutant strains only 3 strains i.e. UV1, UV2 and UV4 exhibited the significant increase in CMCase activity i.e. 17%, 22% and 14%, respectively; as compared to the wild-type strain (0.65 U/mL). Therefore, these three UV mutant strains were selected for genetic stability of mutation. Among all the three UV mutants, only UV2 mutant strain retained 100% of the relative CMCase activity up to 10 generations which suggested that the mutation is genetically stable. Therefore, UV2 strain was selected for the second stage mutagenesis by using a chemical mutagen EMS. EMS dose, 2% (v/v) was sub-lethal for UV2 mutant strain which results only 1% survivors. EMS mutant colonies were also screened in a similar manner as UV mutant colonies screened. Only 4 EMS mutant colonies showed increase in CMCase activity for crude cell free supernatant as compared with the wild-type strain. EMS 2, EMS 3, EMS 7 and EMS 8 mutant strains displayed 3%, 4%, 7% and 3%, respectively, increase in the CMCase activity as compared to the UV2 strain (0.79 U/mL). Among all four EMS mutants, none retained the relative CMCase activity upto 10 generations suggested that the mutation is not genetically stable. Therefore, on the basis of enzyme production and generation

TH-2393_146106034

stability, UV2 mutant strain was the most potential mutant strain and used for further studies. The fermentation profile of UV2 mutant strain showed that maximum CMCase production occur at 48 h of fermentation in the late log phase. The pH of the fermenting medium increased from acidic to alkaline during the growth of UV2 mutant strain, approximately similar to the wild-type strain.

Chapter 3 showed the comparison of purification and biochemical properties of endoglucanase, CMCase-UV2 from UV mutant strain of *Bacillus amyloliquefaciens* SS35 with the CMCase-WT from wild-type strain. The purification of CMCase-UV2 and CMCase-WT enzyme was performed by ammonium sulfate precipitation method which is followed by size exclusion chromatography. The molecular mass of CMCase-UV2 was similar to the CMCase-WT approximately, 37 kDa as also confirmed by Zymogram analysis. The purified CMCase-UV2 showed 210, 161 and 32 U/mg specific activities against barley β-D-Glucan, lichenan and CMC-Na, respectively, which were 1.6, 1.9 and 2.1-fold higher than CMCase-WT. CMCase-UV2 showed pH stability in wider range (4.0-5.5) and retained 100% of the relative CMCase activity whereas CMCase-WT showed stability only at pH 5.0. CMCase-UV2 activity at pH 5.0 and pH stability in 4.0-5.5 range with temperature stability at 30°C reflected the importance of the mutant enzyme in simultaneous saccharification and fermentation process of bioethanol production from lignocellulosic substrate, because many of the ethanol fermenting microorganisms are mesophilic and acidophilic. TLC analysis showed that the mode of action of CMCase-UV2 and CMCase-WT enzyme was endolytic. Unlike CMCase-WT, CMCase-UV2 after 45 min of enzymatic reaction produced also glucose along with cello-oligosaccharides which further confirmed that UV2 mutant strain produced endoglucanase enzyme with increased catalytic efficiency. The efficiency of

TH-2393_146106034

CMCase-UV2 enzyme was also checked by hydrolyzing lignocellulosic biomass Elephant grass. The maximum TRS yield of 154.2 mg/g of pretreated Elephant grass was obtained by CMCase-UV2, which was 1.8-fold higher than CMCase-WT under unoptimized conditions. These results showed that the CMCase-UV2 has high potential for application in enzymatic hydrolysis of lignocellulosic biomass, over CMCase-WT which is an important step in synthesis of bioethanol.

Chapter 4 describes the amplification of genes encoding family 5 glycoside hydrolase (endoglucanase), *Ba*GH5-WT (CMCase-WT) from wild-type strain and the mutant *Ba*GH5-UV2 (CMCase-UV2) from the mutant UV2 strain of *Bacillus amyloliquefaciens* SS35 to identify the genetic changes in the endoglucanase gene of UV2 strain. Degenerate oligonucleotide primers were designed from phylogenetically related spp. *Bacillus amyloliquefaciens* KHG19 for family 5 glycoside hydrolase (GH5) with GenBank accession number: AJK65578.1 from the CAZY database (http://www.cazy.org). The PCR amplified fragments of the genes encoding *Ba*GH5-WT and *Ba*GH5-UV2 showed size 1.5 kb. The genes encoding *Ba*GH5-WT and *Ba*GH5-UV2 proteins were cloned into the pHTP0 cloning vector using NZYEasy cloning kit. The gene sequence alignment of *Ba*GH5-UV2 with *Ba*GH5-WT showed transition mutation at position 698 bp where adenine base was substituted by guanine base. Due to this point mutation in *Ba*GH5-UV2, codon GAC changed to GGC and thus, changed the amino acid residue aspartic acid to glycine. The analysis of *Ba*GH5-WT and *Ba*GH5-UV2 gene sequences in NEBcutter V2.0 showed zero cutter for restriction sites i.e. *Nhe*I and *Xho*I. Thus, with these above restriction sites both the genes were cloned into pET-28a(+) expression vectors for their over-expression. The protein sequence of *Ba*GH5-WT and *Ba*GH5-UV2 showed the molecular architecture

TH-2393_146106034

belonging to the GH5 family. It displayed an N-terminal family 1 glycoside hydrolase (448 amino acid), family 5 glycoside hydrolase (*Ba*GH5, 247 amino acid) followed by family 3 carbohydrate binding module *Ba*CBM3 (82 amino acid) at the C-terminal. The sequence alignment of *Ba*GH5-UV2 and *Ba*GH5-WT displayed the substitution mutation by changing the aspartate (D256) residue to glycine (G256) in the GH5 module. Both the recombinant proteins, *Ba*GH5-WT and *Ba*GH5-UV2 expressed as soluble proteins by *E. coli* BL21 (DE3)pLysS cells at 24°C. Both, the purified *Ba*GH5-WT and *Ba*GH5-UV2 proteins showed homogenous bands of the same molecular size approximately, 58 kDa. The concentrations of purified *Ba*GH5-WT and *Ba*GH5-UV2 proteins were 0.19 mg/mL and 0.18 mg/mL, respectively.

Chapter 5 describes the comparative biochemical characterization of recombinant *Ba*GH5-WT and *Ba*GH5-UV2. The mutant endoglucanase *Ba*GH5-UV2 showed 9.7-fold increase in specific activity against carboxymethyl cellulose sodium salt (CMC-Na) than wild-type *Ba*GH5-WT (4.5 U/mg). Mutant *Ba*GH5-UV2 showed maximum specific activity at temperature 65°C than wild-type *Ba*GH5-WT at 55°C. Both the enzymes showed similar temperature stability up to 45°C and pH optima at 5.5. The comparison of pH stability analysis for endoglucanase activity showed the stability in the wide acidic pH range (5.0-7.0) for *Ba*GH5-UV2 as compared with the narrow basic pH range (7.0-7.5) for *Ba*GH5-WT. The catalytic efficiency of *Ba*GH5-UV2 was increased to approximately, 18.4 and 21.9-fold against β-Glucan and CMC-Na respectively, as compared with *Ba*GH5-WT. The tertiary structure analysis of *Ba*GH5-WT was carried out by RaptorX prediction tools that showed close homology with 3PZT|A from *Bacillus subtilis*, 2L8A|A from *Bacillus subtilis* and 1E5J|A from *Salipaludibacillus agaradhaerens*. The modeled structure of *Ba*GH5-WT validated by

TH-2393_146106034

Ramachandran plot showed approximately, 99.8% residues in allowed region. The tertiary structure of *Ba*GH5-UV2 enzyme was generated after mutating the Asp256 to Gly by using Swiss PDB viewer software. 3-D structure validation of *Ba*GH5-UV2 enzyme by Ramachandran plot also showed approximately, 99.9% residues in the allowed region. The overall modular structures of *Ba*GH5-WT and *Ba*GH5-UV2 showed a classical $(\alpha/\beta)_8$-TIM barrel fold. The modeled structure showed that the mutation occurred in the loop region connecting β_6-sheet with α_6-helix, near the active site cavity at the surface regions of the domain. Molecular Dynamics (MD) simulation studies for 80 ns showed the decreased radius of gyration for *Ba*GH5-UV2 enzyme, suggesting the increase in the compactness of the 3-D structure thus the increased globularity of protein. Root mean square fluctuation for the catalytic residues in *Ba*GH5-UV2 enzyme was decreased showing more rigid catalytic residues. The less fluctuation observed in the SASA of *Ba*GH5-UV2 suggested that the accessibility of the substrate in the catalytic sites was higher as compared with *Ba*GH5-WT. The molecular docking studies of MD simulated structure showed that the subsitution of Asp256 to Gly residue in *Ba*GH5-UV2 showed 1.6-fold and 2.6-fold change in free energy (ΔG°) against cellotriose and cellotetraose, respectively, as compared with *Ba*GH5-WT. These results provide information on the possibility of designing a mutant enzymes of members from glycoside hydrolase family 5 to improving their catalytic efficiency and pH stability by protein engineering.

Chapter 6 elaborates the application of crude endoglucanase *Ba*GH5-UV2 enzyme for saccharification of 1% (v/v) H_2SO_4 (acid) assisted autoclaving and 1% (w/v) NaOH (base) assisted autoclaving pretreated *Sorghum durra* stalk biomass in cocktail with crude *Ct*CBH5A (cellobiohydrolase) and crude *Ct*GH1 (β-glucosidase) at 40°C for

TH-2393_146106034

48 h using the Stat-Ease Design Expert software. The simplex centroid mixture design was carried out to optimize the enzymatic ratio. The response was evaluated through saccharification reaction assays and validated. The optimized enzyme cocktail for base pretreated *Sorghum durra* stalk composed of *Ba*GH5-UV2 fraction (80%), *Ct*CBH5A (10%) and *Ct*GH1 (10%), with a total enzyme loading of 2000 U/g of pretreated biomass producing total reducing sugar (TRS) yield of 165 mg/g of pretreated biomass and glucose yield of 81 mg/g of pretreated biomass at 48 h of saccharification. However, the optimized enzyme cocktail for acid pretreated *Sorghum durra* stalk composed of *Ba*GH5-UV2 fraction (43.1%), *Ct*CBH5A (10%) and *Ct*GH1 (46.9%), with a total enzyme loading of 2000 U/g of pretreated biomass giving TRS yield of 91 mg/g of pretreated biomass and glucose yield of 69 mg/g of pretreated biomass at 48 h of saccharification. The saccharification of acid or base pretreated *Sorghum durra* stalk for optimized enzyme ratios were carried out at 40°C for the different time interval of 24 h, 48 h, 72 h and 96 h, and enzymatic hydrolysate were analyzed for TRS and glucose yield. The base pretreated *Sorghum durra* stalk gave TRS yield 59, 165, 193 and 211 mg/g of pretreated biomass at 24 h, 48 h, 72 h and 96 h of saccharification with a glucose yield of 26, 81, 159 and 174 mg/g of pretreated biomass, respectively. However, the acid pretreated *Sorghum durra* stalk gave TRS yield 38, 91, 112 and 119 mg/g of pretreated biomass at 24 h, 48 h, 72 h and 96 h of saccharification with a glucose yield of 10, 69, 84 and 85 mg/g of pretreated biomass, respectively. Thus, a rational enzyme mixture designed by using the synergistic concept and statistical analysis was capable of improving the biomass saccharification.

TH-2393_146106034

CONTENTS

TH-2393_146106034

TH-2393_146106034

TH-2393_146106034

TH-2393_146106034

TH-2393_146106034

TH-2393_146106034

TH-2393_146106034

Chapter 1

General introduction

1. Importance of biofuels

1.1 Forces for clean fuel production

Environmental safety and energy security are the two most important issues in the current scenario that have heightened the demand for alternative of fossil fuels and eco-friendly biofuels (Vohra *et al.,* 2014). It has been predicted that the increasing consumption rate of fossil fuels will finish the reserves of fossil fuels by 2042 (Shafiee and Topal 2009). The extensive use of fossil fuels result in emission of greenhouse gases that causes global warming, a rise in sea level, climate change, urban pollution and loss of biodiversity (Kundu *et al.,* 2013; Vanhala *et al.,* 2016). Bioethanol is one of the most promising renewable biofuel, which is used as the substitute for fossil fuels. It is produced from the biological organic substance of recent origin, such as agricultural residue and forest residue etc. as compared to the fossil fuels which are produced from the slow geological process (Berner *et al.,* 2003; Zabed *et al.,* 2017). Ethanol (C_2H_5OH) as a biofuel has several advantages over gasoline (C_7H_{17}); 1) The higher octane number of ethanol allows it to be burnt at a higher compression ratio with shorter burning time, resulting in a lower engine

TH-2393_146106034

knocking, 2) Ethanol has a higher heat of vaporization (840 kJ/kg) than that of gasoline (305 kJ/kg) which ensures that the volumetric efficiency of ethanol blending is higher than the efficiency of pure gasoline, thereby improving power output, 3) Bioethanol is an oxygenated fuel and contains 34.7% of oxygen, as compared to the gasoline where oxygen is absent. This results nearly 15% higher combustion efficiency of ethanol than gasoline, thereby decreasing the emission of particulate and nitrogen oxides, 4) Ethanol contains negligible amounts of sulfur as compared with the gasoline, therefore, mixing of these two fuels will decrease the sulfur content in fuel as well as the emission of acid rain causing sulfur oxide (Lynd *et al.*, 1996; Kar *et al.*, 2006; Nigam and Singh 2011). Bioethanol is used as a fuel in cars, either in a neat form (E100 i.e. 100% ethanol with or without a fuel additive) or in a mixture with gasoline (E85 i.e. 85% ethanol and 15% gasoline) (Ahmed *et al.*, 2001).

On the basis of the feedstock and conversion technology, production of bioethanol is usually categorized into first, second and third generation (Sharma *et al.*, 2019). First generation bioethanol is produced from bioenergy crops such as sugarcane, corn, maize, rice, soybean and wheat etc. Brazil was the first country to successfully produce bioethanol in 1975 from sugarcane (*Saccharu*m spp.). In 1978 bioethanol production was started in US from corn as the main substrate (Bothast *et al.*, 2005). According to the United States Department of Economic and Social Affairs, world population will become around 9.6 billion in 2050 and offers direct competition for the food crops to be used as the substrate in bioethanol production (Limayem *et al.*, 2012). Therefore, the cost of first generation bioethanol was increased upto 40-70%. Second generation bioethanol is produced from the lignocellulosic biomass such as agricultural and forestry residues. Moreover, the

TH-2393_146106034

second generation bioethanol has advantage over first generation bioethanol that it does not offer any competition with food or animal feed. Third generation bioethanol is produced from aquatic feedstock (algae) (Demirbas *et al.*, 2010).

Brazil in 2004-2005 became world's largest producer of bioethanol (4.5 billion gallons). According to the agreement by Policy Energy Act (PEA) and the Energy Independence and Security Act (EISA) bioethanol production is estimated to be 36 billion gallons by the year 2022 (Ray *et al.*, 2013). The worldwide production of bioethanol in 2006 was 49 billion liters per year (Licht, 2006) which increased to 117.7 million liters in 2016 (Donato *et al.*, 2019).

1.2 Lignocellulosic biomass

Lignocellulosic biomass is the major renewable source of energy used as the feedstock in bioethanol production (Akhtar *et al.*, 2019). It is easily available in the form of low-cost raw materials such as agricultural and plant residues. Agricultural and plant residues serve as the most auspicious and abundant cellulosic feedstock's because of its whole year decentralized availability, low market value, less prone to microbial contamination and high cellulose content (Cardona *et al.*, 2007). Lignocellulosic biomass contains approximately, 40-60% cellulose, 20-40% hemicellulose and 10-24% lignin (Fig. 1.1) (Putro *et al.*, 2016). Cellulose and hemicellulose are the sources of monomeric sugars such as glucose and xylose, which upon fermentation form bioethanol (Limayem *et al.*, 2012). Major lignocellulosic feedstocks which are used in the production of bioethanol are shown in Table 1.1.

TH-2393_146106034

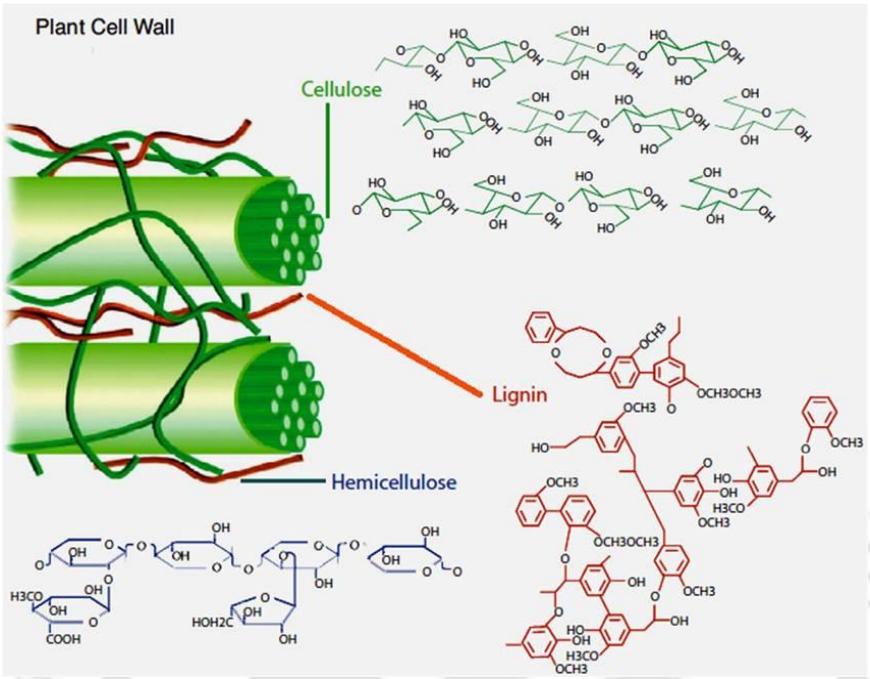


Fig. 1.1 Structural component of lignocellulosic biomass (Alonso *et al.*, 2012).

TH-2393_146106034

Table 1.1 Composition of various lignocellulosic biomass used in bioethanol production

Substrate	Composition (%, w/w)			References
	Cellulose	Hemicellulose	Lignin	
Sugarcane bagasse	40-43	27-38	10-20	Martin *et al.*, 2007
Lantana camera	44	17	32	Kuhad *et al.*, 2010
Hardwood	45	30	20	Balat 2011
Softwood	42	27	28	Balat 2011
Wheat straw	37-41	27-32	13-15	Balat 2011
Grasses	25-40	25-50	10-30	Balat 2011
Bamboo leafy biomass	36	33	27	Das *et al.*, 2012
Jamun leafy biomass	42	30	25	Das *et al.*, 2012
Ashoka leafy biomass	34	47	17	Das *et al.*, 2012
Rice straw	31	22	13	Mood *et al.*, 2013
Barley straw	33-37	22-25	14-16	Mood *et al.*, 2013
Poplar	43	14	29	Mood *et al.*, 2013
Switch grass	39	25	17	Mood *et al.*, 2013
Corn stover	38	25	17	Mood *et al.*, 2013
Poplar leafy biomass	29	42	19	Gupta *et al.*, 2014
Elephant grass	23	21	19	Eliana *et al.*, 2014
Parthenium	26	18	26	Singh *et al.*, 2014
Mango leafy biomass	26	54	17	Das *et al.*, 2014
Newspaper	40-55	25-40	2-30	Saini *et al.*, 2015
Water hyacinth	31	45	19	Das *et al.*, 2016
Forest residues	46	28	26	Gaurav *et al.*, 2017
Sorghum durra stalk	32	23	6	Jamaldheen *et al.*, 2018
Barley husks	22	27	19	Sadhukhan *et al.*, 2019
Sweet *Sorghum* stalk	32	26	5	Thomas *et al.*, 2019
Sorghum durra stem	22	23	2	Thomas *et al.*, 2019

TH-2393_146106034

1.2.1 Cellulose

Cellulose is the most abundant biomolecule on the earth's surface. It is a homopolymer of D-glucose units linked by the β (1→4) glycosidic bond (Fig. 1.2) (Updegraff *et al.*, 1969; Himmel *et al.*, 2007). It is the major structural constituent of plant primary cell wall. Plant wood contains approximately, 40-50% (w/w) of cellulose. Some bacterial species such as *Acetobacter xylinum* and *Dictyostelium discoideum* produce cellulose in the form of biofilms (Goller *et al.*, 2008). The β-(1→4) glycosidic bond in cellulose chain forms linear chain of glucose units and shows stereochemistry, where each glucose residue is rotated to 180° with respect to its neighbouring residue (Srivastava 2002). Numerous cellulose chains are organized together to form a paracrystalline or crystalline lattice, where it is stabilized by the intermolecular (interchain) and intramolecular (intrachain) hydrogen bonds (Srivastava 2002). Intermolecular hydrogen bonds are formed between oxygen atoms and hydroxyl groups on the adjacent chains (Srivastava 2002). Intramolecular hydrogen bonds are formed between the oxygen of one glucose residue and hydroxyl group of C3 residue (Srivastava 2002). Moreover, the van der Waal's forces provide extra stability to the cellulose polymer (Srivastava 2002). Several linear chains of glucose units in cellulose are bonded together and lie against each other to form the cellulose microfibril. These hydrogen bonds and van der Waal's forces in cellulose polymer are responsible for the strength, remarkable inertness and durability of the cellulose (Srivastava 2002).

TH-2393_146106034

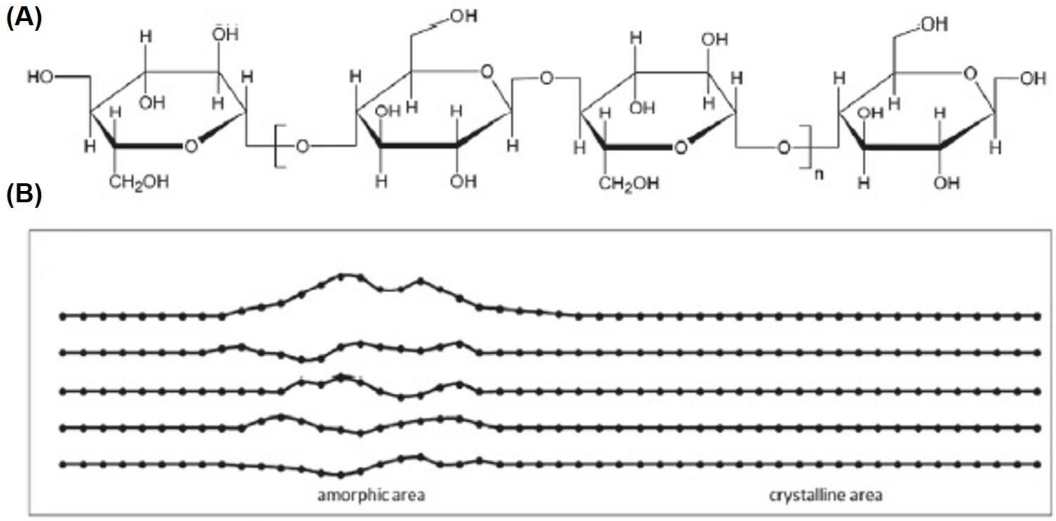


Fig. 1.2 Structure of cellulose. (A) Single chain, (B) fibres (Jaworska and Vogt 2013).

1.2.2 Hemicelluloses

Hemicelluloses are the second most abundant polysaccharides on earth which form approximately, 20-35% of lignocellulosic biomass (Saha 2003). Unlike cellulose, the hemicelluloses are the heterogeneous polysaccharides of hexoses (glucose, galactose, mannose), pentoses (xylose, arabinose) and sugar acids (McMillan 1994). Hemicelluloses in hardwood are mostly xylans whereas in softwood are mostly glucomannans. Xylan is mostly present as heteropolysaccharides with homopolymeric backbone of β-(1→4) xylopyranose units. Moreover, the xylans may also contain the glucuronic acid, arabinose, acetic acid, *p*-coumaric acid and ferulic acid. Xylan is a linear or branched polysaccharide and the frequency of branching depends on its source. Therefore, xylans are characterized as linear homoxylans, glucuronoxylan, arabinoxylan and glucuronoarabinoxylan (Aspinall *et al.,* 1980). The bioconversion of hemicellulosic polysaccharides has gained much attention in different agro-industrial processes, such as in biofuels, paper pulp

TH-2393_146106034

delignification, enhancement of digestibility of animal feedstock, clarification of fruit juices and improvement of beer consistency (Saha 2003).

1.2.3 Lignin

Lignin is the most abundant aromatic polymer in nature which is the essential part of plant lignocellulosic materials whose function is to cement the cellulose fibers (Lora and Glasser, 2002). The percentage of lignin in plants is in a range of approximately, 10-30% depending on the type and origin of plants. Based on the composition of lignin in structural units of lignocellulosic materials, it is divided into three categories such as hardwood lignin, softwood lignin and grass lignin (Pandey and Kim 2011). Hardwood lignin also known as dicotyledonous angiosperm lignin, is made up of sinapyl and coniferyl alcohol units (Carrott *et al.,* 2007). Softwood lignin also known as guaiacyl or coniferous lignin and it is made up of coniferyl or guaiacyl lignin alcohol units (Fig. 1.3). The grass lignin is made up of *p*-coumaryl, sinapyl and coniferyl units (Carrott *et al.,* 2007). Softwood plants contain lignin in a range of 25-31% (w/w) and hardwood plants contain lignin in a range of 16-24% (w/w) (Carrott *et al.,* 2007). Lignin polymer offers the major obstacle in the conversion of polysaccharide into fermentable sugars in biofuel industry (Mukherjee *et al.,* 2016).

TH-2393_146106034

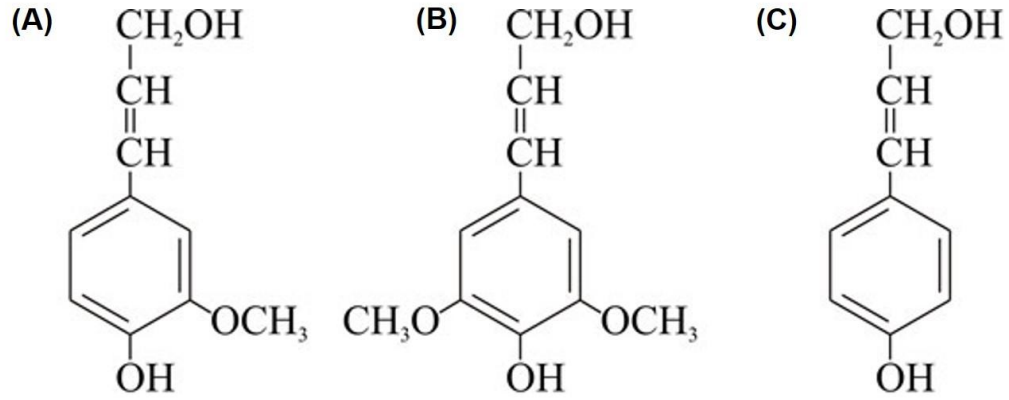


Fig. 1.3 Lignin precursors: (A) coniferyl alcohol, (B) sinapyl alcohol and (C) *p*-coumaryl alcohol (Moore *et al.,* 2011).

1.3 Cellulases

Cellulases are cellulose hydrolyzing enzymes which are synthesized by the large group of microorganisms either extracellularly or cell bound during their growth. Cellulases are classified into three major categories, 1) endoglucanase (EC 3.2.1.4), 2) cellobiohydrolase (EC 3.2.1.91) and 3) β-glucosidase (EC 3.2.1.21) (Singh *et al.,* 2013). These three enzymes act synergistically in a sequence to assist the hydrolysis of cellulose into glucose. Endoglucanase randomly acts on the internal β-1,4 bonds in the amorphous region of cellulose chain to form cello-oligosaccharides. Cellobiohydrolase acts on cello-oligosaccharides from ends to release cellobiose and β-glucosidase hydrolyzes the cellobiose into two molecules of glucose (Shahzadi *et al.,* 2014). Many bacterial, fungal and actinomycetes sp. are the efficient producer of cellulase (Singh *et al.,* 2013) shown in Table 1.2. Among several cellulolytic bacteria *Bacillus* sp. are most preferred because of their diverse range of cellulases stable at extreme environments (Dias *et al.,* 2014 and Patagundi *et al.,* 2014). Few actinomycetes are the potent cellulase producers are shown in Table 1.2. Numerous strain of cellulase producing fungal sp. are given in Table 1.3. Bacterial

TH-2393_146106034

cellulases are getting much more attention than fungal cellulases because, first of all, the production cost of fungal cellulase is higher than bacterial cellulase (Coughlan *et al.*, 1985, Akhtar *et al.*, 2016). Secondly, the growth rate of bacteria is much higher than fungi, it attains high cell density within a short time to produce enzyme. Thirdly, certain bacterial cellulases are expressed in multiple complexes which increase the performance of overall saccharification as they function in synergy. Fourthly, bacteria can inhabit a wide diversity of industrial and ecological niches, because of which they are resistant to environmental stresses. Fifthly, the expression system and genetic manipulation of bacteria are more convenient to achieve the high-level enzyme. Their recombinant enzymes can be hyper-expressed in *E. coli* cells and enhanced cellulase production can be achieved as compared with fungal cellulases, which have RNA with introns and glycosylated proteins (Coughlan *et al.*, 1985).

TH-2393_146106034

Table 1.2 Cellulase producing bacterial and actinomycetes sp.

Microorganism	Type	Species	References
Bacteria	Gram Positive	*Clostridium thermocellum*	Dumitrache *et al.*, 2013
		Bacillus sp.	Deka *et al.*, 2013
		Bacillus amyloliquefaciens SS35	Singh *et al.*, 2013
		Staphylococcus sp.	Ozkan *et al.*, 2016
		Bacillus tequilensis G9	Aslam *et al.*, 2019
		Bacillus cereus	Tabssum *et al.*, 2018
		Bacillus spp.	Ahmad *et al.*, 2019
		Bacillus amyloliquefaciens	Dar *et al.*, 2019
		Bacillus velezensis	Liu *et al.*, 2019
		Ruminococcus sp.	Boonsaen *et al.*, 2019
		Cellulomonas sp.	Kim *et al.*, 2019
		Streptomyces sp.	Saini *et al.*, 2015
		Microbispora bispora	Saini *et al.*, 2015
		Thermomonospora fusca	Saini *et al.*, 2015
		Thermonomospora curvata	Saini *et al.*, 2015
		Sporocytophaga myxococcoides	Leadbetter *et al.*, 2015
Actinomycetes	Gram Negative	*Fibrobacter succinogenes*	Burnet *et al.*, 2015
		Proteus	Osho *et al.*, 2017
		Pseudomonas sp.	Goel *et al.*, 2019
		Serratia	Leo *et al.*, 2019
		Acetivibrio cellulolyticus	Rashid *et al.*, 2019
	Gram Positive	*Streptomyces* sp.	Saini *et al.*, 2015
		Microbispora bispora	Saini *et al.*, 2015
		Thermomonospora fusca	Saini *et al.*, 2015
		Thermonomospora curvata	Saini *et al.*, 2015
		Sporocytophaga myxococcoides	Leadbetter *et al.*, 2015
		Streptomyces sp.	Saini *et al.*, 2015
		Microbispora bispora	Saini *et al.*, 2015

Table 1.3 Cellulase producing fungal sp.

Microorganism	Species	References
Fungi	*Aspergillus terreus*	Sohail *et al.*, 2016
	Trichoderma reesie	Li *et al.*, 2016
	Neocallimastix patriciarum	Kwon *et al.*, 2016
	Penicillium citriviride	Mesa *et al.*, 2016
	Penicillium iriensis	Mesa *et al.*, 2016
	Alternaria alternata	Mesa *et al.*, 2016
	Penicillium funiculosum	Mesa *et al.*, 2016
	Verticillium tricorpus	Mesa *et al.*, 2016
	Trichoderma viride	El Baz *et al.*, 2018

TH-2393_146106034

1.4 Major bottleneck in the production of second generation biofuel from lignocellulosic biomass and possible solutions

The conversion of lignocellulosic biomass into the fermentable sugars and the high production cost of cellulase enzymes are the two major bottleneck for the bioethanol industry. Therefore, efforts are required to overcome the high production cost of cellulase enzyme, low enzyme yield and low enzyme activities (Kim *et al.*, 1987). Various ways to decrease the production cost of cellulase enzymes are:

1.4.1 Isolation of new efficient strains

Isolation and screening of numerous culturable cellulase producing bacterial strains was carried out from different sources such as decaying plant material, forestry or agricultural waste, composting heaps, organic matter, the dung of ruminants, soil and extreme environments like hot-springs (Doi *et al.*, 2008). The isolation of cellulase producing bacterial strains was carried out by efficient plate staining methods. The screening and selection of cellulase activity in microbial isolates (bacteria/fungi) are usually carried out on carboxymethylcellulose sodium salt (CMC-Na) supplemented medium containing pertiplates (Hankin *et al.*, 1977). However, plate screening methods are generally restricted to culturable cellulase producing bacteria and therefore, the whole cellulase potential of the site (non-culturable and culturable microbes) is not examined completely. Earth is rich in microbial biodiversity, but only 1% of the total microorganism present in soil have been cultured with traditional media (Hamaki *et al.*, 2005). Researchers are paying attention to the classification and exploitation of cellulase genes from non-culturable microorganisms found in extreme ecotype in search of novel enzymes. The isolated strain may have specific applications in the biorefining industry owing to its greater

TH-2393_146106034

resistance against harsh environmental surroundings. Nowadays, metagenomic library is used to functionally screen the specific microorganisms with specific characteristics. It is a rapid and highly efficient method to screen a broad population of cellulase producing bacteria from the pulp and paper mill effluent sediments (Xu *et al.*, 2006) and rumen of buffalo (Duan *et al.*, 2009).

1.4.2 Improvement of the biochemical properties of cellulases

Although the wide range of cellulases was being isolated but among them, not a single enzyme is suitable for the complete hydrolysis of lignocellulosic biomass at industrial level (Maki *et al.*, 2009). However, these cellulase enzymes offer excellent initial point for the improvement of the biochemical properties of cellulases and in steps towards improving the whole economics of bioethanol production. Mutagenesis of cellulase enzymes provides a significant route in the improvement of biochemical properties of these enzymes (Chand *et al.*, 2005). It is defined as the change in the genetic information of an organism in a stable manner over the subsequent generations. There are two major strategies for the improvement of a cellulase enzymes: 1) random mutagenesis/directed evolution and 2) rational design and protein engineering (Hutchison *et al.*, 1978; Chen and Arnold 1993; Nath *et al.*, 2019).

1.4.2.1 Random mutagenesis

Random mutagenesis is a strategy used in the directed evolution of microorganism. It resembles the natural evolution process (Leemhuis *et al.*, 2009). Mutation enhances the efficiency of some genes through duplication, deletion and substitution mechanism. Therefore, it can be used for the improvement of industrial enzymes (cellulase) for enhancing the catalytic efficiency, thermal and pH tolerance (Leemhuis *et al.*, 2009; Packer *et al.*, 2015). Directed evolution can be done by using

TH-2393_146106034

error-prone rolling circle amplification of a gene (Vu *et al.*, 2012), genome shuffling, random mutations by UV irradiation and chemical mutagen (Leemhuis *et al.*, 2009). Directed evolution with error-prone rolling circle amplification by designing the gene specific primers is helpful in forming the genetic diversity of a gene whose gene or protein sequence and structural information are available. However, UV and chemical mutagenesis have advantage over error-prone rolling circle amplification that it requires no knowledge of the gene sequence and protein structure to produce the genetic diversity with improved characteristics of cellulase (Leemhuis *et al.*, 2009). Random mutagenesis using UV irradiation and chemical agents is an easy tool to achieve genetic and functional modification in an organism (Singh *et al.*, 2017).

1.4.2.1.1 UV mutagenesis

UV radiation is the mutagen which produces the substitution mutation in DNA by transition and transversion mechanism (Witkin *et al.*, 1976). In DNA, UV light cause changes in the position of hydrogen atoms in adenine, guanine, cytosine and thymine base to form tautomer (Fig. 1.4) as reported earlier by Watson and Crick (1953) and Topal and Fresco (1976). An amino ($-NH_2$) group in adenine base is tautomerized to form imino (═NH) group, because of which this tautomerized adenine base is paired with cytosine instead of thymine base. Hence, in the next round of replication this cytosine base, pairs with guanine base, which results AT to GC transition (Fig. 1.4). UV mutation improved the strains of *Trichoderma viride* (Li *et al.*, 2010), *Trichoderma reesei* (Zhang *et al.*, 2017) and *Streptomyces griseoaurantiacus* (Kumar 2015) for enhanced cellulase activity and *Bacillus subtilis* for enhanced β-glucosidase (Agrawal *et al.*, 2013).

TH-2393_146106034

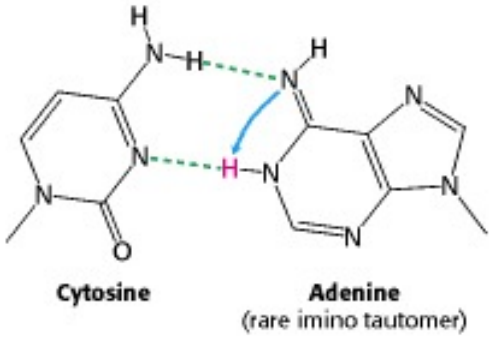


Fig. 1.4 UV-induced imino tautomeric form of adenine base paired with cytosine instead to thymine (Watson 1953 and Topal 1976).

1.4.2.1.2 Chemical mutagenesis by Ethyl methanesulphonate

Ethyl methanesulfonate (EMS) is a mutagen which induces random mutations in DNA by substitution of nucleotide, predominantly by guanine alkylation (Horsfall *et al.,* 1990). It usually causes point mutations in DNA. The ethyl group of EMS reacts with guanine and form O_6-ethylguanine. In DNA replication, in place of cytosine, DNA polymerases catalyze thymine opposite to the O_6-ethylguanine. Hence, in the subsequent round of replication GC to AT transition occurs in DNA (Fig. 1.5). Chandra *et al.,* (2009) mutated *Trichoderma citrinoviride* strain by EMS followed with ethidium bromide treatment. The mutant produced 3.12, IU mL^{-1} endoglucanase with 2-fold higher activity than the wild-type strain. Combined exposure of wild-type strain of *Bacillus subtilis* to EMS and UV mutagens increased β-glucosidase production by 1.2-fold U/mL (Agrawal *et al.,* 2013). Dillon *et al.,* (2006) reported a 1.5-fold increase in cellulase production from *Penicillium echinulatum* by three successive mutagenic treatment steps using UV, EMS and hydrogen peroxide.

TH-2393_146106034

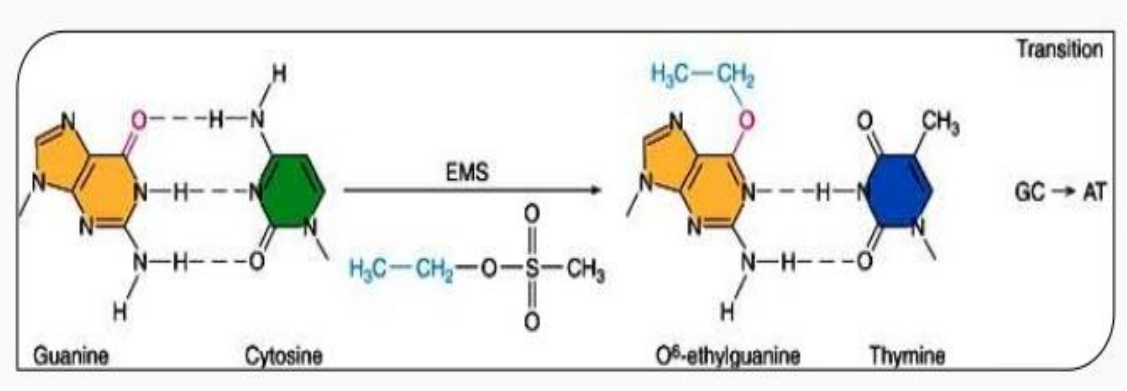


Fig. 1.5 EMS alters a base by guanine alkylation forming O_6-ethylguanine which base pair with thymine instead of cytosine (GC to AT transition) (Gnanamurthy *et al.,* 2014).

1.4.2.2 Rational design/Protein engineering

In rational design, in-depth information of the 3-dimensional structure and function of a protein is required to make changes in amino acid residues for designing the cellulase mutant with desired trait (Zhang *et al.,* 2006). In rational design approach site-directed mutagenesis methods are applied. However, this technique has one major drawback that detailed structural knowledge of most of the proteins is often not available and even if available, it is sometimes difficult to predict the effect of mutations on protein function (Zhang *et al.,* 2006). Here, the researchers examine the feasible amino acid residues to be mutated in the binding pocket of 3-dimensional structure or near the active site of enzyme (Bornscheuer and Pohl, 2001). Site-directed mutagenesis was applied on cellulase enzyme that resulted in significant increase in enzyme activity against insoluble substrates (Escovar-Kousen *et al.,* 2004). Baker *et al.,* (2005) reported that activity of endoglucanase (Cel5A) from *Acidothermus cellulolyticus* increased by 20% against microcrystalline cellulose after mutating Tyr 245 to Gly by site-directed mutagenesis. Vu *et al.,* (2012) reported the enzyme activity of endoglucanase against CMC-Na was increased to 3-fold (2.1 U/mg) by

TH-2393_146106034

Error-Prone Rolling Circle Amplification and to 8-fold (5.5 U/mg) by site-directed mutagenesis. Zhang *et al.,* (2015) reported the 1.8-fold (4.37 U/mL) increase in endoglucanase (Cel12B) activity from *Thermotoga maritima* after site-directed mutagenesis. Nath *et al.,* (2019) reported a 2-fold (40 U/mg) increase in endoglucanase activity of *Ct*GH5 by site-directed mutagenesis of amino acid residue Phy 194 to Ala.

1.4.3 Enhancing the production of cellulase enzyme

In 1993, 50% of the industrial enzyme market was provided by recombinant enzymes using recombinant DNA technology combined with directed evolution to suit the requirements of the users (Hodgson, 1994). Production of natural proteins from microorganism was enhanced by recombinant DNA technology to several folds. The desired gene from particular microbial strain was amplified by designing the degenerate primers from the most similar phylogenetic sp. on the basis of 16S rRNA sequence (Lorenz *et al.,* 2002). The amplified gene was cloned and expressed in particular cloning and expression system (Demain *et al.,* 2009). The wide variety of protein expression systems are available such as cell cultures of bacteria, molds, yeasts, plants, insects, mammals and animals. Protein yield and functionality are the utmost significant factors to consider while selecting the expression system for recombinant protein production. Usually, non-glycosylated proteins are expressed in *E. coli* expression system and N-glycosylated proteins are usually expressed in mammalian cell system (Terpe, 2006). *E. coli* expression system is one of the oldest and widely used hosts for the production of non-glycosylated recombinant proteins (Terpe, 2006). The advantage of this expression system is the ease of culture production, high growth rate and high product yield (Swartz 1996). Moreover, *E. coli*

TH-2393_146106034

genetics is far better understood than those of any other microorganism. This system is used for the mass scale production of commercial proteins due to this high growth rate. The genome of *E. coli* can be easily modified along with plasmid copy number. *E. coli* cells can form recombinant proteins upto 80% of its dry weight and also survive in harsh environmental conditions (Demain *et al.,* 2009). This expression system has few drawbacks, which should be overcome for efficient expression of proteins 1) High cell densities of *E. coli* culture result toxicity to the cells due to acetate formation, which can be overcome by controlling the oxygen level, 2) some recombinant proteins are produced as inclusion bodies that are often insoluble, inactive and require refolding, 3) Proteins with many disulfide bonds (Heterologous proteins) are extremely difficult to express. This problem of can be overcome by the solubilizing the proteins in denaturants which unfold the proteins and reduces the disulfide bonds. Further refolding of a protein is done by the removal of the denaturant/reducing agent which followed by renaturation of the protein in the presence of glutathione reoxidation system (Lavallie *et al.,* 1993).

1.4.4 Engineering of fermenting microbes

Microbes that possess the capability of fermentation at temperatures higher than 50°C can reduce the cost of cellulases by one-half. Currently available fermenting microbes can only meet the desired performance at temperatures, 30-33°C, however, most cellulases act around 40°C-50°C (Wooley *et al.,* 1999). The most advanced method is to develop an organism by genetic engineering in which all biologically mediated steps (e.g., enzyme production, enzymatic cellulose hydrolysis, and biomass sugar fermentation) occur in a single microorganism. This process, also known as

TH-2393_146106034

direct microbial conversion or consolidated bioprocessing can reduce the cost of enzyme production to a significant level (Lynd *et al.*, 1996).

1.5 Different families of cellulase enzyme and their structure

Cellulases are classified as glycoside hydrolase enzymes belonging to families 5-9, 12, 26, 44, 45, 48, 60 and 61 in CAZy database (www.cazy.org) (Schülein *et al.*, 2000; Wang *et al.*, 2010). Glycoside hydrolase family 5 includes endo-β-glucanases, exo-β-glucanases, mannanases and endoceramidases etc. (Davies *et al.*, 1998). Most of the GH5 cellulases are endoglucanase. GH5 enzymes comes under clan A of glycoside hydrolases which are known as 4/7 superfamily. Moreover, the catalytic domain of GH5 enzymes has $(\alpha/\beta)_8$ TIM-barrel fold (Ducros *et al.*, 1995; Santos *et al.*, 2012). Members of this family cleave β-1,4-glycosidic bonds by retaining mechanism through acid-base hydrolysis in which two strictly conserved carboxylates (generally two glutamates) located at the C terminus of β strands 4 and 7, catalyze the reaction (Koshland *et al.*, 1953; Davies *et al.*, 1998). Family GH5 members are divided into 10 subfamilies and have approximately 300-400 amino acid in length which showed 47 Å average diameter in crystal structure (Badieyan *et al.*, 2012). Santos *et al.*, (2012) reported the structural and functional characterization of cellulase, *Bs*Cel5A (Family 5A) from *Bacillus subtilis. Bs*Cel5A consist of two modules, one catalytic module (GH5) and the other CBM3 (family 3 carbohydrate-binding module). Family 6 glycoside hydrolase cellulases contain β/α-barrel with a central β-barrel made up of seven parallel strands. Cel6A (cellobiohydrolase) from *Trichoderma reesei* has Asp221, Asp401 and Asp263 as the catalytic residues whereas Cel6B (endoglucanase) from *H. insolens* has Asp139, Asp316 and Asp92, the catalytic residues (Schülein 2000). Family 7 glycoside hydrolase cellulases

TH-2393_146106034

showed double β-sandwich 3-D topology. Cel7A from *Trichoderma reesei* showed Glu217 and Glu212 catalytic residues. Family 9 cellulases consist of an $(\alpha/\alpha)_6$ helical barrel and family 45 cellulase consists of a six-stranded β-barrel domain (Schülein 2000).

1.6 Homology modeling of a cellulase protein to determine the tertiary structure

Three dimensional shape of a protein is known as the tertiary structure. The tertiary structure of a protein has single polypeptide chain with one or more than one protein domains (Epstein *et al.,* 1963). Many tertiary structures fold to form a quaternary structure. Different online servers and softwares such as Modeller (Baker *et al.,* 2001), EsyPresd3D (Lambert *et al.,* 2002), Robetta (Kim *et al.,* 2004), Lomets (Wu *et al.,* 2007), Swiss-Model (Guex *et al.,* 2009), I-TASSER (Roy *et al.,* 2010), RaptorX (Källberg *et al.,* 2012), Yasara (Balasubramanian *et al.,* 2012) and Phyre2 (Kelley *et al.,* 2015) can be used to predict the tertiary structure of proteins. Swiss-Model is the first online homology protein modeling server on the internet (Guex *et al.,* 2009). Lin *et al.,* (2009) predicted the homology model structure of Cel5A (endoglucanase) and its mutants via an online server Swiss-Model. RaptorX predicts the tertiary structure of proteins which are devoid of close homologs in Protein Data Bank. This server was developed by Xu group (Peng *et al.,* 2010). It is a top server in Critical Assessment of protein Structure Prediction (CASPs) and the fully-automated live benchmark Continuous Automated Model EvaluatiOn (CAMEO, http://raptorx.uchicago.edu/). It models the homology structures of a protein sequence in three steps; 1) template threading, 2) alignment quality assessment and 3) multiple template threading (Källberg *et al.,* 2014). Lugani *et al.,* (2017) construct the tertiary structure of *Bacillus* cellulase by using online server RaptorX

TH-2393_146106034

(http://raptorx.uchicago.edu/). Dadheech *et al.,* (2018) used MODELLER software to design the homology model structure of Cel-1 cellulase from buffalo gut. Liu *et al.,* (2013) predicted the homology protein tertiary structure of GH5 (*Pd*Cel5C) from *Penicillium decumbens* via an online tool (http://pfam.sanger.ac.uk/search).

1.7 Mechanism of action of Glycoside hydrolases

The enzymatic hydrolysis of glycosidic bond was proposed by Koshland (Koshland 1953). This involves normal acid catalysis which requires two important residues: 1) proton donor and 2) nucleophile/base. The hydrolysis of glycosidic bond occurred via two mechanisms i.e. retaining or inverting; and causes the retention or inversion of anomeric configuration, respectively in carbohydrate (Fig. 1.6). In both the retaining and the inverting mechanisms, the position of the proton donor is identical, in other words it is within hydrogen-bonding distance of the glycosidic oxygen. In retaining mechanism, the nucleophilic catalytic base is in close vicinity of the sugar anomeric carbon (Davies *et al.,* 1995; Cairns *et al.,* 2010). This base, however, is more distant in inverting enzymes which must accommodate a water molecule between the base and the sugar. This difference results in an average distance between the two catalytic residues of 5.5 Å in retaining enzymes as opposed to 10 Å in inverting enzymes (McCarter *et al.*, 1994).

TH-2393_146106034

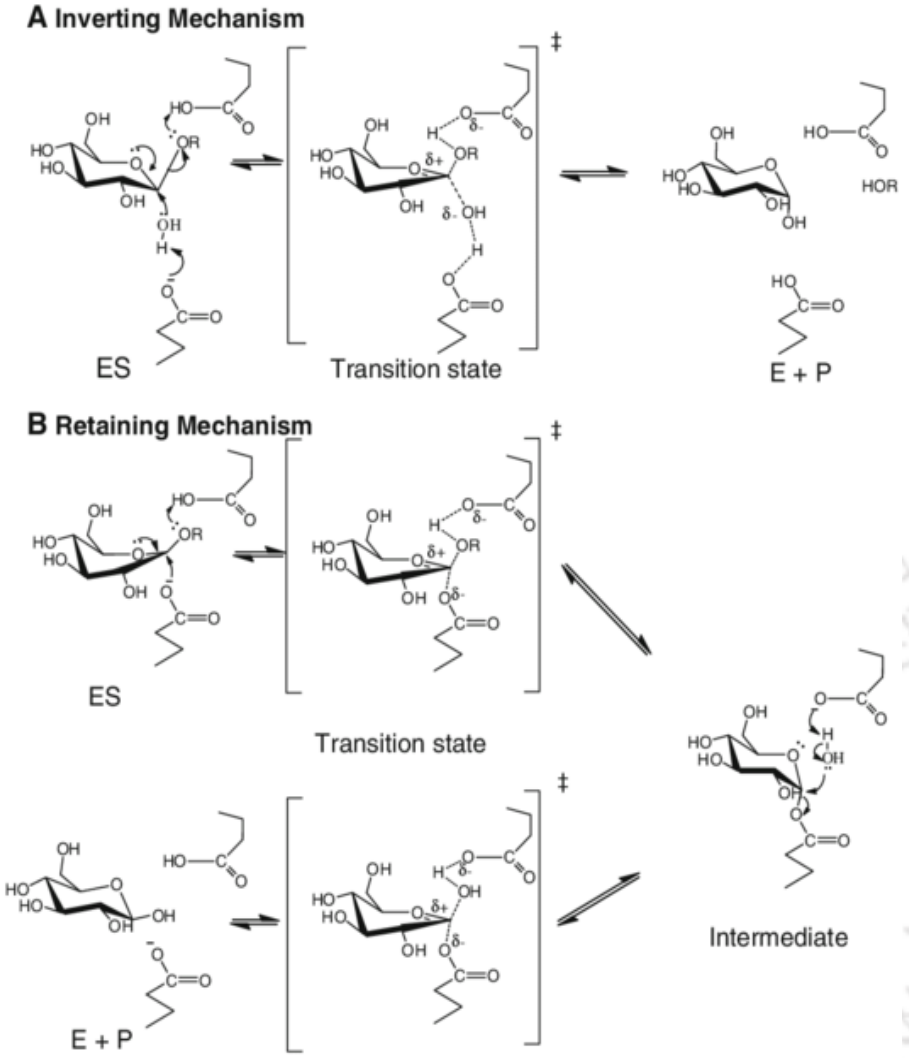


Fig. 1.6 (A) Inverting mechanism (B) Retaining mechanism (Cairns *et al.*, 2010).

1.8 Molecular Dynamic simulation

The relationship of protein flexibility with its activity and stability is quite complex. However, flexibility plays an important role in various proteins in their folded state to accomplish their function such as activity modulation, ligand binding and macromolecular interactions (Yu *et al.*, 2015). Determining the flexibility of a mutant protein would be useful in knowing the role of particular mutated amino acid residues in biological function. Protein molecules possess numerous degrees of flexibility throughout their tertiary structures to perform their function. Some proteins

TH-2393_146106034

generally lose their activity under harsh conditions such as at very low or high temperature, at extreme pHs and high salt concentrations because of the change in protein molecular motion (Smith *et al.*, 2003; Yu *et al.*, 2014). Rigidity in protein is required to maintain the structural integrity in native form, whereas degree of flexibility is required in catalysis. Many computational programs/software are available for the examination of flexibility in protein; 1) atomic positional fluctuations information from B-factors of X-ray crystallography, but there is an important difference between B-factor and solution protein flexibility and 2) Molecular dynamics (MD) simulation is a commonly used method to study solution protein flexibility (Reetz *et al.*, 2007). MD simulation of proteins is performed with computational software, Gromacs (Van Der Spoel *et al.*, 2005). MD simulation studies showed that the mutant of type I coh-Xdoc has a significant increase in flexibility mainly due to the change in hydrogen bonding in the conserved loop region (Xu *et al.*, 2009). Yu *et al.*, (2015) analysed the change in flexibility of a protein luciferase by MD simulation after mutating the amino acid residues Asp to 476 Pro and His 489 to Pro. Niu *et al.*, (2016) performed site-directed mutagenesis of β-glucanase from *Bacillus terquilensis* where 12 lysine residues were substituted by serine residues. In this, MD simulation studies on wild-type and mutant enzyme showed that the rigidity of the β-glucanase got increased with decrease in flexibility because of the overall thermostability of enzyme was increased.

TH-2393_146106034

1.9 Molecular docking analysis of cellulase enzyme

Molecular docking predicts the ligand bound protein complex structure in the protein tertiary structure via computation methods (Meng *et al.,* 2011). Molecular docking is done in two interrelated steps: 1) sampling conformations of ligand into the catalytic cavity of protein; 2) ranking these conformations of ligand according to the scoring function. Algorithm sampling reproduces the experimental binding mode whereas, the scoring function ranks the highest experimental binding mode among all generated conformations. Zheng *et al.,* (2018) analysed the change in hydrogen bond interactions between the mutated residue of *Gt*Cel5 enzyme and its substrate via YASARA software. Mutated residue Asn to 233Ala and Asn to 233 Gly showed extra hydrogen bond formation with cellotetraose complex because of which catalytic efficiency of *Gt*Cel5 enzyme increased. Anbar *et al.*, (2010) reported that substitution of Ser 329 by Gly results loss of hydrogen bond formation with Asp 319 because of which structural flexibility of the protein increases due to which thermostability of Cel8A protein from *Clostridium thermocellum* had increased.

1.10 Application of cellulase enzyme in the saccharification of lignocellulosic biomass

Saccharification of pretreated lignocellulosic biomass into the fermentable sugars (glucose) by cellulase enzymes requires the cocktail combination of three enzymes, i.e. endoglucanase, cellobiohydrolase and β-glucosidase in an appropriate ratio. The rate of enzymatic hydrolysis and saccharification yield from different lignocellulosic biomasses depend on the concentration of enzymes (Marcos *et al.,* 2013). Therefore, the optimization of enzymatic saccharification by using the rational enzyme mixture designed is an essential step in the production of bioethanol from

TH-2393_146106034

lignocellulosic biomass. Bussamra *et al.,* (2012) optimize the enzymatic saccharification for hydrothermally pretreated sugarcane bagasse using a simplex lattice mixture design. Singh *et al.,* (2014) optimized the enzymatic saccharification for acid pretreated delignified *Parthenium hysterophorus* by response surface methodology. Kim *et al.,* (2017) optimize the enzymatic saccharification for acid or alkali pretreated sugarcane bagasse by a rational statistical approach, the mixture design in Minitab 17 software.

1.11 *Sorghum durra* stalk a potential feedstock in bioethanol production

Sorghum is a drought-resistant crop and can grow in adverse conditions of water scarcity, temperature, alkalinity and salinity as compared with sugarcane and maize (Almodares and Hadi, 2009). In India, it is cultivated throughout the year along with the pulse crops with a total production of about 10.62 million tons per year (Jamaldheen *et al.,* 2018). However, only the grain of *Sorghum* is utilized as food and a very little fraction of stalk is utilized as fodder whereas the rest leftover of stalk becomes the agro-waste. *Sorghum durra stalk* has 55% (w/w) holocellulose which makes it best fit as a feedstock for the production of bioethanol. The saccharification of 1% (w/v) NaOH assisted autoclaving pretreated *Sorghum durra* stalk with the combination of endoglucanase (*Ct*GH8) and β-glucosidase (*Ct*GH1) gave a glucose yield of 34 mg/g of raw biomass (Jamaldheen *et al.,* 2018).

TH-2393_146106034

1.12 Objectives of the present study

The bioconversion of lignocellulosic biomass into the fermentable sugars is challenging because of its structural and compositional factors, which hinders the enzymatic saccharification of cellulose fraction. Moreover, only the cellulase production accounts for approximately, 40% of the total cost for bioethanol process. Therefore, significant efforts are required to overcome the challenges of the high production cost of cellulase enzyme. Random mutagenesis of cellulase producing bacterial strains provides a better route for the directed evolution of microorganism with improved of biochemical properties. The change in the structure of mutated protein provides the role of particular amino acid in catalysis and thus help in designing the mutant enzyme by protein engineering for improved catalytic efficiency, temperature stability and pH stability. Among several cellulolytic bacteria *Bacillus* spp. is most preferred because of its diverse range of cellulases which are stable under extreme environments.

In the proposed study, random mutagenesis of *Bacillus amyloliquefaciens* SS35 will be carried out by UV irradiation and EMS. The comparison of biochemical and functional properties of CMCase-WT and CMCase-UV2 enzymes from wild-type and UV2 mutant strain of *Bacillus amyloliquefaciens* SS35, respectively, becomes necessary for elucidating the changes in the endoglucanase activity and the mode of enzyme action. In order to identify the change in amino acid residues, the gene(s) encoding *Ba*GH5-WT (CMCase-WT) from wild-type and *Ba*GH5-UV2 (CMCase-UV2) from UV2 mutant strain of *Bacillus amyloliquefaciens* SS35 will be cloned. The structural, MD simulation and docking studies will be carried out to identify the role of altered amino acid residue(s). The application of *Ba*GH5-UV2 on

TH-2393_146106034

saccharification of previously optimized pretreatment of *Sorghum durra* stalk will be carried out. The following five specific objectives are defined for the proposed thesis work correspond to each of Chapters 2-6.

1.12.1 Specific Objectives

1. Strain improvement of *Bacillus amyloliquefaciens* SS35 by UV and EMS-directed evolution for efficient endoglucanase production.
2. Comparison of purification and biochemical characterization of endoglucanase from UV mutant UV2 with wild-type enzyme from *Bacillus amyloliquefaciens* SS35.
3. Identification of mutation at genetic level and cloning the genes encoding endoglucanase from wild-type *Bacillus amyloliquefaciens* SS35 and its UV2 mutant and their expression and purification.
4. Structure, function and biochemical properties comparison of recombinant *Ba*GH5-WT and *Ba*GH5-UV2.
5. Application of mutant enzyme, *Ba*GH5-UV2 in saccharification of *Sorghum durra* biomass.

TH-2393_146106034

1.13 References

Agrawal, R., Satlewal, A., and Verma, A.K. (2013). Development of a β-glucosidase hyperproducing mutant by combined chemical and UV mutagenesis. *3 Biotech*, 3(5), 381-388.

Ahmad, R.A.Z.A., Bashir, S., and Tabassum, R. (2019). Evaluation of cellulases and xylanases production from *Bacillus* spp. isolated from buffalo digestive system. *Kafkas Universitesi Veteriner Fakultesi Dergisi*, 25(1), 39-46.

Ahmed, I. (2001). Oxygenated diesel: emissions and performance characteristics of ethanol-diesel blends in CI engines. *SAE International*, ISSN 0148-7191, 1-8.

Akhtar, N., Aanchal, Goyal, D., and Goyal, A. (2016). Biodiversity of cellulase producing bacteria and their applications. *Cellulose Chemistry and Technology*, 50(9-10), 983-995.

Akhtar, N., Gupta, K., Goyal, D., and Goyal, A. (2019). Lignocellulosic biomass characteristics for bioenergy applications: An overview. *Environmental Engineering and Management Journal*. 18(2), 367-383.

Almodares, A., and Hadi, M.R. (2009). Production of bioethanol from sweet sorghum: A review. *African Journal of Agricultural Research*, 4(9), 772-780.

Alonso, D.M., Wettstein, S.G., and Dumesic, J.A. (2012). Bimetallic catalysts for upgrading of biomass to fuels and chemicals. *Chemical Society Reviews*, 41(24), 8075-8098.

Anbar, M., Lamed, R., and Bayer, E.A. (2010). Thermostability enhancement of *Clostridium thermocellum* cellulosomal endoglucanase Cel8A by a single glycine substitution. *ChemCatChem*, 2(8), 997-1003.

TH-2393_146106034

Aslam, S., Hussain, A., and Qazi, J.I. (2019). Production of cellulase by *Bacillus amyloliquefaciens*-ASK11 under high chromium stress. *Waste and Biomass valorization*, 10(1), 53-61.

Aspinall, G.O. (1980). Chemistry of cell wall polysaccharides. In carbohydrates: structure and function. *Academic Press*, 3, 473-500.

Badieyan, S., Bevan, D.R., and Zhang, C. (2012). Study and design of stability in GH5 cellulases. *Biotechnology and Bioengineering*, 109(1), 31-44.

Baker, D., and Sali, A. (2001). Protein structure prediction and structural genomics. *Science*, 294(5540), 93-96.

Baker, J.O., McCarley, J.R., Lovett, R., Yu, C.H., Adney, W.S., Rignall, T.R., Vinzant, T.B., and Himmel, M.E. (2005). Catalytically enhanced endocellulase Cel5A from *Acidothermus cellulolyticus*. *Applied Biochemistry and Biotechnology*, 121(1-3), 129-148.

Balasubramanian, A., Das, S., Bora, A., Sarangi, S., and Mandal, A.B. (2012). Comparative analysis of structure and sequences of *Oryza sativa* superoxide dismutase. *American Journal of Plant Sciences*, 3, 1311-1321.

Balat, M. (2011). Production of bioethanol from lignocellulosic materials via the biochemical pathway: a review. *Energy Conversion and Management*, 52(2), 858-875.

Berner, R.A. (2003). The long-term carbon cycle, fossil fuels and atmospheric composition. *Nature*, 426(6964), 323-326.

Boonsaen, P., Poonko, S., Kanjanapruetipong, J., Phiriyangkul, P., and Sawanon, S. (2019). Isolation and partial characterization of *Ruminococcus flavefaciens* from the rumen of swamp buffalo. *Buffalo Bulletin*, 38(2), 311-325.

TH-2393_146106034

Bornscheuer, U.T., and Pohl, M. (2001). Improved biocatalysts by directed evolution and rational protein design. *Current Opinion in Chemical Biology*, 5(2), 137-143.

Bothast, R.J., and Schlicher, M.A. (2005). Biotechnological processes for conversion of corn into ethanol. *Applied Microbiology and Biotechnology*, 67(1), 19-25.

Burnet, M.C., Dohnalkova, A.C., Neumann, A.P., Lipton, M.S., Smith, R.D., Suen, G., and Callister, S.J. (2015). Evaluating models of cellulose degradation by *Fibrobacter succinogenes* S85. *PLoS One*, 10(12), e0143809.

Bussamra, B.C., Freitas, S., and da Costa, A.C. (2015). Improvement on sugar cane bagasse hydrolysis using enzymatic mixture designed cocktail. *Bioresource Technology*, 187, 173-181.

Cairns, J.R.K., and Esen, A. (2010). β-Glucosidases. *Cellular and Molecular Life Sciences*, 67(20), 3389-3405.

Cardona, C.A., and Sanchez, O.J. (2007) Fuel ethanol production: process design trends and integration opportunities. *Bioresource Technology*, 98, 2415-2457.

Carrott, P.J.M., and Carrott, M.R. (2007). Lignin from natural adsorbent to activated carbon: a review. *Bioresource Technology*, 98(12), 2301-2312.

Chand, P., Aruna, A., Maqsood, A.M., and Rao, L.V. (2005). Novel mutation method for increased cellulase production. *Journal of Applied Microbiology*, 98(2), 318-323.

Chandra, M., Kalra, A., Sangwan, N.S., Gaurav, S.S., Darokar, M.P., and Sangwan, R.S. (2009). Development of a mutant of *Trichoderma citrinoviride* for enhanced production of cellulases. *Bioresource Technology*, 100(4), 1659-1662.

TH-2393_146106034

Chen, K., and Arnold, F.H. (1993). Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. Proceedings of the *National Academy of Sciences*, 90(12), 5618-5622.

Coughlan, M.P. (1985). The properties of fungal and bacterial cellulases with comment on their production and application. *Biotechnology and Genetic Engineering Reviews*, 3(1), 39-110.

Dadheech, T., Shah, R., Pandit, R., Hinsu, A., Chauhan, P.S., Jakhesara, S., Kunjadiya, A., Rank, D., and Joshi, C. (2018). Cloning, molecular modeling and characterization of acidic cellulase from buffalo rumen and its applicability in saccharification of lignocellulosic biomass. *International Journal of Biological Macromolecules*, 113, 73-81.

Dar, M.A., Pawar, K.D., Rajput, B.P., Rahi, P., and Pandit, R.S. (2019). Purification of a cellulase from cellulolytic gut bacterium, *Bacillus tequilensis* G9 and its evaluation for valorization of agro-wastes into added value byproducts. *Biocatalysis and Agricultural Biotechnology*, 20, 101219.

Das, S.P., Gupta, A., Das, D., and Goyal, A. (2016). Enhanced bioethanol production from water hyacinth (*Eichhornia crassipes*) by statistical optimization of fermentation process parameters using Taguchi orthogonal array design. *International biodeterioration and biodegradation*, 109, 174-184.

Das, S.P., Ravindran, R., Ahmed, S., Das, D., Goyal, D., Fontes, C.M., and Goyal, A. (2012). Bioethanol production involving recombinant *Clostridium thermocellum* hydrolytic hemicellulase and fermentative microbes. *Applied Biochemistry and Biotechnology*, 167(6), 1475-1488.

TH-2393_146106034

Das, S.P., Ravindran, R., Deka, D., Jawed, M., Das, D., and Goyal, A. (2014). Bioethanol production from leafy biomass of mango (*Mangifera indica*) involving naturally isolated and recombinant enzymes. *Preparative Biochemistry and Biotechnology*, 43(7), 717-734.

Davies, G., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*, 3(9), 853-859.

Davies, G.J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A.M., Schulein, M., and Withers, S.G. (1998). Snapshots along an enzymatic reaction coordinate: analysis of a retaining β-glycoside hydrolase. *Biochemistry*, 37(34), 11707-11713.

Deka, D., Jawed, M., and Goyal, A. (2013). Purification and characterization of an alkaline cellulase produced by *Bacillus subtilis* (AS3). *Preparative Biochemistry and Biotechnology*, 43(3), 256-270.

Demain, A.L., and Vaishnav, P. (2009). Production of recombinant proteins by microbes and higher organisms. *Biotechnology Advances*, 27(3), 297-306.

Demirbas, A. (2010). Use of algae as biofuel sources. *Energy Conversion and Management*, 51(12), 2738-2749.

Dias, P.V.S., Ramos, K.O., Padilha, I.Q.M., Araujob, D.A.M., Santos, S.F.M., and Silva, F.L.H. (2014) Optimization of cellulase production by *Bacillus* sp. isolated from sugarcane cultivated soil. *Chemical Engineering Transactions*, 38, 277-282.

Dillon, A.J., Zorgi, C., Camassola, M., and Henriques, J.A.P. (2006). Use of 2-deoxyglucose in liquid media for the selection of mutant strains of *Penicillium*

TH-2393_146106034

echinulatum producing increased cellulase and β-glucosidase activities. *Applied Microbiology and Biotechnology*, 70(6), 740-746.

Doi, R.H. (2008). Cellulases of mesophilic microorganisms: cellulosome and noncellulosome producers. *Annals of the New York Academy of Sciences*, 1125(1), 267-279.

Donato, P., Finore, I., Poli, A., Nicolaus, B., and Lama, L. (2019). The production of second generation bioethanol: The biotechnology potential of thermophilic bacteria. *Journal of Cleaner Production*, 233, 1410-1417.

Duan, C.J., Xian, L., Zhao, G.C., Feng, Y., Pang, H., Bai, X.L., Ma, Q.S., and Feng, J.X. (2009). Isolation and partial characterization of novel genes encoding acidic cellulases from metagenomes of buffalo rumens. *Journal of Applied Microbiology*, 107(1), 245-256.

Ducros, V., Czjzek, M., Belaich, A., Gaudin, C., Fierobe, H.P., Belaich, J.P., Davies G.J., and Haser, R. (1995). Crystal structure of the catalytic domain of a bacterial cellulase belonging to family 5. *Structure*, 3(9), 939-949.

Dumitrache, A., Wolfaardt, G.M., Allen, D.G., Liss, S.N., and Lynd, L.R. (2013). Tracking the cellulolytic activity of *Clostridium thermocellum* biofilms. *Biotechnology for Biofuels*, 6(1), 175.

El Baz, A.F., Shetaia, Y.M., Shams Eldin, H.A., and ElMekawy, A. (2018). Optimization of cellulase production by *Trichoderma viride* using response surface methodology. *Current Biotechnology*, 7(1), 19-25.

Eliana, C., Jorge, R., Juan, P., and Luis, R. (2014). Effects of the pretreatment method on enzymatic hydrolysis and ethanol fermentability of the cellulosic fraction from elephant grass. *Fuel*, 118, 41-47.

TH-2393_146106034

Epstein, C.J., Goldberger, R.F., and Anfinsen, C.B. (1963). The genetic control of tertiary protein structure: studies with model systems. *Cold Spring Harbor Symposia on Quantitative Biology*, 28, 439-449.

Escovar-Kousen, J.M., Wilson, D., and Irwin, D. (2004). Integration of computer modeling and initial studies of site-directed mutagenesis to improve cellulase activity on Cel9A from *Thermobifida fusca*. *Applied Biochemistry and Biotechnology*, 113(1-3), 287.

Gaurav, N., Sivasankari, S., Kiran, G.S., Ninawe, A., and Selvin, J. (2017). Utilization of bioresources for sustainable biofuels: a review. *Renewable and Sustainable Energy Reviews*, 73, 205-214.

Gnanamurthy, S., and Dhanavel, D. (2014). Effect of EMS on induced morphological mutants and chromosomal variation in Cowpea (*Vigna unguiculata* (L.) Walp). *International Letters of Natural Sciences*, 17, 33-43.

Goel, N., Patra, R., and Verma, S.K. (2019). Purification and characterization of cellulase from *Pseudomonas* sp. isolated from waste dumping site soil. *Journal of Applied Biotechnology and Bioengineering*, 6(3), 118-124.

Goller, C.C., and Romeo, T. (2008). Environmental influences on biofilm development. Bacterial Biofilms, *Springer, Berlin, Heidelberg*, 37-66.

Guex, N., Peitsch, M.C., and Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*, 30(S1), S162-S173.

Gupta, A., Das, S.P., Ghosh, A., Choudhary, R., Das, D., and Goyal, A. (2014). Bioethanol production from hemicellulose rich *Populus nigra* involving

TH-2393_146106034

recombinant hemicellulases from *Clostridium thermocellum*. *Bioresource Technology*, 165, 205-213.

Hamaki, T., Suzuki, M., Fudou, R., Jojima, Y., Kajiura, T., Tabuchi, A., and Shibai, H. (2005). Isolation of novel bacteria and actinomycetes using soil-extract agar medium. *Journal of Bioscience and Bioengineering*, 99(5), 485-492.

Hankin, L., and Anagnostakis, S.L. (1977). Solid media containing carboxymethylcellulose to detect Cx cellulase activity of micro-organisms. *Microbiology*, 98(1), 109-115.

Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W., and Foust, T.D. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuel production. *Science*, 315, 804-807.

Hodgson, J. (1994). The changing bulk biocatalyst market. *Nature Biotechnology*. 12, 789-790.

Horsfall, M.J., Gordon, A.J., Burns, P.A., Zielenska, M., van der Vliet, G.M., and Glickman, B.W. (1990). Mutational specificity of alkylating agents and the influence of DNA repair. *Environmental and Molecular Mutagenesis*, 15(2), 107-122.

Hutchison, C.A., Phillips, S., Edgell, M.H., Gillam, S., Jahnke, P., and Smith, M. (1978). Mutagenesis at a specific position in a DNA sequence. *Journal of Biological Chemistry*, 253(18), 6551-6560.

Jamaldheen, S.B., Sharma, K., Rani, A., Moholkar, V.S., and Goyal, A. (2018). Comparative analysis of pretreatment methods on sorghum (*Sorghum durra*) stalk agrowaste for holocellulose content. *Preparative Biochemistry and Biotechnology*, 48(6), 457-464.

TH-2393_146106034

Jaworska, M., and Vogt, O. (2013). Sorbitol and cellulose derivatives as gelling agents. *Chemik*, 67(3), 242-349.

Källberg, M., Margaryan, G., Wang, S., Ma, J., and Xu, J. (2014). RaptorX server: a resource for template-based protein structure modeling. *Protein Structure Prediction*, 1137, 17-27.

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., and Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. *Nature Protocols*, 7(8), 1511-1522.

Kar, Y., and Deveci, H. (2006). Importance of P-series fuels for flexible-fuel vehicles (FFVs) and alternative fuels. *Energy Sources, Part A*, 28(10), 909-921.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6), 845.

Kim, D.E., Chivian, D., and Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research*, 32(2), W526-W531.

Kim, I.J., Lee, H.J., and Kim, K.H. (2017). Pure enzyme cocktails tailored for the saccharification of sugarcane bagasse pretreated by using different methods. *Process Biochemistry*, 57, 167-174.

Kim, J.M., Kong, I.S., and Yu, J.H. (1987). Molecular cloning of an endoglucanase gene from an alkalophilic *Bacillus* sp. and its expression in *Escherichia coli*. *Applied and Environmental Microbiology*, 53(11), 2656-2659.

TH-2393_146106034

Kim, S.K., Kook, M., Yan, Z.F., Trinh, H., Zheng, S.D., Yang, J.E., Park, S.Y., and Yi, T.H. (2019). *Cellulomonas aurantiaca* sp. nov., isolated from a soil sample from a tangerine field. *Antonie van Leeuwenhoek*, 112, 1623-1632.

Koshland Jr, D.E. (1953). Stereochemistry and the mechanism of enzymatic reactions. *Biological Reviews*, 28(4), 416-436.

Kuhad, R.C., Gupta, R., Khasa, Y.P., and Singh, A. (2010). Bioethanol production from *Lantana camara* (red sage): Pretreatment, saccharification and fermentation. *Bioresource Technology*, 101, 8348-8354.

Kumar, A.K. (2015). UV mutagenesis treatment for improved production of endoglucanase and β-glucosidase from newly isolated thermotolerant actinomycetes, *Streptomyces griseoaurantiacus*. *Bioresources and Bioprocessing*, 2(1), 22-32.

Kundu, A., Sahu, J.N., Redzwan, G., and Hashim, M.A. (2013). An overview of cathode material and catalysts suitable for generating hydrogen in microbial electrolysis cell. *International Journal of Hydrogen Energy*, 38(4), 1745-1757.

Kwon, M., Song, J., Park, H.S., Park, H., and Chang, J. (2016). Characterization of heterologously expressed acetyl xylan esterase1 isolated from the anaerobic rumen fungus *Neocallimastix frontalis* PMA02. *Asian-Australasian Journal of Animal Sciences*, 29(11), 1576-1584.

Lambert, C., Leonard, N., De Bolle, X., and Depiereux, E. (2002). ESyPred3D: Prediction of proteins 3-D structures. *Bioinformatics*, 18(9), 1250-1256.

Lavallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., and McCoy, J.M. (1993). A thioredoxin gene fusion expression system that circumvents

TH-2393_146106034

inclusion body formation in the *E. coli* cytoplasm. *Nature Biotechnology*, 11(2), 187-193.

Leadbetter, E.R. (2015). Sporocytophaga. *Bergey's Manual of Systematics of Archaea and Bacteria*, 17, 1-3.

Leemhuis, H., Kelly, R.M., and Dijkhuizen, L. (2009). Directed evolution of enzymes: library screening strategies. *International Union of Biochemistry and Molecular Biology Life*, 61(3), 222-228.

Leo, V.V., Ramesh, N., and Singh, B.P. (2019). Microorganisms as an Efficient Tool for Cellulase Production: Availability, Diversity, and Efficiency. *New and Future Developments in Microbial Biotechnology and Bioengineering*, 45-61.

Li, F., Wang, C., Li, D., Xu, Z., Yi, Y., Gu, F., Zhu, L., and Guo, W. (2011). Bioinformatics analysis and molecular characteristics of prothymosin alpha encoded by protalhpha gene of giant pandas. *5th International Conference on Bioinformatics and Biomedical Engineering*, 1-7.

Li, X.H., Yang, H.J., Roy, B., Park, E.Y., Jiang, L.J., Wang, D., and Miao, Y.G. (2010). Enhanced cellulase production of the *Trichoderma viride* mutated by microwave and ultraviolet. *Microbiological Research*, 165(3), 190-198.

Li, Y., Liu, C., Bai, F., and Zhao, X. (2016). Overproduction of cellulase by *Trichoderma reesei* RUT C30 through batch-feeding of synthesized low-cost sugar mixture. *Bioresource Technology*, 216, 503-510.

Licht, F.O., 2006. World Ethanol Market: The Outlook to 2015, Tunbridge Wells, Agra Europe Special Report, UK.

TH-2393_146106034

Lim, J., Choi, Y.H., Hurh, B.S., and Lee, I. (2019). Strain improvement of *Aspergillus sojae* for increased l-leucine aminopeptidase and protease production. *Food Science and Biotechnology*, 28(1), 121-128.

Limayem, A., and Ricke, S.C. (2012). Lignocellulosic biomass for bioethanol production: current perspectives, potential issues and future prospects. *Progress in Energy and Combustion Science*, 38(4), 449-467.

Lin, L., Meng, X., Liu, P., Hong, Y., Wu, G., Huang, X., Congcong, Li., Dong, J., Xiao, L., and Liu, Z. (2009). Improved catalytic efficiency of Endo-β-1, 4-glucanase from *Bacillus subtilis* BME-15 by directed evolution. *Applied Microbiology and Biotechnology*, 82(4), 671-679.

Liu, G., Qin, Y., Hu, Y., Gao, M., Peng, S., and Qu, Y. (2013). An endo-1, 4-β-glucanase *Pd*Cel5C from cellulolytic fungus *Penicillium decumbens* with distinctive domain composition and hydrolysis product profile. *Enzyme and Microbial Technology*, 52(3), 190-195.

Liu, Y., Guo, H., Gu, J., and Qin, W. (2019). Optimize purification of a cellulase from *Bacillus velezensis* A4 by aqueous two-phase system (ATPS) using response surface methodology. *Process Biochemistry*, 87, 196-203.

Lora, J.H., and Glasser, W.G. (2002). Recent industrial applications of lignin: a sustainable alternative to nonrenewable materials. *Journal of Polymers and the Environment*, 10(1-2), 39-48.

Lorenz, P., Liebeton, K., Niehaus, F., and Eck, J. (2002). Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Current Opinion in Biotechnology*, 13(6), 572-577.

TH-2393_146106034

Lugani, Y., and Sooch, B.S. (2017). In-silico characterization of Cellulases from genus *Bacillus. International Journal of Current Research and Review,* 9(13), 30.

Lynd, L.R. (1996). Overview and evaluation of fuel ethanol from cellulosic biomass: technology, economics, the environment, and policy. *Annual Review of Energy and the Environment*, 21(1), 403-465.

Lynd, L.R., Elander, R.T., and Wyman, C.E. (1996). Likely features and costs of mature biomass ethanol technology. In *Seventeenth Symposium on Biotechnology for Fuels and Chemicals*, 741-761.

Maki, M., Leung, K.T., and Qin, W. (2009). The prospects of cellulose producing bacteria for the bioconversion of lignocellulosic biomass. *International Journal of Biological Sciences*, 5(5), 500-516.

Marcos, M., García-Cubero, M.T., González-Benito, G., Coca, M., Bolado, S., and Lucas, S. (2013). Optimization of the enzymatic hydrolysis conditions of steam-exploded wheat straw for maximum glucose and xylose recovery. *Journal of Chemical Technology and Biotechnology*, 88(2), 237-246.

Martin, C., Klinke, H.B., and Thomsen A.B. (2007). Wet oxidation as a pretreatment method for enhancing the enzymatic convertibility of sugarcane bagasse. *Enzyme Microbial Technology*, 40, 426-432.

McCarter, J.D., and Withers, G.S. (1994). Mechanisms of enzymatic glycoside hydrolysis. *Current Opinion in Structural Biology*, 4(6), 885-892.

McMillan, J.D. (1994). Enzymatic conversion of biomass for fuels production. Himmel, ME, Baker, JO, Overend, RP, 292-324.

TH-2393_146106034

Meng, X.Y., Zhang, H.X., Mezei, M., and Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current Computer-Aided Drug Design*, 7(2), 146-157.

Mesa, L., Salvador, C.A., Herrera, M., Carrazana, D.I., and González, E. (2016). Cellulases by *Penicillium* sp. in different culture conditions. *Bioethanol*, 2(1), 84-93.

Mood, S.H., Golfeshan, A.H., Tabatabaei, M., Jouzani, G.S., Najafi, G.H., Gholami, M., and Ardjmand, M. (2013). Lignocellulosic biomass to bioethanol, a comprehensive review with a focus on pretreatment. *Renewable and Sustainable Energy Reviews*, 27, 77-93.

Moore, D., Robson, G.D., and Trinci, A.P. (2011). 21st century guidebook to fungi with CD. Cambridge University Press.

Mukherjee, A., Mandal, T., Ganguly, A., and Chatterjee, P.K. (2016). Lignin Degradation in the Production of Bioethanol-A Review. *ChemBioEng Reviews*, 3(2), 86-96.

Nath, P., Dhillon, A., Kumar, K., Sharma, K., Jamaldheen, S.B., Moholkar, V.S., and Goyal, A. (2019). Development of bi-functional chimeric enzyme (*Ct*GH1-L1-*Ct*GH5-F194A) from endoglucanase (*Ct*GH5) mutant F194A and β-1,4-glucosidase (*Ct*GH1) from *Clostridium thermocellum* with enhanced activity and structural integrity. *Bioresource Technology*, 282, 494-501.

Nigam, P.S., and Singh, A. (2011). Production of liquid biofuels from renewable resources. *Progress in Energy and Combustion Science*, 37(1), 52-68.

TH-2393_146106034

Niu, C., Zhu, L., Xu, X., and Li, Q. (2016). Rational design of disulfide bonds increases thermostability of a mesophilic 1, 3-1, 4-β-glucanase from *Bacillus terquilensis*. *PloS One*, 11(4), 1-17.

Osho, M.B., Aruoture, A.O., and Abatan, T.A. (2017). Cellulase production by *Proteus* spp. JC402 from plantain fruits stalk biomass using submerged fermentation. *Nigerian Journal of Biotechnology*, 34(1), 65-70.

Ozkan, B.M., and Ahmet, A. (2016). Isolation, identification and molecular characterization of cellulolytic bacteria from rumen samples collected from Erzurum slaughter house, Turkey. *Research Journal of BioTechnology*, 11(2), 32-38.

Packer, M.S., and Liu, D.R. (2015). Methods for the directed evolution of proteins. *Nature Reviews Genetics*, 16(7), 379-394.

Pandey, M.P., and Kim, C.S. (2011). Lignin depolymerization and conversion: a review of thermochemical methods. *Chemical Engineering & Technology*, 34(1), 29-41.

Patagundi, B.I., Shivasharan, C.T. and Kaliwal, B.B. (2014). Isolation and characterization of cellulase producing bacteria from soil. *International Journal of Current Microbiology and Applied Sciences*, 3, 59-69.

Peng, J., and Xu, J. (2010). Low-homology protein threading. *Bioinformatics*, *26*(12), i294-i300.

Putro, J.N., Soetaredjo, F.E., Lin, S.Y., Ju, Y.H., and Ismadji, S. (2016). Pretreatment and conversion of lignocellulose biomass into valuable chemicals. *Royal Society of Chemistry Advances*, 6(52), 46834-46852.

TH-2393_146106034

Rashid, B., Baba, Z.A., Malik, M.A., Mir, A.H., Akhter, F., Asif, M., Zargar, M.Y., Rashid, N., Rashid, N., and Maqbool, S. (2019). Characterization of Cellulolytic Bacteria from Waste Dumping Sites of Kashmir Himalaya. *International Journal of Current Microbiology and Applied Sciences*, 8(1), 2033-2048.

Ray, D.K., Mueller, N.D., West, P.C., and Foley, J.A. (2013). Yield trends are insufficient to double global crop production by 2050. *PloS one*, 8(6), e66428.

Reetz, M.T., and Carballeira, J.D. (2007). Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nature Protocols*, 2(4), 891-903.

Rodhe, A.V., Sateesh, L., Sridevi, J., Venkateswarlu, B., and Rao, L.V. (2011). Enzymatic hydrolysis of sorghum straw using native cellulase produced by *T. reesei* NCIM 992 under solid state fermentation using rice straw. *3 Biotech*, 1(4), 207-215.

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, 5(4), 725-738.

Ruiz, D.M., Turowski, V.R., and Murakami, M.T. (2016). Effects of the linker region on the structure and function of modular GH5 cellulases. *Scientific Reports*, 6, 28504.

Sadhukhan, J., Martinez-Hernandez, E., Amezcua-Allieri, M.A., and Aburto, J. (2019). Economic and environmental impact evaluation of various biomass feedstock for bioethanol production and correlations to lignocellulosic composition. *Bioresource Technology Reports*, 7, 100230.

TH-2393_146106034

Saha, B.C. (2003). Hemicellulose bioconversion. *Journal of Industrial Microbiology and Biotechnology*, 30(5), 279-291.

Saini, A., Aggarwal, N.K., Sharma, A., and Yadav, A. (2015). Actinomycetes: a source of lignocellulolytic enzymes. *Enzyme Research*, 2015, 279381.

Sangkharak, K., and Prasertsan, P. (2013). The production of polyhydroxyalkanoate by *Bacillus licheniformis* using sequential mutagenesis and optimization. *Biotechnology and Bioprocess Engineering*, 18(2), 272-279.

Santos, C.R., Paiva, J.H., Sforça, M.L., Neves, J.L., Navarro, R.Z., Cota, J., Akao, P. K., and Nogueira, M.L. (2012). Dissecting structure-function-stability relationships of a thermostable GH5-CBM3 cellulase from *Bacillus subtilis* 168. *Biochemical Journal*, 441(1), 95-104.

Schülein, M. (2000). Protein engineering of cellulases. *Biochimica et Biophysica Acta-Protein Structure and Molecular Enzymology*, 1543(2), 239-252.

Shafiee, S., and Topal, E. (2009). When will fossil fuel reserves be diminished?. *Energy Policy*, 37(1), 181-189.

Shahzadi, T., Mehmood, S., Irshad, M., Anwar, Z., Afroz, A., Zeeshan, N., Rashid, U., and Sughra, K. (2014). Advances in lignocellulosic biotechnology: A brief review on lignocellulosic biomass and cellulases. *Advances in Bioscience and Biotechnology*, 5(3), 246-251.

Sharma, H.K., Xu, C., and Qin, W. (2019). Biological pretreatment of lignocellulosic biomass for biofuels and bioproducts: an overview. *Waste and Biomass Valorization*, 10(2), 235-251.

TH-2393_146106034

Singh, A., Patel, A.K., Adsul, M., Mathur, A., and Singhania, R.R. (2017). Genetic modification: a tool for enhancing cellulase secretion. *Biofuel Research Journal*, 4(2), 600-610.

Singh, S., Khanna, S., Moholkar, V.S., and Goyal, A. (2014). Screening and optimization of pretreatments for *Parthenium hysterophorus* as feedstock for alcoholic biofuels. *Applied Energy*, 129, 195-206.

Singh, S., Moholkar, V.S., and Goyal, A. (2013). Isolation, identification, and characterization of a cellulolytic *Bacillus amyloliquefaciens* strain SS35 from Rhinoceros dung. *ISRN Microbiology*, 2013, 728134.

Smith, D.K., Radivojac, P., Obradovic, Z., Dunker, A.K., and Zhu, G. (2003). Improved amino acid flexibility parameters. *Protein Science*, 12(5), 1060-1072.

Sohail, M., Ahmad, A., and Khan, S.A. (2016). Production of cellulase from *Aspergillus terreus* MS105 on crude and commercially purified substrates. *3 Biotech*, 6(1), 103.

Srivastava, L. M. (2002). Cell wall, cell division and cell growth, Elsevier, *Plant Growth and Development: Hormones and Environment*, chapter 2, 23-74.

Swartz, J.R. (1996). *Escherichia coli* recombinant DNA technology. *Escherichia coli and Salmonella: Cellular and Molecular Biology*, *2*, 1693-1711.

Tabssum, F., Irfan, M., Shakir, H.A., and Qazi, J.I. (2018). RSM based optimization of nutritional conditions for cellulase mediated Saccharification by *Bacillus cereus*. *Journal of Biological Engineering*, 12(1), 7-17.

TH-2393_146106034

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*, 72(2), 211-222.

Thomas, H.L., Pot, D., Latrille, E., Trouche, G., Bonnal, L., Bastianelli, D., and Carrère, H. (2019). *Sorghum* biomethane potential varies with the genotype and the cultivation site. *Waste and Biomass Valorization*, 10(4), 783-788.

Topal, M.D., and Fresco, J.R. (1976). Complementary base pairing and the origin of substitution mutations. *Nature*, 263(5575), 285-289.

Updegraff, D.M. (1969). Semimicro determination of cellulose in biological materials. *Analytical Biochemistry*, 32(3), 420-424.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., and Berendsen, H.J. (2005). GROMACS: fast, flexible, and free. *Journal of Computational Chemistry*, 26(16), 1701-1718.

Vanhala, P., Bergstrom, I., Haaspuro, T., Kortelainen, P., Holmberg, M., and Forsius, M. (2016). Boreal forests can have a remarkable role in reducing greenhouse gas emissions locally: land use-related and anthropogenic greenhouse gas emissions and sinks at the municipal level. *Science of the Total Environment*, 557, 51-57.

Vohra, M., Manwar, J., Manmode, R., Padgilwar, S., and Patil, S. (2014). Bioethanol production: feedstock and current technologies. *Journal of Environmental Chemical Engineering*, 2(1), 573-584.

Vu, V.H., and Kim, K. (2012). Improvement of cellulase activity using error-prone rolling circle amplification and site-directed mutagenesis. *Journal of Microbiology and Biotechnology*, 22(5), 607-613.

TH-2393_146106034

Wang, Y., Wang, X., Tang, R., Yu, S., Zheng, B., and Feng, Y. (2010). A novel thermostable cellulase from *Fervidobacterium nodosum*. *Journal of Molecular Catalysis B: Enzymatic*, 66(3-4), 294-301.

Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids. *Nature*, 171(4356), 737-738.

Witkin, E.M. (1976). Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriological Reviews*, 40(4), 869-907.

Wooley, R., Ruth, M., Glassner, D., and Sheehan, J. (1999). Process design and costing of bioethanol technology: a tool for determining the status and direction of research and development. *Biotechnology Progress*, 15(5), 794-803.

Wu, S., and Zhang, Y. (2007). LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic Acids Research*, 35(10), 3375-3382.

Xu, J., and Smith, J.C. (2010). Probing the mechanism of cellulosome attachment to the *Clostridium thermocellum* cell surface: computer simulation of the Type II cohesion-dockerin complex and its variants. *Protein Engineering, Design and Selection*, 23(10), 759-768.

Xu, Y.Q., Duan, C.J., Zhou, Q.N., Tang, J.L., and Feng, J.X. (2006). Cloning and identification of cellulase genes from uncultured microorganisms in pulp sediments from paper mill effluent. *Acta Microbiologica Sinica*, 46(5), 783-788.

Yu, H., and Huang, H. (2014). Engineering proteins for thermostability through rigidifying flexible sites. *Biotechnology Advances*, 32(2), 308-315.

TH-2393_146106034

Yu, H., Zhao, Y., Guo, C., Gan, Y., and Huang, H. (2015). The role of proline substitutions within flexible regions on thermostability of luciferase. *Biochimica et Biophysica Acta-Proteins and Proteomics*, 1854(1), 65-72.

Zabed, H., Sahu, J.N., Suely, A., Boyce, A.N., and Faruq, G. (2017). Bioethanol production from renewable sources: Current perspectives and technological progress. *Renewable and Sustainable Energy Reviews*, 71, 475-501.

Zhang, J., Shi, H., Xu, L., Zhu, X., and Li, X. (2015). Site-directed mutagenesis of a hyperthermophilic endoglucanase Cel12B from *Thermotoga maritima* based on rational design. *PloS one*, 10(7), e0133824.

Zhang, X.Y., Zi, L.H., Ge, X.M., Li, Y.H., Liu, C.G., and Bai, F.W. (2017). Development of *Trichoderma reesei* mutants by combined mutagenesis and induction of cellulase by low-cost corn starch hydrolysate. *Process Biochemistry*, 54, 96-101.

Zhang, Y.H.P., Himmel, M.E., and Mielenz, J.R. (2006). Outlook for cellulase improvement: screenings and selection strategies. *Biotechnology Advances*, 24(5), 452-481.

Zheng, F., Vermaas, J.V., Zheng, J., Wang, Y., Tu, T., Wang, X., Xei, X., Yao, B., Beckham, G.T., and Luo, H. (2018). Characterizing Activity and Thermostability of GH5 Cellulase Chimeras from Mesophilic and Thermophilic Parents. *BioRxiv*, 382069.

TH-2393_146106034

Chapter 2

Strain improvement of *Bacillus amyloliquefaciens* SS35 by UV and EMS-directed evolution for efficient endoglucanase production

2.1 Introduction

The uncertainty in world petroleum supplies, environmental pollution and climate change generates interest in producing sustainable green fuel from bioresources such as lignocellulosic waste. Lignocellulosic biomass is composed of cellulose the most abundant renewable carbon source on the earth's surface (Chen *et al.,* 2013). Thus, it has the potential to serve the worldwide demand of renewable transportation fuel (Sims *et al.,* 2010). However, commercialization of the cellulosic fuels faces several techno-economic challenges in order to compete with fossil fuels (Lynd *et al.,* 2008). A major bottleneck for the biofuel industry is the hydrolysis of plant cell wall polysaccharide, especially the highly recalcitrant cellulose fibers into fermentable sugars (Himmel *et al.,* 2007). Cellulases are the group of enzymes that are well known for bioconversion of lignocellulosic waste into fermentable sugars (Lawford *et al.,* 2003). They are grouped into three categories: (i) endoglucanase (EC 3.2.1.4) which hydrolyzes β-1,4 bonds between glucose units of cellulose, (ii) cellobiohydrolase (EC 3.2.1.91) which cleaves the reducing and non-reducing ends of cellulose fibre to release cellobiose and (iii) β-glucosidase (EC 3.2.1.21) which

TH-2393_146106034

converts cellobiose to glucose (Bayer *et al.,* 2007). Hydrolysis of cellulose into monomeric sugars involves synergistic action of these three cellulase enzymes. Microorganisms such as bacteria, fungi and actinomycetes are the known producers of cellulase enzyme. Aerobic mesophilic bacteria belonging to the genus *Bacillus* are efficient cellulase producers such as *Bacillus* KSM-S237 (Hakamada *et al.,* 1997), *Bacillus subtilis* (Kim *et al.,* 2009), *Bacillus megaterium* (Shobharani *et al.,* 2013) and *Bacillus amyloliquefaciens* SS35 (Singh *et al.,* 2013). However, the use of cellulases at an industrial scale for bioethanol production face challenges because of its high production cost, low enzyme yield and low enzyme activities (Kim *et al.,* 1987).

One way to reduce the production cost of bioethanol is the strain improvement of cellulase producing bacteria for improved cellulase characteristics such as cellulase production, catalytic efficiency and wide range pH stability. This may be achieved by random mutagenesis of natural bacteria leading to directed evolution of the isolate which may result in improved characteristics. Directed evolution by many biotechnological methods is a process of forming highly efficient biocatalyst which is used by industry for forming eco-friendly products. This can be performed by two ways physical and chemical methods. The physical methods are Ultraviolet irradiation (UV), gamma rays, microwaves, X-rays etc. Chemically by N-methyl N'-nitro-N-nitrosoguanidine (NTG), ethidium bromide, ethyl methanesulfonate (EMS) etc. Mutant strains showed several-fold increase in cellulase production and activity compared to the wild-type strain and thus enzyme loading in saccharification of lignocellulosic biomass decreases which increases the efficiency of overall process (Wooley *et al.,* 1999; Sangkharak *et al.,* 2012). Mutagenesis by UV and EMS in microorganisms is simple and easy technique that can be performed in laboratory to

TH-2393_146106034

get the desired product. Together UV and EMS mutagen may be an advantage to induce mutations in desire wild-type strain for producing efficient mutants (Lawrence *et al.*, 1988). Optimum dose of any mutagen is essential for balancing the challenging needs to get high mutation frequency. The highest proportion of mutation frequency per treated cell is usually found at dose that gives 1% survival rate (Patel *et al.*, 2010). There are several reports, on the strain improvement of fungal isolates such as *Trichoderma viride* (Li *et al.*, 2010), *Trichoderma reesei* (Zhang *et al.*, 2017), *Aspergillus nidulans* (Kaur *et al.*, 2014) and *Streptomyces griseoaurantiacus* (Kumar, 2015) by UV irradiation and EMS treatment for producing hyperactive cellulase mutants.

However, very few reports are available on UV and EMS directed evolution of *Bacillus* spp. Therefore, in the present study, the UV-directed evolution of wild-type *Bacillus amyloliquefaciens* SS35 was performed for improved cellulase production. The most potential UV mutant in terms of cellulase production and genetic stability was further improved by EMS mutagen for improved cellulase production. EMS mutants were also screened for the genetic stability. Based on the higher cellulase production and genetic stability the most potential mutant was selected and used for further studies.

TH-2393_146106034

2.2 Materials and Methods

2.2.1 Media and chemicals

Carboxymethylcellulose sodium salt (CMC-Na) (low viscosity) and D-glucose were procured from Sigma Aldrich, USA. Peptone, Yeast extract, K_2HPO_4, NaCl and $MgSO_4.7H_2O$, EMS and Congo red used in this study were procured from HiMedia Pvt. Ltd., India.

2.2.2 Microorganism and culture conditions

Bacillus amyloliquefaciens SS35 isolated from Rhinoceros dung with gyrase A GenBank accession no. KF019284 and 16S rDNA GenBank accession no. JX674030 (Singh *et al.*, 2013) was used in this study. The medium composition for CMCase production from *Bacillus amyloliquefaciens* SS35 was CMC-Na (19.0 g/L), peptone (2 g/L), yeast extract (8 g/L), K_2HPO_4 (1 g/L), NaCl (1 g/L) and $MgSO_4.7H_2O$ (0.2 g/L) at pH 5.6. 7% (v/v) inoculum of cell optical density 1.0 at 600 nm was inoculated in 200 mL medium contained in 500 mL Erlenmeyer flask and incubated at 40ºC in an orbital shaking speed of 120 rpm (Singh *et al.*, 2014). The extracellular enzyme in cell free supernatant of mutant strains was harvested by centrifuging the culture broth at 12,000g for 20 min at 4°C after 48 h.

2.2.3 UV-directed evolution of wild-type *Bacillus amyloliquefaciens* SS35

In UV mutagenesis, the overnight grown culture of the bacterium (cell optical density ~1.0) was spread plated onto a medium as described in section 2.2.2 from 10^{-1} to 10^{-6} serial dilution in 0.85% (w/v) sodium saline under sterile conditions. The effect of UV exposure on wild-type for different time intervals was evaluated at 0.6 J/cm^2 frequency in UV crosslinker (Hoefer, UVC 500) by using direct plate irradiation method (Rapa *et al.*, 2015). The mutation was induced by exposing the petridishes of

TH-2393_146106034

10^{-4} serial dilution to UV light for 5 min, 10 min, 15 min, 20 min, 30 min, 40 min, 50 min, 60 min, 75 min, 90 min, 105 min, 120 min, 135 min, 150 min, 165 min and 180 min to obtain 1% bacterial survival rate (99% lethality). After irradiation petridishes were immediately wrapped by aluminium foil to avoid photoreactivation by SOS mechanism and along with the control plate incubated at 40°C for 12 h.

Bacterial survival rate was calculated by using the following equation:

$$\text{Survival (\%)} = \frac{\text{cfu/ml of colonies grown on UV irradiated plate}}{\text{cfu/ml of colonies grown on control plate}} \times 100$$

2.2.4 Qualitative screening of UV mutant colonies by plate staining method

Different UV mutant colonies were picked up from 1% survival rate exposed plate and streaked on separate CMC-Na supplemented in optimised medium (Singh *et al.*, 2014) plate with grids drawn over it and incubated at 40°C for 12 h. Similarly, replica plates of above mutant colonies were also prepared separately for 0.3% (w/v) Congo red staining (Ruijssenaars and Hartmans, 2001). The replica plates were stained with 0.3% (w/v) Congo red for 15 min. After 15 min Congo red stain was poured off and plates were washed with 1M NaCl until the clear zone around the wild-type mutant colonies was visible (Teather and Wood 1982). The selected wild-type and mutant strains were maintained by preparing the glycerol stocks of the mutant colonies culture in 50% (v/v) glycerol. 1 mL of 18 h grown culture was transferred to 2 mL cryo vials containing 1 mL of 50% (v/v) glycerol and stored at -80ºC.

TH-2393_146106034

2.2.5 Quantitative screening of UV mutant strains for improved CMCase activity

The UV mutant colonies showing clear zone on the plate obtained by staining method were picked up from the master plate and screened further. A loop of culture from UV irradiated plate colonies was inoculated into 5 mL of liquid medium (Singh *et al.*, 2014) and grown for 12 h. After that, 150 µL of culture was inoculated into 15 mL enzyme production medium (Singh *et al.*, 2014) for primary culture. After 12 h, 7% (v/v) inoculum was inoculated in 200 mL enzyme production medium according to the method as described in earlier in section 2.2.2. The broth after 48 h was centrifuged at 4°C 12000g for 20 min. The cell free supernatant was analysed for endoglucanase activity for improved enzyme production. The enzyme activity was estimated by measuring the released reducing sugar by Nelson (1944) and Somogyi (1945) method described in section (2.2.5.2) using D-Glucose standard (10-500 µg/mL). The enzyme assay was carried out by incubating the 10 µL of crude enzyme in 100 µL total reaction volume with 1% (w/v) CMC-Na in 50 mM sodium acetate buffer, pH 5.0 at 65°C for 5 min. To 100 µL of reaction mixture, 100 µL of reagent D was added (Section 2.2.5.1). The solution was mixed and heated for 20 min in the boiling water bath. After 20 min of boiling, the solution was cooled on ice, then and 100 µL of reagent C (Section 2.2.5.1) was added. The colour developed rapidly and completed after the evolution of carbon dioxide stopped. The mixture was diluted by adding distilled water to make up the volume to 1 mL and absorbance at 500 nm (A_{500}) was measured by a UV-Visible spectrophotometer (Varian, Cary 100 Bio). One unit (U) of CMCase activity was defined as the amount of enzyme required to release one µmole of reducing sugar (glucose) per min under mentioned conditions.

TH-2393_146106034

2.2.5.1 Reagents for enzyme assay

Composition of the reagents used in CMCase activity assay is as follows (Nelson 1944; Somogyi 1945):

Reagent A

Sodium carbonate (2.5 g), potassium sodium tartrate tetrahydrate (2.5 g), sodium bicarbonate (2.0 g) and sodium sulphate (20.0 g) dissolved in 100 mL distilled water.

Reagent B

Copper sulphate pentahydrate (4.5 g) and conc. sulphuric acid (1-2 drops) in 30 mL of distilled water.

Reagent C

Ammonium molybdate tetrahydrate (2.5 g) and conc. sulphuric acid (2.1 mL) were added to 45 mL of distilled water.

Solution of sodium arsenate heptahydrate (0.3 g) in 2.5 mL of distilled water was mixed to it and stored in an amber bottle at 37°C for 24 h prior to use.

Reagent D

Mixture of reagents A and B in a ratio of 25:1.

TH-2393_146106034

2.2.5.2 Calculation of CMCase activity

One unit (U) of CMCase activity is defined as the amount of enzyme that liberates 1 µmole of reducing sugar per min at 65°C in 50 mM sodium acetate buffer, pH 5.0. The carboxymethylcellulase activity was calculated as:

$$\text{Enzyme activity (U/ml)} = \frac{\Delta A \times C \times V}{180 \times t \times v}$$

ΔA = change in absorbance at 500 nm

C = 1 OD equivalent glucose concentration (mg/mL) from standard plot

V = volume of the reaction mixture (mL)

t = time of reaction (min)

180 = molecular weight of glucose

v = volume of the enzyme source (mL) for reducing sugar estimation

2.2.5.3 Estimation of protein concentration

The protein concentration of the crude enzyme in cell free supernatant of wild-type and UV mutant strains was determined by the Lowry method described earlier by Lowry *et al.*, (1951). BSA, in 50-500 µg/mL concentration range was used to plot the standard curve. The reaction mixture contains 20 µL of cell free supernatant in 200 µL of reaction volume. To above reaction mixture 1.0 mL of reagent C (described below) was added. After 10 min, 100 µL of Folin's reagent (described below) was added and incubated in dark for 30 min. The absorbance was measured at 660 nm (A_{660}) against a blank.

TH-2393_146106034

2.2.5.4 Reagents composition used in the protein estimation by Lowry method

Reagent A

Sodium carbonate (2.0 g) and Sodium hydroxide (0.4 g) were dissolved in distilled water and the volume made up to 100 mL.

Reagent B1

Potassium sodium tartrate tetrahydrate (2.0 g) dissolved in 100 mL distilled water.

Reagent B2

Copper sulphate pentahydrate (1.0 g) dissolved in 100 mL distilled water.

Reagent C

Freshly prepared by mixing reagent B1, A and B2 in a ratio of B1:A:B2 = 1:100:1.

Folin's reagent : 1 N Folin's reagent.

The concentration of protein was calculated as follows:

$$\text{Protein concentration(mg/ml)} = \frac{\Delta A \times C \times V}{v}$$

ΔA = change in absorbance of the sample at 660 nm

C = 1 absorbance equivalent BSA concentration (mg/mL) from standard plot

V = total volume (mL)

v = volume of the sample (mL)

2.2.6 Genetic stability of the UV mutants

Mutant strains developed after the random mutagenesis have the tendency to revert back to the wild type. Consequently, the genetic stability of a mutant strain is an essential criteria for their application in industry (Molzahn, 1977; Simpson and Caten, 1979; Al-Jailawi *et al.*, 2016). The genetic stability of the UV mutants was determined by measuring the endoglucanase production in batch fermentation at

TH-2393_146106034

shake flask level for the ten successive generations. Enzyme was produced according to the method described in section 2.2.2. UV mutant flasks were covered with aluminum foil to maintain the stability of mutant strains. Endoglucanase production in cell free supernatant was measured by measuring the endoglucanase activity according to the method described earlier by Nelson 1944 and Somogyi 1945. The endoglucanase assay after each generation was carried out by incubating the crude enzyme present in cell free supernatant with 1% (w/v) CMC-Na in 50 mM sodium acetate buffer, pH 5.0 at 65°C for 5 min. The enzyme assay was carried out according to the method described in section 2.2.3.

2.2.7 EMS directed evolution of potential UV mutant

Based on the maximum increase in CMCase activity and stability of UV mutant strain (as described later in Results and discussion section), a UV2 mutant (named after its colony number) was selected for further EMS-directed mutagenesis. 5.0% (v/v) stock solution of EMS was prepared in distilled water and sterilized by passing it through 0.2 µm sterilized syringe filter. 1 mL of 10^{-3} serially diluted 15 h grown culture of potential UV mutant (cell optical density ~1.8, 600 nm) was treated with 0%, 0.5%, 1% and 1.5% EMS dose at 40°C for 20 min at 120 rpm in incubator. The 2% (v/v) EMS dose treatment was performed without the serial dilution of culture at 40°C for 20 min at 120 rpm in incubator. After 20 min, the EMS-treated UV mutant cells were spread plated onto a medium (described in section 2.2.2) supplemented with 1.9% (w/v) CMC-Na and pertidishes were wrapped in aluminium foil and incubated at 40°C in incubator for 18 h. After 18 h EMS mutant colonies were counted in colony counter and bacterial survival rate was calculated according to the equation given in section 2.2.3.

TH-2393_146106034

2.2.8 Qualitative screening of EMS mutant colonies by Plate staining method

Different EMS mutant colonies were picked up from 1% survival rate exposed plate and streaked on separate CMC-Na supplemented in optimised medium (Singh *et al.*, 2014) plate with grids drawn over it and incubated at 40°C for 18 h. Similarly, replica plates of EMS mutant colonies were also prepared separately for 0.3% (w/v) Congo red staining (Ruijssenaars and Hartmans, 2001). Congo red staining of replica plate was performed according to the method described in section 2.2.4. The selected EMS mutant strains were maintained by preparing the glycerol stocks of the mutant colonies culture according to the method described in section 2.2.4.

2.2.9 Quantitative screening of EMS mutants for improved CMCase activity

The EMS mutant colonies showing clear zone in plate staining method were picked up from the master plate and screened further, for improved enzyme production by analyzing the endoglucanase activity (U/mL) of extracellular CMCase enzyme in cell-free culture broth grown in enzyme production medium according to the protocol as described earlier in section 2.2.2. The enzyme activity was estimated by measuring the released reducing sugar by Nelson (1944) and Somogyi (1945) method described in section 2.2.5.

2.2.10 Genetic stability of the EMS mutants

The genetic stability of the EMS mutants was determined by measuring the endoglucanase production in batch fermentation at shake flask level for the ten successive generations according to the method as described in section 2.2.6.

TH-2393_146106034

2.2.11 Fermentation profiles of wild-type *Bacillus amyloliquefaciens* SS35 and its mutant

Based on the maximum increase in CMCase activity and stability of UV and EMS mutant strain (as described later in Results and discussion section), a UV2 mutant after UV mutagenesis was selected as the most potential mutant and used in further studies. The fermentation profiles of wild-type *Bacillus amyloliquefaciens* SS35 and its mutant UV2 were studied by growing in 200 mL optimized medium in 500 mL Erlenmeyer flask according to the method mentioned in section 2.2.2 at 40°C and 120 rpm for 96 h. The endoglucanase activity, cell optical density (600 nm), protein concentration and pH change of the broth were monitored at every 6 h up to 96 h by taking 1 mL aliquot. The endoglucanase assay and calculation of enzyme activity was performed according to the method described in sections 2.2.5, 2.2.5.1 and 2.2.5.2. The wild-type and UV2 mutant growth was monitored by measuring the absorbance at 600 nm using UV-visible spectrophotometer (Perkin Elmer, Lambda-45). Protein concentration was measured according to the method described in section 2.2.5.3. The pH change with time in broth was measured by pH meter.

TH-2393_146106034

2.3 Results and Discussion

2.3.1 UV mutagenesis of wild-type *Bacillus amyloliquefaciens* SS35

Number of UV mutant colonies developed after the UV irradiation of petridish and colonies on control plate (without UV exposure) were enumerated by a colony counter. It was observed that with the increase in UV exposure time there was a gradual decrease in the percentage survivor of bacterial cells. It was found that exposure of wild-type strain at frequency 0.6 J/cm^2 for 180 min was sub-lethal because it results in 99% mortality of bacterium, giving 1% survival rate (Table 2.3.1). According to Li *et al.*, (2010) and Patel *et al.*, (2010) for obtaining the potential mutant strain, wild-type strain was exposed to mutagen at the sub-lethal dose for 1% survival rate. *Streptomyces griseoaurantiacus* tolerates 60 min of UV exposure dose in order to form the UV mutant strain with improved endoglucanase and β-glucosidase activity (Kumar *et al.*, 2015). Fourteen UV mutant colonies of *Bacillus amyloliquefaciens* SS35 were picked up from 180 min UV exposure plate for further qualitative screening.

TH-2393_146106034

Table 2.3.1 Number of mutant colonies of *Bacillus amyloliquefaciens* SS35 obtained after UV exposure

UV exposure (min)	Contol plate (cfu/mL)	UV exposed plate (cfu/mL)	Survival Rate (%)	Lethality Rate (%)
15	$2.99x10^7$	$2.96x10^7$	98.9	1.0
20	$2.99x10^7$	$2.97x10^7$	99.3	0.6
30	$2.99x10^7$	$2.91x10^7$	97.3	2.6
40	$2.99x10^7$	$2.54x10^7$	84.9	15.0
50	$7.01x10^7$	$5.99x10^7$	85.4	14.5
60	$7.01x10^7$	$4.27x10^7$	60.9	39.1
75	$7.01x10^7$	$3.98x10^7$	56.7	43.2
90	$3.70x10^7$	$2.09x10^7$	56.4	43.5
105	$3.70x10^7$	$6.8x10^6$	18.3	81.6
120	$3.70x10^7$	$3.2x10^6$	8.6	91.4
135	$7.01x10^7$	$5.7x10^6$	8.1	91.9
150	$7.01x10^7$	$5.1x10^6$	7.2	92.7
165	$7.01x10^7$	$4.8x10^6$	6.8	93.1
180	$1.85x10^7$	$2.0x10^5$	1.1	98.9

2.3.2 Qualitative screening of UV mutant colonies by plate staining method

All the selected fourteen UV mutant strains, UV1, UV2, UV3, UV4, UV5, UV6, UV7, UV8, UV9, UV10, UV11, UV12, UV13 and UV14 showed the hydrolyzing zone of clearance after qualitative screening by Plate staining method as shown in Fig. 2.3.1. This confirmed that all the UV mutant strains had retained the CMCase activity and none of them produced the inactive enzyme. However, according to the Maki *et al*., (2009) the plate staining method with 0.3% (w/v) Congo red and destaining with 1M sodium chloride is not quantitative because of the poor correlation between CMCase activity and the colony to zone of clearance ratio. Therefore, all the fourteen UV mutant colonies were picked up from the replica plate and screened further on the basis of CMCase activity in enzyme production medium.

TH-2393_146106034

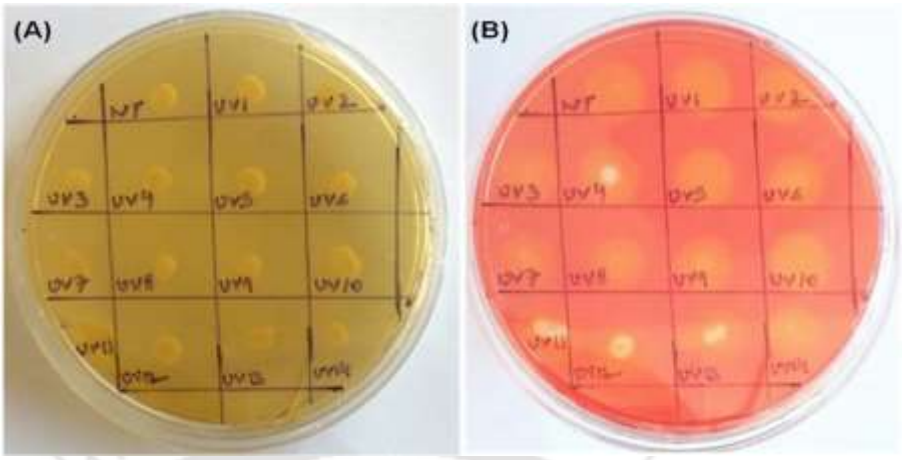


Fig. 2.3.1 (A) Petridish showing UV mutant colonies UV1 to UV14; (B) Replica plate of UV mutant colonies UV1 to UV14 after staining with 0.3% (w/v) Congo red, where, WT represents wild-type strain.

2.3.3 Quantitative screening of UV mutant strains for improved CMCase activity

Initial quantitative screening of UV mutant strains showed that out of 14, only 3 gave the significant increase in CMCase activity for crude cell free supernatant as compared with the wild-type strain as shown in Table 2.3.2. While, some mutant strains also showed the decrease in CMCase activity according to the Gauss distribution (Pelechova *et al.,* 1990). UV1, UV2 and UV4 mutant strains displayed 0.76 U/mL, 0.79 U/mL and 0.74 U/mL, respectively, CMCase activity for crude enzyme, which was 17%, 22% and 14%, respectively, higher than the wild-type strain (0.65 U/mL) (Table 2.3.2). The CMCase activity of these UV mutant strains was also higher than activity of CMCase produced from some identified wild-type strains, for example, *Bacillus* sp. (0.02 U/mL, isolated from Coir retting effluent, Immanuel *et al*. 2006), *Bacillus subtilis* AS3 (0.56 U/mL, Cow dung, Deka *et al*., 2013), *Bacillus pumilus* MGB05 (0.262 U/mL, isolated from midgut of muga silkworm, Bhuyan *et. al.,* 2018), *Bacillus* sp. (0.24 U/mL isolated from Buffalo digestive system, Raza *et al.,* 2019), *Pseudomonas* (0.052 U/mL, isolated from waste damping site, Goel *et. al.,*

TH-2393_146106034

2019) and *Anoxybacillus sp. (0.0*57 U/mL, isolated from North Malaysian tropical mangrove soil, Naresh *et. al.*, 2019). Thus, the UV mutant strains UV1, UV2 and UV3 were the efficient CMCase producing strains. It has been reported that UV mutation improved the strains of *Trichoderma viride* (Li *et al.*, 2010), *Trichoderma reesei* (Zhang *et al.*, 2017) and *Streptomyces griseoaurantiacus* (Kumar 2015) for enhanced cellulase activity. In similar type of study, UV mutant of *Bacillus subtilis* showed 6.5% increase in β-glucosidase production than the wild-type strain (0.675 U/mL) (Agrawal *et al.*, 2013).

Table 2.3.2 CMCase activities of 14 selected UV mutants of isolate *Bacillus amyloliquefaciens* SS35

Strains	Enzyme activity (U/mL)	% increase in Enzyme activity
Control	0.65 ± 0.03	-
UV1	0.76 ± 0.01	17
UV2	0.79 ± 0.02	22
UV3	0.64 ± 0.01	-
UV4	0.74 ± 0.01	14
UV5	0.60 ± 0.07	-
UV6	0.54 ± 0.03	-
UV7	0.57 ± 0.03	-
UV8	0.59 ± 0.04	-
UV9	0.66 ± 0.01	2
UV10	0.59 ± 0.02	-
UV11	0.44 ± 0.05	-
UV12	0.58 ± 0.01	-
UV13	0.66 ± 0.01	2
UV14	0.61 ± 0.04	-

Values are mean ± SE (n=3).

TH-2393_146106034

2.3.4 Genetic stability of UV mutant strains

Genetic stability of selected UV mutant strains UV1, UV2 and UV4 was demonstrated for 10 sequential generations with storage in glycerol stock for about 15 days. Out of 3 UV mutant strains, only 1, UV2 strain was stable for cellulase production upto 10 generations which suggested that this strain is genetically stable. UV2 strain retain almost 100% of the relative CMCase activity for crude enzyme while other two strains UV1 and UV4 showed the decreased CMCase activity after 3 and 4 generations, respectively (Fig. 2.3.2). In a similar type of study, Kostyleva *et al.*, (2018) reported that the UV mutant strain of *Trichoderma reesei* remains stable up to 3 generations for endoglucanase enzyme production. Agrawal *et al.*, (2013) found that UV mutant of *Bacillus subtilis* (PS-UM1) was stable upto the eighth generation. Therefore, UV2 mutant strain was the most potential strain and thus, used in the further mutation studies using the mutagen, EMS.

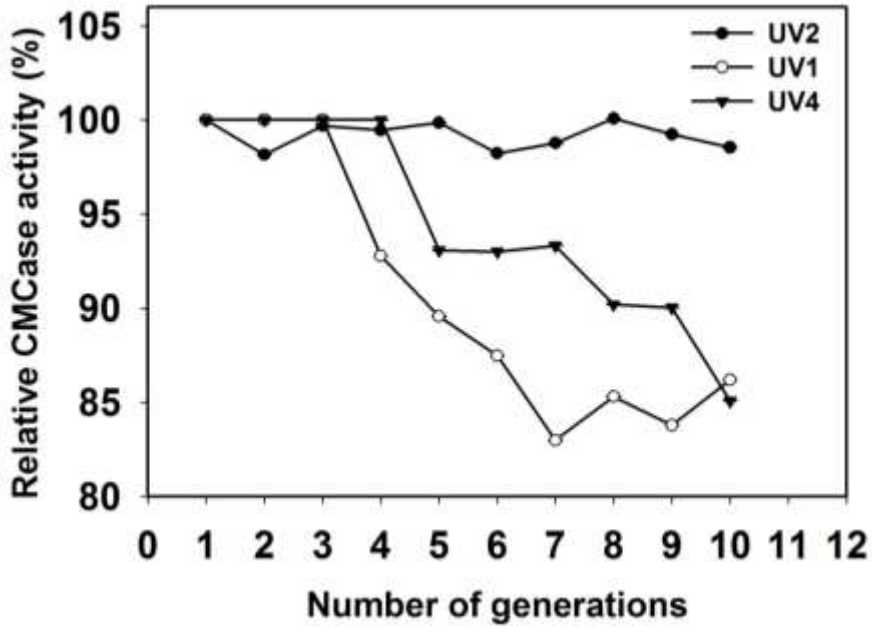


Fig. 2.3.2 Genetic stability of UV mutant strains upto 10 sequential generations.

TH-2393_146106034

2.3.5 EMS directed evolution of potential UV mutant

Effect of EMS treatment on UV2 mutant of *Bacillus amyloliquefaciens* SS35 for different concentrations was evaluated. Numbers of EMS mutant colonies after treating with 0.5% (v/v) to 2.5% (v/v) EMS dose were enumerated by colony counter along with the control (0% EMS). It was observed that, with the increase in EMS concentration there was a gradual decrease in the percentage survivor of UV2 mutant strain cells. It was found that 2% (v/v) EMS dose for 20 min treatment of UV2 strain was sub-lethal and results 99.9% mortality of UV2 strain giving 0.1% survival rate (Table 2.3.3). Rajoka (2004) mutated *Cellulomonas biazotea* with EMS for improved endoglucanase production and found that 160 µg/mL (0.01%) EMS after 60 min treatment yields 90% lethality. Zhang *et al.*, (2017) developed *Trichoderma reesei* mutants for enhanced cellulase production by combined mutagenesis of EMS (50 µL in 10 mL spore suspension) for 24 h followed by plasma irradiation at 24 V, 1.7 A, gap 3 mm for 5 min with lethality rate of 78.38%. Chandra *et al.*, (2009) observed 1 % survival rate of *Trichoderma reesei* spores after mutating it with chemical mutagen NTG for 6 h, followed by 15 min UV exposure. Further, fourteen EMS mutant colonies of UV2 mutant *strain* were picked up from 2% EMS dose for further qualitative screening.

TH-2393_146106034

Table 2.3.3 Number of EMS mutant colonies of UV2 mutant of *Bacillus amyloliquefaciens* SS35 obtained after EMS treatment

EMS dose (%)	Control UV2 (cfu/mL)	EMS treated UV2 plate (cfu/mL)	Survival Rate (%)	Lethality Rate (%)
0.5	6.2×10^4	4.3×10^4	69.2	30.8
1.0	6.2×10^4	2.3×10^4	36.2	63.8
1.5	6.2×10^4	6.5×10^3	10.1	89.9
2.0	6.2×10^4	9.5×10	0.1	99.9
2.5	6.2×10^4	0	0	100

2.3.6 Qualitative and quantitative screening of EMS mutants of UV2 strain for improved CMCase activity

All the selected fourteen EMS mutant strains of UV2 strain, EMS 1, EMS 2, EMS 3, EMS 4, EMS 5, EMS 6, EMS 7, EMS 8, EMS 9, EMS 10, EMS 11, EMS 12, EMS 13 and EMS 14 showed the hydrolyzing zone of clearance after qualitative screening by Plate staining method (Fig. 2.3.3). This, confirmed that all the EMS treated UV mutant strains had retained the CMCase activity and none of them produced the inactive enzyme. Quantitative screening of EMS mutant strains showed that out of 14, only 4 gave the increase in CMCase activity for crude cell free supernatant as compared with the wild-type strain as shown in Table 2.3.4. EMS 2, EMS 3, EMS 7 and EMS 8 mutant strains displayed 0.82 U/mL, 0.83 U/mL, 0.86 U/mL and 0.82 U/mL, respectively, CMCase activity for crude enzyme, which was 3%, 4%, 7% and 3%, respectively, higher than the UV2 strain (0.79 U/mL) (Table 2.3.4). In a similar type of study, Chandra *et al.*, (2009) mutated *Trichoderma citrinoviride* strain by EMS followed with ethidium bromide treatment. The mutant produced 3.12, IU mL^{-1} endoglucanase with 2-fold higher than those in wild-type strain. Combined exposure of wild-type strain of *Bacillus subtilis* to EMS and UV mutagens increased β-glucosidase production by 1.2-fold to 0.81 U/mL (Agrawal *et al.*, 2013). Haq *et al.*, (2010) mutated the *Bacillus amyloliquefaciens* with EMS

TH-2393_146106034

followed by UV irradiation and observed 1.4-fold increase in alpha amylase activity. Dhillon *et al.,* 2006) reported 1.5-fold increase in cellulase production from *Penicillium echinulatum* by three successive mutagenic treatments using UV, EMS and hydrogen peroxide. Chandra *et al.,* (2009) suggested that in random mutagenesis, the increase in CMCase activities might be due to the mutation in gene sequence or increased secretion of modular proteins or co-factors.

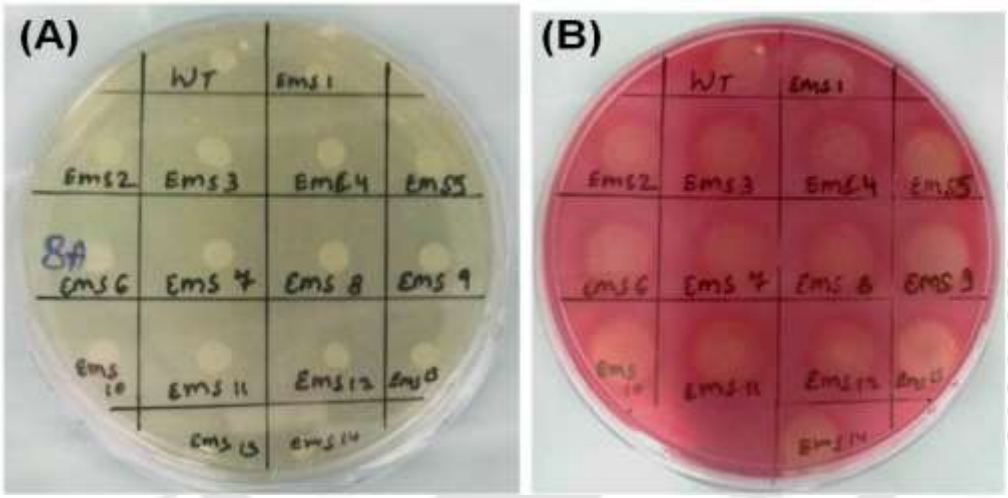


Fig. 2.3.3 (A) Petridish showing EMS mutant colonies of UV2 strain; EMS1 to EMS14; (B) Replica plate of EMS mutant colonies EMS1 to EMS14 after staining with 0.3% (w/v) Congo red, where, WT is wild-type strain.

TH-2393_146106034

Table 2.3.4 Enzyme activities of 14 selected EMS mutants of UV2 mutant of *Bacillus amyloliquefaciens* SS35

Strains	Enzyme activity (U/mL)	Increase in Enzyme activity (%)
Control (UV2)	0.79 ± 0.01	0
EMS1	0.78 ± 0.01	-
EMS2	0.82 ± 0.05	3
EMS3	0.83 ± 0.01	4
EMS4	0.77 ± 0.02	-
EMS5	0.76 ± 0.03	-
EMS6	0.76 ± 0.01	-
EMS7	0.86 ± 0.01	7
EMS8	0.82 ± 0.01	3
EMS9	0.75 ± 0.02	-
EMS10	0.74 ± 0.02	-
EMS11	0.78 ± 0.01	-
EMS12	0.77 ± 0.06	-
EMS13	0.78 ± 0.01	-
EMS14	0.66 ± 0.01	-

2.3.7 Genetic stability of the EMS mutants

Genetic stability of selected EMS mutant strains EMS 2, EMS 3, EMS 7 and EMS 8 was demonstrated for 10 sequential generations with storage in glycerol stock for 15 days. None of the 4 EMS mutant strains showed the stable cellulase production after one or three generations (Fig. 2.3.4). This suggested that these EMS mutant strains were reverting back its mutation to the original strain. From these results, it was concluded that only UV2 strain was the most potential mutant and therefore, was used in further studies. In a similar type of study, Li *et al.*, (2009) mutated the *Trichoderma viride* by 2 successive mutations using microwave and UV irradiation and found that the mutant strain was stable upto 9 generations for increased CMCase activity. Lim *et al.*, (2019) mutated *Aspergillus sojae* by three successive steps of mutagenesis using EMS followed by UV and NTG for increased L-leucine

TH-2393_146106034

aminopeptidase and protease production and found that the final mutant was stable upto 10 generations. Agrawal *et al.,* (2013) found that EMS plus UV mutant strain of *Bacillus subtilis* was stable upto the tenth generation with 78% of initial β-glucosidase production.

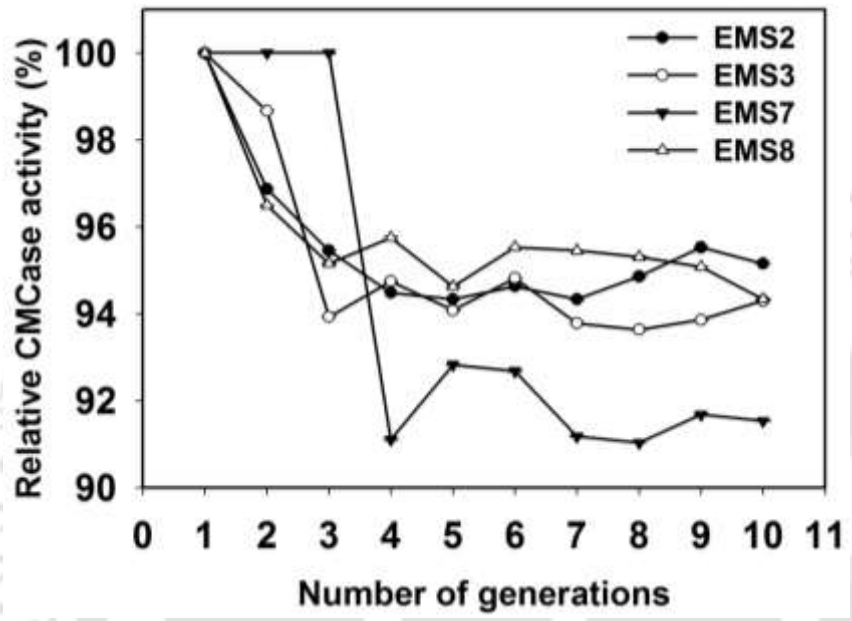


Fig. 2.3.4 Genetic stability of EMS mutant strains upto 10 sequential generations.

2.3.8 Fermentation profiles of wild-type *Bacillus amyloliquefaciens* SS35 and the most potential mutant strain UV2

The fermentation profiles of wild-type *Bacillus amyloliquefaciens* SS35 and its mutant UV2 were studied for 96 h (Fig. 2.3.5 and Table 2.3.5). The cell growth, enzyme activity, protein concentration and pH of broth during fermentation are shown in Fig. 2.3.5. The exponential phase of UV2 strain was upto 12 h followed by the stationary phase upto 84 h followed by the death phase. The growth profile of UV2 strain was similar to the wild-type strain. The CMCase activity of UV2 strain increased by 22% as compared with wild-type strain (0.65 U/mL) at approximately, same optical density of cells (2.7) at 48 h of fermentation. CMCase production in

TH-2393_146106034

UV2 mutant strain was not associated with growth as maximum production occurred after log phase similar to the earlier reports for wild-type strain in unoptimized medium (Singh *et al.,* 2013). The total protein production from UV2 strain was 3.25 mg/mL, approximately similar to the wild-type strain (3.15 mg/mL) at 48 h of fermentation. These observations showed that the increase in cellulase activity is due to the mutation in gene encoding cellulase. The pH of the fermenting medium increased from 5.6 to 9.4 during the growth of UV2 mutant strain, approximately similar to the wild-type strain. The increase in pH was observed in log phase and stationary phase and after 84 h increase in pH was negligible due to the onset of the death phase. This increase in pH value of fermenting medium maybe due to the presence of extracellular proteins in which organic amino compounds are deaminated with the growth of bacterium (Ladeira *et al.*, 2014).

TH-2393_146106034

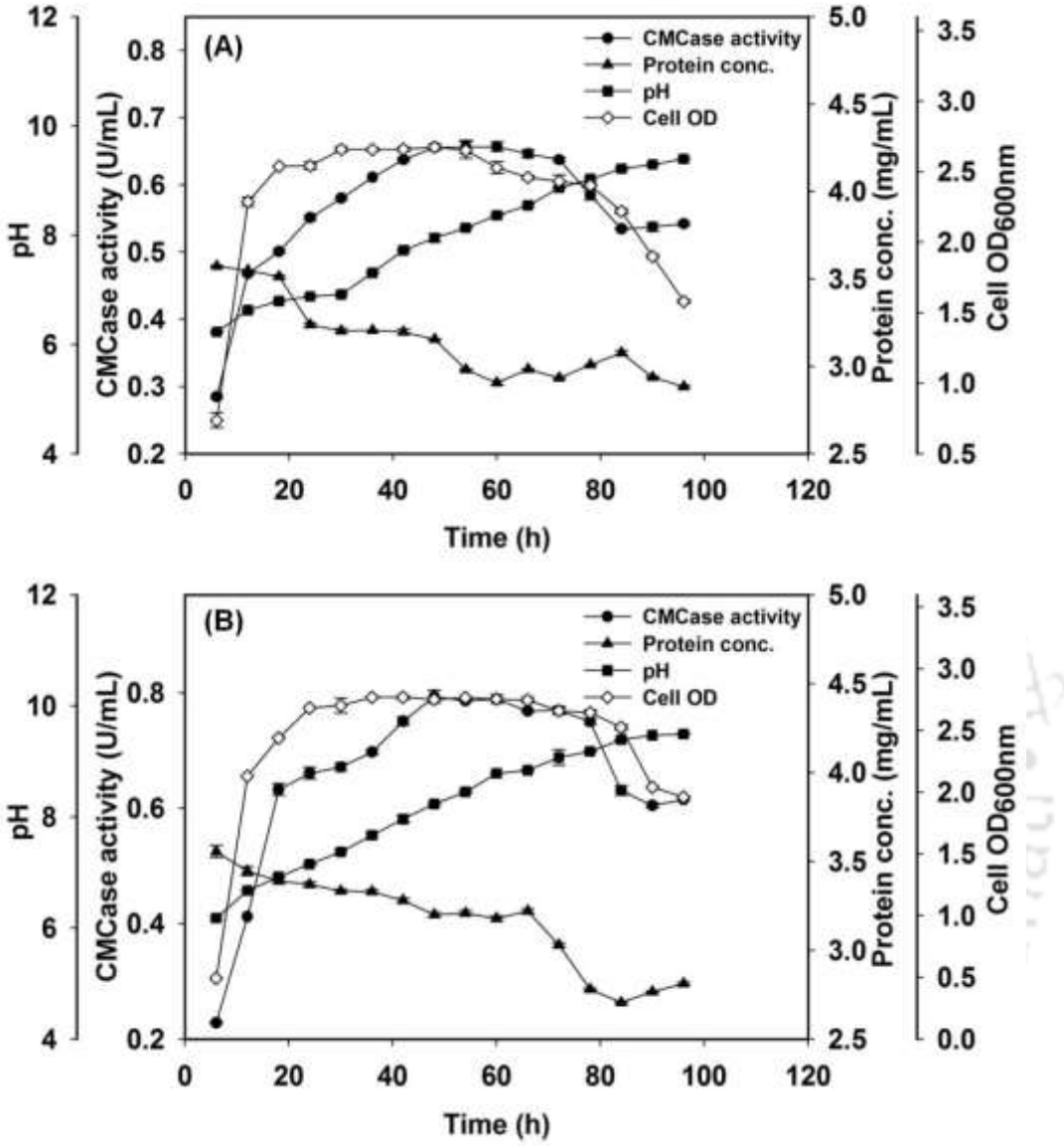


Fig. 2.3.5 Fermentation profiles of (A) Wild-type *Bacillus amyloliquefaciens* SS35 and (B) its mutant UV2 strain. The endoglucanase activity, optical density (600 nm), protein concentration and pH change of the broth were monitored at every 6 h up to 96 h.

Table 2.3.5 Fermentation profile parameters for wild-type and UV2 mutant enzyme after 48 h

	CMCase-WT	CMCase-UV2
Cell OD_{600nm}	2.70 ± 0.02	2.70 ± 0.01
pH	8.00 ± 0.10	8.20 ± 0.02
Enzyme activity (U/mL)	0.65 ± 0.01	0.79 ± 0.01
Protein concentration (mg/mL)	3.15 ± 0.01	3.25 ± 0.01

Values are mean ± SE (n=3).

TH-2393_146106034

2.4 Conclusion

The present study demonstrated that UV/EMS directed evolution of *Bacillus amyloliquefaciens* SS35 enhanced the CMCase production in the cell free supernatant. Wild-type strain of *Bacillus amyloliquefaciens* SS35 was mutated in two stages first by UV irradiation and second EMS. In the first stage of UV mutagenesis, 180 min of UV exposure for wild-type strain (*Bacillus amyloliquefaciens* SS35) was sub-lethal resulting 1% of survivors. UV mutant colonies were screened on the basis of Plate staining method by 0.3% (w/v) Congo red and CMCase production. Out of 14 UV mutant strains only 3 strains i.e. UV1, UV2 and UV4 exhibited the significant increase in CMCase activity i.e. 17%, 22% and 14%, respectively, as compared to the wild-type strain (0.65 U/mL). Therefore, these three UV mutant strains were selected for genetic stability of mutation. Among all the three UV mutants, only UV2 mutant strain retained 100% of the relative CMCase activity upto 10 generations which suggested that the mutation is genetically stable. Therefore, UV2 strain was selected for the second stage mutagenesis by using a chemical mutagen EMS. EMS dose, 2% (v/v) was sublethal for UV2 mutant strain which results only 0.1% survivors. EMS mutant colonies were also screened in a similar manner as UV mutant colonies screened. Only 4 EMS mutant colonies showed increase in CMCase activity for crude cell free supernatant as compared with the wild-type strain. EMS 2, EMS 3, EMS7 and EMS 8 mutant strains displayed 3%, 4%, 7% and 3%, respectively, increase in the CMCase activity as compared to the UV2 strain (0.79 U/mL). Among all four EMS mutants, none retained the relative CMCase activity upto 10 generations suggested that the mutation is not genetically stable. Therefore, on the basis of enzyme production and generation stability, UV2 mutant strain was the most potential mutant

TH-2393_146106034

strain and used for further studies. The fermentation profile of UV2 mutant strain showed that maximum CMCase production occurs at 48 h of fermentation in the late log phase. The pH of the fermenting medium increased from acidic to alkaline during the growth of UV2 mutant strain, approximately similar to the wild-type strain.

TH-2393_146106034

2.5 References

Agrawal, R., Satlewal, A., and Verma, A.K. (2013). Development of a β-glucosidase hyperproducing mutant by combined chemical and UV mutagenesis. *3 Biotech*, 3(5), 381-388.

Al-Jailawi M.H., Al-Shekdhaher, A., Al-Zaiadi, R. (2016). Genetic improvement of *Saccharomyces boulardii* R7 and generate suitable strains for synthesis and expression of recombinant products. *British Biotechnology Journal*, 11, 1-9

Bayer, E.A., Lamed, R., and Himmel, M.E. (2007). The potential of cellulases and cellulosomes for cellulosic waste management. *Current Opinion in Biotechnology*, 18(3), 237-245.

Bhuyan, P.M., Sandilya, S.P., Nath, P.K., Gandotra, S., Subramanian, S., Kardong, D., and Gogoi, D.K. (2018). Optimization and characterization of extracellular cellulase produced by *Bacillus pumilus* MGB05 isolated from midgut of muga silkworm (*Antheraea assamensis* Helfer). *Journal of Asia-Pacific Entomology*, 21(4), 1171-1181.

Caniago, A., Mangunwardoyo, W., Nuswantara, S., and Lisdiyanti, P. (2015). Improvement of endoglucanase activity in *Penicillium oxalicum* ID10-T065 by ultra violet irradiation and ethidium bromide mutation. *Annales Bogorienses*, 19, 2.

TH-2393_146106034

Chandra, M., Kalra, A., Sangwan, N.S., Gaurav, S.S., Darokar, M.P., and Sangwan, R.S. (2009). Development of a mutant of *Trichoderma citrinoviride* for enhanced production of cellulases. *Bioresource Technology*, 100(4), 1659-1662.

Chen, Z., Pereira, J.H., Liu, H., Tran, H.M., Hsu, N.S., Dibble, D., and Simmons, B. A. (2013). Improved activity of a thermophilic cellulase, Cel5A, from *Thermotoga maritima* on ionic liquid pretreated switchgrass. *PLoS One*, 8(11), e79725.

Cherry, J.R., Fidantsef, A. L. Directed evolution of industrial enzymes: an update (2003). *Current Opinion in Biotechnology*,14(4), 438-443.

Deka, D., Jawed, M., and Goyal, A. (2013). Purification and characterization of an alkaline cellulase produced by *Bacillus subtilis* (AS3), *Preparative Biochemistry and Biotechnology*, 43, 256-270.

Dillon, A.J.P., Zorgi, C., Camassola, M., Henriques, J.A. (2006) Use of 2-deoxyglucose in liquid media for the selection of mutant strains of *Penicillium echinulatum* production increased cellulase increased cellulase and β-glucosidase activities. *Applied Microbiology Biotechnology*, 70, 740-746

Goel, N., Patra, R., and Verma, S.K. (2019). Purification and characterization of cellulase from *Pseudomonas* sp. isolated from waste dumping site soil. *Journal of Applied Biotechnology Bioengineering*, 6(3), 118-124.

Hakamada, Y., Koike, K., Yoshimatsu, T., Mori, H., Kobayashi, T., and Ito, S. (1997). Thermostable alkaline cellulase from an alkaliphilic isolate, *Bacillus* sp. KSM-S237. *Extremophiles*, 1(3), 151-156.

TH-2393_146106034

Haq, I., Ali. S., Javed, M.M., Hameed, U., Saleem, A., Adnan, F., Qadeer, M.A. (2010). Production of alpha amylase from a randomly induced mutant strain of *Bacillus amyloliquefaciens* and its application as a desizer in textile industry. *Pakistan Journal of Botany*, 42, 473-484.

Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W., and Foust, T.D. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science*, 315(5813), 804-807.

Immanuel, G., Dhanusha, R., Prema, P., and Palavesam, A. (2006) Effect of different growth parameters on endoglucanase enzyme activity by bacteria isolated from coir retting effluents of estuarine environment. *International Journal of Environmental Science and Technology*, 3, 25-34.

Kaur, B., Oberoi, H.S., and Chadha, B.S. (2014). Enhanced cellulase producing mutants developed from heterokaryotic *Aspergillus* strain. *Bioresource Technology*, 156, 100-107.

Kim, B.K., Lee, B.H., Lee, Y.J., Jin, I.H., Chung, C.H., and Lee, J.W. (2009). Purification and characterization of carboxymethylcellulase isolated from a marine bacterium, *Bacillus subtilis* subsp. subtilis A-53. *Enzyme and Microbial Technology*, 44(6-7), 411-416.

Kim, J.M., Kong, I.S., and Yu, J.H. (1987). Molecular cloning of an endoglucanase gene from an alkalophilic *Bacillus* sp. and its expression in *Escherichia* coli. *Applied and Environmental Microbiology*, 53(11), 2656-2659.

Kostyleva, E.V., Tsurikova, N.V., Sereda, A.S., Velikoretskaya, I.A., Veselkina, T.N., Lobanov, N.S., and Sinitsyn, A.P. (2018). Enhancement of activity of

TH-2393_146106034

carbohydrases with endo-depolymerase action in *Trichoderma reesei* using mutagenesis. *Microbiology*, 87(5), 652-661.

Kotchoni, S.O., Gachomo, E.W., Omafuvbe, B.O., and Shonukan, O.O. (2006). Purification and biochemical characterization of carboxymethyl cellulase (CMCase) from a catabolite repression insensitive mutant of *Bacillus pumilus*. *International Journal of Agriculture and Biological Sciences*, 2, 286-392.

Kumar, A.K. (2015). UV mutagenesis treatment for improved production of endoglucanase and β-glucosidase from newly isolated thermotolerant actinomycetes, *Streptomyces griseoaurantiacus*. *Bioresources and Bioprocessing*, 2(1), 22.

Ladeira, S.A., Cruz, E., Delatorre, A.B., Barbosa, J.B., Leal Martins, M.L. (2015). Cellulase production by thermophilic *Bacillus* sp: SMIA-2 and its detergent compatibility. *Electronic Journal of Biotechnolgy*, 18(2), 110-115.

Lawford, H.G., and Rousseau, J.D. (2003). Cellulosic fuel ethanol. In Biotechnology for Fuels and Chemicals. *Humana Press, Totowa, NJ*, 457-469.

Li, X.H., Yang, H.J., Roy, B., Park, E.Y., Jiang, L.J., Wang, D., and Miao, Y.G. (2010). Enhanced cellulase production of the *Trichoderma viride* mutated by microwave and ultraviolet. *Microbiological Research*, 165(3), 190-198.

Lim, J., Choi, Y.H., Hurh, B.S., and Lee, I. (2019). Strain improvement of *Aspergillus sojae* for increased L-leucine aminopeptidase and protease production. *Food Science and Biotechnology*, 28(1), 121-128.

TH-2393_146106034

Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. (1951). Protein measurement with the Folin phenol reagent. *The Journal of Biological Chemistry*, 193(1), 265-275.

Lynd, L.R., Laser, M.S., Bransby, D., Dale, B.E., Davison, B., Hamilton, R., Himmel, M., Keller, M., McMillan, J.D., Sheehan, J., and Wyman, C.E. (2008). How biotech can transform biofuels. *Nature Biotechnology*, 26(2), 169.

Maki, M., Leung, K.T. and Qin, W. (2009). The prospects of cellulase-producing bacteria for the bioconversion of lignocellulosic biomass. *International Journal of Biological Sciences*, 5, 500-516.

Molzahn, S. (1977). A new approach to the application of genetics to brewing yeast. *Journal of American Soceity of Brewing Chemists*, 35(2), 54-59.

Naresh, S., Kunasundari, B., Gunny, A.A.N., Teoh, Y.P., Shuit, S.H., Ng, Q.H., and Hoo, P.Y. (2019). Isolation and partial characterisation of thermophilic cellulolytic bacteria from North Malaysian tropical mangrove soil. *Tropical Life Sciences Research*, 30(1), 123-147.

Nelson, N. (1944). A photometric adaptation of the Somogyi method for the determination of glucose. *Journal of Biological Chemistry*, 153(2), 375-380.

Patel, S., and Goyal, A. (2010). Isolation, characterization and mutagenesis of exopolysaccharide synthesizing new strains of lactic acid bacteria. *Internet Journal of Microbiology*. 8(1), Internet Scientific Publications LLC, USA.

TH-2393_146106034

Rajoka, M.I. (2005). Double mutants of *Cellulomonas biazotea* for production of cellulases and hemicellulases following growth on straw of a perennial grass. *World Journal of Microbiology and Biotechnology*, 21(6-7), 1063-1066.

Rapa, R.A., Islam, A., Monahan, L.G., Mutreja, A., Thomson, N., Charles, I.G., and Labbate, M. (2015). A genomic island integrated into recA of *Vibrio cholerae* contains a divergent recA and provides multipathway protection from DNA damage. *Environmental Microbiology*, 17(4), 1090-1102.

Raza, A., Bashir, S., and Tabassum, R. (2019). Evaluation of cellulases and xylanases production from *Bacillus* spp. isolated from buffalo digestive system. *Kafkas Universitesi Veteriner Fakultesi Dergisi* 25(1), 39-46.

Ruijssenaars, H.J., and Hartmans, S. (2001). Plate screening methods for the detection of polysaccharase-producing microorganisms. *Applied Microbiology and Biotechnology*, 55, 143-149.

Sangkharak, K., Vangsirikul, P., and Janthachat, S. (2012). Strain improvement and optimization for enhanced production of cellulase in *Cellulomonas* sp. TSU-03. *African Journal of Microbiology Research*, 6(5), 1079-1084.

Shobharani, P., Yogesh, D., Halami, P.M., and Sachindra, N.M. (2013). Potential of cellulase from *Bacillus megaterium* for hydrolysis of sargassum, *Journal of Aquatic Food Product Technology*, 22(5), 520-535.

Simpson, I., and Caten, C. (1979). Recurrent mutation and selection for increased penicillin titre in *Aspergillus nidulans*. *Microbiology*, 113, 209-217.

Sims, R.E., Mabee, W., Saddler, J.N., and Taylor, M. (2010). An overview of second generation biofuel technologies. *Bioresource Technology*, 101(6), 1570-1580.

TH-2393_146106034

Singh, S., Dikshit, P.K., Moholkar, V.S., and Goyal, A. (2014). Purification and characterization of acidic cellulase from *Bacillus amyloliquefaciens* SS35 for hydrolyzing *Parthenium hysterophorus* biomass. *Environmental Progress and Sustainable Energy*, 34(3), 810-818.

Singh, S., Moholkar, V.S., and Goyal, A. (2013). Isolation, identification and characterization of a cellulolytic *Bacillus amyloliquefaciens* strain SS35 from Rhinoceros Dung. *ISRN Microbiology*, 728134, 7.

Singh, S., Moholkar, V.S., and Goyal, A. (2014). Optimization of carboxymethylcellulase production from *Bacillus amyloliquefaciens* SS35. *3 Biotech*, 4(4), 411-424.

Somogyi, M.A., (1945). New reagent for the determination of sugars. *Journal of Biological Chemistry*, 160, 61-68.

Teather, R.M., and Wood, P.J. (1982). Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Applied and Environmental Microbiology*, 43, 777-780.

Witkin, E.M. (1976). Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriological Reviews*, 40(4), 869.

Wooley, R., Ruth, M., Glassner, D., and Sheehan, J. (1999). Process design and costing of bioethanol technology: a tool for determining the status and direction of research and development. *Biotechnology Progress*, 15(5), 794-803.

Zhang, X.Y., Zi, L.H., Ge, X.M., Li, Y.H., Liu, C.G., and Bai, F.W. (2017). Development of *Trichoderma reesei* mutants by combined mutagenesis and

TH-2393_146106034

induction of cellulase by low-cost corn starch hydrolysate. *Process Biochemistry*, 54, 96-101.

TH-2393_146106034

Chapter 3

Comparative characterization of endoglucanase from UV mutant UV2 with wild-type enzyme from *Bacillus amyloliquefaciens* SS35

3.1 Introduction

Cellulose is the most abundant, biodegradable and renewable carbohydrate present in the biosphere. Therefore, lignocellulosic biomass can serve as the most potential substrate for the production of liquid transportation biofuels. Cellulose is a linear polysaccharide composed of β-(1→4) linked β-D-glucopyranose repeating units (Chen 2014). Microbial cellulases are the biocatalysts which convert the cellulosic polysaccharide into the fermentable hexose sugar for production of bioethanol. Moreover, cellulases have other applications in paper and pulp industry for decreasing the fibre coarseness and pulp viscosity and in textile industry for biopolishing of cellulosic fabrics and biostoning of jeans (Kuhad *et al.*, 2010). Nowadays, Bacterial cellulases are getting much more attention than fungal cellulases because production cost of fungal cellulase is higher than bacterial cellulase. The partial purification of extracellular cellulase from *Bacillus subtilis* (Deka *et al.*, 2013) and *Bacillus amyloliquefaciens* SS35 (Singh *et al.*, 2014) by ammonium sulfate precipitation method has reported. Ammonium sulfate is a most frequently used precipitants for

TH-2393_146106034

salting out of proteins, because it has very high solubility upto 4 M, very low heat of solution and stabilizes the protein due to preferential solvation (Englard and Seifter 1990). Partially purified proteins were further purified by gel filtration in order to increase the purity level of protein (Yin *et al.*, 2010). Acidic CMCase from wild-type strain of *Bacillus* spp. such as *Bacillus* sp. M-9 (Bajaj *et al.*, 2009), *Bacillus thuringiensis* (Lin *et al.*, 2012) and *Bacillus amyloliquefaciens* SS35 (Singh *et al.*, 2014) was reported. Thin layer chromatography helps in the determination of mode of action of cellulase i.e. endo/exo on polysaccharides. Mode of action of endoglucanase from wild-type strains of *Bacillus licheniformis* (Bischoff and Rooney 2006), *Bacillus* sp. (Aygan *et al.*, 2008) and *Bacillus circulans* (Kim 1995) was studied by analysing the enzyme hydrolysed products (cellooligosaccharides and glucose) on TLC. In this study, the purification of endoglucanase, CMCase-UV2 from UV2 mutant strain (Chapter 2, Section 2.3.3) of *Bacillus amyloliquefaciens* SS35 was compared with the endoglucanase, CMCase-WT from wild-type strain. The comparison of biochemical characterization of the CMCase-UV2 enzyme with CMCase-WT enzyme in terms of activity staining by zymogram for identifying the active form of the protein of appropriate molecular mass by polyacrylamide gel electrophoresis, change in pH and temperature optima, increase in catalytic efficiency by kinetic studies, thermostability and pH stability was carried out. The mode of action of CMCase-UV2 was compared with CMCase-WT enzyme after TLC analysis. Furthermore, the enzymatic hydrolysis of the pretreated lignocellulosic biomass (Elephant grass) by CMCase-WT and CMCase-UV2 enzyme was carried out.

TH-2393_146106034

3.2 Materials and Methods

3.2.1 Media and chemicals

Carboxymethylcellulose sodium salt (CMC-Na) (low viscosity), lichenan, β-glucan, avicel, birchwood xylan, oat spelt xylan, galactomannan, D-glucose, cellobiose, sodium dodecyl sulfate (SDS) and components of polyacrylamide gel electrophoresis were procured from Sigma-Aldrich Co. LLC., USA. Superdex^{TM} 75pg size exclusion column was purchased from General Electronics Healthcare Biosciences, Sweden. The Page Ruler protein marker was procured from BioBharati Life Science Pvt. Ltd, India. All medium components, chemicals and cellulose powder used in this study were procured from HiMedia Pvt. Ltd., India. Hydrochloric acid, sulphuric acid, n-butanol, glacial acetic acid, silver nitrate and TLC plate were purchased from Merck (India) and α-naphthol was purchased from Fisher Scientific, Germany.

3.2.2 Microorganism and culture conditions

Wild-type strain of *Bacillus amyloliquefaciens* SS35 isolated from Rhinoceros dung (Singh *et al.*, 2013) and its UV mutant strain, UV2 (Chapter 2, Section 2.3.3) was used in this study. The medium composition and culture conditions for CMCase production from UV2 mutant strain was same as of wild-type *Bacillus amyloliquefaciens* SS35 as given earlier in chapter 2 Section 2.2.2. The extracellular enzyme in cell free supernatant of wild-type and UV2 mutant strain was harvested by centrifuging the culture broth at 12,000g for 20 min at 4°C after 48 h.

TH-2393_146106034

3.2.3 Purification of endoglucanase from wild-type and UV2 mutant strain of *Bacillus amyloliquefaciens* SS35

3.2.3.1 Partial purification of enzymes by ammonium sulfate precipitation method

The extracellular CMCase enzyme produced by wild-type strain (CMCase-WT) and UV2 mutant (CMCase-UV2), present in cell-free supernatant was partially purified by ammonium sulfate method as described earlier by Wood (1976). Ammonium sulfate upto 90% saturation was gradually added in 200 mL cell-free supernatant and incubated in stirring condition at 4°C for 12 h. The precipitated protein was separated from the supernatant by centrifugation at 8,000g for 30 min at 4°C. The precipitated pellet of protein was dissolved in 5 mL of 50 mM sodium acetate buffer, pH 5.0. The desalting of the precipitated enzyme was carried out by using 50 mM sodium acetate buffer (pH 5.0) with 4 changes of 2000 mL buffer for 12 h using a 10 kDa cutoff dialysis membrane. Further, the dialysed enzymes were concentrated to 2 mL volume using 15 mL, 10 kDa MWCO concentrators (Amicon Ultra, Merk-Millipore, USA). The reducing sugar was measured by the Nelson (1944) and Somogi (1945) method as described in chapter 2, section 2.2.5 and the CMCase activity of the enzyme(s) were calculated as described in Chapter 2, Section 2.2.5.2. The protein concentration of partially purified enzyme(s) were estimated by the Bradford method using BSA as standard as described in section 3.2.4.

3.2.3.2 Size exclusion chromatography

The partially purified CMCase-WT and CMCase-UV2 enzymes were further purified by using Fast Protein Liquid Chromatography (FPLC, GE Healthcare, USA). The Superdex^{TM} 75pg column (GE Healthcare, 16 mm column ID, 600 mm bed height) was used for protein purification by size exclusion chromatography. The

TH-2393_146106034

Superdex column was pre-equilibrated with 100 mL of 50 mM sodium acetate buffer, pH 5.0. After that 2 mL (1 mg/mL) of the enzyme(s) was loaded on to the column. The enzyme(s) was eluted by using the same buffer at a flow rate of 0.8 mL/min with 1 mL fraction size. The fractions containing CMCase activity (U/mL) were pooled. The CMCase activity of the enzyme(s) was measured as described in chapter 2, section 2.2.5. The protein concentration of purified enzyme was estimated by the Bradford method using BSA as standard as described in section 3.2.4.

3.2.4 Protein estimation by Bradford method

The concentrations of partially purified and purified protein(s) were estimated by the Bradford's method (Bradford, 1976). A concentration range 10-100 µg/mL Bovine serum albumin (BSA) was used for ploting standard curve. To 100 µL of reaction mixture containing 50 µL protein, 1.0 mL of Bradford reagent was added and kept in dark for 15 min. After 15 min, the absorbance at 595 nm (A595) was measured against a blank. Preparation of Bradford reagent is given in subsequent subsection.

3.2.4.1 Preparation of Bradford reagent

The Bradford assay estimates the amount of protein in a solution by using the spectral properties of Coomassie Brilliant Blue G-250 (Bradford, 1976). 100 mg 0.01% (w/v) Coomassie Brilliant Blue G-250 was weighed and dissolved in 50 mL 95% ethanol (in an amber colour bottle). 100 mL 85% (w/v) phosphoric acid was added to it. A magnetic bead was placed inside the bottle and the contents were mixed properly by keeping on magnetic stirrer until the dye completely dissolved. The dye was finally diluted to 1 L with deionized water, filtered (Whatman, No. 1 paper) under dark conditions and stored at 4°C.

TH-2393_146106034

The concentration of proteins were estimated by using the following equation,

$$[\text{Protein}] = \frac{\Delta A_{595} \times V \times C}{v}$$

Where,

A_{595} = change in absorbance of the sample
V = volume of the protein-buffer mixture (mL)
C = 1 OD equivalent of BSA from standard plot (mg/mL)
v = volume of the enzyme used for assay (mL)

3.2.5 Sodium dodecyl sulfate-Polyacrylamide gel electrophoresis (SDS-PAGE) analysis of wild-type and UV mutant protein

3.2.5.1 Preparation of SDS-PAGE gel

The components of a SDS-PAGE gel are acrylamide 30% (w/v), resolving gel (Tris-HCl, pH 8.8), a stacking gel (Tris-HCl, pH 6.8), SDS 10% (w/v), APS 10% (w/v), N,N,N',N'-tetramethylethane-1,2-diamine (TEMED), sample loading buffer (pH 6.8) and electrophoresis running buffer (pH 8.3-8.5). The composition of each component of SDS-PAGE gels and buffers are described below in Sections 3.2.5.2 to 3.2.5.5.

3.2.5.2 Preparation of acrylamide 30% (w/v) solution

0.8 g of bis-acrylamide was weighed and transferred into an amber colour bottle and dissolved in 50 mL of ultra-pure deionized water collected at 18 MΩcm from Milli-Q water purification system (Millipore, USA) using a magnetic stirrer (IKA, C-MAG HS7, Germany). After completely dissolving bis-acrylamide, 29.2 g of acrylamide was added to and stirred on a magnetic stirrer till the solution was clear. The final volume was adjusted to 100 mL with ultra-pure water by keeping the measuring cylinder (100 mL) wrapped in aluminium foil as acrylamide is light sensitive. The acrylamide solution was then filtered (Whatman No. 1) under dark condition and stored at 4°C.

TH-2393_146106034

3.2.5.3 Polymerization of SDS-PAGE gel

The resolving gel and stacking gel were prepared by following the protocols from Sambrook *et al.*, (1989) using the composition as described in Tables 3.2.1 and 3.2.2. The resolving gel was prepared by adding the all the components in the order as mentioned in Table 3.2.1, in a 25 mL beaker, by keeping acrylamide concentration at 12% (w/v). Similarly, the stacking gel (4%, w/v) was prepared by dissolving all the components mentioned in Table 3.2.2.

Table 3.2.1 Composition of SDS-PAGE for preparation of resolving gel.

Components	**12% gel volume (mL)**
Acrylamide solution (30%, w/v)	4.0
Deionized water	0.69
SDS (10%, w/v)	1.0
Glycerol (50%, v/v)	1.0
1.5 M Tris-HCl (pH 8.8)	3.3
Amonium per sulfate (10%, w/v)	0.1
TEMED	0.01

Table 3.2.2 Composition of SDS-PAGE components for preparation of stacking gel.

Components	**4% gel volume (mL)**
Acrylamide solution (30%, w/v)	0.7
Deionized water	2.8
SDS (10%, w/v)	0.5
0.5 M Tris-HCl (pH 6.8)	1.0
Amonium per sulfate (10%, w/v)	0.05
TEMED	0.005

TH-2393_146106034

3.2.5.4 Preparation of SDS-PAGE running buffer

The SDS-PAGE gels were run using a 1x running or tank buffer prepared from the 10x stock solution as described in Table 3.2.3. 30.28 g of Tris free base and 188 g of glycine were dissolved in 800 mL of deionized water. To this 100 mL of 10% (w/v) SDS was added and the final volume was adjusted to 1 L. The final pH of the buffer was adjusted to 8.3. The 10x buffer was filtered (Whatman, Filter No. 1) and stored at 4°C.

Table 3.2.3 Composition of 10x Tris-Glycine, running or tank buffer.

Components	Final concentration (10x buffer)
Tris base	0.25 M
Glycine	2.5 M
SDS	1.0% (w/v)

3.2.5.5 Preparation of sample buffer

5x sample loading buffer was prepared by dissolving the components while keeping the concentration of components as described in Table 3.2.4 and the pH of the buffer was adjusted to 6.8. The components were dissolved in the order as mentioned in Table 3.2.4 to make 5x sample buffer. However, the final concentration while loading to a SDS-PAGE gel was always kept to 1x by mixing 4 volumes of sample (protein) with 1 volume of 5x sample buffer.

Table 3.2.4 Composition of 5x sample loading buffer (Laemmli, 1970).

Components	Final concentration (5x buffer)
Tris-HCl (pH 6.8)	62.5 mM
Glycerol	20.0 (%, v/v)
SDS	2.0 (%, w/v)
Bromophenol Blue	0.025 (%, w/v)
β-mercaptoethanol	5.0 (%, w/v)

TH-2393_146106034

3.2.5.6 Silver staining

3.2.5.6.1 Preparation of reagents for silver staining

Silver staining of protein gel after electrophoresis is 50 times more sensitive than Coomassie Brilliant Blue therefore, provides admirable sensitivity for protein detection in 0.2-2 ng range (Rabilloud, 1992). The solutions used in silver staining are listed in Table 3.2.5 and the staining procedure is described in section 3.2.5.6.2

Table 3.2.5 Composition of reagents used in silver staining

Components	Final concentration (5x buffer)
Fixing solution	40% (v/v) ethanol, 10% acetic acid in water
Sensitizing solution	0.2% (w/v) $Na_2S_2O_3$, 6.8% (w/v) sodium acetate, 30% (v/v) ethanol and 0.025% (v/v) glutaraldehyde
Silver nitrate solution	0.1% (w/v) $AgNO_3$, 0.076% (w/v) formalin
Developing solution	2.5% (w/v) Na_2CO_3, 0.05% (w/v) formaldehyde
Stop solution	1.46% (w/v) EDTA
Preserving solution	30% (v/v) ethanol, 4% (v/v) glycerol

3.2.5.6.2 Silver staining procedure

Silver staining of the gel was performed according to the method as reported earlier by Rabilloud (1992). After the electrophoresis, the gel was immersed in 50 mL of fixing solution at 25°C for at 30 min to fix the protein bands and placed on a shaker at a very gentle speed. After 30 min, the solution was discarded and the gel was washed in 20% ethanol for 20 min to remove the remaining detergent as well as fixing solution from the gel. The gel was then incubated in 100 mL sensitizing solution at 25°C for 20-30 min. After the incubation, the sensitizing solution was discarded and the gel was washed thrice by deionised water at every 10 min intervals. The silver nitrate solution of 100 mL was added to the gel and incubated for 20 min in dark to allow the silver ions to bind with the proteins. After staining, the gel was rinsed with

TH-2393_146106034

100 mL of deionised water for 2-3 min and then 100 mL of developing solution was added to the gel. The reaction was stopped as soon as the protein bands appeared in desired intensity by adding 50 mL of stop solution. The gel was stored in preserving solution.

3.2.6 SDS-PAGE and Zymogram analysis of purified enzymes

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using 12% (w/v) gel was carried out under denaturing conditions according to the protocol as described earlier by Laemmli (1970) to identify the purity and molecular mass of proteins. 0.3% (w/v) CMC-Na was added in the resolving gel for activity staining of CMCase by zymogram. The protein samples were mixed with 5x loading buffer in the ratio of 4:1 and were subjected to heat denaturation for 5 min in boiling water bath. The protein samples for zymogram were mixed with 5x loading buffer (without β-mercaptoethanol) in the ratio of 4:1. The electrophoresis was carried out using 1x running buffer at 60 V. Molecular protein marker (14.3-120 kDa), was used as standard for SDS-PAGE. After the electrophoresis, the gel was cut into two parts, one part was subjected to silver staining protocol as described earlier (Wray et al., 1981) and other part was used for activity staining by zymogram. For zymogram analysis, the gel was renatured in 2.5% (v/v) Triton X-100 in 50 mM sodium acetate buffer (pH 5.0) upto 4 h (Singh *et al.*, 2014). After removal of Triton X-100, the gel was washed with 50 mM sodium acetate buffer (pH 5.0) and incubated at 50°C for 6 h. Thereafter, the gel was stained by 0.3% (w/v) Congo red for 15 min and destained further with 1M NaCl.

TH-2393_146106034

3.2.7 Effect of pH on CMCase-WT and CMCase-UV2 activity and stability

The optimum pH of the purified CMCase-WT and CMCase-UV2 enzymes were determined with 3 µL of each enzyme (0.1 mg/mL) in 1% (w/v) CMC-Na in appropriate buffers; 50 mM sodium acetate (pH 3.5-5.5), 50 mM sodium phosphate buffer (pH 5.8-8.0) and 50 mM Tris-HCl (pH 8-9.5). The reaction mixture at different pH were incubated at 65°C for 5 min and CMCase activity was assayed by reducing sugar estimation Nelson (1944) and Somogyi (1945) method as described in chapter 2, section 2.2.5. The comparison of pH stability of the purified CMCase-WT and CMCase-UV2 enzymes was performed by pre-incubating the enzyme (10 µL of each enzyme, 0.1 mg/mL) at 30°C for 1 h under different buffers at a pH ranging from 3.5 to 9.5. The residual activity for respective samples was determined by Nelson (1944) and Somogyi (1945) method of reducing sugars estimation as described in chapter 2, section 2.2.5.

3.2.8 Effect of temperature on CMCase-WT and CMCase-UV2 activity and stability

The optimum temperature of the purified CMCase-WT and CMCase-UV2 for enzyme assay was determined by reducing sugars estimation as described in chapter 2, section 2.2.5. The reaction mixture containing 3 µL of each enzyme (0.1 mg/mL) at 1% (w/v) CMC-Na in 50 mM sodium acetate buffer, pH 5.0 was incubated in temperature range, 30°C-75°C for 5 min. The thermostability studies of the purified CMCase-WT and CMCase-UV2 was performed by pre-incubating in 50 mM sodium acetate buffer, pH 5.0 at different temperatures ranging from 30 to 75°C for 1 h. The residual enzyme activity for samples was determined by reducing sugar estimation as described in chapter 2, section 2.2.5.

TH-2393_146106034

3.2.9 Comparison of enzyme activity of purified CMCase-WT and CMCase-UV2 against soluble and insoluble substrates

The enzyme activity of purified CMCase-WT and CMCase-UV2 was tested against β-D-glucan, lichenan, CMC-Na, cellulose powder, avicel, birchwood xylan and galactomannan. The enzyme reactions were carried out with 1.0% (w/v) substrate dissolved in 50 mM sodium acetate buffer, pH 5.0 by incubating at 65°C for 5 min. 100 µL of enzyme reaction mixture contained 3 µL of the enzyme (CMCase-WT, 0.1 mg/mL and CMCase-UV2, 0.1 mg/mL). In the case of avicel and oat spelt xylan, the reaction mixture was incubated under shaking conditions (160 rpm, 5 min). The activity of CMCase-WT and CMCase-UV2 was also compared against the pretreated substrate *Parthenium hysterophorus* (Carrot grass) and *Pennisetum purpureum* (Elephant grass). The Carrot grass and Elephant grass biomasses were pretreated by the method described by Singh *et al.*, (2014) and Eliana *et al.*, (2014), respectively. The reactions were carried out with 2.0% (w/v) pretreated biomass dissolved in 50 mM sodium acetate buffer (pH 5.0) by incubating at 65°C for 15 min at 160 rpm. The resulting reducing sugars were determined by Nelson (1944) and Somogyi (1945) method reported earlier as described in chapter 2, section 2.2.5 by measuring the absorbance at 500 nm, on a spectrophotometer (Varian, Cary 100 Bio) using glucose, xylose and mannose as standards.

3.2.10 Comparison of kinetic parameters of CMCase-WT and CMCase-UV2

The kinetic parameters i.e. V_{max} and K_m of CMCase-WT and CMCase-UV2 were determined by fitting the initial rate data to the Michaelis-Menten equation and Lineweaver-Burk double reciprocal plot (Michaelis and Menten 1913). V_{max} is the maximum rate achieved by the system at saturating substrate concentration and K_m is

TH-2393_146106034

the Michaelis constant which define as the substrate concentration at which reaction rate is half of the V_{max} (Michaelis and Menten 1913). The following equation describes the Michaelis-Menten equation

$$v=d[P]/dt=(V_{max}[S])/(K_m+[S])$$

v = Rate of enzyme reaction

d[P]/dt = Rate of product formation

[S] = Concentration of substrate

The enzyme reaction was carried out with the substrates (CMC-Na and β-D-glucan) concentration ranging from 0.01% to 2% (w/v) dissolved in 50 mM sodium acetate buffer (pH 5.0) incubated at 65°C for 5 min. The activity was determined by reducing sugar estimation according to the Nelson (1944) and Somogyi (1945) method as described in chapter 2, section 2.2.5. The data for kinetic parameters were analyzed by GraphPad Prism software (GraphPad Software Inc., San 138 Diego, CA) and the kinetic constants were calculated from the best fit.

3.2.11 Thin Layer Chromatography analysis of hydrolyzed products of CMC-Na and β-glucan by CMCase-WT and CMCase-UV2

The qualitative analysis of hydrolyzed products of CMC-Na and β-D-glucan by CMCase-WT and CMCase-UV2 was performed by Thin Layer Chromatography (TLC) on silica gel-coated aluminium plate (TLC Silica gel 60 F254, 20620 cm, Merck). The enzyme (2.5 mM) catalysed reaction with 1% (w/v) CMC-Na and β-D-glucan in 100 µL reaction mixture was incubated at 50°C in 50 mM sodium acetate buffer (pH 5.0), for time intervals of 5, 10, 15, 30, 45, 60, 120, 180, 360, 720 and 1440 min. To find out the action of both the enzymes on cellobiose, CMCase-WT and CMCase-UV2 enzyme with equimolar concentrations of 2.5 mM were incubated in

TH-2393_146106034

100 µL reaction mixture containing 0.5% (w/v) cellobiose dissolved in 50 mM sodium acetate buffer, pH 5.0 at 50°C for 2 h and reaction products were analysed by TLC. The enzyme reaction(s) were stopped by adding 200 µL of absolute ethanol and the mixture(s) were centrifuged at 13000g for 5 min. The supernatant was transferred to another 1.5 mL micro-centrifuge tube and concentrated by evaporating the absolute ethanol in hot air oven at 70°C for 12 h. Then 1 µL of each reaction mixture for CMC-Na, 0.5 µL of each reaction mixture for β-glucan, 1 µL of each reaction mixture for cellobiose and 1 µL of standard (D-glucose and cellobiose) solutions (1.0 mg/mL) were loaded on the TLC plates. The plates were dried in hot air oven at 70°C for 5 min and kept in the developing chamber saturated with the developing solution consisting of n-butanol-acetic-acid-water (2:1:1, v/v) at 25°C (Jamaldheen *et al.*, 2018). After the run, the released sugars were visualized by staining the TLC plates with a visualizing solution (sulphuric acid/methanol 5:95, v/v; α-naphthol 0.5%, w/v). Further, the TLC plates were dried in hot air oven at 70°C until the migrated sugars were visualized as spots on the TLC plate.

3.2.12 Hydrolysis of *Pennisetum purpureum* by CMCase-WT and CMCase-UV2

The potential of CMCase-WT and CMCase-UV2 enzymes in the hydrolysis of *Pennisetum purpureum* (Elephant grass) pretreated by a physicochemical method (1%, v/v NaOH+20 min autoclaving) as described earlier by Eliana *et al.*, (2014) was evaluated. The reactions with CMCase-WT and CMCase-UV2 enzymes were carried out with 1.0% (w/v) pretreated substrate dissolved in 50 mM sodium acetate buffer, pH 5.0 by incubating at 30°C under shaking condition at 160 rpm for 24 h and 48 h. The enzymatic hydrolysis was carried out at 30°C instead of 65°C, assuming its application in simultaneous saccharification and fermentation (SSF). The 1 mL

TH-2393_146106034

reaction mixtures contained equimolar protein concentration (2.5 mM) of partially purified CMCase-WT or CMCase-UV2. The partially purified CMCase-WT and CMCase-UV2 were used in order to make the enzymatic hydrolysis process more cost-effective. The released reducing sugars were quantified by Nelson (1944) and Somogyi (1945) method as described in chapter 2, section 2.2.5. The qualitative analysis of hydrolyzed products of pretreated Elephant grass by CMCase-WT and CMCase-UV2 was performed by TLC, as describe in section 3.2.11. 1 mL reaction mixtures were stopped by adding 2 mL of absolute ethanol and centrifuged at 13000 g for 5 min. The supernatant was transferred to another micro-centrifuge tube and concentrated by evaporating the absolute ethanol in hot air oven. Then 0.2 µL of each reaction mixture of 24 h and 1 µL of each reaction mixture of 48 h along with 1 µL of standard (D-glucose and cellobiose) solutions (1.0 mg/mL) were loaded on the TLC plates and sugar spots were developed as method described in section 3.2.11.

TH-2393_146106034

3.3 Results and Discussion

3.3.1 Purification of CMCase from wild-type *Bacillus amyloliquefaciens* SS35 and its UV2 mutant

The two-step purification of CMCase-WT and CMCase-UV2 enzymes from the cell-free supernatant included ammonium sulfate precipitation followed by size exclusion chromatography. CMCase-UV2 showed maximum enzyme activity in fractions precipitated by 90% saturation of ammonium sulfate, which was similar to the concentration of ammonium sulfate required for CMCase-WT from wild-type strain as reported earlier by Singh *et al.,* (2014). The partially purified CMCase-UV2 showed 2-fold (15 U/mg) higher specific activity than the wild-type strain (7.5 U/mg) shown in Table 3.3.1. The partially purified enzyme(s) were further purified by size exclusion chromatography and the fraction numbers 71-103 and 75-114 showed CMCase activity for CMCase-WT (Fig. 3.3.1A) and CMCase-UV2 (Fig. 3.3.1B), respectively. The purified CMCase-UV2 enzyme showed 32.5 U/mg specific activity, 2.1-fold higher than the CMCase-WT (15.4 U/mg) against the substrate, CMC-Na (Table 3.3.1). This enzyme activity of CMCase-UV2 was 5.8-fold higher than the mutant generated by site-directed mutagenesis (5.5 U/mg) by Vu *et al.,* (2012). The site-directed mutagenesis is helpful in forming the genetic diversity of a gene, whose sequence is known. However, UV mutagenesis has advantage that, it requires no prior knowledge of the gene sequence to produce genetic diversity with improved characteristics of cellulase. The zymogram analysis for activity staining of CMCase-UV2 showed a single band of CMCase activity of molecular mass approximately, 37 kDa (Fig. 3.3.1D) similar to the CMCase enzyme from wild-type strain (Fig. 3.3.1C)

TH-2393_146106034

which was similar to the earlier reports of Singh *et al.*, (2014) by anion exchange chromatography.

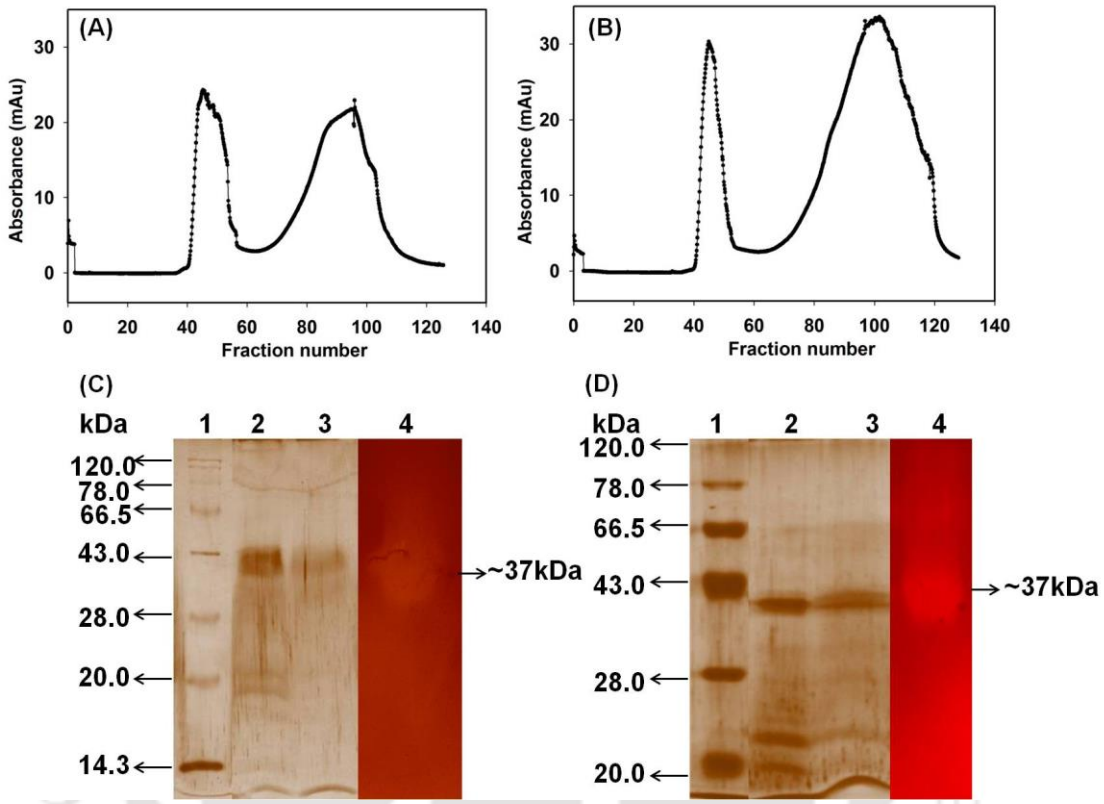


Fig. 3.3.1 Chromatogram from size exclusion chromatography of CMCase purification, (A) CMCase-WT (B) CMCase-UV2. SDS-PAGE analysis using (12%, w/v) gel showing purification and zymogram of CMCase-WT (C) Lanes 1, Protein marker; 2, Partially purified CMCase-WT; 3, Purified CMCase-WT (37 kDa, approx.); 4, Zymogram of purified CMCase-WT and for CMCase-UV2 (D) Lanes 1, Protein marker; 2, Partially purified CMCase-UV2; 3, Purified CMCase-UV2 (37 kDa, approx.); 4, Zymogram of purified CMCase-UV2.

TH-2393_146106034

Table 3.3.1 Comparison of activity of purified CMCase-WT and CMCase-UV2

Purification step	Volume (mL)		Enzyme activity (U/mL)		Total activity (U)		Protein (mg/mL)		Total protein (mg)		Specific activity (U/mg)	
	WT	UV2	WT	UV2	WT	UV2	WT	UV2	WT	UV2	WT	UV2
Culture supernatant	200	200	0.65	0.79	130	158	3.18	3.25	636	650	0.2 ± 0.01	0.24 ± 0.01
60–90% $(NH_4)_2SO_4$	40	69	1.2	1.5	48	104	0.16	0.10	6.4	6.9	7.5 ± 0.1	15.0 ± 0.3
Size exclusion Chromatography	2	2	1.7	3.9	3.4	7.8	0.11	0.12	0.22	0.24	15.4 ± 0.4	32.5 ± 1.2

values are mean SE (n=3).
All the assays were performed at 65°C, 50 mM sodium acetate buffer (pH 5.0), 5 min, 1% (w/v) CMC-Na.

TH-2393_146106034

3.3.2 Effect of pH on CMCase-WT and CMCase-UV2 activity and stability

The optimum pH for CMCase-UV2 using 1% (w/v) CMC-Na was found to be 5.0 (Fig. 3.3.2B), similar to the earlier report by Singh *et al.*, (2014) for CMCase-WT (Fig. 3.3.2A). pH stability showed that CMCase-UV2 retained almost 100% of the relative CMCase activity in the pH range, 4.0-5.5 (Fig. 3.3.2D), whereas, CMCase-WT showed stability only at pH 5.0 and 85% of the relative CMCase activity at pH 4 (Fig. 3.3.2C), after 1h of incubation. Mutation enhanced the stability of CMCase-UV2 in acidic range, which is more suitable for *Saccharomyces cerevisiae* growth (Liu *et al.*, 2014) in SSF process for bioethanol production. Therefore, the use of CMCase-UV2 in SSF may also result in improvement of the ethanol production yield.

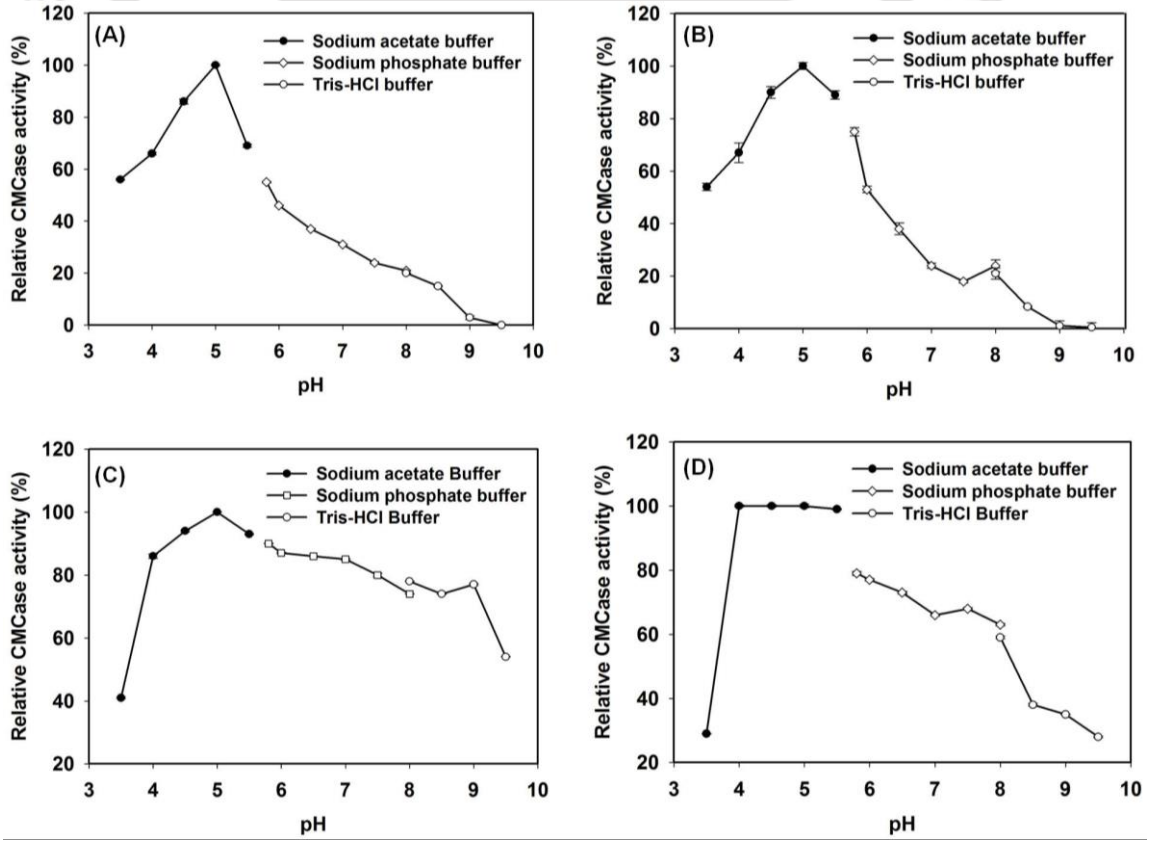


Fig. 3.3.2 Effect of pH on CMCase activity; (A) CMCase-WT and (B) CMCase-UV2; effect of pH on CMCase stability; C) CMCase-WT and D) CMCase-UV2.

TH-2393_146106034

3.3.3 Effect of temperature on CMCase-WT and CMCase-UV2 activity and stability

The optimum temperature of CMCase-UV2 was 65°C, similar to the CMCase-WT (Fig. 3.3.3A) which depicts that mutation happened with a residue that does not have any role related to temperature optima of enzyme. This optimum temperature for CMCase-WT and CMCase UV2 is higher than the CMCase from other reported wild-type strains such as *Bacillus subtilis* subsp. *subtilis* A-53 (50ºC) (Kim *et al.*, 2009), *Bacillus amyloliquefaciens* DL-3 (50ºC) (Lee *et al.*, 2008), *Bacillus* sp. (50ºC) (Sadhu *et al.*, 2013), *Bacillus subtilis* AS3 (45ºC) (Deka *et al.*, 2013) and *Bacillus* sp. M-9 (60ºC) (Bajaj *et al.*, 2009). The thermostability showed that the CMCase-UV2 retained 100% of activity till 45°C in 1h (Fig. 3.3.3B), similar to the CMCase-WT. CMCase-UV2 enzyme retains 100% of the relative CMCase activity at 30°C which makes it a potential enzyme to be used in SSF at 30°C because this temperature (30°C) is suitable for *Saccharomyces cerevisiae* growth in bioethanol production.

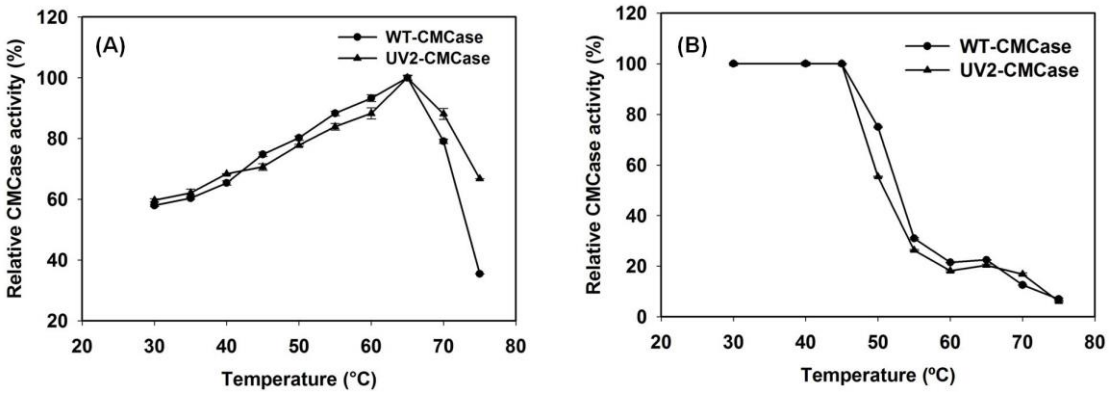


Fig. 3.3.3 (A) Effect of temperature on CMCase-WT and CMCase-UV2 activity and (B) Effect of temperature on CMCase-WT and CMCase-UV2 stability.

TH-2393_146106034

3.3.4 Comparison of activity of purified CMCase-WT and CMCase-UV2

The specific activities of the purified CMCase-WT and CMCase-UV2 enzymes were determined against β-D-glucan, lichenan, CMC-Na, cellulose powder, avicel, birchwood xylan and galactomannan (Table 3.3.2). CMCase-UV2 enzyme showed 1.6-fold to 4-fold higher specific activity against the cellulosic substrates than the wild-type enzyme, displaying enhancement in the catalytic efficiency of the mutant against both soluble and insoluble substrates. The enhancement in the activity of CMCase-UV2 against CMC-Na is 108%, which is significantly higher than that reported for endoglucanase (57%) from *Streptomyces griseoaurantiacus* (Kumar *et al.*, 2015) by UV-irradiation. The CMCase-UV2 mutant gave 1.9-fold (7.1 U/mg) and 1.8-fold (9.8 U/mg) higher enzyme activity against pretreated *Parthenium hysterophorus* and *Pennisetum purpureum*, respectively, (Table 3.3.2) as compared with the wild-type enzyme. These results showed the potential of CMCase-UV2 in saccharification of the complex lignocellulosic biomass. Srikrishnan *et al.*, (2012) reported 1.5 U/mg activity of mutant endoglucanase produced by site-directed mutagenesis of Eg1Y95F against the anhydrous ammonia pretreated corn stover.

TH-2393_146106034

Table 3.3.2. Comparison of enzyme activity of CMCase-WT and CMCase-UV2

Substrate	CMCase-WT (U/mg)	CMCase-UV2 (U/mg)	Fold increase
CMC-Na*	15.4 ± 0.4	32.1 ± 1.2	2.1
β-D-Glucan*	131.7 ± 1.2	210.1 ± 3.1	1.6
Lichenan*	86.5 ± 2.9	161.3 ± 1.1	1.9
Cellulose powder*	5.4 ± 0.2	14.3 ± 0.4	2.6
Avicel*	4.8 ± 0.2	19.5 ± 0.9	4.1
Birchwood xylan*	0.6 ± 0.1	0.78 ± 0.3	1.3
Galactomannan*	ND	ND	-
Pretreated Carrot grass**	3.8 ± 0.9	7.1 ± 0.2	1.9
Pretreated Elephant grass**	5.3 ± 0.5	9.8 ± 0.2	1.8

**All the assays were performed with 1% (w/v) substrate at 65°C, 50 mM sodium acetate buffer (pH 5.0), 5 min, for commercial substrates*
ND = No activity detected.
***All the assays were performed with 2% (w/v) substrate at 65°C, 50 mM sodium acetate buffer (pH 5.0), 15 min, for pretreated substrates.*
± values are mean SE (n=3).

3.3.5 Comparison of kinetic parameters of CMCase-WT and CMCase-UV2

The kinetic parameters (K_m and V_{max}) for CMCase-WT and CMCase-UV2 were determined by Linewaever-Burk plot (Fig. 3.3.4). CMCase-UV2 showed 1.5-fold (238.1 U/mg) and 2.4-fold (41.9 U/mg) higher V_{max} against substrate β-glucan and CMC-Na, respectively, than the CMCase-WT (Table 3.3.3). The V_{max} of CMCase-UV2 against substrate CMC-Na was also higher than other reported V_{max} for CMCase enzymes from wild-type strain *Bacillus* sp. AS3 (3.38 U/mg) (Deka *et al.*, 2013) and *Bacillus* sp. CH43 (0.00093 U/mg) (Mawadza *et al.*, 2000). The mutant enzyme CMCase-UV2 also showed 1.7 and 1.8-fold increase in catalytic efficiency (k_{cat}/K_m) against CMC-Na and β-glucan respectively, as compared with CMCase-WT (Table 3.3.3) showing the potential of CMCase-UV2 in saccharification of the complex lignocellulosic biomass.

TH-2393_146106034

Table 3.3.3 Kinetic parameters of CMCase-WT and CMCase-UV2

Substrate (% w/v)	CMCase-WT				CMCase-UV2				Fold increase (K_{cat}/K_m)*
	K_m ($mgmL^{-1}$)	V_{max} (U/mg)	k_{cat} (min^{-1})	k_{cat}/K_m ($mg^{-1}mL^{1}min^{-1}$)	K_m ($mgmL^{-1}$)	V_{max} (U/mg)	k_{cat} (min^{-1})	K_{cat}/K_m ($mg^{-1}mL^{1}min^{-1}$)	
CMC-Na	0.26 ± 0.01	17.21 ± 0.22	6.37×10^{2}	2.45×10^{3}	0.38 ± 0.04	41.9 ± 1.3	1.55×10^{3}	4.1×10^{3}	1.7
β-Glucan	0.18± 0.00	156.70 ± 2.4	5.79×10^{3}	3.22×10^{4}	0.15 ± 0.01	238.1 ± 4.6	8.80×10^{3}	5.87×10^{4}	1.8

values are mean SE (n=3)
All the assays were performed at 65°C, 50 mM sodium acetate buffer (pH 5.0), 5 min
**catalytic efficiency= K_{cat}/K_m*

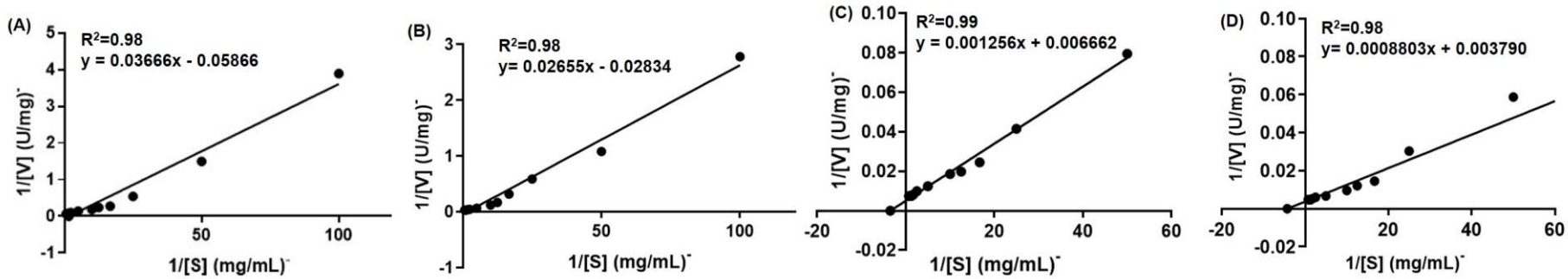


Fig. 3.3.4 Linewaever-Burk plot for determination of kinetics of (A) CMCase-WT against CMC-Na, (B) CMCase-UV2 against CMC-Na, (C) CMCase-WT against β-glucan and (D) CMCase-UV2 against β-glucan.

TH-2393_146106034

3.3.6 Comparative analysis of CMCase-WT and CMCase-UV2 hydrolyzed products of CMC-Na and β-glucan by TLC

The qualitative analysis of CMCase-WT and CMCase-UV2 hydrolysed products of CMC-Na (Fig. 3.3.5A and 3.3.5B) and β-D-glucan (Fig. 3.3.5C and 3.3.5D) was carried out by TLC. The time dependent TLC analysis of CMC-Na and β-D-glucan hydrolysis by CMCase-WT and CMCase-UV2 with equimolar enzyme(s):substrate ratio used, showed the endolytic cleavage pattern, because in the first 30 min, series of cello-oligosaccharides along with cellobiose were formed. However, after 45 min CMCase-UV2 also showed release of glucose (intense spot) along with cellobiose and cello-oligosaccharides (Fig. 3.3.5B, 3.3.5D), whereas, CMCase-WT showed glucose only after 120 min (very faint spot) along with cellobiose and cello-oligosaccharides (Fig. 3.3.5A, 3.3.5C). This displayed enhanced rate and efficiency of the mutant enzyme. This may result into the requirement of less amount of β-glucosidase per gram of biomass when a mixture of CMCase-UV2 with β-glucosidase is employed as against CMCase-WT. Therefore, CMCase-UV2 can help in lignocellulose-based biorefineries for the bioethanol production. The TLC results of CMCase-WT and CMCase-UV2 enzymes with cellobiose showed no release of glucose indicating the absence of β-glucosidase activity (Fig. 3.3.6). These results showed that the glucose appearing on TLC with CMC-Na and β-glucan treatment with CMCase-WT and CMCase-UV2 is released from the higher cello-oligosaccharides and not from cellobiose. The endolytic cleavage of CMC-Na by CMCase from *Bacillus licheniformis* was also reported earlier (Bischoff *et al.*, 2006).

TH-2393_146106034

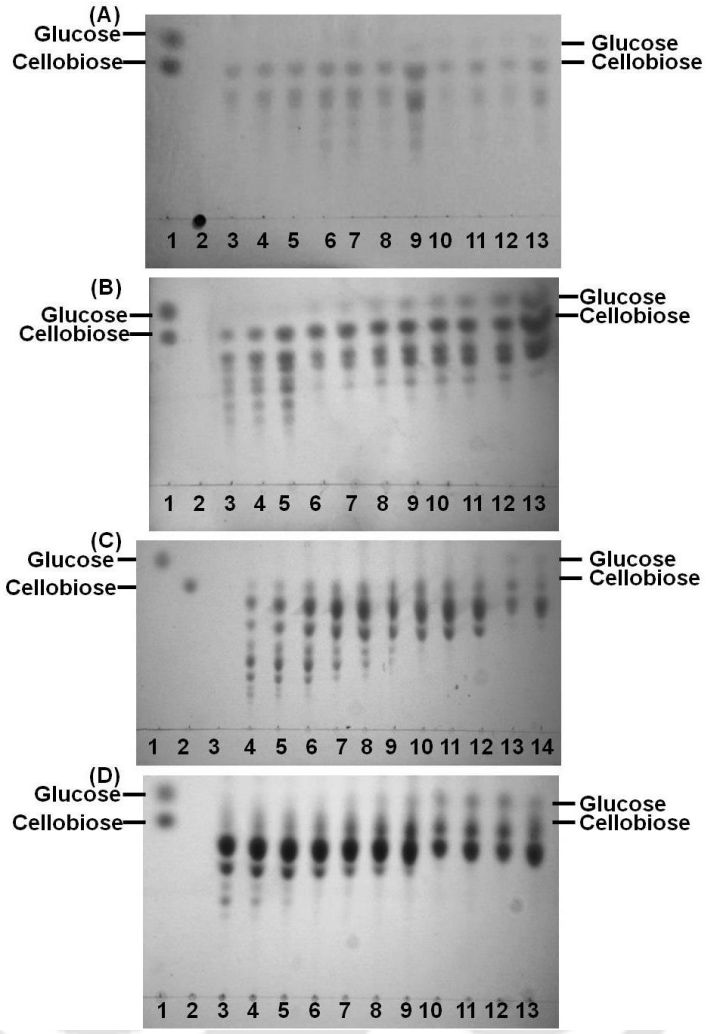


Fig. 3.3.5 (A) TLC analysis of CMCase-WT hydrolysed products of CMC-Na. Lanes; 1, standards (glucose and cellobiose); 2, control (without enzyme); 3, 5 min; 4, 10 min; 5, 15 min; 6, 30 min; 7, 45 min; 8, 60 min; 9, 120 min; 10, 180 min; 11, 360 min; 12, 720 min and 13, 1440 min enzyme reaction products. **(B)** TLC analysis of CMCase-UV2 hydrolysed products of CMC-Na. Lanes; 1, standards (glucose and cellobiose); 2, control (without enzyme); 3, 5 min; 4, 10 min; 5, 15 min; 6, 30 min; 7, 45 min; 8, 60 min; 9, 120 min; 10, 180 min; 11, 360 min; 12, 720 min and 13, 1440 min enzyme reaction products. **(C)** TLC analysis of CMCase-WT hydrolysed products of β-D-glucan. Lanes; 1, standard (glucose); 2, standard (cellobiose); 3, control (without enzyme); 4, 5 min; 5, 10 min; 6, 15 min; 7, 30 min; 8, 45 min; 9, 60 min; 10, 120 min; 11, 180 min; 12, 360 min; 13, 720 min and 14, 1440 min enzyme reaction products reaction products. **(D)** TLC analysis of CMCase-UV2 hydrolysed products of β-D-glucan. Lanes; 1, standards (glucose and cellobiose); 2, control (without enzyme); 3, 5 min; 4, 10 min; 5, 15 min; 6, 30 min; 7, 45 min; 8, 60 min; 9, 120 min; 10, 180 min; 11, 360 min; 12, 720 min and 13, 1440 min enzyme reaction products reaction products.

TH-2393_146106034

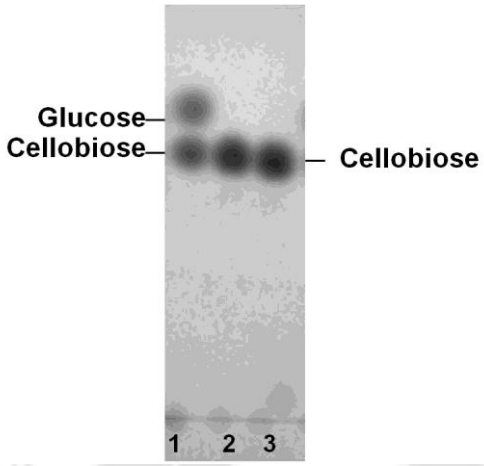


Fig. 3.3.6 TLC analysis of Cellobiose hydrolysed products. Lanes; 1, standards (glucose and cellobiose); 2, CMCase-WT; 3, CMCase-UV2 enzyme reaction products.

3.3.7 CMCase-WT and CMCase-UV2 hydrolysis of *Pennisetum purpureum*

Both the enzymes, CMCase-WT and CMCase-UV2 displayed higher activity against pretreated Elephant grass than pretreated Carrot grass (Table 3.3.2). Therefore, the enzymatic hydrolysis for pretreated Elephant grass was carried out for the release of reducing sugars. The maximum total reducing sugar (TRS) yield of 154.2 mg/g of pretreated Elephant grass was obtained at 48 h hydrolysis by CMCase-UV2. This TRS yield was 1.8-fold higher than that obtained with CMCase-WT (85.6 mg/g of pretreated biomass) under unoptimized conditions of enzyme concentration and hydrolysis time. However, the saccharification by CMCase-UV2 for 24 h gave 108.6 mg/g and CMCase-WT gave 64.7 mg/g pretreated biomass TRS yield. These results showed that CMCase-UV2 enhanced the saccharification efficiency for lignocellulosic biomass. Eliana *et al.,* (2014) reported the maximum TRS yield of 146.9 mg/g of pretreated *Pennisetum purpureum* under optimized conditions of protein concentration and hydrolysis time 45 h. The TLC analysis of CMCase-UV2 hydrolysed mixture of Elephant grass showed the presence of glucose along with the

TH-2393_146106034

cellobiose and cello-oligosaccharides at 24 h, whereas CMCase-WT did not show any glucose at 24 h (Fig. 3.3.7A). However, CMCase-WT produced glucose in addition to cellobiose and cellotriose and cello-oligosaccharides at 48 h of hydrolysis (Fig. 3.3.7B). This further confirmed the increased efficiency of CMCase-UV2 with production of glucose in 24 h as against 48 h for CMCase-WT.

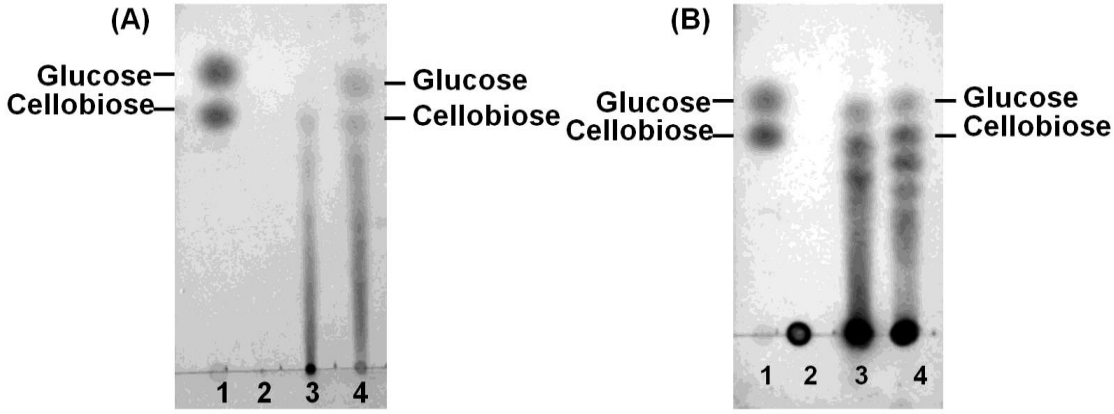


Fig. 3.3.7 TLC analysis of enzyme hydrolysed products from pretreated Elephant grass; (A) After 24 h of hydrolysis. Lanes; 1, standard (glucose and cellobiose); 2, control (without enzyme); 3, CMCase-WT hydrolysed products and 4, CMCase-UV2 hydrolysed products; (B) After 48 h of hydrolysis. Lanes; 1, standard (glucose and cellobiose); 2, control (without enzyme); 3, CMCase-WT hydrolysed products and 4, CMCase-UV2 hydrolysed products.

TH-2393_146106034

3.4 Conclusion

The present study showed the comparison of purification and biochemical properties of endoglucanase, CMCase-UV2 from UV mutant strain of *Bacillus amyloliquefaciens* SS35 with the CMCase-WT from wild-type strain. The purification of CMCase-UV2 and CMCase-WT enzyme was performed by ammonium sulfate precipitation method which is followed by size exclusion chromatography. The molecular mass of CMCase-UV2 was similar to the CMCase-WT approximately, 37 kDa as also confirmed by zymogram analysis. The purified CMCase-UV2 showed 210, 161 and 32 U/mg specific activities against barley β-D-Glucan, lichenan and CMC-Na, respectively which were 1.6, 1.9 and 2.1-fold higher than CMCase-WT. CMCase-UV2 also showed 9.8 U/mg specific activity against pretreated Elephant grass which was 1.8-fold higher than the CMCase-WT. CMCase-UV2 showed pH stability in wider range (4.0-5.5) and retained 100% of the relative CMCase activity whereas CMCase-WT showed stability only at pH 5.0. A UV2-CMCase showed 1.7 and 1.8-fold increase in catalytic efficiency against substrates CMC-Na and β-glucan respectively, as compared with the WT-CMCase. TLC analysis showed the endolytic mode of action of CMCase-WT and CMCase-UV2 enzyme. Unlike CMCase-WT, CMCase-UV2 after 45 min of enzymatic reaction produced also glucose along with cello-oligosaccharides which further confirmed that UV2 mutant strain produced endoglucanase enzyme with increased catalytic efficiency. The efficiency of CMCase-UV2 enzyme was also checked by hydrolysing lignocellulosic biomass, Elephant grass. The maximum TRS yield of 154.2 mg/g of pretreated Elephant grass was obtained by CMCase-UV2, which was 1.8-fold higher than CMCase-WT under unoptimized conditions. These results showed that the

TH-2393_146106034

CMCase-UV2 has high potential for application in enzymatic hydrolysis of lignocellulosic biomass, over CMCase-WT which is an important step in synthesis of bioethanol.

TH-2393_146106034

3.5 References

Aygan, A., Arikan, B., Korkmaz, H., Dinçer, S., and Çolak, O. (2008). Highly thermostable and alkaline α-amylase from a halotolerant-alkaliphilic *Bacillus* sp. AB68. *Brazilian Journal of Microbiology*, 39(3), 547-553.

Bajaj, B.K., Pangotra, H., Wani, M.A., Sharma, P., and Sharma, A. (2009). Partial purification and characterization of a highly thermostable and pH stable endoglucanase from a newly isolated *Bacillus strain* M-9. *Indian Journal of Chemical Technology*, 0975-0991, 382-387

Bayer, E.A., Lamed, R., and Himmel, M.E. (2007). The potential of cellulases and cellulosomes for cellulosic waste management. *Current Opinion in Biotechnology*, 18(3), 237-245.

Bischoff, K.M., Rooney, A.P., Li, X.L., Liu, S., and Hughes, S.R. (2006). Purification and characterization of a family 5 endoglucanase from a moderately thermophilic strain of *Bacillus licheniformis*. *Biotechnology Letters*, 28, 1761-1765.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72(1-2), 248-254.

Chen, Z., Pereira, J.H., Liu, H., Tran, H.M., Hsu, N.S., Dibble, D., and Simmons, B. A. (2013). Improved activity of a thermophilic cellulase, Cel5A, from *Thermotoga maritima* on ionic liquid pretreated switchgrass. *PLOS ONE*, 8(11), e79725.

TH-2393_146106034

Deka, D., Jawed, M., and Goyal, A. (2013). Purification and characterization of an alkaline cellulase produced by *Bacillus subtilis* (AS3), *Preparative Biochemistry and Biotechnology*, 43, 256-270.

Eliana, C., Jorge, R., Juan, P., and Luis, R. (2014). Effects of the pretreatment method on enzymatic hydrolysis and ethanol fermentability of the cellulosic fraction from elephant grass. *Fuel*, 118, 41-47.

Englard, S., and Seifter, S. (1990). [22] Precipitation techniques. In Methods in enzymology, *Academic Press*. 182, 285-300.

Gupta, R., Beg, Q., and Lorenz, P. (2002). Bacterial alkaline proteases: molecular approaches and industrial applications. *Applied Microbiology and Biotechnology*, 59(1), 15-32.

Hakamada, Y., Koike, K., Yoshimatsu, T., Mori, H., Kobayashi, T., and Ito, S. (1997). Thermostable alkaline cellulase from an alkaliphilic isolate, *Bacillus* sp. KSM-S237. *Extremophiles*, 1(3), 151-156.

Hammond, J.B., and Kruger, N.J. (1988). The bradford method for protein quantitation. *New Protein Techniques, Humana Press*, (25-32).

Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W., and Foust, T.D. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science*, 315(5813), 804-807.

Jamaldheen, S.B., Sharma, K., Rani, A., Moholkar, V.S., and Goyal, A. (2018). Comparative analysis of pretreatment methods on sorghum (*Sorghum durra*) stalk agrowaste for holocellulose content. *Preparative Biochemistry and Biotechnology*, 48(6), 457-464.

TH-2393_146106034

Kim, B.K., Lee, B.H., Lee, Y.J., Jin, I.H., Chung, C.H., and Lee, J.W. (2009). Purification and characterization of carboxymethylcellulase isolated from a marine bacterium, *Bacillus subtilis* subsp. subtilis A-53. *Enzyme and Microbial Technology*, 44(6-7), 411-416.

Kim, C.H., and Kim, Y.S. (1995). Substrate specificity and detailed characterization of a bifunctional amylase-pullulanase enzyme from *Bacillus circulans* F-2 having two different active sites on one polypeptide. *European Journal of Biochemistry*, 227(3), 687-693.

Kim, J.M., Kong, I.S., and Yu, J.H. (1987). Molecular cloning of an endoglucanase gene from an alkalophilic *Bacillus* sp. and its expression in *Escherichia* coli. *Applied and Environmental Microbiology*, 53(11), 2656-2659.

Kuhad, R.C., Gupta, R., Khasa, Y.P., and Singh, A. (2010). Bioethanol production from *Lantana camara* (red sage): Pretreatment, saccharification and fermentation. *Bioresource Technology*, 101, 8348-8354.

Kumar, A.K. (2015). UV mutagenesis treatment for improved production of endoglucanase and β-glucosidase from newly isolated thermotolerant actinomycetes, *Streptomyces griseoaurantiacus*. *Bioresources and Bioprocessing*, 2(1), 22.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259), 680.

Lawford, H.G., and Rousseau, J.D. (2003). Cellulosic fuel ethanol. In Biotechnology for Fuels and Chemicals. *Humana Press, Totowa, NJ*, 457-469.

Lee, Y.J., Kim, B.K., Lee, B.H., Jo, K.I., Lee, N.K., Chung, C.H., and Lee, J.W. (2008). Purification and characterization of cellulase produced by *Bacillus*

TH-2393_146106034

amyoliquefaciens DL-3 utilizing rice hull. *Bioresource Technology*, 99(2), 378-386.

Lin, L., Kan, X., Yan, H., and Wang, D. (2012). Characterization of extracellular cellulose-degrading enzymes from *Bacillus thuringiensis* strains. *Electronic Journal of Biotechnology*, 15(3), 2-2.

Lineweaver, H., and Burk, D. (1934). The determination of enzyme dissociation constants. *Journal of the American Chemical Society*, 56(3), 658-666.

Liu, Z.H., Qin, L., Zhu, J.Q., Li, B.Z., and Yuan, Y.J. (2014). Simultaneous saccharification and fermentation of steam-exploded corn stover at high glucan loading and high temperature. *Biotechnology for Biofuels*, 7(1), 167.

Lynd, L.R., Laser, M.S., Bransby, D., Dale, B.E., Davison, B., Hamilton, R., Himmel, M., Keller, M., McMillan, J.D., Sheehan, J., and Wyman, C.E. (2008). How biotech can transform biofuels. *Nature Biotechnology*, 26(2), 169.

Mawadza, C., Hatti-Kaul, R., Zvauya, R., and Mattiason, B. (2000). Purification and characterization of cellulases produced by two *Bacillus* strains, *Journal of Biotechnology*, 83, 177-187.

Michaelis, L., and Menten, M.L. (1913). Die kinetic der invertinwirkung, *Biochemical Journal*, 49, 333-369.

Nelson, N. (1944). A photometric adaptation of the Somogyi method for the determination of glucose. *Journal of Biological Chemistry* 153(2), 375-380.

Rabilloud, T., Brodard, V., Peltre, G., Righetti, P. G., and Ettori, C. (1992). Modified silver staining for immobilized pH gradients. *Electrophoresis*, 13(1), 264-266.

TH-2393_146106034

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Plainview, *Cold Spring Harbor Laboratory Press, Woodbury, New York*. 1.

Schaeffer, H.J., Leykam, J., and Walton, J.D. (1994). Cloning and targeted gene disruption of EXG1, encoding exo-beta 1, 3-glucanase, in the phytopathogenic fungus *Cochliobolus carbonum*. *Applied and Environmental Microbiology*, 60(2), 594-598.

Shobharani, P., Yogesh, D., Halami, P.M., and Sachindra, N.M. (2013). Potential of cellulase from *Bacillus megaterium* for hydrolysis of sargassum, *Journal of Aquatic Food Product Technology*, 22(5), 520-535.

Sims, R.E., Mabee, W., Saddler, J.N., and Taylor, M. (2010). An overview of second generation biofuel technologies. *Bioresource Technology*, 101(6), 1570-1580.

Singh, S., Khanna, S., Moholkar, V.S., and Goyal, A. (2014). Screening and optimization of pretreatments for *Parthenium hysterophorus* as feedstock for alcoholic biofuels. *Applied Energy*, 129, 195-206.

Singh, S., Moholkar, V.S., and Goyal, A. (2013). Isolation, identification and characterization of a cellulolytic *Bacillus amyloliquefaciens* strain SS35 from Rhinoceros Dung. *ISRN Microbiology*, 728134, 7.

Singh, S., Moholkar, V.S., and Goyal, A. (2014). Optimization of carboxymethylcellulase production from *Bacillus amyloliquefaciens* SS35. *3 Biotech*, 4(4), 411-424.

Somogyi, M.A., (1945). New reagent for the determination of sugars. *Journal of Biological Chemistry*, 160, 61-68.

TH-2393_146106034

Srikrishnan, S., Randall, A., Baldi, P., and Da Silva, N.A. (2012). Rationally selected single-site mutants of the *Thermoascus aurantiacus* endoglucanase increase hydrolytic activity on cellulosic substrates. *Biotechnology and Bioengineering*, 109(6), 1595-1599.

van Dyk, J.S., Sakka, M., Sakka, K., and Pletschke, B.I. (2010). Identification of endoglucanases, xylanases, pectinases and mannanases in the multi-enzyme complex of *Bacillus licheniformis* SVD1. *Enzyme and Microbial Technology*, 47(3), 112-118.

Voget, S., Steele, H.L., and Streit, W.R. (2006). Characterization of a metagenome-derived halotolerant cellulase. *Journal of Biotechnology*, 126(1), 26-36.

Vu, V.H., Kim, K. (2012). Improvement of cellulase activity using error-prone rolling circle amplification and site-directed mutagenesis. *Journal of Microbiology Biotechnology*, 22(5), 607-613.

Wood, W.I. (1976). Tables for the preparation of ammonium sulfate solutions. *Analytical Biochemistry*, 73(1), 250-257.

Wray, W., Boulikas, T., Wray, V.P., and Hancock, R. (1981). Silver staining of proteins in polyacrylamide gels. *Analytical Biochemistry*, 118(1), 197-203.

Yin, L.J., Lin, H.H., and Xiao, Z.R. (2010). Purification and characterization of a cellulase from *Bacillus subtilis* YJ1. *Journal of Marine Science and Technology*, 18(3), 466-471.

Zhang, X.Y., Zi, L.H., Ge, X.M., Li, Y.H., Liu, C.G., and Bai, F.W. (2017). Development of *Trichoderma reesei* mutants by combined mutagenesis and induction of cellulase by low-cost corn starch hydrolysate. *Process Biochemistry*, 54, 96-101.

TH-2393_146106034

Chapter 4

Identification of mutation at genetic level and cloning the genes encoding endoglucanase from wild-type *Bacillus amyloliquefaciens* SS35 and its UV2 mutant and their expression and purification

4.1 Introduction

The molecular mechanism for improved efficiency of cellulolytic enzymes from natural strains was understood by amplification of genes encoding cellulase from bacteria by designing degenerate primers from phlylogenetically related species (Schaeffer *et al.*, 1994). Recombinant DNA technology was used to clone various genes in expression vector for getting the enhanced production of natural proteins to several folds (Hodgson, 1994). The target gene from natural microbial strain was amplified by designing the degenerate primers from the most similar phylogenetic sp. on the basis of 16S rRNA/rDNA sequence (Lorenz *et al.*, 2002). The amplified gene than clone and expressed in particular cloning and expression system (Demain *et al.*, 2009). The variety of bacterial, yeasts, plants, insects, mammals and animal cells are used as protein expression systems. Non-glycosylated proteins are expressed in *E. coli* expression system (Terpe, 2006). *E. coli* expression system has several advantages such as easy production of culture, high growth rate, high product yield and *E. coli* genetics are far better understood than those of any other microorganism (Swartz

TH-2393_146106034

1996). This system is used for the massive production of commercialized proteins due to the high growth rate of *E. coli* cells.

In the present study, the family 5 gylcoside hydrolase gene (*Ba*GH5-WT) from wild-type strain of *Bacillus amyloliquefaciens* SS35 for CMCase-WT (Chapter 3, Section 3.2.3) and from its UV2 mutant strain (*Ba*GH5-UV2) for CMCase-UV2 (Chapter 3, Section 3.2.3) was amplified by the degenerate primer based approach, and cloned in pHTP0 cloning vector. The change in the amino acid residue(s) in the mutant strain was identified after sequencing the plasmid DNA(s). The genes encoding *Ba*GH5-WT and *Ba*GH5-UV2 was further cloned in pET-28a(+) expression vector and purified by immobilized metal-ion affinity chromatography (IMAC).

TH-2393_146106034

4.2 Materials and Methods

4.2.1 Chemicals, reagents and kits

The RiPurA Bacterial and Yeast genomic DNA Miniprep purification spin Kit was procured from Himedia, India. NZYEasy cloning kit was procured from NZYTech, Ltd. Genes and Enzymes, Lisbon. Phusion DNA polymerase (ThermoFisher Scientific, USA) for PCR amplification, dNTPs (MP Biomedicals India Pvt. Ltd, India), PCR tubes (0.2 mL) from Axygen, Germany. Restriction enzymes *Nhe*I and *Xho*I were purchased from NEB, UK. The expression vector, pET28a(+) was procured from Novagen, Germany. T_4 DNA ligase and 10x ligase buffer were purchased from Promega, USA. Glacial acetic acid (99.9 % pure), Trizma base (Tris free base), ethidium bromide, Bradford reagent, DNAse-RNAse free water (pH 8.0) and components of polyacrylamide gel electrophoresis were obtained from Sigma-Aldrich Co. LLC, USA. The GenElute miniprep plasmid isolation kit and GenElute gel-extraction kit was obtained Sigma-Aldrich Co. LLC, USA. The DNA was detected using SeaKem® LE agarose (Cambrex Bio Science, USA). The DNA ladder was purchased from NEB, UK. The Protein marker was procured from Biobharti, India. Disodium ethylenediamine tetra acetate salts (EDTA), glucose, sodium hydroxide, sodium dodecyl sulphate (SDS) and LB medium components *viz.* tryptone, yeast extract, sodium chloride and glucose were supplied by HiMedia Laboratories Pvt. Ltd., India. The antibiotics kanamycin and ampicillin were procured from Sigma-Aldrich Co. LLC, USA. The protein staining dye Coomassie Brilliant Blue R250 was procured from Himedia Pvt. Ltd., India and methanol from Merck, India.

TH-2393_146106034

4.2.2 Microorganisms

E. coli (TOP10) competent cells were procured from Invitrogen (USA). *E. coli* BL-21 (DE3)pLysS competent cell for expression of recombinant proteins was obtained from Novagen, EMD4 BioScience, Germany.

4.2.3 Isolation of genomic DNA from wild-type and UV2 mutant strains of *Bacillus amyloliquefaciens* SS35

The genomic DNA from bacterial strains was isolated by using a kit (RiPurA Bacterial and Yeast genomic DNA Miniprep purification spin Kit), following the protocol provided by the manufacturer as described in Section 4.2.3.1. The extracted genomic DNA was eluted in 20 µL DNase free water (Sigma-Aldrich Co. LLC, USA.).

4.2.3.1 Genomic DNA isolation protocol

1. 45 mg/mL of lysozyme solution in micro-centrifuge tubes was freshly prepared.
2. 1.5 mL of 12 h culture (1.7 optical density) was taken in micro-centrifuge tubes and centrifuged at 13,000*g* for 2 min. The supernatant was removed and the above step was repeated thrice with fresh 1.5 mL culture so as to process a total of 4.5 mL for DNA isolation.
3. The cell pellet was resuspended in 200 µL of lysozyme solution and incubated at 37°C for 30 min in waterbath.
4. 20 µL of protinase K solution (20 mg/mL) was added in the reaction mixture and gently mixed by inverting the microcentrifuge tubes.
5. Now, 20 µL of RNAase A solution was added in a reaction and incubated at 25°C for 5 min.

TH-2393_146106034

6. Further, 200 µL of lysis solution was added and the reaction mixture was thoroughly mixed by vortexing for few seconds and incubated at 55°C for 10 min.
7. 200 µL of 95-100% ethanol was added to lysate, the cells and the mixture were thoroughly mixed by vortexing for few seconds.
8. The lysate was transferred into the Hielute Miniprep spin column and centrifuged at ≥6500*g* for 1 min. Flow through liquid was discarded and column was placed in the same 2 mL collection tube.
9. 500 µL of prewash solution was added to the column and centrifuged at 13,000*g* for 1 min. Flow through liquid was discarded and same collection tube was reused with the column.
10. 500 µL of wash solution was added to the column and centrifuged at 13,000*g* for 3 min. The column was transferred into the new collection tube and centrifuged again at 13,000*g* for 1 min to dry the column.
11. The DNA binding column was transferred to a fresh sterile microcentrifuge tube and 20 µL DNase free water was added at the centre of the binding column. The microcentrifuge tube was allowed to stand for 2 min at room temperature and then genomic DNA was eluted by centrifugation at 10000 rpm for 1 min.
12. The genomic DNA then got collected in the sterile microcentrifuge tube and stored at -80°C.

4.2.4 PCR amplification of wild-type endoglucanase gene from wild-type strain and UV mutant endoglucanase gene from UV2 mutant strain of *Bacillus amyloliquefaciens* SS35

The molecular screening of mutation(s) induced in the gene encoding CMCase (endoglucanase, CMCase-UV2) produced from UV2 strains of *Bacillus amyloliquefaciens* SS35, was done by using the degenerate oligonucleotide primers.

TH-2393_146106034

The degenerate oligonucleotide primers were designed against the family 5 glycoside hydrolase gene from the phylogenetic related species. The phylogenetic related species of *Bacillus amyloliquefaciens* SS35 was identified on the basis of 16SrDNA sequence via phylogenetic tree. Polymerase chain reaction (PCR) using degenerate primers was used to identify the family 5 homologues in wild-type *Bacillus amyloliquefaciens* SS35 and mutation in UV2 mutant strain. The genes encoding CMCase-WT (*Ba*GH5-WT) and CMCase-UV2 enzyme (*Ba*GH5-UV2) were amplified from respective genomic DNA (25 ng/µL) using Phusion High-Fidelity DNA polymerase. The PCR amplification was performed using a thermal cycler (Applied Biosystems, GeneAmp® PCR System 9700). The components of 60 µL PCR reaction mixture and the PCR cycles for amplification of *Ba*GH5-WT and *Ba*GH5-UV2 are given in Tables 4.2.1 & 4.2.2. The amplified PCR fragments were run on 0.8 (%, w/v) agarose gel as mentioned in Section 4.2.5.

Table 4.2.1 PCR mixture for amplification of *Ba*GH5-WT and *Ba*GH5-UV2

PCR components	Volume (µL)	Final concentration
5x HF buffer	12.0	1x
25 mM dNTP mix	0.48	1.5 mM
Forward primer (15 mM)	1.8	0.5 µM
Reverse primer (15 mM)	1.8	0.5 µM
Genomic DNA (20.0 ng/µL)	2	
DMSO	1.8	3.0%
Phu DNA Pol (2U/µL)	0.6	0.02 U/µL
Sigma water, pH 8.0	39.52	--
Total	60.0	--

TH-2393_146106034

Table 4.2.2 PCR cycles for amplification of *Ba*GH5WT and *Ba*GH5UV2

Steps		Time
I. Initial denaturation at 98°C		3 min
II. 30 cycles of	i) Denaturation at 98°C	20 s
	ii) Annealing at 62°C	30 s
	iii) Extension at 72°C	1 min
III. Final extension at 72°C		10 min

4.2.5 Agarose gel electrophoresis of PCR amplified DNA of *Ba*GH5-WT and *Ba*GH5-UV2

The PCR amplified products were analysed on 0.8% (w/v) agarose gel prepared in 1x TAE buffer. A stock solution of 10x TAE buffer was prepared using 400 mM Tris-acetate, 10 mM EDTA pH 8.0, according to Sambrook and Russell (2001). 0.8% (w/v) agarose gel was prepared by dissolving 400 mg agarose in 50 mL of 1x TAE buffer by heating in a microwave oven for 5-6 min to get a clear transparent solution. Then 4.0 µL of ethidium bromide (4.0 mg/mL) was added when the solution temperature was around 50°C. The solution was mixed and poured on the casting apparatus, combs were placed and the gel was allowed to set for 30 min. The DNA sample and DNA loading dye (Section 4.2.5.1) were mixed in 4:1 ratio. The PCR amplified DNA was separated under the constant electric field (80 V) for 1h in 1x. The bands were then visualized under UV illumination in a gel documentation system (Kodak, Gel Logic 1500).

4.2.5.1 Preparation of DNA loading dye

The sample or DNA loading dye was prepared by mixing the components as mentioned in Table 4.2.3. A 5x stock solution of DNA loading dye, pH 8.0 was prepared. Stock solution of the DNA loading dye was mixed with 4 volumes of DNA sample to make it final 1x before loading on to the agarose gel.

TH-2393_146106034

Table 4.2.3 Composition of 5x DNA loading dye

Components	Final concentration (5x)
Tris-HCl	50 mM
Glycerol	25% (w/v)
EDTA	5.0 mM
Bromophenol blue	0.2% (w/v)
Xylene cyanol	0.2% (w/v)

4.2.6 Extraction of DNA from agarose gel

The PCR amplified DNA(s) of *Ba*GH5WT and *Ba*GH5UV2 or other plasmid DNA(s) were purified from agarose gel by using a kit (Qiagen, QIAquick® Gel Extraction Kit), following the protocol provided by the manufacturer as described in Section 4.2.6.1. The extracted DNA was eluted in 20 µL DNase free water (Sigma-Aldrich Co. LLC, USA.).

4.2.6.1 DNA gel extraction protocol

1. 1.5 mL sterile, empty microcentrifuge tube was weighed and the weight of the tube was noted.
2. The PCR or plasmid DNA band(s) were excised from gel using sharp sterile scalpel and transferred to an empty microcentrifuge tube. The tube was weighed again and the weight of excised gel was determined by subtracting the empty tube weight (noted above).
3. Three volumes of buffer QG was added to every 1 volume of gel (100 mg ~ 100 µl) in the microcentrifuge tube.
4. The microcentrifuge tube containing excised gel were incubated at 50°C for 10 min (or until the gel slice has completely dissolved). When the gel slice was dissolved completely, the colour of the solution became yellow (similar to Buffer QG without dissolved agarose).

TH-2393_146106034

5. QIAquick spin column (DNA binding column) was placed in 2 mL collection tube provided with the kit. The above solution containing PCR-amplified or plasmid DNA (750 µL) was added to DNA binding columns and centrifuged at 13000***g*** for 1 min at room temperature and the flow through was discarded. If the volume was more than 750 µL, the remaining solution was centrifuged similarly and again the flow through was discarded.
6. 500 µL of buffer QG was added to each QIAquick spin column and the mixtures was centrifuged again at 13000***g*** for 1 min at room temperature, and the flow through was discarded.
7. Now, 750 µL of buffer PE was added to each column containing PCR DNA or recombinant plasmid DNA and the mixture was centrifuged at 13000***g*** for 1 min at room temperature. The flow through was discarded and the column was given an additional spin of 1 min at 13000***g***, to completely remove the residual ethanol.
8. Now the column containing bound DNA was placed in a fresh 1.5 mL sterile microcentrifuge tube. 15 µL of DNAse free water (Sigma-Aldrich Co. LLC, USA) or elution buffer (10 mM Tris-Cl, pH 8.5) was added at the centre of the column. The column was incubated for 2 min at room temperature and centrifuged at 13000***g*** for 1 min.
9. The PCR amplified or plasmid DNA was eluted from QIAquick spin columns was collected in 1.5 mL sterile microcentrifuge tube. The DNA was stored at -80°C for further use.

TH-2393_146106034

4.2.7 Ligase independent cloning of *Ba*GH5-WT and *Ba*GH5-UV2 PCR product by NZYEasy cloning kit in pHTP0 vector

The PCR amplified DNA of *Ba*GH5-WT and *Ba*GH5-UV2 were ligated to pHTP0 cloning vector using NZYEasy cloning kit, following the reaction conditions provided by the manufacturer as described in Table 4.2.4 & 4.2.5. Fig. 4.2.1 showed the vector map of pHTP0 cloning vector. Reaction mixture(s) was prepared and shown in Table 4.2.6 & 4.2.7. The cloning reaction was performed in a thermal cycler programmed with the protocol given in Table 4.2.8.

Table 4.2.4 Optimal amount of the insert DNA, in nanograms, to be used in a cloning reaction

Fragment length (bp)	Optimum DNA quantity for cloning reaction in ng
100	8.3
300	25.0
500	41.5
1000	83.0
2000	166.0
3000	249.0
4000	332.0
Total	10

Table 4.2.5 On ice, in a sterile, nuclease-free microcentrifuge tube, reaction mixture was prepared

Components	Volume (µL)
Purified DNA fragment or positive control	xµL
pHTP0 vector	1
10X reaction buffer	1
NZYEasy enzyme mix	0.5
Nuclease free-water	Upto 10
Total	10

positive control (21 ng/µL) always 2µL

TH-2393_146106034

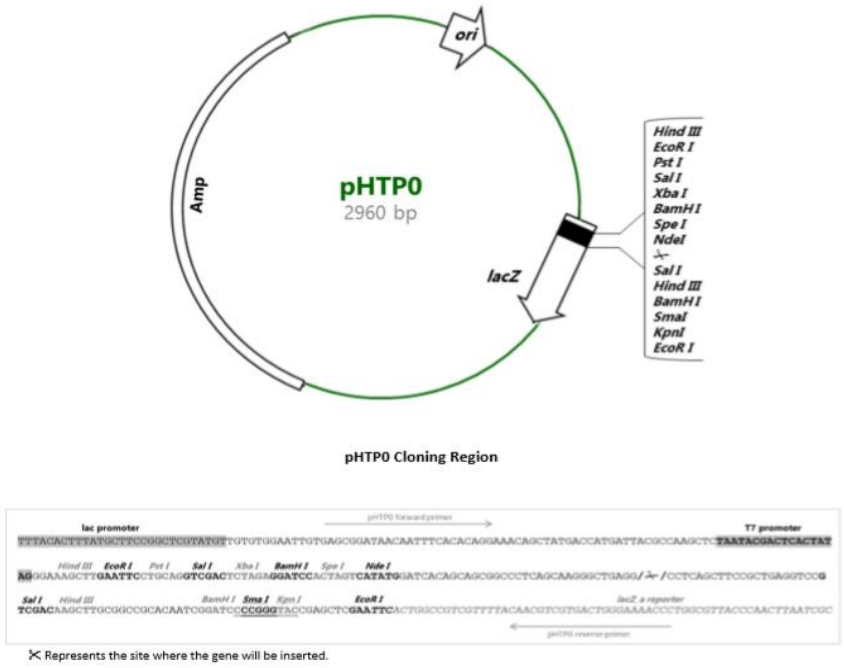


Fig. 4.2.1 Vector map of pHTP0 cloning vector

Table 4.2.6 Preparation of reaction mixture for *Ba*GH5-WT and *Ba*GH5-UV2

Components	Volume (µL)
Purified *Ba*GH5-WT DNA fragment (18 ng/µL)	6.9
pHTP0 vector	1
10X reaction buffer	1
NZYEasy enzyme mix	0.5
Nuclease free-water	0.59
Total	10

Table 4.2.7 Preparation of reaction mixture for positive control

Components	Volume (µL)
Positive control (Kit), ng/µL	2
pHTP0 vector	1
10X reaction buffer	1
NZYEasy enzyme mix	0.5
Nuclease free-water	5.5
Total	10

TH-2393_146106034

Table 4.2.8 Ligation reaction in the thermal cycler

Temperature (°C)	Time (min)
37	60
80	10
30	10
4	∞

4.2.8 Preparation of *E. coli* (TOP10) competent cells by calcium chloride method

Day 1

1. 50 µL culture of *E. coli* (TOP10) competent cells from glycerol stocks were inoculated in 5.0 mL of LB medium (prepared in Section 4.2.9), and grown at 37ºC, 180 rpm for 12 h.
2. 0.1 M $CaCl_2$, 0.1 M $MgCl_2$ and 0.1 M $CaCl_2$ with 15% glycerol solution were prepared and filter-sterilized via 0.22 µm membrane filter (Pall) in laminar air flow and stored at 4°C.

Day 2

3. 1.0 mL of grown culture from above was taken and inoculated into 100 mL LB medium in 250 mL conical flask. The flask was incubated at 37ºC, 180 rpm till cell OD reached 0.3-0.4 at 600 nm.
4. 50 mL micro-centrifuge tubes (round bottom) and micro tips were autoclaved and kept on ice bath placed in a laminar air flow.
5. 50 mL from 100 mL grown culture was transferred aseptically to the two 50 mL round bottom centrifuge tubes (the centrifuge tubes were weighed for balancing before placing on refrigerated centrifuge).
6. The centrifuge tubes were centrifuged at 4ºC with 2810***g*** for 10 min.

TH-2393_146106034

7. The supernatant was discarded and the cell pellet was gently resuspended in 3-4 mL of sterile, ice-chilled 0.1 M $CaCl_2$ solution and the resuspended cells were kept on ice for 10 min.
8. The tubes were centrifuged again at 2810***g*** at 4ºC for 10 min.
9. The supernatant was carefully removed and the pellet was gently resuspended in 3.0 mL of sterile, ice chilled 0.1 M $MgCl_2$ solution and the resuspended cells were kept on ice for 10 min.
10. The tubes were centrifuged again at 2810***g*** at 4ºC for 10 min
11. The supernatant was carefully removed and the pellet was gently resuspended in 1.0 mL of sterile, ice chilled 0.1 M $CaCl_2$ with 15% glycerol solution.
12. 100 µL of competent cells were aliquoted into sterile 1.5 mL microcentrifuge and kept at -80ºC for further use.

4.2.9 Preparation of Luria-Bertani (LB) medium

The composition of LB medium is mentioned in Table 4.2.9. The contents were dissolved in 800 mL of deionized water, the pH was adjusted to 7.2 and the final volume was made upto 1 litre. 100 mL of LB medium was then transferred to each of 250 mL conical flask and the flasks were autoclaved at 121°C at 15 psi for 20 min. It is the most commonly used medium for the growth of recombinant *E. coli* cells. The filter sterilized 100 µg/mL of ampicillin or 50 µg/mL (final concentration) of kanamycin was added to autoclaved and cooled LB medium prior to inoculation.

Table 4.2.9 Composition of Luria-Bertani medium (Sambrook *et al.*, 1989)

Components	**Final concentration (%, w/v)**
Tryptone	1.0
Yeast extract powder	0.5
Sodium chloride	1.0

TH-2393_146106034

4.2.9.1 Preparation of LB-agar medium

LB agar medium was prepared by dissolving the components mentioned in Table 4.2.9 and in addition agar-agar type 1 was also added to a final concentration of 1.8% (w/v). The medium was autoclaved as described in Section 4.2.9. The medium was allowed to cool around 50-55°C and then antibiotics ampicillin 100 µg/mL, final concentration) or kanamycin (50 µg/mL, final concentration) were added and mixed, in respective Erlenmeyer flask. Immediately after mixing 25 mL of medium supplemented with antibiotic was poured in sterile petriplates and the medium was allowed to solidify for 15- 20 min.

4.2.10 Transformation of ligated DNA with gene encoding *Ba*GH5-WT and *Ba*GH5-UV2 in *E. coli* (TOP10) competent cells

The ligated recombinant plasmid DNAs for *Ba*GH5-WT, *Ba*GH5-UV2 and control as described in section 4.2.7 were transformed in *E. coli* (TOP10) competent cells. The steps followed in transformation protocol are as follows:

1. The microcentrifuge tube containing competent cell (100 µL) was taken out from -80°C and kept on ice for 5 min, then 10 µL of ligation mixture was added to it.
2. The microcentrifuge tube was gently tapped 4-5 times and kept on ice for 30 min.
3. The cells were given a heat shock at 42°C for 40s.
4. The cells were immediately transferred to ice for 2-3 min.
5. 1.0 mL of LB medium was added and the cells were incubated at 37°C in a shaking incubator at 200 rpm for 1h.
6. The cells were harvested by centrifugation at 4000***g*** at 25ºC for 3 min.
7. 1.0 mL supernatant was discarded and the cell pellet was then resuspended in remaining 100 µL supernatant.

TH-2393_146106034

8. The cells were spread on LB agar plates supplemented with ampicillin at a final concentration 100 µg/mL.
9. The LB agar plates were incubated at 37°C for overnight.
10. The colonies were picked at random from LB agar plates and inoculated in 5 mL LB medium supplemented with 100 µg/mL ampicillin and incubated at 37ºC, 180 rpm for 12 h for isolation of plasmid DNA to check for positive clones

The colonies were picked at random from LB agar plates and inoculated in 5 mL LB medium supplemented with 100 µg/mL ampicillin and incubated at 37ºC, 180 rpm for 12 h for isolation of plasmid DNA to check for positive clones.

4.2.11 Plasmid isolation protocol (Sigma-Aldrich Co. LLC, USA)

1. 1.5 mL from the each of the grown recombinant culture was taken and was transferred to 1.5 mL microcentrifuge tube aseptically.
2. The cells were then centrifuged at 13000***g*** for 1 min and the process was repeated twice with another 1.5 mL of grown culture.
3. The resulting cell pellets of each recombinant derivative were resuspended in 200 µL resuspension solution by vortex. RNAse at final concentration of 20 µg/mL, was added to the re-suspension solution prior to use.
4. 200 µL of lysis solution was added to each tube and the tubes were gently inverted 5-6 times. The tube was allowed to stand for 2-5 min.
5. 350 µL of neutralization solution was added to the mixture and the tubes were inverted again for 4-6 times to mix properly.
6. The mixture was then centrifuged at 13,000***g*** for 10 min.

TH-2393_146106034

7. The DNA binding columns were prepared or activated by adding 500 µL of column preparation solution to binding column and centrifuging at 13,000***g*** for 1 min. The flow through accumulated in collection tube was discarded.
8. The clear lysate was then transferred to activated DNA binding column and centrifuged at 13,000***g*** for 1 min and the flow through in the collection tube was discarded again.
9. The plasmid DNA bound to column were washed with wash solution and spin at 13,000***g*** for 1 min. The flow through was discarded and the column was given another 1 min spin at 13,000***g*** for removing the wash solution completely.
10. The DNA binding column was transferred to a fresh sterile microcentrifuge tube and 30 µL of DNase free water (Sigma-Aldrich Co. LLC, USA) was added at the centre of binding column. The microcentrifuge tube was allowed to stand for 2 min at room temperature and then plasmid DNA was eluted by centrifugation at 13,000***g*** for 1 min. The plasmid DNA then got collected in the sterile microcentrifuge tube.
11. The eluted plasmid DNA in sterile microcentrifuge tube was stored at -80°C.
12. Concentration of plasmid DNA(s) were measured in nanodrop.

4.2.12 Clone confirmation by PCR using plasmid DNA as template for *Ba*GH5-WT and *Ba*GH5-UV2

Polymerase chain reaction (PCR) was used to identify the positive clones for gene encoding *Ba*GH5-WT and *Ba*GH5-UV2. The gene(s) were amplified from respective cloned plasmid DNA using Phusion High-Fidelity DNA polymerase. The PCR amplification was performed using a thermal cycler (Applied Biosystems, GeneAmp® PCR System 9700). The components of 30 µL PCR reaction mixture and

TH-2393_146106034

the PCR cycles for amplification for *Ba*GH5-WT and*Ba*GH5-UV2 were given in Tables 4.2.10 & 4.2.11. The amplified PCR fragments were run on 0.8 (%, w/v) agarose gel as mentioned in Section 4.2.5.

Table 4.2.10 Clone confirmation by PCR using plasmid DNA as template for *Ba*GH5-WT and *Ba*GH-UV2

PCR components	**Volume (µL)**	**Final concentration**
5x HF buffer	6	1x
25 mM dNTP mix	0.24	1.5 mM
Forward primer (15 mM)	0.9	0.5 µM
Reverse primer (15 mM)	0.9	0.5 µM
Plasmid DNA (109.0 ng/µL)	1.19	-
DMSO	1.8	3.0%
Phu DNA Pol (2U/ µL)	0.6	0.02 U/µL
Sigma water, pH 8.0	18.34	--
Total	30.0	--

Table 4.2.11 PCR cycles for amplification of *Ba*GH5WT and *Ba*GH5UV2

Steps		**Time**
I. Initial denaturation at 98°C		1 min
II. 30 cycles of	i) Denaturation at 98°C	10 s
	ii) Annealing at 62°C	30 s
	iii) Extension at 72°C	1 min
III. Final extension at 72°C		10 min

The plasmids DNA of both strains were sequenced for gene sequence, mutation site and restriction sites identification. For this 20 µL of recombinant plasmid(s) of concentration 109 ng/µL were sent to Eurofins, Bangalore.

TH-2393_146106034

4.2.13 Cloning of gene encoding *Ba*GH5-WT and *Ba*GH5-UV2 in expression vector pET-28a(+)

The oligonucleotide primers were designed with restriction sites *Nhe*I in forward primer and *Xho*I in reverse primer against the family 5 glycoside hydrolase gene sequence for cloning *Ba*GH5-WT and *Ba*GH5-UV2 into the expression vector pET-28a(+). Polymerase chain reaction (PCR) were used to amplify genes encoding *Ba*GH5-WT and *Ba*GH5-UV2 were from respective plasmid DNA using Phusion High-Fidelity DNA polymerase. The PCR amplification was performed using a thermal cycler (Applied Biosystems, GeneAmp® PCR System 9700). The components of 100 µL PCR reaction mixture and the PCR cycles for amplification of *Ba*GH5-WT and *Ba*GH5-UV2 are given in Tables 4.2.12 & 4.2.13, respectively. The amplified PCR fragments were run on 0.8 (%, w/v) agarose gel as mentioned in Section 4.2.5.

Table 4.2.12 PCR mixture for amplification of *Ba*GH5-WT and *Ba*GH5-UV2

PCR components	Volume (µL)	Final concentration
5x HF buffer	20.0	1x
25 mM dNTP mix	0.8	1.5 mM
Forward primer (15 mM)	3.0	0.5 µM
Reverse primer (15 mM)	3.0	0.5 µM
Plasmid DNA (*Ba*GH5-WT or *Ba*GH5-UV2, 109 ng/µL)	3.98	
DMSO	3.0	3.0%
Phu DNA Pol (2U/ µL)	1	0.02 U/µL
Sigma water, pH 8.0	65.22	--
Total	100.0	--

TH-2393_146106034

Table 4.2.13 PCR cycles for amplification of *Ba*GH5-WT and *Ba*GH5-UV2

Steps		Time
I. Initial denaturation at 98°C		1 min
II. 30 cycles of	i) Denaturation at 98°C	10 s
	ii) Annealing at 58°C	30 s
	iii) Extension at 72°C	1.5 min
III. Final extension at 72°C		10 min

4.2.14 Restriction enzyme digestion of the PCR amplified DNA of *Ba*GH5-WT and *Ba*GH5-UV2

The PCR DNA purified after gel extraction was digested with restriction enzymes (RE) *Nhe*I and *Xho*I as per the reaction described in Table 4.2.14. The reaction mixtures were incubated in a water bath (Julabo GmbH, Germany) at 37°C for 90 min. The digested PCR fragments were run on 0.8% (w/v) agarose gel and the desired fragments were extracted using gel extraction kit as mentioned in Section 4.2.6 and eluted in 20 µL of sterile DNase free water pH 8.0 (Sigma-Aldrich Co. LLC, USA).

Table 4.2.14 Restriction enzyme digestion set up of PCR amplified DNAs of *Ba*GH5-WT and *Ba*GH5-UV2

RE digestion set up	1x (µL)
10x reaction buffer	1.8
PCR DNAs (61 ng/µL)	15
DNAse free water	1.82
*Nhe*I (10 U/µL)	0.92
*Xho*I (20 U/µL)	0.46
Total	20

TH-2393_146106034

4.2.15 Restriction digestion of pET-28a(+) expression vector for cloning of *Ba*GH5-WT and *Ba*GH5-UV2 amplified PCR fragments

The pET-28a(+) vector of 5.36 kb was subjected to restriction digestion with *Nhe*I and *Xho*I to prepare it for ligation of PCR amplified and restriction enzymes digested fragments of *Ba*GH5-WT and *Ba*GH5-UV2. The stock solution (100 ng/µL) of pET-28a(+) was prepared from using DNase free water (pH 8.0) and restriction digestion was carried out as described in Table 4.2.15. The digestion mixture was incubated in a water bath at 37°C for 90 min. The *Nhe*I-*Xho*I digested pET-28a(+) vector was run on 0.8 (w/v)% agarose gel and purified as described in Section 4.2.5. The concentration of *Nhe*I-*Xho*I digested pET-28a(+) vector, was 12.95 ng/µL.

Table 4.2.15 Restriction enzyme digestion set up of pET-28a(+)

RE digestion set up	**1x (µL)**
10x reaction buffer	0.5
vector pET-28a(+) (100 ng/µL)	2.5
DNAse free water	1.62
*Nhe*I (10 U/µL)	0.25
*Xho*I (20 U/µL)	0.125
Total	5.0

TH-2393_146106034

4.2.16 Ligation of *Nhe*I-*Xho*I digested PCR fragments into pET-28a(+) vector

The *Nhe*I-*Xho*I digested PCR fragments mentioned in Section 4.2.14 were cloned into pET-28a(+) vector, which was also digested with same restriction enzymes as described in Section 4.2.15. The insert: vector molar ratio was kept at 6:1 for all the fragments. The amounts of *Nhe*I-*Xho*I digested insert DNA of *Ba*GH5-WT and *Ba*GH5-UV2 required for cloning were calculated using the following formula (Engler and Richardson, 1982).

$$\frac{\text{Amount of vector (ng)} \times \text{Size of insert (kb)}}{\text{Size of vector (kb)}} \times \text{insert :vector molar ratio} = \text{amount of insert (ng)}$$

The ligation set up used for digested PCR fragments and pET-28a(+) vector is shown in Table 4.2.16. Each set of ligation reaction was incubated at 24°C, overnight.

Table 4.2.16 Ligation set up of PCR inserts (*Ba*GH5-WT and *Ba*GH5-UV2) and pET-28a(+)

Ligation set up	1x (µL)
10x ligation buffer	0.7
pET-28a(+) vector (12.95 ng/µL)	1.5
RE digested PCR DNA insert (29.4 ng/µL)	3.8
T4 DNA ligase (3 U/µL)	1.0
Total	7

4.2.17 Transformation of ligated pET-28a(+) with gene encoding *Ba*GH5-WT and *Ba*GH5-UV2 in *E. coli* (TOP10) competent cells

The ligated recombinant plasmid DNAs for *Ba*GH5-WT and *Ba*GH5-UV2 and pET-28a(+) as described in section 4.2.16 were transformed in *E. coli* (TOP10) competent cells (preparation described in the section 4.2.8). The protocol for transformation followed was same as described in Section 4.2.10. The cells harboring recombinant plasmids were spread on LB agar plates with 50 µg/mL kanamycin and grown at 37ºC for 12 h. The colonies were randomly picked from LB agar plates

TH-2393_146106034

recombinant cells after 12 h of growth and inoculated into 5.0 mL LB medium supplemented with 50 µg/mL kanamycin.

4.2.18 Isolation of recombinant plasmid DNA by NID miniprep method

The recombinant plasmid DNAs of *Ba*GH5-WT and *Ba*GH5-UV2 were isolated from *E. coli* (TOP10) cells using NID miniprep method (Lezin *et al.*, 2011). The composition of extraction buffer used in NID method was: 20-50 mM EDTA, 5% sucrose, 50 mM Tris pH 8, 20-50 mM $CaCl_2$, 0.75 M NH_4Cl, 0.5% Triton X-100, lysozyme 100 µg/mL, and RNase A 25 µg/mL. A 100x enzyme stock containing 10 mg/mL lysozyme and 2.5 mg/mL of RNase A was prepared in 50% glycerol and 50 mM Tris pH 8 and stored at -20˚C.

4.2.18.1 NID miniprep plamid isolation protocol

1. 2 mL of bacterial cultures were pelleted at 6000-7000 rpm for 1 min.
2. After drawing 150 mL extraction buffer into a pipette tip, the pellet was loosened off the tube wall with the tip without releasing the buffer. Then the extraction buffer was added and the pellet resuspended.
3. The bacterial suspension was incubated at 65˚C for 5 min.
4. Suspensions were centrifuged at maximum rpm for 10 min or until a tight bacterial pellet was formed. The pellet was removed with a toothpick.
5. 100-120 mL isopropanol was added, followed by mixing and centrifugation of the solution at 7000 rpm for 10 min at room temperature.
6. DNA usually forms film-like precipitates that adhere well to tube walls and are invisible in isopropanol solutions. After discarding the supernatant, the DNA was centrifuged after adding 70% ethanol.

TH-2393_146106034

7. Ethanol was removed, and the DNA pellet was dissolved in 20-50 mL TE buffer.
8. The isolated plasmid DNA(s) was run on 0.8% (w/v) agarose gel as described in Section 4.2.5. Plasmid DNA was quantified using nanodrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and further stored at -20˚C.

4.2.19 Screening of recombinant plasmid DNA for identification of positive clones for *Ba*GH5-WT and *Ba*GH5-UV2

10 µL of each recombinant plasmid DNA of *Ba*GH5-WT and *Ba*GH5-UV2, isolated by NID miniprep method as described in Section 4.2.18.1, were taken separately, in a fresh sterile microcentrifuge tube for restriction enzyme digestion analysis. The recombinant DNA of each of the above mentioned derivatives was digested with restriction enzymes, *Nhe*I and *Xho*I, to check for positive clones following the reaction set up as given in the Table 4.2.17. The reaction mixtures for each recombinant derivative were incubated at 37ºC in a water bath (Julabo, GmbH) for 90 min. The digested products were run on 0.8% (w/v) agarose gel as described in Section 4.2.5. The digested fragments *viz.* pET-28a(+) vector and the insert DNA of above mentioned recombinant derivatives were visualized under UV transilluminator. The digested fragments (insert and vector) were analyzed to check whether the size of insert DNA and vector pET-28a(+), were same as determined earlier in Sections 4.2.13. Based on this observation, the positive clones for the respective recombinant derivatives were identified. Glycerol stocks of *E. coli* (TOP10) cells harbouring the recombinant plasmids of each of the above mentioned derivatives were prepared by keeping final concentration of glycerol to 20-25% (v/v). The glycerol stocks in cryo

TH-2393_146106034

vials of 2 mL were stored at -80°C. The positive clones for each of the recombinant derivatives *viz. Ba*GH5-WT and *Ba*GH5-UV2 were identified and were preserved.

Table 4.2.17 Restriction enzyme digestion reaction mixture for clone confirmation of *Ba*GH5-WT and *Ba*GH5-UV2

Digestion set up	**1x (µL)**
10x buffer	2.0
DNAse free water	6.0
Recombinant plasmid DNA (approx. 80 ng)	10.0
*Nhe*I (10 U/µL)	1.0
*Xho*I (20 U/µL)	0.5
Total	20.0

Plasmid DNA showing positive insert for gene encoding *Ba*GH5-WT and *Ba*GH5-UV2, respectively was further isolated by Sigma kit (Sigma-Aldrich Co. LLC, USA) from the respective colonies using method as described in section 4.2.11 and stored in -80°C.

4.2.20 Transformation of recombinant plasmids in *E. coli* BL-21 (DE3)pLysS competent cells for protein expression

2 µL of each of the recombinant plasmid DNA isolated by Sigma miniprep method, was transformed in *E. coli* BL-21 (DE3)pLysS for expression of *Ba*GH5-WT and *Ba*GH5-UV2, respectively. 100 µL of *E. coli* BL-21 (DE3)pLysS competent cells (preparation described in the section 4.2.8). The protocol for transformation followed was same as described in Section 4.2.10. The cells harboring recombinant plasmids were spread on LB agar plates with 50 µg/mL kanamycin and 25 µg/mL chloramphenicol grown at 37ºC for 12h. The colonies were randomly picked from LB agar plates recombinant cells after 12h of and inoculated into 5.0 mL LB medium supplemented with 50 µg/mL kanamycin and 25 µg/mL chloramphenicol as mentioned. The cells containing recombinant plasmids were grown at 37ºC in shaking incubator at 180 rpm for expression analysis of protein for 12 h.

TH-2393_146106034

4.2.21 Hyper-expression of recombinant proteins *Ba*GH5-WT and *Ba*GH5-UV2

The *E. coli* BL-21 (DE3)pLysS cells harbouring recombinant plasmids of *Ba*GH5-WT and *Ba*GH5-UV2, as described in Section 4.2.20 were analysed for protein expression. The protocol for protein expression was followed as described by Taylor *et al.,* (2005). The colonies grown on LB agar plates supplemented with kanamycin (50 µg/mL) and chloramphenicol (25 µg/mL) were picked and inoculated in 5 mL LB medium as stated in Section 4.2.9 containing kanamycin (50 µg/mL) and chloramphenicol (25 µg/mL) and incubated in a shaking incubator at 37ºC and 180 rpm. The cells were allowed to grow up to mid-exponential phase till cell absorbance at 600 nm (A_{600}) reached ~0.6. 1.0 mL of this grown culture containing uninduced cells were taken out for making glycerol stocks (after expression analysis) and for the sample preparation for analysis by SDS-PAGE (12%, w/v). The remaining 4.0 mL culture was then induced with isopropyl-1-thio-β-D-galactopyranoside (IPTG) at 1.0 mM final concentration for hyper-expression of recombinant proteins and further incubated at 24ºC with shaking at 180 rpm for 18 h. The expression of recombinant protein was checked by loading respective uninduced and induced cell samples after sonication and cell free extract after sonication on polyacrylamide gel (SDS-PAGE). SDS-PAGE (12%, w/v) gel and its components, SDS-PAGE running buffer, sample loading was prepared as described in the Chapter 3, section 3.2.5. After running the SDS-PAGE gel it was stained and destained with solutions as described in section 4.2.22 These *E. coli* BL-21 (DE3)pLysS containing recombinant proteins *viz. Ba*GH5-WT and *Ba*GH5-UV2 were preserved at -80°C as glycerol stocks by keeping the final glycerol concentration 20-25% (v/v).

TH-2393_146106034

4.2.22 Preparation of staining and destaining solutions

The staining solution (100 mL) was prepared by dissolving 250 mg or 0.25% (w/v), of Coomassie Brilliant Blue (CBB R-250) dye in 50 mL of deionized water in an amber colour bottle by keeping on a magnetic stirrer for overnight. The solution was filtered (Whatman, Filter No. 1), then 40 mL of methanol and 10 mL of glacial acetic acid were added to finally make the ratio 5:4:1 (deionized water:methanol:glacial acetic acid). The destaining solution was prepared by dissolving deionized water: methanol: glacial acetic in 5:4:1 ratio. The gels were destained by immersing it in destaining solution under gentle shaking condition with change of destaining solution every 30 min, until the protein bands were clear.

4.2.23 Purification of recombinant proteins

The *E. coli* BL-21 (DE3)pLysS cells harbouring recombinant plasmids were grown in 100 mL LB medium in 250 mL flask with kanamycin and chloramphenicol as mentioned. The recombinant proteins contained a His_6 tag at N-terminal in pET-28a(+) vector. A single step purification method based on immobilized metal-ion affinity chromatography (IMAC) for His_6-tag containing proteins was employed using 1.0 mL sepharose columns (GE Healthcare, HiTrap chelating HP). The compositions of binding as well as elution buffers used for affinity column purification of recombinant proteins are mentioned Table 4.2.18. All the buffers (binding, elution and cleaning buffer), 0.1M $NiSO_4$ solution and water was filtered through 0.45 µm filter and degassed to avoid back pressure due to clogging of column.

TH-2393_146106034

Table 4.2.18 Composition of buffers required for purification recombinant proteins by immobilized metal ion affinity chromatography (IMAC).

Buffers	**Composition**
Equilibration buffer	20 mM sodium phosphate, pH 7.0 300 mM NaCl, 50 mM Imidazole
Elution buffer	20 mM sodium phosphate, pH 7.0 300 mM NaCl, 300 mM Imidazole
Cleaning buffer	20 mM sodium phosphate, pH 7.0 300 mM NaCl, 50 mM EDTA

The bacterial cells (100 mL culture) were harvested by centrifugation at 8,000***g*** using a centrifuge (Sigma, 4K15) and the resulting cell pellets were resuspended in 3 mL of 20 mM sodium phosphate buffer pH 7.0. The cells were sonicated (Sonics, Vibra cell) on ice for 16 min (5s on/off pulse; 33% amplitude) and again centrifuged at 13,000***g*** at 4ºC for 50 min to get the crude cell free extract. The cell free extract was passed through a 0.45 µm filter membrane before loading on HiTrap chelating HP column (GE Healthcare, USA). Initially the column was washed with 5 volumes of filtered and degassed water to remove the alcohol. The column was charged with 2-3 mL of 0.1 M $NiSO_4$ solution and the unbound nickel was washed away with 5 volumes of water. Then the column was equilibrated with 5 volumes of equilibration buffer (Table 4.2.18). The filtered cell free extract (4-4.5) mL of recombinant proteins was loaded on to the column using a peristaltic pump (P1, GE Healthcare) at a flow rate of 1 mL/min. The column was then washed with 50-60 column volumes of equilibration buffer to remove the unbound proteins. The retained proteins were then eluted with elution buffer and 1 mL fractions were collected (Carvalho *et al.*, 2004; Taylor *et al.*, 2005). The column was cleaned using cleaning buffer as mentioned in

TH-2393_146106034

Table 4.2.18, washed with 2-5 volumes of water and incubated in 1N NaOH at 4°C for 2h. The column was then washed with 20 volumes of water to remove NaOH, and finally stored in 20% (v/v) ethanol at 4°C. The collected fractions were analyzed for protein content by taking 20 µL of purified protein and 80 µL of Bradford's reagent in a microcentrifuge tube. Development of blue colour complex indicated the presence of protein and usually the purified protein was eluted in third fraction as 1 mL HiTrap column was used. The purified recombinant proteins *Ba*GH5-WT and *Ba*GH5-UV2 were dialyzed (10 kDa Mol wt cut off membrane) against 20 mM sodium phosphate buffer pH 7.0. The purity and molecular mass of recombinant proteins were verified by SDS-PAGE (12%, w/v) as described in chapter 3, section 3.2.5. The purified *Bacillus* recombinant proteins were quantified using Bradford reagent (Bradford, 1976) as described in Chapter 3, Section 3.2.4.

TH-2393_146106034

4.3 Results and Discussion

4.3.1 Genomic DNA isolation from wild-type and UV2 mutant strains of *Bacillus amyloliquefaciens* SS35

The genomic DNA from wild-type and UV2 mutant strain of *Bacillus amyloliquefaciens* SS35 was isolated by using the conditions as mentioned in Section 4.2.3.1. The genomic DNA(s) was detected on 0.8% (w/v) agarose gel (Fig. 4.3.1), genomic DNA from wild-type strain (Fig. 4.3.1, lane 2) and UV2 mutant strain (Fig. 4.3.1, lane 3). The genomic DNA from both the strains were stored at -80°C for subsequent cloning experiments.

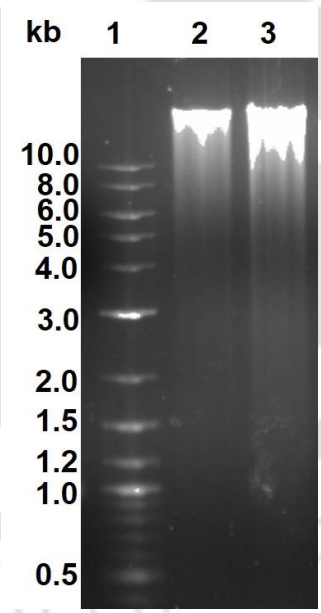


Fig. 4.3.1 Agarose (0.8%, w/v) gel showing genomic DNA of *Bacillus amyloliquefaciens* SS35. Lanes 1, DNA ladder (NEB); 2, genomic DNA from wild-type strain; 3, genomic DNA from UV2 mutant strain.

TH-2393_146106034

4.3.2 Identification of induced mutations in UV2 strain at genetic level with degenerate primer approach

Numerous factors may be associated with the alteration in the enzyme activity of the mutant enzyme produced from the UV2 mutant strain of *Bacillus amyloliquefaciens* SS35. There might be some genetic changes in the UV2 strain. Therefore, the genes encoding endoglucanases i.e. *Ba*GH5-WT and *Ba*GH5-UV2 were amplified from the wild-type strain and UV2 strains, respectively, using degenerate oligonucleotide primers (Table 4.3.1) obtained from phylogenetically related spp. *Bacillus amyloliquefaciens* KHG19 (Fig. 4.3.2) for family 5 glycoside hydrolase (GH5) with GenBank accession number: AJK65578.1 from the CAZy database (http://www.cazy.org) to identify the genetic changes.

Table 4.3.1 Degenerate oligonucleotide primer sequences used for cloning *Ba*GH5-WT from wild-type strain and *Ba*GH5-UV2 from UV2 mutant strain of *Bacillus amyloliquefaciens* SS35

Construct	Primer sequence
***Ba*GH5-WT/ *Ba*GH5-UV2**	Forward: 5'-TCAGCAAGGGCTGAGGATGAAACGGGCAATTTCTATTTT -3' Reverse: 5'-TCAGCGGAAGCTGAGGTAACTAATTGGGTTCTGTTCCC -3'

TH-2393_146106034

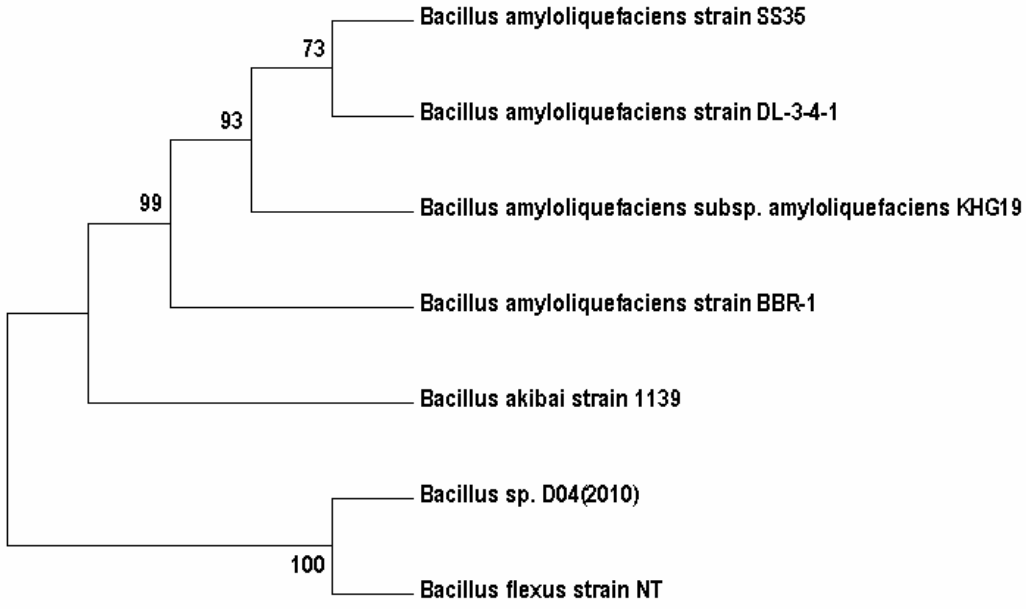


Fig. 4.3.2 Phylogenetic tree analysis of 16S rDNA sequences of *Bacillus amyloliquefaciens* SS35.

4.3.3 PCR amplification of gene encoding family 5 glycoside hydrolase *Ba*GH5-WT and *Ba*GH5-UV2 from wild-type and UV2 mutant strain of *Bacillus amyloliquefaciens* SS35

The gene encoding *Ba*GH5-WT and *Ba*GH5-UV2 were PCR amplified from the genomic DNA of wild-type and UV2 mutant strain of *Bacillus amyloliquefaciens* SS35. The PCR amplified fragments were detected on 0.8% (w/v) agarose gel where *Ba*GH5-WT (Fig. 4.3.3, lane 2) and *Ba*GH5-UV2 (Fig. 4.3.4, lane 2). Both the PCR amplified DNA sequence showed gene size of 1.5 kb. The PCR products were purified from gel using gel extraction kit (Qiagen) as mentioned in Section 4.2.6.1 and were stored at -20°C for subsequent cloning experiments.

TH-2393_146106034

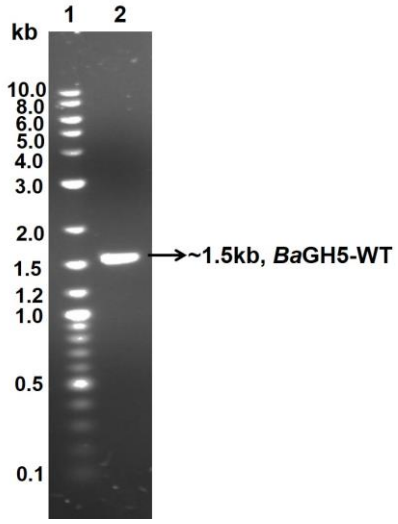


Fig. 4.3.3 Agarose gel electrophoresis (0.8%, w/v) showing amplified PCR DNA of gene encoding *Ba*GH5-WT from genomic DNA. Lanes 1, DNA ladder (NEB); 2, amplified fragment of *Ba*GH5-WT (1.5 kb).

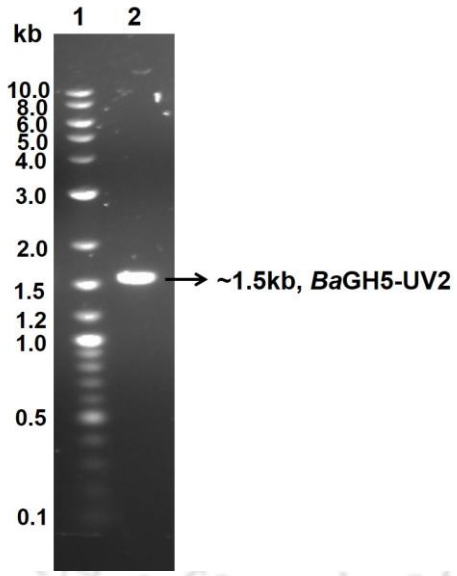


Fig. 4.3.4 Agarose gel electrophoresis (0.8%, w/v) showing amplified PCR DNA of gene encoding *Ba*GH5-UV2 from genomic DNA. Lanes 1, DNA ladder (NEB); 2, amplified fragment of *Ba*GH5-UV2 (1.5 kb).

TH-2393_146106034

4.3.4 Cloning of PCR amplified *Ba*GH5-WT and *Ba*GH5-UV2 product by using NZYEasy cloning kit

4.3.4.1 Ligase independent cloning of BaGH5-WT and BaGH5-UV2 PCR products by NZYEasy cloning kit into the pHTP0 vector

Ligation of gene encoding *Ba*GH5-WT, *Ba*GH5-UV2 and control (kit) into the pHTP0 cloning vector using NZYEasy cloning kit as described in Section 4.2.7 resulted recombinant plasmids and were confirmed after transformation in *E. coli* Top 10 competent cells.

4.3.4.2 Transformation of E. coli (TOP10) competent cells with recombinant plasmid for insert BaGH5-WT and BaGH5-UV2

E. coli (TOP 10) competent cells were transformed with control and recombinant plasmid DNA of each derivative *viz*. *Ba*GH5-WT and *Ba*GH5-UV2 as described in Section 4.2.10. The transformed cells were grown on LB agar plates incubated at 37°C for 12-16 h. Total two colonies was observed on LB agar plate (100 µg/mL ampicillin) for *Ba*GH5-WT, one colony for *Ba*GH5-UV2 and 21 colonies for control.

4.3.4.3 Isolation of plasmid DNA

The plasmid DNA for *Ba*GH5-WT and *Ba*GH5-UV2 recombinant derivative and control was isolated from 5.0 mL LB medium supplemented with ampicillin, by using GeneElite plamid Miniprep kit, method as described in Section 4.2.11. The isolated plasmid DNA for each of the recombinant derivatives *viz.* *Ba*GH5-WT (Fig. 4.3.5, Lane 3) and *Ba*GH5-UV2 (Fig. 4.3.5, Lane 4) along with control (Fig. 4.3.5, Lane 2) were detected on 0.8 % (w/v) agarose gel. Upon comparison with the control it was observed that there was plasmid shift for recombinant plasmids *Ba*GH5-WT

TH-2393_146106034

and *Ba*GH5-UV2. Thus, it was concluded that recombinant plasmid DNA might have the desired gene of insert.

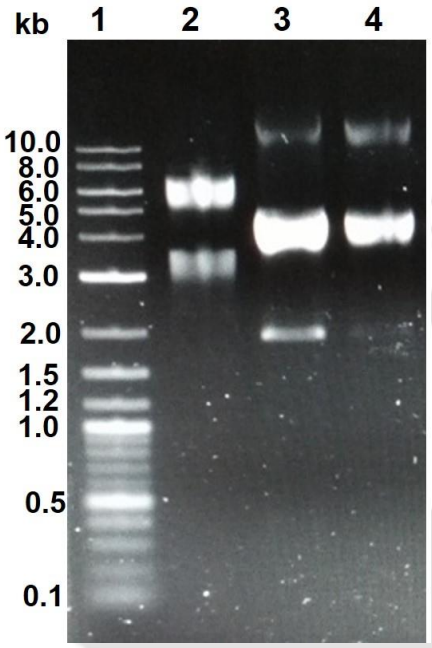


Fig. 4.3.5 Agarose gel (0.8%, w/v) showing recombinant plasmids for insert *Ba*GH5-WT and *Ba*GH5-UV2. Lanes 1, DNA marker (NEB); 2, control; 3, recombinant plasmid for *Ba*GH5-WT; 4, recombinant plasmid of *Ba*GH5-UV2.

4.3.4.4 Screening of recombinant plasmid DNA for the identification of the positive clones of BaGH5-WT and BaGH5-UV2

The positive clones of *Ba*GH5-WT and *Ba*GH5-UV2 were identified by PCR using respective plasmid DNA as template as shown Fig. 4.3.5. The 1.5 kb PCR amplified fragments for gene encoding *Ba*GH5-WT and *Ba*GH5-UV2 were detected on 0.8 %(w/v) agarose gel (Fig. 4.3.6). Thus, it was concluded from the results that each of the clone harbour the desired recombinant plasmid for *Ba*GH5-WT and *Ba*GH5-UV2 gene, respectively. The positive clones of *Ba*GH5-WT and *Ba*GH5-UV2 were further confirmed by sequencing from Eurofin India (Bangalore, India).

TH-2393_146106034

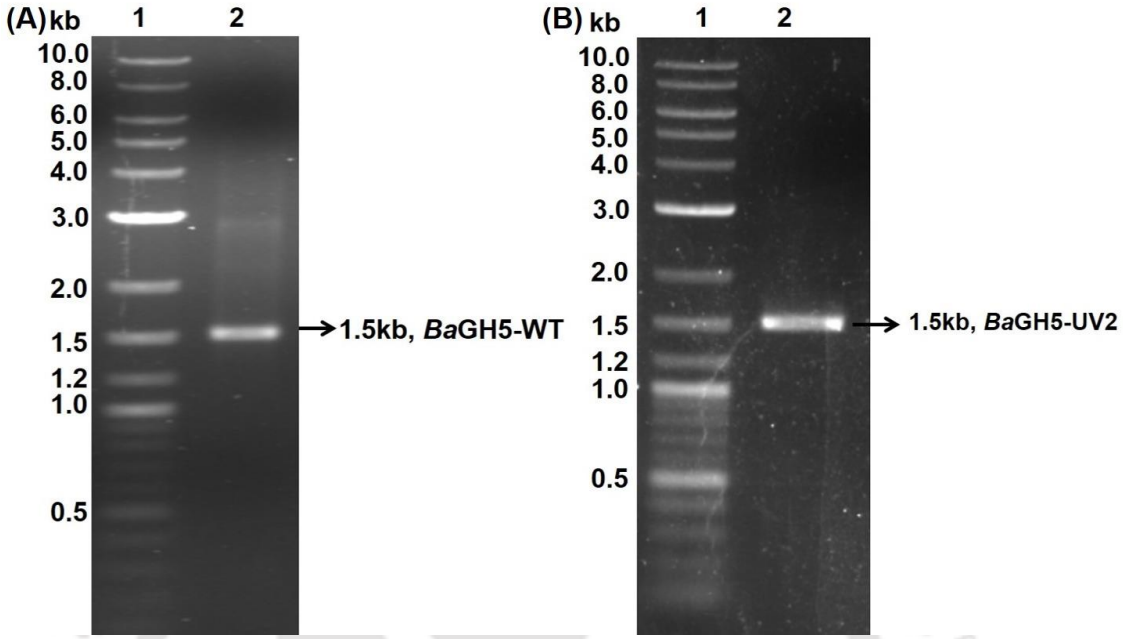


Fig. 4.3.6 Agarose gel electrophoresis (0.8%, w/v) showing the PCR-DNA confirming the clones for *Ba*GH5-WT (A) Lanes: 1, DNA ladder (NEB); 2, amplified PCR fragment of *Ba*GH5-WT (1.5 kb) from recombinant plasmid (positive clone); and for *Ba*GH5-UV2 (B) Lanes: 1, DNA ladder (NEB); 2, amplified PCR fragment of *Ba*GH5-UV2 (1.5 kb) from recombinant plasmid (positive clone).

4.3.5 Sequencing of recombinant plasmid for *Ba*GH5-WT and *Ba*GH5-UV2 gene for mutation site analysis

The gene sequence alignment of *Ba*GH5-UV2 with *Ba*GH5-WT by Clustral Omega showed a single transition mutation at position 698 bp where adenine base was substituted by guanine (Fig. 4.3.7A). In DNA, UV light cause changes in the position of hydrogen atoms in adenine, guanine, cytosine and thymine base to form tautomer as reported earlier by Watson and Crick (1953) and Topal and Fresco (1976). An amino ($-NH_2$) group in adenine base is tautomerized to form imino (=NH) group, because of which this tautomerized adenine base is paired with cytosine instead of thymine base. Hence, in the next round of replication this cytosine base, pairs with guanine base which results AT to GC transition. Due to this above point mutation in *Ba*GH5-UV2, codon GAC changed to GGC and thus, changed the amino

TH-2393_146106034

acid residue aspartic acid to glycine (Fig. 4.3.7B). Cherry *et al.*, (2003) reported the consequences of base substitution mutations in protein coding regions of a gene depend on its locations. Caniago *et al.*, (2015) showed UV irradiation cause change in CCT codon to TCG codon.

Fig. 4.3.7 (A) Nucleotide sequence alignment of *Ba*GH5-UV2 with *Ba*GH5-WT showed mutation at 698 bp; (B) Protein sequence alignment of *Ba*GH5-UV2 with *Ba*GH5-WT.

TH-2393_146106034

4.3.6 Cloning of family 5 glycoside hydrolase of *Ba*GH5-WT and *Ba*GH5-UV2 in pET-28a(+) vector

4.3.6.1 Molecular architecture of family 5 glycoside hydrolase cellulase from Bacillus amyloliquefaciens SS35

The protein sequence of *Ba*GH5-WT and *Ba*GH5-UV2 with 23 amino acid of pET-28a(+) vector sequence at N-terminal showed the molecular architecture of GH5 family upon protein BLAST against protein data bank proteins (Fig. 4.3.8). It displayed an N-terminal family 1 glycoside hydrolase (448 amino acid), family 5 glycoside hydrolase (*Ba*GH5, 247 amino acid) followed by family 3 carbohydrate binding module *Ba*CBM3 (82 amino acid) at the C-terminal. Aspartate residue to glycine substitution was occurred in GH5 region of mutant protein (Fig. 4.3.8).

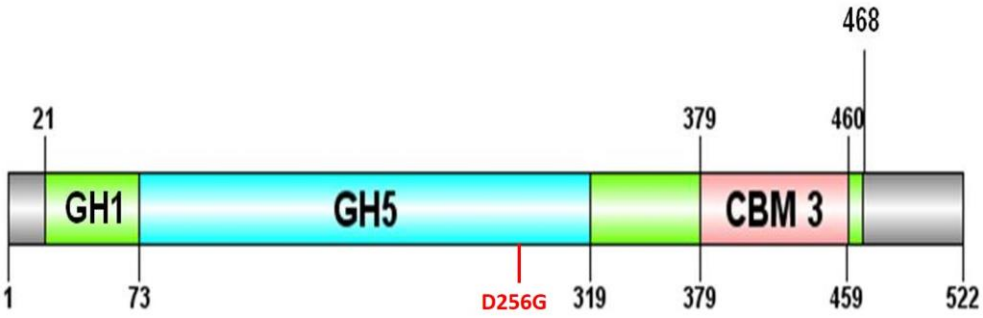


Fig. 4.3.8 Molecular architecture of family 5 glycoside hydrolase from *Bacillus amyloliquefaciens SS35*.

4.3.6.2 PCR amplification of gene encoding sequence

Analysis of *Ba*GH5-WT and *Ba*GH5-UV2 gene sequence in NEBcutter V2.0 provide the information of restriction sites. Both the gene sequence showed zero cutter for restriction sites i.e. *Nhe*I-*Xho*I and thus, help in cloning of these genes into a suitable expression vectors such as pET-28a(+) for their over expression to enhanced their production. Thus, the gene encoding *Ba*GH5-WT and *Ba*GH5-UV2 were amplified with oligonucleotide primers with *Nhe*I-*Xho*I restriction sites by using the

TH-2393_146106034

conditions as mentioned in Section 4.2.13. The PCR fragments detected on 0.8% (w/v) agarose gel were *Ba*GH5-WT (Fig. 4.3.9, lane 2) and *Ba*GH5-UV2 (Fig. 4.3.9, lane 3). The PCR products were purified from gel using gel extraction kit (Qiagen) as mentioned in Section 4.2.6 and were stored at -20°C for subsequent cloning experiments.

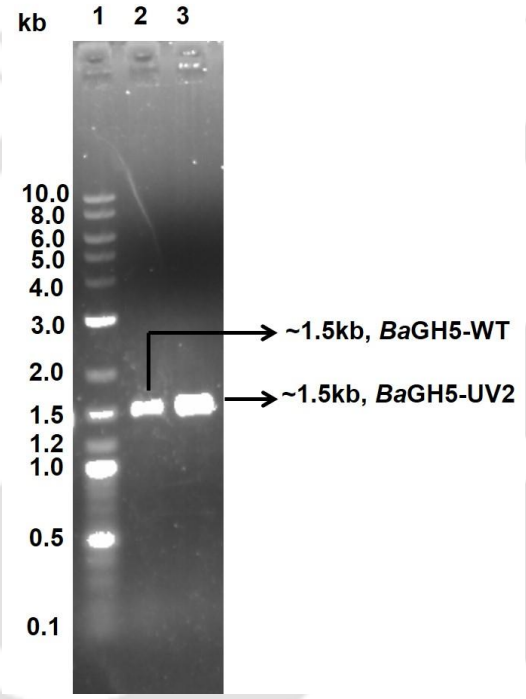


Fig. 4.3.9 Agarose gel electrophoresis (0.8%, w/v) showing amplified PCR DNA of gene encoding *Ba*GH5-WT from recombinant plasmid DNA (pHTP0). Lanes 1, DNA ladder; 2, amplified fragment of *Ba*GH5-WT (1.5 kb); 3, amplified fragment of *Ba*GH5-UV2 (1.5 kb).

4.3.6.3 Restriction enzyme digestion of PCR amplified fragments

The PCR amplified fragments of gene encoding *Ba*GH5-WT and *Ba*GH5-UV2 were digested with *Nhe*I and *Xho*I restriction enzymes for cloning in pET-28a(+) expression vector. Simultaneously, pET-28a(+) vector was also digested using the same restriction enzymes. *Nhe*I-*Xho*I digested *Ba*GH5-WT (Fig. 4.3.10, lane 2), *Ba*GH5-UV2 (Fig. 4.3.11, lane 2) and linear pET-28a(+) vector (5.36 kb, Fig. 4.3.11,

TH-2393_146106034

lane 3) fragments were visualized on 0.8 % (w/v) agarose gel. Each of the digested products were purified from gel using gel extraction kit (Qiagen) as mentioned in Section 4.2.6 and were stored at -20°C for subsequent cloning experiments.

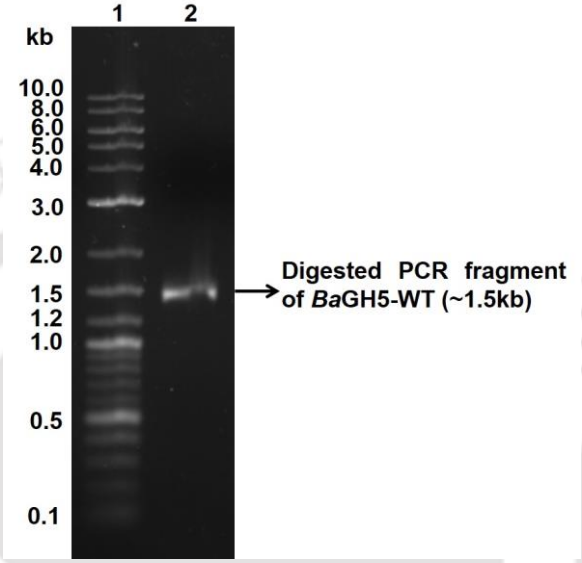


Fig. 4.3.10 Agarose gel (0.8%, w/v) showing *Nhe*I-*Xho*I digested PCR fragments of gene encoding *Ba*GH5-WT. Lanes 1, DNA ladder (NEB); 2, digested *Ba*GH5-WT.

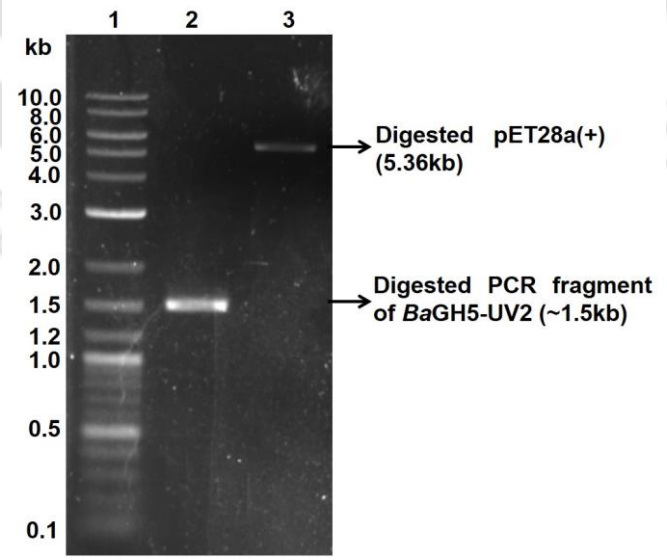


Fig. 4.3.11 Agarose gel (0.8%, w/v) showing *Nhe*I-*Xho*I-digested fragment of gene encoding *Ba*GH5-UV2 and pET-28a(+) vector. Lanes 1, DNA ladder (NEB); 2, digested *Ba*GH5-UV2; 3, digested pET-28a(+) vector (5.36 kb).

TH-2393_146106034

4.3.6.4 Ligation of NheI-XhoI digested fragments of gene encoding BaGH5-WT and BaGH5-UV2 to pET-28a(+) expression vector

The ligation of gene encoding *Ba*GH5-WT and *Ba*GH5-UV2 into the pET-28a(+) vector resulted the recombinant plasmid DNA(s) and were confirmed after the transformation in *E. coli* (TOP) 10 competent cells and digestion with *Nhe*I and *Xho*I restriction enzymes.

4.3.6.4.1 Isolation of recombinant plasmid DNAs from E. coli (TOP) 10 competent cells

The plasmid DNA of each recombinant derivative (*Ba*GH5-WT and *Ba*GH5-UV2) was isolated from 5.0 mL LB medium supplemented with kanamycin, by using the NID miniprep plamid isolation method as described in Section 4.2.18.1. The isolated plasmid DNA for each of the recombinant derivatives viz. *Ba*GH5-WT and *Ba*GH5-UV2 showed shift on 0.8% (w/v) agarose gel as compared with the pET-28a(+) plasmid, shown in Fig 4.3.12. Thus, concluded that each of the recombinant plasmid might have the required gene of interest.

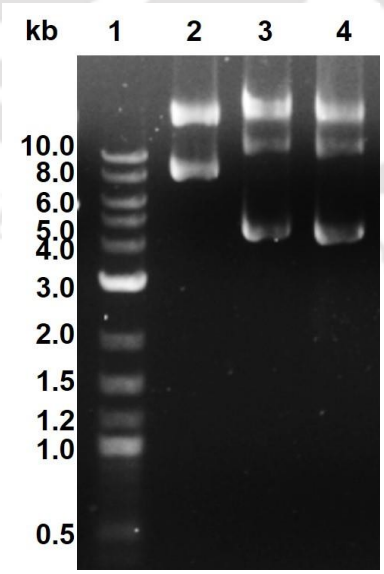


Fig. 4.3.12 Agarose gel (0.8%, w/v) showing recombinant plasmids of *Ba*GH5-WT and *Ba*GH5-UV2. Lanes 1, DNA marker (NEB); 2, pET-28(a+) plasmid; 3, recombinant plasmid of *Ba*GH5-WT; 4, recombinant plasmid of *Ba*GH5-UV2.

TH-2393_146106034

4.3.6.4.2 Screening of recombinant plasmid DNA for the identification of the positive clones of BaGH5-WT and BaGH5-UV2

The positive clones of *Ba*GH5-WT and *Ba*GH5-UV2 were identified by restriction digestion analysis of the recombinant plasmids as shown Fig. 4.3.13.8A-B. The *Nhe*I-*Xho*I digested plasmid DNA of *Ba*GH5-WT and *Ba*GH5-UV2 showed two distinct bands of the 5.36 kb pET-28a(+) vector and 1.5 kb insert (*Ba*GH5WT or *Ba*GH5UV2) on 0.8 (w/v) % agarose gel. Thus, it was concluded from the result that the clones harbour the desired recombinant plasmid for *Ba*GH5-WT or *Ba*GH5-UV2 gene, respectively.

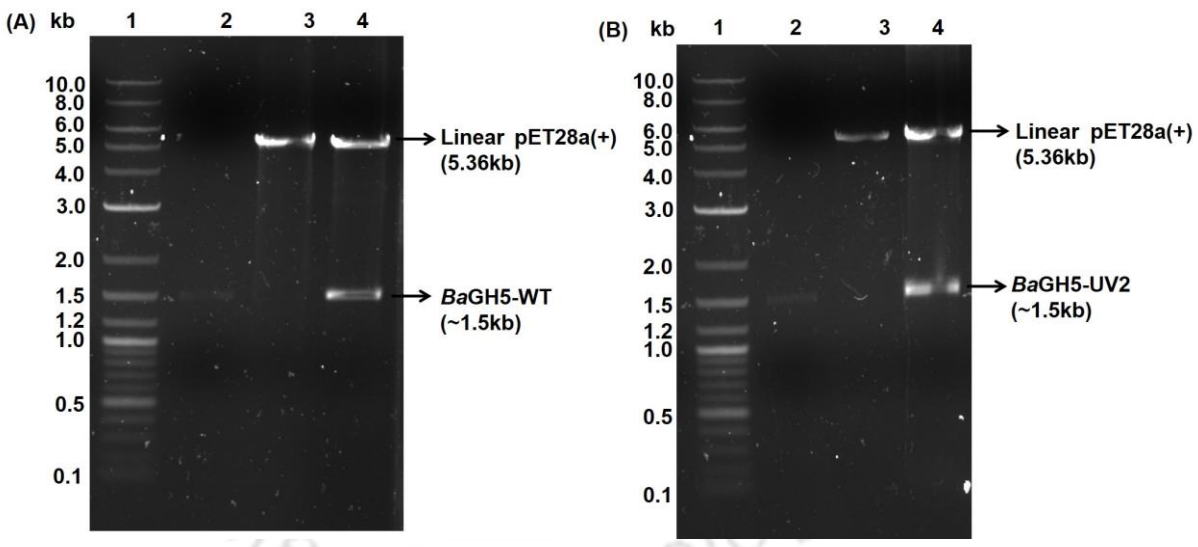


Fig. 4.3.13 Agarose gel (0.8%, w/v) showing *Nhe*I-*Xho*I digested recombinant plasmids containing *Ba*GH5-WT and *Ba*GH5-UV2 fragment. (A) Lanes 1, DNA ladder (NEB); 2, digested 1.5 kb PCR fragment of *Ba*GH5-WT; 3, digested 5.36 kb fragment of pET-28a(+); 4, digested recombinant plasmid showing 1.5 kb fragment of *Ba*GH5-WT (insert) and vector (5.369 kb, pET-28a(+)). (B) Lanes 1, DNA ladder (NEB); 2, digested 1.5 kb PCR fragment of *Ba*GH5-UV2; 3, digested 5.36 kb fragment of pET-28a(+); 4, digested recombinant plasmid showing 1.5 kb fragment of *Ba*GH5-UV2 (insert) and vector (5.369 kb, pET-28a(+)).

TH-2393_146106034

4.3.7 Expression and purification of recombinant *Ba*GH5-WT and *Ba*GH5-UV2 in *E. coli* BL21 (DE3)pLysS competent cells

Recombinant *Ba*GH5-WT and *Ba*GH5-UV2 protein showed hyper-expression in *E. coli* BL21 (DE3)pLysS competent cells at 24°C for 18 h at 180 rpm (Fig. 4.3.14 and 4.3.15, lane 3 and 4). Both the proteins were expressed as soluble proteins in cell free extracts (Fig. 4.3.14 and 4.3.15, lane 4). The purified *Ba*GH5-WT (Fig. 4.3.14, lane 6) and *Ba*GH5-UV2 (4.3.15, lane 6) protein showed homogenous bands of molecular sizes approximately, 57.6 kDa. Each of the recombinant proteins were eluted in 4x1 mL fractions from 100 mL culture. After dialysis the concentration of *Ba*GH5-WT and *Ba*GH5-UV2 protein were 0.19 mg/mL and 0.18 mg/mL, respectively, by Bradford method (Table 4.3.2).

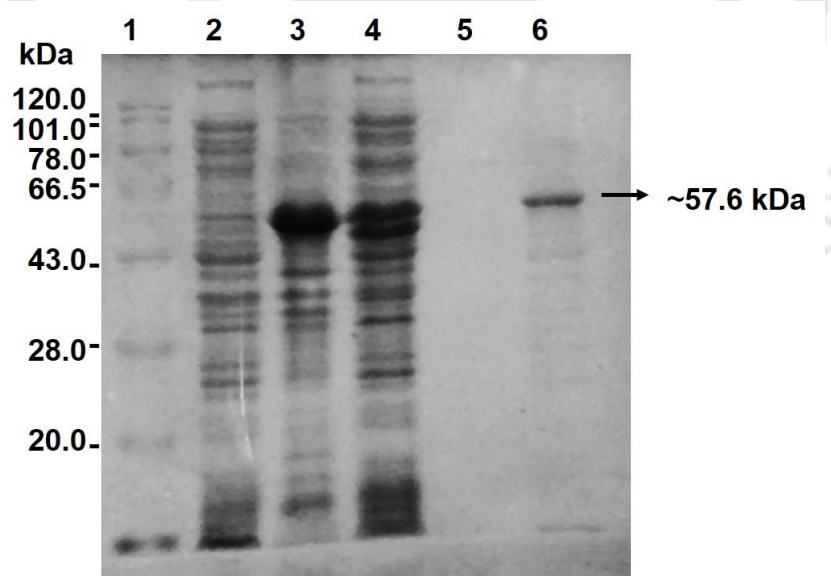


Fig. 4.3.14 SDS-PAGE (12.0%, w/v) gel showing purification of recombinant *Ba*GH5-WT from *E. coli* BL21 (DE3)pLysS cells. Lanes 1, Page Ruler protein marker (Biobharti); 2, Uninduced *Ba*GH5-WT cells; 3, Induced *Ba*GH5-WT cells after sonication; 4, Cell free extract after sonication; 5, last column wash; 6, *Ba*GH5-WT affinity column purified (57.6 kDa) protein.

TH-2393_146106034

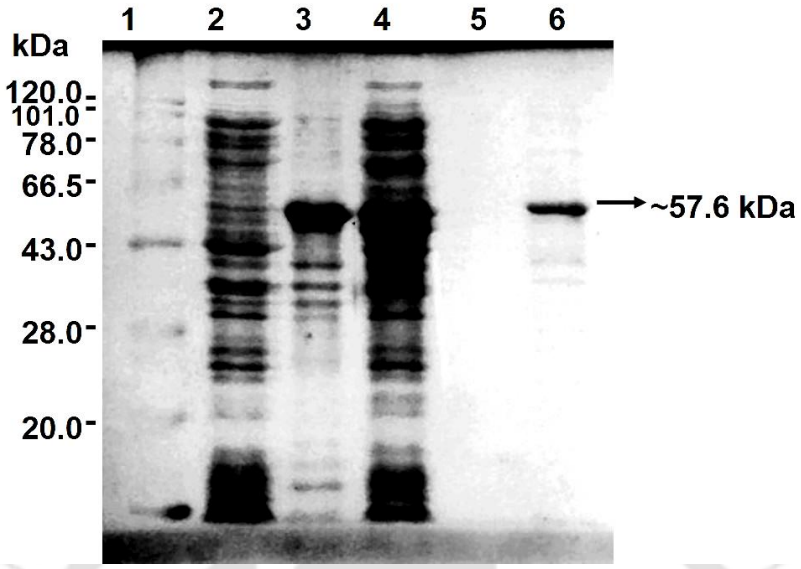


Fig. 4.3.15 SDS-PAGE (12.0%, w/v) gel showing purification of recombinant *Ba*GH5-UV2 from *E. coli* BL21 (DE3)pLysS cells. Lanes 1, Page Ruler protein marker (Biobharti); 2, Uninduced *Ba*GH5-UV2 cells; 3, Induced *Ba*GH5-UV2 cells after sonication; 4, Cell free extract after sonication; 5, last column wash; 6, *Ba*GH5-UV2 affinity column purified (57.6 kDa) protein.

Table 4.3.2 Amount of purified recombinant proteins obtained from 100 mL cultures

Recombinant derivative	Protein concentration (mg/mL)	Volume of purified protein (mL)	Total amount of purified protein (mg)
*Ba*GH5-WT	0.19 ±0.01	4.0	0.76
*Ba*GH5-UV2	0.18 ±0.04	4.0	0.72

± standard deviation (n=3)

TH-2393_146106034

4.4 Conclusion

The genes encoding family 5 glycoside hydrolase (endoglucanase), *Ba*GH5-WT (CMCase-WT) from wild-type strain and the mutant *Ba*GH5-UV2 (CMCase-UV2) from the mutant UV2 strain of *Bacillus amyloliquefaciens* SS35 were amplified to identify the genetic changes in the endoglucanase gene of UV2 strain. Degenerate oligonucleotide primers were designed from phylogenetically related spp. *Bacillus amyloliquefaciens* KHG19 for family 5 glycoside hydrolase (GH5) with GenBank accession number: AJK65578.1 from the CAZy database (http://www.cazy.org). The PCR amplified fragments of the genes encoding *Ba*GH5-WT and *Ba*GH5-UV2 showed size 1.5 kb. The genes encoding *Ba*GH5-WT and *Ba*GH5-UV2 proteins were cloned into the pHTP0 cloning vector using NZYEasy cloning kit. The gene sequence alignment of *Ba*GH5-UV2 with *Ba*GH5-WT showed transition mutation at position 698 bp where adenine base was substituted by guanine. Due to this point mutation in *Ba*GH5-UV2, codon GAC changed to GGC and thus, changed the amino acid residue aspartic acid to glycine. The analysis of *Ba*GH5-WT and *Ba*GH5-UV2 gene sequences in NEBcutter V2.0 showed zero cutter for restriction sites i.e. *Nhe*I-*Xho*I restriction sites. Thus, with these above restriction sites both the genes were cloned into pET-28a(+) expression vectors for their over-expression. The protein sequence of *Ba*GH5-WT and *Ba*GH5-UV2 showed the molecular architecture belonging to the GH5 family. It displayed an N-terminal family 1 glycoside hydrolase (448 amino acid), family 5 glycoside hydrolase (*Ba*GH5, 247 amino acid) followed by family 3 carbohydrate binding module *Ba*CBM3 (82 amino acid) at the C-terminal. The sequence alignment of *Ba*GH5-UV2 and *Ba*GH5-WT displayed the substitution mutation by changing the aspartate (D256) residue to glycine (G256) in the GH5

TH-2393_146106034

module. Both the recombinant proteins, *Ba*GH5-WT and *Ba*GH5-UV2 expressed as soluble proteins by *E. coli* BL21 (DE3)pLysS cells at 24°C. Both, the purified *Ba*GH5-WT and *Ba*GH5-UV2 proteins showed homogenous bands of the same molecular size approximately, 58 kDa. The concentrations of purified *Ba*GH5-WT and *Ba*GH5-UV2 proteins were 0.19 mg/ml and 0.18 mg/ml, respectively.

TH-2393_146106034

4.5 References

Bradford, M. (1976). A Rapid and Sensitive Method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72, 248-254.

Caniago, A., Mangunwardoyo, W., Nuswantara, S., and Lisdiyanti, P. (2015). Improvement of endoglucanase activity in *Penicillium oxalicum* ID10-T065 by ultra violet irradiation and ethidium bromide mutation. *Annales Bogorienses*, 19, 2.

Carvalho, A.L., Goyal, A., Prates, J.A.M., Bolam, D.N., Gilbert, H.J., Pires, V.M.R., Ferreira, L.M.A., Planas, A., Romão, M.J., and Fontes, C.M.G.A. (2004). The family 11 carbohydrate-binding module of *Clostridium thermocellum* Lic26A-Cel5E accommodates β-1,4 and β-1,3-1,4-mixed linked glucans at a single binding site. *Journal of Biological Chemistry*, 279, 34785-34793.

Cherry, J.R., and Fidantsef, A.L. (2003). Directed evolution of industrial enzymes: an update. *Current. Opinion in Biotechnology*, 14(4), 438-443.

Demain, A.L., and Vaishnav, P. (2009). Production of recombinant proteins by microbes and higher organisms. *Biotechnology Advances*, 27, 297-306.

Engler, M.J., and Richardson, D.C. (1982). DNA ligases. In P.D. Boyer, The Enzymes. *Academic Press, San Diego*. 15, 3-30.

Hodgson, J. (1994). The changing bulk biocatalyst market. *Nature Biotechnology*, 12, 789-90.

Lezin, G., Kosaka, Y., Yost, H.J., Kuehn, M.R., and Brunelli, L. (2011). A one-step miniprep for the isolation of plasmid DNA and lambda phage particles. PLoS One, 6(8), e23457.

TH-2393_146106034

Lorenz, P., Liebeton, K., Niehaus, F., and Eck, J. (2002). Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Current Opinion In Biotechnology*, 13(6), 572-577.

Sambrook, J., and Russel, D.W (2001). Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press, Woodbury*, New York, 1-3.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. *Plainview, Cold Spring Harbor Laboratory Press, Woodbury, New York*, 1.

Schaeffer, H.J., Leykam, J., and Walton, J.D. (1994). Cloning and targeted gene disruption of EXG1, encoding exo-beta 1, 3-glucanase, in the phytopathogenic fungus *Cochliobolus carbonum*. *Applied and Environmental Microbiology*, 60(2), 594-598.

Swartz, J.R. (1996). *Escherichia coli* recombinant DNA technology**,** *Cellular and Molecular Biology American Society of Microbiology Press, Washington* DC, 2, 1693-1711.

Taylor, E., Goyal, A., Guerreiro, C.I.P.D., Prates, J.A.M., Money, A.V., Ferry, N., Morland, C., Planas, A., Macdonald, J.A., Stick, R.V., Gilbert, H.J., Fontes, C.M.G.A., and Davies GJ. (2005). How family 26 glycoside hydrolases orchestrate catalysis on different polysaccharides? Structure and activity of a *Clostridium thermocellum* lichenase, *Ct*Lic26A. *Journal of Biological Chemistry*, 280, 32761-32767.

TH-2393_146106034

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*, 72(2), 211.

Topal, M.D., and Fresco, J.R. (1976). Complementary base pairing and the origin of substitution mutations. *Nature*, 263, 285-293.

Watson, J.D., and Crick, F.H.C. (1953). A structure for deoxyribose nucleic acids. *Nature*, 171, 737-738.

TH-2393_146106034

Chapter 5

Structure, function and biochemical properties comparison of recombinant *Ba*GH5-WT and *Ba*GH5-UV2

5.1 Introduction

Cellulases are the group of industrially important biocatalyst that have been broadly used in bioethanol production. These are classified as glycoside hydrolase enzymes belonging to the families 1, 3, 5-9, 12, 26, 30, 44, 45, 48, 51, 60, 61 and 74 as reported in CAZY database (www.cazy.org) (Schülein *et al.,* 2000¸ Wang *et al.,* 2010). The glycoside hydrolase family 5 (GH5) is one of the largest and most functionally diverse families having endo/exo-β-1,4-glucanase (cellulase), endo-β-1,4-xylanase, β-mannosidase, β-glucosylceramidase, glucan β-1,3-glucosidase, endo-β-1,6-glucosidase, mannan endo-β-1,4-mannosidase, β-1,4-cellobiosidase, sterly β-glucosidase, engoglycoceramidase, chitosanase, β-primeverosidase, xyloglucan, endo-β-1,6-galactanase, β-1,3-mannanase, arabinoxylan-specific endo-β-1,4-xylanase, mannan transglycosylase, lichenase, endo-β-1,3-1,4-glucanase, β-glycosidase, laminarinase, chitosanase, β-N-acetylhexosaminidase, β-D-galactofuranosidase and β-galactosylceramidase etc. (http://www.cazy.org/GH5_5.html, Ruiz *et al.,* 2016). Highly efficient endo-β-1,4-glucanase is required to hydrolyse the long chain of

TH-2393_146106034

cellulosic polysaccharide into short oligomers to overcome the cost of bioethanol production process. Directed evolution, site-directed mutagenesis, strain improvement are the promising approaches of protein engineering to improve the catalytic efficiency, broad range pH tolerance and broad range working temperature range of endo-β-1,4-glucanases (cellulases).

The structural studies of endo-β-1,4-glucanase (cellulase) from glycoside hydrolase family 5, GH5 showed that the catalytic domain has $(\beta/\alpha)_8$ TIM-barrel fold, where eight parallel β-strands and eight parallel α-helices are connected by αβ or βα loops (Wierenga 2001; Wang *et al.*, 2010; Tu *et al.*, 2016 and Ruiz *et al.*, 2016). These αβ or βα loops are normally found on the protein surface and are critical for catalytic activity and stability. Tu *et al.*, (2016) reported that the mutation of Tyr residue to Gly residue in loop of cellulase Cel12A from *Thermotoga maritima* increased the specific activity by 1.7-fold. In Cel5B enzyme from *Clostridium thermocellum* conformational changes of $\beta_3\alpha_3$ loop governs the protonation of catalytic residue, Glutamate (Badieyan *et al.*, 2012). In *Ct*Cel5E cellulase from *Clostridium thermocellum*, the mutation of Phe267 to Ala showed 4-fold increase in k_{cat} value due to the breakage of the hydrophobic interactions because the αβ flexible loop was relocated (Yuan *et al.*, 2015). Therefore, the αβ or βα loops connecting the secondary structures in catalytic domain have important role in enzyme catalysis (Zheng *et al.*, 2018).

Molecular dynamics (MD) simulation study of a protein offers a great deal of information regarding the protein flexibility, protein-ligand binding characteristics and stability. The flexibility of a protein plays an important role in their folded state to accomplish their function such as activity modulation, ligand binding and

TH-2393_146106034

macromolecular interactions (Badieyan *et al.*, 2012; Yu *et al.*, 2015). The molecular docking analysis predicts the ligand-receptor complex structure in the protein tertiary structure via computation methods (Ballester and Mitchell 2010). It helps in predicting the change in hydrophilic and hydrophobic interactions in the mutant protein which are responsible for change in catalytic efficiency. Yuan *et al.,* (2015) reported that the mutation of Phe267 to Ala in cellulase, *Ct*Cel5E displayed reduced hydrophobic interactions with the ligand resulting in 4-fold increase in catalytic efficiency of enzyme. In the present study, the comparison of biochemical characterization of purified recombinant *Ba*GH5-UV2 enzyme (Chapter 4, Section 4.3.7) from UV2 mutant bacterial strain with *Ba*GH5-WT enzyme (Chapter 4, Section 4.3.7) from wild-type *Bacillus amyloliquefaciens* SS35 was carried out. The change in temperature optima, pH optima, thermos-stability, pH stability and catalytic efficiency for mutant enzyme, *Ba*GH5-UV2 was determined. The structural, molecular docking and MD simulation studies of *Ba*GH5-WT and *Ba*GH5-UV2 were performed to determine the role of mutated residue aspartate 256 to glycine (Chapter 4, Section 4.3.6.1).

TH-2393_146106034

5.2 Materials and Methods

5.2.1 Substrates and reagents

Carboxymethylcellulose sodium salt (CMC-Na) (low viscosity), lichenan, β-glucan, avicel, birchwood xylan, oat spelt xylan, galactomannan, D-glucose, cellobiose and Bradford's reagent were procured from Sigma-Aldrich Co. LLC., USA. Luria-Bertani (LB) broth, sodium carbonate, sodium potassium tartarate, sodium bicarbonate, sodium sulphate, sodium phosphate (monobasic), sodium phosphate (dibasic), sodium acetate and Tris hydrochloride (Tris-HCl) were procured from HiMedia Laboratories Pvt. Ltd., India. Sodium arsenate, ammonium molybdate, sulphuric acid, hydrochloric acid, ortho-phosphoric acid and acetic acid were purchased from Merck Limited, India.The antibiotics kanamycin and chloramphenicol were procured from HiMedia Laboratories Pvt. Ltd., India.

5.2.2 Culture conditions, expression and purification of recombinant *Ba*GH5-WT and *Ba*GH5-UV2

Recombinant *E. coli* BL21(DE3) pLysS competent cells (Novagen, EMD4 BioScience, Germany) harbouring plasmids *Ba*GH5-WT or *Ba*GH5-UV2 were grown and expressed in LB medium according to the protocol as mentioned earlier in Chapter 4, Section 4.3.7. *Ba*GH5-WT and *Ba*GH5-UV2 proteins were purified according to the method as given in Chapter 4 Section 4.3.7.

5.2.3 Effect of temperature on *Ba*GH5-WT and *Ba*GH5-UV2 activity and stability

The optimum temperature of the purified *Ba*GH5-WT and *Ba*GH5-UV2 for enzyme assay was determined by Nelson (1944) and Somogyi (1945) method of reducing sugars estimation as described in chapter 2, section 2.2.5. The reaction mixture containing 10 µL of *Ba*GH5-WT enzyme (0.1 mg/mL) or 10 µL of *Ba*GH5-

TH-2393_146106034

UV2 enzyme (0.025 mg/mL) at 1% (w/v) CMC-Na in 50 mM sodium acetate buffer, pH 5.5 was incubated in the temperature range, 30°C-75°C for 3 min. The thermostability studies of the purified *Ba*GH5-WT and *Ba*GH5-UV2 was performed by pre-incubating the 50 µL of *Ba*GH5-WT (0.3 mg/mL) or 50 µL of *Ba*GH5-UV2 (0.1 mg/mL) at different temperatures, ranging from 30 to 75°C for 1 h. The residual activity for respective samples was calculated by estimating the released reducing by methods reported by Nelson (1944) and Somogyi (1945) as described in Chapter 2, section 2.2.5.

5.2.4 Effect of pH on *Ba*GH5-WT and *Ba*GH5-UV2 activity and stability

The optimum pH of the purified *Ba*GH5-WT and *Ba*GH5-UV2 enzymes were determined with 10 µL of *Ba*GH5-WT enzyme (0.1 mg/mL) or 10 µL of *Ba*GH5-UV2 enzyme (0.025 mg/mL) in 1% (w/v) CMC-Na in appropriate buffers; 50 mM sodium acetate (pH 3.5-5.5), 50 mM sodium phosphate buffer (pH 5.8-8.0) and 50 mM Tris-HCl (pH 8-9.5). The reaction mixture at different pH for *Ba*GH5-WT were incubated at 55°C for 3 min and for *Ba*GH5-UV2 at 65°C for 3 min and CMCase activity was assayed by reducing sugar estimation using method of Nelson (1944) and Somogyi (1945) as described in Chapter 2, section 2.2.5. The comparison of pH stability of the purified *Ba*GH5-WT and *Ba*GH5-UV2 enzymes was performed by pre-incubating the 10 µL of *Ba*GH5-WT (0.3 mg/mL) or 10 µL of *Ba*GH5-UV2 (0.1 mg/mL) at 35°C for 1 h in different buffers at a pH ranging from 3.5 to 9.5. The residual activity for respective samples was determined by methods of Nelson (1944) and Somogyi (1945) for reducing sugar estimation as described in Chapter 2, section 2.2.5.

TH-2393_146106034

5.2.5 Comparison of enzyme activity of purified *Ba*GH5-WT and *Ba*GH5-UV2 against soluble and insoluble substrates

The enzyme activity of purified *Ba*GH5-WT and *Ba*GH5-UV2 was tested against substrates, β-D-glucan, lichenan, CMC-Na, cellulose powder, avicel, birchwood xylan and galactomannan. The *Ba*GH5-WT enzyme reaction was carried out at 1.0% (w/v) substrate dissolved in 50 mM sodium acetate buffer, pH 5.5 by incubating at 55°C for 3 min. The *Ba*GH5-UV2 enzyme reaction was carried out at 1.0% (w/v) substrate dissolved in 50 mM sodium acetate buffer, pH 5.5 by incubating at 65°C for 3 min. 100 µL of enzyme reaction mixture contained 10 µL of the enzyme (*Ba*GH5-WT, 0.1 mg/mL and *Ba*GH5-UV2, 0.025 mg/mL). In the case of avicel and oat spelt xylan, the reaction mixture was incubated under shaking conditions (160 rpm, 3 min). The resulting reducing sugars were determined by method reported by Nelson (1944) and Somogyi (1945) as described in Chapter 2, section 2.2.5 by measuring the absorbance at 500 nm, on a spectrophotometer (Varian, Cary 100 Bio) using glucose, xylose and mannose as standards. The β-glucosidase activity of *Ba*GH5-WT (0.1 mg/mL) and *Ba*GH5-UV2 (0.025 mg/mL) was determined by GOD-POD method against 1% (w/v) cellobiose (Raabo *et al.,* 1960). The 100 µL of reaction mixture contained 10 µL of the enzyme in 1% (w/v) cellobiose dissolved in 50 mM sodium acetate buffer, pH 5.5. The reaction mixture was incubated at 55°C for *Ba*GH5-WT and at 65°C for *Ba*GH5-UV2 enzyme for 3 min. After that the reaction was stopped by boiling it for 15 min in a boiling water bath. Then 2 µL of reaction mixture was added to a well in 96 well microtitre plate followed by addition of 200 µL of glucose reagent L1 (from kit) and incubated at 37ºC in an incubator for 15 min.

TH-2393_146106034

The β-glucosidase activity was determined by measuring the absorbance at 505 nm, on a spectrophotometer (Varian, Cary 100 Bio) using glucose standards.

5.2.6 Comparison of kinetic parameters of *Ba*GH5-WT and *Ba*GH5-UV2

The kinetic parameters i.e. K_m and V_{max} of *Ba*GH5-WT and *Ba*GH5-UV2 were determined by fitting the initial rate data to the Michaelis–Menten equation by Lineweaver-Burk double reciprocal plot (Michaelis and Menten 1913). The activity of *Ba*GH5-WT was assayed against CMC-Na and β-D-glucan concentration ranging from 0.01% to 2% (w/v) in 50 mM sodium acetate buffer (pH 5.5) at 55°C for 3 min. The activity of *Ba*GH5-UV2 was assayed against CMC-Na and β-D-glucan concentration ranging from 0.01% to 2% (w/v) in 50 mM sodium acetate buffer (pH 5.5) at 65°C for 3 min. The enzyme activity was calculated by estimating the reducing sugar by the methods reported by Nelson (1944) and Somogyi (1945) as described in Chapter 2, section 2.2.5. The data for kinetic parameters were analyzed by GraphPad Prism software (GraphPad Software Inc., San 138 Diego, CA) and the kinetic constants were calculated from the best fit.

5.2.7 Homology modeling of *Ba*GH5-WT and *Ba*GH5-UV2 and quality assessment

The 3-D structure of wild-type *Ba*GH5-WT was predicted by online server RaptorX (http://raptorx.uchicago.edu/ StructPredV2/predict/). The *Ba*GH5-WT protein sequence along with the pET28(a+) vector sequence was uploaded at the online server. The 3-D modeled structure of *Ba*GH5-WT was predicted by using the highest suitable template from BLAST analysis and the modeled structure was generated as single PDB file (Källberg et *al.*, 2012). The 3-D structure of *Ba*GH5-UV2 protein was generated after mutating the Asp 256 residue to Gly in Swiss PDB

TH-2393_146106034

viewer software. The energy of modeled wild-type *Ba*GH5-WT and mutant *Ba*GH5-UV2 structures were minimized by using a software, Swiss PDB viewer. Swiss PDB viewer not only minimizes the energy of the input structure but also repairs the deformities in the 3-D structure by using a GROMOS 43B1 force field (Gunsteren *et al.,* 1996). The quality of energy minimized 3-D structure of *Ba*GH5-WT and *Ba*GH5-UV2 was assessed by submitting the structure of both the proteins at verification server (http://services.mbi.ucla.edu/SAVES/). In this server, structure validation parameters Ramachandran plot using PROCHECK server (Laskowski *et al.,* 1993), Z-score deviation using ProSA server (Wiederstein & Sippl, 2007), verify 3D (Eisenberg *et al*., 1997) and ERRAT (Colovos & Yeates, 1993) were predicted.

5.2.8 Prediction of active site residues, divergence in amino acid at 256 position and mechanism of action

The protein sequence of family 5 GH endoglucanase from *Bacillus subtilis* (PDB ID 3PZT|A), *Salipaludibacillus agaradhaerens* (PDB ID 1E5J|A), *Geobacillus* sp., (PDB ID 4XZB|A), *Cytophaga hutchinsonii* (PDB ID 5IHS|A), *Caldicellulosiruptor saccharolyticus* (PDB ID 5ECU|A) were retrieved from protein-protein blast against PDB for similar homologues. The Multiple Sequence Alignment (MSA) of above mentioned homologous protein sequences was accomplished through Clustal Omega tool (http://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers *et al.*, 2011). The MSA file generated was viewed under ESPript 3.0 (Robert and Gouet, 2014) to identify conserved and semi conserved residues in the family 5 GHs. The divergence in amino acid residues at position 256 were determined by the MSA of 16 homologues (similarity in range of 27%-89% against protein-protein PDB blast) of bacterial cellulases of GH5 family from *Bacillus subtilis* (PDB ID 3PZT|A, PDB ID 1LF1|A), *Salipaludibacillus agaradhaerens* (PDB ID 1E5J|A, PDB ID 1H11|A),

TH-2393_146106034

Bacillus agaradherans (PDB ID 1A3H|A), *Geobacillus* sp. (PDB ID 4XZB|A), *Bacillus halodurans* (PDB ID 4V2X|A), *Cytophaga hutchinsonii* (PDB ID 5IHS|A), *Bacillus sp.* KSM-635 (PDB ID 1G01|A), *Caldicellulosiruptor saccharolyticus* (PDB ID 5ECU|A), Metagenome holotolerant cellulase (PDB ID 5FIP|A), *Thermobifida fusca* (PDB ID 2CKS|A), *Erwinia Chrysanthemi* (PDB ID 1EGZ|A), *Pseudoalteromonas haloplanktis* (PDB ID 1TVN|A) and *Hungateiclostridium thermocellum* (PDB ID 5BYW|A and PDB ID 4U3A|A). The MSA file generated was viewed under ESPript 3.0 to identify the variation or divergence in amino acid residues at 256 position in the family 5 GHs.

Structural alignment of *Ba*GH5-WT and *Ba*GH5-UV2 with PDB structure of endoglucanase 3PZT|A from *Bacillus subtilis,* 1E5J|A from *Salipaludibacillus agaradhaerens* and 5IHS|A from *Cytophaga hutchinsonii* was done in PyMOL 2.0 software to know the catalytically important amino acids residues. The distances among catalytic amino acids of *Ba*GH5-WT was measured by using PyMOL 2.0 (Schrodinger, 2017) software to deduce the hydrolytic mechanism.

5.2.9 Molecular dynamics simulation of modeled *Ba*GH5-WT and *Ba*GH5-UV2 structure

Molecular dynamics (MD) simulation studies for modeled wild-type *Ba*GH5-WT and its mutant *Ba*GH5-UV2 structure was accomplished by using Gromacs v 5.14 to explore the possible mechanism of increase in catalytic efficiency and pH stability in a supercomputer facility (Param-Ishan) available at Indian Institute of Technology Guwahati, Guwahati, Assam. For calculating the protein forces GROMOS96 53a6 was used as a force field. The modeled *Ba*GH5-WT or *Ba*GH5-UV2 was placed inside a cubic box with dimension 6.38 × 7.73 × 10.12 and volume

TH-2393_146106034

1949.8 nm^3 with single point charge (SPC) along with the water molecules. The MD simulation contained 60088 water molecule and 22 Na^+ was used as counter ions to neutralize the negative charges on *Ba*GH5-WT or *Ba*GH5-UV2. The energy of modeled *Ba*GH5-WT and *Ba*GH5-UV2 structure were minimized using steepest descent method using 50,000 iteration steps with cut-off till 1000 $kJmol^{-1}$. Then each system was equilibrated in NVT ensemble for 500 ps at the constant number of particles, volume, and temperature equilibration of the whole system. The production run for each system was carried out for 80 ns with NPT ensemble acquiring 2 fs integration time. The linear constraint solver (LINC) algorithm was applied to constrain the bonds linked with hydrogen atoms and radius of gyration. The modeled *Ba*GH5-WT or *Ba*GH5-UV2 structure was analyzed as a time-dependent function throughout the simulation to verify their stability in the solvent system. The root-mean-square deviation (RMSD) and the root-mean-square fluctuation (RMSF) of the both the simulated system(s) were analyzed by gmx rms command and gmx rmsf command, respectively. The solvent accessible surface area (SASA) and radius of gyration (Rg), of both the simulated systems were analyzed by using the program gmx sasa and gmx gyrate, respectively.

5.2.10 Ligand-binding analysis of simulated *Ba*GH5-WT and *Ba*GH5-UV2 structure by molecular docking

Molecular docking studies were carried out to understand the interactions between enzymes (*Ba*GH5-WT or *Ba*GH5-UV2) and the ligands viz. cellobiose (G2), cellotriose (G3), cellotetraose (G4), cellopentaose (G5) and cellohexaose (G6), respectively, using Autodock 4.2.1. In Autodock 4.2.1, inbuilt Lamarckian Genetic Algorithm (LGA) and empirical free energy scoring function were used to analyse the

TH-2393_146106034

interactions of ligands with the enzyme (Goodsell and Olson 1990). All the ligand molecules were retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov). Input files of all the ligands were prepared by adding polar hydrogens and assigning atomic partial charges to the ligands by Gasteiger method and to protein by Kollman method. This was followed by removing all non-polar hydrogens and merging their charges to the carbon atoms. In Autodock 4.2.1, grid-based method was used to evaluate the energy where grid points contained pre-calculated affinities against different atom of the ligand. Therefore, grid maps were assigned to each atom type present in ligand molecule as well as protein accompanied by electrostatic and desolvation maps using the AutoGrid (Huey *et al*., 2007). Grid box was set at 110, 100, 96 (x,y,z) coordinate points with 0.375 Å grid spacing for modeled simulated *Ba*GH5-WT and 90, 92, 96 (x,y,z) coordinate points with 0.375 Å grid spacing for modeled simulated *Ba*GH5-UV2, where it covered the entire active site pocket. For each ligand, 80 LGA runs were performed using a Lamarckian Genetic Algorithm. The docked confirmation of ligand protein which have the minimum lowest free energy of binding (ΔG°), were saved. The resulting *Ba*GH5-WT-ligand complex or *Ba*GH5-UV2-ligand complex were analyzed in PyMOL 2.1. The 2D protein-ligand complex was generated by LigPlot software for analysing the change in possible hydrophilic and hydrophobic interaction.

TH-2393_146106034

5.3 Results and Discussion

5.3.1 Comparison of effect of temperature on *Ba*GH5-WT and *Ba*GH5-UV2 activity and stability

The purified wild-type endoglucanase, *Ba*GH5-WT gave maximum activity of 4.5 U/mg (considered as 100% CMCase activity) against 1% (w/v) CMC-Na at temperature optima 55°C, whereas, its mutant *Ba*GH5-UV2 with single substitution of Asp to Gly residue at 256 position showed 10°C shift in temperature optima and gave maximum CMCase activity of 43.6 U/mg (considered as 100% CMCase activity) at 65°C (Fig. 5.3.1A). Moreover, the *Ba*GH5-UV2 enzyme displayed 92% and 71% of relative CMCase activity at 70°C and 75°C, respectively, whereas *Ba*GH5-WT enzyme showed 34% and 3% of relative CMCase activity at 70°C and 75°C, respectively (Fig. 5.3.1A). *Ba*GH5-UV2 showed 43% of relative activity whereas *Ba*GH5-WT enzyme showed negligible activity at 80°C (Fig. 5.3.1A). Earlier reports of Liang *et al.,* (2011) also showed that the substitution of single amino acid residue His 82 to Tyr in Cel5A *Thermoanaerobacter tengcongensis* MB4 decreased the temperature optima by 10°C. Zheng *et al.*, 2018 reported the substitution of Asp 233 to Ala and Asn 233 to Gly in *Gt*Cel5A from *Gloeophyllum trabeum* by site-directed mutagenesis showed 10°C decrease in temperature optima. *Ba*GH5-UV2 did not show any effect on thermostability as it retained 100% of relative CMCase activity upto 45°C for 60 min similar to *Ba*GH5-WT (Fig. 5.3.1B). This suggested that the *Ba*GH5-UV2 can be efficiently used in simultaneous saccharification and fermentation (SSF) at 30°C.

TH-2393_146106034

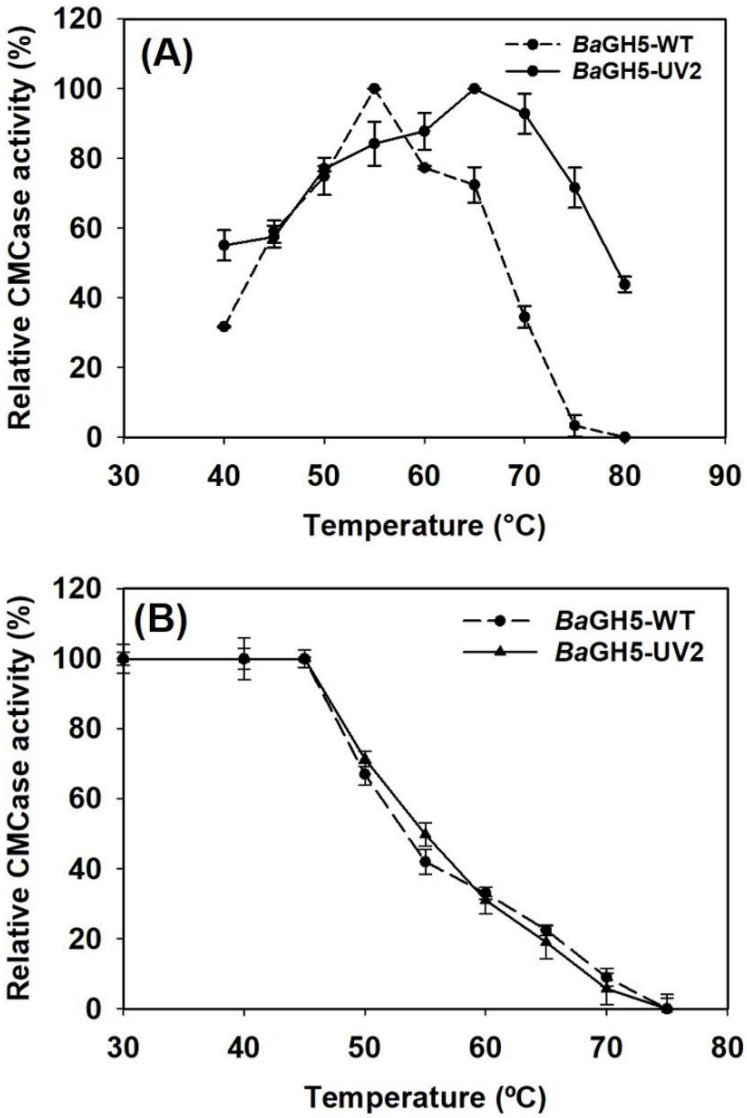


Fig. 5.3.1 (A) Effect of temperature on *Ba*GH5-WT and *Ba*GH5-UV2 activity and (B) Effect of temperature on *Ba*GH5-WT and *Ba*GH5-UV2 stability.

TH-2393_146106034

5.3.2 Comparison of pH effect on *Ba*GH5-WT and *Ba*GH5-UV2 activity and stability

The enzyme, *Ba*GH5-UV2 showed broader range pH optima with 50 mM sodium acetate buffer from pH 5.0 to 6.0 against 1% (w/v) CMC-Na at 65°C (Fig. 5.3.2A) as compared with the wild-type enzyme, *Ba*GH5-WT which showed narrow range pH optima from pH 5.5 and 6.0 against 1% (w/v) CMC-Na at 55°C (Fig. 5.3.2B). *Ba*GH5-UV2 retained 88% of CMCase activity in 50 mM sodium acetate buffer at pH 4.5 whereas *Ba*GH5-WT retained 22% of activity at same pH. This result displayed that substitution of Asp256 to Gly in *Ba*GH5-UV2 shifted the optimum pH range towards the acidic pH. Zheng *et al.*, 2018 reported that the substitution of Asp233 to Val in *Gt*Cel5A from *Gloeophyllum trabeum* showed shift in pH optima towards lower acidic value from pH 4.0 to 5.0.

*Ba*GH5-WT enzyme retained 100% of relative CMCase activity in neutral and slightly alkaline pH range, pH 7.0-7.5 (Fig. 5.3.3A), whereas, its mutant, *Ba*GH5-UV2 retained 100% of CMCase activity in acidic pH range, 5.5-6.0 after 1 h of incubation in 20 mM of respective buffer at 35°C (Fig. 5.3.3B). The enzyme, *Ba*GH5-UV2 retained more than 90% of relative CMCase activity over a broad pH range from 5.0 to 7.0, whereas, *Ba*GH5-WT retained around 80% of relative CMCase activity in narrow pH range from 6.0 to 6.5 after 1 h of incubation in 20 mM of respective buffer at 35°C (Fig. 5.3.3A). Glycine residue is a uncharged residue therefore substitution of aspartate a negatively charged residue to a uncharged residue changed the microenvironment of the catalytic core and increased the pH tolerance. The pH stability shift of *Ba*GH5-UV2 enzyme towards the acidic conditions makes it more favourable for bioethanol production by *Saccharomyces cerevisiae* in SSF process. Zheng *et al.*, 2018 reported that substitution of Asn 233 to Asp and Asn 233 to Gly in

TH-2393_146106034

*Gt*Cel5 increases the thermostability of endoglucanase under both acidic and alkaline conditions.

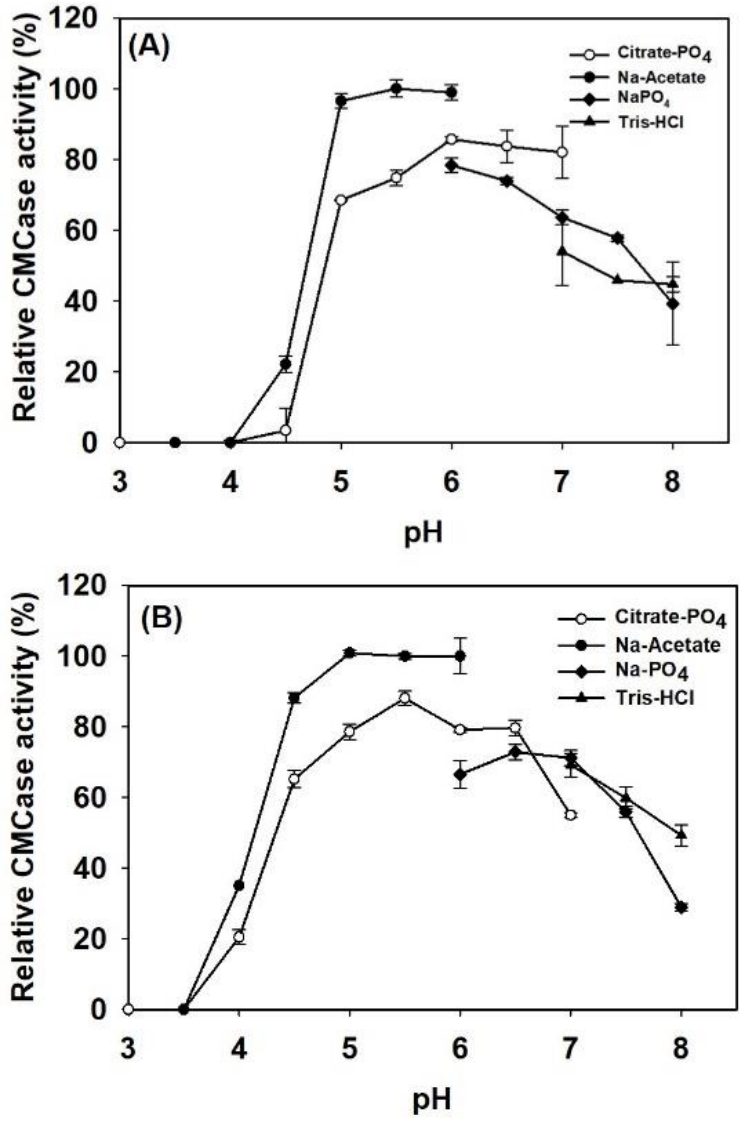


Fig. 5.3.2 Effect of pH on CMCase activity; (A) *Ba*GH5-WT and (B) *Ba*GH5-UV2.

TH-2393_146106034

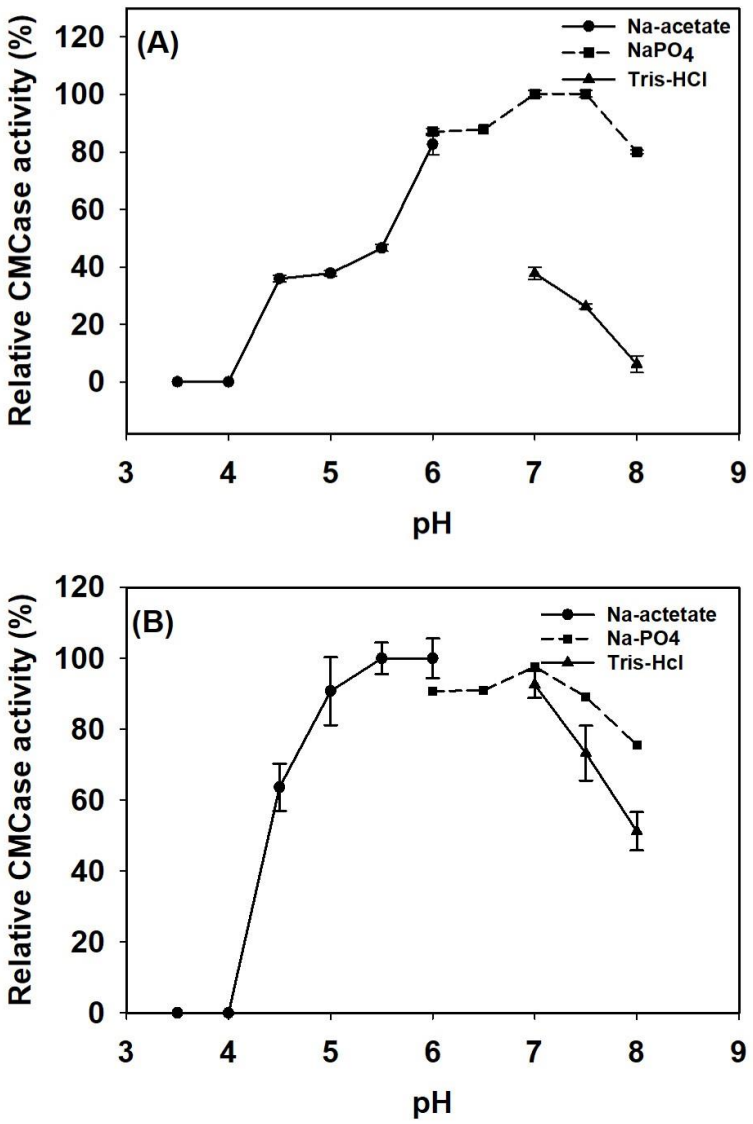


Fig. 5.3.3 Effect of pH on CMCase stability; (A) *Ba*GH5-WT and (B) *Ba*GH5-UV2.

TH-2393_146106034

5.3.3 Comparison of specific activity of purified recombinant *Ba*GH5-WT and *Ba*GH5-UV2 against soluble and insoluble substrates

The specific activities of the purified recombinant *Ba*GH5-WT and *Ba*GH5-UV2 enzymes were determined against CMC-Na, β-D-glucan, lichenan, cellulose powder, avicel, cellobiose, laminarin, birchwood xylan and galactomannan (Table 5.3.1). Mutant *Ba*GH5-UV2 showed highest specific activity against barley β-glucan (94.6 U/mg) followed by lichenan (77.1 U/mg). It showed moderate specific activity against CMC-Na (43.6 U/mg) and lower specific activity against crystalline cellulosic substrate avicel (11.4 U/mg). These results of specific activities of *Ba*GH5-UV2 against all the above four cellulosic substrates were approximately, 10-fold higher than the specific activity displayed by wild-type enzyme *Ba*GH5-WT. Both the wild-type, *Ba*GH5-WT and mutant *Ba*GH5-UV2 enzymes did not show any detectable activity against cellobiose, laminarin, birchwood xylan and galactomannan. Liang *et al.,* (2011) reported the substitution of Leu 321 to Val in Cel5A by directed evolution increased the specific activity of endoglucanase by 1.3-fold. Li *et al.,* (2017) reported an increase in the mannanase activity of *Rm*Man5A from *Rhizomucor miehei* by directed evolution by 3-fold in m*Rm*Man5A mutant by error-prone PCR. Zheng *et al.,* (2018) reported that the mutants Asn 233 to Ala or Asn 233 to Gly of *Gt*Cel5 generated by site-directed mutagenesis show 1.3-or 1.7-fold increase in the specific activity against barley β-glucan.

TH-2393_146106034

Table 5.3.1 Comparison of enzyme activity of purified *Ba*GH5-WT and *Ba*GH5-UV2

Substrate	*Ba*GH5-WT (U/mg)	*Ba*GH5-UV2 (U/mg)	Fold increase
CMC-Na	4.5 ± 0.01	43.6 ± 1.6	9.7
β-D-Glucan	10.2 ± 0.1	94.6 ± 0.2	9.3
Lichenan	7.4 ± 0.1	77.1 ± 2.5	10.4
Cellulose powder	1.6 ± 0.1	4.3 ± 0.38	2.7
Avicel	1.2 ± 0.04	11.4 ± 2.3	9.5
Cellobiose	ND	ND	-
Laminarin	ND	ND	-
Birchwood xylan	ND	ND	-
Galactomannan	ND	ND	-

*All the assays were performed at 1% (w/v) substrate, 50 mM sodium acetate buffer (pH 5.5), 3 min, for Ba*GH5-WT at *55°C and for Ba*GH5-UV2 at *65°C.*

ND = No activity detected. Values are mean SE (n=3).

5.3.4 Comparison of kinetic parameters of *Ba*GH5-WT and *Ba*GH5-UV2

The kinetic parameters K_m and V_{max} for wild-type *Ba*GH5-WT and mutant *Ba*GH5-UV2 enzyme were determined by Lineweaver-Burk plot (Fig. 5.3.4). The mutant enzyme *Ba*GH5-UV2 showed 8.5-fold (100.5 U/mg) and 7-fold (42.05 U/mg) higher V_{max} against β-glucan and CMC-Na, respectively, as compared with the wild-type *Ba*GH5-WT enzyme (Table 5.3.2). *Ba*GH5-UV2 enzyme showed 3.1 and 2.1-fold lower K_m against CMC-Na and β-glucan respectively, as compared with wild-type *Ba*GH5-WT enzyme. A higher V_{max} and lower K_m value is desirable for an enzyme to have higher affinity towards their respective substrates (Singh *et al.*, 2014). The catalytic efficiency (k_{cat}/K_m) of mutant enzyme *Ba*GH5-UV2 increased to 21.9-fold (2.0×10^4 $mg^{-1}mL^1min^{-1}$) and 18.4-fold (8.3×10^4 $mg^{-1}mL^1min^{-1}$) against CMC-Na and β-glucan respectively, as compared with *Ba*GH5-WT (Table 5.3.2). This showed the enhanced potential of mutant enzyme, *Ba*GH5-UV2 in hydrolysis of complex cellulosic polysaccharides into reducing sugars.

TH-2393_146106034

Table 5.3.2 Kinetic parameters of *Ba*GH5-WT and *Ba*GH5-UV2

Substrate (% w/v)	*Ba*GH5-WT				*Ba*GH5-UV2				Fold increase *catalytic efficiency* (K_{cat}/K_m)*
	K_m ($mgmL^{-1}$)	V_{max} (U/mg)	k_{cat} (min^{-1})	k_{cat}/K_m ($mg^{-1}\ mL^{1}min^{-1}$)	K_m ($mgmL^{-1}$)	V_{max} (U/mg)	k_{cat} (min^{-1})	K_{cat}/K_m ($mg^{-1}\ mL^{1}min^{-1}$)	
CMC-Na	0.38 ± 0.11	6.0 ± 0.6	3.46×10^{2}	9.1×10^{2}	0.12 ± 0.05	42.05 ± 3.1	2.4×10^{3}	2.0×10^{4}	21.9
β-Glucan	0.15± 0.04	11.71 ± 0.9	6.74×10^{2}	4.5×10^{3}	0.07 ± 0.01	100.5 ± 8.8	5.7×10^{3}	8.3×10^{4}	18.4

values are mean SE (n=3)
All the assays were performed under 50 mM sodium acetate buffer (pH 5.5), 3 min for BaGH5-WT enzyme at 55°C and for BaGH5-UV2 enzyme at 65°C.

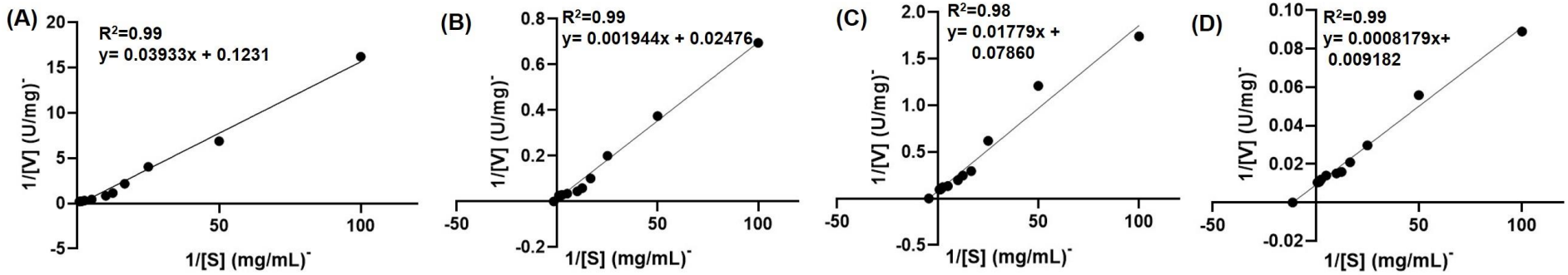


Fig. 5.3.4 Linewaever-Burk plot for determination of kinetics of (A) *Ba*GH5-WT against CMC-Na, (B) *Ba*GH5-UV2 against CMC-Na, (C) *Ba*GH5-WT against β-glucan and (D) *Ba*GH5-UV2 against β-glucan.

TH-2393_146106034

5.3.5 3-D structure comparison of *Ba*GH5-WT and *Ba*GH5-UV2

5.3.5.1 Homology modeling and quality assessment of model of *Ba*GH5-WT

The best modeled 3-D structure of *Ba*GH5-WT was determined by RaptorX server by using PDB ID 3PZT|A (*Bacillus subtilis*), 2L8A|A (*Bacillus subtilis*), and 1E5J|A (*Salipaludibacillus agaradhaerens*) as protein templates with query coverage of 68%, 27% and 56%, respectively with sequence identity 89%, 91.1% and 69.5%, respectively (Fig. 5.3.5A). The energy minimized modeled 3-D structure of *Ba*GH5-WT consists of two domains, a N terminal classical $(\alpha/\beta)_8$-TIM barrel fold glycoside hydrolase family 5 domain and a C-terminal β-sandwich fold CBM3 domain connected by a linker (Fig. 5.3.5A). Surface view of modeled *Ba*GH5-WT structure showed that the loops connecting the $\beta_1\alpha_1$ (L1), $\beta_2\alpha_2$ (L2), $\beta_3\alpha_3$ (L5), $\beta_4\alpha_4$ (L7), $\beta_5\alpha_5$ (L9), $\beta_6\alpha_6$ (L11), $\beta_7\alpha_7$ (L13) and $\alpha_7\beta_8$ (L14) are involved in the formation of catalytic core of TIM barrel fold (confirmed from multiple sequence alignment with structural homologues) (Fig. 5.3.5B). Yan *et al.*, (2016) reported that the endoglucanases belonging to glycoside hydrolase family 5 possess the $(\alpha/\beta)_8$-TIM barrel fold architecture. The quality of energy minimized modeled *Ba*GH5-WT structure was evaluated on the tool available at saves server. The 3-D modeled structure of *Ba*GH5-WT validated by Ramachandran plot through PROCHECK showed approximately, 81.6% residues in most favoured region, 16.3% residues in additional allowed region, 1.9% (Ala60, Thr182, Asn391, Asn181, Leu156, Ala325, Thr 70 and Glu458) residues in generously allowed regions and only 0.2% (Gln 163) residue in disallowed region (Fig. 5.3.6A) out of non-glycine and non-proline residues of modeled structure. These results of Ramachandran plot displayed that the modeled 3-D structure of *Ba*GH5-WT followed the backbone phi (Φ) and psi (Ψ) dihedral angles at favourable positions. The VERIFY 3-D results of modeled *Ba*GH5-

TH-2393_146106034

WT showed 88.66% of amino acid residues possess average 3D-1D score >=0.2 and declared acceptable the modeled structure of *Ba*GH5-WT (Fig. 5.3.6B). The ProSA results of modeled *Ba*GH5-WT showed that the protein resides in the X-ray zone with Z-scores of -7.55 indicated that the modeled structure was error free (Fig. 5.3.6C and D). The above results of *Ba*GH5-WT structural refinement and quality assessment was found admissible to carry out further studies for 3-D structure modeling of *Ba*GH5-UV2, MD simulation, and Molecular docking.

TH-2393_146106034

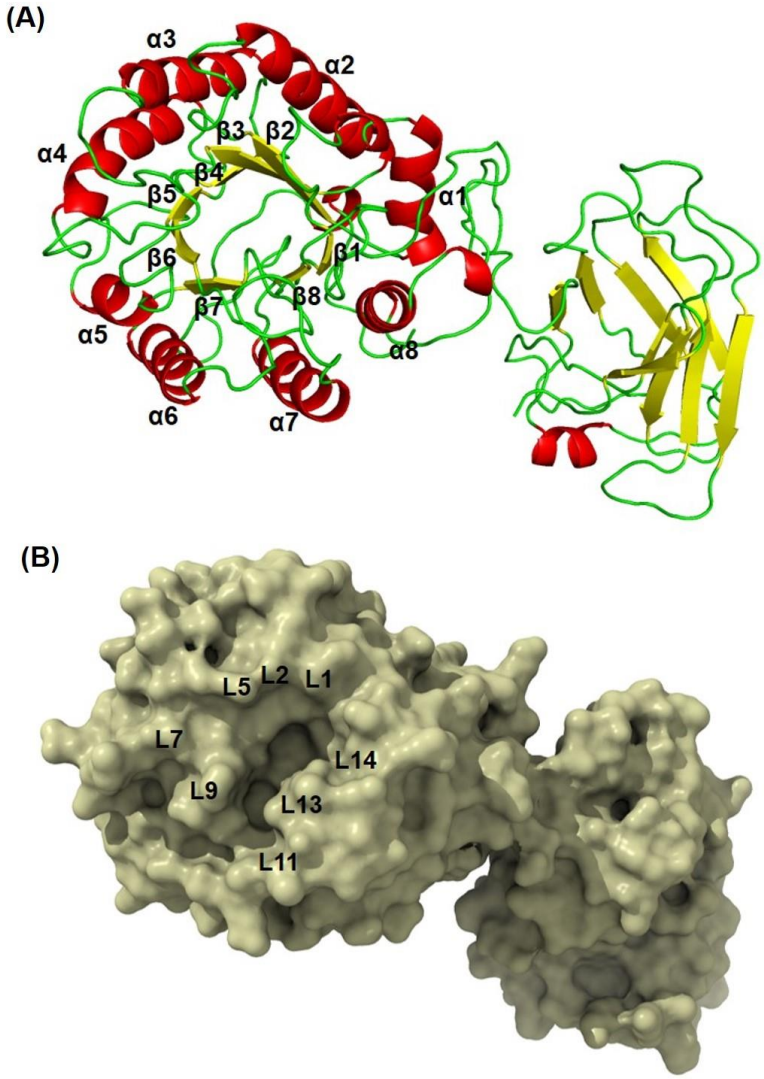


Fig. 5.3.5 3-D structure analysis of *Ba*GH5-WT; A) 3-D cartoon view structure of *Ba*GH5-WT representing TIM barrel $(\alpha/\beta)_8$ fold; B) surface view of 3-D *Ba*GH5-WT showing Loops connecting the $\beta_1\alpha_1$ (L1), $\beta_2\alpha_2$ (L2), $\beta_3\alpha_3$ (L5), $\beta_4\alpha_4$ (L7), $\beta_5\alpha_5$ (L9), $\beta_6\alpha_6$ (L11), $\beta_7\alpha_7$ (L13) and α_7 β_8 (L14) forming the catalytic core.

TH-2393_146106034

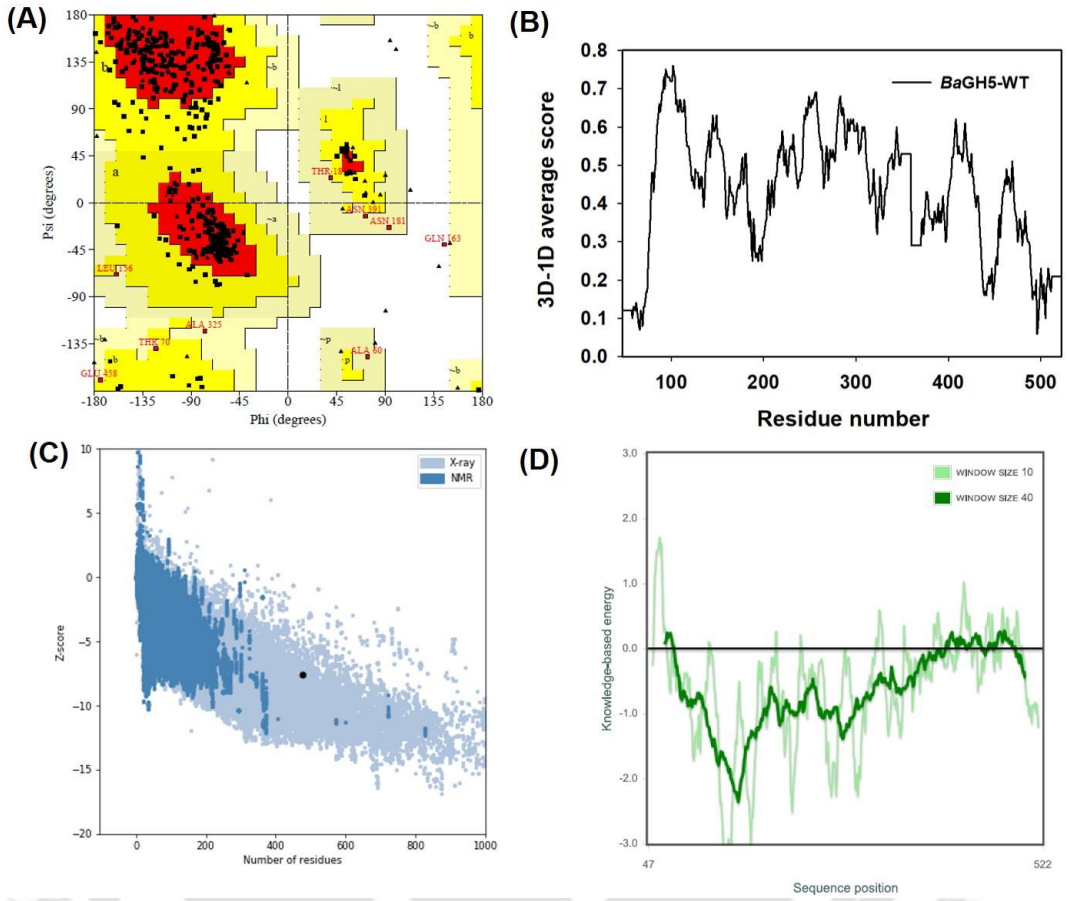


Fig. 5.3.6 Quality assessment of *Ba*GH5-WT 3-D modeled structure A) Ramachandran plot (PROCHECK) showing the amino acid present in most favoured, additional allowed, generously allowed and disallowed regions, (B) VERIFY 3-D showing 88.66% of the residues have averaged average 3D-1D score >=0.2, (C and D) ProSA plot with Z-score -7.55 for overall model quality.

TH-2393_146106034

5.3.5.2 Homology modeling and quality assessment of 3-D model *Ba*GH5-UV2 and its comparison with *Ba*GH5-WT

The 3-D structure of *Ba*GH5-UV2 was constructed after mutating the amino acid residue Asp 256 to Gly (D256G) in 3-D structure of modeled *Ba*GH5-WT protein (before energy minimization structure) by using Swiss-PDB viewer software (Fig. 5.3.7A). The energy minimized structure of *Ba*GH5-UV2 showed the position of D256G mutation was present in loop (L11), which involved in the catalytic core formation as shown in red colour (Fig 5.3.7B). Earlier studies of Lee *et al.*, (2011) also confirmed that the $\beta_6\alpha_6$ loop is involved in the formation of catalytic core of GH5 endoglucanases. Therefore, glycine residue in *Ba*GH5-UV2 might have increased the flexibility of the catalytic pocket because it is the simplest amino acid with no nonpolar side chain and can provide hydrogen bonding to support transition state in catalysis as compared with the aspartic acid which has the larger side chain and hence results in increased catalytic efficiency and pH tolerance. The loop connecting the βα and αβ regions of TIM-barrel enzymes plays vital role in the interactions between catalytic core and substrate (Varrot *et al.*, 1999, Zhai *et al.*, 2015, Tu *et al.*, 2016). In general, most key mutations which lead to the evolution of properties, such as catalytic efficiency, specific activity and optimal condition, are located on the surface of enzyme.

The quality of modeled *Ba*GH5-UV2 structure after energy minimization was evaluated on the tool available at saves server similar to the wild-type *Ba*GH5-WT. The 3-D modeled structure of *Ba*GH5-UV2 validated by Ramachandran plot through PROCHECK showed approximately, 82.0% residues in most favoured region, 15.8% residues in additional allowed region, 1.9% (Ala60, Thr182, Asn391, Asn181, Leu156, Ala325, Thr 70 and Glu458) residues in generously allowed region and 0.2% (Gln 163)

TH-2393_146106034

residues in disallowed region (Fig. 5.3.8A) out of non-glycine and non-proline residues of modeled structure. The VERIFY 3-D results of modeled *Ba*GH5-UV2 showed 89.08% of amino acid residues possesses average 3D-1D score >=0.2 and declared pass the modeled structure of *Ba*GH5-UV2 (Fig. 5.3.8B). The ProSA results of modeled *Ba*GH5-UV2 showed that the protein resides in the X-ray zone with Z-scores of -7.85 indicating that the modeled structure was error free (Fig. 5.3.8C and D). These results depicted that the modeled structure of mutant protein *Ba*GH5-UV2 with only single substitution mutation showed slight increase in the 3-D structure quality in Ramachandran plot, VERIFY 3-D and ProSA which might be responsible for increased catalytic efficiency and pH tolerance.

TH-2393_146106034

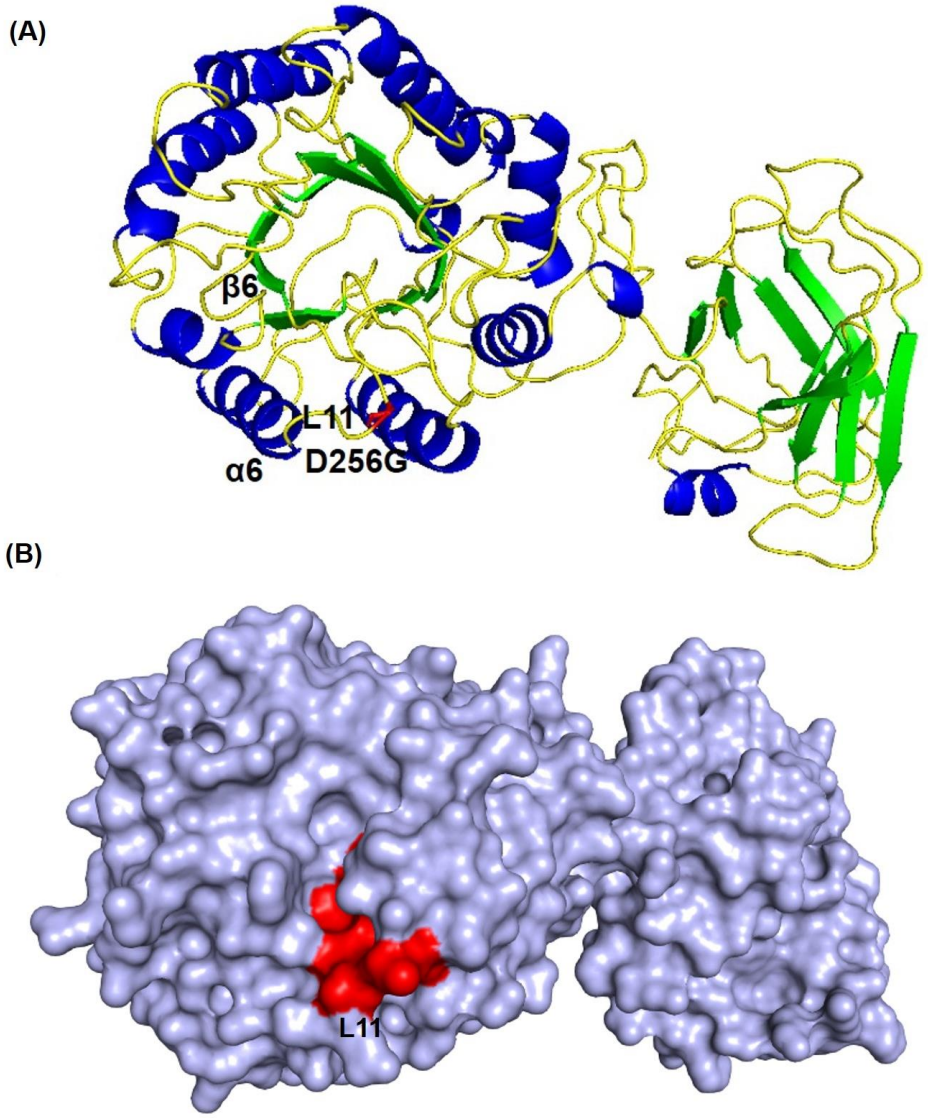


Fig. 5.3.7 3-D structure analysis of *Ba*GH5-UV2; A) 3-D cartoon view structure of *Ba*GH5-UV2 representing mutation D256G in a loop 11 connecting the β_6 to α_6; B) surface view of 3-D *Ba*GH5-UV2 showing loop in red colour having single mutation.

TH-2393_146106034

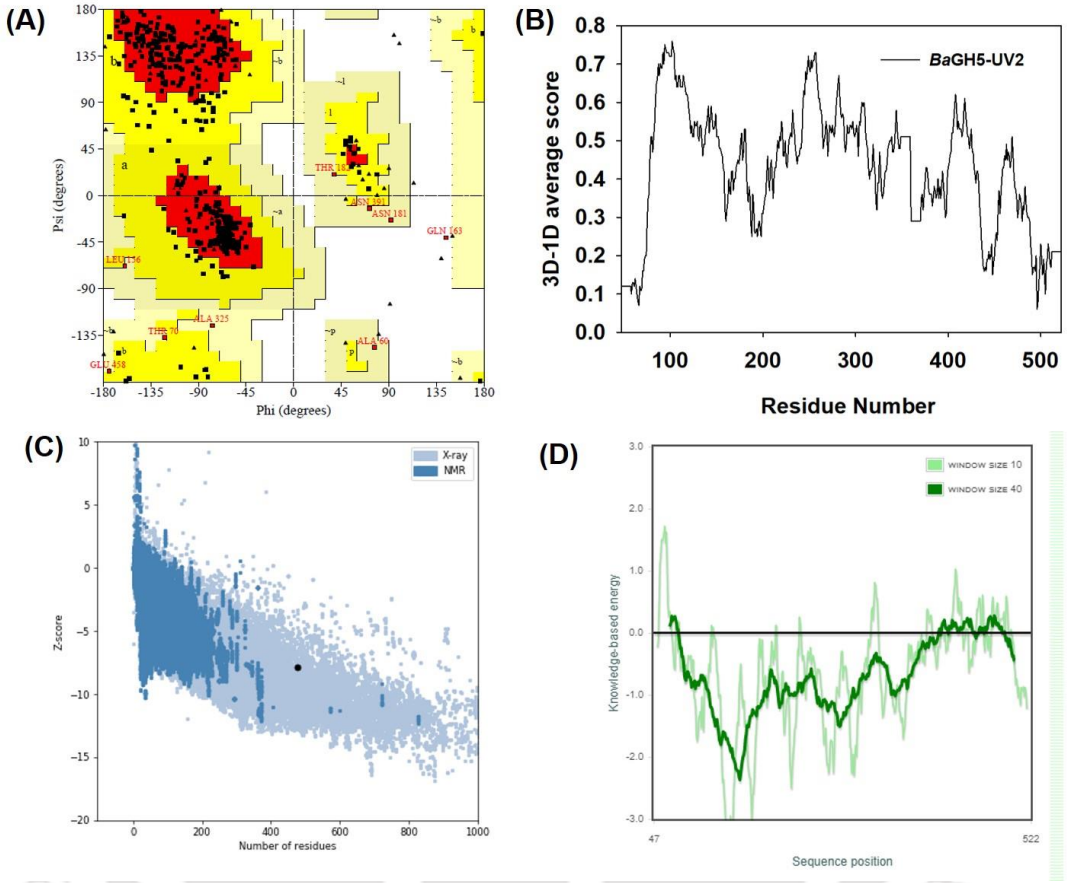


Fig. 5.3.8 Quality assessment of 3-D structure analysis of *Ba*GH5-UV2; A) Ramachandran plot (PROCHECK) showing the amino acid present in most favoured, additional allowed, generously allowed and disallowed regions, (B) VERIFY 3-D showing 89.08% of the residues have averaged average 3D-1D score >=0.2, (C and D) ProSA plot with Z-score -7.85 for overall model quality.

5.3.6 Sequence alignment of *Ba*GH5-WT and *Ba*GH5-UV2 enzyme with other family 5 glycoside hydrolase homologues

The multiple sequence alignment of N-terminal catalytic GH5 domain of *Ba*GH5-WT and *Ba*GH5-UV2 with its homologues in family 5 glycoside hydrolases endoglucanase from *Bacillus subtilis* (PDB ID 3PZV|A), *Salipaludibacillus agaradhaerens* (PDB ID 1E5J|A), *Geobacillus* sp., (PDB ID 4XZB|A), *Cytophaga hutchinsonii (*PDB ID 5IHS|A*), Caldicellulosiruptor saccharolyticus* (PDB ID

TH-2393_146106034

5ECU|A) displayed the highly conserved amino acid residues highlighted in red colour background (Fig. 5.3.9). The semi conserved amino acid residues are shown in red colour highlighted text (Fig. 5.3.9). The residues E192 and E280 are the catalytic residues depicted by black star. Further, the multiple sequence alignment of 16 homologues (sequence similarity in range of 27%-89%) of bacterial cellulases of GH5 family from *Bacillus subtilis*, *Salipaludibacillus agaradhaerens*, *Bacillus agaradherans*, *Geobacillus* sp., *Bacillus halodurans*, *Cytophaga hutchinsonii*, *Bacillus sp*. KSM-635, *Caldicellulosiruptor saccharolyticus*, Metagenome holotolerant cellulase, *Thermobifida fusca*, *Erwinia Chrysanthemi*, *Pseudoalteromonas haloplanktis* and *Hungateiclostridium thermocellum* was carried out to predict the variation in amino acid residues at mutation position 256 of loop, because the loop regions play a vital role in the interactions between catalytic core and the ligand (Fig. 5.3.10). The MSA of 16 homologues showed that the Gly 256 was the conserved residue. It was found that out of 16 homologues of family 5 glycoside hydrolases, endoglucanase, *Ba*GH5-WT, Cel5A (PDB ID 2CKS|A) from *Thermobifida fusca* and Cel5B (PDB ID 4V2X|A) from *Bacillus halodurans* showed divergence in the conserved region at 256 position, where aspartate, alanine and phenylalanine, respectively, were present instead of glycine residue (Fig. 5.3.10). Recombinant mutant enzyme, *Ba*GH5-UV2 from UV2 mutant strain (Singh *et al.*, 2019) of *Bacillus amyloliquefaciens* SS35 have substitution mutation at 256 position leading to D256G mutant due to which catalytic efficiency and pH tolerance was increased. This suggested that the mutation of a divergent residue to a conserved residue near the binding residues increases the catalytic efficiency of enzyme by increasing the stability of catalytic core. Joo *et al.*, (2011) discussed that mutating the divergent amino acid into the consensus amino acid (represented in >50% of the

TH-2393_146106034

homologs) which also fit in structure without destroying the helices or salt bridges nearly at 6 Å from the catalytically important residues improves the catalytic properties of enzymes.

The search by Dali server displayed that the RMSD of *Ba*GH5-WT or *Ba*GH5-UV2 after superposition with GH5 endoglucanase PDB ID 3PZT|A (*Bacillus subtilis*) was 3.1 Å with 298 C_α atoms, with endoglucanase PDB ID 6GJF|A was 3.2 Å with 303 C_α atoms, with endoglucanase PDB ID 1E5J|A (*Salipaludibacillus agaradhaerens*) was 3.2 Å with 302 C_α atoms and with PDB ID 5IHS|A (*Cytophaga hutchinsonii*) was 2.9 Å with 306 C_α atoms showing an excellent alignment. The amino acid residues taking part in the catalysis and substrate binding were identified by superimposing the *Ba*GH5-WT (Red colour) and *Ba*GH5-UV2 (bluewhite colour) modeled structure with PDB structure of endoglucanase 3PZT|A (blue colour) from *Bacillus subtilis*, 1E5J|A (yellow colour) from *Salipaludibacillus agaradhaerens* and 5IHS|A (palegreen) from *Cytophaga hutchinsonii* (Fig. 5.3.11A). The R115, E192, W230, H252, Y254, T279, E280 and W314 are the catalytically important amino acid residues for *Ba*GH5-WT. The superimposition analysis displayed that the E192 acts as a catalytic acid/proton donor and E280 of *Ba*GH5-WT/ *Ba*GH5-UV2 as a catalytic nucleophile located at the crevice formed by the C-terminal ends of *β*-strands (Fig. 5.3.11B). The aromatic amino acid residues W230 and W314 act as the gate keepers and H252, R115 and T279 act as polarizers during the substrate catalysis according to Santos *et al.*, (2012). Moreover, Y254 of loop $\beta_6\alpha_6$ orient and activate the E280 (equal to Glu289, as nucleophile Shirai *et al.*, 2001) because with loop $\beta_8\alpha_8$, it imparts flexibility to the catalytic cavity which causes an induce fit at the time of substrate binding. The substitution of glycine amino acid near the binding site residues H252 and Y254 (Fig. 5.3.11C) increases the

TH-2393_146106034

flexibility/conformation plasticity of loop 11 (L11), that might help in the enhancement of catalytic core flexibility, which resulted in the improvement of the induce fitting of ligand at the cavity. The glycine amino acid residue lacks β-carbon, therefore its substitution in protein backbone increases the conformational flexibility which allows a greater configurational entropy compared to other amino acids and help in increasing the catalytic efficiency of enzymes (Ma *et al.,* 2000 and Liang *et al.,* 2011). The estimated distance between two catalytic residues E192 and E280 of *Ba*GH5-WT and *Ba*GH5-UV2 using PyMOL 2.0 was found to be similar i.e 4.8 Å (Fig. 5.3.11B), this indicated that the *Ba*GH5-WT/*Ba*GH5-UV2 like other glycoside family 5 cellulases (Cel5A), showed retaining type hydrolytic mechanism which cleave β-1,4- glycosidic bond via a double displacement mechanism (Lee *et al.,* 2011).

TH-2393_146106034

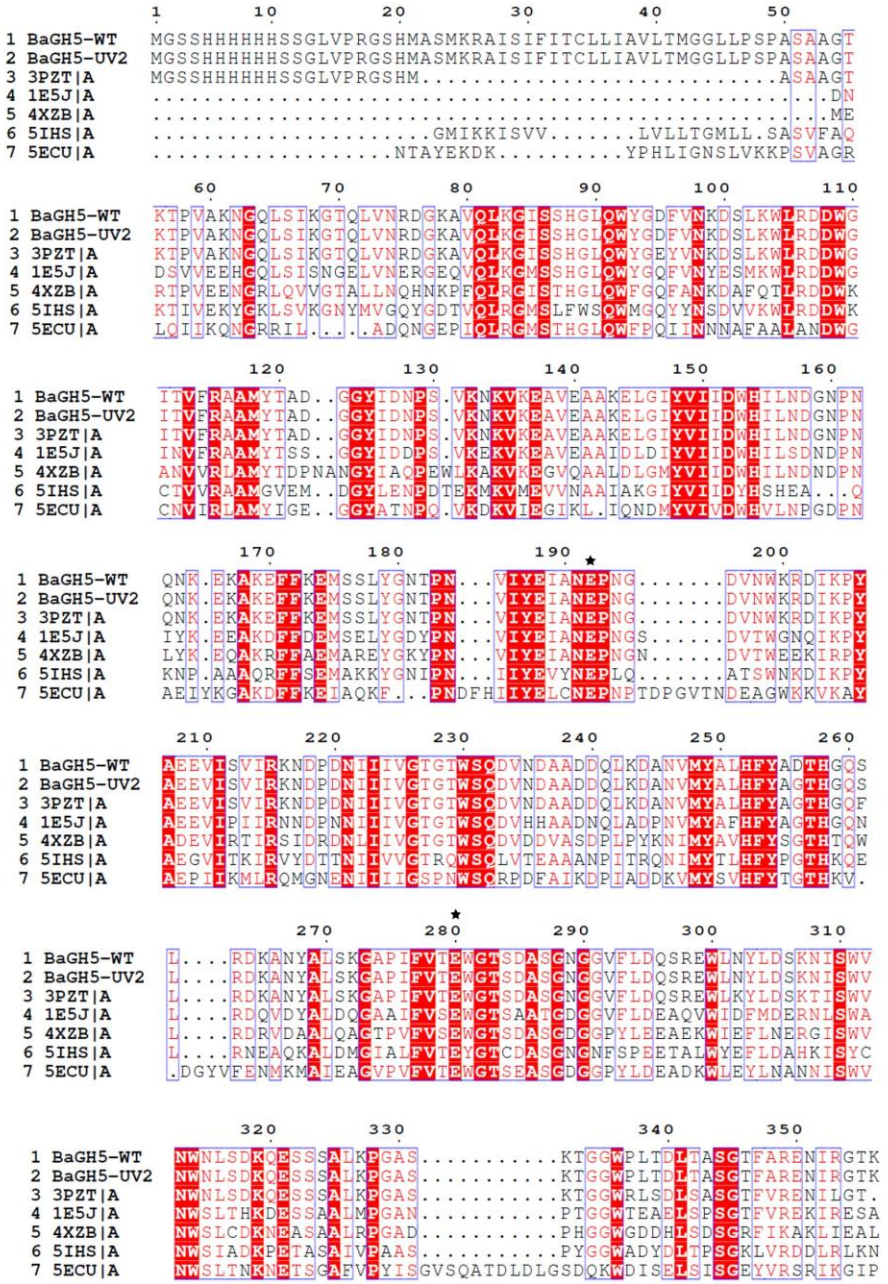


Fig. 5.3.9 Multiple Sequence alignment (MSA) of N-terminal catalytic GH5 domain of *Ba*GH5-WT and *Ba*GH5-UV2 displaying the conserved residues in red background, semi-conserved in red colour letters and dispersive amino acids black colour letter.

TH-2393_146106034

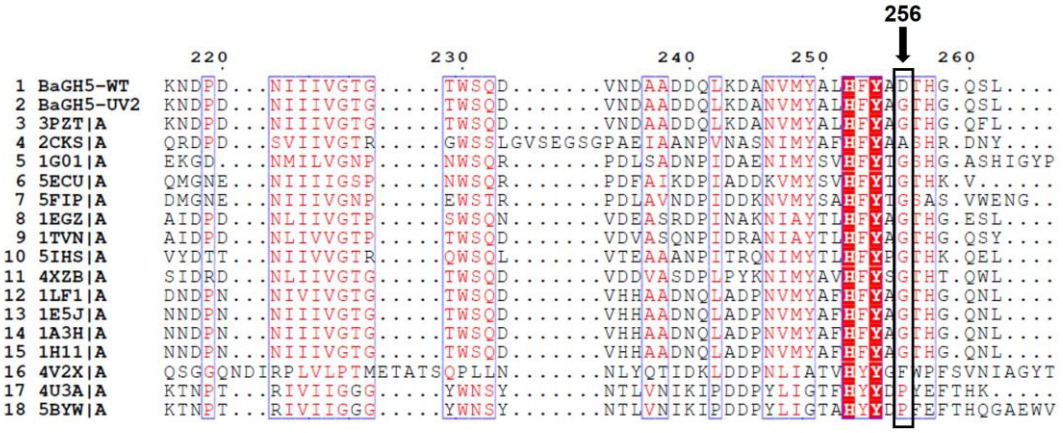


Fig. 5.3.10 Multiple sequence alignments of 16 homologues of *Ba*GH5-WT bacterial endoglucanases of GH5 family. *Ba*GH5-WT, PDB:ID 2CKS|A and PDB:ID 4V2X|A showing divergence at position 256.

TH-2393_146106034

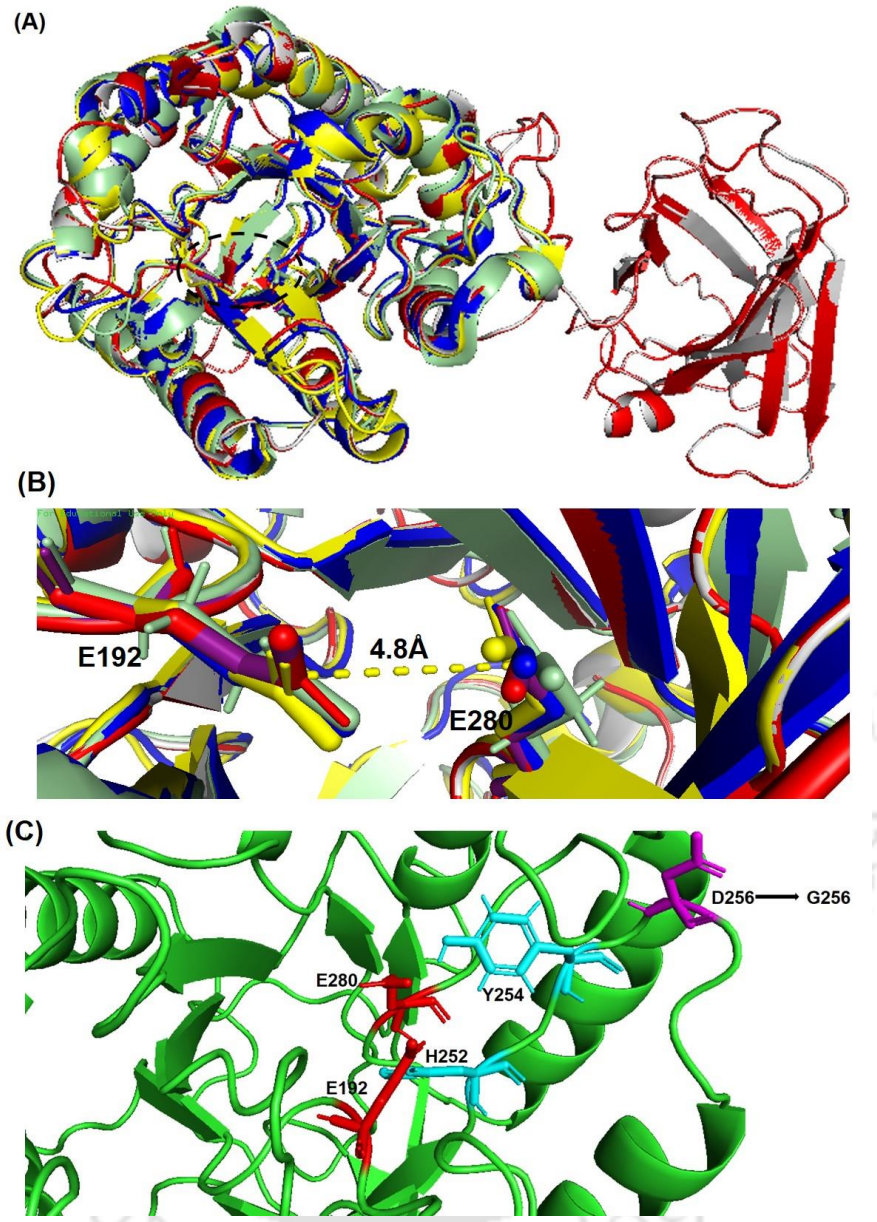


Fig. 5.3.11 A) Superimposition of *Ba*GH5-WT modeled structure (Red colour) and *Ba*GH5-UV2 modeled structure (bluewhite colour) with crystal structures of GH5 endoglucanase of PDB ID 3PZT|A (blue colour), 1E5J|A (yellow colour) and 5IHS|A (palegreen) showed common structure and catalytic residues (in circle), (B) The profile of distance 4.8 Å (from C_α to C_α) between the catalytic acid (E192) and catalytic base (E280) of *Ba*GH5-WT (red colour) by molecular visualization system PyMOL 2 showing hydrolytic mechanism (retaining) and C) Modeled *Ba*GH5-WT structure showing two catalytically important residues H252 and Y254 on loop having substitution mutation of D256G.

TH-2393_146106034

5.3.7 Molecular dynamics simulation of modeled *Ba*GH5-WT and *Ba*GH5-UV2 structure

Molecular Dynamics (MD) simulation studies of modeled *Ba*GH5-WT and *Ba*GH5-UV2 structure were carried out for 80 ns. MD simulation of modeled protein structure over the defined trajectory leads to the more energetically stable conformations of a protein in aqueous solution which can be represented by the RMSD between the starting structure and those obtained during the simulation (Pandey *et al.*, 2019). The average backbone RMSD for wild-type *Ba*GH5-WT reaches a plateau phase value of 0.76 nm at 50 ns and remained stable till 80 ns (Fig. 5.3.12A) whereas, the RMSD profile for mutant *Ba*GH5-UV2 reaches a plateau phase value of 0.91 nm after 69 ns and remained stable till 80 ns (Fig. 5.3.12A). These results indicated that the structure of mutant *Ba*GH5-UV2 undergoes more conformational changes along the simulation period of 69 ns in order to get the final stable conformation. Stable RMSD at the end of the simulation period of 80 ns suggested the MD simulation for *Ba*GH5-WT and *Ba*GH5-UV2 is perfect for further analysis. The overall dimension and global compactness of protein *Ba*GH5-WT and *Ba*GH5-UV2 were determined by using gmx gyrate program during the MD simulation. Rg values showed the compactness of structure and overall dimension of protein. It predicts how regular secondary structures are compactly packed into 3D structure. Therefore, for stably folded protein, relatively steady Rg value is maintained whereas, for not properly folded protein it will change over time (Rogerson 2011). The fluctuation in the R_g value of *Ba*GH5-WT protein was 2.77-3.03 nm in the time scale of 0-54 ns and after that it decreased to 2.84 nm at 69 ns and then remained till 80 ns (Fig. 5.3.12B). However, the fluctuation in R_g value of *Ba*GH5-UV2 protein was 2.72-2.84 nm in the time scale of 0-7.4 ns and after that it

TH-2393_146106034

decreased to 2.55 nm at 60 ns than it became stable till the end of the simulation of 80 ns (Fig. 5.3.12B). These results of R_g values for *Ba*GH5-WT and *Ba*GH5-UV2 proteins depicts the stable conformation after 69 ns and 60 ns, respectively. Moreover, the lower R_g value for *Ba*GH5-UV2 proteins (2.55 nm) as compared with *Ba*GH5-WT protein (2.84 nm) suggested that the compactness of 3-D structure of *Ba*GH5-UV2 protein increased which increased the globularity of protein.

RMSF during MD simulation was calculated by using gmx rmsf program. It measures the displacement of a particular atom or group of atoms with respect to the reference structure. The amino acid residues showing higher RMSF values have highest flexibility for α-carbon among all the amino residues in a polypeptide chain. However, amino acid residues showing lower RMSF values showed lowest flexibility for α-carbon. RMSF values for *Ba*GH5-WT catalytic residues, E192 and E280 were 0.243 nm and 0.164 nm, respectively whereas, for *Ba*GH5-UV2 catalytic residues, E192 and E280 were 0.165 nm and 0.101 nm, respectively (Fig. 5.3.12C). This decrease in RMSF values for catalytic residues in *Ba*GH5-UV2 enzyme showed more rigid catalytic residues. RMSF values for *Ba*GH5-UV2 N-terminal linker connecting the catalytic GH5 domain to CBM3 domain of residues 354-359 (0.449-0.77 nm) was higher than the linker of *Ba*GH5-WT (0.372-0.506 nm) depicting the higher flexibility for the mutant enzyme linker. According to the previous study of Pandey *et al.*, 2019, the catalytic residues should be less flexible and linker should be more flexible in order to maintain the activity of the enzyme. SASA for *Ba*GH5-WT showed higher fluctuations upto the maximum value of 250 nm^2 till 24.5 ns after that it fairly becomes constant with little fluctuation with average value of 242.7 nm^2 between 24.5 ns to 80 ns (Fig. 5.3.12D). However, the SASA of *Ba*GH5-UV2 for time period of simulation

TH-2393_146106034

upto 0-0.92 ns showed maximum fluctuation between 243.28 nm^2 to 257.57 nm^2 after that it decreases to 225 nm^2 till 30 ns and further remained fairly stable till 80 ns with a average value of 225.7 nm^2 (Fig. 5.3.12D). This suggested that the overall exposure of the *Ba*GH5-UV2 towards the solvent area decreases to 7% as compared to the *Ba*GH5-WT which depicted that the accessibility of the solvent in the catalytic sites was lower as compared with *Ba*GH5-WT. According to the earlier reports of Yaacob *et al.,* (2019) for protein stability a lesser protein-solvent contact area is favourable because for larger surface area protein is interacting more with the solvent. This results in water penetration into the interior of a protein, where it forms the hydrogen bonds to the side chains residues and main chain of polypeptide because of which hydrogen bond interactions among the other residues decreases and weaken the internal protein interactions (weaken the internal protein interactions (Balasubramaniam *et al.,* 2019). The superimposed structure of 0 ns and 80 ns MD simulated structure of *Ba*GH5-WT (Fig. 5.3.12E) and *Ba*GH5-UV2 (Fig. 5.3.12F), respectively, depicted the lesser fluctuations over simulation time periods suggesting that both the structure can be further used in ligand binding analysis by Docking.

TH-2393_146106034

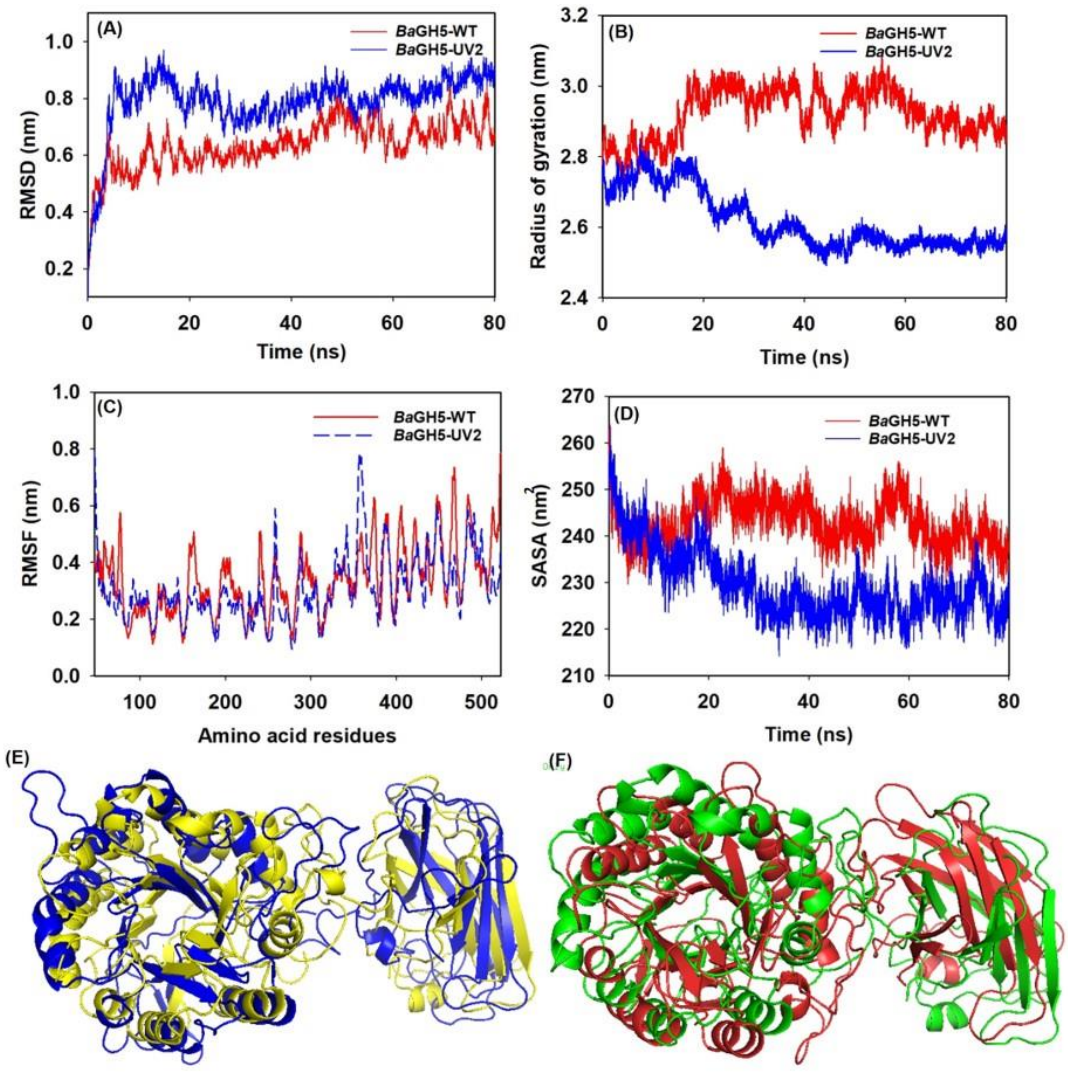


Fig. 5.3.12 Molecular dynamics simulation of *Ba*GH5-WT and *Ba*GH5-UV2 A) RMSD plot; B) Radius of gyration plot; C) RMSF plot; D) SASA plot; E) Superposition of simulated structure of *Ba*GH5-WT (blue) with modeled *Ba*GH5-WT structure (yellow); F) Superposition of simulated structure of *Ba*GH5-UV2 (red) with modeled *Ba*GH5-UV2 structure (green).

TH-2393_146106034

5.3.8 Comparison of ligand binding analysis of *Ba*GH5-WT and *Ba*GH5-UV2

The interactions between the ligands (cellooligosaccharides) and *Ba*GH5-WT or *Ba*GH5-UV2 enzymes were studied by molecular docking analysis of 80 ns MD simulated structures. The enzyme, *Ba*GH5-WT showed binding with cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose with free energy of binding (ΔG°) of -4.28, -3.07, -1.34, -0.25 and -0.33 kcal/mol, respectively, as shown in Table 5.3.3. The enzyme, *Ba*GH5-UV2 showed binding with cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose with free energy of binding (ΔG°) of -4.82, -5.02, -3.53, -0.56 and 1.46 kcal/mol, respectively, as shown in Table 5.3.3. *Ba*GH5-UV2 showed 1.13, 1.63, 2.63 and 2.54-fold decrease in ΔG° against cellobiose, cellotriose, cellotetraose and cellopentaose, respectively, as compared with *Ba*GH5-WT. This showed that the subsitution of Asp256 to Gly residue in *Ba*GH5-UV2 results in maximum decrease in the free energy (ΔG°) against cellotetraose among all cello-oligosaccharides as compared with *Ba*GH5-WT. This depicted that *Ba*GH5-UV2 enzyme can catalyze cellotetraose more efficiently which is responsible for increased efficiency. The amino acid residues of *Ba*GH5-WT and *Ba*GH5-UV2 involved in non-covalent interactions with different ligands (cello-oligosaccharides) are listed in Table 5.3.3. *Ba*GH5-UV2 showed increased interactions with more number of amino acid residues with different hydrogen-bonding network near the active site as compared with the *Ba*GH5-WT (Table 5.3.3). The docking analysis of cellotetraose *Ba*GH5-WT complex (Fig. 5.3.13A and B) and *Ba*GH5-UV2 complex (Fig. 5.3.14A and B) revealed that, *Ba*GH5-UV2 forms more number of hydrogen bonds as compared to *Ba*GH5-WT. This have happened because of the increased flexibility of loop 11 ($\beta_6\alpha_6$) due to the glycine substitution which owing to the altered and enhanced hydrogen bond

TH-2393_146106034

interactions pattern near the active site, might be allowing the ligand to bind more easily and thus catalyzing its hydrolysis more efficiently. According to an earlier report of Wang *et al.*, (2016) hydrogen bonds play a crucial role in substrate recognition and binding. Therefore, increase in number of hydrogen bond formation with its ligand is responsible for the enhancement of kinetic properties for enzymes. Single mutant variants, Asn 233 to Ala and Asn 233 to Gly of *Gt*Cel5 (endoglucanase) from *Gloeophyllum trabeum* at loop 6 formed direct hydrogen bonds with cellotetraose as compared with the wild-type enzyme and showed increased catalytic efficiency (Zheng *et al.*, 2018).

Table 5.3.3 Ligand-binding analysis of simulated *Ba*GH5-WT and *Ba*GH5-UV2 structures by molecular docking

	***Ba*GH5-WT**			***Ba*GH5-UV2**		
Ligand	**Binding energy, ΔG° (kcal/mol)**	**Residues making polar interaction**	**Residues making hydrophobic interaction**	**Binding energy, ΔG° (kcal/mol)**	**Residues making polar interaction**	**Residues making hydrophobic interaction**
Cellobiose	-4.28	E280, D256, Q232, Y254	W230, S261, Q260, F253	-4.82	S324, S323, S284, S322, E321, H88, K319, Q320	W314, A286, D285
Cellotriose	-3.07	Y254, D256, Q232, K265, D264	W230, S261, G259, Q260, V234	-5.02	E192, S287, A286, W314, Y254, D285, H252, N191, E280	W230, H258
Cellotetraose	-1.34	H88, D152, N191, R115, P193, E192, N194	W314, A117	-3.53	N315, E321, S324, S323, S322, H88, D152, H154, L156, D158, N194, E192	S284, A286, D285, Y254, W314, G159
Cellopentaose	-0.25	D256, P193, E192, N191, Y93, G89, H88	Y254, A286, W314, W230, A117, W92	-0.56	N194, I155, H154, D152, N191, H88, K319	D158, L156, G159, A286, E192, W314, Q320, E321
Cellohexaose	-0.33	Y254, Q232, W230, E280, H252, E192	S261, A255, Q260, G259, H258, P193, N191	1.46	S284, S322, W314, E192, N194, D158, N160	G282, Y254, E321, A286, L156, G159, G195

TH-2393_146106034

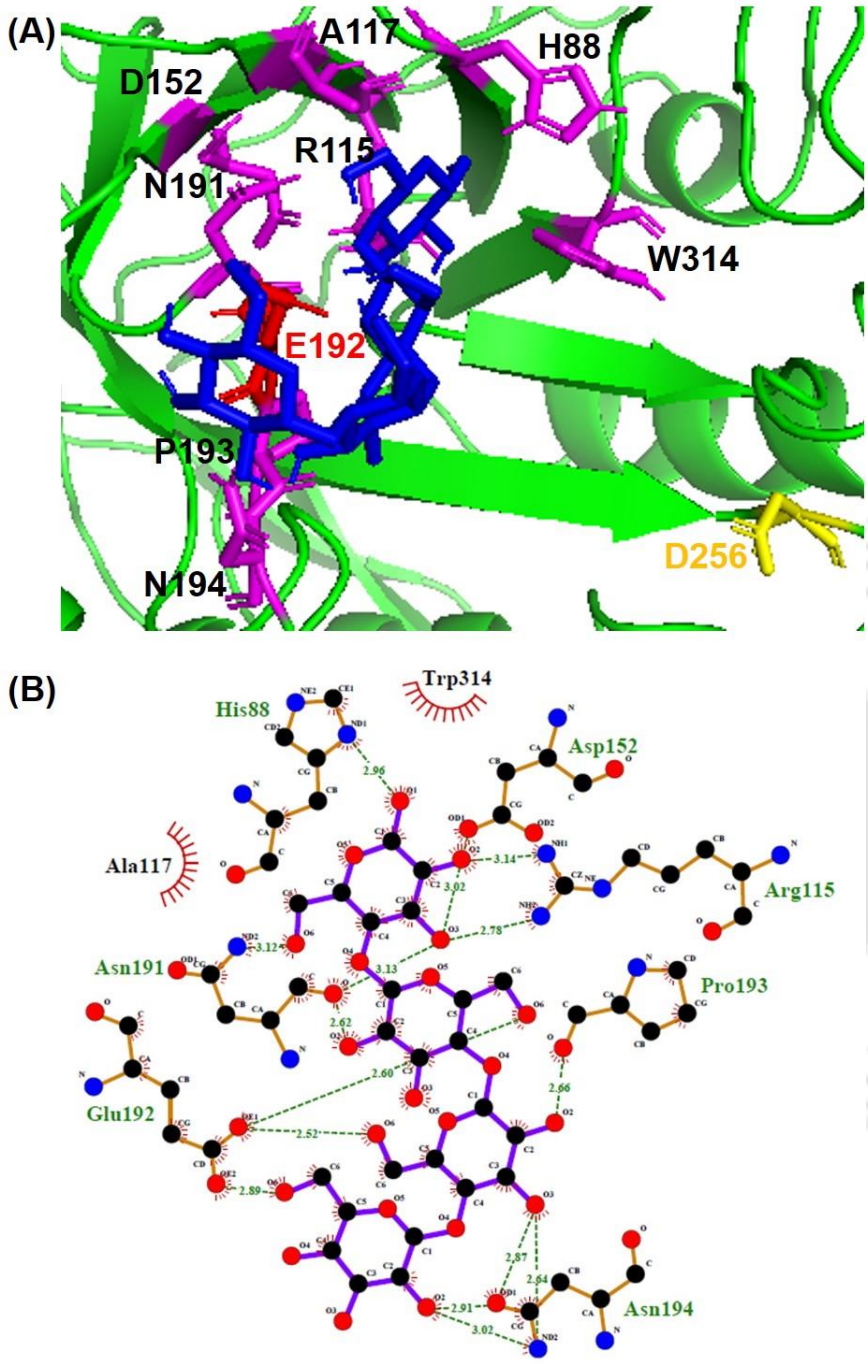


Fig. 5.3.13 Molecular docking analysis of *Ba*GH5-WT. A) the catalytic core of *Ba*GH5-WT showing interaction with the ligand cellotetraose; B) two-dimensional schematic presentation of cellotetraose with the active site residues of *Ba*GH5-WT.

TH-2393_146106034

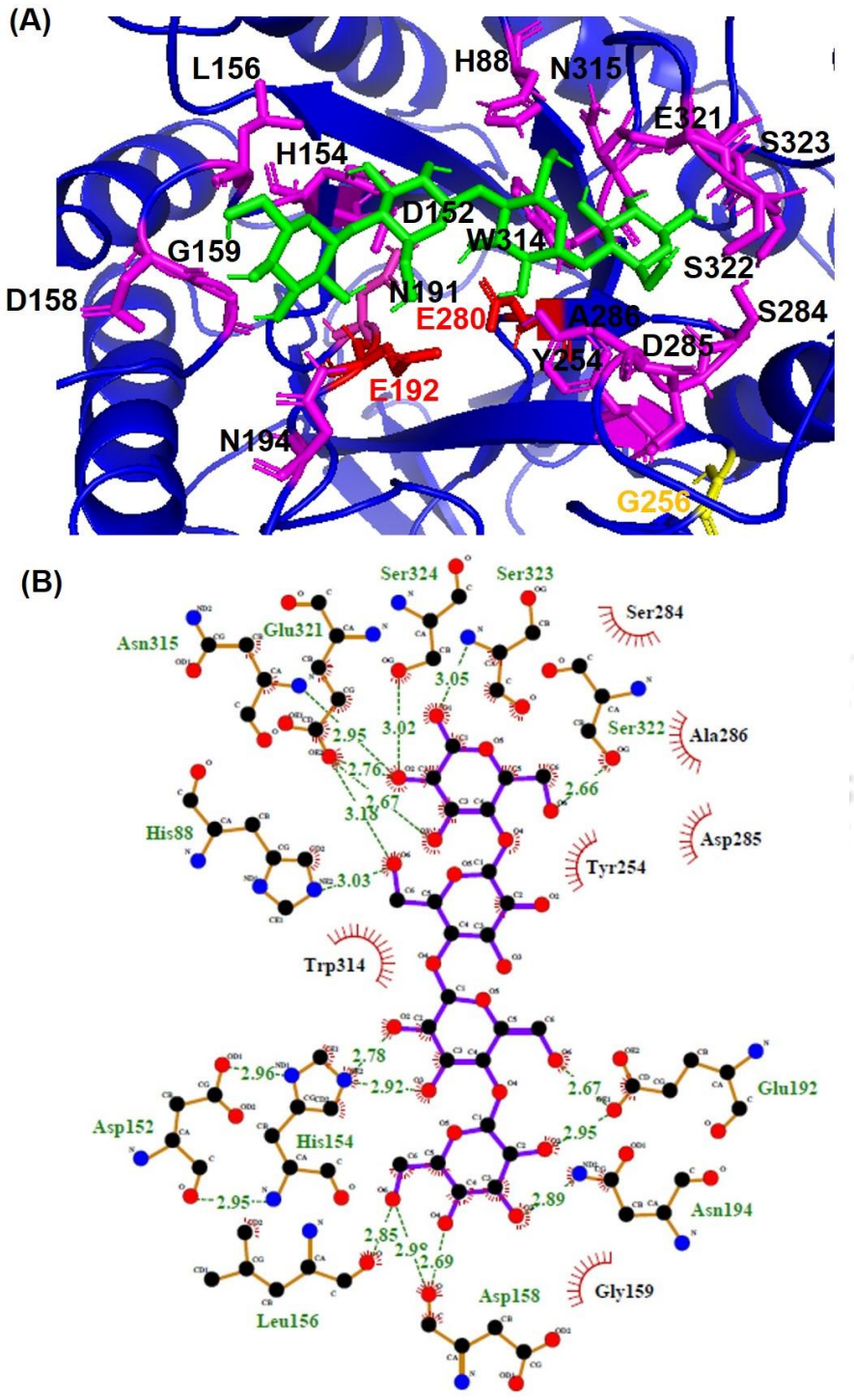


Fig. 5.3.14 Molecular docking analysis of *Ba*GH5-UV2. A) the catalytic core of *Ba*GH5-UV2 showing interaction with the ligand cellotetraose; B) two-dimensional schematic presentation of cellotetraose with the active site residues of *Ba*GH5-UV2.

TH-2393_146106034

5.4 Conclusion

The comparative biochemical characterization of recombinant enzymes *Ba*GH5-WT and *Ba*GH5-UV2 was carried out. The mutant endoglucanase *Ba*GH5-UV2 showed 9.7-fold higher in specific activity than the wild-type *Ba*GH5-WT (4.5 U/mg) against CMC-Na. Mutant *Ba*GH5-UV2 showed maximum specific activity at 65°C, 10°C higher than the wild-type *Ba*GH5-WT which showed at 55°C. Both the enzymes showed similar temperature stability up to 45°C and pH optima at 5.5. The comparison of pH stability analysis for endoglucanase activity showed the stability in the wide acidic pH range (5.0-7.0) for *Ba*GH5-UV2 as compared with the narrow basic pH range (7.0-7.5) for *Ba*GH5-WT. The catalytic efficiency of *Ba*GH5-UV2 enhanced by approximately, 18.4 and 21.9-fold against β-Glucan and CMC-Na respectively, as compared with *Ba*GH5-WT. The tertiary structure analysis of *Ba*GH5-WT was carried out by RaptorX prediction tools that showed close homology with 3PZT|A from *Bacillus subtilis,* 2L8A|A from *Bacillus subtilis* and 1E5J|A from *Salipaludibacillus agaradhaerens.* The modeled structure of *Ba*GH5-WT validated by Ramachandran plot showed approximately, 99.8% residues in the allowed region. The tertiary structure of *Ba*GH5-UV2 enzyme was generated after mutating the Asp 256 to Gly by using Swiss PDB viewer software. 3-D structure validation of *Ba*GH5-UV2 enzyme by Ramachandran plot also showed approximately, 99.9% residues in the allowed region. The overall modular structures of *Ba*GH5-WT and *Ba*GH5-UV2 showed a classical $(\alpha/\beta)_8$-TIM barrel fold. The modeled stcuture showed that the mutation occurred in the loop region connecting α_6-helix with β_6-sheet, near the active site cavity at the surface of the domain. Molecular Dynamics (MD) simulation studies for 80 ns showed the decreased radius of gyration for *Ba*GH5-UV2 enzyme,

TH-2393_146106034

suggesting the increase in the compactness of the 3-D structure thus the increased globularity of the protein. Root mean square fluctuation for the catalytic residues in *Ba*GH5-UV2 enzyme was decreased showing more rigid catalytic residues which is a requirement for enzyme catalysis. The less fluctuation observed in the SASA of *Ba*GH5-UV2 suggested that the accessibility of the substrate in the catalytic sites was higher as compared with *Ba*GH5-WT. The molecular docking studies of MD simulated structure showed that the subsitution of Asp256 to Gly residue in *Ba*GH5-UV2 showed 1.6-fold and 2.6-fold decrease in free energy (ΔG°) against cellotriose and cellotetraose, respectively, as compared with *Ba*GH5-WT. These results provide information on the possibility of designing a mutant enzymes of members from glycoside hydrolase family 5 for improving their catalytic efficiencies and pH stability by protein engineering.

TH-2393_146106034

5.5 References

Badieyan, S., Bevan, D.R., and Zhang, C. (2012). A salt-bridge controlled by ligand binding modulates the hydrolysis reaction in a GH5 endoglucanase. *Protein Engineering, Design and Selection*, 25(5), 223-233.

Badieyan, S., Bevan, D.R., and Zhang, C. (2012). Study and design of stability in GH5 cellulases. *Biotechnology and Bioengineering*, 109(1), 31-44.

Balasubramaniam, K., Sharma, K., and Goyal, A. (2019). Structure and dynamics analysis of a new member heparinase II/III of family 12 polysaccharide lyase from *Pseudopedobacter saltans* by computational modeling and small angle x-ray scattering. *Journal of Biomolecular Structure and Dynamics*, ISSN: 0739-1102, 1538-0254.

Ballester, P.J., and Mitchell, J.B. (2010). A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. *Bioinformatics*, 26(9), 1169-1175.

Colovos, C., and Yeates, T.O. (1993). Verification of protein structures: patterns of non-bonded atomic interactions. *Protein Science*, 2(9), 1511-1519.

Eisenberg, D., Lüthy, R., and Bowie, J.U. (1997). [20] VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods In Enzymology*, 277, 396-404.

Goodsell, D.S., and Olson, A.J. (1990). Automated docking of substrates to proteins by simulated annealing. *Proteins: Structure, Function, and Bioinformatics*, 8(3), 195-202.

TH-2393_146106034

Huey, R., Morris, G.M., Olson, A.J., and Goodsell, D.S. (2007). A semiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry*, 28(6), 1145-1152.

Joo, J.C., Pack, S.P., Kim, Y.H., and Yoo, Y.J. (2011). Thermostabilization of *Bacillus circulans* xylanase: computational optimization of unstable residues based on thermal fluctuation analysis. *Journal of Biotechnology*, 151(1), 56-65.

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., and Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. *Nature Protocols*, 7(8), 1511.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283-291.

Lee, T.M., Farrow, M.F., Arnold, F.H., and Mayo, S.L. (2011). A structural study of *Hypocrea jecorina* Cel5A. *Protein Science*, 20(11), 1935-1940.

Li, Y. X., Yi, P., Yan, Q. J., Qin, Z., Liu, X. Q., and Jiang, Z. Q. (2017). Directed evolution of a β-mannanase from *Rhizomucor miehei* to improve catalytic activity in acidic and thermophilic conditions. *Biotechnology for Biofuels*, 10(1), 143.

Liang, C., Fioroni, M., Rodríguez-Ropero, F., Xue, Y., Schwaneberg, U., and Ma, Y. (2011). Directed evolution of a thermophilic endoglucanase (Cel5A) into highly active Cel5A variants with an expanded temperature profile. *Journal of Biotechnology*, 154(1), 46-53.

Ma, Y.F., Eglinton, J.K., Evans, D.E., Logue, S.J., and Langridge, P. (2000). Removal of the four C-terminal glycine-rich repeats enhances the thermostability and

TH-2393_146106034

substrate binding affinity of barley β-amylase. *Biochemistry*, 39(44), 13350-13355.

Michaelis, L., and Menten, M.L. (1913). Die kinetic der invertinwirkung. *Biochemical Journal*, 49, 333-369.

Nelson, N. (1944). A photometric adaptation of the Somogyi method for the determination of glucose. *The Journal of Biological Chemistry*, 153(2), 375-380.

Pandey, B., Grover, S., Kaur, J., and Grover, A. (2019). Analysis of mutations leading to para-aminosalicylic acid resistance in *Mycobacterium tuberculosis*. *Scientific Reports*, 9(1), 1-15.

Raabo, B.E., and Terkildsen, T.C. (1960). On the enzymatic determination of blood glucose. *Scandinavian Journal of Clinical and Laboratory Investigation*, 12(4), 402-407.

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(W1), W320-W324.

Rogerson, P., and Arteca, G.A. (2011). Molecular size scaling in families of protein native folds. *Journal of Mathematical Chemistry*, 49(8), 1493.

Ruiz, D. M., Turowski, V. R., and Murakami, M. T. (2016). Effects of the linker region on the structure and function of modular GH5 cellulases. *Scientific Reports*, 6, 28504.

Santos, C.R., Paiva, J.H., Sforça, M.L., Neves, J.L., Navarro, R.Z., Cota, J., Akao, P.K., Hoffmam, Z.B., Meza, A.N., Smetana, J.H., Nogueira, M.L., Polikarpov, I., Xavier-Neto, J., Squina, F.M., Ward, R.J., Ruller, R., Zeri, A.C., and Murakami, M. (2012). Dissecting structure-function-stability relationships of a thermostable

TH-2393_146106034

GH5-CBM3 cellulase from *Bacillus subtilis* 168. *Biochemical Journal*, 441(1), 95-104.

Schrodinger, L. (2017). The PyMOL molecular graphics system, Version 2.0. Retrieved from http://www.pymol.org/.

Schülein, M. (2000). Protein engineering of cellulases. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, 1543(2), 239-252.

Shah, R.K., Patel, A.K., Davla, D.M., Parikh, I.K., Subramanian, R.B., Patel, K.C., and Joshi, C.G. (2017). Molecular cloning, heterologous expression, and functional characterization of a cellulolytic enzyme (Cel PRII) from buffalo rumen metagenome. *3 Biotech*, 7(4), 257.

Shirai, T., Ishida, H., Noda, J.I., Yamane, T., Ozaki, K., Hakamada, Y., and Ito, S. (2001). Crystal structure of alkaline cellulase K: insight into the alkaline adaptation of an industrial enzyme. *Journal of Molecular Biology*, 310(5), 1079-1087.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., and Thompson, J.D. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1), 539.

Singh, S., Dhillon, A., and Goyal, A. (2019). Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis. *Renewable Energy*. https://doi.org/10.1016/j.renene.2019.11.105.

Singh, S., Dikshit, P.K., Moholkar, V.S., and Goyal, A. (2015). Purification and characterization of acidic cellulase from *Bacillus amyloliquefaciens* SS35 for

TH-2393_146106034

hydrolyzing *Parthenium hysterophorus* biomass. *Environmental Progress and Sustainable Energy*, 34(3), 810-818.

Somogyi, M. (1945). A new reagent for the determination of sugars. The *Journal of Biological Chemistry*, 160, 61-68.

Tran, F., Billeter, S., Eising, A.A., Hünenberger, P.H., Kruger, P., Mark, E., Scott, R., and Tironi, I.G. (1996). Biomolecular Simulation: The GROMOS96 manual and user guide. Zurich, Switzerland: Verlag der Fachvereine Hochschulverlag AG an der ETH Zurich

Tu, T., Pan, X., Meng, K., Luo, H., Ma, R., Wang, Y., Yao, B. (2016). Substitution of a non-active-site residue located on the T3 loop increased the catalytic efficiency of endo-polygalacturonases. *Process Biochemistry*, 51, 1230-8.

Varrot, A., Schülein, M., and Davies, G.J. (1999). Structural changes of the active site tunnel of *Humicola insolens* cellobiohydrolase, Cel6A, upon oligosaccharide binding. *Biochemistry*, *38*(28), 8884-8891.

Wang, X., Huang, H., Xie, X., Ma, R., Bai, Y., Zheng, F., and Luo, H. (2016). Improvement of the catalytic performance of a hyperthermostable GH10 xylanase from *Talaromyces leycettanus* JCM12802. *Bioresource Technology*, 222, 277-284.

Wang, Y., Wang, X., Tang, R., Yu, S., Zheng, B., and Feng, Y. (2010). A novel thermostable cellulase from *Fervidobacterium nodosum*. *Journal of Molecular Catalysis B: Enzymatic*, 66(3-4), 294-301.

Wiederstein, M., and Sippl, M.J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 35(2), W407-W410.

TH-2393_146106034

Wierenga, R.K. (2001). The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS letters*, 492(3), 193-198.

Yaacob, N., Kamarudin, N.H.A., Leow, A.T.C., Salleh, A.B., Rahman, R.N.Z.R.A., and Ali, M.S.M. (2019). Effects of Lid 1 mutagenesis on lid displacement, catalytic performances and thermostability of cold-active *Pseudomonas* AMS8 lipase in toluene. *Computational and Structural Biotechnology Journal*, 17, 215-228.

Yan, J., Liu, W., Li, Y., Lai, H.L., Zheng, Y., Huang, J.W., Chen, C.C., Chen, Y., Jin, J., Li, H., and Guo, R.T. (2016). Functional and structural analysis of *Pichia pastoris* expressed *Aspergillus niger* 1,4-β-endoglucanase. *Biochemical and Biophysical Research communications*, 475(1), 8-12.

Yu, H., Zhao, Y., Guo, C., Gan, Y., and Huang, H. (2015). The role of proline substitutions within flexible regions on thermostability of luciferase. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1854*(1), 65-72.

Yuan, S.F., Wu, T.H., Lee, H.L., Hsieh, H.Y., Lin, W.L., Yang, B., and Ho, M.C. (2015). Biochemical characterization and structural analysis of a bifunctional cellulase/xylanase from *Clostridium thermocellum*. *Journal of Biological Chemistry*, 290(9), 5739-5748.

Zhai, X., Amyes, T.L., and Richard, J.P. (2015). Role of loop-clamping side chains in catalysis by triosephosphate isomerase. *Journal of the American Chemical Society*, 137(48), 15185-15197.

Zheng, F., Tu, T., Wang, X., Wang, Y., Ma, R., Su, X., and Luo, H. (2018). Enhancing the catalytic activity of a novel GH5 cellulase *Gt*Cel5 from *Gloeophyllum trabeum* CBS 900.73 by site-directed mutagenesis on loop 6. *Biotechnology for Biofuels*, 11(1), 76.

TH-2393_146106034

Chapter 6

Application of mutant endoglucanase *Ba*GH5-UV2 in saccharification of *Sorghum durra* biomass

6.1 Introduction

Cellulases are a group of enzymes, that converts the cellulosic polysaccharides into fermentable sugars for biofuel production. They are classified into the three classes of enzymes, endoglucanases (EC 3.2.1.4), cellobiohydrolases and β-glucosidases (EC 3.2.1.21). Endoglucanase randomly cleaves β-1,4 glycosidic bonds in the cellulose chain and exposes new non-reducing ends and reducing ends for the action of cellobiohydrolase to produce cellobiose. β-Glucosidase converts cellobiose into the fermentable sugar, glucose. All these three enzymes work synergistically to hydrolyze the cellulosic polysaccharide by generating new accessible sites for one other and avoids product inhibition (Glabe and Zacchi, 2002). However, the cost of enzymatic conversion of cellulose into the fermentable sugar is high and is the major bottleneck for the biofuel industries (Hess, 2008). Therefore, the cost-effective enzymatic saccharification of cellulose is desirable for the bio-conversion of lignocellulosic biomasses into the bioethanol (Levine *et al.*, 2011). The rate of saccharification and monomeric sugar yield from lignocellulosic biomass depend on mixture composition of different cellulases (endoglucanase, cellobiohydrolase and β-glucosidase) that rely

TH-2393_146106034

on biomass composition and the type of pretreatment method (Zhao *et al.,* 2009; Marcos *et al.,* 2013). The pretreatment methods used for lignocellulosic biomass depending on the physical characteristics and biomass composition differs. The acid and alkaline pretreatments are the two most widely used cost effective methods for lignocellulosic biomass (Singh *et al.,* 2014; Jamaldheen *et al.,* 2018). After acidic pretreatment with dilute sulfuric acid, the considerable amount of hemicellulosic fraction is solubilized from lignocellulosic biomass (Kim *et al.,* 2013). In contrast, the alkaline pretreatment by diluted sodium hydroxide removes lignin from the lignocellulosic biomass (Kim *et al.,* 2016). Therefore, the tailor-made cocktail of cellulases pertaining to the type, ratio and amount of enzymes, specific to the pretreatment process is crucial for maximization of the saccharification (Banerjee *et al.,* 2010; Kim *et al.,* 2015). The optimum enzymes ratio in a cocktail was designed on the basis of rational mixture design experiments with a concept of synergistic effects of enzymes on biomass (Zhao *et al.,* 2009). Rational mixture design via a mechanistic modeling provides the powerful tool for knowing the substrate specific features and enzymatic reaction steps that affect enzymatic saccharification (Bansal *et al.,* 2009).

In the present study, the application of crude mutant endoglucanase, *Ba*GH5-UV2 in cocktail with crude cellobiohydrolase (*Ct*CBH5A) and crude β-glucosidase (*Ct*GH1) for saccharification of 1% (v/v) H_2SO_4 (acid) assisted autoclaving and 1% (w/v) NaOH (base) assisted autoclaving pretreated *Sorghum durra* stalk biomass was performed. The mixture design, a rational statistical approach, was employed for the pretreatment (acid or base) specific customization of the cellulase cocktails by using the Stat-Ease Design Expert software (Stat-Ease Inc., Minneapolis, MS, USA,

TH-2393_146106034

Version 6.0.4). The simplex centroid mixture design was carried out to optimize the three enzymes (*Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1) ratio in a cocktail. The response was evaluated through saccharification reaction assays and validated. The above approach can be helpful in future for the development of enzyme cocktails, specific to the pretreatment methods.

TH-2393_146106034

6.2 Material and Methods

6.2.1 Media and chemicals

Carboxymethylcellulose sodium salt (CMC-Na) (low viscosity) and D-glucose were procured from Sigma Aldrich, USA. Luria Bertani broth, kanamycin, chloramphenicol, ampicillin, sodium azide, sodium acetate, sodium phosphate monobasic, sodium phosphate dibasic, sodium hydroxide and cellobiose were purchased from Himedia Pvt. Ltd. India. Hydrochloric acid and acetic acid were purchased from Merck, India. Glucose oxidase-Peroxidase (GOD-POD) kit was purchased from the coral clinical system (Tulip Diagnostics, Pvt. Ltd., India). Syringe filters were procured from Axiva (Axiva Sichem Pvt. Ltd., India). Chemicals used for reagent preparation for sugar estimation were purchased from Fisher Scientific, India.

6.2.2 Biomass collection, processing and pretreatment

Sorghum durra stalk biomass was chopped, washed with tap water and dried in hot air oven at 70ºC for 24 h. Ground biomass (particle size ~1 mm) was pretreated with 1% (w/v) NaOH (base) assisted autoclaving at 121ºC, 15 psi for 20 min and 1% (v/v) H_2SO_4 (acid) assisted autoclaving at 121ºC, 15 psi for 20 min, respectively, in separate 500 mL Erlenmeyer flask, respectively, in a solid to liquid ratio 1:10 according to the protocol as reported earlier by Jamaldheen *et al.*, (2018). The pretreated hydrolysate was discarded and the residual pretreated biomass was neutralised with tap water at room temperature till neutral pH then again washed with 500 mL distilled water. The neutralized pretreated *Sorghum durra* stalk biomass was dried in oven at 70ºC for 12 h.

TH-2393_146106034

6.2.3 Recombinant enzymes used in the saccharification of pretreated *Sorghum durra* stalk

The recombinant cellobiohydrolase (*Ct*CBH5A), cloned from *Clostridium thermocellum* was a generous gift from Professor Carlos M.G.A. Fontes, NZYTech Ltd., Lisbon, Portugal. *Ct*CBH5A was expressed in *E. coli* BL21(DE3) cells at 24ºC. β-Glucosidase (*Ct*GH1) cloned from *Clostridium thermocellum* and expressed in *E. coli* BL21 (DE3) cells earlier in our laboratory at 24ºC (Sharma *et al.*, 2019). The crude cell free supernatant of recombinant endoglucanase (*Ba*GH5-UV2), cellobiohydrolase (*Ct*CBH5A) and β-glucosidase (*Ct*GH1) were used in the enzyme cocktail preparation for the saccharification of *Sorghum durra* stalk. The crude cell free supernatant of recombinant endoglucanase (*Ba*GH5-UV2), cellobiohydrolase (*Ct*CBH5A) and β-glucosidase (*Ct*GH1) were used in the enzyme cocktail preparation for the saccharification of *Sorghum durra* stalk. Endoglucanase (*Ba*GH5-UV2) from UV2 mutant strain of *Bacillus amyloliquefaciens* SS35 was expressed in *E. coli* BL21(DE3)pLysS cells according to the method described in Chapter 4, section 4.3.7 The 1000 mL culture of each of the three expressed enzymes after IPTG induction (1mM), were centrifuged at 8000*g* for 15 min at 4ºC and resuspended in 8 mL of 20 mM sodium phosphate buffer, pH 6.5. The cell suspensions were sonicated according to the method as described in chapter 4, section 4.2.21. The cell free extract after sonication of all the three recombinant enzymes were explored, instead of the purified enzymes, in the enzymatic cocktail preparation for the hydrolysis of *Sorghum durra* stalk, in order to make the process of saccharification cost effective. The cell free extracts of all the three enzymes were sterilized by passing through the autoclaved

TH-2393_146106034

syringe filter (0.2µ pore size). The cell free extracts of enzymes were assayed for activity and protein concentration.

6.2.4 Assay of recombinant enzymes and protein estimation

The endoglucanase, *Ba*GH5-UV2 activity was determined under optimized conditions (Chapter 5, section 5.3.1 and 5.3.2) by incubating the 100 µL reaction mixture containing 1% (w/v) CMC-Na in 50 mM sodium acetate buffer, pH 5.5 and 10 µL of crude enzyme at 65°C for 3 min. The cellobiohydrolase, *Ct*CBH5A activity was measured under optimized conditions by incubating 100 µL reaction mixture containing 1% (w/v) CMC-Na in 50 mM sodium phosphate buffer, pH 6.4 and 10 µL of crude enzyme at 65°C for 1 min. The enzyme activity of *Ba*GH5-UV2 and *Ct*CBH5A was determined by reducing sugar estimation method as described earlier in Chapter 2, section 2.2.5. The enzyme activity of β-glucosidase, *Ct*GH1 was determined under optimized conditions by GOD-POD method (Raabo *et al.*, 1960) as described in chapter 5, section 5.2.5 using glucose as a blank (Sharma *et al.*, 2019). The 100 µL reaction mixture contained 1% (w/v) cellobiose in 20 mM citrate phosphate buffer, pH 6.0 and 10 µL of crude *Ct*GH1 and incubated at 65°C for 10 min according to the protocol as described earlier by Sharma *et al.*, (2019). The protein concentrations of crude cell free supernatant of *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 were determined by Lowry method (Lowry *et al*., 1951) using Bovine serum albumin (BSA) standard according to the protocol as described in Chapter 2, section 2.2.5.3.

6.2.5 Estimation of common buffer concentration for saccharification

The enzyme, *Ba*GH5-UV2 retains 100% of relative activity in 50 mM sodium acetate buffer in pH range 5.5 to 6.0 and in sodium phosphate buffer from pH range

TH-2393_146106034

6.0-7.5 as described earlier in Chapter 5, section 5.3.2. *Ct*CBH5A retains 100% of the relative activity in 50 mM sodium phosphate buffer in pH range 6.2 to 6.6 (unpublished data). The *Ct*GH1 protein retains 100% of the relative activity in 20 mM sodium citrate buffer in pH range 4.0 to 6.0 and in 20 mM sodium phosphate in pH range 7-8 as reported earlier by Sharma *et. al.,* (2019). The *Ba*GH5-UV2 protein retains 100% of the relative activity at temperature range, 30-45°C in 50 mM sodium acetate pH 5.5 as described earlier in Chapter 5, section 5.3.1. *Ct*CBH5A retains 100% of relative activity in temperature range, 30-65°C in 50 mM sodium phosphate buffer, pH 6.4 (unpublished data). *Ct*GH1 retains 100% of relative activity in temperature range, 20-40°C in 20 mM sodium citrate buffer, pH 6.0 as reported earlier by Sharma *et. al.*, (2019). In order to perform the enzymatic saccharification of pretreated *Sorghum durra* stalk by *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1, a common pH stable buffer (sodium phosphate, pH 6.5) and temperature (40°C) were chosen based on above results. For common optimum buffer concentration, the stability of the above three enzymes was determined by pre-incubating the crude enzymes (20 µL, 20 mg/mL) at 40°C for 6 h, 12 h and 24 h in 20 mM and 50 mM sodium phosphate buffer, pH 6.5, respectively. The residual enzyme activity of *Ba*GH5-UV2 and *Ct*CBH5A was determined by the method as described earlier by Nelson (1944) and Somogyi (1945) in chapter 2 and section 2.2.5 under optimized conditions given in section 6.2.4. The residual enzyme activity of *Ct*GH1 was determined by reaction conditions given in section 6.2.4 by the GOD-POD method as described earlier in as described in chapter 5, section 5.2.5.

TH-2393_146106034

6.2.6 Customization of limit of enzyme activities for mixture design by Stat-Ease Design Expert software

The limits of enzyme activities for crude enzymes, *Ba*GH5-UV2 or *Ct*CBH5A or *Ct*GH1 was determined against pretreated *Sorghum durra* stalk (acid or base) with a biomass loading of 2% (w/v) (since all the three enzymes showed maximum TRS at 2%, w/v biomass loading). The saccharification using crude enzymes, *Ba*GH5-UV2 or *Ct*CBH5A or *Ct*GH1 separately were carried out in 1 mL reaction volume by varying their loading concentrations *viz.* 200 U/g, 400 U/g, 800 U/g, 1200 U/g, 1600 U/g and 2000 U/g of pretreated biomass in 20 mM sodium phosphate buffer (pH 6.5). Further, to customize the concentration of *Ct*GH1, reaction was carried out in 1 mL reaction volume with a fixed loading concentration of *Ba*GH5-UV2 (800 U/g) and *Ct*CBH5A (400 U/g) of pretreated *Sorghum durra* stalk, respectively, by varying the *Ct*GH1 loading concentrations at 200 U/g, 400 U/g, 800 U/g, 1200 U/g, 1600 U/g, 2000 U/g of pretreated biomass. The reactions were carried out in 2 mL micro-centrifuge tube at 40°C under shaking condition of 150 rpm for 48 h in the triplicate sets. 0.005% (w/v) sodium azide and 100 µg/mL ampicillin was added to the reaction mixture to avoid contamination. After 48 h reactions were stopped by boiling the reaction mixtures in boiling water bath for 10 min. The TRS was estimated by the method as described in Chapter 2, section 2.2.5.

6.2.7 Optimization of enzymatic hydrolysis for acid and base pretreated *Sorghum durra* stalk by simplex centroid mixture design

6.2.7.1 Experimental design, statistical analysis and model fitting

The enzyme cocktail ratios for the three recombinant crude enzymes *Ba*GH5-UV2 (A), *Ct*CBH5A (B) and *Ct*GH1 (C) were optimized for saccharification of acid or base pretreated *Sorghum durra* stalk by using simplex centroid mixture RSM

TH-2393_146106034

design in Design Expert 7.0.0 software. A one-level, 3 parameters Simple-centroid mixture design was employed for the optimization of A, B and C mixture and 10 total runs of experiment (Table 6.6). The total loading amount of the three different enzymes in a mixture was 2000 U/g of pretreated biomass. The loading amounts of A, B, and C in terms of U/g of pretreated biomass were applied as independent variables. Here, units of A, B and C needed per gram of pretreated substrate was varied. The minimum and maximum level of parameters were set as follows, 10% and 80%. Therefore, the constraints of the total composition mixture of these three constituents were set as 100%. A principal advantage of using this above mixture design concept is to set flexible constraints in the total composition range. This showed that if the concentrations of one individual component are expressed as percentage, then it can be varied in a total between 1 and 100%. The TRS yields and glucose yield from acid pretreated biomass and base pretreated biomass, respectively were the response factors.

The reactions were carried out in 2 mL microcentrifuge tube with total reaction volume of 1 mL with enzyme loading according to the mixture ratios given in Table 6.6. 0.005% (w/v) sodium azide and 100 µg/mL ampicillin was added to the reaction mixture to avoid contamination. The microcentrifuge tubes were incubated at 40°C, 150 rpm for 48 h in an incubator shaker (Orbitek, Scigenics Biotech, India). After 48 h, the enzyme reaction mixture was boiled in boiling water bath for 15 min for inactivating the enzymes. The resulting reducing sugars were determined by Nelson (1944) and Somogyi (1945) method reported earlier as described in Chapter 2, section 2.2.5 by measuring the absorbance at 500 nm, on a spectrophotometer (Varian, Cary 100 Bio) using glucose as standards. The resulting hexose sugar glucose, was

TH-2393_146106034

measured by GOD-POD method (Raabo *et al.*, 1960), using glucose as standard as described earlier in as described in chapter 5, section 5.2.5.

To validate the mathematical model and experiment data, analysis of variance (ANOVA) was performed. Regression analysis was carried out for the best-fitted models. Contour plots for the best-fitted models were generated to predict the optimal formulations of the above enzymes (A, B and C). Eventually, the optimized composition mixture of three enzymes for both the pretreated biomass was determined and compared.

6.2.7.2 Time dependent enzymatic hydrolysis of *Sorghum durra* stalk under optimized condition of enzyme cocktail ratio

Time dependent enzymatic hydrolysis of acid or base pretreated *Sorghum durra* stalk(s) was evaluated at the optimal formulations of enzyme ratio in mixture obtained from the quadratic and special cubic models, respectively. 12 mL of each reaction mixture contained optimized mixture ratio of enzymes (*Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1) for respective pretreated biomass in 50 mL Erlenmeyer flask. The reactions were carried out for 2.0% (w/v) pretreated Sds in 20 mM sodium phosphate buffer, pH 6.5 in an incubator shaker at 40°C under shaking at 150 rpm (Orbitek, Scigenics Biotech, India) for 96 h. 20 µL aliquot of saccharification reaction samples were collected after 24 h, 48 h, 72 h and 96 h, in 1 mL microcentrifuge tube. The reactions were stopped by boiling the reaction mixtures in boiling water bath for 5 min. The resulting reducing sugars were determined by Nelson (1944) and Somogyi (1945) method reported earlier as described in Chapter 2, section 2.2.5. The resulting hexose sugar glucose was determined by GOD-POD method (Raabo *et al.*, 1960) according to the protocol as described in chapter 5, section 5.2.5.

TH-2393_146106034

6.3 Results and Discussion

6.3.1 Estimation of protein concentration and enzyme activity of recombinant enzymes and common concentration of buffer for saccharification

The protein concentration of crude *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 in cell free supernatant was 37 mg/mL, 32.1 mg/mL and 21.6 mg/mL, respectively. The crude *Ba*GH5-UV2 in cell free supernatant showed endoglucanase activity, 242 U/mL against CMC-Na in 50 mM sodium acetate buffer, pH 5.5 at 65°C in 3 min. The crude *Ct*CBH5A in cell free supernatant showed cellobiohydrolase activity, 83.2 U/mL against CMC-Na in 50 mM sodium phosphate buffer, pH 6.4, 65°C in 1 min and crude *Ct*GH1 showed β-glucosidase activity 22.5 U/mL against cellobiose in 20 mM citrate phosphate buffer, pH 6.0 and at 65°C in 10 min. The enzymes, *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 retained the maximum (100%) residual enzyme activity in 20 mM sodium phosphate buffer, pH 6.5 at 40°C in 24 h (Table 6.1). Therefore, the buffer concentration for performing the saccharification of pretreated *Sorghum durra* stalk in the cocktail combination of *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 was 20 mM sodium phosphate buffer, pH 6.5 at 40°C.

Table 6.1 Determination of stable buffer concentration for crude *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 at 40°C

Enzyme	**Relative enzyme activity, % (U/mL) 20 mM Sodium phosphate buffer, pH 6.5**			**Relative enzyme activity, % (U/mL) 50 mM, Sodium phosphate buffer, pH 6.5**		
	6h	12h	24h	6h	12h	24h
*Ba*GH5-UV2	100 ± 9.8	100 ± 11.7	100 ± 9.5	100 ± 9.8	98.6 ± 7.9	95 ± 5.6
*Ct*CBH5A	100 ± 5.3	100 ± 6.2	100 ± 3.7	98.2 ± 3.3	96.3 ± 3.5	91.2 ± 2.6
*Ct*GH1	100 ± 1.1	99.8 ± 1.8	99.1 ± 2.8	92.1 ± 1.4	86.5 ± 1.1	60.9 ± 1.9

values are mean SE (n=3).

TH-2393_146106034

6.3.2 Customization of limits of enzyme activities for mixture design by a Stat-Ease Design Expert software

The TRS yield from acid pretreated *Sorghum durra* stalk increases with the increase in *Ba*GH5-UV2 or *Ct*CBH5A concentration. For *Ba*GH5-UV2 and *Ct*CBH5A concentration, 800 U/g of pretreated biomass gave the maximum saccharification yield of 28 mg/g and 9 mg/g of pretreated biomass, respectively (Table 6.2). *Ct*GH1 enzyme alone did not produce any reducing sugars from acid pretreated biomass. However, in cocktail with *Ba*GH5-UV2 and *Ct*CBH5A it produces maximum TRS yield of 46 mg/g of pretreated biomass at *Ct*GH1 concentration, 400 U/g of pretreated biomass (Table 6.3). Therefore, total concentration loading for all the three enzymes used in a mixture was predicted to be 2000 U/g of acid pretreated biomass.

Saccharification yield from base pretreated *Sorghum durra* stalk increases with the increase in *Ba*GH5-UV2 and *Ct*CBH5A concentration. For *Ba*GH5-UV2 and *Ct*CBH5A concentration, 800 U/g of pretreated biomass gave the maximum saccharification yield of 43 mg/g and 16 mg/g of pretreated biomass, respectively (Table 6.4). *Ct*GH1 enzyme alone did not produce any reducing sugars from base pretreated biomass. However, in cocktail with *Ba*GH5-UV2 and *Ct*CBH5A it produces TRS yield of 61 mg/g of pretreated biomass at *Ct*GH1 concentration, 400 U/g of pretreated biomass (Table 6.5). However, further increase in *Ct*GH1 conc. did not cause significant increase in TRS yield. Therefore, total concentration loading for all the three enzymes used in a mixture was predicted to be 2000 U/g of base pretreated biomass.

TH-2393_146106034

Table 6.2 Customization of limits of enzymes to be used in mixture design for acid pretreated *Sorghum durra* stalk

	TRS yield (mg/g of pretreated biomass), 48 h					
Enzymes Units/g of pretreated biomass	200	400	800	1200	1600	2000
*Ba*GH5-UV2	3.4±0.2	11.3±0.9	27.6±1.2	28.1±2.5	28.5±2.1	27.4±3.6
*Ct*CBH5A	1.1±0.01	2.8±0.1	9.4±1.0	9.7±1.5	9.3±0.8	9.8±0.8
*Ct*GH1	0	0	0	0	0	0

± *standard deviation (n=3).*

Table 6.3 Customization of limits for *Ct*GH1 to be used in mixture design for acid pretreated *Sorghum durra* stalk

	TRS yield (mg/g of pretreated biomass), 48 h					
*Ct*GH1 (U/g of pretreated biomass)	200	400	800	1200	1600	2000
(*Ba*GH5-UV2 + *Ct*CBH5A)+*Ct*GH1	39.5±2.6	46.3±3.5	46.3±2.2	47.5±2.9	48.1±4.0	47.8±3.9

At fixed concentration of BaGH5-UV2 (800 U/g of pretreated biomass) and CtCBH5A (400 U/g of pretreated biomass), ± standard deviation (n=3).

Table 6.4 Customization of limits of enzymes to be used in mixture design for base pretreated *Sorghum durra* stalk

	TRS yield (mg/g of pretreated biomass), 48 h					
Enzymes (Units/g of pretreated biomass)	200	400	800	1200	1600	2000
*Ba*GH5-UV2	7.8±0.9	19.9±1.1	42.6±3.5	41.2±4.1	44.7±2.7	43.6±3.4
*Ct*CBH5A	1.4±0.1	9.5±0.8	16.1±1.4	16.9±1.1	16.0±1.7	17.3±2.3
*Ct*GH1	0	0	0	0	0	0

± standard deviation *(n=3).*

Table 6.5 Customization of limits of *Ct*GH1 to be used in mixture design for base pretreated *Sorghum durra* stalk

	TRS yield (mg/g of pretreated biomass), 48 h					
*Ct*GH1 (U/g of pretreated biomass)	200	400	800	1200	1600	2000
(*Ba*GH5-UV2 + *Ct*CBH5A) + *Ct*GH1	51.7±2.5	60.9±5.6	69.8±3.8	70.2±5.3	69.7±5.0	70.9±4.4

At fixed concentration of BaGH5-UV2 (800 U/g of pretreated biomass) and CtCBH5A (400 U/g of pretreated biomass), ± standard deviation (n=3).

TH-2393_146106034

6.3.3 Enzyme ratio optimization in a mixture with TRS yield and glucose yield as the response variable against acid or base pretreated *Sorghum durra* stalk

6.3.3.1. Model data sheet

The optimization of all the three crude enzymes *Ba*GH5-UV2 (A), *Ct*CBH5A (B) and *Ct*GH1 (C) in a mixture within a range of lower limit 200 U/g of pretreated biomass to maximum limit 1600 U/g of pretreated biomass was carried out for getting the maximum TRS yield and glucose yield from acid or base pretreated *Sorghum durra* stalk. On the basis of simplex-centroid model, 10 run data set was generated by simplex centroid mixture design based RSM design approach with varying ratios of (A), (B) and (C) with the total enzyme loading of 2000 U/g of pretreated biomass (Table 6.6). In case of acid pretreatment of *Sorghum durra* stalk run 9 (A45%, B10% and C45%) gave the highest TRS yield and glucose yield of 90.8 mg/g and 69.3 mg/g of pretreated biomass, respectively. Run 1 (A21.6%, B21.7% and C56.6%) gave the second highest TRS yield and glucose yield of 89.5 and 61.4 mg/g of pretreated biomass, respectively at 48 h of saccharification. However, in case of base pretreated *Sorghum durra* stalk the run 4 (A80%, B10% and C10%) gave the highest TRS yield and glucose yield of 165.6 mg/g and 81.7 mg/g of pretreated biomass, respectively. Run 5 (A45%, B45% and C10%) gave the second highest TRS yield and glucose yield of 122.0 and 59.1 mg/g of pretreated biomass, respectively, at 48 h of saccharification. This showed that for acid pretreated *Sorghum durra* stalk, the maximum TRS yield and glucose yield were produced when the proportion of A and C were high and B was relatively low in enzyme cocktail. In case of base pretreated *Sorghum durra* stalk, A has the major role in contributing for TRS yield and glucose yield because with high proportion of A in enzyme mixture, TRS yield and glucose

TH-2393_146106034

yield was maximum. Kim *et al.*, (2017) performed the mixture design experiment for optimizing the endoglucanase, cellobiohydrolase, and exoxylanase ratio in a cocktail for maximum TRS yield from acid and base pretreated Sugarcane bagasse and found that endoglucanase has major role in a cocktail ratio in producing the high amount of reducing sugars.

Table 6.6 Simplex centroid mixture design data inputs

Run no.	Variables % (U/g of biomass)			Acid pretreated *Sorghum durra* stalk		Base pretreated *Sorghum durra* stalk	
	*Ba*GH5-UV2 (A)	*Ct*CBH5A (B)	*Ct*GH1 (C)	TRS yield[a]	Glucose yield[a]	TRS yield[b]	Glucose yield[b]
1	21.6	21.7	56.6	**89.5**	**61.4**	70.1	36.5
2	10	10	80	39.6	26.4	81.2	42.1
3	10	80	10	45.9	29	86.7	43.7
4	80	10	10	26.8	17.9	**165.6**	**81.7**
5	45	45	10	62.1	39.5	**122.0**	**59.1**
6	56.6	21.7	21.6	69.6	49.4	112.1	55.2
7	33.3	33.3	33.3	83.3	53.9	115.9	57.1
8	10	45	45	76.5	56.1	63.4	30.2
9	45	10	45	**90.8**	**69.3**	63.3	29.2
10	21.6	56.7	21.6	71.9	44.5	94.5	44.6

TRS yield and Glucose yield (mean of 3),
(a) yield from acid pretreated Sorghum durra stalk response variables,
(b) yield from base pretreated Sorghum durra stalk response variables.

6.3.3.2 Fitting of the regression models

6.3.3.2.1 Model selection

Different models were compared for selecting the model that explains the best relationship between independent variables (A, B and C) and dependent variable (TRS yield and glucose yield from acid or base pretreated *Sorghum durra* stalk). Various predicted models for TRS yield and glucose yield from acid pretreated biomass are shown in Tables 6.7.1 and 6.7.2 and from base pretreated biomass in Tables 6.8.1 and 6.8.2. For a model to be considered the most suitable, the F-value should be as high and the p-value should be low. In case of acid pretreated biomass,

TH-2393_146106034

when TRS yield and glucose yield were considered as a response factor, the quadratic model showed the highest F value of 55.54 and 29.81, respectively, with lowest p-value of < 0.001 and < 0.0034, respectively. Therefore, among all models, the quadratic model was found to be the best suited model for acid pretreated *Sorghum durra* stalk while considering TRS yield and glucose yield as response variables (Tables 6.7.1 and Table 6.7.2). In case of base pretreated biomass, when TRS yield and glucose yield were considered as a response factor, the special cubic model showed the highest F value of 10.4 and 14.6, respectively, with lowest p-value of < 0.0484 and < 0.0315, respectively. Therefore, the special cubic model was considered to be the best suited model for base pretreated *Sorghum durra* stalk while considering TRS yield and glucose yield as response variables (Table 6.8.1 and Table 6.8.2). Kim *et al.*, (2017) showed that for enzyme mixture (endoglucanase, cellobiohydrolase, β-glucosidase and endoxylanase) hydrolysis of acid and base pretreated sugarcane bagasse, special cubic and full cubic model, respectively, were the best predicted one.

Table 6.7.1 Mixture model design methodology RSM Model output data summary (TRS yield) for acid pretreated biomass

Source	**Sum of squares**	**Df**	**Mean square**	**F Value**	**p-Value Prob > F**	**Comments**
Mean vs total	43124.91	1	43124.91			
Linear vs mean	300.55	2	150.28	0.26	0.7747	
Quadratic vs linear	3878.88	3	1292.96	55.54	0.0010	Suggested
Special cubic vs quadratic	11.95	1	11.95	0.44	0.5538	
Cubic vs special cubic	60.01	2	30.01	1.42	0.5106	Aliased
Residual	21.16	1	21.16			
Total	47397.46	10	4739.75			

TH-2393_146106034

Table 6.7.2 Mixture model design methodology RSM Model output data summary (glucose yield) for acid pretreated biomass

Source	Sum of squares	df	Mean square	F Value	p-Value Prob > F	Comments
Mean vs total	19871.36	1	19871.36			
Linear vs mean	214.36	2	107.18	0.34	0.7224	
Quadratic vs linear	2107.3	3	702.46	29.81	0.0034	**Suggested**
Special cubic vs quadratic	64.09	1	64.09	6.37	0.0858	
Cubic vs special cubic	22.11	2	11.06	1.37	0.5166	Aliased
Residual	8.05	1	8.05			
Total	22287.35	10	2228.74			

Table 6.8.1 Mixture model design methodology RSM Model output data summary (TRS yield) for base pretreated biomass

Source	Sum of squares	df	Mean square	F Value	p-Value Prob > F	Comments
Mean vs total	95074.68	1	95074.68			
Linear vs mean	6238.05	2	3119.03	7.19	0.0201	
Quadratic vs linear	1890.20	3	630.07	2.20	0.2310	
Special cubic vs quadratic	890.33	1	890.33	10.40	0.0484	**Suggested**
Cubic vs special cubic	32.21	2	16.11	0.072	0.9352	Aliased
Residual	224.72	1	224.72			
Total	104350.2	10	10435.02			

Table 6.8.2 Mixture model design methodology RSM Model output data summary (glucose yield) for base pretreated biomass

Source	Sum of squares	df	Mean square	F Value	p-Value Prob > F	Comments
Mean vs total	23022.76	1	23022.76			
Linear vs mean	1396.31	2	698.15	5.66	0.0345	
Quadratic vs linear	530.77	3	176.92	2.12	0.2400	
Special cubic vs quadratic	276.49	1	276.49	14.62	0.0315	Suggested
Cubic vs special cubic	7.16	2	3.58	0.072	0.9348	Aliased
Residual	49.59	1	49.59			
Total	25283.08	10	2528.31			

TH-2393_146106034

6.3.3.2.2 Analysis of variance (ANOVA)

Analysis of variable (ANOVA) for acid pretreated *Sorghum durra* stalk results confirmed the significance of quadratic model for both the response factors, TRS yield and glucose yield. The R^2 and adjusted R^2 values for the generated model for TRS yield as response factor were 0.97 and 0.95, respectively, and for glucose yield as response factor were 0.97 and 0.92, respectively. In case of acid pretreated biomass, for TRS yield as the response factor, interactions of AC variables showed highest F value of 111.58 which was followed by BC variables (F value of 37.63) and AB variables gave the lowest F value of 19.08 (Table 6.9). For glucose yield as the response factor, interactions of AC variables showed highest F value of 65.21 which was followed by BC variables (F value of 20.62) and AB variables gave the lowest F value of 4.38 (Table 6.10). These results suggested that the interactions were maximum between the endoglucanase and β-glucosidase with the highest synergistic effect, whereas, the synergistic effect of endoglucanase and cellobiohydrolase was minimum on acid pretreated *Sorghum durra* stalk. This showed that for the efficient hydrolysis of acid pretreated *Sorghum durra* stalk, the positive synergism (or cooperation) between endo- (*Ba*GH5-UV2) and exo-acting, β-glucosidase (*Ct*GH1), and *Ct*CBH5A and *Ct*GH1 cellulases is required.

The variation in the TRS yield with variation in endoglucanase (A), cellobiohydrolase (B) and β-glucosidase (C) compositions from acid pretreated *Sorghum durra* stalk could be expressed by using the following equations generated from quadratic model:

TRS yield = - 0.34763 x *Ba*GH5UV2 + 0.12190 x *Ct*CBH5A - 0.19983 x *Ct*GH1 + 0.019117 x *Ba*GH5UV2 x *Ct*CBH5A + 0.046233 x *Ba*GH5UV2 x *Ct*GH1 + 0.026849 x *Ct*CBH5A x *Ct*GH1

TH-2393_146106034

The variation in the glucose yield with variation in endoglucanase (A), cellobiohydrolase (B) and β-glucosidase (C) compositions from acid pretreated *Sorghum durra* stalk could be expressed by using the following equations generated from quadratic model:

Glucose yield = - 0.22399 x *Ba*GH5UV2 + 0.074904 x *Ct*CBH5A - 0.19632 x *Ct*GH1 + 0.00921052 x *Ba*GH5UV2 x *Ct*CBH5A + 0.035558 x *Ba*GH5UV2 x *Ct*GH1 + 0.019993 x *Ct*CBH5A x *Ct*GH1

Table 6.9 Analysis of Variance output data summary for TRS yield as response variables for acid pretreated biomass

Source	Sum of squares	df	Mean square	F Value	p-Value Prob > F	Comments
Model	4179.43	5	835.89	35.90	0.0020	Significant
Linear Mixture	300.55	2	150.28	6.46	0.0560	
AB	444.13	1	444.13	19.08	0.0120	
AC	2597.56	1	2597.56	111.58	0.0005	
BC	876.00	1	876.00	37.63	0.0036	
Residual	93.12	4	23.28			
Cor Total	4272.55	9				

Table 6.10 Analysis of Variance output data summary for glucose yield as response variables for acid pretreated biomass

Source	Sum of squares	df	Mean square	F Value	p-Value Prob > F	Comments
Model	2321.74	5	464.35	19.71	0.0064	significant
Linear Mixture	214.36	2	107.18	4.55	0.0933	
AB	103.09	1	103.09	4.38	0.1046	
AC	1536.49	1	1536.49	65.21	0.0013	
BC	485.77	1	85.77	20.62	0.0105	
Residual	94.25	4	23.56			
Cor Total	2416.00	9				

In case of base pretreated biomass, the R^2 and adjusted R^2 values for the generated model for TRS yield as response factor were 0.97 and 0.91, respectively, and for glucose yield as response factor were 0.97 and 0.92, respectively. For TRS yield as the response factor, interactions of AC variables showed highest F value of

TH-2393_146106034

30.05 which was followed by ABC variables (F value of 10.4) and this was followed by BC interactions (F value of 3.47). However, the interactions between the AB variables gave the lowest F value of 0.22 (Table 6.11). For glucose yield as the response factor, interactions of AC variables showed highest F value of 39.15 which was followed by ABC variables (F value 14.62) and this was followed by BC variable (F value 5.92) (Table 6.12). Whereas, the AB variables gave the lowest F value of 0.67 (Table 6.12) thus, showed the negligible synergistic effect. These results suggested that the interactions were maximum between the endoglucanase and β-glucosidase showing the highest synergistic effect whereas the interactions between endoglucanase and cellobiohydrolase showed the lowest synergistic effect on base pretreated *Sorghum durra* stalk. However, all the three enzymes showed very less synergistic effect. These results also showed that in enzymatic hydrolysis of base-pretreated *Sorghum durra* stalk, there was less synergism between endo-acting (*Ba*GH5-UV2) and exo-acting, cellobiohydrolase (*Ct*CBH5A) and β-glucosidase (*Ct*GH1) as compared with the acid pretreated *Sorghum durra* stalk. This might have happened due to the morphological changes in the biomass because acid pretreatment mainly solubilizes hemicellulosic fraction and base pretreatment mainly delignified the biomass. Zhou *et al.,* (2009) optimized a cellulase cocktail using a two-level fractional factorial design. In this type of experimental design, the factor levels vary independently from each other without showing synergistic effect, and consequently, each run presents a different protein load.

The variation in the TRS yield with variation in endoglucanase (A), cellobiohydrolase (B) and β-glucosidase (C) compositions from base pretreated

TH-2393_146106034

Sorghum durra stalk could be expressed by using the following equations generated from special cubic model:

TRS yield = + 2.58753 x *Ba*GH5UV2 + 1.16094 x *Ct*CBH5A + 1.53112 x *Ct*GH1 - 0.032253 x *Ba*GH5UV2 x *Ct*CBH5A - 0.078287 x *Ba*GH5UV2 x *Ct*GH1 - 0.045029 x *Ct*CBH5A x *Ct*GH1 + 0.00279078 x *Ba*GH5UV2 x *Ct*CBH5A x *Ct*GH1

The variation in the glucose yield with variation in endoglucanase (A), cellobiohydrolase (B) and beta glucosidase (C) compositions from base pretreated *Sorghum durra* stalk could be expressed by using the following equations generated from special cubic model:

Glucose yield = + 1.31995 x *Ba*GH5UV2 + 0.61925 x *Ct*CBH5A + 0.83946 x *Ct*GH1 - 0.019097 x *Ba*GH5UV2 x *Ct*CBH5A - 0.042577 x *Ba*GH5UV2 x *Ct*GH1 - 0.02606 x *Ct*CBH5A x *Ct*GH1 + 0.00155522 x *Ba*GH5UV2 x *Ct*CBH5A x *Ct*GH1

Table 6.11 Analysis of Variance output data summary for TRS yield as response variables for base pretreated biomass

Source	Sum of squares	df	Mean square	F Value	p-Value Prob > F	Comments
Model	9018.58	6	1503.10	17.55	0.0195	significant
Linear Mixture	6238.05	2	3119.03	36.42	0.0079	
AB	19.14	1	19.14	0.22	0.6686	
AC	2573.56	1	2573.56	30.05	0.0119	
BC	297.24	1	297.24	3.47	0.1594	
ABC	890.33	1	890.33	10.40	0.0484	
Residual	256.94	3	85.65			
Cor Total	9275.52	9				

Table 6.12 Analysis of Variance output data summary for glucose yield as response variables for base pretreated biomass

Source	Sum of squares	df	Mean square	F Value	p-Value Prob > F	Comments
Model	2203.58	6	367.26	19.42	0.0169	significant
Linear Mixture	1396.31	2	698.15	36.9	0.0077	
AB	12.74	1	12.74	0.67	0.4719	
AC	740.54	1	740.54	39.15	0.0082	
BC	112.07	1	112.07	5.92	0.0930	
ABC	276.49	1	276.49	14.62	0.0315	
Residual	56.75	3	18.92			
Cor Total	2260.33	9				

TH-2393_146106034

6.3.3.2.3 Model graphs

The ternary 3D contour plots were developed to determine the optimal ratio of *Ba*GH5-UV2 (A), *Ct*CBH5A (B) and *Ct*GH1 (C) in the enzyme mixture for the saccharification of acid pretreated *Sorghum durra* stalk on the basis of the quadratic equations generated above in section 6.3.3.2.2 and for base pretreated *Sorghum durra* stalk on the basis of special cubic equations generated above in section 6.3.3.2.2. In case of acid pretreated *Sorghum durra* stalk, the contour plots for TRS and glucose yield responses were developed from the quadratic models. These models showed the areas that outlined the highest TRS yield of 91 mg/g of pretreated biomass (Fig. 6.1) and glucose yields of 69 mg/g of pretreated biomass (Fig. 6.2) were positioned for mixtures with high proportions of *Ba*GH5-UV2 and *Ct*GH1 and a fairly low proportion of *Ct*CBH5A. These results also confirmed the strong positive synergism between endo-acting (*Ba*GH5-UV2) and exo-acting, β-glucosidase (*Ct*GH1) in the hydrolysis of acid pretreated *Sorghum durra* stalk. However, in case of base pretreated *Sorghum durra* stalk, the contour plots for TRS and glucose yield responses were developed from the special cubic models. These models showed the areas that outlined the highest TRS yield of 165.6 mg/g of pretreated biomass (Fig. 6.3) and glucose yields of 81.7 mg/g of pretreated biomass (Fig. 6.4) were positioned for mixtures with high proportions of *Ba*GH5-UV2 and low proportion of *Ct*GH1 and low proportion of *Ct*CBH5A. These results also confirm the less synergism between endo-acting (*Ba*GH5-UV2) and exo-acting, β-glucosidase (*Ct*GH1) and between cellobiohydrolase (*Ct*CBH5A) and *Ct*GH1. Moreover, there was no synergism was observed between *Ba*GH5-UV2 and *Ct*CBH5A in the enzymatic hydrolysis of base-pretreated *Sorghum durra* stalk. Gao *et al.*, 2010 reported the influence of protein

TH-2393_146106034

loading in the ratio of cocktail composition for the saccharification of ammonia fiber expansion pretreated corn stover and observed that the endoglucanase was more important to glucose release with a low protein load of exoglucanase.

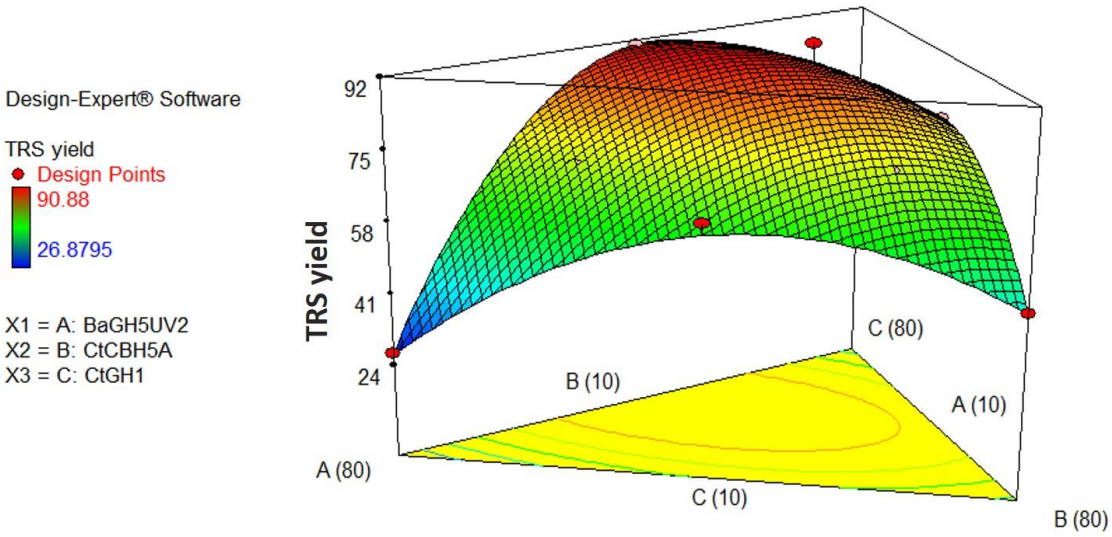


Fig. 6.1 3D contour plots for TRS yield response from acid pretreated *Sorghum durra* stalk.

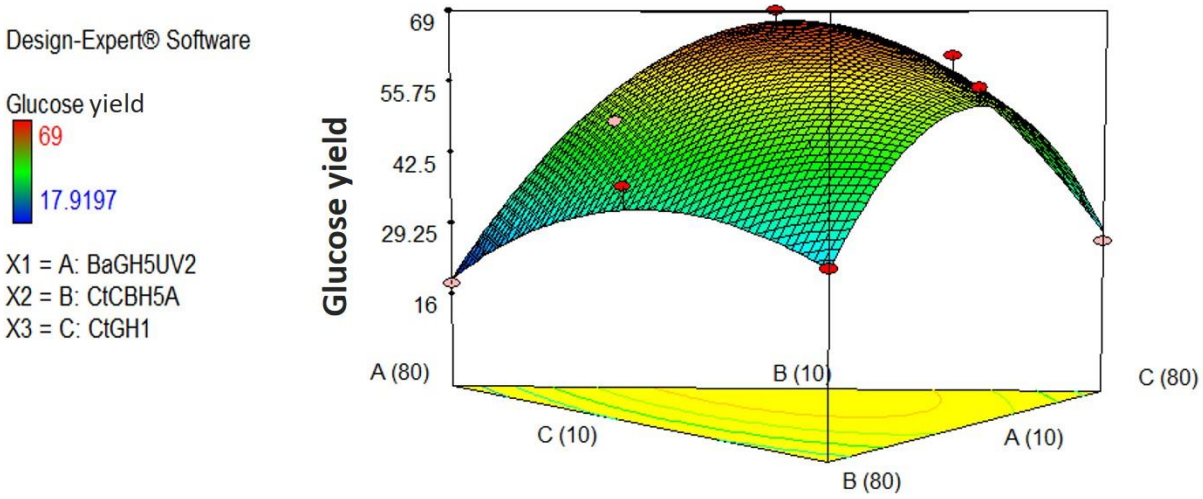


Fig. 6.2 3D contour plots for glucose yield response from acid pretreated *Sorghum durra* stalk.

TH-2393_146106034

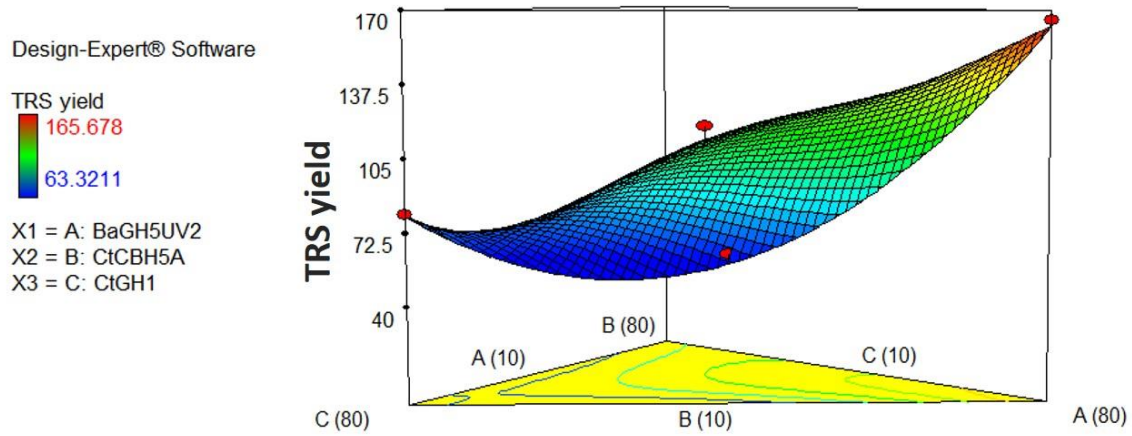


Fig. 6.3 3D contour plots for TRS yield response from base pretreated *Sorghum durra* stalk.

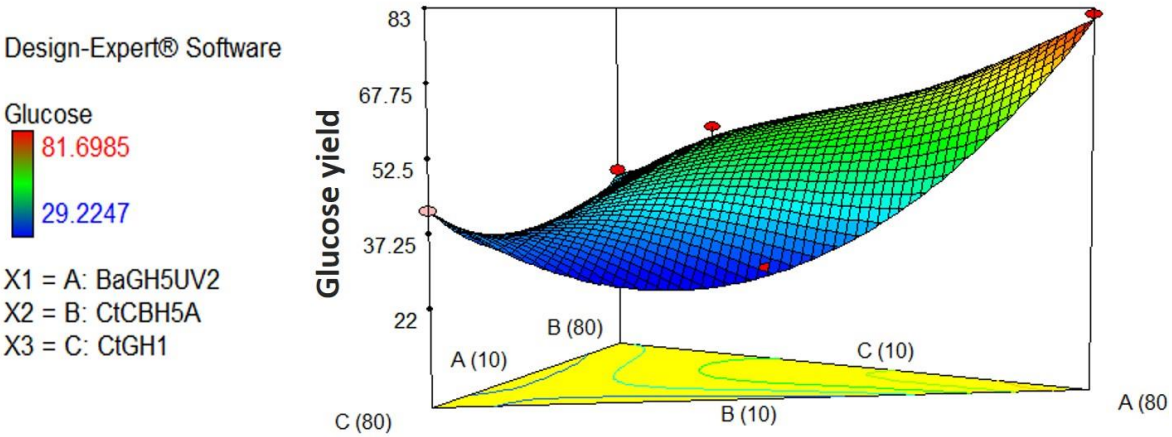


Fig. 6.4 3D contour plots for glucose yield response from base pretreated *Sorghum durra* stalk.

6.3.3.2.4 Optimal mixture ratio composition for acid and base pretreated *Sorghum durra* stalk

The specific optimal ratios leading to the maximal production of TRS and glucose yield were determined by setting the goal limit maximum for TRS yield and glucose yield. In case of acid pretreated *Sorghum durra* stalk the optimal mixture ratios composition of *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 was 43.1%, 10.0%, and 46.9%, respectively, predicted by the quadratic model. In case of base pretreated *Sorghum durra* stalk the optimal mixture ratios composition of *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 was 80%, 10%, and 10%, respectively, predicted by cubic models. Kim *et al.*, (2017) found the optimized predicted ratio for acid pretreated Sugarcane bagasse was 61.25%:38.73%:0.02% (endoglucanase, cellobiohydrolase, and exoxylanase) and for base pretreated Sugarcane bagasse 53.99%:34.60%:11.41% (endoglucanase, cellobiohydrolase, and exoxylanase) by RSM design. Bussamra *et al.*, 2015 optimized the cocktail supplementation of commercial and microbial (*Trichoderma reesei* and *Escherichia coli*) cellulase enzymes for sugar cane bagasse hydrolysis by simplex lattice mixture design. The optimized enzyme mixture, comprised of *T. reesei* fraction (80%), endoglucanase (10%) and β-glucosidase (10%), converted, theoretically, 72% of cellulose present in hydrothermally pretreated bagasse.

6.3.4 Saccharification of acid and base pretreated *Sorghum durra* stalk under optimized conditions

Finally, to validate the statistical prediction, a hydrolysis experiment was performed with acid and base pretreated *Sorghum durra* stalk at the optimal formulations. Therefore, the saccharification of acid or base pretreated *Sorghum durra* stalk for optimized enzyme ratios were carried out at 40°C for the different time

TH-2393_146106034

interval of 24 h, 48 h, 72 h and 96 h and enzymatic hydrolysates were analysed for TRS yield and glucose yield. The base pretreated *Sorghum durra* stalk gave TRS yield 59.2, 165.8, 193.3 and 211 mg/g of pretreated biomass at 24 h, 48 h, 72 h and 96 h of saccharification with a glucose yield of 26.3, 81.7, 159.9 and 174.2 mg/g of pretreated biomass, respectively. However, the acid pretreated *Sorghum durra* stalk gave TRS yield 38.6, 91.1, 112.6 and 119.5 mg/g of pretreated biomass at 24 h, 48 h, 72 h and 96 h of saccharification with a glucose yield of 10.2, 69.5, 84.7 and 85.3 mg/g of pretreated biomass, respectively. The maximum TRS yield and glucose yield for base pretreated *Sorghum durra* stalk were obtained at 96 h. However, in case of acid pretreated *Sorghum durra* stalk the TRS yield and glucose yield was nearly same at 72 h and 96 h. These results showed that the predicted as well as the experimental ratios for the TRS yield and glucose yield from acid and base pretreated biomasses are similar. Jamaldheen *et al.,* (2018) showed that the enzymatic hydrolysis of 1% (w/v) NaOH assisted autoclaving pretreated *Sorghum durra* stalk by a endoglucanase (*Ct*Cel8A: 16 U/g of biomass) with β-glucosidase (*Ct*Bgl1A: 21U/g of biomass) produce a maximum glucose yield of 34 mg/g of raw biomass. Nath *et al.,* (2018) showed that the enzymatic hydrolysis of 1% (w/v) NaOH assisted autoclaving pretreated *Sorghum durra* stalk by chimera (endoglucanase and β-glucosidase) gave a maximum glucose yield of 47.8 mg/g of raw biomass. Thus, a rational enzyme mixture designed by using the synergistic concept and statistical analysis was capable of improving the biomass saccharification.

TH-2393_146106034

6.4 Conclusion

The rational mixture design approach was applied to optimize the ratio of three crude recombinant enzymes, endoglucanase (*Ba*GH5-UV2), cellobiohydrolase (*Ct*CBH5A) and β-glucosidase (*Ct*GH1) in the mixture for saccharification of *Sorghum durra* stalk by using the Stat-Ease Design Expert software (Stat-Ease Inc., Minneapolis, MS, USA, Version 6.0.4). The optimized enzyme cocktail ratio for acid pretreated *Sorghum durra* stalk was *Ba*GH5-UV2 (43.1%), *Ct*CBH5A (10.0%) and *Ct*GH1 (46.9%), which was different from the base pretreated *Sorghum durra* stalk, that was *Ba*GH5-UV2 (80%), *Ct*CBH5A (10%) and *Ct*GH1 (10%) at a total crude enzyme loading of 2000 U/g of pretreated biomass. This difference in the enzyme ratios was due to the difference in their carbohydrate composition after the acid or base pretreatment. The synergism among the cellulase enzymes were higher in acid pretreated *Sorghum durra* stalk than base pretreated *Sorghum durra* stalk which might also depends on the carbohydrate composition of pretreated *Sorghum durra* stalk. Thus, the ratio of cellulases in cocktail and synergism among them depends on the type of pretreatment method. The optimized enzyme ratio for saccharification of acid pretreated *Sorghum durra* stalk gave maximum TRS and glucose yield, 113 mg/g and 85 mg/g, respectively, at 72 h whereas base pretreated *Sorghum durra* stalk gave 211 mg/g and 174 mg/g, respectively, of pretreated *Sorghum durra* stalk at 96 h, respectively. This is the first comprehensive report on the pretreatment specific optimization of cellulase enzymes ratio in a mixture for enzymatic saccharification by using the RSM mixture design approach on substrate *Sorghum durra* stalk. Therefore, a rational enzyme mixture design by using the synergistic concept and statistical

TH-2393_146106034

analysis can help in improving the enzymatic saccharification yield from different pretreatment methods for a biomass.

TH-2393_146106034

6.5 References

Banerjee, G., Car, S., Scott-Craig, J. S., Borrusch, M. S., and Walton, J. D. (2010). Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. *Biotechnology for Biofuels*, 3(1), 22.

Bansal, P., Hall, M., Realff, M.J., Lee, J.H., and Bommarius, A.S. (2009). Modeling cellulase kinetics on lignocellulosic substrates. *Biotechnology Advances*, 27(6), 833-848.

Bussamra, B. C., Freitas, S., and da Costa, A. C. (2015). Improvement on sugar cane bagasse hydrolysis using enzymatic mixture designed cocktail. *Bioresource Technology*, 187, 173-181.

Galbe, M., and Zacchi, G. (2002). A review of the production of ethanol from softwood. *Applied Microbiology and Biotechnology*, 59(6), 618-628.

Gao, D., Chundawat, S.P., Krishnan, C., Balan, V., and Dale, B.E. (2010). Mixture optimization of six core glycosyl hydrolases for maximizing saccharification of ammonia fiber expansion (AFEX) pretreated corn stover. *Bioresource Technology*, 101(8), 2770-2781.

Hess, M. (2008). Thermoacidophilic proteins for biofuel production. *Trends in Microbiology*, *16*(9), 414-419.

Jamaldheen, S. B., Sharma, K., Rani, A., Moholkar, V. S., and Goyal, A. (2018). Comparative analysis of pretreatment methods on sorghum (*Sorghum durra*) stalk agrowaste for holocellulose content. *Preparative Biochemistry and Biotechnology*, 48(6), 457-464.

TH-2393_146106034

Kim, I. J., Jung, J. Y., Lee, H. J., Park, H. S., Jung, Y. H., Park, K., and Kim, K. H. (2015). Customized optimization of cellulase mixtures for differently pretreated rice straw. *Bioprocess and Biosystems Engineering*, 38(5), 929-937.

Kim, I. J., Ko, H. J., Kim, T. W., Nam, K. H., Choi, I. G., and Kim, K. H. (2013). Binding characteristics of a bacterial expansin (BsEXLX1) for various types of pretreated lignocellulose. *Applied Microbiology and Biotechnology*, 97(12), 5381-5388.

Kim, I. J., Lee, H. J., and Kim, K. H. (2017). Pure enzyme cocktails tailored for the saccharification of sugarcane bagasse pretreated by using different methods. *Process Biochemistry*, 57, 167-174.

Kim, J. S., Lee, Y. Y., and Kim, T. H. (2016). A review on alkaline pretreatment technology for bioconversion of lignocellulosic biomass. *Bioresource Technology*, 199, 42-48.

Levine, S. E., Fox, J. M., Clark, D. S., and Blanch, H. W. (2011). A mechanistic model for rational design of optimal cellulase mixtures. *Biotechnology and Bioengineering*, 108(11), 2561-2570.

Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. (1951). Protein measurement with the Folin phenol reagent. *The Journal of Biological Chemistry*, 193(1), 265–275.

Marcos, M., García-Cubero, M. T., González-Benito, G., Coca, M., Bolado, S. and Lucas, S. (2013) Optimization of the enzymatic hydrolysis conditions of steam-exploded wheat straw for maximum glucose and xylose recovery. *Journal of Chemical Technology and Biotechnology*, 88, 237-246.

TH-2393_146106034

Nath, P., Dhillon, A., Kumar, K., Sharma, K., Jamaldheen, S. B., Moholkar, V. S., and Goyal, A. (2019). Development of bi-functional chimeric enzyme (*Ct*GH1-L1-*Ct*GH5-F194A) from endoglucanase (*Ct*GH5) mutant F194A and β-1, 4-glucosidase (*Ct*GH1) from *Clostridium thermocellum* with enhanced activity and structural integrity. *Bioresource technology*, 282, 494-501.

Nelson, N. (1944). A photometric adaptation of the Somogyi method for the determination of glucose. *Journal of Biological Chemistry*, 153(2), 375-380.

Raabo, B. E., and Terkildsen, T. C. (1960). On the enzymatic determination of blood glucose. *Scandinavian Journal of Clinical and Laboratory Investigation*, 12(4), 402-407.

Sharma, K., Thakur, A., Kumar, R., and Goyal, A. (2019). Structure and biochemical characterization of glucose tolerant β-1, 4 glucosidase (*Ht*Bgl) of family 1 glycoside hydrolase from *Hungateiclostridium thermocellum*. *Carbohydrate Research*, 483, 107750.

Singh, S., Khanna, S., Moholkar, V.S., and Goyal, A. (2014). Screening and optimization of pretreatments for *Parthenium hysterophorus* as feedstock for alcoholic biofuels. *Applied Energy*, 129, 195-206.

Somogyi, M. A., (1945). New reagent for the determination of sugars. *Journal of Biological Chemistry* 160, 61–68.

Zhou, J., Wang, Y.H., Chu, J., Luo, L.Z., Zhuang, Y.P., and Zhang, S.L. (2009). Optimization of cellulase mixture for efficient hydrolysis of steam-exploded corn stover by statistically designed experiments. *Bioresource Technology*, 100(2), 819-825.

TH-2393_146106034

Published/accepted (*From Thesis*):

1. **Shweta Singh**, Arun Dhillon and Arun Goyal (2020). Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis. Renewable Energy. 151, 1124-1133. **(JIF 5.439)**
2. **Shweta Singh**, Krishan Kumar, Priyanka Nath and Arun Goyal (2020). Role of glycine 256 residue in improving the catalytic efficiency of mutant endoglucanase of family 5 glycoside hydrolase from *Bacillus amyloliquefaciens* SS35. Biotechnology and Bioengineering. https://doi.org/10.1002/bit.27448. (**JIF 4.26)**
3. **Shweta Singh**, Vikky Rajulapati, Sumitha Banu Jamaldheen, Vijayanand Suryakant Moholkar and Arun Goyal (2020). Statistically designed cellulase mixture for saccharification of pretreated *Sorghum durra* stalk. Industrial Crops and Products. https://doi.org/10.1016/j.indcrop.2020.112678. **(JIF 4.19)**

Other Publications:

1. Ajit Kumar, **Shweta Singh**, Vikky Rajulapati and Arun Goyal (2020). Evaluation of pre-treatment methods for *Lantana camara* stem for enhanced enzymatic saccharification. 3 Biotech. 10, 37. **(JIF 1.786)**
2. Mohanapriya Nedumaran[†], **Shweta Singh[†],** Sumitha Banu Jamaldheen, Priyanka Nath, Vijayanand Suryakant Moholkar and Arun Goyal (2020). Assessment of combination of pretreatment of *Sorghum durra* stalk and production of chimeric enzyme (β-glucosidase and endo β-1,4 glucanase, *Ct*GH1-L1-*Ct*GH5-F194A) and cellobiohydrolase (*Ct*CBH5A) for saccharification to produce bioethanol. Preparative Biochemistry and Biotechnology. https://doi.org/10.1080/10826068.2020.1762214. **(JIF 1.1) [†]Both authors contributed equally.**
3. Priyanka Nath[†], Premeshwori Devi Maibam[†], **Shweta Singh**, Vikky Rajulapati and Arun Goyal (2020). Sequential pretreatment using alkali and organosolv for improving delignification of sugarcane bagasse and cellulose conversion and saccharification by chimera and cellobiohydrolase for bioethanol production. (Submitted). [†]**Both authors contributed equally.**

Book Chapters

1. **Shweta Singh**, Arabinda Ghosh and Arun Goyal (2017). Chapter 12, Manno-oligosaccharides as prebiotic valued products from agro-waste, in Biosynthetic Technology and Environmental Challenges: Energy, Environment, and Sustainability. Springer Book Series by Springer Nature. Pp 205-221. Eds Sunita J. Varjani, Binod Parameswaran, Sunil Kumar and Sunil K. Khare. DOI: 10.1007/978-981-10-7434-9.

TH-2393_146106034

Conferences

Oral Presentation:

From thesis

1. **Shweta Singh**, Priyanka Nath, Krishan Kumar and Arun Goyal (2019). Analysis of mechanism for enhanced catalytic efficiency of CMCase from *Bacillus amyloliquefaciens* SS35 UV2 mutant strain. 8th International Forum on Industrial Bioprocessing (IBA-IFIBiop 2019) "Bridging Sustainability and Industrial Revolution through Green Bioprocessing", 1-5 May, 2019, Imperial Hotel, Miri, Sarawak, Malaysia.

Other

2. Mohanapriya N., **Shweta Singh**, Priyanka Nath, Sumitha Banu and Arun Goyal (2019). Saccharification of *Sorghum durra* by chimeric enzyme (β-glucosidase and endo β-1,4 glucanase, *Ct*GH1-L1-*Ct*GH5-F194A) and cellobiohydrolase (*Ct*CBH5A) from *Clostridium thermocellum* for bioethanol production. 8th International Forum on Industrial Bioprocessing (IBA-IFIBiop 2019) "Bridging Sustainability and Industrial Revolution through Green Bioprocessing", 1-5 May, 2019, Imperial Hotel, Miri, Sarawak, Malaysia.

Poster presentations:

From thesis

1. **Shweta Singh**, Priyanka Nath, Krishan Kumar and Arun Goyal (2019). Mutation of aspartate 256 to glycine enhanced the catalytic efficiency of CMCase from *Bacillus amyloliquefaciens* SS35 UV2 mutant strain. Research Conclave, March 14-17, 2019, IIT Guwahati, Assam.
2. **Shweta Singh**, Arun Dhillon and Arun Goyal (2018). Ultraviolet irradiation of *Bacillus amyloliquefaciens* SS35 for producing hyperactive mutant strain for improved cellulase activity. International Conference on Biotechnological Research and Innovation for Sustainable Development, 15th BRSI convention. CSIR- Indian Institute of Chemical Technology (CSIR-IICT), Nov. 22-25, 2018, Hyderabad, India. **(Best Poster Award)**
3. **Shweta Singh**, Arun Dhillon and Arun Goyal (2018). Strain improvement of *Bacillus amyloliquefaciens* SS35 by ultraviolet radiation for producing hyperactive mutant strain for improved endoglucanase activity. DBT National Workshop on Bioenergy, July 6-7, 2018, IIT Roorkee, Uttarakhand, India.
4. **Shweta Singh** and Arun Goyal (2018). Strain improvement of *Bacillus amyloliquefaciens* SS35 by UV and chemical mutagenesis for producing hyperactive mutant strain for improved β-glucanase and xylanase activities. International Conference on Sustainable Biofuels. February 26-27th, 2018, ICGEB, New Delhi, India.
5. **Shweta Singh**, Arun Dhillon and Arun Goyal (2018). Cloning of wild-type endoglucanase (*Ba*GH5) from *Bacillus amyloliquefaciens* SS35 and its mutant enzyme *Ba*GH5-UV2 from its UV mutant strain and mutant enzyme *B*aGH5-EMS7

TH-2393_146106034

from UV/EMS mutant strain and analysis of induced mutations in the genes. Indo-Japan Bilateral Symposium on Future Perspective of Bioresource Utilization in North-Eastern Region, February 1- 4, 2018, IIT Guwahati.

6. **Shweta Singh** and Arun Goyal (2017). Strain improvement of *Bacillus amyloliquefaciens* SS35 by UV and chemical mutagenesis for producing hyperactive mutant strain for improved β-glucanase and xylanase activities. 2nd International Conference on Sustainable Energy and Environmental Challenges (SEEC-2018). Dec 31 2017- Jan 3, 2018, IISc Bangalore, Bengaluru.
7. **Shweta Singh**, Abhijeet Thakur and Arun Goyal (2016). Strain improvement of *Bacillus amyloliquefaciens* SS35 by UV mutagenesis for producing hyperactive mutants for improved carboxymethyl celluase activity. 57th International Annual Conference of the Association of Microbiologists of India (AMI-2016), Nov 24-27, 2016, Gauhati University and IASST, Guwahati, Assam India.

Others

8. **Shweta Singh** and Arun Goyal (2015). Isolation of bacterial strain efficiently hydrolyzing the cellulosic substrates. 56th International Annual Conference of Association of Microbiologists of India (AMI), December 7-10, 2015, JNU, New Delhi.
9. Ajit Kumar, **Shweta Singh,** Vikky Rajulapati, Arun Goyal (2018). Optimization of pretreatment of *Lantana camara* stem as lignocellulosic biomass for bioethanol. Indo-Japan Bilateral Symposium on Future Perspective of Bioresource Utilization in North-Eastern Region, February 1- 4, 2018, IIT Guwahati.
10. Ashutosh Gupta, **Shweta Singh**, Debasish Das and Arun Goyal (2015). Saccharification of pretreated Napier grass by recombinant cellulase and hemicellulase from *Clostridium thermocellum* for bioethanol production. 56th International Annual Conference of Association of Microbiologists of India (AMI), December 7-10, 2015, JNU Delhi.

Awards:

1. **Best poster award** on the work entitled "Ultraviolet irradiation of *Bacillus amyloliquefaciens* SS35 for producing hyperactive mutant strain for improved cellulase activity*"* in "15th BRSI convention. CSIR- Indian Institute of Chemical Technology (CSIR-IICT), Nov. 22-25, 2018, Hyderabad, India.

TH-2393_146106034

TH-2393_146106034

VITAE

The author was born on April 25, 1991 in Aligarh, (Uttar Pradesh). She passed the Secondary Examination (10th Class) conducted by Board of Indian Certificate of Secondary Education, India, from Three Dots Senior Secondary School Aligarh in 2006 and Higher Secondary Examination (12th Class) conducted by Central Board of Secondary Education, India from Kendriya Vidyalaya Aligarh in 2008. She completed B.Sc. Hons (Botany) from Aligarh Muslim University, Aligarh (Uttar Pradesh) in 2012. She completed M.Sc. (Agriculture Microbiology) from Aligarh Muslim University, Aligarh (Uttar Pradesh) in 2014. Mrs. Shweta Singh joined the Ph.D. program in December, 2014 at Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India. She received Institute Fellowship (IIT Guwahati), for whole Ph.D. duration under the scheme run by the Ministry of Human Resource and Development (MHRD), New Delhi. She delivered the open seminar (Ph.D. Synopsis) on November 21, 2019 and presented her thesis work before the Doctoral Committee and her performance was satisfactory. She submitted the Ph.D. thesis in January 2020.

TH-2393_146106034

TH-2393_146106034

TH-2393_146106034

Renewable Energy xxx (xxxx) xxx

Contents lists available at ScienceDirect

Renewable Energy

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

Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis

Shweta Singh [a, b], Arun Dhillon [a], Arun Goyal [a, b, *]

[a] Carbohydrate Enzyme Biotechnology Laboratory, Department of Biosciences and Bioengineering, Guwahati, 781039, Assam, India
[b] DBT PAN-IIT Center for Bioenergy, Indian Institute of Technology Guwahati, Guwahati, 781039, Assam, India

ARTICLE INFO

Article history:
Received 1 March 2019
Received in revised form
31 August 2019
Accepted 18 November 2019
Available online xxx

Keywords:
Carboxymethylcellulase
Ultraviolet directed evolution
Catalytic efficiency
pH tolerance
Saccharification
Point mutation

ABSTRACT

Bacillus amyloliquefaciens SS35 was subjected to ultraviolet irradiation to improve the enzymatic hydrolysis of lignocellulosic biomass. The resulting mutant, UV2, produced endoglucanase, carboxymethyl cellulase, CMCase-UV2 with 1.6–4.1-fold higher activity against cellulosic substrates than the wild type, CMCase-WT. CMCase-UV2 exhibited wider pH stability in the acidic range than CMCase-WT. The TLC analysis showed that the hydrolysis of CMC-Na and β-glucan by CMCase-UV2 produced glucose along with cello-oligosaccharides and cellobiose in 45 min, whereas, CMCase-WT produced, only cello-oligosaccharides and cellobiose in 120 min by endolytic mode of action. The hydrolysis of pretreated *Pennisetum purpureum* by CMCase-UV2 gave total reducing sugar yield 154.2 mg/g pretreated biomass in 48 h, which was 1.8-fold higher than CMCase-WT. CMCase-UV2 was the promising endoglucanase which improves the saccharification of lignocellulosic biomass therefore, it will improve the efficiency of the process, lignocellulose-based biorefineries for bioethanol production. The gene encoding cellulase was amplified from wild-type and UV2 strains using degenerate primers designed from phylogenetically related spp. *Bacillus amyloliquefaciens* KHG19 for family 5 glycoside hydrolase. Sequences analysis of genes from wild-type and UV2 strains showed the mutation, D233G. These results will provide information for protein engineering in designing mutant of endoglucanase for improved catalytic efficiency and pH stability.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

The ambiguity in world petroleum supplies, concerns about environmental pollution and climate change have generated interest in producing sustainable green biofuels from bioresources such as lignocellulosic biomass. Lignocellulosic biomass is composed of cellulose the most abundant renewable carbon source on the earth's surface [1]. Thus, lignocellulosic biomass has the potential to serve the demand of worldwide renewable transportation fuel [2]. However, commercialization of the cellulosic fuels faces several techno-economic challenges in order to compete with fossil fuels [3]. A major bottleneck in the biofuel industry is the hydrolysis of plant cell wall polysaccharide, especially the highly recalcitrant cellulose fibers into fermentable sugars [4]. Cellulases are the group of enzymes that are well known for bioconversion of lignocellulosic waste into fermentable sugars [5]. They are grouped into three categories: (i) endoglucanase (EC 3.2.1.4) which hydrolyzes β-1,4 bonds of cellulose, (ii) cellobiohydrolase (EC 3.2.1.91) which cleaves the reducing and non-reducing ends of cellulose fibre to release cellobiose and (iii) β-glucosidase (EC 3.2.1.21) which converts cellobiose to glucose [6]. Saccharification of cellulose into monomeric sugars involves synergistic action of these three cellulase enzymes. The aerobic mesophilic bacteria belonging to the genus *Bacillus* are efficient cellulase producers such as *Bacillus* KSM–S237 [7], *Bacillus subtilis* [8], *Bacillus megaterium* [9] and *Bacillus amyloliquefaciens* SS35 [10]. The use of cellulases on an industrial scale for bioethanol production faces challenges because of its high production cost, low enzyme yield and low enzyme activities [11]. One way to reduce the production cost of bioethanol is the strain improvement of cellulase producing bacteria for improved cellulase characteristics such as catalytic efficiency and wide range pH stability. This may be achieved by random mutagenesis using ultraviolet (UV) irradiation, which helps in the UV directed evolution of the natural strain [12]. Directed evolution by

* Corresponding author. Carbohydrate Enzyme Biotechnology Laboratory, Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati, 781039, Assam, India.
E-mail address: arungoyl@iitg.ac.in (A. Goyal).

https://doi.org/10.1016/j.renene.2019.11.105
0960-1481/© 2019 Elsevier Ltd. All rights reserved.

TH-2393_146106034

Please cite this article as: S. Singh et al., Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis, Renewable Energy, https://doi.org/10.1016/j.renene.2019.11.105

many biotechnological methods is a process of forming highly efficient biocatalyst which is used by industry for forming eco-friendly products. There are several reports, on the strain improvement of fungal strain such as *Trichoderma viride* [13], *Trichoderma reesei* [14], *Aspergillus nidulans* [15] and *Streptomyces griseoaurantiacus* [16] by UV irradiation for producing hyperactive mutants. However, bacterial cellulases are getting much more attention than fungal cellulases because, first of all, the production cost of fungal cellulase is higher than bacterial cellulase and therefore bacterial cellulases are getting more attention than fungal cellulases [17,18]. Secondly, because the growth rate of bacteria is much higher than fungi, it attains high cell density within short time to produce enzyme. Thirdly, certain bacterial cellulases are expressed in multiple complexes which increase the performance of overall saccharification as they function synergy. Fourthly, bacteria can inhabit a wide diversity of industrial and ecological niches, because of which they are awfully resistant to environmental stresses. Fifthly, the expression system and genetic manipulation of bacteria is more convenient in order to achieve the high-level enzyme. Their recombinant enzymes can be hyper-expressed in *E. coli* cells and enhanced cellulase production can be achieved as compared with fungal cellulases (which have glycosylated proteins with introns) [18]. There is no report published on strain improvement of *Bacillus* species for improved endoglucanase characteristics by UV irradiation. The molecular mechanism for improved efficiency of cellulolytic enzymes from natural strain was understood by amplification of gene encoding cellulase from bacteria by designing degenerate primers from phylogenetically related species and sequencing [19]. Degenerate primer based PCR approach was used to identify the target gene. In the present study, the wild-type strain of *Bacillus amyloliquefaciens* SS35 [20] was improved by random mutagenesis using UV irradiation. The biochemical characterization of the mutant enzyme from the highly efficient strain for catalytic efficiency and pH tolerance was carried out. The change in the amino acid residue(s) in the mutant strain was identified by the degenerate primer based approach, for amplifying and then cloning the gene encoding cellulase. The catalytic efficiency of mutant cellulase (CMCase) enzyme was increased for cellulosic substrates which support the green and eco-friendly bioconversion of lignocelluloses into reducing sugars.

2. Experimental section

2.1. Media and chemicals

Carboxymethylcellulose sodium salt (CMC-Na) (low viscosity), lichenan, β-glucan, avicel, birchwood xylan, oat spelt xylan and galactomannan were procured from Sigma Aldrich, USA. Superdex™ 75 pg size exclusion column was purchased from General Electronics Healthcare Biosciences, Sweden. All medium components, chemicals, cellulose powder and Bacterial DNA miniprep purification spin kit used in this study were procured from HiMedia Pvt. Ltd., India. Phusion High-Fidelity DNA polymerase was procured from ThermoFisher Scientific (EU, Lithuania). PCR-grade water (Sigma Aldrich, USA), dNTPs were procured from Biobharti Life Science Pvt. Ltd., India. Deoxynucleotide primers were procured from GCC Biotech Pvt. Ltd., India and NZYEasy cloning kit was procured from NZYTech, Portugal.

2.2. Microorganism and culture conditions

Bacillus amyloliquefaciens SS35 isolated from Rhinoceros dung with gyrase A GenBank accession no. KF019284 and 16S rDNA GenBank accession no. JX674030 [20] was used in this study. The medium composition for CMCase production from *Bacillus amyloliquefaciens* SS35 was CMC-Na (19.0 g/L), peptone (2 g/L), yeast extract (8 g/L), K_2HPO_4 (1 g/L), NaCl (1 g/L) and $MgSO_4.7H_2O$ (0.2 g/L) at pH 5.6.7% (v/v) inoculum of cell optical density 1.0 at 600 nm was inoculated in 200 mL medium contained in 500 mL Erlenmeyer flask and incubated at 40 °C in an orbital shaking speed of 120 rpm [21].

2.3. UV mutagenesis of Bacillus amyloliquefaciens SS35 and screening of mutants

Overnight grown culture of the bacterium (cell optical density ~1.0) was spread plated onto a medium supplemented with 1.9% (w/v) CMC-Na [21] from 10^{-1} to 10^{-6} serial dilution under sterile conditions. The mutation was induced by exposing the petri dishes of 10^{-4} serial dilution (with around 100 number of colonies) to UV light for 5 min−180 min using direct plate irradiation method [22] in a UV crosslinker (Hoefer, UVC 500) at 0.6 J/cm^2 frequency. After irradiation petri dishes were immediately wrapped by aluminium foil to avoid photoreactivation by SOS mechanism and incubated at 40 °C for 12 h. Mutant colonies were qualitatively screened for cellulase production by staining the replica plate with 0.3% (w/v) Congo red by method as described earlier [20]. The mutant colonies forming the zone of clearance were further, screened quantitatively by analyzing the endoglucanase activity (U/mL) of extracellular CMCase enzyme in cell-free culture broth containing the crude enzyme. The enzyme activity was estimated by the Nelson and Somogyi method described earlier [23,24] using glucose standard. The enzyme assay was carried out by incubating the enzyme with 1% (w/v) CMC-Na in 50 mM sodium acetate buffer, pH 5.0 at 65 °C for 5 min. The reaction mixture contained 10 μl of the enzyme in 100 μl reaction volume. One unit (U) of CMCase activity was defined as the amount of enzyme required to release one μmol of reducing sugar (glucose) per min under mentioned conditions. The protein concentration for the crude enzyme(s) were estimated by the Lowry method using bovine serum albumin (BSA) as standard [25].

2.4. Fermentation profile of wild-type Bacillus amyloliquefaciens SS35 and mutant UV2

The fermentation profiles of wild-type *Bacillus amyloliquefaciens* SS35 and its mutant UV2 were studied by growing in optimized medium [21] at 40 °C and 120 rpm for 96 h. The endoglucanase activity, cell optical density (600 nm), protein concentration and pH change of the broth were monitored at every 6 h up to 96 h by taking 0.5 mL aliquot.

2.5. Purification of endoglucanase enzyme from wild-type and UV2 mutant strains

2.5.1. Partial purification by ammonium sulfate precipitation method

The extracellular CMCase enzyme produced from wild-type strain (CMCase-WT) and from UV2 mutant (CMCase-UV2) present in cell-free supernatant was partially purified by ammonium sulfate by method described earlier [26,27]. Ammonium sulfate up to 90% saturation was gradually added in 200 mL cell-free supernatant and incubated in stirring condition at 4 °C for 12 h. The precipitated protein was separated from the supernatant by centrifugation at 8,000g for 30 min at 4 °C. The precipitated pellet of protein was dissolved in 5 mL of 50 mM sodium acetate buffer, pH 5.0. The desalting of the precipitated enzyme was carried out by using 50 mM sodium acetate buffer (pH 5.0) for 48 h using a 10 kDa cutoff dialysis membrane. Further, the dialysed enzymes were concentrated to 2 mL volume using 15 mL, 10 kDa MWCO concentrators (Amicon Ultra, Merk-Millipore, USA). The protein

TH-2393_146106034

Please cite this article as: S. Singh et al., Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis, Renewable Energy, https://doi.org/10.1016/j.renene.2019.11.105

concentration of partially purified enzyme(s) were estimated by the Bradford method using BSA as standard [28].

2.5.2. Size exclusion chromatography

The partially purified CMCase-WT and CMCase-UV2 enzymes were further purified by using Fast Protein Liquid Chromatography (GE Healthcare). The Superdex™ 75 pg column (GE Healthcare, 16 mm column ID, 600 mm bed height) was used for protein purification by size exclusion chromatography. The Superdex column was pre-equilibrated with 100 mL of 50 mM sodium acetate buffer, pH 5.0. After that 2 mL (1 mg/mL) of the enzyme(s) was loaded on to the column. The enzyme was eluted by using the same buffer at a flow rate of 0.8 mL/min with 1 mL fraction size. The fractions containing CMCase activity (U/mL) were pooled. The protein concentration of purified enzyme was estimated by the Bradford method using BSA as standard [28].

2.6. SDS-PAGE and zymogram analysis of purified enzymes

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) using 12% (w/v) gel was carried out under denaturing conditions according to the protocol as described earlier to identify the purity and molecular mass of proteins [29]. 0.3% (w/v) CMC-Na was added in the resolving gel for activity staining of CMCase by zymogram. After the electrophoresis the gel was cut into two parts, one part was subjected to silver staining protocol as described earlier [30] and other part was used for activity staining by zymogram. For zymogram analysis, the gel was renatured in 2.5% (v/v) Triton X-100 in 50 mM sodium acetate buffer (pH 5.0) up to 4 h [31]. After removal of Triton X-100, the gel was washed with 50 mM sodium acetate buffer (pH 5.0) and incubated at 50 °C for 6 h. Thereafter, the gel was stained by 0.3% (w/v) Congo red for 15 min, and destained further with 1M NaCl.

2.7. Effect of pH and temperature on CMCase-WT and CMCase-UV2 activity and stability

The optimum pH of the purified CMCase-WT and CMCase-UV2 enzymes were determined with 3 μl of each enzyme (0.1 mg/mL) in 1% (w/v) CMC-Na in appropriate buffers; 50 mM sodium acetate (pH 3.5–5.5), 50 mM sodium phosphate buffer (pH 5.8–8.0) and 50 mM Tris-HCl (pH 8–9.5). The reaction mixture at different pH were incubated at 65 °C for 5 min and CMCase activity was assayed by reducing sugar estimation Nelson and Somogyi method as described earlier [23,24]. The comparison of pH stability of the purified CMCase-WT and CMCase-UV2 enzymes was performed by pre-incubating the enzyme (10 μl of each enzyme, 0.1 mg/mL) at 30 °C for 1 h under different buffers at a pH ranging from 3.5 to 9.5. The residual activity for respective samples was then estimated.

The optimum temperature of the purified CMCase-WT and CMCase-UV2 for enzyme assay was determined. The reaction mixture containing 3 μl of each enzyme (0.1 mg/mL) at 1% (w/v) CMC-Na in 50 mM sodium acetate buffer, pH 5.0 was incubated at the temperature in range, 30°C–75 °C for 5 min. The thermostability studies of the purified CMCase-WT and CMCase-UV2 was performed by pre-incubating in 50 mM sodium acetate buffer, pH 5.0 at different temperatures ranging from 30 to 75 °C for 1 h. The residual activity for respective samples was determined.

2.8. Comparison of enzyme activity of purified CMCase-WT and CMCase-UV2 against soluble and insoluble substrates

The enzyme activity of purified CMCase-WT and CMCase-UV2 was tested against substrates, β-D-glucan, lichenan, CMC-Na, cellulose powder, avicel, birchwood xylan and galactomannan. The reactions were carried out at 1.0% (w/v) substrate dissolved in 50 mM sodium acetate buffer, pH 5.0 by incubating at 65 °C for 5 min 100 μl of enzyme reaction mixture contained 3 μl of the enzyme (CMCase-WT, 0.1 mg/mL and CMCase-UV2, 0.1 mg/mL). In the case of avicel and oat spelt xylan, the reaction mixture was incubated under shaking conditions (160 rpm, 5 min). The activity of CMCase-WT and CMCase-UV2 was also compared against the pretreated substrate *Parthenium hysterophorus* (Carrot grass) and *Pennisetum purpureum* (Elephant grass). The Carrot grass and Elephant grass biomasses were pretreated by the method described by Singh et al. [32] and Eliana et al. [33], respectively. The reactions were carried out with 2.0% (w/v) substrate dissolved in 50 mM sodium acetate buffer (pH 5.0) by incubating at 65 °C for 15 min at 160 rpm. The resulting reducing sugars were determined by Nelson and Somogyi method reported earlier [23,24] by measuring the absorbance at 500 nm, on a spectrophotometer (Varian, Cary 100 Bio) using glucose, xylose and mannose as standards.

2.9. Comparison of kinetic parameters of CMCase-WT and CMCase-UV2

The kinetic parameters i.e. K_m and V_{max} of CMCase-WT and CMCase-UV2 were determined by fitting the initial rate data to the Michaelis–Menten equation by Lineweaver–Burk double reciprocal plot [34,35]. The activity of enzymes was assayed in the substrate (CMC-Na and β-D-glucan) concentration range from 0.01% to 2% (w/v) in 50 mM sodium acetate buffer (pH 5.0) at 65 °C for 5 min. The data for kinetic parameters were analysed by GraphPad Prism software (GraphPad Software Inc., San 138 Diego, CA) and the kinetic constants were calculated from the best fit.

2.10. Thin Layer Chromatography analysis of hydrolysed products of CMC-Na and β-glucan by CMCase-WT and CMCase-UV2

The qualitative analysis of hydrolysed products of CMC-Na and β-D-glucan by CMCase-WT and CMCase-UV2 was performed by Thin Layer Chromatography (TLC) on silica gel-coated aluminium plate (TLC Silica gel 60 F254, 20620 cm, Merck). The enzyme (2.5 mM) catalysed reaction with 1% (w/v) CMC-Na and β-D-glucan in 100 μl reaction mixture was incubated at 50 °C in 50 mM sodium acetate buffer (pH 5.0), for time intervals of 5, 10, 15, 30, 45, 60, 120, 180, 360, 720 and 1440 min. To find out the action of both the enzymes on cellobiose, CMCase-WT and CMCase-UV2 enzyme with equimolar concentrations of 2.5 mM were incubated in 100 μl reaction mixture containing 0.5% (w/v) cellobiose dissolved in 50 mM sodium acetate buffer, pH5.0 at 50 °C for 2 h and reaction products were analysed by TLC as mentioned above. The enzyme reaction(s) were stopped by adding 200 μl of absolute ethanol and the mixture(s) were centrifuged at 13000g for 5 min. The supernatant was transferred to another 1.5 mL micro-centrifuge tube and concentrated by evaporating the absolute ethanol in hot air oven at 70 °C for 12 h. Then 1 μl of each reaction mixture for CMC-Na, 0.5 μl of each reaction mixture for β-glucan, 1 μl of each reaction mixture for cellobiose and standard (D-glucose and cellobiose) solutions (1.0 mg/mL) were loaded on the TLC plates. The plates were dried in hot air oven at 70 °C for 5 min and kept in the developing chamber saturated with the developing solution consisting of n-butanol-acetic-acid-water (2:1:1, v/v) at 25 °C [36]. After the run, the released sugars were visualized by staining the TLC plates with a visualizing solution (sulphuric acid/methanol 5:95, v/v; α-naphthol 0.5%, w/v). Further, the TLC plates were dried in hot air oven at 70 °C until the migrated sugars were visualized as spots on the TLC plate.

TH-2393_146106034

Please cite this article as: S. Singh et al., Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis, Renewable Energy, https://doi.org/10.1016/j.renene.2019.11.105

2.11. Hydrolysis of Pennisetum purpureum by CMCase-WT and CMCase-UV2

The potential of CMCase-WT and CMCase-UV2 enzymes in the hydrolysis of *Pennisetum purpureum* (Elephant grass) pretreated by a physicochemical method (1%, v/v NaOH+20 min autoclaving) as described earlier by Eliana et al. [33] was evaluated. The reactions with CMCase-WT and CMCase-UV2 enzymes were carried out with 1.0% (w/v) pretreated substrate dissolved in 50 mM sodium acetate buffer, pH 5.0 by incubating at 30 °C under shaking condition at 160 rpm for 24 h and 48 h. The enzymatic hydrolysis was carried out at 30 °C instead of 65 °C, assuming its application in simultaneous saccharification and fermentation (SSF). The 1 mL reaction mixture contained equimolar protein concentration (2.5 mM) of partially purified CMCase-WT and CMCase-UV2, respectively. The partially purified CMCase-WT and CMCase-UV2 was used for making the enzymatic hydrolysis process more cost-effective. The released reducing sugars were quantified by using the Nelson and Somogyi method described earlier [23,24]. The qualitative analysis of hydrolysed products of pretreated Elephant grass by CMCase-WT and CMCase-UV2 was performed by TLC by method describe above. The 1 mL reaction mixtures were stopped by adding 2 mL of absolute ethanol and centrifuged at 13000 g for 5 min. The supernatant was transferred to another micro-centrifuge tube and concentrated by evaporating the absolute ethanol in hot air oven. Then 0.2 μl of each reaction mixture of 24 h and 1 μL of each reaction mixture of 48 h along with standard (D-glucose and cellobiose) solutions (1.0 mg/mL) were loaded on the TLC plates and sugar spots were developed as method described above.

2.12. Identification of induced mutations in UV2 strain at genetic level

The molecular screening of mutation(s) induced in the gene encoding CMCase produced from wild-type and UV2 strains of *Bacillus amyloliquefaciens* SS35, was done by using the degenerate oligonucleotide primers. The genomic DNA from bacterial strains was isolated by using bacterial DNA miniprep purification spin kit. Polymerase chain reaction (PCR) using degenerate primers was used to identify the family 5 homologues in *Bacillus amyloliquefaciens* SS35. The genes encoding CMCase-WT (*Ba*GH5-WT) and CMCase-UV2 (*Ba*GH5-UV2) were amplified from respective genomic DNA (25 ng/μl) using 0.02 U/μl of Phusion High-Fidelity DNA polymerase. A 60 μl PCR reaction mixture contained dNTPs (1.5 mM), forward and reverse primers (0.5 μM) and PCR-grade water and DMSO (3% v/v). The primers used for amplifying *Ba*GH5 were, forward primer (5′-TCAGCAAGGGCTGAGGATGAAACGGGCAATTTCTATTTT-3′) and reverse primer (5′-TCAGCGGAAGCTGAGGTAACTAATTGGGTTCTGTTCCC-3′). The PCR cycles were: initial denaturation at 98 °C for 3 min followed by 30 cycles of i) denaturation at 98 °C for 20 s, ii) annealing at 58 °C for 30 s and iii) extension at 72 °C for 1 min followed by a final extension at 72 °C for 10 min. The PCR amplified DNA of *Ba*GH5-WT and *Ba*GH5-UV2 were ligated to pHTP0 cloning vector using NZYEasy cloning kit. The recombinant plasmids DNA were transformed into *E. coli* TOP10 cells. The positive clones were confirmed by colony PCR. The plasmids DNA of both strains were sequenced for identification of sites of mutation in the gene.

3. Results and discussion

3.1. UV mutagenesis and screening of mutants

Effect of UV exposure on wild-type strain in UV crosslinker for different time interval was evaluated. Exposure of wild-type strain at frequency 0.6 J/cm^2 for 180 min was taken sub-lethal that resulted in 99% mortality of bacterium. It was earlier reported that for the strain improvement in order to produce cellulase with improved activity, the wild-type strain was exposed to mutagen at the sub-lethal dose [13]. Fourteen UV mutant strains, UV1 to UV14 showed the hydrolyzing zone of clearance after qualitative screening by staining with 0.3% (w/v) Congo red (Supplementry data, Fig. S1). Further, initial quantitative screening of UV mutant strains showed that the mutant strains UV1, UV2 and UV4 displayed 0.76 U/ml, 0.79 U/ml and 0.74 U/ml, respectively, CMCase activity for crude enzyme, which was 17%, 22% and 14%, respectively, higher than the wild-type strain (0.65 U/ml). In similar kind of study, UV mutant of *Bacillus subtilis* showed 6.5% increase in β-glucosidase activity than the wild-type strain (0.675 U/ml) [37].

3.2. Fermentation profile of wild-type Bacillus amyloliquefaciens SS35 and its UV2 mutant strain

The fermentation profiles of wild-type *Bacillus amyloliquefaciens* SS35 and its mutant UV2 were studied for 96 h. The cell growth, enzyme activity, protein concentration and pH of broth during fermentation are shown in Fig. 1. The exponential phase of UV2 strain was up to 12 h followed by the stationary phase up to 84 h followed by the death phase. The growth profile of UV2 strain was similar to the wild-type strain. The CMCase activity of UV2 strain increased by 22% as compared with wild-type strain (0.65 U/mL) at approximately, same optical density of cells (2.7) at 48 h of fermentation. The developed mutant UV2 produces cellulase, CMCase-UV2 with in 48 h unlike the fungal cellulase that takes 8 days, as reported earlier [38] which leads to industrially uneconomical process. CMCase production in UV2 mutant strain was not associated with growth as maximum production occurred after log phase similar to the earlier reports for wild-type strain in

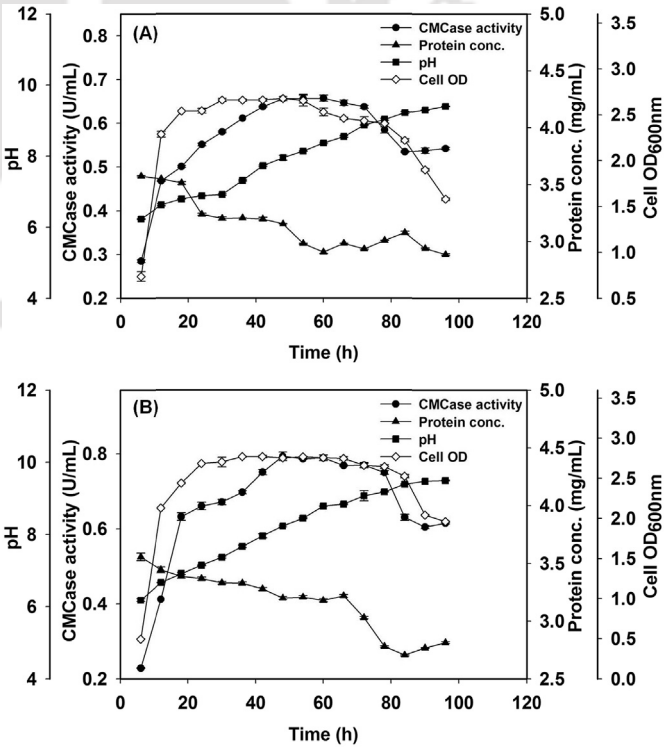


Fig. 1. Fermentation profiles of (A) Wild-type *Bacillus amyloliquefaciens* SS35 and (B) its mutant UV2 strain. The endoglucanase activity, optical density (600 nm), protein concentration and pH change of the broth were monitored at every 6 h up to 96 h.

TH-2393_146106034

Please cite this article as: S. Singh et al., Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis, Renewable Energy, https://doi.org/10.1016/j.renene.2019.11.105

unoptimized medium [20]. The total protein production from UV2 strain was 3.25 mg/mL, approximately similar to the wild-type strain (3.15 mg/mL) at 48 h of fermentation. These observations showed that the increase in cellulase activity is due the mutation in gene encoding cellulase. The pH of the fermenting medium increased from 5.6 to 9.4 during the growth of UV2 mutant strain, approximately similar to the wild-type strain. The increase in pH was observed in log phase and stationary phase and after 84 h increase in pH was negligible due to the onset of the death phase. This increase in pH value of fermenting medium maybe due to the presence of extracellular proteins in which organic amino compounds are deaminated with the growth of bacterium [39].

3.3. Purification of CMCase from wild-type Bacillus amyloliquefaciens SS35 and its UV2 mutant

The two-step purification of CMCase-WT and CMCase-UV2 enzymes from the cell-free supernatant include ammonium sulfate precipitation followed by size exclusion chromatography. CMCase-UV2 showed maximum enzyme activity in fractions precipitated by 90% saturation of ammonium sulfate, which was similar to the wild-type strain as reported earlier [10]. The partially purified CMCase-UV2 showed 2-fold (15 U/mg) higher specific activity than the wild-type strain (7.5 U/mg) shown in Table 1. The partially purified enzyme(s) were further purified by size exclusion chromatography and the fraction numbers 71–103 and 75–114 showed CMCase activity for CMCase-WT (Fig. 2A) and CMCase-UV2 (Fig. 2B), respectively. The purified CMCase-UV2 enzyme showed 32.5 U/mg specific activity, 2.1-fold higher than the CMCase-WT (15.4 U/mg) against the substrate, CMC-Na (Table 1). This enzyme activity of CMCase-UV2 was 5.8-fold higher than the mutant generated by site-directed mutagenesis (5.5 U/mg) by Van Hanh et al. [40]. The site-directed mutagenesis is helpful in forming the genetic diversity of a gene, whose sequence is known. However, UV mutagenesis has advantage that, it requires no prior knowledge of the gene sequence to produce genetic diversity with improved characteristics of cellulase. The zymogram analysis for activity staining of CMCase-UV2 showed a single band of CMCase activity of molecular mass approximately, 37 kDa (Fig. 2C and D) similar to the CMCase enzyme reported from wild-type strain by anion exchange chromatography [10].

3.4. Effect of pH and temperature on CMCase-WT and CMCase-UV2 activity and stability

The optimum pH for CMCase-UV2 using 1% (w/v) CMC-Na was

Table 1
Comparison of enzyme activity of purified CMCase-WT and CMCase-UV2 enzyme.

Purification step	Volume (mL)		Enzyme activity (U/mL)		Total activity (U)		Protein (mg/mL)		Total protein (mg)		Specific activity* (U/mg)	
	WT	UV2	WT	UV2	WT	UV2	WT	UV2	WT	UV2	WT	UV2
Culture supernatant	200	200	0.65	0.79	130	158	3.18	3.25	636	650	0.2 ± 0.01	0.24 ± 0.01
60–90% $(NH_4)_2SO_4$	40	69	1.2	1.5	48	104	0.16	0.10	6.4	6.9	7.5 ± 0.1	15.0 ± 0.3
Size exclusion chromatography	2	2	1.7	3.9	3.4	7.8	0.11	0.12	0.22	0.24	15.4 ± 0.4	32.5 ± 1.2

Values are mean SE (n = 3).
All the assays were performed at 65 °C, 50 mM sodium acetate buffer (pH 5.0), 5 min, 1% (w/v) CMC-Na.

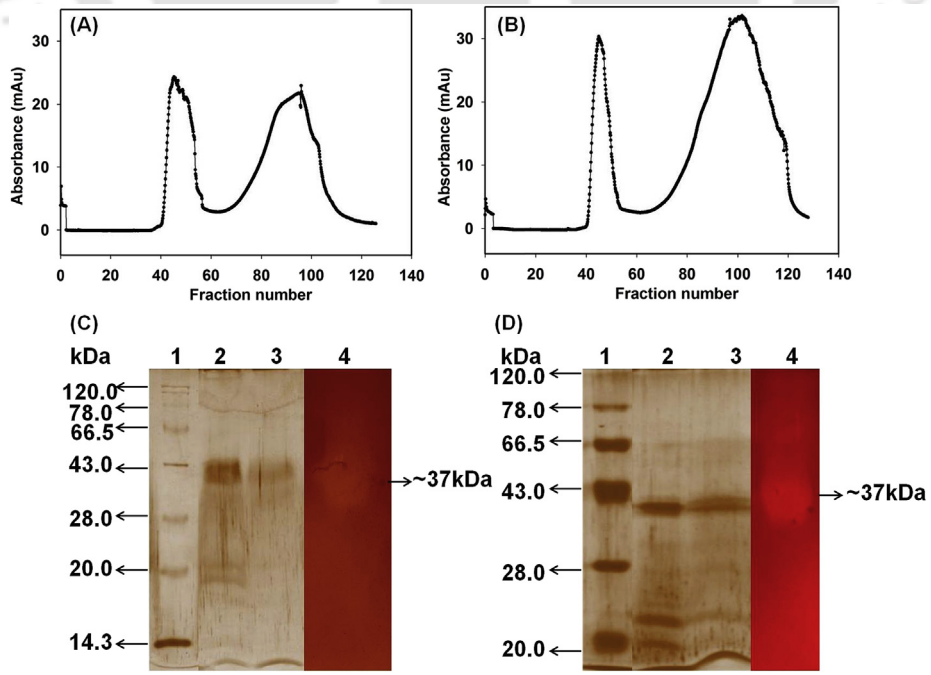


Fig. 2. Chromatogram from size exclusion chromatography of CMCase purification, (A) CMCase-WT (B) CMCase-UV2. SDS-PAGE analysis using (12%, w/v) gel showing purification and zymogram of CMCase-WT (C) Lanes 1, Protein marker; 2, Partially purified CMCase-WT; 3, Purified CMCase-WT (37 kDa, approx.); 4, Zymogram of purified CMCase-WT and for CMCase-UV2 (D) Lanes 1, Protein marker; 2, Partially purified CMCase-UV2; 3, Purified CMCase-UV2 (37 kDa, approx.); 4, Zymogram of purified CMCase-UV2.

TH-2393_146106034

Please cite this article as: S. Singh et al., Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis, Renewable Energy, https://doi.org/10.1016/j.renene.2019.11.105

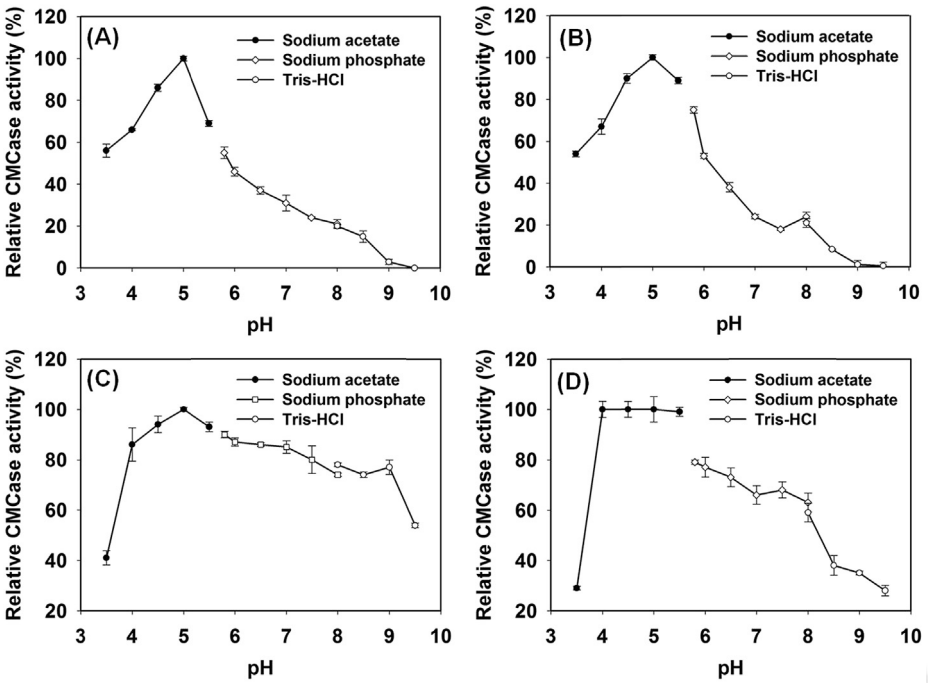


Fig. 3. Effect of pH on CMCase activity; (A) CMCase-WT and (B) CMCase-UV2; effect of pH on CMCase stability; C) CMCase-WT and D) CMCase-UV2.

found to be 5.0 (Fig. 3B), similar to that of CMCase-WT (Fig. 3A) as reported earlier [10]. CMCase-UV2 retained almost 100% of the relative CMCase activity in the pH range, 4.0–5.5 (Fig. 3D), whereas, CMCase-WT showed maximum activity only at pH 5.0 and 85% of the relative CMCase activity at pH 4 (Fig. 3C), after 1 h of incubation. Mutation enhanced the stability of CMCase-UV2 in acidic range, which is more suitable for *Saccharomyces cerevisiae* growth [41] in SSF process for bioethanol production. Therefore, the use of CMCase-UV2 in SSF may also result in improvement of the ethanol production yield. The optimum temperature of CMCase-UV2 was 65 °C similar to the CMCase-WT. The CMCase-UV2 retained 100% of activity till 45 °C for 1 h similar to the CMCase-WT (Supplementary data, Fig. S2), thus can be used efficiently in SSF at 30 °C.

3.5. Comparison of enzyme activity of purified CMCase-WT and CMCase-UV2

The specific activities of the purified CMCase-WT and CMCase-UV2 enzymes were determined against β-D-glucan, lichenan, CMC-Na, cellulose powder, avicel, birchwood xylan and galactomannan (Table 2). CMCase-UV2 enzyme showed 1.6-fold–4-fold higher specific activity against the cellulosic substrates than the wild-type enzyme, displaying enhancement in the catalytic efficiency of the mutant against both soluble and insoluble substrates. The enhancement in the activity of CMCase-UV2 against CMC-Na is 108%, which is significantly higher than that reported for endoglucanase (57%) from *Streptomyces griseoaurantiacus* [16] by UV irradiation. The CMCase-UV2 mutant gave 1.9-fold (7.1 U/mg) and 1.8-fold (9.8 U/mg) higher enzyme activity against pretreated *Parthenium hysterophorus* and *Pennisetum purpureum*, respectively (Table 2), as compared with wild-type enzyme. These results showed the potential of CMCase-UV2 in saccharification of the complex lignocellulosic biomass. Srikrishnan et al. [42] reported 1.5 U/mg activity of mutant endoglucanase produced by site-directed mutagenesis of Eg1Y95F against the anhydrous ammonia pretreated corn stover.

Table 2
Comparison of enzyme activity of CMCase-WT and CMCase-UV2.

Substrate	CMCase-WT (U/mg)[c]	CMCase-UV2 (U/mg)[c]	Fold increase
CMC-Na[a]	15.4 ± 0.4	32.1 ± 1.2	2.1
β-D-Glucan[a]	131.7 ± 1.2	210.1 ± 3.1	1.6
Lichenan[a]	86.5 ± 2.9	161.3 ± 1.1	1.9
Cellulose powder[a]	5.4 ± 0.2	14.3 ± 0.4	2.6
Avicel[a]	4.8 ± 0.2	19.5 ± 0.9	4.1
Birchwood xylan[a]	0.6 ± 0.1	0.78 ± 0.3	1.3
Galactomannan[a]	ND	ND	–
Pretreated Carrot grass[b]	3.8 ± 0.9	7.1 ± 0.2	1.9
Pretreated Elephant grass[b]	5.3 ± 0.5	9.8 ± 0.2	1.8

ND = No activity detected.
[a] All the assays were performed at 1% (w/v) substrate, 65 °C, 50 mM sodium acetate buffer (pH 5.0), 5 min, for commercial substrates.
[b] All the assays were performed at 2% (w/v) substrate, 65 °C, 50 mM sodium acetate buffer (pH 5.0), 15 min, for pretreated substrates.
[c] Values are mean SE (n = 3).

TH-2393_146106034

Please cite this article as: S. Singh et al., Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis, Renewable Energy, https://doi.org/10.1016/j.renene.2019.11.105

Table 3
Kinetic parameters of CMCase-WT and CMCase-UV2.

Substrate (% w/v)	CMCase-WT				CMCase-UV2				Fold increase (K_{cat}/K_m)[a]
	K_m (mgmL^{-1})	V_{max} (μmolmg^{-1}min^{-1})	k_{cat} (min^{-1})	k_{cat}/K_m (mg^{-1} mL^{1}min^{-1})	K_m (mgmL^{-1})	V_{max} (μmolmg^{-1}min^{-1})	k_{cat} (min^{-1})	K_{cat}/K_m (mg^{-1} mL^{1}min^{-1})	
CMC-Na	0.26 ± 0.01	17.21 ± 0.22	6.37×10^2	2.45×10^3	0.38 ± 0.04	41.9 ± 1.3	1.55×10^3	4.1×10^3	1.7
β-Glucan	0.18 ± 0.00	156.70 ± 2.4	5.79×10^3	3.22×10^4	0.15 ± 0.01	238.1 ± 4.6	8.80×10^3	5.87×10^4	1.8

Values are mean SE (n = 3).
All the assays were performed at 65 °C, 50 mM sodium acetate buffer (pH 5.0), 5 min.
[a] Catalytic efficiency = K_{cat}/K_m.

3.6. Comparison of kinetic parameters of CMCase-WT and CMCase-UV2

The kinetic parameters (K_m and V_{max}) for CMCase-WT and CMCase-UV2 were determined by Linewaever–Burk plot. CMCase-UV2 showed V_{max} 238 μmolmg^{-1}min^{-1} and 41.9 μmolmg^{-1}min^{-1} against substrate β-glucan and CMC-Na, respectively, which were higher than the CMCase-WT (156.7 μmolmg^{-1}min^{-1} and 17.2 μmolmg^{-1}min^{-1}, respectively) (Table 3). The V_{max} of CMCase-UV2 against substrate CMC-Na was also higher than other reported V_{max} for CMCase enzymes from wild-type strain *Bacillus* sp. AS3 (3.38 μmolmg^{-1}min^{-1}) [43] and *Bacillus* sp. CH43 (0.00093 μmolmg^{-1}min^{-1}) [44]. The mutant enzyme CMCase-UV2 also showed 1.7 and 1.8-fold increase in catalytic efficiency (k_{cat}/K_m) against CMC-Na and β-glucan respectively, as compared with CMCase-WT (Table 3) showing the potential of CMCase-UV2 in saccharification of the complex lignocellulosic biomass.

3.7. Comparison of CMCase-WT and CMCase-UV2 hydrolysed products of CMC-Na and β-glucan by TLC

The qualitative analyses of CMCase-WT and CMCase-UV2 hydrolysed products of CMC-Na (Fig. 4A and B) and β-D-glucan (Fig. 4C and D) were carried out by TLC. The time dependent TLC analysis of CMC-Na and β-D-glucan hydrolysis by CMCase-WT and CMCase-UV2 with equimolar enzyme(s):substrate ratio used, showed the endolytic cleavage pattern, because in the first 30 min, series of cello-oligosaccharides along with cellobiose were formed. However, after 45 min CMCase-UV2 also showed release of glucose (intense spot) along with cellobiose and cello-oligosaccharides (Fig. 4B and D), whereas, CMCase-WT showed glucose only after 120 min (very faint spot) along with cellobiose and cello-oligosaccharides (Fig. 4A and C). This displayed enhanced rate and efficiency of the mutant enzyme. This may result into the requirement of less amount of β-glucosidase per gram of biomass when a mixture of CMCase-UV2 with β-glucosidase is employed as against CMCase-WT. Therefore, CMCase-UV2 can help in lignocellulose-based biorefineries for the bioethanol production. The TLC results of CMCase-WT and CMCase-UV2 enzymes treated with cellobiose showed no release of glucose indicating the absence of β-glucosidase activity (Supplementary Data, Fig. S3). These results showed that the glucose appearing on TLC with CMC-Na and β-glucan (Fig. 4) treatment with CMCase-WT and CMCase-UV2 is released from the higher cello-oligosaccharides and not from cellobiose. The endolytic cleavage of CMC-Na by CMCase from *Bacillus licheniformis* was also reported earlier [45].

3.8. CMCase-WT and CMCase-UV2 hydrolysis of Pennisetum purpureum

Both the enzymes, CMCase-WT and CMCase-UV2 displayed

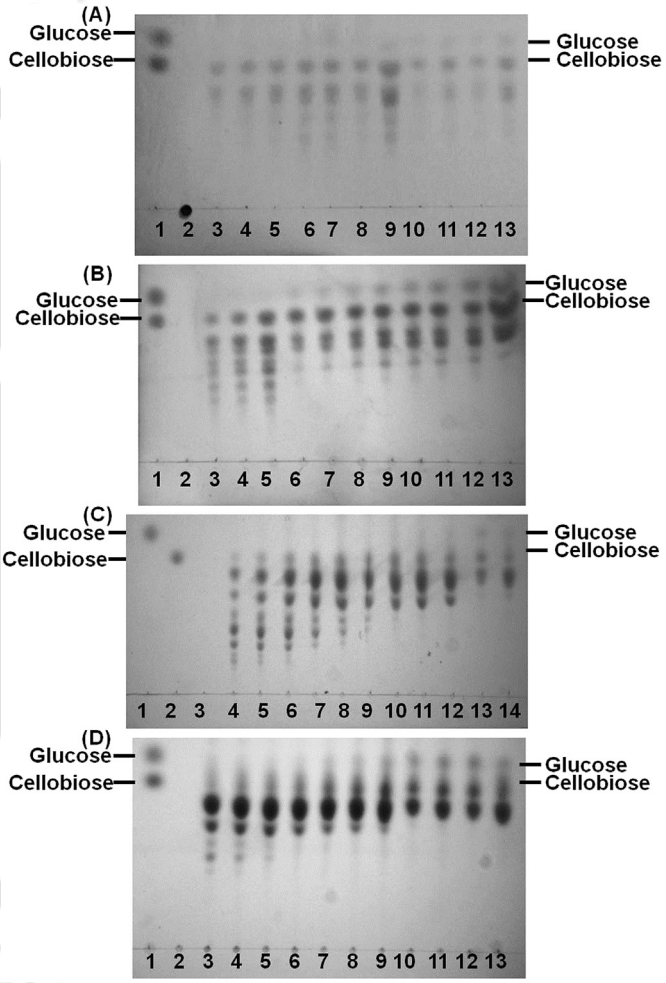


Fig. 4. (A) TLC analysis of CMCase-WT hydrolysed products of CMC-Na. Lanes; 1, standards (glucose and cellobiose); 2, control (without enzyme); 3, 5 min; 4, 10 min; 5, 15 min; 6, 30 min; 7, 45 min; 8, 60 min; 9, 120 min; 10, 180 min; 11, 360 min; 12, 720 min and 13, 1440 min enzyme reaction products. (B) TLC analysis of CMCase-UV2 hydrolysed products of CMC-Na. Lanes; 1, standards (glucose and cellobiose); 2, control (without enzyme); 3, 5 min; 4, 10 min; 5, 15 min; 6, 30 min; 7, 45 min; 8, 60 min; 9, 120 min; 10, 180 min; 11, 360 min; 12, 720 min and 13, 1440 min enzyme reaction products. (C) TLC analysis of CMCase-WT hydrolysed products of β-D-glucan. Lanes; 1, standard (glucose); 2, standard (cellobiose); 3, control (without enzyme); 4, 5 min; 5, 10 min; 6, 15 min; 7, 30 min; 8, 45 min; 9, 60 min; 10, 120 min; 11, 180 min; 12, 360 min; 13, 720 min and 14, 1440 min enzyme reaction products reaction products. (D) TLC analysis of CMCase-UV2 hydrolysed products of β-D-glucan. Lanes; 1, standards (glucose and cellobiose); 2, control (without enzyme); 3, 5 min; 4, 10 min; 5, 15 min; 6, 30 min; 7, 45 min; 8, 60 min; 9, 120 min; 10, 180 min; 11, 360 min; 12, 720 min and 13, 1440 min enzyme reaction products reaction products.

TH-2393_146106034

Please cite this article as: S. Singh et al., Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis, Renewable Energy, https://doi.org/10.1016/j.renene.2019.11.105

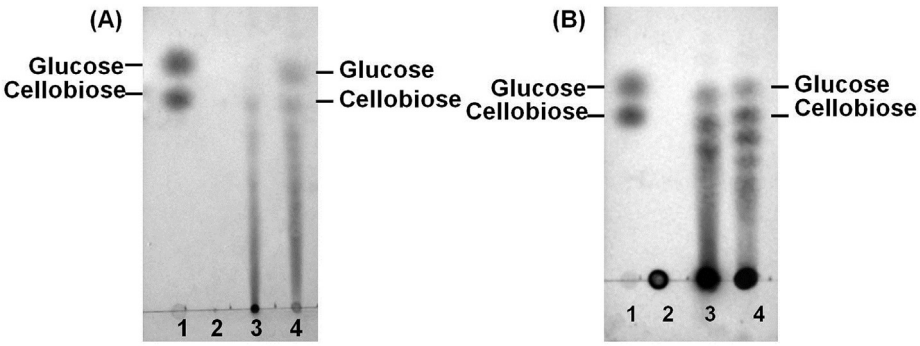


Fig. 5. TLC analysis of enzyme hydrolysed products from pretreated Elephant grass; (A) After 24 h of hydrolysis. Lanes; 1, standard (glucose and cellobiose); 2, control (without enzyme); 3, CMCase-WT hydrolysed products and 4, CMCase-UV2 hydrolysed products; (B) After 48 h of hydrolysis. Lanes; 1, standard (glucose and cellobiose); 2, control (without enzyme); 3, CMCase-WT hydrolysed products and 4, CMCase-UV2 hydrolysed products.

higher activity against pretreated Elephant grass than pretreated Carrot grass (Table 2). Therefore, the enzymatic hydrolysis for pretreated Elephant grass was carried out for the release of reducing sugars. The maximum total reducing sugar (TRS) yield of 154.2 mg/g of pretreated Elephant grass was obtained at 48 h hydrolysis by CMCase-UV2. This TRS yield was 1.8-fold higher than that obtained with CMCase-WT (85.6 mg/g of pretreated biomass) under unoptimized conditions of enzyme concentration and hydrolysis time. However, the saccharification by CMCase-UV2 for 24 h gave 108.6 mg/g and CMCase-WT gave 64.7 mg/g pretreated biomass TRS yield. These results showed that CMCase-UV2 enhanced the saccharification efficiency for lignocellulosic biomass. These results showed that CMCase-UV2 enhanced the saccharification efficiency for lignocellulosic biomass. Eliana et al. [33] reported the maximum TRS yield of 146.9 mg/g of pretreated *Pennisetum purpureum* under optimized conditions of protein concentration and hydrolysis time 45 h. The TLC analysis of CMCase-UV2 hydrolysed mixture of Elephant grass showed the presence of glucose along with the cellobiose and cello-oligosaccharides at 24 h, whereas CMCase-WT did not show any glucose at 24 h (Fig. 5A). However, CMCase-WT produced glucose in addition to cellobiose and cellotriose and cello-oligosaccharides at 48 h of hydrolysis (Fig. 5B). This further confirmed the increased efficiency of CMCase-UV2 with production of glucose in 24 h as against 48 h for CMCase-WT.

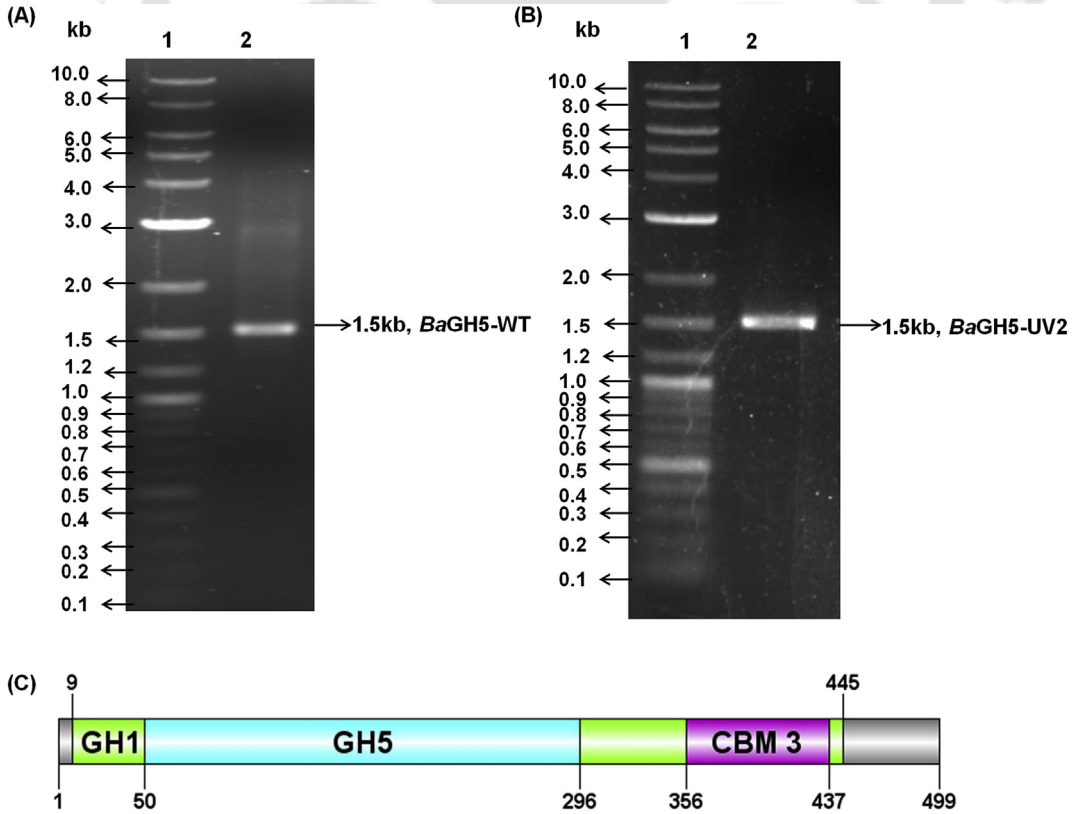


Fig. 6. Agarose gel electrophoresis (0.8%, w/v) showing the PCR-DNA confirming the clones for *Ba*GH5-WT (A) Lanes: 1, DNA ladder; 2, amplified PCR fragment of *Ba*GH5-WT (1.5 kb) from recombinant plasmid (positive clone) and for *Ba*GH5-UV2 (B) Lanes: 1, DNA ladder; 2, amplified PCR fragment of *Ba*GH5-UV2 (1.5 kb) from recombinant plasmid (positive clone). (C) Molecular architecture of family 5 glycoside hydrolase (*Ba*GH5) from *Bacillus amyloliquefaciens SS35.*

TH-2393_146106034

Please cite this article as: S. Singh et al., Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis, Renewable Energy, https://doi.org/10.1016/j.renene.2019.11.105

3.9. Identification of induced mutations in UV2 strain at genetic level

Numerous factors may be associated with the alteration in the enzyme activity of the mutant enzyme produced from the UV2 mutant strain of *Bacillus amyloliquefaciens* SS35. There might be some genetic changes in the UV2 strain. Therefore, the genes encoding endoglucanases i.e. *Ba*GH5-WT and *Ba*GH5-UV2 were amplified from the wild-type strain and UV2 strains, respectively using degenerate oligonucleotide primers obtained from phylogenetically related spp. *Bacillus amyloliquefaciens* KHG19 (Supplementary data, Fig. S4) for family 5 glycoside hydrolase (GH5) with GenBank accession number: AJK65578.1 from the CAZY database (http://www.cazy.org) to identify the genetic changes. The PCR-amplified DNA of *Ba*GH5-WT (~1.5 kb) and *Ba*GH5-UV2 (~1.5 kb) were cloned into a pHTP0 cloning vector using NZYEasy cloning kit (Supplementry data, Fig. S5). The positive clones of *Ba*GH5-WT and *Ba*GH5-UV2 were identified and confirmed by colony PCR (Fig. 6A and B) and sequencing. The sequencing results of positive clones for *Ba*GH5-WT and *Ba*GH5-UV2 showed the molecular architecture of GH5 family (Fig. 6C). It displayed an N-terminal family 1 glycoside hydrolase (437 amino acid), family 5 glycoside hydrolase (*Ba*GH5, 247 amino acid) followed by family 3 carbohydrate binding module *Ba*CBM3 (82 amino acid) at the C-terminal. The gene sequence alignment of enzymes *Ba*GH5-UV2 with *Ba*GH5-WT showed a single transition mutation at position 698 bp where adenine base was substituted by guanine. In DNA, UV light cause changes in the position of hydrogen atoms in adenine, guanine, cytosine and thymine base to form tautomer as reported earlier by Watson and Crick [46] and Topal and Fresco [47]. An amino (-NH_2) group in adenine base is tautomerized to form imino (=NH) group, because of which this tautomerized adenine base is paired with cytosine instead of thymine base. Hence, in the next round of replication this cytosine base, pairs with guanine base which results AT to GC transition. Due to this above point mutation in *Ba*GH5-UV2, codon GAC changed to GGC and thus, changed the amino acid residue aspartic acid 233 (D233) to glycine 233 (G233) (Fig. 7). Cherry et al. [48] reported the consequences of base substitution mutations in protein coding regions of a gene depend on its locations. The substitution in the active site of the enzyme could deactivate enzyme and other than the active site could increase the enzyme activity. Glycine substitution enhanced pH stability in acidic range suggested that this amino acid residue might be the key residue in determining pH stability of enzyme. Its substitution also enhanced the catalytic efficiency of mutant enzyme because it is the simplest amino acid with no nonpolar side chain and can provide hydrogen bonding to support transition state in catalysis as compared with the aspartic acid which has the larger side chain. Caniago et al. [49] showed UV irradiation cause change in CCT codon to TCG codon, changing proline residue into serine resulting in enhanced endoglucanase activity. The gene sequences of *Ba*GH5-WT and *Ba*GH5-UV2 provide information about the restriction sites and thus, may help in cloning of these genes into a suitable expression vector for over expression thereby enhancing their production.

4. Conclusion

The present study demonstrated that UV directed evolution of *Bacillus amyloliquefaciens* SS35 enhanced the CMCase activity along with catalytic efficiency and pH stability in the acidic range. The mutant CMCase-UV2 can be advantageous and economical for SSF in the conversion of lignocellulosic biomass to bioethanol. Substitution of amino acid residue G233 in place of D233 improved the CMCase activity and other properties. This report provides the

```
BaGH5-WT        MKRAISIFITCLLIAVLTMGGLLPSPASAAGTKTPVAKNGQLSIKGTQLVNRDGKAVQLK     60
BaGH5-UV2       MKRAISIFITCLLIAVLTMGGLLPSPASAAGTKTPVAKNGQLSIKGTQLVNRDGKAVQLK     60
                ************************************************************

BaGH5-WT        GISSHGLQWYGDFVNKDSLKWLRDDWGITVFRAAMYTADGGYIDNPSVKNKVKEAVEAAK     120
BaGH5-UV2       GISSHGLQWYGDFVNKDSLKWLRDDWGITVFRAAMYTADGGYIDNPSVKNKVKEAVEAAK     120
                ************************************************************

BaGH5-WT        ELGIYVIIDWHILNDGNPNQNKEKAKEFFKEMSSLYGNTPNVIYEIANEPNGDVNWKRDI     180
BaGH5-UV2       ELGIYVIIDWHILNDGNPNQNKEKAKEFFKEMSSLYGNTPNVIYEIANEPNGDVNWKRDI     180
                ************************************************************

BaGH5-WT        KPYAEEVISVIRKNDPDNIIIVGTGTWSQDVNDAADDQLKDANVMYALHFYADTHGQSLR     240
BaGH5-UV2       KPYAEEVISVIRKNDPDNIIIVGTGTWSQDVNDAADDQLKDANVMYALHFYAGTHGQSLR     240
                ************************************************.*******

BaGH5-WT        DKANYALSKGAPIFVTEWGTSDASGNGGVFLDQSREWLNYLDSKNISWVNWNLSDKQESS     300
BaGH5-UV2       DKANYALSKGAPIFVTEWGTSDASGNGGVFLDQSREWLNYLDSKNISWVNWNLSDKQESS     300
                ************************************************************

BaGH5-WT        SALKPGASKTGGWPLTDLTASGTFARENIRGTKGSTKDSPETPAQDNPTQEKGISVQYKA     360
BaGH5-UV2       SALKPGASKTGGWPLTDLTASGTFARENIRGTKGSTKDSPETPAQDNPTQEKGISVQYKA     360
                ************************************************************

BaGH5-WT        GDGRVNSNQIRPQLHIKNNGNATVDLKDVTARYWYNVKNKGQNFDCDYAQIGCGNLTHKF     420
BaGH5-UV2       GDGRVNSNQIRPQLHIKNNGNATVDLKDVTARYWYNVKNKGQNFDCDYAQIGCGNLTHKF     420
                ************************************************************

BaGH5-WT        VTLHKPKQDADTYLELGFKTGTLSPGASTGNIQLRLHNDDWSNYAQSGDYSFFQSNTFKT     480
BaGH5-UV2       VTLHKPKQDADTYLELGFKTGTLSPGASTGNIQLRLHNDDWSNYAQSGDYSFFQSNTFKT     480
                ************************************************************

BaGH5-WT        TKKITLYHQGKLIWGTEPN      499
BaGH5-UV2       TKKITLYHQGKLIWGTEPN      499
                ********************
```

Fig. 7. Sequence alignment of *Ba*GH5-UV2 with *Ba*GH5-WT protein sequence by Clustral Omega, showing the mutation, D233 to G233.

TH-2393_146106034

Please cite this article as: S. Singh et al., Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis, Renewable Energy, https://doi.org/10.1016/j.renene.2019.11.105

ARTICLE IN PRESS

information for protein engineering to design a mutant enzyme for further improving the catalytic efficiency and pH stability.

Declaration of competing interest

There are no conflicts of interest to declare.

Acknowledgments

The research work was supported by the Centre for Bioenergy DBT-Pan-IIT Grant (BT/EB/PAN-IIT 2012) by the Department of Biotechnology, Ministry of Science and Technology, New Delhi, India. The authors also thankful to the Department of Biosciences and Bioengineering, IIT Guwahati for the instrument facilities. The authors are grateful to Professor Carlos M.G.A. Fontes for fruitful discussion and valuable suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.renene.2019.11.105.

References

[1] Z. Chen, J.H. Pereira, H. Liu, H.M. Tran, N.S. Hsu, D. Dibble, B.A. Simmons, Improved activity of a thermophilic cellulase, Cel5A, from *Thermotoga maritima* on ionic liquid pretreated switchgrass, PLoS One 8 (11) (2013), e79725.
[2] R.E. Sims, W. Mabee, J.N. Saddler, M. Taylor, An overview of second generation biofuel technologies, Bioresour. Technol. 101 (6) (2010) 1570–1580.
[3] L.R. Lynd, M.S. Laser, D. Bransby, B.E. Dale, B. Davison, R. Hamilton, M. Himmel, M. Keller, J.D. McMillan, J. Sheehan, C.E. Wyman, How biotech can transform biofuels, Nat. Biotechnol. 26 (2) (2008) 169.
[4] M.E. Himmel, S.Y. Ding, D.K. Johnson, W.S. Adney, M.R. Nimlos, J.W. Brady, T.D. Foust, Biomass recalcitrance: engineering plants and enzymes for biofuels production, Science 315 (5813) (2007) 804–807.
[5] H.G. Lawford, J.D. Rousseau, Cellulosic fuel ethanol, in: Biotechnology for Fuels and Chemicals, 2003, pp. 457–469.
[6] E.A. Bayer, R. Lamed, M.E. Himmel, The potential of cellulases and cellulosomes for cellulosic waste management, Curr. Opin. Biotechnol. 18 (3) (2007) 237–245.
[7] Y. Hakamada, K. Koike, T. Yoshimatsu, H. Mori, T. Kobayashi, S. Ito, Thermostable alkaline cellulase from an alkaliphilic isolate, *Bacillus* sp, KSM-S237. Extremophiles 1 (3) (1997) 151–156.
[8] B.K. Kim, B.H. Lee, Y.J. Lee, I.H. Jin, C.H. Chung, J.W. Lee, Purification and characterization of carboxymethylcellulase isolated from a marine bacterium, *Bacillus subtilis* subsp. subtilis A-53, Enzym. Microb. Technol. 44 (6–7) (2009) 411–416.
[9] P. Shobharani, D. Yogesh, P.M. Halami, N.M. Sachindra, Potential of cellulase from *Bacillus megaterium* for hydrolysis of sargassum, J. Aquat. Food Prod. Technol. 22 (5) (2013) 520–535.
[10] S. Singh, P.K. Dikshit, V.S. Moholkar, A. Goyal, Purification and characterization of acidic cellulase from *Bacillus amyloliquefaciens* SS35 for hydrolyzing *Parthenium hysterophorus* biomass, Environ. Prog. Sustain. Energy 34 (3) (2014 b) 810–818.
[11] J.M. Kim, I.S. Kong, J.H. Yu, Molecular cloning of an endoglucanase gene from an alkalophilic *Bacillus* sp. and its expression in *Escherichia* coli, Appl. *Environ. Microbiol.* 53 (11) (1987) 2656–2659.
[12] R. Gupta, Q. Beg, P. Lorenz, Bacterial alkaline proteases: molecular approaches and industrial applications, Appl. Microbiol. Biotechnol. 59 (1) (2002) 15–32.
[13] X.H. Li, H.J. Yang, B. Roy, E.Y. Park, L.J. Jiang, D. Wang, Y.G. Miao, Enhanced cellulase production of the *Trichoderma viride* mutated by microwave and ultraviolet, Microbiol. Res. 165 (3) (2010) 190–198.
[14] X.Y. Zhang, L.H. Zi, X.M. Ge, Y.H. Li, C.G. Liu, F.W. Bai, Development of *Trichoderma reesei* mutants by combined mutagenesis and induction of cellulase by low-cost corn starch hydrolysate, Process Biochem. 54 (2017) 96–101.
[15] B. Kaur, H.S. Oberoi, B.S. Chadha, Enhanced cellulase producing mutants developed from heterokaryotic *Aspergillus* strain, Bioresour. Technol. 156 (2014) 100–107.
[16] A.K. Kumar, UV mutagenesis treatment for improved production of endoglucanase and β-glucosidase from newly isolated thermotolerant actinomycetes, *Streptomyces griseoaurantiacus*, Bioresourc. Bioprocess. 2 (1) (2015) 22.
[17] N. Akhtar, Aanchal, D. Goyal, A. Goyal, Biodiversity of cellulase producing bacteria and their applications, Cellul. Chem. Technol. 50 (9–10) (2016) 983–995.
[18] M.P. Coughlan, The properties of fungal and bacterial cellulases with comment on their production and application, Biotechnol. Genet. Eng. Rev. 3 (1) (1985) 39–110.
[19] H.J. Schaeffer, J. Leykam, J.D. Walton, Cloning and targeted gene disruption of EXG1, encoding exo-beta 1, 3-glucanase, in the phytopathogenic fungus *Cochliobolus carbonum*, AEM 60 (2) (1994) 594–598.
[20] S. Singh, V.S. Moholkar, A. Goyal, Isolation, identification and characterization of a cellulolytic *Bacillus amyloliquefaciens* strain SS35 from Rhinoceros Dung, ISRN Microbiol. 728134 (2013) 7.
[21] S. Singh, V.S. Moholkar, A. Goyal, Optimization of carboxymethylcellulase production from *Bacillus amyloliquefaciens* SS35, 3 Biotech. 4 (4) (2014a) 411–424.
[22] R.A. Rapa, A. Islam, L.G. Monahan, A. Mutreja, N. Thomson, I.G. Charles, M. Labbate, A genomic island integrated into recA of *Vibrio cholerae* contains a divergent recA and provides multipathway protection from DNA damage, Environ. Microbiol. 17 (4) (2015) 1090–1102.
[23] N. Nelson, A photometric adaptation of the Somogyi method for the determination of glucose, J. Biol. Chem. 153 (2) (1944) 375–380.
[24] M.A. Somogyi, New reagent for the determination of sugars, J. Biol. Chem. 160 (1945) 61–68.
[25] O.H. Lowry, N.J. Rosebrough, A.L. Farr, R.J. Randall, Protein measurement with the Folin phenol reagent, J. Biol. Chem. 193 (1) (1951) 265–275.
[26] W.I. Wood, Tables for the preparation of ammonium sulfate solutions, Anal. Biochem. 73 (1) (1976) 250–257.
[27] H. Schagger, Chromatographic techniques and basic operations in membrane protein purification, in: A Practical Guide to Membrane Protein Purification, 1994, pp. 23–57.
[28] J.B. Hammond, N.J. Kruger, The Bradford Method for Protein Quantitation. New Protein Techniques, Humana Press, 1988, pp. 25–32.
[29] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (5259) (1970) 680.
[30] W. Wray, T. Boulikas, V.P. Wray, R. Hancock, Silver staining of proteins in polyacrylamide gels, Anal. Biochem. 118 (1) (1981) 197–203.
[31] J.S. van Dyk, M. Sakka, K. Sakka, B.I. Pletschke, Identification of endoglucanases, xylanases, pectinases and mannanases in the multi-enzyme complex of *Bacillus licheniformis* SVD1, Enzym. Microb. Technol. 47 (3) (2010) 112–118.
[32] Pretreated Suchi Singh.
[33] C. Eliana, R. Jorge, P. Juan, R. Luis, Effects of the pretreatment method on enzymatic hydrolysis and ethanol fermentability of the cellulosic fraction from elephant grass, Fuel 118 (2014) 41–47.
[34] L. Michaelis, M.L. Menten, Die kinetic der invertinwirkung, Biochem. J. 49 (1913) 333–369.
[35] H. Lineweaver, D. Burk, The determination of enzyme dissociation constants, J. Am. Chem. Soc. 56 (3) (1934) 658–666.
[36] S.B. Jamaldheen, K. Sharma, A. Rani, V.S. Moholkar, A. Goyal, Comparative analysis of pretreatment methods on Sorghum (*Sorghum durra*) stalk agro-waste for holocellulose content. *Pre*, Biochem. Biotechnol. (2018) 457–464, https://doi.org/10.1080/10826068.2018.1466148.
[37] R. Agrawal, A. Satlewal, A.K. Verma, Development of a β-glucosidase hyper-producing mutant by combined chemical and UV mutagenesis, 3 Biotech. 3 (5) (2013) 381–388.
[38] R. Saini, J.K. Saini, M. Adsul, A.K. Patel, A. Mathur, D. Tuli, R.R. Singhania, Enhanced cellulase production by *Penicillium oxalicum* for bio-ethanol application, Bioresour. Technol. 188 (2015) 240–246.
[39] S.A. Ladeira, E. Cruz, A.B. Delatorre, J.B. Barbosa, M.L. Leal Martins, Cellulase production by thermophilic *Bacillus* sp: SMIA-2 and its detergent compatibility, Electron. J. Biotechnol. 18 (2) (2015) 110–115.
[40] V.H. Vu, K. Kim, Improvement of cellulase activity using error-prone rolling circle amplification and site-directed mutagenesis, J. Microbiol. Biotechnol. 22 (5) (2012) 607–613.
[41] Z.H. Liu, L. Qin, J.Q. Zhu, B.Z. Li, Y.J. Yuan, Simultaneous saccharification and fermentation of steam-exploded corn stover at high glucan loading and high temperature, Biotechnol. Biofuels 7 (1) (2014) 167.
[42] S. Srikrishnan, A. Randall, P. Baldi, N.A. Da Silva, Rationally selected single-site mutants of the *Thermoascus aurantiacus* endoglucanase increase hydrolytic activity on cellulosic substrates, Biotechnol. Bioeng. 109 (6) (2012) 1595–1599.
[43] D. Deka, M. Jawed, A. Goyal, Purification and characterization of an alkaline cellulase produced by *Bacillus subtilis* (AS3), Prep. Biochem. Biotechnol. 43 (2013) 256–270.
[44] C. Mawadza, R. Hatti–Kaul, R. Zvauya, B. Mattiason, Purification and characterization of cellulases produced by two *Bacillus* strains, J. Biotechnol. 83 (2000) 177–0187.
[45] K.M. Bischoff, A.P. Rooney, X.L. Li, S. Liu, S.R. Hughes, Purification and characterization of a family 5 endoglucanase from a moderately thermophilic strain of *Bacillus licheniformis*, Biotechnol. Lett. 28 (2006) 1761–1765.
[46] J.D. Watson, F.H.C. Crick, A structure for deoxyribose nucleic acids, Nature 171 (1953) 737–738.
[47] M.D. Topal, J.R. Fresco, Complementary base pairing and the origin of substitution mutations, Nature 263 (1976) 285–293.
[48] J.R. Cherry, A.L. Fidantsef, Directed evolution of industrial enzymes: an update, Curr. Opin. Biotechnol. 14 (4) (2003) 438–443.
[49] A. Caniago, W. Mangunwardoyo, S. Nuswantara, P. Lisdiyanti, Improvement of endoglucanase activity in *Penicillium oxalicum* ID10-T065 by ultra violet irradiation and ethidium bromide mutation, Ann. Bogor. 19 (2015) 2.

TH-2393_146106034

Please cite this article as: S. Singh et al., Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis, Renewable Energy, https://doi.org/10.1016/j.renene.2019.11.105

Check for updates

Arun Goyal ORCID iD: 0000-0003-3403-9547

Role of glycine 256 residue in improving the catalytic efficiency of mutant endoglucanase of family 5 glycoside hydrolase from *Bacillus amyloliquefaciens* SS35

Shweta Singh[1,2], Krishan Kumar[1], Priyanka Nath[1,2] and Arun Goyal[1,2*]

[1]Carbohydrate Enzyme Biotechnology Laboratory, Department of Biosciences and Bioengineering, [2]DBT PAN-IIT Centre for Bioenergy, Indian Institute of Technology Guwahati, Guwahati-781039, Assam, India.

Corresponding author

Professor Arun Goyal

Carbohydrate Enzyme Biotechnology Laboratory,

Department of Biosciences and Bioengineering,

Indian Institute of Technology Guwahati,

Guwahati-781039, Assam, India

Tel: +91-361-258-2208

Email: arungoyl@iitg.ac.in

Grant numbers: DBT-Pan-IIT Grant (BT/EB/PAN-IIT 2012)

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/bit.27448.

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

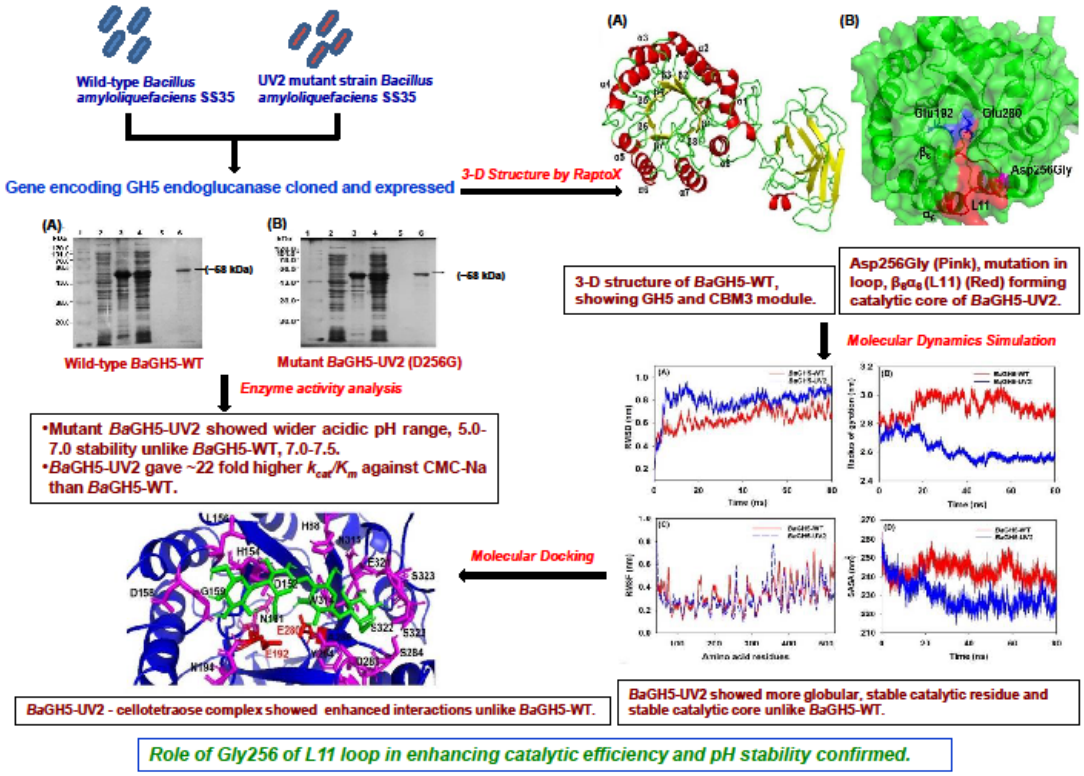
Abstract

Wild-type, *Ba*GH5-WT and mutant, *Ba*GH5-UV2 (aspartate residue mutated to glycine), endoglucanases belonging to glycoside hydrolase family 5 (GH5), from wild-type, and UV2 mutant strain of *Bacillus amyloliquefaciens* SS35, respectively, were earlier cloned in pHTP0 cloning vector. In this study, genes encoding *Ba*GH5-WT or *Ba*GH5-UV2 were cloned into pET28a(+) expression-vector and expressed in *E. coli* BL-21(DE3)pLysS cells. *Ba*GH5-UV2 showed 10-fold (43.6 U/mg) higher specific activity against CMC-Na, higher optimal temperature by 10°C at 65°C, and 22- fold higher catalytic efficiency against CMC-Na, than *Ba*GH5-WT. *Ba*GH5-UV2 showed stability in wider acidic pH range (5.0-7.0) unlike *Ba*GH5-WT in narrow basic pH range (7.0-7.5). *Ba*GH5-UV2 displayed a mutation, Asp256Gly in L11 loop, connecting β_6-sheet with α_6-helix, near active site towards the domain surface of $(\alpha/\beta)_8$-TIM barrel fold. Molecular Dynamics simulation studies showed more stable structure, accessibility of substrate for catalytic site, and increased flexibility of loop L11 of *Ba*GH5-UV2 than wild-type, suggesting enhanced catalysis by *Ba*GH5-UV2. Molecular Docking analysis displayed enhanced hydrogen bond interactions of cello-oligosaccharides with *Ba*GH5-UV2, unlike *Ba*GH5-WT. Thus, Gly256 residue of loop L11 plays an important role in enhancing catalytic efficiency, and pH stability of GH5 endoglucanase. Therefore, these results help in protein engineering of GH5 endoglucanase for improved biochemical properties.

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Graphical Abstract



Keywords: Substitution mutation; Mutant endoglucanase expression; Molecular Dynamics simulation; Molecular Docking.

Introduction

Environmental safety and energy security are the two most important issues that have heightened the demand for alternative of fossil fuels such as eco-friendly biofuel, bioethanol (Vohra *et al.*, 2014). Cellulose is the most abundant, biodegradable, and renewable carbohydrate in biosphere (Akhtar *et al.*, 2019). Therefore, lignocellulosic biomass can serve as the potential substrate for the production of liquid transportation biofuels. Microbial cellulases are the biocatalysts which convert the cellulosic polysaccharide into the fermentable glucose sugar for production of bioethanol. The aerobic mesophilic bacteria of genus *Bacillus* are the efficient producers of cellulase enzymes (Kim *et al.*, 2009; Shobharani *et al.*, 2013;

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

Singh *et al.,* 2013). Cellulases are classified into the three major categories, 1) endoglucanase (EC 3.2.1.4), 2) cellobiohydrolase (EC 3.2.1.91), and 3) β-glucosidase (EC 3.2.1.21) (Singh *et al.,* 2013). The high production cost of cellulase enzymes is the major bottleneck for the bioethanol industry. Therefore, efforts are required to overcome the production cost of cellulase by improving the low enzyme yield, and low catalytic efficiency (Hodgson *et al.,* 1994; Singh *et al.,* 2020). Strain improvement of wild-type bacterial strain *via* random mutagenesis or directed evolution and site-directed mutagenesis are the promising approaches of protein engineering to improve the catalytic efficiency, broad range pH tolerance, and broad range working temperature range of cellulases. These approaches combined with recombinant DNA technology can help in enhancing the production of mutant proteins to several folds (Hodgson *et al.,* 1994).

Cellulases are classified as glycoside hydrolase enzymes belonging to families 5-9, 12, 26, 44, 45, 48, 60, and 61 in CAZy database (www.cazy.org) (Schulein *et al.,* 2000). Most of the Glycoside hydrolase family 5 (GH5) cellulases are endo-β-1,4-glucanase. The structure studies of endo-β-1,4-glucanase (cellulase) from GH5 family showed that the catalytic domain has $(\beta/\alpha)_8$ TIM-barrel fold, where eight parallel β-strands and eight parallel α-helices are connected by αβ or βα loops (Wierenga *et al.,* 2001). These αβ or βα loops are normally found on the protein surface and are critical for catalytic activity, and stability. Thus, a mutation in these loops can change loop flexibility, which led to improve the biochemical properties of cellulase (Badieyan *et al.,* 2012; Yu *et al.,* 2015; Tu *et al.,* 2016; Ruiz *et al.,* 2016; Liang *et al.,* 2018). The relationship of protein flexibility with its activity and stability is quite complex. Molecular dynamics (MD) simulation study of a mutated protein offers a great deal of information regarding the change in protein flexibility, protein-ligand binding

Accepted Article

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

characteristics, and stability (Yu *et al.,* 2015). The molecular docking analysis predicts the ligand-receptor complex structure in the protein tertiary structure *via* computational methods (Ballester *et al.,* 2010). It helps in predicting the change in hydrogen bonds, and hydrophobic interactions in a mutant protein which are responsible for change in the catalytic efficiency.

The gene encoding a GH5, endoglucanase, *Ba*GH5-WT or *Ba*GH5-UV2 from wild-type *Bacillus amyloliquefaciens* SS35 (Singh *et al.,* 2013) and from its UV2 mutant strain (Singh *et al.,* 2020) with enhanced specific activity, respectively, were earlier cloned in pHTP0 cloning vector by degenerate oligonucleotide primers (Singh *et al.,* 2020). The sequencing results of positive clones for *Ba*GH5-WT or *Ba*GH5-UV2 showed that UV-irradiation in UV2 strain caused a single transition mutation of adenine to guanine base at 698 position in *Ba*GH5-UV2 encoding gene. Due to this mutation in *Ba*GH5-UV2, the codon GAC was changed to GGC and replaced the Asp residue by Gly (Singh *et al.,* 2020). The role of this altered amino acid in recombinant *Ba*GH5-UV2 that led to enhanced enzyme activity is needed to be explored, and also its effect on structural changes responsible for changes in biochemical properties of enzyme is to be investigated. In this study, the recombinant endoglucanase, *Ba*GH5-WT, and *Ba*GH5-UV2 were cloned in pET28a(+) expression vector and expressed in *E. coli* BL-21(DE3)pLysS cells. The comparison of biochemical characterization of purified recombinant enzymes, *Ba*GH5-WT and *Ba*GH5-UV2 were carried out. The change in temperature optima, pH optima, thermo-stability, pH stability, and catalytic efficiency for the mutant enzyme, *Ba*GH5-UV2 was determined. The structure, MD simulation, and molecular docking studies of *Ba*GH5-WT and *Ba*GH5-UV2 were performed and compared. This is the first report on studying the role of a mutated residue *via* a UV-directed evolution in GH5 endoglucanase from UV mutant strain.

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

Material and Methods

Bacterial strains and Plasmid

The gene encoding wild-type *Ba*GH5-WT from *Bacillus amyloliquefaciens* SS35 and mutant *Ba*GH5-UV2 from UV2 mutant strain of *Bacillus amyloliquefaciens* SS35 were earlier cloned in cloning vector, pHTP0 (Singh *et al.*, 2020). The pET-28a(+) expression vector was used for expressing the gene encoding *Ba*GH5-WT, and *Ba*GH5-UV2. *E. coli* TOP10 competent cells were used for cloning or plasmid amplification, and *E. coli* BL-21(DE3)pLysS competent cells (Novagen, EMD4 BioScience, Germany) were used as expression host cells.

Chemicals, reagents, substrates, and kits

Deoxynucleotide primers were purchased from GCC Biotech Private Limited, India. dNTPs were purchased from MP Biomedicals Private Limited, India. Phusion DNA polymerase was purchased from ThermoFisher Scientific, USA. Restriction enzymes, *Nhe*I and *Xho*I were purchased from New England Biolabs (NEB), USA. T_4 DNA ligase was purchased from Promega, USA. The expression vector and pET28a(+) was purchased from Novagen, Germany. The substrates, β-glucan, lichenan, carboxymethylcellulose sodium salt (CMC-Na) (low viscosity), avicel, cellobiose, oat spelt xylan, birchwood xylan, and galactomannan were purchased from Sigma-Aldrich Co. LLC., USA. GenElute miniprep plasmid isolation kit, Bradford reagent, and DNAse-RNAse free water (pH 8.0) were procured from Sigma-Aldrich Co. LLC., USA. The Protein marker was procured from Biobharti, Lifescience Pvt. Ltd., India. Glucose oxidase-peroxidase (GOD-POD) kit from the coral clinical system (Tulip Diagnostics, Pvt. Ltd., India). All medium components, chemicals, and antibiotics were purchased from HiMedia Pvt. Ltd., India.

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

Gene amplification and cloning

The amplification of genes encoding *Ba*GH5-WT or *Ba*GH5-UV2 from recombinant plasmid pHTP0 were done by designing the oligonucleotide primers by using the gene sequence(s) detail as described earlier (Singh *et al.*, 2020). The *Nhe*I and *Xho*I restriction sites were absent in wild-type *Ba*GH5-WT and did not show mutation in mutant *Ba*GH5-UV2 (Singh *et al.*, 2020) therefore were used in designing the oligonucleotide primers. The primers used for amplifying the genes encoding *Ba*GH5-WT and *Ba*GH5-UV2 were Forward, 5'-CGCTAGCATGAAACGGGCAATTTCTATTTT-3', and Reverse, 5'-CCCTCGAGTAACTAATTGGGTTCTGTTCCC-3'. A 100 µl PCR reaction mixture contained 3.98 µL template plasmid DNA(s) (4.3 ng/µL), 0.8 µL dNTPs (1.5 µM), 3 µL each forward, and reverse primer (0.5 µM), 3 µL DMSO (3%, v/v), 1 µL Phusion High-Fidelity DNA polymerase (0.02 U/µl), and 65.22 µL PCR-grade water. The thermal cycles employed were: initial denaturation, 98ºC for 1 min, followed by 30 cycles of i) denaturation 98ºC, 10 s, ii) annealing, 62°C, 30 s, and iii) extension, 72°C, 90 s followed by the final extension at 72°C for 10 min on a thermal cycler (Applied Biosystems, ThermoFischer Scientific, US). The PCR amplified DNA of *Ba*GH5-WT or *Ba*GH5-UV2 was ligated to *Nhe*I-*Xho*I digested pET-28a(+) expression vector. The ligated recombinant plasmid DNA of *Ba*GH5-WT or *Ba*GH5-UV2 was transformed into *E. coli* TOP10 cells for plasmid amplification. The transformed cells were spread-plated on Luria Bertani (LB) agar plates containing 50 µg/ml kanamycin and grown at 37°C for 12 h. The positive clones for recombinant plasmid DNA(s) were confirmed by restriction digestion and sequencing and transformed in *E. coli* BL-21 (DE3)pLysS cells for expression. The transformed cells were spread-plated on LB

TH-2393_146106034
This article is protected by copyright. All rights reserved.

agar plates supplemented with chloramphenicol (25 µg/ml), and kanamycin (50 µg/ml), and grown at 37°C for 12 h.

Expression and purification of recombinant proteins, *Ba*GH5-WT and *Ba*GH5-UV2

The *E. coli* BL-21 (DE3)pLysS cells harbouring recombinant plasmids of *Ba*GH5-WT or *Ba*GH5-UV2, were cultured in 100 mL LB medium in 250 mL Erlenmeyer flask for protein expression. The protocol for protein expression was followed as described earlier (Taylor *et al.*, 2005). The cells were grown at 37ºC, and 180 rpm up to mid-exponential phase to cell absorbance at 600 nm (A_{600}) ~0.6. 1.0 mL of this grown culture containing uninduced cells was taken out for sample preparation for SDS-PAGE analysis. To the remaining culture, 1.0 mM final concentration of isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added, and it was further incubated at 24ºC, 180 rpm for 18 h for induction of protein expression. The cells were centrifuged at 8,000 g at 4°C for 10 min. The cell pellet(s) were re-suspended in 4 mL of 20 mM sodium phosphate buffer, pH 7.0, and the cell suspension was sonicated on ice for 15 min by using 5 s on/5 s off pulse and 33% amplitude. The sonicated cell suspension was again centrifuged at 13,000 g at 4°C for 50 min. The recombinant protein(s) from cell free extract was purified by immobilized metal-ion affinity chromatography (IMAC) using 1.0 mL sepharose columns (GE Healthcare, HiTrap chelating HP). The compositions of binding, as well as elution buffers used for affinity column purification of recombinant proteins, are listed in Table S1. The purified recombinant protein(s) were dialyzed against 1000 mL of 20 mM sodium phosphate buffer, pH 7.0 for 24 h with 4 buffer changes after every 2 h. The molecular mass, and purity of recombinant proteins, *Ba*GH5-WT or

TH-2393_146106034
This article is protected by copyright. All rights reserved.

*Ba*GH5-UV2 were confirmed by SDS-PAGE using 12% (w/v) gel. The protein concentration of *Ba*GH5-WT or *Ba*GH5-UV2 was determined by Bradford method (Bradford *et al.*, 1976).

Comparison of temperature optima, and thermal stability of *Ba*GH5-WT and *Ba*GH5-UV2

The temperature optima of the purified recombinant *Ba*GH5-WT or *Ba*GH5-UV2 for enzyme assays was determined by reducing sugars estimation method (Nelson, 1944, Somogyi, 1945). The reaction mixture containing 10 µL of *Ba*GH5-WT enzyme (0.1 mg/mL) or 10 µL of *Ba*GH5-UV2 enzyme (0.025 mg/mL) at 1% (w/v) CMC-Na in 50 mM sodium acetate buffer, pH 5.5 was incubated in the temperature range, 30°C-80°C for 3 min. The temperature stability of purified recombinant *Ba*GH5-WT or *Ba*GH5-UV2 was carried out by pre-incubating the 50 µL of *Ba*GH5-WT (0.3 mg/mL) or 50 µL of *Ba*GH5-UV2 (0.1 mg/mL) at different temperatures (30 to 75°C) for 60 min. The residual enzyme activity for each sample was calculated by estimating the released reducing sugars by method reported by Nelson (1944), and Somogyi (1945).

Comparison of pH effect on *Ba*GH5-WT and *Ba*GH5-UV2 activity, and stability

The pH optima of purified recombinant *Ba*GH5-WT or *Ba*GH5-UV2 enzymes was determined by taking 10 µL of *Ba*GH5-WT enzyme (0.1 mg/mL) or 10 µL of *Ba*GH5-UV2 enzyme (0.025 mg/mL) with 90 µL of 1% (w/v) CMC-Na in different buffers; 50 mM sodium acetate (pH 3.5-6), 50 mM citrate phosphate (pH 3.5-7), 50 mM sodium phosphate buffer (pH 6.0-8.0), and 50 mM Tris-HCl (pH 7-8). The reaction mixture at different pH for *Ba*GH5-WT were incubated at 55°C for 3 min, and for *Ba*GH5-UV2 at 65°C for 3 min and CMCase activity was assayed by reducing

TH-2393_146106034
This article is protected by copyright. All rights reserved.

sugar estimation using the method of Nelson (1944), and Somogyi (1945). The pH stability studies of purified *Ba*GH5-WT or *Ba*GH5-UV2 enzyme(s) was carried out by pre-incubating the 10 µL of *Ba*GH5-WT (0.3 mg/mL) or 10 µL of *Ba*GH5-UV2 (0.1 mg/mL) at 35°C for 60 min against different buffers in the pH range, 3.5 to 8.0. The residual enzyme activity of each sample was calculated by estimating the released reducing sugars by method reported by Nelson (1944), and Somogyi (1945).

Comparison of specific activity of purified recombinant *Ba*GH5-WT and *Ba*GH5-UV2 against commercial substrates

The specific activity of purified recombinant *Ba*GH5-WT and *Ba*GH5-UV2 was determined against commercial substrates, β-D-glucan, lichenan, CMC-Na, avicel, cellulose powder, laminarin, galactomannan, and birchwood xylan. The 100 µL of reaction mixture contained 10 µL of enzyme (0.1 mg/mL *Ba*GH5-WT or 0.025 mg/mL *Ba*GH5-UV2), and 90 µL of 1.0% (w/v) substrate at final concentration dissolved in 50 mM sodium acetate buffer, pH 5.5. The enzyme reaction(s) was incubated for 3 min, for *Ba*GH5-WT at 55°C, and for *Ba*GH5-UV2 at 65°C. The reactions of enzyme with avicel and oat spelt xylan were carried out with agitation at 160 rpm for 3 min. The enzyme activity was calculated by estimating the released reducing sugars by method reported by Nelson (1944), and Somogyi (1945) at 500 nm, on a UV-visible spectrophotometer (Varian, Cary 100 Bio) by using glucose, xylose, and mannose as standards. The β-glucosidase activity of *Ba*GH5-WT (0.1 mg/mL), and *Ba*GH5-UV2 (0.025 mg/mL) was determined by glucose oxidase-peroxidase (GOD-POD) method against 1% (w/v) cellobiose (Raabo *et al.*, 1960). The 100 µL of reaction mixture contained 10 µL of the enzyme in 1% (w/v) cellobiose dissolved in 50 mM sodium acetate buffer, pH 5.5. The reaction mixture was incubated at 55°C for *Ba*GH5-WT, and at 65°C for *Ba*GH5-UV2 enzyme for 3

Accepted Article

TH-2393_146106034

This article is protected by copyright. All rights reserved.

min. After that the reaction was stopped by boiling it for 15 min in a water bath. Then 2 µL of reaction mixture were added to a well in 96 well microtitre plate followed by addition of 200 µL of glucose reagent L1 (from GOD-POD kit) and incubated at 37ºC for 15 min. The β-glucosidase activity was determined by oxidation of glucose by glucose oxidase (GOD) which form hydrogen peroxide (H_2O_2), and the formed H_2O_2 was detected by phenol-aminophenazone (chromogenic oxygen acceptor) in presence of peroxidase (POD) by taking the absorbance at 505 nm by using a UV-visible spectrophotometer (Varian, Cary 100 Bio) using glucose as standard.

Kinetic parameters of purified recombinant *Ba*GH5-WT and *Ba*GH5-UV2

The kinetic parameters, K_m and V_{max} for purified recombinant *Ba*GH5-WT, and *Ba*GH5-UV2 were estimated by fitting the initial rate data of enzyme activity to Michaelis-Menten equation (Michaelis *et al.*, 1913) by using the Lineweaver-Burk double reciprocal plot. The enzyme activity of *Ba*GH5-WT or *Ba*GH5-UV2 was determined against β-D-glucan, and CMC-Na concentration ranging from 0.01% to 2% (w/v) in 50 mM sodium acetate buffer (pH 5.5) for 3 min for *Ba*GH5-WT at 55°C, and for *Ba*GH5-UV2 at 65°C. The enzyme activity was calculated by estimating the reducing sugars by the method reported by Nelson (1944), and Somogyi (1945). The kinetic parameters from the best fit were determined by GraphPad Prism software (GraphPad Software Inc., San 138 Diego, CA).

Homology modeling of *Ba*GH5-WT, and *Ba*GH5-UV2, and quality assessment

The 3-D structure of wild-type *Ba*GH5-WT was best predicted by online server RaptorX (http://raptorx.uchicago.edu/StructPredV2/predict/). The *Ba*GH5-WT protein sequence along with the pET28a(+) vector sequence (23 amino acid residues at N-terminal) was uploaded on the online server. The online server RaptorX

TH-2393_146106034
This article is protected by copyright. All rights reserved.

(Kallberg *et al.,* 2012) predicts the 3-D modeled structure of *Ba*GH5-WT as single PDB file by using the sequence with highest homology was used as a template from BLAST pair-wise alignment. The 3-D modeled structure of *Ba*GH5-UV2 protein was generated after mutating the Asp256 residue to Gly in Swiss PDB viewer software. The energy of modeled wild-type *Ba*GH5-WT, and mutant *Ba*GH5-UV2 structures were minimized by using the Swiss PDB viewer software. Swiss PDB viewer not only minimizes the energy of the input structure but also repairs the deformities in the modeled 3-D structure by using a GROMOS 43B1 force field (Tran *et al.,* 1996). The quality of energy minimized 3-D structure of *Ba*GH5-WT, and *Ba*GH5-UV2 was assessed by submitting the structure of both the proteins at verification server (http://services.mbi.ucla.edu/SAVES/). In this server, the structure validation parameters Ramachandran plot by using PROCHECK server (Laskowski *et al.,* 1993), Z-score deviation by using ProSA server (Wiederstein *et al.,* 2007), and verify 3D (Eisenberg *et al.,* 1997) were predicted.

Prediction of divergence at 256 position, active site residues, and mechanism of enzyme action

The divergence in amino acid residues at position 256 were determined by Multiple Sequence Alignment (MSA) of 16 homologues (similarity in range of 27%-89% against protein-protein PDB blast) of bacterial cellulases of GH5 family from *Bacillus subtilis* (PDB ID: 3PZT|A, and PDB ID: 1LF1|A), *Salipaludibacillus agaradhaerens* (PDB ID: 1E5J|A, and PDB ID: 1H11|A), *Bacillus agaradherans* (PDB ID: 1A3H|A), *Geobacillus* sp. (PDB ID: 4XZB|A), *Bacillus halodurans* (PDB ID: 4V2X|A), *Cytophaga hutchinsonii* (PDB ID: 5IHS|A), *Bacillus sp.* KSM-635 (PDB ID: 1G01|A), *Caldicellulosiruptor saccharolyticus* (PDB ID:5ECU|A),

TH-2393_146106034
This article is protected by copyright. All rights reserved.

Metagenome holotolerant cellulase (PDB ID: 5FIP|A), *Thermobifida fusca* (PDB ID: 2CKS|A), *Erwinia Chrysanthemi* (PDB ID: 1EGZ|A), *Pseudoalteromonas haloplanktis* (PDB ID: 1TVN|A), and *Hungateiclostridium thermocellum* (PDB ID: 5BYW|A, and PDB ID: 4U3A|A) by using Clustal Omega tool (http://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers *et al.*, 2011). The MSA file generated was viewed under ESPript 3.0 (Robert *et al.*, 2014) to identify the variation or divergence in amino acid residues at 256 position. Further, 3-D model structure superposition of *Ba*GH5-WT, and *Ba*GH5-UV2 with PDB structure of endoglucanase PDB ID: 3PZT|A from *Bacillus subtilis*, PDB ID: 1E5J|A from *Salipaludibacillus agaradhaerens* and PDB ID: 5IHS|A from *Cytophaga hutchinsonii* was done in PyMOL 2.0 software to know the catalytically important amino acids residues, and analysed for structure superposition RMSD by Dali server. The distances between the catalytic amino acid residues of *Ba*GH5-WT was measured by using PyMOL 2.0 (Schrodinger *et al.*, 2017) software to infer the hydrolytic mechanism.

MD simulation of modeled *Ba*GH5-WT, and *Ba*GH5-UV2 structure

MD simulation studies for modeled wild-type *Ba*GH5-WT, and its mutant *Ba*GH5-UV2 structure was carried out by using GROMACS v 5.14 software to explore the possible mechanism of increase in catalytic efficiency, and pH stability. The GROMOS96 53a6 was used as a force field for calculating the protein forces. The modeled *Ba*GH5-WT or *Ba*GH5-UV2 was placed inside a cubic box of dimensions 6.38 × 7.73 × 10.12 and volume 1949.8 nm^3 with single point charge, and water molecules. The 17 Na^+ were used as counter ions to neutralize the negative charges on *Ba*GH5-WT or *Ba*GH5-UV2. The energy of modeled *Ba*GH5-WT and *Ba*GH5-UV2 structure was minimized by using the steepest descent method using

TH-2393_146106034
This article is protected by copyright. All rights reserved.

50,000 iteration steps with cut-off till 1000 $kJmol^{-1}$. Then each system was equilibrated in NVT ensemble for 500 ps at the constant number of particles, volume, and temperature equilibration of the whole system. The production run for each system was carried out for 80 ns with NPT ensemble acquiring 2 fs integration time. The modeled *Ba*GH5-WT or *Ba*GH5-UV2 structure was studied as a time-dependent function throughout the simulation to verify their stability in the solvent system. The root-mean-square deviation (RMSD), radius of gyration (Rg), root-mean-square fluctuation (RMSF), and solvent accessible surface area (SASA) of both the simulated structure(s) were analysed by gmx rms, gmx gyrate, gmx rmsf, and gmx sasa program, respectively.

Ligand-binding analysis of simulated *Ba*GH5-WT, and *Ba*GH5-UV2 structure by molecular docking

Molecular docking studies were carried out to understand the interactions between enzyme (*Ba*GH5-WT or *Ba*GH5-UV2), and the ligands (cello-oligosaccharides) *viz.* cellobiose, cellotriose, cellotetraose, cellopentaose, and cellohexaose, respectively, using Autodock 4.2.1. The 3-D modeled structure of *Ba*GH5-WT or *Ba*GH5-UV2 protein after 80 ns of MD simulation was used in molecular docking analysis. In Autodock 4.2.1, inbuilt lamarckian genetic algorithm (GA), and empirical free energy scoring function were used to analyse the interactions of ligand with the enzyme (Goodsell *et al.*, 1990). All the ligand molecules were retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov). Input files of all the ligands were prepared by adding polar hydrogens and assigning atomic partial charge to the ligands by Gasteiger method, and to protein by Kollman method. This was followed by removing all non-polar hydrogens and merging their charges to the

TH-2393_146106034
This article is protected by copyright. All rights reserved.

carbon atoms. In Autodock 4.2.1, grid-based method was used to evaluate the energy where grid points contained pre-calculated affinities against different atom of the ligand. Therefore, grid maps were assigned to each atom type present in ligand molecule as well as protein accompanied by electrostatic, and desolvation maps using the AutoGrid (Huey *et al.*, 2007). Grid box was set at 110, 100, 96 (x,y,z) coordinate points with 0.375 Å grid spacing for modeled simulated *Ba*GH5-WT and 90, 92, 96 (x,y,z) coordinate points with 0.375 Å grid spacing for modeled simulated *Ba*GH5-UV2, where it covered the entire active site pocket. For each ligand, 80 Genetic Algorithm (GA) runs were performed. The docked confirmation of protein-ligand complex which has the minimum lowest Gibb's free energy of binding (ΔG°), were saved. The resulting *Ba*GH5-WT-ligand complexes or *Ba*GH5-UV2-ligand complexes were analysed in PyMOL 2.1. The 2-D protein-ligand complex was generated by LigPlot software for analysing the change in possible hydrogen bonds, and hydrophobic interactions.

Results and Discussion

Cloning, expression, and purification of recombinant *Ba*GH5-WT and *Ba*GH5-UV2

The gene encoding *Ba*GH5-WT and *Ba*GH5-UV2 were cloned in pET28a(+) vector (Fig. S1 and Fig. S2). The recombinant protein sequence of *Ba*GH5-WT and *Ba*GH5-UV2 with extra 23 amino acid residues of pET-28a(+) vector sequence at N-terminus showed the mutation position was at 256 position where residue Asp256 changed to Gly. Recombinant *Ba*GH5-WT and *Ba*GH5-UV2 proteins showed expression in *E. coli* BL21 (DE3)pLysS cells after IPTG induction grown at 24°C for 18 h at 180 rpm (Fig. 1A and 1B, lane 3 and 4). Both the proteins were expressed as

TH-2393_146106034
This article is protected by copyright. All rights reserved.

soluble proteins in the cell free extracts (Fig. 1A and 1B, lane 4). The purified *Ba*GH5-WT (Fig. 1A, lane 6), and *Ba*GH5-UV2 (Fig. 1B, lane 6) protein showed homogeneous bands each of molecular size approximately, 58.0 kDa which were in agreement with the theoretical size of 57.6 kDa. Each of the recombinant proteins were eluted in 4x1 mL fractions from 100 mL culture. After dialysis, the concentration of purified *Ba*GH5-WT, and *Ba*GH5-UV2 proteins were 0.19 mg/mL and 0.18 mg/mL, respectively.

Comparison of temperature optima, and thermal stability of *Ba*GH5-WT and *Ba*GH5-UV2 activity

The purified wild-type endoglucanase, *Ba*GH5-WT gave maximum enzyme activity of 4.5 U/mg (considered as 100% CMCase activity) against 1% (w/v) CMC-Na at temperature optima 55°C, whereas, its mutant *Ba*GH5-UV2 with single substitution of Asp to Gly residue at 256 position showed 10°C shift to a higher temperature optimum (65°C), and also showed enhanced (~10-fold) CMCase activity of 43.6 U/mg (considered as 100% CMCase activity) (Fig. 2A). Moreover, *Ba*GH5-UV2 showed 43% of relative activity at 80°C, whereas *Ba*GH5-WT enzyme showed negligible activity at this temperature. In contrast, Zheng *et al.*, (2018) reported that substitution of Asp233 to Ala, and Asn233 to Gly in *Gt*Cel5A from *Gloeophyllum trabeum* by site-directed mutagenesis showed 10°C decrease in temperature optima. *Ba*GH5-UV2 did not show any difference related to thermostability as it retained 100% of its residual CMCase activity up to 45°C for 60 min, similar to *Ba*GH5-WT (Fig. 2B). This suggested that *Ba*GH5-UV2 can be efficiently used in simultaneous saccharification and fermentation (SSF) at 30°C.

Comparison of pH effect on *Ba*GH5-WT and *Ba*GH5-UV2 activity

TH-2393_146106034
This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

The enzyme, *Ba*GH5-UV2 showed broader pH optima range, 5.0-6.0 with 50 mM sodium acetate buffer against 1% (w/v) CMC-Na at 65°C (Fig. 2D) as compared with the wild-type enzyme, *Ba*GH5-WT which showed narrow range pH optima range, 5.5-6.0 against 1% (w/v) CMC-Na at 55°C (Fig. 2C). *Ba*GH5-UV2 retained 88% of relative CMCase activity in 50 mM sodium acetate buffer at pH 4.5, whereas *Ba*GH5-WT retained only 22% of relative activity at the same pH. This result displayed that substitution of Asp256 to Gly in *Ba*GH5-UV2 shifted the optimum pH range towards slightly more acidic pH. Similar results were reported that earlier the substitution of Asp233 to Val in *Gt*Cel5A from *Gloeophyllum trabeum* showed shift in pH optima towards acidic value from pH 4.0 to 5.0 (Zheng *et al.*, 2018).

Comparison of pH effect on *Ba*GH5-WT and *Ba*GH5-UV2 stability

*Ba*GH5-WT enzyme retained 100% of residual CMCase activity in neutral, and slightly alkaline pH range, pH 7.0-7.5 (Fig. 2E), whereas, its mutant, *Ba*GH5-UV2 retained 100% of residual CMCase activity in acidic pH range, 5.5-6.0 after 60 min of incubation in 20 mM of respective buffer at 35°C (Fig. 2F). The enzyme, *Ba*GH5-UV2 retained more than 90% of residual CMCase activity over a broad acidic pH range from 5.0 to 7.0, whereas, *Ba*GH5-WT retained around 80% of relative CMCase activity in a narrow acidic pH range from 6.0 to 6.5 after 60 min of incubation in 20 mM of respective buffer at 35°C (Fig. 2E and 2F). Glycine residue is an uncharged residue, therefore substitution of aspartate a negatively charged residue to an uncharged residue changed the microenvironment of the catalytic core, and increased the pH stability. The pH stability shift of *Ba*GH5-UV2 enzyme towards the acidic conditions makes it more favourable for bioethanol production by *Saccharomyces cerevisiae* in SSF process. Zheng *et al.*, (2018) reported that substitution of Asn233 to

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Asp or Asn233 to Gly in *Gt*Cel5 increases the stability of endoglucanase under both acidic and alkaline conditions.

Comparison of specific activity of purified recombinant *Ba*GH5-WT and *Ba*GH5-UV2 against commercial substrates

Mutant *Ba*GH5-UV2 showed highest specific activity 94.6 U/mg against barley β-glucan followed by lichenan, 77.1 U/mg (Table 1). It showed moderate specific activity against CMC-Na (43.6 U/mg), and low specific activity against crystalline cellulosic substrate avicel (11.4 U/mg). These results of specific activities of *Ba*GH5-UV2 against all the above four cellulosic substrates were approximately, 10-fold higher than the specific activity displayed by wild-type enzyme *Ba*GH5-WT. Both the wild-type, *Ba*GH5-WT, and mutant *Ba*GH5-UV2 enzymes did not show any detectable activity against cellobiose, laminarin, birchwood xylan, and galactomannan (Table 1). Liang *et al.,* (2011) reported that the substitution of Leu321 to Val in Cel5A by directed evolution increase the specific activity of endoglucanase by 1.3-fold. Li *et al.,* (2017) reported an increase in the mannanase activity of *Rm*Man5A from *Rhizomucor miehei* by directed evolution by 3-fold in m*Rm*Man5A mutant by error-prone PCR. Zheng *et al.,* (2018) reported that the mutants Asn233Ala or Asn233Gly of *Gt*Cel5 generated by site-directed mutagenesis showed 1.3- or 1.7-fold increase in the specific activity, respectively, against barley β-glucan.

Kinetic parameters of purified recombinant *Ba*GH5-WT and *Ba*GH5-UV2

The mutant enzyme *Ba*GH5-UV2 showed 8.5-fold (100.5 U/mg), and 7-fold (42.1 U/mg) higher V_{max} against β-glucan, and CMC-Na, respectively, as compared with the wild-type *Ba*GH5-WT enzyme (Table 2). *Ba*GH5-UV2 enzyme showed 2.1, and 3.1-fold lower K_m against β-glucan (0.07 mg/mL), and CMC-Na (0.12 mg/mL)

TH-2393_146106034
This article is protected by copyright. All rights reserved.

respectively, as compared with wild-type *Ba*GH5-WT enzyme (Table 2). A higher V_{max} and lower K_m value is required for an enzyme to have higher affinity towards their substrates (Singh *et al.*, 2020). The catalytic efficiency (k_{cat}/K_m) of mutant enzyme *Ba*GH5-UV2 increased to 18.4-fold (8.3×10^4 $mg^{-1}mL^1min^{-1}$), and 21.9-fold (2.0×10^4 $mg^{-1}mL^1min^{-1}$) against β-glucan, and CMC-Na, respectively, as compared with *Ba*GH5-WT (Table 2). This showed the increased potential of mutant enzyme, *Ba*GH5-UV2 in hydrolysis of cellulosic polysaccharides into reducing sugars.

3-D structure comparison of *Ba*GH5-WT and *Ba*GH5-UV2

Homology modeling of BaGH5-WT and BaGH5-UV2

The best modeled 3-D structure of *Ba*GH5-WT was determined by RaptorX server by using PDB ID: 3PZT|A (*Bacillus subtilis*), PDB ID: 2L8A|A (*Bacillus subtilis*), and PDB ID: 1E5J|A (*Salipaludibacillus agaradhaerens*) as protein templates with query coverage of 68%, 27% and 56%, respectively, with sequence identity 89%, 91.1% and 69.5%, respectively (Fig. 3A). The energy minimized modeled 3-D structure of *Ba*GH5-WT consists of two domains, an N-terminal classical $(\alpha/\beta)_8$-TIM barrel fold GH5 domain, and a C-terminal β-sandwich fold CBM3 domain connected by a linker (Fig. 3A). Surface view of modeled *Ba*GH5-WT structure showed that the loops connecting the $\beta_1\alpha_1$ (L1), $\beta_2\alpha_2$ (L2), $\beta_3\alpha_3$ (L5), $\beta_4\alpha_4$ (L7), $\beta_5\alpha_5$ (L9), $\beta_6\alpha_6$ (L11), $\beta_7\alpha_7$ (L13), and $\alpha_7\beta_8$ (L14) are involved in the formation of catalytic core of TIM barrel fold (Fig. 3B). Yan [35] reported that the endoglucanases belonging to GH5 possess the $(\alpha/\beta)_8$-TIM barrel fold architecture.

The 3-D structure of *Ba*GH5-UV2 was constructed after mutating the amino acid residue Asp256 to Gly (D256G) in 3-D structure of modeled *Ba*GH5-WT protein (before energy minimization structure) by using Swiss-PDB viewer software. The energy

TH-2393_146106034
This article is protected by copyright. All rights reserved.

minimized structure of *Ba*GH5-UV2 showed the position of D256G mutation was present in loop (L11), which involved in the catalytic core formation as shown in red colour (Fig. 3C). Earlier studies of Lee *et al.*, (2011) also confirmed that the $\beta_6\alpha_6$ loop is involved in the formation of catalytic core of GH5 endoglucanases. The loop connecting the βα and αβ regions of TIM-barrel enzymes plays key role in the interactions between catalytic core, and substrate (Varrot *et al.*, 1999; Zhai *et al.*, 2015; Tu *et al.*, 2016). In general, most key mutations which lead to the evolution of properties, such as catalytic efficiency, specific activity, and optimal condition, are located on loop towards the surface of enzyme (Zheng *et al.*, 2018).

Quality assessment of 3-D modeled structures of BaGH5-WT and BaGH5-UV2

The quality of energy minimized *Ba*GH5-WT, and *Ba*GH5-UV2 structure was evaluated at saves server. The 3-D modeled structure of *Ba*GH5-WT validated by Ramachandran plot through PROCHECK showed approximately, 81.6% residues in most favoured region, 16.3% residues in additional allowed region, 1.9% residues in generously allowed region, and only 0.2% (Gln 163) residue in disallowed region (Fig. S3A) out of non-glycine, and non-proline residues of modeled structure. However, 3-D modeled structure of *Ba*GH5-UV2 validated by Ramachandran plot through PROCHECK showed approximately, 82.0% residues in most favoured region, 15.8% residues in additional allowed region, 1.9% residues in generously allowed region, and 0.2% (Gln 163) residues in the disallowed region (Fig. S4A) out of non-glycine, and non-proline residues of modeled structure. These results of Ramachandran plot displayed that the modeled 3-D structures of *Ba*GH5-WT and *Ba*GH5-UV2 followed the backbone phi (Φ), and psi (Ψ) dihedral angles at favourable positions. The VERIFY 3-D results of modeled *Ba*GH5-WT showed 88.66% of amino acid residues possess average 3D-1D

TH-2393_146106034
This article is protected by copyright. All rights reserved.

score >=0.2, and declared acceptable the modeled structure of *Ba*GH5-WT (Fig. S3B). However, the VERIFY 3-D results of modeled *Ba*GH5-UV2 showed 89.08% of amino acid residues possess average 3D-1D score >=0.2 and declared pass, the modeled structure of *Ba*GH5-UV2 (Fig. S4B). The ProSA results of modeled *Ba*GH5-WT showed that the protein resides in the X-ray zone with Z-scores of -7.55, indicating that the modeled structure was error free (Fig. S3C and D). However, the modeled *Ba*GH5-UV2 showed that the protein resides in the X-ray zone with Z-score of -7.85, indicating the error free modeled structure (Fig. S4C and D). These results showed that the modeled BaGH5-UV2 structure with only single substitution mutation gave improved 3-D structural quality by 0.5% with Ramachandran plot, and VERIFY 3-D analyses, and 3.9% by ProSA analysis, which might be contributing to the increase in catalytic efficiency (22-fold), temperature optima, and pH stability.

Prediction of divergence at 256 position in GH5 endoglucanase from *Bacillus amyloliquefaciens* SS35

The MSA of 16 homologues (sequence similarity in range of 27%-89%) of bacterial cellulases of GH5 family showed that the Gly256 of loop is a semi-conserved residue (Fig. 4). It was found that out of 16 homologues of GH5 endoglucanase, *Ba*GH5-WT from *Bacillus amyloliquefaciens* SS35, Cel5A (PDB ID: 2CKS|A) from *Thermobifida fusca*, and Cel5B (PDB ID: 4V2X|A) from *Bacillus halodurans* showed divergence in the conserved region at 256 position, where aspartate, alanine, and phenylalanine, respectively, were present instead of glycine residue (Fig. 4). Recombinant mutant enzyme, *Ba*GH5-UV2 from UV2 mutant strain (Singh *et al.*, 2020) of *Bacillus amyloliquefaciens* SS35 have substitution mutation at 256 position leading to Asp256Gly due to which catalytic efficiency, temperature optima, and pH

TH-2393_146106034
This article is protected by copyright. All rights reserved.

stability was increased. This suggested that the mutation of a divergent residue to a conserved residue near the ligand binding residues increases the catalytic efficiency of enzyme by increasing the stability of the catalytic core. Joo *et al.*, (2011) discussed that mutating the divergent amino acid into the consensus amino acid (represented in >50% of the homologs) which also fit in structure without destroying the helices or salt bridges nearly at 6 Å from the catalytically important residues improves the catalytic properties of enzymes.

Identification of active site residues, and mechanism of enzyme action

The search by Dali server displayed that the RMSD of *Ba*GH5-WT or *Ba*GH5-UV2 after superposition with GH5 endoglucanase PDB ID: 3PZT|A (*Bacillus subtilis*) was 3.1 Å with 298 C_α atoms, with endoglucanase PDB ID: 6GJF|A was 3.2 Å with 303 C_α atoms, with endoglucanase PDB ID: 1E5J|A (*Salipaludibacillus agaradhaerens*) was 3.2 Å with 302 C_α atoms, and with PDB ID: 5IHS|A (*Cytophaga hutchinsonii*) was 2.9 Å with 306 C_α atoms showing an excellent alignment. The amino acid residues taking part in the catalysis, and substrate binding were identified by superimposing the *Ba*GH5-WT (Red colour), and *Ba*GH5-UV2 (blue white colour) modeled structure with PDB structure of endoglucanase PDB ID: 3PZT|A (blue colour), PDB ID: 1E5J|A (yellow colour), and PDB ID: 5IHS|A (pale green) (Fig. 5A). The superimposition analysis displayed that the residues, R115, E192, W230, H252, Y254, T279, E280, and W314 are catalytically important for *Ba*GH5-WT/ *Ba*GH5-UV2. The residue, E192 acts as a catalytic acid/proton donor and E280 of *Ba*GH5-WT/ *Ba*GH5-UV2 as a catalytic nucleophile (Fig. 5B). The aromatic amino acid residues W230, and W314 act as the gate keepers, and H252, R115, and T279 act as polarizers during the substrate catalysis in *Ba*GH5-WT/*Ba*GH5-UV2 according to the earlier study of Santos *et al.*, (2012).

Accepted Article

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Moreover, Y254 of loop $\beta_6\alpha_6$ orient, and activate the E280 (equivalent to E289, as nucleophile by Shirai *et al.,* 2001), and along with loop $\beta_8\alpha_8$, it imparts flexibility to the catalytic cavity which causes an induced fit at the time of substrate binding. The substitution mutation of Asp 256 to Gly residue was near the binding site residues H252 and Y254 (Fig. 5C). In addition, the mutated residue D256G lying on the loop 11 (L11), increases its flexibility thereby resulting in improvement of fitting of induced ligand at the cavity by increasing the flexibility of catalytic core. The glycine amino acid residue lacks β-carbon, therefore, its substitution in protein backbone increases the conformational flexibility which allows a greater configurational entropy compared to other amino acids and help in increasing the catalytic efficiency of enzymes (Ma *et al.*, 2000). The estimated distance between two catalytic residues E192, and E280 of *Ba*GH5-WT, and *Ba*GH5-UV2 using PyMOL 2.0 was found to be similar *i.e.* 4.8 Å (Fig. 5B), this indicated that the *Ba*GH5-WT/*Ba*GH5-UV2 like other GH5 cellulases (Cel5A), showed retaining type hydrolytic mechanism which cleave β-1,4 glycosidic bond *via* a double displacement mechanism.

MD simulation of modeled *Ba*GH5-WT and *Ba*GH5-UV2 structures

MD simulation studies of modeled *Ba*GH5-WT and *Ba*GH5-UV2 structures were carried out for 80 ns. MD simulation of modeled protein structure over the defined trajectory leads to the more energetically stable conformations of a protein in aqueous solution which can be represented by the RMSD between the starting structure and those obtained during the simulation (Pandey *et al.,* 2011). The average backbone RMSD for wild-type *Ba*GH5-WT reaches a plateau phase value of 0.76 nm at 50 ns and remained stable till 80 ns (Fig. 6A) whereas, the RMSD profile for mutant *Ba*GH5-UV2 reaches a plateau phase value of 0.91 nm after 69 ns, and remained stable till 80 ns (Fig. 6A).

TH-2393_146106034
This article is protected by copyright. All rights reserved.

Accepted Article

These results indicated that the structure of mutant *Ba*GH5-UV2 undergoes more conformational changes along the simulation period of 69 ns in order to get the final stable conformation. Stable RMSD at the end of the simulation period of 80 ns suggested the MD simulation for *Ba*GH5-WT, and *Ba*GH5-UV2 is perfect for further analysis. The overall dimension, and global compactness of protein *Ba*GH5-WT, and *Ba*GH5-UV2 were determined by Rg values during the MD simulation. It predicts how regular secondary structures are compactly packed into 3-D structure. Therefore, for stable folded protein, relatively steady Rg value is maintained whereas, for not properly folded protein it will change over the time (Rogerson *et al.*, 2011). The fluctuation in the R_g value of *Ba*GH5-WT protein was between 2.77-3.03 nm in the time scale of 0-54 ns, and after that it was 2.84 nm at 69 ns, and then remained same till 80 ns (Fig. 6B). However, the fluctuation in R_g value of *Ba*GH5-UV2 protein was 2.72-2.84 nm in the time scale of 0-7.4 ns, and after that it decreased to 2.55 nm at 60 ns, then remain stable till the end up to 80 ns (Fig. 6B). These results of R_g values for *Ba*GH5-WT and *Ba*GH5-UV2 proteins depicts the stable conformation after 69 ns, and 60 ns, respectively. Moreover, the lower R_g value for *Ba*GH5-UV2 proteins (2.55 nm) as compared with *Ba*GH5-WT protein (2.84 nm) suggested that the compactness of 3-D structure of *Ba*GH5-UV2 protein increased which increased the globularity of the protein and thus, improved the kinetic properties of enzyme. RMSF during MD simulation measures the displacement of a particular atom or group of atoms with respect to the reference structure. The amino acid residues showing higher RMSF values have highest flexibility for α-carbon among all the amino residues in a polypeptide chain. However, amino acid residues showing lower RMSF values showed lowest flexibility for α-carbon (Nath *et al.*, 2019). RMSF values for catalytic residues, E192, and E280 were 0.243 nm, and 0.164 nm, respectively for *Ba*GH5-WT, and 0.165 nm, and 0.101 nm, respectively, for *Ba*GH5-UV2 (Fig. 6C). This

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

decrease in RMSF values for catalytic residues of *Ba*GH5-UV2 enzyme showed more rigid and stable catalytic residues for enzyme catalysis. RMSF values for loop L11 increases to 0.58 nm in *Ba*GH5-UV2 than *Ba*GH5-WT, 0.43 nm (Fig. 6C). This confirmed the increased flexibility of loop L11 in *Ba*GH5-UV2 because of which involvement of loop L11 increases in catalysis and results increased catalytic efficiency by 22 fold against CMC-Na, increased temperature optima, and increased pH stability in wide acidic range. According to the previous study of Pandey *et al.*, (2019), the catalytic residues should be less flexible, loop, and linker should be more flexible in order to maintain the activity of the enzyme. SASA for *Ba*GH5-WT showed higher fluctuations up to the maximum value of 250 nm^2 till 24.5 ns, and after that it became constant with little fluctuation with average value of 242.7 nm^2 between 24.5 ns, and 80 ns (Fig. 6D). However, the SASA of *Ba*GH5-UV2 for time period of simulation upto 0-0.92 ns showed maximum fluctuation between 243.28 nm^2 to 257.57 nm^2 after that it decreased to 225 nm^2 till 30 ns, and further remained more or less stable till 80 ns with an average value of 225.7 nm^2 (Fig. 6D). This suggested that the overall exposure of the *Ba*GH5-UV2 towards the solvent area decreases to 7% as compared to the *Ba*GH5-WT which depicted that the accessibility of the solvent in the catalytic sites was lower as compared with *Ba*GH5-WT. According to the earlier reports of Yaacob *et al.*, (2019) for protein stability a lesser protein-solvent contact area is favourable because for larger surface area, protein is interacting more with the solvent. This results in water penetration into the interior of a protein, where it forms hydrogen bonds with the side chain residues, and main chain of the polypeptide, because of which, the hydrogen bond interactions among amino acid residues decrease, and weaken the internal protein interactions (Badieyan *et al.*, 2012). The superposed structure of 3-D modeled, and 80 ns MD simulated structure of *Ba*GH5-WT (Fig. 6E), and *Ba*GH5-UV2 (Fig. 6F), depicted the lesser fluctuations

TH-2393_146106034

This article is protected by copyright. All rights reserved.

over simulation time periods suggesting that both the structure can be further used in ligand binding analysis by molecular docking.

Comparison of ligand binding analysis for *Ba*GH5-WT and *Ba*GH5-UV2

The interactions between the ligands (cello-oligosaccharides), and *Ba*GH5-WT or *Ba*GH5-UV2 enzyme were studied by molecular docking analysis. The enzyme, *Ba*GH5-UV2 showed 1.1, 1.6, 2.6, and 2.5-fold increase in negative Gibb's free energy of binding (ΔG°) against cellobiose, cellotriose, cellotetraose, and cellopentaose, respectively, as compared with *Ba*GH5-WT (Table 3). This showed that the substitution of Asp256 to Gly residue in *Ba*GH5-UV2 results in maximum increase in Gibb's free energy (ΔG°) against cellotetraose (2.6 fold) among all other cello-oligosaccharides as compared with *Ba*GH5-WT. This depicted that *Ba*GH5-UV2 can catalyze cellotetraose more efficiently than *Ba*GH5-WT. The amino acid residues of *Ba*GH5-WT and *Ba*GH5-UV2 involved in non-covalent interactions with different ligands (cello-oligosaccharides) are listed in Table 3. *Ba*GH5-UV2 showed increased interactions with more number of amino acid residues involved in different hydrogen-bonding network with ligands unlike *Ba*GH5-WT (Table 3). The comparison of *Ba*GH5-WT-cellotetraose complex (Fig. 7A and B) with *Ba*GH5-UV2-cellotetraose complex (Fig. 7C and D) showed that, *Ba*GH5-UV2 forms more number of hydrogen bonds as compared to *Ba*GH5-WT. This has happened because of the increased flexibility of loop 11 ($\beta_6\alpha_6$) due to the glycine substitution which causing the altered, and enhanced hydrogen bond interactions near the active site, that might be allowing the ligand to bind more easily, and thus catalysing its hydrolysis more efficiently. According to an earlier report of Wang *et al.*, (2016) hydrogen bonds play a crucial role in ligand recognition and binding. Single mutant variants, Asn233 to Ala or Asn233 to Gly of

TH-2393_146106034
This article is protected by copyright. All rights reserved.

*Gt*Cel5 (endoglucanase) from *Gloeophyllum trabeum* at loop 6 formed direct hydrogen bonds with cellotetraose as compared with the wild-type enzyme, and showed increased catalytic efficiency (Zheng *et al.,* 2018).

Conclusion

The genes encoding endoglucanases *Ba*GH5-WT or *Ba*GH5-UV2 of family GH5 was cloned into pET28a(+) expression vector and expressed in *E. coli* BL-21(DE3)pLysS cells as soluble, and single band proteins. The recombinant mutant *Ba*GH5-UV2 showed over 9-fold higher specific activity than the recombinant wild-type *Ba*GH5-WT against β-glucan (10.2 U/mg), and CMC-Na (4.5 U/mg). The temperature optimal studies depicted the 10°C shift in temperature optima for *Ba*GH5-UV2 than *Ba*GH5-WT. The pH stability studies for endoglucanase activity showed the wide acidic pH range (5.0-7.0) for *Ba*GH5-UV2 as compared with the narrow basic pH range (7.0-7.5) for *Ba*GH5-WT. The catalytic efficiency of *Ba*GH5-UV2 was enhanced by 18 and 22-fold against β-Glucan and CMC-Na, respectively, as compared with *Ba*GH5-WT. The structure analysis of *Ba*GH5-UV2 showed the glycine residue at position 256 is present on $\alpha_6\beta_6$ (L11) loop along with the two ligand binding residues His252, and Tyr254. The quality of the modeled 3-D structure of *Ba*GH5-UV2 showed slightly improved structure than *Ba*GH5-WT. This suggested that the glycine substitution at 256 position results in more stable structure as glycine being smallest amino acid imparts flexibility to the loop leading to enhanced ligand binding. This was confirmed by MD simulation studies which showed decreased Rg for *Ba*GH5-UV2 displaying more globular nature than *Ba*GH5-WT, and the decreased RMSF for catalytic residues showed stable active site for catalysis. The single substitution, Asp256Gly in the loop, L11 showed more hydrogen bond interactions

TH-2393_146106034
This article is protected by copyright. All rights reserved.

between the catalytic core, and cello-oligosaccharides, which increases the substrate affinity, and thus the catalytic efficiency of *Ba*GH5-UV2. Considering the importance of GH5 endoglucanase in biomass conversion to oligosaccharides, the findings are valuable for the protein engineering of GH5 endoglucanase from the viewpoints of research, and industrial applications.

Acknowledgments

The research work was supported by the Centre for Bioenergy DBT-Pan-IIT Grant (BT/EB/PAN-IIT 2012) by the Department of Biotechnology, Ministry of Science and Technology, New Delhi, India. The fellowship to Shweta Singh was supported by funding from the Ministry of Human Resource and Development (MHRD), Government of India. The authors acknowledge the use of Param-Ishan facility for MD simulation studies at IIT Guwahati.

Declaration of interest

The authors declare that they have no conflict of interest.

Appendices A: Supplementary material

References

Akhtar N, Gupta K, Goyal D, Goyal A. 2019. Lignocellulosic biomass characteristics for bioenergy applications: An overview. Environ Eng Manag J 18:367-383.

Badieyan S, Bevan DR, Zhang C. 2012. A salt-bridge controlled by ligand binding modulates the hydrolysis reaction in a GH5 endoglucanase. Protein Eng Des Sel 25:223-233.

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

Badieyan S, Bevan DR, Zhang C. 2012. Study and Design of Stability in GH5 Cellulases. Biotechnol. Bioeng. 109: 31-44.

Ballester PJ, Mitchell JB. 2010. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. Bioinformatics 26:1169-1175.

Bradford M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem 72:248-254.

Eisenberg D, Luthy R, Bowie JU. 1997. [20] VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol 277:396-404.

Goodsell DS, Olsonn AJ. 1990. Automated docking of substrates to proteins by simulated annealing. Proteins: Struct., Funct., Bioinf. 8:195-202.

Hodgson J. 1994. The changing bulk biocatalyst market. Nat. Biotechnol 12:789-90.

Huey R, Morris GM, Olson AJ, Goodsell DS. 2007. A semiempirical free energy force field with charge-based desolvation. J. Comput. Chem. 28:1145-1152.

Joo JC, Pack SP, Kim YH, Yoo YJ. 2011. Thermostabilization of *Bacillus circulans* xylanase: computational optimization of unstable residues based on thermal fluctuation analysis. J. Biotechnol 151:56-65.

Kallberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu J. 2012. Template-based protein structure modeling using the RaptorX web server. Nat. Protoc 7:1511.

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

Kim BK, Lee BH, Lee YJ, Jin IH, Chung CH, Lee JW. 2009. Purification and characterization of carboxymethylcellulase isolated from a marine bacterium, *Bacillus subtilis* subsp. subtilis A-53. Enzyme Microb. Technol. 44:411-416.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallog 26:283-291.

Lee TM, Farrow MF, Arnold FH, Mayo SL. 2011. A structural study of *Hypocrea jecorina* Cel5A. Protein Sci 20:1935-1940.

Li YX, Yi P, Yan QJ, Qin Z, Liu XQ, Jiang ZQ. 2017. Directed evolution of a β-mannanase from *Rhizomucor miehei* to improve catalytic activity in acidic and thermophilic conditions. Biotechnol Biofuels 10:143.

Liang C, Fioroni M, Rodríguez-Ropero F, Xue Y, Schwaneberg U, Ma Y. 2011. Directed evolution of a thermophilic endoglucanase (Cel5A) into highly active Cel5A variants with an expanded temperature profile. J. Biotechnol. 154:46-53.

Liang PH, Lin WL, Hsieh HY, Lin TY, Chen CH, Tewary SK, Lee HL, Yuan SF, Yang B, Yao JY, Ho MC. 2018. A flexible loop for mannan recognition and activity enhancement in a bifunctional glycoside hydrolase family 5. Biochim. Biophys. Acta, Gen. Subj. 1862:513-521.

Ma YF, Eglinton JK, Evans DE, Logue SJ, Langridge P. 2000. Removal of the four C-terminal glycine-rich repeats enhances the thermostability and substrate binding affinity of barley β-amylase. Biochemistry 39:13350-13355.

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Michaelis L, Menten ML. 1913. Die kinetic der invertinwirkung. Biochem J. 49:333-369.

Nelson N. 1944. A photometric adaptation of the Somogyi method for the determination of glucose. J Biol Chem 153:375-380.

Pandey B, Grover S, Kaur J, Grover A. 2019. Analysis of mutations leading to para-aminosalicylic acid resistance in *Mycobacterium tuberculosis*. Sci. Rep 9:1-15.

Raabo BE, Terkildsen TC. 1960. On the enzymatic determination of blood glucose. Scand J Clin Lab Inv 12:402-407.

Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42:W320-W324.

Rogerson P, Arteca GA. 2011. Molecular size scaling in families of protein native folds. J Math Chem. 49:1493.

Ruiz DM, Turowski VR, Murakami MT. 2016. Effects of the linker region on the structure and function of modular GH5 cellulases. Sci. Rep. 6:28504.

Santos CR, Paiva JH, Sforça ML, Neves JL, Navarro RZ, Cota J, Akao PK, Hoffmam ZB, Meza AN, Smetana JH, Nogueira ML, Polikarpov I, Xavier-Neto J, Squina FM, Ward RJ, Ruller R, Zeri AC, Murakami M. 2012. Dissecting structure-function-stability relationships of a thermostable GH5-CBM3 cellulase from *Bacillus subtilis* 168. Biochem J. 441:95-104.

Schrodinger L. 2017. The PyMOL molecular graphics system, Version 2.0. Retrieved from http://www.pymol.org/.

Accepted Article

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Schulein M. 2000. Protein engineering of cellulases. Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol. 1543:239-252.

Shirai T, Ishida H, Noda JI, Yamane T, Ozaki K, Hakamada Y, Ito S. 2001. Crystal structure of alkaline cellulase K: insight into the alkaline adaptation of an industrial enzyme. J. Mol. Biol. 310:1079-1087.

Shobharani P, Yogesh D, Halami PM, Sachindra NM. 2013. Potential of cellulase from *Bacillus megaterium* for hydrolysis of sargassum. J. Aquat. Food Prod. T. 22:520-535.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Thompson JD. 2011. Fast scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 7:539.

Singh S, Dhillon A, Goyal A. 2020. Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis. Renew Energ 151:1124-1133.

Singh S, Moholkar VS, Goyal A. 2013. Isolation, identification, and characterization of a cellulolytic *Bacillus amyloliquefaciens* strain SS35 from Rhinoceros dung. ISRN Microbiology 728134.

Somogyi M. 1945. A new reagent for the determination of sugars. J Biol Chem. 160:61-68.

Taylor E, Goyal A, Guerreiro CIPD, Prates JAM, Money AV, Ferry N, Morland C, Planas A, Macdonald JA, Stick RV, Gilbert HJ, Fontes CMGA, Davies GJ. 2005. How family 26 glycoside hydrolases orchestrate catalysis on different

Accepted Article

TH-2393_146106034

This article is protected by copyright. All rights reserved.

polysaccharides? Structure and activity of a *Clostridium thermocellum* lichenase, *Ct*Lic26A. J Biol Chem. 280:32761-32767.

Tran F, Billeter S, Eising AA, Hunenberger PH, Kruger P, Mark E, Tironi IG. 1996. Biomolecular Simulation: The GROMOS96 manual and user guide.

Tu T, Pan X, Meng K, Luo H, Ma R, Wang Y, Yao B. 2016. Substitution of a non-active-site residue located on the T3 loop increased the catalytic efficiency of endo-polygalacturonases. Process Biochem 51:1230-8.

Varrot A, Schulein M, Davies GJ. 1999. Structural changes of the active site tunnel of *Humicola insolens* cellobiohydrolase, Cel6A, upon oligosaccharide binding. Biochemistry 38:8884-8891.

Vohra M, Manwar J, Manmode R, Padgilwar S, Patil S. 2014. Bioethanol production: feedstock and current technologies. J. Environ. Chem. Eng. 2:573-584.

Wang X, Huang H, Xie X, Ma R, Bai Y, Zheng F, Luo H. 2016. Improvement of the catalytic performance of a hyperthermostable GH10 xylanase from *Talaromyces leycettanus* JCM12802. Bioresource Technol 222:277-284.

Wiederstein M, Sippl MJ. 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35:W407-W410.

Wierenga RK. 2001. The TIM-barrel fold: a versatile framework for efficient enzymes. FEBS Lett. 492:193-198.

Yaacob N, Kamarudin NHA, Leow ATC, Salleh AB, Rahman RNZRA, Ali MSM. 2019. Effects of Lid 1 mutagenesis on lid displacement, catalytic

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

performances and thermostability of cold-active *Pseudomonas* AMS8 lipase in toluene. Comput. Struct. Biotechnol. J. 17:215-228.

Yan J, Liu W, Li Y, Lai HL, Zheng Y, Huang JW, Chen CC, Chen Y, Jin J, Li H, Guo RT. 2016. Functional and structural analysis of *Pichia pastoris* expressed *Aspergillus niger* 1,4-β-endoglucanase. Biochem Bioph Res Co. 475:8-12.

Yu H, Zhao Y, Guo C, Gan Y, Huang H. 2015. The role of proline substitutions within flexible regions on thermostability of luciferase. Biochim. Biophys. Acta, Proteins Proteomics 1854:65-72.

Zhai X, Amyes TL, Richard JP. 2015. Role of loop-clamping side chains in catalysis by triosephosphate isomerase. J. Am. Chem. Soc. 137:15185-15197.

Zheng F, Tu T, Wang X, Wang Y, Ma R, Su X, Luo H. 2018. Enhancing the catalytic activity of a novel GH5 cellulase *Gt*Cel5 from *Gloeophyllum trabeum* CBS 900.73 by site-directed mutagenesis on loop 6. Biotechnol Biofuels 11:76.

Tables

Table 1. Comparison of enzyme activity of purified *Ba*GH5-WT and *Ba*GH5-UV2.

Substrate	*Ba*GH5-WT (U/mg)	*Ba*GH5-UV2 (U/mg)	Fold increase
β-D-Glucan	10.2 ± 0.1	94.6 ± 0.2	9.3
Lichenan	7.4 ± 0.1	77.1 ± 2.5	10.4
CMC-Na	4.5 ± 0.01	43.6 ± 1.6	9.7
Cellulose powder	1.6 ± 0.1	4.3 ± 0.38	2.7
Avicel	1.2 ± 0.04	11.4 ± 2.3	9.5

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

Cellobiose	ND	ND	-
Laminarin	ND	ND	-
Galactomannan	ND	ND	-
Birchwood xylan	ND	ND	-

*All the assays were performed at 1% (w/v) substrate, 50 mM sodium acetate buffer (pH 5.5), 3 min, for Ba*GH5-WT at *55°C, and for Ba*GH5-UV2 at *65°C.*

ND = No activity detected. Values are mean SE (n=3).

Table 2. Kinetic parameters of *Ba*GH5-WT and *Ba*GH5-UV2.

Substrate (% w/v)	***Ba*GH5-WT**				***Ba*GH5-UV2**				**Fold increase catalytic efficiency (K_{cat}/K_m)***
	K_m (mg mL^{-1})	V_{max} (U/mg)	k_{cat} (min^{-1})	k_{cat}/K_m (mL^{1} mg^{-1} min^{-1})	K_m (mg mL^{-1})	V_{max} (U/mg)	k_{cat} (min^{-1})	K_{cat}/K_m (mL^{1} mg^{-1} min^{-1})	
CMC-Na	0.38 ± 0.11	6.0 ± 0.6	3.46×10^{2}	9.1×10^{2}	0.12 ± 0.05	42.05 ± 3.1	2.4×10^{3}	2.0×10^{4}	21.9
β-Glucan	0.15 ± 0.04	11.71 ± 0.9	6.74×10^{2}	4.5×10^{3}	0.07 ± 0.01	100.5 ± 8.8	5.7×10^{3}	8.3×10^{4}	18.4

values are mean SE (n=3)

All the assays were performed under 50 mM sodium acetate buffer (pH 5.5), 3 min for BaGH5-WT enzyme at 55°C, and for BaGH5-UV2 enzyme at 65°C.

Table 3. Ligand-binding analysis of MD simulated *Ba*GH5-WT and *Ba*GH5-UV2 structures by molecular docking.

Ligand	***Ba*GH5-WT**			***Ba*GH5-UV2**			**Increase in ΔG°**
	Binding energy, ΔG°	**Residues making hydrogen**	**Residues making hydrophobic**	**Binding energy,**	**Residues making hydrogen**	**Residues making hydrophobic**	

TH-2393_146106034

This article is protected by copyright. All rights reserved.

	(kcal/mol)	bonds	bic interactions	ΔG° (kcal/mol)	bonds	interactions	
Cellobiose	-4.28	Q232, Y254, D256, E280	W230, F253, Q260, S261	-4.82	H88, S284, K319, Q320, E321, S322, S323, S324,	D285, A286, W314	1.1
Cellotriose	-3.07	E192, Q232,Y254, D256, D264, K265	W230, V234, G259, Q260, S261	-5.02	N191, E192, H252, Y254, E280, D285, A286, S287, W314	W230, H258	1.6
Cellotetraose	-1.34	H88, R115, D152, N191, E192, P193, N194	A117, W314	-3.53	H88, D152, H154, L156, D158, N194, E192, S322, S323, S324	G159, Y254, S284, D285, A286, W314	2.6
Cellopentaose	-0.25	H88, G89, Y93, N191, E192, P193, D256	W92, A117, W230, Y254, A286, W314,	-0.56	H88, E192, D152, H154, I155, N191, N194, K319	L156, D158, G159, A286, W314, Q320, E321	2.5
Cellohexaose	-0.33	E192, W230, Q232, H252, Y254, E280,	N191, P193, A255, G259, H258, Q260, S261	1.46	D158, N160, E192, N194, S284, W314, S322	L156, G159, G195, Y254, A286, G282, E321	-

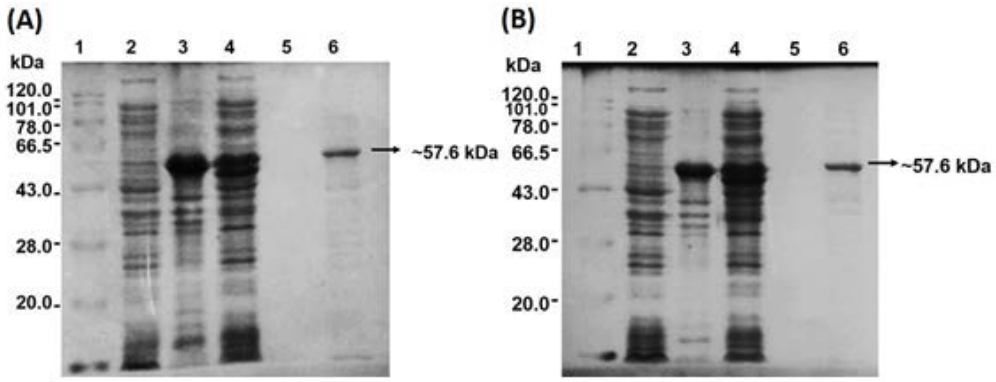
Accepted Article

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Figures

Fig. 1 SDS-PAGE (12.0%, w/v) gel showing purification of recombinant *Ba*GH5-WT from *E. coli* BL21 (DE3)pLysS cells A) Lanes 1, Page Ruler protein marker (Biobharti); 2, Uninduced *Ba*GH5-WT cells; 3, Induced *Ba*GH5-WT cells after sonication; 4, Cell free extract after sonication; 5, last column wash; 6, *Ba*GH5-WT affinity column purified (57.6 kDa) protein; and for recombinant *Ba*GH5-UV2 from *E. coli* BL21 (DE3)pLysS cells B) Lanes 1, Page Ruler protein marker (Biobharti); 2, Uninduced *Ba*GH5-UV2 cells; 3, Induced *Ba*GH5-UV2 cells after sonication; 4, Cell free extract after sonication; 5, last column wash; 6, *Ba*GH5-UV2 affinity column purified (57.6 kDa) protein.

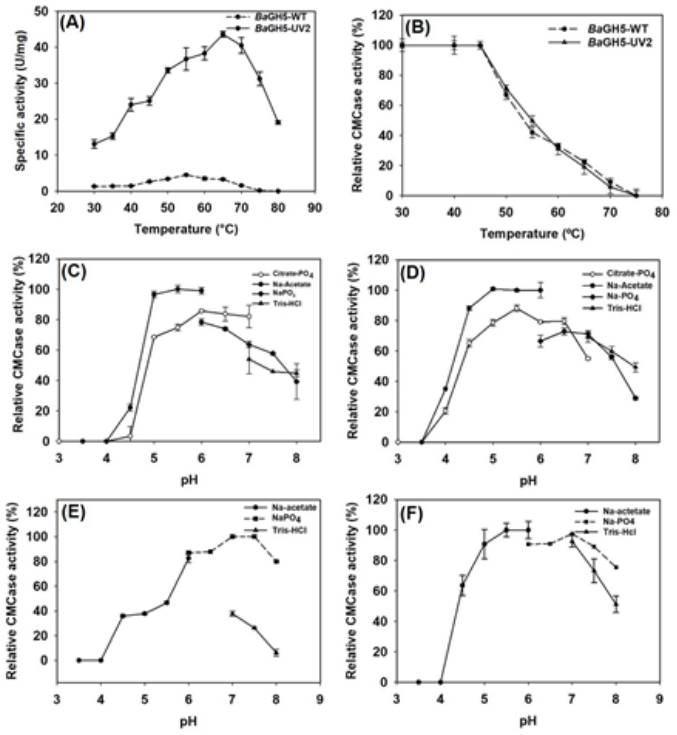


Accepted Article

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Fig. 2 (A) Effect of temperature on *Ba*GH5-WT, and *Ba*GH5-UV2 activity; (B) Effect of temperature on *Ba*GH5-WT, and *Ba*GH5-UV2 stability; (C) Effect of pH on *Ba*GH5-WT activity; (D) Effect of pH on *Ba*GH5-UV2 activity; (E) Effect of pH on *Ba*GH5-WT stability; and (F) Effect of pH on *Ba*GH5-UV2 stability.

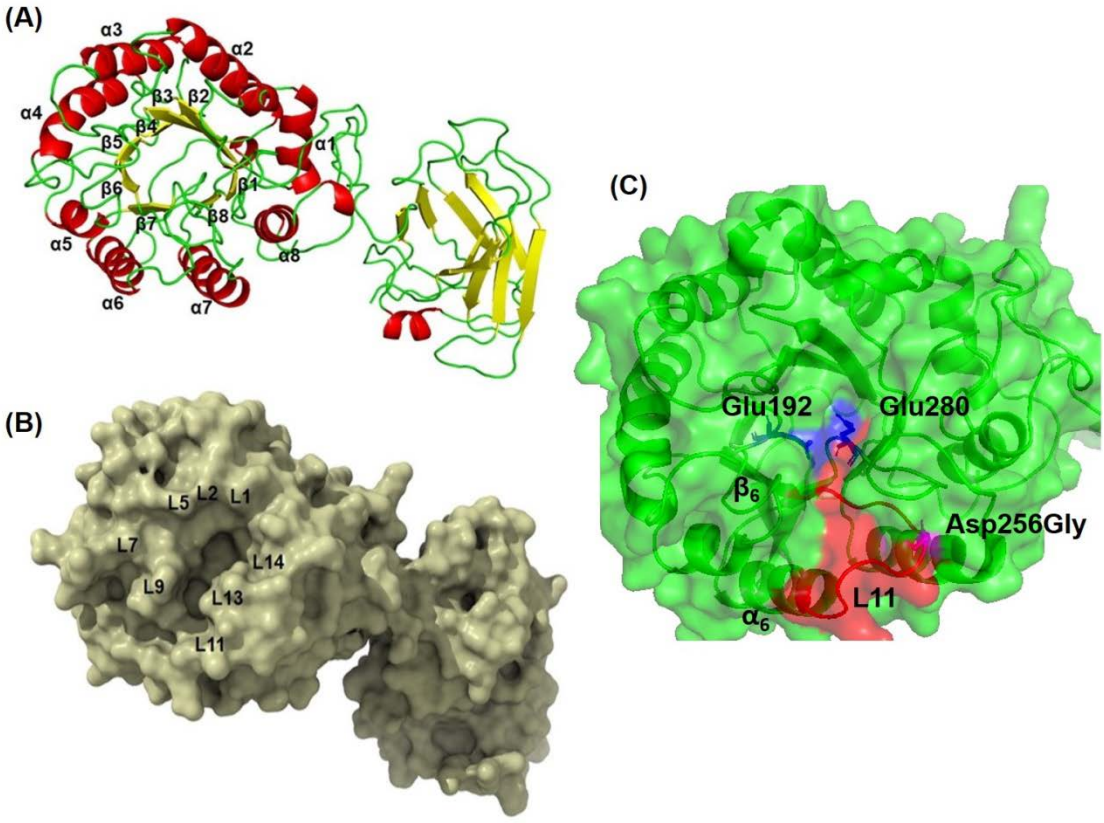


Accepted Article

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Fig. 3 A) 3-D cartoon view structure of modeled *Ba*GH5-WT representing TIM barrel $(\alpha/\beta)_8$ fold; B) Surface view of 3-D *Ba*GH5-WT showing Loops connecting the $\beta_1\alpha_1$ (L1), $\beta_2\alpha_2$ (L2), $\beta_3\alpha_3$ (L5), $\beta_4\alpha_4$ (L7), $\beta_5\alpha_5$ (L9), $\beta_6\alpha_6$ (L11), $\beta_7\alpha_7$ (L13), and α_7 β_8 (L14) forming the catalytic core, and C) 3-D structure analysis of *Ba*GH5-UV2 catalytic core representing mutation Asp256Gly (Pink colour) in a loop 11 connecting the β_6 to α_6 showing loop in red colour and catalytic residues Glu192 and Glu280 (Blue colour).

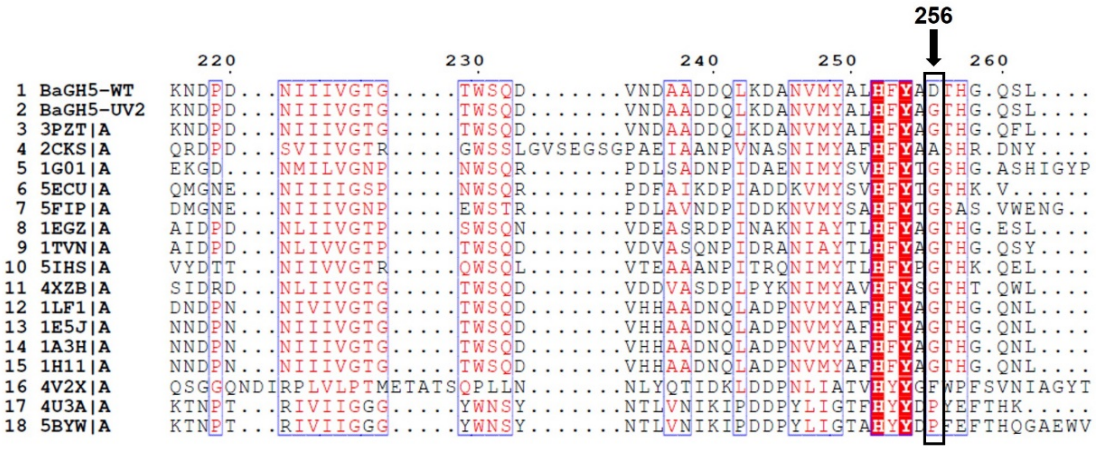


Accepted Article

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Fig. 4 Multiple sequence alignments of 16 homologues of *Ba*GH5-WT bacterial endoglucanases of GH5 family. *Ba*GH5-WT, PDB: ID 2CKS|A, and PDB: ID 4V2X|A showed divergence at position 256.

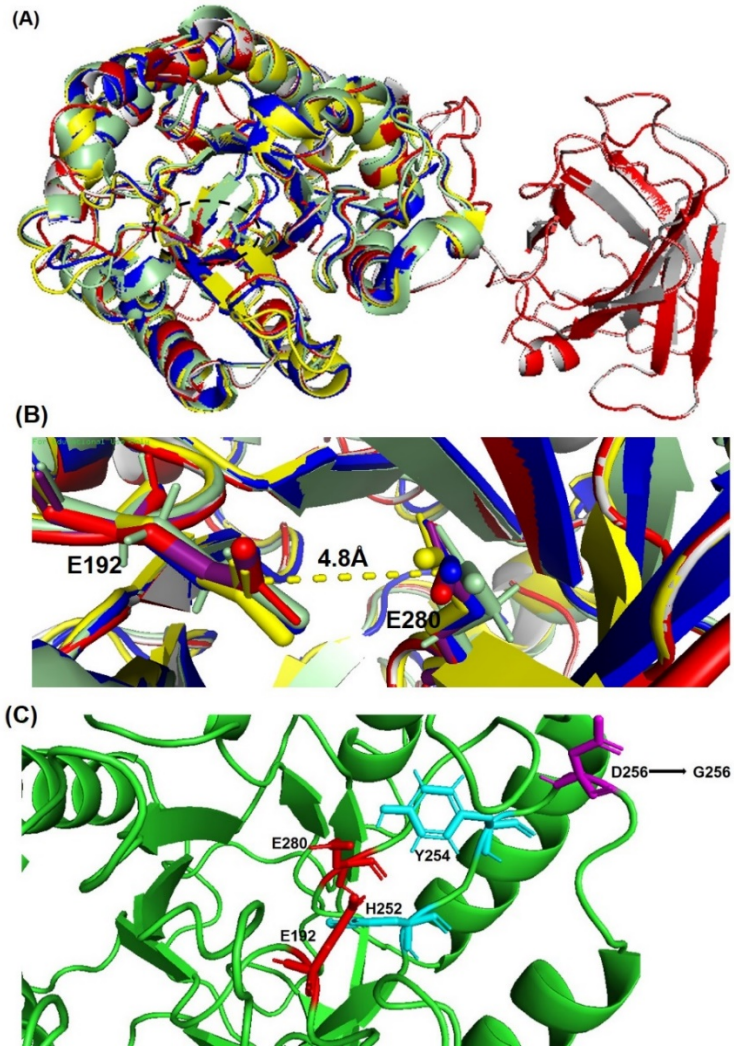


Accepted Article

TH-2393_146106034

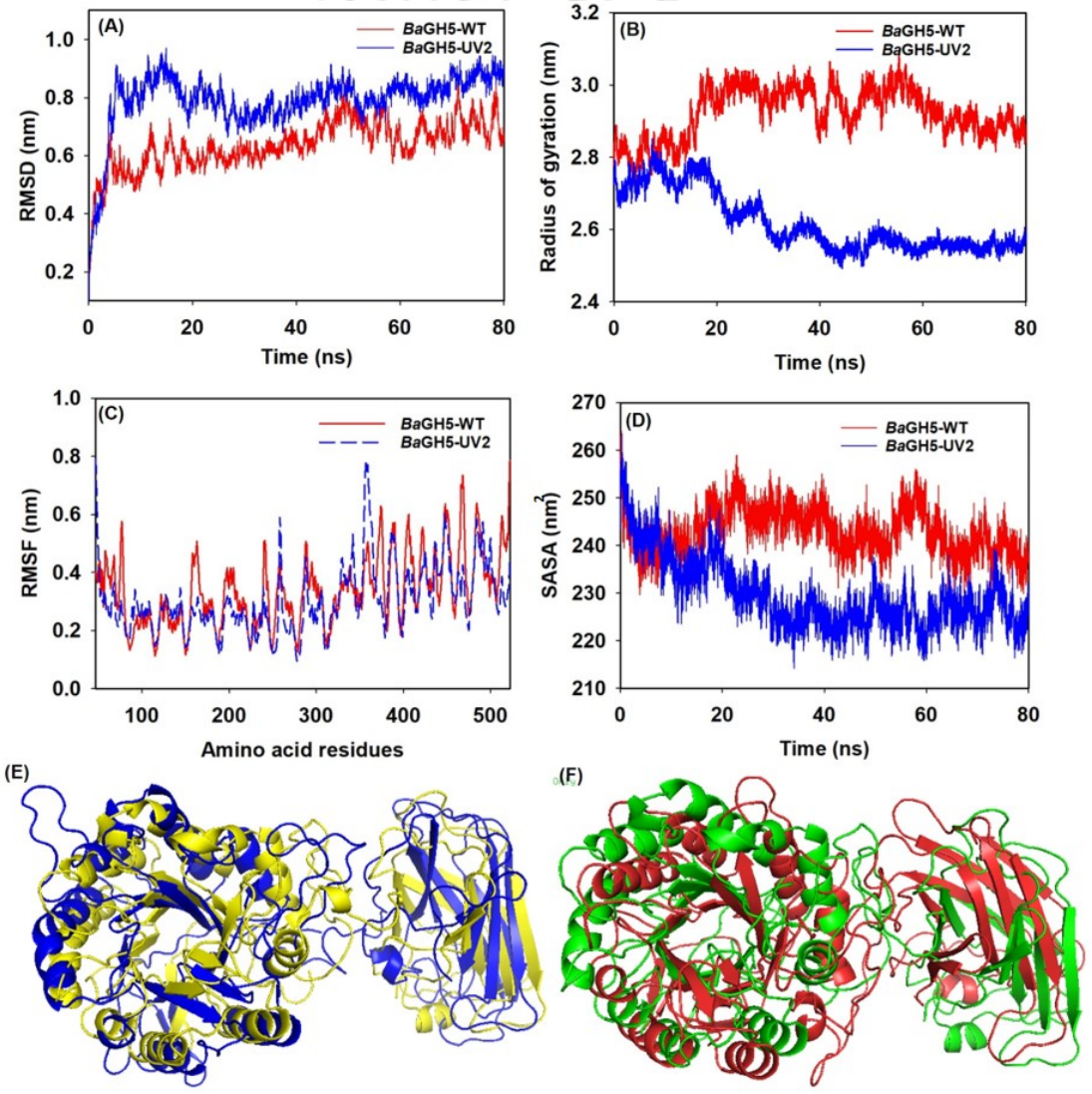
This article is protected by copyright. All rights reserved.

Fig. 5 A) Superposition of *Ba*GH5-WT modeled structure (Red colour), and *Ba*GH5-UV2 modeled structure (blue white colour) with crystal structures of GH5 endoglucanase of PDB ID: 3PZT|A (blue colour), PDB ID: 1E5J|A (yellow colour), and PDB ID: 5IHS|A (pale green) showed common structure, and catalytic residues (in circle); (B) The profile of distance 4.8 Å (from C_α to C_α) between the catalytic acid (E192, red colour), and catalytic base (E280, red colour) of *Ba*GH5-WT by molecular visualization system PyMOL showing hydrolytic mechanism (retaining); and C) Modeled *Ba*GH5-WT structure showing two catalytically important residues H252, and Y254 (blue colour) on loop having substitution mutation of D256G (purple colour).



TH-2393_146106034
This article is protected by copyright. All rights reserved.

Fig. 6 Molecular dynamics simulation of *Ba*GH5-WT, and *Ba*GH5-UV2. A) RMSD plot; B) Radius of gyration plot; C) RMSF plot; D) SASA plot; E) Superposition of 80 ns MD simulated structure of *Ba*GH5-WT (blue) with modeled *Ba*GH5-WT structure (yellow); F) Superposition of 80 ns MD simulated structure of *Ba*GH5-UV2 (red) with modeled *Ba*GH5-UV2 structure (green).

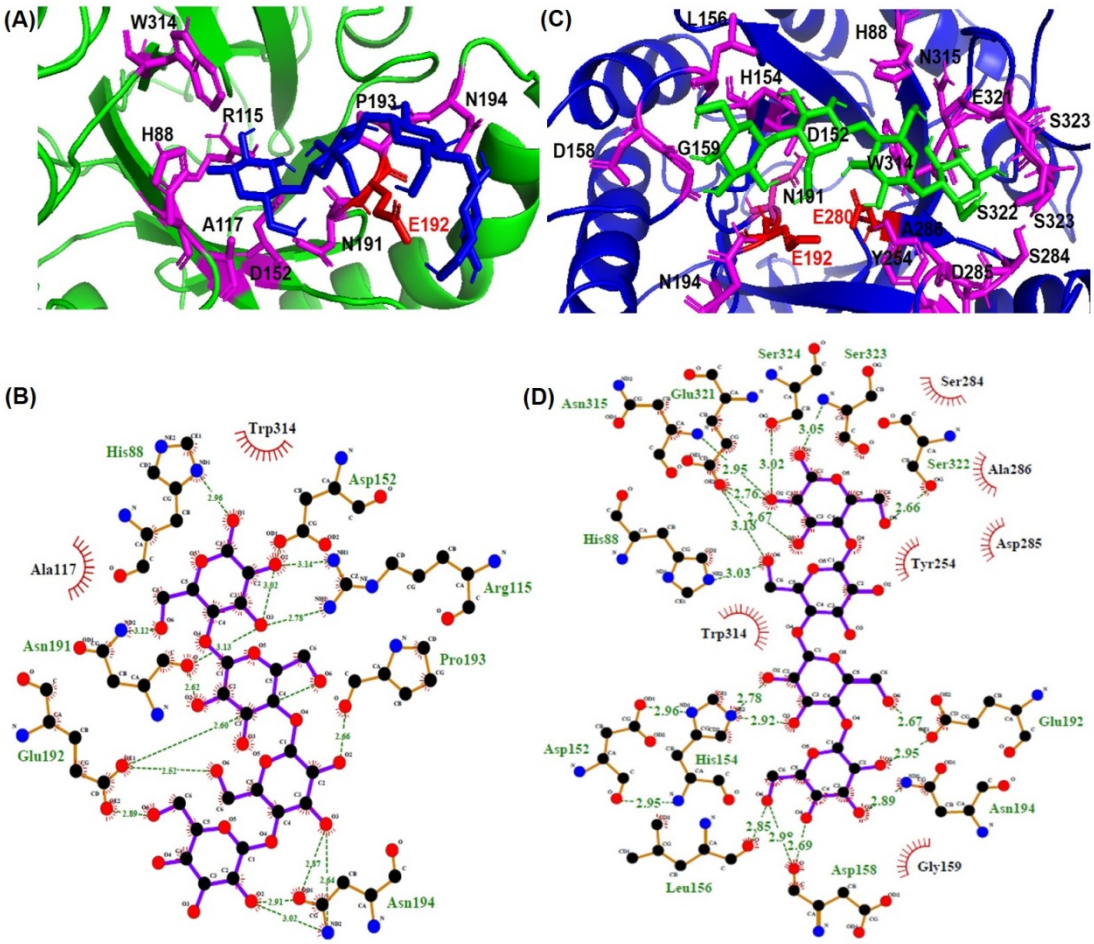


Accepted Article

TH-2393_146106034

This article is protected by copyright. All rights reserved.

Fig. 7 Molecular docking analysis of 80 ns MD simulated *Ba*GH5-WT or *Ba*GH5-UV2 structure. A) the catalytic core of *Ba*GH5-WT showing interaction with the ligand cellotetraose; B) Two-dimensional schematic presentation of cellotetraose with the binding residues of *Ba*GH5-WT; C) The catalytic core of *Ba*GH5-UV2 showing interaction with the ligand cellotetraose; and D) Two-dimensional schematic presentation of cellotetraose with the binding residues of *Ba*GH5-UV2.



TH-2393_146106034

This article is protected by copyright. All rights reserved.

Accepted Article

Industrial Crops & Products 154 (2020) 112678

Contents lists available at ScienceDirect

Industrial Crops & Products

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

Statistically designed cellulase mixture for saccharification of pretreated *Sorghum durra* stalk

Shweta Singh[a,b], Vikky Rajulapati[a], Sumitha Banu Jamaldheen[b,c], Vijayanand Suryakant Moholkar[c,d], Arun Goyal[a,b,c,*]

[a] Carbohydrate Enzyme Biotechnology Laboratory, Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati, 781039, Assam, India
[b] DBT PAN-IIT Centre for Bioenergy, Indian Institute of Technology Guwahati, Guwahati, 781039, Assam, India
[c] Centre for Energy, Indian Institute of Technology Guwahati, Guwahati, 781039, Assam, India
[d] Department of Chemical Engineering, Indian Institute of Technology Guwahati, Guwahati, 781039, Assam, India

ARTICLE INFO

Keywords:
Simplex-centroid mixture design
Acid or base pretreatment
Endoglucanase
Cellobiohydrolase
β-glucosidase
Synergism

ABSTRACT

The potential of least explored *Sorghum durra* as a feedstock for enzymatic saccharification in bioethanol production was evaluated as only one report is available. The pretreatment specific, optimization of saccharification of *Sorghum durra* stalk (Sds) was carried out. The concentrations of crude recombinant endoglucanase (*Ba*GH5-UV2), cellobiohydrolase (*Ct*CBH5A) and β-glucosidase (*Ct*GH1) in cocktail were optimized for saccharification of acid or base pretreated Sds by using mixture design approach. Optimized percentage ratio of *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 with total enzyme loading 2000 U/g of pretreated Sds for acid pretreated Sds was 43.1:10.0:46.9 and for base pretreated Sds was 80:10:10 at 48 h of saccharification. The base pretreated Sds showed 1.8-fold (165.6 mg/g of pretreated Sds) higher total reducing sugar (TRS) yield and 1.2-fold (81.7 mg/g of pretreated Sds) higher glucose yield than acid pretreated Sds. This low sugar yield from acid pretreated Sds was due to the presence of higher acid insoluble lignin content, 11.9 % (w/w) than the base pretreated biomass (2.2 %, w/w). In acid pretreated Sds, ANOVA data showed 3.7-fold higher F value (111.58) for *Ba*GH5-UV2 and *Ct*GH1 interaction than the base pretreated Sds. This indicated that the synergism between *Ba*GH5-UV2 and *Ct*GH1 for acid pretreated Sds was significantly higher. This was due to lower intact hemicellulose content (10.6 %, w/w) in acid pretreated Sds than base pretreated Sds (21.8 %, w/w). The optimized enzyme ratios for saccharification of acid pretreated Sds gave maximum TRS and glucose yield at 72 h, 113 and 85 mg/g of pretreated Sds, respectively, unlike base pretreated Sds at 96 h, 211 and 174 mg/g of pretreated Sds, respectively. These results showed that the ratio of cellulase enzymes in cocktail and their synergism depends on the type of pretreatment used for biomass. Thus, this in-house cocktail design specific to pretreatment can improve the biomass saccharification at large scale.

1. Introduction

The global warming and energy dependence of different countries on non-renewable fossil fuels have turned the world's attention towards the renewable fuels. Lignocellulosic waste is a source of fermentable sugars for bioethanol production (Limayem and Ricke, 2012). In India, 55 % of total sown rainfed area, contributes for more than 75 % of the pulses and millet production (Zalkuwi et al., 2014). *Sorghum durra* is the majorly cultivated millet along with the pulse crops throughout the year in India with a total production of about 10.62 million tons per year (Jamaldheen et al., 2018). Only the grains of Sorghum is utilized as food and a very little fraction of stalk is utilized as fodder, whereas the leftover of stalk becomes the agro-waste and can be used as feedstock in saccharification for fermentable sugar production in bioethanol production (NRAA, 2012). *Sorghum durra* has advantages over other commonly used feedstock for bioethanol production such as rice, sugarcane and wheat crop residue agro-waste. The rice, sugarcane and wheat crop requires continuous supply of water for growth thus, needed

Abbreviation: Sds, *Sorghum durra* stalk; TRS, total reducing sugar; *CrI*, crystallinity index; CMC-Na, carboxymethyl cellulose sodium salt; LB, Luria Bertani; GOD-POD, glucose oxidase-peroxidase; ADL, acid insoluble lignin; RSM, response surface method; ANOVA, analysis of variance; DP, degree of polymerization

* Corresponding author at: Carbohydrate Enzyme Biotechnology Laboratory, Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati, 781039, Assam, India.
E-mail address: arungoyl@iitg.ac.in (A. Goyal).

https://doi.org/10.1016/j.indcrop.2020.112678
Received 6 March 2020; Received in revised form 15 May 2020; Accepted 6 June 2020
0926-6690/ © 2020 Elsevier B.V. All rights reserved.

TH-2393_146106034

a well established irrigated system (Inman-Bamber et al., 2005; NRAA, 2012). Therefore, in order to maintain the continuous operation of bioethanol plant, the assessment of agrowastes from rainfed areas is important at the time of non-availability of irrigation depending feedstocks. Although several reports are available on the pretreatment and saccharification methods for production of bioethanol from other *Sorghum sp*. such as, *Sorghum bicolor* using dilute alkali (McIntosh and Vancov, 2010), Sweet Sorghum bagasse using microwave assisted dilute ammonia (Chen et al., 2012), Sweet Sorghum stalk using organosolv (Ostovareh et al., 2015) and Sweet Sorghum straw using heavy ion beam irradiation (Xu et al., 2019). However, there is only one report available only on the optimization of pretreatment for *Sorghum durra* stalk (Sds) as a feedstock for bioethanol production (Jamaldheen et al., 2018). The raw Sds contains 55 % (w/w) holocellulose content (Jamaldheen et al., 2018), therefore, can be a potential substrate for bioethanol production.

Cellulases are a group of enzymes, that converts the cellulosic polysaccharides into fermentable sugars for biofuel production (Mandels and Weber, 1969). They are classified into the three classes of enzymes i.e. endoglucanases, cellobiohydrolases and β-glucosidases (Cai et al., 1994 and Cai et al., 1999). Endoglucanase randomly cleaves β-1,4 glycosidic bonds in the cellulose chain and expose new ends for the action of cellobiohydrolase to produce cellobiose. β-Glucosidase converts cellobiose into the fermentable sugar, glucose (Koivula et al., 1998; Garcia et al., 2018). All these three enzymes work synergistically to avoid product inhibition (Galbe and Zacchi, 2002). However, the cost of enzymatic conversion of cellulose into the fermentable sugar is high and is the major bottleneck for the biofuel industries (Hess, 2008). Therefore, the cost-effective enzymatic saccharification of cellulose is desirable for the bio-conversion of lignocellulosic biomasses into the bioethanol (Levine et al., 2011). The rate of saccharification and fermentable sugar yield from lignocellulosic biomass depends on the mixture composition of different cellulases (endoglucanase, cellobiohydrolase and β-glucosidase) that rely on biomass composition and the type of pretreatment method (Rosgaard et al., 2007; Zhou et al., 2009; Marcos et al., 2013). The acid and alkali pretreatments are the two most widely used cost effective methods for lignocellulosic biomass (Dagnino et al., 2013; Singh et al., 2014, 2015). The acid pretreatment with dilute sulfuric acid solubilizes the considerable amount of hemicellulosic fraction from lignocellulosic biomass (Kim et al., 2013). In contrast, the alkali pretreatment by diluted sodium hydroxide removes lignin from the lignocellulosic biomass (Kim et al., 2016). Therefore, the tailor-made cocktail of cellulases pertaining to the type, ratio and amount of enzymes, specific to the pretreatment process is crucial for the maximization of saccharification yield (Banerjee et al., 2010; Kim et al., 2015). The optimum enzymes ratio in a cocktail was predicted on the basis of rational mixture design experiments with a concept of synergistic effects of enzymes on biomass (Zhou et al., 2009; Bussamra et al., 2015).

The optimization of pretreatment on Sds by 1% (v/v) acid assisted oven heating/autoclaving or 1% (w/v) NaOH assisted oven heating/autoclaving showed an increase in crystallinity index (*CrI*) of pretreated biomass (Jamaldheen et al., 2018). In the present study, the optimized pretreated Sds biomasses were taken further for enzymtic saccharification by using mixture of cellulases. The mixture of cellulases comprising endoglucanase (*Ba*GH5-UV2), cellobiohydrolase (*Ct*CBH5A) and β-glucosidase (*Ct*GH1) was designed in an optimised ratio for efficient saccharification. Endoglucanase (*Ba*GH5-UV2) from a mutant strain UV2 of *Bacillus amyloliquefaciens* SS35 (Singh et al., 2020a) was cloned and expressed in *E. coli* cells (Singh et al., 2020b) and used. Cellobiohydrolase (*Ct*CBH5A) cloned from *Clostridium thermocellum* was a gift from Professor C.M.G.A. Fontes, NZYTech Ltd., Lisbon, Portugal. β-Glucosidase (*Ct*GH1) from *Clostridium thermocellum* was cloned and expressed earlier and used for the present study (Sharma et al., 2019). The optimization of enzyme ratios in the cellulase cocktail was carried out by a rational statistical approach, simplex-centroid mixture design for pretreatment specific customization. The responses were evaluated through saccharification reaction assays and validated.

2. Materials and methods

2.1. Media and chemicals

Carboxymethyl cellulose sodium salt (CMC-Na) (low viscosity) and D-glucose were procured from Sigma Aldrich, USA. Luria Bertani (LB) broth, kanamycin, chloramphenicol, ampicillin, sodium azide, sodium acetate, sodium phosphate monobasic, sodium phosphate dibasic, sodium hydroxide, sodium carbonate, sodium potassium tartrate, sodium bicarbonate, sodium sulphate and cellobiose were purchased from Himedia Pvt. Ltd. India. Sodium arsenate, ammonium molybdate, sulphuric acid, ortho-phosphoric acid and acetic acid were purchased from Merck, India. Glucose oxidase-Peroxidase (GOD-POD) kit was purchased from the coral clinical system (Tulip Diagnostics, Pvt. Ltd., India). Syringe filters and the PVDF filter membranes (pore size 0.2 µm) were procured from Axiva (Axiva Sichem Pvt. Ltd., India).

2.2. Biomass collection, processing and pretreatment

The Sds's were collected from Kappalpatti, Tamil Nadu, India. The Sds biomass was chopped, washed with water and dried in hot air oven at 70 °C for 24 h. It was ground and sieved to a particle size of approximately, 0.6 mm. 20 g of raw Sds was added to 200 mL of ultra-pure (MilliQ) water in 500 mL Erlenmeyer flask (solid to liquid ratio, 1:10). The raw Sds was pretreated with 1% (v/v) H_2SO_4 assisted autoclaving at 121 °C, 15 psi for 20 min (acid pretreatment) and 1% (w/v) NaOH assisted autoclaving at 121 °C, 15 psi for 20 min (base pretreatment), respectively according to the protocol reported earlier (Jamaldheen et al., 2018). The pretreated biomass(s) was washed with distilled water till neutral pH to remove inhibitors (da Costa et al., 2018). The α-cellulose, hemicellulose and acid insoluble lignin (ADL) composition of both the pretreated biomass was determined by TAPPI (1992) methods in the sets of three.

2.3. Production of recombinant enzymes used in saccharification of pretreated Sds

The recombinant endoglucanase (*Ba*GH5-UV2) from UV2 mutant strain of *Bacillus amyloliquefaciens* SS35 (Singh et al., 2020a) was cloned and expressed using *E. coli* BL-21(DE3) pLysS cells at 24 °C (Singh et al., 2020b). The enzyme cellobiohydrolase (*Ct*CBH5A) was expressed in *E. coli* BL-21(DE3) cells (Nedumaran et al., 2020) and grown at 24 °C. The β-Glucosidase (*Ct*GH1) was expressed in *E. coli* BL-21(DE3) cells as reported earlier (Sharma et al., 2019) and grown at 24 °C. The antibiotic used for the expression of *Ba*GH5-UV2 was kanamycin (50 µg/mL) and chloramphenicol (25 µg/mL) and for *Ct*CBH5A and *Ct*GH1 was kanamycin (50 µg/mL). All the three *E. coli* cultures were grown in one liter LB medium containing respective antibiotics by incubating at 37 °C in shaking incubator at 180 rpm, followed by IPTG (1 mM) induction at OD_{600nm} 0.6 which followed by incubation at 24 °C for 18 h. The cells were centrifuged at 8000*g* for 15 min at 4 °C and the cell pellet(s) were resuspended in 8 mL of 20 mM sodium phosphate buffer, pH 6.5. The cell suspensions were sonicated at 5 s on/off pulse with 33 % amplitude. The sonicated samples were centrifuged at 13,000 *g* for 50 min. The cell free supernatant containing the crude recombinant enzymes *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1, instead of purified were explored in the enzyme cocktail preparation for saccharification of Sds in order to make the process cost effective. The cell free supernatant of all the three enzymes were sterilized by passing through the autoclaved syringe filter using PVDF membrane (0.2 µm pore size). The cell free extracts of enzymes were assayed for activity and protein concentration.

TH-2393_146106034

2.4. Assay of recombinant enzymes and protein estimation

The endoglucanase, *Ba*GH5-UV2 activity was determined under optimized conditions, by incubating the 100 µL reaction mixture containing 1% (w/v) CMC-Na in 50 mM sodium acetate buffer, pH 5.5 and 5 µL of crude enzyme at 65 °C for 3 min (Singh et al., 2020b). The cellobiohydrolase, *Ct*CBH5A activity was measured under optimized conditions (Nedumaran et al., 2020) by incubating 100 µL reaction mixture containing 1% (w/v) CMC-Na in 50 mM sodium phosphate buffer, pH 6.4 and 10 µL of crude enzyme at 65 °C for 2 min. The enzyme activity of *Ba*GH5-UV2 and *Ct*CBH5A was calculated by estimating reducing sugar released method of Nelson (1944) and Somogyi (1945). The reducing sugar was determined by measuring the absorbance at 500 nm, on a spectrophotometer (Varian, Cary 100 Bio) by using glucose standards. The β-glucosidase activity of *Ct*GH1 was determined under optimized conditions as reported by (Sharma et al. 2019) by using GOD-POD method (Raabo and Terkildsen, 1960). The 100 µL reaction mixture contained 1% (w/v) cellobiose dissolved in 20 mM citrate phosphate buffer, pH 6.0 and 10 µL of crude *Ct*GH1 enzyme. The enzyme reaction was incubated at 65 °C for 10 min according to the protocol reported earlier (Sharma et al., 2019). The reaction was stopped by boiling in a boiling water bath for 15 min. After cooling, 2.0 µL of reaction mixture was added to a well in 96 well microtitre plate followed by addition of 200 µL of glucose reagent L1 (from kit) and incubated at 37 °C in an incubator for 15 min. The β-glucosidase activity was determined by measuring the absorbance at 505 nm, on a spectrophotometer (Varian, Cary 100 Bio) using glucose standard. The protein concentrations of the crude cell free supernatants of *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 were determined by Lowry method (Lowry et al., 1951) using Bovine serum albumin as a standard.

2.5. Estimation of common buffer concentration for saccharification

Endoglucanase, *Ba*GH5-UV2 retained 100 % of relative activity in 50 mM sodium acetate buffer in pH range, 5.5–6.0 and in sodium phosphate buffer from pH range, 6.0–7.5 (Singh et al., 2020b). *Ct*CBH5A retained 100 % of relative activity in 50 mM sodium phosphate buffer in pH range, 6.2–6.6 (Nedumaran et al., 2020). The *Ct*GH1 protein retained 100 % of the relative activity in 20 mM sodium citrate buffer in pH range 4.0–6.0 and in 20 mM sodium phosphate buffer in pH range, 7.0–8.0 (Sharma et al., 2019). Therefore, a common buffer sodium phosphate at pH 6.5 was used. *Ba*GH5-UV2 and *Ct*CBH5A protein retained 100 % of relative activity in the temperature range, 30 − 45 °C (Singh et al., 2020b) and 30 − 65 °C (Nedumaran et al., 2020), respectively. *Ct*GH1 retained 100 % relative activity in temperature range, 20 − 40 °C (Sharma et al., 2019). Therefore, a common temperature of 40 °C was used for saccharification. The common buffer concentration was determined by analysing the stability of the above three enzymes by incubating each of the crude enzyme (20 µL, 20 mg/mL) in 20 mM and 50 mM sodium phosphate buffer, pH 6.5 at 40 °C for 6 h, 12 h and 24 h. The residual enzyme activity of *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 was determined as described in Section 2.4.

2.6. Customization of limit of enzyme activities for mixture design by Stat-Ease Design Expert software

The limits of enzyme activities for crude enzymes, *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 were determined against pretreated (acid or base) Sds with a biomass loading of 2% (w/v), as all the three enzymes gave maximum total reducing sugar (TRS) yield (data not shown). The saccharification using crude enzymes, *Ba*GH5-UV2 or *Ct*CBH5A separately were carried out in 1 mL reaction volume by varying their loading concentrations *viz.* 200 U/g, 400 U/g, 800 U/g, 1200 U/g, 1600 U/g and 2000 U/g of pretreated Sds in 20 mM sodium phosphate buffer (pH 6.5). The *Ct*GH1 loading concentrations were customized by varying the loading at 200 U/g, 400 U/g, 800 U/g, 1200 U/g, 1600 U/g and 2000 U/g of pretreated Sds at a fixed loading concentrations of *Ba*GH5-UV2 (800 U/g of pretreated Sds) and *Ct*CBH5A (400 U/g of pretreated Sds). The saccharification was carried out in 20 mM sodium phosphate buffer (pH 6.5) in 1.0 mL reaction volume in 2 mL micro-centrifuge tube by incubating at 40 °C under shaking in an incubator shaker (Orbitek, Scigenics Biotech, India) at 150 rpm for 48 h in triplicate sets. 0.005 % (w/v) sodium azide and 100 µg/mL ampicillin were added to the reaction mixture to prevent contamination. After 48 h reactions were stopped by boiling the reaction mixtures in boiling water bath for 10 min. The TRS yield was estimated by the method of reducing sugar estimation method of Nelson (1944) and Somogyi (1945).

2.7. Optimization of enzymatic hydrolysis for acid and base pretreated Sds by simplex-centroid mixture design

2.7.1. Experimental design, statistical analysis and model fitting

The enzyme cocktail ratios for the three recombinant crude enzymes *Ba*GH5-UV2 (A), *Ct*CBH5A (B) and *Ct*GH1 (C) were optimized for saccharification of acid or base pretreated Sds by using response surface method (RSM) mixture design in Stat-Ease Design Expert software (Stat-Ease Inc., Minneapolis, MS, USA, Version 6.0.4). A three parameters simplex-centroid mixture design was employed for optimization of A, B and C in mixture with total 10 runs of experiments (Table 3). The total loading amount of the three different enzymes in a mixture was 2000 U/g of pretreated biomass. The loading amounts of A, B, and C in terms of U/g of pretreated biomass were applied as independent variables. Here, percentage units of A, B and C needed per gram of pretreated substrate was varied. The minimum and maximum level of parameters were set as follows, 10 % and 80 %. Therefore, the constraints of the total composition mixture of these three constituents were set as 100 %. A principal advantage of using this above mixture design concept is to set flexible constraints in the total composition range. This showed that if the concentrations of one individual component are expressed as percentage, then it can be varied in a total between 1–100%. The TRS yield and glucose yield from acid or base pretreated biomass, respectively were the response factors. The reactions were carried out in 2 mL microcentrifuge tube with total reaction volume of 1 mL with enzyme loading according to the mixture ratios given in Table 3. 0.005 % (w/v) sodium azide and 100 µg/mL ampicillin were added to the reaction mixture to avoid contamination. The microcentrifuge tubes were incubated at 40 °C, 150 rpm for 48 h in an incubator shaker (Orbitek, Scigenics Biotech, India). After 48 h, the enzyme reaction mixture was boiled in boiling water bath for 10 min for inactivating the enzymes. The resulting reducing sugars were determined by the method given by Nelson (1944) and Somogyi (1945) as described in Section 2.4. The resulting hexose sugar, glucose was measured by GOD-POD method (Raabo and Terkildsen, 1960) as mentioned in Section 2.4. To validate the mathematical model and experiment data, analysis of variance (ANOVA) was performed. Regression analysis was carried out for the best-fitted models. Contour plots for the best-fitted models were generated to predict the optimal formulations of the above enzymes (A, B and C) and synergism among them. Eventually, the optimized composition mixture of three enzymes for both the pretreated biomass was determined and compared.

2.7.2. Saccharification of acid or base pretreated Sds under optimized conditions

The statistical prediction was validated by carrying out the enzymatic saccharification with acid or alkali pretreated Sds at the optimal formulations of enzyme ratio in mixture obtained from the quadratic and special cubic models, respectively. 12 mL of each reaction mixture contained optimized mixture ratio of enzymes (*Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1) for respective pretreated biomass in 50 mL Erlenmeyer flask. The reactions were carried out for 2.0 % (w/v) pretreated Sds in 20 mM sodium phosphate buffer, pH 6.5 in an incubator shaker at 40 °C under shaking at 150 rpm (Orbitek, Scigenics Biotech, India) for 96 h.

TH-2393_146106034

20 µL aliquot of saccharification reaction samples were collected after 24 h, 48 h, 72 h and 96 h, in 1 mL microcentrifuge tube. The reactions were stopped by boiling the reaction mixtures in boiling water bath for 5 min. The resulting reducing sugars were determined by the method given by Nelson (1944) and Somogyi (1945). The resulting hexose sugar, glucose was determined by GOD-POD method (Raabo and Terkildsen, 1960).

3. Results and discussion

3.1. Composition analysis of acid and base pretreated Sds

The acid pretreated Sds showed 57.0 % (w/w) cellulose, 10.6 % (w/w) hemicellulose and 11.9 % (w/w) ADL after TAPPI (1992) method analysis. The base pretreated Sds gave 53.1 % (w/w) cellulose, 21.8 % (w/w) hemicellulose and 2.2 % (w/w) ADL. However, an earlier study reported similar ADL content, 13.2 % (w/w) in acid pretreated Sds and 2.3 % (w/w) ADL content in the base pretreated Sds (Jamaldheen et al., 2018).

3.2. Estimation of protein concentration and enzyme activity of recombinant enzymes and common concentration of buffer for saccharification

The protein concentrations of cell free supernatant containing crude *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 was 37.0 mg/mL, 32.1 mg/mL and 21.6 mg/mL, respectively. The crude *Ba*GH5-UV2 and *Ct*CBH5A in cell free supernatant showed endoglucanase activity, 242 U/mL and cellobiohydrolase activity, 83.2 U/mL, respectively, against CMC-Na. The crude *Ct*GH1 showed β-glucosidase activity, 22.5 U/mL against cellobiose. The enzymes, *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 retained 100 % of the residual enzyme activity in 20 mM sodium phosphate buffer, pH 6.5 for 24 h (Table 1). However, in 50 mM sodium phosphate buffer *Ba*GH5-UV2 and *Ct*CBH5A retained 95.0 % and 91.2 % of the residual enzyme activity, respectively, but *Ct*GH1 retained only 60.9 % of relative enzyme activity (Table 1). Therefore, the common buffer concentration used for cocktail mixture design of *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 was 20 mM sodium phosphate buffer, pH 6.5 for performing the saccharification of acid or alkali pretreated Sds. Kim et al. (2017) reported that assessment of common pH and buffer is the most crucial step in mixture of enzyme based enzymatic saccharification.

3.3. Customization of limits of enzyme loading concentration for mixture design by Stat-Ease Design Expert software

The TRS yield from acid pretreated Sds increased from 3.4–27.6 mg/g of pretreated Sds when the loading concentration of crude *Ba*GH5-UV2 was increased from 200 U/g to 800 U/g of pretreated Sds. Similarly, the increase in TRS yield from 1.1–9.4 mg/g of pretreated Sds was observed when loading concentration of crude *Ct*CBH5A was increased from 200 U/g to 800 U/g of pretreated Sds (Table 2). However, further increase in *Ba*GH5-UV2 or *Ct*CBH5A loading concentration did not cause any signification increase in the TRS yield. This saturation in the TRS yield after enzyme loading concentration of 800 U/g of pretreated Sds was due to the end products inhibition produced from the hydrolysis of cellulose according to the earlier reports of Kim et al. (2017). Therefore, the minimum loading concentration of *Ba*GH5-UV2 and *Ct*CBH5A for acid pretreated Sds was 200 U/g of pretreated biomass. The enzyme β-glucosidase, *Ct*GH1 alone did not produce any reducing sugars from acid pretreated Sds. This is because the *Ct*GH1 acts on cellobiose to form two molecules of glucose and it was unable to hydrolyse cellulose polymer (Sharma et al., 2019). Therefore, to optimize the minimum loading concentration of *Ct*GH1, a fixed concentration of *Ba*GH5-UV2, 800 U/g of pretreated Sds and *Ct*CBH5A, 400 U/g of pretreated Sds was used in the saccharification reaction. This is because the enzymes, *Ba*GH5-UV2 and *Ct*CBH5A together produce the end product cellobiose, which is the substrate for *Ct*GH1, that produces glucose. The TRS yield from acid pretreated Sds with fixed *Ba*GH5-UV2, 800 U/g of pretreated Sds and *Ct*CBH5A, 400 U/g of pretreted Sds without *Ct*GH1 loading was 38.1 mg/g of pretreated Sds. This TRS yield increased to 46.3 mg/g of pre-treated Sds with the addition of *Ct*GH1 from concentration 200–400 U/g of pretreated Sds (Table 2). The increase in *Ct*GH1 loading beyond 400 U/g of pretreated Sds did not cause any significant increase in TRS yield which could be due to the end product inhibition. Therefore, the minimum loading concentration for *Ct*GH1 was 200 U/g of pretreated Sds and total loading concentration for all the three enzymes used in a mixture was predicted to be 2000 U/g of acid pretreated Sds.

In case of base pretreated Sds, the TRS yield increased from 7.8–42.6 mg/g of pretreated Sds with the increase in the loading concentration of *Ba*GH5-UV2 from 200 to 800 U/g of pretreated Sds. Similarly, the TRS yield increased from 1.4–16.1 mg/g of pretreated Sds with increased loading of *Ct*CBH5A from 200 to 800 U/g of pretreated Sds (Table 2). Further increase in loading concentration of *Ba*GH5-UV2 or *Ct*CBH5A did not cause any significant increase in TRS yield similar to the case of acid pretreated Sds. Therefore, the minimum loading concentration of *Ba*GH5-UV2 and *Ct*CBH5A for base pretreated Sds was 200 U/g of pretreated biomass. *Ct*GH1 enzyme alone did not produce any reducing sugars also from base pretreated Sds for the same reason as discussed for acid pretreated Sds earlier. Therefore, similar to the acid pretreated Sds, the loading concentration of *Ct*GH1 for base pretreated Sds was also optimized at fixed concentration of *Ba*GH5-UV2, 800 U/g of pretreated Sds and *Ct*CBH5A, 400 U/g of pretreated Sds.

The TRS yield from base pretreated Sds with fixed *Ba*GH5-UV2, 800 U/g of pretreated Sds and *Ct*CBH5A, 400 U/g of pretreted Sds without *Ct*GH1 loading was 49.4 mg/g of pretreated Sds. This TRS yield increased to 69.8 mg/g of pre-treated Sds with the addition of *Ct*GH1 from concentration 200 to 800 U/g of pretreated Sds (Table 2). The further increase in *Ct*GH1 beyond 800 U/g of pretreated Sds did not cause any significant increase in TRS yield might be due the end product inhibition. Therefore, the minimum loading concentration of *Ct*GH1 was predicted to be 200 U/g of pretreated Sds and total loading

Table 1
Determination of common stable buffer concentration for crude recombinant enzymes *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1.

Enzyme	Residual enzyme activity, % 20 mM Sodium phosphate buffer, pH 6.5			Residual enzyme activity, % 50 mM, Sodium phosphate buffer, pH 6.5		
	6 h	12 h	24 h	6 h	12 h	24 h
*Ba*GH5-UV2	100 ± 9.8	100 ± 11.7	100 ± 9.5	100 ± 9.8	98.6 ± 7.9	95.0 ± 5.6
*Ct*CBH5A	100 ± 5.3	100 ± 6.2	100 ± 3.7	98.2 ± 3.3	96.3 ± 3.5	91.2 ± 2.6
*Ct*GH1	100 ± 1.1	99.8 ± 1.8	99.1 ± 2.8	92.1 ± 1.4	86.5 ± 1.1	60.9 ± 1.9

Endoglucanase (*Ba*GH5-UV2) and Cellobiohydrolase (*Ct*CBH5A) activity was determined against 1% (w/v) CMC-Na by method as described earlier by Nelson (1944) and Somogyi (1945).
β-glucosidase activity (*Ct*GH1) was determined against 1% (w/v) cellobiose by GOD-POD method (Raabo et al. 1960).
± values are mean standard deviation (n = 3).

TH-2393_146106034

Table 2
Customization of limits of enzymes loading concentration for mixture design in saccharification of acid or base pretreated *Sorghum durra* stalk.

Pretreatment methods	Enzyme loading concentration (U/g of pretreated biomass)							
	Enzymes	0 U/g	200 U/g	400 U/g	800 U/g	1200 U/g	1600 U/g	2000 U/g
			TRS yield (mg/g of pretreated biomass), 48 h					
Acid	*Ba*GH5-UV2	ND	3.4 ± 0.2	11.3 ± 0.9	27.6 ± 1.2	28.1 ± 2.5	28.5 ± 2.1	27.4 ± 3.6
	*Ct*CBH5A	ND	1.1 ± 0.01	2.8 ± 0.1	9.4 ± 1.0	9.7 ± 1.5	9.3 ± 0.8	9.8 ± 0.8
	*Ct*GH1*	38.1 ± 1.8	39.5 ± 2.6	46.3 ± 3.5	46.3 ± 2.2	47.5 ± 2.9	48.1 ± 4.0	47.8 ± 3.9
Base	*Ba*GH5-UV2	ND	7.8 ± 0.9	19.9 ± 1.1	42.6 ± 3.5	41.2 ± 4.1	44.7 ± 2.7	43.6 ± 3.4
	*Ct*CBH5A	ND	1.4 ± 0.1	9.5 ± 0.8	16.1 ± 1.4	16.9 ± 1.1	16.0 ± 1.7	17.3 ± 2.3
	*Ct*GH1*	49.4 ± 2.3	51.7 ± 2.5	60.9 ± 5.6	69.8 ± 3.8	70.2 ± 5.3	69.7 ± 5.0	70.9 ± 4.4

Acid = 1% (v/v) H_2SO_4 assisted autoclaving pretreatment.
Base = 1% (w/v) NaOH assisted autoclaving pretreatment.
ND = Not determined.
*At fixed concentration of*Ba*GH5-UV2 (800 U/g of pretreated biomass) and *Ct*CBH5A (400 U/g of pretreated biomass) loading (*Ba*GH5-UV2 + *Ct*CBH5A) with varying *Ct*GH1.
± standard deviation (n = 3).

concentration of all the three enzymes was predicted to be 2000 U/g of pretreated Sds, for base pretreated Sds.

3.4. Enzyme ratio optimization in mixture design with TRS yield and glucose yield as the response variable against acid or base pretreated Sds

3.4.1. Model data sheet

The optimization of all the three crude enzymes *Ba*GH5-UV2 (A), *Ct*CBH5A (B) and *Ct*GH1 (C) in a mixture within a range of lower limit 200 U/g (10 %) of pretreated Sds to maximum limit 1600 U/g (80 %) of pretreated Sds at total loading concentration of proteins, 2000 U/g of pretreated Sds was carried out for getting the maximum TRS yield and glucose yield from acid or base pretreated Sds. On the basis of simplex-centroid model, 10 run data set was generated using RSM approach with varying ratios of (A), (B) and (C) (Table 3). In case of acid pretreated Sds run 9 (A45 %, B10 % and C45 %) gave the highest TRS yield and glucose yield of 90.8 mg/g and 69.3 mg/g of pretreated Sds, respectively. Run 1 (A21.6 %, B21.7 % and C56.6 %) gave the second highest TRS yield and glucose yield of 89.5 and 61.4 mg/g of pretreated Sds, respectively at 48 h of saccharification. This showed that the acid pretreated Sds gave maximum TRS yield and glucose yield when the proportion of *Ba*GH5-UV2 and *Ct*GH1 were high and *Ct*CBH5A was relatively low in the enzyme cocktail. Kim et al. (2015) studied that for acid pretreated rice straw, relatively high proportion of cellobiohydrolase (E3, 66.6 %) along with low proportion of endoglucanase (Cel5H, 16.6 %) and xylanase (Xyn10C, 16.6 %) was required for getting the maximum reducing sugars. Kim et al. (2017) observed that for acid pretreated sugarcane bagasse, relatively high proportion of endoglucanase (50 %) and cellobiohydrolase (50 %) is required along with zero proportion of xylanase for getting the maximum reducing sugars.

In case of base pretreated Sds the run 4 (A80 %, B10 % and C10 %) gave the highest TRS yield and glucose yield of 165.6 mg/g and 81.7 mg/g of pretreated Sds, respectively. Run 5 (A45 %, B45 % and C10 %) gave the second highest TRS yield and glucose yield of 122.0 and 59.1 mg/g of pretreated Sds, respectively, at 48 h of saccharification. This showed that in case of base pretreated Sds, *Ba*GH5-UV2 has the major role in contributing for TRS yield and glucose yield because with high proportion of *Ba*GH5-UV2 in enzyme mixture, TRS yield and glucose yield was maximum. Suwannarangsee et al. (2012) showed that for alkali pretreated rice straw relatively equal proportion of commercial cellulase (Celluclast™, 41.4 %) and BCC199 enzyme (crude enzyme extract from *Aspergillus aculeatus,* 37.0 %) along with low proportion of expansin (21.6 %) was required for getting the maximum reducing sugars. Kim et al. (2015) studied that for base pretreated rice straw equal proportion of endoglucanase (Cel5H, 33.3 %), cellobiohydrolase (E3, 33.3 %) and xylanase (Xyn10C, 33.3 %) was required for getting the maximum reducing sugars. Kim et al. (2017) observed that for base pretreated sugarcane bagasse relatively high proportion of endoglucanase (66.7 %) with equal proportion of cellobiohydrolase (16.7

Table 3
Simplex-centroid mixture design predicted ratios for enzymes and experimental data response variables.

Run no.	Predicted enzyme variables % (U/g of pretreated Sds)			Acid pretreated Sds[a]		Base pretreated Sds[b]	
	*Ba*GH5-UV2 (A)	*Ct*CBH5A (B)	*Ct*GH1 (C)	TRS yield (mg/g of pretreated Sds)	Glucose yield (mg/g of pretreated Sds)	TRS yield (mg/g of pretreated Sds)	Glucose Yield (mg/g of pretreated Sds)
1	21.6	21.7	56.6	89.5 ± 4.7	61.4 ± 2.5	70.1 ± 2.3	36.5 ± 2.4
2	10	10	80	39.6 ± 1.4	26.4 ± 1.4	81.2 ± 3.1	42.1 ± 1.9
3	10	80	10	45.9 ± 2.3	29.0 ± 1.6	86.7 ± 3.7	43.7 ± 2.7
4	80	10	10	26.8 ± 1.1	17.9 ± 2.4	165.6 ± 5.4	81.7 ± 3.2
5	45	45	10	62.1 ± 4.6	39.5 ± 2.1	122.0 ± 4.6	59.1 ± 3.2
6	56.6	21.7	21.6	69.6 ± 1.3	49.4 ± 2.0	112.1 ± 4.9	55.2 ± 2.3
7	33.3	33.3	33.3	83.3 ± 2.9	53.9 ± 3.8	115.9 ± 3.8	57.1 ± 3.1
8	10	45	45	76.5 ± 3.0	56.1 ± 2.5	63.4 ± 2.2	30.2 ± 1.4
9	45	10	45	90.8 ± 2.8	69.3 ± 3.1	63.3 ± 2.0	29.2 ± 1.8
10	21.6	56.7	21.6	71.9 ± 2.5	44.5 ± 1.3	94.5 ± 3.1	44.6 ± 2.0

± Standard deviation (n = 3).
Sds = *Sorghum durra* stalk.
(a) 1% (v/v) H_2SO_4 assisted autoclaving pretreated *Sorghum durra* stalk response variables.
(b) 1% (w/v) NaOH assisted autoclaving pretreated *Sorghum durra* stalk response variables.

TH-2393_146106034

%) and xylanase (16.7 %) or equal proportion of all the three enzymes was required for getting the maximum reducing sugars.

The proportion of *Ba*GH5-UV2 required in base pretreated Sds was 2-fold higher than the acid pretreated Sds. This was due to the fact that the base pretreatment displays lower effect on the initial cellulose degree of polymerization (DP) than the acid pretreatment, which results in the cellulosic polysaccharide hydrolysis with less number of free ends for action of cellobiohydrolase (Levine et al. 2011). Therefore, the required proportion of *Ba*GH5-UV2 was relatively more (80 %) for base pretreated Sds than acid pretreated Sds (45 %). The proportion of *Ct*CBH5A for both the acid or base pretreated Sds was less (10 %) or almost same. This was due to the fact that *Ba*GH5-UV2 enzyme form cellobiose in addition to the cellotriose and cellotetraose as the end products upon CMC-Na hydrolysis within 24 h (Singh et al., 2020a), where cellobiose causes substrate inhibition for *Ct*CBH5A. The base pretreated Sds showed 1.8-fold higher TRS yield and 1.2-fold higher glucose yield than acid pretreated Sds. This may be due to the presence of 5.4-fold (11.9 %, w/w) higher ADL content, in acid pretreated Sds than the base pretreated Sds (2.2 %, w/w). Lignin hinders the enzymatic saccharification of cellulose by the non-productive adsorption of cellulases (Ko et al., 2015). In addition to this lignin intrinsically surrounds the cellulose fibres and provides recalcitrance nature to the biomass because of which biomass becomes less accessible for enzymes (Ladeira et al., 2020). Moreover, crystallinity index (*CrI*) is one of the other factors which affects the enzymatic saccharification yield from pretreated biomass (Li et al., 2014; Sun et al., 2014; Kumar et al., 2020). The base pretreated Sds showed higher *CrI*, 58.5 % than that with acid pretreated Sds, *CrI*, 48.5 % (Jamaldheen et al., 2018). Due to this higher *CrI* value, of base pretreated Sds, the TRS yield and glucose yield were higher as compared with acid pretreated Sds. Earlier report of Li et al. (2014) showed that the reducing sugar yield from biomass was linearly proportional to the *CrI* value at optimum loading concentration of enzymes. However, the percentage of cellulose fraction in acid pretreated Sds was higher (57.0 %, w/w) than that in base pretreated Sds (53.1 %, w/w) but, this did not have significant effect on enzymatic saccharification, which may be due to the higher lignin content obtained in acid pretreated Sds. Similarly, Kim et al. (2015) reported that due to composition differences between the acid and base pretreated rice straw, the base pretreated rice straw produced high reducing sugar yield of 21.1 μg than acid pretreated rice straw, 15.6 μg.

3.4.2. Fitting of the regression models

3.4.2.1. Model selection. Different models were compared for selecting the model that explains the best relationship between independent variables (A, B and C) and dependent variable (TRS yield and glucose yield from acid or base pretreated Sds). Various predicted models for TRS yield and glucose yield from acid or base pretreated Sds were shown in (Tables S1–S4). For a model to be considered the most suitable, the F-value should be as high and the p-value should be low (Chakraborty et al., 2018). In case of acid pretreated Sds, the quadratic model showed the highest F value of 55.54 with lowest p-value of < 0.001 for TRS yield as a response factor and highest F value, 29.81with lowest p-value, < 0.0034 for glucose yield as a response factor. Therefore, among all models, the quadratic model was found to be the best suited model for acid pretreated Sds while considering TRS yield and glucose yield as response variables. In case of base pretreated Sds, the special cubic model showed the highest F value of 10.4 with lowest p-value of < 0.0484 for TRS yield as a response factor and highest F value, 14.6 with lowest p-value, < 0.0315 for glucose yield as a response factor. Therefore, the special cubic model was considered to be the best suited model for base pretreated Sds, while considering the TRS yield and glucose yield as response variables. Suwannarangsee et al. (2012) optimized the ternary enzyme complex of enzymes (Celluclast™, BCC 199, expansin) in a mixture for saccharification of alkali pretreated rice straw and found that the full cubic model was the best predicted one for reducing sugar as response factor. Similarly, Kim et al. (2015) optimized the ratio of Cel5H, E3 (CBH) and Xyn10C for formulation of optimal ratio of enzymes in a cocktail for saccharification of acid/alkali pretreated rice straw and found full cubic model was the best predicted one for reducing sugar as response factor. Kim et al. (2017) showed that for enzyme mixture design using endoglucanase, cellobiohydrolase, β-glucosidase and endoxylanase for saccharification of acid and base pretreated sugarcane bagasse, special cubic and full cubic model, respectively, were the best predicted one.

3.4.2.2. Analysis of variance (ANOVA). ANOVA results for acid pretreated Sds confirmed the significance of quadratic model for both the response factors, TRS yield and glucose yield. The R^2 and adjusted R^2 values for the generated model for TRS yield as response factor were 0.97 and 0.95, respectively, and for glucose yield as response factor were 0.97 and 0.92, respectively. In RSM mixture design F value among the variables of mixture component help in predicting the interactions among them. A high F value between two variables suggested high interactions and low F value suggested low interactions between them (Chakraborty et al., 2018). In case of acid pretreated Sds, for TRS yield as the response factor, interactions of AC variables showed highest F value of 111.58 which was followed by BC variables, F value 37.63 and AB variables gave the lowest F value of 19.08 (Table 4). For glucose yield as the response factor, interactions of AC variables showed highest F value of 65.21 which was followed by BC variables, F value 20.62 and AB variables gave the lowest F value of 4.38 (Table 4). These results suggested that the interactions were maximum between the

Table 4
Analysis of Variance output data summary for TRS yield and glucose yield as response variables for acid* pretreated *Sorghum durra* stalk.

Response variables	Analysis of Variance output data						
	Source	Sum of squares	df	Mean square	F Value	p-Value Prob > F	Comments
TRS yield	Model	4179.43	5	835.89	35.90	0.0020	Significant
	Linear Mixture	300.55	2	150.28	6.46	0.0560	
	AB	444.13	1	444.13	19.08	0.0120	
	AC	2597.56	1	2597.56	111.58	0.0005	
	BC	876.00	1	876.00	37.63	0.0036	
	Residual	93.12	4	23.28			
	Cor Total	4272.55	9				
Glucose yield	Model	2321.74	5	464.35	19.71	0.0064	Significant
	Linear Mixture	214.36	2	107.18	4.55	0.0933	
	AB	103.09	1	103.09	4.38	0.1046	
	AC	1536.49	1	1536.49	65.21	0.0013	
	BC	485.77	1	85.77	20.62	0.0105	
	Residual	94.25	4	23.56			
	Cor Total	2416.00	9				

acid* = 1% (v/v) H_2SO_4 assisted autoclaving pretreatment.

TH-2393_146106034

Table 5
Analysis of Variance output data summary for TRS yield and glucose yield as response variables for base* pretreated *Sorghum durra* stalk.

Response variables	Analysis of Variance output data						
	Source	Sum of squares	df	Mean square	F Value	p-Value Prob > F	Comments
TRS yield	Model	9018.58	6	1503.10	17.55	0.0195	Significant
	Linear Mixture	6238.05	2	3119.03	36.42	0.0079	
	AB	19.14	1	19.14	0.22	0.6686	
	AC	2573.56	1	2573.56	30.05	0.0119	
	BC	297.24	1	297.24	3.47	0.1594	
	ABC	890.33	1	890.33	10.40	0.0484	
	Residual	256.94	3	85.65			
	Cor Total	9275.52	9				
Glucose yield	Model	2203.58	6	367.26	19.42	0.0169	Significant
	Linear Mixture	1396.31	2	698.15	36.9	0.0077	
	AB	12.74	1	12.74	0.67	0.4719	
	AC	740.54	1	740.54	39.15	0.0082	
	BC	112.07	1	112.07	5.92	0.0930	
	ABC	276.49	1	276.49	14.62	0.0315	
	Residual	56.75	3	18.92			
	Cor Total	2260.33	9				

base* = 1% (w/v) NaOH assisted autoclaving pretreatment.

endoglucanase, *Ba*GH5-UV2 (A) and β-glucosidase, *Ct*GH1 (C) with the highest synergistic effect, whereas, the synergistic effect of endoglucanase (*Ba*GH5-UV2) and cellobiohydrolase, *Ct*CBH5A (B) was minimum on acid pretreated Sds. These results also corroborated the data of enzyme proportion in a mixture, where the addition of 45 % of *Ct*GH1 in 45 % of *Ba*GH5-UV2 and 10 % of *Ct*CBH5A, the TRS yield and glucose yield was 90.8 and 69.3 mg/g of pretreated Sds, respectively, whereas the enzyme mixture having 80 % of *Ba*GH5-UV2 and a low proportion of *Ct*GH1 (10 %) along with *Ct*CBH5A (10 %) gave TRS yield, 26.8 mg/g of pretreated Sds and glucose yield, 17.9 mg/g of pretreated Sds. This 3.4- and 3.8-fold increase in TRS and glucose yields suggested strong synergism between *Ba*GH5-UV2 and *Ct*GH1 enzymes. Similarly, upon addition of 45 % of *Ba*GH5-UV2 in 45 % of *Ct*CBH5A along with very low proportion of *Ct*GH1 (10 %) the TRS yield and glucose yield was 62.1 and 39.5, respectively, whereas with the enzyme mixture having 10 % of *Ba*GH5-UV2, 80 % of *Ct*CBH5A and 10 % of *Ct*GH1 have TRS and glucose yield 45.9 and 29.0 mg/g of pretreated Sds. Although in the former case there was 1.35- and 1.36-fold higher TRS and glucose yields than the latter, but the synergy between *Ba*GH5-UV2 and *Ct*CBH5A enzymes was also low. This was due to the fact the enzyme mixture having less proportion of *Ba*GH5-UV2 (10 %) with high proportion of *Ct*CBH5A (45 %) and *Ct*GH1 (45 %) produced 1.7- fold higher TRS yield and 1.9-fold higher glucose yield than mixture having 10 % of *Ba*GH5-UV2, 80 % of *Ct*CBH5A and 10 % of *Ct*GH1. Therefore, there was moderate synergy between *Ct*CBH5A and *Ct*GH1 instead of *Ba*GH5-UV2 and *Ct*CBH5A.

The variation in the TRS yield with variation in A (*Ba*GH5-UV2), B (*Ct*CBH5A) and C (*Ct*GH1) compositions from acid pretreated Sds could be expressed by using the following equations generated from quadratic model:

TRS yield (mg/g of pretreated Sds) = −0.34763 A + 0.12190 B − 0.19983 C + 0.019117 AB + 0.046233 AC + 0.026849 BC

The variation in the glucose yield with variation in A (*Ba*GH5-UV2), B (*Ct*CBH5A) and C (*Ct*GH1) compositions from acid pretreated Sds could be expressed by using the following equations generated from quadratic model:

Glucose yield (mg/g of pretreated Sds) = −0.22399 A + 0.074904 B − 0.19632 C + 0.00921052 AB + 0.035558 AC + 0.019993 BC

ANOVA results for base pretreated Sds confirmed the significance of special cubic model for both the response factors, TRS yield and glucose yield. In case of base pretreated Sds, the R^2 and adjusted R^2 values for the generated model for TRS yield as response factor were 0.97 and 0.91, respectively, and for glucose yield as response factor were 0.97 and 0.92, respectively. For TRS yield as the response factor, interactions between AC variables showed highest F value of 30.05 which was followed by ABC variables, F value 10.4 and this was followed by BC variables, F value 3.47. However, the interactions between the AB variables gave the lowest F value of 0.22 (Table 5). For glucose yield as the response factor, interactions of AC variables showed highest F value of 39.15 which was followed by ABC variables, F value 14.62 and this was followed by BC variable, F value 5.92 (Table 5). Whereas, the AB variables gave the lowest F value of 0.67 (Table 5) thus, showed the negligible synergistic effect. These results suggested that the interactions were maximum between the endoglucanase (*Ba*GH5-UV2) and β-glucosidase (*Ct*GH1) showing the highest synergistic effect whereas the interactions between endoglucanase (*Ba*GH5-UV2) and cellobiohydrolase (*Ct*CBH5A) showed the negligible synergistic effect on base pretreated Sds. But all the three enzymes (*Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1) in mixture act synergistically. Depending on the F value, these results also showed that in enzymatic hydrolysis of base pretreated Sds, there was less synergism between endo-acting (*Ba*GH5-UV2) and exo-acting, cellobiohydrolase (*Ct*CBH5A) and β-glucosidase (*Ct*GH1) as compared with the acid pretreated Sds. This might have happened due to the composition differences in the biomass because acid pretreated Sds has lower, 10.6 % (w/w) hemicellulose whereas, base pretreated Sds has higher, 21.8 % (w/w) of hemicellulose. The above results also corroborated the data of enzyme proportion in a mixture where 80 % of *Ba*GH5-UV2 along with significantly less proportion of *Ct*CBH5A (10 %) and *Ct*GH1 (10 %) produced the highest TRS yield of 165.6 and glucose yield of 81.7 mg/g of pretreated Sds, whereas, the enzyme mixture with 45 % of *Ba*GH5-UV2, 45 % of *Ct*GH1 along with low proportion of *Ct*CBH5A (10 %) gave TRS yield, 63.3 mg/g of pretreated Sds and glucose yield, 29.2 mg/g of pretreated Sds. Here with increasing the proportion of *Ct*GH1 decrease in TRS yield and glucose yield was observed. Therefore, these results suggested that there was less synergy between enzymes (*Ba*GH5-UV2 and *Ct*GH1, *Ba*GH5-UV2 and *Ct*CBH5A) for base pretreated Sds as compared with the acid pretreated Sds. Therefore, the synergism between the enzymes also depends on pretreatment method. Suwannarangsee et al. (2012) found similar synergism when 50:50 ratio of Celluclast™ to BCC enzyme extract showed 70 % high reducing sugar yield than Celluclast™ alone from alkali pretreated rice straw. Jia et al. (2015) reported that synergism between cellulase and xylanase for the hydrolysis of sugarcane bagasse depends on the type of pretreatment method used. Kim et al. (2017) also predicts that synergism between endoglucanase and CBH or endoglucanase,

TH-2393_146106034

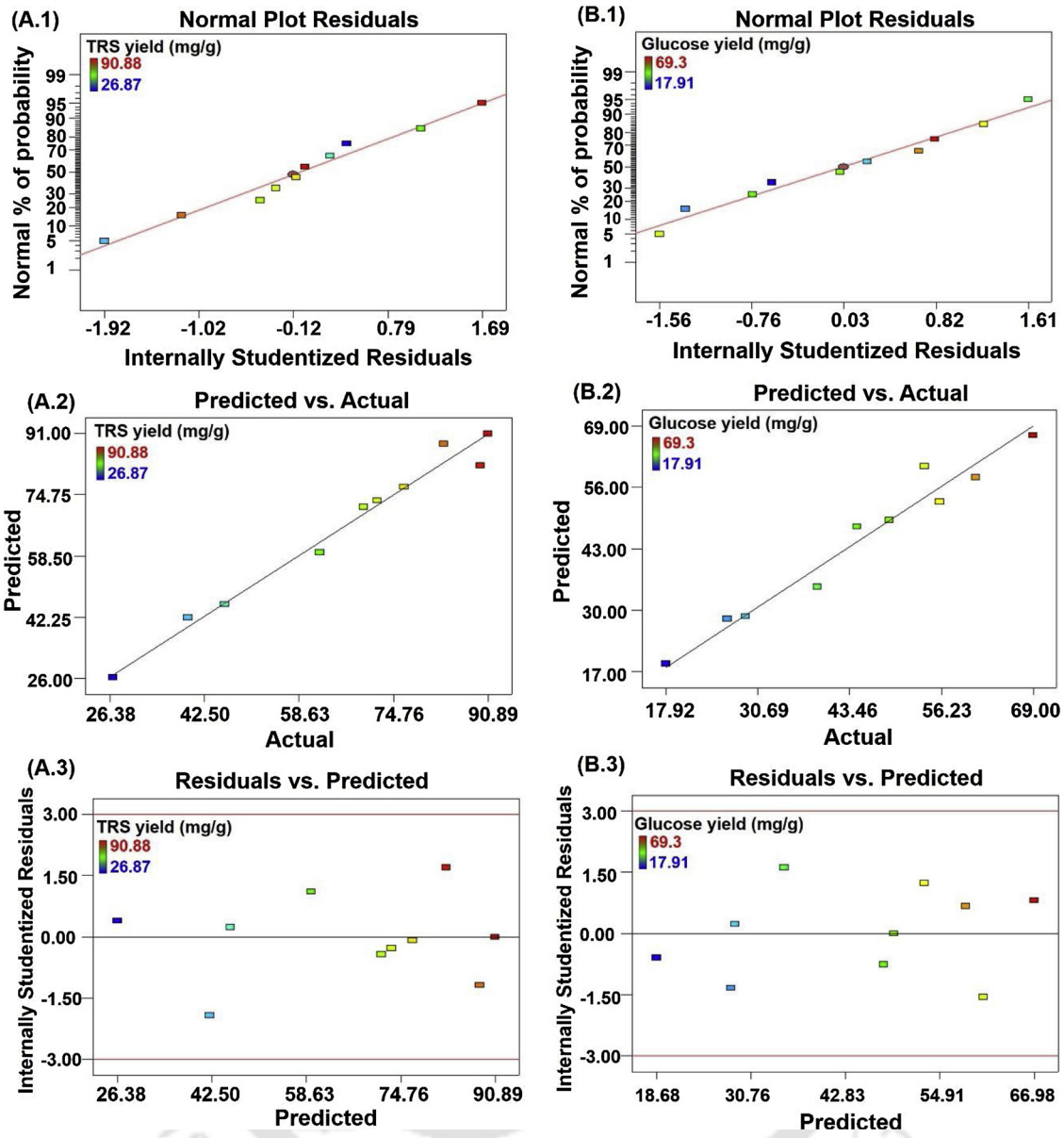


Fig. 1. RSM response plots for TRS yield (mg/g of pretreated *Sorghum durra* stalk) from acid pretreated *Sorghum durra* stalk: (A.1) Normal Plot of Residuals, (A.2) Predicted vs Actual plot, (A.3) Residuals vs Predicted Plot. RSM response plots for glucose yield (mg/g of pretreated *Sorghum durra* stalk) from acid pretreated *Sorghum durra* stalk: (B.1) Normal Plot of Residuals, (B2) Predicted vs Actual plot, (B3) Residuals vs Predicted Plot.

CBH and xylanase depends on the type of pretreatment, acid or alkali due to the composition differences between them.

The variation in the TRS yield with variation in A (*Ba*GH5-UV2), B (*Ct*CBH5A) and C (*Ct*GH1) compositions from base pretreated Sds could be expressed by using the following equations generated from special cubic model:

TRS yield (mg/g of pretreated Sds) = + 2.58753 A + 1.16094 B + 1.53112 C − 0.032253 AB − 0.078287 AC − 0.045029 BC + 0.00279078 ABC

The variation in the glucose yield with variation in A (*Ba*GH5-UV2), B (*Ct*CBH5A) and C (*Ct*GH1) compositions from base pretreated Sds could be expressed by using the following equations generated from special cubic model:

Glucose yield (mg/g of pretreated Sds) = + 1.31995 A + 0.61925 B + 0.83946 C − 0.019097 AB − 0.042577 AC − 0.02606 BC + 0.00155522 ABC

3.4.2.3. Diagnostics. The diagnostics of TRS yield and glucose yield as a response factor from acid pretreated Sds shown in Fig. 1 and base pretreated Sds shown in Fig. 2. The figures of the normal plots residual showed all the data points very close to the straight line (Figs. 1A.1, and B.1, 2 A.1 and B.1) suggesting the non-normality of data points as also reported earlier (Chakraborty et al., 2018). The predicted vs actual plots showed the data points were located very close to the diagonal line (Figs. 1A.2, and B.2, 2 A.2 and B.2). Thus, displayed the locus of similarity between actual and predicted values as reported also earlier (Chakraborty et al., 2018). The residual vs predicted plots showed the existence of linear patterns (Figs. 1A.3, and B.3, 2 A.3 and B.3), which is ensured by observing scattered data across the residual (horizontal central line) according to an earlier report of Chakraborty et al. (2018). Therefore, the expressed quadratic model for acid pretreated Sds and special cubic model for base pretreated Sds, were suitable to represent the appropriate variations in the response variables as functions of variation in the cocktail composition of *Ba*GH5-UV2, *Ct*GH5A and *Ct*GH1.

TH-2393_146106034

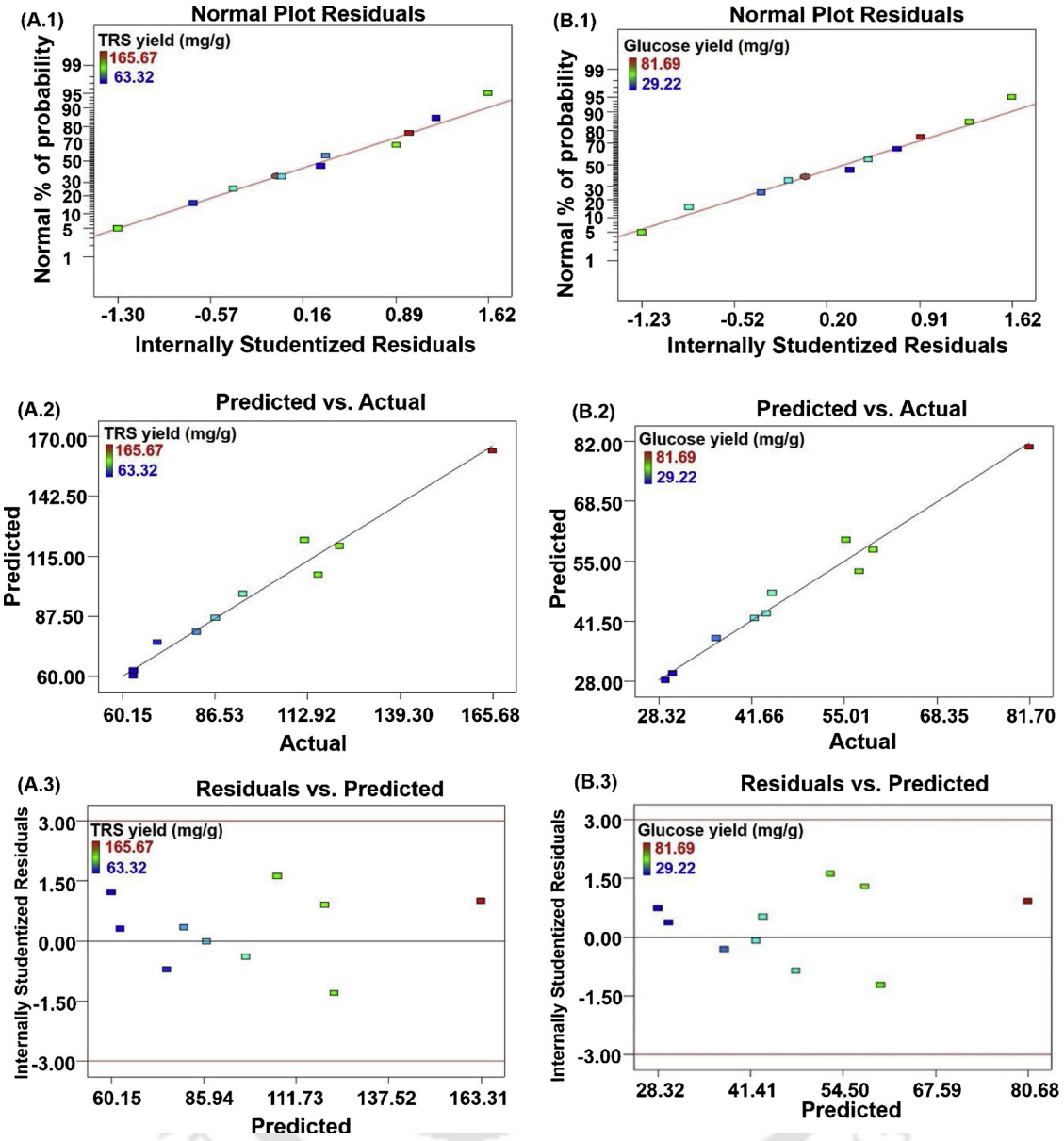


Fig. 2. RSM response plots for TRS yield (mg/g of pretreated *Sorghum durra* stalk) from base pretreated *Sorghum durra* stalk: (A.1) Normal Plot of Residuals, (A.2) Predicted vs Actual plot, (A.3) Residuals vs Predicted Plot. RSM response plots for glucose yield (mg/g of pretreated *Sorghum durra* stalk) from base pretreated *Sorghum durra* stalk: (B.1) Normal Plot of Residuals, (B.2) Predicted vs Actual plot, (B.3) Residuals vs Predicted Plot.

3.4.3. Model graphs

The ternary 3D contour plots were developed to determine the optimal ratio of *Ba*GH5-UV2 (A), *Ct*CBH5A (B) and *Ct*GH1 (C) in the enzyme mixture for the saccharification of acid pretreated Sds on the basis of the quadratic equations generated above in Section 3.4.2.2 and for base pretreated Sds on the basis of special cubic equations generated above in Section 3.4.2.2. In case of acid pretreated Sds, the contour plots for TRS yield and glucose yield responses from the quadratic models showed the areas that outlined the highest TRS yield of 90.8 mg/g of pretreated Sds (Fig. 3A) and glucose yields of 69.3 mg/g of pretreated Sds (Fig. 3B) were positioned for mixtures with high proportions of *Ba*GH5-UV2 and *Ct*GH1 and a fairly low proportion of *Ct*CBH5A. These results also confirmed the strong positive synergism between endo-acting (*Ba*GH5-UV2) and exo-acting, β-glucosidase (*Ct*GH1) and low synergism between *Ba*GH5-UV2 and *Ct*CBH5A in the hydrolysis of acid pretreated Sds. However, in case of base pretreated Sds, the contour plots for TRS yield and glucose yield responses were developed from the special cubic models. These models showed the areas that outlined the highest TRS yield of 165.6 mg/g of pretreated Sds (Fig. 4A) and glucose yields of 81.7 mg/g of pretreated Sds (Fig. 4B) were positioned for mixtures with high proportions of *Ba*GH5-UV2 and low proportion of *Ct*CBH5A and *Ct*GH1. These results also confirm the less synergism between endo-acting (*Ba*GH5-UV2) and exo-acting, β-glucosidase (*Ct*GH1) as compared with the acid pretreated Sds. Moreover, there no synergism observed between *Ba*GH5-UV2 and *Ct*CBH5A in the enzymatic hydrolysis of base-pretreated Sds. This could be because endoglucanase, *Ba*GH5-UV2 from UV2 mutant strain of *Bacillus amyloliquefaciens* SS35 releases also cellobiose along with the higher cello-oligosaccharides as reported earlier (Singh et al., 2020a) which results end product inhibition for *Ct*CBH5A. Synergism of *Ct*GH1 with *Ba*GH5-UV2 and *Ct*CBH5A depends on the end product, cellobiose released from the *Ba*GH5-UV2 and *Ct*CBH5A which depends on the composition of pretreated Sds. Gao et al. (2010) reported the influence of protein loading in the ratio of cocktail composition for the saccharification of ammonia fibre expansion pretreated corn stover and observed that the endoglucanase was more important to glucose release

TH-2393_146106034

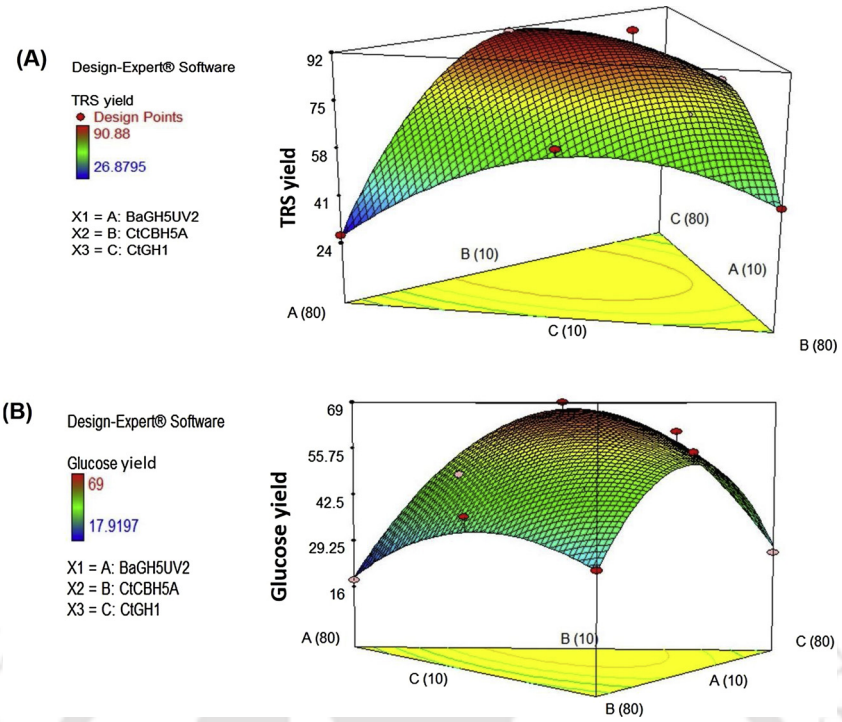


Fig. 3. (A) 3D contour plot for TRS yield (mg/g of pretreated *Sorghum durra* stalk) response from acid pretreated *Sorghum durra* stalk; (B) 3D contour plot for glucose yield (mg/g of pretreated *Sorghum durra* stalk) response from acid pretreated *Sorghum durra* stalk.

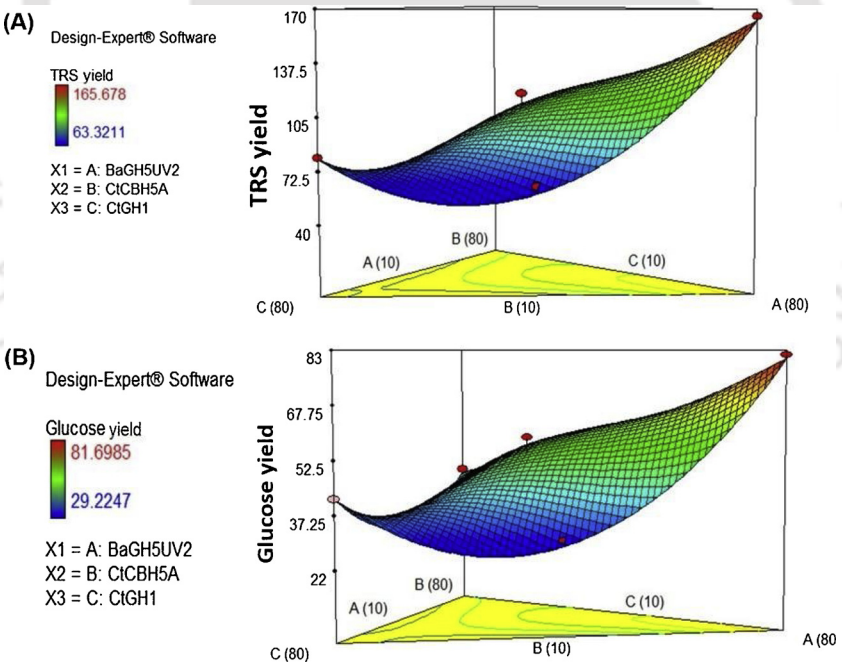


Fig. 4. (A) 3D contour plot for TRS yield (mg/g of pretreated *Sorghum durra* stalk) response from base pretreated *Sorghum durra* stalk; (B) 3D contour plot for glucose yield (mg/g of pretreated *Sorghum durra* stalk) response from base pretreated *Sorghum durra* stalk.

with a low protein load of exoglucanase.

3.4.4. Optimal mixture ratio composition for acid and base pretreated Sds

The specific optimal ratios leading to the maximal production of TRS yield and glucose yield were determined by setting the goal limit maximum for TRS yield and glucose yield. In case of acid pretreated Sds the optimal mixture ratios composition of *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 was 43.1 %, 10.0 %, and 46.9 %, respectively, predicted by the quadratic model. In case of base pretreated Sds the optimal mixture ratios composition of *Ba*GH5-UV2, *Ct*CBH5A and *Ct*GH1 was 80 %, 10 %, and 10 %, respectively, predicted by special cubic models. This slight change in optimized percentage ratio for acid pretreated Sds from 45 %:10 %:45 % (*Ba*GH5-UV2:*Ct*CBH5A:*Ct*GH1) was finally confirmed the strong synergism between the enzymes whereas constant ratio for base pretreated Sds (80 %:10 %:10 %,) confirmed the week synergism between the enzymes which corroborates the prediction of interaction from F value. Kim et al. (2015) proposed the pretreatment specific customization of Cel5H, E3 (CBH) and Xyn10C enzymes ratios in a

TH-2393_146106034

Table 6
TRS and glucose yield from acid or base pretreated *Sorghum durra* stalk for different time interval.

Pretreatment	Time (h)	TRS yield (mg/g of pretreated Sds)	Glucose yield (mg/g of pretreated Sds)
Acid pretreated Sds	24	38.6 ± 1.5	10.2 ± 1.1
	48	91.1 ± 2.2	69.5 ± 1.5
	72	112.6 ± 3.4	84.7 ± 1.1
	96	119.5 ± 3.4	85.3 ± 1.9
Base pretreated Sds	24	59.2 ± 1.3	26.3 ± 1.2
	48	165.8 ± 2.1	81.7 ± 1.4
	72	193.3 ± 1.1	159.9 ± 2.0
	96	211.0 ± 3.8	174.2 ± 1.8

Sds = *Sorghum durra* stalk.
± Standard deviation (n = 3).
Acid pretreated *Sorghum durra* stalk = 1% (v/v) H_2SO_4 assisted autoclaving.
Base pretreated *Sorghum durra* stalk = 1% (w/v) NaOH assisted autoclaving.

cocktail for efficient saccharification of acid and base pretreated rice straw. The optimal predicted ratio of Cel5H, E3 (CBH) and Xyn10C for acid pretreated rice straw was 23.23 %:76.77 %:0, whereas for base pretreated rice straw was 30.04 %:26.54 %:43.43 %. Bussamra et al. (2015) optimized the cocktail supplementation of commercial and microbial (*Trichoderma reesei* and *Escherichia coli*) cellulase enzymes for sugarcane bagasse hydrolysis by simplex lattice mixture design. The optimized enzyme mixture, comprised *T. reesei* fraction (80 %), endoglucanase (10 %) and β-glucosidase (10 %). Kim et al. (2017) found the optimized predicted ratio for acid pretreated Sugarcane bagasse was 61.25 %:38.73 %:0.02 % (endoglucanase, cellobiohydrolase, and exoxylanase) and for base pretreated Sugarcane bagasse 53.99 %:34.60 %:11.41 % (endoglucanase, cellobiohydrolase, and exoxylanase) by RSM design.

3.4.5. Saccharification of acid and base pretreated Sds under optimized conditions

The statistical prediction was validated by carrying out the enzymatic saccharification of acid or base pretreated Sds under optimized enzyme ratios (given in Section 3.4.4) and the enzymatic hydrolysates were analysed for TRS yield and glucose yield (Table 6). The acid pretreated Sds after 48 h of enzymatic saccharification gave TRS yield and glucose yield, 91.1 mg/g and 69.5 mg/g of pretreated Sds, respectively and base pretreated Sds gave 165.8 mg/g and 81.7 mg/g of pretreated Sds, respectively (Table 6). These results showed that the scale up validation of experimental cocktail ratios (given in Section 3.4.4) for TRS yield and glucose yield from acid or base pretreated biomass was correct as it gave the TRS and glucose yield similar to the 1 mL enzymatic saccharification reaction volume. Further, in case of acid pretreated Sds, the maximum TRS yield and glucose yield were 112.6 mg/g (2.25 g/L) and 84.7 mg/g (1.7 g/L) of pretreated Sds, respectively, at 72 h which remained almost constant thereafter (Table 6). However, in case of base pretreated Sds, the TRS yield, 211.0 mg/g (4.2 g/L) of pretreated Sds and glucose yield, 174.2 mg/g (3.5 g/L) of pretreated Sds reached maximum at 96 h (Table 6). This glucose yield was much higher than 34 mg/g of pretreated Sds obtained from unoptimized enzyme ratios reported earlier (Jamaldheen et al., 2018). Another report showed a maximum TRS yield of 47.8 mg/g of pretreated Sds by unoptimized enzyme concentration (Nath et al., 2019).

3.5. Conclusion

The rational mixture design approach was applied to optimize the ratio of three crude recombinant enzymes, endoglucanase (*Ba*GH5-UV2), cellobiohydrolase (*Ct*CBH5A) and β-glucosidase (*Ct*GH1) in the mixture for saccharification of Sds. The carbohydrate composition analysis showed that the acid pretreated Sds has low hemicellulose 10.6 % (w/w) and high ADL 11.9 % (w/w) unlike the base pretreated Sds which has high hemicellulose 21.8 % (w/w) and low ADL 2.2 % (w/w) content. The optimized enzyme cocktail ratio for acid pretreated Sds was *Ba*GH5-UV2 (43.1 %), *Ct*CBH5A (10.0 %) and *Ct*GH1 (46.9 %), which was different from the base pretreated Sds, that was *Ba*GH5-UV2 (80 %), *Ct*CBH5A (10 %) and *Ct*GH1 (10 %). This difference in the enzyme ratios was due to the difference in their composition after the pretreatment. The synergism among the cellulase enzymes were higher in acid pretreated Sds than base pretreated Sds. This may be due to the low hemicellulose content in acid pretreated Sds. Thus, the ratio of cellulases in cocktail and synergism among them depends on the type of pretreatment method. The optimized enzyme ratio for saccharification of acid and base pretreated Sds gave maximum glucose yield, 85 mg/g at 72 h and 174 mg/g of pretreated Sds at 96 h, respectively. The cocktail formulation of crude recombinant enzymes specific to the acid or base pretreatment saves the purification cost of enzyme and time. Moreover, this is the first comprehensive report on the pretreatment specific optimization of cellulase enzymes ratio in a mixture for enzymatic saccharification by using the RSM mixture design approach on substrate Sds. Therefore, a rational enzyme mixture design by using the synergistic concept and statistical analysis can help in improving the enzymatic saccharification yield from different pretreatment methods for a biomass.

Credit author statement

AG designed the study. AG and VSM design the pretreatment method. SJ optimized the pretreatment method. SS and VR performed the pretreatment experiment. SS and VR together performed the cocktail optimization of cellulase enzymes using mixture design approach by Stat-Ease Design Expert software. AG and SS analyse all the experimental data. AG and SS together wrote the paper.

Declaration of Competing Interest

Authors declare no conflict of Interest.

Acknowledgments

The research work was supported by the Centre for Bioenergy DBT-Pan-IIT Grant (BT/EB/PAN-IIT 2012) and DBT-Pan-IIT Grant (Phase II) by the Department of Biotechnology, Ministry of Science and Technology, New Delhi, India. The authors also thankful to the Department of Biosciences and Bioengineering, IIT Guwahati for the instrument facilities.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.indcrop.2020.112678.

References

Banerjee, G., Car, S., Scott-Craig, J.S., Borrusch, M.S., Walton, J.D., 2010. Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. Biotechnol. Biofuels 3 (1), 22.
Bussamra, B.C., Freitas, S., da Costa, A.C., 2015. Improvement on sugarcane bagasse hydrolysis using enzymatic mixture designed cocktail. Bioresour. Technol. 187, 173–181.
Cai, Y.J., Buswell, J.A., Chang, S.T., 1994. Production of cellulases and hemicellulases by the straw mushroom, *Volvariella volvacea*. Mycol. Res. 98 (9), 1019–1024.
Cai, Y.J., Chapman, S.J., Buswell, J.A., Chang, S.T., 1999. Production and distribution of endoglucanase, cellobiohydrolase, and β-glucosidase components of the cellulolytic system of *Volvariella volvacea*, the edible straw mushroom. Appl. Environ. Microbiol. 65 (2), 553–559.
Chakraborty, S., Uppaluri, R., Das, C., 2018. Optimal fabrication of carbonate free kaolin based low cost ceramic membranes using mixture model response surface methodology. Appl. Clay Sci. 162, 101–112.
Chen, C., Boldor, D., Aita, G., Walker, M., 2012. Ethanol production from sorghum by a microwave-assisted dilute ammonia pretreatment. Bioresour. Technol. 110, 190–197.

TH-2393_146106034

da Costa Nogueira, C., de Araújo Padilha, C.E., de Sá Leitão, A.L., Rocha, P.M., de Macedo, G.R., dos Santos, E.S., 2018d. Enhancing enzymatic hydrolysis of green coconut fiber-Pretreatment assisted by tween 80 and water effect on the post-washing. Ind. Crop Prod. 112, 734–740.
Dagnino, E.P., Chamorro, E.R., Romano, S.D., Felissia, F.E., Area, M.C., 2013. Optimization of the acid pretreatment of rice hulls to obtain fermentable sugars for bioethanol production. Ind. Crop Prod. 42, 363–368.
Galbe, M., Zacchi, G., 2002. A review of the production of ethanol from softwood. Appl. Microbiol. Biotechnol. 59 (6), 618–628.
Gao, D., Chundawat, S.P., Krishnan, C., Balan, V., Dale, B.E., 2010. Mixture optimization of six core glycosyl hydrolases for maximizing saccharification of ammonia fiber expansion (AFEX) pretreated corn stover. Bioresour. Technol. 101 (8), 2770–2781.
Garcia, N.F.L., da Silva Santos, F.R., Bocchini, D.A., da Paz, M.F., Fonseca, G.G., Leite, R.S.R., 2018. Catalytic properties of cellulases and hemicellulases produced by *Lichtheimia ramosa*: potential for sugarcane bagasse saccharification. Ind Crop Prod. 122, 49–56.
Hess, M., 2008. Thermoacidophilic proteins for biofuel production. Trends Microbiol. 16 (9), 414–419.
Inman-Bamber, N.G., Smith, D.M., 2005. Water relations in sugarcane and response to water deficits. Field Crops Res. 92 (2-3), 185–202.
Jamaldheen, S.B., Sharma, K., Rani, A., Moholkar, V.S., Goyal, A., 2018. Comparative analysis of pretreatment methods on sorghum (*Sorghum durra*) stalk agrowaste for holocellulose content. Prep. Biochem. Biotechnol. 48 (6), 457–464.
Jia, L., Gonçalves, G.A., Takasugi, Y., Mori, Y., Noda, S., Tanaka, T., Ichinose, H., Kamiya, N., 2015. Effect of pretreatment methods on the synergism of cellulase and xylanase during the hydrolysis of bagasse. Bioresour. Technol. 185, 158–164.
Kim, I.J., Ko, H.J., Kim, T.W., Nam, K.H., Choi, I.G., Kim, K.H., 2013. Binding characteristics of a bacterial expansin (*Bs*EXLX1) for various types of pretreated lignocellulose. Appl. Microbiol. Biotechnol. 97 (12), 5381–5388.
Kim, I.J., Jung, J.Y., Lee, H.J., Park, H.S., Jung, Y.H., Park, K., Kim, K.H., 2015. Customized optimization of cellulase mixtures for differently pretreated rice straw. Bioproc. Biosyst. Eng. 38 (5), 929–937.
Kim, J.S., Lee, Y.Y., Kim, T.H., 2016. A review on alkaline pretreatment technology for bioconversion of lignocellulosic biomass. Bioresour. Technol. 199, 42–48.
Kim, I.J., Lee, H.J., Kim, K.H., 2017. Pure enzyme cocktails tailored for the saccharification of sugarcane bagasse pretreated by using different methods. Process Biochem. 57, 167–174.
Ko, J.K., Ximenes, E., Kim, Y., Ladisch, M.R., 2015. Adsorption of enzyme onto lignins of liquid hot water pretreated hardwoods. Biotechnol. Bioeng. 112 (3), 447–456.
Koivula, A., Kinnari, T., Harjunpaa, V., Ruohonen, L., Teleman, A., Drakenberg, T., Teeri, T.T., 1998. Tryptophan 272: an essential determinant of crystalline cellulose degradation by *Trichoderma reesei* cellobiohydrolase Cel6A. FEBS Lett. 429 (3), 341–346.
Kumar, A., Singh, S., Rajulapati, V., Goyal, A., 2020. Evaluation of pre-treatment methods for *Lantana camara* stem for enhanced enzymatic saccharification. 3 Biotech 10 (2), 1–11.
Ladeira Ázar, R.I., Bordignon-Junior, S.E., Laufer, C., Specht, J., Ferrier, D., Kim, D., 2020. Effect of lignin content on cellulolytic saccharification of liquid hot water pretreated sugarcane bagasse. Molecules 25 (3), 623.
Levine, S.E., Fox, J.M., Clark, D.S., Blanch, H.W., 2011. A mechanistic model for rational design of optimal cellulase mixtures. Biotechnol. Bioeng. 108 (11), 2561–2570.
Li, L., Zhou, W., Wu, H., Yu, Y., Liu, F., Zhu, D., 2014. Relationship between crystallinity index and enzymatic hydrolysis performance of celluloses separated from aquatic and terrestrial plant materials. BioResources 9 (3), 3993–4005.
Limayem, A., Ricke, S.C., 2012. Lignocellulosic biomass for bioethanol production: current perspectives, potential issues and future prospects. Prog. Energy Combust. Sci. 38 (4), 449–467.
Lowry, O.H., Rosebrough, N.J., Farr, A.L., Randall, R.J., 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. 193 (1), 265–275.
Mandels, M., Weber, J., 1969. The production of cellulases. Adv. Chem. 95, 391–414.
Marcos, M., García-Cubero, M.T., González-Benito, G., Coca, M., Bolado, S., Lucas, S., 2013. Optimization of the enzymatic hydrolysis conditions of steam-exploded wheat straw for maximum glucose and xylose recovery. J. Chem. Technol. Biotechnol. 88, 237–246.
McIntosh, S., Vancov, T., 2010. Enhanced enzyme saccharification of *Sorghum bicolor* straw using dilute alkali pretreatment. Bioresour. Technol. 101 (17), 6718–6727.
Nath, P., Dhillon, A., Kumar, K., Sharma, K., Jamaldheen, S.B., Moholkar, V.S., Goyal, A., 2019. Development of bi-functional chimeric enzyme (*Ct*GH1-L1-*Ct*GH5-F194A) from endoglucanase (*Ct*GH5) mutant F194A and β-1, 4-glucosidase (*Ct*GH1) from *Clostridium thermocellum* with enhanced activity and structural integrity. Bioresour. Technol. 282, 494–501.
Nedumaran, M., Singh, S., Jamaldheen, S.B., Nath, P., Moholkar, V.S., Goyal, A., 2020. Assessment of combination of pretreatment of Sorghum durra stalk and production of chimeric enzyme (β-glucosidase and endo β-1,4 glucanase, CtGH1-L1-CtGH5-F 194A) and cellobiohydrolase (CtCBH5A) for saccharification to produce bioethanol. Prep. Biochem. Biotech. https://doi.org/10.1080/10826068.2020.1762214.
Nelson, N., 1944. A photometric adaptation of the Somogyi method for the determination of glucose. J. Biol. Chem. 153 (2), 375–380.
NRAA, 2012. Prioritization of Rainfed Areas in India. Study Report 4. NRAA, New Delhi, India, pp. 1–100.
Ostovareh, S., Karimi, K., Zamani, A., 2015. Efficient conversion of sweet sorghum stalks to biogas and ethanol using organosolv pretreatment. Ind. Crop Prod. 66, 170–177.
Raabo, B.E., Terkildsen, T.C., 1960. On the enzymatic determination of blood glucose. Scand. J. Clin. Lab. Invest. Suppl. 12 (4), 402–407.
Rosgaard, L., Pedersen, S., Langston, J., Akerhielm, D., Cherry, J.R., Meyer, A.S., 2007. Evaluation of minimal *Trichoderma reesei* cellulase mixtures on differently pretreated barley straw substrates. Biotechnol. Prog. 23, 1270–1276.
Sharma, K., Thakur, A., Kumar, R., Goyal, A., 2019. Structure and biochemical characterization of glucose tolerant β-1, 4 glucosidase (*Ht*Bgl) of family 1 glycoside hydrolase from *Hungateiclostridium thermocellum*. Carbohydr. Res. 483, 107750.
Singh, S., Khanna, S., Moholkar, V.S., Goyal, A., 2014. Screening and optimization of pretreatments for *Parthenium hysterophorus* as feedstock for alcoholic biofuels. Appl. Energy 129, 195–206.
Singh, J., Suhag, M., Dhaka, A., 2015. Augmented digestion of lignocellulose by steam explosion, acid and alkaline pretreatment methods: a review. Carbohydr. Polym. 117, 624–631.
Singh, S., Dhillon, A., Goyal, A., 2020a. Enhanced catalytic efficiency of *Bacillus amyloliquefaciens* SS35 endoglucanase by ultraviolet directed evolution and mutation analysis. Renew. Energy 151, 1124–1133.
Singh, S., Kumar, K., Nath, P., Goyal, A., 2020b. Role of glycine 256 residue in improving the catalytic efficiency of mutant endoglucanase of family 5 glycoside hydrolase from Bacillus amyloliquefaciens SS35. Biotechnol. Bioeng. https://doi.org/10.1002/bit.27448.
Somogyi, M.A., 1945. New reagent for the determination of sugars. J. Biol. Chem. 160, 61–68.
Sun, S.L., Wen, J.L., Ma, M.G., Sun, R.C., 2014. Enhanced enzymatic digestibility of bamboo by a combined system of multiple steam explosion and alkaline treatments. Appl. Energy 136, 519–526.
Suwannarangsee, S., Bunterngsook, B., Arnthong, J., Paemanee, A., Thamchaipenet, A., Eurwilaichitr, L., Laosiripojana, N., Champreda, V., 2012. Optimisation of synergistic biomass-degrading enzyme systems for efficient rice straw hydrolysis using an experimental mixture design. Bioresour. Technol. 119, 252–261.
TAPPI, 1992. Technical Association of Pulp and Paper Industry. Atlanta, GA. .
Xu, F., Wang, J., Dong, M., Wang, S., Xiao, G., Li, Q., Chen, J., Li, W., Hu, W., Liu, J., 2019. Enhancing enzymatic hydrolysis yield of sweet sorghum straw polysaccharides by heavy ion beams irradiation pretreatment. Carbohydr. Polym. 222, 114976.
Zalkuwi, J., Singh, R., Bhattarai, M., Singh, O.P., Dayakar, B., 2014. Profitability analysis of Sorghum production in India. Int. J. Commerce Bus. Manage. 3, 707–714.
Zhou, J., Wang, Y.H., Chu, J., Luo, L.Z., Zhuang, Y.P., Zhang, S.L., 2009. Optimization of cellulase mixture for efficient hydrolysis of steam-exploded corn stover by statistically designed experiments. Bioresour. Technol. 100 (2), 819–825.

TH-2393_146106034