Synthetic Biology and Biocomputational Approaches for Improving Microbial Endoglucanases toward Their Innovative Applications

Mandeep, Hao Liu, and Pratyooosh Shukla*

1. INTRODUCTION

Lignocellulosic feedstocks are mainly composed of cellulose, and therefore, hydrolytic complexes are required for their bioconversion processes primarily composed of cellulases. Cellulases are the third largest set of enzymes that are used for industrial purposes. Cellulases include endoglucanases, β-glucosidase, and exoglucanases. Enzymatic hydrolysis along with ionic liquids (ILs) based on imidazolium is a superior technique for pretreatment of lignocellulosic materials. Cellulose amorphization is one of the specialties of IL pretreatment. For that reason, enzymes showing hydrolytic action to the amorphous cellulose, mainly endoglucanases (EGs), become important and significant enzymes in the hydrolytic complex. Among these, endoglucanase is the most promising cellulase for lignocellulosic degradation as it starts cellulose hydrolysis and is used in various industries. Endoglucanases (3.2.1.4) are related to the wider group of enzymes called glycosyl hydrolases (GHs). According to the CAZy enzyme database (http://www.cazy.org), they are part of 14 distinct GH families that are present in diverse organisms such as bacteria, archaea, fungi, and eukaryotes. Fungal endoglucanases have been categorized into 8 GH families together with other enzymes. These 8 GH families are 5, 6, 7, 9, 12, 45, 48, and 74. Fascinatingly, GH45 endoglucanases are normally differentiated by an inverting stereochemical mechanism and a low molecular weight. They show superior activity on amorphous cellulose. Cellulosic substrates are generally converted to cello-oligosaccharides as end products without any release of glucose as known for endoglucanases of other families. They also play the most important role predominantly in the initial liquefaction phase during biomass conversion. Endoglucanase is formed by diverse filamentous fungi such as Fusarium, Aspergillus, Penicillium, and Trichoderma genera. Among them, the Trichoderma genus is broadly utilized to produce cellulase, and the genus Aspergillus has acquired more interest due to its enormous capability to exude cellulases.

The industrial application of endoglucanase requires a change in their chemical behavior to attain superior productivity. Generally, enzymes attain their chemical specificity due to the presence of different amino acids. The chemical nature of amino acids present in enzymes decides the reactions carried out by them because enzymes have evolved from these basic building blocks. Enzymes execute a huge range of highly complex chemical reactions found in nature. They do this in different physiological states with extraordinary yields and wonderful stereoselectivity and regioselectivity. The high yield of organic

Received: November 25, 2020
Accepted: February 3, 2021
Published: February 26, 2021
Table 1. Endoglucanases Obtained from Different Micro-organisms Using Synthetic Biology Tools with Enhanced Properties for Industrial Application

| Sr. no. | Enzyme                     | Microorganism       | Outcome of the study                          | Reason for altered behavior                        | Ref   |
|---------|----------------------------|---------------------|-----------------------------------------------|---------------------------------------------------|-------|
| 1       | Endoglucanase II (Cel5A)   | Trichoderma reeseri | Thermal stability of enzyme at 80 °C           | Change of cysteine99 with valine and cysteine323 with histidine | 10    |
| 2       | Glycoside hydroxylase family S | Tricholoma matsutake | Degradation of a variety of substrates such as barley β-glucan, lichenan, and CMC-Na | High molecular weight due to the presence of more amino acids | 7     |
| 3       | Acidic and thermostable endoglucanase | Aspergillus niger | Better thermostability and higher melting temperature in ethanol solution | Combination with cellobiohydrolase | 12    |
| 4       | Endoglucanase egGH45       | Phialaphora sp. G5  | Thermostable enzyme at 60 °C                  | Use of rice products as a substrate                | 13    |
| 5       | Endoglucanase              | Rhizopus oryzae CCT 7560 | Thermostable enzyme at 20–100 °C              | Increased stable duration of the enzyme for 120 min at 60 °C | 14    |
| 6       | Endoglucanase GH12 family  | Aspergillus terreus | Increased stable duration of the enzyme for 120 min at 60 °C | Serine, threonine, hydroxylysine, and asparagine amino acids for substrate binding and glycosylation | 15    |
| 7       | Thermostable endoglucanase (CTendo45) GH45 family | Chaetomium thermophilum | Both cellulose and pectin hydrolysis by enzyme | | |

chemical processes is continuously increasing to meet the requirements of the masses. This requires the enzymes to perform the same chemical changes under very harsh conditions of industries. Apart from the increased rate of substrate degradation and synthesis in industries, enzymes are also accountable for the uptake, amalgamation, and breakdown of chemicals (e.g., pesticides) in our bodies. In humans, cellular reactions are controlled by thousands of enzymes. Enzymes represent 36% of all drug targets, i.e., 4906 out of 13382 of total drug targets listed in ChEMBL (ChEMBLDB, http://www.ebi.ac.uk/chembl, accessed 16 November 2020). The catalytic site is the enzyme’s workshop where the catalytic reaction occurs. This catalytic site contains a relatively small number of amino acids that are utilized in substrate binding (and/or cofactor). This site can be explored with the help of computational tools to harness the binding efficiency for enhanced enzyme action. The amino acids present in the catalytic site may be altered with the help of bioengineering techniques to increase the surface area for substrate attachment. Even smaller subsets of the amino acid present in the catalytic site are critical to the catalytic function of enzymes. Protein surface engineering has been utilized to improve the thermostability and activity of endoglucanase II from Penicillium verruculosum. The engineered enzyme was able to release 10–20% more reducing sugars from aspen wood. The experimentation of computational methods can be applied in vitro to produce better productive enzymes. Akcapinar et al. used molecular dynamics and molecular mechanics to design endoglucanase I from Trichoderma reeseri. They developed the model computationally by inserting glycine and lysine amino acid loops. The predicted β-loop mutation was inserted by site-directed mutagenesis. Thus, the use of computational tools and synthetic biology goes hand in hand for enzyme improvement. This review highlights the use of endoglucanases in industries and the techniques utilized to engineer or modify different enzymes for superior industrial utilization.

2. INDUSTRIAL APPLICATION OF ENDOGLUCANASES

Endoglucanases are a very important set of enzymes for industrial uses. They are used for the degradation of lignocellulosic complexes and plant biomass. Li et al. in 2019 used a consortium of endoglucanase and exoglucanase for the degradation of lignocellulosic material. They fused endoglucanase and exoglucanase to prepare a fusion complex. They found that the fused protein was able to hydrolyze the natural substrates more effectively with a 2–14-fold increase in specific activities from the parent enzymes. The lignocellulosic substrate can also be used in biorefineries for the production of bioethanol. The enzyme saccharifies the biomass to monosugars (glucose). These sugars can be fermented to ethanol and envisioned as a source of future fuel and biobased chemicals. Xue et al. in 2018 produced ethanol-tolerant endoglucanase from isolated Aspergillus niger. The acidic, thermostable, and ethanol-tolerant endoglucanase was able to carry out saccharification and fermentation simultaneously. Thus, it is helpful for the production of bioethanol in the industry. It is a secure replacement for conservative fuels with the additional profit of being readily accessible. It would considerably lessen the environmental toxic waste and the health dangers coupled with the blazing of conventional fuels. Endoglucanases are also used in the food and feed industry for the deconstruction of glucans. They are thus useful for the production of prebiotics and functional foods. The endoglucanase polypeptides have shown enhanced performance in textile industries, particularly in biostoning and biofinishing, in detergent functions, and in fabric protection and color preservation, mainly in deterrence and exclusion of fuzz and pills, in color preservation and restoration. Endoglucanases are widely used in the pulp and paper industry for the production of cellulose nanocrystals and fermentable sugars. Yang et al. in 2020 used four different endoglucanases for the production of cellulose nanocrystals. They found that Cel7B (endoglucanase belonging to GH7 family) and Cel5B (endoglucanase belonging to GH5 family) when used in optimum conditions gave a 13.16% yield of cellulose nanocrystals. Attributes of cellulose nanocrystals can be improved with the use of engineered enzymes. Endoglucanase is also used in the bioremediation of toxic effluents from the pulp and paper industry. Table 1 gives a list of the endoglucanases obtained from different microorganisms showing enhanced characteristics for industrial application. The endoglucanases used in their native form are not suitable for industrial uses due to their chemical instability and activity loss due to denaturation in harsh conditions. So to provide stability and get better results from enzyme hydrolysis, synthetic biology tools and engineering techniques are utilized.
3. USE OF SYNTHETIC BIOLOGY TOOLS AND ENZYME IMMOBILIZATION

Synthetic biology is a newly emerged field of science that has taken over the conventional practices of engineering. With the advent of synthetic biology, chemical alterations in enzymes have made it possible to carry out desired functions. The chemistry of enzymes has been changed due to alterations in their structure and functional groups. Many researchers have tried to enhance the production of enzymes from microorganisms by increasing their stability and activity with the help of synthetic biology and other artificial intelligence tools.

In an interesting study, Akbarzadeh and colleagues identified the differences between the sequence similarities of efficient thermostable endoglucanase from *Thermoascus aurantiacus* and high catalytically active endoglucanase of *Trichoderma reesei*. They found the presence of an extra disulfide bond in *Trichoderma reesei*. They eliminated the disulfide bonds with the help of site-directed mutagenesis. Their work demonstrated that the enzyme activity and thermal stability of the recombinant enzyme were increased approximately 2.4-fold in both cases.

Similarly, Nath and colleagues in 2019 developed a chimeric enzyme after site-directed mutagenesis of endoglucanase. They converted phenylalanine194 (Phe194) to alanine. The chimera showed enzyme activity of both fused endoglucanase and glucosidase. It was thermally stable and structurally integral. In another experiment, Chahed et al. used recombinant plasmid pPICZαA to transform endoglucanase. The recombinant endoglucanase showed better thermal stability, a wide pH range, and different substrate disintegration. This was due to the glycosylation of the recombinant enzyme. The glycosylation process protects the protein from degradation and hence increased its thermal stability and activity.

Dotsenko et al. carried out protein surface engineering to increase enzyme stability. They carried out this through rational design. An ionic liquid, 1-butyl-3-methylimidazolium chloride ([Bmim]Cl), is responsible for endoglucanase II stability in the bioconversion process. Structure analysis with simulations helped in the identification of enzyme pockets for surface modification and enzyme engineering. The redesigned enzyme thus obtained was able to hydrolyze aspen wood effectively. Similarly, endoglucanase modification with the substitution of proline in the amino acid chain increased the thermostability of the enzyme. This occurred due to an increase in hydrogen bonding and changes in the flexibility of the enzyme.

Sasaki et al. used clustered regularly interspaced short palindromic repeat (CRISPR)-δ integration as well as multiple promoters shuffling for endoglucanase engineering. They found that the simultaneous use of these two techniques was able to increase the activity of the enzyme by 17.3-fold compared to that with the conventional recombination technique.
enzyme engineering techniques have paved the way for the production of efficient enzymes for industrial purposes. Enzyme immobilization has proven economically and industrially important in recent times. Carli and co-workers immobilized two recombinant enzymes (glucosidase and endoglucanase) on ferromagnetic nanoparticles. The synergistic action of enzyme immobilization proved better for substrate hydrolysis and production of fermentable sugars. This was due to the catalytic accessibility of the enzyme to the substrate. Nanoparticles provided better disintegration and stable contact with the substrate, thus providing faster action. Figure 1 depicts the advantages of using combined enzyme technology and nanotechnology over the independent techniques. Apart from synthetic biology tools, protein can be modified by engineering techniques and computational tools for better functioning, as discussed further below.

3.1. Protein Bioengineering for Enzymes. Enzyme fabrication from micro-organisms is a better technique to attain product sustainability. Energy is intended for the fabrication of qualitatively superior enzymes. Protein bioengineering is rooted in considering the folding and existence of substrate binding sites of proteins known as catalytic sites. Biologically, proteins are very important as they serve various complex life-sustaining purposes and have evolved enormously for making our survival possible in a changing environment. Among the abundant proteins that subsist, structural proteins hold meticulous attention for designing nanomaterials for bioengineering. Protein evolution and functions can be enhanced with the help of protein design by studying the naturally occurring proteins and their structures. Mammalian and insect-derived proteins have been studied using these approaches for drug delivery and tissue engineering. Plant-based designed proteins (PBPs) composed of simple polypeptides possess the host protein properties, e.g., elasticity similar to elastin and strength similar to silk. The entirely synthetic PBPs are of enormous importance as they permit the integration of functionalities anomalous to their natural corresponding protein, e.g., silk or hydroxyapatite (HAP)-binding elastin and cell-adhesive resiliency.

The hasty progress in genome sequencing has widely opened a myriad of designs of proteins that have developed naturally for a variety of functions. The customization of these naturally occurring proteins for appropriate industrial applications remains a challenge. Complementary interfaces are formed in natural protein complexes which enhance the catalytic efficiency through substrate binding and increased enzyme availability at the local site. Researchers have tried to bring the catalytic site in close proximity because the formation of interfaces is very difficult. Various techniques have been used to imitate protein complexes such as matrix encapsulation, arbitrary chemical conjugation, and physisorption. However, denaturation and disruption of protein function possess a limitation to these methods. Fusion protein preparation is the simplest way to generate protein complex with specific composition and order. The utilization of biological scaffolds is a striking substitute for large complex formation in which smaller fusion proteins are fused by adding definite binding domains to every protein. Chemical conjugation to protein surfaces faces problems such as the presence of a small number of highly reactive side groups, naturally limited to lysines and cysteine surfaces. Increasing the existing conjugation reactions with the help of unnatural amino acids (UAA) is a striking alternative. Though protein engineering is a useful method to increase the effectiveness of enzymes, it is tedious to generate such enzymes directly in vivo. Computational tools have provided the next step to this synthesis by modulating the energy of enzymes and microorganisms to get better output from designed enzymes.

3.2. Biocomputational Tools for Enzyme Improvement. Characterization of enzyme topology and constituent amino acids has provided new insights into the development of efficient enzymes. Computational tools help in designing enzymes with reduced misfolding of residues and uncovering of catalytic sites for better substrate binding. Researchers have observed that the free spaces and catalytic domains of enzymes accommodate sugars for their bioconversion. Thus, the biological activity of enzymes can be increased with the use of biocomputational tools. Biocomputational tools also help in understanding the complexity of enzymes. The MACIE (mechanism, annotation, and classification in enzymes) tool can provide the complex chemical mechanism carried out by enzymes. It unravels both the structural and functional attributes of enzymes. MACIE provides information about the number of steps involved in a chemical reaction, the catalytic site of enzymes, an amino acid present in the enzyme, cofactors available for enzymatic action, and chemical changes occurring during the reaction. Apart from MACIE, several other tools are important for understanding the enzyme mechanism. SPLD (structure function linkage database), InterPro signature, and EzCatDb (database of enzyme catalytic mechanisms) are the other tools for the bulk prediction of enzyme mechanism. De Ferrari and Mitchell performed a machine learning experiment with MACIE and InterPro signatures and inferred a conclusion that InterPro signatures are very significant for correct enzyme mechanism prediction as they are computational representations of evolutionarily conserved sequence patterns. Nath et al. used the Psi-Pred version 3.3 web server and RaptorX computational web server tools to determine the structural integrity of a chimera produced by the fusion of endoglucanase and glucosidase. They analyzed loop flexibility and bond energy to determine the stability of the enzyme complex. Molecular dynamics simulation (MDS) is a helpful tool to find the binding affinity of the enzyme to substrates or ligands and the stability of the enzymes. Compactness and stability of the modeled chimera were studied using MDS. There was a variation in the initial chimera and the chimera obtained from MDS. There was a difference in loop regions and root mean square deviation of 1.244 Å. The N-terminal and C-terminal residues showed the highest flexibility, whereas the amino acid in the core regions showed lesser flexibility owing to their stability. This showed the stability of the core region for enzyme activity and stability. The interaction of catalytic cleft with cello-oligosaccharides was studied using molecular docking analysis. Both the catalytic sites can accommodate cello-oligosaccharides up to 7 degrees of polymerization. Hydrophobic interaction and hydrogen bonding were mainly observed in the complexes. The MDS analyses of contact with ligands of the chimera and each enzyme demonstrated that the endoglucanase module of the chimera bound stably to cellohexaose and released efficiently the hydrolyzed product (cellobiose). It was followed by the accomplishment of the combined β-glucosidase module of the chimera with altered conformation owing to its enhanced catalytic effectiveness.

Researchers have analyzed the structural and functional characteristics of endoglucanase obtained from Neospora cressa OR74A. The molecular dynamics simulation of this enzyme revealed the role of a loop 6 in the substrate binding,
which was earlier not known. Thus, computational tools help in revealing the new functionalities and structural insights of enzyme and substrate interactions. Identification of binding sites with the help of computational tools along with homology modeling and molecular dynamic simulation provides a superior method for the production of industrially valuable enzymes. Molecular dynamics simulation and homology modeling provide superior enzyme confirmation over the molecular analysis of enzymes. The analysis of enzymes and proteins by molecular biology is time- and energy-consuming compared to the computational analysis. Development of different conformations of the enzymes by the molecular approach is more difficult than the homology-based models which can predict the minimized energy structures for enzyme model generation. It is noticed that the formation of hydrogen bonds helps in the efficient binding of the substrate and enzyme. Cellotetraose substrate forms a hydrogen bond with cellulase enzyme and binds proficiently. Puhl and colleagues carried out the structural

Figure 2. Techniques employed for enzymes and micro-organisms engineering for enhanced enzyme production and to study enzyme functioning. xln, cbh, axh, and egl refers to xylanase cellobiohydrolase, α-L-arabinofuranosidase, and endoglucanase in STRING network of enzymes involved in carbohydrate metabolism of *Aspergillus niger*. A–C refers to endoglucanase ribbon and ball-and-stick structures obtained by SWISS-MODEL. D,E refers to surface topology studies and ligand interaction, respectively, studied by SWISS-MODEL. KEGG pathway highlights (dark color) the carbohydrate metabolism and carbon utilization pathway of *Aspergillus niger*. 
and computational analysis for finding the hydrolysis pattern of endoglucanases GH5 (XccCel5A) obtained from Xanthomonas campestris pv campestris. This is done by virtual hydrolysis using computational tools to find out the action of the endoglucanase enzyme. The virtual hydrolysis was carried out by a newly designed un-named computational tool written in code Fortran90. The virtual hydrolysis was able to give the result in agreement with the capillary zone electrophoresis. Thus, virtually the hydrolysis products of the enzymatic reaction can be observed, and useful products can be obtained by the desired hydrolysis pattern. It is found with the help of this experiment that four binding subsites are present in endoglucanases for the substrate position.

Apart from the substrate utilization, de novo design of the enzyme with the help of computational tools has also been practiced by researchers to develop efficient enzymes and to decipher the enzyme catalysis. Hong and co-workers in 2018 studied the multiple active site evolution in designed enzymes with the help of computational tools and laboratory evolution. They used Kemp eliminase for their study and found that computational tools can help in identifying mutations that can alter enzyme catalysis and stability. They found that the mutations Ile7Asp, Lys146Glu, Gly202Arg, and Asn224Asp considerably progressed the catalytic ability of enzymes, mainly by enthalpic effects. The reason for increased activity might be the increased basicity of the catalytic group or enhanced TS stabilization. They also noticed that point mutations far from the active site are also responsible for the catalytic efficiency of the enzyme. This may be due to the stability of the active site due to backbone adjustment and cavity filling by the mutated amino acids. Their study established that the outer shell amino acids are also responsible for providing catalytic power to enzymes. Apart from mutational studies, computational tools can also be utilized for deciphering the catalytic reactions. Oanca et al. carried out a study on the catalytic mechanism of cysteine protease. This enzyme is grouped like endoglucanases in several families. Using computational analysis they found that the decacylation step might have a major role in the rate-limiting step, which is supported by other enzymatic studies. The catalytic mechanism of endoglucanases can also be understood properly using such computational tools. Thus, with the help of computational analysis, conformation of substrate binding with an enzyme can be calculated and improved with the help of enzyme engineering techniques. Another interesting tool to alter the micro-organism for increased flux toward the desired product is metabolic pathway engineering, as discussed below.

4. METABOLIC PATHWAY ENGINEERING

Production of enzymes from micro-organisms and then their extraction needs special handling. To reduce the energy of the micro-organisms and increase the effectiveness to get the final product, metabolic pathway engineering has come up as an extremely valuable field. By taking uptake and utilization of carbon in the metabolic reaction, various metabolic pathways have been designed for the production of biofuels, useful polymer acid materials, food additives, and many other chemicals. PATH\textsuperscript{40}, OptKnock, OptReg, and OptStrain are a few of the useful computational tools for metabolic engineering. Development of computational tools to design micro-organisms for the production of desired products is cumbersome. The desired products may be utilized in other competing reactions, and hence, its yield is decreased. Thus, the design of computational tools must take into account these reactions and be required to recognize the most suitable chemical reactions to increase the desired product yield. Motwali et al. in 2020 developed the PATH\textsuperscript{40} tool to search heterologous biosynthetic routes. They constructed a universal metabolic network and predicted the pathway score to predict the most favorable biosynthetic pathway. This tool is able to predict pathways among different organisms (bacteria, fungi, cyanobacteria). Retrieval of the biosynthetic pathway of isoprene and cocaine provided the ability of this tool in a confident retrieval mechanism.

Metabolic mechanisms occurring in organisms can be understood with the help of computational tools developed from genome-scale metabolic models. Interesting metabolites can be overproduced with the help of optimal knockout strategies developed with the help of flux balance analysis. OptKnock maximizes biochemical production by maximizing the production of biomass in the restrictions obligatory by gene knockouts. It follows bilevel optimization. The actual flux can be minimized further by understanding the closeness and robustness of the levels involved. Successful manipulation of metabolism is reported in many organisms. For example, in \textit{Saccharomyces cerevisiae}, the yield of isobutanol was fruitfully enhanced following the intracellular metabolism manipulation by investigating the cytosolic and mitochondrial sections. The isobutanol produced by yeast should be dehydrated and converted to isobutene before using as a jet fuel. Researchers are trying to manipulate the metabolic pathway of microbes to directly produce isobutene. After building artificial pathways, the investigation and assessment of heterologous enzymes are required to progress the fabrication of the desired product. This requires the study of the kinetic parameters of enzymes. Similar manipulation can be predicted in other fungi with the help of computational tools to overproduce endoglucanases and for increased production of bioethanol from waste biomass.

Figure 2 shows various techniques used to study the enzymatic characteristics using SWISS-MODEL (https://swissmodel.expasy.org), STRING (https://string-db.org) and KEGG pathway (https://www.genome.jp/kegg/pathway.html) tools and their improvement thereof by metabolic engineering and synthetic biology.

System biology can provide all possible pathways that can be rationally engineered to increase protein production and decrease byproduct formation. An in silico system biology approach can be used to carry out pathway editing. Researchers have used metabolic pathway engineered \textit{Saccharomyces cerevisiae}, which resulted in considerable augmentation in the production of fatty acids, particularly monounsaturated fatty acids. In the engineered $\Delta idh1/2$ strain having heterologous ATP-citrate lyase, there was a 92 and 77% increase in 9-hexadecanoic acid (C16:1) and 9-octadecenoic acid (C18:1), respectively. The augmented fabrication of monounsaturated fatty acids is advantageous for the cold-flow property of biodiesel fuels. Similarly other groups of scientists have studied the metabolic pathway engineering for the production of flavonoids from microbes. They found that the use of micro-organisms is superior over the plant flavonoids due to easy extraction and stable production. Like fatty acids and flavonoids, the production of enzymes can also be enhanced by studying the metabolic pathways in micro-organisms. BNICE, DESHARKY, RetroPath, and GEM-Path are a few of the tools to design novel metabolic pathways. BNICE categorizes probable enzymes to catalyze new reactions founded on substrate resemblance and molecular simulations, whereas DESHARKY includes the augmentation.
probability of the cell with contemplation to cellular cargo for production of the enzyme. RetroPath utilizes molecular signatures to mechanically sieve the appropriate biochemical reaction operators to those with comparable substrates or products. GEM-Path is utilized for recognition of those products which could be prepared using restricted steps from native metabolites. The host metabolism is modified after recognition of a suitable pathway. Modification may involve deletion and up/downregulation of products and intermediates.34

The problems faced in metabolic pathway engineering arise due to a lack of understanding in the thorough expression condition, the enzyme degradation, the location of enzyme and substrate, and the kinetic property of every expressed pathway enzyme in the host cell. This lack of knowledge hinders pathway engineering at the enzymatic level for enhanced function and a superior manufacturing yield. An increase in understanding the enzymatic reaction in the cellular environment will lead to the use of enzyme-based pathway engineering for industrial purposes.35 A thorough understanding of enzyme kinetics and reactions involved in a cellular environment is required to accommodate the metabolic engineering pathway in the desired microorganism. Thus, with the help of metabolic pathway engineering tools, we can use micro-organisms as a biocatalyst for the production of useful chemicals.

5. CONCLUSION

Enzymes are useful in every walk of life. With the increase of sustainable biological methods for industrial purposes, there is an increase in the enzyme-mediated process. Endoglucanases have widely been used for cellulose biomass degradation. Biologically synthesized enzymes are used widely to reduce any chemical exposure to the environment. The stability of enzymes can be increased with the help of suitable synthetic biology tools and protein engineering methods. Computational tools have been widely developed for designing a pathway based on flux balance analysis. These computational tools focus on the substrate interaction with the enzymes and their stability during the reaction. However, metabolic pathway engineering, which is an emerging field, can prove to be a stepping stone for the modification of enzymes and micro-organisms for industrial use. Many tools have been developed to design metabolic pathways. By studying these pathways and strain development method, enhanced production can be achieved. The development of such pathways for industrially important enzymes is still limited. Scientists need to apply such an algorithm and machine-learning-based tools on commercially important strains and enzymes to make their utilization in an improved manner. Though due to the limitations of understanding cellular behavior, the use of metabolic pathway engineering is less now, but in the coming time it will be the most crucial field with the addition of more data in the databases. This will be augmented by biocomputational analysis and synthetic biology tools. Biocomputational tools will make the use of synthetic biology more precise based on localized modifications.

■ AUTHOR INFORMATION

Corresponding Author

Pratyooosh Shukla — Enzyme Technology and Protein Bioinformatics Laboratory, Department of Microbiology, Maharshi Dayanand University, Rohtak 124001, Haryana, India; School of Biotechnology, Institute of Science, Banaras Hindu University, Varanasi 221005, India; orcid.org/0000-0002-9307-4126; Email: pratyoosh.shukla@gmail.com, shuklap@bhu.ac.in

Authors

Mandeep — Enzyme Technology and Protein Bioinformatics Laboratory, Department of Microbiology, Maharshi Dayanand University, Rohtak 124001, Haryana, India
Hao Liu — State Key Laboratory of Pulp and Paper Engineering, South China University of Technology, Guangzhou 510640, China

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.0c05744

Author Contributions
*Mandeep and H.L. are joint first authors.

Notes

The authors declare no competing financial interest.

Biographies

Mandeep graduated in Botany (H) from University of Delhi, New Delhi, India in 2016 and did his postgraduation in Botany in 2018 from University of Delhi, New Delhi, India. He is pursuing his doctoral degree from the Department of Microbiology, Maharshi Dayanand University, Rohtak, India. His research interests are in enzyme technology, systems biology, and protein bioinformatics.

Dr. Hao Liu is an Associate Professor in the State Key Laboratory of Pulp and Paper Engineering, South China University of Technology, Guangzhou, China. His research covers a range of enzymatic cellulose hydrolysis, enzyme-assisted pulp bleaching, and refining, direct fermentation of lignocellulose to ethanol, focusing on lignocellulolytic enzymes including cellulases, xylanases, and laccases. He obtained his Ph.D. in 2011 at South China University of Technology, China, and during 2008–2010, he was a visiting scholar in the Department of Biological Systems Engineering, University of Wisconsin at Madison, Wisconsin 53706.
USA, working on the mechanism of lignin—cellulase nonproductive binding. He also served as a DAAD short-term visiting scholar in the Enzyme Process Engineering Group, Department of Chemical Engineering, RWTH Aachen University, Germany in 2015. Dr. Liu has nearly 21-year research experience in the field of pulp and paper engineering and enzyme biotechnology.

Prof. Pratyush Shukla received his Doctor of Science (D.Sc.) in Microbiology from Barkatullah University, Bhopal, India, with specialization in computational enzyme engineering, and a Ph.D. in Microbiology from Faculty of Medicine, A.P.S. University, Rewa, India. He is presently serving as Professor in the School of Biotechnology, Institute of Science, Banaras Hindu University (BHU), Varanasi, India. Prior to joining BHU, he served as Professor and Head of Department of Microbiology at Maharsi Dayanand University, Rohtak, India. He served as Indo-USA Visiting Research Professor at the University of Cincinnati, College of Medicine, Cincinnati, Ohio, USA in 2014. He is also serving as a Guest Professor at South China University of Technology, Guangzhou, China. His research interests are in the areas of enzyme technology and protein bioinformatics and other interdisciplinary areas of biotechnology.

**ACKNOWLEDGMENTS**

P.S. acknowledges the infrastructural support from Department of Science and Technology, New Delhi, Govt. of India, through a FIST grant (Grant No. 1196 SR/FST/LS-I/2017/4), BTIS-NET - Sub-Distributed Information Centre, funded by DBT, Govt. of India at the School of Biotechnology, Banaras Hindu University, Varanasi, India and Department of Biotechnology, Government of India (Grant No. BT/PR27437/BCE/8/1433/ 2018). Mandeep acknowledges the Junior Research Fellowship from CSIR, India (Award No. 09/382(0211)/2019-EMR-I). H.L. and P.S. acknowledge State Key Laboratory of Pulp and Paper Engineering, South China University of Technology, Guangzhou, China for Guest Professorship to P.S. P.S. alsoacknowledges Ministry of Human Resource Development, Government of India for funding the Global Initiative of Academic Networks (GIAN) Course on ‘Current Challenges in Commercial Production of Cellulosic Ethanol’ [Course Code: 176021H04; Foreign Faculty: Dr. Hao Liu].

**REFERENCES**

(1) Dotsenko, A. S.; Rozhkov, A. M.; Zorov, I. N.; Sinitsyn, A. P. Protein surface engineering of endoglucanase *Penicillium verruculosum* for improvement in thermostability and stability in the presence of 1-butyl-3-methylimidazolium chloride ionic liquid. *Bioresour. Technol.* 2020, 296, 122370.

(2) Chahed, H.; Boumaiza, M.; Ezine, A.; Marzouki, M. N. Heterologous expression and biochemical characterization of a novel thermostable *Sclerotinia sclerotiorum* GH45 endoglucanase in *Pichia pastoris*. *Int. J. Biol. Macromol.* 2018, 106, 629–635.

(3) Holliday, G. L.; Fischer, J. D.; Mitchell, J. B.; Thornton, J. M. Characterizing the complexity of enzymes on the basis of their mechanisms and structures with a bio-computational analysis. *FEBS J.* 2011, 278 (20), 3835–3845.

(4) Bartlett, G. J.; Porter, C. T.; Borkakoti, N.; Thornton, J. M. Analysis of catalytic residues in enzyme active sites. *J. Mol. Biol.* 2002, 324 (1), 105–121.

(5) Akcapinar, G. B.; Gucl, O.; Sezerman, U. O. From silica in to vitro: modelling and production of *Trichoderma reesei* endoglucanase 1 and its mutant in *Pichia pastoris*. *J. Biotechnol.* 2012, 159 (1–2), 61–68.

(6) Li, X.; Xia, J.; Zhu, X.; Bilal, M.; Tan, Z.; Shi, H. Construction and characterization of bifunctional cellulases: *Caldicellulosi* ruptor-sourced endoglucanase, CBB, and exoglucanase for efficient degradation of lignocellulose. *Biochem. Eng. J.* 2019, 151, 107363.

(7) Xue, D. S.; Zeng, X.; Lin, D.; Yao, S. Ethanol tolerant endoglucanase from *Aspergillus niger* isolated from wine fermentation cellar. *Biotatal. Agric. Biotechnol.* 2018, 15, 19–24.

(8) Ontanon, O. M.; Ghio, S.; Marrero Diaz de Villegas, R.; Garrido, M. M.; Talia, P. M.; Feher, C.; Campos, E. A thermostable GH48 endoglucanase of *Enterobacter sp.* R1 is suitable for β-glucan deconstruction. *Food Chem.* 2019, 298, 124999.

(9) Yang, T.; Li, X.; Guo, Y.; Peng, S.; Liu, G.; Zhao, J. Effect of endoglucanases from different glycoside hydrolase families on enzymatic preparation of cellulose nanocrystal. *Ind. Crops Prod.* 2020, 155, 112755.

(10) Akbarzadeh, A.; Pourzardosht, N.; Dehnavi, E.; Ranaei Siadat, S. O.; Zamani, M. R.; Motallebi, M.; Nikzad Jamnani, F.; Aghaeeoor, M.; Barshan Tashnizi, M. Disulfide bonds elimination of endoglucanase II from *Trichoderma reesei* by site-directed mutagenesis to improve enzyme activity and thermal stability: An experimental and theoretical approach. *Int. J. Biol. Macromol.* 2018, 120, 1572–1580.

(11) Onuma, H.; Hara, K.; Sugita, K.; Kano, A.; Fukuta, Y.; Shirasaka, N. Purification and characterization of a glycoside hydrolase family 5 endoglucanase from *Tricholoma matsutake* grown on barley based solid-state medium. *J. Biosci. Bioeng.* 2019, 128 (6), 669–676.

(12) Zhao, J.; Shi, P.; Li, Z.; Yang, P.; Luo, H.; Bai, Y.; Wang, Y.; Yao, B. Two neutral thermostable cellulases from *Phialophora* sp. GS act synergistically in the hydrolysis of filter paper. *Bioresour. Technol.* 2012, 121, 404–410.

(13) Kupski, L.; Pagnussatt, F. A.; Buffon, J. G.; Furlong, E. B. Endoglucanase and total cellulase from newly isolated *Rhizopusoryzae* and *Trichoderma reesei*: production, characterization, and thermal stability. *Appl. Biochem. Biotechnol.* 2014, 172 (1), 458–468.

(14) Narra, M.; Dixit, G.; Divecha, J.; Kumar, K.; Madamwar, D.; Shah, A. R. Production, purification and characterization of a novel GH 12 family endoglucanase from *Aspergillus terreus* and its application in enzymatic degradation of delignified rice straw. *Int. Biodeterior. Biodegrad.* 2014, 88, 150–161.

(15) Zhou, Q.; Ji, P.; Zhang, J.; Li, X.; Han, C. Characterization of a novel thermostable GH45 endoglucanase from *Chaetomium thermophilum* and its biodegradation of pectin. *J. Biosci. Bioeng.* 2017, 124 (3), 271–276.

(16) Shukla, P. Synthetic Biology Perspectives of Microbial Enzymes and Their Innovative Applications. *Ind. J. Microbial.* 2019, 59, 401–409.

(17) Mandeep; Kumar Gupta, G.; Shukla, P. Insights into the resources generation from pulp and paper industry wastes: challenges, perspectives and innovations. *Bioresour. Technol.* 2020, 297, 122496.

(18) Nath, P.; Dhillon, A.; Kumar, K.; Sharma, K.; Jamaldheen, S. B.; Moholkar, V. S.; Goyal, A. Development of bi-functional chimeric enzyme (CtGH1-L-CtGH5-F194A) from endoglucanase (CtGH5) mutant F194A and β-1, 4-glucosidase (CtGH1) from *Clastridium thermocellum* with enhanced activity and structural integrity. *Bioresour. Technol.* 2019, 282, 494–501.

(19) Dotsenko, A. S.; Pramanik, S.; Gusakov, A. V.; Rozhkov, A. M.; Zorov, I. N.; Sinitsyn, A. P.; Davari, M. D.; Schwanenberg, U. Critical
effect of proline on thermostability of endoglucanase II from *Penicillium verruculosum*. Biochem. Eng. J. 2019, 152, 107395.

(20) Sasaki, Y.; Mitsui, R.; Yamada, R.; Ogino, H. Secretory overexpression of the endoglucanase by *Saccharomyces cerevisiae* via CRISPR-δ-integration and multiple promoter shuffling. Enzyme Microb. Technol. 2019, 121, 17–22.

(21) Carlí, S.; de Campos Carneiro, L. A. B.; Ward, R. J.; Meleiro, L. P. Immobilization of a β-glucosidase and an endoglucanase in ferromagnetic nanoparticles: A study of synergistic effects. Protein Expression Purif. 2019, 160, 28–35.

(22) Desai, M. S.; Lee, S. W. Protein-based functional nanomaterial design for bioengineering applications. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol*. 2015, 7 (1), 69–97.

(23) Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res.* 2018, 46 (W1), W363–W367.

(24) De Ferrari, L.; Mitchell, J. B. From sequence to enzyme mechanism using multi-label machine learning. *BMC Bioinf.* 2014, 15 (1), 150.

(25) Nath, P.; Sharma, K.; Kumar, K.; Goyal, A. Combined SAXS and computational approaches for structure determination and binding characteristics of Chimera (CtGH1-L1-CtGH5-F194A) generated by assembling β-glucosidase (CtGH1) and a mutant endoglucanase (CtGH5-F194A) from *Clostridium thermocellum*. *Int. J. Biol. Macromol.* 2020, 148, 364–377.

(26) Kadowaki, M. A. S.; Polikarpov, I. Structural insights into the hydrolysis pattern and molecular dynamics simulations of GH45 subfamily a endoglucanase from *Neurospora crassa* OR74A. *Biochimie* 2019, 165, 275–284.

(27) Puhl, A. C.; Prates, E. T.; Rosseto, F. R.; Manzine, L. R.; Stankovic, I.; de Araujo, S. S.; Alvarez, T. M.; Squina, F. M.; Skaf, M. S.; Polikarpov, I. Crystallographic structure and molecular dynamics simulations of the major endoglucanase from *Xanthomonas campestris* pv. campestris shed light on its oligosaccharide products release pattern. *Int. J. Biol. Macromol.* 2019, 136, 493–502.

(28) Hong, N.-S.; Petrovic, D.; Lee, R.; Grynová, G.; Purg, M.; Saunders, J.; Bauer, P.; Carr, P. D.; Lin, C.-Y.; Mabbutt, P. D.; Zhang, W.; Altamore, T.; Easton, C.; Coote, M. L.; Kamerlin, S. C. L.; Jackson, C. J. The evolution of multiple active site configurations in a designed enzyme. *Nat. Commun.* 2018, 9, 3900.

(29) Oanca, G.; Asadi, M.; Saha, A.; Ramachandran, B.; Warshel, A. Exploring the Catalytic Reaction of Cysteine Proteases. *J. Phys. Chem. B* 2020, 124, 11349.

(30) Motwalli, O.; Uludag, M.; Mijakovic, I.; Alazmi, M.; Bajic, V. B.; Gojborov, T.; Gao, X.; Essack, M. PATHcre8: A Tool That Facilitates the Searching for Heterologous Biosynthetic Routes. *ACS Synth. Biol.* 2020, 9, 3217.

(31) Apaydin, M.; Zeng, B.; Qian, X. A reliable alternative of OptKnock for desirable mutant microbial strains. In *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*; IEEE, 2016; pp 573–576.

(32) Dasgupta, A.; Chowdhury, N.; De, R. K. Metabolic pathway engineering: Perspectives and applications. *Comput. Meth. Prog. Bio.* 2020, 192, 105436.

(33) Hammer, S. K.; Avalos, J. L. Uncovering the role of branched-chain amino acid transaminases in *Saccharomyces cerevisiae* isobutanol biosynthesis. *Metab. Eng.* 2017, 44, 302–312.

(34) Long, M. R.; Ong, W. K.; Reed, J. L. Computational methods in metabolic engineering for strain design. *Curr. Opin. Biotechnol.* 2015, 34, 135–141.

(35) Zha, J.; Wu, X.; Gong, G.; Koffás, M. A. Pathway enzyme engineering for flavonoid production in recombinant microbes. *Metab. Eng. Comm.* 2019, 9, No. e00104.