The Development of Primers Specific to Bacterial Species that Produce Cellulase Enzymes Using the Tools of Bioinformatics

Steven R. Ohrmund,* and Susan Elrod
Environmental Biotechnology Institute (EBI), California Polytechnic State University, San Luis Obispo, California 93407

Bacteria play a major role in the maintenance of life on earth. Life on earth is a dynamic process where solar energy and nutrients from soil lead to the growth of plants, and energy from plants to the growth of animals. This process is cyclic. Bacteria help in cycling nutrients and energy to sustain this growth. Bacterial species are particularly helpful in the breakdown of dead organic matter to make the resources from dead trees, other plants, and animals available again to other living organisms. The enzymes responsible for this breakdown process can be referred to collectively as cellulase enzymes. Determining which bacterial species contributes to the recycling of organic matter would reveal which bacteria are vital to soil health. Rapid characterization of microbial communities is not easy however. This is due to the extreme diversity of the microbial world and the complexity of microbial communities. Recently, many techniques have been developed for community analysis. Most of these take advantage of the molecular phylogeny derived from 16s rRNA comparative sequence analysis. The research done at the Environmental Biotechnology Institute (EBI) takes advantage of terminal restriction fragment length polymorphism (T-RFLP) technology to perform *in silico* comparative analyses of fragments based on phylogeny and/or by size. T-RFLP allows specific genes of a genome to be targeted, instead of analysis of rRNA alone. In order to target specific genes for T-RFLP, specific primers for the gene in question were created. Two primer sequences were designed around multiple sequence analysis (MSA) of two genes involved in the breakdown of cellulose, ß-glucosidase (EC 3.2.1.21) and 6-phospho-ß-glucosidase (EC 3.2.1.86). The forward primer consisted of fourteen bases with 42.86 percent G/C content, while the reverse primer consisted of nineteen bases with 52.63 percent G/C content. These genes were used to develop primers based on sequence similarity of twelve different bacterial species. These primers were used to amplify DNA isolated from *Bacillus subtilis*. These initial results of primer development did not show success. With further PCR optimization and primer development, these primers may prove useful for use with T-RFLP technology in testing the soil health of soils from Los Alamos, New Mexico. This analysis would make a thorough characterization of the bacterial species contributing to soil health possible.

Carbohydrates are one of the most diverse groups of organic molecules found on this planet. A reducing hexasaccharide is estimated to form more than 10^{12} oligosaccharide stereoisomers, although many of these may not be found in nature (2, 3). Carbohydrates can be classified into four general groups. They may occur as simple sugars (e.g. glucose, galactose, etc.), oligosaccharides (e.g. cellobiose, sucrose, etc.), polymers (e.g. cellulose, starch, etc.), and as conjugates (e.g. many O-, N-, and S-glycosides) to other compounds that may be referred to as aglycones (6). Such diversity related to carbohydrate structure correlates with the diverse array of functions for carbohydrates produced by living organisms. These functions are required for signal transduction, cell adhesion, fertilization, transport, immunity, defense and metabolism. Glyco-materials are produced for industrial products such as, emulsifiers, thickeners, and gelling agents (9). These materials are also used in the manufacturing of foodstuffs and any other use that may involve the absorption of water (4, 6, 9). Glucose, xylose, and N-acetylglucosamine are common building blocks for many disaccharides, oligomers, and polymers. These simple sugars provide a highly diverse array of carbohydrates

* Corresponding author. E-mail: sohrmund@calpoly.edu Phone: (805) 712-8848

when combined with other molecules. Structural architecture is the main purpose for many of these carbohydrates (10).

The most recognized carbohydrate involved with structural integrity is cellulose. Cellulose is a linear polysaccharide that comprises 40 to 60% of all plant woody tissues (9). In nature, cellulose is usually associated with other polysaccharides such as xylan or lignin. It is made of glucose residues that are connected by ß-1,4 linkages (1, FIG. 1). These linkages allow cellulose to form micro fibrils that lead to the structural framework of plant cell walls. The strength provided by cellulose is due to an insoluble, crystalline structure of densely packed glucose units. Up to 10,000 glucose units can be tightly bound by numerous hydrogen bonds (1, FIG. 2).

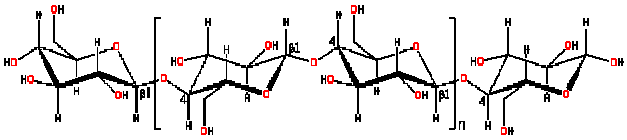

FIG. 1. Cellulose is a linear polymer of ß-(1,4)–D-glucopyranose units (1).

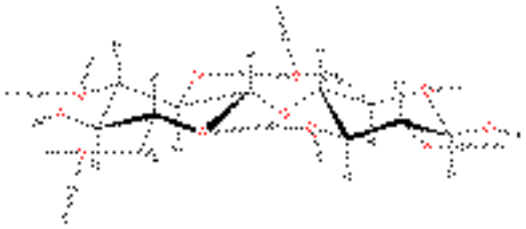

FIG. 2. Hydrogen bonds occur within the linear chain, as well as, with separate linear chains of cellulose polymers (1).

Hydrogen bonds in large numbers incur the bonding strength needed to form strong tissues. Strong bonds correlate with a high-energy yield when these bonds are broken. Cellulose is the most abundant organic source of food, fuel and chemicals (6). The usefulness of cellulose depends upon its hydrolysis to glucose (4, 6). In nature, acidic conditions and high temperatures can be prevalent. These conditions are unsatisfactory for efficient hydrolysis because the resulting sugars are decomposed and contain many impurities. How is it that organisms are able to break down the strong bonds of cellulose and other carbohydrate derivatives in order to harvest calories?

Over the years it has been found that microorganisms such as, fungi and bacteria, are efficient degraders of starch, chitin and polysaccharides in plant cell walls. Attempts to purify hydrolase enzymes led to the realization that microorganisms may produce a multiplicity of enzymes, together making a system for the efficient utilization of polysaccharides (21). The large number of hydrolase enzymes is directly related to the diversity of carbohydrates in the biosphere. Every carbohydrate needs a specific enzyme to facilitate its hydrolysis. These enzymes can be collectively referred to as cellulase enzymes.

Cellulase enzymes exist in complex systems in order to efficiently harvest energy from polysaccharides. These systems are complex for two reasons. First, many of the enzymes themselves are complex. The enzymes are modular proteins that are comprised of one or more catalytic domains and one or more substrate binding domains (21). Secondly, the systems are complex and may contain from a few to twenty or more enzymes (21). All of these enzymes hydrolyze a particular substrate. Researchers have proposed that there are at least two steps involved with hydrolysis. The first step involves a prehydrolytic step where anhydroglucose chains are swollen and hydrated. The second step involves hydrolytic cleavage of the hydrated polymer by one or many enzymes and/or steps (2).

These enzymes occur universally in all living organisms from taxa of the Archaea to humans (7). With the advent of gene libraries and DNA/protein databases, over 400 genes of the cellulase family of enzymes have been cloned and sequenced (21). Currently, a nomenclature system has been implemented and these enzymes are now classified into 82 families based on amino acid sequence and functional similarity (7). The accessibility of enzyme sequences through genetic and protein databases has made the production of possible primer sequences an easier task.

The process begins with creating multiple sequence alignments (MSA) of genetic/protein sequences. Sequence information comes from other researchers who have submitted data to the National Center for Biotechnology Institute (NCBI) or another relevant database. Primers are designed around regions that have high sequence similarity, a low degree of degeneracy's, and the presence of a G/C clamp at the 5' end. It is also

important that the primers do not form secondary structures like primer dimers or hairpins. For the purposes of T-RFLP, primers must amplify a DNA fragment of 500 – 650 base pairs.

T-RFLP is a nucleic acid-based approach that can provide insight into the structure and function of microbial communities. T-RFLP employs the polymerase chain reaction (PCR) using a primer pair that targets 2 highly conserved regions of the gene. Each primer is 5' labeled with a different fluorescent dye. The termini of the resulting amplicons are then distinguishable by their fluorescent tags. Amplicons are identified by size determination of the labeled terminal fragments by capillary electrophoresis with laser-induced fluorescence detection (18). Comparison of the lengths of terminal fragments with those predicted based on analyses of GenBank submissions, entries into reliable databases, or comparison with clone libraries, leads to presumptive identification of the species (11,18). The technique has both high sensitivity and throughput making it ideal for comparative analyses (16, 17).

The Environmental Biotechnology Institute (EBI) has been creating primers for fungi and bacteria that can detect cellulase enzymes. This research is being done in order to characterize species in soil nearby the Los Alamos National Laboratories in Los Alamos, New Mexico using terminal restriction fragment length polymorphism (T-RFLP) technology. The goal of this project was to design primers based on a subset of bacterial cellulase enzymes to be used for T-RFLP analysis. Primers were designed to amplify DNA for a broad spectrum of bacterial species. ß-glucosidase (EC 3.2.1.21) and 6-phospho-ß-glucosidase (EC 3.2.1.86) sequences were used as the subset of cellulase enzymes to be considered for primer development.

There are three types of ß-glycosidases. Bacteria are able to produce an extracellular, cell-wall bound, and intracellular type of ß-glycosidase based on the carbon source (9). The common motif for the enzyme is that it has two active components. One of these sites is the cellulose-binding domain (CBD) and the other is the catalytic site for cleavage of beta-glucosides and ß-galactosides (9, FIG. 3).

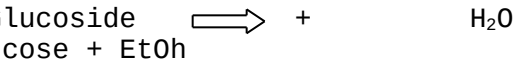

FIG. 3. ß-Glycosidase catalyzes the hydrolysis of glucosides to Glucose (3, 9).

ß-glycosidase cleaves ß-D-glycosides to produce two D-glucose units (FIG. 3). ß-glycosidase does not act upon cellulose directly but is important for thorough degradation of cellulose. ß-D-glycosides act as negative feedback inhibitors to cellulase enzymes that degrade cellulose prior to the evolution of ß-D-glycosides (9). ß-glycosidases are important to reduce inhibition of cellulases and their mechanism of action produces glucose units from carbohydrate polymers.

The specific mechanism of these enzymes begins with a glycosylation step that binds the anomeric carbon of the ß-D-glycoside with the CBD (15). This forms a covalent glycosyl-enzyme intermediate. A deglycosylation step performs a nucleophilic attack on the covalent bond between the glycone and the enzyme (9, 15). The result is the release of the glycone and the evolution of two glucose units. It is postulated that a binding domain near the C-terminal end of the enzyme is responsible for the catalytic activity of the protein (9, 15). The CBD is believed to be located near the N-terminal end of the enzyme (7, FIG. 5).

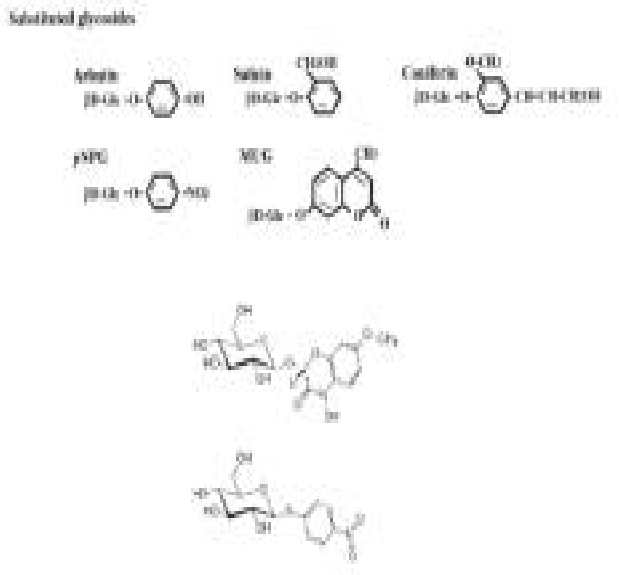

FIG. 4. The two common ligands for ß-D-glycosidases and some examples of substituted glycosides (9).

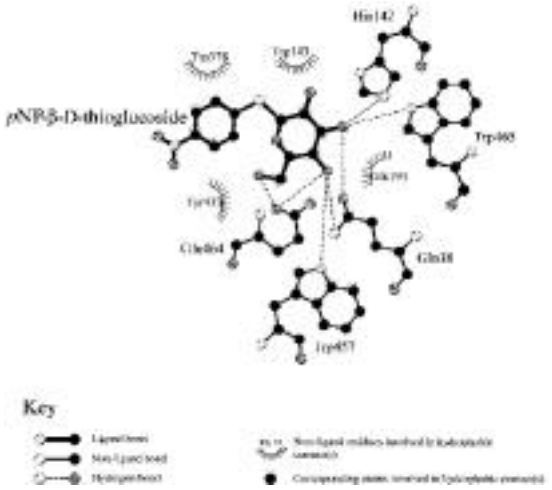

FIG. 5. The proposed binding site of a ß-D-glycosidase enzyme with reference to substrate and AA residue position (7).

The molecular weight (MW) of ß-glycosidases range from 96 Kda to 190 Kda, depending on the carbon source and purpose (2). The optimum pH is 4.4 – 6.0 depending on its purpose (2). $HgCl_2$ and other heavy metals inhibit the enzyme (2). Recently, a crystalline structure of this protein has been described (7, FIG. 6).

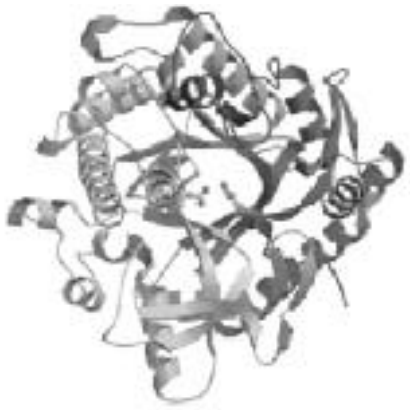

FIG. 6. Ribbon structure of *Zea mays* ß-glucosidase with 2.5 Å resolution elucidated by x-ray crystallography (7).

ß-glycosidases constitute a major group amongst cellulase enzyme families. These enzymes have been found to play a part in fundamental biological processes. These processes include growth and development, chemical defense, host-parasite interactions, cellulolysis, lignification, glycosylceramide and vitamin B_6 metabolism, signaling, etc. These enzymes have also been used for biotechnological processes such as biomass conversion, food detoxification and beverage quality enhancement. With it's involvement in many biological processes and it's use in biotechnological research and development, it is apparent that the class 3 families of cellulases are an important link to understanding bacterial communities and their roles in biotechnology and the subterranean ecosystem. Biotechnological study is important for understanding host-microbe interactions and the many options for industrial processing and development.

The results presented in this paper will show how two primer sequences were determined from a multiple sequence alignment of ß-glucosidase and 6-phospo-ß-glucosidase from 12 different bacterial species. All of the species can be found in soil. Although the primer sequences presented have low degeneracy, their discrepancies in length and melting temperature made them difficult to develop initially. PCR results showed that no ß-glucosidase DNA was amplified, but primer dimers were present. Primers that are found to amplify ß-glucosidase DNA successfully will assist in the rapid characterization of bacterial species contributing to the energy cycling of cellulose in the soil. T-RFLP analysis will make this rapid characterization possible.

MATERIALS AND METHODS

Database Searching for Cellulase Enzymes. All database searching was done through the website of the National Center for Biotechnology Institute (NCBI) at http://www.ncbi.nlm.nih.gov/. Sequences collected were catalogued by accession number, length, DNA or protein, bacterial species, and the type of enzyme within the family. This information was saved in a laboratory notebook, zip disk and on the LMMR hard drive.

Local Alignments of ß-glycosidases. Nucleotide sequences from NCBI were saved as GenBank and FASTA files. Using the LaserGene Navigator suite of software, FASTA files were copied into EditSeq files. Twenty-five EditSeq files per alignment were then entered into the MegAlign program. Multiple Sequence alignments (MSA) were performed using the ClustalW algorithm. Stringency was varied to achieve an alignment with a small number of gaps and mismatches. Altering the stringency was also done to yield as many regions with a high degree of sequence similarity as possible. MSA's were consolidated based on obvious discrepancies (i.e. the presence of a thermophilic bacterium) and a lack of sequence similarity to the consensus. The lack of sequence similarity was measured subjectively and on a percent

similarity basis when needed. Consolidated trials were then aligned with each other and sequences with low similarity were discarded.

Peptide sequences from NCBI were saved as GenPept files. These files were also transferred from EditSeq to MegAlign. In MegAlign, the peptide sequences were reverse translated into genetic code. The degeneracy of unambiguous codons led to unsuccessful searches for primers. However, BLASTing protein sequences provided accession numbers to sequences that would prove useful with development of successive alignments.

BLAST searching for related sequences. The consensus sequences of consolidated alignments were submitted to a BLAST search using the NCBI web page at http://www.ncbi.nlm.nih.gov/. Sequences that were found to be of pertinent taxa, with high similarity (i.e. high bit scores), and low E-values were chosen to be included into a final consolidated MSA. This MSA was used to determine the two primer sequences that were ordered.

Primers. Two primers were ordered from Sigma-Genosys at 1442 Lake Front Circle, The Woodlands, TX 77380. Lyophilized preparations were delivered and diluted with nanopure PCR water to a 100 µM stock concentration. Expected fragment sizes were between 776 and 806 bp for the species used in constructing the alignment.

Primer Development. Primers were tested for efficacy by doing PCR with *Bacillus subtilis* DNA (31 ng/µl) and a negative control. Primers were tested with a gradient PCR machine from 47 – 59 ºC to test for varying annealing temperatures. Concentrations of $MgCl_2$ were varied from 1.0 mM to 4.0 mM. 10 ng of DNA was used per reaction tube. Reaction volumes of 50 µl consisted of 5 µl 10x Buffer, 10 mM dNTPs, 20 µg/ml BSA, 5U/µl *Taq* polymerase, 10 µM forward and reverse primers and enough nanopure water to fill reaction volume to 50 µl.

The PCR began with a 94.0 ºC hot start for 10:00 minutes. The PCR cycles consisted of a 94.0 ºC melting temp. for 00:30 seconds/cycle, a 47-59 ºC annealing temp. for 00:30 seconds/cycle, and a 72.0 ºC polymerase elongation step for 1:00 minute/cycle. The PCR ended with a 72.0ºC elongation for 10:00 minutes and a holding period at 4.0 ºC for infinite time.

Samples were loaded into a 1.6% agarose gel. 1 Kb DNA ladders were loaded in 5 µl volumes, while 7 µl of sample was loaded with 2 µl of loading dye. The gel was allowed to run for ~2 hours at 60 v. Test results were visualized with a Gel Doc and EtBr staining.

RESULTS AND CONCLUSIONS

This experiment was performed in order to develop primers for the Environmental Biotechnology Institute (EBI). The developed primers would be able to amplify DNA from bacterial species that produce a subset of cellulase enzymes. The results presented here show that two primer sequences were determined to amplify a subset of the cellulase gene family. Although the sequences were determined, the primers must be further developed to assess their worth as a working primer set.

Primers designed to amplify ß-glucosidase were chosen after analysis of over 160 DNA and protein sequences. These sequences were chosen after searching the NCBI database for bacterial cellulase enzymes that included cellobiose hydrolase, amylase, ß-mannosidase, endo-1,4-ß-glucosidase, 6-phospho-ß-glucosidase and ß-glucosidase. Protein and nucleotide sequences of ß-glucosidase and 6-phospho-ß-glucosidase were chosen instead of other cellulases based on their close similarities and relative abundance. These similarities were based on MSA's that showed local alignments with the consensus of greater than 80 % similarity.

The multiple sequence alignment that led to the determination of two primer sequences encompassed twelve different bacterial species and eighteen sequences overall (FIG. 8). All of these species can be found in soil. These species include members of *Lactobacillus*, *Lactococcus*, *Streptococcus*, *Clostridium*, *Listeria*, *Staphylococcus*, and even *Yersinia pestis* (FIG. 8). The consensus sequence determined from a multiple sequence alignment (MSA) is shown below (FIG. 7).

The forward primer consisted of fourteen bases and it lacked degeneracy at all positions (5' TTTGCTGAAATGGG 3') (FIG. 7, FIG. 8). This sequence has a G/C clamp at the 3' end, a melting temperature of 50.8 ºC, and it is located ~300 bases past the 5' end of the coding sequence. The sequences in the MSA showed 91.6% similarity with the consensus sequence.

The determined reverse primer consisted of nineteen bases and some degeneracy at the 15th position (5' GGATCAATTTGCCANCCCC 3') (FIG. 7, FIG. 8). This sequence has a G/C clamp at the 3' end and it is located ~425 bases before the 3' end of the coding sequence. The sequences in the MSA showed 88.9% similarity with the consensus sequence.

Primers were tested with a gradient PCR machine from 47 – 59 ºC to test for varying annealing temperatures. Concentrations of $MgCl_2$ were varied from 1.0 mM to 4.0 mM in order to test the primers specificity at varying salt

```
1       ATGACATCATTTTCCAAAAGATTTTTTATGGGGTGGAGCTGTTGCTGCTAATCAAGTTGAAGGTGCATATAA
73      TGAAGATGGTAAGGGATTATCTGTTGCAGATGTTATACCAAAAGGTGGAATGGGAGCCCATAGATTACAGA
144     GAAACCAATAGAAGATATATTATCCAAATCATGGGGCATTGATTTTTATCATCGTTATAAAGAAGATATTGC
216     ACTTTTTGCTGAAATGGGCTTTAAATTTTTTCGTACTTCTATTGCCTGGACTCGTATTTTTCCAAATGGTGATG
290     AAGAAGAACCAAATGAAGCAGATTGCAATTTTATGATGAATTATTTGATGAATTGTTGAAGTATGGTATTG
361     AACCGTTAGTTACTTTATCTCATTATGAAACTCCTTTACATTTAGCTAAAAAATATGATGGTTGGACAAATA
433     GAAATTGATTGATTTTTTTGAAAATTTTGCTGAAACAGTATTTGAACGTTATAAAGATAAAGTTAAATATTG
505     GATGACTTTTAATGAAATCAATTCAGTATTGCAATTCCATTTACTAGTGCAGGTATTATATGGATAATGAAA
577     AATTAAGAGACAAGATTTATAGCAAGCAGTTCATCATGAATTAGTTGCATGTGCTCGAGCAGTAAAAATTG
648     GTCATGAAATTAATCCTGATTTTAAAATTGGTTGTATGATTGCTAGGTGTGCCTGCTTATCCATTGACTTGTA
721     ATCCAAAAGATATATTGGCTGCTAATGAATCTGAGAGATAAATTTTTCTTTGCAGATGTTCATGTTAGAGGT
793     AAATATCCTAGCTATGTTAAAAGTATTTTAAAGAAAATGATATTAATATTGATATTGCAGAAGAAGATGAA
864     GAGATATTACAAGAAGATACGGTTGATTTTATATCTTTTAGTTATTACATGAGTTCATCAAGCTGTCGATTCC
937     AAAGAAAAATAAATATGGGCAGGCAATGTTATTGGTGGTTTGTCAAAATCCTCATTTAAAAGCTTCTGAAT
1008    GGGGTTGGCAAATTGATCCAGTAGGTCTACGTTTTACACTGAATTAGTTCTATGATCGCTATCAAATACCAT
1080    TATTTATTGTTGAAAATGGTTTTGGTGCTGTAGATGAAGTTAAAGAAGATGGCATGGTTCATGATTATCGAA
1152    TTGATTATTTAAGTGACCATATAGTGCAAATGGAAGAAGCTATAGAAGATGGTGTTGAACTAATGGGTTAT
1223    ACTCCTTGGGGATGTATTGATTTTGTTTCTGTTTCTACCGCTGAAATGAATAAACGTTATGGTTTTATTTATGT
1297    TGATAGAGATGATGATGGAAATGGTACATTAAAACGATATAAGAAAAAATCTTTTGATTGGTATAAAAAA
1367    GTTATTGAATCTAATGGAGAAGACTTAAAAGATAA
```

FIG. 7. Consensus nucleotide sequence from MSA of eighteen bacterial ß-glucosidase and 6-phospho-ß-glucosidase sequences. The regions chosen for the determined primer set are bold and underlined. The forward primer is in blue and the reverse primer is seen in red.

3.2.1.86	(L. gasseri)	TTTGCTGAAATGGG	GGGGCTGGGCAACTGATCC
3.2.1.21	(L. lactis)	TTTGATGAATTGGG	GGGGCTGGCAAATTGATCC
3.2.1.86	(S. pyogenes)	TTTGCCGAGATGGG	GGGGTTGGCAGATAGACCC
3.2.1.86	(S. pneumoniae)	TTTGCTGAAATGGG	GGGAGTGGCAAATTGATCC
3.2.1.86	(S. pneumoniae)	TTTTCTGAAATGGG	GGGGATGGCAAATTGATCC
3.2.1.86	(C. acetobutylicum)	TTTGCAGAGATGGG	GGGGATGGCAAATAGATCC
3.2.1.21	(C. acetobutylicum)	TTTGCTGAAATGGG	GGGGATGGCAAATAGATCC
3.2.1.86	(S. pneumoniae)	TTTGCTGAAATGGG	GGGAGTGGCAAATTGATCC
3.2.1.86	(S. pneumoniae)	TTTTCTGAAATGGG	GGGGATGGCAAATTGATCC
3.2.1.21	(S. pyogenes)	TTTGCCGAGATGGG	GGGGTTGGCAGATAGACCC
3.2.1.21	(S. mutans)	TTTGCTGAAATGGG	GGGGTTGGCAAATTGATCC
3.2.1.21	(L. plantarum)	TTTGCTGAGATGGG	GGGGTTGGCAAATTGATCC
3.2.1.86	(Y. pestis)	TTTGCCGAAATGGG	GGGGGTGGCAGATTGATCC
3.2.1.21	(L. monocytogenes)	CTTGCTGAAATGGG	GGGGCTGGGCAGTCGATCC
3.2.1.21	(L. innocua)	ATGGCAGAAATGGG	GGGGTTGGCAAATTGATCC
3.2.1.21	(L. innocua)	CTTGCTGAAATGGG	GGGGCTGGGCTGTGGATCC
3.2.1.86	(S. aureus)	TTTAAAGAAATGGG	GGGGTTGGGCGATTGATCC
3.2.1.21	(C. longisporum)	TTTGCAGAAATGGG	GGGGATGGCAAATTGATCC

FIG. 8. Forward and reverse primer sequence data, showing the degeneracy in comparison to the consensus sequence, as determined by multiple sequence alignment (MSA). E.C. 3.2.1.21 corresponds to the enzyme ß-glucosidase and E.C. 3.2.1.86 corresponds to the enzyme 6-phospho-ß-glucosidase. Degenerate bases are seen in red, while shaded regions are matches with the consensus.

concentrations. The Gel Doc photo of the PCR results showed that there wasn't a 775 – 800 base pair fragment present (FIG. 9). The only results were that primer-dimers appear to be present on the gel (FIG. 9).

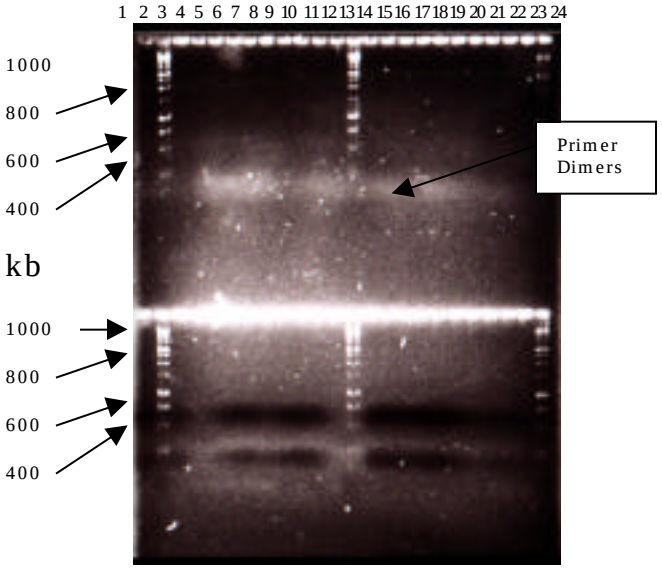

FIG. 9. Gradient PCR results with the determined primer set and genomic DNA from *B. subtilis* as seen on an agarose gel. Lanes 1-2, 12-13, and 23-24 on both halves of the gel contain a negative control (2.0 mM $MgCl_2$ on top half and 4.0 mM $MgCl_2$ on the bottom half) and a 1 kb ladder at 47ºC, 51ºC and 55ºC respectively.

DISCUSSION

ß-glucosidase and 6-phospho-ß-glucosidase were chosen as the target genes for primer development due to their abundant representation amongst many bacterial species in comparison to other cellulase enzymes. These enzymes had the most extensive amount of study in relation to structure and function. This was discovered when searching the literature and sequence databases available. This would lead the researcher to conclude that the sequence data presented may be the most reliable. By using two enzyme types the developed primers would be able to recognize bacterial species that produce one, the other, both, or a similarly related enzyme. This breadth would prove useful for subterranean analysis of diverse bacterial communities.

The significance of the alignment containing twelve different bacterial species was that the developed primers would have a better chance of amplifying community DNA. This means that our primers may encompass a broader range of species to be recognized by T-RFLP analysis. These species could be bacteria that we had not considered during our development process. The primers were developed around bacterial species commonly found in soil so that they could be used to amplify community DNA extracted from soil near Los Alamos, New Mexico.

The forward primer sequence is good due to its legitimate G/C clamp at the 3' end, it's moderate melting temperature, and it's location past the 5' end of the coding sequence. Being located ~300 bases past the 5' end ensures that the primer is not exclusively associated with a promoter sequence or a membrane binding portion of the enzyme. This sequence is problematic because of its short length, which may lead to non-specific binding, and it's large difference from the reverse primer characteristics.

The reverse primer sequence is good because of it's G/C clamp at the 3' end and it's position ~425 bases before the 3' end of the coding sequence. Being located ~425 bases before the 3' end ensures that the resulting fragment from PCR would be ~775 base pairs long. This sequence is problematic because of its high melting temperature (66.2 ºC), and it's large difference from the forward primer's characteristics.

Primer development using a gradient PCR machine tested the ability of these primers with varying annealing temperatures and salt concentrations. Gradient PCR was used in order to hasten the process. It was also used to test for varying conditions at the same time, in order to ensure consistency. For all the good qualities of these two primers they did not produce any data upon PCR analysis.

The high degree of similarity between the sequences chosen, through many rounds of search and refinement, implies that the primers are specific for the ß-glucosidase gene. Since these primers were designed around bacterial species found in soil we can conclude that they must be specific for a certain amount of bacteria

in soil. These results show that perhaps the discrepancy in primer length was too much to overcome when choosing an optimal annealing temperature for these primers in concert. Although these primers haven't worked together through one round of primer development, with further development they may prove to be useful. They may not be useful for amplification together, but perhaps more primers can be developed to compliment those presented here.

Further work could be done with primer development of the primers presented here. Perhaps the optimal conditions for both of these primers to work together have not been tested. Primers could also be designed around the other cellulase enzymes discovered during research.

It would be significant if one or both of these primers proves to amplify ß-glucosidase because the primer(s) could then be used for T-RFLP technology. This would allow EBI to test soil communities for the bacterial species present that contribute to soil health. A thorough characterization of these bacterial species would pinpoint the species that could be best used for biotechnological purposes.

ACKNOWLEDGEMENT

This research was supported by the Environmental Biotechnology Institute (EBI), California Polytechnic State University, San Luis Obispo, California 93407. This project owes special thanks to Dr. Susan Elrod, Darci Spiker, and Dr. Chris Kitts.

REFERENCES

1. http::/ / www.sbu.ac.uk / water / hycel.html

2. http:/ / www.worthington-biochem.com / manual / C / CEL.html

3. http:/ / www.worthington-biochem.com / manual / G / E.html

4. **Artymiuk, P.J., Rice, D.W., Poirette, A.R., and Willett, P.** 1995. Struct. Biology. **2:**117-120.

5. **Boschker, H.T.S., and T.E. Cappenberg.** 1994. A Sensitive Method Using 4-Methylumbelliferyl-(beta)-Cellobiose as a Substrate To Measure (1,4)-(beta)-Glucanase Activity in Sediments. Appl. Environmental Microbiology. **60(10):** 3592-3596.

6. **Clark, Anthony J.** Biodegradation of Cellulose: enzymology and biotechnology. Technomic Publishing Co. 1996.

7. **Czjek, Mirjam, M. Cicek, V. Zamboni, W.P. Burmeister, D.R. Bevans, B. Henrissat, and A. Esens.** 2001. Crystal Structure of a Monocotyledon (maize ZMGlu1) ß-Glucosidase and a Model of its Complex with p-nitropehnyl ß-D-thioglucoside. Biochemical Journal. **354:**37-46.

8. **Ducros, V., M . Czjzek, A . Belaich, C . Gaudin, et al.** 1995. Crystal structure of the catalytic domain of a bacterial cellulase belonging to family 5. Structure. **3(9):**939-49.

9. **Faure, Denis.** 2002. The Family-3 Glycoside Hydrolases: from Housekeeping Functions to Host-Microbe Interactions. Appl. Environmental Microbiology. **68(4):**1485-1490.

10. **Holm, L., and Sander, C.** 1995. EMBO J. **14:**1287-1293.

11. **Kitts, C.L.** 2001. Terminal restriction fragment patterns: a tool for comparing microbial communities and assessing community dynamics. Curr. Issues Intest. Microbiology. **2(1):** 17-25.

12. **Kowalchuk, G.A., J.R. Stephen, W. De Boer, J.I. Prosser, T.M. Embley, and J.W. Woldendorp.** 1997. Analysis of ammonia-oxidizing bacteria of the beta subdivision of the class Proteobacteria in coastal sand dunes by denaturing gradient gel electrophoresis and sequencing of PCR-amplified 16S ribosomal DNA fragments. Appl. Environmental Microbiology. **63(4):** 1489-1497.

13. **Le Coq, Dominique, C. Lindner, S. Kruger, M. Steinmetz, and J. Stulke.** 1995. New ß-Glucoside (*bgl*) Genes in *Bacillus subtilis*: the *bglP* Gene Product Has Both Transport and Regulatory Functions Similar to Those of BglF, Its *Escheria coli* Homolg. Journal of Bacteriology. **177(6):**1527-1535.

14. **Li, Bin, and V. Renganathan.** 1998. Gene Cloning and Characterization of a Novel Cellulose-Binding ß-Glucosidase from *Phanerochaete chrysosporium*. Appl. Environmental Microbiology. **64(7):**2748-2754.

15. **Lymar, Elena S., B. Li, and V. Renganathan.** 1995. Purification and Characterization of a Cellulose-Binding ß-Glucosidase from Cellulose-Degrading Cultures of *Phanerochaete chrysosporium*. Appl. Environmental Microbiology. **61(8):** 2976-2980.

16. **Marsh, T.L.** 1999. Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterizing diversity among homologous populations of amplification products. Curr. Opin. Microbiology. **2(3):** 323-327.

17. **Marsh, Terence L., P. Saxman, J. Cole, and J. Tiedje.** 2000. Terminal Restriction Fragment Length Polymorphism Analysis Program, a Web-Based Research Tool for Microbial Community Analysis. Appl. Environmental Microbiology. **66(8):** 3616-3620.

18. **Nilsson W.B., Strom M.S.** 2002. Detection and identification of bacterial pathogens of fish in kidney tissue using terminal restriction fragment length polymorphism (T-RFLP) analysis of 16S rRNA genes. Dis. Aquat. Organ. **5; 48(3):** 175-85.

19. **Ritchie, Nancy J., M. E. Schutter, R. P. Dick, and D. D. Myrold.** 2000. Use of Length Heterogeneity PCR and Fatty Acid Methyl Ester Profiles To Characterize Microbial Communities in Soil. Appl. Environmental Microbiology. **66(4):** 1668-1675.

20. **Singh, Ajay. K. Hayashi.** 1995. Construction of Chimeric ß-Glucosidases with Improved Enzymatic Properties. Journal of Biological Chemistry. **270(37):**21928-21933.

21. **Warren, R.A.J.** 1996. Microbial Hydrolysis of Polysaccharides. Annual Reviews of Microbiology. **50(1):**153-212.