GENES ENCODING CELLULASE

Polynucleotide sequences are provided encoding a thermostable cellulase and directing its increased expression are provided, and the use of the thermostable cellulase in hydraulic fracturing methods and the treatment of flowback fluids.

Specification includes a Sequence Listing.
Figure 1

[Graph showing Viscosity vs. Time with two curves: one labeled 'No Enzyme' and another labeled 'SEQ ID NO: 2 @ 200ppm']
Figure 2

Legend

Lanes 1, 5, 9 = negative control
Lanes 2, 6, 10 = SEQ ID NO:3
Lane 3, 7, 11 = SEQ ID NO:4
Lane 4, 8, 12 = SEQ ID NO:1
Lane 13 = Molecular Weight Ladder
Figure 3

Improved Glucanase Expression

Legend

Lanes 1 = negative control
Lanes 2, 5, 8, 11 = SEQ ID NO:3
Lane 3, 6, 9, 12 = SEQ ID NO:4
Lane 4, 7, 10, 13 = SEQ ID NO:1
Figure 4

SEQ ID NO:1

ATGGGGCGCTCGATCCGTCTTTGAAACGTAAACAAATCTTTGGGCGCGGATTTAATATTCGCGC
AATGCAGCTTCAGAGCACCAGAAAATGAAGGCGGACTGGGGAGTGGTGATAAAAGATGAGTTTC
TTGGCACATTATAAGAAAGGCGGTTTCTCTCATGTTCGAACTCCAAATAAGATGGAGT
ACGCAGCTCAGCGGTCTTCTCTTTATRAGGAGCCTTTCTTCTCTAATAAGAGTG
GATAAGTGAATAGCAGGACCCCTGAAAAGAGCGAGACTGGGCTGTCTGTTGTTATAACTATTATCT
CACTACGAGAGGTAAAATGAGAATCGAGAAGAAGAAAAACAGAGAAAGATTCTTCTGCTTCT
TTGGAAACAAAATTGCTGAGTCGTTATTAAAAGAATCTATCCCGAAACTCTATTAAAAATGAGA
CTCTAAAGTTATAGATAACATTTGCAAAAAGACACACTCAGGATTATAGGCGACAGCT
GAATGGGGGCGCTGATATCTGCCATGAAAAACTGAGCGGTCCCAAATGGAGAAAAAT
GCAGATGTTAGTTAATTTGACTATACATTACAAATCCTTTTGAAATTTATCCATCAAGGACTGAG
TGTTGGCTCGAGCTGAGAARTGTTTGAGGAAAGAAGTGGGATGCTCAGGAGATGATCG
AAAACTTTGAGAAGAAATTTTATTTTATAGAAGAATGGGCAAAAAAAGAAACAAGAA
CCAAATTACATATGGTGGTTGGTGGCTCAGAAAAGCGACCTTCTCAATGAAATCAAATA
AAATGGGACCTCCTTGGTCGTCGCCAGACCAGAAAAGGGGGAGTAGGAGACTGGGCAAGC
TGGAATTTTATTGTTCCCAGCTTTATTGATCCCTCTGAGAAAGAATAGTGGAATAAA
GATCTTTTAGAGCTTATAGGAGATAGCATTGAAATGA
SEQ ID NO:2

MGVDPFERNKILGRGIGNALEAPNEGDWGVVIKDEFFDIKEAGFSHVRIPIRS
THAQAFFPYKIEPSFFKRVDEVINGALKRGLAVVINIHHEELMNDEEHEKERFLAL
WKQIADRYKDFEFTPTEILNEPPHNLTPEKWNELLEALKVRISDIDKHTVIGTA
EWGGSASEALEKLPKKEKNAIVTIHYYNPFEFTHGQAEGWPCGSEKWLGRKWSFDDQ
KHLIEEFNFIEEWSKKNRPPIYIGEFGAYRKLDSRIMWTSCFVVREAERKGSWAY
WEFCSEFGVYDPKQQNKLLEALIGDSIE
ATGGGTGTGGATCTTTCCTTTGGAAAGGAGCAAATATTGGAAGGACATTGATATAGGAAAATGCGC
TTGAAAGCACCACATGAGGGAGACGTGGGGAGTGTTGATTAAGGAGTACCTCTCGACATTTAA
AGAAGCCTGTTTCCTCATGTTTGCAATTAGAGATGAGACCGAGCCTGAGGGGTTTCTTCT
CTTATAAACATGAGCTGCTTTTCTCTTCTCTAAAAAGAGTGGAAGTGATAAAACGGAGCCCTGAAAAG
GAGGACTGGCTGTGGTTTAPAAATATTCAATCACTACCCGAGAGTTGATGAGTCCAGAAGAAACA
CAAGGAAAAGATTTCTCTCCTTTTGAAACAAATTTGCTGATTCTATTAAGACTATCCCGAAGACT
CTATTTTTTGAAATTTCTGATAATGCGAAATCTTTACTCGGAAATGATGGAATCGTGCTAGGAA
TTGAAAGCTCTAAAGTTTAAAGATCAATTGACAAAGACACGTGTATTATAGGAGACAGC
TGAAAATGGGGGTTATATCTGGCGCCTTTGGAAAACCTAGGAGGTCACACCAATGGAGAAAATTCGGAT
GTTAACAATCACTACTCAATCCTTCGGAATTACCAATCPTAGGACGCTGGGCTGGGCGCTGGAT
CTGAGAAATGGGTCTGGGAGAAAGTGGGGATCCAGATGACTGAAACATGGATGAAAGAATT
CAATTTTAGAGAAGATGGTCAAAGAAAGCAACAAAGACCCATTTACATGATGATGTGCGCC
TACCAGAAAGCTGAGGCTTTGGAATCAAGAATAAATGGACCTCTTTTGCTTTTCGGAAGCCGAGA
AAAGGGGTTGGAGCCTGGGATATCGGGAATTTGTTCTCGGTATTTTGTGTTTATGATCCTCTGAG
AAAAAGTGGGATAAAGATCTTTTATGAGCTTTTAATAGGAGAGATGCAAGTTGAAA

Figure 6
SEQ ID NO: 3

ATGGGTGTGGATCTTTCCTTTGGAAAGGAGCAAATATTGGAAGGACATTGATATAGGAAAATGCGC
TTGAAAGCACCACATGAGGGAGACGTGGGGAGTGTTGATTAAGGAGTACCTCTCGACATTTAA
AGAAGCCTGTTTCCTCATGTTTGCAATTAGAGATGAGACCGAGCCTGAGGGGTTTCTTCT
CTTATAAACATGAGCTGCTTTTCTCTTCTCTAAAAAGAGAGTGGAAGTGATAAAACGGAGCCCTGAAAAG
GAGGACTGGCTGTGGTTTAPAAATATTCAATCACTACCCGAGAGTTGATGAGTCCAGAAGAAACA
CAAGGAAAAGATTTCTCTCCTTTTGAAACAAATTTGCTGATTCTATTAAGACTATCCCGAAGACT
CTATTTTTTGAAATTTCTGATAATGCGAAATCTTTACTCGGAAATGATGGAATCGTGCTAGGAA
TTGAAAGCTCTAAAGTTTAAAGATCAATTGACAAAGACACGTGTATTATAGGAGACAGC
TGAAAATGGGGGTTATATCTGGCGCCTTTGGAAAACCTAGGAGGTCACACCAATGGAGAAAATTCGGAT
GTTAACAATCACTACTCAATCCTTCGGAATTACCAATCPTAGGACGCTGGGCTGGGCGCTGGAT
CTGAGAAATGGGTCTGGGAGAAAGTGGGGATCCAGATGACTGAAACATGGATGAAAGAATT
CAATTTTAGAGAAGATGGTCAAAGAAAGCAACAAAGACCCATTTACATGATGATGTGCGCC
TACCAGAAAGCTGAGGCTTTGGAATCAAGAATAAATGGACCTCTTTTGCTTTTCGGAAGCCGAGA
AAAGGGGTTGGAGCCTGGGATATCGGGAATTTGTTCTCGGTATTTTGTGTTTATGATCCTCTGAG
AAAAAGTGGGATAAAGATCTTTTATGAGCTTTTAATAGGAGAGATGCAAGTTGAAA
Figure 7
SEQ ID NO: 4

TCTA TCTAGTGGAGTTATGTGCTCATCCGTATTGCCATGTAACAAATCTTGAGGCGCC
GGCATTAATAGGGCATGCGCTGCAAGCAACAAATTGAGACGGCTGAGGAGTGAGTGATGAA
ATGAGTCTTTGCAATTAAGAAGACGCGCTTTTCTCTCAGTTGCAATCCAAATAGATGAG
TACGCACGCTAAAGGCTTTCTCCTTATAAAATCGACCTTCTTCTTCTTCCAAAGGTGAGGCTA
GTGATACAGAGCCCTGAAAGGAGGACTGCGTGGTGGTTTTATATAAATATCTACACTGAGGAT
TATGCAATGATCCAGAAGACACAGAAGAAAGATTTCTTGCTCTTTGGAAACAAATGCTGATCG
TTTAAAGACTATCCCGAAAACTCACTTTTGGAAATTCTGGAACACCTCAGGAAATCTTACT
CCGGAAAANTGGAATGAACTGCTTTGGAGGAAGCTCTAAAAGGATATAAGATCAATTGCAAAAGGC
ACACTGTAATTAGGCACAGCTGAATGGGGGTATATCTGCCCTTGGAAACACGCTGAGGTCC
AATGAGGAAAAATGGGATAGTTAACATCTACACTACATACATCCCTTGAAATTTACCCATCAA
GGAGCGTAGGCTGGTGGTGGCTGGAATGTTGGGAAGAAAGGTTGGGATCTCCAGATGATC
AGAAACATTTGATAAGAAGATCCAATTATAGATAGAGAATGTCAAAAGAAGAAGAAGACCAAT
TTACATAGGTTGAGTCTTGCTCAAGAAAGACCTGACCTTGAATCAAGAAATAAAAATACATCC
TTTGTCTGTGCGCAAAGCGGAAAGAGGGGCTGAGGCTGGAATTTTGTCTCGGTT
TTGCTGGTTATGATCCCTGAGAAACAGTTGGAATAAAGATCTTTTAGAAGCTTTATAGGAGG
AGATAGCAATTGGAATGA
Figure 8

SEQ ID NO: 5

ATGGGTGTGGATCTTTTGGAGAGGAGCAAAATATTGGAAGAGGCAATATAATAGGAAAATGGC
TGAAAGCACCAGATTAGGAGAGACTGGGAGTGGGATAAAAGAGAGAGTTCTCGACATTATAA
AGAAGCCGTTCTCTCATGTTCAATCCCAATAAAGATGGAGTACGCCGCCTAAGCGCTTTCT
CCTTAATAATCATGGATCGCTTCCTTTGAARAGGTGAGTGAAGGGATCAAAGGAGCTGTG
GAGATCCTGGCTTTATAAATATTCCATCCCTACGGAGGTATATGATGATCCCAAGARGAAC
CAAGGAAAGATTTCTGCTTTTTGAAAACAAAATTGCTGATCGTTATAAAGACTATCCCGAAACT
CTTTTTTGGAAAATCTGAAATGAACTTCACGGAATCTTACTCCGGAAAATGAAATGAACTGC
TGAGGGAAGCTCTAAAGATTATAGATCAAATGGCAAAAAAGCAGACACTATAATTATAGGCACAGC
TGATGGGCGGTATATCTGCGCCCTTGGAAAAACTGTCGCTGGCCAAAGATTGGAAGAAAAATTCTATA
GTACATTCTACTAAATCTTTTGGGGATCTTATTTACACCATCAAGGAGCTGAGTGGGTGGGAAAGAT
CTGAGAAATGTGGTGGGAAGAAGTGGGAGATCTCGAGATGATAAATTTGATAGAAAGATT
CAATTTATAGAAAGATGGCATAAAAGAAGCCAAAGAAGAAAATGACTATACATGGATGGTTGGCC
TACAGAAAGCTGACCTTGACCTACAGAAATAAAAGGAGGCTCTTTTGTGTCTCGCGAAATGGAGA
AAAGGAGATTGGGCTGGCCATACGTTGGGATTTTGTGTCTCGGAGTTTATGATACCTCTGAG
AAAACCTGGGAATAAAGATCTTTTAGAAGCTTTATAGGAGAGATAGCATTGGAAATTA
SEQ ID NO: 6

MGVDPFERNKILGRCIIHCAGNEAGSHVRIPIKWSTRHAYAFPYKINDRFKRVDEVINGALKRGLAVVINIHHEELMNDEEHKREFLAWQIADHYKDYPETLFEEILNEPHGNLPEKWNELLEALKVIRSDKKHTIIGTAEWGIGSALEKLSVPKWEKNSIVTIHYYNPFETHQGAEWVEGSEKWLGKPDDQKHLIEEFNFIEEWSKKNKRPYIGEFGAYRKADLESRTIWTSSVREMKEKRWSWAYFECSGFGVYDTAKTKWNKDLLEALIGSDTE
SEQ ID NO:7

ATGGAAACAGTCAAGTACGTAAGTGAATCGCCATCTGCAATTGAATAACACAACAATAGTGATGTA
AAGCAGTAAATATTGGAAAATGCTTTAGAAAGCTCTCTTTGGGGCTAACGAGAATGGAATGGGA
GTGAAATATTTTGAGTATAATAAGAGAAGGGGATTTGATTCTGTGTAGATTTCCATAGAGATGG
TCAGCACTCATATCCGAAAAGCCACCACTATGATATTGACAGGAATTTTCTCGAAGAGGTTAACC
ATGTGCGATAGGGCTCTGGAAATATTTCAAGTAAATCATATAATACACCACTATTGAAGAGA
ACTCTATCAGAAACCAGTGATTTAACGCAGTGTGTTTGGTGGAAAATTGGAGACAGATTGCAAA
TTCTTAAAGGATATTACCCGAAAATCTGGTTTTGAAATCTTCACAGACGGCTGCTGACTTTGA
CAGCTGAAAAATGGAAGCGACTTTATCCAAAGATTGGCTCAATATCGGGAGAGCAATCCAAC
CCGGATTGTCATTATCGATGCTCCAATCGGCAACATATACGCGATGAGAATCTGAAATTA
GTCAACGAAACAGCATTTGTTCTTTCCATTACGAACTTTCAAATTCACACATCAGG
GTGGAGAAATGGAAGTTATCCACATCCACACGTTAGGTTAAGTGAATGGCAGGAAATGGGAAAT
TACCACAAATCAAGATCTTCACTTTGAGTCAGGACAGTTGGCAAGAACTTAAACCTTACCAATC
TTCTTGCGTGAATTCGCGTCTATTTCAAAGCAGACATGCACTCAAGGTTGTAAGTGCAACCAGA
GTGTGAGAAAATGCGCAGGAATTTGGATTTCATACGCGATATGGGAAATTTTGCTGCAAGATT
TGGGATATAGGATAGTGGCTCAAAAATGCGAACCATTGGCAACAGCTGGTGGTCAGAC
GGCAAGAGAAGATA
SEQ ID NO:9

ATGGGTGTGATCCCTTTTGAAGGAACAAATATTGGGAGGACATTGAATATAGGAAAAGCGC
TTGAAGACACAAATGAGGAGACTGGGAGTTGTCATAAAGAATGAGATTTCGACATTATAAA
AGAAGCCGGTTTCTCTCTATGTCGATATCCATATAGATGGAATACGACGCTACGGGCTTCCT
CCGTATAAAATCGAGATGCTCTTTCAAAAAAGATTGGATGAAGTGAATAAGGGAGCCCTGAA
AGAGCTGCCTGTTGTTATAAATCAGCTCCTCAGAGAGTTAATGAAATGATCCAGAAGAACA
CAAGGAAGATTCCTTGTCTTTTGAACAAAATTGCTGACTGTTAATAAGACTATACCCGAACCT
CTATTTTTGAAATTTCGATGGACACCTACGGGAAATCTTACTCAGGGAATATGGAATGAACTGC
TTGAGGAAAGCCTCTAAAACATTTTATAAGTACATTGACAAAGAACACACTATAATTAACAGGCACAGC
TGAATGGGGGGATATCTGCCGCTTTGGAACATGGAGGCTCCAAATATGGGAAAAATTCGATA
GTACACATTCACTACTACAAATCCTTTTCGAATTTACCCATCAAGAGCTGATGGGTTGAAGGAT
CTGAGAAATGTTGCGAGAAAGTGGGCTTCATCAGAGATCAGAAACATTTTGTAGTAGAGAATT
CAATTTATAGAAGAATGTTGCAAAAAGGACAAAGGACAAATACATAGGGATTTGGATTTGTGCC
TACAGGAAAGCTGACCTTTGAGATCAAGATAAAAGAGGACCTCCCTTGTGTCGTGGGACAGCTGAG
AAGAGGAGATGGAGCTGCGGTGATCTGGGAAATTTGTTTGCCTGTGTTTTGTTATGATACCTGGAG
AAAAACCTGGGAAATTAAAGACTTTTTGAAAGCTTTATAAGGAGAGATAGCATTTGGAATAACACCAT
TCCAGAGATGGCGTG
SEQ ID NO: 11

ATGGGTCTTGATCTTTTGAAGGAAAATATTGGGAAAGGCTTAAATATAGGAAATGGCC
TGAAGGATCCAAATGAGGAGACGAGGAGTGTGATGATAAAAAGATGATTTTGGACATTATAAA
AGAAGCCGTTCTCTCCTGTGAAATTCAAATAGATGAGTACGCGACGCTACGGCTTTCCCT
CCTTATAAATTCAGAGATTTCTTAATCTTTTCTCCAAAGAGGTGGATGAAATGATAAAGCGGAGCCCTGGAAA
GAGGCTGCTGGTTTTAAAATTACGACCTACGAGGATTTATGATGATCCAGAAGAAACA
CAAGGAAAGATTTTCTGCTCTTTGAAAACAATTTGGCTGATCCTTATAAAAGACTATATCCGAAACT
CTATTTTGAAATTCTGAATGAGCTGACGAAAATCTTACTCCGGAAAAATGGGAATGAACTGC
TGGAGGAAAGCTCTAAAAGTTTTAACAGATGAAATTGCCAAAAGAAGCACAATGTGATTATAAGGCCAGCC
TGATGAGGGGGGTATATCTGTCCCTGGAAAAATGGGAAAAATGGCATA
GGTCACATTGACTCTACACTACATCCATTTTCCGATTTACCCCATCAAGGAGCTGATGGSCTGCTGGAT
CTGGRAAATGTTGGGAAGGAAGTTGGSCTTCCAGATGRTCAAGAACRTGATGAGAAATT
CAATTATTAGAGAGATATGACAGAGAAAAAGGCAATTACATTACAGGTGATTTGGTGCC
TACAGAAAAAGCTGACCTTTGAGATCAGATAAAAATGGAACCTCTTGGTCGCTCCGGAAGCCGAGA
AAAGGGGGTTGAGCTGCGCAATCGGGAATTTTTGTCCTGGTTTTGGTGTTTATGATCCCTCTGAG
AAAACAGTGGAATAAGAGATCTTTTGGAGAAGGAGATAGCAATTTGAAAT
Figure 16

SEQ ID NO: 13

ATGGGTGTTGATCCCTTTGAAAGGAAAGAAATATTTGGGAAGAAGGCATTTATATAGGAAATGCGC
TTGAAAGCACCATAATGAGGGAGAGCTGGGAGAGTGTTGATCAAAGATGAGTTCTTGTGACATATAAA
AGAGCACGTTTCTCTCATGTTCGAATTTCCATATAGTGATACGCACGCCTCTCAAGGTTCTT
CCCTATAAATATCGAGATCTTTACTCCTCAAAAAGAAGTGGATGAGTAGATAAAGCGAGCCTGA
GAGGACTGCTGTGGTGTTATAAATCAAGTTCACTACGAGGAGTTAATGAGATTTGATCCAGAAGAA
CAGGAAAGATTCTTTGCTTTGGAAAACAAATTTGCTGATGTATATAAAGACTATCTCCGAAAC
CTATTTTTTGAAATCCTGAAATGACCTACGGAACAATCTCTCCGGAATAGGAATGACTG
CTTGAGGAAGCTCTAAAAGTTTAATAAGATCAAATTGACAAGAACACTGTGATTATAGGCCAAC
GCTGATTGGGGGGTTATATCTGCTCTTTGAAATAGTTGGGGTCCAAATGAGGAAATGCGATA
GTTCACATTCTACTACACTCTCCATTTGAAAAGGCTGCAGATGGTGTGCTGCTGGAT
CTGAGAATACTGGTGGGAAGAAATGTTGGGAGACTTCCAGATGATCAAGAACATTTGGATAGGAGAT
CTATTATAGAAGATATGCTCAAAAAAGAACAAAAAGACCAATTTACATAGGTGAGTCGCC
TACAGGAAAGCTGACCTGTGAAATGAAATATAAAGATGNGCTCTCTTTGCTGTTCGCCGAAGCCGA
AARGGGGTGGGAGCTGCGTGGACCAGAGATTTTGTTTCCCGTCTTGTGTGTATGATTCCTCTGGAG
AAAACAGTGGAAATAAGATCTTTTAGGATGAGGAGATAGCAATTGGAATA
ATGGCCAAAGTACTCCGAGCTGGAAAGGCGCGGCTCATATTCGAGGCGTCTACTGGGACGTGC
CTTCGAGGGAATATGGTGGAACACAATACGGCCGAAGATACCGGAGTGTACGATGCCGAAT
CTCCGCAATATGGATATCCCGCGCTGGAAGCAGCGGCTGAGGCGCTGACGATCGCCGAC
CGTCGACCTCCCTTGGACGCTGACTGCTAGAAGGACGAGGGTGGAGAAGCGGCTGCTGCTC
CAAGCACGGAGTCTGTGAACTGATAAACCACGCCCGCGCCTATGCGCTAAGATAGCTGAT
ATGTCACCAACCCCGCCCGCGGCCTGACCTGGAGTAGGGACGGATGGAACCCCCCTCGTGACGAAGAATACCT
GGACCGACTTCTCAANGTCGGGTGGAATACCGCCAACTACCTCGACTTCACCGGA
CGACGCTCCATGCGGGCGATCTCCCGATTTTAGGCTGAGATGCCCATATGCCCAGCAATGCGAAGGACG
CTGGAGCGCATGCTGCTGACTGCTAAGGCTATGGCTCCTGGGATGCTAGACGTGAA
CTGATGCGCGCCCTCACCCTACGACATGGCGGTGGAGATCTGCGGTGATGAGCAGCAGCTGCGTGG
GCATACGGAGCTGGTCGAAGGCTCTTGGCTGCTCCCGCTACTACAGAGTAGGAGGGCTCTTG
ACAAACAAATCCCTACAGGCGCTGTCCTGGCGCTCCCTACGAGCGCCGGGCAAGGCTGCTGGCCTGCGA
CCGTCGACCGTTAATCTTTGTTGAGACACACACCACGACACCGGATATAATGCGAACCAGAATCCA
GCCGCACTGCTGATCTTCCATACCTACGAGAGCGCCGGACAGATATCTACCGCGACTACGAGGAGT
GGCTCAACAGGATAAAGCTCAAGAAGCTGCTGCTGACATGGAACCTCGCCCGGAGAGACG
CGACATAGCTACTACGATACGAAGCTAGCATCTTGGCTGCGAGACGCTGCGGTGGAGAGGCCC
GGCTATACACCCTACCTACAGCTCGTGGACAGGGCGGAGGTGGTTATTGTGGCCGAA
GGCTGGCGGCACGTCCTCCACAGGATATCTGCTAACCCTGCGCGCTGACCATGGTGGAGACAGTACCGCT
CTCACAAGGCGCTGGTACAGGATTACCTCGGCGGTGGTAGAACAGTACGGTAC
CTCCGTGCGCTACTGCGGCGTGGGCTGAC
Figure 19

SEQ ID NO:16

MAKYSELEKGGVIMQAFYWVDPVQGGINWDTIYQKIFWYDAGISAIAWIFPSKKGMGAYSMGYD
FYDFTDLQGYSQKQTVBGSAQELVMINTAHAYGKVIADINGVAGRGGDLEWNPVNDYT
WTDFSKVASGKYTANYLDHPELHALDSGT GGYPDICHDKSWDQYVLWASQESYAYLTD
IDAWRFYVKGYA9PWVVCWLDNWWGGNAVGEYWDNVDAYLNWASYSSGAKVDFALYYKMDEAF
DNMIAPAVLSALSNGQTQVSRDFPKATFCFVANHDTDIWNKRFAYAFLTYEGGFTFFYRYED
WNKDKLKLNIWHENLAGSNDIVYDDNDELIFVRNGYGDKGLITYINLGGSKAGRNVYVF
FGACIHEYTVNLGGWVDKYVSSGWVTLEAPAYDPANGQYGSVWSYCGVG
GENES ENCODING CELLULASE

RELATED APPLICATIONS

[0001] This application is a divisional of U.S. application Ser. No. 14/389,339, which is the United States National Phase filing of International Application No.: PCT/US2013/030527, filed Mar. 12, 2013, which designated the United States, was published in English and claims priority of U.S. Provisional Applications 61/618,610, filed Mar. 30, 2012 and 61/704,368, filed Sep. 21, 2012. The contents of the aforementioned applications are hereby expressly incorporated by reference in their entirety.

FIELD OF THE INVENTION

[0002] Polynucleotide sequences encoding a cellulase are provided. In particular, the polynucleotide sequences may provide increased expression of a specific, thermostable, thermotolerant, pressure stable enzyme such as a cellulase.

SEQUENCE LISTING

[0003] This application is being filed electronically via the USPTO EFS-WEB server, as authorized and set forth in MPEP § 502.05 and this electronic filing includes an electronically submitted sequence listing; the entire content of this sequence listing is hereby incorporated by reference into the specification of this application. The sequence listing is identified on the electronically filed ASCII (.txt) text file as follows:

<table>
<thead>
<tr>
<th>File Name</th>
<th>Date of Creation</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2480-IN_SEQLISTING</td>
<td>Mar. 11, 2013</td>
<td>31.4 KB</td>
</tr>
</tbody>
</table>

BACKGROUND OF THE INVENTION

[0004] O-Glycosyl hydrolases (EC 3.2.1.-) are a widespread group of naturally-occurring enzymes that hydrolyze the glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety. The International Union of Biochemistry and Molecular Biology (IUBMB) enzyme nomenclature of glycosyl hydrolases (or glycohydrolases) is based principally on their substrate specificity and occasionally on their molecular mechanism (Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB), Accessed Oct. 24, 2011).

[0005] IUBMB Enzyme Nomenclature EC 3.2.1.4 has been designated for a subgroup of glycosylhydrolase-type enzymes termed “cellulas.” Other names used for enzymes belonging to this group include: endoglucanase, endo-1,4-beta-glucanase, carboxymethyl cellulase, and beta-1,4-glucanase. The reaction catalyzed by enzymes belonging to this group is the endo-hydrolysis of 1,4-beta-D-glycosidic linkages in cellulose, lichenin, and cereal beta-D-glucans (such as barley beta-glucan). Since the predominant activities of the disclosed cellulase of the present invention are the endo-hydrolysis of barley beta-glucan and carboxymethyl cellulose, it is appropriately ascribed the IUBMB Enzyme Nomenclature EC 3.2.1.4.

[0006] An alternative classification of glycosyl hydrolases is based on amino acid sequence similarities (Henrissat, B. Accessed at UniProt Oct. 26, 2011). According to this classification scheme, glycosyl hydrolases can be divided into more than 70 families. Based on a comparison of the primary amino acid sequence of the disclosed cellulase of the present invention with the sequences of other glycosyl hydrolases contained in public databases, the disclosed cellulase of the present invention may be assigned to glycosyl hydrolase Family 5. This family contains more than 20 endoglucanases (IUBMB Enzyme Nomenclature EC 3.2.1.4) whose predominant catalytic activity is the endo-hydrolysis of beta-1,4-glycosidic linkages in cellulosic substrates. Using this second way of classifying enzymes provides further support for the conclusion that the disclosed cellulase of the present invention should be ascribed the IUBMB Enzyme Nomenclature EC 3.2.1.4.

[0007] Cellulases are used for a variety of industrial and commercial purposes including but not limited to oil and gas exploration, food and beverage, alcohol production potable or fuel, e.g. brewing, ethanol, wine, flavor, fragrance, textile, detergents, paper, pulp, environmental, and agriculture, as well as in research purposes. (Rebecca S. Bryant, Erle C. Donaldson, Teh Fu Yen, George V. Chilingarian, Chapter 14 Microbial Enhanced Oil Recovery, In: Erle C. Donaldson, George V. Chilingarian and Teh Fu Yen, Editor(s), Developments in Petroleum Science, Elsevier, 1989, Volume 17(B):423-450) (M. Karmakar and R. R. Ray, 2011. Current Trends in Research and Application of Microbial Cellulases. Research Journal of Microbiology; 6:41-53.)

[0008] A typical corollary activity and expense to oil and gas discovery and drilling operations is the treatment of fluids used and or produced by such operations. For example, the drilling of wells, the washing and prepping of wells (“well completion”), hydraulic fracturing operations, and oil and gas processing, all typically produce thousands of gallons of a contaminated by-product fluid. Often the by-product fluids created by such operations are called “flowback fluids,” as the liquids typically flow back out of the wellbore to the surface. The by-product fluid, or flowback, must typically be treated for either disposal or reuse.

Need for More Efficient Means to Treat Flowback Fluids

[0009] As the treatment of flowback fluids in the gas discovery and drilling industry requires considerable resources and time, a need exists for efficient methods or compositions to treat flowback fluid. Additionally, as gas discovery and drilling operations typically require fresh (e.g. cleaned or filtered) fluids, a significant need exists for methods to treat flowback fluids to allow the re-use of such fluids for additional gas discovery and drilling operations.

SUMMARY

[0010] Enzymes are proteins that act as catalysts. Proteins are polymers of amino acids linked in dehydration reactions by peptide bonds. The identity of the amino acids and the order in which they are linked to form proteins determines a given protein’s activity. This order in which amino acids are assembled into proteins (the protein “sequence”) is ultimately determined by the sequence of a DNA strand which “encodes” the protein.

[0011] The three-nucleotide sequence that specifies a given amino acid to be assembled into a protein is called a “codon.” The 20 amino acids built into proteins are collectively encoded by 64 tri-nucleotide codon sequences. The series of codons which specifies a protein is called an “Open
Reading Frame.” An amino acid may be specified by as few as one or as many as six distinct codons. A change (or mutation) in the trinucleotide sequence of a codon that does not affect the amino acid specified is called a “silent” mutation.

As a result, there are many DNA sequences capable of encoding the same protein, because the DNA sequences differ from one another only through “silent” mutations. By altering one or more of the codons which encode a given protein, it may be possible to greatly increase the amount of protein which a gene produces without affecting the sequence of the protein that is encoded.

In some embodiments, the invention comprises SEQ ID NO:1. In some embodiments, the invention comprises the polynucleotide sequence of SEQ ID NO:1. In some embodiments, this sequence encodes a protein. In some embodiments, this protein is an enzyme having cellulase activity.

The improved nucleotide sequence disclosed herein is given as SEQ ID NO:1 and encodes a previously disclosed cellulase enzyme (SEQ ID NO:2) that was evolved from a parent cellulase enzyme isolated from a DNA library originating from *Thermotoga maritima* strain MS88. The disclosed cellulase of SEQ ID NO:2 is described in PCT Publication No. WO 2009/020459, as SEQ ID NO:9 of that reference (encoded by the polynucleotide SEQ ID NO:8 of the same publication, described herein as SEQ ID NO:3). In some embodiments, the invention comprises the polynucleotide sequence of SEQ ID NO:1, or fragments thereof. In some embodiments, these sequences encode a protein. In some embodiments, the protein is an enzyme having cellulase activity.

The invention comprises multiple nucleotide base changes with respect to SEQ ID NO:3. These changes are silent as to the encoded protein. The 14 base changes are set forth below. “Position” indicates the number of the nucleotide within the Open Reading Frame of SEQ ID NO:1, with the first nucleotide of the first codon numbered as 1. In the event that the Open Reading Frame of SEQ ID NO:1 is joined to another nucleic acid sequence at its 5′ end so that the Open Reading Frame extends beyond the 5′ end of SEQ ID NO:1, the “Position” will continue to refer to the bases as numbered from the 5′ end of the Open Reading Frame of SEQ ID NO:1. Similarly, if the Open Reading Frame of SEQ ID NO:1 is truncated so that the Open Reading Frame does not begin at the 5′ end of a sequence related to SEQ ID NO:1, the numbering system will continue to originate from the 5′ end of said sequence corresponding to the 5′ end of SEQ ID NO:1.


The base changes which distinguish SEQ ID NO:1 from prior reported sequences encoding the disclosed cellulase, collectively and individually, result in an Open Reading Frame which leads to a higher level of protein expression than previously employed nucleotide sequences encoding the same protein.


59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary nucleic acid used to practice the invention, including SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, and/or SEQ ID NO:15 over a region of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, or more residues, wherein these nucleic acids encode at least one polypeptide having a cellulase activity in particular, the genus based on the exemplary SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 and/or at least one polypeptide having any amylose activity in particular, the genus based on the exemplary SEQ ID NO:16.

[0027] The compositions and methods disclosed herein comprise use of isolated, synthetic, or recombinant polypeptides having a cellulase activity in particular, the genus based on the exemplary ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, and/or amylose activity in particular, the genus based on the exemplary SEQ ID NO:16.

[0028] In some aspects, the polypeptides used to practice this invention have an amylose or cellulase activity, which is thermostable. The polypeptide can retain an amylose or cellulase activity under conditions comprising a temperature range from about –100°C to about –80°C, about –80°C to about –60°C, about –60°C to about –40°C, about –40°C to about –20°C, about –20°C to about 0°C, about 0°C to about 37°C, about 37°C to about 4°C, about 4°C to about 15°C, about 15°C to about 25°C, about 25°C to about 37°C, about 37°C to about 45°C, about 45°C to about 55°C, about 55°C to about 70°C, about 70°C to about 75°C, about 75°C to about 85°C, about 85°C to about 90°C, about 90°C to about 100°C, about 100°C to about 110°C, about 110°C to about 120°C, about 120°C to about 130°C, about 130°C to about 140°C, about 140°C to about 150°C, about 150°C to about 160°C, about 160°C to about 170°C, about 170°C to about 180°C, about 180°C to about 190°C, about 190°C to about 200°C, about 200°C to about 210°C, about 210°C to about 220°C, about 220°C to about 230°C, about 230°C to about 240°C, about 240°C to about 250°C, about 250°C to about 260°C, about 260°C to about 270°C, about 270°C to about 280°C, about 280°C to about 290°C, about 290°C to about 300°C, or any temperature or range of temperatures in between or above.

[0029] In some aspects, the polypeptides used to practice this invention have an amylose, or cellulase activity, which is thermostable. The polypeptides can retain an amylose, or a cellulase activity after exposure to a temperature in the range from greater than 37°C to about 95°C or anywhere in the range from greater than 55°C to about 85°C. The polypeptide can retain an amylose, or a cellulase activity after exposure to a temperature in the range between about 100°C to about 120°C, about 120°C to about 140°C, about 140°C to about 160°C, about 160°C to about 180°C, about 180°C to about 200°C, about 200°C to about 220°C, about 220°C to about 240°C, about 240°C to about 260°C, about 260°C to about 280°C, about 280°C to about 300°C, or any temperature or range of temperatures in between or above.
an amino acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 and/or SEQ ID NO:16, or an enzymatically active fragment thereof, wherein such fragments are of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350 or more residues, or over the full length of the polypeptide. In some aspects, the polypeptide has cellulase or amylase activity.

[0032] Exemplary polypeptide or peptide sequences used to practice this invention include SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, and/or SEQ ID NO:16, including subsequences (enzymatically active fragments) thereof and variants thereof, e.g., including fragments of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350 or more residues, or over the full length of an enzyme.


[0034] In some embodiments, a polypeptide used to practice this invention, whether alone or with a “cocktail” disclosed herein, includes amylases that can catalyze the hydrolysis of polysaccharides comprising glucose monomers, such as starch (a polymer of glucose monomers joined by 1,4-alpha or 1,6-alpha linkages). In some aspects, the polypeptide has an amylase activity, e.g., an alpha amylase activity, endoamylase activity, or a glucoamylase activity; and the term “amylase” as used herein also includes enzyme activity which catalyzes the hydrolysis of a polysaccharide, e.g., a starch. Amyloses used to practice the invention include polypeptides having an alpha-amylase activity, a alpha-amylase activity, a glucoamylase activity, a 1,4-alpha-D-glucan glucohydrolase activity, an exoamylase activity, a glucan alpha-maltotetrahydrolase activity, a maltase activity, an isomaltase activity, a glucan 1,4, 1,4-glucosidase activity, a alpha-glucosidase activity, a sucrase activity or an agarase activity (e.g., a alpha-agarase activity). For example, an amylase used to practice this invention includes polypeptides having alpha-amylase activity, including the ability to hydrolyze internal alpha-1,4-glucosidic linkages in starch to produce smaller molecular weight maltodextrins. In some aspects, the alpha-amylase activity includes hydrolyzing internal alpha-1,4-glucosidic linkages in starch at random. An amylase used to practice this invention includes polypeptides having glucoamylase activity, such as the ability to hydrolyze glucose polymers linked by alpha-1,4- and alpha-1,6-glucosidic bonds. In some aspects, an amylase used to practice this invention includes polypeptides having glucoamylase activity, hydrolyzing internal alpha-1,4-glucosidic linkages to yield smaller molecular weight maltodextrins. An amylase used to practice this invention includes polypeptides having glucon 1,4-alpha-glucosidase activity, or 1,4-alpha-D-glucan glucohydrolase, commonly called glucoamylase but also called amyloglucosidase and alphamyelase that, in one aspect, releases alpha-D-glucose from 1,4-alpha-, 1,6-alpha- and 1,3-alpha-linked glucans. An amylase used to practice this invention includes polypeptides having exo-amylase activity.

[0035] In some embodiments, a polypeptide used to practice this invention, whether alone or with a “cocktail” disclosed herein, includes for example a cellulase or cellulas that can catalyze the hydrolysis of polysaccharides comprising glucose monomers, such as guar gum (a polymer of glucose monomers joined by 1,4-alpha or 1,6-alpha linkages). In some aspects, the polypeptide used to practice this invention, whether alone or with a “cocktail” disclosed herein, includes cellulase enzymes described herein, possess gluconase, e.g., endogluconase, mannanase, xylanase activity or a combination of these activities. In some aspects, the gluconase activity is an endogluconase activity (e.g., endo-1,4-beta-D-glucan 4-glucano hydrolase activity) and comprises hydrolysis of 1,4-beta-D-glycosidic linkages in cellulose, cellulose derivatives (e.g., carboxy methyl cellulose and hydroxyl ethyl cellulose) lichenin, beta-1,4 bonds in mixed beta-1,3 glucans, such as cereal beta-1,3 glucans and xylanoglucans and other plant material containing cellulosic parts. In alternative aspects, these glucanases, e.g., endogluconases, mannanases, xylanases have increased activity and stability, including thermostolerance or thermostability, at increased or decreased pHS and temperatures.

[0036] Examples of suitable polysaccharide substrates include galactomannan gums, guars, derivatized gums, cellulose and cellulose derivatives, starch, starch derivatizes, xanthan, derivatized xanthan, and mixtures thereof. Specific examples also include, but are not limited to, guar gum, guar gum derivative, locust bean gum, karaya gum, xanthan gum, cellulose, and cellulose derivatives, etc. Typical polymeric viscosifiers or gelling agents to which the disclosed enzymes may be directed include guar gum, hydroxypropyl guar,羧基甲基羟丙基 guar, hydroxyethyl cellulose, carboxymethyl hydroxyethyl cellulose, carboxymethyl cellulose, dialkyl carboxymethyl cellulose, etc. Other examples of polymers include, but are not limited to, polysiliconnons, serylglucans, dextrans and other types of polymers. In some embodiments, a polymer substrate is carboxymethyl hydroxypropyl guar. In some embodiments, a disclosed enzyme may also be effective in hydrolyzing biogums (e.g., succinoglycan biogums made from date syrup or sucrose). In some embodiments, a disclosed enzyme may be used to hydrolyze cellulose-containing or derivatized cellulose-containing polymers—typically, the enzymes attack glucosidic linkages of the cellulose backbone. The disclosed enzymes may be suitable for degrading the polymer into mostly monosaccharide units, in some cases, by specifically hydrolyzing the exo(1,4)-1,6-glucosidic and endo(1,4)-1,4-gluc-
cosidic linkages between monosaccharide units and the cellulose backbone in the (1,4)-\(\beta\)-D-glucosidic linkages of any cellulobiose fragments.

[0037] The enzyme-comprising compositions disclosed herein can comprise one polysaccharide-degrading enzyme as described herein, or can comprise a mixture (a “cocktail”) of one two, three, four, or more of any of the polysaccharide-degrading polypeptides described herein, including the genus based on SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 and/or SEQ ID NO:16. A composition used to practice the invention can comprise one, two, three or more polypeptides described herein, including the genus based on SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, and/or SEQ ID NO:16 and any combination of other enzymes, such as trypophanases or tyrosine decarboxylases, laccases, catalases, laccases, other cellulases, endoglucanases, endo-beta-1,4-laccases, amyloglucosidases, other glucosidases, glucose isomerases, glycosyltransferases, lipases, phospholipases, lipooxygenases, beta-laccases, endo-beta-1,3(4)-laccases, cutinases, peroxidases, amylases, xanthanases, glucomylases, pectinases, reductases, oxidases, decarboxylases, phenoloxidases, ligninases, pullulanases, arabinanases, hemicellulases, mannanases, xylolaccases, xylanases, pectin acetyl esterases, rhamnogalacturonan acetyl esterases, proteases, peptidases, proteinases, polygalacturonases, rhamnogalacturonases, galactanases, pectin lyases, transglucosaminases, pectin methyltransferases, other cellulolytic enzymes, and/or transglucosaminases.

[0038] The compositions and methods disclosed herein comprise use of isolated, synthetic or recombinant polypeptides comprising these polypeptides (e.g., the genus of polypeptides as described above), and a signal sequence. The signal sequence can be derived from another amylase, xanthanase, and/or glycosidase, e.g., cellulase or endoglucanase, cellulohydrolase, mannanase and/or beta-glucosidase enzyme or a noncellulase, e.g., non-endoglucanase, non-cellobiohydrolase, and/or non-beta-glucosidase enzyme (a heterologous) enzyme.

[0039] The compositions and methods disclosed herein comprise use of isolated, synthetic, or recombinant polypeptides not containing a signal sequence, or lacking all or part of a signal sequence, or comprising a heterologous signal sequence, such as a heterologous amylase, xanthanase, or glycosidase, or cellulase, e.g., endoglucanase, cellulohydrolase mannanase, and/or beta-glucosidase enzyme signal sequence or, non-amylase, non-xanthanase, or noncellulase, e.g., non-endoglucanase, non-cellobiohydrolase, and/or non-beta-glucosidase enzyme signal sequence.

[0040] The compositions and methods disclosed herein comprise use of isolated, synthetic, or recombinant chimeric proteins comprising a first domain comprising a signal sequence and at least a second domain comprising the genus of polypeptides as described above. The protein can be a fusion protein. The second domain can comprise several enzymes or activities. The enzyme can be a non-enzyme. The compositions and methods disclosed herein comprise use of isolated, synthetic, or recombinant chimeric proteins comprising the genus of polypeptides as described above and a signal peptide (SP), a prepro sequence and/or a catalytic domain (CD), and in an alternative embodiment, at least another domain comprising a heterologous polypeptide or peptide, wherein the heterologous polypeptide or peptide is not naturally associated with the signal peptide (SP), prepro sequence and/or catalytic domain (CD). In some aspects, the heterologous polypeptide or peptide is not an amylase, or xanthanase, or cellulase, e.g., endoglucanase, cellulohydrolase, mannanase and/or beta-glucosidase enzyme. The heterologous polypeptide or peptide can be amino terminal to, carboxy terminal to, or on both ends of the signal peptide (SP), prepro sequence and/or catalytic domain (CD).

[0041] In some aspects, an amylase and/or cellulase used to practice this invention can retain enzyme activity under conditions comprising about pH 6.5, pH 6.5, pH 5.5, pH 5, pH 4.5, pH 4.0, pH 3.5, pH 3.0 or less (more acidic); or, can retain activity under conditions comprising about pH 7.5 pH 8.0, pH 8.5, pH 9, pH 9.5, pH 10, pH 10.5, pH 11.0, pH 11.5, pH 12, pH 12.5 or more (more basic); or, can retain enzyme activity after exposure to conditions comprising about pH 6.5, pH 6, pH 5.5, pH 5, pH, 4.5, pH 4.0, pH 3.5, pH 3.0 or less (more acidic); or, can retain enzyme activity after exposure to conditions comprising about pH 7, pH 7.5 pH 8.0, pH 8.5, pH 9, pH 9.5, pH 10, pH 10.5, pH 11.0, pH 11.5, pH 12, pH 12.5 or more (more basic); or, can retain activity at under alkaline conditions. In certain aspects, the pH of the reaction is conducted in the range of about 3.0 to about 9.0. In other aspects, the pH is about 4.5 or the pH is about 7.5 or the pH is about 9. Reaction conditions conducted under alkaline conditions also can be advantageous, e.g., in some industrial applications of enzymes disclosed herein.

[0042] The invention provides protein preparations comprising any member of the several genuses of polypeptides (including peptides) described herein, wherein the protein preparation comprises a liquid, a solid or a gel; and any member of the several genuses of polypeptides (including peptides) used to practice this invention can be a heterodimer comprising a polypeptide as described herein, e.g., where the second member of the heterodimer can be a different amylase, or xanthanase, or cellulase, e.g., endoglucanase, cellulohydrolase, mannanase and/or beta-glucosidase enzyme, a different enzyme or another protein. In some aspects, the second domain can be a polypeptide and the heterodimer can be a fusion protein. In some aspects, the second domain can be an epitope or a tag. In some aspects, the invention provides homodimers comprising a polypeptide used to practice the invention.

[0043] The invention can be practiced using immobilized polypeptides (including peptides) having amylase and/or cellulase enzyme activity as described herein; and a polypeptide can have at least one additional (a second) domain. In some aspects, the polypeptide can be immobilized on a cell, a metal, a resin, a polymer, a ceramic, a glass, a microelectrode, a graphic particle, a bead, a gel, a plate, an array or a capillary tube.

[0044] The amylase and/or cellulase used to practice this invention can be prepared by expression of a polynucleotide encoding the enzyme in any organism, e.g., a bacterium, a yeast, a plant, an insect, a fungus and/or an animal. The organism can be, for example, a P. fluorescens, S. pombe, S. cerevisiae, Pichia pastoris, E. coli, Streptomycyes sp., Bacillus sp. or a Lactococci sp.

[0045] The amylase and/or cellulase enzyme, used to practice this invention can be formulated in any enzyme delivery matrix, e.g., comprising a thermostable recombinant enzyme; e.g., as an enzyme delivery matrix in the form
of pellets comprising a granulate carrier and a thermostable recombinant enzyme, wherein the pellets readily disperse the enzyme contained therein into aqueous media, and administering the enzyme delivery matrix into the desired environment, e.g., a flowback liquid.

[0046] The invention provides compositions and enzymes used in a variety of forms and formulations. For example, purified polypeptides can be used in enzyme preparations deployed in drilling or fracturing applications.


[0048] In another embodiment of the present invention, the invention comprises a nucleotide sequence from *Thermotoga maritima* having at least one mutation and having an increased expression level of a protein encoded by said nucleotide sequence compared to a *Thermotoga maritima* wild-type genomic sequence, wherein optionally, said mutation/s is silent.

[0049] In another embodiment of the present invention, the invention comprises a first nucleotide sequence encoding the polypeptide of SEQ ID NO:2 wherein said nucleotide sequence has been mutated with respect to a second nucleotide sequence encoding SEQ ID NO:2 such that the expression level of said protein is increased relative to that of said protein encoded by said second nucleotide sequence.


[0051] In another embodiment of the present invention, any of the proteins of the invention are expressed in bacterial expression systems, wherein the bacteria expression system is a gram-negative bacteria expression system, e.g., *Pseudomonas, E. coli*, *Ralstonia*, or *Caulobacter* expression system.

[0052] In another embodiment of the present invention, expression of the cellulase of the cellulase is produced at least 1.0 g/L, 2.0 g/L, 3.0 g/L, 4.0 g/L, 5.0 g/L, 6.0 g/L, 7.0 g/L, 8.0 g/L, 9.0 g/L, 10.0 g/L, 11.0 g/L, 12.0 g/L, 13.0 g/L, 14.0 g/L, 15.0 g/L, 16.0 g/L, 17.0 g/L, 18.0 g/L, 19.0 g/L, 20.0 g/L, 21.0 g/L, 22.0 g/L, 23.0 g/L, 24.0 g/L, 25.0 g/L, 26.0 g/L, 27.0 g/L, 28.0 g/L, 29.0 g/L, 30.0 g/L, 31.0 g/L, 32.0 g/L, 33.0 g/L, 34.0 g/L, or 35.0 g/L.

[0053] In another embodiment of the present invention, a cellulase of the present invention is combined with a second enzyme wherein the second enzyme is selected from the group consisting of: a lactase, a lipase, a protease, a catalase, a xylanase, a cellulase, a glucanase, a mannanase, an amylose, an amylase, an epoxide hydrolase, an esterase, phospholipase, transaminase, an amine oxidase, cellobiohydrolase, an ammonia lyase, or any combination thereof.

[0054] In another embodiment of the present invention, the invention comprises an isolated, recombinant, or synthetic nucleotide, having a nucleic acid sequence comprising SEQ ID NO:1, wherein the nucleic acid sequence encodes a polypeptide having a cellulase activity.

[0055] In another embodiment of the present invention, the invention comprises an isolated, recombinant, or synthetic nucleotide, comprising a nucleic acid sequence of SEQ ID NO:1, wherein the nucleic acid sequence encodes a polypeptide having a cellulase activity and the polypeptide comprises an amino acid sequence of SEQ ID NO:2, or an enzymatically active fragment thereof.

[0056] In another embodiment of the present invention, the invention comprises, an isolated, recombinant, or synthetic nucleic acid sequence comprising SEQ ID NO:1 that encodes a polypeptide having a cellulase activity, wherein the polypeptide comprises an amino acid sequence of SEQ ID NO:2 and the polypeptide is produced in recombinant *Pseudomonas fluorescens* expression systems.

[0057] In another embodiment of the present invention, the invention comprises a method for treating flowback fluids, used in or produced by oil or gas operations comprising: (a) providing an enzyme or enzyme treatment to a flowback fluid; (b) allowing the enzyme or enzyme treatment to degrade a polysaccharide- or starch-consuming material in the flowback fluid, wherein the enzyme or enzyme treatment is effective to break down or hydrolyze the polysaccharide- or starch-consuming material in the flowback fluid, wherein optionally the enzyme is a cellulase or an amylase. In further embodiment of the method for treating flowback fluids, the enzyme or enzyme treatment comprises and amylase wherein the amylase comprises a polypeptide having an amino acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to SEQ ID NO:16, or an enzymatically active fragment thereof. In further embodiment of the method for treating flowback fluids, the enzyme or enzyme treatment comprises and cellulase wherein wherein the cellulase comprises a polypeptide having an amino acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to SEQ ID NO:3, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12 and/or SEQ ID NO:14, or an enzymatically active fragment thereof.
In another embodiment of the present invention, the invention comprises, a composition comprising a polymeric viscosifier, a surfactant, a thermostabilizer, and an enzyme breaker comprising a wild-type cellulase derived from a hyperthermophilic bacterium or a mutated variant thereof. In further embodiment of the composition, the viscosifier is a guar gel comprising a linear guar, a cross-linked guar, or mixtures thereof. In further embodiment of the composition the enzyme breaker specifically hydrolyzes β-1,4 glycosidic bonds in the guar gel. In further embodiment of the composition the enzyme breaker does not specifically hydrolyze α-1,6 glycosidic bonds in the guar gel. In further embodiment of the composition the enzyme breaker retains its ability to hydrolyze β-1,4 glycosidic bonds in the guar gel at temperatures up to about 275°F. In further embodiment of the composition wherein the enzyme breaker retains its ability to hydrolyze β-1,4 glycosidic bonds in the guar gel at a pH of up to about 11. In further embodiment of the composition, the enzyme breaker has SEQ ID NO. 2. In further embodiment of the composition, the enzyme breaker is encoded by a polynucleotide having SEQ ID NO. 1. In further embodiment of any of the above compositions, the enzyme breaker is a mutated variant of the wild-type cellulase, and has a melting temperature that is at least 20°F greater than the melting temperature of the wild type cellulase at about pH 6.5 and at least 10°F greater than the melting temperature of the wild type cellulase at about pH 10.5. In further embodiment of any of the above compositions, further comprising an ester wherein optionally the ester is selected from the group comprising ethylacetate, 2-ethoxyethyl acetate, ethyl acetacetate, methylbenzoate, ethylformate, methylacetate, and dimethylphthalate.

In another embodiment, the present invention comprises an animal feed or animal feed additive comprising the polypeptide encoded by SEQ ID NO 1. In another embodiment the animal feed or animal feed additive assists or aids in the digestion of foodstuffs.

All publications, patents, patent applications, GenBank sequences and ATCC deposits, cited herein are hereby expressly incorporated by reference for all purposes in their respective entitites.

DETAILED DESCRIPTION

*Thermotoga maritima* is a thermophilic eubacteria characterized by its ability to grow in extreme salt concentrations (i.e., from 0.25% NaCl to 6.00% NaCl). *Thermotoga maritima* belongs to the order Thermotogales whose members are thermophilic, rod-shaped, anaerobic and gram-negative. The minimum temperature for growth is around 55°C, optimum is 80°-85°C, and maximum is about 90°C. In some embodiments, the minimum temperature is less than 55°C and the maximum temperature is greater than 90°C. These bacteria have slowly evolved from one of the deepest branches in the kingdom of eubacteria. Members of Thermotogales have been described as “wide-spread and cosmopolitan” (Huber, R. et al., 2006), thriving in active geothermal areas. *Thermotoga maritima* is closely related to the species *Thermotoga neapolitana*, *Thermotoga petrophila*, and *Thermotoga naphthophila*. Specimens of *Thermotoga maritima* have been obtained from sea floors in Vulcano, Italy; Ribera Quente and Sao Miguel Island, Azores; Sangeang Island, Indonesia; and Fiji Island. (Huber, R. et al., 2006).

Strain MSB8 was isolated from a geothermally heated marine sediment at Vulcano, Italy (Huber, 1986). The temperature at the collection site ranged from 70-100°C, with a pH of 6.5-7.0. The strain has been deposited at the Deutsche Sammlung von Mikroorganismen as DSM 3100 and at ATCC as ATCC 43589 (Huber, R. et al., 2006).

*Thermotoga maritima* strain MSB8 has been studied for its enzyme encoding genes due to the exceptional thermostability of the enzymes it produces. Liebl (Liebl, W. et al., 1996) has published an “Analysis of a Thermotoga maritima DNA fragment encoding two similar thermostable cellulosases, CelA and CelB, and characterization of the recombinant enzymes.” Additionally, genes for amyloglycotic enzymes (Bibel, M. et al., 1998), reverse gyrase (Bouthier de la Tour, C. et al., 1998), alpha-amylyase (Liebl, W. et al., 1997), alpha-glucuronidase (Ruile, P. et al., 1997), xylanase (Winterhalter, C. et al., 1995), beta-glucosidase (Liebl, W. et al., 1994), glucanotransferase (Liebl, W. et al., 1992) have been isolated and analyzed. A study by Bronnenmeier (Bronnenmeier, K. et al., 1995), “Purification of Thermotoga maritima enzymes for the degradation of cellulosic materials” has shown that these enzymes are of value for degrading cellulosate and xylan.

Expression Systems

In some embodiments, the DNA encoding the cellulase of the present invention may be introduced, either on a plasmid or stably transformed into the genome of, for example, any number of gram negative bacterial systems such as *E. coli*, *Pseudomonas* species such as *fluorescens*, *Pseudomonas putida*, *Pseudomonas aeruginosa*. *Ralstonia* species, or *Caulobacter* species. Similarly, the cellulase may be introduced into any number of gram positive bacterial expression systems such as *Bacillus* species such as *Bacillus subtilis*, *Bacillus megaterium*, *Bacillus brevis*, *Lactococcus* species such as *Lactococcus lactis*, *Lactobacillus* species, *Streptomyces* species such as *Streptomyces lividans*. Other gram negative, gram positive or unrelated eubacterial or archaeal expression systems may be used to express the cellulase.

Polypeptides used to practice this invention can be expressed in a microorganism using procedures known in the art. In other aspects, the polypeptides used to practice this invention can be immobilized on a solid support prior to use in the methods disclosed herein. Methods for immobilizing enzymes on solid supports are commonly known in the art, for example J. Mol. Cat. B: Enzymatic 6 (1999) 29-39; Chivata et al. Biocatalysis: Immobilized cells and enzymes, J Mol. Cat. 37 (1986) 1-24: Sharma et al., Immobilized Biomaterials Techniques and Applications, Angew. Chem. Int. Ed. Engl. 21 (1982) 837-15 54: Laskin (Ed.), Enzymes and Immobilized Cells in Biotechnology. Polypeptides used to practice this invention can be recombinantly expressed. Polypeptides used to practice this invention can include recombinant proteins encoded by a genus of nucleic acids based on SEQ ID NO 1, SEQ ID NO 5, SEQ ID NO 7, SEQ ID NO 9, SEQ ID NO 11, SEQ ID NO 13 and/or SEQ ID NO 15 (to encode, e.g., SEQ ID NO 2, SEQ ID NO 5, SEQ ID NO 7, SEQ ID NO 9, SEQ ID NO 12, SEQ ID NO 14 and/or SEQ ID NO 16).
In some embodiments, SEQ ID NO:1 is used to direct an increased level of expression in a number of systems in which the disclosed cellulase protein may be expressed. SEQ ID NO:1 may be introduced into any number of expression systems to express the disclosed cellulase at an improved accumulation level. For example, SEQ ID NO:1 may be introduced, either on a plasmid or stably transformed into the genome of, for example, any number of gram negative bacterial systems such as E. coli, Pseudomonas species such as fluorescent, Pseudomonas putida, Pseudomonas aeruginosa,Ralstonia species, or Caulobacter species. Similarly, SEQ ID NO:1 may be introduced into any number of gram positive bacterial expression systems such as Bacillus species such as Bacillus subtilis, Bacillus megaterium, Bacillus brevis, Lactococcus species such as Lactococcus lactis, Lactobacillus species, Streptomyces species such as Streptomyces lividus. Other gram negative, gram positive or unrelated eubacterial or archaeal expression systems may be used to express SEQ ID NO:1. In a further embodiment, SEQ ID NO:1 may be introduced into any number of eukaryotic expression systems such as Saccharomyces, Schizosaccharomyces pombe, Pichia pastoris, and Hansenula polymorpha.

More specifically, SEQ ID NO:1 may be introduced into a plasmid to direct its expression. Plasmids which SEQ ID NO:1 may be introduced include, for example, E. coli expression vectors of the families pQE, pET, and pASK; Pseudomonas expression vectors of the families pCN51 LI8, RSFI10, pWZ112T, and pMYC; Bacillus expression vectors of the families pBAX, pHT101, and pHTIS1525; Streptomyces expression vectors of the families pl6021 and pJ2460; and Lactococcus: expression vectors of the families pNZ9530 and pNZ8148, for example. These examples are for demonstrative purposes and do not represent a complete set of vectors in which the polynucleotide sequence of SEQ ID NO:1 can be expressed.

In some embodiments, the expression system could be any Pseudomonas fluorescens expression system known in the art, for example, the Pseudomonas fluorescens expression system that is commercially available from Dow Global Technologies Inc., strain DC454 (US Patent PUB. APP. NO. 20050130160 and US Patent PUB. APP. NO. 20050186666). A nucleic acid sequence encoding the cellulase enzyme or polypeptide is inserted either in the pMYC vector (Dow Global Technologies Inc., US Patent PUB. APP. NO. 20050130160) or in the pDOW1169 vector (Dow Global Technologies Inc., US Patent PUB. APP. NO. 20080058282) and then introduced into the Pseudomonas fluorescens host by electroporation. Those skilled in the art will know alternative vectors that can be used as embodiments of this invention.

In some embodiments, the cellulase will be expressed at least at the following expression levels: 1.0 g/L, 2.0 g/L, 3.0 g/L, 4.0 g/L, 5.0 g/L, 6.0 g/L, 7.0 g/L, 8.0 g/L, 9.0 g/L, 10.0 g/L, 11.0 g/L, 12.0 g/L, 13.0 g/L, 14.0 g/L, 15.0 g/L, 16.0 g/L, 17.0 g/L, 18.0 g/L, 19.0 g/L, 20.0 g/L, 21.0 g/L, 22.0 g/L, 23.0 g/L, 24.0 g/L, 25.0 g/L, 26.0 g/L, 27.0 g/L, 28.0 g/L, 29.0 g/L, 30.0 g/L, 31.0 g/L, 32.0 g/L, 33.0 g/L, 34.0 g/L, 35.0 g/L, or more.

Nucleic Acid

The invention provides isolated, synthetic, or recombinant nucleic acids comprising sequences completely complementary to the nucleic acid sequences disclosed herein (complementary (non-coding) and coding sequences also hereinafter collectively referred to as nucleic acid sequences disclosed herein).

The invention provides isolated, synthetic, or recombinant nucleic acids comprising a nucleic acid encoding at least one polypeptide having a cellulytic activity, wherein the nucleic acid comprises a sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity (homology) to an exemplary nucleic acid disclosed herein, including the sequence of SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, and/or SEQ ID NO:15. For example, the invention provides isolated, synthetic or recombinant nucleic acids comprising a nucleic acid sequence SEQ ID NO:1 (an exemplary polynucleotide sequence of this invention). The invention provides isolated, synthetic, or recombinant nucleic acids encoding a polypeptide comprising a sequences as set forth in SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 and/or SEQ ID NO:16 (exemplary polypeptide sequences of this invention), and enzymatically active fragments thereof.

Polypeptide

Polyepptides and peptides disclosed herein are isolated, synthetic, or recombinant polypeptides. Peptides and proteins can be recombinantly expressed in vitro or in vivo. The peptides and polypeptides disclosed herein can be made and isolated using any method known in the art. Polypeptides and peptides disclosed herein can also be synthesized, whole or in part, using chemical methods well known in the art. For example, cellulase polypeptides can be produced in a standard recombinant expression system (as described herein), chemically synthesized, or purified from organisms in which they are naturally expressed.

The invention provides isolated, synthetic, or recombinant polypeptides having cellulytic activity comprising an amino acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or has 100% (complete) sequence identity to an exemplary amino acid sequence disclosed herein (e.g., SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 and/or SEQ ID NO:16), or an enzymatically active fragment thereof.
The invention provides isolated, synthetic, or recombinant polypeptides comprising a sequence as set forth in SEQ ID NO: 2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 and/or SEQ ID NO:16, and enzymatically active fragments thereof, and variants thereof.

In alternative embodiments, the invention provides polypeptides (and the nucleic acids that encode them) having cellulolytic activity but lacking a signal sequence, a prepro domain, a dockerin domain, and/or a carbohydrate binding module (CBM); and in one aspect, the carbohydrate binding module (CBM) comprises, or consists of, a cellulose binding module, a lignin binding module, a xylan binding module, a xylose binding module, a mannan binding module, a xyleoglucan-specific module, and/or a arabinofuranoside binding module.

In alternative embodiments, the invention provides polypeptides (and the nucleic acids that encode them) having a cellulolytic activity further comprising a heterologous sequence; and in one aspect, the heterologous sequence comprises, or consists of, a sequence encoding: (i) a heterologous signal sequence, a heterologous carbohydrate binding module, a heterologous dockerin domain, a heterologous catalytic domain (CD), or a combination thereof; (ii) the sequence of (i), wherein the heterologous signal sequence, carbohydrate binding module or catalytic domain (CD) is derived from a heterologous enzyme; or, (iii) a tag, an epitope, a targeting peptide, a cleavable sequence, a detectable moiety or an enzyme; and in one aspect, the heterologous carbohydrate binding module (CBM) comprises, or consists of, cellulose binding module, a lignin binding module, a xylan binding module, a xylose binding module, a mannan binding module, a xyleoglucan-specific module and/or a arabinofuranoside binding module; and in one aspect, the heterologous signal sequence targets the encoded protein to a vacuole, the endoplasmic reticulum, a chloroplast or a starch granule.

Enzymatic Activity

The enzymatic hydrolysis of pNP-β-D-lactopyranoside by the disclosed cellulase can be used as a measure of activity of an enzyme disclosed herein such as SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12. The liberation of p-nitrophenol can be followed spectrophotometrically at 405 nm. The increase in absorbance at 405 nm can be converted to umoles of p-nitrophenol by using a standard absorbance at those defined conditions. One unit of activity is defined as the quantity of enzyme required to liberate 0.42 umole of p-nitrophenol from 2 mM pNP-β-D-lactopyranoside during one minute at pH 7.00 and 80°C. (Advances in Carbohydrate Chemistry and Biochemistry, Academic Press, 1999)

Thermostability

Thermotolerance

In some aspects, the recombinant nucleic acid of the present invention encodes a polypeptide having a cellulolytic activity that is thermostolerant. For example, a polypeptide disclosed herein, SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 and/or SEQ ID NO:16, or the variant evolved enzymes disclosed herein can be thermostolerant. In some aspects, the cellulolytic activity is thermostolerant, e.g., wherein the polypeptide retains cellulolytic activity after exposure to a temperature in the range from greater than 37°C to about 95°C, or between about 55°C to about 85°C, or between about 70°C to about 75°C, or between about 70°C to about 95°C, or between about 90°C to about 95°C, or between about 95°C to about 105°C, or between about 95°C to about 110°C. In some aspects, wherein the polypeptide can retain binding and/or enzymatic activity, e.g., cellulolytic activity, under conditions comprising 1°C to about 5°C, between about 5°C to about 15°C, between about 15°C to about 25°C, between about 25°C to about 37°C. In some aspects polypeptides disclosed herein can retain binding and/or enzymatic activity, e.g., cellulolytic activity, under conditions comprising 90°C, 91°C, 92°C, 93°C, 94°C, 95°C, 96°C, 97°C, 98°C, 99°C, 100°C, 101°C, 102°C, 103°C, 103.5°C, 104°C, 105°C, 107°C, 108°C, 109°C, or 110°C, or more. In some embodiments, the thermostable polypeptides according to the invention retains activity, e.g., a cellulolytic activity at a temperature in the ranges described above, under acidic conditions comprising about pH 6.5, pH 6, pH 5.5, pH 5, pH 4.5 or pH 4 or less (more acidic), or retain a cellulolytic activity after exposure to acidic conditions comprising about pH 6.5, pH 6, pH 5.5, pH 5, pH 4.5 or pH 4 or less (more acidic); or, retain activity under basic conditions comprising about pH 7, pH 7.5, pH 8, pH 8.5, pH 9, pH 9.5, pH 10, pH 10.5, pH 11, pH 11.5, pH 12, or more (more basic) or, retain a cellulolytic activity after exposure to basic conditions comprising about pH 7, pH 7.5, pH 8, pH 8.5, pH 9, pH 9.5, pH 10, pH 10.5, pH 11, pH 11.5, pH 12, or more (more basic).

Thermotolerance

In some aspects, the recombinant nucleic acid of the present invention encodes a polypeptide having a cellulolytic activity that is thermostolerant. For example, a polypeptide disclosed herein, SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 and/or SEQ ID NO:16, or the variant evolved enzymes disclosed herein can be thermostolerant. In some aspects, the cellulolytic activity is thermostolerant, e.g., wherein the polypeptide retains cellulolytic activity after exposure to a temperature in the range from greater than 37°C to about 95°C, or between about 55°C to about 85°C, or between about 70°C to about 75°C, or between about 70°C to about 95°C, or between about 90°C to about 95°C, or between about 95°C to about 105°C, or between about 95°C to about 110°C. In some aspects, wherein the polypeptide can retain binding and/or enzymatic activity, e.g., cellulolytic activity, under conditions comprising 1°C to about 5°C, between about 5°C to about 15°C, between about 15°C to about 25°C, between about 25°C to about 37°C. In some aspects polypeptides disclosed herein can retain a cellulolytic activity under acidic conditions comprising about pH 6.5, pH 6, pH 5.5, pH 5, pH 4.5 or pH 4 or less (more acidic), or retain a cellulolytic activity after exposure to acidic conditions comprising about pH 6.5, pH 6, pH 5.5, pH 5, pH 4.5 or pH 4 or less (more acidic); or, retain activity under basic conditions comprising about pH 7, pH 7.5, pH 8, pH 8.5, pH 9, pH 9.5, pH 10, pH 10.5, pH 11, pH 11.5, pH 12, or more (more basic).
Cellulosic Digestion

In some aspects, the compositions and methods disclosed herein are used in the enzymatic digestion of biomass and can comprise use of many different enzymes, including the cellulases and hemicellulases. Cellulases used to practice the invention can digest cellulose to glucose. In some aspects, compositions used to practice the invention can include mixtures of enzymes, e.g., xylanases, xylosidases (e.g., β-xylosidases), cellobiohydrolases, and/or arabinofuranosidases, or other enzymes that can digest hemicelluloses, cellulose, and lignocellulosic material, to fermentable sugars and/or to monomer sugars.

Enzymes, e.g., endoglucanases, disclosed herein are used to digest cellulose or any beta-1,4-linked glucon-comprising synthetic or natural material, including those found in any plant material. Enzymes, e.g., endoglucanases, disclosed herein are used as commercial enzymes to digest cellulose from any source, including all biological sources, such as plant biomasses, e.g., corn, grains, grasses (e.g., Indian grass, such as Sorghastrum nutans; or, switch grass, e.g., Panicum species, such as Panicum virgatum), or, woods or wood processing byproducts, e.g., in the wood processing, pulp and/or paper industry, in textile manufacture and in household and industrial cleaning agents, and/or in biomass waste processing.

Dietary

In some embodiments, the cellulase of the present invention may be used to pre-treat, modify, or digest a food, food additive, or dietary supplement for animals or human beings. In some embodiments, the cellulase of the present invention may be used as a food, food additive, or dietary supplement for animals or human beings. In some aspects, the cellulase will treat or will act as a prophylaxis for digestive disorders. In some aspects of the present invention the cellulase will alter or enhance digestion. In some aspects of the present invention the cellulase will enhance, alter, or aid in the digestion of foodstuffs. In further aspect disclosed herein the cellulase will enhance, aid, or alter the nutrient value of foodstuffs. In a further aspect, the cellulase is active in the digestive tract, e.g., in a stomach and/or intestine, for example.

In some embodiments, the cellulase disclosed herein may be used as an animal feed or an animal feed additive. In some embodiments the thermostability and or thermotolerance of the cellulase allows for the formation of pellets without the need for a secondary agent such as salt or wax. An animal feed comprising a cellulase can be provided to an animal in any formulation known to those skilled in the art. Examples of animal feed formulations include, but are not limited to a delivery matrix, a pellet, a tablet, a gel, a liquid, a spray, ground grain, or a powder.

The invention provides edible enzyme delivery matrix comprising a thermostable recombinant cellulase enzyme, e.g., a polypeptide disclosed herein. The invention provides methods for delivering a cellulase supplement to an animal, the method comprising: preparing an edible enzyme delivery matrix in the form of pellets comprising a granulate edible carrier and a thermostable recombinant cellulase enzyme, wherein the pellets readily disperse the cellulase enzyme contained therein into aqueous media, and administering the edible enzyme delivery matrix to the animal. The recombinant cellulase enzyme can comprise a polypeptide disclosed herein. The granulate edible carrier can comprise a carrier selected from the group consisting of a grain germ, a grain germ that is spent of oil, a hay, an alfalfa, a timothy, a soy hull, a sunflower seed meal and a wheat middl. The edible carrier can comprise grain germ that is spent of oil. The cellulase enzyme can be glycosylated to provide thermostability at pelleting conditions. The delivery matrix can be formed by pelleting a mixture comprising a grain germ and a cellulase. The pelleting conditions can include application of steam. In some embodiments, the pelleting conditions can comprise application of a temperature in excess of about 80°C, for about 5 minutes and the enzyme retains a specific activity of at least 350 to about 900 units per milligram of enzyme.

Methods of Making Ethanol

The invention provides methods for making ethanol comprising contacting a starch-comprising composition with a polypeptide having a cellulosytic activity, such as the enzyme of SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, and/or SEQ ID NO:14 wherein the polypeptide has a sequence disclosed herein, or the polypeptide is encoded by a nucleic acid comprising a sequence disclosed herein, or an enzymatically active fragment thereof. The invention provides compositions comprising a starch and a polypeptide having a cellulosytic activity, wherein the polypeptide has a sequence disclosed herein, or the polypeptide is encoded by a nucleic acid comprising a sequence disclosed herein, or an enzymatically active fragment thereof.

Brewing and Fermenting

The invention provides methods of brewing (e.g., fermenting) beer comprising the cellulase disclosed herein. In one exemplary process, starch-containing raw materials are disintegrated and processed to form a malt. An enzyme disclosed herein is used at any point in the fermentation process. The cellulase disclosed herein can be used in the brewing industry for the degradation of beta-glucans. In some aspects, the cellulases disclosed herein are used in the brewing industry for the clarification of the beverage. Enzymes disclosed herein can be used in the beverage industry in improving filterability of wort or beer, as described, e.g., in U.S. Pat. No. 4,746,517.

In some aspects, the cellulase disclosed herein is used in mashing and conversion processes. In the brewing and fermentation industries, mashing and conversion processes are performed at temperatures that are too low to promote adequate degradation of water-soluble glucans, mannan, arabinoxylans or xylans, or other polysaccharides. These polymers form gummy substrates that can cause increased viscosity in the mashing wort, resulting in longer mash run-off, residual haze and precipitates in the final beer product due to inefficient filtration and low extraction yield.

In some aspects, the cellulase disclosed herein are used in malhouse operations. e.g., glucanase is added to the process water, to shorten germination times and/or to encourage conversion of poor quality barley to acceptable mashing. In some aspects, enzymes disclosed herein are used
for mashing, e.g., they are added to increase wort filterability and/or improve lautering (separating the wort from the mash). In some aspects, enzymes disclosed herein are used in the fermentor and/or settling tank to, e.g., assist in haze clearing and/or to improve filtration. In some aspects, enzymes disclosed herein are used in adjunct brewing, e.g., a gluconase disclosed herein is added to breakdown glucans, mannan, arabinoxylans or xylans, or other polysaccharides from barley, wheat, and/or other cereals, including glycans in malt. In some aspects, enzymes disclosed herein are used in malt brewing, e.g., a gluconase disclosed herein is added to modify poor malts with high glucon content.

[0090] The cellulase disclosed herein can be used in any beer or alcoholic beverage producing process, as described, e.g., in U.S. Pat. Nos. 5,762,991; 5,536,650; 5,405,624; 5,021,246; 4,788,066, each of which is hereby incorporated by reference in its entirety.

Treating Foods and Food Processing

[0091] The cellulases disclosed herein have numerous applications in food processing industry. For example, in one aspect, the enzymes disclosed herein are used to improve the extraction of oil from oil-rich plant material, e.g., oil-rich seeds, for example, soybean oil from soybeans, olive oil from olives, rapeseed oil from rapeseed and/or sunflower oil from sunflower seeds.

[0092] The cellulase disclosed herein can be used for separation of components of plant cell materials. For example, enzymes disclosed herein can be used in the separation of glucan-rich material (e.g., plant cells) into components. In some aspects, enzymes disclosed herein can be used to separate glucan-rich or oil-rich crops into valuable protein and oil and hull fractions. The separation process may be performed by use of methods known in the art.

[0093] The cellulase disclosed herein can be used in the preparation of fruit or vegetable juices, syrups, extracts and the like to increase yield. The enzymes disclosed herein can be used in the enzymatic treatment (e.g., hydrolysis of glucan-comprising plant materials) of various plant cell wall-derived materials or waste materials, e.g. from cereals, grains, wine or juice production, or agricultural residues such as vegetable hulls, bean hulls, sugar beet pulp, olive pulp, potato pulp, and the like. The enzymes disclosed herein can be used to modify the consistency and appearance of processed fruit or vegetables. The enzymes disclosed herein can be used to treat plant material to facilitate processing of plant material, including foods, facilitate purification or extraction of plant components. The cellulase disclosed herein can be used to improve feed value, decrease the water binding capacity, improve the degradability in waste water plants and/or improve the conversion of plant material to ensilage, and the like. The cellulase disclosed herein can also be used in the fruit and brewing industry for equipment cleaning and maintenance.

Detergent Compositions

[0094] The invention provides detergent compositions comprising one or more polypeptides disclosed herein and methods of making and using these compositions. The invention incorporates all methods of making and using detergent compositions, see, e.g., U.S. Pat. Nos. 6,413,928; 6,399,561; 6,365,561; 6,380,147, each of which is hereby incorporated by reference in its entirety. The detergent compositions can be a one and two part aqueous composition, a non-aqueous liquid composition, a cast solid, a granular form, a particulate form, a compressed tablet, a gel and/or a paste and a slurry form. The invention also provides methods capable of a rapid removal of gross food soils, films of food residue and other minor food compositions using these detergent compositions. Enzymes disclosed herein can facilitate the removal of starchy stains by means of catalytic hydrolysis of the starch polysaccharide. Enzymes disclosed herein can be used in dishwashing detergents in textile laundering detergents. The actual active enzyme content depends upon the method of manufacture of a detergent composition and is not critical, assuming the detergent solution has the desired enzymatic activity. In some aspects, the amount of glucoamylase present in the final solution ranges from about 0.001 mg to 0.5 mg per gram of the detergent composition. The particular enzyme chosen for use in the process and products of this invention depends upon the conditions of final utility, including the physical product form, use pH, use temperature, and soil types to be degraded or altered. The enzyme can be chosen to provide optimum activity and stability for any given set of utility conditions. The detergents disclosed herein can comprise, for example, anionic, semi-polar nonionic, or zwitterionic surfactants; or, mixtures thereof.

[0095] The present invention provides cleaning compositions including detergent compositions for cleaning hard surfaces, detergent compositions for cleaning fabrics, dishwashing compositions, oral cleaning compositions, denture cleaning compositions, and contact lens cleaning solutions. In some aspects, the invention provides a method for washing an object comprising contacting the object with a polypeptide disclosed herein under conditions sufficient for washing a polypeptide disclosed herein may be included as a detergent additive. The detergent composition disclosed herein may, for example, be formulated as a hand or machine laundry detergent composition comprising a polypeptide disclosed herein. A laundry additive suitable for pre-treatment of stained fabrics can comprise a polypeptide disclosed herein. A fabric softener composition can comprise a polypeptide disclosed herein. Alternatively, a polypeptide disclosed herein can be formulated as a detergent composition for use in general household hard surface cleaning operations.

Oil and Gas Exploration and Clean-Up

[0096] To increase the productivity of oil and gas wells and shale gas reservoirs, a highly specialized technique called “hydraulic fracturing” is being increasingly utilized. In a typical hydraulic fracturing operation, large volumes of guar-based fluid (in gel form and referred to as “fracturing fluid”) are pumped into the wellbore under very high hydrostatic pressure. The pressurized fluid creates new fissures and fractures in the formation surrounding the wellbore. The sand particles contained in the fracturing fluid move and settle into the newly-created fractures and function to prop these channels open thus increasing oil and gas flow. Once the sand is deposited into the fractures, the gel has to be degraded (i.e., broken down) and brought back up to the surface so as to remove any blockage to the flow of oil or gas. Industry uses viscosity breakers (such as oxidizers, acids, or enzymes) to degrade the fracturing fluid and to remove any solid gel residue from the fissures and fractures.
During typical oil and gas drilling operations, liquid is pumped through the drill shaft and exits above the drill bit, such liquid is commonly referred to as “drilling fluid.” The drilling fluid serves to cool the bit, add pressure to the bit, lubricate the drill bit, and to remove debris away from the drilling site. Drilling debris is carried back to the surface by the fluid as it circulates back to the surface outside the drill shaft. The drilling fluid carrying the debris is often referred to as “mud,” “sludge,” or “flowback.”

Common materials found in the mud, sludge, or flowback are rock and sand, and a variety of hydrocarbons such as oil and petroleum present in the drilling fluid. The mud or sludge often has a high salt content dependent upon where the drilling takes place. The salt content of the drilling fluid may often be near or even higher than an average salinity found in the ocean (approximately 35 parts per thousand). Furthermore, the mud or sludge has been found to contain toxins and heavy metals which also contaminate the sludge.

The average drilling process can generate 300,000 barrels of mud, sludge, flowback per day over a two week period, or the equivalent of 4,200,000 barrels of mud or sludge for each drilled well. The mud, sludge, or flowback generated in drilling processes is typically shipped, temporarily stored, treated, and/or disposed of in an insertion well.

In some embodiments, the disclosed cellulase will be used as a high temperature viscosity breaker to enhance oil and gas operations. More specifically, the disclosed cellulase of the present invention will be applied to a fracturing fluid when hydraulic fracturing is performed in oil or gas wells.

The enzyme encoded by SEQ ID NO:1, as well as cellulases and amylases encoded by other polynucleotides disclosed herein, such as SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, and/or SEQ ID NO:15, or obtainable by methods disclosed herein, may potentially be used to hydrolyze a broad spectrum of polysaccharides—many of which are useful in oil and gas drilling, fracturing and well clean-up operations. The disclosed cellulases exhibit broad spectrum β-glucosidase activity, e.g., against guar, hydroxypropyl guar, carboxymethyl guar, carboxymethyl hydroxypropyl guar, carboxymethyl cellulose, barley β-glucan, and locust bean gum. The enzyme activity pattern is preferably both endo and exo, allowing effective reduction in the viscosity of polysaccharides, e.g., guar and derivatized guar solutions, by cleaving within long polysaccharide chains and also by cleaving disaccharide units from the ends of the polymers. Besides the aforementioned polysaccharides, other substrates of the disclosed enzymes include those capable of forming linear or cross-linked gels. Examples of suitable polysaccharide substrates include starch and/or guar gum, guar, derivatized guar, cellulose and cellulose derivatives, starch, starch derivatives, xanthan, derivatized xanthan and mixtures thereof. Specific examples also include, but are not limited to, guar gum, guar gum derivative, locust bean gum, karaya gum, xanthan gum, cellulose and cellulose derivatives, etc.

Typical polymers or gelling agents to which the disclosed enzymes may be directed include guar gum, hydroxypropyl guar, carboxymethyl hydroxypropyl guar, hydroxyethyl cellulose, carboxymethyl hydroxyethyl cellulose, carboxymethyl cellulose, dialkyl carboxymethyl cellulose, etc. Other examples of polymers include, but are not limited to, phosphonamides, scleroglucans, dextran and other types of polymers. In some embodiments, a polymer substrate is carboxymethyl hydroxypropyl guar. In some embodiments, a disclosed enzyme may also be effective in hydrolyzing biogums (e.g., succinoglycan biogums made from date syrup or sucrose). In some embodiments, a disclosed enzyme may be used to hydrolyze cellulose-containing or derivatized cellulose-containing polymers—typically, the enzymes attack glucosidic linkages of the cellulose backbone. The disclosed enzymes may be suitable for degrading the polymer into mostly monosaccharide units, in some cases, by specifically hydrolyzing the exo(1,4)-β-D-glucosidic and endo(1,4)-β-D-glucosidic linkages between monosaccharide units and the cellulose backbone in the (1,4)-β-D-glucosidic linkages of any cellulose fragments.

In each fracturing job that uses the disclosed cellulases, field operators will generally first perform an enzyme dose optimization study in an industrial lab. Such studies may include dilution of the cellulase to a concentration of 10–400 ppm and mixed with linear or cross-linked guar gum (25–60 lb/1,000 gal). Depending on the application conditions, guar gum may be cross-linked using a cross-linker, especially for wells where higher temperature, pressure, and pH conditions are present. The enzyme dose information resulting from such optimization studies may then be used in the actual fracturing job.

The unique activity of the disclosed cellulase allows for the hydrolysis of guar-based fracturing fluids in a smooth and controlled manner in deep wells, where high temperature and high pH conditions are present. Compared to chemical breakers, the disclosed cellulase of the present invention provides a non-corrosive and environmentally benign alternative to the harsh and non-selective chemical breakers.

In some embodiments of the present invention the cellulase may be used to treat, clean, or alter fluids used in oil and gas exploration activities. In a further aspect the cellulase of this invention will treat or alter the fluids, in part, or completely, so that the fluids may be used again, or recycled, for use in additional oil and gas exploration activities or to be disposed of in an environmentally friendly way.

The invention provides compositions and methods of using polysaccharide degrading enzymes to treat flowback from oil and gas exploration and drilling operations. In some aspects, the compositions and methods disclosed herein are used to degrade polysaccharides, which may include starch and/or guar gum, present in flowback fluids, by adding polysaccharide-degrading enzymes to the flowback fluids.

In some embodiments, the enzymes used in the method are a cellulase, or an amylase, or a combination thereof to treat flowback produced by oil and gas exploration and drilling operations.

In some embodiments, the compositions disclosed herein, including amylases and cellulases, as described herein, are added to a flowback fluid. In some embodiments the compositions and methods disclosed herein include an environmental trigger step, (e.g., pH, salinity, or mechanical treatment system step) which activates the compositions disclosed herein, including the amylases and cellulases described herein.

The enzyme comprising compositions disclosed herein can be formulated in a variety of forms, e.g., as
liquids, gels, pills, tablets, sprays, powders, pellets or encapsulated forms, including nanoencapsulated forms.

Well Completion

[0109] The walls of oil and gas formations are exposed during the process of drilling a borehole. The successful completion of a well bore requires the deposit of a low-permeable filter cake on the walls of the well bore to seal the permeable formation exposed by the drilling bit. A filter cake can limit drilling fluid losses from the well bore and protect the natural formation from possible damage by the fluids permeating into the well bore. Solids in the flowback fluid may also damage the formation, particularly drilling fines. The suspension of fine particles that enters the formation while the cake is being established is known as “mud spurt” and the liquid that enters subsequently is known as “filtrate.”

For a filter cake to form, the drilling fluid must contain some particles of a size only slightly smaller than the pore openings of the formation. These particles are known as bridging particles and are trapped in surface pores, thereby forming a bridge over the formation pores. Filter cake building fluids can also contain polymers for suspension of solids and for reducing liquid loss through the filter cake by encapsulating the bridging particles. These can be either natural or synthetic polymers. The polymers can include one polymer such as xanthan selected for its rheological properties and a second polymer, a starch for example, selected for reduction of fluid loss. At completion of the drilling or other well servicing, the filter cake must be removed to allow production of the formation fluids or bonding of cement to the formation at the completion stage. Removal of the deposited filter cake should be as complete as possible to recover permeability within the formation. Typically, when the filter cake is removed from the well bore, some of the polymers used to create the filter cake remain intact and are carried to the surface in the flowback fluid.

Hydraulic Fracturing

[0110] In a hydraulic fracturing process, aqueous fracturing fluid is injected under pressure into the bore hole. The pressure drives the fluid into cracks, fissures, and fractures in the formation, forcing such openings to become larger and propagate. Propellant material contained within wedges into the expanded cracks, fissures and fractures to help hold them open when the pressure is reduced and to provide improved formation permeability. The injected fracturing fluid mixes with groundwater, gas, and other materials present in the subterranean environment.

[0111] When the pressure is removed, this fluid mixture flows back to the surface and gas is extracted therefrom. The fracturing fluid mixture after extraction is referred to as “flowback fluid,” the recovered water and fracturing fluid which flow back from oil or gas well drilling fracturing operation. This flowback fluid typically can be anywhere from 10-60% percent of the volume of fluid that is injected into the well, and it flows back over a period of several days to several weeks or longer after fracturing. A significant amount of fracturing fluid can remain in the formation. At a certain point there is a transition between primarily recovering fracturing fluid to that of produced water. A typical fracturing job on a Marcellus shale formation could require 20,000 barrels to 150,000 barrels of fracturing fluid, depending upon the number of stages pumped. For a project pumping 40,000 barrels of fracturing fluid, the load recovery could be 50% or 20,000 barrels of flowback fluid. After the initial several weeks post-fracturing recovery, an additional 10,000 to 30,000 barrels of flowback liquid may flow from the well for two years.

[0112] The formation may consist of water, the fracturing chemicals that were injected into the well, including but not limited to guar gum, proppant, and cross-linker, as well as any contaminants that are present in the rock formation water. In addition to natural salinity of water in the formation, any fresh water that is injected into the well during the fracturing process will tend to dissolve salts in the formation, thus increasing the salinity of the flowback liquid.

Treatment Systems

[0113] U.S. Pat. Nos. 4,536,293; 5,093,008; 6,132,619; 4,896,665; 6,198,656; 6,110,382; 4,465,598; 7,754,080, all of which are hereby incorporated by reference in their respective entireties, disclose methods of treating flowback liquids from drilling and oil discovery processes. This method, and other filtration methods used to treat flowback fluids have the potential of becoming clogged. Flowback treatment systems, and/or filters, including but not limited to reverse osmosis filters, have the tendency to become clogged or inefficient in processing flowback fluid because of the viscosity and/or flow rate of the flowback fluid.

Treating Flowback Fluids

[0114] The invention provides methods using one or more enzymes or enzyme cocktails as described herein, wherein the method treats, or is a step in the treatment, of flowback fluids produced by drilling and exploration operations by degrading viscous, starch containing, or polysaccharide components of the flowback fluid. Thus, this method decreases the viscosity and/or flow rate of flowback fluids.

[0115] In some aspects, the invention provides for formulating an enzyme treatment (using an enzyme used to practice the invention) by blending together an aqueous fluid and a polypeptide used to practice the invention; adding the enzyme treatment to the flowback fluids; allowing the enzyme treatment to degrade the viscous polysaccharide-materials materials in the flowback fluid, wherein the enzyme treatment is effective to break down or hydrolyze the starch and/or polysaccharide components of such fluids.

[0116] In some embodiments, polypeptide used to practice the invention may be capable of breaking bonds within the recovered fluid or flowback. In some embodiments, the enzymes may be capable of reducing the viscosity or increasing the flow rate of the flowback fluids.

[0117] In some embodiments, polypeptide used to practice the invention will reduce the likelihood of clogging in systems used to treat flowback fluid.

[0118] In some embodiments polypeptide used to practice the invention will be added to a system or device used to treat flowback fluids. In some embodiments the polypeptide used to practice the invention will be added to the flowback fluid to enhance previously known flowback fluid treatments.

[0119] In some embodiments, the polypeptide used to practice the invention disclosed herein will be used in combination with microbes used to treat flowback fluid.
In some embodiments, the enzymes will be used to break down materials in the flowback fluid. In another embodiment, the enzymes will be used to break down filter cake materials or fracturing materials in the flowback fluid. In some embodiments, polypeptide used to practice the invention described may be encapsulated to stabilize the enzyme, improve thermostability and alkaline pH tolerance, and provide controlled release. Examples of breaker encapsulation compositions and methods are provided in U.S. Pat. Nos. 5,164,099, 6,163,766, 5,373,901, 5,437,331, and 6,357,527, the disclosures of each of which are incorporated herein by reference thereto.

In some embodiments, polypeptide used to practice the invention is encapsulated, having a coating or membrane that hydrolytically degrades allowing better control of release time and ease of handling not previously afforded. For example, because the polypeptide used to practice the invention is encapsulated in a material that reacts with water, rather than simply dissolves or disintegrates in water, the release can be controlled through the reaction rate of the coating with water. Likewise, by insulating the polypeptide used to practice the invention from the harsh conditions (high temperature and pH) for some period of time, can provide delayed degradation. Those skilled in the art will appreciate that the reaction rate of the coating (and therefore the breaker release profile) can be varied broadly depending on the encapsulating polymer chemistry employed.

In some embodiments, the disclosed enzymes such as cellulases and amylases are thermostable and/or thermostable; for example, the enzyme can retain at least 75% residual activity (e.g., glucanase activity) after 2 minutes at 95°C; and in another aspect, retains 100% activity after heating for 30 minutes at 95°C. In yet another aspect, the enzyme retains 100% activity after heating for 30 minutes at 96°C, 97°C, 98°C or 99°C. In yet another aspect, the disclosed cellulases retain at least 90% activity after heating for 30 minutes at 100°C.

In some embodiments, the cellulase enzymes described herein possess glucanase, e.g., endoglucanase, mannanase, xylanase activity or a combination of these activities. In some aspects, the glucanase activity is an endoglucanase activity (e.g., endo-1,4-beta-D-glucan 4-glucano hydrolase activity) and comprises hydrolysis of 1,4-beta-D-glycosidic linkages in cellulose, cellulose derivatives (e.g., carboxymethyl cellulose and hydroxy ethyl cellulose) lignin, beta-1,3 bonds made of mixed beta-1,3 glucans, such as cereal beta-D-glucans or xylglucans and other plant material containing cellulosic parts. In alternative aspects, these glucanases e.g., endoglucanases, mannanases, xylanases have increased activity and stability, including thermostability or thermostability, at increased or decreased pHs and temperatures.

Examples of suitable polysaccharide substrates of some of the enzymes disclosed herein include galactomannan gums, guar, derivatized guar, cellulose and cellulose derivatives, starch, starch derivatives, xanthan, derivatized xanthan and mixtures thereof. Specific examples also include, but are not limited to, guar gum, guar gum derivative, locust bean gum, karaya gum, xanthan gum, cellulose and cellulose derivatives, etc. Typical polymeric viscosifiers or gelling agents to which the disclosed enzymes may be directed include guar gum, hydroxypropyl guar, carboxymethyl hydroxypropyl guar, hydroxyethyl cellulose, carboxymethyl hydroxyethyl cellulose, carboxymethyl cellulose, dialkyl carboxymethyl cellulose, etc. Other examples of polymers include, but are not limited to, phosphonaprons, sceroglucans, dextran and other types of polymers. In some embodiments, a polymer substrate is carboxymethyl hydroxypropyl guar. In some embodiments, a disclosed enzyme may also be effective in hydrolyzing biogums (e.g., succinoglycan biogums made from date syrup or sucrose). In some embodiments, a disclosed enzyme may be used to hydrolyze cellulose-containing or derivatized cellulose-containing polymers—typically, the enzymes attack glucosidic linkages of the cellulose backbone. The disclosed enzymes may be suitable for degrading the polymer into mostly monosaccharide units, in some cases, by specifically hydrolyzing the exo(1,4)-beta-D-glucosidic and endo(1,4)-beta-D-glucosidic linkages between monosaccharide units and the cellulose backbone in the (1,4)-beta-D-glucosidic linkages of any cellbiose fragments.

DESCRIPTION OF THE FIGURES

FIG. 1 depicts an image of a rheology graph displaying the viscometry of two guar solutions against time, as described in Example 1.

FIG. 2 depicts an image of SDS PAGE gel electrophoresis displaying various levels of protein expression, as described in Example 3.

FIG. 3 depicts a bar graph showing the level of activity of protein preparations, as described in Example 4.


FIG. 5 depicts SEQ ID NO:2, the polypeptide encoded by SEQ ID NO:1, 3, and 4.

FIG. 6 depicts SEQ ID NO:3, the unmodified parent polynucleotide sequence of SEQ ID NO:1 and 4.

FIG. 7 depicts SEQ ID NO:4, the polynucleotide with 14 silent mutations: T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C, as compared to SEQ ID NO:3, plus one additional point mutation upstream from the start codon (additional upstream sequence shown).

FIG. 8 depicts the nucleic acid of SEQ ID NO:5.

FIG. 9 depicts the polypeptide of SEQ ID NO:6.

FIG. 10 depicts the nucleic acid of SEQ ID NO:7.

FIG. 11 depicts the polypeptide of SEQ ID NO:8.

FIG. 12 depicts the nucleic acid of SEQ ID NO:9.

FIG. 13 depicts the polypeptide of SEQ ID NO:10.

FIG. 14 depicts the nucleic acid of SEQ ID NO:11.

FIG. 15 depicts the polypeptide of SEQ ID NO:12.

FIG. 16 depicts the nucleic acid of SEQ ID NO:13.

FIG. 17 depicts the polypeptide of SEQ ID NO:14.

FIG. 18 depicts the nucleic acid of SEQ ID NO:15.

FIG. 19 depicts the polypeptide of SEQ ID NO:16.

DEFINITION OF TERMS

“cellulase” refers to enzymes having cellulase, endoglucanase, β-cellobiohydrolase, β-glucosidase, xylanase, mannanase, β-xylosidase, arabinofuranosidase, and/or oligomerase activity.
“cellulolytic activity” is an enzyme having cellulase, endoglucanase, cellobiohydrolase, beta-glucosidase, xylanase, mannanase, beta-xylanase, arabinoxylosidase, and/or oligosaccharide activity.

A “codon” is a three polynucleotide sequence that specifies the identity of an amino acid to be added to a protein.

A “silent mutation” is a mutation in a codon that does not result in the specification of a different amino acid.

“Open Reading Frame” is a series of codons that specifies the sequence of amino acids in a protein.

A base “position” is the numerical location of a base in a polynucleotide sequence, counted consecutively from the start of the open reading frame or some other reference marker.

To “encode” a protein means to specify the amino acid sequence of that protein.

A “mutation” is a change in a nucleotide sequence or an amino acid sequence compared to a reference.

“Nucleotide” refers to one of the four bases which comprise DNA sequence—Adenine (A), Thymidine (T), Guanidine (G), and Cytosine (C).

“Thermotoga maritima genomic sequence” refers to the Thermotoga maritima strain MS88 genomic sequence specified by GenBank Accession No. AE000512.

An “Expression level” for a given protein is the amount of protein generated by an expression system, such as a transformed cell culture as measured per unit volume of cell culture.

An “Expression level” for a given enzyme is the amount of enzyme activity generated by an expression system, such as a transformed cell culture as measured per unit volume of cell culture.

“Wild-type” refers to a protein or nucleic acid sequence that can be obtained in nature.

Example 1—Use of SEQ ID NO. 2 with Ester

Rheology testing was performed using a Grace M5600 HPHT Rheometer, two samples were assayed, one with the cellulase encoded by SEQ ID NO. 2, and the other, a control, with no enzyme. The assay conditions were as follows: sample 1, included the cellulase encoded by SEQ ID NO. 2 at 200 ppm, 0.25 ppt ester, 25 ppt cross-linked guar, at pH 10.5, at 180 degrees Fahrenheit; sample 2, included 0.25 ppt ester, 25 ppt cross-linked guar, at pH 10.5, at 180 degrees Fahrenheit. As displayed in FIG. 1, the sample including cellulase encoded by SEQ ID NO. 2 and ester at pH 10.5, and 180 degrees Fahrenheit (bottom line) reached nearly 0 centipoise (cP), whereas the control sample with no enzyme maintained a viscosity of 500 cP (top line).

Example 2—Method of Making Enhanced Expression Variants

Two variants (SEQ ID NO:1 and NO: 4) were designed based on SEQ ID NO:3 to mutate at the DNA level to improve the gene expression. The design takes into account of many factors that may influence gene expression. The mutations were introduced on the PCR primers using PCR techniques known of those of skill in the art. Both genes were PCR-amplified and cloned into the Pseudomonas vector pDOW1169 (DOW AgроSciences, IN) using standard molecular cloning techniques. The resulting expression constructs were transformed into Pseudomonas fluorescens DC454 (DOW AgроSciences, IN). A transformant with the SEQ ID NO:1 was designated as the lead as it showed the most enhanced expression.

Example 3—Using SDS-PAGE Gel Electrophoresis and Nonspecific Protein Staining to Visualize Expression Levels of the SEQ ID NO:2

Polypeptide Expressed by Constructs Comprising SEQ ID NOS:1, 3, and 4

Criterion™ precast Tris-HCl polyacrylamide gel (Bio-rad Laboratories, Inc.) was used to separate proteins. The gel was run at 150V using Tris-glycerine buffer (see FIG. 1). Protein loading was normalized to load proteins from 0.33 OD260 cells for each lane. SeeBlue® pre-stained protein standard was used (Life Technologies). The gel was stained with a nonspecific dye, and each lane was visually inspected for the presence of a band at the size of SEQ ID NO:2, about 37 kilodaltons.

The results indicate that there is a single band having an accumulation level which varies across samples and which is absent from the negative control. This band has a size expected for SEQ ID NO:2.

The accumulation level of this band is significantly higher in lanes corresponding to protein extracts from cells harboring constructs comprising SEQ ID NO:1, and to a lesser extend SEQ ID NO:4, that SEQ ID NO:3 or the negative control.

Example 4—Method of Determining Relative Expression Levels for Variants

Nucleic acid sequence comprising SEQ ID NO:1, SEQ ID NO:3 and SEQ ID NO:4 gene were transformed into a suitable host cell for expression of the protein of SEQ ID NO:2. The cells were cultured in flasks so that the encoded protein would be expressed. The cultures were grown at 30° C. and 220 rpm to an OD600 of 0.9 in a designed complex medium, and induced with 0.3 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) for 24 hours. Cells were harvested and lysed either by sonication or heat-treatment at 80°C. for 1 hour. Cellulase activity was measured by a p-Nitrophenyl (pNP) based assay using pNP-β-D-lactopyranoside as substrate. (Advances in Carbohydrate Chemistry and Biochemistry, Academic Press, 1999). Activity levels were measured in U/ml as shown in FIG. 2 to determine relative expression levels from each culture.

The results indicate that cells harboring the construct comprising SEQ ID NO:1 demonstrated significantly more SEQ ID NO:2 activity than those harboring SEQ ID NO:4, and that both SEQ ID NOS:1 and 4 yielded a greater amount of activity of the expressed protein than the cells harboring SEQ ID NO:3.
FEATURE: Open reading frame encoding SEQ ID NO: 2 conveying enhanced protein expression

SEQUENCE: 1

```
ATGGGCTCG ATCTGTTGGA ACCTAACAAA ATCTGGGCC GGCACATTAA TATGGCAAT  
GCGTCAAAG CACCAATAGA AGGCGACTGG GAGTTGGTGAA TAAAGATGAA TCTTCTGACG  
ATTATAAGA AGGGGGTTT CTCTCTGTGT CGAATCACA TAAAGTGGG TACGCAAGCT  
CAGGCGTCTC CTTCTTAGAA ATTCGACGCT TCTTCTCTCA AAAAGTGGA TGAAGTGATA  
AAGCAGCCCG TGAAGGAAGG ACTGGCTTGT GTTATAAAT AAATCATCTTA CAGAAGTGT  
ATGAATGAC CAGAGAAACA CAGGAAAGG TTTCTGTGCT TTGGAAACAA ATGGCTGAT  
CGTATAAGA ACTAATCGCA AAATCTTATTT TTGGAAATTC GAAATGAAACC TCAAGGAAT  
CTTACCGGAA AAAAAATGGA TGAATGTGGG GAGGAGCTAC TAAATGTATAA AATGCAAT  
GACAACAGAC ACAGGTGAT TATAGGCACA GCTGAAGGAG GGGGTATATAC TGCCCTTGG  
AAGCTGAGG TCCAAAATGG GCAGAAGAAT GGCTAGATGA CAATCCTCA TACATCCCTC  
TCTGAAATTT CCATCAGAAG GCAGATGCGG GTGGCTGGAT CTTGGAATATG GTTGGGGAAG  
AAGTGGGGAT CTCTCGATGC TGAAGACAT TTTGATAGAC AATCAATTTC TATAGAAAGA  
TTGTCACAAC AAACAGAAGG ACAAATTC ATAGGGTGAT TGGCGCTGTA CAGAAAGCCT  
GACGTTGAAT CAGAGAAACA ATGGACCTCC TTTCTGCTG CAGAAGCGA GAAAGGGGAG  
TGAGAGCTGCG CATACTCGGA ATTTGGTCC GGTGGTGGT TTTATGATCC TCTGAGAAGA  
cGCTGAAATAAAGATTTT TTTAAGCTTA TTAGGAGG ATAGCATTGA ATGA  
```

SEQ ID NO 2

LENGTH: 317

TYPE: PRT

ORGANISM: Thermotoga maritima

SEQUENCE: 2

```
Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile  
1  5  10  15  
Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val  
20  25  30  
Val Ile Lys Asp Glu Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser  
35  40  45  
His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Glu Ala Phe Pro  
50  55  60  
Pro Tyr Lys Ile Glu Pro Ser Phe Phe Lys Arg Val Asp Glu Val Ile  
65  70  75  80  
Asn Gly Ala Leu Gly Arg Leu Ala Val Ile Arg Asn His His  
85  90  95  
Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Gly Arg Phe Leu  
100 105 110  
Ala Leu Trp Lys Gln Ile Ala Asp Arg Tyr Lys Asp Tyr Pro Glu Thr  
115 120 125  
Leu Phe Phe Glu Ile Leu Asn Glu Pro His Gly Asn Leu Thr Pro Glu  
130 135 140  
Lys Trp Asn Glu Leu Leu Gly Ala Leu Lys Val Ile Arg Ser Ile  
145 150 155 160  
```
<210> SEQ ID NO 3
<211> LENGTH: 954
<212> TYPE: DNA
<213> ORGANISM: Thermotoga maritima

<400> SEQUENCE: 3

atggttggtg atcttttttga aaggaacaaa atatgggaa gaggcattaa tataggaatatggttgagg ccaccaaagtg gggaga ctggatgta taaaagatg gaaggcattaa gattttcgag tttatgtatt ccctgagaa atcgagt aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt 60
gtttttgcct cagtttattt tttatgtatt ccctgagaa atcgagt aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt atggttggtg 120
attataaag aagccc gttctgtattt tttatttgcct cagtttattt tttatgtatt ccctgagaa atcgagt aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt 180
cagtttattt tttatgtatt ccctgagaa atcgagt aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt atggttggtg 240
aagggagcct tgaagattg atgttctgag agtgatgttt tatttttttt tataagt ggttttggtg aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt 300
cagtttattt tttatgtatt ccctgagaa atcgagt aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt atggttggtg 360
cagtttattt tttatgtatt ccctgagaa atcgagt aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt atggttggtg 420
cagtttattt tttatgtatt ccctgagaa atcgagt aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt atggttggtg 480
aagggagcct tgaagattg atgttctgag agtgatgttt tatttttttt tataagt ggttttggtg aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt 540
cagtttattt tttatgtatt ccctgagaa atcgagt aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt atggttggtg 600
aagggagcct tgaagattg atgttctgag agtgatgttt tatttttttt tataagt ggttttggtg aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt 660
aagggagcct tgaagattg atgttctgag agtgatgttt tatttttttt tataagt ggttttggtg aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt 720
tggttttttt tataagt ggttttggtg aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt atggttggtg 780
tggttttttt tataagt ggttttggtg aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt atggttggtg 840
tggttttttt tataagt ggttttggtg aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt atggttggtg 900
cagtttattt tttatgtatt ccctgagaa atcgagt aagatctttt atggttggtg aatggggaga tggtaaagct tatttc tatttttttt tataagt ggttttggtg aagatctttt atggttggtg 964

<210> SEQ ID NO 4
<211> LENGTH: 976
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
OTHER INFORMATION: 5' region and open reading frame encoding SEQ ID NO:2 having modest increased expression

SEQUENCE: 4

ttactagttt aggaggttaa tattggtgcgt cagatcgttt gaacyttaaaca aatcttgggg 60
cgcgcggcatt aatactggcga atgcgcctgca agcaccacatt gaaagcggct ggagagttgt 120
gatataaatg gatatcttgcg acataataaa agaagcccgt ttcttcctatg ttctgaattcc 180
aatagatgag a tagcacagc ctcagaccgt gttctctttat aataatggcg cttctctcttt 240
caaaagagtc gatgaagttga taatacggagc cctgaaaaga ggapctggtgc ttgatataaa 300
tatctccac taagagagct taattatga taattataaagg aacacgaaagaa gatccttcttc 360
tccttggaaaa caaattgctt atccttttaaa agaatctgcc gaaacccctat tttttataaat 420
tctgaatgaa cccacaccgg atccttaacctc ggaaaaatgag aatacttgcg ttcgtggagaac 480
tctaaaagtt ataagataca taattcacaac ggaaattctg attataaggca cagctgtaatg 540
ggggggtata ttcgctcctgt aaaaaacttag ggtccccaaa tggssssssaaa atgcctatagt 600
tataactttc aatcataatct ctttcgtaa ttaataaatgaa gaacggtgatg gggtgcctgg 660
atctgagaa tggtagttggg gaacacctgg atctcaagttat gatcagaaac aatttataag 720
agaattcaa tttatagaaag aatgcaccaaa aaagacaaca aacaggedtta aatagatgta 780
gtctgtcgcc taacaagacc aagcactttg aataaatga aatacatgac ccttttgtctg 840
tgctcacaagc gagaaacgag ggttggacat ggcatactgg gaattttgtc cggcttttgtg 900
tggttattggt cctctgaaaa aacaagttgaa taagatatttat ttagaaggtta ttagaaggg 960
agaatgcctt ggatga 976

SEQ ID NO: 5

LENGTH: 564

ORGANISM: Artificial Sequence

FEATURES:

- OTHER INFORMATION: Open reading frame encoding cellulase

SEQUENCE: 5

atgggttcttg aatctttttgga aaggaacaaa atatrgggaa gagggcattaa tataggaatt 60
gotctttgaa cagcataacaa gggaagacagta gagaatgttaa taaaatcagaa gttctttgacc 120
attataaag aacgcctgtt ctcctctggt cgaatctccaa taagatggag taaccaagct 180
taccgttcttc cttctctatgcatc atctgccaa aacgattct gtaaggata 240
aaccaggcc ttggaaagacg atggctctttgt tttctttata ttcatcaact ocagggagtt 300
atagatgacg caagaaacga caagaaagac tttctttgctc tttgaaacca aatggctgtat 360
cgtttataag aacgctcttattttt tggaaacctc taatggaacct caaggaasaa 420
cctactctgg aaaaatggaa tgaatgcttt taggaagcct ttaaaattatg aagatcactt 480
gacaaaaagc acactataa tatagggcaca gctgaatggg gggttatatt tcgctttgga 540
aaactgttctt tccccaaaaag gaaaaaaatt ttataatgtaa cattcaacta cttacactatc 600
tctgaattta cccataaagc aagctgtgtgg gttgaaaggt aagcaaaaat gttggaagaa 660
aagctgttttgctc ttcacagatac ttcgaaacact tttgagaag aatcttttaaa tataggaasaa 720
tgtctaaaaggcagaaaaag accaatcttac atagggtagt tcggcccota cagaaaagct 780
gaccttggt aacagaaat aatggaactcc tttgtcggttc ggaatgaga gaaaaggaga 840
<210> SEQ ID NO 6
<211> LENGTH: 317
<212> TYPE: PRT
<213> ORGANISM: Thermotoga maritima

<400> SEQUENCE: 6

Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile
1 5 10 15

Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val
20 25 30

Val Ile Lys Asp Glu Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser
35 40 45

His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Tyr Ala Phe Pro
50 55 60

Pro Tyr Lys Ile Met Asp Arg Phe Phe Lys Arg Val Asp Glu Val Ile
65 70 75 80

Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Val Ile Asn Ile His His
85 90 95

Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Glu Arg Phe Leu
100 105 110

Asn Thr Lys Gln Ile Ala Asp Arg Tyr Asp Tyr Pro Glu Thr
115 120 125

Leu Phe Phe Glu Ile Leu Asn Glu Pro Asp Gly Leu Thr Pro Glu
130 135 140

Lys Thr Asp Glu Leu Leu Glu Ala Leu Lys Val Ile Arg Ser Ile
145 150 155 160

Asp Lys Lys His Thr Ile Ile Lys Thr Ala Glu Trp Gly Gly Ile
165 170 175

Ser Ala Leu Glu Lys Leu Ser Val Pro Lys Trp Glu Lys Asn Ser Ile
180 185 190

Val Thr Ile His Tyr Thr Asp Pro Phe Glu Phe Thr His Gln Gly Ala
195 200 205

Glu Trp Val Glu Gly Ser Glu Lys Trp Leu Gly Arg Lys Trp Gly Ser
210 215 220

Pro Asp Asp Gln Lys His Leu Ile Glu Glu Phe Asn Phe Ile Glu Glu
225 230 235 240

Trp Ser Lys Asp Asn Met Lys Arg Pro Ile Tyr Ile Gly Glu Phe Gly Ala
245 250 255

Tyr Arg Lys Ala Asp Leu Glu Ser Arg Ile Lys Thr Thr Ser Phe Val
260 265 270

Val Arg Glu Met Glu Lys Arg ASp Thr Met Asp Thr Tyr Trp Glu Phe
275 280 285

Cys Ser Gly Phe Gly Val Tyr Asp Thr Leu Arg Lys Thr Trp Asn Lys
290 295 300

Asn Leu Leu Glu Ala Leu Ile Gly Gly Asp Ser Ile Glu
305 310 315

<210> SEQ ID NO 7
<211> LENGTH: 972
-continued

<210> SEQ ID NO 8
<211> LENGTH: 323
<212> TYPE: PRT
<213> ORGANISM: Thermotoga maritima

<400> SEQUENCE: 8

Met Glu Gln Ser Val Ala Glu Ser Asp Ser Asn Ser Ala Phe Glu Tyr
1      5      10      15

Asn Lys Met Val Gly Lys Gly Val Asn Ile Gly Asn Ala Leu Glu Ala
20     25     30

Pro Phe Glu Gly Ala Trp Gly Val Arg Ile Glu Asp Gly Tyr Phe Glu
35     40     45

Ile Ile Lys Arg Gly Phe Asp Ser Val Arg Ile Pro Ile Arg Trp
50     55     60

Ser Ala His Ile Ser Glu Lys Pro Pro Tyr Asp Ile Asp Arg Asn Phe
65     70     75     80

Leu Glu Arg Val Asn His Val Val Asp Arg Ala Leu Glu Asn Asn Leu
85     90     95

Thr Val Ile Ile Asn Thr His His Phe Glu Glu Leu Tyr Glu Glu Pro
100    105    110

Asp Lys Tyr Gly Asp Val Leu Val Glu Ile Trp Arg Glu Ile Ala Lys
115    120    125

Phe Phe Lys Asp Tyr Pro Glu Asn Leu Phe Phe Glu Ile Tyr Asn Glu
130    135    140
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro Ala Gln Asn Leu Thr Ala Glu Lys Trp Asn Ala Leu Tyr Pro Lys</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>145</td>
<td>150</td>
<td>155</td>
</tr>
<tr>
<td>Val Leu Lys Val Ile Arg Glu Ser Asn Pro Thr Arg Ile Val Ile Ile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>165</td>
<td>170</td>
<td>175</td>
</tr>
<tr>
<td>Asp Ala Pro Asn Trp Ala His Tyr Ser Ala Val Arg Ser Leu Lys Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>180</td>
<td>185</td>
<td>190</td>
</tr>
<tr>
<td>Val Asn Asp Lys Arg Ile Ile Val Ser Phe His Tyr Tyr Glu Pro Phe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>195</td>
<td>200</td>
<td>205</td>
</tr>
<tr>
<td>Lys Phe Thr His Gln Gly Ala Glu Trp Val Asn Pro Ile Pro Asp Val</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>210</td>
<td>215</td>
<td>220</td>
</tr>
<tr>
<td>Arg Val Lys Trp Arg Gly Glu Trp Glu Ile Asn Gln Ile Arg Ser</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>225</td>
<td>230</td>
<td>235</td>
</tr>
<tr>
<td>His Phe Lys Tyr Val Ser Asp Trp Ala Lys Gln Asn Asn Val Pro Ile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>240</td>
<td>245</td>
<td>250</td>
</tr>
<tr>
<td>Phe Leu Gly Glu Phe Gly Ala Tyr Ser Lys Ala Asp Met Asp Ser Arg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>260</td>
<td>265</td>
<td>270</td>
</tr>
<tr>
<td>Val Lys Trp Thr Glu Ser Val Lys Met Ala Glu Glu Phe Gly Phe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>270</td>
<td>275</td>
<td>280</td>
</tr>
<tr>
<td>Ser Tyr Ala Tyr Trp Glu Phe Cys Ala Gly Phe Gly Ile Tyr Asp Arg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>290</td>
<td>295</td>
<td>300</td>
</tr>
<tr>
<td>Trp Ser Gln Asn Trp Ile Glu Pro Leu Ala Thr Ala Val Val Gly Thr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>305</td>
<td>310</td>
<td>315</td>
</tr>
</tbody>
</table>

Gly Lys Glu

<210> SEQ ID NO 9
<211> LENGTH: 974
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Cellulase 7xGene site saturation mutagenesis product and 3' sequence
<400> SEQUENCE: 9

atgggtgtc atcttctttg aaggaacaaa atatggtgg aaacgcattaa tataggaat

60
gcctgtaagg ccaacaaaag gggagacgct ggagttgctga taaagagtgta gttatccagc

120
attatataag aagcctgttt cctctatgtt cgaattccaa taaatggag taacgcagct

180
caggcgttct cttctttaaa aatcgaggtg cgtcttttca aagatgagga tggagtgata

240
aacggagccc tgaacagagc actgtgtgtt gtatataaac agatcactta cagaatgta

300
atgactgtac tccagagaca caagagaaga tttcttgctt tttgaaaca aatgtctgtat

360
cgttattag actatctcgg aactatattt tttgaatattc tgaatgaccc ttaccaggaat

420
cttacctcggg aaaaaaggg tgaacaagcc cggagaagct ttaaagagt aagatcacta

480
gcagaacc gcacactata taagaccca cggtaaggg ggggtatatc tcggacctgg

540
aaactggagg tcccaaatgt gaaaaaagtt cgctagttgt taaatcata cctcatacct

600
ttccaatatta ccatctacgg agctgaggg tgggagagag ctggagaaat gttgggaagagg

660
aagttgggtat ctcacagtca tccagacaat tttgtagaag acatcaatattt ttagaagaac

720
tggctaaaaa aagacaaaaag acaaatattcc ataggtgaag tttggaagct tcaaaaaaggtct

780
gaccttac gacaacaaa atggacaccc tttctcgcgg cggagacgct gaaaaaggg

840
tggagctgg ggacactgga atttgtgtcc gttttgggtg ttataagata ctctgaaaga

900
Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile
1 5 10 15
Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val
20 25 30
Val Ile Lys Asp Glu Tyr Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser
35 40 45
His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Gln Ala Phe Pro
50 55 60
Pro Tyr Lys Ile Glu Asp Arg Phe Lys Arg Val Asp Glu Val Ile
65 70 75 80
Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Val Ile Asn Gln His His
85 90 95
Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Gly Arg Phe Leu
100 105 110
Ala Leu Trp Lys Gln Ile Ala Asp Arg Tyr Lys Asp Tyr Pro Glu Thr
115 120 125
Leu Phe Phe Glu Ile Leu Asn Leu Pro His Gly Asn Leu Thr Pro Glu
130 135 140
Lys Trp Asn Glu Leu Leu Glu Ala Leu Lys Val Ile Arg Ser Ile
145 150 155 160
Asp Lys Lys His Thr Ile Ile Ile Gly Thr Ala Glu Trp Gly Gly Ile
165 170 175
Ser Ala Leu Glu Lys Leu Arg Val Pro Lys Trp Glu Lys Asn Ala Ile
190 195 200 205
Val Thr Ile His Tyr Tyr Asn Pro Phe Glu Phe Thr His Gln Gly Ala
196 200 205
Glu Trp Val Glu Gly Ser Glu Lys Trp Leu Gly Arg Lys Trp Gly Ser
210 215 220
Pro Asp Asp Glu Lys His Leu Ile Glu Glu Phe Asn Phe Ile Glu Glu
225 230 235 240
Trp Ser Lys Asn Lys Arg Pro Ile Tyr Ile Gly Glu Phe Gly Ala
245 250 255
Tyr Arg Lys Ala Asp Leu Glu Ser Arg Ile Lys Trp Thr Ser Phe Val
260 265 270
Val Arg Glu Ala Glu Lys Arg Arg Ile Ser Trp Ala Tyr Trp Glu Phe
275 280 285
Cys Ser Gly Phe Gly Val Tyr Asp Thr Leu Arg Lys Thr Trp Asn Lys
290 295 300
Asp Leu Leu Glu Ala Leu Ile Gly Gly Asp Ser Ile Glu His His Ser
305 310 315 320
Lys Met Ala

<210> SEQ ID NO 11
<210> SEQ ID NO 12
<211> LENGTH: 317
<212> TYPE: PRT
<213> ORGANISM: Thermotoga maritima

<400> SEQUENCE: 12
Met Gly Val Asp Pro Phe Glu Arg Ann Lys Ile Leu Gly Arg Gly Ile
1   5      10     15
Ann Ile Gly Ann Ala Leu Glu Ala Pro Ann Glu Gly Asp Trp Gly Val
20  25  30
Val Ile Lys Asp Glu Tyr Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser
35  40
His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Gln Ala Phe Pro
50  55  60
Pro Tyr Lys Ile Glu Asp Ser Phe Phe Arg Val Asp Glu Val Ile
65  70  75  80
Ann Gly Ala Leu Lys Arg Gly Leu Ala Val Ile Ann Ile His His
85  90  95
Tyr Glu Glu Leu Met Ann Asp Pro Glu Glu His Lys Gly Arg Phe Leu
100 105 110
Ala Leu Trp Lys Glu Ile Ala Asp Arg Tyr Lys Asp Tyr Pro Glu Thr
115 120 125
Leu Phe Phe Glu Ile Leu Ann Glu Pro His Gly Ann LeuThr Pro Glu
130 135 140
<table>
<thead>
<tr>
<th>Amino Acid Sequence</th>
<th>Start</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lys Trp Asn Glu Leu Leu Glu Glu Ala Leu Lys Val Ile Arg Ser Ile</td>
<td>148</td>
<td>160</td>
</tr>
<tr>
<td>Asp Lys Lys His Thr Val Ile Ile Gly Thr Ala Glu Trp Gly Gly Ile</td>
<td>165</td>
<td>175</td>
</tr>
<tr>
<td>Ser Ala Leu Glu Lys Leu Arg Val Pro Lys Trp Glu Lys Asn Ala Ile</td>
<td>180</td>
<td>190</td>
</tr>
<tr>
<td>Val Thr Ile His Tyr Tyr Asn Pro Phe Glu Phe Thr His Gln Gly Ala</td>
<td>195</td>
<td>205</td>
</tr>
<tr>
<td>Glu Trp Val Pro Gly Ser Lys Trp Leu Gly Arg Lys Trp Gly Ser</td>
<td>210</td>
<td>220</td>
</tr>
<tr>
<td>Pro Asp Asp Gin Lys His Val Ile Glu Glu Phe Asn Phe Ile Glu Glu</td>
<td>225</td>
<td>240</td>
</tr>
<tr>
<td>Trp Ser Lys Lys Asn Lys Arg Pro Ile Tyr Ile Gly Glu Phe Gly Ala</td>
<td>245</td>
<td>255</td>
</tr>
<tr>
<td>Tyr Arg Lys Ala Asp Leu Glu Ser Arg Ile Lys Trp Thr Ser Phe Val</td>
<td>260</td>
<td>270</td>
</tr>
<tr>
<td>Val Arg Glu Ala Glu Lys Arg Gly Trp Ser Trp Ala Tyr Trp Glu Phe</td>
<td>275</td>
<td>285</td>
</tr>
<tr>
<td>Cys Ser Gly Phe Gly Val Tyr Asp Pro Leu Arg Lys Gin Trp Asn Lys</td>
<td>290</td>
<td>300</td>
</tr>
<tr>
<td>Asp Leu Leu Glu Ala Leu Lys Gly Asp Ser Ile Glu</td>
<td>305</td>
<td>315</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 13
<211> LENGTH: 954
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURES:
<223> OTHER INFORMATION: Cellulase 12X-1 combined Gene site saturation mutagenesis product

<400> SEQUENCE: 13

```
atgggttgtg atcccttttg aaggaacaaa atatgggaa gaggctattaa tataggaat
60
gcccttgaag cacaatgtg ggagactggg ggaagggagtt gaaagttologia
120
attataaag aacgccggtt cttccatagtt gaatcactc aacaatcagctg
180
caggggttcc cttccctataa aatcggagat cttcctttctt aaggaaatgta
240
aacgggacct tgaagaaggg aatgggctgtt gtttataacct agcatcacta cggaggtta
300
atgaatgtgg cagaagaga caaggaaga tttcctgtcc ttgggaaaca aatggcgtat
360
cgtttataaag acctacgctt gscatcgttt tttggggaaa gctgaatagg taggaatggg
420
cctctgctgg aaaaatggga tgaattgctt gcaggaagtt caaagttatat aagatcaatt
480
gacaaaaagc aacatgctg tattagccac gactgaatgg ggagaattgg cggcccttgaa
540
aasctagggg gccccaaaatg gaaaaaatggcagatat cactcaactatcctatctctc
600
ttcgaattta cccatcaag agctgagtgt ggtcatgttt cgctagaaatg ttggaagaag
660
aagttgggag cttccggtat gcagaaatct ttgtataagg aatctaatatt tataggaag
720
tgtgcaaaaaa gcaaaaaag accaatattac aatgtagtattaggctc taagacagct
780
gacctttattt cagaaatatt atagaggccct cttgctcgttt gcagaaagcct gaaaaaggggg
840
tggaagcttcttccagcatagtattttgtttgtttt tttatatctc cttgaagaaa
900
cagttgagaaa aagatcttttt ataggagag aatcaggtttag aataa
954
```
<210> SEQ ID NO: 14
<211> LENGTH: 317
<212> TYPE: PRT
<213> ORGANISM: Thermotoga maritima

<400> SEQUENCE: 14

Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile
1  5  10  15
Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val
20 25 30
Val Ile Lys Asp Glu Phe Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser
35 40 45
His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Gln Ala Phe Pro
50 55 60
Pro Tyr Lys Ile Glu Asp Ser Phe Phe Lys Arg Val Asp Glu Val Ile
65 70 75 80
Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Val Ile Asn Gln His His
85 90 95
Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Glu Arg Phe Leu
100 105 110
Ala Leu Trp Lys Gln Ile Ala Asp Arg Tyr Lys Asp Tyr Pro Glu Thr
115 120 125
Leu Phe Phe Glu Ile Leu Asn Glu Pro His Gly Asn Leu Thr Pro Glu
130 135 140
Lys Trp Asn Glu Leu Leu Glu Ala Leu Lys Val Ile Arg Ser Ile
145 150 155 160
Asp Lys Lys His Thr Val Ile Ile Gly Thr Ala Glu Trp Gly Gly Ile
165 170 175
Ser Ala Leu Glu Lys Leu Arg Val Pro Lys Trp Glu Lys Asn Ala Ile
180 185 190
Val Thr Ile His Tyr Tyr Asn Pro Phe Glu Phe Thr His Gln Gly Ala
195 200 205
Glu Trp Val Pro Gly Ser Glu Lys Trp Leu Gly Arg Lys Trp Gly Ser
210 215 220
Pro Asp Asp Gln Lys His Leu Ile Glu Glu Phe Asn Phe Ile Glu Glu
225 230 235 240
Trp Ser Lys Asn Lys Arg Pro Ile Tyr Ile Gly Glu Phe Gly Ala
245 250 255
Tyr Arg Lys Ala Asp Leu Glu Ser Arg Ile Lys Trp Thr Ser Phe Val
260 265 270
Val Arg Glu Ala Glu Lys Arg Gly Trp Ser Ser Ala Tyr Trp Glu Phe
275 280 285
Cys Ser Gly Phe Gly Val Tyr Asp Pro Leu Arg Lys Gln Glu Trp Asn Lys
290 295 300
Asp Leu Leu Glu Ala Leu Ile Gly Gly Asp Ser Ile Glu
305 310 315

<210> SEQ ID NO: 15
<211> LENGTH: 1310
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

<220> FEATURE: Other information: Open reading frame encoding SEQ ID NO: 16
Amylase
<400> SEQUENCE: 15
atggccacagt atcccagagt ggaaaggggc ggggtcataa tgcaagcgcgg ctactcgggac 60
gtcotacag gagaatagtgtgagcagcataa atacgccaga agatacgcaga gtggtaacatg 120
gccgaacctct gcacataagc atcctctcccg ggacgcaagg gcagtgccagg cgctttcatt 180
tagctctcag aacctacagc tttttggac tctctgagta gcaaccagacag ggaacggttac 240
agacgcgttt ttcacctcag cagagcctcgg tgaacatgat aacacccgcc cagcccttagt 300
gcatgaggt aatacogatata agatcaaaat caccacgacg eggcggtcag ctggagttgga 360
accctctttgg gaagcacaat accttgacgg ccactttcagag ggtgctgctgg ggttaataac 420
cggcaacctc atgaagtcct caccagagcg agtagatcgg gcggagctgg ggaacattgg 480
gagacctc gcagatcgc gcagcacaag cgcgtgacagct ggccggcagc 540
agagagctc acggctctaa tccagggctc tcggcgtgatg agctctcggc ttcgggactc 600
toaagctcgg agtgccgcaag atgcgtgagc atggctggttc ggtgaggggcc 660
tccagagctc ctcggtgac ccctctcttc ggtgtcagtag aagctgctct 720
ccoagctttg tcaagctgcc ccctctcttc gatgcggatc gcgtggcgg ccacaaccaac 780
tccagcttccgctctc ttttctagc caccagagcg cgcgtgacagct ggtggcagcgc 840
caggggcttc ccctctcttc ggtgtcagtag aagctgctct 900
acggtggcttc ccctctcttc ggtgtcagtag aagctgctct 960
ggtggcagc gcggcagcgc gcgtgacagct ggtggcagcgc 1020
cacggtggcttc ccctctcttc ggtgtcagtag aagctgctct 1080
acoagctttg tcaagctgcc ccctctcttc gatgcggatc gcgtggcgg ccacaaccaac 1140
aticggtggcttc ccctctcttc ggtgtcagtag aagctgctct 1200
tgacgtggtgcc gcggcagcgc gcgtgacagct ggtggcagcgc 1260
ccacoagctttg tcaagctgcc ccctctcttc gatgcggatc gcgtggcgg ccacaaccaac 1320

<210> SEQ ID NO 16
<211> LENGTH: 436
<212> TYPE: PRT
<213> ORGANISM: Thermotoga maritima

<400> SEQUENCE: 16
Met Ala Lys Tyr Ser Glu Leu Glu Gly Gly Val Ile Met Gln Ala
1   5   10  15
Phe Tyr Trp Amp Val Pro Ser Gly Gly Ile Trp Thr Ile Arg
20  25  30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
35  40  45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
50  55  60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65  70  75  80
Glu Thr Arg Phe Gly Ser Lys Gin Glu Leu Val Asn Met Ile Asn Thr
85  90  95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
100 105 110
What is claimed is:

1. A nucleotide sequence of SEQ ID NO:1.