MICROORGANISMS HAVING ENHANCED TOLERANCE TO INHIBITORS AND STRESS

Inventors: Steven D. Brown, Knoxville, TN (US); Shihui Yang, Knoxville, TN (US)

Correspondence Address:
Scully Scott Murphy & Presser PC
400 Garden City Plaza, Suite 300
Garden City, NY 11530 (US)

Assignee: UT-BATTELLE, LLC, Oak Ridge, TN (US)

Filed: Jun. 8, 2010

Related U.S. Application Data

Provisional application No. 61/184,961, filed on Jun. 8, 2009.

ABSTRACT

The present invention provides genetically modified strains of microorganisms that display enhanced tolerance to stress and/or inhibitors such as sodium acetate and vanillin. The enhanced tolerance can be achieved by increasing the expression of a protein of the Sm-like superfamily such as a bacterial Hfq protein and a fungal Sm or Lsm protein. Further, the present invention provides methods of producing alcohol from biomass materials by using the genetically modified microorganisms of the present invention.
Anaerocellum_thermophilum
Caldicellulosiruptor_saccharol
Clostridium_thermocellum
Thermoanaerobacter_sp._X514_
Zymomonas_mobilis_ZM4
Escherichia_coli_K12

--MAKGSLNLQLDFLNLQQLKEKVKNTFLLSFGFQLKGVKGDNFNTLVPE 48
--MAKGNLNQLFLNLQQLKEKVKNTFLLSFGFQLKGVKGDNFNTLVVE 48
MVS-KNNINLQDFVNLQVRKHEPVTVYLNFGQLKGVKGDNFNTVVLQ 49
MASSKAAINLQDIFLNQVRKHEPVTVYLNFGQLKGVKGDNFNTVVLQ 50
--MAEKVNLQLDFVLNLQTRPVTMFLVQLQVKTQVTWFDNSILLR 48
--MAKGQSLQDPLFLNALRERPVPSLYLVNGIKLQOGQAESFDQFVILK 47

 Anaerocellum_thermophilum
Caldicellulosiruptor_saccharol
Clostridium_thermocellum
Thermoanaerobacter_sp._X514_
Zymomonas_mobilis_ZM4
Escherichia_coli_K12

TDNKQQLIYKHAISIMPS^KP-------------------INYMAQA 77
TENNKQQLIYKHAISSILPS^KP-------------------INYMAQV 77
SEGRQQ-LIYKAISTIPMK-------------------IVSLIF 75
SEN^QQLLIYKHAISTIPQK-------------------VIFSAAD 79
RDGSQ-LVYKHAIISTIPAHPLEQLRESRSKSLQAERKSSLQDVFVLSAM 97
--NTVSQMVYKHAISTVPVSMP-------------------VSXHJS 72

Anaerocellum_thermophilum
Caldicellulosiruptor_saccharol
Clostridium_thermocellum
Thermoanaerobacter_sp._X514_
Zymomonas_mobilis_ZM4
Escherichia_coli_K12

QNQQ---QAQQSNNNQG------------------- 92
QNSQVNTASQQSNMQNQQESK------------------- 99
NDNN------RSE------------------- 82
KDEK------REE------------------- 86
QQQEPVTLINGVMLQGEIAAFDLFVCVLTRNDAOLVYKHAVSTVQP147
NNAQ---GGTSSNHYHGSSAQN-------------------TSAQQDS 98

-------------------
-------------------
-------------------
-------------------
-------------------
-------------------

Fig. 1E
Fig. 2A

Fig. 2B
Anaerobic growth in RM

Fig. 3A-3B
Anaerobic growth in RM with 16 g/L NaAc

*Fig. 3C*
195 mM KAc (RM, pH 5.0)

Fig. 4E
Fig. 5A

RM, pH 5.0

Fig. 5B

1g/L vanillin (RM, pH 5.0)
Fig. 5C

1 g/L furfural (RM, pH 5.0)

Time (h)

Log10(OD600nm)

ZM4
AcR
AcRIM0347
AcRIM0347(p42-0347)

Fig. 5D

1 g/L HMF (RM, pH 5.0)

Time (h)

Log10(OD600nm)
Anaerobic growth in RM with 0.001% $\text{H}_2\text{O}_2$

Fig. 5E
A Growth of yeast deletion mutant strains with sodium acetate

B Growth of yeast gene over-expression strains with sodium acetate

Figs. 6A-6B
Fig. 7A

![Graph CM+Glucose, pH5.5](image)

Fig. 7B

![Graph 305 mM NaCl (CM+Glucose, pH5.5)](image)
305 mM KAc (CM+Glucose, pH5.5)

Fig. 7E

0.75 g/L vanillin (CM+Glucose, pH 5.5)

Fig. 7F
1.5 g/L furfural (CM+Glucose, pH 5.5)

Fig. 7G

1.5 g/L HMF (CM+Glucose, pH 5.5)

Fig. 7H
CM+Glucose+2% Galactose-Uracil, pH 5.5

305 mM NaCl (CM+Glucose+2% Galactose-Uracil, pH 5.5)
305 mM NaAc (CM+Glucose+2% Galactose-Uracil, pH 5.5)

![Graph 7K](image)

305 mM NH₄OAc (CM+Glucose+2% Galactose-Uracil, pH 5.5)

![Graph 7L](image)
305 mM KAc (CM+Glucose+2% Galactose-Uracil, pH 5.5)

![Graph showing growth curve for different samples over time.](image)

Fig. 7M

0.75 g/L vanillin (CM+Glucose+2% Galactose, pH 5.5)

![Graph showing growth curve for different samples over time.](image)

Fig. 7N
1.5 g/L furfural (CM+Glucose+2% Galactose, pH 5.5)

Fig. 7O

1.5 g/L HMF (CM+Glucose+2% Galactose, pH 5.5)

Fig. 7P
MICROORGANISMS HAVING ENHANCED TOLERANCE TO INHIBITORS AND STRESS

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of priority from U.S. Provisional Application No. 61/184,961, filed on Jun. 8, 2009, the content of which in its entirety is incorporated herein by reference.

[0002] This invention was made with government support under Contract Number DE-AC05-00OR22725 between the United States Department of Energy and UT-Battelle, L.L.C. The U.S. Government has certain rights in this invention.

FIELD OF THE INVENTION

[0003] This invention generally relates to the field of microorganism and genetic modification thereof. In particular, the invention relates to microorganisms that display enhanced tolerance to stress and inhibitors as a result of increased expression of a protein of the Sm-like superfamily such as bacterial Hfq and yeast Sm or Lsm proteins. Such microorganisms are advantageous for use in fermentation of biomass materials to produce biofuels such as ethanol.

BACKGROUND OF THE INVENTION

[0004] Biomass-based bioenergy is crucial to meet the goal of makingcellulosic biofuels cost-competitive with gasoline. Lignocellulosic materials represent an abundant feedstock for cellulosic-biofuel production. A core challenge in converting cellulosic material to biofuels such as ethanol and butanol is the recalcitrance of biomass to breakdown. Because of the complex structure of lignocellulosic biomass, pretreatment is necessary to make it accessible for enzymatic attack. Severe biomass pretreatments are required to release the sugars, which along with by-products of fermentation can create inhibitors in the production of ethanol or butanol, for example. During the pretreatment processes, a range of inhibitory chemicals are formed that include sugar degradation products such as furfural and hydroxymethyl furfural (HMF); weak acids such as acetic, formic, and levulinic acids; lignin degradation products such as the substituted phenolics vanillin and lignin monomers. In addition, the metabolic byproducts such as ethanol, lactate, and acetate also impact the fermentation by slowing and potentially stopping the fermentation prematurely. The increased lag phase and slower growth increases the ethanol cost due to both ethanol production rate and total ethanol yield decreases (Takahashi et al. 1999; Kadar et al. 2007).

[0005] Efficient conversion of lignocellulosic hydrolysates to biofuel requires high-yield production and resistance to industrially relevant stresses and inhibitors. To overcome the issue of inhibition caused by pretreatment processes, there are two approaches, one is to remove the inhibitor after pretreatment from the biomass physically or chemically, which requires extra equipment and time leading to increased costs. A second approach utilizes inhibitor tolerant microorganisms for efficient fermentation of lignocellulosic material to ethanol and their utility is considered an industrial requirement (Almeida et al. 2007).

[0006] Zymomonas mobilis are gram-negative facultative anaerobic bacteria with a number of desirable industrial characteristics, such as high-specific productivity and ethanol yield, unique anaerobic use of the Entner-Doudoroff pathway that results in low cell mass formation, high ethanol tolerance (12%), pH 3.5-7.5 range for ethanol production and has been generally regarded as safe (GRAS) status (Swings and De Ley 1977; Rogers et al. 1984; Gunasekaran and Raj 1999; Dien et al. 2003; Panesar et al. 2006; Rogers et al. 2007). One drawback to using wild-type Z. mobilis is its narrow substrate utilization range. However, recombinant Z. mobilis strains have been developed to ferment pentose sugars such as xylose and arabinose (Zhang et al. 1995; Deanda et al. 1996; Mohagheghi et al. 2002). On the other hand, low tolerance to acetic acid and decreased ethanol tolerance have been reported in recombinant strains (Ranatunga et al. 1997; Lawford and Rousseau 1998; Lawford et al. 2001; Dien et al. 2003).

[0007] Acetic acid is an inhibitor produced by the deacetylation of hemicelluloses during biomass pretreatment. At pH 5.0, 36% of acetic acid is in the uncharged and undissociated form (HAc) and is able to permeate the Z. mobilis plasma membrane (Lawford and Rousseau 1993). The inhibition mechanism has been ascribed to the ability of the undissociated (protonated) form to cross the cell membrane leading to uncoupling and anion accumulation causing cytoplasmic acidification. Its importance comes from the significant concentrations of acetate that are produced relative to fermentable sugars (McMillan 1994) and the ratio of acetic to fermentable sugars is particularly high in material from hardwoods (Lawford and Rousseau 1993). Acetate may reach inhibitory levels when pretreated biomass hydrolysates are concentrated to generate high final ethanol concentrations or where process water is recycled. Acetate removal processes have been described but they are energy or chemical-intensive and their impact on processing costs have yet to be determined (McMillan 1994).

[0008] An acetate tolerant Z. mobilis mutant (AcR) has been generated by a random mutagenesis and selection strategy (Joachimsthal and Rogers 1998). The AcR mutant was capable of efficient ethanol production in the presence of 20 g/L sodium acetate while the parent ZM4 was inhibited significantly above 12 g/L sodium acetate under the same conditions. A number of studies have characterized the performance of recombinant Z. mobilis strains able to utilize both C-5 and C-6 sugars, including under acetate stress conditions (Lawford et al. 1999; Joachimsthal and Rogers 2000; Lawford and Rousseau 2001). Acetic acid was shown to be strongly inhibitory to wild-type derived strain ZM4(pZBS) on xylose medium and nuclear magnetic resonance studies indicated intracellular deenergization and acidification appeared to be the major inhibition mechanisms (Kim et al. 2000). A recombinant strain able to utilize both xylose (a C-5 sugar) and glucose (a C-6 sugar) with increased acetate resistance was generated by transforming plasmid pZBS into the AcR background (Jeon et al. 2002). Mohagheghi et al. (2004) reported a recombinant Zymomonas mobilis 8b tolerant up to 16 g/L acetic acid and achieved 82%-87% (w/w) ethanol yields from pure glucose/xylose solutions.

[0009] Acetic acid bacteria are used for the industrial production of vinegar and are intrinsically resistant to acetic acid. Although the resistance mechanism is not completely understood, progress toward this goal has been made in recent years. Spontaneous acetic acid bacteria mutants for Acetobacter aceti (Okumura et al. 1985) and several Acetobacter pasteurianus strains (Takemura et al. 1991; Chinnawiri-pisan et al. 2003) showed growth defects in the presence of acetic acid, which was associated with loss of alcohol dehy-
drogenase activity. Fukaya et al (1990) identified the aarA, aarB, and aarC gene cluster as being important for conferring acetic acid resistance using a genetic approach (Fukaya et al. 1990). aarA encodes citrate synthase and aarC encodes a protein that is involved in acetate assimilation (Fukaya et al. 1993), and the three aar genes have been suggested to support increased flux through a complete but unusual citric acid cycle to lower cytoplasmic acetate levels (Mullins et al. 2008). The presence of a proton motive force-dependent efflux system for acetic acid has been demonstrated as being important in A. acetii acetic acid resistance, although the genetic determinant(s) remain to be identified (Matsushita et al. 2005). In E. coli, over-expression of the ATP-dependent helicase RecG has been reported to improve resistance to weak organic acids including acetate (Steiner and Sauer 2003). Baumlert et al. (2006) describe the enhancement of acid tolerance in Z. mobilis by the expression of a proton-buffering peptide in acidified TSB (pH 3.0) or acetic acid (pH 3.5), glyoxylate-HCl buffer (pH 3.0) and sodium acetate-acetic acid buffer (pH 3.5) (Baumlert et al. 2006). Baumlert et al. (2006) also note that the presence of the antibiotic also significantly increased acid tolerance by an unknown mechanism.

0010] Aerobic, stationary phase conditions were found to produce a number of inhibitory secondary metabolites from Z. mobilis when compared to anaerobic conditions at the same time point. The Z. mobilis global regulator gene hq is has been identified as associated with stress responses generated under aerobic stationary phase conditions (Yang et al., 2009). Hq is a bacterial member of the Sm family of RNA-binding proteins, which act by base-pairing with target mRNAs and functions as a chaperone for non-coding small RNA (sRNA) in E. coli (Valentin-Hansen et al. 2004; Zhang et al. 2002; Zhang et al. 2003). E. coli Hq is involved in regulating various processes and deletion of hq has pleiotropic phenotypes, including slow growth, osmosensitivity, increased oxidation of carbon sources, and altered patterns of protein synthesis in E. coli (Valentin-Hansen et al. 2004; Tsui et al. 2004). E. coli Hq has also been reported to affect genes involved in amino acid biosynthesis, sugar uptake, metabolism, and energetics (Guisbert et al. 2007). The expression of thirteen ribosomal genes was down-regulated in hq mutant background in E. coli (Guisbert et al. 2007). Hq also up-regulated sugar uptake transporters and enzymes involved in glycolysis and fermentation such as pgk and pykA, and adhE (Guisbert et al. 2007). E. coli Hq is also involved in regulation of general stress responses that are mediated by alternative sigma factors such as RpoS, RpoE and RpoH. Cells lacking Hq induce the RpoE-mediated envelope stress response and rpoH is also induced in cells lacking Hq (Guisbert et al. 2007), which is consistent with our results that Z. mobilis hq was less abundant in aerobic fermentation condition in ZM4 at 26 h post-inoculation and was rpoH induced (Yang et al. 2009).

SUMMARY OF THE INVENTION

0011] It has been identified in accordance with the present invention that increased expression of a protein of the Sm-like superfamily in a microorganism confers enhanced tolerance to stress and inhibitors such as sodium acetate, ammonium acetate, potassium acetate, vanillin, furfural, hydroxymethylfurfural (HMF) and H2O2. In accordance with the present invention, microorganisms can be genetically modified to increase the expression of a protein of the Sm-like superfamily to achieve enhanced tolerance to stress and inhibitors. Such genetically modified microorganisms are particularly useful for production of biofuels based on fermentation of biomass materials.

0012] In one aspect, the invention is directed to genetically modified microorganisms that display enhanced tolerance to stress and/or inhibitors as a result of increased expression of a protein of the Sm-like superfamily in the microorganisms.

0013] In one embodiment, the microorganism is a genetically engineered bacterial strain, and the protein being expressed on an elevated level is a bacterial Hq protein.

0014] Bacteria contemplated by the present invention include both Gram-negative and Gram positive bacteria. Examples of bacteria of particular interest include Acetobacterium, Bacillus, Streptococcus, Clostridium (e.g., C. thermocellum), Zymomonas sp. (e.g., Z. mobilis), Anaerocellum (e.g., Anaerocellum thermophilum), Caldicellulosiruptor (e.g., C. saccharolyticus), Thermoaerobacter (e.g., Thermoaerobacter sp. X514), Gluconobacter, and E. coli.

0015] Bacterial strains that display enhanced tolerance to stress and/or inhibitors can be generated, e.g., by introducing to a bacterial strain an expression vector which includes the coding sequence of a bacterial Hq protein. The expression vector directs the expression of the Hq protein as a replicative plasmid, or mediates the integration of the coding sequence into the host genome to achieve chromosomal expression. Preferably, the bacterial Hq protein in the vector is identical with or substantially homologous with an endogenous Hq protein of the recipient bacterial strain.

0016] In specific embodiments, the expression vector includes the coding sequence of a bacterial Hq protein having an amino acid sequence selected from the group consisting of SEQ ID NO: 2 (Z. mobilis ZM4), SEQ ID NO: 4 (E. coli), SEQ ID NO: 6 (Clostridium thermocellum), SEQ ID NO: 8 (Anaerocellum thermophilum), SEQ ID NO: 10 (Caldicellulosiruptor saccharolyticus), SEQ ID NO: 12 (Thermoaerobacter sp. X514), and functional derivatives thereof.

0017] In a further embodiment, the invention is directed to genetically engineered fungal strains that display enhanced tolerance to stress and/or inhibitors. Examples of fungi include Saccharomyces sp. (e.g., S. cerevisiae), Kluyveromyces sp., Pichia sp. (e.g., Pichia pastoris), Candida sp., and Debrosaccharomyces sp.

0018] Such fungal strains can be generated, e.g., by introducing to a fungal strain an expression vector which includes the coding sequence of a fungal protein of the Sm-like superfamily. Similarly, the expression vector can be a replicative vector or integrative vector. Preferably, the fungal protein of the Sm-like superfamily in the expression vector is identical with or substantially homologous with an endogenous Sm-like protein of the fungal strain.

0019] In specific embodiments, the expression vector includes the coding sequence of a fungal protein of the Sm-like superfamily having an amino acid sequence selected from the group consisting of SEQ ID NO: 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 and 50 (representing 19 S. cerevisiae Sm and Lsm proteins) and functional derivatives thereof.

0020] The genetically modified microorganisms that display enhanced tolerance to stress and inhibitors can be additionally modified as appropriate, for example, by transformation with additional recombinant genes or sequences suitable for fermentation and production of ethanol. For example, the
bacterial and fungal strains can be additionally modified so as to have the ability to utilize C5 sugars such as xylose and arabinose in addition to C6 sugars.

**[0021]** In a further aspect, the present invention provides a method of producing biofuels from cellulosic biomass based on use of the microbial strains that are able to grow at elevated concentrations of inhibitors and/or under stress conditions.

**BRIEF DESCRIPTION OF THE DRAWINGS**

**[0022]** FIGS. 1A-1E. Domain and motif sites of Z. mobilis HfQ (A), E. coli Hfq (B), S. cerevisiae Sm (C), and S. cerevisiae Lsm1 (D) proteins. Bacterial Hfq alignment and Clustal W (E). Residues that are identical across the species are indicated by **“”**, and residues that are not identical but conserved in function across the species are indicated by **“”**.

**[0023]** FIG. 2A. Graphic map of the low copy number Gateway® compatible plasmid pBRR3-DEST42. Tc(R): Tetracycline resistance gene tet; Cm: chloramphenicol resistance gene cat, attR1 and attR2 are recombination sites allowing recombinational cloning of the gem of interest from an entry clone; ccdB is ccdB gene allowing negative selection of expression clones.

**[0024]** FIG. 2B. The insertion position and complementation region of ZMO0347 as well as the primers and mutation position. ZMO0346, ZMO0347, and ZMO0348 are Z. mobilis ZM4 genes. hFq_MF and hFq_MR are primers used for insertional mutation construction using pkNOCK mutagenesis system. Hfq_Cf and Hfq_CR are primers used to clone the hFq gene into pBRR3-DEST42 for complementation, which resulted in a plasmid called as p42-0347. The primer sequences are: hFq_MF: cggagagttgtgattgcaac (SEQ ID NO: 51); Hfq_Cf: atggccgaaagatctgca (SEQ ID NO: 53); Hfq_CR: atctgacgctgcctgcttc (SEQ ID NO: 54).

**[0025]** FIGS. 3A-3C. HfQ is responsible for sodium acetate tolerance of Z. mobilis. Z. mobilis strains were grown in RM (pH 5.0) overnight, 20-µL culture were then transferred into 250-µL RM media in the Bioscreen plate. The growth differences of different strains were monitored by Bioscreen (GrowthCurve, Mass.) under anaerobic conditions in RM (pH 5.0) containing 0, 12, and 16 g/L NaAc (A, B, C respectively). Strains included in this study are: ZM4: Zymomonas mobilis ZM4 wild-type; AcR: ZM4 acetate tolerant mutant (Joachimsthal 1998); ZM4 (p42-0347); ZM4 containing a gateway plasmid p42-0347 over-expressing ZM4 gene ZMO0347; AcRIM0347 (p42-0347): AcR insertional mutant of ZMO0347; AcRIM0347 (p42-0347): AcRIM0347 containing gateway plasmid p42-0347. This experiment has been repeated at least three times with similar result. Triplicates were used for each condition.

**[0026]** FIGS. 4A-4E. HfQ contributes to Z. mobilis acetate tolerance. Z. mobilis strains were grown in RM (pH 5.0) overnight, 5-µL culture were then transferred into 250-µL RM media in the Bioscreen plate. The growth differences of different strains were monitored by Bioscreen (GrowthCurves USA, NJ) under anaerobic conditions; in RM, pH 5.0 (A), RM with 195 mM NaCl, pH 5.0 (B), 195 mM NaAc, pH 5.0 (C), 195 mM NH₄OAc, pH 5.0 (D), or 195 mM KAc, pH 5.0 (E). Strains included in this study are: ZM4: Zymomonas mobilis ZM4 wild-type; AcR: ZM4 acetate tolerant mutant; ZM4 (p42-0347): ZM4 containing a gateway plasmid p42-0347 to express ZM4 gene ZMO0347; AcRIM0347: AcR insertional mutant of ZMO0347; AcRIM0347 containing gateway plasmid p42-0347. This experiment has been repeated at least three times with similar result. Duplicate biological replicates were used for each condition.

**[0027]** FIGS. 5A-5E. Z. mobilis Hfq conferred tolerance to different classes of pretreatment inhibitors. Z. mobilis strains were grown in RM (pH 5.0) overnight, 2-µL culture were then transferred into 250-µL RM media in the Bioscreen plate. The growth differences of different strains were monitored by Bioscreen (GrowthCurves USA, NJ) under anaerobic conditions in RM, pH 5.0 (A), RM with 1 g/L vanillia, pH 5.0 (B), 1 g/L furfural, pH 5.0 (C), 1 g/L HMF; pH 5.0 (D) and 0.001% H₂O₂ (E). Strains included in this study are: ZM4: Zymomonas mobilis ZM4 wild-type; AcR: ZM4 acetate tolerant mutant; AcRIM0347: AcR insertional mutant of ZMO0347; AcRIM0347 (p42-0347): AcRIM0347 containing gateway plasmid p42-0347 over-expressing ZM4 gene ZMO0347. This experiment has been repeated at least three times with similar result for hydrogen peroxide growth and in duplicate for the vanillia growth.

**[0028]** FIGS. 6A-6B. Lsm-like proteins in S. cerevisiae are responsible for sodium acetate tolerance. S. cerevisiae strains were grown in CM with 2% glucose for wild-type BY4741 and the deletion mutants, CM with 2% glucose minus uracil for GST over-expression strains. Five-µL culture was then transferred into 300-µL CM broth in the Bioscreen plate. The growth differences of different strains were monitored by Bioscreen (Growth Curve USA, NJ) containing 40 g/L sodium acetate for yeast deletion mutants (A) and GST over-expression strains (B). This experiment has been repeated at least three times with similar result.

**[0029]** FIGS. 7A-7F. Lsm proteins in S. cerevisiae are involved in multiple inhibitor tolerance. S. cerevisiae strains were grown in CM with 2% glucose (CM+glucose) for wild-type BY4741 and the deletion mutants, CM with 2% glucose and 2% galactose minus uracil (CM+glucose+2% galactose) for GST over-expression strains. A 5-µL culture was then transferred into 250-µL CM broth in the Bioscreen plate. The growth differences of different deletion mutant strains were monitored by Bioscreen (Growth Curves USA, NJ) in CM+glucose at pH 5.5 (A), CM+glucose with 305 mM NaCl, pH 5.5 (B), 305 mM NaAc, pH 5.5 (C), 305 mM NH₄OAc, pH 5.5 (D), and 305 mM KAc, pH 5.5 (E), 0.75 g/L vanillia, pH 5.5 (F), 1.5 g/L furfural, pH 5.5 (G), and 1.5 g/L HMF, pH 5.5 (H). The growth differences of different GST over-expressing strains were monitored by Bioscreen (Growth Curves USA, NJ) in CM+glucose+2% galactose at pH 5.5 (I), CM+glucose+2% galactose with 305 mM NaCl, pH 5.5 (J), 305 mM NaAc, pH 5.5 (K), 305 mM NH₄OAc, pH 5.5 (L), 305 mM KAc, pH 5.5 (M), 0.75 g/L vanillia, pH 5.5 (N), 1.5 g/L furfural, pH 5.5 (O), and 1.5 g/L HMF, pH 5.5 (P). Strains included in this study are listed in table 1. This experiment has been repeated at least three times with similar result.

**DETAILED DESCRIPTION OF THE INVENTION**

**[0030]** It has been identified in accordance with the present invention that increased expression of a protein of the Sm-like superfamily in a microorganism confers enhanced tolerance to stress and inhibitors. Based on this discovery, the present invention provides strains of microorganisms displaying enhanced tolerance to stress and/or inhibitors, which are particularly advantageous for use in fermentation of biomass materials to produce biofuels.
In one aspect, the invention is directed to genetically modified strains of microorganisms that display enhanced tolerance to stress and/or growth inhibitor as a result of increased expression of a protein of the Sm-like superfamily in the microorganisms.

Sm-like superfamily proteins are a highly conserved family of proteins found in eukaryotes, archaea and bacteria, and are characterized by an Sm-like superfamily domain having two conserved motifs referred to as Sm1 motif and Sm2 motif. The Sm1 and Sm2 motifs were first defined for human Sm snRNP proteins (Hermann et al. 1995), and were subsequently found to be highly conserved in other Sm and Lsm (Sm-like) proteins in eukaryotes including plant, drosophila, C. elegans, and S. cerevisiae. Eukaryotic Sm and Lsm proteins are integral to RNA processing and mRNA degradation complexes. Subsequently, the E. coli global response regulator Hfq was reported to be a homolog of the Sm and Lsm proteins (Zhang et al. 2002). The bacterial Hfq proteins contain a first region that shares significant similarity with the Sm1 motif found in eukaryotes, and a second region of particularly high conservation among the bacterial proteins which contains a number of conserved hydrophobic residues that align with hydrophobic residues found in the Sm2 motif of eukaryotic cells (Zhang et al. 2002). Similar to the eukaryotic Sm and Lsm proteins, the E. coli Hfq protein also forms a multimeric ring and is believed to also function to enhance RNA-RNA pairing.

As used herein, the term “Sm-like superfamily” includes both Sm and Lsm proteins of eukaryotes and archaea, and Hfq proteins of bacteria. A eukaryotic protein is considered to be a protein of the Sm-like superfamily in the context of the present invention if the protein contains an Sm-like superfamily domain characterized by the Sm1 motif and Sm2 motif defined by Hermann et al. (1995). Specifically, the Sm1 motif typically spans 32 amino acids, with positions 13 and 23 being Gly and Asn, respectively, positions 1, 3, 11, 15, 18 and 26 being a hydrophobic residue, and positions 19 and 31 being an acidic amino acid (Asp or Glu). The Sm2 motif typically spans only 14 amino acids, and has the consensus sequence (I or L)R or K(G or C) at positions 6-8, with positions 1, 4, 11, 13 and 14 being a hydrophobic residue, and positions 9-10 being a hydrophilic residue. Examples of eukaryotic proteins of the Sm-like superfamily include S. cerevisiae Sm B, Sm D1, Sm D2, Sm D3, Sm E, Sm F, Sm G, Lsm1, Lsm2, Lsm3, Lsm4, Lsm5, Lsm6, Lsm7, Lsm8, Lsm9, Lsm 13, and Lsm16 (SEQ ID NOS: 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 and 50, respectively). The locations of Sm1 and Sm2 motifs are illustrated for Sm B and Lsm1 proteins in FIGS. 1C-1D.

In the context of the present invention, a bacterial protein is considered to be an Hfq protein and therefore a protein of the Sm-like superfamily if the bacterial protein contains an Sm-like superfamily domain characterized by a first motif similar to the Sm1 motif of eukaryotic proteins defined above, and a second highly conserved region. Generally, the bacterial Sm1 motif spans 26 amino acids, and like the eukaryotic Sm1 motif, has Gly at position 13, an acidic amino acid at position 19 (typically Asp), a hydrophobic residue at positions 1 (preferably V), 3, 11 (preferably L), 15 and 18 (preferably F). Additionally, the second highly conserved region (“the bacterial Sm2 motif”) generally spans 12 amino acids, and has a “KIAA” sequence at positions 6-8, and preferably, with Y, I and S at positions 5, 9, and 10, respectively, and with a hydrophobic residue at positions 3-4 and 11. Examples of bacterial Hfq proteins include SEQ ID NO: 2 (Z. mobilis ZM4), SEQ ID NO: 4 (E. coli), SEQ ID NO: 6 (Clostridium thermocellum), SEQ ID NO: 8 (Anaerocellum thermophilum), SEQ ID NO: 10 (Caldicellulosiruptor saccharolyticus), and SEQ ID NO: 12 (Thermoanaerobacter sp. X514). Alignment of these bacterial Hfq proteins is provided in FIG. 1B, and the locations of the Sm-like superfamily domain including the Sm1 and Sm2 motifs of Z. mobilis and E. coli proteins are illustrated in FIG. 1A. It is clear that these six bacterial Hfq proteins share significant homologies, having conserved V, L, G, L, G, F, D and F at positions 1, 5, 8, 11, 13, 18, 19 and 21 of the Smal motif, and Y, K, H, A, I, and S at positions 5-10 of the Sm2 motif.

Functional derivatives and homologs of a given protein of the Sm-like superfamily are also suitable for use in the present invention. As used herein, “functional derivatives” and “homologs” of a protein of the Sm-like superfamily refer to proteins that share at least 45% identity or similarity, or preferably at least 50%, 60%, 75%, or 85% identity or similarity, or more preferably 90%, 95%, 98%, or 99% identity or similarity, with the protein of the Sm-like superfamily. Similarity between two protein sequences can be determined, for example, using the well known Lipman-Pearson Protein Alignment program with the following choice of parameters: Ktupple 2, Gap Penalty 4, and Gap Length Penalty 12. Preferably, the derivatives and homologs share consensus motifs of the Sm-like superfamily, which are believed to be critical to the function of the proteins.

A functional derivative of a given protein includes derivatives where modifications are made to non-conserved residues, as well as a functional or enzymatically active fragment of the protein. The term “functional fragment” or “enzymatically active fragment” means a polypeptide fragment of a full length protein, which substantially retains the activity of the full-length protein. By “substantially” it is meant at least about 50%, or preferably at least 70%, or even 80% or more of the activity of the full-length protein is retained.

The genetically engineered microbial strains of the present invention display enhanced tolerance to stress and/or one or more inhibitors as a result of increased expression of a protein of the Sm-like superfamily.

The term “stress”; as used herein, refers generally to environmental stress, i.e., stress received from the environment, such as high temperatures, low temperatures, low pH, oxidation (i.e., the presence of reactive oxidative species such as H2O2), osmotic, drought, the presence of inhibitors, or nutrient limit such as starvation, among others. For example, “cold stress” is stress on microorganism due to exposure to environments below the minimum optimal growth temperature of the microorganism. “Drought stress” is stress due to exposure of the microorganism to environments under the minimum optimal growth moisture concentration. “Osmotic stress” is stress on microorganisms due to exposure of the microorganisms to environments over or under the maximum or minimum optimal growth osmotic of the microorganisms.

The term “inhibitors” as used herein refer particularly to inhibitory chemical compounds that are formed during biomass pretreatment, including sugar degradation products such as furfural and hydroxymethyl furfural (HMF), weak acids such as acetic, formic, and levulinic acids, lignin degradation products such as the substituted phenolics vanillin and lignin monomers, reactive oxidative species generating hydrogen peroxide (H2O2) and vanillin, as well as meta-
bolic byproducts such as ethanol, lactate, and acetate. A particularly desirable trait of microorganisms is an enhanced tolerance to sodium and acetate ions, e.g., in the form of sodium acetate, ammonium acetate, and potassium acetate. [0041] In the present invention, microorganisms with enhanced tolerance to stress and/or one or more inhibitors refer to microorganisms which, as a result of genetic modification to increase the level of proteins of the Sm-like superfamily in the microorganisms, demonstrate improved tolerance as compared to microorganisms without the genetic modification. Improved tolerance can be determined by an improved growth profile (either as a shorter lag phase, a shorter doubling time, or a higher maximum density) under a given stress condition or inhibitor concentration. Alternatively, improved tolerance can be determined by an increase in the concentration of an inhibitory molecule which the microorganisms can tolerate.

For example, microorganisms having an elevated expression of Sm-like superfamily proteins exhibit enhanced tolerance to acetate. "Tolerance to acetate" is meant herein to include resistance to acetate salts including, for example, sodium acetate, ammonium acetate and potassium acetate, and/or to acetic acid. Tolerance of a strain to acetate can be determined by assessing the growth of the strain in media containing various concentrations of acetate (e.g., sodium acetate). The microbial strains containing a desirable genetic modification of the present invention are able to grow in media containing a higher concentration of acetate (e.g., sodium acetate) than the unmodified strains. For example, the concentration of sodium acetate that can be tolerated by a strain can be increased by 15%, 20%, 30%, or 50% or higher, as a result of a genetic modification. As demonstrated herein below, wild type Z. mobilis strain ZM4 is unable to grow in media containing 16 g/L (195 mM) sodium acetate, while ZM4-p42-0SP47 (expressing additional ZM4 hflQ proteins) is able to grow at this concentration. Alternatively, "enhanced tolerance" can be measured by a shorter lag time (e.g., shortened by 10%, 20%, 30% or 50% or greater), a shorter doubling time (e.g., shortened by 10%, 20%, 30% or 50% or greater) or a higher cell density reached at the end of the exponential growth phase (e.g., 25%, 50%, 75%, 100%, 150%, 200%, 500%, or even 1000% or higher cell density). See FIGS. 3A-3C.

Microorganisms encompassed within the scope of the present invention include both bacteria and fungi. In accordance with the present invention, bacterial strains having enhanced tolerance to stress and inhibitors as a result of increased expression of Sm-superfamily proteins include both Gram-positive and Gram-negative bacteria. Examples of Gram-positive bacteria include those from the genus of *Clostridium*, particularly strains of *Acetobacterium*, *Bacillus*, *Streptococcus*, *Clostridium* (e.g., *C. thermocellum*), *Anaerococcus* (e.g., *Anaerococcus thermophili*), *Caldicellulosiruptor* (e.g., *C. saccharolyticus*), and *Thermoanaerobacter* (e.g., *Thermoanaerobacter sp. X514*). Examples of Gram-negative bacteria of particular interest include those generally considered medically safe, such as *Zymomonas* sp. (e.g., *Z. mobilis*), *E. coli*, *Gluconobacter* sp. (e.g., *Gluconobacter oxydans*, previously known as *Acetobacter suboxydans*), *Cyanobacteria*, *Green sulfur* and *Green non-sulfur bacteria*.

Fungal strains contemplated by the present invention include filamentous and unicellular fungal species, particularly the species from the class of *Ascomycota*, for example, *Saccaromyces* sp., *Kluyveromyces* sp., *Pichia* sp., *Candida* sp., and *Schizosaccharomyces* sp. Preferred fungal strains contemplated by the present invention are *S. cerevisiae*, *S. pombe* and *Pichia pastoris*. Where the fungal strains are *S. cerevisiae*, additional genetic modifications are preferred besides the genetic modification that results in an increased expression of a Sm-like superfamily protein. For example, *S. cerevisiae* is also modified such that the strain is able to utilize C5 sugars. Strains of microorganisms that display enhanced tolerance to stress and/or inhibitors as a result of increased expression of a Sm-like superfamily protein can be made using any of the known genetic engineering techniques. For example, the 5' upstream regulatory region of an endogenous Sm-like superfamily gene can be modified to achieve enhanced expression of the encoded endogenous Sm-like superfamily protein.

In one embodiment, a microbial strain having enhanced tolerance is created by introducing an exogenous expression vector into the strain which contains the coding sequence of a protein of the Sm-like superfamily. In a preferred embodiment, the protein encoded by the expression vector is identical with an endogenous protein of the Sm-like superfamily or a functional derivative thereof, even though homologs from other related species can also be utilized.

Generally, the nucleotide sequence coding for a protein of the Sm-like superfamily is placed in an operably linkage to a promoter and a 3' termination sequence that are functional in a recipient microbe. The promoter can be a constitutive promoter or an inducible promoter. The promoter can be the native promoter of the Sm-like superfamily gene being expressed, or a heterologous promoter from a different gene. Promoters suitable for use in expression in a bacterial host include, for example, lac, T7 and T3 promoters and SP6 phage RNA polymerase promoters. Specific examples of promoters suitable for use in expression in *Zymomonas* species include *Z. mobilis* pdc promoter and adhB promoter. Specific examples of promoters suitable for use in expression in yeast including *S. cerevisiae* include ADE2 (constitutive high expression), HSP1 (carbon source responsive), a tetracycline-repressible system based on the CaMV promoter, and the mntl (no message in thiamine) promoter. These and other examples of promoters are well documented in the art.

A variety of vector backbones can be used for purposes of the present invention. Choices of vectors suitable for transformation and expression in bacteria and fungi have been well documented in the art. For example, numerous plasmids have been reported for transformation and expression in *Zymomonas*, including, e.g., pCB serial plasmids developed based on *Zymomonas* cryptic plasmid, as described in U.S. Patent Nos. 5,712,133, 5,726,053, and 5,843,760, and a cloning-compatible broad-host-range destination vector described by Pelletier et al. (2008), among many others.

In addition to the Sm-like superfamily protein expression unit, the expression vector can include other sequences where appropriate, such as sequences for maintenance and selection of the vector, e.g., a selection marker gene and a replication origin. The selection marker gene can be a gene that confers resistance to antibiotics such as ampicillin resistance (Amp'), tetracycline resistance (Tet'), neomycin resistance, hygromycin resistance, and zeocin resistance (Zeo') genes, or a gene that provides selection based on media supplement and nutrition.

The vector can be a replicative vector (such as a replicating circular plasmid), or an integrative vector which mediates the introduction of the vector into a recipient cell and subsequent integration of the vector into the host genome for chromosomal expression.
For industrial applications, the inhibitors generated from the biomass pretreatments will select for plasmid maintenance where hfaq expression confers an advantage to the strain (i.e., enhanced tolerance to inhibitors) in the absence of additional marker or antibiotic selection. The vectors can also be modified to include the parDE genes to enhance plasmid stability in bacteria in the absence of selection using standard molecular biology approaches, as described in the art (Brown et al., 2002; Pecota et al., 1997). Alternatively and preferably, the desired expression unit (such as an hfaq coding sequence operably linked to a promoter) is integrated into the chromosome of the microorganism for expression and enhanced stability. Methods for chromosomal integration in bacteria include modified homologous Campbell-type recombination (Kalogeraki et al. 1997) or transposition (Koch et al. 2001). Methods for chromosomal integration in yeast are well known and are described in Amberg et al. (2005).

An expression vector can be introduced into a microbial host by various approaches known in the art, including transformation (e.g., chemical reagent based transformation), electroporation and conjugation.

The genetic modification to a microbial strain results in an increased expression of a Sm-like superfamily protein. Where the exogenously introduced expression unit codes for a protein identical with an endogenous protein, the level of such protein (expressed from both the native sequence and the exogenous sequence) is increased. Where the exogenously introduced expression unit codes for a protein that is not identical with any endogenous protein but is a functional derivative of or most homologous to an endogenous protein, the collective level of the endogenous protein and the exogenous protein is increased as compared to the unmodified strain. The extent of increase in expression contemplated by the present invention is at least 40%, 50%, 75%, 100% (i.e., twice the level of parental strain), or more preferably at least four or five times, or even more preferably at least ten to fifteen times, the level of parental strain. As a practical matter, the level of expression can be assessed both at the mRNA level and at the protein level.

Pretreatment of biomass by chemical or enzymatic methods yields a mixture of hexose sugars (C6 sugars, primarily glucose and mannose) and pentose sugars (C5 sugars, primarily xylose and arabinose). The fermentation of almost all the available C6 and C5 sugars to ethanol or other liquid biofuel is critical to the overall economics of these processes. Most microorganisms are able to ferment glucose but few have been reported to utilize xylose efficiently and even fewer ferment this pentose to ethanol.

The genetically modified strains of microorganisms of the present invention, which display enhanced tolerance to stress and/or one or more inhibitors as a result of increased expression of a Sm-like superfamily protein, can be additionally modified as appropriate. For example, Z. mobilis strains overexpressing Z. mobilis Hfaq can be additionally modified in order to expand the range of substrates that can be utilized by the strains for efficient ethanol production. For instance, Z. mobilis strains over-expressing Hfaq can also be introduced with additional genes so that the strains can ferment xylose, arabinose or other pentose sugars as the sole carbon source to produce ethanol. See, e.g., U.S. Pat. No. 5,514,583. Additionally, yeast strains over-expressing a Sm or Lsm protein, particularly S. cerevisiae strains, can be additionally modified to have an enhanced ability to ferment xylose, arabinose or other pentose sugars to produce ethanol. For example, yeast cells can be modified to overexpress (via transformation with additional expression unit) xylose reductase, xylulokinase, or xylose isomerase; or modified to have reduced expression of xyitol dehydrogenase, PHO13 or a PHO13 ortholog. See, e.g., U.S. Pat. No. 7,285,403, US 20060234364 A1, and US 20080254524 A1, the teachings of which are incorporated herein by reference.

The isolated or genetically modified microbial strains of the present invention are particularly useful for production of biofuels based on fermentation of biomass materials. Therefore, in a further aspect, the present invention provides a method of producing biofuels from cellullosic biomass based on use of the microbial strains of the present invention that are able to grow at elevated concentrations of acetate.

Biofuels contemplated by the present invention include in particular the types of biologically produced fuels, such as bioalcohols, based on the action of microorganisms and enzymes through fermentation of biomass materials. Examples of bioalcohols include ethanol, butanol, and propanol.

In a typical cellululosic biomass to alcohol process, raw cellullosic biomass material is pretreated in order to convert, or partially convert, cellullosic and hemicellulosic components into enzymatically hydrolyzable components (e.g., poly- and oligo-saccharides). The pretreatment process also serves to separate the cellullosic and hemicellulosic components from solid lignin components also present in the raw cellullosic material. The pretreatment process typically involves reacting the raw cellullosic biomass material, often as a finely divided mixture or slurry in water, with an acid, such as sulfuric acid. Other common pretreatment processes include, for example, hot water treatment, wet oxidation, steam explosion, elevated temperature (e.g., boiling), alkaline treatment and/or ammonia fiber explosion. The pretreated biomass is then treated by a saccharification step in which poly- and oligo-saccharides are enzymatically hydrolyzed into simple sugars. The free sugars and/or oligosaccharides produced in the saccharification step are then subjected to fermentation conditions for the production of ethanol or butanol, for example. Fermentation can be accomplished by combining one or more fermenting microorganisms with the produced sugars under conditions suitable for fermentation.

One can also add enzyme to the fermentor to aid in the degradation of substrates or to enhance alcohol production. For example, cellulases can be added to degrade cellulose to glucose simultaneously with the fermentation of glucose to ethanol by microorganisms in the same fermentor. Similarly, a hemicellulase can be added to degrade hemicellulose.

Because the pretreatment processes and by-products of fermentation can create a range of inhibitors including acetate, it is especially advantageous to utilize the genetically modified microbial strains described herein which display enhanced resistance to acetate and are able to continue fermentation despite acetate present in the fermentation broth, either in the fermentation substrate carried over from pretreatment of biomass material, or built up as a byproduct of fermentation.

For purpose of fermentation, one strain or a mixture of several strains, some or all of which display enhanced tolerance to stress and/or inhibitors, can be used.

Specific fermentation conditions can be determined by those skilled in the art, and may depend on the particular feedstock or substrates, the microorganisms chosen and the type of biofuel desired. For example, when Zymomonas mobilis is employed, the optimum pH conditions range from about 3.5 to about 7.3; substrate concentrations of up to about 25% (based on glucose), and even higher under certain conditions, may be used; and no oxygen is needed at any stage for
microorganism survival. Agitation is not necessary but may enhance availability of substrate and diffusion of ethanol. **[0065]** After fermentation, alcohol is separated from the fermentation broth by any of the many conventional techniques known to separate alcohol from aqueous solutions, including evaporation, distillation, solvent extraction and membrane separation. Particles of substrate or microorganisms may be removed before separation to enhance separation efficiency.

**[0066]** Table 1. List all the sequence identifiers for the nucleotide and protein sequences of the Sm-like superfamily molecules exemplified in the present application.

**[0067]** The present invention is further illustrated and by no means limited by the following examples.

### TABLE 1

<table>
<thead>
<tr>
<th>SEQ ID NO</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Zymomonas mobilis hq nucleotide</td>
</tr>
<tr>
<td>2</td>
<td>Zymomonas mobilis Hq amino acid</td>
</tr>
<tr>
<td>3</td>
<td>E. coli hq nucleotide</td>
</tr>
<tr>
<td>4</td>
<td>E. coli Hq amino acid</td>
</tr>
<tr>
<td>5</td>
<td>Clostridium thermocellum hq nucleotide</td>
</tr>
<tr>
<td>6</td>
<td>Clostridium thermocellum Hq amino acid</td>
</tr>
<tr>
<td>7</td>
<td>Azotobacter vinelandii hq nucleotide</td>
</tr>
<tr>
<td>8</td>
<td>Azotobacter vinelandii Hq amino acid</td>
</tr>
<tr>
<td>9</td>
<td>Caldicellulosiruptor saccharolyticus hq nucleotide</td>
</tr>
<tr>
<td>10</td>
<td>Caldicellulosiruptor saccharolyticus Hq amino acid</td>
</tr>
<tr>
<td>11</td>
<td>Thermotoga maritima sp. X85 hq nucleotide</td>
</tr>
<tr>
<td>12</td>
<td>Thermotoga maritima sp. X85 Hq amino acid</td>
</tr>
<tr>
<td>13</td>
<td>S. cerevisiae SNM1 nucleotide</td>
</tr>
<tr>
<td>14</td>
<td>S. cerevisiae Sn B amino acid</td>
</tr>
<tr>
<td>15</td>
<td>S. cerevisiae SNM1 nucleotide</td>
</tr>
<tr>
<td>16</td>
<td>S. cerevisiae Sn D1 amino acid</td>
</tr>
<tr>
<td>17</td>
<td>S. cerevisiae SM2 nucleotide</td>
</tr>
<tr>
<td>18</td>
<td>S. cerevisiae Sm D2 amino acid</td>
</tr>
<tr>
<td>19</td>
<td>S. cerevisiae SM3 nucleotide</td>
</tr>
<tr>
<td>20</td>
<td>S. cerevisiae Sm I3 amino acid</td>
</tr>
<tr>
<td>21</td>
<td>S. cerevisiae SME1 nucleotide</td>
</tr>
<tr>
<td>22</td>
<td>S. cerevisiae Sm E amino acid</td>
</tr>
<tr>
<td>23</td>
<td>S. cerevisiae SMX3 nucleotide</td>
</tr>
<tr>
<td>24</td>
<td>S. cerevisiae Sm F amino acid</td>
</tr>
<tr>
<td>25</td>
<td>S. cerevisiae SMX2 nucleotide</td>
</tr>
<tr>
<td>26</td>
<td>S. cerevisiae Sm G amino acid</td>
</tr>
<tr>
<td>27</td>
<td>S. cerevisiae LSM1 nucleotide</td>
</tr>
<tr>
<td>28</td>
<td>S. cerevisiae Lsm1 amino acid</td>
</tr>
<tr>
<td>29</td>
<td>S. cerevisiae LSM2 nucleotide</td>
</tr>
<tr>
<td>30</td>
<td>S. cerevisiae Lsm2 amino acid</td>
</tr>
<tr>
<td>31</td>
<td>S. cerevisiae LSM3 nucleotide</td>
</tr>
<tr>
<td>32</td>
<td>S. cerevisiae Lsm3 amino acid</td>
</tr>
<tr>
<td>33</td>
<td>S. cerevisiae LSM4 nucleotide</td>
</tr>
<tr>
<td>34</td>
<td>S. cerevisiae Lsm4 amino acid</td>
</tr>
<tr>
<td>35</td>
<td>S. cerevisiae LSM5 nucleotide</td>
</tr>
<tr>
<td>36</td>
<td>S. cerevisiae Lsm5 amino acid</td>
</tr>
<tr>
<td>37</td>
<td>S. cerevisiae LSM6 nucleotide</td>
</tr>
<tr>
<td>38</td>
<td>S. cerevisiae Lsm6 amino acid</td>
</tr>
<tr>
<td>39</td>
<td>S. cerevisiae LSM7 nucleotide</td>
</tr>
<tr>
<td>40</td>
<td>S. cerevisiae Lsm7 amino acid</td>
</tr>
<tr>
<td>41</td>
<td>S. cerevisiae LSM8 nucleotide</td>
</tr>
<tr>
<td>42</td>
<td>S. cerevisiae Lsm8 amino acid</td>
</tr>
<tr>
<td>43</td>
<td>S. cerevisiae LSM9 nucleotide</td>
</tr>
<tr>
<td>44</td>
<td>S. cerevisiae Lsm9 amino acid</td>
</tr>
<tr>
<td>45</td>
<td>S. cerevisiae LSM12 nucleotide</td>
</tr>
<tr>
<td>46</td>
<td>S. cerevisiae Lsm12 amino acid</td>
</tr>
<tr>
<td>47</td>
<td>S. cerevisiae LSM13 nucleotide</td>
</tr>
<tr>
<td>48</td>
<td>S. cerevisiae Lsm13 amino acid</td>
</tr>
<tr>
<td>49</td>
<td>S. cerevisiae LSM16 nucleotide</td>
</tr>
<tr>
<td>50</td>
<td>S. cerevisiae Lsm16 amino acid</td>
</tr>
</tbody>
</table>

**Strains and Culture Conditions**

**[0069]** Bacterial strains and plasmids used in this study are listed in Table 2. E. coli strains were cultured using Luria-Bertani (LB) broth or plates. E. coli WM3064 was supplemented with 100 µg/mL dianimopinolic acid (DAP). Z. mobilis ZM4 was obtained from the

**[0070]** American Type Culture Collection (ATCC31321) and the Z. mobilis acetate tolerant strain AcrB has been described previously (Joachimshal et al. 2000). ZM4 and AcrB were cultured in RM medium at 30°C. C. cerevisiae wild-type, deletion mutant and GST-fusion ORF over-expression strains were obtained through Open Biosystems (Huntsville, Ala.). S. cerevisiae strains were cultured in rich YPD media. CM media with 2% glucose was used for S. cerevisiae wild-type and S. cerevisiae deletion mutants, CM media with 2% glucose minus uracil was used for S. cerevisiae GST-over expressing strains, 2% galactose was used to induce the GST-fusion strains. Plasmid-containing strains were routinely grown with antibiotics at the following concentrations (µg/mL): kanamycin of 50 for E. coli and 200 for ZM4; tetracycline, 10 for E. coli and 20 for ZM4; and gentamicin, 10 for E. coli. G418 of 200 for S. cerevisiae YKO deletion mutants. Growth was monitored turbidometrically by measuring optical density at 600 nm periodically with the Bioscreen C automated microbiology growth curve analysis system (Growth Curve USA, Piscataway, N.J.).

**PCR and DNA Manipulations**

**[0071]** Genomic DNA from Z. mobilis was isolated using a Wizard Genomic DNA purification kit, following the manufacturer’s instructions (Promega, Madison, Wis.). The QIAprep Spin Miniprep and HiSpeed Plasmid Midi kits (Qiagen, Valencia, Calif.) were used for plasmid isolation, respectively. PCR, restriction enzyme digestion, DNA ligation, DNA cloning, and DNA manipulations were done following standard molecular biology approaches (Sambrook 2000).

**Construction of the Novel Tetracycline Resistant Gateway Entry Vector and ZM00347 Over-Expression Plasmid**

**[0072]** The construction of the broad-host-range, tetracycline resistant Gateway® compatible destination plasmid vector pBBR3DEST42 (FIG. 2A) was carried out essentially as described previously (Pelletier et al. 2008), except that pBBR MCS-3 tetracycline resistance cassette was used in this study instead of pBBRMCS-5 gentamicin resistance cassette used to construct pBBR3DEST42. Briefly, pBBR MCS-3 plasmid DNA was restricted with the KpnI and PvuI enzymes, treated with calf intestine alkaline phosphatase and purified using a Qiagen gel purification kit according to the manufacturer’s instructions (Qiagen, Valencia, Calif.). The recombinant region on pET-DEST42 vector DNA (Invitrogen, Carlsbad, Calif.) was PCR-amplified using the primers 42F and 42R that include KpnI and PvuI restriction sites as described previously (Pelletier et al. 2008). The gel-purified PCR product was ligated with pBBR MCS-3 KpnI/PvuI fragment with Fast-Link™ DNA Ligation Kit (Epigence, Madison, Wis.). Ligation products were transformed into E. coli DB3.1 chemically competent cells (Invitrogen, Carlsbad, Calif.) and the transformants were selected by plating on LB agar plates containing tetracycline. Individual colonies were grown overnight in LB containing 30 µg/mL chloramphenicol and 10 µg/mL tetracycline, and plasmid DNA was pre-
pared using QIAprep spin miniprep or HiSpeed Plasmid Midi Kit following the manufacturer’s protocol (Qiagen, Valencia, Calif.). Plasmid DNA was digested with KpnI and PvuII and digestion products were analyzed on an agarose gel to confirm the presence of products of the expected sizes.

[0073] The construction of entry vector and expression clone of target gene hfq (ZMO00347) was carried out as described previously (Pelletier et al. 2008). Briefly, target gene hfq (ZMO00347) was PCR amplified using AcR genomic DNA as template and primer hfq3_26 MF and hfq3_26 MR as primers. PCR products were then cloned into Gateway® entry clone pDONR221 using BP Clonase II enzyme mix following the manufacturer’s protocol (Invitrogen, Carlsbad, Calif.), and then transformed into chemically competent DH5α cells (Invitrogen, Carlsbad, Calif.) and plated onto LB with appropriate antibiotic selection. The inserts were confirmed by sequencing using M13 forward and reverse primers (Integrated DNA Technologies, Inc., Coralville, Iowa). The confirmed entry clone vector was then recombined with the destination vector pBHR3DEST42 using LR Clonase II enzyme mix (Invitrogen Carlsbad, Calif.) to create the expression vector as described previously (Pelletier et al. 2008). The resulting expression vector construct was designed as p42-0347. The plasmid construct p42-0347 was confirmed by sequencing using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems Inc., Foster City, Calif.).

Mutant Plasmid Construction

[0074] Briefly, a 262-bp hfq internal part PCR product was purified and cloned into pKnock-Km suicide vector (Alexeyev 1999) digested with XbaI and HindIII restriction enzymes followed by de-phosphorylation. The plasmid construct named as pKm-0347 was then sequenced to confirm the presence of the target gene fragment, which was then electroporated into E. coli WM3064 strain. The transformant E. coli WM3064 (pKm-0347) was verified by PCR and sequencing for the presence of correct plasmid construct pKm-0347. E. coli WM3064 (pKm-0347) was then conjugated with AcR. The conjugant of potential hfq mutant grown on RM plate with kanamycin concentration of 200 μg/mL and no DAP was selected based on PCR size shift by comparing the PCR size of wild-type AcR and conjugants using primer hfq_OCF and hfq_OCR (table 2). Wild-type AcR has a 1050-bp PCR product and hfq mutant candidates have a 2.9-kb PCR product. The PCR product was sequenced for mutant confirmation.

[0075] The internal part of the Z. mobilis hfq gene (ZMO00347) was amplified by PCR using primers hfq3_26 MF and hfq3_26 MR supplied by MWG-Biotech (Huntsville, Ala.). The hfq gene and the primer positions used for mutant construction and an hfq gene-expressing vector are shown in FIG. 2B. The 262-bp hfq internal part PCR product was then purified and cloned into pCR2.1-TOPO and then transformed into E. coli TOP10 one competent cell (Invitrogen, Carlsbad, Calif.). Transformants containing the correct construct were confirmed by PCR and sequencing. The plasmid was then extracted using Qiagen Midiprep and digested with XbaI and HindIII restriction enzyme, the 262-bp hfq internal part was then purified by Qiagen Gel purification kit. Similarly, pKnock-Km suicide vector was also digested with XbaI and HindIII restriction enzyme followed by de-phosphorylation, and then ligated with 262-bp purified hfq internal part using Fast-Link™ DNA Ligation Kit (Epigene, Madison, Wis.). The ligation product (pKm-0347) was then transformed into TransformMax EC100D pir-116 Electropotent E. coli competent cells (Epigene, Madison, Wis.) by electroporation. Transformants containing plasmid pKm-0347 were selected on LB agar plate with 50 μg/mL kanamycin. The plasmid was then extracted from the transformants, sequenced to confirm the presence of the target gene fragment, and was then electroporated into E. coli WM3064 strain. Transformants were verified by PCR and sequencing using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems Inc., Foster City, Calif.) for the presence of the correct plasmid construct pKm-0347.

Plasmid Transformation of Z. mobilis

[0076] Z. mobilis wild-type ZM4 and acetate tolerant strain AcR cultures were grown aerobically at 30°C in RM, and E. coli WM3064 containing plasmid pKm-0347 or p42-0347 cultures were grown at 37°C in LB containing 100 μg/mL DAP and 10 μg/mL tetracycline to exponential phase. E. coli WM3064 cells containing plasmid pKm-0347 or p42-0347 were washed with RM for three times by centrifugation at 13,000 rpm for 1 min and resuspended in RM. AcR cells were mixed with E. coli WM3064 (pKm-0347) cells in different ratios (1:3, 1:1, and 3:1). Similarly, ZM4 or AcR cells were mixed with E. coli WM3064 (p42-0347) cells in different ratios (1:3, 1:1, and 3:1). The mixtures of cells were plated on RM agar plates with 100 μg/mL DAP and 10 μg/mL tetracycline for plasmid p42-0347 conjugation or 50 μg/mL kanamycin for plasmid pKm-0347 conjugation. The cells were incubated at 30°C overnight. Conjugants were selected by plating on RM agar plates containing 20 μg/mL tetracycline for p42-0347 plasmid conjugants or 200 μg/mL kanamycin for pKm-0347 plasmid conjugants at 30°C. The conjugants were confirmed for the presence of correct plasmid constructs by PCR and sequencing using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems Inc., Foster City, Calif.).

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacterial strains, plasmids and primers used in this application</strong></td>
</tr>
<tr>
<td><strong>Strain, plasmid, or primer</strong></td>
</tr>
<tr>
<td><strong>E. coli</strong></td>
</tr>
<tr>
<td>K-12</td>
</tr>
<tr>
<td>DH5α</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
# TABLE 2 - continued

<table>
<thead>
<tr>
<th>Strain, plasmid, or primer</th>
<th>Genotype, phenotype, or sequence of primer (5' to 3')</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB3.1</td>
<td>p' gyrA462 endA1A (sr1-recA) mcrB mrr hsdR20 (rK-, mK-) supE44 ara-14 galK2 lacY1 proA2 rpsL20 (Sm^r) xyl1-5A-leu mtl1</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>WM0664</td>
<td></td>
<td>Denef et al. (2006)</td>
</tr>
<tr>
<td>BL21 (DE3)</td>
<td>p-ompT hsdSB (rB-mB-) gal dcm (DE3)</td>
<td>Invitrogen</td>
</tr>
</tbody>
</table>

**Zymomonas mobilis**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZM4</td>
<td>ATCC31821</td>
<td></td>
</tr>
<tr>
<td>AcR</td>
<td>ZM4 acetate tolerant strain generated by random mutagenesis</td>
<td>Joachimstahl et al. (1998)</td>
</tr>
<tr>
<td>ZM4(p42-0347)</td>
<td>ZM containing plasmid p42-0347</td>
<td>This application</td>
</tr>
<tr>
<td>AcRIM0347</td>
<td>Insertional mutant of AcR gene ZM00347</td>
<td>This application</td>
</tr>
<tr>
<td>AcRIM0347 (p42-0347)AcRIM0347 containing plasmid p42-0347</td>
<td>This application</td>
<td></td>
</tr>
</tbody>
</table>

**S. cerevisiae**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype, phenotype, or sequence of primer (5' to 3')</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BY4741</td>
<td>MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0-s288c background</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>YSC1021-547768</td>
<td>Yeast: Yeast Knock Out Strain, HHA1</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 14095</td>
<td>Accession: YLR136W</td>
<td></td>
</tr>
<tr>
<td>YSC1021-551633</td>
<td>Yeast: Yeast Knock Out Strain, VNX1</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 1123</td>
<td>Accession: YNL321W</td>
<td></td>
</tr>
<tr>
<td>YSC1021-553567</td>
<td>Yeast: Yeast Knock Out Strain, ARR3</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 5616</td>
<td>Accession: YPR201W</td>
<td></td>
</tr>
<tr>
<td>YSC1021-555633</td>
<td>Yeast: Yeast Knock Out Strain, NHX1</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 4290</td>
<td>Accession: YDR456W</td>
<td></td>
</tr>
<tr>
<td>YSC1021-551475</td>
<td>Yeast: Yeast Knock Out Strain, transporter</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 610</td>
<td>Accession: YMR034C</td>
<td></td>
</tr>
<tr>
<td>YSC1021-551260</td>
<td>Yeast: Yeast Knock Out Strain, PSR1</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 1498</td>
<td>Accession: YLL010C</td>
<td></td>
</tr>
<tr>
<td>YSC1021-551318</td>
<td>Yeast: Yeast Knock Out Strain, PSR2</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 1574</td>
<td>Accession: YLR019W</td>
<td></td>
</tr>
<tr>
<td>YSC1021-555189</td>
<td>Yeast: Yeast Knock Out Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 2341</td>
<td>Accession: YJR005W</td>
<td></td>
</tr>
<tr>
<td>YSC1021-554440</td>
<td>Yeast: Yeast Knock Out Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 1301</td>
<td>Accession: YNL124C</td>
<td></td>
</tr>
<tr>
<td>YSC1021-552226</td>
<td>Yeast: Yeast Knock Out Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 4214</td>
<td>Accession: YDR378C</td>
<td></td>
</tr>
<tr>
<td>YSC1021-556031</td>
<td>Yeast: Yeast Knock Out Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td>Clone Id: 7383</td>
<td>Accession: YNL147W</td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 2—continued

<table>
<thead>
<tr>
<th>Strain, plasmid, or primer</th>
<th>Genotype, phenotype, or sequence of primer (5' to 3')</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>YSC1021-552677</strong></td>
<td>Yeast Knock Out Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC1021-552563</strong></td>
<td>Yeast Knock Out Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC1021-552280</strong></td>
<td>Yeast Knock Out Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC1021-553518</strong></td>
<td>Yeast Knock Out Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC1021-553919</strong></td>
<td>Yeast Knock Out Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC4515-98809240</strong></td>
<td>Yeast GST-Tagged Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC4515-98810980</strong></td>
<td>Yeast GST-Tagged Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC4515-98807049</strong></td>
<td>Yeast GST-Tagged Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC4515-98805426</strong></td>
<td>Yeast GST-Tagged Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC4515-98809076</strong></td>
<td>Yeast GST-Tagged Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC4515-98808930</strong></td>
<td>Yeast GST-Tagged Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC4515-98806813</strong></td>
<td>Yeast GST-Tagged Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC4515-98806389</strong></td>
<td>Yeast GST-Tagged Strain</td>
<td>Open Biosystems</td>
</tr>
<tr>
<td><strong>YSC4515-98805850</strong></td>
<td>Yeast GST-Tagged Strain</td>
<td>Open Biosystems</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmids</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pENK2-Km</td>
<td>Km', mob, broad host range cloning vector, 1.8 kb</td>
<td>Alexeyev (1999)</td>
</tr>
<tr>
<td>pET-DEST42</td>
<td>Ap', Cm', C-terminal 6xHis and V5 epitope</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pBBR1MCS-3</td>
<td>Tc', mob, broad host range cloning vector</td>
<td></td>
</tr>
<tr>
<td>pBBR3DEST42</td>
<td>Cm'Tc', C-terminal 6xHis and V5 epitope</td>
<td>This application</td>
</tr>
<tr>
<td>pDONR221</td>
<td>Km', gateway entry vector Qm', N-terminal GST</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>p42-0347</td>
<td>pBBR3DEST42 containing ZM4 ZMD0347</td>
<td>This application</td>
</tr>
</tbody>
</table>
TABLE 2-continued

<table>
<thead>
<tr>
<th>Strain, plasmid, or primer</th>
<th>Genotype, phenotype, or sequence of primer (5' to 3')</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hfg MF</td>
<td>cggagaposgggtcagtcaca (SEQ ID NO: 51)</td>
<td>1050-bp</td>
</tr>
<tr>
<td>hfg MR</td>
<td>ttotgtgtcagcaatacg (SEQ ID NO: 52)</td>
<td>262-bp</td>
</tr>
<tr>
<td>hfg CF</td>
<td>atggcagaaaaggtcaca (SEQ ID NO: 53)</td>
<td>483-bp</td>
</tr>
<tr>
<td>hfg CR</td>
<td>tcaaatctcgctctgccct (SEQ ID NO: 54)</td>
<td></td>
</tr>
<tr>
<td>hfg GCF</td>
<td>ccaggtcagtcagtcaccagcc (SEQ ID NO: 55)</td>
<td>1050-bp</td>
</tr>
<tr>
<td>hfg OCR</td>
<td>caggtgacatcagtcaccagcc (SEQ ID NO: 56)</td>
<td></td>
</tr>
</tbody>
</table>

Example 2

[0077] This example describes the results of the experiments showing that overexpression of the *Zymomonas mobilis* global regulator gene *hfg* confers enhanced tolerance to sodium acetate.

[0078] The *Z. mobilis* hfg gene (SEQ ID NO: 1) was cloned into the vector pBBR3-DEST42 (FIG. 2A) and the resulting plasmid construct p42-0347 was transformed into the wild-type strain ZM4 and acetate mutant AcR through conjugation. In addition, an insertional mutant of *Z. mobilis* strain AcR hfg gene (ZMO0347) was generated using the pKNOCK system (Brown 2006; Alexeyev 1999) and complemented with plasmid p42-0347. The hfg gene and the primer positions used for mutant construction and hfg gene over-expression are shown in FIG. 2B. An insertional mutant of hfg (ZMO0347) was generated in the AcR background and designated as strain "AcRIM0347". The hfg gene was over-expressed via plasmid p42-0347 in both wild-type ZM4 and the acetate mutant AcR backgrounds, their susceptibilities to sodium acetate and other stressors were tested in growth assays along with strains ZM4 and AcR. The AcR acetate tolerant mutant has more tolerance to sodium acetate than its wild-type ZM4 parental strain (Joachimsthal et al. 1998); however, the insertional inactivation of the hfg gene in AcR reduces its sodium acetate tolerance (FIGS. 3A-3C). These strains were tested for their growth responses in four different concentrations of sodium acetate: 0, 12 g/L, 16 g/L (195 mM) and 20 g/L (FIGS. 3A-3C).

[0079] The sum of these data show that hfg expression contributed to sodium acetate tolerance. The AcRIM0347 mutant strain grew slightly more slowly in RM medium compared to the parental strain, i.e., in the absence of the sodium acetate stressor (FIG. 3A). Strains ZM4 and AcR with intact hfg genes grew faster than the hfg mutant AcRIM0347 strain in the presence of 12 g/L sodium acetate (FIG. 3B). The AcRIM0347 mutant phenotype was mostly restored by hfg expression and complementation via plasmid p42-0347. Similar, but more dramatic growth phenotypes were observed for sodium acetate of 16 g/L with the wild-type strain unable to grow at this concentration (FIG. 3C).

Example 3

[0080] This example describes the experiments performed to compare the negative effects of pretreatment inhibitors on *Z. mobilis* growth, and to demonstrate that Hfg overexpression confers tolerance to pretreatment inhibitors.

[0081] Pretreatment Inhibitors had Negative Effects on *Z. mobilis* Growth

[0082] The growth of *Z. mobilis* strains was reduced in the presence of acetate, vanillin, furfural, or HMF with increased lag phases and/or slower growth rates and/or final bacterial cell densities depending on the respective condition and strain (Tables 3-4; FIGS. 4A-4E and 5A-5E). Among the different forms of acetate counter-ions tested, sodium acetate had the most significant inhibitory effect on wild-type *Z. mobilis* growth. This was followed by potassium acetate and ammonium acetate, and sodium chloride had the least negative influence on wild-type *Z. mobilis* growth (Table 3; FIGS. 4A-4E). Wild-type ZM4 growth was completely inhibited when RM medium was supplemented with 195 mM sodium acetate (Table 3; FIG. 4C). Among the pretreatment inhibitors of vanillin, furfural, and HMF, vanillin had the most significant inhibitory effect on *Z. mobilis*, while HMF had the least effect (Table 4). It took *Z. mobilis* a longer period of time to complete active growth and reach the stationary phase, which was about 16, 19 or 21 h in the presence of HMF, furfural or vanillin, respectively, as compared to 11 h without any inhibitor present in the medium (FIGS. 5A-5D).

| TABLE 3 |

| Growth rate and final cell density of different *Z. mobilis* strains in the absence or presence of different sodium and acetate ions. |
|---------------------------|--------------------------------------------------|---------------------------|
|                   | ZM4                          | AcR                          | AcRIM0347 (p42-0347) | ZM4                          | AcR                          | AcRIM0347 (p42-0347) |
| Growth (hour⁻¹) | RM 0.42 ± 0.01               | 0.39 ± 0.01                  | 0.32 ± 0.003             | 0.33 ± 0.002              | 0.38 ± 0.003           |
|                   | RM (NaCl) 0.24 ± 0.008       | 0.29 ± 0.005                 | 0.21 ± 0.008             | 0.22 ± 0.009              | 0.25 ± 0.008           |
|                   | RM (NH₄Cl) 0.20 ± 0.008      | 0.19 ± 0.005                 | NA                       | 0.22 ± 0.002              | 0.19 ± 0.007           |
|                   | RM (KCl) 0.15 ± 0.004        | 0.12 ± 0.006                 | 0.09 ± 0.003             | 0.12 ± 0.006              | 0.12 ± 0.006           |
|                   | RM (NaAc) 0.29 ± 0.004       | 0.12 ± 0.004                 | 0.16 ± 0.002             | 0.27 ± 0.004              | 0.27 ± 0.004           |
**TABLE 3-continued**

Growth rate and final cell density of different *Z. mobilis* strains in the absence or presence of different sodium and acetate ions.

<table>
<thead>
<tr>
<th>ZM4</th>
<th>AcR</th>
<th>AcRIMO347 (pA2-0347)</th>
<th>ZM4 (pA2-0347)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final RM</td>
<td>0.95 ± 0.006</td>
<td>1.01 ± 0.006</td>
<td>0.94 ± 0.004</td>
</tr>
<tr>
<td>Cell (NaCl)</td>
<td>0.73 ± 0.01</td>
<td>0.96 ± 0.01</td>
<td>0.73 ± 0.03</td>
</tr>
<tr>
<td>Density (NH4OAc)</td>
<td>0.43 ± 0.01</td>
<td>0.42 ± 0.006</td>
<td>NA</td>
</tr>
<tr>
<td>(OD600 nm)</td>
<td>0.42 ± 0.002</td>
<td>0.40 ± 0.000</td>
<td>NA</td>
</tr>
<tr>
<td>RM (NaAc)</td>
<td>NA</td>
<td>0.63 ± 0.02</td>
<td>0.25 ± 0.001</td>
</tr>
</tbody>
</table>

*NA* indicates that the data are not available due to the lack of growth in that condition. The concentration for all the chemicals (NaCl, NH4OAc, KAc, NaAc) supplemented into the RM is 195 mM; NaCl; sodium chloride; NH4OAc; ammonium acetate; KAc; potassium acetate; NaAc; sodium acetate. Strains included in this study are: ZM4: *Zymomonas mobilis* ZM4 wild-type; AcR: *Z. mobilis* acetate tolerant mutant; ZM4 (pA2-0347); ZM4 containing a gateway plasmid pA2-0347 to express *Z. mobilis* grown ZM00347; AcRIM0347 (pA2-0347); AcRIM0347 (pA2-0347). This experiment has been repeated at least three times with similar result. Duplicate biological replicates were used for each condition.

**TABLE 4**

Growth rate and final cell density of different *Z. mobilis* strains in the absence or presence of different pretreatment inhibitors.

<table>
<thead>
<tr>
<th>ZM4</th>
<th>AcR</th>
<th>AcRIMO347 (pA2-0347)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth rate</td>
<td>RM</td>
<td>0.48 ± 0.03</td>
</tr>
<tr>
<td>(hour^-1)</td>
<td>HMF</td>
<td>0.36 ± 0.02</td>
</tr>
<tr>
<td>Furfural</td>
<td>0.31 ± 0.01</td>
<td>0.30 ± 0.005</td>
</tr>
<tr>
<td>Vanillin</td>
<td>0.26 ± 0.001</td>
<td>0.26 ± 0.01</td>
</tr>
<tr>
<td>Final RM</td>
<td>0.91 ± 0.01</td>
<td>0.98 ± 0.006</td>
</tr>
<tr>
<td>Cell</td>
<td>HMF</td>
<td>0.93 ± 0.003</td>
</tr>
<tr>
<td>Density</td>
<td>Furfural</td>
<td>0.88 ± 0.006</td>
</tr>
<tr>
<td>Vanillin</td>
<td>0.69 ± 0.006</td>
<td>0.71 ± 0.01</td>
</tr>
</tbody>
</table>

The concentration for the inhibitor supplemented into the RM is: HMF: 0.75 g/L, fufural, or vanillin: 1 g/L. Strains included in this study are: ZM4: *Zymomonas mobilis* ZM4 wild-type; AcR: *Z. mobilis* acetate tolerant mutant; ZM4 (pA2-0347); ZM4 containing a gateway plasmid pA2-0347 to express *Z. mobilis* grown ZM00347; AcRIM0347 (pA2-0347). This experiment has been repeated at least three times with similar result. Duplicate biological replicates were used for each condition.

**[0083]** Hfq Contributes to Sodium and Acetate Ion Tolerances

**[0084]** Although the final cell density of hfq mutant AcRIM0347 was similar to that of AcR parental strain (Table 3; FIG. 5A), the growth rate of AcRIM0347 was reduced by about one-fifth even without any inhibitor in the RM, indicating that hfq plays a central role in normal *Z. mobilis* physiology. Wild-type ZM4 that contained pA2-0347 was able to grow in the presence of 195 mM sodium acetate and had a similar growth rate and final cell density to that of acetate tolerant strain AcR (Table 3; FIG. 4C). The wild-type ZM4 was unable to grow under this condition.

**[0085]** The inactivation of the hfq gene in AcR decreased this acetate tolerant strain’s resistance to both sodium ion (sodium chloride) and acetate ion (ammonium acetate or potassium acetate) (Table 3; FIGS. 4A-4E). hfq mutant AcRIM0347 was unable to grow in the presence of 195 mM ammonium acetate or potassium acetate (Table 3; FIGS. 4D-4E). Both the growth rate and final cell density of hfq mutant AcRIM0347 were reduced by at least a quarter in the presence of 195 mM sodium chloride, and about 60% in the presence of 195 mM sodium acetate compared to that of the parental strain AcR (Table 3; FIGS. 4B-4C). The AcRIM0347 hfq mutation was complemented by the introduction of an hfg-expressing plasmid (pA2-0347) into the strain. The complemented mutant strain recovered at least half of the parental strains growth rate and 70% of its final cell density in the presence of 195 mM acetate ion (whether as sodium, ammonium or potassium acetate) (Table 3; FIGS. 4A-4E).

**[0086]** Hfq Contributes to Vanillin, Furfural, HMF and H2O2 Tolerances

**[0087]** AcRIM0347 growth rates were lower than that of ZM4 and AcR under all conditions tests, except for growth in RM broth (Table 4; FIGS. 5A-5D). AcRIM0347 also achieved lower final cell densities compared to ZM4 and AcR (Table 4; FIGS. 5A-5D). When AcRIM0347 was provided functional Z. mobilis Hfq via pA2-0347, growth rates under all conditions were largely unchanged (Table 4). However, shorter lag phases were observed for AcRIM0347 (pA2-0347) grown with vanillin, fufural or HMF and increases in final cell densities were also observed under these conditions (Table 4; FIGS. 5A-5D). These data indicate that hfg is important for optimal *Z. mobilis* growth and its ability to resist furfural, HMF and vanillin toxicity.

**[0088]** Hfq also contributed to tolerance of other stress such as the reactive oxidative species generating hydrogen peroxide (H2O2). hfg mutant AcRIM0347 was sensitive to hydrogen peroxide H2O2 and no observable growth was detected in RM medium with 0.001% H2O2 (FIG. 5E). The wild-type strain ZM4 and acetate tolerant strain AcR grew well at this concentration. Complementation of the hfg mutant strain allowed strain AcRIM0347 (pA2-0347) to grow in RM medium with 0.001% H2O2.
Example 4

This Example describes experiments to show that Yeast Lsm proteins contribute to pretreatment inhibitor tolerance.

Lsm Protein and Yeast Tolerance to Sodium and Acetate Ions

*S. cerevisiae* Sm and Sm-like (Lsm) proteins are similar to *Z. mobilis* Hfq at the level of protein sequence. Growth of yeast Lsm deletion mutants and Lsm over-expressing strains in 305 mM ammonium acetate, potassium acetate, or sodium acetate was assessed to test whether *S. cerevisiae* Lsm proteins and ZM4 Hfq had functionally similar roles.

Deletion of seven Lsm genes affecting three Lsm heterohexameric ring components (Lsm1, Lsm6, Lsm7) and four other Lsm proteins containing an Sm domain (Lsm9, Lsm12, Lsm13, Lsm16), was shown to have negative effects on the growth of *S. cerevisiae* in the presence of sodium acetate 40 g/L (FIG. 6A). On the other hand, six Lsm protein over-expressing *S. cerevisiae* strains (Lsm1, Lsm6, Lsm9, Lsm12, Lsm13, Lsm16) displayed enhanced growth in the presence of sodium acetate 40 g/L (FIG. 6B).

Growth differences between the Lsm mutants and yeast wild-type BY4741 in the CM broth without the addition of acetate or with 305 mM NaCl were not observed (FIGS. 7A-7B, respectively). *S. cerevisiae* Lsm proteins involved in RNA processing ring complex formation (Lsm1, 6, 7), especially Lsm6, played a role in acetate tolerance (FIGS. 7C-7E, 7K-7M). Lsm protein deletion mutants Lsm1, 6, and 7 showed decreased acetate tolerance compared to the wild-type control strain, especially in early growth stages for acetate with sodium, ammonium and potassium counter-ions (FIGS. 7C-7E). The Lsm overexpression strains grew similarly to wild-type BY4741 without the addition of acetate or with 305 mM NaCl (FIGS. 7F, 7G), but each of the Lsm protein overexpression strains showed enhanced acetate tolerance compared to the wild-type strain with sodium, ammonium or potassium counter-ions (FIGS. 7K-7M).

Lsm Proteins and Yeast Tolerance to Vanillin, Furfural and HMF

The effect of Lsm proteins on *S. cerevisiae* tolerance to pretreatment inhibitors vanillin, furfural, and HMF was also investigated using the seven Lsm deletion mutants and six Lsm overexpression strains described above. Each yeast deletion mutant and each overexpression strain showed similar growth profiles compared to wild-type strain BY4741 in the absence of inhibitors (FIGS. 7A, 7I). Deletion mutants for Lsm1, 6 and 7 proteins were unable to grow or showed extended lag phases before recovery from the inhibitory effects of pretreatment inhibitors (FIGS. 7F-7H). Overexpression of Lsm proteins provided a slight growth advantage in the presence of 1.5 g/L HMF and furfural (FIGS. 7O-7P). However, a detrimental effect on growth was observed for overexpression strains when cultured in the presence of 0.75 g/L vanillin (FIG. 7N). The data indicated that Lsm proteins Lsm1, 6, and 7 especially Lsm6, which are the components of yeast RNA processing ring complex, play a role in tolerance to the model inhibitors used in this study.

REFERENCES

- Brown S D: Analysis of a 16 kb region of the *Mesorhizobium loti* R7A symbiosis island encoding bio and nad loci and a novel member of the lad family Dunedin: University of Otago, 2002.
- Joachimsthal E L, Rogers P L: Characterization of a high-productivity recombinant strain of *Zymomonas mobi-


SEQUENCE LISTING

<160> NUMBER OF SEQ ID NO(S): 56

<210> SEQ ID NO 1
<211> LENGTH: 486
<212> TYPE: DNA
<213> ORGANISM: Zymomonas mobilis

<400> SEQUENCE: 1

atggccgaaa aggtaacaa ttctcagat tttttcttta ataccttgcg caagaaggccg 60
acacgcgta cgatgttctt ggtaaaaggt ctcaattac aggcctgat cacctggatt 120
gacaatattt ctattctctg gcggagagat ggtcagctca acgctgctta taacacagct 180
atctcacc tttctcggcg gcacagctgc gcgaagaagcg caagtctcgtg 240
gctagaacgtacttcaggt gttctccttta ogcgattat gcagcagcaaa 300
gacagcgcgt cactgttttt gtaaaaggg gctagtctgc gaaagctaat tcgctccttc 360
gatttactt gacgctgtt gacgcgtgg cgctgtaatt taaacagtgg 420
gttctcaacag tgcactcctgta gagactcagta sgagcagcag 480
gatgga 486

<210> SEQ ID NO 2
<211> LENGTH: 161
<212> TYPE: PRT
<213> ORGANISM: Zymomonas mobilis

<400> SEQUENCE: 2

Met Ala Glu Lys Val Asn Asn Leu Glu Asp Phe Phe Leu Asn Thr Leu 1 5 10 15
Arg Lys Thr Arg Thr Pro Val Thr Met Phe Leu Val Lys Gly Val Lys 20 25 30
Leu Glu Gly Val Ile Thr Trp Phe Asp Asn Phe Ser Ile Leu Arg 35 40 45
Arg Asp Gly Gln Ser Gln Val Leu Tyr Lys His Ala Ile Ser Thr Ile 50 55 60
Ile Pro Ala His Pro Leu Glu Gln Glu Arg Glu Ser Arg Ser Leu Met 65 70 75 80
Ala Glu Arg Lys Ser Ser Leu Leu Glu Asp Val Phe Leu Ser Ala Ile 85 90 95
Met Gln Glu Gln Glu Pro Val Thr Met Phe Leu Ile Asn Gly Val Met 100 105 110
Leu Gln Gly Glu Ile Ala Ala Phe Leu Phe Cys Val Leu Leu Thr 115 120 125
Arg Asn Asp Ala Gln Val Leu Tyr Lys His Ala Val Ser Thr Val 130 135 140
Gln Pro Val Lys Ser Val Asp Leu Thr Met Thr Glu Arg Arg Asp Glu 145 150 155 160

<210> SEQ ID NO 3
<211> LENGTH: 309
<212> TYPE: DNA
<213> ORGANISM: Escherichia coli

Dec. 9, 2010
-continued

<400> SEQUENCE: 3
atggttaag gcgatctttt acaagatcgc ttctgaaacg cactgcgctcg ggaagtgtt
15
ccagttctct tttatgtggt gaaatgtaat aagctgcaag ggcaatcgaac gctgtttgat
30
cagtcgag tgcctgtgaa acaaacggct gccacagtg gttacacgaa ccgcgtttctc
45
acgtgtgccc ctgcctgccc gttttcctat ccacgtaaca acgcoctgtgg cgggtcaccgc
60
agtaacctac atcctgttag cagcgcgcag aatctttcctc gcaacacgga cagcgaacag
75
acggttat
90

<210> SEQ ID NO 4
<211> LENGTH: 102
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli
<400> SEQUENCE: 4
Met Ala Lys Gly Gln Ser Leu Gln Asp Pro Phe Leu Aen Ala Leu Arg
1 5 10 15
Arg Gly Arg Val Pro Val Ser Ile Tyr Leu Val Aen Gly Ile Lys Leu
20 25 30
Gln Gly Gln Ile Glu Ser Aep Pro Gln Phe Val Ile Leu Leu Gly Aen
35 40 45
Thr Val Ser Gln Met Val Tyr His Ala Ile Ser Thr Val Val Pro
50 55 60
Ser Arg Pro Val Ser His His Ser Aen Ala Gly Gly Gly Thr Ser
65 70 75 80
Ser Aen Tyr His His Gly Ser Aen Leu Thr Ser Ala Gln Gln
85 90 95 100
Aep Ser Glu Glu Thr Glu

<210> SEQ ID NO 5
<211> LENGTH: 249
<212> TYPE: DNA
<213> ORGANISM: Clostridium thermoalcaliphilum
<400> SEQUENCE: 5
gtgtgagca aatataataat taatttcacag gcagtttttt taacaccaggt aagaaagas
15 30
catatcgg ccactgttta tcttaccaac ggatcagag ttaagaagac gcggtaagga
45 60
ctttgcaatt ttacctgtgc gcccacagct tagaggaaggc acgaactgtat ttataaacat
75 90
gcattttcc caaagcccc catgaaatt gttagccctga ttttaacgga caataacaga
105 120
tcgaataaa
125 140

gtgtgagca aatataataat taatttcacag gcagtttttt taacaccaggt aagaaagas
15 30
catatcgg ccactgttta tcttaccaac ggatcagag ttaagaagac gcggtaagga
45 60
tttgcaatt ttacctgtgc gcccacagct tagaggaaggc acgaactgtat ttataaacat
75 90
gcattttcc caaagcccc catgaaatt gttagccctga ttttaacgga caataacaga
105 120
tcgaataaa
125 140

gtgtgagca aatataataat taatttcacag gcagtttttt taacaccaggt aagaaagas
15 30
catatcgg ccactgttta tcttaccaac ggatcagag ttaagaagac gcggtaagga
45 60
tttgcaatt ttacctgtgc gcccacagct tagaggaaggc acgaactgtat ttataaacat
75 90
gcattttcc caaagcccc catgaaatt gttagccctga ttttaacgga caataacaga
105 120
tcgaataaa
125 140

<210> SEQ ID NO 6
<211> LENGTH: 82
<212> TYPE: PRT
<213> ORGANISM: Clostridium thermoalcaliphilum
<400> SEQUENCE: 6
Met Val Ser Lys Asn Ile Aen Leu Gln Asp Val Phe Leu Aen Gln
1 5 10 15
Val Arg Lys Glu His Ile Pro Val Thr Val Tyr Leu Thr Aen Gly Phe
20 25 30 35
Gln Leu Lys Gly Thr Val Lys Gly Phe Aep Aen Phe Thr Val Val Leu

-continued

Asp Ser Gly Arg Glu Glu Leu Ile Tyr Lys His Ala Ile Ser Thr
50 55 60
Ile Ser Pro Met Lys Ile Val Ser Leu Ile Phe Asn Asp Asn Asn Arg
65 70 75 60
Ser Glu

<210> SEQ ID NO 7
<211> LENGTH: 279
<212> TYPE: DNA
<213> ORGANISM: Anaerocellum thermophilum

<400> SEQUENCE: 7
gtggcgaaag gaaagtttaaa cttgccagcag ttttttttaa atctagttaa aaaaaaasaa 60
gtgaatgctt caaatctttct gttcagcctt tttcgatta aagaggttat caagggtttt 120
gtcaacctt cacggtctgt agagagctgac atacacacgc agcaacactt ttcagcatc 180
gtccttttta caactctgac cctcaagccc aaataataaa tgggtcagct acaggaatat 240
cacaagctt cccaacactt aaataataaa caaggtttaa 279

<210> SEQ ID NO 8
<211> LENGTH: 92
<212> TYPE: PRT
<213> ORGANISM: Anaerocellum thermophilum

<400> SEQUENCE: 8
Met Ala Lys Gly Ser Leu Asn Leu Glu Asp Leu Phe Leu Asn Glu Leu
1 5 10 15
Arg Lys Glu Lys Val Asn Val Thr Ile Phe Leu Leu Ser Gly Phe Glu
20 25 30
Leu Lys Gly Val Ile Lys Gly Phe Asn Asn Thr Leu Ile Val Glu
35 40 45
Thr Asp Asn Asn Lys Glu Leu Ile Tyr Lys His Ala Ile Ser Ser
50 55 60
Ile Met Pro Ser Lys Pro Ile Asn Tyr Met Ala Glu Ala Glu Asn Asn
65 70 75 80
Gln Glu Ala Ser Glu Ser Asn Asn Asn Asn Glu Gly
85 90

<210> SEQ ID NO 9
<211> LENGTH: 300
<212> TYPE: DNA
<213> ORGANISM: Caldicellulosiruptor saccharolyticus

<400> SEQUENCE: 9
gtggcgaaga gaaatgctaa cttgtagcat ttttttaa aaccagcttcg aaaaaaaasaa 60
gtcaacgctt caaatctttct actgagcctg ttccaagattt aagaggttat caagggtttt 120
gtcaacctt cacggtctgt agagagctgac atacacacgc agcaacactt ttcagcatc 180
gcaaccttt caactctgac cctcaagccc aaataataaa tgggtcagct acaggaatat 240
cagaagctt cccaacactt aaataataaa caaggtttaa 300

<210> SEQ ID NO 10
<211> LENGTH: 59
<212> TYPE: PRT
ORGANISM: Caldicellulosiruptor saccharolyticus

Val Ala Lys Gly Asn Leu Asn Gln Asp Leu Phe Leu Asn Gln Leu
1  5  10  15

Arg Lys Glu Lys Val Asn Val Thr Ile Phe Leu Leu Ser Gly Phe Gln
20 25 30

Leu Lys Val Ile Lys Gly Phe Asn Asn Phe Thr Leu Val Val Glu
35 40 45

Thr Glu Asn Asn Lys Gln Gln Leu Ile Tyr Lys His Ala Ile Ser Ser
50 55 60

Ile Leu Pro Ser Lys Pro Ile Asn Tyr Met Ala Gln Val Gln Asn Ser
65 70 75 80

Gln Val Gln Asn Thr Ala Ser Gln Gln Ser Asn Asn Gln Gln Glu
85 90 95

Glu Ser Lys

ORGANISM: Thermoanaerobacter sp.

atggcagagg caaaaagac gattaattta caggtatct tttaaatca aagtgaagaa 60
gagcatgttc atcataatc atcaggttc aatnaggg tttaggaas 120
ggattgata atcttacagt ggtgtgag tggagaacg acagcaact tcttatctac 180
aaacatcga tttgtacaat tcatcctca aagcctgtga tttctcttc ttctgataaa 240
gatgcagaa gagaagagt g 261

ORGANISM: Thermoanaerobacter sp.

Met Ala Ser Ser Lys Ala Ala Ala Asn Leu Glu Asp Ile Phe Leu Asn
1  5  10  15

Gln Val Arg Lys Glu His Val Pro Val Thr Val Tyr Leu Ile Asn Gly
20 25 30

Phe Gln Leu Lys Gly Leu Val Lys Gly Phe Asp Asn Phe Thr Val Val
35 40 45

Leu Glu Ser Glu Asn Lys Gln Gln Leu Ile Tyr Lys His Ala Ile
50 55 60

Ser Thr Ile Thr Pro Gln Lys Pro Val Ile Phe Ser Ala Ser Asp Lys
65 70 75 80

Asp Glu Lys Arg Glu Glu
85

ORGANISM: Saccharomyces cerevisiae

SEQ ID NO 13

LENGTH: 591

TYPE: DNA
-continued

```
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
</table>
| atgagc ... | /
| aggtt ... | /
| atgaat ... | /
| ttggat ... | /
| ttggac ... | /
| cgacta ... | /
| cgctact ... | /
| aacact ... | /
| aacgac ... | /
| agacgc ... | /
| cgaagc ... | /
| cgcccc ... | /

SEQ ID NO 14
LENGTH: 196
OREGANISM: Saccharomyces cerevisiae
```

```
| Met Ser Lys Ile Gln Val Ala His Ser Ser Arg Leu Ala Asn Leu Ile | 1-15 |
| Val Leu Arg Val Leu Thr Gln Asp Gly Arg Val Tyr Ile Gly | 20-30 |
| Gln Leu Met Ala Phe Asp Lys His Met Asn Val Leu Asn Glu Cys | 35-45 |
| Ile Glu Glu Arg Val Pro Lys Thr Gln Leu Asp Lys Leu Arg Pro Arg | 50-60 |
| Lys Asp Ser Lys Asp Gly Thr Leu Asn Ile Lys Val Glu Lys Arg | 65-80 |
| Val Leu Gly Leu Thr Ile Leu Arg Gly Glu Gln Ile Leu Ser Thr Val | 85-95 |
| Val Glu Asp Lys Pro Leu Leu Ser Lys Gly Arg Leu Val Arg Asp | 100-110 |
| Lys Lys Gly Lys Gln Ala Gln Gly Gln Thr Lys Leu Arg Lys Glu | 115-125 |
| Lys Gly Lys Lys Pro Gly Lys Ile Ala Lys Pro Asn Thr Ala Asn Ala | 130-140 |
| Lys His Thr Ser Ser Asn Ser Arg Glu Ile Ala Gln Pro Ser Ser Ser | 145-160 |
| Arg Tyr Asp Gly Asn Asp Ile Gly Ala Asn Arg Ser Arg Phe | 165-175 |
| Asn Asn Glu Ala Pro Pro Gln Thr Lys Phe Gly Pro Pro Pro Gly | 180-190 |
| Phe Lys Arg Lys | 195 |
```

```
<210> SEQ ID NO 15
<211> LENGTH: 441
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae
```
-continued

```
aacggtacca ccttgggg tacaactgcc tggcttacac ccaaaatgaa tgcctatcctta 120
agtacgtgta agttgacocct accacaacc ccgactaaata aattgacacg taatgtatt 180
ggcatggtct gctgtaatt gactgagaga caggcaacctta atgcaattga caacatagca 240
agttgcaac atctaaacat taggggcaat accataagac agataatactt acctgtccct 300
ttgaacgctg attacaacctt ggtgaccaaa aagcaactta atcccctaag aagatcgggt 360
caaaagccaa atgcacccca ccaaagagaga aggccgcaatt taggtgcacc agccaaatat 420
agagacgaag gaggctcatg a 441
```

```
<210> SEQ ID NO 16
<211> LENGTH: 146
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae
<400> SEQUENCE: 16

Met Lys Leu Val Asn Phe Leu Lys Leu Arg Asn Glu Gln Val Thr
1 5 10 15
Ile Glu Leu lys Asn Gly Thr Thr Val Trp Gly Thr Leu Gln Ser Val
20 25 30
Ser Pro Gln Met Asn Ala Ile Leu Thr Asp Val Leu Thr Leu Pro
35 40 45
Gln Pro Arg Leu Asn Leu Asn Ser Asn Gly Ile Ala Met Ala Ser
50 55 60
Leu Tyr Leu Thr Gly Gln Gln Pro Thr Ala Ser Asp Asn Ile Ala
65 70 75 80
Ser Leu Gln Tyr Ile Asn Ile Arg Gly Asn Thr Ile Arg Gln Ile
85 90 95
Leu Pro Asp Ser Leu Asn Leu Asp Ser Leu Val Asp Gln Lys Gln
100 105 110
Leu Asn Ser Leu Arg Arg Ser Gly Ile Ala Asn Asp Pro Ser Lys
115 120 125
Lys Arg Arg Arg Asp Phe Gly Ala Pro Ala Asn Lys Arg Pro Arg Arg
130 135 140
Gly Leu
145
```

```
<210> SEQ ID NO 17
<211> LENGTH: 333
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae
<400> SEQUENCE: 17

atgtttcccc aaatatctga tgggttcataa atgatctctc atagagcagc attaggga 60
catgagaagt ttagttccaa acatgtctcga atgtgctgtgta taatgtatgc ttaggtgaca 120
agaaacocct tgatatctc attaagaac acatatcaaa taatagcagc aatggaagct 180
ttcgacgggc atgttagatt ggttttagaa aatgtgaagag gctttggac agagaagag 240
ggcactaatg taatatactg gaagatcctt ataagtaac atatttaaag aagttgattc 300
gttatggtt gttaaaaccc ctggtgtgag ta 333

<210> SEQ ID NO 18
<211> LENGTH: 110
Continued

```plaintext
<210> SEQ ID NO 19
<211> LENGTH: 306
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 19

atgactatga atggaattc agtgaatcc ttaatgagg cacagggca cacagcgttct  60
cagctatcct cagctatcct cagctatcct cagctatcct cagctatcct cagctatcct  120
apgtacagc aacagggac aacagggac aacagggac aacagggac aacagggac  180
cattagagt tcattaagagt tcattaagagt tcattaagagt tcattaagagt tcattaagagt  240
gccacattat tcacacacac acacacacac acacacacac acacacacac acacacacac  300
agctgta  306

<210> SEQ ID NO 20
<211> LENGTH: 101
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 20

Met Thr Met Asn Gly Ile Pro Val Lys Leu Leu Leu Asn Glu Ala Glu Gly  1  5 10 15
His Ile Val Ser Leu Glu Thr Thr Gly Ala Thr Tyr Arg Gly Lys  25 30
Leu Val Glu Ser Glu Asp Ser Met Asn Val Glu Arg Asp Val Ile  40 45
Ala Thr Glu Pro Gln Gly Ala Val Thr His Met Asp Gln Ile Phe Val  50 55 60
Arg Gly Ser Glu Ile Lys Phe Ile Val Val Pro Asp Leu Leu Lys Asn  70 75 80
Ala Pro Leu Phe Lys Asn Ser Ser Arg Pro Met Pro Pro Ile Arg  85 90 95
Gly Pro Lys Arg Arg  100
```
<210> SEQ ID NO: 21
<211> LENGTH: 285
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 21
atgctgaca acagtgaaac cagcgccatg gtcgacacaa taaaattgcat atttaacctc 60
ttcaacacg aacaccaagt aacgatatgg ttattgagc aatcggcat aagaatcaag 120
ggttaaatag ttgagttgga taagttgtca tggatgaagc cyttgaaatt 180
cctgtaata tgcagcatgg taaagaagat gttgagaaag ggacgctcct gggaagatct 240
cctgtaagc ggtgataat cacattgata acatcagcgg acctg 285

<210> SEQ ID NO: 22
<211> LENGTH: 94
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 22
Met Ser Asn Lys Val Lys Thr Lys Ala Met Val Pro Ile Asn Cys
1 5 10 15
1 Leu Phe Asn Phe Leu Gln Gln Gln Thr Pro Val Thr Ile Trp Leu Phe
20 25 30
Glu Glu Ile Gly Ile Arg Ile Lys Gly Lys Ile Val Gly Phe Asp Glu
35 40 45
Phe Met Asn Val Val Ile Asp Glu Ala Val Glu Ile Pro Val Asn Ser
50 55 60
Ala Asp Gly Lys Glu Asp Val Glu Lys Gly Thr Pro Leu Gly Lys Ile
65 70 75 80
Leu Leu Lys Gly Asp Asn Ile Thr Leu Ile Thr Ser Ala Asp
85 90

<210> SEQ ID NO: 23
<211> LENGTH: 261
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 23
atgacgacg gacatgtatat cattgcogatg cagccogatg acgccagcc gtctctcaaa 60
ggcctgtgca accatctgtt aagctgcaag cttaagtca acagcaacgt atataggt 120
cagtcggtgt ccaagcagaa ctagtttaac ctagcgtgta acagaagcaga agatgttggt 180
gcgggtgtct cgccagcaac cttggcgag atattctcc gttgaatata cgttctgtac 240
atcagggag tcgcgaacta a 261

<210> SEQ ID NO: 24
<211> LENGTH: 86
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 24
Met Ser Glu Ser Ser Asp Ile Ser Ala Met Gln Pro Val Asn Pro Lys
1 5 10 15
1 Pro Phe Leu Lys Gly Leu Val Asn His Arg Val Gly Val Lys Leu Lys
20 25 30
Phe Asn Ser Thr Glu Tyr Arg Gly Thr Leu Val Ser Thr Asp Asn Tyr
<table>
<thead>
<tr>
<th>Seq ID No</th>
<th>Length</th>
<th>Type</th>
<th>Organism</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>234</td>
<td>DNA</td>
<td>Saccharomyces</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>519</td>
<td>DNA</td>
<td>cerevisiae</td>
<td></td>
</tr>
</tbody>
</table>

**Continued**

35

<table>
<thead>
<tr>
<th>Residue</th>
<th>35</th>
<th>40</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>50</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>Gln</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>55</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>Ala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>65</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

His 65

<table>
<thead>
<tr>
<th>Residue</th>
<th>65</th>
<th>70</th>
<th>75</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>65</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>Gly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>70</td>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td>Val</td>
<td>75</td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>Ala</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ile 95

**<210> SEQ ID NO 25**

**<211> LENGTH: 234**

**<212> TYPE: DNA**

**<213> ORGANISM: Saccharomyces cerevisiae**

**<400> SEQUENCE:**

```
atggttctta cccctgaact gagaatat atggacaaga agatattgc gagataaat 60
ggatcaga aatggtcagag aatttggcga ggtacaagata ttttctaaaa cgtggtttct 120
gattgtcag taagagataa tgggagac ccgccgga ctaccacgct aaggttgcag 180
acctggccata ggcggcgct ctagaggtc tagagcctatat 234
```

**<210> SEQ ID NO 26**

**<211> LENGTH: 77**

**<212> TYPE: PRT**

**<213> ORGANISM: Saccharomyces cerevisiae**

**<400> SEQUENCE:**

```
Met Val Ser Thr Pro Glu Leu Lys Tyr Met Asp Lys Lys Ile Leu 1 5 10 15
Leu Asn Ile Asn Gly Ser Arg Lys Val Ala Gly Ile Leu Arg Gly Tyr 20 25 30
Asp Ile Phe Leu Asn Val Leu Asp Asp Ala Met Gly Ile Asn Gly 35 40 45
Glu Asp Pro Ala Asn His Gly Leu Gly Leu Gly Thr Val Ile Arg 50 55 60
Gly Asn Ser Ile Ile Ser Leu Gln Leu Asp Ala Ile 65 70 75
```

**<210> SEQ ID NO 27**

**<211> LENGTH: 519**

**<212> TYPE: DNA**

**<213> ORGANISM: Saccharomyces cerevisiae**

**<400> SEQUENCE:**

```
agtctgcaa atgctcaacg acaacagct aatggtcagag aatttggcga ggtacaagata ttttctaaaa cgtggtttct 60
aattttccaa aagagatttc aagaaggtgag ggcgattttat atctgacca gttacttcttct 120
actaaccacgg ctgtgttttt gactcagta gaccgtaaaa ttttctgggtttggttgat 180
gggagagtc tctctggtt gcaattgtt actacaacct tttgctatt ttttgtgtcttttg 240
gcattctgttggctta aacttggtctg gatatttttt ctgttct gtaatcttccagcc 300
atagaatgtaaacttggtta gttcagtggctatgcttct gttcagtggctatgcttct gtaatcttccagcc 360
gatcagcccc ttgaggctcgagagcata ccatttasaggg aaggttgc caccagcaga 420
agagaatgtaaaagagactttc taaaggggaa acacccaaag gccagatt cgtataaat gcccgccgca 480
gggtactttt actatctgac atgtactaat 519
```
<210> SEQ ID NO 28
<211> LENGTH: 172
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 28

Met Ser Ala Asn Ser Lys Asp Arg Asn Gin Ser Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin In...
**Sequence 31:**

<table>
<thead>
<tr>
<th>Lys</th>
<th>Lys</th>
<th>Tyr</th>
<th>Pro</th>
<th>His</th>
<th>Leu</th>
<th>Gly</th>
<th>Ser</th>
<th>Val</th>
<th>Arg</th>
<th>Asn</th>
<th>Ile</th>
<th>Phe</th>
<th>Ile</th>
<th>Arg</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Sequence 32:**

<table>
<thead>
<tr>
<th>Met</th>
<th>Glu</th>
<th>Thr</th>
<th>Pro</th>
<th>Leu</th>
<th>Asp</th>
<th>Leu</th>
<th>Leu</th>
<th>Lys</th>
<th>Leu</th>
<th>Asn</th>
<th>Leu</th>
<th>Asp</th>
<th>Glu</th>
<th>Arg</th>
<th>Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Sequence 33:**

<table>
<thead>
<tr>
<th>Met</th>
<th>Val</th>
<th>Phe</th>
<th>Ile</th>
<th>Arg</th>
<th>Gly</th>
<th>Ala</th>
<th>Val</th>
<th>Glu</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Organism:** Saccharomyces cerevisiae

**Type:** DNA

**Length:**
- Sequence 31: 270
- Sequence 32: 89
- Sequence 33: 564
<210> SEQ ID NO 34
<211> LENGTH: 187
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 34

Met Leu Pro Leu Tyr Leu Leu Thr Asn Ala Lys Gly Gln Gln Met Gln
1      5     10  15

Ile Glu Leu Lys Ann Gly Glu Ile Ile Glu Gly Ile Leu Thr Asn Val
20     25    30

Asp Asn Trp Met Asn Leu Thr Leu Ser Asn Val Thr Glu Tyr Ser Glu
35     40    45

Glu Ser Ala Ile Asn Ser Glu Asp Ala Glu Ser Ser Lys Ala Val
50     55    60

Lys Leu Ann Glu Ile Tyr Ile Arg Gly Thr Phe Ile Lys Phe Ile Lys
65     70    75

Leu Gln Asp Ann Ile Ile Asp Val Lys Gln Gln Ile Asn Ser Asn
85     90    95

Ann Asn Ser Ann Ser Ann Gly Pro Gly His Lys Arg Tyr Tyr Ann Ann
100   105   110

Arg Asp Ser Ann Ann Arg Gly Ann Tyr Ann Arg Arg Ann Asn Ann
115   120   125

Ann Gly Ann Ser Ann Arg Pro Tyr Ser Glu Ann Arg Gln Tyr Ann
130   135   140

Ann Ser Ann Ser Ann Ile Ann Ann Ile Ann Ser Ile Ann Ser
145   150   155

Ann Ann Gln Ann Met Ann Ann Gly Leu Gly Gly Ser Val Gln His His
165   170   175

Phe Asn Ser Ser Ser Ser Pro Gln Lys Val Glu Phe
180   185

<210> SEQ ID NO 35
<211> LENGTH: 262
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 35

atgatctac cgggagtttt ggctttggsa gtcata gata aacaattaa cacaagaagtg 60
ttgattgtgc tcagtc gaag cgggacgt tagttgtgtt cgacagactc 120
gtcaagttta tacggaaga cgtgctcag tgggtttgat atcg gga cagagcaga 180
aatgagaaag ttacgaca ccaatgagca atgctttaa cggcga cacaatttt 240
tgtgctcag gcggcagaaa gcac ccactgcag ggctttgat aa 282

<210> SEQ ID NO 36
<211> LENGTH: 93
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 36

Met Ser Leu Pro Glu Ile Leu Pro Leu Glu Val Ile Asp Lys Thr Ile
-continued

```
1  5  10  15
Asn Glu Lys Val Leu Ile Val Leu Glu Ser Asn Arg Glu Phe Glu Gly
Thr Leu Val Gly Phe Asp Phe Val Asn Val Ile Leu Glu Asp Ala
Val Glu Trp Leu Ile Asp Pro Glu Glu Ser Arg Asn Glu Lys Val
Met Glu His His Gly Arg Met Leu Leu Ser Gly Asn Asn Ile Ala Ile
Leu Val Pro Gly Gly Lys Thr Pro Thr Glu Ala Leu

<210> SEQ ID NO: 37
<211> LENGTH: 261
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 37
atgctggaaa aagcttttat acagggtagc gttacttccg agtttcttc tggatatcatt
  60
ggtgaacag tgaacgtcaaa acttgcttgc ggttacctct acacagggaa attggaatcc
  120
atgtagtt gataaggt tggctacttgc agtgcacctg aacacttacg gaagtaaatc
  180
aataaggttc ttaaatagtt caataaggt gttcttttta ggggcacggca ggtcatgtat
  240
atcagtgac aaaaaatata g
  261

<210> SEQ ID NO: 38
<211> LENGTH: 86
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 38
Met Ser Gly Lys Ala Ser Thr Glu Gly Ser Val Thr Glu Phe Leu
  1  5  10  15
Ser Asp Ile Ile Gly Lys Thr Val Asn Val Lys Leu Ala Ser Gly Leu
  20  25  30
Leu Tyr Ser Gly Arg Leu Glu Ser Ile Asp Gly Phe Met Asn Val Ala
  35  40  45
Leu Ser Ser Ala Thr Glu His Tyr Glu Ser Asn Asn Asn Lys Leu Leu
  50  55  60
Ann Lys Phe Ann Ser Asp Val Phe Leu Arg Gly Thr Glu Val Met Tyr
  65  70  75  80
Ile Ser Glu Gln Lys Ile
  85

<210> SEQ ID NO: 39
<211> LENGTH: 348
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 39
atgctggaaa aacactcga atcagagataa aacactcga aacactcga aacactcga
  60
ggcccctaaa gagaagcttat ctctgtatat aggcttataaa aatccgcgctct
  120
aataaactg gttgtaaatt agttataatt gttcctaaaag gttagtagctc atgtagac
  180
ttggtactttg atgtagatag aataaatatg tttatatctg atgtagaaaa caacactcga
  240
```
ctgatttctta aaaaagcaag aagctaggt tggacgctca taagaggtac tattttggtc
ctctttgaagacctagagctgctgta ctatatagc aaaaatag

<210> SEQ ID NO 40
<211> LENGTH: 115
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 40
Met His Gln Gln His Ser Lys Ser Glu Asn Lys Pro Gln Gln Gln Arg
Lys Lys Phe Glu Gly Pro Lys Arg Glu Ala Ile Leu Asp Leu Ala Lys
Tyr Lys Asp Ser Lys Ile Arg Val Lys Leu Met Gly Gly Lys Leu Val
Ile Gly Val Leu Lys Gly Tyr Asp Gln Leu Met Asn Leu Val Leu Asp
Asp Thr Val Glu Tyr Met Ser Asn Pro Asp Glu Asn Asn Thr Glu
Leu Ile Ser Lys Asn Ala Arg Lys Leu Gly Leu Thr Val Ile Arg Gly
Thr Ile Leu Val Ser Leu Ser Ser Ala Glu Gly Ser Asp Val Leu Tyr
Met Gln Lys

<210> SEQ ID NO 41
<211> LENGTH: 330
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 41
atgtagcaaccttgaagactcatattaataaaaagagttg ttataatacaagttgaagcgc
agagtgcctaacagcccttgaatctgcttctaaaaagacc catctttcat ctacactatgtt
ctactataacaagaatctatcttatctattgca ctacactaat ttatcatat

<210> SEQ ID NO 42
<211> LENGTH: 109
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 42
Met Ser Ala Thr Leu Lys Asp Tyr Leu Asn Lys Arg Val Val Ile Ile
Lys Val Asp Gly Gnu Cys Leu Ile Ala Ser Leu Asn Gly Phe Asp Lys
Asn Thr Asn Leu Phe Ile Thr Asn Val Phe Asn Arg Ile Ser Lys Glu
Phe Ile Cys Lys Ala Gln Leu Leu Arg Gly Ser Glu Ile Ala Leu Val
-continued

Gly Leu Ile Amp Ala Glu Asn Asp Ser Leu Ala Pro Ile Asp Glu
65 70 75 80
Lys Lys Val Pro Met Leu Lys Asp Thr Lys Asn Ile Glu Asn Glu
85 90 95
His Val Ile Trp Glu Lys Val Tyr Glu Ser Lys Thr Lys
100 105

<210> SEQ ID NO 43
<211> LENGTH: 267
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 43
atggaactct tgaacctgct agatatttat ggaaactttt taatagtttc cttacagaa
60

gatgatattt tagttgggaag cttggtgtgc gttagatgcac aaaggaatct gtattgatag
120
cagcgttgg gccttgatgc ctccagatag aagtaggtgg gcctgtgc agcttctgag
180
cggctcgta agaccaaat aattgataag cttggtcttg acggagttac tggcattaa
240
gtgaatgtag tggcataaat ttgtagg
267

<210> SEQ ID NO 44
<211> LENGTH: 86
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 44
Met Asp Ile Leu Lys Leu Ser Asp Phe Ile Gly Asn Thr Leu Ile Val
1 5 10 15
Ser Leu Thr Glu Asp Arg Ile Leu Val Gly Ser Leu Val Ala Val Asp
20 25 30
Ala Gln Met Asn Leu Leu Leu Asp His Val Glu Glu Arg Met Gly Ser
35 40 45
Ser Ser Arg Met Met Gly Leu Val Ser Val Pro Arg Arg Ser Val Lys
50 55 60
Thr Ile Met Ile Asp Lys Pro Val Leu Glu Leu Thr Ala Asn Lys
65 70 75 80
Val Glu Leu Met Ala Asn Ile Val
85

<210> SEQ ID NO 45
<211> LENGTH: 564
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 45
atgaggtgca gctttgagca aacgctggga ttcagataaa aagttacgaac cgtgtggatt
60
gtgattctg aggagagggt ctattggttc acattccacacaacattctagctatccca
120
acaacaagaga gatcatcaac tccacacaaat tccagggtga taaatgtgac attctacag
180
cattttggg tcattttgca taaacccctg tttaaatcctc taaaaagac caaattccca
240
ccccctcatt tccacagggac aagagggttg aagctttttag aagagaggtg atacagatctt
300
aaaaagcgaa aaccttaaag ggccaggggt gtgaagtgca agggtcgatc cattttcttt
gatttctca agaccaatat tttgcaggct ctggggctta aagacatcatattcttgat
360
420
<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-10</td>
<td>gacgtaaagg tgcaacctcc atacaaggtc gaagatatca aagtgctaca tgagggagt</td>
<td>-continued</td>
</tr>
<tr>
<td>11-20</td>
<td>aaccaatcct tataaataa taagaagaaa gcttgagagca gctgagaacaa</td>
<td></td>
</tr>
<tr>
<td>21-30</td>
<td>gacgatggtg ggaaggtgg atag</td>
<td></td>
</tr>
</tbody>
</table>

**<210> SEQ ID NO 46**

**<211> LENGTH: 187**

**<212> TYPE: PRT**

**<213> ORGANISM: Saccharomyces cerevisiae**

**<400> SEQUENCE: 46**

Met Ser Val Ser Leu Glu Gln Thr Leu Gly Phe Arg Ile Lys Val Thr
1  5  10  15

Asn Val Leu Asp Val Val Thr Glu Arg Leu Tyr Ser Phe Asn Ser
20 25 30

Ser Asn Asn Thr Leu Thr Ile Gln Thr Thr Lys Asn Gln Ser Pro
35 40 45

Gln Asn Phe Lys Val Ile Lys Cys Thr Phe Ile Lys His Leu Glu Val
50 55 60

Ile Gly Asp Lys Pro Ser Phe Asn Ser Phe Lys Gly Gln Gln Ile Lys
65 70 75 80

Pro Ser Tyr Val Asn Val Glu Arg Val Glu Leu Lys Glu Ser
85 90 95

Val Ile Ala Ser Lys Lys Gly Leu Leu Arg Gly Lys Gly Val Ser
100 105 110

Ala Glu Gly Gln Phe Ile Phe Asp Gln Ile Phe Lys Thr Ile Gly Asp
115 120 125

Thr Lys Trp Val Ala Lys Asp Ile Ile Leu Asp Val Lys Val
130 135 140

Gln Pro Pro Tyr Lys Val Glu Asp Ile Lys Val Leu His Glu Gly Ser
145 150 155 160

Asn Gln Ser Ile Thr Leu Ile Gln Arg Ile Val Glu Arg Ser Trp Glu
165 170 175

Gln Leu Glu Gln Asp Asp Gly Arg Lys Gly Gly
180 185

**<210> SEQ ID NO 47**

**<211> LENGTH: 1050**

**<212> TYPE: DNA**

**<213> ORGANISM: Saccharomyces cerevisiae**

**<400> SEQUENCE: 47**

atgctgcagt acattggtaa aactattttc ttaatcttg tsgcttgaca cagataatgtg
1-60

gggctgttag aagatatgta cttgaaaag ggtaccgtga cttgaaaag agttgctgtg
61-120

tttttgtcag aagttgagaa gaagttggtg ctggaagaaa tttatcggaa toctcaagta
121-180
tacaaatttt taatgctaca cggctagttg gtcagattt taaccatttt agtgctgtac
181-240

tactagcaca taacggtcct tgtccttaca atagtgacac ccgctttcaca atccctctct
241-300
taccagcata cttcctccac ccaacgctca cgtcagcagc acgctgaacaa aaccccacca
301-360
gttcacaacc ccgaatcaca tgtgcaagca gctggtgttt tgtaccacca
361-420

acagcagac aacccgtcag tgcctgttag aagataaaga gcaactctcag aagaaaca
421-480
gtaaactgca aagagttgaa aaggtggtaa atagctggca cattgccccatatgaga
481-540
<table>
<thead>
<tr>
<th>Val</th>
<th>Asp</th>
<th>Thr</th>
<th>Phe</th>
<th>Gly</th>
<th>Gln</th>
<th>Ala</th>
<th>Ser</th>
<th>Ala</th>
<th>Ser</th>
<th>Ala</th>
<th>Arg</th>
<th>Pro</th>
<th>Arg</th>
<th>Phe</th>
<th>His</th>
<th>Ser</th>
<th>Arg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>275</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gly</th>
<th>Leu</th>
<th>Gly</th>
<th>Arg</th>
<th>Gly</th>
<th>Arg</th>
<th>Asn</th>
<th>Tyr</th>
<th>Arg</th>
<th>Gly</th>
<th>Arg</th>
<th>Arg</th>
<th>Gly</th>
<th>Asn</th>
<th>Arg</th>
<th>Gly</th>
<th>Asn</th>
<th>Arg</th>
<th>Arg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>295</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gly</th>
<th>Arg</th>
<th>Gly</th>
<th>Gln</th>
<th>Gly</th>
<th>Arg</th>
<th>Gly</th>
<th>Asn</th>
<th>Tyr</th>
<th>Gln</th>
<th>Arg</th>
<th>Gly</th>
<th>Arg</th>
<th>Gln</th>
<th>Arg</th>
<th>Gly</th>
<th>Asn</th>
<th>Tyr</th>
<th>Gln</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Arg | Asp | Ser | Gly | Ala | Tyr | Gln | Asn | Arg | Ala | Pro | Ser | Arg | Tyr | Ser | Arg | Pro |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

| Ala | Asn | Gln | Phe | Ser | Gln | Pro | Pro | Ser | Arg | Asn | Val | Glu | Phe |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     |     |     |     |     |     |     |     |     |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |

| <210> SEQ ID NO 49 |
| <211> LENGTH: 1656 |
| <212> TYPE: DNA |
| <213> ORGANISM: Saccharomyces cerevisiae |
| <400> SEQUENCE: 49 |

| atgctcaaat gtagtggttt cggagtacaa gtagggctaa aagatgggaa gtctatcag |
| gggaaaaatt ccagaaacac tccaaaaagga ggctactttta atgacgtctca tttgctgctat |
| gtagtgtaaat cttaaggtt caaagtgac ggtctaaagc ttaagctcct aaagttgtta |
| actctgtgtt ttaacctgg gaaaggaag cgaaaaagcg cagcagcgac aaacaagcat |
| tataatoaa atggcggctga acaatattgt tggcgaagat atgagtttag taagatataaa |
| cagcagagag atgactggat ccaagaaat tgggtcgttg ttaacaaaaa agacgtcctcc |
| gcacacatgg atgcaacttta ctgagagttt gagctctggg acataactcca |
| aagctcctcc aatgcagctaa aataaattat caaatgctag atgattttgt ttcagatgca |
| aagcagagat catcgccacca gattctcata cgaagtcgac cagcagcagc ccaatcctcag |
| cggcagcagc atgctctaa gtagctgttt ctggagatgct atgacatgctg |
| gtagctactg cagcagacta cttaactctg cttgagattt ttcacattta |
| attctgactGA aactcactG aatattcctg cttgagattt ttcacattta |
| gtagctacacG gtacatctag aacatggcact tgtttgctggt ttcgatgctg |
| aacctttctG cgggtggcctgc aatattcctg cttgagattt ttcacattta |
| cctctggtta gaaaggggag cagcagacca aatattcctg cttgagattt ttcacattta |
| catccttggc aagcagagaa ctcgctgttc atacattctt ttcacattt ctaaaattgga |
| ttcctggtta ctcgctgttc atacattctt ttcacattt ctaaaattgga |
| catccttggc aagcagagaa ctcgctgttc atacattctt ttcacattt ctaaaattgga |
| ttcctggtta ctcgctgttc atacattctt ttcacattt ctaaaattgga |
| ttcctggtta ctcgctgttc atacattctt ttcacattt ctaaaattgga |

| 60  |
| 120 |
| 180 |
| 240 |
| 300 |
| 360 |
| 420 |
| 480 |
| 540 |
| 600 |
| 660 |
| 720 |
| 780 |
| 840 |
| 900 |
| 960 |
| 1020 |
| 1080 |
| 1140 |
| 1200 |
| 1260 |
| 1320 |
| 1380 |
| 1440 |
| 1500 |
| 1560 |
- continued

staccacagg gcgcctattt acagagaact tctttgctga agttccaaas ctgctgtcct

<210> SEQ ID NO 50
<211> LENGTH: 551
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 50

Metc Ser Gln Phe Val Gly Phe Gly Val Gln Val Glu Leu Lys Asp Gly
1   5   10   15
Lys Leu Ile Gln Gly Lys Ile Ala Lys Ala Thr Ser Lys Gly Leu Thr
20  25  30
Leu Asn Asp Val Gln Phe Gly Asp Gly Gly Lys Ser Glu Ala Phe Lys
35  40  45
Val Arg Ala Ser Arg Leu Lys Asp Leu Lys Val Leu Thr Val Ala Ser
50  55  60
Gln Ser Gly Lys Arg Lys Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Asp
65  70  75  80
Tyr Asn Gln Asn Arg Gly Glu His Ile Asp Trp Glu Asp Asp Asp Val
85  90  95
Ser Lys Ile Lys Gln Gln Glu Asp Phe Asp Phe Glu Arg Asn Leu Gly
100 105 110
Met Phe Asn Lys Asp Val Phe Ala Asn Leu Lys Gln Gln Gln Gln Gln Gln Asp
115 120 125
Ile Leu Pro Glu Arg Leu Gln Gly His Asn Arg Lys Gln Thr Gln
130 135 140
Leu Gln Gln Asn Thr Tyr Glu Arg Asp Glu Leu Val Ile Pro Asp Ala
145 150 155 160
Lys Lys Thr Ser Trp Asn Lys Ile Ser Ser Arg Ser Thr Gln Ser Thr
165 170 175
His Gln Ser Gln Pro Gln Gln Gln Ser Ala Gln Asp Leu Val Leu Glu
180 185 190
Asp Asp Glu His Glu Tyr Asp Val Asp Asp Ile Asp Asp Pro Lys Tyr
195 200 205
Leu Pro Ile Thr Glu Ser Leu Arg Ile Thr His Leu Ile His Ser Ala
210 215 220
Thr Asn Ser Pro Ser Ile Asn Arg Lys Thr Lys Thr Val Ile Asn
225 230 235 240
Asp Lys Asp Gln Val Leu Ala Lys Leu Gly Gln Met Ile Ile Ser Gln
245 250 255
Ser Arg Ser Asn Ser Thr Ser Leu Pro Ala Ala Ala Asn Lys Glu Thr Thr
260 265 270
Ile Arg Ser Lys Ser Thr Gln Asn Ile Pro Arg Met Ala Thr Pro Val
275 280 285
Gln Leu Leu Leu Met Glu Ser Ile Thr Ser Glu Phe Phe Ser Ile Asn
290 295 300
Ser Ala Gly Leu Leu Gln Gln Ala Val Asn Ala Ser Phe Phe Leu
305 310 315 320
Lys Gln Lys Leu Gly Arg Ala Arg Leu Arg Leu Gln Asn Ser Asn
325 330 335
Pro Glu Pro Leu Val Val Ile Leu Ala Ser Asp Ser Asn Arg Ser Gly
 Ala Leu Ala Leu Ala Leu Gly Arg His Leu Cys Gin Thr Gly His Ile 355 360 365
Arg Val Ile Thr Leu Phe Thr Cys Ser Gin Asn Leu Gin Asp Ser 370 375 380
Met Val Lys Lys Gin Thr Asp Ile Tyr Lys Cys Gly Gly Lys Ile 395 390 395 400
Val Asn Ser Val Ser Ser Leu Glu Ser Ala Met Glu Thr Leu Asn Ser 405 410 415
Pro Val Glu Ile Val Ile Asp Ala Met Gin Gly Tyr Asp Cys Thr Leu 420 425 430
Ser Asp Leu Ala Gly Thr Ser Glu Val Ile Glu Ser Arg Ile Lys Ser 435 440 445
Met Ile Ser Trp Cys Asn Lys Gin Arg Gly Ser Thr Lys Val Trp Ser 450 455 460
Leu Asp Ile Pro Asn Gly Phe Asp Ala Gly Ser Glu Met Pro Asp Ile 465 470 475 480
Phe Phe Ser Asp Arg Ile Glu Ala Thr Gly Ile Ile Cys Ser Gly Trp 485 490 495
Pro Leu Ile Ala Ile Asn Leu Ile Ala Asn Leu Pro Ser Leu Glu 500 505 510
Asp Ala Val Leu Ile Asp Ile Gly Ile Pro Gin Gly Ala Tyr Ser Gin 515 520 525
Arg Thr Ser Leu Arg Lys Phe Gin Asn Cys Asp Leu Phe Val Thr Asp 530 535 540
Gly Ser Leu Leu Leu Asp Leu
545 550

-continued

<210> SEQ ID NO 51
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide primer
<400> SEQUENCE: 51
cggagagatgcgtcatacag

<210> SEQ ID NO 52
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide primer
<400> SEQUENCE: 52
tttgtcgcttgacatactcg

<210> SEQ ID NO 53
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide primer
<400> SEQUENCE: 53
atgcccgaagaagtcacatg
What is claimed is:

1. A genetically modified microorganism, wherein said genetic modification comprises introduction of an expression vector comprising the coding sequence of a protein of the Sm-like superfamily, and wherein said genetic modification results in elevated tolerance to stress or at least one inhibitor as compared to without the genetic modification.

2. The microorganism of claim 1, wherein said microorganism is selected from bacteria or fungi.

3. The microorganism of claim 2, wherein said microorganism is a bacterium selected from the group consisting of Acetobacterium, Bacillus, Streptococcus, Clostridium, Zymomonas, Anaerocellum, Caldicellulosiruptor, Thermoanaerobacter, Gluconobacter, and E. coli.

4. The microorganism of claim 3, wherein said bacterium is selected from the group consisting of C. thermocellum, Z. mobilis, Anaerocellum thermophilum, Caldicellulosiruptor saccharolyticus), Thermoaeronaerobacter sp. X514, and E. coli.

5. The microorganism of claim 2, wherein said microorganism is a bacterium and wherein said protein of the Sm-like superfamily is a bacterial Hfq protein.

6. The microorganism of claim 5, wherein said Hfq protein comprises an amino acid sequence selected from SEQ ID NO: 2, 4, 6, 8 or 10, or a functional derivative or homolog thereof that shares at least 95% sequence identity therewith.

7. The microorganism of claim 5, wherein said microorganism is Z. mobilis.

8. The microorganism of claim 7, wherein said Hfq protein comprises the sequence as set forth in SEQ ID NO: 2.

9. The microorganism of claim 2, wherein said microorganism is a fungal species selected from Saccharomyces sp., Kluyveromyces sp., Pichia sp., Candida sp., and Schizosaccharomyces sp.

10. The microorganism of claim 9, wherein said fungal species is yeast selected from S. cerevisiae or P. pastoris.

11. The microorganism of claim 10, wherein said protein of the Sm-like superfamily is a yeast Sm or Lsm protein.

12. The microorganism of claim 11, wherein said protein comprises an amino acid sequence selected from any one of SEQ ID NOS: 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 or 50, or a functional derivative or homolog thereof that shares at least 95% sequence identity therewith.

13. The microorganism of claim 1, wherein said expression vector is a replicative vector or an integrative vector.

14. The microorganism of claim 1, wherein said stress is environmental stress selected from the group of high temperatures, low temperatures, low pH, oxidation, osmotic, and drought.

15. The microorganism of claim 1, wherein said at least one inhibitor is selected from the group consisting of an acetate
16. The microorganism of claim 15, wherein said acetate salt is selected from the group consisting of sodium acetate, ammonium acetate and potassium acetate.

17. The microorganism of claim 1, wherein said enhanced tolerance is characterized by ability to grow in a media containing sodium acetate at a concentration of 195 mM.

18. A method of producing alcohol from a cellulosic biomass material, comprising adding a genetically modified microorganism according to any one of claims 1-17 to a fermentation mixture comprising a cellulosic biomass material and/or fermentation substrates derived from said cellulosic biomass material, allowing said microorganism to ferment and produce alcohol, and recover alcohol produced.