TITLE OF INVENTION

REGULATED GENE EXPRESSION SYSTEMS AND CONSTRUCTS THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benfit of priority to U.S. Provisional Application Serial No. 61/467,873, March 24, 2011, which is incorporated herein by reference in its entirety.

MATERIAL INCORPORATED-BY-REFERENCE

The Sequence Listing, which is a part of the present disclosure, includes a computer readable form comprising nucleotide or amino acid sequences of the present invention. The subject matter of the Sequence Listing is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

Large scale production of materials by fermentation is often limited by the costs associated with regulating gene expression. Often the strategies employed require expensive sugar derivatives or other small molecules to trigger protein production leading to desired products.

Promoters are nucleic acid molecules which comprise the 5' regulatory elements that play an integral part in the overall expression of genes in living cells. Isolated promoters that function in host cells or organisms are useful for controlling the expression of operably linked transgenes and thereby modifying host organism or cell phenotypes through the methods of genetic engineering.

NtcA is the primary controller of the nitrogen regulon of cyanobacteria, having a DNA-binding sequence ($GTAN_8TAC$). Depending upon placement relative to the promoter and the transcription start position, NtcA binding can either activate or repress transcription (reviewed in Herrero et al. 2001 and Muro-Pastor et al. 2005). NtcA-mediated control of transcription is influenced by the nitrogen to-carbon ratio, with 2-oxoglutarate functioning as the effector molecule. Emlyn-Jones et al. (2003) disclosed the use of the nitrite reductase promoter for mutS expression, allowing control of the mutation frequency of Synechococcus elongatus PCC 7942 by varying the nitrogen source.

The ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCo) promoter from cyanobacteria is thought to be one of the strongest known promoters, since RuBisCo is one of the proteins expressed in highest abundance in the biosphere.

SUMMARY OF THE INVENTION

Among the various aspects of the present disclosure is the provision of a nitrogen-sensitive expression system for the expression of a transcribable nucleic acid molecule in a host cell.

One aspect provides nitrogen-sensitive expression system. In some embodiments, the expression system includes a transcription factor region comprising an NtcA binding site; and a core promoter region comprising a RuBisCo promoter or a variant or a functional fragment thereof. In some embodiments, the core promoter region comprises or is operably linked to the transcription factor region. Some embodiments include a transcribable nucleic acid molecule operably linked to the 3' transcription termination nucleic acid molecule.

Another aspect provides a method for expressing a transcribable nucleic acid molecule in a host cell. In some embodiments, a host cell is stably transformed with an expression system described herein and the host, or progeny having the expression system, is grown under conditions whereby the transcribable nucleic acid molecule is expressed. The expression system can provide for nitrogen-regulated expression of the transcribable nucleic acid molecule in a host cell. In some embodiments, expression of the transcribable nucleic acid molecule is repressed when nitrate is a predominant nitrogen source and the transcribable nucleic acid molecule is expressed when urea or ammonia is the predominant nitrogen source.

Another aspect provides an expression cassette. In some embodiments, the expression cassette includes (a) a transcription factor region comprising an NtcA binding site; (b) a core promoter region comprising a RuBisCo promoter or a variant or a functional fragment thereof; (c) a transcribable nucleic acid molecule; and (d) a 3' transcription termination nucleic acid molecule; wherein, the core promoter region comprises or is operably linked to the transcription factor region; the core promoter region is operably linked to the transcribable nucleic acid molecule; the transcribable nucleic acid molecule is operably linked to the 3' transcription termination nucleic acid molecule; and elements (a)-(d) are positioned in relation to each other such that expression of the expression cassette in a host organism results in production of a polypeptide sequence encoded by the transcribable nucleic acid molecule. In some embodiments, the expression cassette includes a polynucleotide sequence encoding an NtcA polypeptide or an NtcB polypeptide positioned such that expression of the expression cassette in the host organism results in production of the NtcA polypeptide or the NtcB polypeptide.

Another aspect provides a transgenic host cell. In some embodiments, the transgenic host cell, or progeny thereof, includes an expression system described herein. In some embodiments, the transgenic host cell is produced by the method for expressing a transcribable nucleic acid molecule described herein, or the progeny thereof, comprising the expression system. In some embodiments, the transgenic host cell includes the expression cassette described herein.

In some embodiments, the transgenic host cell expresses the polypeptide sequence encoded by the transcribable nucleic acid molecule in the presence of nitrate; and represses expression of the polypeptide sequence encoded by the transcribable nucleic acid molecule in the presence of ammonia.

In some embodiments, the transgenic host cell expresses the polypeptide sequence encoded by the transcribable nucleic acid molecule in the presence of ammonia; and represses expression of the polypeptide sequence encoded by the transcribable nucleic acid molecule in the presence of nitrate.

Another aspect provides a kit including one or more of an expression system or expression cassette described herein and, optionally, instructions for introducing the expression cassette into a host cell.

Other objects and features will be in part apparent and in part pointed out hereinafter.

## DESCRIPTION OF THE DRAWINGS

Those of skill in the art will understand that the drawings, described below, are for illustrative purposes only. The drawings are not intended to limit the scope of the present teachings in any way.

FIG. 1 is a cartoon depicting the structure of plasmid pLybAL50, a derivative of pLybAL19, where the primary differences are that the GTG start codon of *asf* has been replaced with the more conventional ATG codon and a His6-tag has been appended to the C-terminus of asf. Further details regarding methodology are disclosed in Example 3.

FIG. 2 is a cartoon depicting the "RuBisCo Swap." The top group of arrows shows the *Synechocystis* sp. PCC 6803 RuBisCo operon. The bottom group of arrows shows the synthetic operon comprised of the *Synechocystis* sp. PCC 6803 sps and spp genes, as found in plasmid pLybAL67. The synthetic *Synechocystis* sp. PCC 6803 sps/spp operon should have the same transcriptional and translational initiation and termination signals as the original

*Synechocystis* sp. PCC 6803 RuBisCo operon. Plasmid pLybAL66 is the same as pLybAL67, except the XmaI/SpeI fragment of the sps gene is absent. Further details regarding methodology are disclosed in Example 3.

FIG. 3A is a graphical and polynucleotide sequence depiction of the Nostoc sp. PCC 7120 promoter. The sequence covers the region between the rbcLXS (which begins 3 bp after the BamHI site) and all1523 (which is transcribed in the direction opposite that of the rbcLXS operon) and is shown in capital letters. Numbering is relative to the transcription start site. The nucleotides upstream and downstream from the core (which contains the promoter, transcription start, and the NtcA and Factor 2 binding sites) that have been deleted from the various constructs are highlighted. The constructs in which the region downstream of the core has been deleted leave the ribosome binding site intact. Further details regarding methodology are disclosed in Example 4. FIG. 3B shows a blown up portion of FIG. 3A.

## DETAILED DESCRIPTION OF THE INVENTION

The present disclosure relates to the use of a defined gene regulatory sequence to control the expression of functional genes according to culture media composition.

The present disclosure is based at least in part on the recognition that gene expression under a strong promoter, such as a RuBisCo promoter, may require regulation to provide for adequate or optimal organism growth. The present disclosure is also based at least in part on the recognition that a promoter containing an NtcA binding site (e.g., *rbcL*) can be used to regulate target gene expression where the target gene is repressed when nitrate is the sole nitrogen source, but expressed when grown with ammonical nitrogen, or other non-nitrate reduced nitrogen sources, such as urea. This system is understood to be an NtcA-repressed system. It is also recognized that the expression systems described herein can be used to regulate target gene expression where the target gene is expressed when nitrate is the sole nitrogen source, but repressed when grown with ammonical nitrogen, or other non-nitrate nitrogen sources, such as urea. This system is understood to be an NtcA-activated system. As reported herein, while a nitrite reductase promoter containing an NtcA binding site can regulate gene expression, the nitrite reductase promoter is relatively weak leading to unacceptably low protein production.

Various embodiments provide compositions and methods that utilize a nitrogen source of a culture medium to control gene regulation and protein expression.

Further provided is modulated gene expression by providing a construct comprising a

nitrogen-sensitive transcription factor region and a core promoter region, or variant or functional fragment thereof. Such a construct can afford high levels of protein production in the system at defined points of a cultivation process. Various systems described herein can be applied to production of protein materials including, but not limited to, sugar biosynthetic enzymes or other industrial enzymes, such as ester hydrolases. A system described herein can be applied to produce small molecules (including, but not limited to, sugars) derived from the action corresponding enzymes.

Except as otherwise noted herein, compositions and processes of the present disclosure can be carried out in accordance with compositions and processes described in US App Pub No. 2009/0181434, filed January 5, 2009, incorporated herein by reference in its entirety.

A nitrogen regulated expression system as described herein can be particularly advantageous when used in conjunction with a bioreactor, such as a solid phase bioreactor. When used in conjunction with a bioreactor, on-the-fly modulation of protein expression can be obtained by altering the nitrogen source within the feed. A solid phase bioreactor can be according to that disclosed in US App Pub No. 2009/0181434, filed January 5, 2009, incorporated herein by reference in its entirety.

The following definitions and methods are provided to better define the present invention and to guide those of ordinary skill in the art in the practice of the present invention. Unless otherwise noted, terms are to be understood according to conventional usage by those of ordinary skill in the relevant art.

The terms "heterologous DNA sequence", "exogenous DNA segment" or "heterologous nucleic acid," as used herein, each refer to a sequence that originates from a source foreign to the particular host cell or, if from the same source, is modified from its original form. Thus, a heterologous gene in a host cell includes a gene that is endogenous to the particular host cell but has been modified through, for example, the use of DNA shuffling. The terms also include non-naturally occurring multiple copies of a naturally occurring DNA sequence. Thus, the terms refer to a DNA segment that is foreign or heterologous to the cell, or homologous to the cell but in a position within the host cell nucleic acid in which the element is not ordinarily found. Exogenous DNA segments are expressed to yield exogenous polypeptides. A "homologous" DNA sequence is a DNA sequence that is naturally associated with a host cell into which it is introduced.

Expression vector, expression construct, plasmid, or recombinant DNA construct is

generally understood to refer to a nucleic acid that has been generated via human intervention, including by recombinant means or direct chemical synthesis, with a series of specified nucleic acid elements that permit transcription or translation of a particular nucleic acid in, for example, a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector can include a nucleic acid to be transcribed operably linked to a promoter.

A "promoter" is generally understood as a nucleic acid control sequence that directs transcription of a nucleic acid. An inducible promoter is generally understood as a promoter that mediates transcription of an operably linked gene in response to a particular stimulus. A promoter can include necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter can optionally include distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription.

A "transcribable nucleic acid molecule" as used herein refers to any nucleic acid molecule capable of being transcribed into a RNA molecule. Methods are known for introducing constructs into a cell in such a manner that the transcribable nucleic acid molecule is transcribed into a functional mRNA molecule that is translated and therefore expressed as a protein product. Constructs may also be constructed to be capable of expressing antisense RNA molecules, in order to inhibit translation of a specific RNA molecule of interest. For the practice of the present disclosure, conventional compositions and methods for preparing and using constructs and host cells are well known to one skilled in the art (see e.g., Sambrook and Russel (2006) Condensed Protocols from Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, ISBN-10: 0879697717; Ausubel et al. (2002) Short Protocols in Molecular Biology, 5th ed., Current Protocols, ISBN-10: 0471250929; Sambrook and Russel (2001) Molecular Cloning: A Laboratory Manual, 3d ed., Cold Spring Harbor Laboratory Press, ISBN-10: 0879695773; Elhai, J. and Wolk, C. P. 1988. Methods in Enzymology 167, 747-754).

The "transcription start site" or "initiation site" is the position surrounding the first nucleotide that is part of the transcribed sequence, which is also defined as position +1. With respect to this site all other sequences of the gene and its controlling regions can be numbered. Downstream sequences (i.e., further protein encoding sequences in the 3' direction) can be denominated positive, while upstream sequences (mostly of the controlling regions in the 5' direction) are denominated negative.

"Operably-linked" or "functionally linked" refers preferably to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a regulatory DNA sequence is said to be "operably linked to" or "associated with" a DNA sequence that codes for an RNA or a polypeptide if the two sequences are situated such that the regulatory DNA sequence affects expression of the coding DNA sequence (i.e., that the coding sequence or functional RNA is under the transcriptional control of the promoter). Coding sequences can be operably-linked to regulatory sequences in sense or antisense orientation. The two nucleic acid molecules may be part of a single contiguous nucleic acid molecule and may be adjacent. For example, a promoter is operably linked to a gene of interest if the promoter regulates or mediates transcription of the gene of interest in a cell.

A "construct" is generally understood as any recombinant nucleic acid molecule such as a plasmid, cosmid, virus, autonomously replicating nucleic acid molecule, phage, or linear or circular single-stranded or double-stranded DNA or RNA nucleic acid molecule, derived from any source, capable of genomic integration or autonomous replication, comprising a nucleic acid molecule where one or more nucleic acid molecule has been operably linked.

A constructs of the present disclosure can contain a promoter operably linked to a transcribable nucleic acid molecule operably linked to a 3' transcription termination nucleic acid molecule. In addition, constructs can include but are not limited to additional regulatory nucleic acid molecules from, e.g., the 3'-untranslated region (3' UTR). Constructs can include but are not limited to the 5' untranslated regions (5' UTR) of an mRNA nucleic acid molecule which can play an important role in translation initiation and can also be a genetic component in an expression construct. These additional upstream and downstream regulatory nucleic acid molecules may be derived from a source that is native or heterologous with respect to the other elements present on the promoter construct.

The term "transformation" refers to the transfer of a nucleic acid fragment into the genome of a host cell, resulting in genetically stable inheritance. Host cells containing the transformed nucleic acid fragments are referred to as "transgenic" cells, and organisms comprising transgenic cells are referred to as "transgenic organisms".

"Transformed," "transgenic," and "recombinant" refer to a host cell or organism such as a bacterium, cyanobacterium, animal or a plant into which a heterologous nucleic acid molecule has been introduced. The nucleic acid molecule can be stably integrated into the genome as generally known in the art and disclosed (Sambrook 1989; Innis 1995; Gelfand

1995; Innis & Gelfand 1999). Known methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially mismatched primers, and the like. The term "untransformed" refers to normal cells that have not been through the transformation process.

"Wild-type" refers to a virus or organism found in nature without any known mutation.

***EXPRESSION SYSTEM COMPONENTS***

One aspect provides an expression regulation system sensitive to a nitrogen source of a media, such as a growth media or fermentation media.

In various embodiments, an expression regulation system includes a transcription factor region and a core promoter region operably linked to a target sequence.

A transcribable nucleic acid molecule sequence can be regulated at the transcriptional level by changing a nitrogen source in the media. For example, mRNA transcription of a transcribable nucleic acid molecule sequence can be switched off in the presence of nitrate or switched on in the presence of a non-nitrate source of reduced nitrogen, such as ammonia or urea. Thus, an expression regulation system described herein can provide controllable expression of a heterologous gene using the inexpensive repressor nitrate and inducer ammonium.

As another example, mRNA transcription of a transcribable nucleic acid molecule sequence can be switched off in the presence of a non-nitrate source of reduced nitrogen, such as ammonia or urea, or switched on in the presence of nitrate. Thus, an expression regulation system described herein can provide controllable expression of a heterologous gene using the inexpensive inducer nitrate and repressor ammonium.

Inclusion of a termination region control sequence is optional, and if employed, then the choice is be primarily one of convenience, as the termination region is relatively interchangeable. The termination region may be native to the transcriptional initiation region (the promoter), may be native to the DNA sequence of interest, or may be obtainable from another source.

A promoter of the present disclosure can be incorporated into a construct using marker genes as described and tested for an indication of gene expression in a stable host

system. As used herein the term "marker gene" refers to any transcribable nucleic acid molecule whose expression can be screened for or scored in some way.

***TRANSCRIPTION FACTOR REGION***

Provided herein is an expression regulation system that can include a nitrogen sensitive transcription factor region. Using an embodiment of a nitrogen regulation system as described herein, a switch from nitrate to ammonical nitrogen (or another non-nitrate reduced form of nitrogen, such as urea) in a culture medium (*e.g.*, a fermentation broth) can result in a loss of repression prohibiting expression of an operably linked gene, thereby effectively turning on protein production at defined times within the cultivation process. In another embodiment, a switch from ammonical nitrogen (or another non-nitrate reduced form of nitrogen, such as urea) to nitrate to in a culture medium can result in a loss of repression prohibiting expression of an operably linked gene, thereby effectively turning on protein production at defined times within the cultivation process.

The present disclosure is based, at least in part, on observations related to NtcA, a primary controller of the nitrogen regulon of cyanobacteria. Results presented herein show that genes operably linked to a nitrite reductase promoter (which contains an NtcA binding site), isolated from *Synechocystis* sp. PCC 6803 and Synechococcus elongatus PCC 7942, are expressed when nitrate is the sole nitrogen source, but repressed when grown with ammonical nitrogen (see Example 2). Such a system is consistent with an NtcA activated promoter system (see Hererro 2001 J Bacteriol 183(2) 411-425, 416). These results demonstrate that controlled protein expression, and thus natural product production, can be achieved by alteration of inexpensive sources of nitrogen. While prior reports, such as Emlyn-Jones et al. (2003), have reported use of nitrite reductase promoter for gene regulation, the Inventors have found that while some degree of controlled expression may be achieved, the nitrite reductase promoter in prior systems is relatively weak, which leads to unacceptably low protein production (see Example 2).

The transcription factor region can be operably linked to a core promoter region. The present disclosure is also based, at least in part, on observations from the use of an NtcA binding sequence operably linked to a strong promoter, such as a RuBisCo promoter. Results presented herein show that genes operably linked to a rbcL promoter (which contains an NtcA binding site), or functional fragments thereof, isolated from Nostoc sp. PCC 7120 are expressed when ammonia is the sole nitrogen source, but repressed when grown with

ammonical nitrogen (see Example 2). Such a system is consistent with an NtcA repressed promoter system (see Hererro 2001 J Bacteriol 183(2) 411-425, 420).

A transcription factor region can have various positions related to a core promoter region. For example, the transcription factor region can be upstream or downstream of a core promoter region. The transcription factor region can be proximate to a core promoter region. The transcription factor region can be distal to a core promoter region. The transcription factor region can be contained within a core promoter region. For example, the transcription factor region can be integrated within the sequence of a core promoter region.

A transcription factor region can comprise a nitrogen sensitive transcription factor. Transcription factors generally can play a role in regulating gene expression in various organisms by activating or repressing expression by, for example, binding to DNA at sites typically upstream of a transcription start site. A transcription factor can exert their function by recruiting polymerases to the coding DNA molecule. A transcription factor, or sequence-specific DNA-binding factor, is generally understood as a protein that binds to specific DNA sequences, thereby controlling the transcription of genetic information from DNA to mRNA. A transcription factor can perform this function alone or with other proteins in a complex by promoting (activator), or blocking (repressor), the recruitment of RNA polymerase which is the enzyme that performs the transcription of genetic information from DNA to RNA to specific genes. One feature of a transcription factor is that they contain one or more DNA-binding domains, which can attach to specific sequences of DNA adjacent to the genes that they regulate.

A nitrogen sensitive expression regulation system can include a promoter, or a functional fragment or variant thereof, associated with a nitrogen metabolism related gene. A nitrogen sensitive expression regulation system can include a promoter, or a functional fragment or variant thereof, of a nitrite reductase gene. For example, a nitrogen sensitive expression regulation system can include a promoter, or a functional fragment or variant thereof, of a nitrite reductase gene from a cyanobacteria. As a further example, a nitrogen sensitive expression regulation system can include an NtcA-regulated nitrite reductase promoter isolated from *Synechocystis* or *Synechococcus*, or a functional fragment or variant thereof. The NtcA binding site in the endogenous nitrite reductase promoter is understood to function as an activator site by way of its distance from the transcription start site (see Herrero 2001 J Bacteriol 183(2) 411-425). As a further example, a nitrogen sensitive expression regulation system can include an NtcB-regulated nitrite reductase promoter.

A nitrogen sensitive expression regulation system can include a promoter, or a functional fragment or variant thereof, of a nitrogen-regulated gene. A nitrogen sensitive expression regulation system can include a promoter, or a functional fragment or variant thereof, of a nitrogen-regulated RuBisCo gene. For example, a nitrogen sensitive expression regulation system can include a promoter comprising one or more of an NtcA binding site or NtcA binding site consensus sequence. As another example, a nitrogen sensitive expression regulation system can include a promoter comprising one or more of an NtcB binding site or NtcB binding site consensus sequence. As another example, a nitrogen sensitive expression regulation system can include a promoter comprising one or more of an NtcA binding site and an NtcB binding site or an NtcA binding site consensus sequence and an NtcB binding site consensus sequence. As a further example, a nitrogen sensitive expression regulation system can include a promoter of a nitrogen-regulated RuBisCo gene from an algae or a cyanobacteria. As a further example, a nitrogen sensitive expression regulation system can include a promoter of a nitrogen-regulated RuBisCo gene from *Nostoc*, *Synechocystis* or *Synechococcus*, or a functional fragment or variant thereof. The NtcA binding site in the endogenous *rbcL* promoter from Nostoc is understood to function as a repressor site by way of its proximity to the transcription start site (see Herrero 2001 J Bacteriol 183(2) 411-425).

NtcA is understood to be a helix-turn-helix transcriptional regulator that binds to promoters of nitrogen-regulated genes at various binding domains (see Su et al. 2005 Nucleic Acids Research 33(16), 5156-5171; Llacer et al. 2010 Proc Natl Acad Sci USA 107(35), 15397-402). In an endogenous cyanobacterial NtcA regulon, where ammonium is a preferred nitrogen source, expression of genes for the assimilation of other sources of nitrogen is suppressed in an ammonium-replete environment; while in an ammonium-limited environments, genes involved in the uptake and metabolism of alternative sources of nitrogen can be induced through the binding of NtcA to the cis-regulatory elements in the promoter regions of the genes.

A transcription factor region can comprise one or more of an NtcA nitrogen sensitive transcription factor binding site, an NtcB nitrogen sensitive transcription factor binding site, or a variant or functional fragment thereof. A transcription factor region can comprise at least one each of an NtcA nitrogen sensitive transcription factor binding site and an NtcB nitrogen sensitive transcription factor binding site, or a variant or functional fragment thereof. For example, a transcription factor region can comprise the NtcA nitrogen sensitive transcription factor binding site and the NtcB nitrogen sensitive transcription factor binding

site present in SEQ ID NO: 279 (which sequence is a NirA promoter operably linked to NtcA and NtcB), or a fragment thereof, or a sequence having at least at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity thereto.

A transcription factor region can comprise a polynucleotide having a sequence that includes an NtcA binding site consensus sequence or an NtcB binding site consensus sequence. A transcription factor region can comprise a polynucleotide having a sequence that includes both an NtcA binding site consensus sequence and an NtcB binding site consensus sequence. A transcription factor region can comprise a polynucleotide having a sequence that includes both an NtcA binding site consensus sequence and an NtcB binding site consensus sequence present in SEQ ID NO: 279 (which sequence is a NirA promoter operably linked to NtcA and NtcB), or a fragment thereof, or a sequence having at least at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity thereto.

An Ntca binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding sequence functionally associated with nitrogen-regulated genes of organisms such as Nostoc, Synechocystis, or Synechococcus. For example, an NtcA binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding sequence functionally associated with promoter regions of nitrogen-regulated genes (e.g., nitrogen-regulated cyanobacterial genes). Promoter regions for *nir* operon, *nirB-ntcB*, *ntcA*, *glnA*, *glnB*, *amt1*, *urt* operon, *ntcB*, *hetC*, *dev*BCA, *icd*, *rpo*D2-V, *nrtP*, *glnN*, *nifP*, *petH*, *nifH*, *nif* ORF1, *vnf*DG, and *nif*HDK are understood or predicted to act as NtcA-activated promoters (see Herrero 2001 J Bacteriol 183(2) 411-425, 417, 418). Promoter regions for *rbcL*, *hanA gor*, gi*f*A, and *gifB* are understood or predicted to act as NtcA-repressed promoters (see Herrero 2001 J Bacteriol 183(2) 411-425, ). Other promoter regions containing an NtcA binding site include *xisA*, *glbN*, and *nifH*. Any of the above sequences, or a variant or functional fragment thereof containing an NtcA binding site can be included in an expression system described herein.

For example, an NtcA binding site sequence of a construct described herein can have a sequence the same as or similar to an NtcA binding sequence functionally associated with promoter regions of nirA (SEQ ID NO: 26, SEQ ID NO; 32), or sequence having at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity thereto, or a fragement thereof.

For example, an NtcA binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding site in one or more of the following genomes: Gloeobacter violaceus PCC 7421(PCC7421); Nostoc sp. PCC 7120(PCC7120); Prochlorococcus marinus CCMP1375(PCC1375); Prochlorococcus marinus MED4(MED4); Prochlorococcus marinus MIT9313 (MIT9313); Synechococcus elongatus PCC 6301(PCC6310); Synechococcus sp. WH8102(WH8102); Synechocystis sp. PCC 6803(PCC6803) and Thermosynechococcus elongates BF-1(thermosynechococcus). As another example, an NtcA binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding site disclosed in Su et al. 2005 Nucleic Acids Research 33(16), 5156-5171, incorporated herein by reference. As another example, an NtcA binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding site disclosed in Jiang et al. 1997 Biochem J 327, 513-517, incorporated herein by reference.

As another example, an NtcA binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding site contained in a RuBisCo promoter sequence from *Nostoc* sp. PCC 7120. As another example, an NtcA binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding site contained in any of SEQ ID NO: 180, SEQ I DNO: 181, SEQ ID NO: 182, SEQ ID NO: 183 (Nsp7120 rbc Promoter), SEQ ID NO: 227 (Ssp6803 RuBisCo promoter), SEQ ID NO: 231 (Selo7942 RuBisCo promoter), or SEQ ID NO: 234 (Nsp7120 rbc promoter).

As another example, an NtcA binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding site contained in plasmids pLybDB2 (SEQ ID NO: 188), pLybDB3 (SEQ ID NO: 189), pLybDB4 (SEQ ID NO: 190), pLybDB5 (SEQ ID NO: 191), pLybDB6 (SEQ ID NO: 235), pLybDB7 (SEQ ID NO: 236), or pLybDB9 (SEQ ID NO: 237). As another example, an NtcA binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding site

contained in plasmid pLybDB4 (SEQ ID NO: 190).

As another example, an NtcA binding site sequence or an NtcB binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding site contained in plasmids pLybAL106 (SEQ ID NO: 272) or pLybAL107 (SEQ ID NO: 273). For example, an NtcA binding site sequence of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to an NtcA binding site contained in plasmid pLybAL106 (SEQ ID NO: 272).

A transcription factor region can comprise a polynucleotide having a sequence that includes a $GTAN_{11}C$ consensus sequence, which can be a canonical binding site consensus sequence for NtcA. The GTA triplet is considered to be the most conserved region of this consensus sequence. As a component of the presently described expression system, the NtcA binding site consensus sequence can be positioned such that binding of NtcA effects control of expression of a transcribable nucleic acid molecule. For example, the $GTAN_{11}C$ domain is known to be centered at 39.5 to 40.5 nucleotides upstream of the transcription start site, and can also found in some promoters at upstream positions -109.5 and -180.5. It is understood that such positioning is exemplary and other positions that maintain functionality are contemplated.

A transcription factor region can comprise a polynucleotide having a sequence that includes a $GTAN_8TAC$ consensus sequence. The $GTAN_8TAC$ binding domain is considered a canonical binding site consensus sequence for NtcA.

A transcription factor region can comprise a polynucleotide having a sequence that includes a $GTAN_8TGC$ consensus sequence.

A transcription factor region can comprise a polynucleotide having a sequence that includes a $GTN_{10}AC$ consensus sequence.

A transcription factor region can comprise a polynucleotide having a sequence that includes a $TGTN_9ACA$ consensus sequence. A transcription factor region can comprise a polynucleotide having a sequence that includes a $TGTN_{10}ACA$ consensus sequence.

A transcription factor region can comprise an additional binding domain, such as a -10 like box in the form of $TAN_3T/A$. For example, a transcription factor region can $TAN_3T$ binding domain. As another example, a transcription factor region can $TAN_3A$ binding

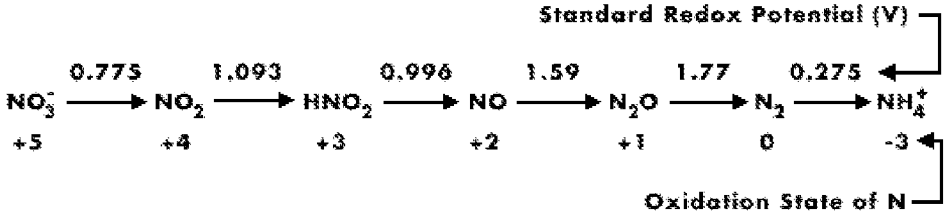
domain. Such a binding domain can be similar to a -10, E. coli $\sigma^{70}$ -like box in the form of $TAN_3T$ with an Ntca binding site (which can be as discussed above) replacing the -35 box present in E. coli $\sigma^{70}$ -type3 promoters (see Su et al. 2005 Nucleic Acids Research 33(16), 5156-5171). In some embodiments, the -10 like box can be present in Ntca-activated promoter constructs (see Herrero 2001 J Bacteriol 183(2) 411-425, 418, 419). A $TAN_3T$ box binding domain can lie, for example, five to six nucleotides upstream from a transcription start site. A $TAN_3T$ box binding domain can lie, for example, 22 or 23 nucleotides upstream from an NtcA binding site consensus sequence. A $TAN_3T$ box binding domain can lie, for example, about 21, about 22, about 23, or about 24 nucleotides upstream from an NtcA binding site consensus sequence (see Herrero 2001 J Bacteriol 183(2) 411-425, 419).

In some embodiments, an NtcA binding site as discussed above can be about 21 to about 24, e.g., 22 or 23, nucleotides downstream of a -10 like box, for example, in constructs where NtcA binding acts as an activator. In some embodiments, an NtcA binding site as discussed above can be near the transcription start site or overlap the -10 like box, for example, in constructs where NtcA binding acts as a repressor.

A transcription factor region can comprise one or more of an NtcA binding site sequence or an NtcA binding site consensus sequence. For example, transcription factor region can comprise at least two, at least three, at least four, at least five, or more, of any NtcA binding site sequence or an NtcA binding site consensus sequence discussed above.

The choice of nitrogen source for inducer and repressor in an embodiment of the expression control system comprising a nitrogen-sensitive transcription factor region can be according to compounds known to function with that transcription factor (see Herrero et al. 2001 J Bacteriol 183(2), 411-425, incorporated herein by reference). For example, in some embodiments, in the presence of nitrate (having a nitrogen oxidation state of +5), expression of transcribable nucleic acid molecule can occur when operably linked to an NtcA activated promoter system; while in the presence of a reduced nitrogen source such as ammonia (having a nitrogen oxidation state of -3), expression from such system is repressed. Conversely, in other embodiments, in the presence of nitrate (having a nitrogen oxidation state of +5), expression of transcribable nucleic acid molecule is repressed when operably linked to an NtcA repressed promoter system; while in the presence of a reduced nitrogen source such as ammonia (having a nitrogen oxidation state of -3), expression from such system is activated. As such, in some embodiments, a nitrogen source having a lower oxidation state than nitrate can have an opposite effect on an NtcA regulated system. In other

words, a nitrogen source having a lower oxidation state than nitrate (e.g., ammonia, urea) can act as an expression repressor in an NtcA activated promoter system; and a nitrogen source having a lower oxidation state than nitrate (e.g., ammonia, urea) can act as an expression inducer in an NtcA repressed promoter system.




A nitrate can include a salt thereof, such as potassium nitrate or sodium nitrate. A nitrogen compound having a lower oxidation state than nitrate can include, but is not limited to, urea, cyanate, ammonia, ammonium sulfate, amino acids, nitrogen dioxide, nitric oxide, or nitrous oxide. Preferably, a nitrogen compound having a lower oxidation state than nitrate is a soluble nitrogen compound such as ammonia or urea.

*NITROGEN REGULON REGULATORY PROTEINS*

In some embodiments, addition of excess nitrogen regulon regulatory proteins to the system can further reduce or eliminate expression under non-inducing conditions. As discussed above, an expression regulation system can include a nitrogen sensitive transcription factor region sensitive to nitrogen regulon regulatory proteins (e.g., nitrate inducible nitrogen regulon regulatory proteins). In some embodiments, addition of excess nitrogen regulon regulatory proteins (e.g., NtcA or NtcB) can further reduce or eliminate expression of the target nucleotide sequence under non-inducing conditions (see e.g., Example 9).

A construct described herein can include a nucleotide sequence encoding one or more copies of NtcA or NtcB, or both. For example, a construct described herein can include a nucleotide sequence of SEQ ID NO: 279, or a fragment thereof containing portions encoding NtcA or NtcB or both, or a nucleotide having at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity to SEQ ID NO: 279 and encoding NtcA or NtcB or both, or a fragment thereof. As another example, a construct including a nucleotide containing portions encoding NtcA or NtcB can be pLybAL98 (SEQ ID NO: 265).

A nucleotide sequence encoding one or more copies of NtcA or NtcB can be upstream of a transcription factor region. A nucleotide sequence encoding one or more copies of NtcA or NtcB can be downstream of a transcription factor region.

A nucleotide sequence encoding one or more copies of NtcA or NtcB can be positioned within the construct described herein such that expression of the construct results in expression of NtcA polypeptide or NtcB polypeptide. A nucleotide sequence encoding one or more copies of NtcA or NtcB can be upstream of a core promoter region. A nucleotide sequence encoding one or more copies of NtcA or NtcB can be downstream of a core promoter region. For example, nucleotide sequence encoding one or more copies of NtcA or NtcB can be downstream of a nirA promoter (see e.g., Example 9).

***CORE PROMOTER REGION***

Provided herein is a nitrogen-sensitive expression regulation system that can include a core promoter region. Using an embodiment of a nitrogen regulation system as described herein, a core promoter region (e.g., from a RuBisCo promoter) operably linked to a transcription factor region (e.g., an NtcA binding sequence in an Ntca-repressed configuration) can result in regulated expression of a transcribable nucleic acid molecule, where nitrate in a culture medium (*e.g.*, a fermentation broth) can effectively turn off expression and a non-nitrate reduced nitrogen source, such as ammonical nitrogen or urea, can effectively turn on protein production at defined times within the cultivation process. In other embodiments,

In another embodiment, a core promoter region operably linked to a transcription factor region (e.g., an NtcA binding sequence in an NtcA-activated configuration) can result in regulated expression of a transcribable nucleic acid molecule upon a switch from a non-nitrate nitrogen source, such as ammonical nitrogen or urea, to nitrate in a culture medium (*e.g.*, a fermentation broth), thereby effectively turning on protein production at defined times within the cultivation process.

The present disclosure is based, at least in part, on observations from the use of a strong RuBisCo promoter. While prior reports have reported use of a strong RuBisCo promoter, the Inventors have found that a RuBisCo promoter in the absence of a regulon (such as the NtcA regulon) results in uncontrolled expression of a transcribable nucleic acid molecule at the expense of host cell growth and viability. Results presented herein show that genes operably linked to a *Synechocystis* sp. PCC 6803 or *Synechococcus elongatus* PCC

7942 RuBisCo promoter are strongly expressed but the host organism exhibited extremely slow growth and loss of viability over time (see Example 4).

The present disclosure is also based, at least in part, on observations from the use of NtcA and its DNA-binding sequence operably linked to a strong promoter, such as a RuBisCo promoter. Initial experiments with an intact RuBisCo promoter comprising an NtcA binding site results in little to no target gene expression. Results presented herein show that genes operably linked to a (non-truncated) Nostoc sp. PCC 7120 RuBisCo promoter afforded little to no target gene expression (see Example 6).

Surprisingly, it has been discovered that upstream or downstream truncation of a RuBisCo promoter comprising an NtcA binding site can result significant nitrogen source regulated target gene expression. Results presented herein show that genes operably linked to a Nostoc sp. PCC 7120 RuBisCo promoter with either the leading or trailing regions of the promoter sequence deleted (i.e., upstream or downstream truncation) individually demonstrated significant target gene expression in the presence of ammonia (see Example 6).

Yet, truncation of both upstream and downstream regions of a RuBisCo promoter comprising an NtcA binding site results in little to no target gene expression. Results presented herein show that genes operably linked to a Nostoc sp. PCC 7120 RuBisCo promoter missing both flanking regions of sequence (i.e., upstream and downstream truncated) afforded little to no target gene expression (see Example 6).

A core promoter region can be operably linked to a transcription factor region. For example, a core promoter region can be upstream or downstream of a transcription factor region. A core promoter region can comprise a transcription factor region. For example, the transcription factor region can be integrated within the sequence of a core promoter region. Location of the transcription factor region with respect to the core promoter region can influence whether the system is an activated or repressed system in the presence of reduced ammonia, as described further herein. For example, where the transcription factor region is an NtcA binding site that is proximate to a core promoter region, as in the *rbcL* promoter, the expression system can be an NtcA repressor system, where the presence of a reduced nitrogen source, such as ammonia or urea, can turn on expression of an operably linked transcribable nucleic acid molecule, and the presence of nitrate can turn off expression.

A core promoter region can comprise a RuBisCo promoter sequence, or variant or functional fragment thereof.

A core promoter region can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to a RuBisCo promoter sequence of organisms such as Nostoc, Synechocystis, or Synechococcus. For example, a core promoter region of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to a RuBisCo promoter sequence in one or more of the following genomes: Gloeobacter violaceus PCC 7421(PCC7421); Nostoc sp. PCC 7120(PCC7120); Prochlorococcus marinus CCMP1375(PCC1375); Prochlorococcus marinus MED4(MED4); Prochlorococcus marinus MIT9313 (MIT9313); Synechococcus elongatus PCC 6301(PCC6310); Synechococcus sp. WH8102(WH8102); Synechocystis sp. PCC 6803(PCC6803) and Thermosynechococcus elongates BF-1(thermosynechococcus).

As another example, a core promoter region of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to a RuBisCo (*rbcLS*) promoter sequence in *Nostoc* sp. PCC 7120. As another example, a core promoter region of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to a RuBisCo (*rbcLS*) promoter sequence in *Synechocystis* sp. PCC 6803. As another example, a core promoter region of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to a RuBisCo (*rbcLS*) promoter sequence in *Synechococcus elongatus* PCC 7942. As another example, a core promoter region of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to a RuBisCo (*rbcLS*) promoter sequence in Anabaena PCC 7120. As another example, a core promoter region of a construct described herein can be a variant or functional fragment of a RuBisCo (*rbcLS*) promoter sequence.

A core promoter region of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to a RuBisCo (*rbcLS*) promoter sequence, or variant or functional fragment thereof. For example, a core promoter region of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to a rbc promoter region of Nostoc sp. PCC 7120, such as SEQ ID NO. 180, SEQ ID NO. 181, SEQ ID NO. 182, or SEQ ID NO. 183. As another example, a core promoter region of a construct described herein can have a sequence the same as or similar (e.g., at

least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to the Nsp7120 rbc promoter region of SEQ ID NO: 234 as follows:

```
      cctgcaggct ccagccatta gcgaaactga ccaaaggtta atctcagtct tcccttagtt  60
      gggatttttt aaacatatgg ggattgggga ttgggaaaag gcaggggggca gaggataagt
120
      agaggaggca ggggaagtaa agaagaatga ctatggacta ttgaccaatg actattggca
180
      actgacaact gactcctata cagtcattga tacattttgt aactgattgt taacaaaacg
240
      tttaaaactt tatgtaataa caaatttaaa tatgtaagtt aagaactttc aaagaataac
300
      ttatgccatt tcttgatata ttgtgagaca agttacaaat tacgtggtgt gcaatttttt
360
      catcttgcgc tgattactct actaaatatc cgtcaagtaa attggctctt agctcgtctc
420
      ctgtcaataa aggaggtcgg caagagtgca gaagcgggaa tgtgtgaaaa ctaacccaat
480
      tcattaaata ccccgaaata taggggaatc atctcatact ttccgtaaac cgcgaaggtc
540
      gtgaagggat aaaagcaatt tagtgggtga gaagaacaga taaaaaagaa ttttttaact
600
      atggcaagag gaaaaagtaa aagcgttaac ttatgcactc ctagatgagc aagacactgg
660
      tgaagaggat taccactaaa gctaagtgtt agttgcagaa aggtcgctga cctctaccaa
720
      aagattattc ctgtttttct cctggctgat agggaggtag ggcaattgtg agaggaaatt
780
gtaccaaaac gtgatttgat aagtaaaaag agtgacatct tggaaggatc catatg      836
```

For example, core promoter region of a construct described can have at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, and at least about 99% sequence identity to any one of SEQ ID NO: 234, SEQ ID NO. 180, SEQ ID NO. 181, SEQ ID NO. 182, or SEQ ID NO. 183.

A core promoter region of a construct described herein can have a sequence the same as or similar (e.g., at least about 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity) to a core promoter region contained in plasmids pLybDB2 (SEQ ID NO: 188), pLybDB3 (SEQ ID NO: 189), pLybDB4 (SEQ ID NO: 190), pLybDB5 (SEQ ID NO: 191), pLybDB6 (SEQ ID NO: 235), pLybDB7 (SEQ ID NO: 236), or pLybDB9 (SEQ ID NO: 237).

A core promoter region of a construct described herein can be a functional fragment of any promoter sequence discussed herein. For example, a core promoter region of a construct described herein can be a functional fragment of RuBisCo (e.g., *rbcLS*) promoter sequence. A functional fragment of a RuBisCo promoter sequence can be a nucleic acid sequence having at least about 50, at least about 75, at least about 95, at least about 100, at

least about 125, at least about 150, at least about 175, at least about 200, at least about 250, at least about 300, at least about 350, at least about 400, at least about 450, at least about 500, at least about 550, at least about 600, at least about 650, at least about 700, at least about 750, or at least about 800 contiguous bases of RuBisCo promoter sequence (e.g., *rbcLS* promoter from *Nostoc* sp. PCC 7120) that retain promoter activity.

A core promoter region of a construct described herein can be a truncated form of a RuBisCo (e.g., *rbcLS*) promoter sequence. For example, a core promoter region of a construct described herein can be a truncated form of a nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120.

Truncation of a RuBisCo promoter sequence can be an upstream truncation. For example, a RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 can be upstream truncated from position -319 to -162 (numbering relative to the transcription start site, see FIG. 3) (see Example 6, Example 7). Upstream truncation of a RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 can occur through about position -300, about position -250, about position -200, about position -190, about position -180, about position -170, about position -160, about position -150, about position -140, about position -130, about position -120, about position -110, about position -100, about position -90, or about position -80. The -319 to -162 position of *Nostoc* sp. PCC 7120 as depicted in FIG. 3 corresponds to position 1 to 176 of SEQ ID NO:234.

Truncation of a RuBisCo promoter sequence can be an downstream truncation. For example, a RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 can be downstream truncated from position +26 to +491 (numbering relative to the transcription start site, see FIG. 3) (see Example 6, Example 7). Downstream truncation of a RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 can occur at about position +10, about position +20, about position +30, about position +40, about position +50, about position +60, about position +70, about position +80, about position +90, about position +100, about position +150, about position +200, about position +250, about position +300, about position +350, about position +400, about position +450, about position +480, or about position +490. Downstream truncation of a RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 can occur through, for example, a restriction site present at or near the first codon of *rbc*L (e.g., BamHI restriction site at position +499). The +26 to +491 position of *Nostoc* sp. PCC 7120 as depicted in FIG. 3 corresponds to position 353 to 818 of SEQ ID NO:234.

*TRANSCRIBABLE NUCLEIC ACID MOLECULE*

As described herein, a target nucleotide sequence, such as a transcribable nucleic acid molecule, can be operably linked to an expression regulation system sensitive to one or more components of a media, such as a growth media or fermentation media. For example, a transcribable nucleic acid molecule can be operably linked to an expression regulation system sensitive to nitrogen source present in the media.

Exemplary transcribable nucleic acid molecules for incorporation into constructs of the present disclosure include, for example, nucleic acid molecules or genes from a species other than a host species, or even genes that originate with or are present in the same species, but are incorporated into recipient cells by genetic engineering methods rather than classical reproduction or breeding techniques. Exogenous gene or genetic element is intended to refer to any gene or nucleic acid molecule that is introduced into a recipient cell. The type of nucleic acid molecule included in the exogenous nucleic acid molecule can include a nucleic acid molecule that is already present in the host cell, a nucleic acid molecule from another organism, a nucleic acid molecule from a different organism, or a nucleic acid molecule generated externally, such as a nucleic acid molecule containing an antisense message of a gene, or a nucleic acid molecule encoding an artificial or modified version of a gene.

A transcribable nucleic acid molecule can be any sequence encoding a polypeptide of interest. For example, a transcribable nucleic acid molecule can be a gene encoding a polypeptide having a particular activity of interest.

A transcribable nucleic acid molecule can encode a polypeptide having a sucrose biosynthetic activity. A transcribable nucleic acid molecule can encode a polypeptide having sucrose phosphate synthase activity. A transcribable nucleic acid molecule can encode a polypeptide having sucrose phosphate phosphatase activity. A transcribable nucleic acid molecule can encode a polypeptide having sucrose phosphate synthase activity and sucrose phosphate phosphatase activity.

A transcribable nucleic acid molecule can be an active *sps/spp* fusion (*asf*) gene from *Synechococcus elongatus* PCC 7942 that produces an *asf* gene product, ASF, which has both SPS and SPP biosynthetic functions (*see e.g.*, US App Pub No. 2009/0181434, Example 5). In some embodiments, a target ASF-encoding nucleotide sequence is cloned from its native source (e.g., *Synechococcus elongatus* PCC 7942). In some embodiments, a transcribable nucleic acid molecule comprises an *asf* polynucleotide of SEQ ID NO: 1. In some embodiments, a transcribable nucleic acid molecule comprises a nucleotide sequence

encoding ASF polypeptide of SEQ ID NO: 2. In further embodiments, a transcribable nucleic acid molecule comprises a nucleotide sequence having at least about 80% sequence identity to SEQ ID NO: 1 or a nucleotide sequence encoding a polypeptide having *sps* and *spp* activity and at least about 80% sequence identity to SEQ ID NO: 2. As an example, a transcribable nucleic acid molecule can comprise a nucleotide sequence having at least about 85%, at least about 90%, at least about 95%, or at least about 99% sequence identity to SEQ ID NO: 1, wherein the expressed sequence exhibits ASF, SPS, or SPP activity. As an example, a transcribable nucleic acid molecule can comprise a nucleotide sequence encoding a polypeptide having at least about 85%, at least about 90%, at least about 95%, or at least about 99% sequence identity to SEQ ID NO: 2, wherein the expressed sequence exhibits ASF, SPS, or SPP activity. As another example, a transcribable nucleic acid molecule can comprise a nucleotide sequence that hybridizes under stringent conditions to SEQ ID NO: 1 over the entire length of SEQ ID NO: 1, and which encodes an active SPS/SPP fusion (ASF) polypeptide. As a further example, a transcribable nucleic acid molecule can comprise the complement to any of the above sequences.

In some embodiments, a transcribable nucleic acid molecule comprises a sucrose phosphate synthase (*sps*) (see e.g., SEQ ID NO: 3 encoding *sps* gene and SEQ ID NO: 4 encoding SPS polypeptide), or homologue thereof. For example, a transcribable nucleic acid molecule can comprise a nucleotide having a sequence of SEQ ID NO: 3 so as to express sucrose phosphate synthase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide having at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% percent identity to SEQ ID NO: 3 encoding a polypeptide having sucrose phosphate synthase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide sequence encoding a polypeptide having at least about 85%, at least about 90%, at least about 95%, or at least about 99% sequence identity to SEQ ID NO: 4, wherein the expressed sequence exhibits SPS activity.

In some embodiments, a transcribable nucleic acid molecule comprises a sucrose phosphate phosphatase (*spp*) (see e.g., SEQ ID NO: 5 encoding *spp* gene and SEQ ID NO: 6 encoding SPP polypeptide), or homologue thereof. For example, a transcribable nucleic acid molecule can comprise a nucleotide having a sequence of SEQ ID NO: 5 so as to express sucrose phosphate phosphatase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide having at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% percent identity to SEQ ID NO: 5 encoding a

polypeptide having sucrose phosphate phosphatase activity. As another example, a transcribable nucleic acid molecule can comprise a nucleotide sequence encoding a polypeptide having at least about 85%, at least about 90%, at least about 95%, or at least about 99% sequence identity to SEQ ID NO: 6, wherein the expressed sequence exhibits SPP activity.

In some embodiments, a transcribable nucleic acid molecule can comprise one or more of *asf*, *sps*., or *spp*. For example, a transcribable nucleic acid molecule can comprise *asf* and *sps*; *asf* and *spp*; *sps* and *spp*; or *asf*, *sps*, and *spp*.

A transcribable nucleic acid molecule can encode a polypeptide having trehalose phosphate synthase activity or trehalose phosphate phosphatase activity; gluocosylglycerol phosphate synthase activity or gluocosylglycerol phosphate phosphatase activity; or mannosylfructose phosphate synthase activity or mannosylfructose phosphate phosphatase activity.

In some embodiments, a transcribable nucleic acid molecule comprises a trehalose phosphate synthase (*tps*) (see e.g., SEQ ID NO: 76 encoding *tps* gene and SEQ ID NO: 77 encoding TPS polypeptide), or homologue thereof. For example, a transcribable nucleic acid molecule can comprise a nucleotide having a sequence of SEQ ID NO: 76 so as to express trehalose phosphate synthase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide having at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% percent identity to SEQ ID NO: 76 encoding a polypeptide having trehalose phosphate synthase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide sequence encoding a polypeptide having at least about 85%, at least about 90%, at least about 95%, or at least about 99% sequence identity to SEQ ID NO: 77, wherein the expressed sequence exhibits TPS activity.

In some embodiments, a transcribable nucleic acid molecule comprises a trehalose phosphate phosphatase (*tpp*) (see e.g., SEQ ID NO: 78 encoding *tpp* gene and SEQ ID NO: 79 encoding TPP polypeptide), or homologue thereof. For example, a transcribable nucleic acid molecule can comprise a nucleotide having a sequence of SEQ ID NO: 78 so as to express trehalose phosphate phosphatase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide having at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% percent identity to SEQ ID NO: 78 encoding a polypeptide having trehalose phosphate phosphatase activity. As another example, a transcribable nucleic acid molecule can comprise a nucleotide sequence encoding

a polypeptide having at least about 85%, at least about 90%, at least about 95%, or at least about 99% sequence identity to SEQ ID NO: 79, wherein the expressed sequence exhibits TPP activity.

In some embodiments, a transcribable nucleic acid molecule comprises a glucosylglycerolphosphate synthase (*gps*) (see e.g., SEQ ID NO: 80 encoding *gps* gene and SEQ ID NO: 81 encoding GPS polypeptide), or homologue thereof. For example, a transcribable nucleic acid molecule can comprise a nucleotide having a sequence of SEQ ID NO: 80 so as to express glucosylglycerolphosphate synthase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide having at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% percent identity to SEQ ID NO: 80 encoding a polypeptide having glucosylglycerolphosphate synthase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide sequence encoding a polypeptide having at least about 85%, at least about 90%, at least about 95%, or at least about 99% sequence identity to SEQ ID NO: 81, wherein the expressed sequence exhibits GPS activity.

In some embodiments, a transcribable nucleic acid molecule comprises a glucosylglycerolphosphate phosphatase (*gpp*) (see e.g., SEQ ID NO: 82 encoding *gpp* gene and SEQ ID NO: 83 encoding GPP polypeptide), or homologue thereof. For example, a transcribable nucleic acid molecule can comprise a nucleotide having a sequence of SEQ ID NO: 82 so as to express glucosylglycerolphosphate phosphatase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide having at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% percent identity to SEQ ID NO: 82 encoding a polypeptide having glucosylglycerolphosphate phosphatase activity. As another example, a transcribable nucleic acid molecule can comprise a nucleotide sequence encoding a polypeptide having at least about 85%, at least about 90%, at least about 95%, or at least about 99% sequence identity to SEQ ID NO: 83, wherein the expressed sequence exhibits GPP activity.

In some embodiments, a transcribable nucleic acid molecule comprises a mannosylfructose phosphate synthase (*mps*) (see e.g., SEQ ID NO: 84 encoding *mps* gene and SEQ ID NO: 85 encoding MPS polypeptide), or homologue thereof. For example, a transcribable nucleic acid molecule can comprise a nucleotide having a sequence of SEQ ID NO: 84 so as to express mannosylfructose phosphate synthase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide having at least about 80%, at

least about 85%, at least about 90%, at least about 95%, or at least about 99% percent identity to SEQ ID NO: 84 encoding a polypeptide having mannosylfructose phosphate synthase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide sequence encoding a polypeptide having at least about 85%, at least about 90%, at least about 95%, or at least about 99% sequence identity to SEQ ID NO: 85, wherein the expressed sequence exhibits MPS activity.

In some embodiments, a transcribable nucleic acid molecule comprises a mannosylfructose phosphate phosphatase (*mpp*) (see e.g., SEQ ID NO: 86 encoding *mpp* gene and SEQ ID NO: 87 encoding MPP polypeptide), or homologue thereof. For example, a transcribable nucleic acid molecule can comprise a nucleotide having a sequence of SEQ ID NO: 86 so as to express mannosylfructose phosphate phosphatase. As another example, a transcribable nucleic acid molecule can comprise a nucleotide having at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% percent identity to SEQ ID NO: 86 encoding a polypeptide having mannosylfructose phosphate phosphatase activity. As another example, a transcribable nucleic acid molecule can comprise a nucleotide sequence encoding a polypeptide having at least about 85%, at least about 90%, at least about 95%, or at least about 99% sequence identity to SEQ ID NO: 87, wherein the expressed sequence exhibits MPP activity.

Alternatively, a transcribable nucleic acid molecule can effect a host cell or organism phenotype by encoding a RNA molecule that causes the targeted inhibition of expression of an endogenous gene, for example via antisense, inhibitory RNA (RNAi), or co-suppression-mediated mechanisms. The RNA could also be a catalytic RNA molecule (i.e., a ribozyme) engineered to cleave a desired endogenous mRNA product. Thus, any nucleic acid molecule that encodes a protein or mRNA that expresses a phenotype or morphology change of interest may be useful for the practice of the present disclosure.

***HOSTS***

A host organism or a host cell can be transformed with a construct including a transcribable nucleic acid molecule operably linked to an expression regulation system sensitive to one or more media components. For example, a host organism or a host cell can be transformed with a construct including a transcribable nucleic acid molecule operably linked to an expression regulation system sensitive to nitrogen source present in the media.

A construct described herein can be plasmid based or intergrated into the host

genome. For example, a construct described herein (e.g., plasmid pLybDB4, SEQ ID NO: 190) can be present in the host as a plasmid (see e.g., Example 11). As another example, a construct described herein (e.g., plasmid pLybAL98, SEQ ID NO: 265) can be integrated into the genome of the host (e.g., strain LYB511) (see e.g., Example 9, Example 11, Example 12). In some embodiments, integration into the genome of the host can increase inducible expression of the target nucleotide (compare Example 11 and Example 12).

A transformed host organism or a host cell can be analyzed for the presence of a gene of interest and the expression level or profile conferred by the expression system of the present disclosure. Those of skill in the art are aware of the numerous methods available for the analysis of transformed hosts. For example, methods for host analysis include, but are not limited to Southern blots or northern blots, PCR-based approaches, biochemical analyses, phenotypic screening methods, and immunodiagnostic assays.

A host organism can be a eukaryotic or a prokaryotic organism.

A host organism can be a photosynthetic microorganism. A host organism can be, for example, a naturally photosynthetic microorganism, such as a cyanobacterium, or an engineered photosynthetic microorganism, such as an artificially photosynthetic bacterium. Exemplary microorganisms that are either naturally photosynthetic or can be engineered to be photosynthetic include, but are not limited to, bacteria; fungi; archaea; protists; microscopic plants, such as a green algae; and animals such as plankton, planarian, and amoeba. Examples of naturally occurring photosynthetic microorganisms include, but are not limited to, *Spirulina maximum*, *Spirulina platensis*, *Dunaliella salina*, *Botrycoccus braunii*, *Chlorella vulgaris*, *Chlorella pyrenoidosa*, *Serenastrum capricomutum*, *Scenedesmus auadricauda*, *Porphyridium cruentum*, *Scenedesmus acutus*, *Dunaliella* sp., *Scenedesmus obliquus*, *Anabaenopsis*, *Aulosira*, *Cylindrospermum*, *Synechoccus* sp., *Synechocystis* sp., or *Tolypothrix*.

Preferably, the host photosynthetic microorganism is a cyanobacterium. Cyanobacteria, also known as blue-green algae, are a broad range of oxygenic photoautotrophs. The host cyanobacterium can be any photosynthetic microorganism from the phylum Cyanophyta. The host cyanobacterium can have a unicellular or colonial (*e.g.*, filaments, sheets, or balls) morphology. Preferably, the host cyanobacterium is a unicellular cyanobacterium. Examples of cyanobacteria that can be a host organism include, but are not limited to, the genus *Synechocystis*, *Synechococcus*, *Thermosynechococcus*, *Nostoc*, *Prochlorococcu*, *Microcystis*, *Anabaena*, *Spirulina*, and *Gloeobacter*. Preferably the host

cyanobacterium is a *Synechocystis* spp. or *Synechococcus* spp. More preferably, the host cyanobacterium is *Synechococcus elongatus* PCC 7942 (ATCC 33912) or *Synechocystis* spp. PCC 6803 (ATCC 27184).

The host cell or organism can comprise an endogenous nitrogen control system. The genome of the host cell or organism can encode an NtcA polypeptide. The host cell or organism can express an NtcA polypeptide according to a nitrogen control system. Alternatively, a host cell or organism can be engineered to encode an NtcA polypeptide or express an NtcA polypeptide according to a nitrogen source.

As described herein, unregulated expression of a transcribable nucleic acid molecule can decrease growth rate of a host cell or organism or reduce viability of a host cell or organism (see Example 4). Various embodiments of the expression system described herein can provide for nitrogen-sensitive regulation of expression of a transcribable nucleic acid molecule, thereby providing increased growth rates or viability of the transformed host as compared to unregulated expression. Using an expression system described herein, a transformed host cell or organism can have less than about a 20% lower growth rate (e.g., less than about a 15%, 10%, or 5% lower growth rate) than a non-transformed host cell or organism under the same or substantially similar conditions. Using an expression system described herein, a transformed host cell or organism can have more than about a 5% higher growth rate (e.g., more than about a 10%, 15%, or 20% higher growth rate) than a host cell or organism transformed with a non-regulated construct grown under the same or substantially similar conditions. The above discussion applies equally to viability, where viability is substituted for growth rate.

***KITS***

Also provided are kits. Such kits can include an agent or composition described herein and, in certain embodiments, instructions for administration. Such kits can facilitate performance of the methods described herein. When supplied as a kit, the different components of the composition can be packaged in separate containers and admixed immediately before use. Components include, but are not limited to an expression system or expression cassette described herein, or components or sequences thereof. Such packaging of the components separately can, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the composition. The pack may, for example, comprise metal or plastic foil such as a blister pack. Such packaging of the

components separately can also, in certain instances, permit long-term storage without losing activity of the components.

Kits may also include reagents in separate containers such as, for example, sterile water or saline to be added to a lyophilized active component packaged separately. For example, sealed glass ampules may contain a lyophilized component and in a separate ampule, sterile water, sterile saline or sterile each of which has been packaged under a neutral non-reacting gas, such as nitrogen. Ampules may consist of any suitable material, such as glass, organic polymers, such as polycarbonate, polystyrene, ceramic, metal or any other material typically employed to hold reagents. Other examples of suitable containers include bottles that may be fabricated from similar substances as ampules, and envelopes that may consist of foil-lined interiors, such as aluminum or an alloy. Other containers include test tubes, vials, flasks, bottles, syringes, and the like. Containers may have a sterile access port, such as a bottle having a stopper that can be pierced by a hypodermic injection needle. Other containers may have two compartments that are separated by a readily removable membrane that upon removal permits the components to mix. Removable membranes may be glass, plastic, rubber, and the like.

In certain embodiments, kits can be supplied with instructional materials. Instructions may be printed on paper or other substrate, or may be supplied as an electronic-readable medium, such as a floppy disc, mini-CD-ROM, CD-ROM, DVD-ROM, Zip disc, videotape, audio tape, and the like. Detailed instructions may not be physically associated with the kit; instead, a user may be directed to an Internet web site specified by the manufacturer or distributor of the kit.

***MOLECULAR ENGINEERING***

Design, generation, and testing of the variant nucleotides, and their encoded polypeptides, having the above required percent identities and retaining a required activity of the expressed protein is within the skill of the art. For example, directed evolution and rapid isolation of mutants can be according to methods described in references including, but not limited to, Link et al. (2007) Nature Reviews 5(9), 680-688; Sanger et al. (1991) Gene 97(1), 119-123; Ghadessy et al. (2001) Proc Natl Acad Sci USA 98(8) 4552-4557. Thus, one skilled in the art could generate a large number of nucleotide or polypeptide variants having, for example, at least 95-99% identity to the reference sequence described herein and screen such for desired phenotypes according to methods routine in the art. Generally, conservative

substitutions can be made at any position so long as the required activity is retained.

Nucleotide or amino acid sequence identity percent (%) is understood as the percentage of nucleotide or amino acid residues that are identical with nucleotide or amino acid residues in a candidate sequence in comparison to a reference sequence when the two sequences are optimally aligned (with appropriate insertions, deletions, or gaps totaling less than about 20 percent of the reference sequence over the window of comparison). To determine percent identity, sequences are aligned and if necessary, gaps are introduced to achieve the maximum percent sequence identity. Sequence alignment procedures to determine percent identity are well known to those of skill in the art and include tools such as the local homology algorithm of Smith and Waterman, the homology alignment algorithm of Needleman and Wunsch, the search for similarity method of Pearson and Lipman, and preferably by computerized implementations (e.g., GAP, BESTFIT, FASTA, and TFASTA) of these algorithms. Often publicly available computer software such as BLAST, BLAST2, ALIGN2 or Megalign (DNASTAR) software is used to align sequences. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-length of the sequences being compared. When sequences are aligned, the percent sequence identity of a given sequence A to, with, or against a given sequence B (which can alternatively be phrased as a given sequence A that has or comprises a certain percent sequence identity to, with, or against a given sequence B) can be calculated as: percent sequence identity = X/Y100, where X is the number of residues scored as identical matches by the sequence alignment program's or algorithm's alignment of A and B and Y is the total number of residues in B. If the length of sequence A is not equal to the length of sequence B, the percent sequence identity of A to B will not equal the percent sequence identity of B to A.

As used herein, the term "substantial percent sequence identity" refers to a percent sequence identity of at least about 70% sequence identity, at least about 80% sequence identity, at least about 90% sequence identity, or even greater sequence identity, such as about 98% or about 99% sequence identity. Thus, one embodiment is a nucleic acid molecule that has at least about 70% sequence identity, at least about 80% sequence identity, at least about 90% sequence identity, or even greater sequence identity, such as about 98% or about 99% sequence identity with a nucleic acid sequence described herein. Nucleic acid molecules that are capable of regulating transcription of operably linked transcribable nucleic acid molecules and have a substantial percent sequence identity to the nucleic acid sequences of

the expression systems provided herein are encompassed within the scope of this disclosure.

"Highly stringent hybridization conditions" are defined as hybridization at 65 °C in a 6 X SSC buffer (*i.e.,* 0.9 M sodium chloride and 0.09 M sodium citrate). Given these conditions, a determination can be made as to whether a given set of sequences will hybridize by calculating the melting temperature ($T_m$) of a DNA duplex between the two sequences. If a particular duplex has a melting temperature lower than 65°C in the salt conditions of a 6 X SSC, then the two sequences will not hybridize. On the other hand, if the melting temperature is above 65 °C in the same salt conditions, then the sequences will hybridize. In general, the melting temperature for any hybridized DNA:DNA sequence can be determined using the following formula: $T_m = 81.5\ °C + 16.6(\log_{10}[Na^+]) + 0.41(\text{fraction G/C content}) - 0.63(\%\ \text{formamide}) - (600/l)$. Furthermore, the $T_m$ of a DNA:DNA hybrid is decreased by 1-1.5°C for every 1% decrease in nucleotide identity (see e.g., Sambrook and Russel, 2006).

Host cells can be transformed using a variety of standard techniques known to the art (*see, e.g.*, Sambrook and Russel (2006) Condensed Protocols from Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, ISBN-10: 0879697717; Ausubel et al. (2002) Short Protocols in Molecular Biology, 5th ed., Current Protocols, ISBN-10: 0471250929; Sambrook and Russel (2001) Molecular Cloning: A Laboratory Manual, 3d ed., Cold Spring Harbor Laboratory Press, ISBN-10: 0879695773; Elhai, J. and Wolk, C. P. 1988. Methods in Enzymology 167, 747-754). Such techniques include, but are not limited to, viral infection, calcium phosphate transfection, liposome-mediated transfection, microprojectile-mediated delivery, receptor-mediated uptake, cell fusion, electroporation, and the like. The transfected cells can be selected and propagated to provide recombinant host cells that comprise the expression vector stably integrated in the host cell genome.

Host strains developed according to the approaches described herein can be evaluated by a number of means known in the art (*see e.g.*, Studier (2005) Protein Expr Purif. 41(1), 207–234; Gellissen, ed. (2005) Production of Recombinant Proteins: Novel Microbial and Eukaryotic Expression Systems, Wiley-VCH, ISBN-10: 3527310363; Baneyx (2004) Protein Expression Technologies, Taylor & Francis, ISBN-10: 0954523253).

Methods of down-regulation or silencing genes are known in the art. For example, expressed protein activity can be down-regulated or eliminated using antisense oligonucleotides, protein aptamers, nucleotide aptamers, and RNA interference (RNAi) (e.g., small interfering RNAs (siRNA), short hairpin RNA (shRNA), and micro RNAs (miRNA)

(see e.g., Fanning and Symonds (2006) Handbook Exp Pharmacol. 173, 289-303G, describing hammerhead ribozymes and small hairpin RNA; Helene, C., et al. (1992) Ann. N.Y. Acad. Sci. 660, 27-36; Maher (1992) Bioassays 14(12): 807-15, describing targeting deoxyribonucleotide sequences; Lee et al. (2006) Curr Opin Chem Biol. 10, 1-8, describing aptamers; Reynolds et al. (2004) Nature Biotechnology 22(3), 326 – 330, describing RNAi; Pushparaj and Melendez (2006) Clinical and Experimental Pharmacology and Physiology 33(5-6), 504-510, describing RNAi; Dillon et al. (2005) Annual Review of Physiology 67, 147-173, describing RNAi; Dykxhoorn and Lieberman (2005) Annual Review of Medicine 56, 401-423, describing RNAi). RNAi molecules are commercially available from a variety of sources (*e.g.*, Ambion, TX; Sigma Aldrich, MO; Invitrogen). Several siRNA molecule design programs using a variety of algorithms are known to the art (*see e.g.*, Cenix algorithm, Ambion; BLOCK-iT™ RNAi Designer, Invitrogen; siRNA Whitehead Institute Design Tools, Bioinformatics & Research Computing). Traits influential in defining optimal siRNA sequences include G/C content at the termini of the siRNAs, Tm of specific internal domains of the siRNA, siRNA length, position of the target sequence within the CDS (coding region), and nucleotide content of the 3' overhangs.

Definitions and methods described herein are provided to better define the present disclosure and to guide those of ordinary skill in the art in the practice of the present disclosure. Unless otherwise noted, terms are to be understood according to conventional usage by those of ordinary skill in the relevant art.

In some embodiments, numbers expressing quantities of ingredients, properties such as molecular weight, reaction conditions, and so forth, used to describe and claim certain embodiments of the present disclosure are to be understood as being modified in some instances by the term "about." In some embodiments, the term "about" is used to indicate that a value includes the standard deviation of the mean for the device or method being employed to determine the value. In some embodiments, the numerical parameters set forth in the written description and attached claims are approximations that can vary depending upon the desired properties sought to be obtained by a particular embodiment. In some embodiments, the numerical parameters should be construed in light of the number of reported significant digits and by applying ordinary rounding techniques. Notwithstanding that the numerical ranges and parameters setting forth the broad scope of some embodiments of the present disclosure are approximations, the numerical values set forth in the specific examples are reported as precisely as practicable. The numerical values presented in some

embodiments of the present disclosure may contain certain errors necessarily resulting from the standard deviation found in their respective testing measurements. The recitation of ranges of values herein is merely intended to serve as a shorthand method of referring individually to each separate value falling within the range. Unless otherwise indicated herein, each individual value is incorporated into the specification as if it were individually recited herein.

In some embodiments, the terms "a" and "an" and "the" and similar references used in the context of describing a particular embodiment (especially in the context of certain of the following claims) can be construed to cover both the singular and the plural, unless specifically noted otherwise. In some embodiments, the term "or" as used herein, including the claims, is used to mean "and/or" unless explicitly indicated to refer to alternatives only or the alternatives are mutually exclusive.

The terms "comprise," "have" and "include" are open-ended linking verbs. Any forms or tenses of one or more of these verbs, such as "comprises," "comprising," "has," "having," "includes" and "including," are also open-ended. For example, any method that "comprises," "has" or "includes" one or more steps is not limited to possessing only those one or more steps and can also cover other unlisted steps. Similarly, any composition or device that "comprises," "has" or "includes" one or more features is not limited to possessing only those one or more features and can cover other unlisted features.

All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (*e.g.* "such as") provided with respect to certain embodiments herein is intended merely to better illuminate the present disclosure and does not pose a limitation on the scope of the present disclosure otherwise claimed. No language in the specification should be construed as indicating any non-claimed element essential to the practice of the present disclosure.

Groupings of alternative elements or embodiments of the present disclosure disclosed herein are not to be construed as limitations. Each group member can be referred to and claimed individually or in any combination with other members of the group or other elements found herein. One or more members of a group can be included in, or deleted from, a group for reasons of convenience or patentability. When any such inclusion or deletion occurs, the specification is herein deemed to contain the group as modified thus fulfilling the written description of all Markush groups used in the appended claims.

All publications, patents, patent applications, and other references cited in this application are incorporated herein by reference in their entirety for all purposes to the same extent as if each individual publication, patent, patent application or other reference was specifically and individually indicated to be incorporated by reference in its entirety for all purposes. Citation of a reference herein shall not be construed as an admission that such is prior art to the present disclosure.

Having described the present disclosure in detail, it will be apparent that modifications, variations, and equivalent embodiments are possible without departing the scope of the present disclosure defined in the appended claims. Furthermore, it should be appreciated that all examples in the present disclosure are provided as non-limiting examples.

EXAMPLES

The following non-limiting examples are provided to further illustrate the present disclosure. It should be appreciated by those of skill in the art that the techniques disclosed in the examples that follow represent approaches the inventors have found function well in the practice of the present disclosure, and thus can be considered to constitute examples of modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments that are disclosed and still obtain a like or similar result without departing from the spirit and scope of the present disclosure.

***EXAMPLE 1: REMOVAL OF INVERTASE ACTIVITY FROM THE CELL***

To accumulate sucrose, *Synechocystis* sp. PCC 6803 was engineered to remove the invertase (β-fructofuranosidase) activity from the cell since invertase hydrolyzes sucrose to glucose and fructose that are re-assimilated into normal metabolic functions within the cell. The gene encoding invertase, *lim17* (SEQ ID NO: 140), was deleted from ATCC 27184 (LYB426) to yield LYB471. It was also deleted from the wild-type strain (LYB467) to yield LYB472. Over time, differences between these strains have developed such the loss of motility by LYB426 (Kamei et al. 2001).

The invertase-encoding gene *lim17* was first deleted from LYB426 using the neomycin resistance marker to generate LYB443. Using sequential PCR, a deletion in *lim17* (SEQ ID NO: 140) was created to leave large stretches of flanking DNA. The first PCR reaction was performed on LYB426 whole cell genomic DNA template using Sspinvdel-F

(SEQ ID NO: 142) and Sspinvdel-R (SEQ ID NO: 143) as primers. The product of this reaction (Ssp6803 Invertase Deletion PCR 1; SEQ ID NO: 144) was used as a primer in the second reaction (with Sspinvdel-R2 (SEQ ID NO: 145) again using LYB426 whole cell genomic DNA as the template to create Ssp6803 Invertase Deletion PCR 2 (SEQ ID NO: 146). The secondary PCR product was digested with *XbaI* and *SphI* and then ligated into similarly digested pUC19 (SEQ ID NO: 139), creating pLybAL38 (SEQ ID NO: 147). A DNA fragment containing an aminoglycoside 3'-phosphotransferase gene (aph) and the uracil phosphoribosyl transferase gene (*upp*) from Bacillus subtilis 168 (Bs upp aph PCR) (SEQ ID NO: 220) was amplified from plasmid pLybAL8f (SEQ ID NO: 69) with the oligonucleotides bsuppkanmcs–F and bsuppkanPstI-R. This PCR product was digested with *SalI* and *PstI* and then ligated into pLybAL38 (SEQ ID NO: 147) that had been digested with *SbfI* and *SalI*, which are found between the *lim17* (SEQ ID NO: 140) flanking sequences, creating pLybAL41 (SEQ ID NO: 221). Plasmid pLybAL41 was then linearized by restriction endonucleases digestion with *AatII*, which is found within the pUC 19 (SEQ ID NO: 139) backbone, and then transformed into LYB426 by the protocol of Eaton-Rye (2004) with minor modifications. In particular, the transformation mix was plated directly onto the agar plates instead of a membrane overlayed onto the agar plate. Antibiotic was then administered by lifting the agar, adding the antibiotic to the bottom of the Petri dish and then lowering the agar back into the dish. This modification was made solely to reduce the number of agar plates required by three-fold. Integration was selected on BG11-A plates containing 25 μg/ml neomycin. Colonies were analyzed for proper integration and complete segregation throughout all of the copies of the chromosome by amplification of the genomic DNA from whole cells using the oligonucleotides Sspinvdelscreen-F and Sspinvdelscreen-R. Wild type amplification yields a product of 2.8 kbp, whereas the deletion of *lim17* (SEQ ID NO: 140) with the neomycin resistance marker and the B. subtilis *upp* gene yields a 3.7 kbp product.

Then the neomycin resistance marker in LYB443 was swapped with the marker for spectinomycin resistance (the added *upp* gene was also removed) and the invertase was deleted from LYB467. The spectinomycin resistance marker was amplified from plasmid pBSL175 (Alexeyev et al. 1995) with the oligonucleotides specSalI-F and specSbfI-R. The product of this reaction was then digested with *SalI* and *SbfI*, and then ligated into similarly digested pLybAL41 to generate pLybAL58. Double homologous recombination was performed, as described above, with strains LYB443 and LYB467 to yield strains LYB471 and LYB472, respectively. Integration was selected on BG11-A plates with 25 μg/ml spectinomycin. Proper integration and complete segregation of colonies were again

examined by amplification of genomic DNA from whole cells using the oligonucleotides Sspinvdelscreen-F and Sspinvdelscreen-R. Deletion of *lim17* (SEQ ID NO: 140) with the spectinomycin resistance marker yields a 2.6 kbp product.

Deletion of the invertase allowed sucrose to as evidenced by 1) comparison between LYB426 (wild type) and LYB443 run under conditions of salt exposure resulted in no measurable sucrose in the wild type but 3-fold increase over background for LYB443 and 2) LYB443 transformed with the *asf* gene and run under control of functional promoters indicated that sucrose had accumulated within cells and in the culture media.

### *EXAMPLE 2: REDUCTION OF EXOPOLYSACCHARIDE EXPRESSION*

In addition, *Synechocystis* sp. PCC 6803 is known to produce significant quantities of an exopolysaccharide (EPS) in response to normal aging of the culture or under conditions of environmental stress on the cells (Panoff et al. 1988). The large amounts of EPS represent a considerable allocation of carbon and overall metabolic flux that could be shunted to sucrose production. The genetics of exopolysaccharide production in *Synechocystis* sp. PCC 6803 are relatively unknown, although EPS mutants have been selected (Panoff and Joset 1989). A large cluster of genes (almost 18 kbp in size) on the endogenous plasmid pSYSM have been speculated to be largely responsible for EPS production (Kaneko et al. 2003), based upon their homologies to known genes. To this end, the putative genes considered to be responsible for key biosynthetic pathways associated with EPS biosynthesis were removed from LYB472 to create LYB476.

A strategy similar to the deletion of the invertase was employed for the deletion of the presumed EPS locus. Similar to the above, sequential PCR was performed with the primary oligonucleotides epsko-F and epsko-R and then the secondary oligonucleotide pSYSM-R2. The secondary PCR product was digested with *XbaI* and *SphI*, and then ligated into similarly digested pLybAL41 to yield pLybAL61. The erythromycin resistance marker from plasmid pE194 (Horinouchi and Weisblum 1982) was amplified with the oligonucleotides MLSSalI-F and MLSSbfI-R. The resultant product was digested with *SalI* and *SbfI* and then ligated into similarly digested pLybAL61, yielding pLybAL62. Double homologous recombination of LYB472 with linearized pLybAL62 was performed as described above, creating strain LYB476. Colonies were analyzed for proper integration by PCR of whole cell DNA with the oligonucleotides EPSKOscreen-F and EPSKOscreen-R. Deletion results in a 3.2 kbp product, instead of the 18.7 kbp product of the wild type strain. The 3.2 kbp deletion could

easily be observed, but amplifying the 18.7 kbp wild-type product presented difficulties, making amplification unsuitable for determining if there was complete segregation. Complete segregation was instead assessed by the loss of multiple PCR products using the oligonucleotide pairs EPSKOscreen-F/EPSKOint-R1, EPSKOint-F2/ EPSKOint-R2, EPSKOint-F3/EPSKOint-R3, EPSKOint-F4/EPSKOint-R4 and EPSKOint-F1/ EPSKOscreen-R, for which one or both of the oligonucleotides are found within the deleted region. Amplification of wild type DNA with these oligonucleotides yields the products EPS KO Sreen_Border 1, EPS KO int PCR 2, EPS KO int PCR 3, EPS KO int PCR 4 and EPS KO Sreen_Border 2, respectively.

Results showed that deletion of the above EPS genes resulted in a new engineered organism that displayed a >80% loss in EPS production as determined by the extraction and colorimetric assay described by Panoff and Joset, 1989 and Laurentin and Edwards, 2003. Furthermore, it appears that the modified organisms with reduced EPS biosynthesis have no reduction in growth rate or overall viability.

### *EXAMPLE 3: ASF EXPRESSION FROM NITRITE REDUCTASE PROMOTER*

The following example shows ASF expression from the *Synechocystis* sp. PCC 6803 or Synechococcus elongatus PCC 7942 nitrite reductase promoters.

LYB476, the strain of *Synechocystis* sp. bearing the plasmid comprising the nitrite reductase promoter from *Synechococcus elongatus* PCC7942 in front of the *asf* gene (pLybAL18) but with reduced EPS expression, was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25µg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. (Plasmid pLybAL18, and generation thereof, is as described in U.S. App. Pub. No. 2009/0181434, incorporated herein by reference in its entirety. Plasmid pLybAL18 was introduced into LYB476 by triparental conjugation, as also described in U.S. App. Pub. No. 2009/0181434.) The cultivation was performed at 28°C in shake flasks at 250RPM under 100µE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed into 50ml fresh BG11 medium containing 20mM potassium nitrate and 25mg/ml chloramphenicol pH 9.0. The culture was shaken at 250RPM and 28°C for 24 hours under 100mE illumination. The culture media was separated from the cells by centrifugation and the media was assayed for sucrose concentration using an enzyme

based colorimetric method (Biovision Sucrose Assay Kit #K626). Cells were subjected to detergent lysis and the clarified supernatant was assayed for sucrose concentration using the enzyme based method described above.

LYB476, bearing the *asf* gene via the pLybAL18 plasmid, expressed sucrose into the culture media and composed between 30 and 50 percent by weight of the dry biomass of cells in the original culture. The concentration of sucrose liberated from the cells upon lysis corresponded to less than 10% by weight of the total sucrose observed in the culture.

LYB476 was alternatively transformed with a different plasmid carrying a nitrite reductase promoter from *Synechocystis* sp.PCC 6803 in front of the *asf* gene, pLybAL16, and cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25µg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. (Plasmid pLybAL16, and generation thereof, is as described in U.S. App. Pub. No. 2009/0181434, incorporated herein by reference in its entirety. Plasmid pLybAL16 was introduced into LYB476 by triparental conjugation, as described in U.S. App. Pub. No. 2009/0181434.) The cultivation was performed at 28°C in shake flasks at 250RPM under 100µE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed into 50ml fresh BG11 medium containing 20mM potassium nitrate and 25mg/ml chloramphenicol pH 9.0. The culture was shaken at 250RPM and 28°C for 24 hours under 100mE illumination. The culture media was separated from the cells by centrifugation and the media was assayed for sucrose concentration using an enzyme based colorimetric method (Biovision Sucrose Assay Kit #K626). Cells were subjected to detergent lysis and the clarified supernatant was assayed for sucrose concentration using the enzyme based method described above.

Results from LYB476 bearing the plasmid pLybAL16 showed sucrose was observed primarily in the culture media composing between 30 and 50 percent by weight of the dry biomass of cells in the original culture. The concentration of sucrose liberated from the cells upon lysis corresponded to less than 10% by weight of the total sucrose observed in the culture.

***EXAMPLE 4: ASF EXPRESSION FROM THE SYNECHOCYSTIS SP. PCC 6803 OR SYNECHOCOCCUS ELONGATUS PCC 7942 RUBISCO PROMOTERS, AS WELL AS SPS AND SPP CO-EXPRESSION FROM THE SYNECHOCYSTIS SP. PCC 6803 RUBISCO PROMOTER***

The ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCo) promoter from cyanobacteria is thought to be one of the strongest known promoters, since RuBisCo is one of the proteins expressed in highest abundance in the biosphere. Isolation and cloning of this promoter from both *Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* PCC 7942 and placement in front of the sucrose producing asf gene was successfully accomplished and verified. The results suggested that production of sucrose by expressing asf under RuBisCo promoters control leads to growth inhibition, presumably due to the over allocation of resources toward sucrose production. These results further suggest that the regulation sucrose production during periods of biomass buildup is required.

The $\lambda_{PR}$ promoter and CI from plasmid pLybAL19 (as described in U.S. App. Pub. No. 2009/0181434) were replaced by the RuBisCo promoters from *Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* PCC 7942. The *Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* PCC 7942 RuBisCo promoters were amplified from whole cell genomic DNA using the oligonucleotide pairs SsprbcL-F/SsprbcL-R and SelorbcL-F/SelorbcL-R, respectively. The PCR products were digested with XbaI and AflII and ligated into similarly digested pLybAL19 to create pLybAL42 and pLybAL43 for the *Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* PCC 7942 RuBisCo promoters, respectively.

Triparental conjugation (as described in U.S. App. Pub. No. 2009/0181434) was employed to insert plasmids pLybAL42 and pLybAL43 into strain LYB472. But transconjugants would appear and then turn brown and die off upon their restreaking, to purify the plasmid bearing strain away from the E. coli donor and helper strains. It was thought that this could be due to extreme sucrose production.

Also produced was a homologous recombinant sucrose production system where the *Synechocystis* sp. PCC 6803 sps and spp genes was placed behind the *Synechocystis* sp. PCC 6803 RuBisCo promoter in the context of the *Synechocystis* sp. PCC 6803 RuBisCo operon (rbcLXS) (SEQ ID NO: 233), and terminated by the presumed rbcLXS terminator. A C-terminal $His_6$-tagged sps open reading frame started at the start of rbcL and ended at the end of rbcX. A C-terminal $His_6$-tagged spp open reading frame replaced rbcS. To assemble this construct, the artificial operon, except for a 1.3 kbp XmaI/SpeI fragment within sps, with flanking XbaI and PmeI sites was synthesized (Blue Heron, Bothell, WA). This XbaI/PmeI

fragment was placed into plasmid pLybAL50, (a derivative pLybAL19) digested with XbaI and ClaI (blunt-ended with T4 DNA polymerase) creating plasmid pLybAL66. The XmaI/SpeI fragment of sps was amplified by PCR from LYB467 whole cell genomic DNA with the oligonucleotides Ssp6803spsXS-F and Ssp6803spsXS-R. The product was digested with XmaI and SpeI and ligated into similarly digested pLybAL66, creating pLybAL67.

Both pLybAL66 and pLybAL67 were transferred into LYB472 by triparental conjugation. Transconjugants of pLybAL66 (having an incomplete sps) appeared healthy. But transconjugants of pLybAL67 (having a complete sps) behaved the same as those of pLybAL42 and pLybAL43. A few green colonies of LYB472 with pLybAL67 did appear over time. Some of these were grown in BG11-A with 25 μg/ml chloramphenicol. Plasmid DNA was purified with the Wizard Plus SV Miniprep Kit (Promega, Madison, WI), with the addition of 1 minute of treatment with glass beads in a bead beater after resuspension of the pellet. The low quantity of purified plasmid DNA was amplified by transformation into E. coli NEB5alpha (NEB, Ipswich, MA). The DNA was again miniprepped and subjected to restriction analysis, where some isolates did not match the original plasmid DNA. Isolates that looked correct by restriction analysis presumably contained either very small deletions/insertions or point mutations. In either case, further analyses of the sequences of these plasmids were not pursued.

Some pLybAL67 transconjugants, however, were qualitatively analyzed for sucrose production. Briefly, small isolated colonies of LYB472 bearing either plasmid pLybAL66 or pLybAL67 were harvested from an agar plate using a sterile drawn glass whisker tube and transferred to 50 microliters of BG11A medium supplemented with chloramphenicol. This microculture was then positioned in a sterile multiwell tray as a hanging drop above 1 ml of a pool of cell free culture media to in order to prevent evaporative loss of the hanging drop solution. After 7 days the drops were collected, biomass was removed by centrifugation and the supernatant (spent culture broth) was assayed for the presence of sucrose. Strong colorimetric response was observed in cultures growing LYB472 bearing pLybAL67 while cultures of LYB472 with no plasmid or LYB472 pLybAL66 displayed no significant response. The quantitative assay of sucrose in these samples was not possible given the extremely low level of biomass in the samples, however, qualitative analysis suggested that significant sucrose production was observed with a fully functional RuBisCo promoter upstream of the asf gene for LYB472 bearing pLybAL67. Attempts to scale LYB472 bearing pLybAL67 resulted in extremely slow growth that ceased after several days resulting in pale

yellow to whitish cells that appeared to lose viability over time.

### *EXAMPLE 5: ESTERASE EXPRESSION FROM VARIATIONS OF THE NOSTOC SP. PCC 7120 RUBISCO PROMOTER*

The following example shows Esterase expression from variations of the Nostoc sp. PCC 7120 RuBisCo promoter.

The results of pLybAL42, pLybAL43, pLybAL66 and pLybAL67 suggested that the RuBisCo promoter may be able to synthesize enough enzyme for sucrose production, but production needed to be down-regulated during the biomass accumulation phase (see Example 3). In addition, regulation should be inexpensive to economically produce a commodity product. While prior reports have showed NtcA used to control expression (see e.g., Jiang et al. 1997; Herrero et al., 2001; Muro-Pastor et al., 2005), use in *Synechocystis* sp. PCC 6823 had not been reported.

The Ntca system in *Synechocystis* sp. PCC 6823 was first tested on a simple enzyme system, thermophilic carboxylesterase, that could be easily assayed. The thermophilic carboxylesterase gene was amplified by PCR from plasmid pCE020R using the oligonucleotides E020-F and E020-R, which also appended a C-terminal $His_6$-tag to the open reading frame. The PCR product was digested with NdeI and ClaI and ligated into similarly digested pLybAL50, thus creating pLybAL68 where the *asf* gene of LybAL50 was replaced with the esterase gene.

Next, CI and the lambda PR promoter were replaced with various fragments of the Nostoc sp PCC 7120 rbcLXS promoter. Four different constructs were made, containing the entire region, the core plus the downstream region, the core plus the upstream region and the core alone. These fragments were obtained by PCR amplification of Nostoc sp. PCC 7120 purified chromosomal DNA (ATCC # 27893D-5) the oligonucleotide pairs Nsp7120rbcprom-F1/Nsp7120rbcprom-R1, Nsp7120rbcprom-F2/Nsp7120rbcprom-R1, Nsp7120rbcprom-F1/Nsp7120rbcprom-R2 and Nsp7120rbcprom-F2/Nsp7120rbcprom-R2, respectively. The PCR products were digested with SbfI and BamHI, and then ligated into similarly digested pLybAL68 to create pLybAL69, pLybAL70, pLybAL71 and pLybAL72, respectively.

Plasmids were introduced into LYB476 by triparental conjugation (as described in U.S. App. Pub. No. 2009/0181434).

*Synechocystis* sp. (LYB476) bearing the plasmid pLybAL69 (nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the *LYB E020* gene) was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25μg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28°C in shake flasks at 250RPM under 100μE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. The culture was shaken at 250RPM and 28°C for 24 hours under 100mE illumination. The culture media was separated from the cells by centrifugation and the cells were subjected to detergent lysis and the clarified supernatant was assayed for esterase activity using p-nitrophenol acetate colorimetric assay. Briefly, 100μM p-nitrophenylacetate (from a 1mM stock in acetonitrile) was dispersed in 50mM phosphate buffer pH 7. The reaction was initiated by addition of 0.1 μl crude lysate and the formation of p-nitrophenate liberated from the hydrolysis reaction was followed at 348nm (isosbestic point for the nitrophenol/nitrophenylate ion).

A strain of *Synechocystis* sp. (LYB476) bearing the plasmid pLybAL70 (nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the *LYB E020* gene) was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25μg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28°C in shake flasks at 250RPM under 100μE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. The culture was shaken at 250RPM and 28°C for 24 hours under 100mE illumination. The culture media was separated from the cells by centrifugation and the cells were subjected to detergent lysis and the clarified supernatant was assayed for esterase activity using p-nitrophenol acetate colorimetric assay as described above.

A strain of *Synechocystis* sp. (LYB476) bearing the plasmid pLybAL71(nitrogen

regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the *LYB E020* gene)was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25µg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28°C in shake flasks at 250RPM under 100µE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. The culture was shaken at 250RPM and 28°C for 24 hours under 100mE illumination. The culture media was separated from the cells by centrifugation and the cells were subjected to detergent lysis and the clarified supernatant was assayed for esterase activity using p-nitrophenol acetate colorimetric assay as described above.

A strain of *Synechocystis* sp. (LYB476) bearing the plasmid pLybAL72 (nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the *LYB E020* gene)was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25µg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28°C in shake flasks at 250RPM under 100µE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. The culture was shaken at 250RPM and 28°C for 24 hours under 100mE illumination. The culture media was separated from the cells by centrifugation and the cells were subjected to detergent lysis and the clarified supernatant was assayed for esterase activity using p-nitrophenol acetate colorimetric assay as described above.

Results showed that the expression of the carboxyesterase gene corresponds to a thermophilic esterase that was cloned and engineered to contain a 6XHisidine tag on the carboxy- terminus of the polypeptide. The enzyme is extremely stable and shown to be well expressed as a highly soluble active enzyme in a host of recombinant organisms. The colorimetric assay for enzyme activity is an extremely sensitive and quantitative means to determine protein expression yields. To this end, testing of plasmid constructs containing the

*NE020* gene provided a clear response to overall protein expression of the constructs tested. Using initial rates of enzyme turnover correlated to total protein measured in the each system provided a measure of enzyme produced. Furthermore, Western blotting analysis of the lysed biomass using antibodies raised against the 6XHis tag attached to the protein provided insight into the levels of protein expressed in each sample and also provided a means to validate the kinetic assay data.

### *EXAMPLE 6: ASF EXPRESSION FROM VARIATIONS OF THE NOSTOC SP. PCC 7120 RUBISCO PROMOTER*

The following example shows ASF expression from variations of the Nostoc sp. PCC 7120 RuBisCo promoter.

The carboxyesterase (E020) gene from plasmids pLybAL69, pLybAL70, pLybAL71 and pLybAL72 was replaced with the *asf* gene (bearing a C-terminal His6-tag) from plasmid pLybAL50. The *asf* gene in pLybAL50 was removed by digestion with BamHI and ClaI and placed into similarly digested pLybAL69, pLybAL70, pLybAL71 and pLybAL72 to created pLybDB2 (SEQ ID NO: 188), pLybDB3 (SEQ ID NO: 189), pLybDB4 (SEQ ID NO: 190) and pLybDB5 (SEQ ID NO: 191), respectively.

Plasmids were introduced into LYB476 by triparental conjugation (as described in U.S. App. Pub. No. 2009/0181434).

*Synechocystis* sp. (LYB476) bearing the plasmid pLybDB2 (SEQ ID NO: 188) (nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the *asf* gene) was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25μg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28°C in shake flasks at 250RPM under 100μE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The cultivation was performed at 28°C in shake flasks at 250RPM under 100μE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. The culture was shaken at 250RPM and 28°C for 48 hours under 100mE illumination. Cells

were separated from culture medium and sucrose content in the spent culture medium was analyzed using a coupled enzyme bioassay from Biovision. The quantity of sucrose in the spent culture media was compared to similar sucrose measurements performed on retained samples of the ammonical spent culture media employed during the biomass accumulation phase of the cultivation. Control experiments were performed to measure residual sucrose content within cells following final induced cultivation and subsequent lysis of the final cell pellet.

*Synechocystis* sp. (LYB476) bearing the plasmid pLybDB3 (SEQ ID NO: 189) (truncated nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the *asf* gene) was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25µg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28°C in shake flasks at 250RPM under 100µE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The cultivation was performed at 28°C in shake flasks at 250RPM under 100µE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. Cells were separated from culture medium and sucrose content in the spent culture medium was analyzed using a coupled enzyme bioassay from Biovision. The quantity of sucrose in the spent culture media was compared to similar sucrose measurements performed on retained samples of the ammonical spent culture media employed during the biomass accumulation phase of the cultivation. Control experiments were performed to measure residual sucrose content within cells following final induced cultivation and subsequent lysis of the final cell pellet.

*Synechocystis* sp. (LYB476) bearing the plasmid pLybDB4 (SEQ ID NO: 190) (truncated nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the *asf* gene) was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25µg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28°C in shake flasks at 250RPM under 100µE illumination with white LEDs. Following four days

of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. The culture was shaken at 250RPM and 28°C for 48 hours under 100mE illumination. Cells were separated from culture medium and sucrose content in the spent culture medium was analyzed using a coupled enzyme bioassay from Biovision. The quantity of sucrose in the spent culture media was compared to similar sucrose measurements performed on retained samples of the ammonical spent culture media employed during the biomass accumulation phase of the cultivation. Control experiments were performed to measure residual sucrose content within cells following final induced cultivation and subsequent lysis of the final cell pellet.

*Synechocystis* sp. (LYB476) bearing the plasmid pLybDB5 (SEQ ID NO: 191) (truncated nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the *asf* gene) was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25µg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28°C in shake flasks at 250RPM under 100µE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. The culture was shaken at 250RPM and 28°C for 48 hours under 100mE illumination. Cells were separated from culture medium and sucrose content in the spent culture medium was analyzed using a coupled enzyme bioassay from Biovision. The quantity of sucrose in the spent culture media was compared to similar sucrose measurements performed on retained samples of the ammonical spent culture media employed during the biomass accumulation phase of the cultivation. Control experiments were performed to measure residual sucrose content within cells following final induced cultivation and subsequent lysis of the final cell pellet.

Structure of the RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 was according to Table 1.

TABLE 1: Regions of Nostoc sp. PCC 7120 promoter (see FIG. 3)

| Element | Location |
|---|---|
| Region Depicted | -327 to +509 |
| SbfI Site | -327 to -320 |
| BamHI Site | +499 to +504 |
| Upstream Deletion | -319 to -162 |
| Downstream Deletion | +26 to +491 |
| Transcription Start | +1 |
| -10 Consensus | -11 to -6 |
| Promoter | -34 to -6 |
| Translation Start (rbcL) | +507 |
| Translation Start (all1523) | -310 |
| NtcA Consensus Site 1 | -5 to +9 |
| NtcA Binding Site 1 | -10 to +12 |
| NtcA Consensus Site 2 | -54 to -42 |
| NtcA Binding Site 2 | -54 to -42 |
| Factor 2 Binding Site | -77 to -15 |

The region depicted in FIG. 3 starts from the point where the *Sbf*I site was introduced, through the first codon of *rbcL* (RuBisCo large subunit). The DNA sequence is taken from the published genome (Kaneko et al. 2001 DNA Res 8, 205-213). Numbering is relative to the start of transcription (+1). Lower case letters indicate mutations made to introduce the *Sbf*I and *Bam*HI restriction endonucleases sites. The sequence identified as the promoter covers the area that would be normally identified as a typical *E. coli*-type $\sigma^{70}$ promoter. A -10 consensus sequence (TATAAT) can be identified relative to the start of transcription. DNA footprint analysis was used to show that NtcA binds the DNA at two locations (see Kaneko et al. 2001 DNA Res 8, 205-213). The NtcA Consensus Site 1 highlighted in FIG. 3 is $GTN_{10}AC$. The NtcA Consensus Site 2 highlighted in FIG. 3 is $GTN_9AC$. Both are flanked by As and Ts. DNA footprint analysis suggests that a second protein (Factor 2) binds

to the promoter region (see Ramasubramanian et al. 1994 J Bacteriol 176, 1214-1223).

Sucrose measurements were performed on cell free spent culture media and normalized to total dry biomass and corrected for background glucose in each sample.

Results showed that the promoter construct composed of the entire segment region surrounding the nitrogen regulated RuBisCo promoter afforded little to no detectable sucrose as did the construct missing both flanking regions of sequence. The constructs with either the leading or trailing regions of sequence deleted individually demonstrated significant sucrose production (between 0.15 and 0.9 grams of sugar/gram of dry biomass.) The sucrose yields were shown to be dependent on the nitrogen source in the culture media with ammonia containing cultures affording 3-5 fold increase in sugar yields compared to nitrate composed culture solutions. Western Blot analysis performed to detect the presence of the *asf* (containing a 6XHistidine tag) protein demonstrated measurable quantities of protein for the constructs with single leading or trailing sequence regions truncated which is consistent with the sucrose production findings. Control experiments in which ammonia or nitrate containing culture media demonstrated that the organisms tolerated nitrogen source switches without noticeable changes in the growth properties of the organism.

### *EXAMPLE 7: SPS/SPP TWO-GENE OPERON EXPRESSION FROM VARIATIONS OF THE NOSTOC SP. PCC 7120 RUBISCO PROMOTER*

The carboxyesterase (E020) gene from plasmids pLybAL69, pLybAL70, pLybAL71 and pLybAL72 was replaced with the sps and spp genes (each bearing a C-terminal His6-tag) in their operon structure from plasmid pLybAL67. A fragment bearing the sps and spp genes was amplified by PCR from the pLybAL67 template with the oligonucleotides SPSSPP Forward #2 and SPSSPP Reverse. The PCR product was digested with BglII and NarI and placed into plasmids pLybAL69, pLybAL70 and pLybAL72 that had been digested with BamHI and ClaI to create plasmids pLybDB6 (SEQ ID NO: 235), pLybDB7 (SEQ ID NO: 236) and pLybDB9 (SEQ ID NO: 237), respectively. The sps/spp equivalent of plasmids pLybAL71 (E020) and pLybDB4 (SEQ ID NO: 190) (asf) which was to be named pLybDB8 was abandoned due to difficulties encountered during its construction.

Plasmids were introduced into LYB476 by triparental conjugation (as described in U.S. App. Pub. No. 2009/0181434).

*Synechocystis* sp. (LYB476) bearing the plasmid pLybDB6 (SEQ ID NO: 235) (nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the

*Synechocystis* sp. PCC 6803 *sps/spp* genes) was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25µg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28°C in shake flasks at 250RPM under 100µ E illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. The culture was shaken at 250RPM and 28°C for 48 hours under 100mE illumination. Cells were separated from culture medium and sucrose content in the spent culture medium was analyzed using a coupled enzyme bioassay from Biovision. The quantity of sucrose in the spent culture media was compared to similar sucrose measurements performed on retained samples of the ammonical spent culture media employed during the biomass accumulation phase of the cultivation. Control experiments were performed to measure residual sucrose content within cells following final induced cultivation and subsequent lysis of the final cell pellet.

*Synechocystis* sp. (LYB476) bearing the plasmid pLybDB7 (SEQ ID NO: 236) (truncated nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the *Synechocystis* sp. PCC 6803 *sps/spp* genes) was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25µg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28°C in shake flasks at 250RPM under 100µE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. The culture was shaken at 250RPM and 28°C for 48 hours under 100mE illumination. Cells were separated from culture medium and sucrose content in the spent culture medium was analyzed using a coupled enzyme bioassay from Biovision. The quantity of sucrose in the spent culture media was compared to similar sucrose measurements performed on retained samples of the ammonical spent culture media employed during the biomass accumulation phase of the cultivation. Control experiments were performed to measure residual sucrose content within cells following final induced cultivation and subsequent lysis of the final cell pellet.

*Synechocystis* sp. (LYB476) bearing the plasmid pLybDB9 (SEQ ID NO: 237) (truncated nitrogen regulated RuBisCo promoter sequence from *Nostoc* sp. PCC 7120 in front of the *Synechocystis* sp. PCC 6803 *sps/spp* genes) was cultured in 50ml of BG11 medium adjusted to pH 8.0 with 5mM HEPES buffer supplemented with 25μg/ml chloramphenicol and 4mM ammonium chloride substituted for potassium nitrate as a nitrogen source. The cultivation was performed at 28˚C in shake flasks at 250RPM under 100μE illumination with white LEDs. Following four days of fermentation the cells were harvested aseptically by centrifugation, washed once with 1x volume of sterile deionized water. The pelleted cells were redispersed in minimal sterile water and equally split and inoculated into 25ml fresh BG11 medium containing either 4mM ammonium chloride or 20mM potassium nitrate and 25mg/ml chloramphenicol at pH 8.0 or pH 9.0 respectively. The culture was shaken at 250RPM and 28˚C for 48 hours under 100mE illumination. Cells were separated from culture medium and sucrose content in the spent culture medium was analyzed using a coupled enzyme bioassay from Biovision. The quantity of sucrose in the spent culture media was compared to similar sucrose measurements performed on retained samples of the ammonical spent culture media employed during the biomass accumulation phase of the cultivation. Control experiments were performed to measure residual sucrose content within cells following final induced cultivation and subsequent lysis of the final cell pellet.

Sucrose measurements were performed on cell free spent culture media and normalized to total dry biomass and corrected for background glucose in each sample.

In accord with the results observed with the *asf* constructs, the promoter construct composed of the entire segment region surrounding the nitrogen regulated RuBisCo promoter upstream of the sps/spp genes afforded little detectable sucrose. The construct with both flanking regions of sequence deleted demonstrated modest sucrose yields, which is in variance to the results obtained with the *asf* constructs.

The construct with the leading segment of sequence deleted will be generated; however, the trailing deleted region of sequence significant sucrose production (between 0.15 and 0.3 grams of sugar/gram of dry biomass). The sucrose yields were shown to be insensitive to nitrogen source with both nitrate and ammonical nitrogen sources providing similar levels of sugar production. The yields of sucrose are significantly reduced in the sps/spp constructs relative as well as the overall regulatory control effectively realized.

***EXAMPLE 8: PLASMID CONSTRUCTION***

The low copy vector pSMARTGC LK (Lucigen, Middleton, WI) (SEQ ID NO: 192), which is described by the manufacturer to be approximately 20 copies per cell, was used in the initial plasmids as the backbone for construction in *Escherichia coli* NEB5α (NEB, Ipswich, MA). Subsequent to that, pSMARTGC LK was replaced with pLG338 (SEQ ID NO: 202) (6 to 8 copies per cell). All PCR amplification were done using the Phusion polymerase as described by the manufacturer (NEB, Ipswich, MA).

The *upp* locus of *Synechocystis* sp. PCC 6803 was amplified by PCR from genomic DNA with the oligonucleotides Sspuppins-F (SEQ ID NO: 193) and Sspuppins-R (SEQ ID NO: 194) and cloned into the pSMARTGC LK vector as described by the manufacturer, creating pLybAL73f (SEQ ID NO: 195). The gene was divided and a multiple cloning site added by digesting pLybAL73f with *Fsp*I and *Kpn*I, then inserting the phosphorylated annealed oligonucleotides SspuppMCS-F (SEQ ID NO: 196) and SspuppMCS-R (SEQ ID NO: 197), yielding plasmid pLybAL74f (SEQ ID NO: 198). Plasmid pLybAL75f (SEQ ID NO: 201) was then made from pLybAL74f by digesting pLybAL74f with *Fsp*I and *Sph*I and ligating in the phosphorylated annealed oligonucleotides Selo7942rbcTerm-F (SEQ ID NO: 199) and Selo7942rbcTerm–R (SEQ ID NO: 200). This placed the *Synechococcus elongatus* PCC 7942 rubisco operon transcription terminator in front of the multiple cloning site to prevent expression of *asf* (SEQ I DNO: 1) from read-through transcripts produced by the *upp* promoter. Efforts were then made to clone *asf* along with the respective promoters from pLybDB3 (SEQ ID NO: 189) and pLybDB4 (SEQ ID NO: 190) into pLybAL75f so that their respective products could be linearized and then transformed into LYB476 for integration into the chromosome, but the correct products were never obtained.

An even lower copy number vector was then chosen, pLG338 (SEQ ID NO: 202) (6-8 copies per cell). Plasmid pLybAL78A (SEQ ID NO: 203) was constructed from pLG338 by partial digestion with *Sph*I, followed by digestion with *Eco*RV, treatment with T4 polymerase to blunt the *Sph*I site, and then religation of the 5.4 kbp vector. Plasmid pLybAL78A was then digested with *Psh*AI and *Nhe*I, and then the *Hin*dIII (treated with T4 polymerase) – *Xba*I fragment bearing the *Synechocystis* sp. PCC 6803 *upp* locus with the intervening transcription terminator and multiple cloning site from pLybAL75f (SEQ ID NO: 201) was inserted, creating pLybAL79 (SEQ ID NO: 204). The *Sph*I – *Cla*I (treated with T4 polymerase) fragments bearing *asf* along with the respective promoters from pLybDB3 (SEQ ID NO: 189) and pLybDB4 (SEQ ID NO: 190) were then successfully inserted into

pLybAL79 digested with *Sph*I and *Not*I (treated with T4 polymerase) to create pLybAL80 (SEQ ID NO: 205) and pLybAL81 (SEQ ID NO: 206), respectively.

To allow for selection by co-integration of an antibiotic resistance marker, plasmids pLybEA8 (SEQ ID NO: 238), pLybEA9 (SEQ ID NO: 239) and pLybEA10 (SEQ ID NO: 240) were made from pLybAL79, pLybAL80 and pLybAL81, respectively, by insertion of the chloramphenicol acetyltransferase gene from pKD32 (SEQ ID NO: 241) amplified using the oligonucleotides asfcmlint-F (SEQ ID NO: 242) and asfcmlint-R (SEQ ID NO: 243) and digested with *Sac*II, into the *Sac*II site. This yielded plasmids in which the *asf* gene and the downstream chloramphenicol acetyltransferase gene are convergently transcribed, without an intervening transcription terminator. Plasmids containing the chloramphenicol resistance marker in the opposite orientation of that in plasmids pLybEA8, pLybEA9 and pLybEA10 were contructed by their digestion with *Sac*II and religation to create plasmids pLybAL84 (SEQ ID NO: 244), pLybAL85 (SEQ ID NO: 245) and pLybAL86 (SEQ ID NO: 246), respectively.

Vector construction for integration at the invertase locus began with plasmid pLybAL87b (SEQ ID NO: 247). Plasmid pLybAL87b was made from pLybEA8 by inserting a PCR product bearing the invertase locus with an intervening multiple cloning site. This PCR product was made by the successive PCR of *Synechocystis* sp. PCC 6803 wild-type chromosomal DNA with the oligonucleotides Sspinvint-F (SEQ ID NO: 248) and Sspinvint-R (SEQ ID NO: 249), followed by a secondary amplification using the product of the first reaction and the oligonucleotide Sspinvint-R2 (SEQ ID NO: 250) as primers. This final PCR product was digested with *Eco*RI and the 1.4 kbp product was inserted into the 4.9 kbp *Eco*RI fragment of pLybEA8, thus replacing the *upp* locus along with its intervening multiple cloning site and chloramphenicol resistance marker. Again, the *Synechococcus elongatus* PCC 7942 rubisco operon transcription terminator was placed upstream of the multiple cloning site to prevent read-through transcription of the *asf* gene by upstream promoters. Plasmid pLybAL87b was digested with *Not*I and *Sph*I and the phosphorylated annealed oligonucleotides Selo7942rbcTerm-F2 (SEQ ID NO: 251) and Selo7942rbcTerm-R2 (SEQ ID NO: 252) were inserted, creating pLybAL88b (SEQ ID NO: 253). To prevent unnecessary additional expression of the chloramphenicol resistance marker from promoters placed in front of the *asf* gene in the final constructs, a transcription terminator was also placed between the *asf* gene and the cloramphenicol resistance marker in plasmid pLybAL85 (SEQ ID NO: 245). Plasmid pLybAL85 was partially digested with both *Sac*II and *Bam*HI,

and then the phosphorlyated annealed oligonucleotides Selo7942AIITerm-F (SEQ ID NO: 254) and Selo7942AIITerm-R (SEQ ID NO: 255) containing the *Synechococcus elongatus* PCC 7942 *psbAII* transcription terminator were inserted into the 9.5 kbp vector, creating pLybAL89 (SEQ ID NO: 256). Plasmids pLybAL88b and pLybAL89 were combined to make pLybAL90 (SEQ ID NO: 257). The 3.9 kbp *Sph*I – *Kpn*I fragment from pLybAL89 was ligated to the 6.3 kbp *Sph*I – *Kpn*I fragment from pLybAL88b. The 0.37 kbp *Sph*I – *Bam*HI fragment from plasmid pLybAL86 was then combined with the 9.6 kbp *Sph*I – *Bam*HI fragment from plasmid pLybAL90 to create pLybAL91 (SEQ ID NO: 258). Plasmid pLybAL91 contains the invertase locus split by a transcription terminator, the *Nostoc* sp PCC7120 rubisco promoter fragment from pLybDB4, *asf*, another transcription terminator and then the chloramphenicol resistance marker with its own promoter transcribed in the same direction as *asf*.

The invertase locus of plasmid pLybAL88b was swapped with the exopolysaccharide locus in two steps. First, *Synechocystis* sp. PCC 6803 wild-type chromosomal DNA was amplified with the oligonucleotides SspEPSint-F (SEQ ID NO: 259) and SspEPSint-R (SEQ ID NO: 260). The 0.67 kbp product of this reaction was digested with *Eco*RI and *Not*I and ligated to the similarly digested 5.7 kbp fragment of pLybAL88b, forming pLybAL93 (SEQ ID NO: 261). Second, *Synechocystis* sp. PCC 6803 wild-type chromosomal DNA was amplified with the oligonucleotides SspEPSint-F2 (SEQ ID NO: 262) and SspEPSint-R2 (SEQ ID NO: 263). The 1.1 kbp product of this reaction was digested with *Eco*RI and *Kpn*I and ligated to the similarly digested 5.6 kbp fragment of pLybAL93, forming pLybAL94 (SEQ ID NO: 264). Plasmid pLybAL98 (SEQ I DNO: 265) was then created by ligating the 6.7 kbp *Sph*I – *Kpn*I fragment of pLybAL94 with the 3.6 kbp fragment of similarly digested pLybAL91. Plasmid pLybAL98 contains the exopolysaccharide locus split by a transcription terminator, the *Nostoc* sp PCC 7120 rubisco promoter fragment from pLybDB4, *asf*, another transcription terminator and then the chloramphenicol resistance marker with its own promoter transcribed in the same direction as *asf*.

The exopolysaccharide locus of plasmid pLybAL94 was swapped with the *sps* locus of *Synechocystis* sp. PCC 6803 in two steps. First, *Synechocystis* sp. PCC 6803 wild-type chromosomal DNA was amplified with the oligonucleotides Sspspsint-F (SEQ ID NO: 266) and Sspspsint-R (SEQ ID NO: 267). The 1.3 kbp product of this reaction was digested with *Rsr*II and *Kpn*I and ligated to the similarly digested 5.6 kbp fragment of pLybAL94, forming pLybAL95 (SEQ ID NO: 268). Second, *Synechocystis* sp. PCC 6803 wild-type chromosomal

DNA was amplified with the oligonucleotides Sspspsint-F2 (SEQ ID NO: 269) and Sspspsint-R2 (SEQ ID NO: 270). The 0.92 kbp product of this reaction was digested with *Mre*I and *Not*I and ligated to the similarly digested 6.3 kbp fragment of pLybAL95, forming pLybAL96 (SEQ ID NO: 271).

Plasmid pLybAL106 (SEQ ID NO: 272) and pLybAL107 (SEQ ID NO: 273) were constructed to integrate nitrate inducible nitrogen regulon regulatory proteins and invertase genes, respectively, along with the kanamycin resistance marker into the *sps* locus of *Synechocystis* sp. PCC 6803. First, the kanamycin resistance marker of pLybAL90 was deleted by digestion with MluI and relegation of the 8.7 kbp fragment, creating pLybAL100 (SEQ ID NO: 274). The chloramphenicol resistance marker of pLybAL100 was then replaced with the kanamycin resistance marker from pKD13 (SEQ ID NO: 275). The kanamycin resistance marker from pKD13 was amplified with the oligonucleotides neoflpint-F (SEQ ID NO: 276) and asfcmlint-R (a) (SEQ ID NO: 277). The 1.3 kbp product was digested with *Pac*I and *Sac*II and ligated to the 7.7 kbp fragment of similarly digested pLybAL100, creating pLybAL101 (SEQ ID NO: 278). Fragments containing the *nirA* promoter (SEQ ID NO: 32) from *Synechococcus elongatus* PCC 7942 followed by either the *ntcA* and *ntcB* (PnirA_ntcA_ntcB) (SEQ ID NO: 279) or invertase (PnirA_lim17) (SEQ ID NO: 280) genes from *Synechococcus elongatus* PCC 7942 were synthesized by Blueheron (Bothell, WA). The genes from *Synechococcus elongatus* PCC 7942 were used instead of those from *Synechocystis* sp. PCC 6803 to limit erroneous chromosomal recombination. These fragments (1.9 and 1.7 kbp) were digested with *Sbf*I and *Pme*I and placed into the 6.2 kbp fragment of pLybAL101 to create pLybAL102 (SEQ ID NO: 281) and pLybAL103 (SEQ ID NO: 282), respectively. The nitrate inducible genes and kanamycin resistance markers from pLybAL102 and pLybAL103 were combined with the pLybAL96, for integration into the *sps* locus. The 3.2 and 3.1 kbp *Sbf*I – *Sac*II fragments from pLybAL102 and pLybAL103 were ligated to the 7.2 kbp fragment of similarly digested pLybAL96 to yield pLybAL106 (SEQ ID NO: 272) and pLybAL107 (SEQ ID NO: 273), respectively.

### *EXAMPLE 9: INTEGRATION*

For integration into strains of *Synechocystis* sp. PCC 6803, cells were grown in BG11-A (buffered with either HEPES at pH 8.0 or CHES at pH 9.0) to an $OD_{730}$ of approximately 0.4, pelleted by centrifugation and then resuspended in fresh medium to an $OD_{730}$ of 2.5. Better growth is observed when buffered with CHES at pH 9.0, but when the

medium is supplemented with base sensitive compounds like neomycin and 5-fluoruracil, HEPES at pH 8.0 was used. Transformation was performed by mixing 2-10 µg of DNA (linearized with AflII and MluI) with 500 µl of this cell suspension, incubated under light in a clear plastic tube for 3hrs, mixed and then incubated for another 3hrs. The cell suspension was then applied to BG11-A agar plates containing 0.3% sodium thiosulfate. The plates were incubated overnight under constant illumination. After 24hrs from the time the cells were plated, the agar on the plates was lifted and 50% of the antibiotic applied underneath the agar. The other 50% of the antibiotic was applied after 36 hrs. The final concentrations of chloramphenicol, neomycin and 5-fluorouracil were 25, 25 and 1 µg/ml, respectively. Candidates would be restreaked onto antibiotic containing plates, and then screened by colony PCR. This process would be repeated until complete segregation was observed. Upon complete segregation, the candidates would again be restreaked, this time in the absence of antibiotic. Candidates would again be screened by colony PCR to determine that complete segregation was maintained.

Integration was analyzed by PCR with oligonucleotides outside of the region of recombination. Integration at the *upp* locus was screened with the oligonucleotides Sspuppintscrn-F (SEQ ID NO: 283) and Sspuppintscrn-R (SEQ ID NO: 284). Wild-type *upp*, as found in strain LYB476, would yield a 0.89 kbp product. Proper integration with pLybAL81 (SEQ D NO: 206), pLybEA8 (SEQ ID NO: 238), pLybEA10 (SEQ ID NO: 240), pLybAL84 (SEQ ID NO: 244) and pLybAL86 (SEQ ID NO: 246) would yield products of 3.4, 2.0, 4.5, 2.0 and 4.5 kbp, respectively. Integration at the invertase (*lim17*) locus was screened with the oligonucleotides Sspinvdelscreen-F (SEQ ID NO: 153) and Sspinvdelscreen-R (SEQ I DNO: 154). Amplification of the invertase locus from strain LYB476 would yield a 2.6 kbp product. Proper integration with pLybAL91 (SEQ I DNO: 258) would yield a product of 5.3 kbp. Integration at the exopolysaccharide locus was screened with the oligonucleotides EPSKOscreen-F (SEQ I DNO: 167) and EPSKOscreen-R (SEQ ID NO: 168). Amplification of the exopolysaccharide locus from strain LYB476 would yield a 3.2 kbp product. Proper integration with pLybAL98 (SEQ ID NO: 265) would yield a product of 5.6 kbp. Integration at the *sps* locus was screened with the oligonucleotides Ssp6803spsscreen-F (SEQ ID NO: 285) and Ssp6803spsscreen-R (SEQ ID NO: 286). Wild-type *sps*, as found in strain LYB476, would yield a 4.0 kbp product. Proper integration with plasmids pLybAL106 (SEQ ID NO: 272) and pLybAL107 (SEQ ID NO: 273) would yield products of 5.6 kbp and 5.4 kbp, respectively.

Plasmids pLybAL106 and pLybAL107 were linearized and transformed into strain LYB476 to yield LYB509 and LYB510, respectively. Plasmid pLybAL98 was linearized and transformed into LYB509 and LYB510 to yield LYB511 and LYB512, respectively.

Integration of *asf* at the *upp* locus.

Attempts were made to integrate *asf* into the *Synechocystis* sp. PCC 6803 *upp* locus using pLybAL81, but were not successful. Plasmid pLybAL81 (SEQ D NO: 206) was linearized with *Afl*II and *Mlu*I, transformed into LYB476, and then transformants were selected on BG11-A plates containing 1 μg/ml 5-fluorouracil. Candidates were screened by PCR amplification from their genomic DNA using the oligonucleotides Sspuppintscrn-F (SEQ ID NO: 283) and Sspuppintscrn-R (SEQ ID NO: 284). Proper integration would yield a PCR product of 3.4 kbp, instead of the wild-type 0.89 kbp. PCR products from the amplification of genomic DNA from the 5-fluorouracil resistant candidates obtained were the size of the wild-type DNA. These candidates could result from the spontaneous selection of 5-fluorouracil resistant mutants at either the *upp* locus or another locus, instead of *asf* integration. They were not characterized, however.

It was decided to select for the integration of *asf* at the *upp* locus by co-integration of an antibiotic resistance marker using plasmids pLybEA8 (SEQ ID NO: 238) and pLybEA10 (SEQ ID NO: 240). Plasmid pLybEA8 contains only the chloramphenicol resistance marker, whereas pLybEA10 also contains the *asf* gene transcribed from the *Nostoc* PCC 7120 rubisco promoter fragment found in plasmid pLybDB4 (SEQ ID NO: 190). Colonies were only obtained from transformation of LYB476 with linearized pLybEA8, the plasmid devoid of the *asf* gene. The *upp* locus was analyzed by PCR using the primers sspuppintscrn-F (SEQ ID NO: 283) and sspuppintscrn-R (SEQ ID NO: 284). Proper integration should produce a 2.0 kbp DNA fragment, instead of the wild-type 0.89 kbp. Both wild-type and integrant fragments were detected, indicating incomplete segregation. Repeated streaking with antibiotic selection failed to produce complete segregation. No colonies were obtained for the pLybEA10 transformations, the *asf* containing plasmid.

To ensure proper expression of the chloramphenicol resistance marker was not a problem for obtaining integration of *asf*, the orientation of the marker of plasmids pLybEA8 and pLybEA10 was reversed in pLybAL84 (SEQ ID NO: 244) and pLybAL86 (SEQ ID NO: 246), respectively. It was speculated that strong expression of *asf* might result in diminished expression of the downstream convergent *cat* gene. However, similar results were obtained.

Integration of *asf* at the invertase and exopolysaccharide loci.

Integration at the *upp* locus was abandoned in favor of integration at the invertase and exopolysaccharide loci, for which it was known from previous work that the complete segregation of deletions of these genes was not problematic. LYB476 was transformed with the linearized pLybAL91 (SEQ ID NO: 258) and pLybAL98 (SEQ ID NO: 265) to integrate the *asf* gene into the invertase (*lim17*) locus on the chromosome and the exopolysaccharide locus located on the large plasmid pSYSM. Integration procedures afforded very small colonies which would begin to appear then, however, quickly die.

Integration of asf at the invertase and exopolysaccharide loci in modified hosts.

The lack of successful integration was suspected to be the result of sucrose production by ASF, which is expressed even when grown in nitrate-containing medium. Previous results with plasmid pLybDB4 (SEQ ID NO: 190) showed that expression was not completely shut down under non-inducing conditions. This led to efforts to reduce sucrose toxicity (which probably results from severe osmotic imbalance) by either reducing the basal level of *asf* expression (addition of excess NtcA and NtcB) or degradation of the sucrose causing the toxicity (addition of invertase activity). The genes encoding these proteins were placed behind the *Synechococcus elongatus* PCC 7942 *nirA* promoter, which is induced during growth on nitrate, but repressed during growth on ammonia. Plasmid pLybAL106 (SEQ ID NO: 272) and pLybAL107 (SEQ ID NO: 273) were used to integrate these constructs at the sps locus, leading to strains LYB509 and LYB510. These strains were constructed without any difficulty.

Integration of *asf* at the invertase locus of strains LYB509 and LYB510 with plasmid pLybAL91 (SEQ ID NO: 258) was still not successful. However, integration of *asf* at the exopolysaccharide locus in these strains with plasmid pLybAL98 (SEQ ID NO: 265) was, yielding LYB511 and LYB512, respectively.

In accordance with previous data from LYB476 bearing plasmid pLybDB4 (SEQ ID NO: 190), which showed increased sucrose production in response to urea induction. Urea induced cells secreted approximately 3-fold more sucrose than the cells grown in nitrate. Our results show that the LYB511 *asf* integrant also secretes sucrose in response to urea induction. LYB511 displayed a 4-5 fold increase in sucrose secretion in response to urea over nitrate. Interestingly, LYB511 treated with urea secreted 4 fold more sucrose than LYB476 bearing pLybDB4 treated with urea. In addition, LYB511 treated with nitrate also secreted 3-fold more sucrose than LYB476 bearing pLybDB4 treated with nitrate. These results indicate that the *asf* integrant produces more sucrose overall compared to expression

of *asf* exogenous from the plasmid. The amount of sucrose obtained from the cell lysates was similar between all samples. These results indicate that the cells can only tolerate a finite concentration of internal sucrose prior to secretion. The sucrose assay results also infer stable endogenous expression of the *asf* gene.

The LYB512 strain does not show increased sucrose production in response to urea. The nitrogen regulatory control of the additional invertase expression must either be leaky or the invertase is very stable over time.

***EXAMPLE 10: INTEGRATED NITROGEN REGULATED SUCROSE PRODUCTION IN A LIQUID STATE PHOTOBIOREACTOR***

Strain LYB511 was cultivated on BG11 agar culture media and single colonies were transferred to 50 ml of liquid BG11 culture media and shaken at 250 RPM at 30°C under 50 microeinsteins white light until the culture is determined to enter log phase growth as determined by optical density measuremebnts at 730 nm (nominally 3-4 days). The log phase culture is aseptically transferred into 500 ml of stirred BG11 culture broth maintained at 30°C, aerated with 1 volume of filtered air per minute under 150 microeinsteins white fluorescent light. Once the culture had achieved a mid log phase growth the cells were aseptically harvested by centrifugation and washed once with deionized water. The cell pellet was dispersed into 10 ml of deionized water and 5 ml was introduced into 500 ml of BG11 culture broth supplemented with 20mM sodium nitrate or 10 mM urea. Cultures were grown at 30°C with stirring and aeration at 1 volume air per minute under 150 microeinsteins of white fluorescent light. 10 ml of culture broth were removed daily over the course of 4 days and pH, biomass density and sucrose concentration were measured.

Results showed that, for the culture including sodium nitrate as the nitrogen source, biomass accumulation during the 4 day trial increased 2.25 fold with a peak sucrose concentration of 9.3 micromolar. For the culture including urea as the nitrogen source the biomass accumulation during the 4 day trial increased only 1.2 fold and the peak sucrose concentration was 83 micromolar. In both cultures the pH slowly increased from 7 to 7.8.

***EXAMPLE 11: PLASMID BASED NITROGEN REGULATED SUCROSE PRODUCTION IN SOLID STATE PHOTOBIOREACTOR***

Strain LYB476 bearing plasmid pLybDB4 was cultivated on BG11 agar culture media and single colonies were transferred to 50 ml of liquid BG11 culture media and shaken

at 250 RPM at 30°C under 50 microeinsteins white light until the culture is determined to enter log phase growth as determined by optical density measurements at 730 nm (nominally 3-4 days). The log phase culture is aseptically transferred into 3000 ml of stirred BG11 culture broth maintained at 30°C, aerated with 1 volume of filtered air per minute under 150 microeinsteins white fluorescent light.

Once the culture had achieved a mid log phase growth the cells were aseptically transferred to a sterile trough. Solid state photobioreactor (as described in US Patent Application 20090181434) fabric measuring 6 inches square was submerged in the trough and allowed to incubate overnight in the dark at room temperature. The fabric was aseptically installed into the solid state photobioreactor with gas and culture media plumbing attached. The gas source was ambient air filtered and introduced at 0.1 liters per minute. Culture media was introduced into the reactor at regular 4 hour intervals with 15 minutes active pumping at a flow rate of 0.2 ml per min. Cultures were grown at 30°C under 150 microeinsteins of white fluorescent light. The initial culture media consisted of BG11 with 20 mM sodium nitrate as a nitrogen source which was transitioned after 5 days to BG11 containing 10 mM urea as nitrogen source and the fermentation continued over the course of 7 days. 10 ml of culture broth were removed daily over the course of the trial with pH and sucrose concentration measured. During the culture the pH remained at 7.2 in the collected effluent.

Results showed that the peak sucrose production during the nitrate phase feeding afforded 0.16 millimoles at a concentration of 8.8 mM sucrose. The peak sucrose production during the urea phase feeding afforded 0.94 millimoles at a concentration of 52mM sucrose.

### *EXAMPLE 12: INTEGRATED NITROGEN REGULATED SUCROSE PRODUCTION IN A SOLID STATE PHOTOBIOREACTOR*

Strain LYB511 was cultivated on BG11 agar culture media and single colonies were transferred to 50 ml of liquid BG11 culture media and shaken at 250 RPM at 30°C under 50 microeinsteins white light until the culture is determined to enter log phase growth as determined by optical density measurements at 730 nm (nominally 3-4 days). The log phase culture is aseptically transferred into 3000 ml of stirred BG11 culture broth maintained at 30°C, aerated with 1 volume of filtered air per minute under 150 microeinsteins white fluorescent light. Once the culture had achieved a mid log phase growth the cells were aseptically transferred to a sterile trough. Solid state photobioreactor (as described in Patent Application 20090181434) fabric measuring 6 inches by 6 inches was submerged in the

trough and allowed to incubate overnight in the dark at room temperature. The fabric was aseptically installed into the solid state photobioreactor with gas and culture media plumbing attached. The gas source was ambient air filtered and introduced at 0.1 liters per minute. Culture media was introduced into the reactor at regular 4 hour intervals with 15 minutes active pumping at a flow rate of 0.2 ml per min. Cultures were grown at 30C under 150 microeinsteins of white fluorescent light. The initial culture media consisted of BG11 with 20 mM sodium nitrate as a nitrogen source which was transitioned after 5 days to BG11 containing 10 mM urea as a nitrogen source and the fermentation continued over the course of 7 days. 10 ml of culture broth were removed daily over the course of the trial with pH and sucrose concentration measured. During the culture the pH remained at 7.2 in the collected effluent.

Results showed that the peak sucrose production during the nitrate phase feeding afforded 0.32 millimoles at a concentration of 18 mM sucrose. The peak sucrose production during the urea phase feeding afforded 1.6 millimoles at a concentration of 91 mM sucrose.

REFERENCES

Herrero, A, Muro-Pastor, AM and E Flores. 2001. Nitrogen control of cyanobacteria. J Bact. 183:411-425.

Muro-Pastor, MI, Reyes, JC and FJ Florencio. 2005. Ammonium assimilation in cyanobacteria. Photosynth Res. 83:135-150.

Emlyn-Jones, D, Price, GD and TJ Andrews. 2003. Nitrogen-regulated hypermutator strain of Synechococcus sp. For use in in vivo artificial evolution. Appl Environ Microbiol. 69:6427-6433.

Kamei, A, Yuasa, T, Orikawa, K, Geng, XX and M Ikeuchi. 2001. A eukaryotic type protein kinase, SpkS, is required for normal motility of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. J Bact. 183:1505-1510.

Panoff, J-M, Priem, B, Morvan, H and F Joset. 1988. Sulphated exopolysaccharides produced by two unicellular strains of cyanobacteria, *Synechocystis* PCC 6803 and 6714. Arch Microbiol. 150:558-563.

Panoff, J-M., Joset, F. 1989, Selection by anion-exchange chromatography of exopolysaccharide mutants of the cyanobacterium *Synechocystis* PCC 6803. Appl. Environ. Microbiol. 55, 1452-1456.

Laurentin, A., Edwards, C.A., 2003. A microtiter modification of the anthrone-sulfuric acid colorimetric assay for glucose-based carbohydrates. Anal. Biochem. 315, 143-145.

Kaneko T *et al.* 2001 Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium Anabaena sp. strain PCC 7120. DNA Res 8:205-213.

Nierzwicki-Bauer SA, Curtis SE and Haselkorn R. 1984. Cotranscription of genes encoding the small and large subunits of ribulose-1,5-bisphosphate carboxylase in the cyanobacterium Anabaena 7120. Proc Natl Acad Sci USA 81:5961-5965.

Ramasubramanian TS, Wei T-F and Golden JW. 1994. Two Anabaena sp. strain PCC 7120 DNA-binding factors interact with vegetative cell- and heterocyst-specific genes. J Bacteriol 176:1214-1223.

Vázquez-Bermúdez MF, Flores E and Herrero A. 2002. Analysis of binding sites for the nitrogen-control transcription factor NtcA in the promoters of *Synechococcus* nitrogen-regulated genes. Biochemica et Biophysica Acta 1578:95-98.

5

CLAIMS

Claim 1. A nitrogen-sensitive expression system comprising:

a transcription factor region comprising an NtcA binding site; and

a core promoter region comprising a RuBisCo promoter or a variant or a functional fragment thereof;

wherein the core promoter region comprises or is operably linked to the transcription factor region.

Claim 2. The expression system of claim 1 further comprising:

a transcribable nucleic acid molecule and a 3' transcription termination nucleic acid molecule;

wherein,

the core promoter region is operably linked to the transcribable nucleic acid molecule; and

the transcribable nucleic acid molecule is operably linked to the 3' transcription termination nucleic acid molecule.

Claim 3. The expression system of any one of claims 1-2, wherein the transcription factor region comprises at least one of: a $GTAN_{11}C$ consensus sequence; a $GTAN_8TAC$ consensus sequence; a $GTAN_8TGC$ consensus sequence; a $GTN_{10}AC$ consensus sequence; a $TGTN_9ACA$ consensus sequence; or a $TGTN_{10}ACA$ consensus sequence.

Claim 4. The expression system of any one of claims 1-3, wherein the transcription factor region comprises at least one of a $TAN_3T/A$ box.

Claim 5. The expression system of any one of claims 1-4, wherein the core promoter region comprises a polymerase binding site of a RuBisCo promoter.

Claim 6. The expression system of any one of claims 1-5, wherein the core promoter region comprises at least about 70% sequence identity to a RuBisCo (*rbcLS*) promoter sequence from *Nostoc* sp. PCC 7120; *Synechocystis* sp. PCC 6803; or Anabaena PCC 7120, or a functional fragment thereof.

Claim 7. The expression system of any one of claims 1-6, comprising an isolated nucleic acid molecule having promoter activity and at least about 70% sequence identity to any one of SEQ ID NO: 180, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 183, or SEQ ID NO: 234.

Claim 8. The expression system of any one of claims 1-7, comprising an isolated nucleic acid molecule having promoter activity and

(a) at least 95 contiguous bases of any one of SEQ ID NO: 180, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 183, or SEQ ID NO: 234; or

(b) a functional fragment of any one of SEQ ID NO: 180, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 183, or SEQ ID NO: 234.

Claim 9. The expression system of any one of claims 1-8, comprising

(a) a polynucleotide sequence encoding an NtcA polypeptide or an NtcB polypeptide, wherein the NtcA polypeptide or the NtcB polypeptide is operably linked to the core promoter region of the construct; or

(b) a polynucleotide sequence of SEQ ID NO: 279, encoding an NtcA polypeptide and an NtcB polypeptide, operably linked to the core promoter region.

Claim 10. The expression system of any one of claims 1-9, comprising an:

an isolated nucleic acid molecule of nucleotide positions 177 to 836 of SEQ ID NO: 234, or at least about 70% sequence identity thereto having promoter activity;

an isolated nucleic acid molecule of nucleotide positions 1 to 352 of SEQ ID NO: 234, or at least about 70% sequence identity thereto having promoter activity; or

an isolated nucleic acid molecule of nucleotide positions 177 to 352 of SEQ ID NO: 234, or at least about 70% sequence identity thereto having promoter activity.

Claim 11. The expression system of any one of claims 1-10, wherein the transcribable nucleic acid molecule encodes a polypeptide having: sucrose phosphate synthase activity; sucrose phosphate phosphatase activity; sucrose phosphate synthase and sucrose phosphate phosphatase activity; trehalose phosphate synthase activity; trehalose phosphate phosphatase activity; gluocosylglycerol phosphate synthase activity; gluocosylglycerol phosphate phosphatase activity; mannosylfructose phosphate synthase activity; or mannosylfructose phosphate phosphatase activity.

Claim 12. The expression system of any one of claims 1-10, wherein the transcribable nucleic acid molecule comprises:

SEQ ID NO: 1, or at least 80% sequence identity thereto and encoding a polypeptide having sucrose phosphate synthase and sucrose phosphate phosphatase activity; or

a nucleic acid sequence encoding a polypeptide of SEQ ID NO: 2, or a polypeptide at least 80% similar thereto having sucrose phosphate synthase and sucrose phosphate phosphatase activity.

Claim 13. A method for expressing a transcribable nucleic acid molecule in a host cell comprising:

stably transforming a host cell with an expression system of any one of claims 1-12; and

growing the host cell, or progeny thereof comprising the expression system under conditions whereby the transcribable nucleic acid molecule is expressed;

wherein expression of the transcribable nucleic acid molecule is repressed when nitrate is a predominant nitrogen source and the transcribable nucleic acid molecule is expressed when urea or ammonia is the predominant nitrogen source.

Claim 14. The method of claim 13, wherein

the growth rate of the transformed host organism is no more than about 20% lower than a non-transformed host organism under the same or substantially similar conditions; or

the growth rate of the transformed host organism is more than about a 5% higher growth rate (e.g., more than about a 10%, 15%, or 20% higher growth rate) than a host organism transformed with a non-regulated construct grown under the same or substantially similar conditions.

Claim 15. The method of any one of claims 13-14, wherein

(a) the host cell endogenously expresses an NtcA polypeptide or an NtcB polypeptide under a nitrogen control system;

(b) the host cell is engineered to express an NtcA polypeptide or an NtcB polypeptide under a nitrogen control system;

(c) the expression system comprises a polynucleotide sequence encoding an NtcA polypeptide or an NtcB polypeptide and the NtcA polypeptide or the NtcB polypeptide is operably linked to the core promoter region of the construct.

Claim 16. The method of any one of claims 13-15, wherein the host cell is a photosynthetic microorganism.

Claim 17. The method of any one of claims 13-16, wherein the host cell is a photosynthetic microorganism selected from the group consisting of: *Spirulina maximum*, *Spirulina platensis*, *Dunaliella salina*, *Botrycoccus braunii*, *Chlorella vulgaris*, *Chlorella pyrenoidosa*, *Serenastrum capricomutum*, *Scenedesmus auadricauda*, *Porphyridium cruentum*, *Scenedesmus acutus*, *Dunaliella* sp., *Scenedesmus obliquus*, *Anabaenopsis*, *Aulosira*, *Cylindrospermum*, *Synechoccus* sp., *Synechocystis* sp., and *Tolypothrix.*

Claim 18. An expression cassette comprising:

(a) a transcription factor region comprising an NtcA binding site;

(b) a core promoter region comprising a RuBisCo promoter or a variant or a functional fragment thereof;

(c) a transcribable nucleic acid molecule;

(d) a 3' transcription termination nucleic acid molecule; and

(e) optionally, a polynucleotide sequence encoding an NtcA polypeptide or an NtcB polypeptide;

wherein,

the core promoter region comprises or is operably linked to the transcription factor region;

the core promoter region is operably linked to the transcribable nucleic acid molecule;

the transcribable nucleic acid molecule is operably linked to the 3' transcription termination nucleic acid molecule;

elements (a)-(d) are positioned in relation to each other such that expression of the expression cassette in a host organism results in production of a polypeptide sequence encoded by the transcribable nucleic acid molecule; and

element (e) when present is positioned such that expression of the expression cassette in the host organism results in production of the NtcA polypeptide or the NtcB polypeptide.

Claim 19. A transgenic host cell:

comprising the expression system of any one of claims 1-12;

produced by the method of any one of claims 13-17, or the progeny thereof, comprising the expression system; or

comprising the expression cassette of claim 18;

wherein

expression of the polypeptide sequence encoded by the transcribable nucleic acid molecule is repressed in the presence of nitrate; and

expression of the polypeptide sequence encoded by the transcribable nucleic acid molecule is induced in the presence of urea or ammonia.

Claim 20. A kit comprising the expression system of any one of claims 1-13 or the expression cassette of claim 18 and at least one reagent for introducing the expression cassette into a host cell and, optionally, instructions for introducing the expression cassette into a host cell.