MICROORGANISMS FOR THE PRODUCTION OF ANILINE

BACKGROUND OF THE INVENTION

This invention relates generally to *in silico* design of organisms and engineering of organisms, more particularly to organisms having aniline biosynthesis capability.

Aniline is an organic compound with the formula C_6H_7N and is a precursor to numerous complex chemicals. Aniline is usually produced industrially in two steps from benzene. First, benzene is nitrated using a concentrated mixture of nitric acid and sulfuric acid at 50 to 60 °C, to provide nitrobenzene. In the second step, nitrobenzene is hydrogenated, typically at 600 °C in presence of a nickel catalyst to give aniline. In an alternative process, aniline is prepared from phenol and ammonia as described in U.S. Patent Number 3,965,182. The phenol, in turn, is derived from the cumene process.

The main application of aniline is in the manufacture of polyurethane. Aniline also has value in the production of dyestuffs. In addition to its use as a precursor to dyestuffs, it is a starting-product for the manufacture of many drugs, such as paracetamol (acetaminophen, Tylenol). Currently, the largest market for aniline is preparation of methylene diphenyl diisocyanate (MDI), some 85% of aniline serving this market. Other uses include rubber processing chemicals (9%), herbicides (2%), and dyes and pigments (2%).

When polymerized, aniline can be used as a type of nanowire for use as a semiconducting electrode bridge in, for example, nano-scale devices such as biosensors. These polyaniline nanowires can be doped in order to achieve certain semiconducting properties.

It is desirable to develop a method for production of aniline by alternative means that substitute renewable for petroleum-based feedstocks, while also using less energy- and capital-intensive processes. The present invention satisfies this need and provides related advantages as well.

SUMMARY OF THE INVENTION

In some aspects, embodiments disclosed herein relate to a non-naturally occurring microbial organism having an aniline pathway that includes at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline. The aniline pathway includes an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase.

In some aspects, embodiments disclosed herein relate to a method for producing aniline, that includes culturing a non-naturally occurring microbial organism having an aniline pathway that includes at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline, under conditions and for a sufficient period of time to produce aniline. The aniline pathway includes an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase.

In some aspects, embodiments disclosed herein relate to a non-naturally occurring microbial organism having an aniline pathway that includes at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline. The aniline pathway includes an anthranilate synthase and an anthranilate decarboxylase.

In some aspects, embodiments disclosed herein relate to a method for producing aniline, that includes culturing a non-naturally occurring microbial organism having an aniline pathway that includes at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline, under conditions and for a sufficient period of time to produce aniline. The aniline pathway includes an anthranilate synthase and an anthranilate decarboxylase.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the metabolic pathway to chorismate. E4P is erythrose-4-phosphate, PEP is phosphoenolpyruvate, DAHP is 3-deoxy-D-arabino-heptulosonic acid-7-phosphate.

Figure 2 shows metabolic pathways for the production of aniline. E4P is erythrose-4-phosphate, PEP is phosphoenolpyruvate, DAHP is 3-deoxy-D-arabino-heptulosonic acid-7-phosphate.

DETAILED DESCRIPTION OF THE INVENTION

This invention is directed, in part, to the design and production of cells and microbial organisms incorporating biosynthetic pathways for the production of aniline. Enzymes useful for the production of aniline from the central metabolism precursors erythrose-4-phosphate (E4P) and phosphoenolpyruvate (PEP), via multiple routes, are depicted in Figure 2. Such organisms can utilize renewable feedstocks, providing an alternative to petroleum based aniline production. The maximum theoretical yield of aniline from glucose as the carbon source is 0.857 mole/mole glucose based on the equation 1 below.

$$7C_6H_{12}O_6 + 6\ NH_3 \rightarrow 6\ C_6H_5NH_2 + 6\ CO_2 + 30\ H_2O \quad \text{equation 1}$$

Engineering these pathways into a microorganism involves cloning an appropriate set of genes encoding a set of enzymes into a production host described herein, optimizing fermentation conditions, and assaying product formation following fermentation. To engineer a production host for the production of aniline, one or more exogenous DNA sequence(s) can be expressed in a microorganism. In addition, the microorganism can have endogenous gene(s) functionally disrupted, deleted or overexpressed.

In some embodiments, the invention provides a non-naturally occurring microbial organism having an aniline pathway that includes at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline. The aniline pathway includes an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase, as depicted in Figure 2.

In some embodiments, the invention provides a non-naturally occurring microbial organism having an aniline pathway that includes at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline. The aniline pathway includes an anthranilate synthase and an anthranilate decarboxylase, as depicted in Figure 2.

In some embodiments, the invention provides a method for producing aniline that includes culturing a non-naturally occurring microbial organism having an aniline pathway. The pathway includes at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline, under conditions and for a sufficient period of time to produce aniline. In some embodiments, the aniline pathway includes an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase. In other embodiments the aniline pathway includes an anthranilate synthase and an anthranilate decarboxylase.

As used herein, the term "non-naturally occurring" when used in reference to a microbial organism or microorganism of the invention is intended to mean that the microbial organism has at least one genetic alteration not normally found in a naturally occurring strain of the referenced species, including wild-type strains of the referenced species. Genetic alterations include, for example, modifications introducing expressible nucleic acids encoding metabolic polypeptides, other nucleic acid additions, nucleic acid deletions and/or other functional disruption of the microbial organism's genetic material. Such modifications include, for example, coding regions and functional fragments thereof, for heterologous, homologous or both heterologous and homologous polypeptides for the referenced species. Additional modifications include, for

example, non-coding regulatory regions in which the modifications alter expression of a gene or operon. Exemplary metabolic polypeptides include enzymes or proteins within an aniline biosynthetic pathway.

A metabolic modification refers to a biochemical reaction that is altered from its naturally occurring state. Therefore, non-naturally occurring microorganisms can have genetic modifications to nucleic acids encoding metabolic polypeptides or, functional fragments thereof. Exemplary metabolic modifications are disclosed herein.

As used herein, the term "isolated" when used in reference to a microbial organism is intended to mean an organism that is substantially free of at least one component as the referenced microbial organism is found in nature. The term includes a microbial organism that is removed from some or all components as it is found in its natural environment. The term also includes a microbial organism that is removed from some or all components as the microbial organism is found in non-naturally occurring environments. Therefore, an isolated microbial organism is partly or completely separated from other substances as it is found in nature or as it is grown, stored or subsisted in non-naturally occurring environments. Specific examples of isolated microbial organisms include partially pure microbes, substantially pure microbes and microbes cultured in a medium that is non-naturally occurring.

As used herein, the terms "microbial," "microbial organism" or "microorganism" is intended to mean any organism that exists as a microscopic cell that is included within the domains of archaea, bacteria or eukarya. Therefore, the term is intended to encompass prokaryotic or eukaryotic cells or organisms having a microscopic size and includes bacteria, archaea and eubacteria of all species as well as eukaryotic microorganisms such as yeast and fungi. The term also includes cell cultures of any species that can be cultured for the production of a biochemical.

As used herein, the term "substantially anaerobic" when used in reference to a culture or growth condition is intended to mean that the amount of oxygen is less than about 10% of saturation for dissolved oxygen in liquid media. The term also is intended to include sealed chambers of liquid or solid medium maintained with an atmosphere of less than about 1% oxygen.

"Exogenous" as it is used herein is intended to mean that the referenced molecule or the referenced activity is introduced into the host microbial organism. The molecule can be introduced, for example, by introduction of an encoding nucleic acid into the host genetic material such as by integration into a host chromosome or as non-chromosomal genetic material

such as a plasmid. Therefore, the term as it is used in reference to expression of an encoding nucleic acid refers to introduction of the encoding nucleic acid in an expressible form into the microbial organism. When used in reference to a biosynthetic activity, the term refers to an activity that is introduced into the host reference organism. The source can be, for example, a homologous or heterologous encoding nucleic acid that expresses the referenced activity following introduction into the host microbial organism. Therefore, the term "endogenous" refers to a referenced molecule or activity that is present in the host. Similarly, the term when used in reference to expression of an encoding nucleic acid refers to expression of an encoding nucleic acid contained within the microbial organism. The term "heterologous" refers to a molecule or activity derived from a source other than the referenced species whereas "homologous" refers to a molecule or activity derived from the host microbial organism. Accordingly, exogenous expression of an encoding nucleic acid of the invention can utilize either or both a heterologous or homologous encoding nucleic acid.

It is understood that when more than one exogenous nucleic acid is included in a microbial organism that the more than one exogenous nucleic acid refers to the referenced encoding nucleic acid or biosynthetic activity, as discussed above. It is further understood, as disclosed herein, that such more than one exogenous nucleic acids can be introduced into the host microbial organism on separate nucleic acid molecules, on polycistronic nucleic acid molecules, or a combination thereof, and still be considered as more than one exogenous nucleic acid. For example, as disclosed herein a microbial organism can be engineered to express two or more exogenous nucleic acids encoding a desired pathway enzyme or protein. In the case where two exogenous nucleic acids encoding a desired activity are introduced into a host microbial organism, it is understood that the two exogenous nucleic acids can be introduced as a single nucleic acid, for example, on a single plasmid, on separate plasmids, can be integrated into the host chromosome at a single site or multiple sites, and still be considered as two exogenous nucleic acids. Similarly, it is understood that more than two exogenous nucleic acids can be introduced into a host organism in any desired combination, for example, on a single plasmid, on separate plasmids, can be integrated into the host chromosome at a single site or multiple sites, and still be considered as two or more exogenous nucleic acids, for example three exogenous nucleic acids. Thus, the number of referenced exogenous nucleic acids or biosynthetic activities refers to the number of encoding nucleic acids or the number of biosynthetic activities, not the number of separate nucleic acids introduced into the host organism.

The non-naturally occurring microbial organisms of the invention can contain stable genetic alterations, which refers to microorganisms that can be cultured for greater than five generations without loss of the alteration. Generally, stable genetic alterations include modifications that persist greater than 10 generations, particularly stable modifications will persist more than about 25 generations, and more particularly, stable genetic modifications will be greater than 50 generations, including indefinitely.

Those skilled in the art will understand that the genetic alterations, including metabolic modifications exemplified herein, are described with reference to a suitable host organism such as *E. coli* and their corresponding metabolic reactions or a suitable source organism for desired genetic material such as genes for a desired metabolic pathway. However, given the complete genome sequencing of a wide variety of organisms and the high level of skill in the area of genomics, those skilled in the art will readily be able to apply the teachings and guidance provided herein to essentially all other organisms. For example, the *E. coli* metabolic alterations exemplified herein can readily be applied to other species by incorporating the same or analogous encoding nucleic acid from species other than the referenced species. Such genetic alterations include, for example, genetic alterations of species homologs, in general, and in particular, orthologs, paralogs or nonorthologous gene displacements.

An ortholog is a gene or genes that are related by vertical descent and are responsible for substantially the same or identical functions in different organisms. For example, mouse epoxide hydrolase and human epoxide hydrolase can be considered orthologs for the biological function of hydrolysis of epoxides. Genes are related by vertical descent when, for example, they share sequence similarity of sufficient amount to indicate they are homologous, or related by evolution from a common ancestor. Genes can also be considered orthologs if they share three-dimensional structure but not necessarily sequence similarity, of a sufficient amount to indicate that they have evolved from a common ancestor to the extent that the primary sequence similarity is not identifiable. Genes that are orthologous can encode proteins with sequence similarity of about 25% to 100% amino acid sequence identity. Genes encoding proteins sharing an amino acid similarity less that 25% can also be considered to have arisen by vertical descent if their three-dimensional structure also shows similarities. Members of the serine protease family of enzymes, including tissue plasminogen activator and elastase, are considered to have arisen by vertical descent from a common ancestor.

Orthologs include genes or their encoded gene products that through, for example, evolution, have diverged in structure or overall activity. For example, where one species encodes a gene

product exhibiting two functions and where such functions have been separated into distinct genes in a second species, the three genes and their corresponding products are considered to be orthologs. For the production of a biochemical product, those skilled in the art will understand that the orthologous gene harboring the metabolic activity to be introduced or disrupted is to be chosen for construction of the non-naturally occurring microorganism. An example of orthologs exhibiting separable activities is where distinct activities have been separated into distinct gene products between two or more species or within a single species. A specific example is the separation of elastase proteolysis and plasminogen proteolysis, two types of serine protease activity, into distinct molecules as plasminogen activator and elastase. A second example is the separation of mycoplasma 5'-3' exonuclease and *Drosophila* DNA polymerase III activity. The DNA polymerase from the first species can be considered an ortholog to either or both of the exonuclease or the polymerase from the second species and vice versa.

In contrast, paralogs are homologs related by, for example, duplication followed by evolutionary divergence and have similar or common, but not identical functions. Paralogs can originate or derive from, for example, the same species or from a different species. For example, microsomal epoxide hydrolase (epoxide hydrolase I) and soluble epoxide hydrolase (epoxide hydrolase II) can be considered paralogs because they represent two distinct enzymes, co-evolved from a common ancestor, that catalyze distinct reactions and have distinct functions in the same species. Paralogs are proteins from the same species with significant sequence similarity to each other suggesting that they are homologous, or related through co-evolution from a common ancestor. Groups of paralogous protein families include HipA homologs, luciferase genes, peptidases, and others.

A nonorthologous gene displacement is a nonorthologous gene from one species that can substitute for a referenced gene function in a different species. Substitution includes, for example, being able to perform substantially the same or a similar function in the species of origin compared to the referenced function in the different species. Although generally, a nonorthologous gene displacement will be identifiable as structurally related to a known gene encoding the referenced function, less structurally related but functionally similar genes and their corresponding gene products nevertheless will still fall within the meaning of the term as it is used herein. Functional similarity requires, for example, at least some structural similarity in the active site or binding region of a nonorthologous gene product compared to a gene encoding the function sought to be substituted. Therefore, a nonorthologous gene includes, for example, a paralog or an unrelated gene.

Therefore, in identifying and constructing the non-naturally occurring microbial organisms of the invention having aniline biosynthetic capability, those skilled in the art will understand with applying the teaching and guidance provided herein to a particular species that the identification of metabolic modifications can include identification and inclusion or inactivation of orthologs. To the extent that paralogs and/or nonorthologous gene displacements are present in the referenced microorganism that encode an enzyme catalyzing a similar or substantially similar metabolic reaction, those skilled in the art also can utilize these evolutionally related genes.

Orthologs, paralogs and nonorthologous gene displacements can be determined by methods well known to those skilled in the art. For example, inspection of nucleic acid or amino acid sequences for two polypeptides will reveal sequence identity and similarities between the compared sequences. Based on such similarities, one skilled in the art can determine if the similarity is sufficiently high to indicate the proteins are related through evolution from a common ancestor. Algorithms well known to those skilled in the art, such as Align, BLAST, Clustal W and others compare and determine a raw sequence similarity or identity, and also determine the presence or significance of gaps in the sequence which can be assigned a weight or score. Such algorithms also are known in the art and are similarly applicable for determining nucleotide sequence similarity or identity. Parameters for sufficient similarity to determine relatedness are computed based on well known methods for calculating statistical similarity, or the chance of finding a similar match in a random polypeptide, and the significance of the match determined. A computer comparison of two or more sequences can, if desired, also be optimized visually by those skilled in the art. Related gene products or proteins can be expected to have a high similarity, for example, 25% to 100% sequence identity. Proteins that are unrelated can have an identity which is essentially the same as would be expected to occur by chance, if a database of sufficient size is scanned (about 5%). Sequences between 5% and 24% may or may not represent sufficient homology to conclude that the compared sequences are related. Additional statistical analysis to determine the significance of such matches given the size of the data set can be carried out to determine the relevance of these sequences.

Exemplary parameters for determining relatedness of two or more sequences using the BLAST algorithm, for example, can be as set forth below. Briefly, amino acid sequence alignments can be performed using BLASTP version 2.0.8 (Jan-05-1999) and the following parameters: Matrix: 0 BLOSUM62; gap open: 11; gap extension: 1; x_dropoff: 50; expect: 10.0; wordsize: 3; filter: on. Nucleic acid sequence alignments can be performed using BLASTN version 2.0.6 (Sept-16-1998) and the following parameters: Match: 1; mismatch: -2; gap open: 5; gap

extension: 2; x_dropoff: 50; expect: 10.0; wordsize: 11; filter: off. Those skilled in the art will know what modifications can be made to the above parameters to either increase or decrease the stringency of the comparison, for example, and determine the relatedness of two or more sequences.

In some embodiments, the invention provides a non-naturally occurring microbial organism, comprising a microbial organism having an aniline pathway comprising at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline, said aniline pathway comprising an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase. In some embodiments, such a non-naturally occurring microbial organism can further include a DAHP synthase and in still further embodiments, the non-naturally occurring microbial organism can further include a 3-dehydroquinate synthase.

In some embodiments, the non-naturally occurring microbial organism includes two exogenous nucleic acids each encoding an aniline pathway enzyme, while in other embodiments the non-naturally occurring microbial organism includes three exogenous nucleic acids each encoding an aniline pathway enzyme. For example, in some embodiments, the non-naturally occurring microbial organism can include three exogenous nucleic acids encoding an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase.

In some embodiments, the non-naturally occurring microbial organism can include four exogenous nucleic acids each encoding an aniline pathway enzyme. For example, a non-naturally occurring microbial organism having four exogenous nucleic acids can encode a DAHP synthase, an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase. The DAHP synthase, which can be endogenous to the non-naturally occurring microbial organism, can be overexpressed, for example, by insertion of additional copies of the gene and/or through the use of exogenous regulatory genes and removing feedback regulation by the aromatic amino acids.

In still further embodiments, the non-naturally occurring microbial organism can include five exogenous nucleic acids each encoding an aniline pathway enzyme. For example, the non-naturally occurring microbial organism having five exogenous nucleic acids can encode a 3-dehydroquinate synthase, a DAHP synthase, an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylase. The 3-dehydroquinate

synthase, which can be endogenous to the non-naturally occurring microbial organism, can also be overexpressed, for example, by insertion of additional copies of the gene and/or through the use of exogenous regulatory genes.

Moreover, any one or more of the other enzymes that are in a pathway en route to chorismate, which can be endogenous in some embodiments, can be overexpressed to increase the production of chorismate. These include, for example, a 3-dehydroquinate dehydratase (EC 4.2.1.10), a shikimate dehydrogenase (1.1.1.25), a quinate/shikimate dehydrogenase (1.1.1.282), a shikimate kinase (2.7.1.71), a 3-phosphoshikimate-1-carboxyvinyltransferase (2.5.1.19), and a chorismate synthase (4.2.3.5). These enzymes constitute the pathway for making chorismate from DAHP in prokaryotes and most eukaryotes. An alternative pathway for formation of 3-dehydroquinate (the steps from 3-dehydroquinate to chorismate are the same in all organisms, including arachea) includes the following enzymatic steps: triosephosphate isomerase, frustose-1,6-bisphosphate aldolase, 2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate synthase, and dehydroquinate synthase.

In some embodiments, the non-naturally occurring microbial organisms described above can have at least one exogenous nucleic acid which is a heterologous nucleic acid. Moreover, the non-naturally occurring microbial organisms described above, can be provided in a substantially anaerobic culture medium.

In some embodiments, the present invention also provides a non-naturally occurring microbial organism having an aniline pathway that includes at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline, in which the aniline pathway includes an anthranilate synthase and an anthranilate decarboxylase. Such a non-naturally occurring microbial organism can further include a DAHP synthase, as described above. In some embodiments, such a non-naturally occurring microbial organism can further include a 3-dehydroquinate synthase.

In some embodiments, this non-naturally occurring microbial includes two exogenous nucleic acids each encoding an aniline pathway enzyme. For example, the two exogenous nucleic acids can encode an anthranilate synthase and an anthranilate decarboxylase. In some embodiments the microbial organism can include three exogenous nucleic acids each encoding an aniline pathway enzyme. For example, the three exogenous nucleic acids can encode a DAHP synthase, an anthranilate synthase and an anthranilate decarboxylase. In some embodiments, the microbial organism includes four exogenous nucleic acids each encoding an aniline pathway enzyme. For

example, the four exogenous nucleic acids can encode a 3-dehydroquinate synthase, a DAHP synthase, an anthranilate synthase and an anthranilate decarboxylase.

As described above, any one or more of the other enzymes that are in a pathway en route to chorismate, which can be endogenous in some embodiments, can be overexpressed to increase the production of chorismate. These include, for example, a 3-dehydroquinate dehydratase, a shikimate dehydrogenase, a quinate/shikimate dehydrogenase, a shikimate kinase, a 3-phosphoshikimate-1-carboxyvinyltransferase, and a chorismate synthase.

In some embodiments, such non-naturally occurring microbial organisms described above can include at least one exogenous nucleic acid is a heterologous nucleic acid. In some embodiments, such non-naturally occurring microbial organisms are in a substantially anaerobic culture medium.

In an additional embodiment, the invention provides a non-naturally occurring microbial organism having an aniline pathway, wherein the non-naturally occurring microbial organism comprises at least one exogenous nucleic acid encoding an enzyme or protein that converts a substrate to a product selected from the group consisting of chorismate to 4-amino-4-deoxychorismate, 4-amino-4-deoxychorismate to *p*-aminobenzoate, and *p*-aminobenzoate to aniline.

In an additional embodiment, the invention provides a non-naturally occurring microbial organism having an aniline pathway, wherein the non-naturally occurring microbial organism comprises at least one exogenous nucleic acid encoding an enzyme or protein that converts a substrate to a product selected from the group consisting of chorismate to anthranilate, and anthranilate to aniline.

One skilled in the art will understand that these are merely exemplary and that any of the substrate-product pairs disclosed herein suitable to produce a desired product and for which an appropriate activity is available for the conversion of the substrate to the product can be readily determined by one skilled in the art based on the teachings herein. Thus, the invention provides a non-naturally occurring microbial organism containing at least one exogenous nucleic acid encoding an enzyme or protein, where the enzyme or protein converts the substrates and products of an aniline pathway, such as that shown in Figure 1.

While generally described herein as a microbial organism that contains an aniline pathway, it is understood that the invention additionally provides a non-naturally occurring microbial

organism comprising at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce an intermediate of an aniline pathway. For example, as disclosed herein, an aniline pathway is exemplified in Figure 1. Therefore, in addition to a microbial organism containing an aniline pathway that produces aniline, the invention provides a non-naturally occurring microbial organism comprising at least one exogenous nucleic acid encoding an aniline pathway enzyme, where the microbial organism produces an aniline pathway intermediate, for example, DAHP, chorismate, anthranilate, 4-amino-4-deoxychorismate, or *p*-aminobenzoate.

It is understood that any of the pathways disclosed herein, as described in the Examples and exemplified in the Figures, including the pathways of Figure 1, can be utilized to generate a non-naturally occurring microbial organism that produces any pathway intermediate or product, as desired. As disclosed herein, such a microbial organism that produces an intermediate can be used in combination with another microbial organism expressing downstream pathway enzymes to produce a desired product. However, it is understood that a non-naturally occurring microbial organism that produces an aniline pathway intermediate can be utilized to produce the intermediate as a desired product.

The invention is described herein with general reference to the metabolic reaction, reactant or product thereof, or with specific reference to one or more nucleic acids or genes encoding an enzyme associated with or catalyzing, or a protein associated with, the referenced metabolic reaction, reactant or product. Unless otherwise expressly stated herein, those skilled in the art will understand that reference to a reaction also constitutes reference to the reactants and products of the reaction. Similarly, unless otherwise expressly stated herein, reference to a reactant or product also references the reaction, and reference to any of these metabolic constituents also references the gene or genes encoding the enzymes that catalyze or proteins involved in the referenced reaction, reactant or product. Likewise, given the well known fields of metabolic biochemistry, enzymology and genomics, reference herein to a gene or encoding nucleic acid also constitutes a reference to the corresponding encoded enzyme and the reaction it catalyzes or a protein associated with the reaction as well as the reactants and products of the reaction.

As shown in Figure 1, the first step of an aniline pathway is an aldol-type condensation that combines one molecule of E4P and one molecule of PEP to form the intermediate, 3-deoxy-D-arabino-heptulosonic acid 7-phosphonate (DAHP). This enzyme is referred to as DAHP synthase, or equivalently, 2-dehydro-3-deoxyphosphoheptonate aldolase. This reaction (EC #

2.5.1.54) is the first committed step in the shikimate pathway and is required for the biosynthesis of aromatic amino acids, folates, quinones and other secondary metabolites in bacteria, fungi and plants. DAHP synthases have been categorized into AroAI and AroAII classes (Wu et al., *J. Biol. Chem.* 281:4042-4048 (2006)). The former class comprises of mainly microbial proteins while the latter is comprised of primarily plant proteins.

In *Escherichia coli*, the function is catalyzed by three genes: aroFGH. Each of these encodes for an isozyme and is feedback regulated by a different aromatic amino acid. In contrast, some other organisms, such as *Bacillus subtilis* and *Porphyromonas gingivalis* are bifunctional enzymes. The aroA gene encodes for DAHP synthase activity and aroQ gene encodes for chorismate mutase activity in *B. subtilis*. However, these activities can be separated by domain truncation ((Wu et al., *J. Biol. Chem.* 281:4042-4048 (2006)). The *B. subtilis* enzyme is sensitive to the downstream intermediates, chorismate and prephanate. The DAHP synthase from *Corynebacterium glutamicum* is feedback sensitive to both, phenylalanine and tyrosine (Wu et al., *J. Biol. Chem.* 278:27525-27531 (2003)).

These enzymes are metalloenzymes and their mechanisms of regulation are well-understood by those skilled in the art. The crystal structures of the *E. coli* and *S. cerevisiae* DAHP synthases have been solved and reveal structures consisting of $(\beta/\alpha)_8$ barrel. There are several enzymes that however, don't have regulatory domains and belong to organisms such as *Pyrococcus furiosus* and *Nostoc sp*. Exemplary genes are summarized below in Table 1.

Table 1.

Gene	GI number	GenBank ID	Organism
aroF	16130522	NP_417092.1	*Escherichia coli K12 MG1655*
aroG	16128722	NP_415275.1	*Escherichia coli K12 MG1655*
aroH	16129660	NP_416219.1	*Escherichia coli K12 MG1655*
aroA	16080027	NP_390853.1	*Bacillus subtilis*
aroA	34396967	AAQ66031.1	*Porphyromonas gingivalis W83*
Aro4	6319726	NP_009808.1	*Saccharomyces cerevisiae*
aroG	1168513	P44303.1	*Haemophilus influenza*
aroF	16765985	NP_461600.1	*Salmonella typhimurium*
aroG	21903376	P35170.2	*Corynebacterium glutamicum*
PF1690	18893851	AAL81814.1	*Pyrococcus furiosus*
alr3050	17132144	BAB74749	*Nostoc sp*

Chorismate is a known intermediate for aromatic amino acid biosynthesis in Gram positive, Gram-negative bacteria and in archaea. It is also a precursor for the production of folic acid, ubiquinone, menaquinone and enterocholein in some microorganisms. DAHP can be converted into 3-dehydroquinate and that can be subsequently converted into chorismate via multiple well-known steps. In *E. coli*, DAHP can be converted into 3-dehydroquinate by 3-dehydroquinate synthase. The synthase in *E. coli* is understood to catalyze an oxidation, a β-elimination, an intramolecular aldol condensation and a reduction (Frost et al., *Biochemistry* 23:4470-4475 (1984); Maitra et al., *J. Biol. Chem.* 253:5426-5430 (1978)). The enzyme requires catalytic amounts of NAD^+ and Co^{2+} (Maitra et al., *J. Biol. Chem.* 253:5426-5430 (1978)). Enzymes useful for the production of chorismate include, for example, a 3-dehydroquinate dehydratase, a shikimate dehydrogenase, a quinate/shikimate dehydrogenase, a shikimate kinase, a 3-phosphoshikimate-1-carboxyvinyltransferase, and a chorismate synthase, as described above

The conversion of chorismate into 4-amino-4-deoxychorismate can be accomplished by aminodeoxychorismate synthase (EC# 2.6.1.85), also referred to as chorismate L-glutamine aminotransferase. In *E. coli*, the function is catalyzed by two genes, pabA and pabB. The pabA polypeptide is a conditional glutaminase which requires a 1:1 complex with pabB for activity. The pabB enzyme uses the nascent ammonia released by this reaction to transform chorismate to 4-amino-4-deoxychorismate (in the presence of Mg^{2+}). The pabB reaction is fully reversible. In the absence of pabA, pabB utilizes NH_3 at significantly reduced rates (Roux and Walsh, *Biochemistry* 32:3763-3768 (1993); Roux and Walsh, *Biochemistry* 31:6904-6910 (1992)).

A similar enzyme complex formed by pabA and pabB, catalyzes the conversion of chorismate into 4-amino-4-dexoychorismate in *Streptomyces venezuelae* (Brown et al., *Microbiology* 142(pt 6): 1345-1355 (1996)). This organism is known to have more than one set of pabAB genes (Chang et al., *Microbiology* 147:2113-2126 (2001)). The gene with the aforementioned function has been identified in *Arabidopsis thaliana* and *Solanum lycopersicum* also. The protein sequences of the PabA and PabB genes of *E. coli* were used to isolate the cDNA encoding the aminodeoxychorismate synthase (ADCS) in *Arabidopsis thaliana* (Basset et al., *Proc. Natl. Acad. Sci. U.S.A.* 101:1496-1501 (2004)). The enzyme was recombinantly expressed in *E. coli* demonstrating the formation of 4-amino-4-deoxychorismate. No feedback inhibition of the enzyme has been reported for either *p*-aminobenzoate or folate. The corresponding genes, along with their GenBank ids are listed below in Table 2:

Table 2

pabAB	710438	AAB30312.1	*Streptomyces venezuelae*
pabA	16131239	NP_417819.1	*Escherichia coli K12 MG1655*
pabB	16129766	NP_416326.1	*Escherichia coli K12 MG1655*
pabA	152972254	YP_001337400.1	*Klebsiella pneumoniae*
pabB	152970875	YP_001335984.1	*Klebsiella pneumoniae*
pabA	118467576	YP_884448.1	*Mycobacterium smegmatis*
pabB	118473035	YP_889684.1	*Mycobacterium smegmatis*

In several organisms, the gene encoding for aminodeoxychorismate lyase (EC# 4.1.3.38) is typically coupled with *pabB* and *pabA* to catalyze the conversion of aminodexoychorismate into *p*-aminobenzoate, with the release of a pyruvate molecule. In both *E. coli* (Green et al., *J. Bacteriol.* 174:5317-5323 (1992); Green and Nichols, *J. Biol. Chem.* 266:12971-12975 (1991)) and *S. venezuelae*, *pabC* catalyzes this reaction. Recently, 4-amino-4-deoxychorismate lyase was functionally characterized in two more species of *Streptomcyes*, namely FR-008 and *griseus* (Zhang et al., *Microbiology* 155:2450-2459 (2009)). Aminodeoxychorismate synthase and aminodeoxychorismate lyase are typically part of folate biosynthesis in most organisms and facilitate the conversion of chorismate into para-aminobenzoate. Aminodeoxychorismate lyase is a pyridoxal-phosphate dependent protein. A putative enzyme has been found in *A. thaliana* ((Basset et al., *Proc. Natl. Acad. Sci. U.S.A.* 101:1496-1501 (2004)) and *B. subtilis* (Schadt et al., *J. Am. Chem. Soc.* 131:3481-3493 (2009); Slock et al., *J. Bacteriol.* 172:7211-7226 (1990)), as part of the folate operon. Some exemplary genes are shown below in Table 3:

Table 3

pabC	16129059	NP_415614.1	*Escherichia coli K12 MG1655*
pabC	16077144	NP_387957.1	*Bacillus subtilis*
pabC	29828105	NP_822739.1	*Streptomyces avermitilis*
pabC-1	224831591	AAQ82550.2	*Streptomyces sp. FR-008*
pabC-2	219879202	ACL50980.1	*Streptomyces sp. FR-008*

Anthranilate synthase (EC: 4.1.3.27), also known by the systematic name chorismate pyruvate-lyase (amino-accepting; anthranilate-forming) or by the synonym glutamine amidotransferase, is the first step in the tryptophan synthesis pathway from chorismate. The formation of anthranilate is accompanied by the transfer of an amine group from glutamine and leading to the formation of glutamate. Pyruvate is also released during the reaction. In *E. coli,* this reaction is catalyzed by a tetrameric enzyme complex comprised of two monomers of TrpD and two

monomers of TrpE. TrpE on its own can carry out an alternate version of this reaction, using ammonium sulfate rather than glutamine as an amino donor (Ito et al., *Acta. Pathol. Jpn.* 19:55-67 (1969); Ito and Yanofsky, *J. Bacteriol.* 97:734-742 (1969)). However, TrpD increases the affinity of TrpE for glutamine over TrpE alone. The enzyme is feedback regulated by tryptophan. This feedback regulation is also observed for the enzyme complex in the hyperthermophilic *Sulfolobus solfataricus*. The enzyme complex from this organism has been expressed in *E. coli* (Tutino et al., *Biochem. Biophys. Res. Commun.* 230:306-310 (1997)). The thermodynamics of the reaction catalyzed by anthranilate synthase has been described in *Salmonella typhimurium* (Byrnes et al., *Biophys. Chem.* 84:45-64 (2000)). The subunits of the enzyme complex have also been described in *Thermotoga maritima (*Kim et al., *J. Mol. Biol.* 231:960-981 (1993)*)*. A summary of these genes is shown below in Table 4.

Table 4

TrpD	16129224	NP_415779.1	*Escherichia coli K12 MG1655*
TrpE	16129225	NP_415780.1	*Escherichia coli K12 MG1655*
TrpE	15897780	NP_342385.1	*Sulfolobus solfataricus*
TrpGD	15897781	NP_342386.1	*Sulfolobus solfataricus*
trpD	16765068	NP_460683.1	*Salmonella typhimurium*
trpE	16765067	NP_460682.1	*Salmonella typhimurium*
trpE	15642916	NP_227957.1	*Thermotoga maritima*
trpGD	15642915	NP_227956.1	*Thermotoga maritima*

The decarboxylation of *p*-aminobenzoate and anthranilate can be catalyzed by an aminobenzoate carboxylyase (McCullough et al., *J. Am. Chem. Soc.* 79:628-630 (1957)). It has been indicated that the cell free enzyme obtained from *E. coli* 0111:B4 was capable of decarboxylating both of these molecules. The activity of the enzyme was found to be dependent on pyridoxal phosphate and iron (III). The conversion of *p*-aminobenzoate to aniline in some extracts of *Mycobacteria* has been described (Sloane et al., *J. Biol. Chem.* 193:453-458 (1951)). New strains have been identified that are capable of degrading aniline anaerobically (Kahng et al., *FEMS Microbiol. Lett.* 190:215-221 (2000); Schnell et al., Arch. Microbiol. 152:556-563 (1989)). These strains first carboxylate aniline to 4-aminobenzoate. In the strain, *Desulfobacterium anilini*, the rate of aniline degradation is dependent on the presence of CO_2 in the medium. GC analysis of aniline culture supernatant of strain HY99 under anaerobic, denitrifying conditions showed the presence of 4-aminobenzoate (Kahng et al., *FEMS Microbiol. Lett.* 190:215-221 (2000)).

Numerous other studies have been conducted on decarboxylation of aromatic compounds, primarily hydroxyl aromatics. For example, a 4-hydroxybenzoate decarboxylase has been identified from the facultative anaerobe, *Enterobacter cloacae* (Matsui et al., *Arch. Microbiol.* 186:21-29 (2006)). The corresponding gene has been sequenced. The enzyme has been tested for activity on multiple substrates and was shown to be induced by both 4-hydroxybenzoic acid and 4-aminobenzoic acid. Another decarboxylase has been reported in *Clostridium theromaceticum* that can remove CO_2 from *p*-hydroxy benzoate (Hsu et al., *J. Bacteriol.* 172:5901-5907 (1990)). The enzyme has broad substrate specificity and can act on *p*-hydroxy benzoate with varied functional group substituents at the meta-position. These include hydroxyl, chloro, fluoro, and methoxy groups. The enzyme was not repressed by glucose or other external energy sources. *Klebsiella aerogens* was also reported to be able to carry out non-oxidative decarboxylation of para-hydroxy benzoate, 2,5-dihydroxybenzoate, 3,4-dihydroxybenzoate and 3,4,5- trihydroxybenzoate (Grant et al., *Antonie Van Leeuwenhoek* 35:325-343 (1969)). A reversible 4-hydroxybenzoate decarboxylase was purified from *Clostridium hydroxybenzoicum* (now called *Sedimentibacter hydroxybenzoicus*). This enzyme is encoded by three clustered genes, *shdB*, *C* and *D*. The enzyme can act on both 4-hydroxybenzoate and 3,4-dihydroxybenzoate. The enzyme activity was not affected by metal ions or other cofactor (He et al., *Eur. J. Biochem.* 229:77-82 (1995)). *Bacillus subtilis* was recently demonstrated to have a hydroxyarylic acid decarboxylase activity. Three genes *bcdB, C,* and *D* were cloned in *E. coli* and showed activity on 4-hydroxybenzoate and vanillate (Lupa et al., *Can. J. Microbiol.* 54:75-81 (2008)). These decarboxylases have been reported in several other organisms (Lupa et al., *Genomics* 86:342-351 (2005)) and gene candidates for some of these are listed below in Table 5.

Table 5

shdB	67462197	AAY67850.1	*Sedimentibacter hydroxybenzoicus*
shdC	5739200	AAD50377.1	*Sedimentibacter hydroxybenzoicus*
shdD	67462198	AAY67851.1	*Sedimentibacter hydroxybenzoicus*
	110331749	BAE97712.1	*Enterobacter cloacae*
bsdB	13124411	P94404.1	*Bacillus subtilis*
bsdC	6686207	P94405.1	*Bacillus subtilis*
bsdD	239977069	C0H3U9.1	*Bacillus subtilis*
STM292	16766227	NP_461842.1	*Salmonella typhimurium LT2*
STM2922	16766228	NP_461843.1	*Salmonella typhimurium LT2*
STM2923	16766229	NP_461844.1	*Salmonella typhimurium LT2*
kpdB	206580833	YP_002236894.1	*Klebsiella pneumoniae 342*
kpdC	206576360	YP_002236895.1	*Klebsiella pneumoniae 342*
kpdD	206579343	YP_002236896.1	*Klebsiella pneumoniae 342*
pad1	15832847	NP_311620.1	*Escherichia coli O157*
yclC	15832846	NP_311619.1	*Escherichia coli O157*
yclD	15832845	NP_311618.1	*Escherichia coli O157*

It is understood that modifications which do not substantially affect the activity of the various embodiments of this invention are also included within the definition of the invention provided herein. Accordingly, the following examples are intended to illustrate but not limit the present invention.

The non-naturally occurring microbial organisms of the invention can be produced by introducing expressible nucleic acids encoding one or more of the enzymes or proteins participating in one or more aniline biosynthetic pathways. Depending on the host microbial organism chosen for biosynthesis, nucleic acids for some or all of a particular aniline biosynthetic pathway can be expressed. For example, if a chosen host is deficient in one or more enzymes or proteins for a desired biosynthetic pathway, then expressible nucleic acids for the deficient enzyme(s) or protein(s) are introduced into the host for subsequent exogenous expression. Alternatively, if the chosen host exhibits endogenous expression of some pathway genes, but is deficient in others, then an encoding nucleic acid is needed for the deficient enzyme(s) or protein(s) to achieve aniline biosynthesis. Thus, a non-naturally occurring microbial organism of the invention can be produced by introducing exogenous enzyme or

protein activities to obtain a desired biosynthetic pathway or a desired biosynthetic pathway can be obtained by introducing one or more exogenous enzyme or protein activities that, together with one or more endogenous enzymes or proteins, produces a desired product such as aniline.

Host microbial organisms can be selected from, and the non-naturally occurring microbial organisms generated in, for example, bacteria, yeast, fungus or any of a variety of other microorganisms applicable to fermentation processes. Exemplary bacteria include species selected from *Escherichia coli, Klebsiella oxytoca, Anaerobiospirillum succiniciproducens, Actinobacillus succinogenes, Mannheimia succiniciproducens, Rhizobium etli, Bacillus subtilis, Corynebacterium glutamicum, Gluconobacter oxydans, Zymomonas mobilis, Lactococcus lactis, Lactobacillus plantarum, Streptomyces coelicolor, Clostridium acetobutylicum, Pseudomonas fluorescens*, and *Pseudomonas putida*. Exemplary yeasts or fungi include species selected from *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces lactis, Kluyveromyces marxianus, Aspergillus terreus, Aspergillus niger*, *Pichia pastoris, Rhizopus arrhizus, Rhizobus oryzae*, and the like. *E. coli* is a particularly useful host organism since it is a well characterized microbial organism suitable for genetic engineering. Other particularly useful host organisms include yeast such as *Saccharomyces cerevisiae*. It is understood that any suitable microbial host organism can be used to introduce metabolic and/or genetic modifications to produce a desired product.

Depending on the aniline biosynthetic pathway constituents of a selected host microbial organism, the non-naturally occurring microbial organisms of the invention will include at least one exogenously expressed aniline pathway-encoding nucleic acid and up to all encoding nucleic acids for one or more aniline biosynthetic pathways. For example, aniline biosynthesis can be established in a host deficient in a pathway enzyme or protein through exogenous expression of the corresponding encoding nucleic acid. In a host deficient in all enzymes or proteins of an aniline pathway, exogenous expression of all enzyme or proteins in the pathway can be included, although it is understood that all enzymes or proteins of a pathway can be expressed even if the host contains at least one of the pathway enzymes or proteins. For example, exogenous expression of all enzymes or proteins in a pathway for production of aniline can be included, such as a 3-dehydroquinate synthase, a DAHP synthase, an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase or a 3-dehydroquinate synthase, a DAHP synthase, an anthranilate synthase and an anthranilate decarboxylase.

Given the teachings and guidance provided herein, those skilled in the art will understand that the number of encoding nucleic acids to introduce in an expressible form will, at least, parallel the aniline pathway deficiencies of the selected host microbial organism. Therefore, a non-naturally occurring microbial organism of the invention can have one, two, three, four, five, up to all nucleic acids encoding the enzymes or proteins constituting an aniline biosynthetic pathway disclosed herein. In some embodiments, the non-naturally occurring microbial organisms also can include other genetic modifications that facilitate or optimize aniline biosynthesis or that confer other useful functions onto the host microbial organism. One such other functionality can include, for example, augmentation of the synthesis of one or more of the aniline pathway precursors such as chorismate, anthranilate, 4-amino-4-deoxychorismate, and *p*-aminobenzoate.

Generally, a host microbial organism is selected such that it produces the precursor of an aniline pathway, either as a naturally produced molecule or as an engineered product that either provides *de novo* production of a desired precursor or increased production of a precursor naturally produced by the host microbial organism. For example, chorismate is produced naturally in a host organism such as *E. coli*. A host organism can be engineered to increase production of a precursor, as disclosed herein. In addition, a microbial organism that has been engineered to produce a desired precursor can be used as a host organism and further engineered to express enzymes or proteins of an aniline pathway.

In some embodiments, a non-naturally occurring microbial organism of the invention is generated from a host that contains the enzymatic capability to synthesize aniline. In this specific embodiment it can be useful to increase the synthesis or accumulation of an aniline pathway product to, for example, drive aniline pathway reactions toward aniline production. Increased synthesis or accumulation can be accomplished by, for example, overexpression of nucleic acids encoding one or more of the above-described aniline pathway enzymes or proteins. Over expression of the enzyme or enzymes and/or protein or proteins of the aniline pathway can occur, for example, through exogenous expression of the endogenous gene or genes, or through exogenous expression of the heterologous gene or genes. Therefore, naturally occurring organisms can be readily generated to be non-naturally occurring microbial organisms of the invention, for example, producing aniline, through overexpression of one, two, three, four, five, up to all nucleic acids encoding aniline biosynthetic pathway enzymes or proteins. In addition, a non-naturally occurring organism can be generated by mutagenesis of an endogenous gene that results in an increase in activity of an enzyme in the aniline biosynthetic pathway.

In particularly useful embodiments, exogenous expression of the encoding nucleic acids is employed. Exogenous expression confers the ability to custom tailor the expression and/or regulatory elements to the host and application to achieve a desired expression level that is controlled by the user. However, endogenous expression also can be utilized in other embodiments such as by removing a negative regulatory effector or induction of the gene's promoter when linked to an inducible promoter or other regulatory element. Thus, an endogenous gene having a naturally occurring inducible promoter can be up-regulated by providing the appropriate inducing agent, or the regulatory region of an endogenous gene can be engineered to incorporate an inducible regulatory element, thereby allowing the regulation of increased expression of an endogenous gene at a desired time. Similarly, an inducible promoter can be included as a regulatory element for an exogenous gene introduced into a non-naturally occurring microbial organism.

It is understood that, in methods of the invention, any of the one or more exogenous nucleic acids can be introduced into a microbial organism to produce a non-naturally occurring microbial organism of the invention. The nucleic acids can be introduced so as to confer, for example, an aniline biosynthetic pathway onto the microbial organism. Alternatively, encoding nucleic acids can be introduced to produce an intermediate microbial organism having the biosynthetic capability to catalyze some of the required reactions to confer aniline biosynthetic capability. For example, a non-naturally occurring microbial organism having an aniline biosynthetic pathway can comprise at least two exogenous nucleic acids encoding desired enzymes or proteins, such as the combination of aminodoxychorismate synthase and aminodeoxychorismate lyase, or aminodeoxychorismate lyase and 4-aminobenzoate carboxylase, or aminodeoxychorismate synthase and 4-amionbenzoate carboxylase, and the like. Thus, it is understood that any combination of two or more enzymes or proteins of a biosynthetic pathway can be included in a non-naturally occurring microbial organism of the invention. Similarly, it is understood that any combination of three or more enzymes or proteins of a biosynthetic pathway can be included in a non-naturally occurring microbial organism of the invention, and so forth, as desired, so long as the combination of enzymes and/or proteins of the desired biosynthetic pathway results in production of the corresponding desired product. Similarly, any combination of four, or more enzymes or proteins of a biosynthetic pathway as disclosed herein can be included in a non-naturally occurring microbial organism of the invention, as desired, so long as the combination of enzymes and/or proteins of the desired biosynthetic pathway results in production of the corresponding desired product.

In addition to the biosynthesis of aniline as described herein, the non-naturally occurring microbial organisms and methods of the invention also can be utilized in various combinations with each other and with other microbial organisms and methods well known in the art to achieve product biosynthesis by other routes. For example, one alternative to produce aniline other than use of the aniline producers is through addition of another microbial organism capable of converting an aniline pathway intermediate to aniline. One such procedure includes, for example, the fermentation of a microbial organism that produces an aniline pathway intermediate. The aniline pathway intermediate can then be used as a substrate for a second microbial organism that converts the aniline pathway intermediate to aniline. The aniline pathway intermediate can be added directly to another culture of the second organism or the original culture of the aniline pathway intermediate producers can be depleted of these microbial organisms by, for example, cell separation, and then subsequent addition of the second organism to the fermentation broth can be utilized to produce the final product without intermediate purification steps.

In other embodiments, the non-naturally occurring microbial organisms and methods of the invention can be assembled in a wide variety of subpathways to achieve biosynthesis of, for example, aniline. In these embodiments, biosynthetic pathways for a desired product of the invention can be segregated into different microbial organisms, and the different microbial organisms can be co-cultured to produce the final product. In such a biosynthetic scheme, the product of one microbial organism is the substrate for a second microbial organism until the final product is synthesized. For example, the biosynthesis of aniline can be accomplished by constructing a microbial organism that contains biosynthetic pathways for conversion of one pathway intermediate to another pathway intermediate or the product. Alternatively, aniline also can be biosynthetically produced from microbial organisms through co-culture or co-fermentation using two organisms in the same vessel, where the first microbial organism produces an aniline intermediate and the second microbial organism converts the intermediate to aniline.

Given the teachings and guidance provided herein, those skilled in the art will understand that a wide variety of combinations and permutations exist for the non-naturally occurring microbial organisms and methods of the invention together with other microbial organisms, with the co-culture of other non-naturally occurring microbial organisms having subpathways and with combinations of other chemical and/or biochemical procedures well known in the art to produce aniline.

Sources of encoding nucleic acids for an aniline pathway enzyme or protein can include, for example, any species where the encoded gene product is capable of catalyzing the referenced reaction. Such species include both prokaryotic and eukaryotic organisms including, but not limited to, bacteria, including archaea and eubacteria, and eukaryotes, including yeast, plant, insect, animal, and mammal, including human. Exemplary species for such sources include, for example, *Escherichia coli*, *Sedimentibacter hydroxybenzoicus*, and *Bacillus subtilis*, as well as other exemplary species disclosed herein or available as source organisms for corresponding genes. However, with the complete genome sequence available for now more than 550 species (with more than half of these available on public databases such as the NCBI), including 395 microorganism genomes and a variety of yeast, fungi, plant, and mammalian genomes, the identification of genes encoding the requisite aniline biosynthetic activity for one or more genes in related or distant species, including for example, homologues, orthologs, paralogs and nonorthologous gene displacements of known genes, and the interchange of genetic alterations between organisms is routine and well known in the art. Accordingly, the metabolic alterations allowing biosynthesis of aniline described herein with reference to a particular organism such as *E. coli* can be readily applied to other microorganisms, including prokaryotic and eukaryotic organisms alike. Given the teachings and guidance provided herein, those skilled in the art will know that a metabolic alteration exemplified in one organism can be applied equally to other organisms.

In some instances, such as when an alternative aniline biosynthetic pathway exists in an unrelated species, aniline biosynthesis can be conferred onto the host species by, for example, exogenous expression of a paralog or paralogs from the unrelated species that catalyzes a similar, yet non-identical metabolic reaction to replace the referenced reaction. Because certain differences among metabolic networks exist between different organisms, those skilled in the art will understand that the actual gene usage between different organisms may differ. However, given the teachings and guidance provided herein, those skilled in the art also will understand that the teachings and methods of the invention can be applied to all microbial organisms using the cognate metabolic alterations to those exemplified herein to construct a microbial organism in a species of interest that will synthesize aniline.

Methods for constructing and testing the expression levels of a non-naturally occurring aniline-producing host can be performed, for example, by recombinant and detection methods well known in the art. Such methods can be found described in, for example, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Third Ed., Cold Spring Harbor Laboratory, New

York (2001); and Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, MD (1999).

Exogenous nucleic acid sequences involved in a pathway for production of aniline can be introduced stably or transiently into a host cell using techniques well known in the art including, but not limited to, conjugation, electroporation, chemical transformation, transduction, transfection, and ultrasound transformation. For exogenous expression in *E. co*li or other prokaryotic cells, some nucleic acid sequences in the genes or cDNAs of eukaryotic nucleic acids can encode targeting signals such as an N-terminal mitochondrial or other targeting signal, which can be removed before transformation into prokaryotic host cells, if desired. For example, removal of a mitochondrial leader sequence led to increased expression in *E. coli* (Hoffmeister et al., *J. Biol. Chem.* 280:4329-4338 (2005)). For exogenous expression in yeast or other eukaryotic cells, genes can be expressed in the cytosol without the addition of leader sequence, or can be targeted to mitochondrion or other organelles, or targeted for secretion, by the addition of a suitable targeting sequence such as a mitochondrial targeting or secretion signal suitable for the host cells. Thus, it is understood that appropriate modifications to a nucleic acid sequence to remove or include a targeting sequence can be incorporated into an exogenous nucleic acid sequence to impart desirable properties. Furthermore, genes can be subjected to codon optimization with techniques well known in the art to achieve optimized expression of the proteins.

An expression vector or vectors can be constructed to include one or more aniline biosynthetic pathway encoding nucleic acids as exemplified herein operably linked to expression control sequences functional in the host organism. Expression vectors applicable for use in the microbial host organisms of the invention include, for example, plasmids, phage vectors, viral vectors, episomes and artificial chromosomes, including vectors and selection sequences or markers operable for stable integration into a host chromosome. Additionally, the expression vectors can include one or more selectable marker genes and appropriate expression control sequences. Selectable marker genes also can be included that, for example, provide resistance to antibiotics or toxins, complement auxotrophic deficiencies, or supply critical nutrients not in the culture media. Expression control sequences can include constitutive and inducible promoters, transcription enhancers, transcription terminators, and the like which are well known in the art. When two or more exogenous encoding nucleic acids are to be co-expressed, both nucleic acids can be inserted, for example, into a single expression vector or in separate expression vectors. For single vector expression, the encoding nucleic acids can be operationally linked to one

common expression control sequence or linked to different expression control sequences, such as one inducible promoter and one constitutive promoter. The transformation of exogenous nucleic acid sequences involved in a metabolic or synthetic pathway can be confirmed using methods well known in the art. Such methods include, for example, nucleic acid analysis such as Northern blots or polymerase chain reaction (PCR) amplification of mRNA, or immunoblotting for expression of gene products, or other suitable analytical methods to test the expression of an introduced nucleic acid sequence or its corresponding gene product. It is understood by those skilled in the art that the exogenous nucleic acid is expressed in a sufficient amount to produce the desired product, and it is further understood that expression levels can be optimized to obtain sufficient expression using methods well known in the art and as disclosed herein.

In some embodiments, the present invention provides a method for producing aniline that includes culturing a non-naturally occurring microbial organism having an aniline pathway in which at least one exogenous nucleic acid encoding an aniline pathway enzyme is expressed in a sufficient amount to produce aniline, under conditions and for a sufficient period of time to produce aniline. The aniline pathway includes an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase. In some embodiments, the pathway further includes a DAHP synthase. In some embodiments, the pathway further includes a 3-dehydroquinate synthase. A method for producing aniline, includes culturing the non-naturally occurring microbial organism under conditions and for a sufficient period of time to produce aniline. Moreover, the non-naturally occurring microbial organism can be cultured in a substantially anaerobic culture medium.

Methods of the invention can include culturing a microbial organism having two exogenous nucleic acids each encoding an aniline pathway enzyme. In some embodiments, the cultured microbial organism can include three exogenous nucleic acids each encoding an aniline pathway enzyme. For example, the three exogenous nucleic acids can encode an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylase. In some embodiments, the cultured microbial organism can include four exogenous nucleic acids each encoding an aniline pathway enzyme. For example, the four exogenous nucleic acids can encode a DAHP synthase, an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase.

In still further embodiments, the cultured microbial organism can include five exogenous nucleic acids each encoding an aniline pathway enzyme. For example, the five exogenous nucleic acids

can encode a 3-dehydroquinate synthase, a DAHP synthase, an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase.

Any of the cultured organisms described above can have at least one exogenous nucleic acid that is a heterologous nucleic acid.

In some embodiments, the present invention provides a method for producing aniline, that includes culturing a non-naturally occurring microbial organism having an aniline pathway in which at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline, under conditions and for a sufficient period of time to produce aniline. In some embodiments, the aniline pathway includes an anthranilate synthase and an anthranilate decarboxylase. In some embodiments, such an organism can further include a DAHP synthase. In some embodiments, such an organism can further include a 3-dehydroquinate synthase, 3-dehydroquinate dehydratase, a shikimate dehydrogenase or a quinate/shikimate dehydrogenase, a shikimate kinase, a 3-phosphoshikimate-1-carboxyvinyltransferase, and a chorismate synthase. In some embodiments, the cultured non-naturally occurring microbial organism is cultured in a substantially anaerobic culture medium.

In some embodiments, the above cultured microbial organism can include two exogenous nucleic acids each encoding an aniline pathway enzyme. For example, the two exogenous nucleic acids can encode an anthranilate synthase and an anthranilate decarboxylase. In some embodiments, the cultured microbial organism can include three exogenous nucleic acids each encoding an aniline pathway enzyme. For example, the three exogenous nucleic acids encode a DAHP synthase, an anthranilate synthase and an anthranilate decarboxylase. In still further embodiments, the cultured microbial organism can include four exogenous nucleic acids each encoding an aniline pathway enzyme. For example, the four exogenous nucleic acids encode a 3-dehydroquinate synthase, a DAHP synthase, an anthranilate synthase and an anthranilate decarboxylase. Any of the at least one exogenous nucleic acids can be provided as a heterologous nucleic acid.

Suitable purification and/or assays to test for the production of aniline can be performed using well known methods. Suitable replicates such as triplicate cultures can be grown for each engineered strain to be tested. For example, product and byproduct formation in the engineered production host can be monitored. The final product and intermediates, and other organic compounds, can be analyzed by methods such as HPLC (High Performance Liquid Chromatography), GC-MS (Gas Chromatography-Mass Spectroscopy) and LC-MS (Liquid

Chromatography-Mass Spectroscopy) or other suitable analytical methods using routine procedures well known in the art. The release of product in the fermentation broth can also be tested with the culture supernatant. Byproducts and residual glucose can be quantified by HPLC using, for example, a refractive index detector for glucose and alcohols, and a UV detector for organic acids (Lin et al., *Biotechnol. Bioeng.* 90:775-779 (2005)), or other suitable assay and detection methods well known in the art. The individual enzyme or protein activities from the exogenous DNA sequences can also be assayed using methods well known in the art (McCullough et al., *J. Am. Chem. Soc.* 79:628-630 (1957)).

Aniline can be separated from other components in the culture using a variety of methods well known in the art. Such separation methods include, for example, extraction procedures as well as methods that include continuous liquid-liquid extraction, pervaporation, membrane filtration, membrane separation, reverse osmosis, electrodialysis, distillation, crystallization, centrifugation, extractive filtration, ion exchange chromatography, size exclusion chromatography, adsorption chromatography, and ultrafiltration. All of the above methods are well known in the art.

Any of the non-naturally occurring microbial organisms described herein can be cultured to produce and/or secrete the biosynthetic products of the invention. For example, the aniline producers can be cultured for the biosynthetic production of aniline.

For the production of aniline, the recombinant strains are cultured in a medium with carbon source and other essential nutrients. It is highly desirable to maintain anaerobic conditions in the fermenter to reduce the cost of the overall process. Such conditions can be obtained, for example, by first sparging the medium with nitrogen and then sealing the flasks with a septum and crimp-cap. For strains where growth is not observed anaerobically, microaerobic conditions can be applied by perforating the septum with a small hole for limited aeration. Exemplary anaerobic conditions have been described previously and are well-known in the art. Exemplary aerobic and anaerobic conditions are described, for example, in United State Patent application serial No. 11/891,602, filed August 10, 2007. Fermentations can be performed in a batch, fed-batch or continuous manner, as disclosed herein.

If desired, the pH of the medium can be maintained at a desired pH, in particular neutral pH, such as a pH of around 7 by addition of a base, such as NaOH or other bases, or acid, as needed to maintain the culture medium at a desirable pH. The growth rate can be determined by

measuring optical density using a spectrophotometer (600 nm), and the glucose uptake rate by monitoring carbon source depletion over time.

The growth medium can include, for example, any carbohydrate source which can supply a source of carbon to the non-naturally occurring microorganism. Such sources include, for example, sugars such as glucose, xylose, arabinose, galactose, mannose, fructose, sucrose and starch. Other sources of carbohydrates include, for example, renewable feedstocks and biomass. Exemplary types of biomasses that can be used as feedstocks in the methods of the invention include cellulosic biomass, hemicellulosic biomass and lignin feedstocks or portions of feedstocks. Such biomass feedstocks contain, for example, carbohydrate substrates useful as carbon sources such as glucose, xylose, arabinose, galactose, mannose, fructose and starch. Given the teachings and guidance provided herein, those skilled in the art will understand that renewable feedstocks and biomass other than those exemplified above also can be used for culturing the microbial organisms of the invention for the production of aniline.

In addition to renewable feedstocks such as those exemplified above, the aniline microbial organisms of the invention also can be modified for growth on syngas as its source of carbon. In this specific embodiment, one or more proteins or enzymes are expressed in the aniline producing organisms to provide a metabolic pathway for utilization of syngas or other gaseous carbon source.

Synthesis gas, also known as syngas or producer gas, is the major product of gasification of coal and of carbonaceous materials such as biomass materials, including agricultural crops and residues. Syngas is a mixture primarily of H_2 and CO and can be obtained from the gasification of any organic feedstock, including but not limited to coal, coal oil, natural gas, biomass, and waste organic matter. Gasification is generally carried out under a high fuel to oxygen ratio. Although largely H_2 and CO, syngas can also include CO_2 and other gases in smaller quantities. Thus, synthesis gas provides a cost effective source of gaseous carbon such as CO and, additionally, CO_2.

The Wood-Ljungdahl pathway catalyzes the conversion of CO and H_2 to acetyl-CoA and other products such as acetate. Organisms capable of utilizing CO and syngas also generally have the capability of utilizing CO_2 and CO_2/H_2 mixtures through the same basic set of enzymes and transformations encompassed by the Wood-Ljungdahl pathway. H_2-dependent conversion of CO_2 to acetate by microorganisms was recognized long before it was revealed that CO also could be used by the same organisms and that the same pathways were involved. Many

acetogens have been shown to grow in the presence of CO_2 and produce compounds such as acetate as long as hydrogen is present to supply the necessary reducing equivalents (see for example, Drake, *Acetogenesis*, pp. 3-60 Chapman and Hall, New York, (1994)). This can be summarized by the following equation:

$$2\ CO_2 + 4\ H_2 + n\ ADP + n\ Pi \rightarrow CH_3COOH + 2\ H_2O + n\ ATP$$

Hence, non-naturally occurring microorganisms possessing the Wood-Ljungdahl pathway can utilize CO_2 and H_2 mixtures as well for the production of acetyl-CoA and other desired products.

The Wood-Ljungdahl pathway is well known in the art and consists of 12 reactions which can be separated into two branches: (1) methyl branch and (2) carbonyl branch. The methyl branch converts syngas to methyl-tetrahydrofolate (methyl-THF) whereas the carbonyl branch converts methyl-THF to acetyl-CoA. The reactions in the methyl branch are catalyzed in order by the following enzymes or proteins: ferredoxin oxidoreductase, formate dehydrogenase, formyltetrahydrofolate synthetase, methenyltetrahydrofolate cyclodehydratase, methylenetetrahydrofolate dehydrogenase and methylenetetrahydrofolate reductase. The reactions in the carbonyl branch are catalyzed in order by the following enzymes or proteins: methyltetrahydrofolate:corrinoid protein methyltransferase (for example, AcsE), corrinoid iron-sulfur protein, nickel-protein assembly protein (for example, AcsF), ferredoxin, acetyl-CoA synthase, carbon monoxide dehydrogenase and nickel-protein assembly protein (for example, CooC). Following the teachings and guidance provided herein for introducing a sufficient number of encoding nucleic acids to generate an aniline pathway, those skilled in the art will understand that the same engineering design also can be performed with respect to introducing at least the nucleic acids encoding the Wood-Ljungdahl enzymes or proteins absent in the host organism. Therefore, introduction of one or more encoding nucleic acids into the microbial organisms of the invention such that the modified organism contains the complete Wood-Ljungdahl pathway will confer syngas utilization ability.

Accordingly, given the teachings and guidance provided herein, those skilled in the art will understand that a non-naturally occurring microbial organism can be produced that secretes the biosynthesized compounds of the invention when grown on a carbon source such as a carbohydrate. Such compounds include, for example, aniline and any of the intermediate metabolites in the aniline pathway. All that is required is to engineer in one or more of the required enzyme or protein activities to achieve biosynthesis of the desired compound or intermediate including, for example, inclusion of some or all of the aniline biosynthetic

pathways. Accordingly, the invention provides a non-naturally occurring microbial organism that produces and/or secretes aniline when grown on a carbohydrate or other carbon source and produces and/or secretes any of the intermediate metabolites shown in the aniline pathway when grown on a carbohydrate or other carbon source. The aniline producing microbial organisms of the invention can initiate synthesis from an intermediate, for example, chorismate, anthranilate, 4-amino-4-deoxychorismate, or *p*-aminobenzoate.

The non-naturally occurring microbial organisms of the invention are constructed using methods well known in the art as exemplified herein to exogenously express at least one nucleic acid encoding an aniline pathway enzyme or protein in sufficient amounts to produce aniline. It is understood that the microbial organisms of the invention are cultured under conditions sufficient to produce aniline. Following the teachings and guidance provided herein, the non-naturally occurring microbial organisms of the invention can achieve biosynthesis of aniline resulting in intracellular concentrations between about 0.1-200 mM or more. Generally, the intracellular concentration of aniline is between about 3-150 mM, particularly between about 5-125 mM and more particularly between about 8-100 mM, including about 10 mM, 20 mM, 50 mM, 80 mM, or more. Intracellular concentrations between and above each of these exemplary ranges also can be achieved from the non-naturally occurring microbial organisms of the invention.

In some embodiments, culture conditions include anaerobic or substantially anaerobic growth or maintenance conditions. Exemplary anaerobic conditions have been described previously and are well known in the art. Exemplary anaerobic conditions for fermentation processes are described herein and are described, for example, in U.S. publication 2009/0047719, filed August 10, 2007. Any of these conditions can be employed with the non-naturally occurring microbial organisms as well as other anaerobic conditions well known in the art. Under such anaerobic conditions, the aniline producers can synthesize aniline at intracellular concentrations of 5-10 mM or more as well as all other concentrations exemplified herein. It is understood that, even though the above description refers to intracellular concentrations, aniline producing microbial organisms can produce aniline intracellularly and/or secrete the product into the culture medium.

In addition to the culturing and fermentation conditions disclosed herein, growth condition for achieving biosynthesis of aniline can include the addition of an osmoprotectant to the culturing conditions. In certain embodiments, the non-naturally occurring microbial organisms of the invention can be sustained, cultured or fermented as described herein in the presence of an osmoprotectant. Briefly, an osmoprotectant refers to a compound that acts as an osmolyte and helps a microbial organism as described herein survive osmotic stress. Osmoprotectants

include, but are not limited to, betaines, amino acids, and the sugar trehalose. Non-limiting examples of such are glycine betaine, praline betaine, dimethylthetin, dimethylslfoniopropionate, 3-dimethylsulfonio-2-methylproprionate, pipecolic acid, dimethylsulfonioacetate, choline, L-carnitine and ectoine. In one aspect, the osmoprotectant is glycine betaine. It is understood to one of ordinary skill in the art that the amount and type of osmoprotectant suitable for protecting a microbial organism described herein from osmotic stress will depend on the microbial organism used. The amount of osmoprotectant in the culturing conditions can be, for example, no more than about 0.1 mM, no more than about 0.5 mM, no more than about 1.0 mM, no more than about 1.5 mM, no more than about 2.0 mM, no more than about 2.5 mM, no more than about 3.0 mM, no more than about 5.0 mM, no more than about 7.0 mM, no more than about 10mM, no more than about 50mM, no more than about 100mM or no more than about 500mM.

The culture conditions can include, for example, liquid culture procedures as well as fermentation and other large scale culture procedures. As described herein, particularly useful yields of the biosynthetic products of the invention can be obtained under anaerobic or substantially anaerobic culture conditions.

As described herein, one exemplary growth condition for achieving biosynthesis of aniline includes anaerobic culture or fermentation conditions. In certain embodiments, the non-naturally occurring microbial organisms of the invention can be sustained, cultured or fermented under anaerobic or substantially anaerobic conditions. Briefly, anaerobic conditions refer to an environment devoid of oxygen. Substantially anaerobic conditions include, for example, a culture, batch fermentation or continuous fermentation such that the dissolved oxygen concentration in the medium remains between 0 and 10% of saturation. Substantially anaerobic conditions also includes growing or resting cells in liquid medium or on solid agar inside a sealed chamber maintained with an atmosphere of less than 1% oxygen. The percent of oxygen can be maintained by, for example, sparging the culture with an N_2/CO_2 mixture or other suitable non-oxygen gas or gases.

The culture conditions described herein can be scaled up and grown continuously for manufacturing of aniline. Exemplary growth procedures include, for example, fed-batch fermentation and batch separation; fed-batch fermentation and continuous separation, or continuous fermentation and continuous separation. All of these processes are well known in the art. Fermentation procedures are particularly useful for the biosynthetic production of commercial quantities of aniline. Generally, and as with non-continuous culture procedures, the

continuous and/or near-continuous production of aniline will include culturing a non-naturally occurring aniline producing organism of the invention in sufficient nutrients and medium to sustain and/or nearly sustain growth in an exponential phase. Continuous culture under such conditions can include, for example, growth for 1 day, 2, 3, 4, 5, 6 or 7 days or more. Additionally, continuous culture can include longer time periods of 1 week, 2, 3, 4 or 5 or more weeks and up to several months. Alternatively, organisms of the invention can be cultured for hours, if suitable for a particular application. It is to be understood that the continuous and/or near-continuous culture conditions also can include all time intervals in between these exemplary periods. It is further understood that the time of culturing the microbial organism of the invention is for a sufficient period of time to produce a sufficient amount of product for a desired purpose.

Fermentation procedures are well known in the art. Briefly, fermentation for the biosynthetic production of aniline can be utilized in, for example, fed-batch fermentation and batch separation; fed-batch fermentation and continuous separation, or continuous fermentation and continuous separation. Examples of batch and continuous fermentation procedures are well known in the art.

In addition to the above fermentation procedures using the aniline producers of the invention for continuous production of substantial quantities of aniline, the aniline producers also can be, for example, simultaneously subjected to chemical synthesis procedures to convert the product to other compounds or the product can be separated from the fermentation culture and sequentially subjected to chemical conversion to convert the product to other compounds, if desired.

In some embodiments, methods for producing aniline include a step of isolating aniline from the fermentation broth. This can be achieved by means of standard extraction, distillation, salt crystallization techniques, and combinations of these techniques and those described above. For a basic product such as aniline, a salt crystallization can include the formation of an acid salt of a Bronsted or Lewis acid. Exemplary acid salts include, without limitation, acetate, aspartate, benzoate, bicarbonate, carbonate, bisulfate, sulfate, chloride, bromide, benzene sulfonate, methyl sulfonate, phosphate, biphosphate, lactate, maleate, malate, malonate, fumarate, lactate, tartrate, borate, camsylate, citrate, edisylate, esylate, formate, fumarate, gluceptate, glucuronate, gluconate oxalate, palmitate, pamoate, saccharate, stearate, succinate, tartrate, tosylate and trifluoroacetate salts.

To generate better producers, metabolic modeling can be utilized to optimize growth conditions. Modeling can also be used to design gene knockouts that additionally optimize utilization of the pathway (see, for example, U.S. patent publications US 2002/0012939, US 2003/0224363, US 2004/0029149, US 2004/0072723, US 2003/0059792, US 2002/0168654 and US 2004/0009466, and U.S. Patent No. 7,127,379). Modeling analysis allows reliable predictions of the effects on cell growth of shifting the metabolism towards more efficient production of aniline.

One computational method for identifying and designing metabolic alterations favoring biosynthesis of a desired product is the OptKnock computational framework (Burgard et al., *Biotechnol. Bioeng.* 84:647-657 (2003)). OptKnock is a metabolic modeling and simulation program that suggests gene deletion or disruption strategies that result in genetically stable microorganisms which overproduce the target product. Specifically, the framework examines the complete metabolic and/or biochemical network of a microorganism in order to suggest genetic manipulations that force the desired biochemical to become an obligatory byproduct of cell growth. By coupling biochemical production with cell growth through strategically placed gene deletions or other functional gene disruption, the growth selection pressures imposed on the engineered strains after long periods of time in a bioreactor lead to improvements in performance as a result of the compulsory growth-coupled biochemical production. Lastly, when gene deletions are constructed there is a negligible possibility of the designed strains reverting to their wild-type states because the genes selected by OptKnock are to be completely removed from the genome. Therefore, this computational methodology can be used to either identify alternative pathways that lead to biosynthesis of a desired product or used in connection with the non-naturally occurring microbial organisms for further optimization of biosynthesis of a desired product.

Briefly, OptKnock is a term used herein to refer to a computational method and system for modeling cellular metabolism. The OptKnock program relates to a framework of models and methods that incorporate particular constraints into flux balance analysis (FBA) models. These constraints include, for example, qualitative kinetic information, qualitative regulatory information, and/or DNA microarray experimental data. OptKnock also computes solutions to various metabolic problems by, for example, tightening the flux boundaries derived through flux balance models and subsequently probing the performance limits of metabolic networks in the presence of gene additions or deletions. OptKnock computational framework allows the construction of model formulations that allow an effective query of the performance limits of metabolic networks and provides methods for solving the resulting mixed-integer linear

programming problems. The metabolic modeling and simulation methods referred to herein as OptKnock are described in, for example, U.S. publication 2002/0168654, filed January 10, 2002, in International Patent No. PCT/US02/00660, filed January 10, 2002, and U.S. publication 2009/0047719, filed August 10, 2007.

Another computational method for identifying and designing metabolic alterations favoring biosynthetic production of a product is a metabolic modeling and simulation system termed SimPheny®. This computational method and system is described in, for example, U.S. publication 2003/0233218, filed June 14, 2002, and in International Patent Application No. PCT/US03/18838, filed June 13, 2003. SimPheny® is a computational system that can be used to produce a network model *in silico* and to simulate the flux of mass, energy or charge through the chemical reactions of a biological system to define a solution space that contains any and all possible functionalities of the chemical reactions in the system, thereby determining a range of allowed activities for the biological system. This approach is referred to as constraints-based modeling because the solution space is defined by constraints such as the known stoichiometry of the included reactions as well as reaction thermodynamic and capacity constraints associated with maximum fluxes through reactions. The space defined by these constraints can be interrogated to determine the phenotypic capabilities and behavior of the biological system or of its biochemical components.

These computational approaches are consistent with biological realities because biological systems are flexible and can reach the same result in many different ways. Biological systems are designed through evolutionary mechanisms that have been restricted by fundamental constraints that all living systems must face. Therefore, constraints-based modeling strategy embraces these general realities. Further, the ability to continuously impose further restrictions on a network model via the tightening of constraints results in a reduction in the size of the solution space, thereby enhancing the precision with which physiological performance or phenotype can be predicted.

Given the teachings and guidance provided herein, those skilled in the art will be able to apply various computational frameworks for metabolic modeling and simulation to design and implement biosynthesis of a desired compound in host microbial organisms. Such metabolic modeling and simulation methods include, for example, the computational systems exemplified above as SimPheny® and OptKnock. For illustration of the invention, some methods are described herein with reference to the OptKnock computation framework for modeling and simulation. Those skilled in the art will know how to apply the identification, design and

implementation of the metabolic alterations using OptKnock to any of such other metabolic modeling and simulation computational frameworks and methods well known in the art.

The methods described above will provide one set of metabolic reactions to disrupt. Elimination of each reaction within the set or metabolic modification can result in a desired product as an obligatory product during the growth phase of the organism. Because the reactions are known, a solution to the bilevel OptKnock problem also will provide the associated gene or genes encoding one or more enzymes that catalyze each reaction within the set of reactions. Identification of a set of reactions and their corresponding genes encoding the enzymes participating in each reaction is generally an automated process, accomplished through correlation of the reactions with a reaction database having a relationship between enzymes and encoding genes.

Once identified, the set of reactions that are to be disrupted in order to achieve production of a desired product are implemented in the target cell or organism by functional disruption of at least one gene encoding each metabolic reaction within the set. One particularly useful means to achieve functional disruption of the reaction set is by deletion of each encoding gene. However, in some instances, it can be beneficial to disrupt the reaction by other genetic aberrations including, for example, mutation, deletion of regulatory regions such as promoters or cis binding sites for regulatory factors, or by truncation of the coding sequence at any of a number of locations. These latter aberrations, resulting in less than total deletion of the gene set can be useful, for example, when rapid assessments of the coupling of a product are desired or when genetic reversion is less likely to occur.

To identify additional productive solutions to the above described bilevel OptKnock problem which lead to further sets of reactions to disrupt or metabolic modifications that can result in the biosynthesis, including growth-coupled biosynthesis of a desired product, an optimization method, termed integer cuts, can be implemented. This method proceeds by iteratively solving the OptKnock problem exemplified above with the incorporation of an additional constraint referred to as an integer cut at each iteration. Integer cut constraints effectively prevent the solution procedure from choosing the exact same set of reactions identified in any previous iteration that obligatorily couples product biosynthesis to growth. For example, if a previously identified growth-coupled metabolic modification specifies reactions 1, 2, and 3 for disruption, then the following constraint prevents the same reactions from being simultaneously considered in subsequent solutions. The integer cut method is well known in the art and can be found described in, for example, Burgard et al., *Biotechnol. Prog.* 17:791-797 (2001). As with all

methods described herein with reference to their use in combination with the OptKnock computational framework for metabolic modeling and simulation, the integer cut method of reducing redundancy in iterative computational analysis also can be applied with other computational frameworks well known in the art including, for example, SimPheny®.

The methods exemplified herein allow the construction of cells and organisms that biosynthetically produce a desired product, including the obligatory coupling of production of a target biochemical product to growth of the cell or organism engineered to harbor the identified genetic alterations. Therefore, the computational methods described herein allow the identification and implementation of metabolic modifications that are identified by an *in silico* method selected from OptKnock or SimPheny®. The set of metabolic modifications can include, for example, addition of one or more biosynthetic pathway enzymes and/or functional disruption of one or more metabolic reactions including, for example, disruption by gene deletion.

As discussed above, the OptKnock methodology was developed on the premise that mutant microbial networks can be evolved towards their computationally predicted maximum-growth phenotypes when subjected to long periods of growth selection. In other words, the approach leverages an organism's ability to self-optimize under selective pressures. The OptKnock framework allows for the exhaustive enumeration of gene deletion combinations that force a coupling between biochemical production and cell growth based on network stoichiometry. The identification of optimal gene/reaction knockouts requires the solution of a bilevel optimization problem that chooses the set of active reactions such that an optimal growth solution for the resulting network overproduces the biochemical of interest (Burgard et al., *Biotechnol. Bioeng.* 84:647-657 (2003)).

An *in silico* stoichiometric model of *E. coli* metabolism can be employed to identify essential genes for metabolic pathways as exemplified previously and described in, for example, U.S. patent publications US 2002/0012939, US 2003/0224363, US 2004/0029149, US 2004/0072723, US 2003/0059792, US 2002/0168654 and US 2004/0009466, and in U.S. Patent No. 7,127,379. As disclosed herein, the OptKnock mathematical framework can be applied to pinpoint gene deletions leading to the growth-coupled production of a desired product. Further, the solution of the bilevel OptKnock problem provides only one set of deletions. To enumerate all meaningful solutions, that is, all sets of knockouts leading to growth-coupled production formation, an optimization technique, termed integer cuts, can be implemented. This entails iteratively solving

the OptKnock problem with the incorporation of an additional constraint referred to as an integer cut at each iteration, as discussed above.

As disclosed herein, a nucleic acid encoding a desired activity of an aniline pathway can be introduced into a host organism. In some cases, it can be desirable to modify an activity of an aniline pathway enzyme or protein to increase production of aniline. For example, known mutations that increase the activity of a protein or enzyme can be introduced into an encoding nucleic acid molecule. Additionally, optimization methods can be applied to increase the activity of an enzyme or protein and/or decrease an inhibitory activity, for example, decrease the activity of a negative regulator.

One such optimization method is directed evolution. Directed evolution is a powerful approach that involves the introduction of mutations targeted to a specific gene in order to improve and/or alter the properties of an enzyme. Improved and/or altered enzymes can be identified through the development and implementation of sensitive high-throughput screening assays that allow the automated screening of many enzyme variants (for example, $>10^4$). Iterative rounds of mutagenesis and screening typically are performed to afford an enzyme with optimized properties. Computational algorithms that can help to identify areas of the gene for mutagenesis also have been developed and can significantly reduce the number of enzyme variants that need to be generated and screened. Numerous directed evolution technologies have been developed (for reviews, see Hibbert et al., *Biomol.Eng* 22:11-19 (2005); Huisman and Lalonde, *In* Biocatalysis in the pharmaceutical and biotechnology industries pgs. 717-742 (2007), Patel (ed.), CRC Press; Otten and Quax. *Biomol.Eng* 22:1-9 (2005).; and Sen et al., *Appl Biochem.Biotechnol* 143:212-223 (2007)) to be effective at creating diverse variant libraries, and these methods have been successfully applied to the improvement of a wide range of properties across many enzyme classes. Enzyme characteristics that have been improved and/or altered by directed evolution technologies include, for example: selectivity/specificity, for conversion of non-natural substrates; temperature stability, for robust high temperature processing; pH stability, for bioprocessing under lower or higher pH conditions; substrate or product tolerance, so that high product titers can be achieved; binding (K_m), including broadening substrate binding to include non-natural substrates; inhibition (K_i), to remove inhibition by products, substrates, or key intermediates; activity (kcat), to increases enzymatic reaction rates to achieve desired flux; expression levels, to increase protein yields and overall pathway flux; oxygen stability, for operation of air sensitive enzymes under aerobic conditions; and anaerobic activity, for operation of an aerobic enzyme in the absence of oxygen.

A number of exemplary methods have been developed for the mutagenesis and diversification of genes to target desired properties of specific enzymes. Such methods are well known to those skilled in the art. Any of these can be used to alter and/or optimize the activity of an aniline pathway enzyme or protein. Such methods include, but are not limited to EpPCR, which introduces random point mutations by reducing the fidelity of DNA polymerase in PCR reactions (Pritchard et al., *J Theor.Biol.* 234:497-509 (2005)); Error-prone Rolling Circle Amplification (epRCA), which is similar to epPCR except a whole circular plasmid is used as the template and random 6-mers with exonuclease resistant thiophosphate linkages on the last 2 nucleotides are used to amplify the plasmid followed by transformation into cells in which the plasmid is re-circularized at tandem repeats (Fujii et al., *Nucleic Acids Res.* 32:e145 (2004); and Fujii et al., *Nat. Protoc.* 1:2493-2497 (2006)); DNA or Family Shuffling, which typically involves digestion of two or more variant genes with nucleases such as Dnase I or EndoV to generate a pool of random fragments that are reassembled by cycles of annealing and extension in the presence of DNA polymerase to create a library of chimeric genes (Stemmer, *Proc Natl Acad Sci USA* 91:10747-10751 (1994); and Stemmer, *Nature* 370:389-391 (1994)); Staggered Extension (StEP), which entails template priming followed by repeated cycles of 2 step PCR with denaturation and very short duration of annealing/extension (as short as 5 sec) (Zhao et al., *Nat. Biotechnol.* 16:258-261 (1998)); Random Priming Recombination (RPR), in which random sequence primers are used to generate many short DNA fragments complementary to different segments of the template (Shao et al., *Nucleic Acids Res* 26:681-683 (1998)).

Additional methods include Heteroduplex Recombination, in which linearized plasmid DNA is used to form heteroduplexes that are repaired by mismatch repair (Volkov et al, *Nucleic Acids Res.* 27:e18 (1999); and Volkov et al., *Methods Enzymol.* 328:456-463 (2000)); Random Chimeragenesis on Transient Templates (RACHITT), which employs Dnase I fragmentation and size fractionation of single stranded DNA (ssDNA) (Coco et al., *Nat. Biotechnol.* 19:354-359 (2001)); Recombined Extension on Truncated templates (RETT), which entails template switching of unidirectionally growing strands from primers in the presence of unidirectional ssDNA fragments used as a pool of templates (Lee et al., *J. Molec. Catalysis* 26:119-129 (2003)); Degenerate Oligonucleotide Gene Shuffling (DOGS), in which degenerate primers are used to control recombination between molecules; (Bergquist and Gibbs, *Methods Mol.Biol* 352:191-204 (2007); Bergquist et al., *Biomol.Eng* 22:63-72 (2005); Gibbs et al., *Gene* 271:13-20 (2001)); Incremental Truncation for the Creation of Hybrid Enzymes (ITCHY), which creates a combinatorial library with 1 base pair deletions of a gene or gene fragment of interest (Ostermeier et al., *Proc. Natl. Acad. Sci. USA* 96:3562-3567 (1999); and Ostermeier et al., *Nat.*

Biotechnol. 17:1205-1209 (1999)); Thio-Incremental Truncation for the Creation of Hybrid Enzymes (THIO-ITCHY), which is similar to ITCHY except that phosphothioate dNTPs are used to generate truncations (Lutz et al., *Nucleic Acids Res* 29:E16 (2001)); SCRATCHY, which combines two methods for recombining genes, ITCHY and DNA shuffling (Lutz et al., *Proc. Natl. Acad. Sci. USA* 98:11248-11253 (2001)); Random Drift Mutagenesis (RNDM), in which mutations made via epPCR are followed by screening/selection for those retaining usable activity (Bergquist et al., *Biomol. Eng.* 22:63-72 (2005)); Sequence Saturation Mutagenesis (SeSaM), a random mutagenesis method that generates a pool of random length fragments using random incorporation of a phosphothioate nucleotide and cleavage, which is used as a template to extend in the presence of "universal" bases such as inosine, and replication of an inosine-containing complement gives random base incorporation and, consequently, mutagenesis (Wong et al., *Biotechnol. J.* 3:74-82 (2008); Wong et al., *Nucleic Acids Res.* 32:e26 (2004); and Wong et al., *Anal. Biochem.* 341:187-189 (2005)); Synthetic Shuffling, which uses overlapping oligonucleotides designed to encode "all genetic diversity in targets" and allows a very high diversity for the shuffled progeny (Ness et al., *Nat. Biotechnol.* 20:1251-1255 (2002)); Nucleotide Exchange and Excision Technology NexT, which exploits a combination of dUTP incorporation followed by treatment with uracil DNA glycosylase and then piperidine to perform endpoint DNA fragmentation (Muller et al., *Nucleic Acids Res.* 33:e117 (2005)).

Further methods include Sequence Homology-Independent Protein Recombination (SHIPREC), in which a linker is used to facilitate fusion between two distantly related or unrelated genes, and a range of chimeras is generated between the two genes, resulting in libraries of single-crossover hybrids (Sieber et al., *Nat. Biotechnol.* 19:456-460 (2001)); Gene Site Saturation Mutagenesis™ (GSSM™), in which the starting materials include a supercoiled double stranded DNA (dsDNA) plasmid containing an insert and two primers which are degenerate at the desired site of mutations (Kretz et al., *Methods Enzymol.* 388:3-11 (2004)); Combinatorial Cassette Mutagenesis (CCM), which involves the use of short oligonucleotide cassettes to replace limited regions with a large number of possible amino acid sequence alterations (Reidhaar-Olson et al. *Methods Enzymol.* 208:564-586 (1991); and Reidhaar-Olson et al. *Science* 241:53-57 (1988)); Combinatorial Multiple Cassette Mutagenesis (CMCM), which is essentially similar to CCM and uses epPCR at high mutation rate to identify hot spots and hot regions and then extension by CMCM to cover a defined region of protein sequence space (Reetz et al., *Angew. Chem. Int. Ed Engl.* 40:3589-3591 (2001)); the Mutator Strains technique, in which conditional *ts* mutator plasmids, utilizing the *mutD5* gene, which encodes a mutant subunit of DNA polymerase III, to allow increases of 20 to 4000-X in random and natural mutation frequency during selection and

block accumulation of deleterious mutations when selection is not required (Selifonova et al., *Appl. Environ. Microbiol.* 67:3645-3649 (2001)); Low et al., *J. Mol. Biol.* 260:359-3680 (1996)).

Additional exemplary methods include Look-Through Mutagenesis (LTM), which is a multidimensional mutagenesis method that assesses and optimizes combinatorial mutations of selected amino acids (Rajpal et al., *Proc. Natl. Acad. Sci. USA* 102:8466-8471 (2005)); Gene Reassembly, which is a DNA shuffling method that can be applied to multiple genes at one time or to create a large library of chimeras (multiple mutations) of a single gene (Tunable GeneReassembly™ (TGR™) Technology supplied by Verenium Corporation), *in Silico* Protein Design Automation (PDA), which is an optimization algorithm that anchors the structurally defined protein backbone possessing a particular fold, and searches sequence space for amino acid substitutions that can stabilize the fold and overall protein energetics, and generally works most effectively on proteins with known three-dimensional structures (Hayes et al., *Proc. Natl. Acad. Sci. USA* 99:15926-15931 (2002)); and Iterative Saturation Mutagenesis (ISM), which involves using knowledge of structure/function to choose a likely site for enzyme improvement, performing saturation mutagenesis at chosen site using a mutagenesis method such as Stratagene QuikChange (Stratagene; San Diego CA), screening/selecting for desired properties, and, using improved clone(s), starting over at another site and continue repeating until a desired activity is achieved (Reetz et al., *Nat. Protoc.* 2:891-903 (2007); and Reetz et al., *Angew. Chem. Int. Ed Engl.* 45:7745-7751 (2006)).

Any of the aforementioned methods for mutagenesis can be used alone or in any combination. Additionally, any one or combination of the directed evolution methods can be used in conjunction with adaptive evolution techniques, as described herein.

EXAMPLE I

Aniline Biosynthesis using p-aminobenzoate as the precursor

This Example describes the generation of a microbial organism capable of producing aniline using chorismate as the precursor.

Escherichia coli is used as a target organism to engineer the pathway using the enzymes aminodeoxychorismate synthase, aminodeoxychorismate lyase, and 4-aminobenzoate carboxylyase as shown in Figure 2. *E. coli* provides a good host for generating a non-naturally occurring microorganism capable of producing aniline. *E. coli* is amenable to genetic manipulation and is known to be capable of producing various products, like ethanol, acetic

acid, formic acid, lactic acid, and succinic acid, effectively under anaerobic or microaerobic conditions.

To generate an *E. coli* strain engineered to produce aniline, nucleic acids encoding the enzymes utilized in the disclosed pathway, as described previously, are expressed in *E. coli* to the desired extent using well known molecular biology techniques (see, for example, Sambrook, *supra*, 2001; Ausubel *supra*, 1999; Roberts et al., *supra*, 1989).

The native enzymes in *E. coli* can be modified or heterologous enzymes can be introduced to produce significant quantities of p-aminobenzoate. Further, 4-aminobenzoate carboxylyase activity can be incorporated into the strain by introducing the appropriate genes , such as *shdB, C* and *D* from *Sedimentibacter hydroxybenzoicus. The genes* are cloned into the pZE13 vector (Expressys, Ruelzheim, Germany) under the PA1/lacO promoter. The plasmid is transformed into the recombinant *E. coli* strain producing p-aminobenzoate to express the proteins and enzymes required for aniline synthesis from this metabolite.

The resulting genetically engineered organism is cultured in glucose containing medium following procedures well known in the art (see, for example, Sambrook et al., *supra*, 2001). The expression of the pathway genes is corroborated using methods well known in the art for determining polypeptide expression or enzymatic activity, including, for example, Northern blots, PCR amplification of mRNA, immunoblotting. Enzymatic activities of the expressed enzymes are confirmed using assays specific for the individually activities. The ability of the engineered *E. coli* strain to produce aniline is confirmed using HPLC, gas chromatography-mass spectrometry (GCMS) or liquid chromatography-mass spectrometry (LCMS).

Microbial strains engineered to have a functional aniline synthesis pathway are further augmented by optimization for efficient utilization of the pathway. Briefly, the engineered strain is assessed to determine whether any of the exogenous genes are expressed at a rate limiting level. Expression is increased for any enzymes expressed at low levels that can limit the flux through the pathway by, for example, introduction of additional gene copy numbers.

To generate better producers, metabolic modeling is utilized to optimize growth conditions. Modeling is also used to design gene knockouts that additionally optimize utilization of the pathway (see, for example, U.S. patent publications US 2002/0012939, US 2003/0224363, US 2004/0029149, US 2004/0072723, US 2003/0059792, US 2002/0168654 and US 2004/0009466, and in U.S. Patent No. 7,127,379). Modeling analysis allows reliable predictions of the effects on cell growth of shifting the metabolism towards more efficient production of aniline. One

modeling method is the bilevel optimization approach, OptKnock (Burgard et al., *Biotechnol. Bioengineer.* 84:647-657 (2003)), which is applied to select gene knockouts that collectively result in better production of aniline. Adaptive evolution also can be used to generate better producers of, for example, the intermediate, chorismate or the product, aniline. Adaptive evolution is performed to improve both growth and production characteristics (Fong and Palsson, *Nat. Genet.* 36:1056-1058 (2004); Alper et al., *Science* 314:1565-1568 (2006)). Based on the results, subsequent rounds of modeling, genetic engineering and adaptive evolution can be applied to the aniline producer to further increase production.

For large-scale production of aniline, the recombinant organism is cultured in a fermenter using a medium known in the art to support growth of the organism under anaerobic conditions. Fermentations are performed in either a batch, fed-batch or continuous manner. Anaerobic conditions are maintained by first sparging the medium with nitrogen and then sealing culture vessel (e.g., flasks can be sealed with a septum and crimp-cap). Microaerobic conditions also can be utilized by providing a small hole for limited aeration. The pH of the medium is maintained at a pH of 7 by addition of an acid, such as H_2SO_4. The growth rate is determined by measuring optical density using a spectrophotometer (600 nm), and the glucose uptake rate by monitoring carbon source depletion over time. Byproducts such as undesirable alcohols, organic acids, and residual glucose can be quantified by HPLC (Shimadzu) with an HPX-087 column (BioRad), using a refractive index detector for glucose and alcohols, and a UV detector for organic acids (Lin et al., *Biotechnol Bioeng.* 90:775-779 (2005)).

Throughout this application various publications have been referenced. The disclosures of these publications in their entireties, including GenBank and GI number publications, are hereby incorporated by reference in this application in order to more fully describe the state of the art to which this invention pertains. Although the invention has been described with reference to the examples provided above, it should be understood that various modifications can be made without departing from the spirit of the invention.

Although the invention has been described with reference to the disclosed embodiments, those skilled in the art will readily appreciate that the specific examples and studies detailed above are only illustrative of the invention. It should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

What is claimed is:

1. A non-naturally occurring microbial organism, comprising a microbial organism having an aniline pathway comprising at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline, said aniline pathway comprising a 4-aminobenzoate carboxylyase, an aminodeoxychorismate synthase, and an aminodeoxychorismate lyase.

2. The non-naturally occurring microbial organism of claim 1, further comprising a DAHP synthase.

3. The non-naturally occurring microbial organism of claim 1, further comprising an enzyme selected from a 3-dehydroquinate synthase, a 3-dehydroquinate dehydratase, a shikimate dehydrogenase or a quinate/shikimate dehydrogenase, a shikimate kinase, a 3-phosphoshikimate-1-carboxyvinyltransferase, and a chorismate synthase.

4. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises two exogenous nucleic acids each encoding an aniline pathway enzyme.

5. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises three exogenous nucleic acids each encoding an aniline pathway enzyme.

6. The non-naturally occurring microbial organism of claim 5, wherein said three exogenous nucleic acids encode an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase.

7. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises four exogenous nucleic acids each encoding an aniline pathway enzyme.

8. The non-naturally occurring microbial organism of claim 7, wherein said four exogenous nucleic acids encode a DAHP synthase, an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase.

9. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises five exogenous nucleic acids each encoding an aniline pathway enzyme.

10. The non-naturally occurring microbial organism of claim 9, wherein said five exogenous nucleic acids encode, a DAHP synthase, an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase and an enzyme selected

from a 3-dehydroquinate synthase, a 3-dehydroquinate dehydratase, a shikimate dehydrogenase or a quinate/shikimate dehydrogenase, a shikimate kinase, a 3-phosphoshikimate-1-carboxyvinyltransferase, and a chorismate synthase.

11. The non-naturally occurring microbial organism of claim 1, wherein said at least one exogenous nucleic acid is a heterologous nucleic acid.

12. The non-naturally occurring microbial organism of claim 1, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

13. A method for producing aniline, comprising culturing the non-naturally occurring microbial organism of claim 1 under conditions and for a sufficient period of time to produce aniline.

14. The method of claim 13 further comprising a DAHP synthase.

15. The method of claim 13, further comprising an enzyme selected from a 3-dehydroquinate synthase, a 3-dehydroquinate dehydratase, a shikimate dehydrogenase or a quinate/shikimate dehydrogenase, a shikimate kinase, a 3-phosphoshikimate-1-carboxyvinyltransferase, and a chorismate synthase.

16. The method of claim 13, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

17. The method of claim 13, wherein said microbial organism comprises two exogenous nucleic acids each encoding an aniline pathway enzyme.

18. The method of claim 13, wherein said microbial organism comprises three exogenous nucleic acids each encoding an aniline pathway enzyme.

19. The method of claim 18, wherein said three exogenous nucleic acids encode a 4-aminobenzoate carboxylyase, an aminodeoxychorismate synthase, and an aminodeoxychorismate lyase.

20. The method of claim 13, wherein said microbial organism comprises four exogenous nucleic acids each encoding an aniline pathway enzyme.

21. The method of claim 20, wherein said four exogenous nucleic acids encode a DAHP synthase, an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, and a 4-aminobenzoate carboxylyase.

22. The method of claim 13, wherein said microbial organism comprises five exogenous nucleic acids each encoding an aniline pathway enzyme.

23. The method of claim 22, wherein said five exogenous nucleic acids encode a DAHP synthase, an aminodeoxychorismate synthase, an aminodeoxychorismate lyase, a 4-aminobenzoate carboxylyase and an enzyme selected from a 3-dehydroquinate synthase, a 3-dehydroquinate dehydratase, a shikimate dehydrogenase or a quinate/shikimate dehydrogenase, a shikimate kinase, a 3-phosphoshikimate-1-carboxyvinyltransferase, and a chorismate synthase.

24. The method of claim 13, wherein said at least one exogenous nucleic acid is a heterologous nucleic acid.

25. A non-naturally occurring microbial organism, comprising a microbial organism having an aniline pathway comprising at least one exogenous nucleic acid encoding an aniline pathway enzyme expressed in a sufficient amount to produce aniline, said aniline pathway comprising an anthranilate synthase and an anthranilate decarboxylase.

26. The non-naturally occurring microbial organism of claim 25, further comprising a DAHP synthase.

27. The non-naturally occurring microbial organism of claim 25, further comprising an enzyme selected from a 3-dehydroquinate synthase, a 3-dehydroquinate dehydratase, a shikimate dehydrogenase or a quinate/shikimate dehydrogenase, a shikimate kinase, a 3-phosphoshikimate-1-carboxyvinyltransferase, and a chorismate synthase.

28. The non-naturally occurring microbial organism of claim 25, wherein said microbial organism comprises two exogenous nucleic acids each encoding an aniline pathway enzyme.

29. The non-naturally occurring microbial organism of claim 28, wherein said two exogenous nucleic acids encode an anthranilate synthase and an anthranilate decarboxylase.

30. The non-naturally occurring microbial organism of claim 25, wherein said microbial organism comprises three exogenous nucleic acids each encoding an aniline pathway enzyme.

31. The non-naturally occurring microbial organism of claim 30, wherein said three exogenous nucleic acids encode a DAHP synthase, an anthranilate synthase and an anthranilate decarboxylase.

32. The non-naturally occurring microbial organism of claim 25, wherein said microbial organism comprises four exogenous nucleic acids each encoding an aniline pathway enzyme.

33. The non-naturally occurring microbial organism of claim 32, wherein said four exogenous nucleic acids encode a 3-dehydroquinate synthase, a DAHP synthase, an anthranilate synthase and an anthranilate decarboxylase.

34. The non-naturally occurring microbial organism of claim 25, wherein said at least one exogenous nucleic acid is a heterologous nucleic acid.

35. The non-naturally occurring microbial organism of claim 25, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

36. A method for producing aniline, comprising culturing the non-naturally occurring microbial organism of claim 25 under conditions and for a sufficient period of time to produce aniline.

37. The method of claim 36 further comprising a DAHP synthase.

38. The method of claim 36, further comprising an enzyme selected from a 3-dehydroquinate synthase, a 3-dehydroquinate dehydratase, a shikimate dehydrogenase or a quinate/shikimate dehydrogenase, a shikimate kinase, a 3-phosphoshikimate-1-carboxyvinyltransferase, and a chorismate synthase.

39. The method of claim 36, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

40. The method of claim 36, wherein said microbial organism comprises two exogenous nucleic acids each encoding an aniline pathway enzyme.

41. The method of claim 40, wherein said two exogenous nucleic acids encode an anthranilate synthase and an anthranilate decarboxylase.

42. The method of claim 36, wherein said microbial organism comprises three exogenous nucleic acids each encoding an aniline pathway enzyme.

43. The method of claim 42, wherein said three exogenous nucleic acids encode a DAHP synthase, an anthranilate synthase and an anthranilate decarboxylase.

44. The method of claim 36, wherein said microbial organism comprises four exogenous nucleic acids each encoding an aniline pathway enzyme.

45. The method of claim 44, wherein said four exogenous nucleic acids encode a DAHP synthase, an anthranilate synthase, an anthranilate decarboxylase and an enzyme selected from a 3-dehydroquinate synthase, a 3-dehydroquinate dehydratase, a shikimate dehydrogenase or a quinate/shikimate dehydrogenase, a shikimate kinase, a 3-phosphoshikimate-1-carboxyvinyltransferase, and a chorismate synthase.

46. The method of claim 36, wherein said at least one exogenous nucleic acid is a heterologous nucleic acid.