

(51) International Patent Classification:
C12P 5/02 (2006.01)(21) International Application Number:
PCT/US2015/028296(22) International Filing Date:
29 April 2015 (29.04.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/986,669 30 April 2014 (30.04.2014) US(71) Applicants: **THE REGENTS OF THE UNIVERSITY OF CALIFORNIA** [US/US]; 1111 Franklin Street, 12th Floor, Oakland, California 94607 (US). **TOTAL MARKETING SERVICES** [FR/FR]; 24, Cours Michelet, 92800 Puteaux (FR).(72) Inventors: **MUKHOPADHYAY, Aindrila**; 3425 Louise Street, Oakland, California 94608 (US). **MINGARDON, Florence**; 6363 Christie Avenue, Emeryville, California 94608 (US). **CHANAL, Angelique**; 38 rue d'Hauteville, 75010 Paris (FR).(74) Agents: **LOCKYER, Jean M.** et al.; Kilpatrick Townsend and Stockton, LLP, Two Embarcadero Center, Eighth Floor, San Francisco, California 94111 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

— with sequence listing part of description (Rule 5.2(a))

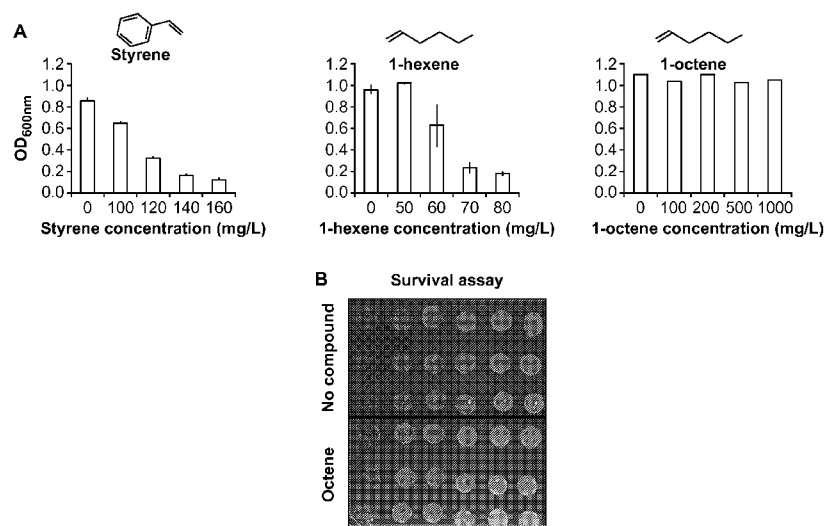
(54) Title: MODIFIED HOST CELLS HAVING TOLERANCE TO α -OLEFINS

FIG. 1

(57) Abstract: The present invention provides methods for improving microbial tolerance to alpha olefin compounds, host cells having increased tolerance to such compounds, and method of using the host cells to produce alpha olefin compounds.

Modified Host Cells Having Tolerance to α -Olefins

CROSS-REFERENCE TO RELATE APPLICATIONS

[0001] This application claims benefit of priority to U.S. provisional application no. 61/986,669, filed April 30, 2014, which is herein incorporated by reference.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

[0002] The invention was made with government support under Contract No. DE-AC02-05CH11231 awarded by the U.S. Department of Energy. The government has certain rights in the invention.

REFERENCE TO SUBMISSION OF A SEQUENCE LISTING AS AN ASCII TEXT FILE

[0003] This application includes a Sequence Listing as a text file named "SEQTEXT_77429-943405_ST25" created on April 29, 2015 and containing 50,738 bytes. The material contained in this text file is hereby incorporated by reference in its entirety for all purposes.

BACKGROUND OF THE INVENTION

[0004] Recent interest in biofuels and alternative sources of chemicals has made microbial production of bulk chemicals an important focus of development. However, many chemicals of industrial interest are toxic to the microorganisms producing them. Microbial production levels are thus constrained by the toxicity limits of the organism to the final product. Host engineering to improve strain tolerance towards target compounds is thus important to increase production levels.

[0005] Different compounds may require different host engineering approaches, as toxicity relates to the mode of inhibition, chemical properties, and hydrophobicity of a compound. Hydrophobic solvent-like compounds are hypothesized to impact membrane permeability and fluidity, diminish energy transduction and interfere with membrane protein function, affecting a range of essential cellular processes (Wagner et al, 2007). The cell defense mechanisms that respond to these compounds include induction of chaperones, modification

of membrane composition and cellular morphology, and induction of active membrane efflux pump transporter proteins that export the compounds out of the cells and the membrane. Of these, efflux pumps have recently emerged as an important target in engineering of host cells (Doshi et al. 2013; Dunlop et al. 2011; Dunlop et al. 2010; Fisher et al. 2013; Foo and Leong 2013).

[0006] Efflux pump proteins, also referred to herein simply as “pump” proteins, provide the general mechanism for the export of toxic compounds from cells (Nikaido 2009; Takatsuka et al. 2010). One of the best-characterized families of pumps, and also the key pumps in gram negative bacteria for tolerance towards solvent-like compounds, belong to the hydrophobe/amphiphile efflux (HAE1) family of resistance-nodulation-division (RND) pumps (Nikaido and Takatsuka 2009; Ramos et al. 2002; Tseng et al. 1999). RND efflux pumps are composed of 3 subunits: (i) an inner membrane unit proton antiporter that binds the substrate and transports it through (ii) the outer membrane channel subunit, and (iii) a periplasmic subunit that connects and stabilizes the inner and outer membrane units (Nikaido and Takatsuka 2009).

[0007] The *E. coli* AcrAB-TolC efflux pump is a member of the HAE1 family. AcrAB-TolC is composed of AcrB (inner membrane protein), TolC (outer membrane protein) and AcrA (periplasmic protein) and has been extensively studied (Murakami et al. 2002; Tikhonova et al. 2011). There are many reports in the scientific literature characterizing the mechanism of action of AcrB, including its rotational conformation changes (Seeger et al. 2006; Seeger et al. 2008; Sennhauser et al. 2007; Takatsuka and Nikaido 2009; Takatsuka and Nikaido 2010) as well as potential binding pockets and the amino acids involved in substrate recognition (Eicher et al. 2012; Husain and Nikaido 2010; Vargiu and Nikaido 2012). The reported substrate entry points have been located in the periplasm (Takatsuka and Nikaido 2007), the membrane and/or in the cytoplasm side of the pump (Eicher et al. 2012; Husain and Nikaido 2010; Murakami et al. 2004; Murakami and Yamaguchi 2003; Sennhauser et al. 2007). Although AcrB functions with AcrA and TolC in *E. coli*, AcrB also has efflux pump activity in the absence of AcrA and TolC (*see, e.g.,* Kapoor & Wendell, *Nano Lett.* 2013, 13, 2189–2193, 2013).

[0008] AcrB has broad substrate specificity that ranges from detergents to antibiotics and solvents. This pump is reported to play a major role in the secretion of various alkanes such as hexane, heptane, octane, and nonane (Takatsuka et al. 2010) and also in imparting tolerance to various terpene based biofuel compounds (Dunlop et al. 2011).

[0009] Production of α -olefins by microorganisms is hampered by toxicity of the compound to the host cells. There is therefore a need to improve tolerance. This invention addresses that need, in part, by providing host cells with genetic modifications to an *acrB* gene, or homolog, that result in improved tolerance to α -olefins and accordingly, increased α -olefin yields from modified host cells that produce the compounds.

SUMMARY OF THE INVENTION

[0010] The present invention is based, in part, on the discovery that an AcrB polypeptide can be engineered such that expression of the modified AcrB polypeptide in a host cell that produces an α -olefin compound improves tolerance of the host cell to the compound. This section summarizes illustrative embodiments of the invention, but does not limit the invention to these embodiments.

[0011] In one aspect, the invention provides a microorganism host cell comprising a recombinant non-naturally occurring variant AcrB-like polypeptide that comprises at least one amino acid sequence mutation relative to a native AcrB-like polypeptide sequence, wherein the non-naturally occurring AcrB-like polypeptide reduces accumulation of α -olefin compounds in the host cell in comparison to a wild-type host cell that expresses the native AcrB-like polypeptide at a native level. The sequence mutation may be an amino acid substitution relative to the native AcrB-like polypeptide sequence, *e.g.*, an amino acid other than the corresponding amino acid of SEQ ID NO:1. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide amino acid sequence is substantially identical to SEQ ID NO:1 and comprises one or more amino acid substitutions within 2 to 8 Angstroms of, or at, a position corresponding to amino acid 279, amino acid 584, amino acid 617, amino acid 822, amino acid 927, or amino acid 1033, as numbered with reference to SEQ ID NO:1. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide comprises an amino acid other than alanine at a position corresponding to amino acid 279. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide comprises a threonine at a position corresponding to amino acid 279. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide comprises an amino acid other than a glutamine at a position corresponding to amino acid 584. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide comprises an arginine at a position corresponding to amino acid 584. In some embodiments, the recombinant non-naturally occurring AcrB-like

polypeptide comprises an amino acid other than a leucine at a position corresponding to amino acid 617. In some embodiments, the recombinant non-naturally occurring AcrB-like polypeptide comprises a phenylalanine at a position corresponding to amino acid 617. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide comprises an amino acid other than a leucine at a position corresponding to amino acid 822. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide comprises proline at a position corresponding to amino acid 822. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide comprises an amino acid residue other than a phenylalanine at a position corresponding to amino acid 927. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide comprises a serine at a position corresponding to amino acid 927. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide comprises an amino acid residue other than a phenylalanine at a position corresponding to amino acid 1033. In some embodiments, the recombinant non-naturally occurring variant AcrB-like polypeptide comprises a tyrosine at a position corresponding to amino acid 1033. In some embodiments, the alpha-olefin compounds comprise 1-hexene, or a derivative thereof. In such embodiments, the host cell typically produces a greater amount of the 1-hexene, or derivative thereof, as compared to a wild-type host cell that expresses the native AcrB-like polypeptide at a native level. In additional embodiments, the host cell further comprises the proteins necessary to produce 1-hexene, or a derivative thereof. The host cell may be a bacterial cell, *e.g.*, either a gram negative or a gram positive host cell. In some embodiments, the bacterial host cell is a gram negative cell, such as an *E. coli* cell. In some embodiments, the host cell is a yeast host cell.

[0012] In a further aspect, the invention provides a recombinant variant AcrB polypeptide, wherein the recombinant variant AcrB polypeptide has at least 70% identity, or at least 80%, 85%, 90%, or 95% identity, to SEQ ID NO:1 and comprises one or more amino acid substitutions within 2 to 8 Angstroms of, or at, a position corresponding to amino acid 279, amino acid 584, amino acid 617, amino acid 822, amino acid 927, or amino acid 1033 of SEQ ID NO:1. In some embodiments, the variant AcrB polypeptide comprises an amino acid other than alanine at a position corresponding to amino acid 279. In some embodiments, the recombinant variant AcrB polypeptide comprises a threonine at a position corresponding to amino acid 279. In some embodiments, the recombinant variant AcrB polypeptide comprises an amino acid other than a glutamine at a position corresponding to amino acid 584. In some embodiments, the recombinant variant AcrB polypeptide comprises an arginine at a position

corresponding to amino acid 584. In some embodiments, the recombinant variant AcrB polypeptide comprises an amino acid other than a leucine at a position corresponding to amino acid 617. In some embodiments, the recombinant variant AcrB polypeptide comprises a phenylalanine at a position corresponding to amino acid 617. In some embodiments, the recombinant variant AcrB polypeptide comprises an amino acid other than a leucine at a position corresponding to amino acid 822. In some embodiments, the recombinant variant AcrB polypeptide comprises proline at a position corresponding to amino acid 822. In some embodiments, the recombinant variant AcrB polypeptide comprises an amino acid residue other than a phenylalanine at a position corresponding to amino acid 927. In some embodiments, the recombinant variant AcrB-like polypeptide comprises a serine at a position corresponding to amino acid 927. In some embodiments, the recombinant variant AcrB polypeptide comprises an amino acid residue other than a phenylalanine at a position corresponding to amino acid 1033. In some embodiments, the recombinant variant AcrB polypeptide comprises a tyrosine at a position corresponding to amino acid 1033. In some embodiments, the recombinant variant AcrB polypeptide comprises an amino acid sequence of SEQ ID NO:1 or SEQ ID NO:5 in which one or more substitutions at amino acid positions corresponding to amino acids 279, 584, 617, 822, 927, or 1033 of SEQ ID NO:1 are substituted relative. In typical embodiments, the recombinant variant AcrB polypeptide reduces accumulation of alpha-olefin compounds in the host cell in comparison to a wild-type host cell that expresses native AcrB at a native level.

[0013] In a further aspect, the invention provides a recombinant variant AcrB-like polypeptide, wherein the recombinant variant AcrB polypeptide has at least 70% identity, or at least 80%, 85%, 90%, or 95% identity, to SEQ ID NO:2, SEQ ID NO:3, or SEQ ID NO:4 and comprises one or more amino acid substitutions within 2 to 8 Angstroms of, or at, a position corresponding to amino acid 279, amino acid 584, amino acid 617, amino acid 822, amino acid 927, or amino acid 1033, as numbered with reference to SEQ ID NO:1. In typical embodiments, the recombinant variant AcrB-like polypeptide reduces accumulation of alpha-olefin compounds, *e.g.*, 1-hexene, in the host cell in comparison to a wild-type host cell that expresses a native AcrB-like polypeptide at a native level.

[0014] In a further embodiment, the invention provides a nucleic acid encoding a recombinant variant AcrB or AcrB-like polypeptide as described hereinabove and a microorganism host cell comprises the nucleic acid. The host cell may be a bacterial host cell, *e.g.*, either a gram positive or gram negative host cell. In some embodiments, the host

cell is a gram negative host cell, such as *E. coli*. In some embodiments, the host cell is a yeast host cell. In some embodiments, the host cell further comprises the proteins necessary to produce 1-hexene, or a derivative thereof.

[0015] In a further aspect the invention provides a method of producing an alpha-olefin compound, comprising culturing a host cell as described herein above under conditions sufficient for the host cell to produce the alpha-olefin compound. In some embodiments, the alpha-olefin compound is hexene or a derivative of hexene. In some embodiments, the host cell produces a greater amount of the 1-hexene, or derivative thereof, as compared to a wild-type host cell that expresses the native AcrB-like polypeptide at a natural level

[0016] In a further aspect, the invention provides a method for producing an alpha-olefin compound, comprising: culturing a microorganism host cell that overexpresses an AcrB-like polypeptide under conditions sufficient for the host cell to produce the alpha-olefin compound, wherein the host cell produces a greater amount of the alpha-olefin compound compared to a corresponding host cell that expresses the native AcrB-like polypeptide at native levels. In some embodiments, the alpha-olefin compound is 1-hexene, or a derivative thereof. In some embodiments, the AcrB-like polypeptide has at least 70% identity, or at least 80%, 85%, 90%, or 95% identity, to the amino acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5. In some embodiments, the AcrB-like polypeptide comprises the amino acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5. In some embodiments, the AcrB-like polypeptide comprises the amino acid sequence of SEQ ID NO:1 or SEQ ID NO:5. In some embodiments, the host cell is a bacterial host cell such as a gram negative or gram positive host cell. In some embodiments, the bacterial host cell is a gram negative cell, *e.g.*, an *E. coli* host cell. In some embodiments, the host cell is a yeast host cell. In some embodiments, the host cell further comprises the proteins necessary to produce 1-hexene, or a derivative thereof.

[0017] In a further aspect, the invention provides a cell culture medium comprising an alpha-olefin compound and a microorganism host cell that overexpresses an AcrB-like polypeptide, wherein the host cell produces a greater amount of the alpha-olefin compound compared to a corresponding host cell that expresses the native AcrB-like polypeptide at native levels. In some embodiments, the alpha-olefin compound is 1-hexene, or a derivative thereof. In some embodiments, the AcrB-like polypeptide has at least 70% identity, or at least 80%, 85%, 90%, or 95% identity, to the amino acid sequence of SEQ ID NO:1, SEQ ID

NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5. In some embodiments, the AcrB-like polypeptide comprises the amino acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5. In some embodiments, the AcrB-like polypeptide comprises the amino acid sequence of SEQ ID NO:1 or SEQ ID NO:5. In some embodiments, the host cell is a bacterial host cell such as a gram negative or gram positive host cell. In some embodiments, the bacterial host cell is a gram negative cell, *e.g.*, an *E. coli* host cell. In some embodiments, the host cell is a yeast host cell. In some embodiments, the host cell further comprises the proteins necessary to produce 1-hexene, or a derivative thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

[0018] Figure 1. Impact of styrene, 1-hexene and 1-octene on cell growth. A: The olefin compounds were added to different concentrations as indicated to an *E. coli* K12 culture at an $OD_{600}=0.2$. OD was monitored after 2 hour at 37 °C. The bars represent standard deviation between duplicates B: A survival test in presence of 1-octene was performed. Six different dilutions of a culture (10^8 cells to 10^2 cells, left to right) of *E. coli* K12 were spotted in triplicate on an agar plate. An overlay of 1-octene was added over the cells (controls had no overlay). Plates were incubated at room temperature for 2 days and then opened to allow evaporation of the olefin.

[0019] Figure 2: Impact of AcrAB production on growth, in presence of styrene 1-hexene, 1-octene and 1-nonene on cell growth. A: Different concentrations of styrene were added to an $OD_{600}=0.2$ culture of *E. coli* K12 Δ *acrAB* containing the pABc (white) or the control plasmid p0c (black). OD density was monitored after 2 hours of growth at 37°C. The bars represent standard deviation between duplicates. B: A survival test in presence of 1-hexene, 1-octene, or 1-nonene saturated atmosphere was performed. Six different dilutions of cultures (10^8 cells to 10^2 cells, left to right on the figure) were spotted in triplicate on an agar plates. Plates were then exposed to the aforementioned chemicals.

[0020] Figure 3: Impact of over-production of AcrA, AcrB and TolC on tolerance to 1-hexene. A. Survival tests of *E. coli* K12 wild type containing pABc or p0c and either pTolCa or p0a in a 1-hexene saturated atmosphere were performed. Six different dilutions of cultures (10^8 cells to 10^2 cells, left to right on the picture) were spotted in triplicate on an agar plate. Plates were then exposed to 1-hexene.

[0021] Figure 4: Impact of beneficial mutations on AcrB on tolerance to 1-hexene. A: Survival test of *E. coli* strain producing either the AcrB wild type or variants in a 1-hexene saturated atmosphere was performed. Two dilutions of cultures (Dilution 1: 10^5 cells and Dilution 2: 10^4 cells) were spotted in triplicate on an agar plate. B: Localization of the 6 beneficial mutations represented on the structure of a single monomer of AcrB or C: on the structure of the AcrB trimer.

[0022] Figure 5: Impact of beneficial AcrB mutations on growth, in presence of 1-hexene and styrene. Different concentrations of 1-hexene (A), styrene (B), or bile salts (C) were added to an $OD_{600}=0.2$ culture of *E. coli* K12 $\Delta acrAB$ containing either the AcrB wild type (pABc) or the AcrB variants. OD densities were monitored after 2 hours of culture at 37 °C. The bars represent standard deviation between duplicates.

[0023] Figure 6: Impact of beneficial mutations on AcrB on tolerance to 1-hexene. A: Survival test of *E. coli* strain producing either the AcrB Wild type or variants with 1, 2 or several beneficial mutations, in an 1-hexene saturated atmosphere was performed. Two dilutions of cultures (Dilution 2: 10^4 cells and Dilution 3: 10^3 cells) were spotted in triplicate on an agar plate. B: Soluble fraction of total proteins from a culture of strains producing either the wild type, either the variant F617L, either the variant 6 (A279T, Q584R, F617L) were analyzed by western blot using anti-His antibody.

[0024] Figure 7: Impact of homologous pumps on growth in presence of 1-hexene. Survival test of *E. coli* DH1 $\Delta acrAB$ producing either the AcrAB wild type or AcrAB homologs in a 1-hexene saturated atmosphere was performed. One dilution of cultures (10^6 cells) was spotted on agar plate. This experiment was repeated and showed similar results. From left to right and from top to bottom of the picture, are culture of strains producing the pumps *A. vinelandii* (YP_002797425), *A. vinelandii* (MexF), *A. vinelandii* (YP_002801502), *P. putida* (TtgB), *P. putida* (NP_745594), *P. putida* (MexF), *P. putida* (NP_743067), *P. aeruginosa* (MexB), *P. aeruginosa* (NP_250708), *P. aeruginosa* (MexF), *P. aeruginosa* (NP_253065), *P. fluorescens* (YP_258460), *P. fluorescens* (YP_260376), *P. fluorescens* (YP_258159), *P. fluorescens* (YP_258289), *R. metallidurans* (YP_586997), *R. metallidurans* (YP_583276), *R. metallidurans* (YP_583852), *M. aqueolei* (YP_960752), *M. aqueolei* (YP_957870), *P. syringae* (YP_276144), *P. syringae* (YP_275089), *P. syringae* (MdtB), *P. syringae* (YP_273025), *P. syringae* (YP_273463), *P. syringae* (YP_272632), *G. metallireducens* (YP_383777), *G. metallireducens* (YP_384620), *G. metallireducens* (YP_385464), *G. metallireducens* (YP_384608), *G. metallireducens* (YP_386066), *B.*

pertussis (NP_880739), *D. vulgaris* (YP_009662), *S. putrefaciens* (YP_001181626), *A. borkumensis* (YP_692684), *B. borkumensis* (YP_694242), *P. haloplanktis* (AcrB), *S. halifaxensis* (YP_001672433), *S. halifaxensis* (YP_001673423), *S. halifaxensis* (YP_001674385), *E. coli* (AcrEF), *E. coli* (AcrAB), *E. coli* (AcrD).

[0025] Figure 8 provides an alignment of SEQ ID NOs:1, 2, 3, 4, and 5. Transmembrane domains and amino acids involved in binding or H⁺ transport are shown in SEQ ID NO:1. As understood in the art, not all of the amino acids of SEQ ID NOS: 2, 3, 4, and 5 are explicitly shown in the alignments to SEQ ID NO:1 so that the overall homologies of the sequences to SEQ ID NO:1 are more readily apparent.

DETAILED DESCRIPTION OF THE INVENTION

[0026] Before the present invention is described, it is to be understood that this invention is not limited to particular embodiments described, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

Definitions

[0027] The terms "optional" or "optionally" as used herein mean that the subsequently described feature or structure may or may not be present, or that the subsequently described event or circumstance may or may not occur, and that the description includes instances where a particular feature or structure is present and instances where the feature or structure is absent, or instances where the event or circumstance occurs and instances where it does not.

[0028] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are now described. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited.

[0029] As used herein and in the appended claims, the singular forms "a", "and", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example,

reference to "an efflux pump" includes a plurality of such efflux pumps; reference to a "host cell" includes a plurality of such host cells.

[0030] As used herein, the term " α -olefin-tolerant" or "alpha-olefin-tolerant" in the context of a host cell refers to a host cell that has been engineered to have an AcrB-like polypeptide, *e.g.* AcrB, that provides for increased survival of the host cell, *i.e.*, the cell has increased tolerance, in the presence of an α -olefin, such as 1-hexene, compared to the counterpart host cell that produces the wild type AcrB-like polypeptide, *e.g.*, wild type AcrB, at a native level, *i.e.*, at a level that occurs in nature. As used herein, a host cell that is "more tolerant" to an α -olefin compound, *e.g.*, 1-hexene, refers to a host cell that has increased α -olefin tolerance relative to the counterpart wild type host cell.

[0031] As used herein, the term "AcrB polypeptide" refers to an *E. coli* K12 AcrB membrane polypeptide comprising the amino acid sequence of SEQ ID NO:1, or a variant thereof, that functions as an efflux pump to transport compounds out of a host cell. An "AcrB polypeptide" in the context of the present invention has at least 65% identity, typically at least 70%, 75%, 80%, 85%, or 90% identity, or greater, *e.g.*, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater, amino acid sequence identity, preferably over a region of at least about 100, 200, 300, 400, 500, 600, 700, 800, 900, or more amino acids, or over the total length of the polypeptide, to SEQ ID NO:1. A nucleic acid that encodes an AcrB polypeptide refers to a gene, cDNA, pre-mRNA, mRNA, and the like, including nucleic acids that encode variants of a naturally occurring AcrB amino acid sequence as described herein. AcrB polypeptide structure is well known (*see, e.g.*, Husain and Nikaido 2010; Murakami et al. 2002). In the present invention, although an AcrB polypeptide can function as a component of a HAE1 family efflux pump, *e.g.*, with AcrA and TolC, in some embodiments, an AcrB polypeptide functions independent of other HAE1 efflux pump components.

[0032] As used herein, the term "AcrF polypeptide" refers to an *E. coli* HAE1 family member membrane polypeptide comprising the amino acid sequence of SEQ ID NO:5 or a variant thereof, that functions as an efflux pump to transport compounds out of a host cell. An "AcrF polypeptide" in the context of the present invention has at least 65% identity, typically at least 70%, 75%, 80%, 85%, or 90% identity, or greater, *e.g.*, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater, amino acid sequence identity, preferably over a region of at least about 100, 200, 300, 400, 500, 600, 700, 800, 900, or more amino acids, or over the total length of the polypeptide, to SEQ ID NO:5. A nucleic acid that encodes an AcrF polypeptide refers to a gene, cDNA, pre-mRNA, mRNA, and the like, including nucleic

acids that encode variants of a naturally occurring AcrF amino acid sequence as described herein. HAE1 family polypeptide structure is well known (*see, e.g.*, Husain and Nikaido 2010; Murakami et al. 2002, which provide an illustrative AcrB structure). In the present invention, although an AcrF polypeptide can function as a component of a HAE1 family efflux pump, *e.g.*, AcrE and TolC, in some embodiments, an AcrF polypeptide functions independent of other HAE1 efflux pump components.

[0033] As used herein, the term “*Pseudomonas putida* KT2440 polypeptide NCBI NP_743544” refers to a *Pseudomonas putida* HAE1 family member membrane polypeptide comprising the amino acid sequence of SEQ ID NO:2, or a variant thereof, that functions as an efflux pump to transport compounds out of a host cell. A “*Pseudomonas putida* KT2440 polypeptide NCBI NP_743544” in the context of the present invention has at least 65% identity, typically at least 70%, 75%, 80%, 85%, or 90% identity, or greater, *e.g.*, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater, amino acid sequence identity, preferably over a region of at least about 100, 200, 300, 400, 500, 600, 700, 800, 900, or more amino acids, or over the total length of the polypeptide, to SEQ ID NO:2. A nucleic acid that encodes a *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544 refers to a gene, cDNA, pre-mRNA, mRNA, and the like, including nucleic acids that encode variants of a naturally occurring *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544 amino acid sequence as described herein. HAE1 family polypeptide structure is well known (*see, e.g.*, Husain and Nikaido 2010; Murakami et al. 2002, which describe AcrB structures). In the present invention, although a *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544 polypeptide can function as a component of a HAE1 family efflux pump, in some embodiments, a *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544 polypeptide functions independent of other HAE1 efflux pump components.

[0034] As used herein, the term “*Marinobacter aqueolei* EPL polypeptide NCBI YP_960752” refers to a *Marinobacter aqueolei* EPL HAE1 family member membrane polypeptide comprising the amino acid sequence of SEQ ID NO:3, or a variant thereof, that functions as an efflux pump to transport compounds out of a host cell. A *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752” in the context of the present invention has at least 65% identity, typically at least 70%, 75%, 80%, 85%, or 90% identity, or greater, *e.g.*, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater, amino acid sequence identity, preferably over a region of at least about 100, 200, 300, 400, 500, 600, 700, 800, 900, or more amino acids, or over the total length of the polypeptide, to SEQ ID NO:3. A

nucleic acid that encodes a *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752 refers to a gene, cDNA, pre-mRNA, mRNA, and the like, including nucleic acids that encode variants of a naturally occurring *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752 amino acid sequence as described herein. HAE1 family polypeptide structure is well known (*see, e.g.*, Husain and Nikaido 2010; Murakami et al. 2002, which describe AcrB structures). In the present invention, although a *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752 can function as a component of a HAE1 family efflux pump, in some embodiments, a *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752 functions independent of other HAE1 efflux pump components.

[0035] As used herein, the term “*Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810” refers to a *Pseudoalteromonas haloplanktis* TAC125 HAE1 family member membrane polypeptide comprising the amino acid sequence of SEQ ID NO:4, or a variant thereof, that functions as an efflux pump to transport compounds out of a host cell. A “*Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810” in the context of the present invention has at least 65% identity, typically at least 70%, 75%, 80%, 85%, or 90% identity, or greater, *e.g.*, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater, amino acid sequence identity, preferably over a region of at least about 100, 200, 300, 400, 500, 600, 700, 800, 900, or more amino acids, or over the total length of the polypeptide, to SEQ ID NO:4. A nucleic acid that encodes a *Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810 refers to a gene, cDNA, pre-mRNA, mRNA, and the like, including nucleic acids that encode variants of a naturally occurring *Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810 amino acid sequence as described herein. HAE1 family polypeptide structure is well known (*see, e.g.*, Husain and Nikaido 2010; Murakami et al. 2002, which describe AcrB structures). In the present invention, although a *Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810 can function as a component of a HAE1 family efflux pump, in some embodiments, a *Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810 functions independent of other HAE1 efflux pump components.

[0036] As used herein, the term “AcrB-like polypeptide” encompasses an *E. coli* AcrB polypeptide, an AcrF polypeptide, a *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544, a *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752, and a *Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810, and variants of the polypeptide as described herein. An AcrB-like polypeptide has at least 50% amino acid

sequence identity, or at least 55%, 60%, or 65% identity, preferably at least 70%, 75%, 80%, 85%, or 90% identity, or greater, *e.g.*, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater, amino acid sequence identity, preferably over a region of at least about 100, 200, 300, 400, 500, 600, 700, 800, 900, or more amino acids, or over the total length of the polypeptide, to SEQ ID NO:1. A nucleic acid that encodes an AcrB-like polypeptide refers to a gene, cDNA, pre-mRNA, mRNA, and the like, including nucleic acids that encode variants of a naturally occurring AcrB, AcrF, *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544, *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752, or *Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810 amino acid sequences described herein. AcrB polypeptide structure is well known (*see, e.g.*, Husain and Nikaido 2010; Murakami et al. 2002). In the present invention, although an AcrB-like polypeptide can function as a component of a HAE1 family efflux pump, *e.g.*, with AcrA and TolC, in some embodiments, AcrB-like polypeptides are used in embodiments in which the AcrB-like polypeptide functions independent of other HAE1 efflux pump components.

[0037] The terms “wild type”, “native”, “endogenous”, and “naturally occurring” are used herein in the context of an AcrB-like polypeptide or polynucleotide to refer to an AcrB or AcrB homolog such as AcrF, *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544, *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752, or *Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810, that is present in a Gram negative bacteria in nature.

[0038] A “variant” as used herein with respect to an AcrB-like polypeptide or AcrB-like polynucleotide; comprises one or more modifications relative to a native AcrB-like polypeptide or comprises one or more modifications relative to a native AcrB-like polynucleotide. Thus, a “variant” of an AcrB polypeptide or AcrB polynucleotide comprises one or more modifications relative to a native AcrB polypeptide or nucleic acid. Similarly, a “variant” with respect to an AcrF polypeptide or AcrF polynucleotide comprises one or more modifications relative to a native AcrF polypeptide or comprises one or more modifications relative to the native AcrF polynucleotide; and a “variant” with respect to a *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544 or polynucleotide, *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752 or polynucleotide, or *Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810 or polynucleotide refers to a polypeptide that comprises one or more modifications relative to the native polypeptide or a polynucleotide that comprises one or more modifications relative to the native polynucleotide.

Modifications include substitutions, insertions, deletions and/or truncations of one or more specific amino acid residues or of one or more specific nucleotides or codons in the polypeptide or polynucleotide.

[0039] In the context of this invention, the term “mutant” with respect to a mutant polypeptide or mutant polynucleotide is used interchangeably with “variant”. A “non-naturally” occurring AcrB-like polypeptide refers to a variant or mutant AcrB-like polypeptide that is not present in a cell in nature and that is produced by genetic modification, *e.g.*, using genetic engineering technology or mutagenesis techniques, of a native AcrB-like polynucleotide, such as a native AcrB or AcrF polynucleotide, or a native polynucleotide that encodes *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544, *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752, or *Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810.

[0040] A “sequence mutation” in the context of the present invention refers to a substitution, deletion, or insertion of one more amino acids in an amino acid sequence in comparison to a native polypeptide.

[0041] The terms “numbered with reference to”, or “corresponding to,” or “determined with reference to” when used in the context of the numbering of a given amino acid sequence, refers to the numbering of the residues of a specified reference sequence when the given amino sequence is maximally aligned and compared to the reference sequence. For example, a residue in a variant AcrB-like polypeptide “corresponds to” an amino acid at a position in SEQ ID NO:1 when the residue aligns with the amino acid in a comparison of SEQ ID NO:1 and the variant protein.

[0042] The terms “polynucleotide” and “nucleic acid” are used interchangeably and refer to a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, nucleic acid analogs may be used that may have alternate backbones, comprising, *e.g.*, phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages (see Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press); positive backbones; non-ionic backbones, and non-ribose backbones. Thus, nucleic acids or polynucleotides may also include modified nucleotides that permit correct read-through by a polymerase.

“Polynucleotide sequence” or “nucleic acid sequence” includes both the sense and antisense

strands of a nucleic acid as either individual single strands or in a duplex. As will be appreciated by those in the art, the depiction of a single strand also defines the sequence of the complementary strand; thus the sequences described herein also provide the complement of the sequence. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses variants thereof (*e.g.*, degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and combinations of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine hypoxanthine, isocytosine, isoguanine, etc.

[0043] The terms "peptide," "polypeptide" and "protein" are used interchangeably in the context of the present invention.

[0044] The term "substantially identical," used in the context of two nucleic acids or polypeptides, refers to a sequence that has at least 50% sequence identity with a reference sequence. Percent identity can be any integer from 50% to 100%. Some embodiments include at least: 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%, compared to a reference sequence, *e.g.*, across a region of 500 amino acids or 750 amino acids or more, or across the length of, a reference sequence such as SEQ ID NO:1, or one of SEQ ID NOS:2 to 5 using the programs described herein; preferably BLAST using standard parameters.

[0045] Percent identity with respect to amino acid or nucleotide sequences is defined herein as the percentage of amino acid or nucleotide residues in the candidate sequence that are identical with the known polypeptides, after aligning the sequences for maximum percent identity and introducing gaps, if necessary, to achieve the maximum percent homology. Identity at the nucleotide or amino acid sequence level may be determined using methods known in the art, including but not limited to BLAST (Basic Local Alignment Search Tool) analysis using the algorithms employed by programs such as the BLAST programs blastp, blastn, blastx, tblastn and tblastx (Altschul (1997), *Nucleic Acids Res.* 25, 3389-3402, and Karlin (1990), *Proc. Natl. Acad. Sci. USA* 87, 2264-2268), which are tailored for sequence similarity/identity searching.

[0046] Nucleic acid or protein sequences that are substantially identical to a reference sequence include "conservatively modified variants." One of skill will recognize that

individual substitutions, in a nucleic acid, peptide, polypeptide, or protein sequence that alters a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Examples of amino acid groups defined in this manner can include: a "charged/polar group" including Glu (Glutamic acid or E), Asp (Aspartic acid or D), Asn (Asparagine or N), Gln (Glutamine or Q), Lys (Lysine or K), Arg (Arginine or R) and His (Histidine or H); an "aromatic or cyclic group" including Pro (Proline or P), Phe (Phenylalanine or F), Tyr (Tyrosine or Y) and Trp (Tryptophan or W); and an "aliphatic group" including Gly (Glycine or G), Ala (Alanine or A), Val (Valine or V), Leu (Leucine or L), Ile (Isoleucine or I), Met (Methionine or M), Ser (Serine or S), Thr (Threonine or T) and Cys (Cysteine or C). Within each group, subgroups can also be identified. For example, the group of charged/polar amino acids can be sub-divided into sub-groups including: the "positively-charged sub-group" comprising Lys, Arg and His; the "negatively-charged sub-group" comprising Glu and Asp; and the "polar sub-group" comprising Asn and Gln. In another example, the aromatic or cyclic group can be sub-divided into sub-groups including: the "nitrogen ring sub-group" comprising Pro, His and Trp; and the "phenyl sub-group" comprising Phe and Tyr. In another further example, the aliphatic group can be sub-divided into sub-groups including: the "large aliphatic non-polar sub-group" comprising Val, Leu and Ile; the "aliphatic slightly-polar sub-group" comprising Met, Ser, Thr and Cys; and the "small-residue sub-group" comprising Gly and Ala. Examples of conservative mutations include amino acid substitutions of amino acids within the sub-groups above, such as, but not limited to: Lys for Arg or vice versa, such that a positive charge can be maintained; Glu for Asp or vice versa, such that a negative charge can be maintained; Ser for Thr or vice versa, such that a free --OH can be maintained; and Gln for Asn or vice versa, such that a free --NH₂ can be maintained.

[0047] A polynucleotide or polypeptide is "heterologous" to an organism or a second polynucleotide or polypeptide sequence if it originates from a foreign species, or, if from the same species, is modified from its original form. For example, when a polynucleotide encoding a polypeptide sequence is said to be operably linked to a heterologous promoter, it means that the polynucleotide coding sequence encoding the polypeptide is derived from one species whereas the promoter sequence is derived from another, different species; or, if both are derived from the same species, the coding sequence is not naturally associated with the promoter. As an example of a "heterologous" sequence, an AcrB polypeptide containing an

amino acid substitution with reference to SEQ ID NO:1 is heterologous to an *E. coli* host cell, as it has been modified from the native AcrB polypeptide sequence of SEQ ID NO:1.

[0048] A "host cell" as used herein refers to a cell that is genetically modified. The term includes a parental cell that is genetically modified as well as progeny and genetically modified derivatives. Genetic modification can be achieved by any suitable genetic engineering technique and/or mutagenesis techniques (*e.g.*, chemical or UV mutagenesis and subsequent selection). Using recombinant technology, nucleic acid molecules can be introduced, deleted, inhibited or otherwise modified in the host cell. A "host cell" encompasses a cell from any prokaryote or eukaryote. A "microorganism host cell" as used here refers to a cell from a single-cell organism, *e.g.*, a bacteria; a single-cell fungi, such as a yeast; an archaea, a cyanobacteria, and the like.

[0049] A "recombinant nucleic acid" or "recombinant polynucleotide" as used herein refers to a polymer of nucleic acids wherein at least one of the following is true: (a) the sequence of nucleic acids is foreign to (*i.e.*, not naturally found in) a given host cell; (b) the sequence may be naturally found in a given host cell, but in an unnatural (*e.g.*, greater than expected) amount; or (c) the sequence of nucleic acids comprises two or more subsequences that are not found in the same relationship to each other in nature. For example, regarding instance (c), a recombinant nucleic acid sequence will have two or more sequences from unrelated genes arranged to make a new functional nucleic acid. A polypeptide is "recombinant" when it is artificial or engineered, or derived from or contains an artificial or engineered protein. For example, a polynucleotide that is inserted into a vector or any other heterologous location, *e.g.*, in a genome of a recombinant organism, such that it is not associated with nucleotide sequences that normally flank the polynucleotide as it is found in nature is a recombinant polynucleotide. A protein expressed *in vitro* or *in vivo* from a recombinant polynucleotide is an example of a recombinant polypeptide. Likewise, a polynucleotide sequence that does not appear in nature, for example a variant of a naturally occurring gene, is recombinant. A "recombinant" host cell contains a recombinant nucleic acid or recombinant protein.

[0050] The term "promoter," as used herein, refers to a polynucleotide sequence capable of driving transcription of a DNA sequence in a cell. Thus, promoters used in the polynucleotide constructs of the invention include *cis*- and *trans*- acting transcriptional control elements and regulatory sequences that are involved in regulating or modulating the timing and/or rate of transcription of a gene. For example, a promoter can be a *cis*- acting transcriptional control element, including an enhancer, a repressor binding sequence and the

like. These cis-acting sequences typically interact with proteins or other biomolecules to carry out (turn on/off, regulate, modulate, *etc.*) gene transcription. Most often the core promoter sequences lie within 1-2 kb of the translation start site, more often within 1 kbp and often within 500 bp or 200 bp or fewer, of the translation start site. By convention, promoter sequences are usually provided as the sequence on the coding strand of the gene it controls. In the context of this application, a promoter is typically referred to by the name of the gene for which it naturally regulates expression. A promoter used in an expression construct of the invention is referred to by the name of the gene. Reference to a promoter by name includes a wild type, native promoter as well as variants of the promoter that retain the ability to induce expression. Reference to a promoter by name is not restricted to a particular species, but also encompasses a promoter from a corresponding gene in other species.

[0051] A "constitutive promoter" in the context of this invention refers to a promoter that is capable of initiating transcription under most conditions in a cell, *e.g.*, in the absence of an inducing molecule. An "inducible promoter" initiates transcription in the presence of an inducer molecule.

[0052] The term "operably linked" refers to a functional relationship between two or more polynucleotide (*e.g.*, DNA) segments. Typically, it refers to the functional relationship of a transcriptional regulatory sequence to a transcribed sequence. For example, a promoter or enhancer sequence is operably linked to a DNA sequence if it stimulates or modulates the transcription of the DNA sequence in an appropriate host cell or other expression system. Generally, promoter transcriptional regulatory sequences that are operably linked to a transcribed sequence are physically contiguous to the transcribed sequence, *i.e.*, they are cis-acting. However, some transcriptional regulatory sequences, such as enhancers, need not be physically contiguous or located in close proximity to the coding sequences whose transcription they enhance.

[0053] The term "expression cassette" or "DNA construct" or "expression construct" refers to a nucleic acid construct that, when introduced into a host cell, results in transcription and/or translation of an RNA or polypeptide, respectively. One example of an expression cassette is a polynucleotide construct that comprises a polynucleotide sequence encoding an AcrB-like polypeptide of the invention protein operably linked to a promoter, *e.g.*, an inducible promoter, where the expression cassette is introduced into a microorganism, such as a bacterial cell. In some embodiments, an expression cassette comprises a polynucleotide sequence encoding an AcrB-like polypeptide of the invention, *e.g.*, an AcrB variant, where

the polynucleotide is targeted to a position in the genome of a microorganism such that expression of the polynucleotide sequence is driven by a promoter that is present in the microorganism. The promoter may be a native promoter, or a non-native promoter that has been introduced into the microorganism.

[0054] In instances when a NCBI/GenBank Accession No. (such as YP_957870) is employed to refer to the inner membrane protein of an efflux pump, it is understood in the art that the entire nucleotide sequence of the operon, and the known/deduced amino acid sequences of the proteins forming the efflux pump, are available from the literature, including the NCBI database ([http://www site ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/)).

[0055] This invention employs various routine recombinant nucleic acid techniques. Generally, the nomenclature and the laboratory procedures in recombinant DNA technology described below are well known and commonly employed in the art. Many manuals that provide direction for performing recombinant DNA manipulations are available, *e.g.*, Sambrook & Russell, *Molecular Cloning, A Laboratory Manual* (3rd Ed, 2001); and *Current Protocols in Molecular Biology* (Ausubel, *et al.*, John Wiley and Sons, New York, 2009, supplements through 2014).

Introduction

[0056] The present invention is based, in part, on the discovery of a mechanism for manipulating an efflux pump polypeptide to improve tolerance of a microorganism, such as a yeast or bacteria, towards α -olefin compounds, *e.g.*, 1-hexene, or derivatives thereof, by genetically modifying an AcrB-like polypeptide, *e.g.*, an AcrB polypeptide of SEQ ID NO:1. In some embodiments, the microorganism has improved tolerance to 1-hexene. In some embodiments, the microorganism has improved tolerance to 1-butene, 1-pentene, 1-heptene, or 1-octene. In some embodiments, the microorganism has improved tolerance to multiple α -olefin compounds, such as one or more of 1-hexene, 1-pentene, 1-butene, 1-heptene, or 1-octene. In some embodiments, the microorganism may also exhibit improved tolerance to compounds in addition to an α -olefin, for example, improved tolerance to a hydrocarbon such as pentane, hexane, heptane; or to an alcohol such as hexanol or heptanol as well as improved tolerance to the α -olefin.

[0057] In one aspect, the invention thus provides nucleic acids encoding non-naturally occurring AcrB-like polypeptide and recombinant host cells engineered to express such polypeptides. A microorganism modified to express a variant AcrB-like polypeptide as

described herein can be employed in methods of producing an α -olefin, *e.g.*, 1-hexene. The yield of the α -olefin is increased in comparison to the yield obtained from a corresponding microorganism that expresses the naturally occurring AcrB-like polypeptide, *e.g.*, SEQ ID NO:1 or SEQ ID NO:5; or SEQ ID NO:2, 3, or 4, at native levels, *i.e.*, the corresponding microorganism has not been engineered to express a variant AcrB-like polypeptide as described herein or engineered to over-express a naturally occurring AcrB-like polypeptide.

[0058] In a further aspect, the invention also provides methods of producing α -olefin compounds, *e.g.*, 1-hexene, using a host cell that has been genetically modified to overexpresses a naturally occurring AcrB-like polypeptide or a variant thereof at a level greater than the native amount of polypeptide produced by a counterpart host cell that does not have the genetic modification.

[0059] The following sections describe AcrB-like polypeptide and nucleic acid sequences in accordance with the invention, methods of engineering a microorganism to express the polypeptides, and methods of using such microorganism to produce an α -olefin, *e.g.*, 1-hexene or 1-butene, 1-pentene, 1-heptene, or 1-octene.

AcrB-like polypeptide and nucleic acid sequences

[0060] In one aspect, the invention provides nucleic acids encoding non-naturally occurring AcrB-like polypeptides and recombinant host cells engineered to express such polypeptides. Such polypeptides have one or more amino acid substitutions compared to a native AcrB-like polypeptide sequence, such as SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, or SEQ ID NO:5.

[0061] In the context of the present invention, illustrative substitution mutations are typically described with reference to the *E. coli* K12 AcrB polypeptide sequence (SEQ ID NO:1):

```
MPNFFIDRPIFAWVIAIIIMLAGGLAILKLPVAQYPTIAPPAVTISASYPGADAKTVQDT
VTQVIEQNMNGIDNLMYMSSNSDSTGTVQITLTFESGTDADIAQVQVQNKQLQLAMP
LLPQEVQQQGVSVSEKSSSSFLMVVGVIINTDGTMTQEDISDYVAANMKDAISRSGV
GDVQLFGSQYAMRIWMNPNELNKFQLTPVDVITAIKAQNAQVAAGQLGGTPPVKG
QQLNASIIAQTRLTSTEEFGKILLKVNQDGSRVLLRDVAKIELGGENYDIIAEFNGQPA
SGLGIKLATGANALDTAAAIRAELAKMEPFFPSGLKIVYPYDTTPFVKISIEHVVKTLV
EAIILVFLVMYFLQNFRATLIPTIAVPVLLGTFVLAAFGFSINTLTMFGMVLAIGL
```

LVDDAIVVVENVVERVMAEEGLPPKEATRKSMGQIQGALVGIAMVLSAVFVPMMAFFG
 GSTGAIYRQFSITIVSAMALSVLVALILTPALCATMLKPIAKGDHGEGKKGGFFGWFN
 MFEKSTHHYTDSVGGILRSTGRYLVLYLIHVVGMAVYLFVRLPSSFLPDEDQGVFMTM
 VQLPAGATQERTQKVLNEVTHYYLTKEKNNVESVFAVNGFGFAGRGQNTGIAFVSL
 KDWADRPGEENKVEAITMRATRAFSQIKDAMVFAFNLPAIVELGTATGDFDFELIDQA
 GLGHEKLTQARNQLLAEAAKHPDMLTSVRPNGLEDTPQFKIDIDQEKAQALGVSIND
 INTTLGAAWGGSYVNDFIDRGRVKKVYVMSEAKYRMLPDDIGDWYVRAADGQMV
 PFSAFSSSRWEYGSRLERYNGLPSMEILGQAAPGKSTGEAMELMEQLASKLPTGVG
 YDWTGMSYQERLSGNQAPSLYAISLIVVFLCLAALYESWSIPFSVMLVPLGVIGALL
 AATFRGLTNDVYFQVGLLTTIGLSAKNAILIVEFAKDLMDKEGKGLIEATLDAVRMR
 LRPILMTSLAFILGVMPLVISTGAGSGAQNNAVGTGVMGGMVTATVLAIFFVPVFFVV
 VRRRFSRKNEDIEHSHTVDHH. Thus, where SEQ ID NO:1 is the reference sequence, the position designation of a mutated amino acid in a variant polypeptide aligned to the reference sequence is the number of the position in SEQ ID NO:1. Illustrative transmembrane domains of SEQ ID NO:1 and SEQ ID NO:1 amino acids identified from structural studies as being important for binding or for H⁺ transport are shown in Figure 8. The positions of illustrative substitutions in SEQ ID NO:1 that improve tolerance are also shown in Figure 8.

[0062] In some embodiments, an α -olefin-tolerant host cell comprises a nucleic acid that encodes a variant AcrB-like polypeptide that has an amino acid sequence that has at least 50% amino acid sequence identity, or at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino acid sequence identity over a region of at least 1,000 or more amino acids, or over the length of the entire polypeptide, to an amino acid sequence of SEQ ID NO:1, where the mutant protein confers tolerance to the host cell relative to a counterpart host cell that has not been modified to express the variant protein or to overexpress native AcrB polypeptide. In some embodiments, the variant AcrB-like polypeptide comprises one or more amino acid substitutions as determined with reference to SEQ ID NO:1. Thus, in some embodiments, the invention provides a variant AcrB-like polypeptide, wherein the variant polypeptide is substantially identical to SEQ ID NO:1 and has one or more amino acid substitutions at a position corresponding to residue 279, 584, 617, 822, 927, or 1033 of SEQ ID NO:1. In some embodiments, a polypeptide of the invention is at least 70% identical, or at least 80%, at least 85%, at least 90% or at least 95% identical, to SEQ ID NO:1 and has at least one amino acid substitution at a position corresponding to residue 279, 584, 617, 822, 927, or 1033 of SEQ ID NO:1. In some embodiments, a polypeptide of the invention is at least 70%

identical, or at least 80%, at least 85%, at least 90% or at least 95% identical, to SEQ ID NO:1, and has an amino acid substitution, *e.g.*, alanine to threonine, at a position corresponding to position 279 of SEQ ID NO:1. In some embodiments, a polypeptide of the invention is at least 70% identical, or at least 80%, at least 85%, at least 90% or at least 95% identical, to SEQ ID NO:1, and has an amino acid substitution, *e.g.*, glutamine to arginine, at a position corresponding to position 584 of SEQ ID NO:1. In some embodiments, a polypeptide of the invention is at least 70% identical, or at least 80%, at least 85%, at least 90% or at least 95% identical, to SEQ ID NO:1, and has an amino acid substitution, *e.g.*, phenylalanine to leucine, at a position corresponding to position 617 of SEQ ID NO:1. In some embodiments, a polypeptide of the invention is at least 70% identical, or at least 80%, at least 85%, at least 90% or at least 95% identical, to SEQ ID NO:1, and has an amino acid substitution, *e.g.*, leucine to proline, at a position corresponding to position 822 of SEQ ID NO:1. In some embodiments, a polypeptide of the invention is at least 70% identical, or at least 80%, at least 85%, at least 90% or at least 95% identical, to SEQ ID NO:1, and has an amino acid substitution, *e.g.*, phenylalanine to serine, at a position corresponding to position 927 of SEQ ID NO:1. In some embodiments, a polypeptide of the invention is at least 70% identical, or at least 80%, at least 85%, at least 90% or at least 95% identical, to SEQ ID NO:1, and has an amino acid substitution, *e.g.*, from phenylalanine to tyrosine, at a position corresponding to position 1033 of SEQ ID NO:1.

[0063] In further embodiments, an AcrB-like polypeptide in accordance with the invention has at least two amino acid substitutions, or at least three amino acid substitutions, selected from positions 279, 584, 617, 822, 927 and 1033, as numbered with reference to SEQ ID NO:1. In some embodiments, an AcrB-like polypeptide in accordance with the invention has 3 or fewer substitutions at positions 279, 584, 617, 822, 927, or 1033 as numbered with reference to SEQ ID NO:1.

[0064] In some embodiments, an AcrB-like polypeptide in accordance with the invention has a mutation, typically a substitution, at a residue that is in close proximity, *e.g.*, within 2 to within 8 Angstroms, or in some embodiments, within 2 to 5 Angstroms or 2 to 6 Angstroms, with respect to the secondary structure of the AcrB-like polypeptide, to one of residues 279, 584, 617, 822, 927, and 1033 as numbered with reference to SEQ ID NO:1. The structure of the AcrB polypeptide has been well characterized (see, *e.g.*, Murakami et al, 2002 and additional references cited in the Examples section). One of skill can thus identify residues that are in close proximity and of likely functional relevance. For example, in some

embodiments, a variant AcrB-like polypeptide conferring increased tolerance of a host cell to an α -olefin may have a mutation, *e.g.*, a substitution, at a position in close proximity to residue 279, *e.g.*, at a position corresponding to residue 141, 151, 178, 179, 277, 278, 280, 281, 283, 284, 285, 286, 287, 288, 610, 611, 612, 613, or 614 as numbered with reference to SEQ ID NO:1. In some embodiments, the position in close proximity to residue 279 is within 3 to 5 Angstroms and corresponds to position 178, 278, 280, 285, 286, 287, or 612 as numbered with reference to SEQ ID NO:1. In some embodiments, a variant AcrB-like polypeptide conferring increased tolerance of a host cell to an α -olefin may have a mutation, *e.g.*, a substitution, at a position in close proximity to residue 584, *e.g.*, at a position corresponding to residue 275, 276, 277, 278, 582, 583, 585, 586, 587, 588, 589, 613, 614, 621, 622, or 623 as numbered with reference to SEQ ID NO:1. AcrB functions as a trimer; thus in some embodiments, a residue in close proximity to residue 584, *e.g.*, at a position corresponding to residue 220, 221, 222, 223, 228, 229, or 231 as determined with reference to SEQ ID NO:1, is present in another AcrB polypeptide molecule that is part of in the trimer. In some embodiments, the position in close proximity to residue 584 is within 3 to 6 Angstroms and corresponds to position 222, 223, 275, 276, 277, 278, 582, 583, 585, 586, 587, 588, 613, 614, or 622 as numbered with reference to SEQ ID NO:1. In some embodiments, the position in close proximity to residue 584 is within 3 to 5 Angstroms corresponds to residue 222, 276, 277, 278, 583, 585, 586, 587, 588, or 622 as numbered with reference to SEQ ID NO:1. In some embodiments, a variant AcrB polypeptide conferring increased tolerance of a host cell to an α -olefin may have a mutation, *e.g.*, a substitution, at a position in close proximity to residue 617, *e.g.*, at a position corresponding to residue 134, 135, 136, 573, 575, 615, 616, 618, 619, 620, 624, 625, or 626 as numbered with reference to SEQ ID NO:1. In some embodiments, the position in close proximity to residue 617 is within 3 to 6 Angstroms and corresponds to residue 134, 135, 136, 615, 616, 618, or 619 as numbered with reference to SEQ ID NO:1. In some embodiments, the position in close proximity to residue 617 is within 3 to 5 Angstroms and corresponds to residue 616, 618, or 619 as numbered with reference to SEQ ID NO:1. In some embodiments, a variant AcrB-like polypeptide conferring increased tolerance of a host cell to an α -olefin may have a mutation, *e.g.*, a substitution, at a position in close proximity to residue 822, *e.g.*, at a position corresponding to residue 79, 80, 160, 161, 165, 313, 685, 686, 687, 688, 689, 817, 818, 819, 820, 821, 823, 824, or 825 as numbered with reference to SEQ ID NO:1. In some embodiments, the position in close proximity to residue 822 is within 3 to 5 Angstroms and corresponds to residue 687, 818, 819, 820, 821, 823, or 824 as numbered with reference to

SEQ ID NO:1. In some embodiments, a variant AcrB-like polypeptide conferring increased tolerance of a host cell to an α -olefin may have a mutation, *e.g.*, a substitution, at a position in close proximity to residue 927, *e.g.*, at a position corresponding to residue 556, 559, 561, 909, 912, 913, 914, 915, 916, 917, 921, 922, 923, 924, 925, 926, 928, 929, 930, 931, 932, 933, 1002, 1003, 1005, 1006, or 1007 as numbered with reference to SEQ ID NO:1. In some embodiments, the position in close proximity to residue 927 is within 3 to 5 Angstroms and corresponds to residue 912, 913, 923, 924, 925, 926, 928, 929, 930, 931, or 1006 as numbered with reference to SEQ ID NO:1. In some embodiments, a variant AcrB-like polypeptide conferring increased tolerance of a host cell to an α -olefin may have a mutation, *e.g.*, a substitution, at a position in close proximity to residue 1033, *e.g.*, at a position corresponding to residue 896, 953, 958, 959, 960, 961, 963, 1028, 1029, 1030, 1031, 1032, 1034, 1035, or 1036 as numbered with reference to SEQ ID NO:1. In some embodiments, the position in close proximity to residue 1033 is within 3 to 5 Angstroms and corresponds to residue 894, 895, 896, 960, 1031, 1032, 1034, 1035, or 1036 as numbered with reference to SEQ ID NO:1.

[0065] In addition, based on the structural information available, one of skill can additionally identify residues of an AcrB-like polypeptide that would likely be tolerant to substitution. For example, an AcrB-like polypeptide variant having a mutation at one or more positions corresponding to 279, 584, 617, 822, 927 or 1033 of SEQ ID NO:1 may comprise substitutions at other positions that do not have an effect on the ability of the polypeptide to confer α -olefin tolerance to a host cell and do not otherwise impair pump activity. Accordingly, one of skill can identify variant AcrB-like polypeptides having a mutation described herein and the specified percent identity to the reference sequence, *e.g.*, SEQ ID NO:1, that would retain the desired function.

[0066] In additional embodiments, a mutation as described herein can be introduced in an AcrB-like polypeptide that has another mutation in a domain that enhances efflux of a compound, for example at a site for enlargement in the entrance of the cleft that facilitates conformational changes or improve the affinity for the substrate (Fisher et al. 2013; Foo and Leong 2013).

[0067] In some embodiments, an AcrB-like polypeptide of the invention is a homolog of an *E. coli* AcrB efflux protein that is modified at a site corresponding to one or more of positions 279, 584, 617, 822, 927 or 1033 of SEQ ID NO:1, to improve tolerance of an organism to an α -olefin. For example, *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544,

Marinobacter aqueolei EPL NCBI YP_960752, and *Pseudoalteromonas haloplanktis* TAC125 NCBI YP_341810 have AcrB homologs designated by the accession number provided in parenthesis. The polypeptide sequences are shown in SEQ ID NOS:2, 3, and 4. These homologs have 66%, 59%, and 62% identity, respectively, to AcrB SEQ ID NO:1 when the full-length sequences are aligned. One or more mutations, *e.g.*, a substitution at one or more positions selected from the positions in SEQ ID NO:2, 3, or 4 corresponding to a position within 2 to 8 Angstroms, or at, positions 279, 584, 617, 822, 927 or 1033 of SEQ ID NO:1, may be introduced into such homologs to obtain a variant that confers α -olefin tolerance, *e.g.*, 1-hexene tolerance, to a host cell. Further, one of skill understands that a structural model of an AcrB homolog such as SEQ ID NO:2, 3, or 4 can be generated and compared to the structure of AcrB to determine residues of SEQ ID NOS:2, 3, or 4 that correspond to the regions within 2 to 8 Angstroms of positions 279, 584, 617, 822, 927 or 1033 of SEQ ID NO:1. Such modified homologs may contain other mutations, *e.g.*, conservative substitutions, relative to the native sequence that do not have an effect on ability of the polypeptide to confer α -olefin tolerance to a host cell and do not otherwise impair pump activity.

[0068] In some embodiments, a native AcrB-like polypeptide modified in accordance with the invention to obtain a variant that confer α -olefin tolerance, *e.g.*, 1-hexene tolerance, to a host cell is *E. coli* AcrF (SEQ ID NO:5), which has 78% identity to SEQ ID NO:1. Thus, in some embodiments, an AcrB-like polypeptide of the invention may be a variant of SEQ ID NO:5 that has one or more substitutions at position 279, 584, 617, 734, 822, 927, or 1033 as numbered with reference to SEQ ID NO:1, *i.e.*, at position 279, 583, 616, 733, 821, 926, or 1032 of SEQ ID NO:5; or at one or more positions within 2 to 8 Angstroms of position 279, 583, 616, 821, 926, or 1032 of SEQ ID NO:5. Such modified homologs may contain other mutations, *e.g.*, conservative substitutions, relative to the native sequence that do not have an effect on ability of the polypeptide to confer α -olefin tolerance to a host cell and do not otherwise impair pump activity.

[0069] In a further aspect, the invention also provides a host cell that is modified to overexpress an AcrB-like polypeptide to confer α -olefin tolerance to the host cell. Such a host cell can be used to produce the α -olefin. For example, a host cell can be modified to overexpress a native AcrB-like polypeptide, *e.g.*, a polypeptide comprising the amino acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, or SEQ ID NO:5. In some embodiments, a host cell can be modified to overexpress a variant AcrB-like

polypeptide. In some embodiments, a host cell may be modified to overexpress an AcrB polypeptide of SEQ ID NO:1. In some embodiments, a host cell may be modified to overexpress a native AcrB-like homolog polypeptide, such as an AcrB-like polypeptide of SEQ ID NO:2, 3, or 4. In some embodiments, a host cell may be modified to over express a native AcrF polypeptide of SEQ ID NO:5.

[0070] In the context of this invention, the product of a gene is considered to be overexpressed when the level of protein is increased by at least 5%, at least 10%, at least 20%, at least 30%, or at least 50% or greater in comparison to the natural level of AcrB or AcrB-like polypeptide that is produced by a counterpart wild type host cell of the same genetic background that has not been genetically modified to overexpress the protein.

[0071] An α -olefin-tolerant host cell of the present invention typically exhibits at least 5%, more often at least 10%, at least 15%, at least 20%, at least 25%, or at least 50%, or greater survival compared to the counterpart host cell that comprises a wild type AcrB or AcrB-like polypeptide expressed at natural levels when tested in the presence of an α -olefin, *e.g.*, 1-hexene.

[0072] Tolerance to an α -olefin, *e.g.*, 1-hexene, can be assessed using known assays. For example, a gene encoding a variant of an AcrB polypeptide of SEQ ID NO:1 can be introduced into a host cell, *e.g.*, a bacterial host cell such as *E. coli*, and survival of the host cell in the presence of an α -olefin can be assessed. Illustrative assays are provided in the Examples section. In one illustrative assay, a plasmid-borne nucleic acid sequence encoding an AcrB-like variant polypeptide (along with a wild-type copy of *acrA*) is introduced into a Δ *acrAB E. coli* host cell and colonies are selected and grown overnight. A control strain may comprise a plasmid-borne copy of a nucleic acid sequence encoding the native *acrB* sequence (along with a wild type copy of *acrA*). In this example, the polynucleotide encoding the AcrB sequence is operably linked to an inducible promoter, such as a lac-inducible promoter. Dilutions are then prepared and the cells are plated onto agar plates with 1-hexene exposure using sufficient 1-hexene to saturate the atmosphere and expression of the variant protein (or for the control, wild type protein) is induced. An α -olefin tolerant strain expressing an AcrB-like variant shows greater growth in the presence of 1-hexene relative to strains expressing the wild type (control) sequence of AcrB at native levels. This phenomenon is more readily observed at lower cell dilutions in this assay.

Nucleic acids encoding a variant polypeptide

[0073] Isolation or generation of polynucleotide sequences to express a variant AcrB-like, polypeptide and/or to overexpress a native AcrB-like polypeptide of the invention can be accomplished by any number of techniques well known in the art, *e.g.*, using PCR as described herein. The sequences of native AcrB-like nucleic acids are known in the art and can be obtained using the protein accession numbers and sequences provided herein. Appropriate primers and probes can be generated from comparisons of the nucleic acid sequences to amplify a nucleic acid encoding the polypeptide sequence of interest. Alternatively, a library can be screened to obtain a nucleic acid encoding the polypeptide. A desired mutation can be introduced into a sequence using known techniques, *e.g.*, PCR followed by assembly of the gene by overlap extension.

[0074] Alternatively, a nucleic acid sequence encoding a non-naturally occurring AcrB-like polypeptide that confers enhanced α -olefin tolerance to a host cell can be generated using mutagenesis. For example, in some embodiments, one of skill can prepare a library of variants of an *acrB* gene, *e.g.*, by error prone PCR, and introduce the library into host cells, such as an *E. coli*. Transformants can then be evaluated for tolerance to an α -olefin and evaluated to identify mutations in the *acrB* gene that confer tolerance.

[0075] Nucleic acid sequences encoding an AcrB-like polypeptide that confers increased α -olefin tolerance, *e.g.*, increased 1-hexene tolerance, to a host cell, may additionally be codon-optimized for expression in a desired host cell. Methods and databases that can be employed are known in the art. For example, preferred codons may be determined in relation to codon usage in a single gene, a set of genes of common function or origin, highly expressed genes, the codon frequency in the aggregate protein coding regions of the whole organism, codon frequency in the aggregate protein coding regions of related organisms, or combinations thereof. See, *e.g.*, Henaut and Danchin in "Escherichia coli and Salmonella," Neidhardt, *et al.* Eds., ASM Pres, Washington D.C. (1996), pp. 2047-2066; Nucleic Acids Res. 20:2111-2118; Nakamura *et al.*, 2000, Nucl. Acids Res. 28:292.

Preparation of recombinant expression constructs

[0076] To use isolated sequences in the above techniques, recombinant DNA vectors suitable for transformation of host cells are prepared. Preparation of recombinant vectors is well known in the art. For example, a DNA sequence encoding an AcrB-like polypeptide of the invention can be combined with transcriptional and other regulatory sequences that direct

the transcription of the sequence from the gene in the intended cells. Control elements include promoters, optionally containing operator sequences, and other elements depending on the nature of the host. In some embodiments, an expression vector that comprises an expression cassette that comprises the gene also comprises a promoter operably linked to the gene. Such a promoter can be a promoter from a native gene that encodes AcrB or an AcrB homolog or can be a heterologous promoter. In other embodiments, a promoter and/or other regulatory elements that direct transcription of the gene are endogenous to the microorganism and an expression cassette comprising the gene encoding the AcrB-like polypeptide is introduced, *e.g.*, by homologous recombination, such that the heterologous gene is operably linked to an endogenous promoter and expression is driven by the endogenous promoter. In some embodiments, a promoter that is endogenous microorganism can also be genetically modified or replaced to enhance production of an AcrB-like polypeptide when the polynucleotide encoding the AcrB-like polypeptide is integrated in a host chromosome at a position such that the polynucleotide is operably linked to the promoter.

[0077] As noted above, expression of the gene encoding an AcrB-like polypeptide can be controlled by a number of regulatory sequences including promoters, which may be either constitutive or inducible; and, optionally, repressor sequences. Examples of suitable promoters, especially in a bacterial host cell, are the promoters obtained from the *E. coli* lac operon and other promoters derived from genes involved in the metabolism of other sugars, *e.g.*, galactose and maltose. Additional examples include promoters derived from biosynthetic enzymes such as for tryptophan (*trp*), the β -lactamase (*bla*), bacteriophage lambda PL, and T5. In addition, synthetic promoters, such as the *tac* promoter (U.S. Pat. No. 4,551,433), can be used. Further examples of promoters include *Streptomyces coelicolor* agarase gene (*dagA*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus subtilis* *xylA* and *xylB* genes. Suitable promoters are also described in Ausubel and Sambrook & Russell, both *supra*.

[0078] Suitable promoters of use in a yeast host cell include promoters obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH1, ADH2/GAP), *Saccharomyces cerevisiae* triose phosphate isomerase (TPI), *Saccharomyces cerevisiae* metallothionein (CUP1), and *Saccharomyces*

cerevisiae 3-phosphoglycerate kinase. Other useful promoters for yeast host cells are described by Romanos et al., 1992, *Yeast* 8: 423-488 or are otherwise known in the art.

[0079] An expression vector may also comprise additional sequences that influence expression of a gene encoding the AcrB-like polypeptide. Such sequences include enhancer sequences, a ribosome binding site, or other sequences such as transcription termination sequences, and the like.

[0080] A vector expressing a nucleic acid encoding an AcrB-like polypeptide of the invention may be an autonomously replicating vector, i.e., a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Thus, an expression vector may additionally contain an element(s) that permits integration of the vector into the host's genome.

[0081] An expression vector of the invention preferably contains one or more selectable markers which permit easy selection of transformed hosts. For example, an expression vector may comprise a gene that confers antibiotic resistance (e.g., ampicillin, kanamycin, chloramphenicol or tetracycline resistance) to the recombinant host organism, e.g., a bacterial cell such as *E. coli*. Suitable markers for other microbial host cells, such as yeast host cell are also well known and include, for example, ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3.

[0082] Although any suitable expression vector may be used to incorporate the desired sequences, readily available bacterial expression vectors include, without limitation: plasmids such as pSC101, pBR322, pBBR1MCS-3, pUR, pEX, pMR100, pCR4, pBAD24, pUC19, or plasmids derived from these plasmids; and bacteriophages, such as M13 phage and λ phage. Of course, such expression vectors may only be suitable for particular host cells. One of ordinary skill in the art, however, can readily determine through routine experimentation whether any particular expression vector is suited for any given host cell. For example, the expression vector can be introduced into the host cell, which is then monitored for viability and expression of the sequences contained in the vector.

[0083] Expression vectors of the invention may be introduced into the host cell using any number of well-known methods, including calcium chloride-based methods, electroporation, or any other method known in the art.

[0084] A nucleic acid encoding an AcrB-like polypeptide may optionally also be introduced into a host cell that expresses other components of an AcrAB-TolC efflux pump, or homologous tripartite pump. However, in some embodiments, an AcrB-like polypeptide functions independent of the tripartite pump. Accordingly, the AcrA and TolC, or the homologs of AcrA and TolC, need not be expressed in the host cell.

[0085] In some embodiments, the invention provides methods of producing an α -olefin compound, *e.g.*, 1-hexene, wherein the method employs a host cell, *e.g.*, *E. coli*, genetically modified to overexpress a naturally occurring AcrB or AcrB-like polypeptide, such as a polypeptide having a sequence of any one of SEQ ID NOS:1, 2, 3, 4, or 5, or a variant thereof. Such host cells can be obtained using known methods, such as those detailed above, for generating a construct comprising a nucleic acid encoding the naturally occurring AcrB or AcrB-like polypeptide that can be introduced into a host cell and expressed at the desired level.

[0086] In some embodiments, a promoter that influences expression of the native AcrB or AcrB-like polypeptide may be modified to increase expression. For example, an endogenous AcrB promoter may be replaced by a promoter that provides for increased expression compared to the native AcrB promoter.

[0087] In alternative embodiments, such a host cell that overexpresses a naturally occurring AcrB or AcrB-like polypeptide can be obtained by other techniques, *e.g.*, by mutagenizing a microorganism, *e.g.*, *E. coli* cells, and screening cells to identify a cell that expresses the native AcrB or AcrB-like polypeptide, *e.g.*, SEQ ID NO:1, 2, 3, 4, or 5, at a level higher than native (*e.g.*, at a level greater than that produced by the counterpart cell prior to mutagenesis).

Host Cells

[0088] The present invention provides for a modified host cell comprising a nucleic acid encoding any non-naturally occurring variant AcrB-like polypeptide as described herein. Further, the invention provides a genetically modified host cell that overexpresses a native AcrB polypeptide compared to counterpart host cells that express native levels of native AcrB; or that overexpresses a native AcrF or AcrB homolog *Pseudomonas putida* KT2440 polypeptide NCBI NP_743544, *Marinobacter aqueolei* EPL polypeptide NCBI YP_960752,

or *Pseudoalteromonas haloplanktis* TAC125 polypeptide NCBI YP_341810 compared to counterpart host cells that express native levels of the native AcrF or AcrB homolog.

[0089] In some embodiments, a host cell has further genetic modifications. For example, in some embodiments, an *E. coli* host cell may have a deletion of an endogenous *acrB* gene or *acrAB* gene locus, or an alternative bacterial host cell may have a deletion of a homologous gene or gene locus that corresponds to *acrB* or *AcrAB*.

[0090] In some embodiments, a host cell is genetically engineered to produce an α -olefin of interest, e.g., 1-hexene, or a derivative of the α -olefin. Thus, for example, a host cell may be engineered to express one or more genes that encode enzymes for production of an α -olefin compound, e.g., a C5-C12, α -olefin compound such as pentene or hexene (see, e.g., WO/2012/050931, WO/2012/109601, each incorporated by reference).

[0091] A host cell comprising an AcrB-like polypeptide as described herein is typically a microorganism, such as a bacterial or yeast host cell. In some embodiments, the host cell is a bacterial host cell, such as a Gram-negative bacterial host cell. Alternatively, the bacterial host cell need not be Gram-negative. In some embodiments of the invention, the host cell is a proteobacteria. In some embodiments of the invention, the bacterium is any bacterium that produces an α -olefin. In some embodiments of the invention, the bacterium is an enteric bacterium. In some embodiments of the invention, the bacterium is a species of the genus *Planctomyces*, *Bradyrhizobium*, *Rhodobacter*, *Rhizobium*, *Myxococcus*, *Klebsiella*, *Azotobacter*, *Escherichia*, *Salmonella*, *Pseudomonas*, *Caulobacter*, *Chlamydia*, *Acinetobacter*, *Acetobacter*, *Enterobacter*, *Sinorhizobium*, *Vibrio*, or *Zymomonas*. In some embodiments, the host cell is *E. coli*. In some embodiments, the host cell is *Pseudomonas putida*, e.g., *Pseudomonas putida* KT2440; *Marinobacter aqueolei*, e.g., *Marinobacter aqueolei* EP; *Pseudoalteromonas haloplanktis*, *Pseudoalteromonas haloplanktis* TAC125, or *Zymomonas mobilis*. In some embodiments, the host cells include species assigned to *Azotobacter*, *Erwinia*, *Bacillus*, *Clostridium*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Proteus*, *Serratia*, *Shigella*, *Staphylococcus*, *Streptococcus*, *Streptomyces*, *Vitreoscilla*, *Synechococcus*, *Synechocystis*, and *Paracoccus* taxonomical classes. In some embodiments, the host cells are cyanobacteria.

[0092] In some embodiments, the host cell is a yeast. Examples of yeast host cells include, without limitation, *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* host cells. In some embodiments, the yeast host cell is a

Saccharomyces carlsbergensis, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, or *Saccharomyces oviformis* cell. In some embodiments, the yeast host cell is a *Kluyveromyces lactis* cell. In another embodiment, the yeast host cell is a *Yarrowia lipolytica* cell.

Screening of host cells for α -olefin tolerance

[0093] Host cells can be screened for tolerance to an α -olefin as described herein. In a further aspect, the invention provides methods of screening for α -olefin-tolerant host cells.

[0094] In one embodiment, a high-throughput screening assay can be employed to evaluate α -olefin tolerance of a library of candidate cells that have been genetically modified, *e.g.*, by mutagenesis of an AcrB-like gene. In this embodiment, following transformation of a library, prepared in a vector where expression of the AcrB-like gene is under control of an inducible promoter, into suitable host cells, *e.g.*, *E. coli* K12 Δ acrAB, colonies are collected and inoculated into a small volume liquid culture, *e.g.*, 96 deep well plates, and cultured for sufficient time to allow the cells to reach stationary phase, *e.g.*, overnight at 37 °C with agitation. Cells are then diluted, *e.g.*, 100-fold, and an aliquot of the dilution is spotted onto growth medium agar plates supplemented with an inducer and exposed to 1-hexene. Exposure to 1-hexene is performed by saturating the atmosphere. This assay also preferably comprises a step in which duplicate plates are not exposed to 1-hexene. Growth of various dilutions of the cells can then be monitored to determine whether the cells exhibit increased tolerance to 1-hexene compared to control cells that have a wild type AcrB-like gene, or that have not been mutagenized. Colonies that exhibit increased growth in the presence of 1-hexene compared to the control cells exposed to 1-hexene are then selected and further evaluated. Optionally, the screening method may further comprise a “pre-screen” step in which library transformants are initially plated onto agar plates containing an inducer and are assessed for growth in the presence or absence of a saturating atmosphere of 1-hexene. Colonies that grow in the presence of 1-hexene compared to a control may then be selected, grown in a liquid culture, and the dilution analysis screening assay performed.

Methods of producing an α -olefin

[0095] A host cell comprising a variant AcrB-like polypeptide of the invention and/or that overexpresses an AcrB-like polypeptide, such as a native AcrB polypeptide or native AcrB polypeptide homolog, *e.g.*, a native polypeptide having a sequence of SEQ ID NO:1, 2, 3, 4, or 5, can be employed to produce an α -olefin compound or a derivative of an α -olefin

compound. In some embodiments, the host cell produces 1-hexene; or 1-butene, 1-pentene, 1-heptene, or 1-octene. To produce the α -olefin, a host cell genetically modified to express an AcrB-like polypeptide as described herein can be cultured under conditions suitable to allow expression of the polypeptide and expression of genes that encode the enzymes that are used to produce the α -olefin or a derivative of the α -olefin. A host cell modified to express an AcrB-like polypeptide as described herein provides a higher yield of α -olefin relative to a non-modified counterpart host cell that expresses a native AcrB or AcrB homolog at native levels. A host cell modified to overexpress a native AcrB or AcrB homolog such as a polypeptide having the amino acid sequence of SEQ ID NO:1, 2, 3, 4 or 5, also provides a higher yield of α -olefin relative to a non-modified counterpart host cell that expresses a polypeptide comprising SEQ ID NO:1, 2, 3, 4, or 5 at native levels.

[0096] The α -olefin can then be separated and purified using known techniques. As used herein, the term “purify,” “purified,” or “purification” means the removal or isolation of a molecule from its environment by, for example, isolation or separation from the cells and/or cell culture medium that is producing the α -olefin. As used herein, “purified” also refer to the removal of contaminants from a sample. For example, when α -olefins are produced in a host cell, the α -olefins can be purified by the removal of host cell proteins.

[0097] The present invention will be described in greater detail by way of specific examples. The following examples are offered for illustrative purposes, and are not intended to limit the invention in any manner. Those of skill in the art will readily recognize a variety of noncritical parameters, which can be changed or modified to yield essentially the same results.

EXAMPLES

Example 1. Toxicity of olefins and AcrB involvement

Toxicity of olefins, and AcrB involvement.

[0098] In this example, toxicity of olefin compounds was evaluated. A toxic dose was defined as the concentration of compound that decreased growth by half. This concentration for *E. coli* K12 is 65 mg/L, 110 mg/L and >1 g/L for 1-hexene, styrene and 1-octene, respectively (Figure 1A) for the media and growth conditions used. The compound 1-octene

is less toxic than the two other olefins tested (Figure 1B). The involvement of the tripartite pump, AcrAB-TolC in styrene or alpha-olefin tolerance was then assessed. *AcrA* and *acrB* mutants (*E. coli* K12 Δ *acrAB*) were constructed and transformed into a strain with a plasmid-borne *acrA* and *acrB* in an operon under the control of an arabinose inducible promoter (pABc) or a control plasmid (p0c). Exogenously added styrene was tested in liquid culture (Figure 2 A). The *E. coli* K12 Δ *acrAB* cells transformed to express AcrA and AcrB exhibited enhanced survival in the presence of styrene compared to those *E. coli* K12 Δ *acrAB* cells transformed with the control plasmid. A survival assay on agar plate in an atmosphere saturated with the volatile compound was developed for 1-hexene, 1-octene, and 1-nonene. For 1-hexene, 1-octene, and 1-nonene, the expression of the pump gave a strong advantage in the survival and growth of the strain (Figure 2 B).

AcrAB involvement on olefin production.

[0099] Among olefins used for plastic production, styrene is the only one that has been reported to be produced at levels that high enough to pose toxicity to *E. coli* growth (McKenna et al, 2011). Accordingly, the involvement of the AcrAB-TolC tripartite pump in growth in cells that produced styrene was evaluated. *AcrAB* was deleted in the L-phenylalanine over-producing *E. coli* strain (*E. coli* NST74 Δ *acrAB*) and then introduced on a plasmid with the plasmid encoding the styrene production genes. Two different IPTG induced promoters for *fdc1* and *pall* (P_{trc}, a strong promoter and P_{lacUV5}, a weak promoter, respectively) were used to generate p_{trc}Sty and p_{lac}Sty. These plasmids, and the corresponding control plasmids (p_{trc}0 and p_{lac}0) were introduced in the strain containing plasmids encoding the pump (pABc and control, p0c). It was found that *acrAB* was important for growth, and consequently production, in the styrene production strain. The absence of the pump had a small negative impact on growth when styrene production was not induced; however, the absence of the pump had a strong negative impact when styrene production was induced. Further, the quantity of styrene produced by the cells was significantly higher when the pump-containing plasmid was present. These data thus further confirmed that AcrAB-TolC played an important role in growth in the presence of a toxic compound such as styrene.

Improvement of tolerance by pump overproduction.

[0100] The results described above indicated that overexpression of the AcrAB-TolC pump provided an advantage in presence of 1-hexene. The plasmid pABc encoding AcrA and

AcrB (or the control p0c) and the plasmid pTolCa encoding the outer membrane protein TolC (or its control p0a) were introduced into the *E. coli* K12 wild type strain and the survival of the strain tested in presence of 1-hexene. As this alpha-olefin is extremely volatile, a survival assay on agar plates in an atmosphere saturated in 1-hexene (see methods) was employed. It was observed that over-expressing *acrA* and *acrB* and TolC in the wild type strain led to an increased tolerance to 1-hexene (Figure 3). Thus, expression of this pump over and above its native levels provided additional benefit towards this solvent-like compound. Figure 3 shows that the over expression of only TolC in the wild type does not improve the survival. However, the overexpression of the genes *acrA* and *acrB* led to a strain with 1-hexene tolerance greater than the wild type, suggesting than TolC is not the limiting protein of the complex and is in excess in comparison to AcrA and AcrB. This imbalance can be explained by the fact that TolC is also involved in other membrane complexes (Ramos et al, 2002). The results also show that there was an upper limit to inducing the expression the pump, and induction of pump expression by adding more inducer (10, 20, 30, 40, and 50 mM of arabinose) caused a reduction in growth (data not shown). Thus, while an increase in the levels of AcrAB-TolC improved tolerance to 1-hexene, further increase in tolerance to the olefin compound is achieved using a more efficient pump.

Evaluating RND efflux pumps from other organisms.

[0101] An existing library of heterologous efflux pumps in *E. coli* (Dunlop et al, 2011) was used to test if another pump may provide better tolerance to 1-hexene than the native *E. coli* AcrAB-TolC system. Using the agar plate-based method, this strain library constructed in a DH1 Δ *acrAB* background was tested and the survival compared in presence of 1-hexene relative to the strain expressing an equivalent plasmid-borne copy of the native *acrAB* genes. The strains containing pumps from *Pseudomonas putida* KT2440, *Marinobacter aqueolei* EPL and *Pseudoalteromonas haloplanktis* TAC125 showed equivalent tolerance to 1-hexene relative to the native *E. coli* *acrAB*-expressing strain (Figure 7).

Directed evolution on AcrB to improve tolerance to 1-hexene.

[0102] To identify AcrB variants that were more efficient at providing tolerance than the wild type AcrB, a high throughput screening using the 1-hexene saturated atmosphere assay (see, methods section) was developed. A library of AcrB variants was generated by introducing random mutations in the gene. *AcrB* variant genes were cloned in a low copy plasmid, in an operon with the *acrA* gene to maintain stoichiometry of the pump complex. The library of plasmids was introduced into *E. coli* K12 Δ *acrAB*. The library contained an

average of 2.2 amino acid mutations per gene. Initially, 2500 clones expressing an AcrB variant were tested. It was observed that 1.1 % of the clones had better survival in presence of 1-hexene relative to the strain encoding the wild type pump. A second library was then generated with a selection pressure directly after transformation, by incubating the freshly transformed cells in a 1-hexene atmosphere. In this condition, only 1 cell out of 8 survived and among the 500 cells tested in presence of 1-hexene, 9% were more tolerant than the strain expressing the wild type pump. The 20 most efficient variants were re-tested in duplicate, and 80% of them maintained better survival than the strain expressing the wild type AcrB, confirming the reliability of the screen. Plasmids were isolated from these strains and re-cloned them into fresh cells of the starting strain (K12 *ΔacrAB*). Of the freshly transformed cells, 63% maintained the increased tolerance relative to wild type, confirming that the 1-hexene tolerance phenotype tracked with the plasmids rather than in the genome of the mutant strain. Since most variants contained more than one mutation, the coding regions of the *acrB* variant genes were sequenced to identify the mutations. Each individual mutation was generated to determine the ones that are responsible for the improved phenotype. Six amino acid mutations in AcrB were identified and confirmed that provided improvement in tolerance over the wild type sequence: A279T, Q584R, F617L, L822P, F927S and F1033Y (Figure 4A). These mutations are localized all along the protein sequence (Figure 4 B and C). Among the beneficial mutations, A279T, Q584R, F617L and L822P were more efficient.

[0103] To determine if the mutations could act synergistically, selected mutations were combined in a single *acrB* gene generating 8 new variants with 2, 3 or 6 beneficial mutations (Var1: A279T and F617L, Var2: F617L and L822P, Var3: A279T and L822P, Var4: A279T and Q584R, Var5: A279T, F617L and L822P, Var6: A279T, Q584R and F617L, Var7: A279T, Q584R and F617L, Var8: A279T, Q584R, F617L, L822P, F927S and F1033Y). Most of the strains containing the AcrB with two beneficial mutations were found to be more tolerant to 1-hexene relative to an AcrB with one mutation (Figure 6). However the combination of three or more mutations in the same protein did not result in additional advantage. It was also determined that the level of the AcrB variants in the soluble fraction of *E. coli* was less relative to the AcrB wild type (data not shown). Additionally, among the AcrB mutants, an increase in the number of beneficial mutations results in a decrease in the levels of the protein in the soluble fraction. These data suggest that the improvement in tolerance is likely due to an improvement in the efficiency or in the mode of action of the pump rather than due to improved protein production or stability.

[0104] The evolved AcrB strains were evaluated in liquid cultures containing 1-hexene, bile salts, or styrene. Growth of selected strains containing the AcrB variants in presence of different concentration of toxic olefins (Figure 5) was analyzed. In presence of 1-hexene in the liquid cultures, it was observed that several strains producing evolved AcrB variants grew better than the wild type (Figure 5A) confirming that these mutations were not specific to the agar plate assay. In presence of bile salts, the mutation Q584R improved the growth of the strain. None of the evolved AcrB variants that were tested provided benefit in presence of styrene relative to wild type AcrB (Fig 5B). In contrast the evolved AcrB variants conferred tolerance to 1-hexene, relative to wild type AcrB (Fig 5A). Thus, although AcrAB-TolC had been shown to be important to the growth of cells that produce styrene and the strains containing the AcrB variants were not tolerant to the olefin styrene, surprisingly, the strains exhibited tolerance to α -olefins, *e.g.*, hexene.

Materials and Methods

Media, Chemicals, and Strains

[0105] Chemicals: styrene (99%), 1-hexene (99%), 1-octene (98%), nonene (96%), isopropyl b-D-1-thiogalactopyranoside (IPTG), dodecane, arabinose, bile salts, chloramphenicol (Cm), kanamycin (Km), ampicillin (Amp) are from Sigma-Aldrich (St. Louis, MO); dimethylformamide (DMF) is from Fisher Scientific (Hampton, NH); and glucose is from VWR (Westchester, PA).

[0106] All strains, plasmids and oligonucleotides are listed in Table 1. Oligonucleotides synthesis and sequencing were conducted at Integrated DNA Technologies (Coralville, IA) and Quintara (Albany, CA) respectively. Strains were cultured in Luria-Bertani (LB) broth supplemented with appropriate antibiotics, or inducers as necessary, and with 15 g/L agar (BD, Sparks, MD) for plate culture and assays.

[0107] *E. coli* DH10B (Table 1) was used for sub-cloning, *E. coli* K12 BW25113 (Baba et al., 2001) was used for toxicity assay and *E. coli* NST74 was used for styrene production. *Arabidopsis thaliana* (ecotype Col-0) cDNA and *Saccharomyces cerevisiae* S288C were used to amplify genes required for styrene production: PAL2 and FDC1 (McKenna et al, 2011). The efflux pump library containing 43 different pumps from heterologous sources was in an *E. coli* DH1 Δ acrAB background (Dunlop et al, 2011).

Construction of Gene Deletion Mutants

[0108] The *E. coli* K12 BW25113 Δ *acrAB* and *E. coli* NST74 Δ *acrAB* strains were constructed using the one-step inactivation of chromosomal genes using published methods (Datsenko et al, 2000). In order to generate a Δ *acrAB* strain, the *E. coli* K12 BW25113 Δ *acrA* was used. The chromosomally-encoded kanamycin (Km) marker in the Δ *acrA* strain was removed using the pCP20, and the plasmid pKD46 was introduced into the resulting strain. The PCR product used to mediate gene replacement for *acrB* was amplified from the *E. coli* K12 BW25113 Δ *acrB* strain using the primers DelB F and DelB R and introduced in the Km-free *E. coli* K12 BW25113 Δ *acrA* strain (pKD46). Km^R colonies were tested for the loss of *acrB*. After curing the pKD46, the Km^R gene was removed generating the strain *E. coli* K12 Δ *acrAB*. For the *E. coli* NST74 Δ *acrAB*, the PCR used to mediate the replacement of *acrA* and *acrB* in the *E. coli* NST74 was amplified from *E. coli* K12 BW25113 Δ *acrAB* Km^R using the primers DelAB F and DelAB R, and introduced in the *E. coli* NST74 carrying pKD46. Km^R colonies were tested for the loss of operon encoding *acrA* and *acrB*. Here again, after curing the pKD46, the Km^R gene was removed generating the final strain *E. coli* NST74 Δ *acrAB*.

Construction of plasmids

[0109] High fidelity phusion polymerase (Thermo Scientific, Waltham, MA) was used for all gene amplifications, except for generating the *acrB* variant library. All restriction enzymes were fast-digest enzymes from Thermo Scientific (Waltham, MA). *tolC* was amplified using the primers TolC F and TolC R and cloned in the pBbS5a-RFP (p0a) with the restriction sites *NdeI* and *XhoI* generating the plasmid pTolCa. *acrA* was amplified using the primers AcrA F and AcrA R and cloned into the pBbA8c-RFP (p0c) using the restriction sites *BglII* and *NdeI* generating the plasmid pAc. *acrB* was amplified using the primers AcrB F and AcrB R and cloned in pAc using the restriction site *NdeI* and *XhoI* generating the plasmid pABc.

[0110] Genes encoding PAL2 and FDC1, required for styrene production, were amplified from *Arabidopsis thaliana* and *Saccharomyces cerevisiae* S288C cDNA respectively.

[0111] To generate the *acrB* variant library by error prone PCR, Dream Taq polymerase (Thermo Scientific, Waltham, MA) was used. PCR products were obtained by amplifying *acrB* from pABc (0.5 ng/ μ L), using the primers Cam F and Cam R (0.1 μ M), the kit buffer (containing MgCl₂ at 2 mM final concentration), dNTP mix at 0.2 mM (Thermo Scientific,

Waltham, MA), 1 % DMSO and 0.1 unit/ μ L. No extra Mn^{2+} was used to induce mutation. Twenty cycles were performed resulting in a pool of mutated *acrB* sequences containing 2.2 amino acid mutations, on an average. PCR fragments were gel-extracted, cloned using the Gibson cloning method (Gibson et al, 2009) into pAc digested with *NdeI* and *XhoI*, and then introduced into the *E. coli* K12 Δ *acrAB* strain.

[0112] The *acrB* variants with desired mutations were generated by overlap PCR using the high fidelity phusion polymerase (Thermo Scientific, Waltham, MA). For each single mutation variant, two PCR products were obtained by amplification from the pABc vector as listed in Table 1: PCR 1 and 2 were extracted for every variant and used to generate long fragments containing the full length genes. The resulting products were introduced with the Gibson assembly method (Gibson et al, 2009) into pAc digested by *NdeI* and *XhoI*, and transformed in *E. coli* DH10B. Plasmids with confirmed sequences were then introduced into the *E. coli* K12 Δ *acrAB* strain. The variants with multiple mutations were obtained by introducing additional mutation(s) in a given variant, using the same method described previously (overlapping PCR products). PyMOL ([www website pymol.org/](http://www.pymol.org/)) was used to visualize the structure of AcrB and the localization of the individual mutations.

Growth impact of AcrA-acrB overexpression.

[0113] Colonies of *E. coli* K12 Δ *acrAB* containing the plasmid pABc were used to inoculate a 6 mL LB-Cm pre-culture tube. 100 μ L of cells at $OD_{600}=1$ were used to inoculate 10 mL LB-Cm supplemented with glucose (2 g/L) and various concentration of arabinose (0, 5, 10, 20, 30, 40 and 50 mM). Cell growth at 37 °C was followed by OD_{600} measurements using a spectrophotometer (Beckman Coulter DU800, Brea, CA).

Impact of alpha-olefins and styrene on growth.

[0114] Cells were first inoculated in 6 mL LB-Cm with glucose (2 g/L) and arabinose (5 mM). At $OD_{600}=1$ at 37 °C, 100 μ L of culture were transferred to 15 mL glass culture tubes (VWR, Westchester, PA) with a septum lid (VWR, Westchester, PA) containing 10 mL of LB, supplemented with the appropriate antibiotic and inducer. Septum lids were required to limit evaporation of chemicals. DMF was used as cosolvent for all three compounds to prepare stock solutions at different concentrations (styrene at 10, 12, 14 and 16 g/L, 1-hexene at 5, 6, 7 and 8 g/L and 1-octene at 10, 20, 50 and 100 g/L). At an OD_{600} of 0.2, 100 μ L of the olefin stock solutions was added. After 2 hours at 37 °C, OD_{600} was measured. All measurements are in duplicate.

[0115] For agar plate based assays, cells were first streaked overnight and few colonies were used to inoculate a preculture of 6 mL of LB-Cm with glucose (2 g/L) and arabinose (5 mM). At $OD_{600}=1$ at 37 °C, the cultures were used to prepare a dilution series. Dilution 0 (D0) corresponds to the $OD_{600}=1$, and contains $\sim 10^9$ cells/mL. Dilution 1 (D1) is a 10-fold dilution of D0, dilution 2 (D2) a 10-fold dilution of D1 up to dilution 5 (D5), which contains $\sim 10^4$ cells/mL. 10 μ L of each dilution was spotted in triplicate on LB-agar plates (single well plates, Thermo Scientific, Waltham, MA) containing 30 mL of LB-agar, Cm (30 mg/L), glucose (2 g/L) and arabinose (5 mM). Small containers (microfuge tube lids) were placed on the open lid of the agar plate and filled with 400 to 600 μ L of 1-hexene or 1-octene or with 75 μ L of 1-nonene. The agar plates were then inverted over the lid, cells facing the olefin compound, sealed with parafilm, placed in the plastic bag (the original packaging for the petri dishes) and incubated for 2 days at room temperature. For 1-octene, a second condition, via overlaying 2 mL of compound on the top of the agar submerging the cells, was used. TolC overexpression impact was examined using the same method, but in this case, the LB media and the LB-agar were also supplemented with ampicillin (100 mg/L) and IPTG (100 μ M).

[0116] For the high throughput screening of AcrB variants, two methods were used. After transformation of the library in *E. coli* K12 Δ acrAB, cells were plated on LB-agar-Cm plate either without a 1-hexene selection pressure or with 1-hexene exposure using 300 μ L to saturate the atmosphere. Clones were collected and inoculated in 96 deep well plates (USA Scientific, Ocala, FL) using 800 μ L LB (supplemented with chloramphenicol 30 mg/L, glucose 2 g/L, arabinose 5 mM) and cultured overnight at 37 °C with agitation. Cells were then diluted 100 fold and 5 μ L of the dilution was spotted on agar plates (Thermo Scientific, Waltham, MA) containing 25 mL of LB-agar-Cm and inducer, using a liquid handler robot (Biomek FX, Beckman Coulter, Brea, CA). 600 μ L of 1-hexene was used per plate to create a saturated atmosphere. Plates were sealed and incubated for 2 days as described in the previous paragraph.

[0117] To test the heterologous efflux pump library, the same method was used. Clones were grown in 96 DWP overnight at 37 °C in presence of IPTG 10 μ M, and 5 μ L of a 100x dilution of the overnight culture was spotted on agar plate using the liquid handler robot. 600 μ L of 1-hexene was used to generate a toxic atmosphere.

Measuring the toxic effect of styrene

[0118] *E. coli* NST74 Δ *acrAB* containing pABc or p0c and either p_{lac}Sty, p_{trc}Sty, p_{lac}0 or p_{trc}0, were inoculated into 6 mL LB (supplemented with 15 g/L glucose, and 10 mM arabinose and appropriate antibiotics). At OD₆₀₀ =1, 5 μ L cells were inoculated in 95 μ L of LB-antibiotic, 15 g/L glucose, 10 mM arabinose, biolog redox dye (Mix A) solution (Biolog, Hayward, CA) and various concentration of IPTG (0, 10, 20, 50 100, 200 and 300 μ M). The experiment was performed in 96 well plates (Biolog, Hayward, CA), at 37 °C and growth was monitored using the Omnilog system (Biolog, Hayward, CA). In this condition all the strains containing p0c (no AcrAB production) did not grow. For the strains possessing the p_{trc}Sty, the impact of styrene production on growth was also tested on LB-agar plate containing 25 mg/L Km, 15 g/L glucose, 10 mM arabinose and 100 μ M IPTG). For glass tubes based growth of *E. coli* NST74 Δ *acrAB* (with pABc or p0c and either p_{lac}Sty or p_{lac}0), colonies were inoculated in 10 mL of LB (supplemented with 100 mg/L ampicillin, 15 g/L glucose, 10 mM arabinose) and grown overnight at 37 °C. From this culture, cells were inoculated in glass tubes (starting OD₆₀₀ of 0.02) containing 10 mL of LB (supplemented with 100 mg/L ampicillin, 15 g/L glucose, 10 mM arabinose and 100 μ M IPTG. Growth was monitored using a spectrophotometer (Beckman Coulter DU800, Brea, CA).

Western blotting

[0119] Cultures were incubated at 37 °C in LB-Cm, 2 g/L glucose and 5 mM arabinose. At OD₆₀₀ =1, the pellet of 2 mL of culture was resuspended and incubated for 15 min in 30 μ L of B-Per buffer (Thermo Scientific, Waltham, MA) supplemented with 1 mg/ml lysozyme (Sigma-Aldrich, St. Louis, MO). Each sample was sonicated (SP Scientific, Gardiner, NY) for 5 seconds at the lowest power setting at 4 °C, then centrifuged at 4 °C for 30 min at 15 rpm. Total protein quantity present in the soluble fraction was determined using the DC Lowry Reagent according the protocol of the supplier (Bio-Rad, Hercules, CA). All samples were normalized to the same quantity of total soluble protein. 20 μ L of normalized samples was mixed with SDS loading dye buffer (Life technologies, Carlsbad, CA) and DTT (Life technologies, Carlsbad, CA) and incubated at 98 °C for 20 min. 10 μ L of samples were run on 4-12% Bis-Tris gel (Life technologies, Carlsbad, CA) at 165 V under SDS MES Buffer (Life technologies, Carlsbad, CA). Proteins were transferred to a PVDF membrane (Life technologies, Carlsbad, CA) using iBlot transfer system (Life technologies, Carlsbad, CA). The membrane was washed in PBS buffer (20 mM NaPhosphate pH=7.4, 50 mM NaCl) and blocked overnight at 4 °C with 25 mL of PBS-Tween-20 (EMD Millipore, Billerica, MA) supplemented with 3% of BSA (Sigma-Aldrich, St. Louis, MO). The primary antibody is a

monoclonal mouse anti-His (Sigma-Aldrich, St. Louis, MO) diluted 1:5000, the secondary antibody is a goat anti-mouse HRP conjugated (Sigma-Aldrich, St. Louis, MO) diluted 1:10,000. Both antibodies were suspended in 1% BSA and PBS-Tween-20, and consecutively added to the membrane for 1 hour at room temperature. Three washes in PBS-Tween20 were performed after incubation of each antibody. Membrane was then incubated in 2 mL of HRP detection solution (Bio-Rad, Hercules, CA) for 10 min.

Summary of illustrative examples

[0120] In these examples, it was demonstrated that the tripartite RND efflux pump in *E. coli* composed of AcrA, AcrB and TolC subunits played a major role in the tolerance of *E. coli* cells to alpha-olefins, and is important in reaching high production levels of these toxic compounds. *E. coli* has been engineered to manufacture many bulk chemicals (Nakamura and Whited 2003; Yim et al. 2011). In these, examples, the benefit of increasing the levels of the AcrAB-TolC efflux pump was first evaluated. Increasing the expression of AcrAB-TolC had a positive impact on tolerance to 1-hexene, although in these experiments, this was limited by the toxicity linked to overexpression of membrane proteins (Wagner et al. 2007). Directed evolution of AcrB was then used to improve tolerance. Two recent studies have used directed evolution to improve AcrB tolerance to n-butanol (Fisher et al. 2013) and α -pinene and n-octane (Foo and Leong 2013).

[0121] These examples further illustrated that engineering the AcrB subunit of this efflux pump improved tolerance to 1-hexene. AcrB variants having improved tolerance to 1-hexene were identified and the primary domains responsible for 1-hexene tolerance were characterized. It has been shown that the entrance of substrates into the efflux pump and the amino acids involved in the substrate binding depends on the properties and the size of the compounds (Eicher et al. 2012; Takatsuka et al. 2010). To better understand the domain involved in 1-hexene tolerance, several beneficial mutations obtained in the first round of evolution. Several mutations (A279T, Q584R, F617L, L822P, F927S and F1033Y) were identified that resulted in improved tolerance to 1-hexene. These mutations had an additive effect, but in some experiments, it was also observed that multiple changes may be disruptive for protein solubility. Most of the AcrB variants in these experiments resulted in decreased protein in the soluble fraction, which indicates the improvement in the tolerance conferred by these variants is due to higher pump efficiency.

[0122] Crystallographic and site-directed mutagenesis studies for AcrB have determined the role of domains and of several amino acids in the protein (Husain and Nikaido 2010; Murakami et al. 2002). Not to be bound by theory, the crystal structure and the data available provide a basis to analyze the effect of beneficial mutations. Of the many amino acid changes identified, F927S and more particularly F1033Y were in the most unexpected positions. F927S is localized in the transmembrane domain, at the top of the TM10 α -helix. This helix contains an important amino acid (K940) involved in the proton transport (Su et al. 2006) and a mutation in this helix could impact the rotational movements leading to compound export. However the mutations A279T, Q584R, F617L and L822P were found in positions known to be important for pump function. The amino acid L822 is positioned between the 2 β -sheets C β 13 belonging to the pore domain PN1 and C β 14 from the PC2. The mutation of a leucine to a proline is localized between these 2 β -sheets at the “ceiling” of the vestibule, suggested to be a highly probable substrate entrance point, and could change the flexibility and/or the opening of the vestibule facilitating 1-hexene entrance into the pore. The amino acid A279 is localized at the binding pocket in which residues Glu273, Asn274, Asp276, Ile277, play an important roles (Husain et al. 2011). Q584 is localized in a position such that it may be involved in the trimer assembly. AcrB subunit is reported to fold independently, and then assemble into a trimer (Yu et al. 2011). This mutation is also clearly beneficial to the growth of *E. coli* in presence of bile salts confirming that its role is not specific to 1-hexene, although in these experiments, we could not detect any positive impact in presence of styrene. Finally, the amino acid F617 has been shown to be located in the switch loop of the hydrophobic binding pocket and reported crystal structures suggest that this amino acid could directly interact with various substrates (Bohnert et al. 2008; Eicher et al. 2012; Vargiu et al. 2011). Mutating this amino acid F617 to an alanine has been reported to have a direct impact on substrate uptake and was responsible for a substantial decrease in transport of novobiocin, but a minor effect on the transport of oxacillin and various other macrolides (Bohnert et al. 2008). In this study, a mutation to a leucine at this position had a positive impact on the transport of 1-hexene.

[0123] The experiments performed for this study, as well as other reports in the literature (Bohnert et al. 2008) suggest that the AcrB protein is destabilized when more than three mutations are introduced in the sequence. However, additional rounds of mutations, starting with variants with one or two mutations, could stabilize the protein and allow the introduction of more mutations to further improve function. Additionally, analyses of more variants

showing a positive impact for 1-hexene tolerance could give a better understanding of the path(s) of this compound through the pump. Directed evolution of AcrB has also been used to improve tolerance and secretion of other compounds. It has also been shown that AcrB can be improved for n-butanol, α -pinene and n-octane (Fisher et al. 2013; Foo and Leong 2013). These studies resulted in beneficial mutations in different positions relative to the ones identified in our study. Specifically, N189H, T678S, Q737L and M844L improved the efflux of α -pinene, n-octane, while I466F, M355L, and S880P improved the growth on n-butanol. These authors proposed corresponding improvements in the AcrB-TolC interaction and enlargement in the entrance of the cleft that facilitate conformational changes or improve the affinity for the substrate. The results provided in the illustrative experiments presented in these examples thus also suggest that several domains can be targeted for mutations to lead to a more efficient pump.

Table 1

strains	Description
<i>E. coli</i> DH10B	<i>ara</i> Δ139 Δ(<i>ara</i> , <i>leu</i>)7697 <i>fhuA lacX74 galK16 galE15 mcrA f80d(lacZ</i> Δ <i>M15)</i> <i>recA1 relA1 endA1 nupG rpsL rph spoT1Δ(mrr-hsdRMS-mcrBC)</i>
<i>E. coli</i> K12 BW25113	<i>lacIq rrnBT14 ΔlacZw16 hsdR514 ΔaraBA-DAH33 ΔrhaBADLD78</i>
<i>E. coli</i> NST74	<i>aroH367, tyrR366, tna-2, lacY5, aroF394(fbr), malT384, pheA101(fbr),</i> <i>pheO352, aroG397(fbr)</i>
<i>E. coli</i> K12 BW25113 Δ <i>acrA</i>	<i>lacIq rrnBT14 ΔlacZw16 hsdR514 ΔaraBA-DAH33 ΔrhaBADLD78 ΔacrA KmR</i>
<i>E. coli</i> K12 BW25113 Δ <i>acrB</i>	<i>lacIq rrnBT14 ΔlacZw16 hsdR514 ΔaraBA-DAH33 ΔrhaBADLD78 ΔacrB KmR</i>
<i>E. coli</i> K12 BW25113 Δ <i>acrAB</i>	<i>lacIq rrnBT14 ΔlacZw16 hsdR514 ΔaraBA-DAH33 ΔrhaBADLD78 ΔacrAB KmR</i>
<i>E. coli</i> NST74 Δ <i>acrAB</i>	<i>aroH367, tyrR366, tna-2, lacY5, aroF394(fbr), malT384, pheA101(fbr),</i> <i>pheO352, aroG397(fbr)</i>
<i>Saccharomyces cerevisiae</i> S288C	source of <i>fdcI</i>
strain library of heterologous pumps	Description
<i>E. coli</i> DH1 Δ <i>acrAB</i>	<i>endA1 recA1 gyrA96 thi-1 glnV44 relA1 hsdR17(rK- mK+) λ- ΔacrA KmR</i>
cDNA	Description
<i>Arabidopsis thaliana</i> (ecotype Col-0)	source of <i>pul2</i>
plasmids	Description
pCP20	oriR101 w/repA101ts, AmpR CmR, flp recombinase

primers (5'-3')	sequence (SEQ ID NOS:6-33)
DelB F	CTGAACAGTCCAAAGTCTTAACTTAAACAGGAGC
DelB R	CATACTAGCACTAAACCCGCAGCAGG
DelAB F	CGTGCCATATGTTGCTGAATTTACAGCG
DelAB R	CATACTAGCACTAAACCCGCAGCAGG
TolC F	AAAGATCTTTTAAAGAAGGAGATATACATATGAAGAATGCTCCCCATTTATATGGCC
TolC R	ACTCGAGGGATCCTCAGTGGTGGTGGTGGTGAACGGAAAGGTTATGACCGTTAC
AcrA F	TAAAGATCTTTTAAAGAAGGAGATATACAGATGAACAAAAACAGAGGGTTTACGCC
AcrA R	AAAACATATGTTAAACGGCTCCGTTTAAGTTAAGACTTGGACTGTTTCAGGCTGAGC
AcrB F	CCAAAGTCTTAACTTAAACAGGAGCGGTTAACATATG
AcrB R	GATCCTTACTCGAGGGATCCCTCAGTGG
FdcI F	AAAGATCTTTTAAAGAAGGAGATATACATATGAGGAAGCTAAATCCAGCTTTAGAAATTAG
FdcI R	TCCTTCTTAAACCCGGTCCCTTATTTATATCCGTACCTTTTCCAAATTTTCATTACTTTG
Pal2 F	ATAAGGACCCGGGTTTAAAGAAGGAGATATACAAATCGATCAAAATCGAAGCAATGTTGTG
Pal2 R	CTTACTCGAGTTAGCAAAATCGGAATCGGAGCTCCGTTTC
Cam F	CGGTGCTCAGCCTGAACAGTCCCAAGTC

Cam R	TCTTTCGACTGAGCCTTTCGTTTATTGATGCC
A279T R	GCCGTTAAACTCTGTGATGATGTCGTAG
A279T F	CTACGACATCATCACAGAGTTTAAACGGC
Q584R R	GTGTACGTTCCCGCGTTGCACC
Q584R F	GGTGCAACCGGGGAACGTACAC
F617L R	CTGACCACTGCCCGCAAGGCCGAAGCCGTTAACG
F617L F	CGTTAAACGGCTTCGGCCTTGCGGGACGTGGTCAG
L822P R	CATGGATGGCGGCGCGTTGTAAC
L822P F	GTTACAAAGGCCCGCCATCCATG
F927S R	GGCCTACCTGGGAGTAAACGTC
F927S F	GACGTTTACTCCAGGTAGGCC
F1033Y R	CATTCTTGGGCTATAGCGGGGCG
F1033Y F	GCCGCCGTATAGCCGCAAGATG

AcrB variants	AcrB variants have been made by overlapping 2 PCRs fragments below
variant A279T	PCR1 (primers: Cam F + A279T R) + PCR2 (primers: A279T F + Cam R)
variant Q584R	PCR1 (primers: Cam F + Q584R R) + PCR2 (primers: Q584R F + Cam R)
variant F617L	PCR1 (primers: Cam F + F617L R) + PCR2 (primers: F617L F + Cam R)
variant L822P	PCR1 (primers: Cam F + L822P R) + PCR2 (primers: L822P F + Cam R)
variant F927S	PCR1 (primers: Cam F + F927S R) + PCR2 (primers: F927S F + Cam R)
variant F1033Y	PCR1 (primers: Cam F + F1033Y R) + PCR2 (primers: F1033Y F + Cam R)

Reference listing for references cited by author and publication year

- Bohnert JA, Schuster S, Seeger MA, Fahrnich E, Pos KM, Kern WV. 2008. Site-directed mutagenesis reveals putative substrate binding residues in the *Escherichia coli* RND efflux pump AcrB. *Journal of bacteriology* 190(24):8225-9.
- Doshi R, Nguyen T, Chang G. 2013. Transporter-mediated biofuel secretion. *Proceedings of the National Academy of Sciences of the United States of America* 110(19):7642-7.
- Dunlop MJ, Dossani ZY, Szmidt HL, Chu HC, Lee TS, Keasling JD, Hadi MZ, Mukhopadhyay A. 2011. Engineering microbial biofuel tolerance and export using efflux pumps. *Molecular Systems Biology* 7:487.
- Dunlop MJ, Jay D, Keasling J A, Mukhopadhyay A. 2010. A model for improving microbial biofuel production using a synthetic feedback loop *Systems and Synthetic Biology*.
- Eicher T, Cha HJ, Seeger MA, Brandstatter L, El-Delik J, Bohnert JA, Kern WV, Verrey F, Grutter MG, Diederichs K and others. 2012. Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop. *Proceedings of the National Academy of Sciences of the United States of America* 109(15):5687-92.
- Fisher MA, Boyarskiy S, Yamada MR, Kong N, Bauer S, Tullman-Ercek D. 2013. Enhancing Tolerance to Short-Chain Alcohols by Engineering the *Escherichia coli* AcrB Efflux Pump to Secrete the Non-native Substrate n-Butanol. *ACS synthetic biology*.
- Foo JL, Leong SS. 2013. Directed evolution of an *E. coli* inner membrane transporter for improved efflux of biofuel molecules. *Biotechnology for biofuels* 6(1):81.
- Husain F, Bikhchandani M, Nikaido H. 2011. Vestibules are part of the substrate path in the multidrug efflux transporter AcrB of *Escherichia coli*. *Journal of bacteriology* 193(20):5847-9.
- Husain F, Nikaido H. 2010. Substrate path in the AcrB multidrug efflux pump of *Escherichia coli*. *Molecular microbiology* 78(2):320-30.
- Jarboe LR, Zhang X, Wang X, Moore JC, Shanmugam KT, Ingram LO. 2010. Metabolic engineering for production of biorenewable fuels and chemicals: contributions of synthetic biology. *Journal of biomedicine & biotechnology* 2010:761042.
- McKenna R, Nielsen DR. 2011. Styrene biosynthesis from glucose by engineered *E. coli*. *Metabolic engineering* 13(5):544-54.
- Murakami S, Nakashima R, Yamashita E, Yamaguchi A. 2002. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* 419(6907):587-93.
- Murakami S, Tamura N, Saito A, Hirata T, Yamaguchi A. 2004. Extramembrane central pore of multidrug exporter AcrB in *Escherichia coli* plays an important role in drug transport. *The Journal of biological chemistry* 279(5):3743-8.
- Murakami S, Yamaguchi A. 2003. Multidrug-exporting secondary transporters. *Current opinion in structural biology* 13(4):443-52.
- Nakamura CE, Whited GM. 2003. Metabolic engineering for the microbial production of 1,3-propanediol. *Current opinion in biotechnology* 14(5):454-9.
- Nikaido H. 2009. Multidrug resistance in bacteria. *Annual review of biochemistry* 78:119-46.
- Nikaido H, Takatsuka Y. 2009. Mechanisms of RND multidrug efflux pumps. *Biochimica et biophysica acta* 1794(5):769-81.

- Ramos JL, Duque E, Gallegos MT, Godoy P, Ramos-Gonzalez MI, Rojas A, Teran W, Segura A. 2002. Mechanisms of solvent tolerance in gram-negative bacteria. Annual review of microbiology 56:743-68.
- Schubert P, Steinbuchel A, Schlegel HG. 1988. Cloning of the *Alcaligenes eutrophus* genes for synthesis of poly-beta-hydroxybutyric acid (PHB) and synthesis of PHB in *Escherichia coli*. Journal of bacteriology 170(12):5837-47.
- Seeger MA, Schiefner A, Eicher T, Verrey F, Diederichs K, Pos KM. 2006. Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. Science 313(5791):1295-8.
- Seeger MA, von Ballmoos C, Eicher T, Brandstatter L, Verrey F, Diederichs K, Pos KM. 2008. Engineered disulfide bonds support the functional rotation mechanism of multidrug efflux pump AcrB. Nature structural & molecular biology 15(2):199-205.
- Sennhauser G, Amstutz P, Briand C, Storchenegger O, Grutter MG. 2007. Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors. PLoS biology 5(1):e7.
- Sikkema J, de Bont JA, Poolman B. 1995. Mechanisms of membrane toxicity of hydrocarbons. Microbiological reviews 59(2):201-22.
- Su CC, Li M, Gu R, Takatsuka Y, McDermott G, Nikaido H, Yu EW. 2006. Conformation of the AcrB multidrug efflux pump in mutants of the putative proton relay pathway. Journal of bacteriology 188(20):7290-6.
- Takatsuka Y, Chen C, Nikaido H. 2010. Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America 107(15):6559-65.
- Takatsuka Y, Nikaido H. 2007. Site-directed disulfide cross-linking shows that cleft flexibility in the periplasmic domain is needed for the multidrug efflux pump AcrB of *Escherichia coli*. Journal of bacteriology 189(23):8677-84.
- Takatsuka Y, Nikaido H. 2009. Covalently linked trimer of the AcrB multidrug efflux pump provides support for the functional rotating mechanism. Journal of bacteriology 191(6):1729-37.
- Takatsuka Y, Nikaido H. 2010. Site-directed disulfide cross-linking to probe conformational changes of a transporter during its functional cycle: *Escherichia coli* AcrB multidrug exporter as an example. Methods in molecular biology 634:343-54.
- Tikhonova EB, Yamada Y, Zgurskaya HI. 2011. Sequential mechanism of assembly of multidrug efflux pump AcrAB-TolC. Chemistry & biology 18(4):454-63.
- Tseng TT, Gratwick KS, Kollman J, Park D, Nies DH, Goffeau A, Saier MH, Jr. 1999. The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. Journal of molecular microbiology and biotechnology 1(1):107-25.
- Vargiu AV, Collu F, Schulz R, Pos KM, Zacharias M, Kleinekathofer U, Ruggerone P. 2011. Effect of the F610A mutation on substrate extrusion in the AcrB transporter: explanation and rationale by molecular dynamics simulations. Journal of the American Chemical Society 133(28):10704-7.
- Vargiu AV, Nikaido H. 2012. Multidrug binding properties of the AcrB efflux pump characterized by molecular dynamics simulations. Proceedings of the National Academy of Sciences of the United States of America 109(50):20637-42.

- Wagner S, Baars L, Ytterberg AJ, Klussmeier A, Wagner CS, Nord O, Nygren PA, van Wijk KJ, de Gier JW. 2007. Consequences of membrane protein overexpression in *Escherichia coli*. Molecular & cellular proteomics : MCP 6(9):1527-50.
- Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, Khandurina J, Trawick JD, Osterhout RE, Stephen R and others. 2011. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. Nature chemical biology 7(7):445-52.
- Yu L, Lu W, Wei Y. 2011. AcrB trimer stability and efflux activity, insight from mutagenesis studies. PloS one 6(12):e28390.

[0124] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, accession number, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

Illustrative native AcrB-like sequences:

Pseudomonas putida KT2440 (NCBI NP_743544), SEQ ID NO:2; synonyms: ttgB, PP_1385

```

1 mskffidrpi fawvialvim lvgalsilk1 pingypsiap paiaiaavtyp gasaqtqvqdt
61 vvqvieqqln gidnlryvss esnsdgsmti tatfeggtntp dtaqvqvqnk lnlatpllpq
121 evqqggirvt kavknfllvi glvsedgsmt kddlanyivs nmqdpisrta gvgdfqvfga
181 qyamriwldp aklnkfqltp vdvktavaaq nvqvssgqlg glpalpgtql natiightrl
241 qtaeqfesil lkvnkdgsqv rlgdvaqvg1 ggenyavsag fngkpasgla vklatganal
301 dtakalreti kglepffppg vkavfpydt1 pvvtesisgv ihtlieavvl vflvmlyflq
361 nfratiittm tvpvvllgtf gilaaagfsi ntltmfamvl aigllvddai vvvenvervm
421 seeglppeka tkrsmeqigg alvgialvls avllpmaffg gstgviyrqf sitivsamgl
481 svlvaliftp alcatmlkpl kkgelhtakg gffgwfnnrf drsvngyers vgailrnkvp
541 fllayalivv gmiwlfarip taflpeedqg vlfaqvqtpa gssaertqv1 vdqmreyllk
601 deadtvs1vf tvngfnfagr gqsegmafim lkpwderske nsvfalagra qghfftfrda
661 mvfafappav lelgnatgfd vflqdrggvg heklmearnq flakaaqski lsavrpngln
721 depqyqltid derasalgtv iadinntlsi algasyvndf idrgrvkkvy iggepsarms
781 pedlqkwyvr ngagemvpfs sfakgewtyg spklsryngv eameilgapa pgystgeama
841 everiagelp sgigfswtgm syeeklsgsq mpalfalsvl fvflclaaly eswsipiavv
901 lvvp1giiga liatslrgls ndvyflvgll ttiglaakna ilivefakel heqgrslyda
961 aieacrmrlr piimtslafi lgvvpl1tias gagagsqhai gtgviggmis atvlaifwvp
1021 lffvavsslf gskepekdv1t penpryeagq

```

Marinobacter aqueolei EPL (NCBI YP_960752), SEQ ID NO:3

```

1 mprffierpi fawvvalmlm lggglavknl avnqfpdvap paiaalsvny1p gasaqtqvqdt
61 vvqvieqqln gldglryiss esnsdgsmti iatfeggt1dp diaqvqvqnk lqlanpllpe
121 evqrqgirvs kykvnfft1vf altspdgkyt qgdldadyivs niqdpvartq gvgdfllf1gs
181 qyamrlwldp eklnsyqltp qdvinsvraq nvqvsagqlg glptaegvq1l qatvigkqrm
241 ktaeefenil lkvnpdgsqv rladvaevnl gnenyattgk yngapaagma lrlatganq1l
301 etagr1vketl aelerflpeg veivfpydt1t pvvsasietv amt1lieavvl vflvmfl1flq
361 swrati1ptl avpvvllatf gvlyafgftv nvmtmfamvl aigllvddai vvvenver1lm
421 eeeg1spkea akksmdqisg allgig1lvis avflpmaffg gstgviyrqf svtiisamsf

```

481 svlvafiftp alcatllkpg dqhvrkgffg wfnrtfdrsa dryksqvsyl ikrkgrfmgv
 541 ylllvvavgf lfkglptafl pdedqgvmlv mvqlptnatg erteavlaea gnylleese
 601 vvksvmsvrg fnfagrgqns gilfvdikpf adresfaqsv falagrsgar faqikdaivf
 661 pivppailel gnatgfdlyl kdngaighha lmaatnefis ranaapelnm vrhnglpdep
 721 qyqviiiddek arllqvsiad inatmsaawg ssyvndflhn grvkkvyvqg kpdsrlaped
 781 fdkwfvrnaq gemvpfaafa tgewvfgspr lgryqglpat qiqgapangy stgdamaale
 841 riaadlpqgl gleytglsfe ekqagnqamm lyllsilvvf lclaalyesw sipfavimlv
 901 plgvlgavla tmarglsndv ffqvgmlttm glaaknaili vefarqlyeq egkpllgata
 961 eaarlrlrpi imtslafifg vlpmaiasga ssasghaigt avvggtlaat ilaiffvplf
 1021 yvfvvgtgk rksadd

Pseudoalteromonas haloplanktis TAC125 (NCBI YP_341810), SEQ ID NO:4

1 msryfidkpi fawvlaivm lagilavksl piaqypsiap paisitanyp gasaqtleds
 61 vtqvieqkmk gldgilymss tsesngsatl tltfnadtdp diaqvqvqnk latatpllpe
 121 evqrqgvvva kaarnfllvl afvskdgsmt nidigdyvas nvqdivsrvd gvgeaqlfgs
 181 qyamriwldp aklqnfklt p ndigaaitag naqvsagqlg gmpaiagqql natitaqsr
 241 qtpeqfenil vktnsdgs v rldkdarvel ggenygvvar yngdpaagig iklasganal
 301 dtaegvkkal edlkpffpeg ldvaipydt p fvslsiekv vhtlieavil vfvvmylflq
 361 nfratlipti avpvllgtf ailytfgy si ntltmfamvl aigllvddai vvvenverlm
 421 teeklsalda trksmdaikg alvgiamvls avfipmaffs gstgviyrqf sitivsamgl
 481 svlvaliltp alcatllkps hvhdksfig rffsgfnrgf dktnrgaggi vgrmisqskr
 541 ylliygvlv gmvvfvsslp taflpdedqg ilfnqvmlpa gstieqtlgv vekvenhyln
 601 dqaeavgsif tvtgfsfags gqnsaigfvn lkhwderqrd dlsnavagk gmgfstike
 661 afvfafp ppa iveigtangf nmflqdrvg l ghdelnarn mllgmasksp vlagvrpnqg
 721 edmpelql di dlakaealgv tqtdinstls tawgsryvnd fidrgrvkv ymggeadsm
 781 vpedlnkwyv rnnngdmvpf aafasshwsy gsprleryng fsameiqgsa apgystgqam
 841 demerlvkql pngiasewsg isygerssgg qapillygl l fvfllclaal yeswsvpfav
 901 miivplgfig aimaalignl sndiylqvgl lttiglaskn ailivefaih kmeeglsldv
 961 aaiaavklrl rpvlmtslaf icgvvplaia ssagsagqna lgisiiggti aastlvvlfv
 1021 plffvlvrkt fssksavatk ese

E. coli AcrF (SEQ ID NO:5)

1 manffirrpi fawvlaiilm magalailql pvaqyptiap pavsvsanyp gadaqtvqdt
 61 vtqvieqnmn gidnlmymss tsdsagsvti tltfqsgtdp diaqvqvqnk lqlatpllpq
 121 evqqqgisve kssssylmva gfvsdnpgtt qddisdyvas nvkdtlsrln gvgdvqlfga
 181 qyamriwlda dllnkyklt p vdvingikvq ndqiaagqlg gtpalpqqql nasiaqtrf
 241 knpeefgkvt lrvnsdgs v rldkdarvel ggenynviar ingkpaaglg iklatganal
 301 dtakaikakl aelqpffp qg mkvlypydt p fvlqslsihev vktlfeaiml vflvmylflq
 361 nmratlipti avpvllgtf ailaafigy si ntltmfgmvl aigllvddai vvvenvervm
 421 medklppkea tekmsqigq alvgiamvls avfipmaffg gstgaiyrqf sitivsamal

481 svlvaliltp alcatllkpv saehhenkgy ffgwfnttfd hsvnhytnsv gkilgstgry
541 lliyalivag mvvlflrrips sflpeedqgv fltmiqlpag atqertqkvl dqvtdyyln
601 ekanvesvft vngfsfsgqa qnagmafvsl kpweerngde nsaeavihra kmelgkirdg
661 fvipfnmpai velgtatgfd felidqaglg hdaltqarnq llgmaaqhpa slvsvrpngl
721 edtaqfklev dgekaqalgv slsdingtis talggtyvnd fidrgrvkk1 yvqadakfrm
781 lpedvdklyv rsangemvpf safttshwvy gsprleryng lpsmeiqgea apgtssgdam
841 almenlaskl pagigydwgt msyqerlsgn qapalvaisf vvvflclaal yeswsipvsv
901 mlvvplgivg vllaatlfnq kndvyfmvgl lttiglsakn aillivefakd lmekegkgvv
961 eatlmavrmr lrpilmtsla filgvlplai sngagsgagn avgigvmggm vsatllaiff
1021 vpvffvvirr cfkq

WHAT IS CLAIMED IS:

- 1 1. A microorganism host cell comprising a recombinant non-naturally
2 occurring variant AcrB-like polypeptide that comprises at least one amino acid sequence
3 mutation relative to a native AcrB-like polypeptide sequence, wherein the non-naturally
4 occurring AcrB-like polypeptide reduces accumulation of alpha-olefin compounds in the host
5 cell in comparison to a wild-type host cell that expresses the native AcrB-like polypeptide at
6 a native level.
- 1 2. The host cell of claim 1, wherein the at least one sequence mutation is
2 an amino acid substitution.
- 1 3. The host cell of claim 1, wherein the recombinant non-naturally
2 occurring variant AcrB-like polypeptide amino acid sequence is substantially identical to
3 SEQ ID NO:1 and comprises one or more amino acid substitutions within 2 to 8 Angstroms
4 of, or at, a position corresponding to amino acid 279, amino acid 584, amino acid 617, amino
5 acid 822, amino acid 927, or amino acid 1033, as numbered with reference to SEQ ID NO:1.
- 1 4. The host cell of claim 3, wherein the recombinant non-naturally
2 occurring variant AcrB-like polypeptide comprises an alanine to threonine substitution at a
3 position corresponding to amino acid 279.
- 1 5. The host cell of claim 3, wherein the recombinant non-naturally
2 occurring variant AcrB-like polypeptide comprises a glutamine to arginine substitution at a
3 position corresponding to amino acid 584.
- 1 6. The host cell of claim 3, wherein the recombinant non-naturally
2 occurring AcrB-like polypeptide comprises a phenylalanine to leucine amino acid
3 substitution at a position corresponding to amino acid 617.
- 1 7. The host cell of claim 3, wherein the recombinant non-naturally
2 occurring variant AcrB-like polypeptide comprises a leucine to proline amino acid
3 substitution at a position corresponding to amino acid 822.
- 1 8. The host cell of claim 3, wherein the recombinant non-naturally
2 occurring variant AcrB-like polypeptide comprises a phenylalanine to serine amino acid
3 substitution at a position corresponding to amino acid 927.

- 1 9. The host cell of claim 3, wherein the recombinant non-naturally
2 occurring variant AcrB-like polypeptide comprises a phenylalanine to tyrosine amino acid
3 substitution at a position corresponding to amino acid 1033.
- 1 10. The host cell of any one of claims 1 to 9, wherein the alpha-olefin
2 compounds comprise 1-hexene, or a derivative thereof.
- 1 11. The host cell of claim 10, wherein the host cell further comprises the
2 proteins necessary to produce 1-hexene, or a derivative thereof.
- 1 12. The host cell of claim 11, wherein the host cell produces a greater
2 amount of the 1-hexene, or derivative thereof, as compared to a wild-type host cell that
3 expresses the native AcrB-like polypeptide at a native level.
- 1 13. The host cell of any one of claims 1 to 12, wherein the host cell is a
2 bacterial cell.
- 1 14. The host cell of claim 13, wherein the bacterial cell is a gram negative
2 bacterial cell.
- 1 15. The host cell of claim 14, wherein the gram negative bacterial cell is an
2 *E. coli* cell.
- 1 16. The host cell of any one of claims 1 to 12, wherein the host cell is a
2 yeast cell.
- 1 17. A method for producing an alpha-olefin compound, comprising:
2 a) providing a host cell of any one of claims 1 to 16; and
3 b) culturing said host cell under conditions sufficient for the host cell to
4 produce an alpha-olefin compound.
- 1 18. The method of claim 17, further comprising purifying the alpha-olefin
2 compound.
- 1 19. A recombinant variant AcrB polypeptide, wherein the recombinant
2 variant AcrB polypeptide has at least 70% identity to SEQ ID NO:1 and comprises one or
3 more amino acid substitutions within 2 to 8 Angstroms of, or at, a position corresponding to

4 amino acid 279, amino acid 584, amino acid 617, amino acid 822, amino acid 927, or amino
5 acid 1033 of SEQ ID NO:1.

1 20. The recombinant variant AcrB polypeptide of claim 19, wherein the
2 recombinant variant AcrB polypeptide comprises a threonine at a position corresponding to
3 amino acid 279.

1 21. The recombinant variant AcrB polypeptide of claim 19, wherein the
2 recombinant variant AcrB polypeptide comprises an arginine at a position corresponding to
3 amino acid 584.

1 22. The recombinant variant AcrB polypeptide of claim 19, wherein the
2 recombinant variant AcrB polypeptide comprises a leucine at a position corresponding to
3 amino acid 617.

1 23. The recombinant variant AcrB polypeptide of claim 19, wherein the
2 recombinant variant AcrB polypeptide comprises a proline at a position corresponding to
3 amino acid 822.

1 24. The recombinant variant AcrB polypeptide of claim 19, wherein the
2 recombinant variant AcrB polypeptide comprises a phenylalanine at a position corresponding
3 to amino acid 927.

1 25. The recombinant variant AcrB polypeptide of claim 19, wherein the
2 recombinant variant AcrB polypeptide comprises a tyrosine amino acid at a position
3 corresponding to amino acid 1033.

1 26. The recombinant variant AcrB polypeptide of any one of claims 19 to
2 25, wherein the polypeptide has at least 80% identity to SEQ ID NO:1.

1 27. The recombinant variant AcrB polypeptide of any one of claims 19 to
2 25, wherein the polypeptide has at least 90% identity to SEQ ID NO:1.

1 28. The recombinant variant AcrB polypeptide of any one of claims 19 to
2 25, wherein the polypeptide has at least 95% identity to SEQ ID NO:1.

1 29. The recombinant variant AcrB polypeptide of any one of claims 19 to
2 28, wherein the variant AcrB polypeptide reduces accumulation of alpha-olefin compounds in

3 the host cell in comparison to a wild-type host cell that expresses native AcrB at a native
4 level.

1 30. The recombinant variant AcrB polypeptide of any one of claims 19 to
2 25, wherein the mutant AcrB polypeptide comprises an amino acid sequence of SEQ ID
3 NO:1 or SEQ ID NO:5 that has one or more substitutions at amino acid positions
4 corresponding to amino acids 279, 584, 617, 822, 927, or 1033 of SEQ ID NO:1.

1 31. A recombinant polynucleotide comprising a nucleic acid sequence
2 encoding a variant AcrB polypeptide of any one of claims 19 to 30.

1 32. A recombinant microorganism host cell comprising the polynucleotide
2 of claim 31.

1 33. The host cell of claim 32, wherein the host cell is a yeast host cell.

1 34. The host cell of claim 32, wherein the host cell is a bacterial host cell.

1 35. The host cell of claim 34, wherein the bacterial host cell is a gram-
2 negative host cell.

1 36. The host cell of claim 35, wherein the gram-negative host cell is an *E.*
2 *coli* host cell.

1 37. The host cell of any one of claims 32 to 36, wherein the host cell
2 further comprises the proteins necessary to produce 1-hexene, or a derivative thereof.

1 38. A recombinant variant AcrB-like polypeptide, wherein the
2 recombinant variant AcrB-like polypeptide has at least 70% identity to SEQ ID NO:2, SEQ
3 ID NO:3, or SEQ ID NO:4 and comprises one or more amino acid substitutions within 2 to 8
4 Angstroms of, or at, a position corresponding to amino acid 279, amino acid 584, amino acid
5 617, amino acid 822, amino acid 927, or amino acid 1033, as numbered with reference to
6 SEQ ID NO:1.

1 39. The recombinant variant AcrB-like polypeptide of claim 38, wherein
2 the polypeptide has at least 80% identity to SEQ ID NO:2, SEQ ID NO:3, or SEQ ID NO:4.

1 40. The recombinant variant AcrB-like polypeptide of claim 38, wherein
2 the polypeptide has at least 90% identity to SEQ ID NO:2, SEQ ID NO:3, or SEQ ID NO:4.

1 41. The recombinant variant AcrB-like polypeptide of claim 38, wherein
2 the polypeptide has at least 95% identity to SEQ ID NO:2, SEQ ID NO:3, or SEQ ID NO:4.

1 42. The recombinant variant AcrB polypeptide of any one of claims 38 to
2 41, wherein the variant AcrB-like polypeptide reduces accumulation of alpha-olefin
3 compounds in the host cell in comparison to a wild-type host cell that expresses a native
4 AcrB-like polypeptide at a native level.

1 43. A recombinant polynucleotide comprising a nucleic acid sequence
2 encoding a variant AcrB polypeptide of any one of claims 38 to 42.

1 44. A recombinant microorganism host cell comprising the polynucleotide
2 of claim 31.

1 45. The host cell of claim 44, wherein the host cell is a yeast host cell.

1 46. The host cell of claim 44, wherein the host cell is a bacterial host cell.

1 47. The host cell of claim 46, wherein the bacterial host cell is a gram-
2 negative host cell.

1 48. The host cell of claim 47, wherein the gram-negative host cell is an *E.*
2 *coli* host cell.

1 49. The host cell of any one of claims 44 to 48, wherein the host cell
2 further comprises the proteins necessary to produce 1-hexene, or a derivative thereof.

1 50. A method for producing an alpha-olefin compound, comprising:
2 a) providing a host cell of any one of claims 32 to 37 or 44 to 49; and
3 b) culturing said host cell under conditions sufficient for the host cell to
4 produce an alpha-olefin compound.

1 51. The method of claim 50, wherein the alpha-olefin compound is hexene
2 or a derivative of hexene.

1 52. The method of claim 51, wherein the host cell produces a greater
2 amount of the 1-hexene, or derivative thereof, as compared to a wild-type host cell that
3 expresses the native AcrB-like polypeptide at a natural level

- 1 53. A method for producing an alpha-olefin compound, comprising:
2 a) providing a microorganism host cell that overexpresses an AcrB-like
3 polypeptide; and
4 b) culturing said host cell under conditions sufficient for the host cell to
5 produce an alpha-olefin compound, wherein the host cell produces a greater amount of the
6 alpha-olefin compound compared to a corresponding host cell that expresses the native AcrB-
7 like polypeptide at a native level.
- 1 54. The method of claim 53, wherein the alpha-olefin compound is 1-
2 hexene, or a derivative thereof.
- 1 55. The method of claim 53 or 54, wherein the AcrB-like polypeptide has
2 at least 70% identity to the amino acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID
3 NO:3, SEQ ID NO:4 or SEQ ID NO:5.
- 1 56. The method of claim 55, wherein the AcrB-like polypeptide has at
2 least 80%, or at least 90% identity to the amino acid sequence of SEQ ID NO:1, SEQ ID
3 NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5.
- 1 57. The method of claim 56, wherein the AcrB-like polypeptide comprises
2 the amino acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or
3 SEQ ID NO:5.
- 1 58. The method of claim 57, wherein the AcrB-like polypeptide comprises
2 the amino acid sequence of SEQ ID NO:1 or SEQ ID NO:5
- 1 59. The method of claim 58, wherein the AcrB-like polypeptide comprises
2 the amino acid sequence of SEQ ID NO:1.
- 1 60. The method of any one of claims 53 to 59, wherein the host cell is a
2 bacterial host cell.
- 1 61. The method of claim 60, wherein the bacterial host cell is a gram
2 negative cell.
- 1 62. The method of claim 61, wherein the gram negative host cell is an *E.*
2 *coli* host cell.

1 63. The method of any one of claims 53 to 59, wherein the host cell is a
2 yeast host cell.

1 64. The method of any one of claims 53 to 63, wherein the host cell further
2 comprises the proteins necessary to produce 1-hexene, or a derivative thereof.

1 65. A cell culture medium comprising an alpha-olefin compound and a
2 microorganism host cell that overexpresses an AcrB-like polypeptide, wherein the host cell
3 produces a greater amount of the alpha-olefin compound compared to a corresponding host
4 cell that expresses the native AcrB-like polypeptide at a native level.

1 66. The cell culture medium of claim 65, wherein the alpha-olefin
2 compound is 1-hexene, or a derivative thereof.

1 67. The cell culture medium of claim 65 or 66, wherein the AcrB-like
2 polypeptide has at least 70% identity to the amino acid sequence of SEQ ID NO:1, SEQ ID
3 NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5.

1 68. The cell culture medium of claim 67, wherein the AcrB-like
2 polypeptide has at least 80%, or at least 90% identity to the amino acid sequence of SEQ ID
3 NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5.

1 69. The cell culture medium of claim 68, wherein the AcrB-like
2 polypeptide comprises the amino acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID
3 NO:3, SEQ ID NO:4 or SEQ ID NO:5.

1 70. The cell culture medium of claim 69, wherein the AcrB-like
2 polypeptide comprises the amino acid sequence of SEQ ID NO:1 or SEQ ID NO:5

1 71. The cell culture medium of claim 70, wherein the AcrB-like
2 polypeptide comprises the amino acid sequence of SEQ ID NO:1.

1 72. The cell culture medium of any one of claims 65 to 71, wherein the
2 host cell is a bacterial host cell.

1 73. The cell culture medium of claim 72, wherein the bacterial host cell is
2 a gram negative cell.

1 74. The cell culture medium of claim 73, wherein the gram negative host
2 cell is an *E. coli* host cell.

1 75. The cell culture medium of any one of claims 65 to 71, wherein the
2 host cell is a yeast host cell.

1 76. The cell culture medium of any one of claims 65 to 75, wherein the
2 host cell further comprises the proteins necessary to produce 1-hexene, or a derivative
3 thereof.

1

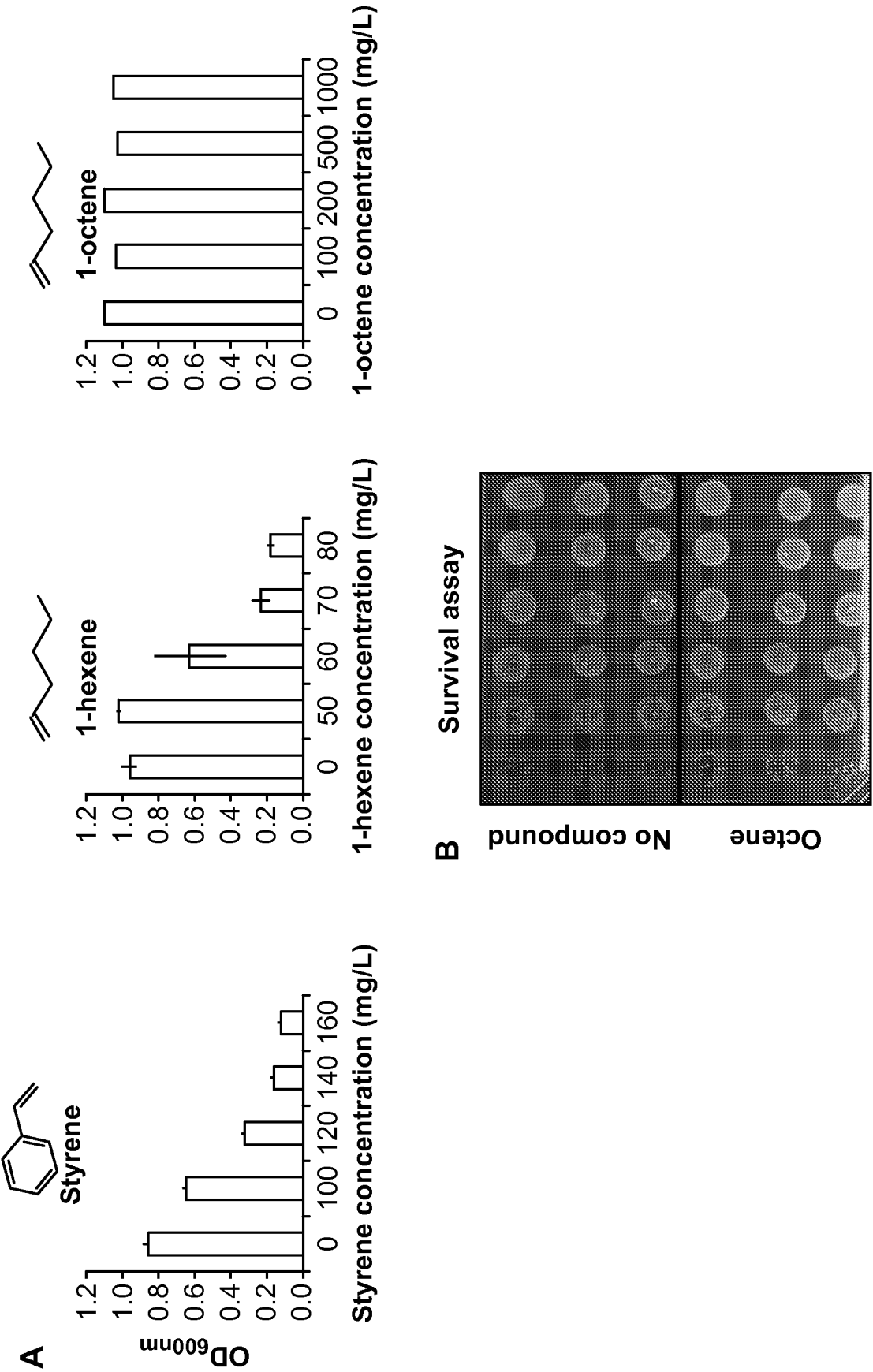


FIG. 1

2 / 11

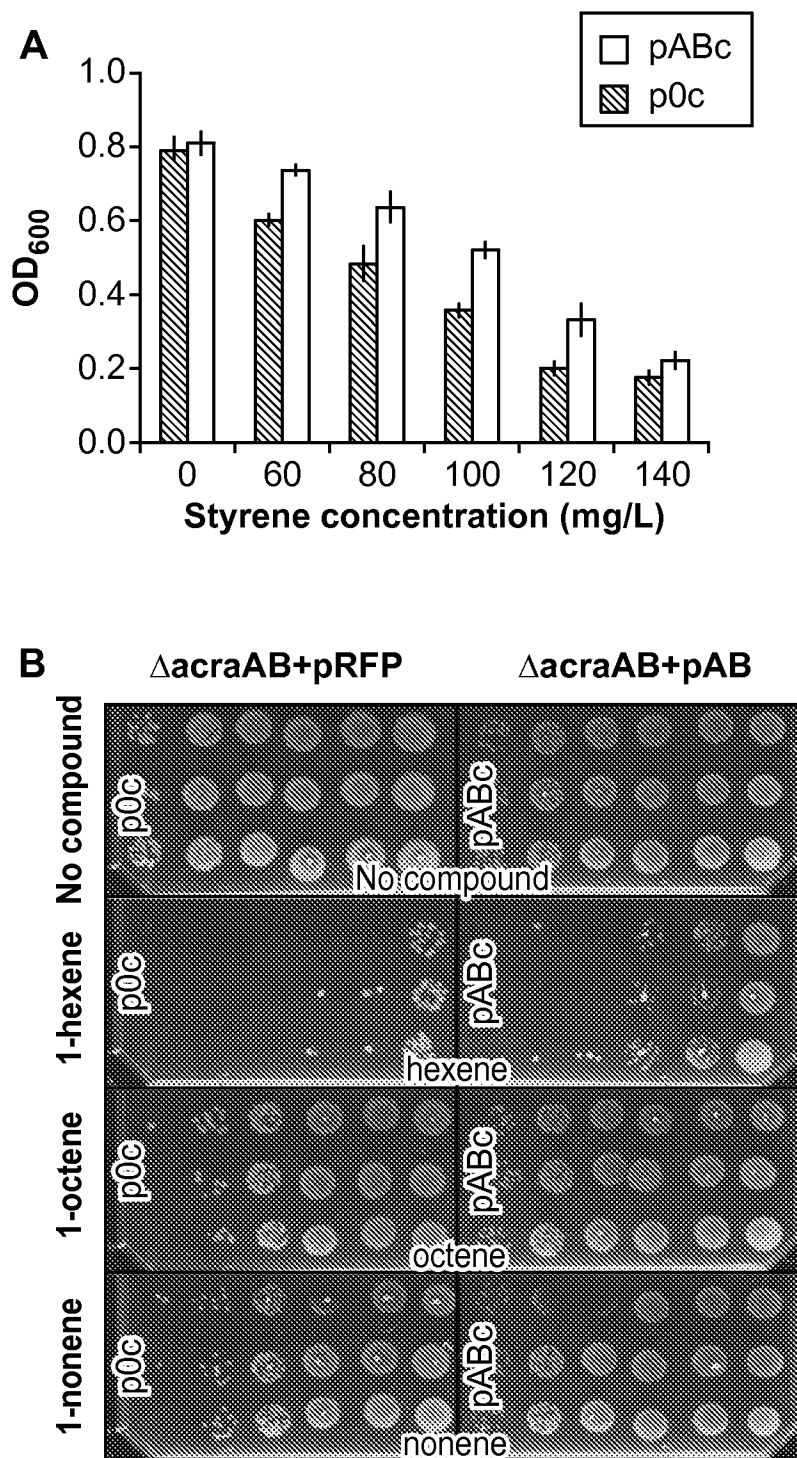
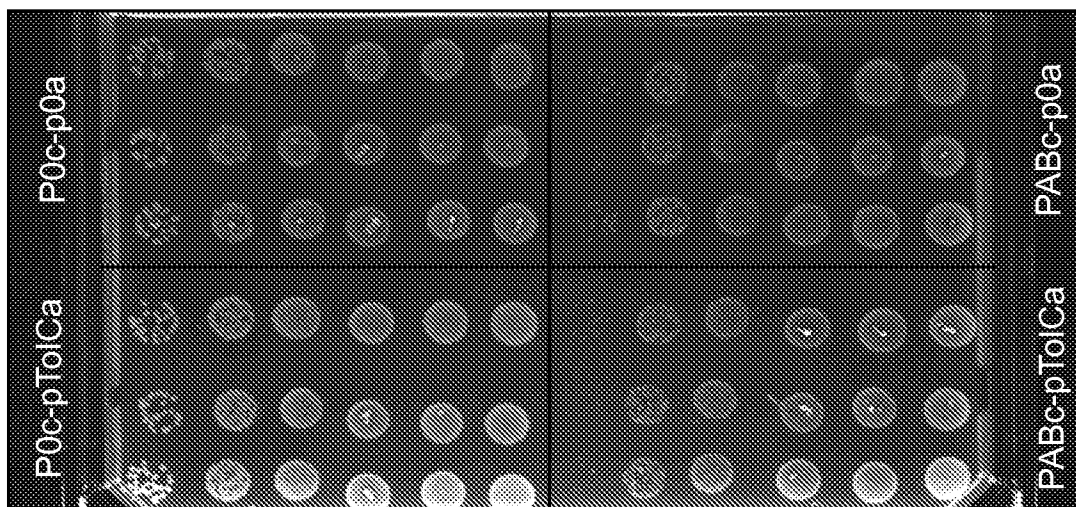


FIG. 2

3 / 11

No compound



1-hexene

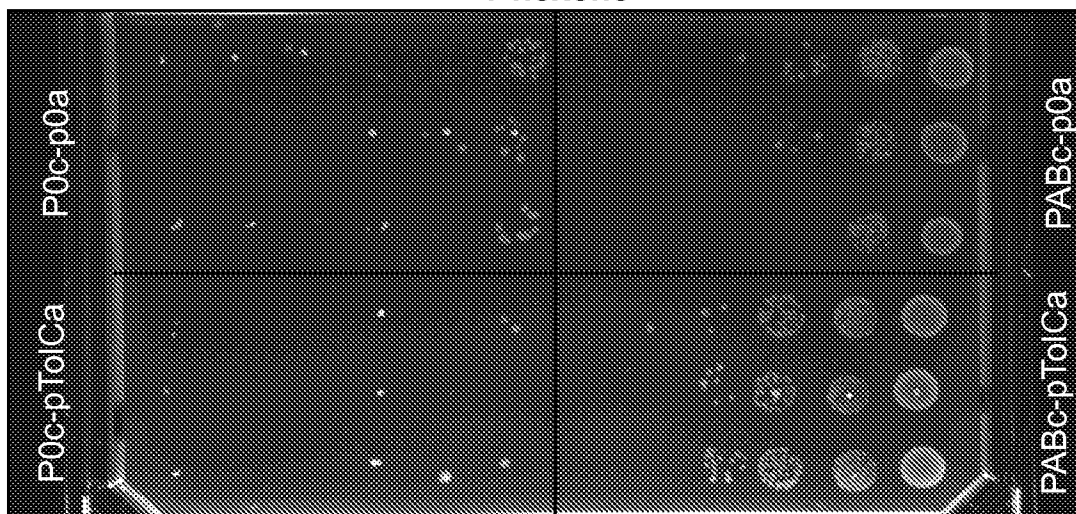


FIG. 3

4 / 11

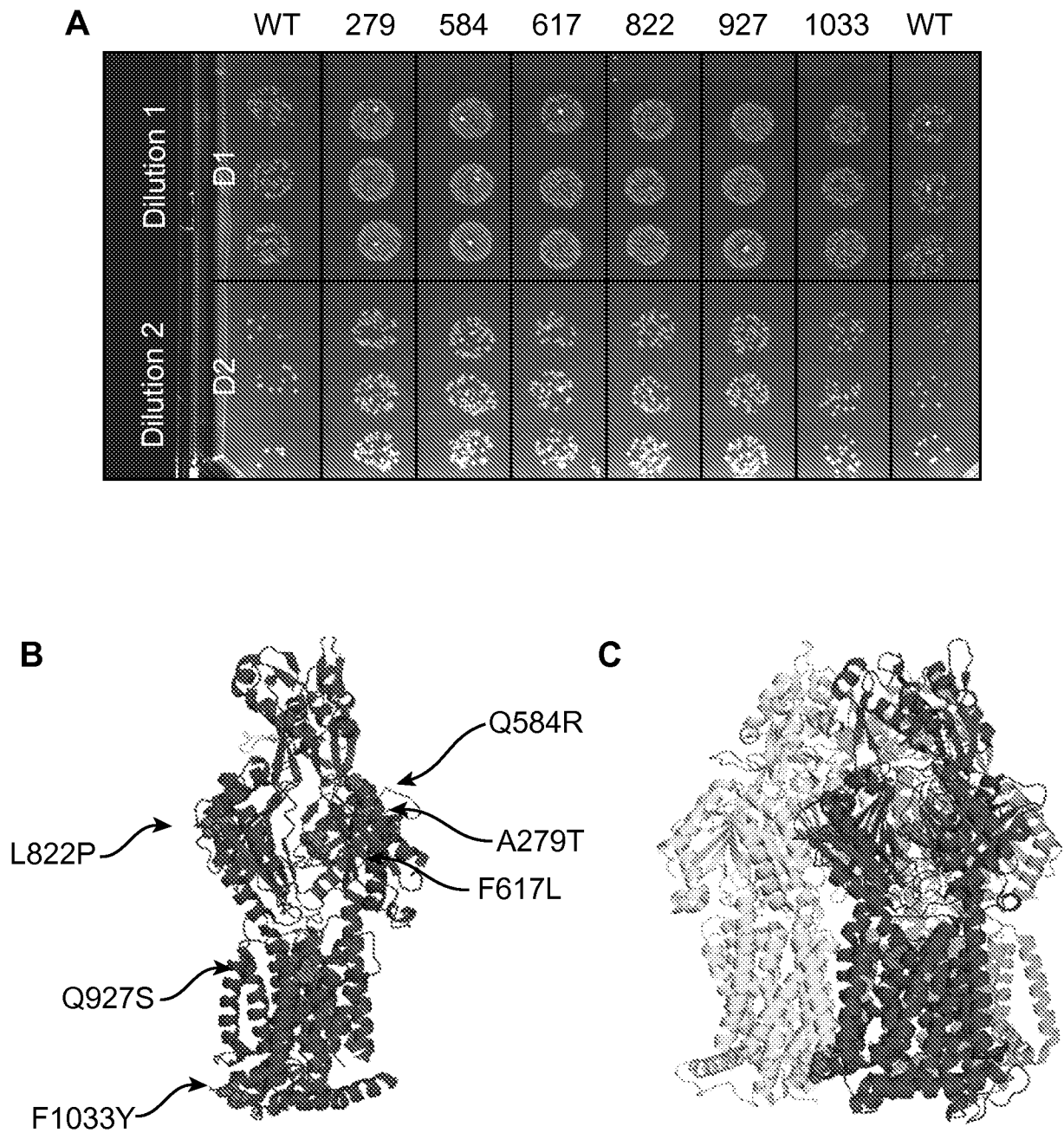


FIG. 4

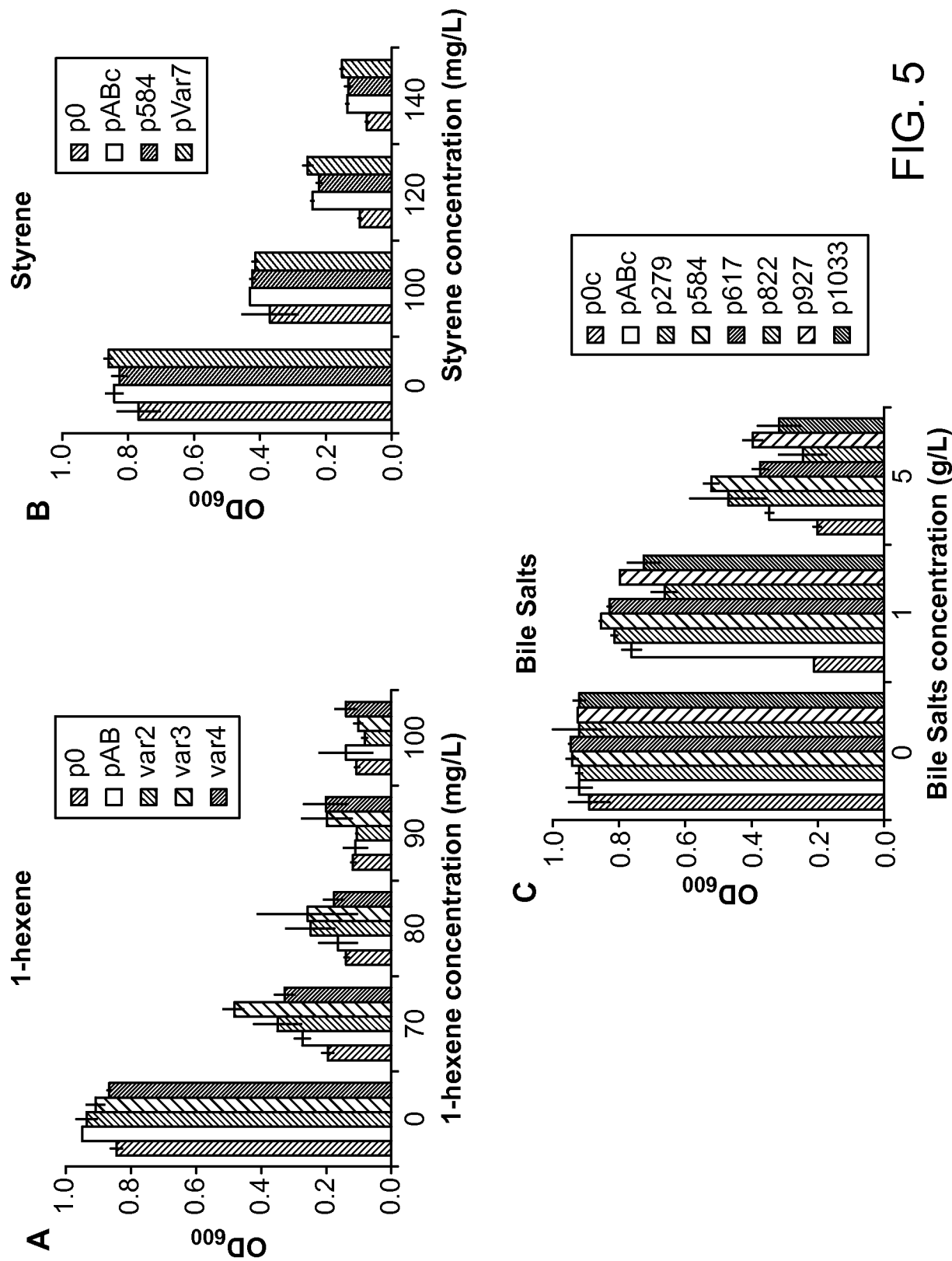


FIG. 5

6 / 11

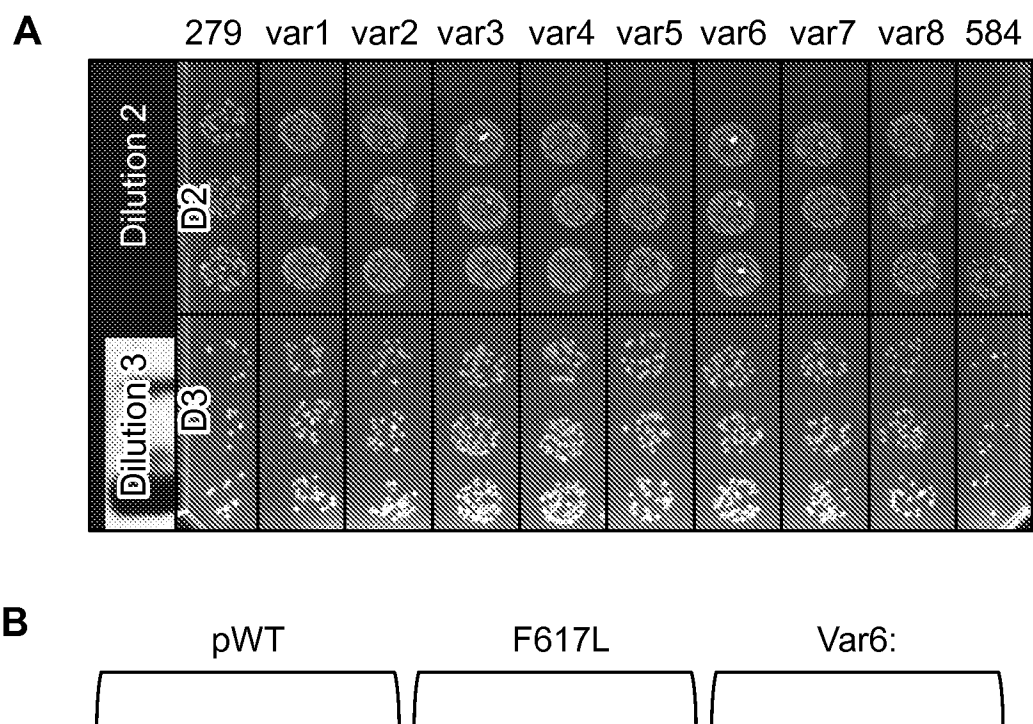


FIG. 6

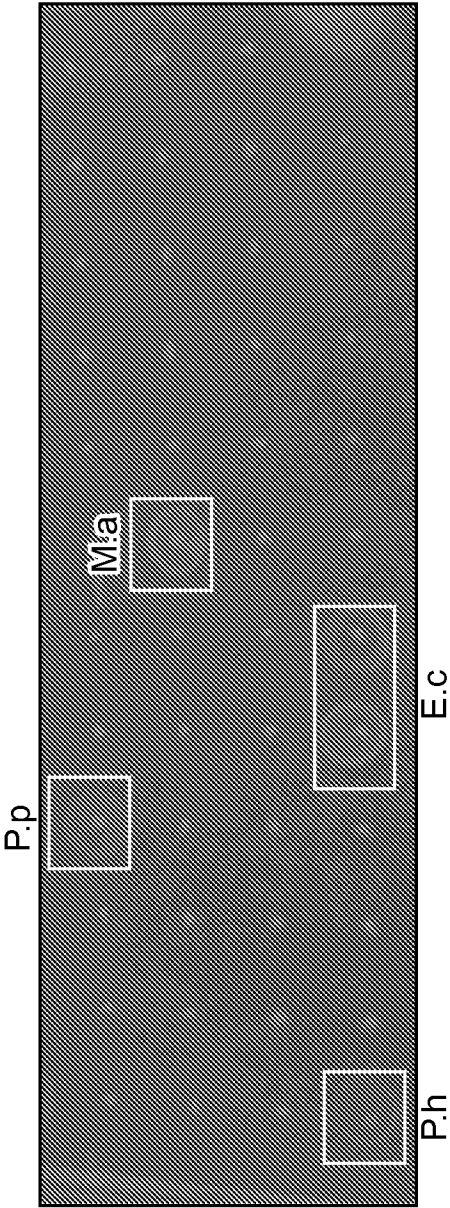


FIG. 7

sequences alignments, illustrative AcrB domains

SEQ ID NO:1: AcrB *E. coli*
SEQ ID NO:2: *Pseudomonas putida* 66% identity with AcrB
SEQ ID NO:3: *Marinobacter aqueolei* 59% identity with AcrB
SEQ ID NO:4: *Pseudoalteromonas haloplanktis* 62% of identity with AcrB
SEQ ID NO:5: AcrF 78% of identity with AcrB

Illustrative domain structures of SEQ ID NO:1:
transmembrane domains--highlighted
Italics, underlined, increased font size: illustrative mutations improving tolerance
Bolded, underlined: Amino acids important for binding
Underlined: Amino acids important for H+ transport

SEQ ID NO:1	MPNFFIDRPIFAWVIAIIIMLAGGLAILKLPVAQYPTIAPPAVTISASYPGADAKTVQDT	60
SEQ ID NO:2	M FFIDRPIFAWVIA++IML G L+ILKLP+ QYP+IAPPA+ I+ +YPGA A+TVQDT	
SEQ ID NO:3	MP FFI+RPIFAWV+A+++ML GGLA+ L V Q+P +APPA+ +S +YPGA A+TVQDT	
SEQ ID NO:4	M +FID+PIFAWV+AII+MLAG LA+ LP+AQYP+IAPPA++I+A+YPGA A+T++D+	
SEQ ID NO:5	M NFFI RPIFAWV+AII+M+AG LAIL+LPVAQYPTIAPPAV++SA+YPGADA+TVQDT	
SEQ ID NO:1	VTQVIEQNMNGIDNLMYMSSNSDSTGTVQITLTFESGTDADIAQVQVQNKQLQAMP LLPQ	120
SEQ ID NO:2	V QVIEQ +NGIDNL Y+SS S+S G++ IT TFE GT+ D AQVQVQNKL LA PLLPQ	
SEQ ID NO:3	V QVIEQ +NG+D L Y+SS S+S G++ I TFE GTD DIAQVQVQNKQLQA PLLP+	
SEQ ID NO:4	VTQVIEQ M G+D L+YMSS S+S G+ +TLTF + TD DIAQVQVQNKL A PLLP+	
SEQ ID NO:5	VTQVIEQNMNGIDNLMYMSS SDS G+V ITLTF+SGTD DIAQVQVQNKQLQA PLLPQ	
SEQ ID NO:1	EVQQQGVSV EKSSSFLMVVGVIN TDGTM TQEDISDYVAANMKDAISRTSGVG DVQLFGS	180
SEQ ID NO:2	EVQQQG+ V K+ +FL+V+G+++ DG+MT++D+++Y+ +NM+D ISRT+GVGD Q+FG+	
SEQ ID NO:3	EVQ+QG+ V K +F V + + DG TQ D++DY+ +N++D ++RT GVG D LFGS	
SEQ ID NO:4	EVQ+QGV V K++ +FL+V+ ++ DG+MT DI DYVA+N++D +SR GVG+ QLFGS	
SEQ ID NO:5	EVQQQG+SVEKSSSS+LMV G ++ + TQ+DISDYVA+N+KD +SR +GVGDVQLFG+	

FIG. 8

SEQ ID NO:1	QYAMRIWMNP	ELNKFQLT	PVDVITAIKAQNAQVAAGQLGGT	PPVKGQQLNASIIAQTRL	240
SEQ ID NO:2	QYAMRIW++P	+LNKFQLT	PVDV TA+ AQN QV++GQLGG P + G Q	LNA+II +TRL	
SEQ ID NO:3	QYAMR+W++P	+LN +QLTP	DVI +++AQN QV+AGQLGG P +G QL	A++I + R+	
SEQ ID NO:4	QYAMRIW++P	+L F+LTP	D+ AI AQNAQV+AGQLGG P + GQQLNA+I	AQ+RL	
SEQ ID NO:5	QYAMRIW++	+ LNK++LTP	VDVI +K QN Q+AAGQLGGTP + GQQLNASIIAQTR		
SEQ ID NO:1	TSTEEFGKILLKVNQD	GSRVLLRDVAKIELGGENY	DI	AEFNGQPASGLGIKLIATGANAL	300
SEQ ID NO:2	+ E+F	ILLKVN+DGS+V L DVA++	LGGENY + A+FNG+PASGL +KLATGANAL		
SEQ ID NO:3	+ EEF	ILLKVN DGS+V L DVA++	LG ENY ++NG PA+G+ ++LATGAN L		
SEQ ID NO:4	+ E+F	IL+K N DGS V L+DVA++	ELGGENY ++A +NG PA+G+GIKLA+GANAL		
SEQ ID NO:5	+ EEF	GK+ L+VN DGS V L+DVA++	ELGGENY++IA NG+PA+GLGIKLIATGANAL		
SEQ ID NO:1	DTAAAIRAELAKMEP	FFPSGLKIVYP	Y	DTTTPFVKIS	360
SEQ ID NO:2	DTA A+R	+ +EPFFP	G+K V+PYD	TPP V SI V+ TL+EA++LVFLVMYLF	LQ
SEQ ID NO:3	+TA ++	LA++E F P G++IV+PYD	TPP V SI V TL+EA++LVFLVM+L	FLQ	
SEQ ID NO:4	DTA ++	L ++PFFP	GL + PYD	TPFV +SI +VV TL+EA+ILVF+VMYLF	LQ
SEQ ID NO:5	DTA AI+A+LA++	+PFFP	G+K++YPYD	TPFV++SIHEVVKTLEAI+LVFLVMYLF	LQ
SEQ ID NO:1	NFRAT	LIPTIAVPVLLGTFA	V	LA	420
SEQ ID NO:2	NFRAT+I	T+ VPVLLGT	F +LAA	GFSIN	TL
SEQ ID NO:3	++RAT+IPT+AVP	VLL	TF VL	AFGF++N +TMF	MVLAIGLLVDDAIVV
SEQ ID NO:4	NFRAT	LIPTIAVPVLLGTFA+L	FG+SIN	TL	TMF
SEQ ID NO:5	N	RAT	LIPTIAVPVLLGTFA+L	AAFG+SIN	TL
SEQ ID NO:1	AEEGLPPKEAT	RKSMGQI	Q	GALVGIAMVLSAVF	VPMA
SEQ ID NO:2	+EEGLPPKEAT++	SM	QIQGALV	GIA+VLSAV +PMAFF	GGSTG
SEQ ID NO:3	EEGL	PKEA +KSM	QI	GAL+GI +V+SAVF+PMAFF	GGSTG
SEQ ID NO:4	EE L	+ATR	KSM +I+GALVGIAMVLSAVF+PMAFF	GSTG	IYRQFSITIVS
SEQ ID NO:5	E+	LPPKEAT	KSM	QIQGALVGIAMVLSAVF+PMAFF	GGSTG

FIG. 8 (Cont. 1)

SEQ ID NO:1	<u>SVLVALILTP</u> <u>A</u> LCATMLKPIAKGDHGEKKGGFFGWFNRMFEKSTHHYTDSVGGILRST <u>GR</u>	540
SEQ ID NO:2	SVLVALI TPALCATMLKP+ KG+H K GFFGWFNR F++S + Y SVG ILR+	
SEQ ID NO:3	SVLVA I TPALCAT+LKP GD +KGGFFGWFNR F++S Y V +++ GR	
SEQ ID NO:4	SVLVALILTPALCAT+LKP D FF FNR F+K+ VG ++ + R	
SEQ ID NO:5	SVLVALILTPALCAT+LKP++ +H E K GFFGWFN F+ S +HYT+SVG IL STGR	
SEQ ID NO:1	<u>YLVL</u> YLIIVVGMAYL <u>FV</u> RLPSSFLP <u>DE</u> DQGV <u>F</u> TMVQLPAGAT <u>Q</u> ERTQKVLNEVTHYYLT	600
SEQ ID NO:2	+L+ Y +IVVGM +LF R+P++FLP+EDQGV VQ PAG++ ERTQ V++++ Y L	
SEQ ID NO:3	++ +YL++VV + +LF LP++FLPDEDQGV + MVQLP AT ERT+ VL E +Y L	
SEQ ID NO:4	YL++Y ++V GM Y+F LP++FLPDEDQG+ V LPAG+T E+T V+ +V ++YL	
SEQ ID NO:5	YL++Y +IV GM LF+RLPSSFLP+EDQGVF+TM+QLPAGAT <u>Q</u> ERTQKVL++VT YYL	
SEQ ID NO:1	KEKNNVESVFAVNGFG <u>F</u> AGRGQNTGIA <u>F</u> VSLKDWADRPGEENKVEAITMRA ⁶⁶⁰ TRAFSQIKD	660
SEQ ID NO:2	E + V SVF VNGF FAGRGQ++GAF+ LK W +R +EN V A+ RA + F +D	
SEQ ID NO:3	+E V+SV +V GF FAGRGQ+GI FV LK +ADR V A+ R+ F+QIKD	
SEQ ID NO:4	+ V S+F V GF FAG GQN+ I FV+LK W +R ++ V A+ + FS IK+	
SEQ ID NO:5	EK NVESVF VNGF F+G+ QN G+AFVSLK W +R G+EN EA+ RA +I+D	
SEQ ID NO:1	AMV <u>F</u> <u>A</u> FN <u>L</u> PAIV <u>E</u> LG <u>T</u> ATGDFD <u>F</u> ELIDQAGLGHEKLTQARNQLLAEEAAKHPDMLTSV <u>R</u> PNG	720
SEQ ID NO:2	AMVFAF PA++ELG ATGFD L D+ G+GHEKL +ARNQ LA+AA+ +L++VRPNG	
SEQ ID NO:3	A+VF PAI+ELG ATGFD L D +GH L A N+ ++ A P+ L VR NG	
SEQ ID NO:4	A VFAF PAIVELGTA GF+ L D+ GLGH++L ARN LL A+K P +L VRPNG	
SEQ ID NO:5	V FN+PAIVELGTATGDFD <u>F</u> ELIDQAGLGH+ LTQARNQLL AA+HP L SVRPNG	
SEQ ID NO:1	LEDTPQFKIDIDQEKAAQALGVSINDINTTLGAAGGS <u>Y</u> VNDFIDGRVK <u>K</u> VYVMSEAKY <u>R</u>	780
SEQ ID NO:2	L D PQ+++ ID E+A ALGV+I DIN TL A G SYVNDFIDGRVKKY+ E R	
SEQ ID NO:3	L D PQ+++ ID EKA+ L VSI DIN T+ AAWG SYVNDF+ GRVKKYV + R	
SEQ ID NO:4	ED P+ ++DID KA+ALGV+ DIN+TL AWG YVNDFIDGRVKKY+ EA R	
SEQ ID NO:5	LEDT QFK+++DQEKAAQALGVS++DIN T+ A GG+YVNDFIDGRVKK+YV ++AK+R	

FIG. 8 (Cont. 2)

SEQ ID NO:1	MLPDDIGDWYVRAADGQMPVFSAFSSSRWEYGSPLRYNGLPSMEI <u>L</u> GQAAPGK <u>S</u> TGEA	840
SEQ ID NO:2	M P+D+ WYVR G+MVDFS+F+ W YGSP+L RYNG+ +MEILG APG STGEA	
SEQ ID NO:3	+ P+D W+VR A G+MVPF+AF++ W +GSPRL+RY GLP+ +I G A G STG+A	
SEQ ID NO:4	M+P+D+ WYVR +G MVPF+AF+SS W YGSPRLERYNG +MEI G AAPG STG+A	
SEQ ID NO:5	MLP+D+ YVR+A+G+MVDFSAF++S W YGSPRLERYNGLPSMEI G+AAPG S+G+A	
SEQ ID NO:1	MEILMEQ ¹ LASKLPTGVGYDWTGMSYQERL <u>SGN</u> <u>Q</u> APSLYAISLIVVFLCLAAALYESWSIP <u>F</u> S	900
SEQ ID NO:2	M +E++A +LP+G+G+ WTGMSY+E+LSG+Q P+L+A+S++ VFLCLAALYESWSIP +	
SEQ ID NO:3	M +E++A+ LP G+G ++TG+S++E+ +GNQA LY +S+++VFLCLAALYESWSIPF+	
SEQ ID NO:4	M+ ME+L +LP G+ +W+G+SYQER SG QAP LY +SL+ VFLCLAALYESWS+PF+	
SEQ ID NO:5	M LME LASKLP G+GYDWTGMSYQERLSGNQAP+L AIS +VVFLCLAALYESWSIP S	
SEQ ID NO:1	VMLVPLGVIGALLAAT <u>F</u> RGLTNDVY <u>FQ</u> VGLLTITIGLSAKNAI <u>L</u> LIVEFAKDLMDKEGKGL	960
SEQ ID NO:2	V+LVVPLG+IGAL+A + RGL+NDVYF VGLLTITIGL+AKNAI <u>L</u> LIVEFAK+L ++G+ L	
SEQ ID NO:3	V+++VPLGV+GA+LA RGL+NDV+FQVG+LTT+GL+AKNAI <u>L</u> LIVEFA+ L ++EGK L	
SEQ ID NO:4	VM++VPLG+ GA++AA L+ND+Y QVGLLTITIGL++KNAILIVEFA M+ EG L	
SEQ ID NO:5	VMLVPLG++G LLAAT NDVYF VGLLTITIGLSAKNAI <u>L</u> LIVEFAKDLM+KEGK+	
SEQ ID NO:1	IEATLDVVRML <u>R</u> PIILMTSLAFILGVMLPLVIS <u>T</u> GAGSC <u>Q</u> ANAVGTGVMGGMVTATVLA <u>I</u> IF	1020
SEQ ID NO:2	+A ++A RMRLRPI+MTSLAFILGV+PL I++GAG+G+Q+A+GTGV+GGM++ATVLAIF	
SEQ ID NO:3	++AT +A R+RLRPI+MTSLAFI GV+P+ I++GA S +Q+A+GT V+GG + AT+LAIF	
SEQ ID NO:4	++A + AV++RLRP+LMTSLAFI GV+PL I++ AGSGAQNA+G ++GG + A+ L +	
SEQ ID NO:5	+EATL AVRMLRPIILMTSLAFILGV+PL IS GAGSGAQNAVG GVMGMV+AT+LAIF	
SEQ ID NO:1	FVPVFFVVRRR <u>R</u> FSRKNEIDIEHSHTVDDH	1049
SEQ ID NO:2	+VP+FFV V <u>F</u> K +	
SEQ ID NO:3	FVP+F+V VV RK+ D	
SEQ ID NO:4	FVP+FFV+VR+ FS K+	
SEQ ID NO:5	FVPVFFV+RR F	

FIG. 8 (Cont. 3)