(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 March 2004 (04.03.2004)

PCT

(10) International Publication Number
WO 2004/018635 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number:
PCT/US2003/026413

(22) International Filing Date: 20 August 2003 (20.08.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/405,196 21 August 2002 (21.08.2002) US

(71) Applicant *(for all designated States except US)*: **KOSAN BIOSCIENCES, INC.** [US/US]; 3832 Bay Center Place, Hayward, CA 94545 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **JULIEN, Bryan** [US/US]; 4514 Fleming Avenue, Oakland, CA 94619 (US).

(74) Agents: **APPLE, Randolph, Ted** et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: TRANSFORMATION SYSTEM BASED ON THE INTEGRASE GENE AND ATTACHMENT SITE FOR MYXOCOCCUS XANTHUS BACTERIOPHAGE Mx9

(57) Abstract: The invention provides a transformation system based on bacteriophage Mx9, a temperate phage that infects Myxococcus xanthus. Vectors containing an integrase encoding gene and a phage attachment site (attP) integrate into a chromosomal attB site and can be used to alter or introduce genes into a variety of host cells.

WO 2004/018635 A2

PATENT APPLICATION

Transformation System Based on the Integrase Gene and Attachment Site for *Myxococcus xanthus* Bacteriophage Mx9

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit of U.S. provisional patent application no. 60/405,196, filed August 21, 2002, the entire contents of which are incorporated by reference.

FIELD OF THE INVENTION

[0002] The invention relates to methods and materials for transforming host bacterial cells using a bacteriophage Mx9 system. The invention finds application in the fields of molecular biology and drug development.

BACKGROUND OF THE INVENTION

[0003] Mx9 is a general transducing phage that infects the Gram-negative bacterium *Myxococcus xanthus* (9). The phage particle has a polyhedral head with a very short tail. Structurally it resembles Mx8, which also infects *M. xanthus*.

[0004] The integrase gene and attachment site for Mx8 have been characterized (7, 8, 11). Integration of Mx8 by site-specific recombination requires a single phage protein, Int, and the phage attachment site, *attP*. Unlike most temperate bacteriophage, the Mx8 *attP* site is contained within the *int* gene and upon insertion into the *M. xanthus* chromosome, the 3' end of the *int* gene is altered. This modified *int* gene produces a protein, IntX, with lower specific integrase activity (8).

[0005] Because no natural replicating plasmids have been identified for *M. xanthus*, or for any other myxobacteria, phage attachment sites provide an efficient and stable alternative for introducing new genes or adding additional copies of existing ones into the cell. The Mx8 *int* and attachment site can be used to integrate DNA into the chromosome, but expression of many genes is affected by insertion into the Mx8 *attB*

sites; many developmental as well as two constitutive promoters, *mgl* and *pilA*, have reduced activity at this site (2, 6). There remains a need for more effective and reliable transformation systems that will enable insertion of DNA into the chromosome of *M. xanthus* and other bacteria. The present invention meets these and other needs.

SUMMARY OF THE INVENTION

[0006] The present invention provides methods and materials for transforming host cells using a bacteriophage Mx9 transformation system. In another aspect, the present methods, materials, host cells and vectors are directed to enhancing the production of a useful compound, including but not limited to a polyketide, through the introduction of one or more genes into the DNA of a variety of bacterial host cells.

[0007] In one aspect, the invention provides a method for modification of a DNA of a bacterial cell comprising in its genome a first attachment site recognized by a protein with Mx9 integrase activity, comprising introducing a Mx9 transformation system into the cell, said system comprising (a) a gene encoding a protein with Mx9 integrase activity protein operably linked to a promoter active in the host cell, and (b) a DNA vector comprising a second attachment site recognized by the integrase protein, which may be the same as the first attachment site.

[0008] These and other embodiments of the invention are described in more detail in the following description, examples, and claims set forth below.

BRIEF DESCRIPTION OF THE FIGURES

[0009] Figure 1 presents a physical map of the int region from Mx9. Boxes represent putative open reading frames. The hatched box in *int* designates the position of *attP*.

[0010] Figure 2 presents the nucleotide sequence of the Mx9 *int* gene [SEQ ID NO:1] and the deduced amino acid sequence [SEQ ID NO:2]. Amino acids are in one-letter code underneath the DNA sequence. The sequence in bold [SEQ ID NO:5] is the

Mx9 *attP* core site. Arrows represent inverted repeats. A previous version of this sequence had the following differences: 504 A-->T and 505 G-->A.

[0011] Figure 3 presents (A) Nucleotide sequence of the Mx9 *attB1* site [SEQ ID NO:3] and (B) Nucleotide sequence of the Mx9 *attB2* site [SEQ ID NO:4]. Nucleotides in bold are the 42 bp [SEQ ID NO:5] identical in the Mx9 *attP* site. Underlined nucleotides encode $tRNA^{gly}$. Arrows; inverted repeat within *attB2*. (C) Nucleotide sequence of the native Mx9 *attB1* [SEQ ID NO:6]. Nucleotides in bold indicate the partial core sequence. (D) Nucleotide sequence of the *attP* site [SEQ ID NO:7]. Arrows; inverted repeat.

[0012] Figure 4 presents the predicted cloverleaf secondary structure for tRNAgly from *M. xanthus* [bases 1397 to 1428 of SEQ ID NO:1]. The bases that are contained within the core *attB* sequence are outlined.

[0013] Figure 5 shows an agarose gel of PCR amplified DNA fragments. Lanes 1. 100 bp ladder from New England Biolabs. Lane 2. PCR amplification reactions for detection of *attB2* in the wild type strain DZ1. Lanes 3 and 4. PCR amplification reactions for detection of *attB2* in two independent isolates that contain a plasmid integrated at *attB1*. Lanes 5 and 6. PCR amplification reactions for detection of *attB2* in two independent isolates that contain a plasmid integrated at *attB2*.

[0014] Figure 6A shows the *lacZ* gene transcribed from the *pilA* promoter integrated at the either the *pilA* chromosomal location, Mx9 *attB1* or *attB2,* or the Mx8 *attB* sites. Figure 6B and Figure 6C show the *lacZ* gene transcribed from the *mgl* promoter integrated at the either the *mgl* chromosomal location, Mx9 *attB1* or *attB2,* or the Mx9 *attB* sites.

[0015] Figure 7 shows the consensus sequence of a *Chrysoperla carnea* transposase gene [SEQ ID NO:19].

DETAILED DESCRIPTION OF THE INVENTION

[0016] The present invention provides methods and materials for transforming bacterial cells using a bacteriophage Mx9 transformation system (also called an Mx9 enzyme system). In one aspect, the invention provides an Mx9 transformation system

that may be used to introduce DNA into a bacterial cell comprising an *attB* site. The Mx9 transformation system comprises (1) a gene encoding a protein with Mx9 integrase activity and (2) a DNA vector comprising an attachment site (*attP*) recognized by the attachment site. The *int* gene product catalyses recombination between the *attP* and *attB* sites, resulting in integration of DNA sequences from the DNA vector. Proteins with Mx9 integrase activity, the *attP* site, and *attB* site are described in detail below.

[0017] In one embodiment of the invention, the *attB* site comprises the 42-b core sequence [SEQ ID NO:5]. The *attB* site may further include at least a portion of the sequences flanking the *attB1* and/or *attB2* site core sites (e.g., *attR* and *attL*, discussed below, which comprise portions of SEQ ID NOS: 3, 4 and 6). In an embodiment, the *attP* site comprises the 42-b core sequence [SEQ ID NO:5]. The *attP* site may further include at least a portion of the sequences flanking the core sequence, e.g., as shown in Figure 3D. In an embodiment, the protein with Mx9 integrase activity (hereinafter, "*int* protein") is the product of the *int* gene having the sequence of SEQ ID NO:2. It will be apparent to the reader that the *attB* site, *attP* site and *int* protein used in the practice of the invention need not be identical to those of the naturally occurring Mx9-*Myxococcus xanthus* system and that the invention can be practiced using an having sequences substantially identical to those of the naturally occurring sequences. For example, the *int* protein can differ from SEQ ID NO:2 by conservative amino acid replacements or other substitutions, so long as it has Mx9 integrase activity, i.e. catalyses recombination between *attP* and *attB* sites having the sequences of SEQ ID NO:7 and 4, respectively (see Figure 3). Conversely, the *attP* and *attB* sites can differ from naturally occurring sites (and may comprise only a fraction of SEQ ID NO:7, 3, 4, or 6), so long as they are recognized by the *int* protein having a sequence of SEQ ID NO:2.

[0018] In one embodiment, the protein with Mx9 integrase activity has the sequence shown in Figure 2 [SEQ ID NO:2], or has a substantially identical sequence. In this context, substantial sequence identity means at least about 70%, more often at least about 80%, most often at least about 90% identity. Sequence identity can be calculated according to the method of Pearson and Lipman, 1988, *Proc. Natl. Acad. Sci.* U.S.A. 85:2444 using default parameters. In an aspect the invention provides an integrase

having the sequence shown in Figure 2 [SEQ ID NO:2] or having a substantially identical sequence and having integrase activity (e.g., when substrates are the sequence of *attP* and *attB2* sites shown in Figure 3). In an aspect, the integrase is encoded by a DNA having the sequence of SEQ ID NO:1 or a substantially identical sequence, e.g., at least about 70%, at least about 80%, at least about 90%, or at least about 95% identical (which can be calculated for nucleic acids using the method of Altschul, 1990, *J. Mol. Biol.* 215:403-10 using default parameters). In another aspect, the invention provides an isolated or recombinant DNA molecule comprising the sequence of SEQ ID NO:1 or a substantially identical sequence (e.g., at least about 70%, more often at least about 80%, most often at least about 90% identity). In a related aspect, the invention provides an isolated or recombinant DNA molecule comprising a sequence encoding SEQ ID NO:2 or a substantially identical sequence (e.g., at least about 70%, at least about 80%, or at least about 90% identity). In some embodiments the isolated or recombinant DNA is less than 5000, less than 1000, less than 5000 or less than 2000 bases in length. In one aspect, the invention provides a recombinant vector comprising an integrase encoding gene. In an embodiment, the gene is operably linked to a promoter that functions in a host cell, so that upon introduction into a cell the integrase is expressed in a host cell.

[0019] In an aspect, the *attP* and *attB* sites comprise the 42-base core sequence, and may also comprise at least about 10, at least about 20, at least about 30, at least about 40, at least about 50, at least about 100, or all, of one or more of the flanking sequences shown for *attP*, *attB1* or *attB2* in Figure 3 [e.g., SEQ ID NOS:7, 3, and 4 respectively], or a substantially identical sequence. The *attB* and *attP* core sequences may be sufficient for recombination. Alternatively, at least a portion of the flanking sequence(s) may be necessary for recombination or improve recombination frequency. The precise extent of sequence required for efficient recombination can easily be determined using routine assays for recombination using a series of constructs comprising different amounts of sequence.

[0020] In an aspect, the invention provides an isolated or recombinant DNA molecule comprising a sequence selected from a sequence comprising the Mx9 *attB1* site [SEQ ID NO:3]; the Mx9 *attB2* site [SEQ ID NO:4]; the Mx9 native *attB1* site [SEQ ID

NO:6], the *attR* site of *attB1* [nucleotides 205-360 of SEQ ID NO:3], the *attR* site of *attB2* [nucleotides 207-360 of SEQ ID NO:4], the *attL* site of *attB1* [nucleotides 1-162 of SEQ ID NO:3] or the *attL* site of *attB2* [nucleotides 1-164 of SEQ ID NO:4], or, alternatively, at least about 10, at least about 20, at least about 30, at least about 40, at least about 50, at least about 100, from, or all of, an aforementioned sequence. In some embodiments the isolated or recombinant DNA is less than 5000, less than 1000, less than 500 or less than 200 bases in length. In an aspect, the invention provides an isolated or recombinant DNA molecule comprising a 42 base sequence corresponding to nucleotides 165-206 of SEQ ID NO:4, i.e., SEQ ID NO:5. In an aspect, the invention provides an isolated or recombinant DNA molecule comprising an *attP* sequence. In one embodiment the *attP* sequence consists of or comprises SEQ ID NO:5, or alternatively, SEQ ID NO:7, or at least 50, at least 100, or at least 150 bases of SEQ ID NO:7 (generally including the core sequence). The invention provides recombinant vectors comprising any of the aforementioned DNA molecules.

[0021] In one aspect the *attB* and *attP* sites comprise identical sequences, e.g., 42 base pair core sequences. In an embodiment, the *attB* site is located within the 5′ region of the $tRNA^{gly}$ gene of the host cell. In another aspect, the one or more *attB* sites are comprised of *attB1* and/or *attB2*. In an embodiment, the present invention provides methods wherein the target DNA for the Mx9 transformation system comprises flanking sites *attR* and *attL*, and the integrase protein, when expressed, is an enzyme that facilitates site-specific recombination through binding to the *attP* and *attB* sites.

[0022] The *int* gene and *attP* site may be situated on the same vector. However, the integrase can function in *trans* and, accordingly, the sites can be introduced on different vectors. In another embodiment of the invention, the vector comprising an *attP* site is introduced into a recombinant cell expressing the *int* gene (e.g., a cell stably transformed with *int* protein encoding gene). As used herein, "vector" has its usual meaning in the art, and refers to polynucleotide elements that are used to introduce recombinant nucleic acid into cells for either expression or replication. Exemplary vector classes include recombinant DNA or RNA constructs, such as a plasmid, a phage, recombinant virus or other vectors. An "expression vector" is a vector capable of

expressing DNAs that are operatively linked with regulatory sequences, such as promoter regions. It will be appreciated by those of skill that the vectors may contain additional elements for selection (e.g., antibiotic resistance markers), cloning (e.g., polylinkers), replication, and the like. Appropriate expression vectors are well known to those of skill in the art and include those that are replicable in prokaryotic cells, and those that remain episomal or those which integrate into the host cell genome (the term "host" cell refers to the cell into which the *attP* containing vector is introduced). It will be appreciated that a naturally occurring (non-recombinant) Mx9 phage is not itself a vector, although a recombinant Mx9 phage modified to carry a heterologous DNA would be considered a vector.

[0023] The integrase gene of the Mx9 transformation system is operably linked to a promoter that functions in the intended host. Numerous prokaryotic, viral and synthetic promoters are known in the art and include, for example *act* promoters, *tcm* promoters, promoters derived from sugar metabolizing enzymes, such as galactose, lactose (*lac*) and maltose, promoters derived from biosynthetic enzymes such as for tryptophan (*trp*), the β-lactamase (*bla*), bacteriophage lambda PL and T5, synthetic promoters, such as the *tac* promoter (U.S. Patent No. 4,551,433), and mariner-type promoters may be used Exemplary promoters for Myxococcus cells include the native *int* gene promoter, the *pilA* promoter and the *mgl* promoter (see Wu and Kaiser, 1997, "Regulation of expression of the pilA gene in *Myxococcus xanthus*" *J. Bacteriol.* 179:7748-7758 and GenBank accession number AF377950).

[0024] The methods of the present invention may be used to transform any of a variety of host cells that comprise an *attB* attachment site recognized by the *int* gene product. Importantly, cells that lack a required integration or attachment site can be genetically engineered to contain one or more such sites, and the integrase gene can be placed under the control of a desired promoter. Thus, the invention can be applied to virtually any host cell. The invention is particularly suited for Myxobacteria, such as Sorangium or Myxococcus. In certain embodiments, the host cells of the present invention may be *Sorangium* cells (e.g., Sorangium cellulosum), *Myxococcus* cells (e.g.,

Myxococcus xanthus), Cystobactera, bacteria of order Stigmatella (e.g., *S. erecta* and *S. aurantiaca*), *Pseudomonas* cells, or *Streptomyces* cells.

[0025] Methods for introducing the recombinant vectors and exogenous DNA molecules of the present invention into suitable hosts are known to those of skill in the art and typically include the use of $CaCl_2$ or other agents, such as divalent cations, lipofection, DMSO, protoplast transformation, conjugation, or electroporation. References herein to "transformation" and its grammatical equivalents is intended to encompass any method of introducing an exogenous DNA into a cell.

[0026] In one aspect, the present invention is directed to methods of transforming deoxyribonucleic acid (DNA) into a bacterial host cell to effectuate or improve polyketide expression. In one embodiment, the method comprises a) introducing a gene to the DNA of a bacteriophage Mx9 transformation system, said system comprising a gene encoding an integrase protein (*int*) and an attachment site (*attP*); b) introducing said bacteriophage Mx9 transformation system to a host cell that contains a nucleotide sequence encoding a polyketide and one or more integration sites (*attB*) located in the DNA of said host cell; and c) transforming said host cell with said gene by site-specific recombination at the one or more *attB* sites.

[0027] As noted, the invention provides materials and methods useful for insertion of a gene or genes into a host cell, even if that host cell lacks an Mx9 attachment site. Thus, in accordance with the methods of the invention, such host cells can be modified to include the required attachment site. One useful method for modifying host cells to include an Mx9 attachment site is transposon-based transformation (see provisional patent application no. 60/403,290 (filed August 13, 2002) and U.S. patent application no. 10/_____, filed August 13, 2003, entitled "Transposon-Based Transformation System," having attorney docket number 30062-2009800). In one embodiment, a transposon vector comprising (1) inverted terminal repeat sequences (ITRs) comprising the sequence ACAGGTTGGCTGATAAGTCCCCGGTCT [SEQ ID NO:17] GGATCCAGACCGGGGACTTATCAGCCAACCTGT [SEQ ID NO:18] and (2) a gene encoding a transposase having a sequence shown in Fig. 7, optionally comprising an E137K mutation, operably linked to a T7A1 promoter (Lanzer et al., 1988,

Proc. Nat'l Acad Sci 85:8973-77) is used. In one embodiment, an *attB* site is introduced into a bacterial cell genome by a) transforming the cell with a transposon vector comprising inverted repeat sequences and a nucleotide sequence comprising a bacteriophage Mx9 integration site (*attB*), whereby the transposon vector transposes into the DNA of said cell; b) introducing a gene to the bacteriophage Mx9 transformation system, said system comprising a gene encoding an integrase protein (*int*) and an attachment site (*attP*); c) introducing said bacteriophage Mx9 transformation system to a host cell; and d) transforming said host cell with said gene by site-specific recombination at said *attB* site. In one aspect, the invention provides a method for a) transforming a cell that contains a nucleotide sequence encoding a polyketide synthase with a transposon vector comprising inverted repeat sequences and a nucleotide sequence comprising a bacteriophage Mx9 integration site (*attB*), whereby the transposon vector transposes into the DNA of said cell; b) introducing a gene into a bacteriophage Mx9 transformation system, said system comprising a gene encoding an integrase protein (*int*) and an attachment site (*attP*); c) introducing said bacteriophage Mx9 transformation system to a host cell; and d) transforming said host cell with said gene by site-specific recombination at said *attB* site.

[0028] In another aspect, vectors useful for introducing genes into host cells containing an Mx9 integration site are provided. In a particular aspect, vectors of the present disclosure include (1) vectors (including bacteriophage and plasmid vectors) comprising DNA encoding an Mx9 phage attachment site (*attP*), and another gene, and (2) vectors comprising DNA encoding an integrase protein, an Mx9 phage attachment site (*attP*), and another gene. The other gene can be any DNA sequence that is desired to be introduced into the target cell, whether encoding a protein or not. As described below, in some embodiments, the gene changes or improves polyketide production in a polyketide producing cell.

[0029] In another aspect, the present invention provides host cells, including *e.g.*, *M. xanthus* host cells, comprising genes introduced by the described methods. In one embodiment, the present methods, materials, host cells and vectors are directed to enhancing the production of a useful compound, including but not limited to a polyketide,

through the introduction of one or more genes into the DNA of a variety of bacterial host cells. Thus, in one aspect, transformed host cells are provided that are produced by the claimed methods, which host cells comprise one or more genes integrated to effectuate or improve polyketide expression by the cell. For example, *M. xanthus* may be used, for example, for the production of epothilone (4; US Pat. No. 6,410,301 "Myxococcus host cells for the production of epothilones") and genes may be introduced into such epothilone-producing cells to affect the amount, structure or other characteristics of the polyketide produced. In one embodiment, host cells of the present invention are epothilone-producing cells, wherein the epothilone produced is generally selected from epothilone A, B, C, and D.

[0030] In one aspect, a gene that improves polyketide production upon functional integration into the DNA of a host cell is introduced into a cell that expresses, or can be engineered to express, a polyketide synthase. In one aspect, the genes introduced into a host cell by the methods of the invention comprise an operon of a *prpE* gene, *accA*, and *pccB* genes to produce increased quantities of malonyl-CoA and/or methylmalonyl-CoA. The genes can be under the control of a suitable promoter, such as a PKS promoter, *i.e.*, from epothilone (U.S. Pat. No. 6,303,342; U.S. Patent Application Serial No. 09/957,483, filed September 19, 2001), soraphen (U.S. Pat. No. 5,716,849, incorporated herein by reference), or tombamycin (U.S. Patent Application Serial No. 09/942,025, filed August 28th, 2001, and U.S. Pat. Nos. 6,280,999, and 6,090,601, each of which is incorporated herein by reference) gene clusters. The gene or genes are inserted in a recombinant bacteriophage Mx9 of the invention and then integrated into the DNA of the host cell. In one aspect the *prpE* gene, *accA*, and *pccB* genes are inserted into a *Myxococcus xanthus* cell.

[0031] In another aspect, the genes inserted into the host cell may comprise a *matB* gene or an operon comprising *matB* and *matC* genes, such as those from *Rhizobium leguminosarum* bv. *trifolii,* which respectively encode a ligase that can attach a CoA group to malonic or methylmalonic acid and a transporter molecule to transport malonic or methylmalonic acid into the host cell respectively, to produce increased quantities of malonyl-CoA and methylmalonyl-CoA (U.S. patent application Serial Nos. 09/687,555,

filed October 13, 2000; 09/798,033, filed February 28, 2001; and 10/087,451, filed February 28, 2002; each of which is incorporated herein by reference).

[0032] In another aspect, vectors useful for introducing genes into host cells containing an Mx9 integration site are provided. In a particular aspect, vectors of the present disclosure include bacteriophage vectors comprising DNA encoding an integrase protein, an Mx9 phage attachment site (*attP*), and another gene. In an embodiment, the vector is a plasmid vector. In a related aspect, the invention provides a vector selected from the group consisting of pKOS35-93, pKOS35-117.9.7, pKOS249-12, pKOS249-23, and pKOS249-31. In one aspect of the invention, an Mx9 transformation system is used to introduce DNA into a host chromosome.

[0033] In related aspects, the invention provides a method of transforming a bacterial host cell, said method comprising the steps of a) introducing a first gene into a bacteriophage Mx9 transformation system, said system comprising a second gene encoding an integrase protein (*int*) and an attachment site (*attP*); b) introducing said bacteriophage Mx9 transformation system to a host cell that contains one or more integration sites (*attB*) located in the DNA of said host cell; and c) transforming said host cell with said first gene by site-specific recombination at the one or more *attB* sites. In an embodiment, the one or more *attB* sites are comprised of *attB1* (SEQ ID NO:3), *attB2* (SEQ ID NO:4), or a combination thereof. In an embodiment, the cells are *Myxococcus* cells, for example epothilone-producing cells. In an embodiment, the epothilone is selected from the group consisting of epothilone C and D. In some embodiments, the first gene is selected from the group consisting of *prpE*, *accA*, *pccB*, *matB* and *matC* genes. In an embodiment of the invention, the *attB* and *attP* sites are comprised of identical sequences, which may be identical 42 base pair sequences corresponding to nucleotides 1394-1435 of SEQ ID NO:1. In an embodiment, the *attB* site is located within the 5′ region of the $tRNA^{gly}$ gene. In an embodiment of the method, DNA from said *attR* site is deleted upon transformation of said host cell. In an embodiment, the gene encoding an integrase protein is altered upon transformation of said host cell.

[0034] The invention also provides a transformed bacterial host cell produced by an aforementioned method. In an embodiment, the host cell produces an epothilone

selected from epothilone A, B, C, and D. Optionally, the first gene is selected from the group consisting of *prpE*, *accA*, *pccB*, *matB* and *matC* genes.

[0035] In an aspect, the invention provides a method of transforming a bacterial host cell that lacks a bacteriophage Mx9 integration site (*attB*) to improve polyketide expression, said method by a) transforming a host cell with a transposon vector comprising inverted repeat sequences and a nucleotide sequence comprising a bacteriophage Mx9 integration site (*attB*), whereby the transposon vector transposes into the DNA of said cell; b) introducing a first gene to a bacteriophage Mx9 transformation system, said system comprising a second gene encoding an integrase protein (*int*) and an attachment site (*attP*); c) introducing said bacteriophage Mx9 transformation system to the host cell; and d) transforming said host cell with said first gene by site-specific recombination at said *attB* site. According to this method, the host cells may be *Sorangium* cells, *Myxococcus* cells, *Pseudomonas* cells, or *Streptomyces* cells as well as others. In embodiments, the host cells produce epothilone selected from epothilone A, B, C, and D and/or the first gene is selected from the group consisting of *prpE*, *accA*, *pccB*, *matB* and *matC* genes and/or the attB site comprises flanking sites *attR* and *attL*, and said integrase protein, when expressed, is an enzyme that facilitates said site-specific recombination through binding to *attB* and *attP* sites. The invention further provides a transformed bacterial host cell produced by this method, which optionally may produce an epothilone selected from epothilone A, B, C, and D.

[0036] The invention also provides a bacteriophage Mx9 vector comprising DNA encoding an integrase protein, an Mx9 phage attachment site (*attP*), and another gene.

Experimental Results and Discussion

Materials and Methods

[0037] Bacteria, Phage, and plasmids. DZ1 is a nonmotile strain of *M. xanthus* and was used for plating Mx9 and for characterization of the Mx9 attachment sites (12). DK816 is the natural *M. xanthus* isolate lysogenic for Mx9 (9). *M. xanthus* strains were grown in CYE medium (1) or 1% CTS (1% casitone, 0.2% $MgSO_4 \cdot 7H_2O$, 50 mM HEPES pH 7.6). Phleomycin (Cayla) was used at a concentration of 30 μg/ml. The Mx9

phage was reisolated from DK816 by growing a culture to stationary phase, pelleting the cells, and plating dilutions of the supernatant onto DZ1. High titer stocks of Mx9 were made by coring a plaque and placing it in phage buffer (10 mM MOPS [pH7.6], 4 mM $MgCl_2$, 2 mM $CaCl_2$). The eluted phage were diluted and mixed with 0.5 ml of DZ1 in early stationary phase. After incubating the cells and phage at room temperature for 20 minutes, 2.5 ml of top agar was added and the suspension was poured onto phage plates (1% BBL trypticase, 0.1% $MgSO_4 \cdot 7H_2O$, 1% agar, 10 mM MOPS pH 7.6). The plates that gave confluent lysis after 2 days of incubation at 30°C were overlayed with 5 ml of phage buffer and incubated at 4°C overnight. The eluted phage were stored at 4°C. Phage stocks greater than 1 $x10^9$ pfu/ml were obtained with this method. Plasmids used are described in Table 1.

Table 1

Plasmid	Characteristics
pKOS35-117.9.9	amp^r kan^r colEI, 4.6 kb fragment from Mx9
pKOS139-29	amp^r, colEI, P_{T7A1} Mx8 *int attP-*
pKOS139-47	tc^r, p15A, P_{mgl} *lacZ*, Mx8 *attP*
pKOS178-86	tc^r, p15A, P_{pilA} *lacZ*, Mx8 *attP*
pKOS178-177	tc^r, p15A, P_{pilA} *lacZ*, Mx9 *int attP*
pKOS178-188	tc^r, p15A, P_{mgl} *lacZ*, Mx9 *int attP*
pKOS249-31	amp^r $bleo^r$ colEI, P_{T7A1} Mx9 *int attP*

[0038] Isolation of phage DNA. The phage from a high titer stock were pelleted by centrifuging in an SS-34 rotor at 28,000 rpm for 3 hours and then resuspended in TE (10 mM Tris [pH7.6] 1 mM EDTA). The phage proteins were removed by extracting twice with phenol and twice with phenol/chloroform/isoamylalcohol. The DNA was precipitated and resuspended in TE.

[0039] Isolation and sequence of the phage attachment site. To isolate the phage attachment site, phage DNA was partially cleaved with *Hin*PI and the fragments were ligated into pKOS35-93 cleaved with *Acc*I. The plasmid pKOS35-93 is pBluescriptII SK+ with the kanamycin resistance from Tn*5* ligated into the *Sma*I and *Eco*RI sites. One

plasmid, pKOS35-117.9.7, integrated efficiently into the chromosome. The insert from this plasmid was sequenced

[0040] Isolation of the bacterial attachment site. The bacterial attachment site (*attB*) was isolated by electroporating pKOS35-117.9.7 into DZ1, making chromosomal DNA, and then recovering the plasmid with flanking chromosomal DNA. Six kanamycin resistant colonies were picked and chromosomal DNA was prepared from each. The DNA was cleaved with either *Pst*I or *Xho*I, ligated, and then transformed into *E. coli*. Three colonies from each of the electroporations were picked and the recovered plasmids were cleaved with *Pst*I or *Xho*I. One plasmid from each was sequenced using either primer 183-66.3 (GAAGGAGGCACCATGCACGG [SEQ ID NO:8] or 183-66.4 (CTCACTGAGAGTGAAGCCGC [SEQ ID NO:9]).

[0041] PCR amplification of the Mx9 *attB*. Primers were designed to PCR amplify *attB1* and *attB2*. Primers 183-99.4 (CGAGGTCCGGGACGCGCGCA [SEQ ID NO:10]) and 183-99.6 (TGCCAGGGCTTACGGCTTC [SEQ ID NO:11]) were used to amplify a 285 bp *attB1* fragment and 183-99.5 (TATCCCAGCAACCGCCGGAG [SEQ ID NO:13]) with primer 183-99.4 was used to amplify a 373 bp *attB2* fragment. To amplify the native *attB1* site primers 183-99.6 and 249-179.7 (CAGCACGGGTGCAGCAAC [SEQ ID NO:14]) were used to amplify a 250 bp fragment. PCR reactions were done using chromosomal DNA from DZ1 and the FailSafe™ PCR system from Epicentre. Amplification conditions were 96°C for two minutes and then 30 cycles of 94°C 30 seconds, 55°C for 1 minute, 72°C for 2 minutes.

[0042] Construction of a minimal integration plasmid. The *int* gene was PCR amplified from pKOS35-117.9.7 using the primers 111-74.4 (CCCAATTGGCTCAGGGCAGCGGCTCATT [SEQ ID NO:15]) and 111-82.5 (CCCCATGGCGCTCAGGGGTGCGTCGGACGCC [SEQ ID NO:16]). PCR amplification conditions were those previously described. The amplified fragment was ligated into the *Eco*RV site of pLitmus 28 (New England Biolabs) to create pKOS249-12. The *int* gene was removed from this plasmid by cleaving with *Eco*RI, the DNA ends were made blunt with the Klenow fragment of DNA polymerase followed by cleaving with *Nco*I. The fragment was ligated with pUHE24-2B (3) that was cleaved with *Pst*I, the

DNA ends were made blunt with the Klenow fragment of DNA polymerase I and cleaved with *Nco*I. The resulting plasmid, pKOS249-23, contains the *int* gene under the control of the *E. coli* phage T7 A1 promoter that has been engineered to contain 2 LacI binding sites to repress transcription. The bleomycin resistance gene was added to this plasmid by isolating the bleomycin resistance gene from pKOS183-112 as a *Bam*HI to *Hin*dIII fragment, the DNA ends were made blunt with the Klenow fragment of DNA polymerase I and ligating it with pKOS249-23, which was cleaved with *Xho*I and the DNA ends were made blunt with the Klenow fragment of DNA polymerase I. This plasmid is designated pKOS249-31.

[0043] β-galactosidase assays. Seed cultures of two isolates for each integration site were grown in 1% CTS (5 ml) to mid to late log phase. To start the assay cultures, 35 ml of CTS was inoculated with 1 ml of seed culture at an OD_{600} of 0.073. β-galactosidase assays were performed by removing an aliquot of cells and adding them to Z buffer for a combined volume of 1 ml. The cells were lysed by adding one drop of 0.1% SDS, two drops of chloroform, and vortexing the sample for 5 seconds. The assay was initiated by the addition of 0.1 ml of O-nitrophenyl β-D-galactopyranoside (8 mg/ml) and mixing. The reactions were stopped by the addition of 0.5 ml of 1 M Na_2CO_3. The OD_{600} of the cell culture and the OD_{420} of the enzyme reactions were determined using a SpetraMax 250 plate reader. Miller units were determined as previously described (10).

[0044] Accession numbers. The Mx9 sequence has been assigned the accession number AY247757. The accession numbers for *attB1* and *attB2* are AY297770 and AY297771, respectively.

[0045] Identification of the Mx9 *int* and attachment site. To identify the *int* gene and attachment site, a library of 5-8 kb fragments of Mx9 was made, and a clone that was able to integrate into the *M. xanthus* chromosome was identified. The insert in this plasmid, pKOS35-117.9.7, was sequenced. Five complete and one partial open reading frames (orf) were identified in the 4.6 kb fragment (Fig. 1). Orf 1 was the only reading frame that showed amino acid similarity with other known integrase genes, and therefore was given the gene designation *int*. The other orfs resembled orfs from Mx8; orf 2, orf3, orf4, orf5, and orf6 showed similarity to P15, P14, P16, P17, and P18, respectively from

Mx8. From the degree of similarity of these orfs between, it appears that Mx8 and Mx9 are very similar phages.

[0046] The Mx9 *int* gene was examined for sequences that would indicate an attachment site. Analysis revealed a DNA segment within the *int* gene (nt 1397-1428 (Figure 2)) that had sequence similarity to $tRNA^{gly}$ from various organisms. Since Mx8 integrates into the $tRNA^{Asp}$ gene of *M. xanthus*, the sequence that showed similarity with $tRNA^{gly}$ was predicted to serve as the site of integration for Mx9.

[0047] To test this prediction, chromosomal DNA from six integrants containing pKOS35-117.9.7 were cleaved with restriction enzymes, ligated, and transformed into *E. coli* to recover the plasmid along with flanking chromosomal DNA. Sequencing, using primers adjacent to the proposed attachment site, revealed that the point of recombination was indeed that of the putative $tRNA^{gly}$. Furthermore, the sequence of flanking chromosomal DNA showed that there were two *attB* sites. It appeared from the number of integrants at each site, 3 for *attB1* and 3 for *attB2*, that both served equally well as the insertion site (Figure 3).

[0048] Structure of the two *attB* sites. Figure 3 shows 360 bp from each of the *attB* sites. Both have a common 42 bp core sequence that is also found within the Mx9 *int* gene. In addition, there are 22 bp 5' to both *attB* sites that are identical in 21 positions. There is a putative inverted repeat that may play a role in Integrase protein binding at the *attB* and *attP* (Fig. 3b). The site of integration within *attB2* lies in the 5' end of $tRNA^{gly}$ gene, which is underlined in Figure 3b. However, the sequence of *attB1* does not contain a complete $tRNA^{gly}$ gene. Figure 4 shows the predicted folding of this segment of *attB2* into a corresponding tRNA.

[0049] Analysis of the *attR* and *attL* half-sequences for both *attB* sites reveals the two *attR* are identical whereas the *attL* differ. This is also the case with the two Mx8 *attB* sites (7). Plasmids containing the Mx8 *int* gene preferentially integrate at *attB1,* and this integration often is accompanied by a deletion between *attB1* and *attB2* (8).

[0050] To determine if the identical *attR* sites are due to the presence of two *attB* sites containing with identical *attR* sites or due to the deletion of the DNA between the

two *attB* sites after integration into one of them, PCR analysis was performed using either primer pair 183-99.4 and 183-99.6 for *attB1* or 183-99.4 and 183-99.5 for *attB2*.

[0051] A PCR fragment was detected using primers specific for *attB2* but none was detected using primers specific for *attB1* (data not shown). This suggests that a deletion may occur upon integration of *attB1* but to be certain that the lack of a PCR product was not due to the failure to PCR amplify the DNA fragment, further experiments were performed.

[0052] Next, the genomic sequence of *M. xanthus* strain DK1622, generated by Monsanto and available at the TIGR web site, was examined for the two *attB* sites (www.TIGR.org). The *attB2* sequence was almost identical to that previously identified (Fig. 3B) but only the first 178 bp of the *attB1* site from Figure 3A was present before the sequence diverged. Using this sequence information for *attB1*, a primer was designed that was approximately 100 bp downstream from the point at which the sequence diverged (249-179.7). Using this primer along with 183-99.6, the one 5' to the *attB1* site, and DZ1 genomic DNA, a PCR product of approximately 250 bp was isolated and sequenced. The PCR product was identical to that obtained from the DK1622 genomic sequence (Fig 3C). Analysis of this sequence reveals that only 16 bp of the 42 bp core *att* site are present in the native *attB1*site.

[0053] Final proof that a deletion does occur between *attB1* and *attB2* is shown in Figure 5. Using the primer pair 183-99.4 and 183-99.5, the ones that amplify the *attB2* site, PCR amplification was performed using genomic DNA from the wild type strain or strains harboring a plasmid integrated at either *attB1* or *attB2*. Using chromosomal DNA from DZ1, a strain with no plasmids integrated at either *attB* site, a 372 bp PCR product containing the *attB2* site was detected in lane 2 figure 5. Two strains that contain insertions at *attB2*, lanes 5 and 6 (Fig. 5) do not give the 372 bp band and should not amplify the *attB2* due to the presence of a plasmid integrated at that site. If a deletion does occur between attB1 and *attB2*, then there should be no detectable amplification of *attB2* when a plasmid integrates at *attB1*. Lanes 3 and 4 (Fig. 5) shows that no *attB2* PCR product is detected, indicating a deletion of DNA between *attB1* and *attB2* when an integration occurs at *attB1*.

[0054] Integration results in the alteration of the carboxy terminus of the Mx9 Int protein. Because *attP* lies within the *int* gene, integration into the chromosome should alter the 3' end of *int* gene is altered. From the 1160 bp of *attR* that has been sequenced, no stop codon has been identified (data not shown). Thus 70 amino acids from Int should be removed and more than 389 amino acids should be added to the Int protein that is synthesized after integration into the chromosome. These additional amino acids presumably will reduce the enzymatic activity of Int because the IntX protein of Mx8 has lost 112 residues and added 13 amino acids, and is a less active at site specific recombination (8).

[0055] Mx9 Int is the only phage protein required for integration. To determine whether *int* is necessary and sufficient for integration, the *int* gene was PCR amplified and ligated into an *E. coli* expression vector that uses an engineered phage T7 A1 promoter. The plasmid pKOS249-31, when electroporated into DZ1, integrated efficiently into the chromosome; approximately $1x10^4$ colonies were obtained per microgram of DNA. Thus, the Mx9 *int* gene is the only phage encoded protein required for integrative recombination into the bacterial chromosome.

[0056] Transcription from the *pilA* and the *mgl* promoters integrated at the two Mx9 *attB* sites. To find a phage attachment site on the *M. xanthus* chromosome that supported efficient expression of genes from a variety of promoters, fusions of *lacZ* to the *mgl* or *pilA* promoters were constructed and transcription from these promoters at the two Mx9 *attB*, the Mx8 *attB*, and the native chromosomal location was analyzed. Figure 6A shows the expression level of the *pilA* promoter (P_{pilA}) at the four different locations. Surprisingly, there was little transcription when the P_{pilA} plasmid was integrated by homologous recombination at the *pilA* location (pKOS178-86). This suggests that there may be a deletion in the *pilA* promoter region that abolishes activation of the *pilA* promoter in DZ1 since there was no expression in several isolates that were examined. As we have observed previously, little transcription from P_{pilA} is seen when integrated at Mx8 *attB* site (pKOS178-86 + pKOS139-29). However, the Mx9 sites show high levels of transcription from P_{pilA} (pKOS178-177) and they are fairly similar at both sites, although *attB2* had high variability of expression from the two isolates examined. In

addition, the regulation at both sites was similar; transcription from P_{pilA} increased during late log and stationary phases.

[0057] The results of transcription from the *mgl* promoter (P_{mgl}) are shown in Figure 6B. Transcription from P_{mgl} at the two Mx9 *attB* (pKOS178-188) sites was better than at the Mx8 site (pKOS139-47 + pKOS139-29) but not as high when integrated by homologous recombination at the chromosomal *mgl* location (pKOS139-47). However, this lower expression at the two Mx9 sites may be vector dependent. Using a plasmid that contained only the *attP* site and integrating it by supplying the *int* gene in trans, P_{mgl} functions just as well at both Mx9 sites as it does at the chromosomal *mgl* location (see Fig. 6C). In this experiment, a plasmid was constructed that contained the mgl promoter fused to lacZ and harbored only the Mx9 attP site. This plasmid was integrated into the Mx9 attB1 or attB2 by co-electroportating it with a second plasmid that expressed the *int* gene. β-galactosidase assays with cells containing this plasmid reveals that the levels of expression from the mgl promoter is as good, if not better, than the native mgl chromosomal location. Thus expression from the mgl promoter at the Mx9 attB locations may be vector dependent. The conclusion from these studies indicates that the Mx9 *attB* sites are good for expression of foreign or native genes.

[0058] The Mx9 *int* gene and attachment site have been identified, along with the site of integration into the *M. xanthus* chromosome. The analysis reveals remarkable similarity to the *int* gene and attachment site from the myxophage Mx8 (7, 8, 11). Both contain the *attP* within the *int* gene and integrate within a tRNA gene. They have two *attB* sites and it appears that adjacent chromosomal DNA is deleted when integration occurs at one of the sites. For both, Int is the only phage-encoded protein needed for integration.

[0059] A difference between the Mx8 and Mx9 phage integration systems is the length of their respective core sequences. The core sequence for Mx8 integration is smaller, composed of 29 bp. The *attB2* site has two nucleotides that differ at one end, which may account for the preference of Mx8 for inserting at *attB1*. The *att* core region for Mx9 is 42 bp, but of the two integration sites only *attB2* contains all 42 bases. The *attB1* site contains only 16 bases of the core sequence. The lack of a complete core

sequence in *attB1* may explain why there is always a deletion between *attB1* and *attB2* when integration occurs at *attB1*. The Int protein may bind to the inverted repeat within the 42 bp core. Binding of the λ Int protein to its *att* sites has been shown (5). Since the *attB1* contains half of the inverted repeat, only half of the necessary protein complex can form, but once it has assembled, it may interact with the complementary half of proteins from *attB2* to allow for integration. This would result in a looping out of the DNA between *attB1* and *attB2,* and its subsequent loss upon integration of DNA.

[0060] In our PCR reactions to detect *attB1* with primers 183-99.4 & 183-99.6, the conditions were such that if the distance between *attB1* and *attB2* was less than 2 kb, then a PCR product should have been detected. Since no product was observed, this suggests that the distance between the two sites is greater than 2 kb. Analysis of the DK1622 sequence shows that the two *attB* sites are 6.7 kb apart. Partial analysis of this sequence shows a couple open reading frames that have sequence similarity to transposase genes, suggesting the presence of a transposon. The other reading frame that was identified reveals high sequence similarity to proteins of unknown functions. Clearly, the open reading frames encoded in between the two *attB* sites are not critical for growth under laboratory conditions since strains with integrations at *attB1* have no visible growth defects.

[0061] References

1. Campos, J. M., and D. R. Zusman. 1975. Regulation of development in Myxococcus xanthus: effect of 3':5'-cyclic AMP, ADP, and nutrition. Proc. Natl. Acad. Sci. U S A 72:518-22.

2. Fisseha, M., M. Gloudemans, R. E. Gill, and L. Kroos. 1996. Characterization of the regulatory region of a cell interaction-dependent gene in Myxococcus xanthus. J Bacteriol 178:2539-50.

3. Julien, B., and R. Calendar. 1995. Purification and characterization of the bacteriophage P4 delta protein. J Bacteriol 177:3743-51.

4. Julien, B., and S. Shah. 2002. Heterologous expression of the epothilone biosynthetic genes in Myxococcus xanthus. Antimicrobial Agents Chemotherapy 46:2772-2778.

5. Landy, A. 1989. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. Annual Review of Biochemistry 58:913-949.

6. Li, S. F., and L. J. Shimkets. 1988. Site-specific integration and expression of a developmental promoter in Myxococcus xanthus. J Bacteriol 170:5552-6.

7. Magrini, V., C. Creighton, and P. Youderian. 1999. Site-specific recombination of temperate Myxococcus xanthus phage Mx8: genetic elements required for integration. J Bacteriol 181:4050-61.

8. Magrini, V., M. L. Storms, and P. Youderian. 1999. Site-specific recombination of temperate Myxococcus xanthus phage Mx8: regulation of integrase activity by reversible, covalent modification. J Bacteriol 181:4062-70.

9. Martin, S., E. Sodergren, T. Masuda, and D. Kaiser. 1978. Systematic isolation of transducing phages for Myxococcus xanthus. Virology 88:44-53.

10. Miller, J. H. 1992. A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria. Cold Spring Harbor Laboratory Press, NY.

11. Tojo, N., K. Sanmiya, H. Sugawara, S. Inouye, and T. Komano. 1996. Integration of bacteriophage Mx8 into the Myxococcus xanthus chromosome causes a structural alteration at the C-terminal region of the IntP protein. J Bacteriol 178:4004-11.
12. Zusman, D. R., D. M. Krotoski, and M. Cumsky. 1978. Chromosome replication in Myxococcus xanthus. J. Bacteriol. 133:122-129.

[0062] Numerous modifications may be made to the foregoing systems without departing from the basic teachings thereof. Although the present invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the claims which follow. All publications and patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0063] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents.

CLAIMS

We claim:

1. A method for modification of a DNA of a bacterial cell comprising in its genome a first attachment site recognized by a protein with Mx9 integrase activity, comprising introducing a Mx9 transformation system into the cell, said system comprising

a) a gene encoding a protein with Mx9 integrase activity protein operably linked to a promoter active in the host cell, and

b) a DNA vector comprising a second attachment site recognized by the integrase protein, which may be the same as the first attachment site.

2. The method of claim 1 wherein the cell is *Myxococcus* or *Sorangium*.

3. The method of claim 1 wherein the protein has a sequence at least substantially identical to SEQ ID NO:2.

4. The method of claim 3 wherein the protein has a sequence of SEQ ID NO:2.

5. The method of claim 4 wherein the protein is encoded by a gene comprising the sequence of SEQ ID NO:1.

6. The method of claim 1 wherein said first attachment site comprises SEQ ID NO:5.

7. The method of claim 6 wherein said first attachment site is *attB2*.

8. The method of claim 1 wherein said second attachment site comprises SEQ ID NO:5.

9. The method of claim 3 wherein said first attachment site has been recombinantly introduced into the cell genome.

10. The method of claim 1 wherein said DNA vector further comprises an exogenous gene.

11. The method of claim 10 wherein the exogenous gene is selected from the group consisting of *prpE*, *accA*, *pccB*, *matB*, *matC* and beta-galactosidase genes.

12. The method of claim 6 wherein the first and second attachment sites are comprised of identical sequences.

13. The method of claim 2 wherein the cell is *Myxococcus xanthus*.

14. The method of claim 13 wherein the cell produces an epothilone.

15. The method of claim 14, wherein the epothilone is selected from the group consisting of epothilone C and D.

16. A bacterial host cell produced by the method of claim 10.

17. The cell of claim 16 wherein that produces an epothilone selected from epothilone A, B, C, and D.

18. The cell of claim 17, wherein said exogenous gene is selected from the group consisting of *prpE*, *accA*, *pccB*, *matB* and *matC* genes.

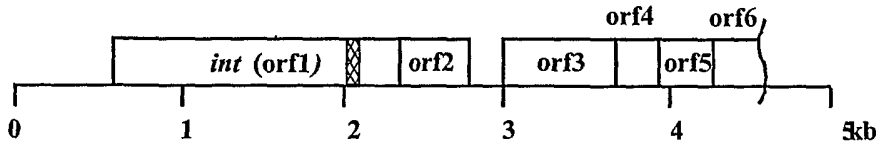

Figure 1

Figure 2

```
        10         20         30         40         50         60         70         80         90        100        110        120
GTGGCGCTCAGGGGTGCGTCGGACGCCACTACCAACCCCTCTCGACTTGTGCAGTCCGTCGCCGCCGGCCCGCGTGCGACTCCGTGGGGTGTCAGTGCGTCGTGGTACCTGCTAGGGCGT
CACCGCGAGTCCCCACGCAGCCTGCGGTGATGGTTGGGGAGAGCTGAACACGTCAGGCAGCGGCGGCCGGGCGCACGCTGAGGCACCCCACAGTCACGCAGCACCATGGACGATCCCGCA
 V  A  L  R  G  A  S  D  A  T  T  N  P  S  R  L  V  Q  S  V  A  A  G  P  R  A  T  P  W  G  V  S  A  S  W  Y  L  L  G  R

        130        140        150        160        170        180        190        200        210        220        230        240
ACAGCAACGGGGGAGTACATCGTGAGTAGCGACGCGGCGAAGAAGGGCCATCCAATGGCAACTGCGGCGGAGCGGTTGCCGACGTCACCAATCGACGTCAACGCTCTGGCGCTGGAGGTG
TGTCGTTGCCCCCTCATGTAGCACTCATCGCTGCGCCGCTTCTTCCCGGTAGGTTACCGTTGACGCCGCCTCGCCAACGGCTGCAGTGGTTAGCTGCAGTTGCGAGACCGCGACCTCCAC
 T  A  T  G  E  Y  I  V  S  S  D  A  A  K  K  G  H  P  M  A  T  A  A  E  R  L  P  T  S  P  I  D  V  N  A  L  A  L  E  V

        250        260        270        280        290        300        310        320        330        340        350        360
GCCCGGCTTGTGGCCCTCCAGCAGCAAAGTGCGACGCCGCCATCGTCCGGCCGCACTTTCGGCGCGGTGGCGGATGACTGGCTCATCACTGAGGCCAAGCGCCTCGTGTGCCCCGACAAT
CGGGCCGAACACCGGGAGGTCGTCGTTTCACGCTGCGGCGGTAGCAGGCCGGCGTGAAAGCCGCGCCACCGCCTACTGACCGAGTAGTGACTCCGGTTCGCGGAGCACACGGGGCTGTTA
 A  R  L  V  A  L  Q  Q  Q  S  A  T  P  P  S  S  G  R  T  F  G  A  V  A  D  D  W  L  I  T  E  A  K  R  L  V  C  P  D  N

        370        380        390        400        410        420        430        440        450        460        470        480
GAGCGCCGCCATCTTCGCCATATGGAGGCGCTCTGGGGCATGACGGATGTGGAGCTCACGCCGCGCGTCGTGAAGGCGCACCTGGCGGGACTTCTCAAGCCAGAGGGGCCGCTGAGCGCA
CTCGCGGCGGTAGAAGCGGTATACCTCCGCGAGACCCCGTACTGCCTACACCTCGAGTGCGGCGCGCAGCACTTCCGCGTGGACCGCCCTGAAGAGTTCGGTCTCCCCGGCGACTCGCGT
 E  R  R  H  L  R  H  M  E  A  L  W  G  M  T  D  V  E  L  T  P  R  V  V  K  A  H  L  A  G  L  L  K  P  E  G  P  L  S  A

        490        500        510        520        530        540        550        560        570        580        590        600
GCCACCGTCAATAAGGTGCGCTCTACCGGCAAGCGCATCATCAAGGCGGCGCAAATCAACGGCGAGTGGGGCCCGGTGAATCCTTTCGGCGTGCTCGACCGCGAAAAAGAGGCGAAGGCC
CGGTGGCAGTTATTCCACGCGAGATGGCCGTTCGCGTAGTAGTTCCGCCGCGTTTAGTTGCCGCTCACCCCGGGCCACTTAGGAAAGCCGCACGAGCTGGCGCTTTTTCTCCGCTTCCGG
 A  T  V  N  K  V  R  S  T  G  K  R  I  I  K  A  A  Q  I  N  G  E  W  G  P  V  N  P  F  G  V  L  D  R  E  K  E  A  K  A

        610        620        630        640        650        660        670        680        690        700        710        720
GAGCGCCTCACGCTGACGGCAGCGGAGTGCCGGGCGGTGCTCCCGCACTTCCGCGCGGACCGGCGCCGCGAGTTTCTCTTCCAGGTCTTTCTGGGGCCACGCCCCGGCGAAGAGAAGGCG
CTCGCGGAGTGCGACTGCCGTCGCCTCACGGCCCGCCACGAGGGCGTGAAGGCGCGCCTGGCCGCGGCGCTCAAAGAGAAGGTCCAGAAAGACCCCGGTGCGGGGCCGCTTCTCTTCCGC
 E  R  L  T  L  T  A  A  E  C  R  A  V  L  P  H  F  R  A  D  R  R  R  E  F  L  F  Q  V  F  L  G  P  R  P  G  E  E  K  A

        730        740        750        760        770        780        790        800        810        820        830        840
CTCCTCAAGGAAGATGTGGACGTCGAGGCGCGCACCGTCATTTTCCGGCGCAGCAATGGACGAGACACGACAAAGACGGGACGCGAGCGTCGCGTGCCGGTGCCGGATGAGTTGTGGCCC
GAGGAGTTCCTTCTACACCTGCAGCTCCGCGCGTGGCAGTAAAAGGCCGCGTCGTTACCTGCTCTGTGCTGTTTCTGCCCTGCGCTCGCAGCGCACGGCCACGGCCTACTCAACACCGGG
 L  L  K  E  D  V  D  V  E  A  R  T  V  I  F  R  R  S  N  G  R  D  T  T  K  T  G  R  E  R  R  V  P  V  P  D  E  L  W  P

        850        860        870        880        890        900        910        920        930        940        950        960
GTGCTCCTCGATGCGATGCAGGCCAGTCCGTCTGACCTCGTTTTCCCGAACGCGAAGGGTGAGAGGCAGCGCGCAGACACGAAGATGACGCGCGTGCTGCGCACTGCGCTATCCGCGGCT
CACGAGGAGCTACGCTACGTCCGGTCAGGCAGACTGGAGCAAAAGGGCTTGCGCTTCCCACTCTCCGTCGCGCGTCTGTGCTTCTACTGCGCGCACGACGCGTGACGCGATAGGCGCCGA
 V  L  L  D  A  M  Q  A  S  P  S  D  L  V  F  P  N  A  K  G  E  R  Q  R  A  D  T  K  M  T  R  V  L  R  T  A  L  S  A  A

        970        980        990       1000       1010       1020       1030       1040       1050       1060       1070       1080
GGTGTCGTGGTGGGCTGGGATTACATCTGCCGCACGCAGGGCTGCGGCTACCGAGATGTGCAGTCTGGTGGCGCGCGCCAGGAGCGTCGGTGCCCCGCCTGCGACAAGCGCATGTGGGCC
CCACAGCACCACCCGACCCTAATGTAGACGGCGTGCGTCCCGACGCCGATGGCTCTACACGTCAGACCACCGCGCGCGGTCCTCGCAGCCACGGGGCGGACGCTGTTCGCGTACACCCGG
 G  V  V  V  G  W  D  Y  I  C  R  T  Q  G  C  G  Y  R  D  V  Q  S  G  G  A  R  Q  E  R  R  C  P  A  C  D  K  R  M  W  A
```

```
    1090      1100      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
AGTGGTCGCCCCAAACCCGCCGTCTGGTACGGGCTCCGTCACACCGCGGCGACACTGCACAGGAAGGCGGGCTGCGACCCGCTCGTCATCAAGCTCGTGCTGGGGCATGCGGCTGTCGAC
TCACCAGCGGGGTTTGGGCGGCAGACCATGCCCCGAGGCAGTGTGGCGCCGCTGTGACGTGTCCTTCCGCCCGACGCTGGGCGAGCAGTAGTTCGAGCACGACCCCGTACGCCGACACGCTG
  S  G  R  P  K  P  A  V  W  Y  G  L  R  H  T  A  A  T  L  H  R  K  A  G  C  D  P  L  V  I  K  L  V  L  G  H  A  A  V  D

    1210      1220      1230      1240      1250      1260      1270      1280      1290      1300      1310      1320
ACCACGGACGACGTGTACACGCACCTCGACGAGGACTACTGCCGCGCCGAACTTAACAAGTTGTCGCTGAAGGCCCCGCCGCCACCACCTACTCACCAGGGAGGAAGTGACGGCGGCCCT
TGGTGCCTGCTGCACATGTGCGTGGAGCTGCTCCTGATGACGGCGCGGCTTGAATTGTTCAACAGCGACTTCCGGGGCGGCGGTGGTGGATGAGTGGTCCCTCCTTCACTGCCGCCGGGA
  T  T  D  D  V  Y  T  H  L  D  E  D  Y  C  R  A  E  L  N  K  L  S  L  K  A  P  P  P  P  P  T  H  Q  G  G  G  S  D  G  G  P

    1330      1340      1350      1360      1370      1380      1390      1400      1410      1420      1430      1440
GACTCAGGACGCAACACCTACGGTGAAGGAGGCACCATGCACGGATTGGGAGATTTGCAGCATCACCGGGCGAGAGCTTGGGAAGCTCGTGCTCTACCAACTGAGCTACCACCGCGGAAC
CTGAGTCCTGCGTTGTGGATGCCACTTCCTCCGTGGTACGTGCCTAACCCTCTAAACGTCGTAGTGGCCCGCTCTCGAACCTTCGAGCACGAGATGGTTGACTCGATGGTGGCGCCTTG
  D  S  G  R  N  T  Y  G  E  G  G  T  M  H  G  L  G  D  L  Q  H  H  R  A  R  A  W  E  A  R  A  L  P  T  E  L  P  P  R  N

    1450      1460      1470      1480      1490      1500      1510      1520      1530      1540      1550      1560
TTGGCCGGGGGTATACCGGCGCCGCTGCTGAGCGTCAAGGACGTTGCGGCTTCACTCTCAGTGAGCACGGCGAAGGTGTACCAGCTCCTCGCCGCCGGCGTCCTGCCTACCGTGTGGGTG
AACCGGCCCCCATATGGCCGCGGCGACGACTCGCAGTTCCTGCAACGCCGAAGTGAGAGTCACTCGTGCCGCTTCCACATGGTCGAGGAGCCGGCCGCAGGACGGATGGCACACCCAC
  L  A  G  G  I  P  A  P  L  L  S  V  K  D  V  A  A  S  L  S  V  S  T  A  K  V  V  Y  Q  L  L  A  A  G  V  L  P  T  V  W  V

    1570      1580      1590      1600      1610      1620      1630      1640
GGCCAGTCGCGCCGCGTCAAGCGTGAGGACCTGGACGCCTACATCGCCCGCGCGACGGCCACCGGCGGGAAGCGGGGTGGCAAATGA
CCGGTCAGCGCGGCGCAGTTCGCACTCCTGGACCTGCGGATGTAGCGGGCGCGCTGCCGGTGGCCGCCCTTCGCCCCACCGTTTACT
  G  Q  S  R  R  V  K  R  E  D  L  D  A  Y  I  A  R  A  T  A  T  G  G  K  R  G  G  G  K  *
```

Figure 3

A.

```
      10        20        30        40        50        60        70        80        90       100       110       120
GTGAGCTGACCTCAACGGTTTGTTGGGTGGGAGCGGGACACGCGGACCACATGGTGCCAGGGCTTACGGCTTCGCACACGGGGCTGGCGATGCTGAACGGAGCGTCCCATGTCCACGCG
CACTCGACTGGAGTTGCCAAACAACCCACCCTCGCCCTGTCGCCTGGTGTACCACGGTCCCGAATGCCGAAGCGTGTGCCCCGACCCGCTACGACTTGCCTCGCAGGGTACAGGTGCGC

      130       140       150       160       170       180       190       200       210       220       230       240
ATGCCGCCTGGCTTGCACATAGGGATTCGAAACCTCGACCCCGAGCTTGGGAAGCTCGTGCTCTACCAACTGAGCTACCACCGCAGGCGAAGCAGGGCGCAAAGTACGGGCCGCCCTGTG
TACGGCGGACCGAACGTGTATCCCTAAGCTTTGGAGCTGGGGCTCGAACCCTTCGAGCACGAGATGGTTGACTCGATGGTGGCGTCCGCTTCGTCCCGCGTTTCATGCCCGGCGGGACAC

      250       260       270       280       290       300       310       320       330       340       350       360
GCTTGTCAACGGGAAGTGAGGTGCTACTCCGTCTCCTCGACGGTGAGCTGGTACGAGTCCTGGAAGTTGGACTCGCGGTTGCGCGCGTCCCGGACCTCGAAGAGGTAGACGCCTGGCTCG
CGAACAGTTGCCCTTCACTCCACGATGAGGCAGAGGAGCTGCCACTCGACCATGCTCAGGACCTTCAACCTGAGCGCCAACGCGCGCAGGGCCTGGAGCTTCTCCATCTGCGGACCGAGC
```

B.

```
      10        20        30        40        50        60        70        80        90       100       110       120
CGAGCCGGGGACGGAGCGGCGGGACCGGCTTCGCGCCGTTTACAGCATCCTTGCTGCAAGACGCCCCGAGGCCCCGAAAAGACGAAGGCCGGCAGTCCCGAGTTCCTCAAGGACTACCG
GCTCGGCCCCTGCCCTCGCCGCCTGGCCGAAGCGCGGCAAATGTCGTAGGAACGACGTTCTGCGGGCTCCGGGCTTTTCTGCTTCCGGCCGTCAGGGCTCAAAGGAGTTCCTGATGGC

      130       140       150       160       170       180       190       200       210       220       230       240
GCCTTCATGGGTGAGCGGCGGAAGGGATTCGAACCCTCGACCCCGAGCTTGGGAAGCTCGTGCTCTACCAACTGAGCTACCACCGCAGGCGAAGCAGGGCGCAAAGTACGGGCCGCCCTG
CGGAAGTACCCACTCGCCGCCTTCCCTAAGCTTGGGAGCTCGGGGCTCGAACCCTTCGAGCACGAGATGGTTGACTCGATGGTGGCGTCCGCTTCGTCCCGCGTTTCATGCCCGGCGGGAC

      250       260       270       280       290       300       310       320       330       340       350       360
TGGCTTGTCAACGGGAAGTGAGGTGCTACTCCGTCTCCTCGACGGTGAGCTGGTACGAGTCCTGGAAGTTGGACTCGCGGTTGCGCGCGTCCCGGACCTCGAAGAGGTAGACGCCTGGCT
ACCGAACAGTTGCCCTTCACTCCACGATGAGGCAGAGGAGCTGCCACTCGACCATGCTCAGGACCTTCAACCTGAGCGCCAACGCGCGCAGGGCCTGGAGCTTCTCCATCTGCGGACCGA
```

C.

```
        10        20        30        40        50        60        70        80        90       100       110       120
TGCCAGGGCTTACGGCTTCGCACACGGGGCTGGGCGATGCTGAACGGAGCGTCCCATGTCCACGCGATGCCGCCTGGCTTGCACATAGGGATTCGAAACCTCGACCCCGAGCTTGGGAAG
ACGGTCCCGAATGCCGAAGCGTGTGCCCCGACCCGCTACGACTTGCCTCGCAGGGTACAGGTGCGCTACGGCGGACCGAACGTGTATCCCTAAGCTTTGGAGCTGGGGCTCGAACCCTTC

       130       140       150       160       170       180       190       200       210       220       230       240
CTCGGCCTCGACCCGTCCAGGCGTTATCAGCCGTTCGCAAACCCTTACTTCGCCTTGGGGATTCCGGGCCGGGGGCCTGTCCATCCGTCGCAGCGGTAGCAGGGAGTCTCAGGGGGGTT
GAGCCGGAGCTGGGCAGGTCCGCAATAGTCGGCAAGCGTTTGGGAATGAAGCGGAACCCCTAAGGCCCGGCCCCCGGACAGGTAGGCAGCGTCGCCCATCGTCCCTCAGAGTCCCCCCAA
```

D.

```
        10        20        30        40        50        60        70        80        90       100       110       120
CGCCACCACCTACTCACCAGGGAGGAAGTGACGGCGGCCCTGACTCAGGACGCAACACCTACGGTGAAGGAGGCACCATGCACGGATTGGGAGATTTGCAGCATCACCGGGCGAGAGCTT
GCGGTGGTGGATGAGTGGTCCCTCCTTCACTGCCGCCGCGGACTGAGTCCTGCGTTGTGGATGCCACTTCCTCCGTGGTACGTGCCTAACCCTCTAAACGTCGTAGTGGCCCGCTCTCGAA

       130       140       150       160       170       180       190       200       210       220       230       240
GGGAAGCTCGTGTGCTCTACCAACTGAGCTACCACGCGGAACTTGGCCGGGGGTATACCGGCGCCGCTGCTGAGCGTCAAGGACGTTGCGGCTTCACTCTCAGTGAGCACGGCGAAGGTGT
CCCTTCGAGCACGAGATGGTTGACTCGATGGTGGCGCCTTGAACCGGCCCCCATATGGCCGCGGCGACGACTCGCAGTTCCTGCAACGCCGAAGTGAGAGTCACTCGTGCCGCTTCCACA
                  ←                  ←
       250
ACCAGCTCCTCGCCGCC
TGGTCGAGGAGCGGCGG
```

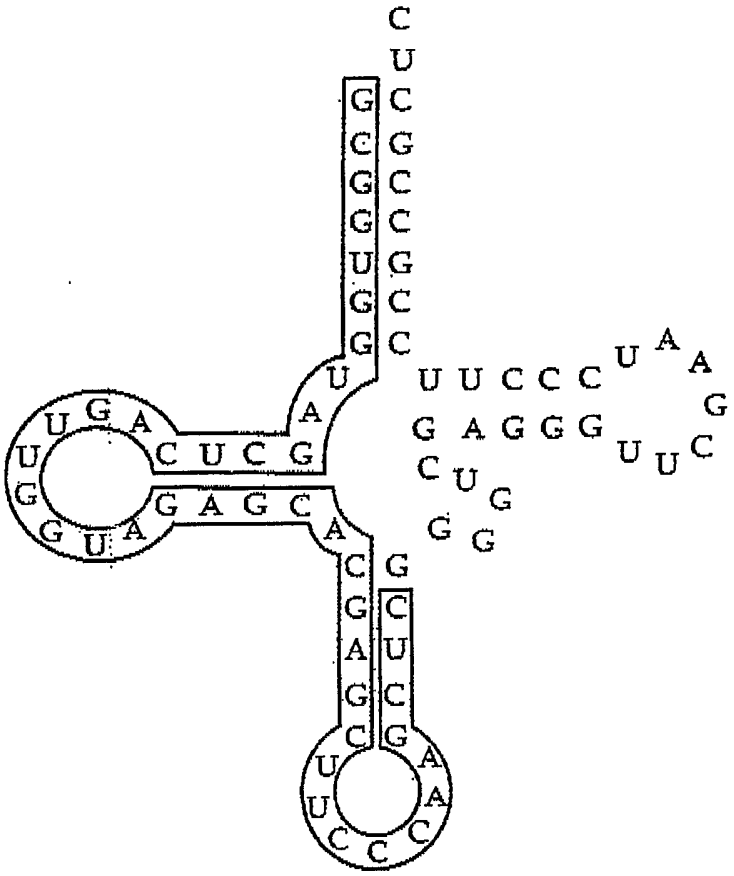

Figure 4

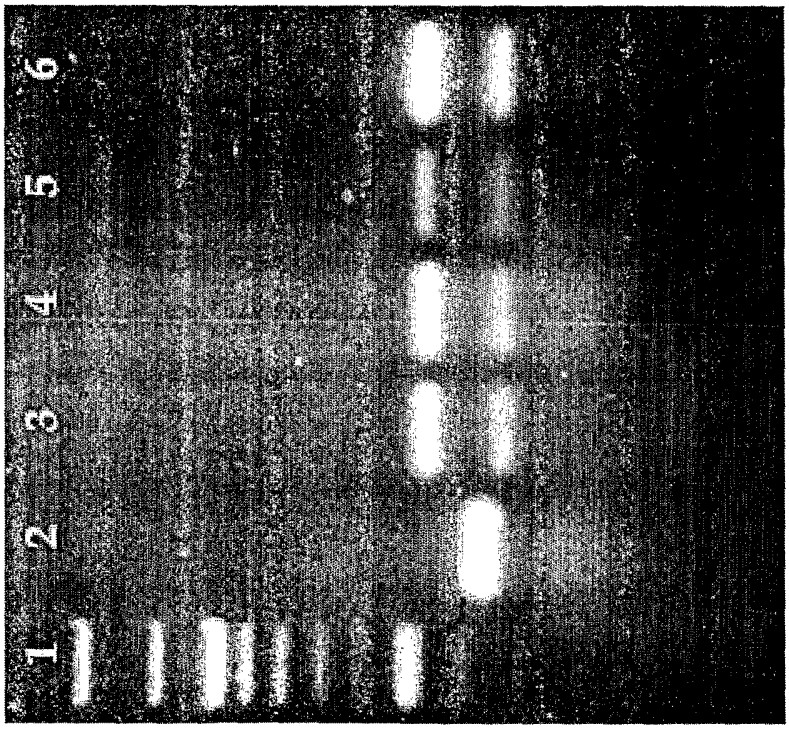

FIGURE 5

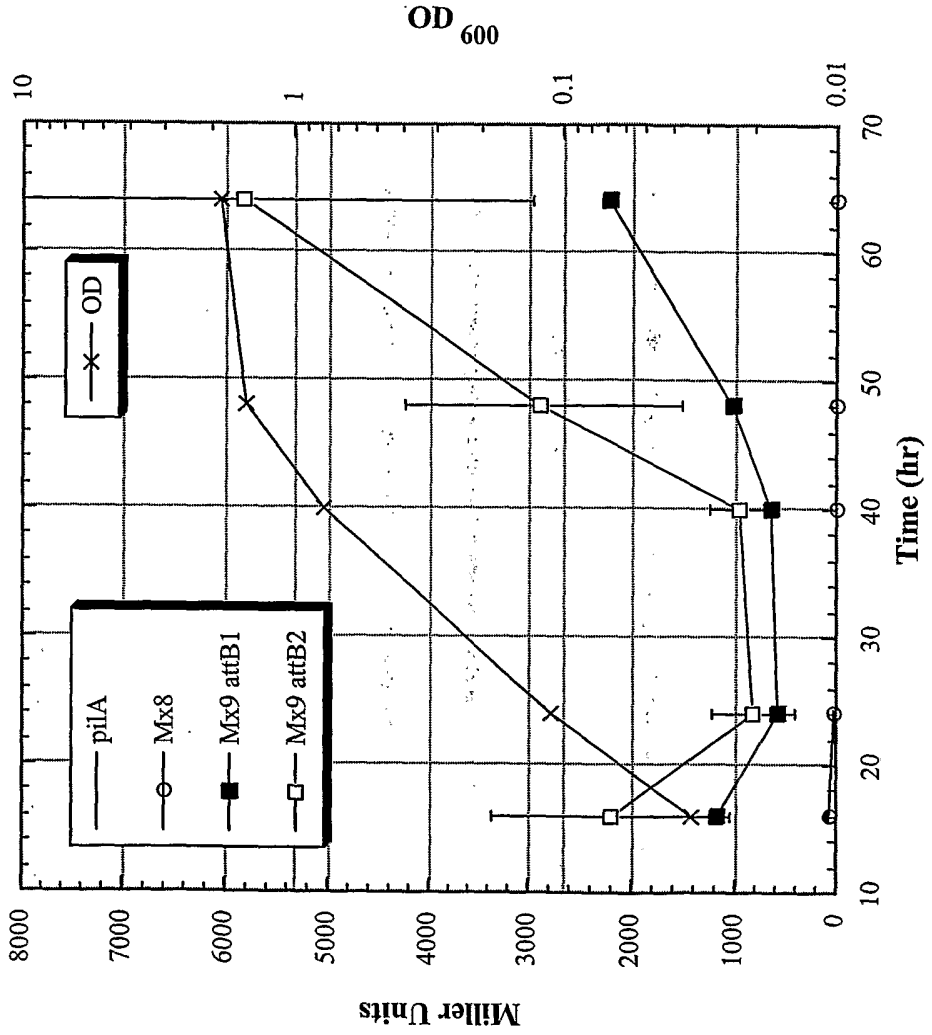
OD 600
10
1
0.1
0.01
OD
pilA
Mx8
Mx9 attB1
Mx9 attB2
Miller Units
8000
7000
6000
5000
4000
3000
2000
1000
0
10
20
30
40
50
60
70
Time (hr)

FIGURE 6A

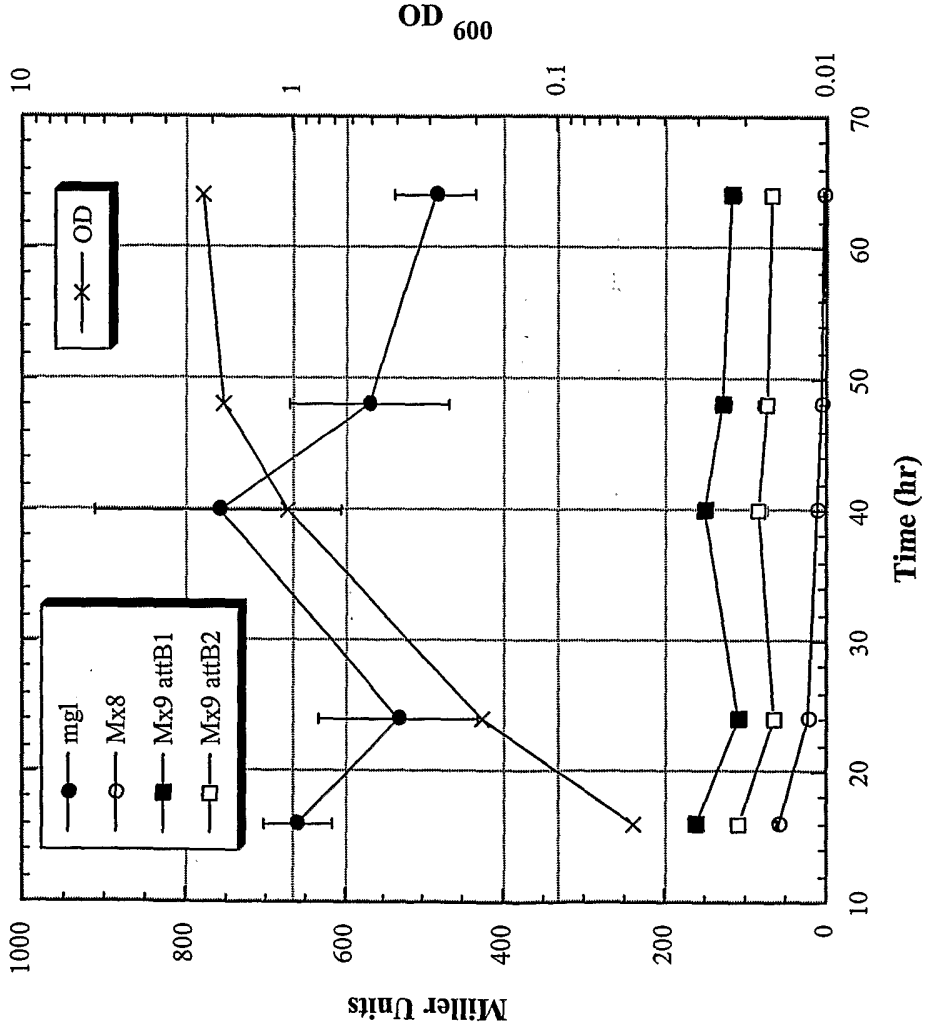
OD 600
10
1
0.1
0.01
OD
mgl
Mx8
Mx9 attB1
Mx9 attB2
Miller Units
1000
800
600
400
200
0
10
20
30
40
50
60
70
Time (hr)

FIGURE 6B

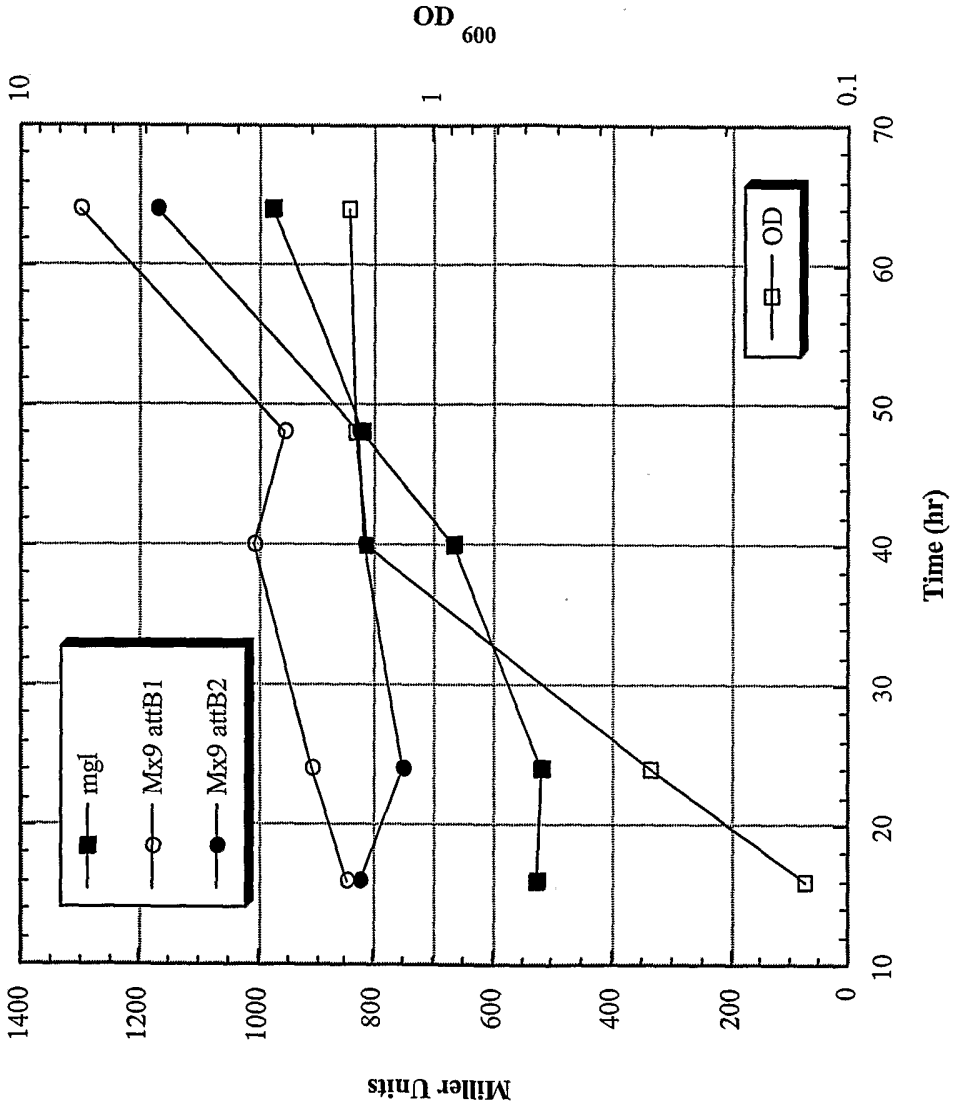
OD 600
10
1
0.1
70
60
50
40
30
20
10
Time (hr)
OD
mgl
Mx9 attB1
Mx9 attB2
1400
1200
1000
800
600
400
200
0
Miller Units

FIGURE 6C

Figure 7

```
    10          20          30          40          50          60
ATG GAA AAA AAG GAA TTT CGT GTT TTG ATA AAA TAC TGT TTT CTG AAG GGA AAA AAT ACA
TAC CTT TTT TTC CTT AAA GCA CAA AAC TAT TTT ATG ACA AAA GAC TTC CCT TTT TTA TGT
Met Glu Lys Lys Glu Asn Arg Val Leu Ile Lys Tyr Cys Asn Leu Lys Gly Lys Asn Thr

          70          80          90          100         110         120
GTG GAA GCA AAA ACT TGG CTT GAT AAT GAG TTT CCG GAC TCT GCC CCA GGG AAA TCA ACA
CAC CTT CGT TTT TGA ACC GAA CTA TTA CTC AAA GGC CTG AGA CGG GGT CCC TTT AGT TGT
Val Glu Ala Lys Thr Trp Leu Asp Asn Glu Asn Pro Asp Ser Ala Pro Gly Lys Ser Thr

          130         140         150         160         170         180
ATA ATT GAT TGG TAT GCA AAA TTC AAG CGT GGT GAA ATG AGC ACG GAG GAC GGT GAA CGC
TAT TAA CTA ACC ATA CGT TTT AAG TTC GCA CCA CTT TAC TCG TGC CTC CTG CCA CTT GCG
Ile Ile Asp Trp Tyr Ala Lys Phe Lys Arg Gly Glu Met Ser Thr Glu Asp Gly Glu Arg

          190         200         210         220         230         240
AGT GGA CGC CCG AAA GAG GTG GTT ACC GAC GAA AAC ATC AAA AAA ATC CAC AAA ATG ATT
TCA CCT GCG GGC TTT CTC CAC CAA TGG CTG CTT TTG TAG TTT TTT TAG GTG TTT TAC TAA
Ser Gly Arg Pro Lys Glu Val Val Thr Asp Glu Asn Ile Lys Lys Ile His Lys Met Ile

          250         260         270         280         290         300
TTG AAT GAC CGT AAA ATG AAG TTG ATC GAG ATA GCA GAG GCC TTA AAG ATA TCA AAG GAA
AAC TTA CTG GCA TTT TAC TTC AAC TAG CTC TAT CGT CTC CGG AAT TTC TAT AGT TTC CTT
Leu Asn Asp Arg Lys Met Lys Leu Ile Glu Ile Ala Glu Ala Leu Lys Ile Ser Lys Glu

          310         320         330         340         350         360
CGT GTT GGT CAT ATC ATT CAT CAA TAT TTG GAT ATG CGG AAG CTC TGT GCA AAA TGG GTG
GCA CAA CCA GTA TAG TAA GTA GTT ATA AAC CTA TAC GCC TTC GAG ACA CGT TTT ACC CAC
Arg Val Gly His Ile Ile His Gln Tyr Leu Asp Met Arg Lys Leu Cys Ala Lys Trp Val

          370         380         390         400         410         420
CCG CGC GAG CTC ACA TTT GAC CAA AAA CAA CAA CGT GTT GAT GAT TCT GAG CGG TGT TTG
GGC GCG CTC GAG TGT AAA CTG GTT TTT GTT GTT GCA CAA CTA CTA AGA CTC GCC ACA AAC
Pro Arg Glu Leu Thr Asn Asp Gln Lys Gln Gln Arg Val Asp Asp Ser Glu Arg Cys Leu

          430         440         450         460         470         480
CAG CTG TTA ACT CGT AAT ACA CCC GAG TTT TTC CGT CGA TAT GTG ACA ATG GAT GAA ACA
GTC GAC AAT TGA GCA TTA TGT GGG CTC AAA AAG GCA GCT ATA CAC TGT TAC CTA CTT TGT
Gln Leu Leu Thr Arg Asn Thr Pro Glu Asn Phe Arg Arg Tyr Val Thr Met Asp Glu Thr

          490         500         510         520         530         540
TGG CTC CAT CAC TAC ACT CCT GAG TCC AAT CGA CAG TCG GCT GAG TGG ACA GCG ACC GGT
ACC GAG GTA GTG ATG TGA GGA CTC AGG TTA GCT GTC AGC CGA CTC ACC TGT CGC TGG CCA
Trp Leu His His Tyr Thr Pro Glu Ser Asn Arg Gln Ser Ala Glu Trp Thr Ala Thr Gly

          550         560         570         580         590         600
GAA CCG TCT CCG AAG CGT GGA AAG ACT CAA AAG TCC GCT GGC AAA GTA ATG GCC TCT GTT
CTT GGC AGA GGC TTC GCA CCT TTC TGA GTT TTC AGG CGA CCG TTT CAT TAC CGG AGA CAA
Glu Pro Ser Pro Lys Arg Gly Lys Thr Gln Lys Ser Ala Gly Lys Val Met Ala Ser Val
```

```
          610         620         630         640         650         660
TTT TTC GAT GCG CAT GGA ATA ATT TTT ATC GAT TAT CTT GAG AAG GGA AAA ACC ATC AAC
AAA AAG CTA CGC GTA CCT TAT TAA AAA TAG CTA ATA GAA CTC TTC CCT TTT TGG TAG TTG
Asn Phe Asp Ala His Gly Ile Ile Asn Ile Asp Tyr Leu Glu Lys Gly Lys Thr Ile Asn

          670         680         690         700         710         720
AGT GAC TAT TAT ATG GCG TTA TTG GAG CGT TTG AAG GTC GAA ATC GCG GCA AAA CGG CCC
TCA CTG ATA ATA TAC CGC AAT AAC CTC GCA AAC TTC CAG CTT TAG CGC CGT TTT GCC GGG
Ser Asp Tyr Tyr Met Ala Leu Leu Glu Arg Leu Lys Val Glu Ile Ala Ala Lys Arg Pro

          730         740         750         760         770         780
CAT ATG AAG AAG AAA AAA GTG TTG TTC CAC CAA GAC AAC GCA CCG TGC CAC AAG TCA TTG
GTA TAC TTC TTC TTT TTT CAC AAC AAG GTG GTT CTG TTG CGT GGC ACG GTG TTC AGT AAC
His Met Lys Lys Lys Lys Val Leu Phe His Gln Asp Asn Ala Pro Cys His Lys Ser Leu

          790         800         810         820         830         840
AGA ACG ATG GCA AAA ATT CAT GAA TTG GGC TTC GAA TTG CTT CCC CAC CCA CCG TAT TCT
TCT TGC TAC CGT TTT TAA GTA CTT AAC CCG AAG CTT AAC GAA GGG GTG GGT GGC ATA AGA
Arg Thr Met Ala Lys Ile His Glu Leu Gly Phe Glu Leu Leu Pro His Pro Pro Tyr Ser

          850         860         870         880         890         900
CCA GAT CTG GCC CCC AGC GAC TTT TTC TTG TTC TCA GAC CTC AAA AGG ATG CTC GCA GGG
GGT CTA GAC CGG GGG TCG CTG AAA AAG AAC AAG AGT CTG GAG TTT TCC TAC GAG CGT CCC
Pro Asp Leu Ala Pro Ser Asp Asn Phe Leu Phe Ser Asp Leu Lys Arg Met Leu Ala Gly

          910         920         930         940         950         960
AAA AAA TTT GGC TGC AAT GAA GAG GTG ATC GCC GAA ACT GAG GCC TAT TTT GAG GCA AAA
TTT TTT AAA CCG ACG TTA CTT CTC CAC TAG CGG CTT TGA CTC CGG ATA AAA CTC CGT TTT
Lys Lys Asn Gly Cys Asn Glu Glu Val Ile Ala Glu Thr Glu Ala Tyr Asn Glu Ala Lys

          970         980         990        1000        1010        1020
CCG AAG GAG TAC TAC CAA AAT GGT ATC AAA AAA TTG GAA GGT CGT TAT AAT CGT TGT ATC
GGC TTC CTC ATG ATG GTT TTA CCA TAG TTT TTT AAC CTT CCA GCA ATA TTA GCA ACA TAG
Pro Lys Glu Tyr Tyr Gln Asn Gly Ile Lys Lys Leu Glu Gly Arg Tyr Asn Arg Cys Ile

         1030        1040
GCT CTT GAA GGG AAC TAT GTT GAA TAA
CGA GAA CTT CCC TTG ATA CAA CTT ATT
Ala Leu Glu Gly Asn Tyr Val Glu ***
```