PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/52, 15/62, 9/02, 9/04, 9/10, 9/88, C12P 17/06, 19/62, C07K 19/00

A2

(11) International Publication Number: **WO 00/01827**

(43) International Publication Date: 13 January 2000 (13.01.00)

(21) International Application Number: PCT/GB99/02158

(22) International Filing Date: 6 July 1999 (06.07.99)

(30) Priority Data:
9814622.8 6 July 1998 (06.07.98) GB

(71) Applicants *(for all designated States except US)*: BIOTICA TECHNOLOGY LIMITED [GB/GB]; 112 Hills Road, Cambridge CB2 1PH (GB). PFIZER INC. [US/US]; 235 East 42nd Street, New York, NY 10017 (US).

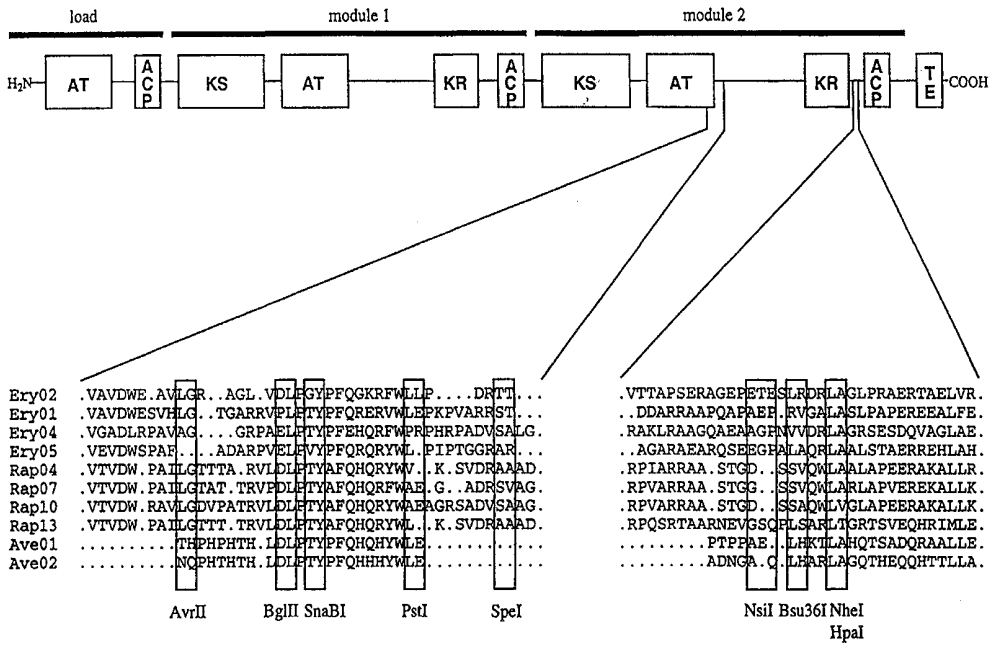
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: KELLENBERGER, Johannes, Laurenz [CH/CH]; Hohlweg 7, CH–4125 Riehen (CH). LEADLAY, Peter, Francis [GB/GB]; 17 Clarendon Road, Cambridge CB2 2BH (GB). STAUNTON, James [GB/GB]; 29 Porson Road, Cambridge CB2 2ET (GB). STUTZMAN–ENGWALL, Kim, Jonelle [US/US]; 547 Boston Post Road, East Lyme, CT 06333 (US). McARTHUR, Hamish, Alastair, Irvine [GB/US]; 19 Pheasant Run Drive, Gales Ferry, CT 06335 (US).

(74) Agents: STUART, Ian et al.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, 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, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: POLYKETIDES, THEIR PREPARATION, AND MATERIALS FOR USE THEREIN



(57) Abstract

Nucleic acid molecules encoding at least part of a Type I polyketide synthase, and having a polylinker with multiple restriction enzyme sites in place of one or more PKS genes encoding enzymes associated with reduction, optionally further including nucleic acid incorporated into the polylinker, the further nucleic acid encoding one or more reductive enzymes; plasmids incorporating such nucleic acids; host cells transfected with such plasmids; methods relating thereto.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

POLYKETIDES, THEIR PREPARATION, AND MATERIALS FOR USE THEREIN

The present invention relates to polyketides, their preparation, and materials for use therein.

Polyketides are a large and structurally diverse class of natural products that includes many compounds possessing antibiotic or other pharmacological properties, such as erythromycin, tetracyclines, rapamycin, avermectin, polyether ionophores, and FK506.

In particular, polyketides are abundantly produced by Streptomyces and related actinomycete bacteria. They are synthesised by the repeated stepwise condensation of acylthioesters in a manner analogous to that of fatty acid biosynthesis. The greater structural diversity found among natural polyketides arises from the selection of (usually) acetate or propionate as "starter" or "extender" units; and from the differing degree of processing of the β-keto group observed after each condensation. Examples of processing steps include reduction to β-hydroxyacyl-, reduction followed by dehydration to 2-enoyl-, and complete reduction to the saturated acylthioester. The stereochemical outcome of these processing steps is also specified for each cycle of chain extension.

The biosynthesis of polyketides is initiated by a group of chain-forming enzymes known as polyketide synthases (PKSs). Two classes of polyketide synthase have been described in actinomycetes. However the novel

polyketides and processes which are the subject of the present invention relate mainly to Type I PKSs, represented by the PKSs for the macrolides erythromycin, rapamycin and avermectin. Type I PKSs contain a different set or "module" of enzymes for each cycle of polyketide chain extension (Cortes, J. et al. Nature (1990) 348:176-178; Donadio, S. et al. Science (1991) 252:675-679; MacNeil, D. J. et al. Gene (1991) 115:119-125; Schwecke, T. et al. Proc.Natl. Acad. Sci. USA (1995) 92:7839-7843 and see e.g. Figure 1 herein, or Figures 2a and 3 of WO98/01546); whereas Type II PKSs are represented by the synthases for aromatic compounds and contain only a single set of enzymatic activities for chain extension. These are re-used as appropriate in successive cycles.

A complete module dictating full reduction contains a ketoacyl-ACP synthase (KS) domain; an acyl carrier protein domain (ACP); an acyl-CoA:ACP acyltransferase (AT) for loading of the extender unit; and a ketoreductase (KR), a dehydratase (DH) and an enoylreductase (ER) domain for accomplishment of the processing of the β-keto group. Since these domains have enzymic activity, they may also be referred to herein as "enzymes", though this is not intended to imply anything about their structural relationship to other PKS domains. Similarly, the nucleic acid sequences encoding such domains may also be referred to as "genes", though this is not intended to imply anything about the presence or otherwise of separate regulatory regions for the different domains of a PKS.

The present invention particularly relates to processes for preparing polyketides by replacing the reductive loop (the segment from the end of the AT to the beginning of the ACP comprising either a KR or a KR and a DH or a KR, a DH and a ER) in a selected module of a Type I polyketide synthase gene cluster by the equivalent segment from the same or from a different PKS gene cluster, or by a mutated or synthetic segment, thereby generating new hybrid polyketide synthases that produce polyketides with different extent of reduction and/or stereochemistry in a predictable way.

For the avoidance of doubt, the term "extension module", as used hereinafter, refers to a set of domains of a Type I PKS, each having enzymic activity, which participate in one cycle of polyketide chain extension. More particularly, an extension module comprises KS, AT, a reductive loop (comprising one or more of KR, DH and ER), and ACP.

Rarely, the reductive loop may include other domains. For example yersiniabacter, which possesses a mixed PKS and polypeptide synthase, possesses a methyl transferase domain.

It has been reported that replacement of the reductive loop of module 2 in DEBS1TE with the equivalent segment of module 3 of the (Type I) erythromycin PKS gene yields a triketide ketolactone when expressed in S. coelicolor CH999 (Bedford, D. et al.Chemistry and Biology (1996) 3:827-831).

Similarly, replacement of the reductive loop of module 2 in DEBS1TE with the equivalent segment of module 5 of the erythromycin PKS yields a triketide lactone with the predicted structure and stereochemistry when expressed in S. coelicolor CH999 (McDaniel, R. et al.Chemistry and Biology (1997) 4:667-674). On the contrary, when the same experiment was carried out using the reductive loop of module 6 of the erythromycin PKS only a ketolactone could be isolated (McDaniel, R. et al.Chemistry and Biology (1997) 4:667-674).

In a further experiment it has been shown, that the reductive loop of module 2 in a trimodular system comprising the loading domain, the first, second and third extension module and the TE of the ery gene can also be substituted by the equivalent segment of module 4 of the rapamycin PKS comprising a KR and DH domain yielding a tetraketide with the predicted double bond when expressed in S. coelicolor CH999 (McDaniel, R. et al. J. Am. Chem. Soc. (1997) 119:4309-4310). In the same system the reductive loop of module 2 has been replaced by the equivalent segment of module 1 of the rapamycin PKS comprising a KR a DH and a ER domain yielding a tetraketide with the predicted oxidation level at C-5 when expressed in S. coelicolor CH999 (Kao, C. M. et al. J. Am. Chem. Soc. (1997) 119:11339-11340). On the contrary, when using the corresponding segment of module 4 of the erythromycin PKS gene only a polyketide with a double bond at the relevant position could be observed and not, as one would predict, full reduction (Kao, C. M. et al. J. Am. Chem. Soc. (1997) 119:11339-11340).

In two similar experiments the reductive loop of module 2 in the trimodular system has been substituted by the corresponding segment of module 2 of the rapamycin PKS containing a KR and an inactive DH domain and by the KR domain of module 4 of the rap PKS (the reductive loop of rap module 4 contains a KR and a DH domain). Both constructs are reported to yield a triketide lactone with a different stereochemistry at C-3 (Kao, C. M. et al. J. Am. Chem. Soc. (1998) 120:2478-2479).

In all the examples described above the same restriction sites, PstI and XbaI, have been used to join the DNA fragments (the location of the PstI site is identical to the PstI site used in the system described below and the XbaI site is in the same place as the Bsu36I site).

A model has been proposed for the structure of the DEB synthase, where the reductive domains form a loop which lies outside the core formed by the KS, AT and the ACP domains (Staunton et al. Nature structural biology (1996) 3:188-192). In addition it has been found that DEBS1 is hydrolysed by proteolytic enzymes at specific locations which mark the boundaries of the domains (Aparicio, J. F. et al. J. Biol. Chem. (1994) 269: 8524-8528). These proteolytic sites are found mainly in linker regions and it seems therefore ideal to join the fragments in close neighbourhood to these sites. Examples of this are documented in WO98/01546.

In one aspect the invention provides nucleic acid (particularly DNA) encoding at least part of a Type I polyketide synthase (PKS), said part comprising at least

part of an extension module, wherein the nucleic acid has a polylinker with multiple restriction enzyme sites in place of one or more genes encoding enzymes associated with reduction.

In another aspect the invention provides nucleic acid (particularly DNA) encoding at least part of a Type I polyketide synthase, said part comprising at least part of an extension module, wherein the nucleic acid has a polylinker with multiple restriction enzyme sites which connects nucleic acid encoding (at least part of) AT to nucleic acid encoding (at least part of) ACP.

Such nucleic acids may have an additional nucleic acid, which encodes one or more reductive enzymes, inserted into the polylinker as described in more detail below. Such insertion is preferably performed following digestion of the polylinker-containing nucleic acids by two restriction enzymes. In order to provide a choice of insertion sites, the polylinker preferably includes at least three restriction sites, more preferably at least four, and further preferably at least six or eight restriction sites.

The polylinker may be provided by introducing exogenous (usually synthetic) nucleic acid into the Type I PKS-encoding nucleic acid, or may be provided by engineering the existing sequence of the Type I PKS-encoding nucleic acid. For example, to achieve the latter, restriction sites may be engineered (e.g. by site-directed mutagenesis) into sequences up- and/or downstream (preferably both) of where the absent reductive enzyme-

encoding sequence would normally lie, particularly into sequences which encode polypeptide linkers between the reductive enzyme(s) and adjacent domains.

The polylinker desirably includes at least some of the following restriction sites: AvrII, BglII; SnaBI; PstI; SpeI; NsiI; Bsu36l; NheI; and HpaI. More desirably the polylinker includes at least four of these sites.

Preferably at least some of the restriction sites included in the polylinker are absent from the remainder of the nucleic acid into which it is incorporated. Desirably at least some of the sites included in the polylinker are uncommon in or absent from naturally occurring nucleic acid sequences which encode reductive enzymes of other (preferably Type I) PKSs. Desirably at least two of the sites are absent from at least about half, more desirably at least about three quarters, of known nucleic acid sequences encoding reductive enzymes of PKSs. Preferably the restriction sites are rich in A and T residues, since PKS genes tend to be rich in G and C residues.

Desirably the nucleic acids of the invention encode a loading module and/or one or more extension modules. More detail concerning varieties of loading modules may be found in our copending international patent application, entitled "Polyketides and their synthesis", filed 29 June 1999.

In another aspect the invention provides nucleic acid generally as indicated above but having further nucleic

acid encoding one or more reductive enzymes (e.g. KR and/or DH and/or ER) inserted into the polylinker. The inserted nucleic acid may encode one or more reductive enzymes of the same polyketide synthase as that of the nucleic acid into which the polylinker is inserted, but from a different extension module. Alternatively the inserted nucleic acid may be exogenous, encoding one or more reductive enzymes from a different natural PKS or fatty acid synthase, or may be synthetic or may be mutated from a naturally occurring nucleic acid which encodes one or more reductive enzymes of a PKS or fatty acid synthase. Preferably, the inserted nucleic acid encodes one or more reductive enzymes from the same or another Type I PKS or fatty acid synthase, but alternatively it may encode one or more reductive enzymes from a Type II PKS or fatty acid synthase.

The genes encoding numerous examples of Type I PKSs have been sequenced and these sequences disclosed in publicly available DNA and protein sequence databases including Genbank, EMBL, and Swissprot. For example the sequences are available for the PKSs governing the synthesis of, respectively, erythromycin (Cortes, J. et al. Nature (1990) 348:176-178; accession number X62569, Donadio, S. et al. Science (1991) 252:675-679; accession number M63677); rapamycin (Schwecke, T. et al. Proc.Natl. Acad. Sci. USA (1995) 92:7839-7843; accession number X86780); rifamycin (August et al. (1998); accession number AF040570); and tylosin (Eli Lily, accession number U78289), among others. Furthermore, figure 7 herein shows the nucleic acid sequence encoding the first two modules of the avermectin PKS from S. avermitilis; this

may be used as an alternative source for the inserts used in certain of the examples.

It is apparent to those skilled in the art that the overall sequence similarity between the nucleic acids encoding comparable domains or modules of different Type I PKSs is sufficiently high, and the domain organisation of different Type I PKSs so consistent between different polyketide-producing microorganisms, that the processes for obtaining novel hybrid polyketides described in the present invention will be generally applicable to all natural modular Type I PKSs or their derivatives.

In further aspects, the present invention provides vectors, such as plasmids or phages (preferably plasmids), including nucleic acids as defined in the above aspects and host cells (particularly of *Streptomyces* species) transfected with such nucleic acids or constructs.

In a still further aspect, the present invention provides polyketide synthases expressible by host cells as defined above. Such polyketide synthases may if desired be isolated from the host cells by routine methods, though it is usually preferable not to do so.

In further aspects the invention provides methods of creating novel functional PKS's and nucleic acids encoding them by means of insertion of nucleic acid encoding reductive enzymes into polylinkers as indicated above; and novel polyketides as produced by such PKS's.

In a still further aspect, the present invention provides novel processes for the specific or preferential production of particular polyketides, using the materials and methods as defined in previous aspects. For example, the present invention provides processes for the generation by direct fermentation of C22-C23 dihydroavermectins, such as ivermectin (see e.g. Examples 25 and 26), and of B1 avermectins substantially free of B2 avermectins (see e.g. Examples 27 and 28).

In another aspect, the present invention provides novel polyketides and novel stereoisomers of polyketides, such as particular polyketides produced in accordance with one or more of the Examples.

In order to enable the exchange of the reductive loop in module 2 of the erythromycin PKS gene in the DEBS1TE system (Cortes J. et al. (1995) 268:1487-1489) a polylinker (multiple cloning site (mcs)) has been inserted in place of the reductive loop of module 2 thereby generating a minimal module comprising a KS, an AT and an ACP. (This system is still functional and produces a ketolactone (see examples 2 and 4).) The mcs contains unique recognition sites for 9 restriction enzymes.

These new restriction sites are situated partly in DNA encoding a linker region near positions where the polyketide synthase is hydrolysed by proteolytic enzymes (vide supra). While some of the restriction sites lie in DNA encoding regions of low homology, others are situated in DNA encoding highly conserved regions (Figure 1). The

introduction of recognition sites for the enzymes AvrII, BglII, Bsu36I and NheI does not change the amino acid sequence in DEBS module 2. In the other five cases (SnaBI, PstI, SpeI, Nsi, HpaI) the amino acid sequence is changed (Figure 2). These changes do not affect the activity of the protein (see example 6).

Because two of the restriction sites cover the same bases it was decided to construct two plasmids containing different mcs (pJLK114 and pJLK117).

The use of an mcs offers the following advantages over a single restriction site on each side of the reductive loop:

1) suitable positions to join the DNA fragments (20 different combinations) can be chosen for every different reductive loop thereby avoiding unfavourable changes in the amino acid sequence

2) enzymes that cut within the loop can be avoided; and

3) loop insertion may be performed in a combinatorial way.

The present inventors have made the further surprising discovery that different results may be obtained using the same polylinker-containing nucleic acid and the same nucleic acid encoding one or more reductive enzymes, when the nucleic acid encoding one or more reductive enzymes is incorporated at different sites in the polylinker.

For example, in Examples 7 and 8, the reductive loop of the rapamycin module 13 was inserted into ery module 2 to bring about complete reduction of the polyketide chain as the outcome of the second extension module. The desired triketide lactone products were obtained in good yield. However, in Examples 37 and 38, the same reductive loop, or set of domains, from rap module 13 was inserted into essentially the same position in ery module 2 as in examples 7 and 8, save that different restriction sites of the polylinker were used (AvrII and HpaI instead of BglII and NsiI) and significant amounts of by-products were obtained. Such by-products included triketide lactones in which C-3 was either keto or hydroxy, showing that simply altering the sites used for swapping the reductive loop made the difference between obtaining the desired product and obtaining an undesirable mixture of the desired product with the products of incomplete reduction.

Similarly, in Examples 31 and 32, when the sites PstI and Bsu36I were used to insert the reductive domains of avermectin module 1 (plasmid pGMS2) in place of the reductive loop of ery module 2, the expected product was produced, but also a substantial amount of ketolactone. In the experiment of Examples 29 and 30, when the sites BglII and NheI were used (plasmid pJLK30) hardly any ketolactone byproduct was produced, although the amounts of lactone were in a similar range in each case.

When, entirely analogously to the Examples 29 and 30, in Example 14 the same sites BglII and NheI were used to replace the reductive loop of ery module 2 with the reductive loop of tylosin module 1 (plasmid pJLK35), the

same target triketide lactones were produced as in Examples 30 and 32 but with much higher yield, albeit accompanied by some ketolactone, demonstrating that different reductive loops may be most advantageously inserted into different restriction sites.

In Examples 33 and 34, when the sites BglII and NheI were used to insert the reductive domains of avermectin module 2 (plasmid pJLK31) the expected products were produced as the major products. In the experiment of Examples 35 and 36, when the sites SnaBI and Bsu36I were used (plasmid pGMS4) only trace amounts of a triketide lactone mixture could be obtained.

Thus, the present invention provides the opportunity, should the desired and predicted products not be obtained when a particular reductive loop is inserted into a particular PKS, of simple adjustment of the insertion site by use of different restriction enzymes having sites in the polylinker. As demonstrated by the above comparative examples, such readjustment can dramatically affect the outcome and yield of polyketide synthesis.

Example 1

Construction of plasmid pJLK114

Plasmid pJLK114 is a pCJR24 based plasmid containing a PKS gene comprising the ery loading module, the first and the second extension modules of the ery PKS and the ery chain-terminating thioesterase except that the DNA segment between the end of the acyltransferase and the beginning of the ACP of the second ery extension module has been substituted by a synthetic oligonucleotide

linker containing the recognition sites of the following restriction enzymes: AvrII, BglII, SnaBI, PstI, SpeI, NsiI, Bsu36I and HpaI. It was constructed via several intermediate plasmids as follows (Figure 3).

Construction of plasmid pJLK02

The approximately 1.47 kbp DNA fragment of the eryAI gene of S. erythraea was amplified by PCR using as primers the synthetic oligonucleotides:
5'-TACCTAGGCCGGGCCGGACTGGTCGACCTGCCGGGTT-3' and
5'-ATGTTAACCGGTCGCGCAGGCTCTCCGTCT-3' and plasmid pNTEP2 (Oliynyk, M. et al., Chemistry and Biology (1996) 3:833-839; WO98/01546) as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK02 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK03

The approximately 1.12 kbp DNA fragment of the eryAI gene of S. erythraea was amplified by PCR using as primers the synthetic oligonucleotides:
5'-ATGTTAACGGGTCTGCCGCGTGCCGAGCGGAC-3' and
5'-CTTCTAGACTATGAATTCCCTCCGCCCAGC-3' and plasmid pNTEPH as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli

DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK03 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK04

Plasmid pJLK02 was digested with PstI and HpaI and the 1.47 kbp insert was ligated with plasmid pJLK03 which had been digested with PstI and HpaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK04 was identified by its restriction pattern.

Construction of plasmid pJLK05

Plasmid pJLK01 (PCT/GB97/01819) was digested with PstI and AvrII and the 460 bp insert was ligated with plasmid pJLK04 which had been digested with PstI and AvrII. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK05 was identified by its restriction pattern.

Construction of plasmid pJLK07

Plasmid pJLK05 was digested with ScaI and XbaI and plasmid pNTEP2 was digested with NdeI and ScaI and these two fragments were ligated with plasmid pCJR24 which had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK07 was identified by its restriction pattern.

Construction of plasmid pJLK114

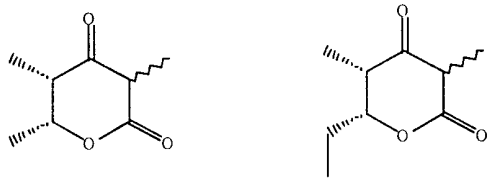
The two synthetic oligonucleotides Plf and Plb (Figure 4) were each dissolved in TE-buffer. 10 μl of each solution (0.5nmol/μl) were mixed and heated for 2 minutes to 65C and then slowly cooled down to room temperature. Plasmid pJLK07 was digested with AvrII and HpaI and ligated with the annealed oligonucleotides. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK114 was identified by its restriction pattern.

Example 2

Use of plasmid pJLK114 for construction of S. erythraea JC2/pJLK114 and the production of TKL derivatives

Approximately 5 μg plasmid pJLK114 were used to transform protoplasts of S. erythraea JC2 (strain deposited as No. NCIMB 40802. WO98/01546.) and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE gene. JC2/pJLK114 was plated onto SM3 agar (5.0 g glucose, 50.0 g MD30E maltodextrin, 25.0 g Arkasoy soya flour, 3.0 g molasses (beet), 0.25 g K_2HPO_4, 2.5 g $CaCO_3$ 22.0 g agar distilled water to 1 litre pH=7.0) containing 50 μg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1 cm^2 (500μl) of the plate was homogenised and extracted with a mixture of 1.2 ml ethyl acetate and 20 ul formic acid. The solvent was decanted and removed by evaporation and the residue dissolved in

methanol and analysed by GC/MS and electrospray mass spectroscopy. The major products were identified as (4S, 5R)-5-hydroxy-2,4-dimethyl-3-oxo-n-hexanoic acid-δ-lactone and as (4S, 5R)-5-hydroxy-2,4-dimethyl-3-oxo-n-heptanoic acid-δ-lactone.



Example 3

Construction of plasmid pJLK117

Plasmid pJLK117 is a pCJR24 based plasmid containing a PKS gene comprising the ery loading module, the first and the second extension modules of the ery PKS and the ery chain-terminating thioesterase except that the DNA segment between the end of the acyltransferase and the beginning of the ACP of the second ery extension module has been substituted by a synthetic oligonucleotide linker containing the recognition sites of the following restriction enzymes. AvrII, BglII, SnaBI, PstI, SpeI, NsiI, Bsu36I and NheI.

It was constructed via several intermediate plasmids as follows (Figure 3).

Construction of plasmid pJLK115

Plasmid pJLK114 was digested with NdeI and XbaI and the

approximately 9.9 kbp insert was ligated with plasmid pUC18 which had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK115 was identified by its restriction pattern.

Construction of plasmid pJLK116

Plasmid pJLK13 (PCT/GB97/01819) was digested with Bsu36I and XbaI and the 1.1 kbp fragment was ligated with plasmid pJLK115 which had been digested with Bsu36I and XbaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK116 was identified by its restriction pattern.

Construction of plasmid pJLK117

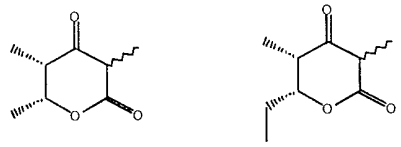
Plasmid pJLK116 was digested with NdeI and XbaI and the 9.9 kbp fragment was ligated with plasmid pCJR24 which had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK117 was identified by its restriction pattern.

Example 4

Use of plasmid pJLK117 for construction of S. erythraea JC2/pJLK117 and the production of TKL derivatives

Approximately 5 μg plasmid pJLK117 were used to transform

protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE. JC2/pJLK117 was plated onto SM3 agar containing 50 μg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1 cm^2 (0.5 ml) of the plate was homogenised and extracted with a mixture of 1.2 ml ethyl acetate and 20 ul formic acid. The solvent was decanted and removed by evaporation and the residue dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. The major products were identified as (4S, 5R)-5-hydroxy-2,4-dimethyl-3-oxo-n-hexanoic acid-δ-lactone and as (4S, 5R)-5-hydroxy-2,4-dimethyl-3-oxo-n-heptanoic acid-δ-lactone.



Example 5

Construction of plasmid pJLK25

Plasmid pJLK25 is a pJLK114 based plasmid except that the DNA fragment encoding the reductive loop of the second module of the erythromycin PKS gene has been inserted into the mcs.

It was constructed via several intermediate plasmids as follows.

Construction of plasmid pJLK118

The approximately 1.4 kbp DNA fragment of the eryAI gene of S. erythraea encoding the reductive loop of module 2 was amplified by PCR using as primers the synthetic oligonucleotides:
5'-ATACTAGTCCTCGTGACGAGCTCGACGG-3' and
5'-TAATGCATCCGGTTCTCCGGCCCGCTCGCT-3' and pNTEP2 as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK118 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK23

Plasmid pJLK118 was digested with SpeI and NsiI and the 1.4 kbp fragment was ligated with plasmid pJLK115 which had been digested with SpeI and NsiI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK23 was identified by its restriction pattern.

Construction of plasmid pJLK25

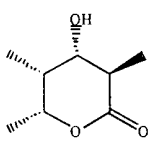
Plasmid pJLK23 was digested with NdeI and XbaI and the approximately 11.2 kbp fragment was ligated with plasmid pCJR24 which had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked

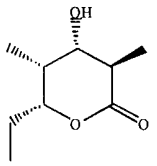
for their plasmid content. The desired plasmid pJLK25 was identified by its restriction pattern.

Example 6

Use of plasmid pJLK25 for construction of S. erythraea JC2/pJLK25 and the production of triketides

Approximately 5 μg plasmid pJLK25 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE. JC2/pJLK25 was plated onto SM3 agar containing 50 μg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1 cm^2 (0.5 ml) of the plate was homogenised and extracted with a mixture of 1.2 ml ethyl acetate and 20 ul formic acid. The solvent was decanted and removed by evaporation and the residue dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. The major products were identified (by comparison with authentic material) as (2R, 3S, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-hexanoic acid δ-lactone and as (2R, 3S, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-heptanoic acid δ-lactone.





Example 7

Construction of plasmid pJLK28

Plasmid pJLK28 is a pJLK117 based plasmid except that the DNA fragment encoding the reductive loop of module 13 of the rap PKS has been inserted into the mcs. It was constructed via several intermediate plasmids as follows. (Figure 5)

Construction of plasmid pJLK120

The approximately 3.2 kbp DNA segment of the rapC gene of S. hygroscopicus encoding the reductive loop of module 13 was amplified by PCR using as primers the synthetic oligonucleotides:
5'-TAAGATCTTCCGACCTACGCCTTCCAAC-3' and
5'-TAATGCATCGACCTCGTTGCGTGCCGCGGT-3' and cosmid cos 31 (Schwecke, T. et al. (1995) Proc. Natl. Acad. Sci. USA 92:7839-7843) as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK120 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK28

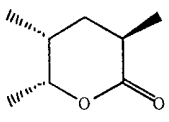
Plasmid pJLK120 was digested with BglII and NsiI and the 3.2 kbp fragment was ligated with plasmid pJLK117 which had been digested with BglII and NsiI. The ligation

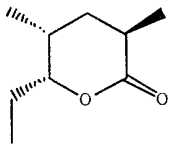
mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK28 was identified by its restriction pattern.

Example 8

Use of plasmid pJLK28 for construction of JC2/pJLK28 and the production of triketides

Approximately 5 μg plasmid pJLK28 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE. JC2/pJLK28 was plated onto SM3 agar containing 50 μg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1 cm^2 (0.5 ml) of the plate was homogenised and extracted with a mixture of 1.2 ml ethyl acetate and 20 ul formic acid. The solvent was decanted and removed by evaporation and the residue dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. The major products were identified (by comparison with authentic material) as (2R, 4S, 5R)-2,4-dimethyl-5-hydroxy-n-hexanoic acid δ-lactone and as (2R, 4S, 5R)-2,4-dimethyl-5-hydroxy-n-heptanoic acid δ-lactone.





Example 9

Construction of plasmid pJLK41

Plasmid pJLK41 is a pJLK117 based plasmid except that the DNA fragment encoding the reductive loop of module 4 of the ery PKS has been inserted into the mcs. It was constructed via several intermediate plasmids as follows. (Figure 5)

Construction of plasmid pJLK32.3

The approximately 3.2 kbp DNA segment of the eryAII gene of S. erythraea encoding the reductive loop of module 4 was amplified by PCR using as primers the synthetic oligonucleotides:
5'-ATAGATCTGCCTACGTACCCGTTCGAACACCAGCGCTTC-3' and
5'-ATCCTCAGGTTCGGCCCTGCCGCCTCGGCCTGCCCGGCGGCGCGCAGCTT-3'
and cosmid cos4B (cosmid containing the erythromycin PKS) as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK32.3 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK38

Plasmid pJLK32.3 was digested with BglII and Bsu36I and the 3.2 kbp fragment was ligated with plasmid pJLK116

which had been digested with BglII and Bsu36I. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK38 was identified by its restriction pattern.

Construction of plasmid pJLK41

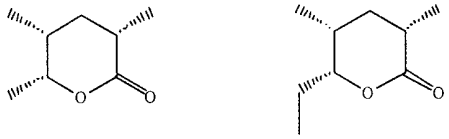
Plasmid pJLK38 was digested with NdeI and XbaI and the approximately 13 kbp fragment was ligated with plasmid pCJR24 which had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK41 was identified by its restriction pattern.

Example 10

Use of plasmid pJLK41 for construction of JC2/pJLK41 and the production of triketides

Approximately 5 μg plasmid pJLK41 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE. JC2/pJLK41 was plated onto SM3 agar containing 50 μg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1 cm^2 (0.5 ml) of the plate was homogenised and extracted with a mixture of 1.2 ml ethyl acetate and 20 ul formic acid. The solvent was decanted and removed by evaporation and the residue dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. The major products were identified (by

comparison with authentic material) as (2S, 4S, 5R)-2,4-dimethyl-5-hydroxy-n-hexanoic acid δ-lactone and as (2S, 4S, 5R)-2,4-dimethyl-5-hydroxy-n-heptanoic acid δ-lactone.



Example 11

Construction of plasmid pJLK29

Plasmid pJLK29 is a pJLK117 based plasmid except that the DNA fragment encoding the reductive loop of module 10 of the rap PKS has been inserted into the mcs. It was constructed via several intermediate plasmids as follows. (Figure 5)

Construction of plasmid pJLK121.1

The approximately 2.2 kbp DNA segment of the rapB gene of S. hygroscopicus encoding the reductive loop of module 10 was amplified by PCR using as primers the synthetic oligonucleotides:
5'-TAAGATCTTCCGACGTACGCGTTCCAGC-3' and
5'-ATGCTAGCCACTGCGCCGACGAATCACCGGTGG-3' and as template an approximately 7 kbp fragment, which has been obtained by digestion of cosmid cos 26 (Schwecke, T. et al. (1995) Proc. Natl. Acad. Sci. USA 92:7839-7843) with ScaI and SphI. The PCR product was treated with T4 polynucleotide

kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK121.1 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK29

Plasmid pJLK121.1 was digested with BglII and NheI and the 2.2 kbp fragment was ligated with plasmid pJLK117 which had been digested with BglII and NheI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK29 was identified by its restriction pattern.

Example 12

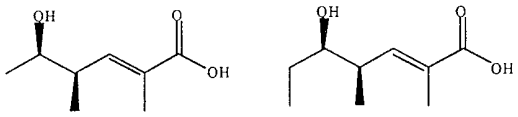
Use of plasmid pJLK29 for construction of S. erythraea JC2/pJLK29 and the production of triketides

Approximately 5 μg plasmid pJLK29 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE. JC2/pJLK29 was used to inoculate 30 ml of SM3 medium containing 5 μg/ml thiostrepton in a 250 ml flask with a single spring to reduce clumping, shaken at 300 rpm and at 30°C. After 8 days the broth was centrifuged, the supernatant adjusted to pH 3 and extracted three times with an equal volume of ethyl acetate. The solvent was

removed by evaporation and the residue dissolved in methanol and analysed by HPLC and electrospray mass spectroscopy and, after conversion to the methyl ester with trimethylsilyl-diazomethane by GC/MS. The major products were identified (by comparison with authentic material) as

(4S, 5R)- 5-hydroxy-2,4-dimethyl-n-hex-2-enoic acid and as

(4S, 5R)-5-hydroxy-2,4-dimethyl0-n-hept-2-enoic acid.



Example 13

Construction of plasmid pJLK35

Plasmid pJLK35 is a pJLK117 based plasmid except that the DNA fragment encoding the reductive loop of module 1 of the tylosin PKS has been inserted into the mcs. It was constructed via several intermediate plasmids as follows. (Figure 5)

Construction of plasmid pJLK33.1

The approximately 1.6 kbp DNA segment of the tylosin PKS gene of S. fradiae encoding the reductive loop of module 1 was amplified by PCR using as primers the synthetic

oligonucleotides:
5'-TAAGATCTCCCTACGTACCCCTTCAACCAC-3' and
5'-GCTAGCCGCCGCGCCAGCTCGGGC-3' and cosmid 6T (cosmid containing the tylosin-producing PKS genes) as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK33.1 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK35

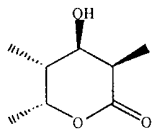
Plasmid pJLK33.1 was digested with BglII and NheI and the 1.6 kbp fragment was ligated with plasmid pJLK117 which had been digested with BglII and NheI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK35 was identified by its restriction pattern.

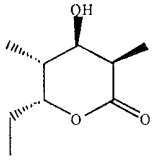
Example 14

Use of plasmid pJLK35 for construction of S. erythraea JC2/pJLK35 and the production of triketides

Approximately 5 μg plasmid pJLK35 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE. JC2/pJLK35 was plated onto SM3 agar

containing 50 μg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1 cm^2 (0.5 ml) of the plate was homogenised and extracted with a mixture of 1.2 ml ethyl acetate and 20 ul formic acid. The solvent was decanted and removed by evaporation and the residue dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. The major products were identified (by comparison with authentic material) as (2R, 3R, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-hexanoic acid δ-lactone and as (2R, 3R, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-heptanoic acid δ-lactone.





Example 15

Construction of plasmid pRIF7

Plasmid pRIF7 is a pJLK117 based plasmid except that the DNA fragment encoding the reductive loop of module 7 of the rifamycin PKS has been inserted into the mcs. It was constructed via several intermediate plasmids as follows. (Figure 5)

Construction of plasmid pUCRIF7

The approximately 2.1 kbp DNA segment of the rifamycin PKS gene of Amycolatopsis mediterranei encoding the reductive loop of module 7 was amplified by PCR using as primers the

synthetic oligonucleotides:
5'-CCTACGTACGCCTTCGACCACCAGCACTT-3' and
5'-CGGCTAGCGGGCGTTCCAGGCCGCCGTCCT and cosmid 6 (cosmid starting at 35727 and going beyond 76199, numbers according to accession number AF040570) as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pUCRIF7 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pRIF7

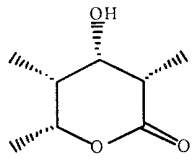
Plasmid pUCRIF7 was digested with SnaBI and NheI and the 2.1 kbp fragment was ligated with plasmid pJLK117 which had been digested with SnaBI and NheI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pRIF7 was identified by its restriction pattern.

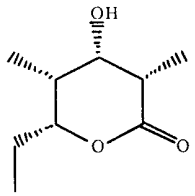
Example 16

Use of plasmid pRIF7 for construction of S. erythraea JC2/pRIF7 and the production of triketides

Approximately 5 μg plasmid pRIF7 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated

into the TE. JC2/pRIF7 was plated onto SM3 agar containing 50 μg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1 cm^2 of the plate was homogenised and extracted with a mixture of 1.2 ml ethyl acetate and 20 ul formic acid. The solvent was decanted and removed by evaporation and the residue dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. The major products were identified (by comparison with authentic material) as (2S, 3S, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-hexanoic acid δ-lactone and as (2R, 3R, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-heptanoic acid δ-lactone.





Example 17

Construction of plasmid pJLK52

Plasmid pJLK52 is a pJLK35 based plasmid containing a PKS gene comprising the ery loading module, the first, the second and the third extension modules of the ery cluster and the ery chain-terminating thioesterase except that the DNA segment between the end of the acyltransferase and the beginning of the ACP of the second ery extension module

has been substituted by the equivalent segment of module 1 of the tylosin PKS.
It was constructed via several intermediate plasmids as follows.

Construction of plasmid pJLK50

The approximately 6.1 kbp DNA segment of the erythromycin PKS gene cluster of S. erythraea encoding the DNA fragment from the beginning of the ACP of module 2 to the beginning of the ACP of module 3 was amplified by PCR using as primers the synthetic oligonucleotides:
5'-TACCTGAGGGACCGGCTAGCGGGTCTGCCGCGTG-3' and
5'-ATGCTAGCCGTTGTGCCGGCTCGCCGGTCGGTCC-3' and plasmid pBAM25 as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK50 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK52

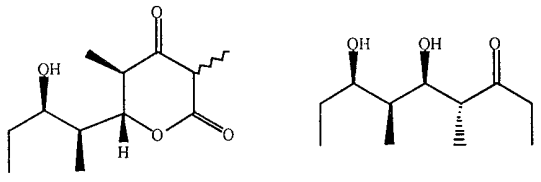
Plasmid pJLK50 was digested with NheI and the 6.1 kbp insert was ligated with plasmid pJLK35 which had been digested with NheI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK52 was identified by its restriction pattern.

Example 18

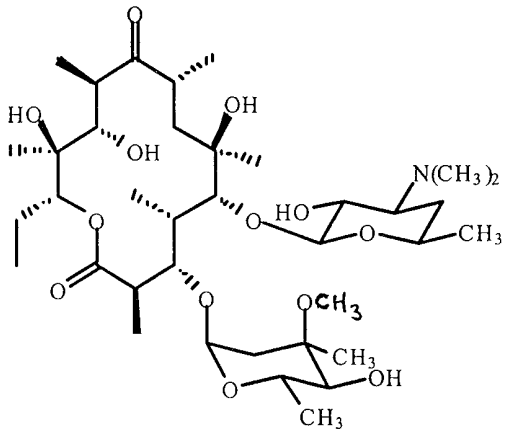
Use of plasmid pJLK52 for construction of S. erythraea NRRL2338/pJLK52 and the production of tetraketides and macrolides

Approximately 5 μg plasmid pJLK52 were used to transform protoplasts of S. erythraea NRRL2338 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA is obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE.

S. erythraea NRRL2338/pJLK52 was used to inoculate SM3 medium containing 5 μg/ml thiostrepton and allowed to grow for seven to twelve days at 28-30°C. After this time the broth was centrifuged and the pH of the supernatant adjusted to pH=9.5. The supernatant was then extracted three times with an equal volume of ethyl acetate and the solvent was removed by evaporation. The residue was dissolved in methanol and analysed by GC/MS by HPLC/MS and MS-MS. Tetraketides were identified by GC/MS. The major components were



The following macrolide was identified by HPLC/MS, MS-MS and 1H-NMR (it was accompanied by products of incomplete processing by post-PKS enzymes)



Example 19

Construction of plasmid pJLK53

Plasmid pJLK53 is a pJLK28 based plasmid containing a PKS gene comprising the ery loading module, the first, the second and the third extension modules of the ery cluster and the ery chain-terminating thioesterase except that the DNA segment between the end of the acyltransferase and the beginning of the ACP of the second ery extension module has been substituted by the equivalent segment of module 13 of the rapamycin PKS. It was constructed as follows.

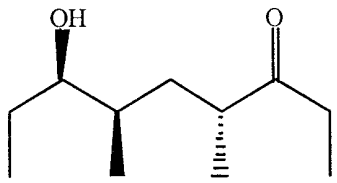
Plasmid pJLK50 was digested with NheI and the 6.1 kbp insert was ligated with plasmid pJLK28 which had been digested with NheI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK53 was identified by its restriction pattern.

Example 20

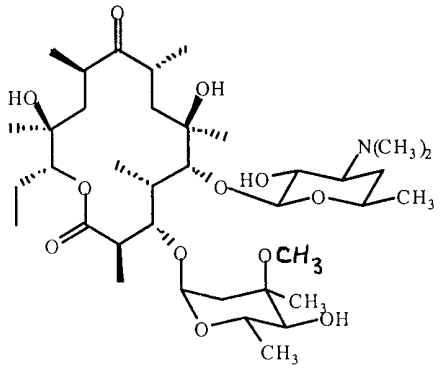
Use of plasmid pJLK53 for construction of S. erythraea NRRL2338/pJLK53 and the production of TKL derivatives

Approximately 5 μg plasmid pJLK53 were used to transform protoplasts of S. erythraea NRRL2338 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA is obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE.

S. erythraea NRRL2338/pJLK53 was used to inoculate SM3 medium containing 5 μg/ml thiostrepton and allowed to grow for seven to ten days at 28-30°C. After this time the broth was centrifuged and the pH of the supernatant adjusted to pH=9.5. The supernatant was then extracted three times with an equal volume of ethyl acetate and the solvent was removed by evaporation. The residue was dissolved in methanol and analysed by GC/MS by HPLC/MS and MS-MS. Tetraketides were identified by GC/MS. The major component was



The following macrolide was identified by HPLC/MS, MS-MS and 1H-NMR (it was accompanied by products of incomplete processing by post-PKS enzymes)



Example 21

Construction of plasmid pJLK54

Plasmid pJLK54 is a pJLK29 based plasmid containing a PKS gene comprising the ery loading module, the first, the second and the third extension modules of the ery cluster and the ery chain-terminating thioesterase except that the DNA segment between the end of the acyltransferase and the beginning of the ACP of the second ery extension module has been substituted by the equivalent segment of module 10 of the rapamycin PKS.
It was constructed as follows.

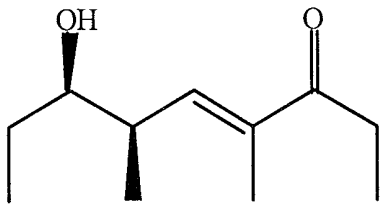
Plasmid pJLK50 was digested with NheI and the 6.1 kbp insert was ligated with plasmid pJLK29 which had been digested with NheI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK54 was identified by its restriction pattern.

Example 22

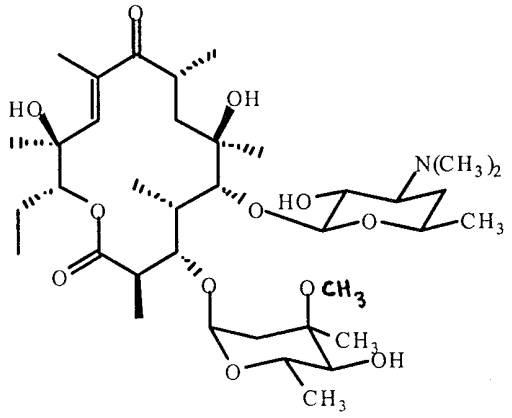
Use of plasmid pJLK54 for construction of S. erythraea NRRL2338/pJLK54 and the production of tetraketide derivatives and macrolides

Approximately 5 μg plasmid pJLK54 were used to transform protoplasts of S. erythraea NRRL2338 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA is obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE.

S. erythraea NRRL2338/pJLK54 was used to inoculate SM3 medium containing 5 μg/ml thiostrepton and allowed to grow for seven to ten days at 28-30°C. After this time the broth was centrifuged and the pH of the supernatant adjusted to pH=9.5. The supernatant was then extracted three times with an equal volume of ethyl acetate and the solvent was removed by evaporation. The residue was dissolved in methanol and analysed by GC/MS by HPLC/MS and MS-MS. Tetraketides were identified by GC/MS. The major component was



The following macrolide was identified by HPLC/MS, MS-MS and 1H-NMR (it was accompanied by products of incomplete processing by post-PKS enzymes)



Avermectins

Example 23

Construction of pJLK136

Plasmid pJLK136 is a pWHM3 based plasmid comprising the upstream and the downstream flanking region of the reductive loop of module 2 of the avermectin PKS gene and the erythromycin resistance gene inserted into the mcs which connects these two fragments. Plasmid pWHM3 is described in Vara J et al, J Bacteriol 1989, 171: 5872-5881. Plasmid pJLK136 was constructed via several intermediate plasmids as follows (Figure 6).

Construction of pJLK130

The approximately 2.4 kbp DNA segment of the avermectin PKS gene of S. avermitilis encoding the region upstream of the reductive loop of module 2 was amplified by PCR using as primers the synthetic oligonucleotides:
5'-GACGCCGAATTCTTCGGCATCAGCCCCCGCGAAG-3' and
5'-
GAGCTAGCAGGTGGGGAGATCTAGGTGGGTGTGGGTGTGGGGTTGGTTGTGGTGGTGG

GTGTA-3' and plasmid pIG22 (Galloway, I. S. (1998) Thesis, University of Cambridge, UK) as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK130 was identified by its restriction pattern and DNA sequencing.

Construction of pJLK131

The approximately 2.0 kbp DNA segment of the avermectin PKS gene of S. avermitilis encoding the region downstream of the reductive loop of module 2 was amplified by PCR using as primers the synthetic oligonucleotides: 5'-GCCCGGCTAGCCGGCCAGACACACGAACAACAGC-3' and 5'-GGGAATTCCTCGAGGATGACGTGGGCGTTGGTGC-3' and plasmid pIG25 (Galloway, I. S. (1998) Thesis, University of Cambridge, UK) as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK131 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK132

Plasmid pJLK130 was digested with NheI and XbaI and the approximately 2.4 kbp insert was ligated with plasmid pJLK131 which had been digested with NheI and XbaI. The ligation mixture was used to transform electrocompetent E.

coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK132 was identified by its restriction pattern.

Construction of plasmid pJLK133

Plasmid pJLK117 was digested with BglII and NheI and the approximately 0.1 kbp insert was ligated with plasmid pJLK132 which had been digested with BglII and NheI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK132 was identified by its restriction pattern.

Construction of pJLK134

The approximately 1.9 kbp DNA segment of the erythromycin gene cluster of S. erythraea encoding the erythromycin resistance was amplified by PCR using as primers the synthetic oligonucleotides:
5'-TAAGATCTAGCGCTCCGAGGTTCTTGCCCG-3' and
5'-ATGCTAGCCTACCGCTGCCCGGGTCCGCCG-3' and plasmid pRH3 (Dhillon, N, et al. Molecular Microbiology (1989) 3:1405-1414) as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK134 was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK135

Plasmid pJLK134 was digested with BglII and NheI and the approximately 1.9 kbp insert was ligated with plasmid pJLK133 which had been digested with BglII and NheI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK135 was identified by its restriction pattern.

Construction of plasmid pJLK136

Plasmid pJLK135 was digested with EcoRI and the approximately 6.3 kbp insert was ligated with plasmid pWHM3 which had been digested with EcoRI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK136 was identified by its restriction pattern.

Example 24

Use of plasmid pJLK136

Approximately 10 μg plasmid pJLK136 were used to transform protoplasts of S. avermitilis (MacNeil, D.J. and Klapko, C.M. Journal of Industrial Microbiology (1987) 2:209-218) and stable thiostrepton and erythromycin resistant colonies were isolated. Individual colonies were selected and subcultured four times in non-selective liquid medium followed by preparation and regeneration of protoplasts (media according to MacNeil T. et al J. Bacteriol. (1993) 175:2552-2563) Thiostrepton sensitive and erythromycin resistant colonies were isolated and characterised by Southern blot hybridisation. One such colony was

designated S. avermitilis/JLK1.

Example 25

Construction of plasmid pJLK137

Plasmid pJLK120 was digested with BglII and NsiI and the approximately 3.2 kbp insert was ligated with plasmid pJLK133 which had been digested with BglII and NsiI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK137 was identified by its restriction pattern.

Construction of plasmid pJLK138

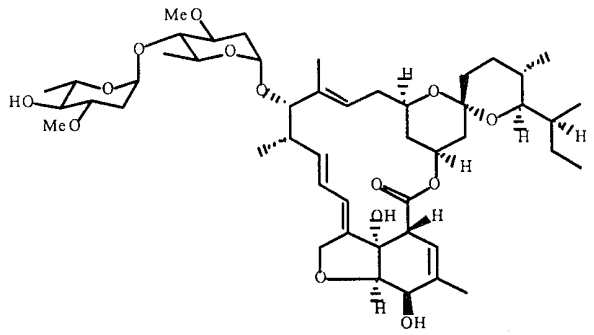
Plasmid pJLK137 was digested with EcoRI and the approximately 7.6 kbp insert was ligated with plasmid pWHM3 which had been digested with EcoRI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK138 was identified by its restriction pattern.

Example 26

Use of plasmid pJLK138

Approximately 10 μg plasmid pJLK138 were used to transform protoplasts of S. avermitilis (MacNeil, D.J. and Klapko, C.M. Journal of Industrial Microbiology (1987) 2:209-218) and stable thiostrepton and erythromycin resistant colonies were isolated. Individual colonies were selected

and subcultured four times in non-selective liquid medium followed by preparation and regeneration of protoplasts (media according to MacNeil T. et al J. Bacteriol. (1993) 175:2552-2563) Thiostrepton and erythromycin sensitive colonies were isolated and characterised by Southern blot hybridisation. One colony of S. avermitilis/pJLK138 was used to inoculate liquid media (fermentation according to Pang, C-H. et al J. of Antibiotics (1995) 48:59-66). the cultures were harvested and the products isolated and purified as described in the literature (Pang, C-H. et al J. of Antibiotics (1995) 48:59-66). The products were analysed by HPLC/MS and 1H-NMR and the following compound could be identified:



Example 27

Construction of plasmid pJLK139

Plasmid pJLK121.1 was digested with BglII and NheI and the 2.2 kbp fragment was ligated with plasmid pJLK133 which had been digested with BglII and NheI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK139 was

identified by its restriction pattern.

Construction of plasmid pJLK140

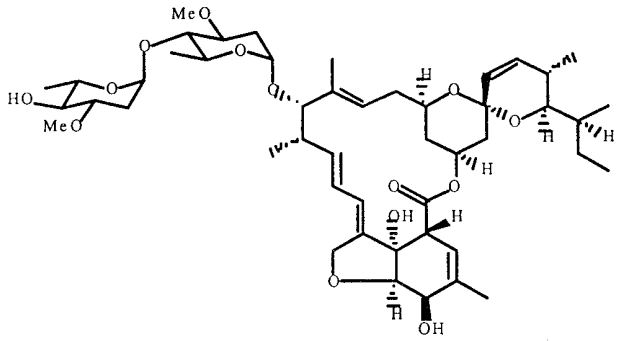
Plasmid pJLK139 was digested with EcoRI and the approximately 6.6 kbp insert was ligated with plasmid pWHM3 which had been digested with EcoRI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK140 was identified by its restriction pattern.

Example 28

Use of plasmid pJLK140

Approximately 10 μg plasmid pJLK140 were used to transform protoplasts of S. avermitilis (MacNeil, D.J. and Klapko, C.M. Journal of Industrial Microbiology (1987) 2:209-218) and stable thiostrepton and erythromycin resistant colonies were isolated. Individual colonies were selected and subcultured four times in non-selective liquid medium followed by preparation and regeneration of protoplasts (media according to MacNeil T. et al J. Bacteriol. (1993) 175:2552-2563) Thiostrepton and erythromycin sensitive colonies were isolated and characterised by Southern blot hybridisation. One colony of S. avermitilis/pJLK140 was used to inoculate liquid media (fermentation according to Pang, C-H. et al J. of Antibiotics (1995) 48:59-66). the cultures were harvested and the products isolated and purified as described in the literature (Pang, C-H. et al J. of Antibiotics (1995) 48:59-66). The products were analysed by HPLC/MS and 1H-NMR and the following compound

could be identified:



Example 29

Construction of plasmid pJLK30

pJLK30 is a pJLK117 based plasmid except that the DNA encoding the reductive loop of module 1 of the avermectin PKS has been inserted into the polylinker using the restriction sites BglII and NheI. It was constructed via several intermediate plasmids.

Construction of plasmid pIG67

The approximately 1.7 kbp DNA segment of the gene of the avermectin PKS of S. avermitilis encoding the reductive loop of module 1 was amplified by PCR using the following synthetic oligonucleotides as primers:
5'-CCTAGATCCGCCCACCTACCCCTTCCAACACCAG-3' and
5'-TGGGCTAGCGTTTTGTGCAACTCCGCCGGTGGAGTG-3' and as template either plasmid pIG155, which contains the first two modules of the avermectin PKS cloned into plasmid pT7-7,

or chromosomal DNA of Streptomyces avermitilis. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pIG67 was identified by its restriction pattern and by DNA sequencing.

Construction of plasmid pJLK30

Plasmid pIG67 was digested with BglII and NheI and the 1.7 kbp fragment was ligated with plasmid pJLK117 which had been digested with BglII and NheI. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK30 was identified by its restriction pattern.

Example 30

Use of plasmid pJLK30 for the construction of S. erythraea JC2/pJLK30 and the production of triketides.

Approximately 5 mg of plasmid pJLK30 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation to confirm that the plasmid had integrated into the TE. S. erythraea JC2/pJLK30 was plated onto SM3 agar containing 50 mg/ml thiostrepton and

allowed to grow for twelve days at 30°C. 1cm^2 of the plate was homogenized and extracted with a mixture of 1.2 ml ethyl acetate with 20 ml formic acid. The solvent was decanted and evaporated. The residue was dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. The major products were identified as (2R, 3R, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-hexanoic acid δ-lactone and as (2R, 3R, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-heptanoic acid δ-lactone (total of 25 mg/l). Almost none of the corresponding 3-ketolactone could be detected.

Example 31

Construction of plasmid pGMS2

pGMS2 is a pJLK117 based plasmid except that the DNA encoding the reductive loop of module 1 of the avermectin PKS has been inserted into the polylinker using the restriction sites PstI and Bsu36I. It was constructed via several intermediate plasmids.

Construction of plasmid pIG68

The approximately 1.7 kbp DNA segment of the gene of the avermectin PKS of S. avermitilis encoding the reductive loop of module 1 was amplified by PCR using the following synthetic oligonucleotides as primers:
5'-TGGCTGCAGAGCTCACAGCCGGGTGCCGGATCCGGTT-3' and
5'-TTTCCTCAGGTCCGCCGGTGGAGTGGGGCGCTGGAC-3' and as template either plasmid pIG155, which contains the first two modules of the avermectin PKS cloned into plasmid pT7-7,

or chromosomal DNA of Streptomyces avermitilis. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pIG68 was identified by its restriction pattern and by DNA sequencing.

Construction of plasmid pGMS1

Plasmid pIG68 was digested with PstI and Bsu36I and the 1.7 kbp fragment was ligated with plasmid pJLK116 which had been digested with PstI and Bsu36I. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pGMS1 was identified by its restriction pattern.

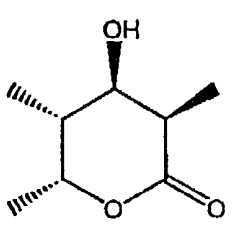
Construction of plasmid pGMS2

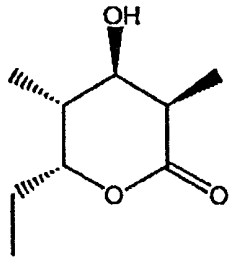
Plasmid pGMS1 was digested with NdeI and XbaI and the approximately 11.5 kbp fragment was ligated with plasmid pCJR24 which had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pGMS2 was identified by its restriction pattern.

Example 32

Use of plasmid pGMS2 for the construction of S. erythraea JC2/pGMS2 and the production of triketides.

Approximately 5mg of plasmid pGMS2 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation to confirm that the plasmid had integrated into the TE. S. erythraea JC2/pGMS2 was plated onto SM3 agar containing 50μg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1cm^2 of the plate was homogenized and extracted with a mixture of 1.2 ml ethyl acetate with 20 ml formic acid. The solvent was decanted and evaporated. The residue was dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. The products were identified as (2R, 3R, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-hexanoic acid δ-lactone and as (2R, 3R, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-heptanoic acid δ-lactone (total of 17 mg/l), and also a substantial amount of the corresponding 3-ketolactone (5.5 mg/l).





Example 33

Construction of plasmid pJLK31

pJLK31 is a pJLK117 based plasmid except that the DNA encoding the reductive loop of module 2 of the avermectin PKS has been inserted into the polylinker using the restriction sites BglII and NheI. It was constructed via several intermediate plasmids.

Construction of plasmid pIG69

The approximately 2.4 kbp DNA segment of the gene of the avermectin PKS of S. avermitilis encoding the reductive loop of module 2 was amplified by PCR using the following synthetic oligonucleotides as primers: 5'-CCTAGATCTCCCCACCTACCCCTTCCAACACCACCACTACTG-3' and 5'-CCGGCTAGCCGGGCGTGCAGCTGGGCGCCGTTGTCCGCAC-3' and as template either plasmid pIG155, which contains the first two modules of the avermectin PKS cloned into plasmid pT7-7, or chromosomal DNA of Streptomyces avermitilis. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pIG69 was identified by its restriction pattern and by DNA sequencing.

Construction of plasmid pJLK31

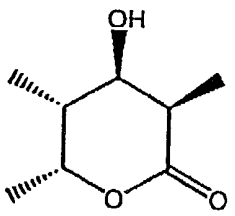
Plasmid pIG69 was digested with BglII, NheI and DraI and

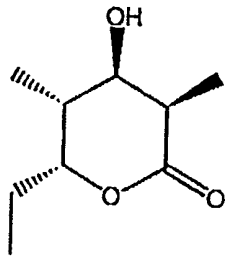
the 2.4 kbp fragment was ligated with plasmid pJLK117 which had been digested with BglII and NheI. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK31 was identified by its restriction pattern.

Example 34

Use of plasmid pJLK31 for the construction of S. erythraea JC2/pJLK31 and the production of triketides.

Approximately 5 mg of plasmid pJLK31 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation to confirm that the plasmid had integrated into the TE. S. erythraea JC2/pJLK31 was plated onto SM3 agar containing 50 mg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1cm^2 of the plate was homogenized and extracted with a mixture of 1.2 ml ethyl acetate with 20 ml formic acid. The solvent was decanted and evaporated. The residue was dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. The major products were identified as (2R, 3R, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-hexanoic acid δ-lactone and as (2R, 3R, 4S, 5R)-5,3-dihydroxy-2,4-dimethyl-n-heptanoic acid δ-lactone (total of 30 mg/litre).





Example 35

Construction of plasmid pGMS4

pGMS4 is a pJLK117 based plasmid except that the DNA encoding the reductive loop of module 2 of the avermectin PKS has been inserted into the polylinker using the restriction sites SnaBI and Bsu36I. It was constructed via several intermediate plasmids.

Construction of plasmid pIG70

The approximately 2.4 kbp DNA segment of the gene of the avermectin PKS of S. avermitilis encoding the reductive loop of module 2 was amplified by PCR using the following synthetic oligonucleotides as primers:
5'-CCCTACGTACCCCTTCCAACACCACTACTGGCTCGAAAG-3' and
5'-GGCCCTCAGGTGGGCGCCGTTGTCCGCACCACCGGTA-3' as template either plasmid pIG155, which contains the first two modules of the avermectin PKS cloned into plasmid pT7-7, or chromosomal DNA of Streptomyces avermitilis. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by

digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pIG70 was identified by its restriction pattern and by DNA sequencing.

Construction of plasmid pGMS3

Plasmid pIG70 was digested with SnaBI, Bsu36I and DraI and the 2.4 kbp fragment was ligated with plasmid pJLK116 which had been digested with SnaBI and Bsu36I. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pGMS3 was identified by its restriction pattern.

Construction of plasmid pGMS4

Plasmid pGMS2 was digested with NdeI and XbaI and the approximately 12.4 kbp fragment was ligated with plasmid pCJR24 which had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pGMS4 was identified by its restriction pattern.

Example 36

Use of plasmid pGMS4 for the construction of S. erythraea JC2/pGMS4 and the production of triketides.

Approximately 5 mg of plasmid pGMS4 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation to confirm that the plasmid had integrated into the TE. S. erythraea JC2/pGMS4 was plated onto SM3 agar containing 50 mg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1cm^2 of the plate was homogenized and extracted with a mixture of 1.2 ml ethyl acetate with 20 ml formic acid. The solvent was decanted and evaporated. The residue was dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. Only traces of putative triketide products were detected.

Example 37

Construction of plasmid pJLK27

Plasmid pJLK27 is a pJLK114 based plasmid except that the DNA fragment encoding the reductive loop of module 13 of the rap PKS has been inserted into the mcs. It was constructed via several intermediate plasmids as follows.

Construction of plasmid pJLK120a

The approximately 3.2 kbp DNA segment of the rapC gene of S. hygroscopicus encoding the reductive loop of module 13 was amplified by PCR using as primers the synthetic oligonucleotides:
5'-TACCTAGGCACCACCACAACCCGGGTA-3' and
5'-TACAATTGGCCCGCGAGTCCCCGACGCT-3' and cosmid cos 31 (Schwecke, T. et al. (1995) Proc. Natl. Acad. Sci. USA

92:7839-7843) as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK120a was identified by its restriction pattern and DNA sequencing.

Construction of plasmid pJLK27

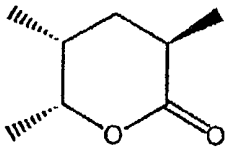
Plasmid pJLK120a was digested with AvrII and HpaI and the 3.2 kbp fragment was ligated with plasmid pJLK114 which had been digested with AvrII and HpaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK27 was identified by its restriction pattern.

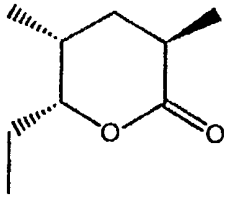
Example 38

Use of plasmid pJLK27 for construction of JC2/pJLK27 and the production of triketides

Approximately 5 mg plasmid pJLK27 were used to transform protoplasts of S. erythraea JC2 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE. JC2/pJLK27 was plated onto SM3 agar containing 50 mg/ml thiostrepton and allowed to grow for twelve days at 30°C. 1 cm^2 (0.5 ml) of the plate was

homogenised and extracted with a mixture of 1.2 ml ethyl acetate and 20 ml formic acid. The solvent was decanted and removed by evaporation and the residue dissolved in methanol and analysed by GC/MS and electrospray mass spectroscopy. The major products were identified (by comparison with authentic material) as (2R, 4S, 5R)-2,4-dimethyl-5-hydroxy-n-hexanoic acid δ-lactone and as (2R, 4S, 5R)-2,4-dimethyl-5-hydroxy-n-heptanoic acid δ-lactone (total of 41 mg/l), with some of the corresponding 3-ketolactones (total of 12 mg/l) and 3-hydroxylactones (total of 2.8 mg).





All documents and sequence deposits referred to herein are explicitly and individually incorporated herein by reference.

CLAIMS

1. A nucleic acid molecule encoding at least part of a Type I polyketide synthase, said part comprising at least part of an extension module, wherein the nucleic acid has, in place of one or more genes encoding enzymes associated with reduction, a polylinker with multiple restriction enzyme sites.

2. A nucleic acid according to claim 1 wherein the polylinker is in place of all genes encoding enzymes which are associated with reduction and which are normally included in said extension module.

3. A nucleic acid encoding at least part of a Type I polyketide synthase, said part comprising at least part of an extension module, wherein the nucleic acid has a polylinker with multiple restriction enzyme sites, which polylinker connects nucleic acid encoding at least part of an acyl transferase enzyme to nucleic acid encoding at least part of an acyl carrier protein.

4. A nucleic acid according to any preceding claim wherein at least some of the restriction sites included in the polylinker are absent from the Type I polyketide synthase-encoding nucleic acid.

5. A nucleic acid according to any preceding claim wherein at least some of the restriction sites included in the polylinker are uncommon in or absent from other naturally occurring nucleic acid sequences which encode reductive enzymes of Type I polyketide synthases.

6. A nucleic acid according to any preceding claim

wherein the polylinker includes at least some of the following restriction sites: AvrII, BglII; SnaBI; PstI; SpeI; NsiI; Bsu36l; NheI; and HpaI.

7. A nucleic acid according to any preceding claim which additionally encodes a loading module

8. A nucleic acid according to any preceding claim which additionally encodes one or more further extension modules.

9. A nucleic acid according to any preceding claim further including a nucleic acid sequence incorporated into the polylinker, which incorporated nucleic acid encodes one or more reductive enzymes.

10. A nucleic acid according to claim 9 wherein said one or more reductive enzymes is/are a β-ketoreductase, a dehydratase and/or an enoyl reductase.

11. A nucleic acid according to claim 10 wherein said one or more reductive enzymes include(s) at least a β-ketoreductase.

12. A nucleic acid according to claim 10 or claim 11 wherein at least one of said one or more reductive enzymes is from a different extension module of the same polyketide synthase as said at least part of a Type I polyketide synthase.

13. A nucleic acid according to any one of claims 10 to 12 wherein at least one of said one or more reductive enzymes is from a different polyketide synthase.

14. A vector including a nucleic acid as defined in any preceding claim.

15. A host cell transfected, transformed or conjugated with a nucleic acid or vector as defined in any preceding claim.

16. A host cell according to claim 15 which is a cell of a *Streptomyces* species.

17. A host cell according to claim 16 which is a cell of *S. erythraea* or *S. avermitilis*.

18. A method for producing a nucleic acid encoding a novel polyketide synthase, the method including the steps of:

i. providing a nucleic acid as defined in any one of claims 1 to 8; and

ii. incorporating into said nucleic acid a nucleic acid sequence which encodes at least one reductive enzyme.

19. A method according to claim 18 wherein said nucleic acid sequence encoding at least one reductive enzyme is as defined in any one of claims 9 to 13.

20. A method for producing a fermentation product containing a polyketide, the method including the step of culturing a host cell as defined in claim 15.

21. A fermentation product containing a C22-C23 dihydroavermectin, substantially free of other macrolides.

22. A fermentation product according to claim 21 wherein the dihydroavermectin is ivermectin.

23. A fermentation product containing a B1 avermectin substantially free of B2 avermectins.

24. A method for producing a polyketide, the method including the steps of:

i. providing a fermentation product resulting from the method of claim 20, or a fermentation product according to any of claims 21-23; and

ii. at least partially purifying a polyketide from said fermentation product.

25. A method according to claim 24 wherein the polyketide is an avermectin.

26. A method according to claim 25 wherein the avermectin is a B1 avermectin.

27. A method according to claim 25 wherein the avermectin is a B1 avermectin.

Fig. 1

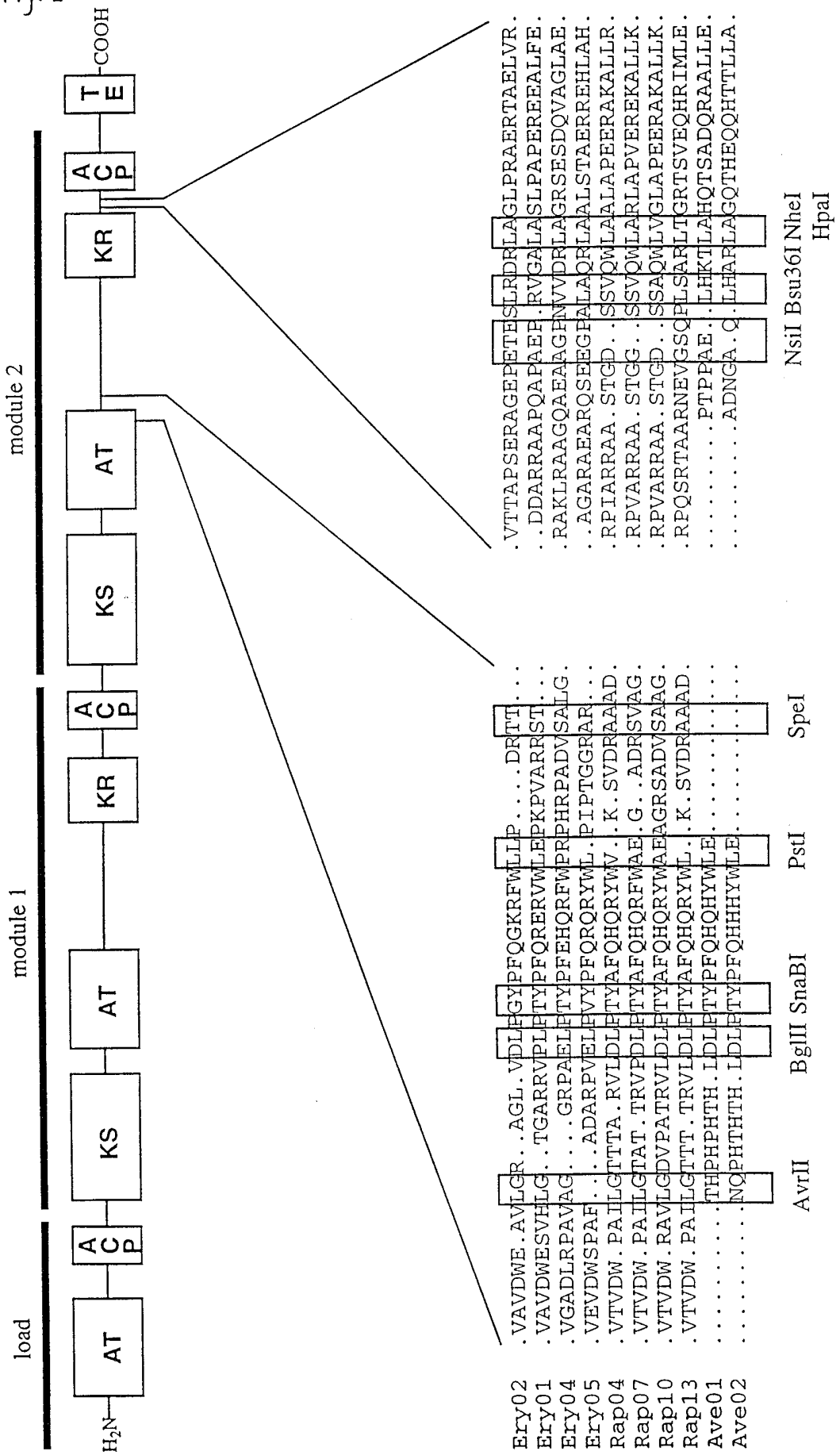

load
module 1
module 2
H_2N
AT
ACP
KS
AT
KR
ACP
KS
AT
KR
ACP
TE
COOH
AvrII
BglII SnaBI
PstI
SpeI
NsiI Bsu36I NheI
HpaI
Ery02
Ery01
Ery04
Ery05
Rap04
Rap07
Rap10
Rap13
Ave01
Ave02

Fig. 2

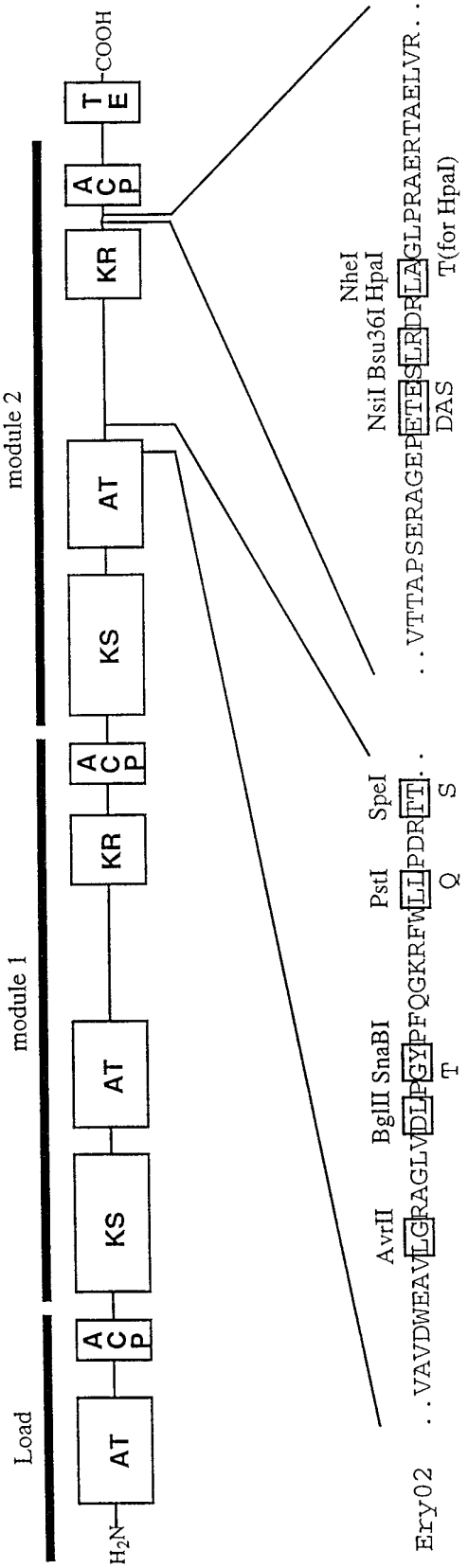

Load
module 1
module 2
H_2N
AT
ACP
KS
AT
KR
ACP
KS
AT
KR
ACP
TE
COOH
Ery02
. . VAVDWEAVLGRAGLVDLLPGYPFQGKRFWLLPDRTT. .
. . VTTAPSERAGEPETESLRDRLAGLPRAERTAELVR. .
AvrII
BglII SnaBI
PstI
SpeI
NsiI Bsu36I HpaI
NheI
T
Q
S
DAS
T(for HpaI)

Fig. 3

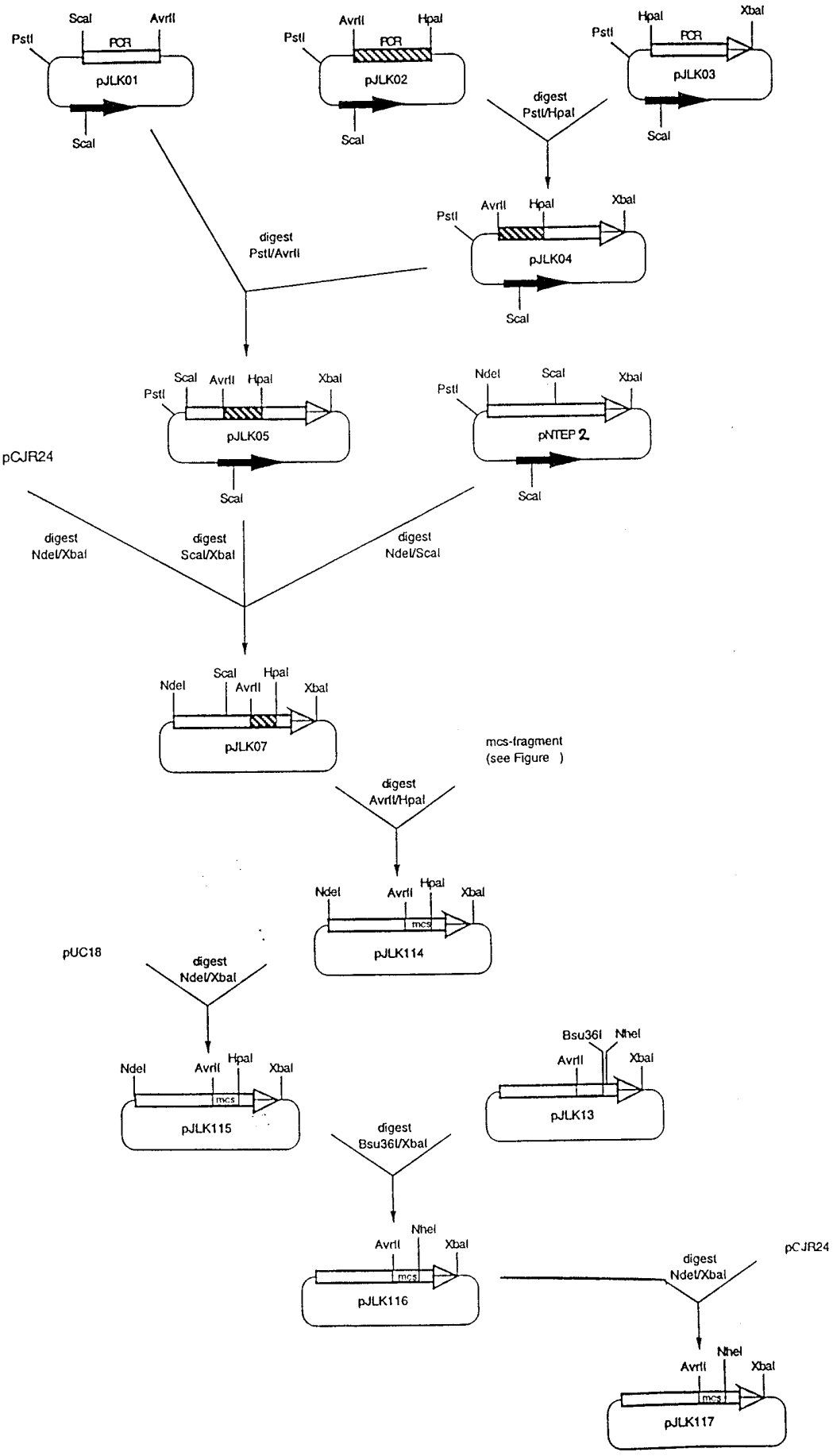

PstI
ScaI
AvrII
PCR
pJLK01
ScaI
PstI
AvrII
HpaI
PCR
pJLK02
ScaI
digest
PstI/HpaI
PstI
HpaI
XbaI
PCR
pJLK03
ScaI
digest
PstI/AvrII
PstI
AvrII
HpaI
XbaI
pJLK04
ScaI
pCJR24
PstI
ScaI
AvrII
HpaI
XbaI
pJLK05
ScaI
PstI
NdeI
ScaI
XbaI
pNTEP 2
ScaI
digest
NdeI/XbaI
digest
ScaI/XbaI
digest
NdeI/ScaI
NdeI
ScaI
AvrII
HpaI
XbaI
pJLK07
mcs-fragment
(see Figure)
digest
AvrII/HpaI
NdeI
AvrII
HpaI
XbaI
mcs
pJLK114
pUC18
digest
NdeI/XbaI
NdeI
AvrII
HpaI
XbaI
mcs
pJLK115
Bsu36I
NheI
AvrII
XbaI
pJLK13
digest
Bsu36I/XbaI
NheI
AvrII
XbaI
mcs
pJLK116
pCJR24
digest
NdeI/XbaI
NheI
AvrII
XbaI
mcs
pJLK117

Fig. 4

forward (Plf):

5'-CTA GGC CGG GCC GGA CTG GTA GAT CTG CCT ACG TAT CCT TTC CAG
GGC AAG CGG TTC TGG CTG CAG CCG GAC CGC ACT AGT CCT CGT GAC GAG
GGA GAT GCA TCG AGC CTG AGG GAC CGG TT-3'

backward (Plb):

5'-AAC CGG TCC CTC AGG CTC GAT GCA TCT CCC TCG TCA CGA GGA CTA GTG
CGG TCC GGC TGC AGC CAG AAC CGC TTG CCC TGG AAA GGA TAC GTA
GGC AGA TCT ACC AGT CCG GCC CGG C-3'

oligos annealed:

```
CTAGGCCGGGCCGGACTGGTAGATCTGCCTACGTATCCTTTCCAGGGCAAGCGGTTCTGGCTGCAG...
    CGGCCCGGCCTGACCATCTAGACGGATGCATAGGAAAGGTCCCGTTCGCCAAGACCGACGTC...
-----                ------   ------                          ------
AvrII                BglII    SnaBI                           PstI
```

```
...CCGGACCGCACTAGTCCTCGTGACGAGGGAGATGCATCGAGCCTGAGGGACCGGTT
...GGCCTGGCGTGATCAGGAGCACTGCTCCCTCTACGTAGCTCGGACTCCCTGGCCAA
          ------                 ------    -------     ---
          SpeI                   NsiI      Bsu36I      HpaI
```

Fig. 5a

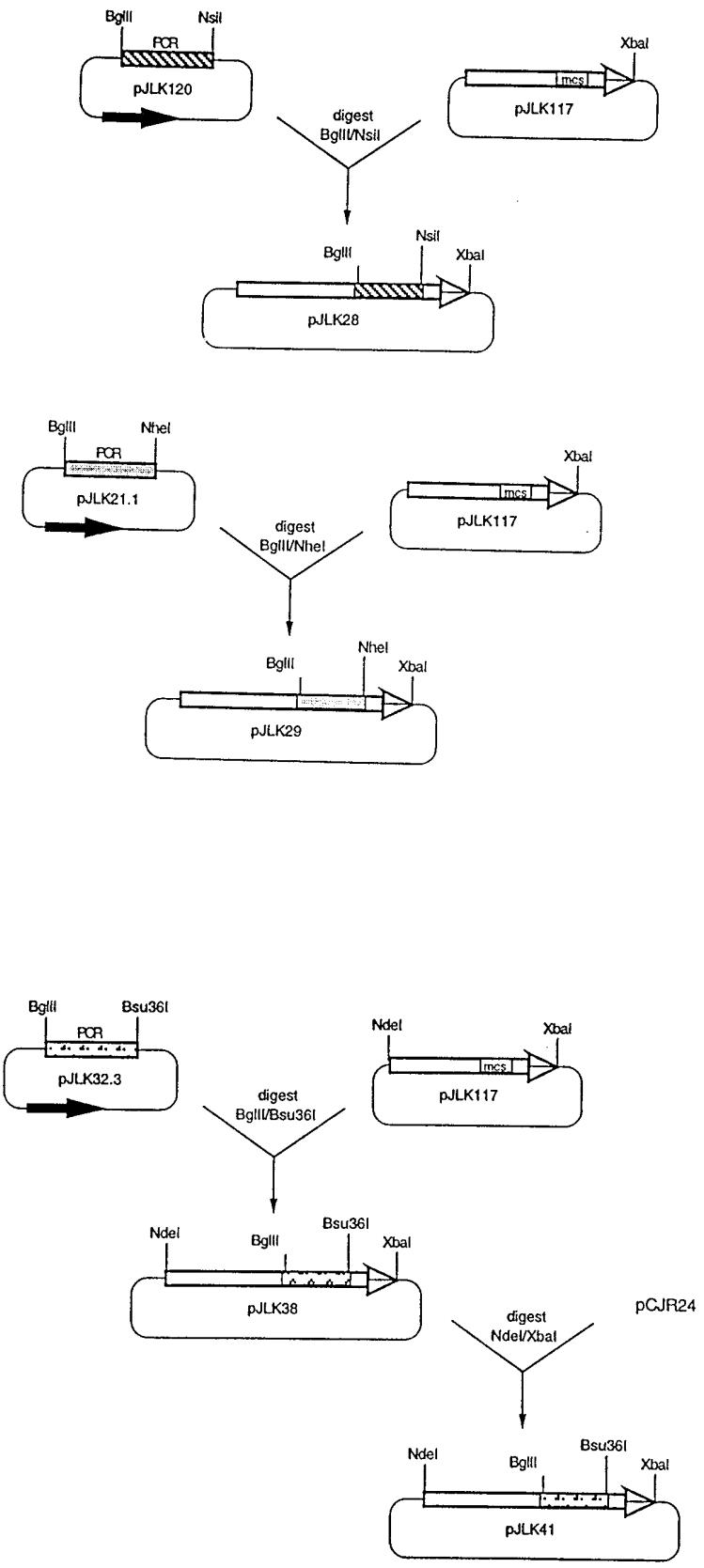
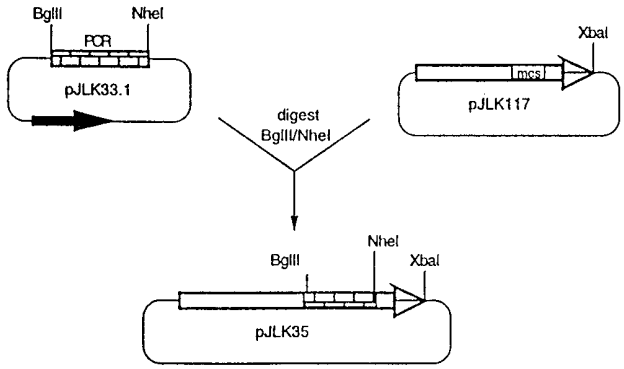

BglII
NsiI
PCR
pJLK120
digest
BglII/NsiI
XbaI
mcs
pJLK117
BglII
NsiI
XbaI
pJLK28
BglII
NheI
PCR
pJLK21.1
digest
BglII/NheI
XbaI
mcs
pJLK117
BglII
NheI
XbaI
pJLK29
BglII
Bsu36I
PCR
pJLK32.3
digest
BglII/Bsu36I
NdeI
XbaI
mcs
pJLK117
NdeI
BglII
Bsu36I
XbaI
pJLK38
digest
NdeI/XbaI
pCJR24
NdeI
BglII
Bsu36I
XbaI
pJLK41

Fig. 5b


BglII
NheI
PCR
pJLK33.1
digest
BglII/NheI
XbaI
mcs
pJLK117
NheI
BglII
XbaI
pJLK35

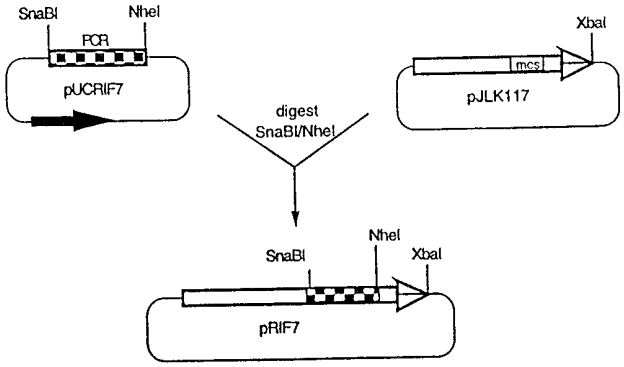

SnaBI
NheI
PCR
pUCRIF7
digest
SnaBI/NheI
XbaI
mcs
pJLK117
NheI
SnaBI
XbaI
pRIF7

Fig. 6

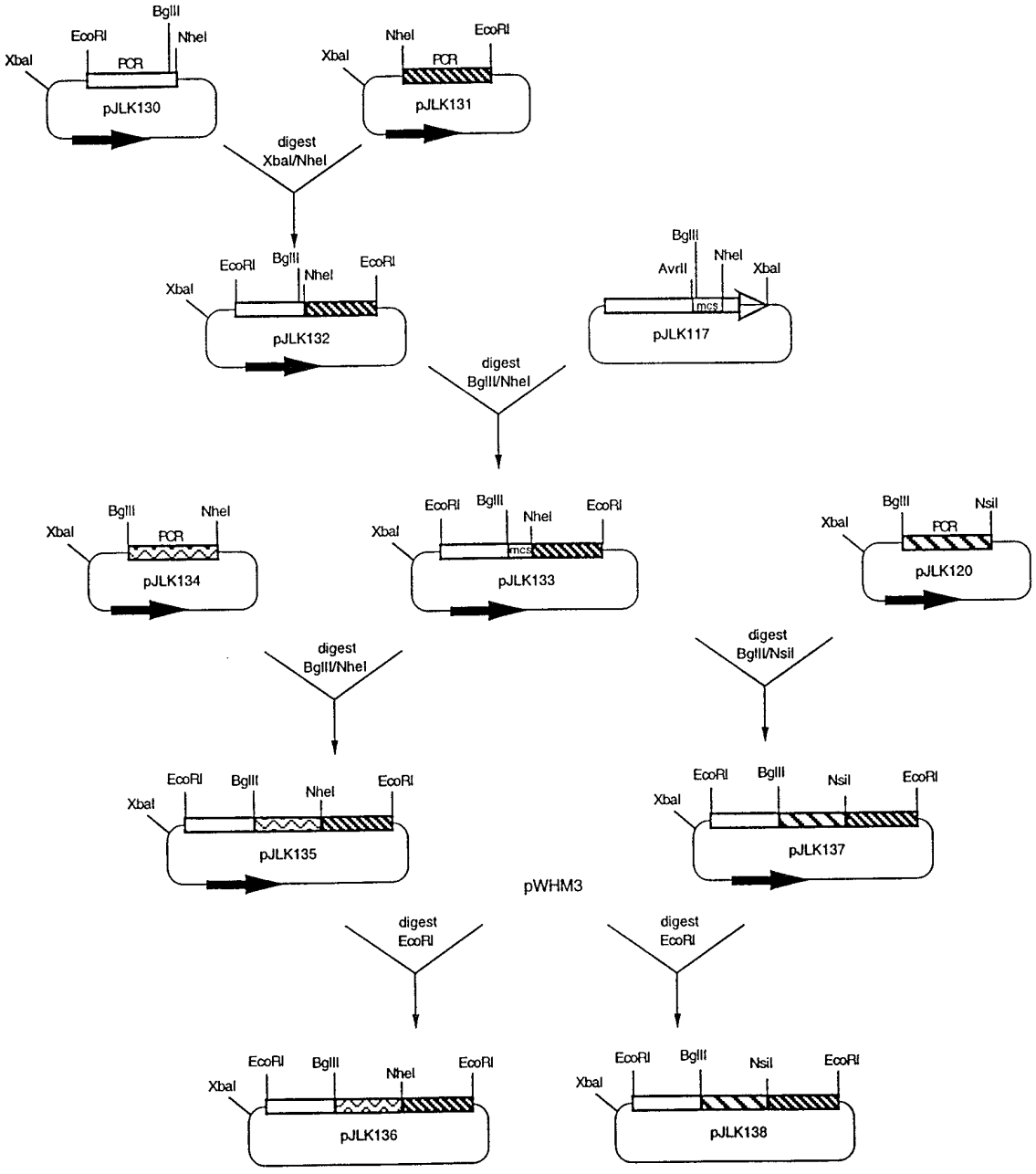

BglII
EcoRI
NheI
XbaI
PCR
pJLK130
pJLK131
digest
XbaI/NheI
pJLK132
AvrII
mcs
pJLK117
digest
BglII/NheI
pJLK134
pJLK133
NsiI
pJLK120
pJLK135
pJLK137
pWHM3
pJLK136
pJLK138

Fig. 7a

```
   1  CCCGGGCGAT CTCCCGGATC ACCTGTGCGG GGCTGGGCAT GTGCAGGAGA
  51  CACTCCAGGG CCCACGCCGC GTCGAAGGAC CCGTCGGGAA ACGGCAGTTC
 101  CATCGCGTCG GCACGGGTGA ACACGACCCG GTCCGCCACG TGCGACTGCT
 151  TCGCGAGAGC GGTCGCCAGC CCGACCTGAA CCTCGCTCAC CGTCACGCCG
 201  ACGACATCGA CGGGCGCGCT CAGGGCGAGC CGCACCGCCG GCTTTCCGGA
 251  ACCGCAGCCG ACGTCCAGGA CCCGGCGGCC CGTGATGCCT CTCAGCTTGC
 301  CGATGAGGAG ATCGGTGAGC CGGTCGGCGG CCTTGCCCGG TGAACTGCCG
 351  TCCCCCGGCT GCGGCCAGTA TCCGAGGTGG GTGTTCCCAC CCAGCGCACG
 401  ATTCATGAGG TCGGTCAAAC GGTCGTAGTA GTCCCCCACT TCCAGGGAAG
 451  AGGGCGGGGT CTGCTCCGGG ACGGCCATCA TGGTCGGGAA CCTCCGCAAT
 501  CCGGGCCGGG CGGCCCAGCT GTCGTGGCGA TCTACTCCAG GAAACGTCGA
 551  CCTTTTTCTG CCACTTGTCC GAGCTATGCA GACACCCCGA TCCCCTAAGA
 601  AATGAACACC CTTGGGAACG GCACAGCCCA GGGGTGGATA GGGGTATTCG
 651  CCGCCGCCGC GCCGTCATTA GCTTTGAAGA GTTGAAGACG TTCAAGACAT
 701  TGATGCCCGG CCGTCAGCGG ATTTCTCGCG CTCCTTTCAT TCTTCGACGC
 751  TGCATTGCAG CTCTCATCAT GTCCGCACGG CCGCCGAGCA TTGCCTAGCG
 801  GTGAGGACAC AGCTCAGGTG CAGAGGATGG ACGGCGGGGA AGAACCCCGC
 851  CCTGCGGCAG GGGAGGTCCT CGGAGTGGCC GACGAGGCGG ACGGCGGCGT
 901  CGTCTTCGTT TTTCCCGGGC AGGGCCCGCA ATGGCCGGGC ATGGGAAGGG
 951  AACTTCTCGA CGCTTCCGAC GTCTTCCGGG AGAGCGTCCG CGCCTGCGAA
1001  GCCGCGTTCG CGCCCTACGT CGACTGGTCG GTGGAGCAGG TGTTGCGGGA
1051  CTCGCCGGAC GCTCCCGGGC TGGACCGGGT GGACGTCGTC CAGCCGACCC
1101  TGTTCGCCGT CATGATCTCC CTGGCCGCCC TCTGGCGCTC GCAAGGGGTC
1151  GAGCCGTGCG CGGTGCTGGG ACACAGCCTG GGCGAGATCG CGGCAGCCCA
1201  CGTCTCGGGA GGCCTGTCCC TGGCCGACGC CGCACGCGTG GTGACGCTTT
1251  GGAGCCAGGC ACAGACCACC CTTGCCGGGA CCGGCGCGCT CGTCTCCGTC
1301  GCCGCCACGC CGGATGAGCT CCTGCCCCGA ATCGCTCCGT GGACCGAGGA
1351  CAACCCGGCG CGGCTCGCCG TCGCAGCCGT CAACGGACCC CGGAGCACAG
1401  TCGTTTCCGG TGCCCGCGAG GCCGTCGCGG ACCTGGTGGC CGACCTCACC
1451  GCCGCGCAGG TGCGCACGCG CATGATCCCG GTGGACGTTC CCGCCCACTC
1501  CCCCCTGATG TACGCCATCG AGGAACGGGT CGTCAGCGGC CTGCTGCCCA
1551  TCACCCCACG CCCCTCCCGC ATCCCCTTCC ACTCCTCGGT GACCGGCGGC
1601  CGCCTCGACA CCCGCGAGCT AGACGCGGCG TACTGGTACC GCAACATGTC
1651  GAGCACGGTC CGGTTCGAGC CCGCCGCCCG GCTGCTTCTG CAGCAGGGGC
1701  CCAAGACGTT CGTCGAGATG AGCCCGCACC CGGTGCTGAC CATGGGCCTC
1751  CAGGAGCTCG CCGCGGACCT GGGCGACACC ACCGGCACCG CCGACACCGT
1801  GATCATGGGC ACGCTGCGCC GCGGCCAGGG CACCCTGGAC CACTTCCTGA
1851  CGTCTCTCGC CCAACTACGG GGGCATGGTG AGACGTCGGC GACCACCGTC
1901  CTCTCGGCAC GCCTGACCGC GCTGTCCCCC ACGCAGCAGC AGTCGCTGCT
1951  CCTGGACCTG GTGCGCGCCC ACACCATGGC GGTGCTGAAC GACGACGGAA
2001  ACGAGCGCAC CGCGTCGGAT GCCGGCCCAT CGGCGAGTTT CGCCCACCTC
2051  GGCTTCGACT CCGTCATGGG TGTCGAACTG CGCAACCGCC TCAGCAAGGC
2101  CACGGGCCTG CGGTTGCCCG TGACGCTCAT CTTCGACCAC ACCACGCCGG
2151  CCGCGGTCGC CGCGCGCCTT CGGACCGCGG CGCTCGGCCA CCTCGACGAG
2201  GACACCGCGC CCGTACCGGA CTCACCCAGC GGCCACGGAG GCACGGCAGC
2251  GGCGGACGAC CCGATCGCCA TCATCGGCAT GGCATGCCGT TTCCCGGGCG
```

Fig. 7b

```
2301  GAGTCCGGTC CCCGAAGGAC CTGTGGGAGC TGCCCGCCTC GGGCGGAGAC
2351  GCCATCGGGC CGTTCCCCAC CGACCGCGGA TGGCCCACGG AACAGCGTCA
2401  CGCCCAGGAC CCCACGCAGC CCGGCACGTT CTATCCGCAG GGAGGCGGGT
2451  TCCTTCACGA CGCGGCGCAC TTCGACGCCG GCTTCTTCGG AATCAGTCCA
2501  CGTGAGGCAC TGGCGATGGA TCCGCAGCAG CGGCTGCTGC TGGAGACGTC
2551  CTGGGAGGCG TTCGAGCGGG CGGGAATCGA TCCGCTGTCG GTACGCGGGT
2601  CCCGTACGGG CGTCTTCGCG GGCGCCCTCT CCTTCGACTA CGGCCCGCGT
2651  ATGGACACCG CGTCGTCGGA GGGCGCCGCG GACGTGGAGG GCCACATCCT
2701  CACCGGTACC ACGGGCAGCG TCCTGTCGGG CCGTATCGCC TACAGCTTCG
2751  GGCTGGAAGG GCCGGCGATC ACCGTGGACA CGGGGTGCTC GGCATCGCTC
2801  GTGACGCTGC ATCTGGCGTG CCAGTCGCTG CGGTCGGGTG AGTGCACGCT
2851  CGCGCTGGCC GGCGGCGTCT CGGTCATGTC CACCCTCGGC ATGTTCATCG
2901  AGTTCTCCCG GCAGCGCGGG CTGTCGGTGG ACGGCAGGTG CAAGGCGTAC
2951  TCGGCTGCAG CCGACGGCAC CGGCTGGGGC GAGGGCGTCG GGATGCTGTT
3001  GGTGGAGCGG TTGTCGGATG CGGTGCGGCT GGGGCATCGG GTGCTGGCGG
3051  TGGTACGCGG CAGTGCGGTC AACCAGGACG GTGCGTCGAA TGGGCTGACG
3101  GCGCCGAACG GTCCGGCTCA GGAGCGGGTG ATCCGGCAGG CGTTGGCGAA
3151  CGCGGGGTTG TCCGTGGCGG ATGTGGATGT GGTGGAGGGG CACGGGACGG
3201  GCACGACGCT GGGTGATCCG ATCGAGGCAC AGGCGTTGCT CGCCACGTAC
3251  GGGCAGCGGG CCGGTGACAG GCCGCTGTGG CTGGGGTCTC TGAAGTCCAA
3301  CATCGGGCAC ACCATGGCTG CCGCGGGTGT GGGTGGGGTC ATCAAGATGG
3351  TGATGGCGTT GCGGGAGGGG GTGTTGCCGC GGACGTTGCA TGTGGATGAG
3401  CCGTCGCCGC AGGTGGACTG GTCCGCGGGG GCGGTGCGGC TGCTGACGGA
3451  GGCGGTGCCG TGGCCGGGGG ACGCGGCAGG GCGGTTGCGG CGGGCGGGAG
3501  TGTCGTCGTT CGGGATCGGC GGCACGAATG CGCATGTGAT TTTGGAGGAG
3551  GCGCCGGCGG CGGGGGGCTG TGTTGCCGGG GGTGGGGTGT TGGAGGGTGC
3601  TCCGGGTCTT GCCATTTCGG TGGCTGAGTC GGTGGCCGCT CCAGTGGCTG
3651  TGTCTGCGCC GGTGGCTGAG TCGGTGCCGG TGCCGGTGCC GGTGCCGGTT
3701  CCTGTGCCGG TGTCGGCTAG GTCTGAGGCT GGGTTGCGGG CGCAGGCGGA
3751  GGCGTTGCGT CAGTACGTGG CAGTCCGGCC GGACGTTTCG CTTGCCGATG
3801  TGGGTGCGGG TCTGGCCTGT GGGCGGGCTG TGCTGGAGCA TCGTGCGGTC
3851  GTCCTGGCCG CGGACCGTGA GGAGCTGGTG CAAGGGTTGG GGGCGCTGGC
3901  GGCGGGTGAG CCGGATCGGC GGGTGACCAC GGGTCATGCG CCGGGTGGTG
3951  ACCGGGGCGG TGTCGTCTTC GTGTTTCCCG GACAGGGTGG GCAGTGGGCC
4001  GGGATGGGTG TGCGTCTGCT CGCCTCCTCT CCGGTGTTCG CCCGGCGGAT
4051  GCAGGCGTGC GAGGAGGCTC TGGCGCCGTG GGTGGACTGG TCTGTGGTGG
4101  ACATCCTGCG CCGGGACGCG GGGGATGCGG TGTGGGAGCG GGCCGATGTG
4151  GTCCAGCCTG TGCTGTTCAG CGTCATGGTG TCTTTGGCTG CTCTGTGGCG
4201  TTCCTACGGT ATCGAACCCG ACGCGGTCCT TGGCCATTCC CAGGGCGAGA
4251  TCGCGGCCGC GCATGTGTGT GGGGCGCTGA GCCTGAAGGA CGCGGCGAAG
4301  ACTGTTGCGC TGCGCAGCCG GGCGCTGGCC GCTGTGCGGG GCCGGGGCGG
4351  CATGGCCTCA GTGCCGCTGC CTGCCCAGGA GGTGGAGCAG CTCATTGGTG
4401  AGCGGTGGGC GGGGCGGTTG TGGGTGGCGG CGGTCAACGG CCCCCGCTCC
4451  ACCGCCGTCT CGGGGGATGC CGAGGCGGTG GACGAGGTGC TGGCGTACTG
4501  TGCCGGCACC GGGGTGCGGG CCCGGCGGAT CCCGGTCGAC TATGCCTCGC
4551  ACTGCCCCCA TGTGCAGCCC CTGCGGGAGG AGTTGCTGGA GCTGCTGGGG
```

Fig. 7c

```
4601  GACATCAGCC CGCAGCCGTC CGGCGTGCCG TTCTTCTCCA CGGTGGAGGG
4651  CACCTGGCTG GACACCACAA CCCTGGACGC CGCCTACTGG TACCGCAACC
4701  TGCACCAGCC GGTCCGTTTC AGCGATGCCG TCCAGGCCCT GGCGGATGAC
4751  GGACACCGCG TCTTCGTCGA AGTCAGCCCC CACCCCACCC TCGTCCCCGC
4801  CATCGAAGAC ACCACCGAAG ACACCGCCGA AGACGTCACC GCGATCGGCA
4851  GCCTCCGCCG CGGCGACAAC GACACCCGCC GCTTCCTCAC CGCCCTCGCC
4901  CACACCCATA CCACCGGCAT CGGCACACCC ACCACCTGGC ACCACCACTA
4951  CACCCACCAC CACACCCACC CCCACCCCCA CACGCACCTC GACCTGCCCA
5001  CCTACCCCTT CCAACACCAG CACTACTGGC TCGAGAGCTC ACAGCCGGGT
5051  GCCGGATCCG GTTCGGGTGC CGGTGCCGGT TCGGGTGCCG GTTCCGGGCG
5101  GGCAGGGACT GCGGGCGGGA CGGCAGAGGT GGAGTCGCGG TTCTGGGACG
5151  CGGTGGCCCG CCAGGACCTG GAAACGGTCG CGACCACACT CGCCGTGCCC
5201  CCCTCCGCCG GCCTGGACAC GGTGGTGCCC GCACTCTCCG CCTGGCACCG
5251  CCACCAACAC GACCAAGCCC GCATCAACAC CTGGACCTAC CAGGAAACCT
5301  GGAAACCCCT CACCCTCCCC ACCACCCACC AACCCCACCA AACCTGGCTC
5351  ATCGCCATCC CCGAAACCCA GACCCACCAC CCCCACATCA CCAACATCCT
5401  CACCAACCTC CACCACCACG GCATCACCCC CATCCCCCTC ACCCTCAACC
5451  ACACCCACAC CAACCCCCAA CACCTCCACC ACACCCTCCA CCACACCCGA
5501  CAACAAGCCC AAAACCACAC CACCGGAGCC ATCACCGGCC TGCTCTCCCT
5551  CCTCGCCCTC GACGAAACAC CCCACCCCCA CCACCCCCAC ACACCCACCG
5601  GCACCCTCCT CAACCTCACC CTCACCCAAA CCCACACCCA AACCCACCCA
5651  CCAACCCCCC TCTGGTACGC CACCACCAAC GCCACCACCA CCCACCCCAA
5701  CGACCCCCTC ACACACCCCA CCCAAGCCCA AACCTGGGGA CTCGCCCGCA
5751  CCACCCTCCT CGAACACCCC ACCCACACCG CCGGAATCAT CGACCTCCCC
5801  ACCACCCCCA CCCCCCACAC CCTCCACCAC CTCACCCAAA CCCTCACCCA
5851  ACCCCACCAC CAAACCCAAC TCGCCATCCG CACCACCGGC ACCCACACCC
5901  GCCGCCTCAC CCCCACCACC CTCACCCCCA CACACCAACC ACCCACCCCC
5951  ACCCCCCACG GAACCACCCT CATCACCGGC GGAACCGGCG CCCTCGCCAC
6001  CCACCTCACC CACCACCTCA CCACCCACCA ACCCACCCAA CACCTCCTCC
6051  TCACCAGCCG AACCGGCCCC CACACCCCCC ACGCACAACA CCTCACCACC
6101  CAACTCCAAC AAAAAGGCAT CCACCTCACC ATCACCACCT GCGACACCAG
6151  CAACCCAGAC CAACTCCAAC ATCTCCTCAA CACCATCCCC CCACAACACC
6201  CCCTCACCAC CGTCATCCAC ACCGCAGGCA TCCTCGACGA CGCCACCCTC
6251  ACCAACCTCA CCCCCACCCA ACTCAACAAC GTCCTCCGCG CCAAAGCCCA
6301  CAGCGCCCAC CTCCTCCACC AACTCACCCA ACACACCCCC CTCACCGCCT
6351  TCGTCCTCTA CTCCTCCGCC GCCGCCACCT TCGGCGCACC CGGCCAAGCC
6401  AACTACGCCG CAGCCAACGC CTACCTCGAC GCCCTCGCCC ACCACCGCCA
6451  CACCCACCAC CTCCCCGCCA CCAGCATCGC CTGGGGCACC TGGCAAGGAA
6501  ACGGACTCGC TGATTCGGAC AAGGCCCGCG CATATCTCGA CCGCCGCGGG
6551  TTTCGACCCA TGTCACCCGA GTTGGCCACG GCAGCGGTCA CGCAGGCGAT
6601  CGCGGACACC GAACGGCCGT ATGTCGTCAT CGCCGACATC GACTGGAGCA
6651  AGATCGAACA CACCTCTCAG ACCAGCGACC TGGTGAGCGC GGCCCGGGAA
6701  AGGGAGCCAG CTGTCCAGCG CCCCACTCCA CCGGCGGAGT TGCACAAAAC
6751  GCTGGCCCAT CAGACGTCGG CCGACCAACG GGCCGCATTG CTCGAGCTCG
6801  TACGAGACCA TGTGGCGGCA GTGCTCCGGC ACGCGGACCC GAAAGCCATC
6851  GCGCCCGACC AGTCGTTCCG TGCACTCGGC TTCGATTCAC TCACGGCCGT
```

Fig. 7d

```
6901  CGAGTTCCGA AACCTGCTGA TCAAGGCAAC AGGACTCCGC CTTCCTGTCT
6951  CGCTGGTCTT CGACCACCCG ACCCCTGCCA AACTCGCCGT ACACCTGCAG
7001  AACCAACTGC GGGGCACAGC AGCGGAGTCG GCTCCTTCAG CGGCAGCCGT
7051  TACCGCCGAG GCTTCTGTCA CCGAGCCGAT CGCCATCGTT GGCATGGCCT
7101  GTCGTTTCCC CGGCGGAGTG ACCTCGGCGG ACGACTTCTG GGATCTGATC
7151  TCCTCCGAGC AGGACGCGAT CGGCGGATTC CCCACCGACC GCGGCTGGGA
7201  CCTGGACACG CTCTACGACC CCGACCCCGA CCACCCCGGC ACCTGCTACA
7251  CCCGAAACGG CGGATTCCTC TACGACGCAG GCCACTTCGA CGCCGAATTC
7301  TTCGGCATCA GCCCCCGCGA AGCCCTCGCC ATGGACCCCC AGCAACGACT
7351  CCTCCTCGAA ACCGCCTGGG AAACCATCGA ACACGCCGGC ATCAACCCCC
7401  ACACCCTCCA CGGCACCCCC ACCGGAGTCT TCACCGGCAC CAACGGACAG
7451  GACTACGCAC TTCGCGTGCA CAACGCGGGC CAGTCAACCG ATGGTTTCGC
7501  ACTGACCGGA ACCGCCGGCA GCGTCATCTC CGGTCGTATC TCGTACACGT
7551  TTGGTTTTGA GGGTCCTGCG GTGTCGGTGG ACACGGCTTG TTCCTCGTCG
7601  TTGGTGGCTT TGCATCTGGC CTGTCAGGCG TTGCGTGCGG GTGAGTGCTC
7651  GATGGCGCTT GCCGGGGGTG TGACGGTGAT GTCGTCTCCG GGTGCCTTCG
7701  TGGAGTTTTC GCGGCAGCGG GGTCTGGCCG CGGACGGGCA TTGCAAGGCG
7751  TTCTCGGCGG CGGCGGACGG GACCGGCTGG GGTGAGGGTG TGGGGATGCT
7801  GCTGGTGGAG CGGCTCTCCG ACGCCCATCG CAACGGTCAC CGTGTCCTGG
7851  CCGTGGTGCG TGGCAGTGCG GTCAACCAGG ACGGTGCGAG CAACGGTCTG
7901  ACCGCGCCCA ACGGGCCGTC CCAGCAGCGT GTCATCCGCC AGGCCCTCGC
7951  CAACGCCGGC TTGTCGGCCG GTGATGTCGA CGCGGTGGAG GCCCACGGCA
8001  CCGGCACCAC TTTGGGCGAC CCGATCGAGG CCCAGGCCCT CCTCGCGACC
8051  TACGGACAGG ACCGTGCCGG CGAGGGGCCG CTGTGGCTGG GCTCGGTCAA
8101  GTCCAATGTC GGTCACACAC AGGCTGCCGC GGGCGTCGCC GGGGTGATCA
8151  AGATGGTGAT GGCGCTGCGG CATGGTCTGC TGCCGCGGAC GTTGCATGTG
8201  GATGAGCCGT CGCCGCATGT GGACTGGTCC GCGGGTGCGG TGCAGCTGCT
8251  GACGGAGACG GTGCCCTGGC CCGGCGGGGA GGGGCGGCTA CGGCGGGCAG
8301  GAGTGTCATC ATTCGGCGTC AGCGGCACCA ACGCCCACGT CATCCTCGAA
8351  GAAGCACCCG CCGACGACGT TCCGGGGGGA CCACCCGCCG GCGAGGGTGA
8401  CGCGGGCAGC GACGATGAGG CTGCTGCCGG CAGTCCTGGG GTGTGGCCGT
8451  GGCTGGTGTC GGCCAAGTCG CAGCCGGCCC TGCGCGCCCA GGCCCAGGCC
8501  CTGCACGCCC ACCTCACCGA CCACCCCGGC CTCGACCTCG CGGATGTCGG
8551  ATACACCCTC GCCCACGCCC GCGCCGTGTT CGACCACCGC GCCACCCTCA
8601  TCGCCGCGGA CCGCGACACG TTCCTGCAAG CACTCCAGGC ACTCGCCGCA
8651  GGCGAGCCCC ACCCCGCCGT CATCCACAGC AGCGCCCCGG GCGGGACCGG
8701  GACCGGGGAG GCCGCAGGAA AGACCGCATT CATCTGCTCC GGACAGGGCA
8751  CCCAACGCCC CGGCATGGCC CACGGCCTCT ACCACACCCA CCCCGTCTTC
8801  GCCGCCGCAC TCAACGACAT CTGCACCCAC CTCGACCCCC ACCTCGACCA
8851  CCCCCTCCTC CCCCTCCTCA CCCAAAACGA CAACGACAAC GAGGACGCGG
8901  CCGCACTGCT CCAGCAGACC CGCTACGCCC AGCCCGCCCT CTTCGCCTTC
8951  CAGGTCGCCC TCCACCGCCT CCTCACCGAC GGCTACCACA TCACCCCCCA
9001  CTACTACGCC GGACACTCCC TCGGCGAAAT CACCGCCGCC CACCTCGCCG
9051  GCATCCTCAC CCTCACCGAC GCCACCACCC TCATCACCCA ACGCGCCACC
9101  CTCATGCAAA CCATGCCCCC CGGCACCATG ACCACCCTCC ACACCACCCC
9151  CCACCACATC ACCCACCACC TCACCGCCCA CGAAAACGAC CTCGCCATCG
```

Fig. 7e

```
 9201  CCGCCATCAA CACCCCCACC TCCCTCGTCA TCAGCGGCAC CCCCCACACC
 9251  GTCCAACACA TCACCACCCT CTGCCAACAA CAAGGCATCA AAACCAAAAC
 9301  CCTCCCCACC AACCACGCCT TCCACTCCCC CCACACCAAC CCCATCCTCA
 9351  ACCAACTCCA CCAGCACACC CAAACCCTCA CCTACCACCC ACCCCACACC
 9401  CCCCTCATCA CCGACAACAC CCCACCCGAC CAACTCCTCA CCCCCCACTA
 9451  CTGGACCCAA CAAGCCCGCA ACACCGTCGA CTACGCCACC ACCACCCAAA
 9501  CCCTCCACCA ACACGGCGTC ACCACCTACA TCGATCTCGG ACCCGACAAC
 9551  ACCCTCACCA CCCTCACCCA CCACAACCTC CCCAACACCC CCACCACCAC
 9601  CCTCACCCTC ACCCACCCCC ACCACCACCC CCAAACCCAC CTCCTCACCA
 9651  ACCTCGCCAA AACCACCACC ACCTGGCACC CCCACCACTA CACCCACCAC
 9701  CACAACCAAC CCCACACCCA CACCCACCTC GACCTCCCCA CCTACCCCTT
 9751  CCAACACCAC CACTACTGGC TCGAAAGCAC ACAGCCCGGT GCCGGCAACG
 9801  TGTCAGCAGC CGGACTCGAC CCCACCGAAC ACCCCCTACT CGGCGCCACA
 9851  TTGGAACTGG CGACTGACGG TGGAGCGCTT CTTGCAGGGC GCTTGTCTTT
 9901  GAGGTCGCAT CCGTGGCTGG CTGACCATGC CGTCGGCGGC ACGGTGCTGC
 9951  TGTCGGGCGC CACCTTCCTC GAACTCGCCC TTCATGCGGG CACATACGTG
10001  GGCTGCGACC GAGTGGATGA GCTGACGCTG CATGCGCCGC TGGTGGTTCC
10051  TGTGGATGGG GGTGTGAGTG TGCAGGTTGG GGTTGCGGCT GCGGATGGGG
10101  AGGGGCGGCG TTTGGTGAGT GTGTATGCGC GGGGTGGGAG TGCTTGTGGT
10151  GGGGGTGGTG CGTCGGGTGG GGTGTGGACG TGTCATGCCT CGGGGGTGCT
10201  GGTTGAGGCT GCTGCTGGTG GTGTGGTGGT GGATGGTCTG GCGGGGGTGT
10251  GGCCGCCGCG GGGTGCGGTG GCGGTGGATG TCGATGGTGT CCGTGACCGT
10301  TTGGCTGGGG CTGGTTGTGT TTTGGGGCCG GTGTTTTCGG GGCTGCGTGC
10351  GGTGTGGCGT GATGGGGGGG ATTTGCTGGC TGAGGTGTGT CTGCCGGAGG
10401  AGGCGTGGGG TGATGCGGCT GGTTTTGGGC TGCATCCGGC GTTGCTGGAT
10451  GGTGTGGTCC AGCCGTTGTC GGTGTTGCTT CCGGGTGGGA CGGGGTTTGG
10501  GGAGGGGGCG GGGTTCGGGG AGGGTGTTCG GGTGCCGGCT GTGTGGGGTG
10551  GTGTGTCGCT TCACCGGGCG GGTGTGACCG GTGTGCGGGT GCGTGTGTCG
10601  GCTGTCGGGC GGGGCGGCGG GCGTGAGGCG GTGTCGGTCG TGGTCGGGGA
10651  TGAGGCGGGT GTGCCGGTGG CGTCGGTCGA TCGTCTTGAG TTGCGGCCTG
10701  TGGATATGGG TCAGTTGCGT GCTGTCTCGG TTTCGGCGGG GCGGCGGGGT
10751  TCGCTGTATG CGGTGCAGTG GGCTGAGGTG GGTCCTGTGC CGGTGTGTGG
10801  GCAGGCGTGG GCGTGGCACG AGGACGTGGG TGAGAGCGGT GGTGGGCCTG
10851  TGCCGGGGGT GGTGGTGTTG CGGTGCCCGG ATGCCGGTGC CGGTGGCGGT
10901  GGCGGTGGCG GTGGTGGCGG TGGTGTGGGT GAGGTTGTTG GTGGGGTGTT
10951  GGGTGTGGTG CAGGGGTGGC TGGGGCTGGA GCGGTTTGCG GGTTCGCGGC
11001  TGGTGGTGGT GACCCGGGGT GCGGTGGTGG CCGGCCCGGA GGACGGCCCG
11051  GTGGATGTGG TGGGTGCGTC GGTGTGGGGG CTGGTGCGTT CGGCGCAGGC
11101  TGAGCATCCG GACCGGTTTG TCCTCCTCGA CCTCGACACC GACACCGGCA
11151  CCGACCTCGA CACCGGTGCT GGTGCTGGTT GGGGCGTGGA TGGTGGGCGT
11201  GTGGCGGCGG TGGTGGCGTG TGGTGAGCCG CAGTTGGCGG TGCGTGGGGA
11251  GCGGTTGCTG GCCGCACGCC TGACACGACT TGAGTCATCC GGTGATGTTC
11301  CAGCCCAGCG GTCCGGTGAC ACACGAGCCC GGCGGTCCGA CGTGCCTGCC
11351  CAGCGCTCCG GTGGCGTGCC TGCTCGGCGG TCGGTTGATG TATCGGGTCG
11401  GGAGGTGTTG CCGTGGTTGT CGGGTGGGTC GGTGTTGGTG ACGGGTGGGA
11451  CGGGTGTGCT GGGTGCGGCG GTGGCGCGGC ATCTGGCTGG TGTGTGTGGG
```

Fig. 7f

```
11501  GTGCGGGATC TGCTGTTGGT GAGCCGGCGT GGTCCGGATG CTCCGGGTGC
11551  GGAGGGTCTG CGGGCGGAGC TGGCCGCGTT GGGGGCGGAG GTGCGGATTG
11601  TTGCGTGTGA TGTGGGGGAG CGGCGGGAGG TGGTCCGGCT GCTGGAGGGT
11651  GTTCCTGCCG GGTGTCCGCT GACGGGTGTC GTGCATGCGG CTGGTGTGCT
11701  GGACGATGCG ACGATCGCCT CTCTCACGCC CGAGCGGCTG GGCACGGTGT
11751  TCGCGGCCAA GGTGGATGCC GCTCTTTTGC TGGATGAGCT GACGCGGGGT
11801  ATGGAGCTGT CGGCGTTCGT GCTGTTCTCC TCGGCCGCGG GGATCCTGGG
11851  GTCGGCCGGG CAGGGCAACT ACGCCGCGGC CAATGCCGCT CTGGACGCGC
11901  TGGCGTACCG GCGGCGGGCG GCGGGTCTGC CGGGGGTGTC GCTGGCGTGG
11951  GGGCTGTGGG AAGAGGCCAG CGGGATGACC GGGCACCTGG CCGGCACCGA
12001  CCACCGGCGC ATCATCCGTT CCGGTCTGCA TCCCATGTCG ACCCCGGACG
12051  CACTGGCCCT CTTCGATGCG GCCCTGGCTC TGGACCGGCC GGTCCTGCTG
12101  CCCGCCGACC TGCGTCCCGC CCCGCCCCTG CCGCCCCTGC TGCAGGACCT
12151  CCTGCCCGCC ACCCGCCGCC GCACCACCCG CACCACCACT ACCGGTGGTG
12201  CGGACAACGG CGCCCAGCTG CACGCCCGGC TGGCCGGCCA GACACACGAA
12251  CAACAGCACA CCACCCTCCT CGCCCTGGTC CGCTCCCACA TCGCCACCGT
12301  CCTGGGCCAC ACCACCCCCG ACACCATCCC CCCCGACCGC GCGTTCCGCG
12351  ACCTCGGCTT CGACTCCCTC ACCGCCGTCG A
```