

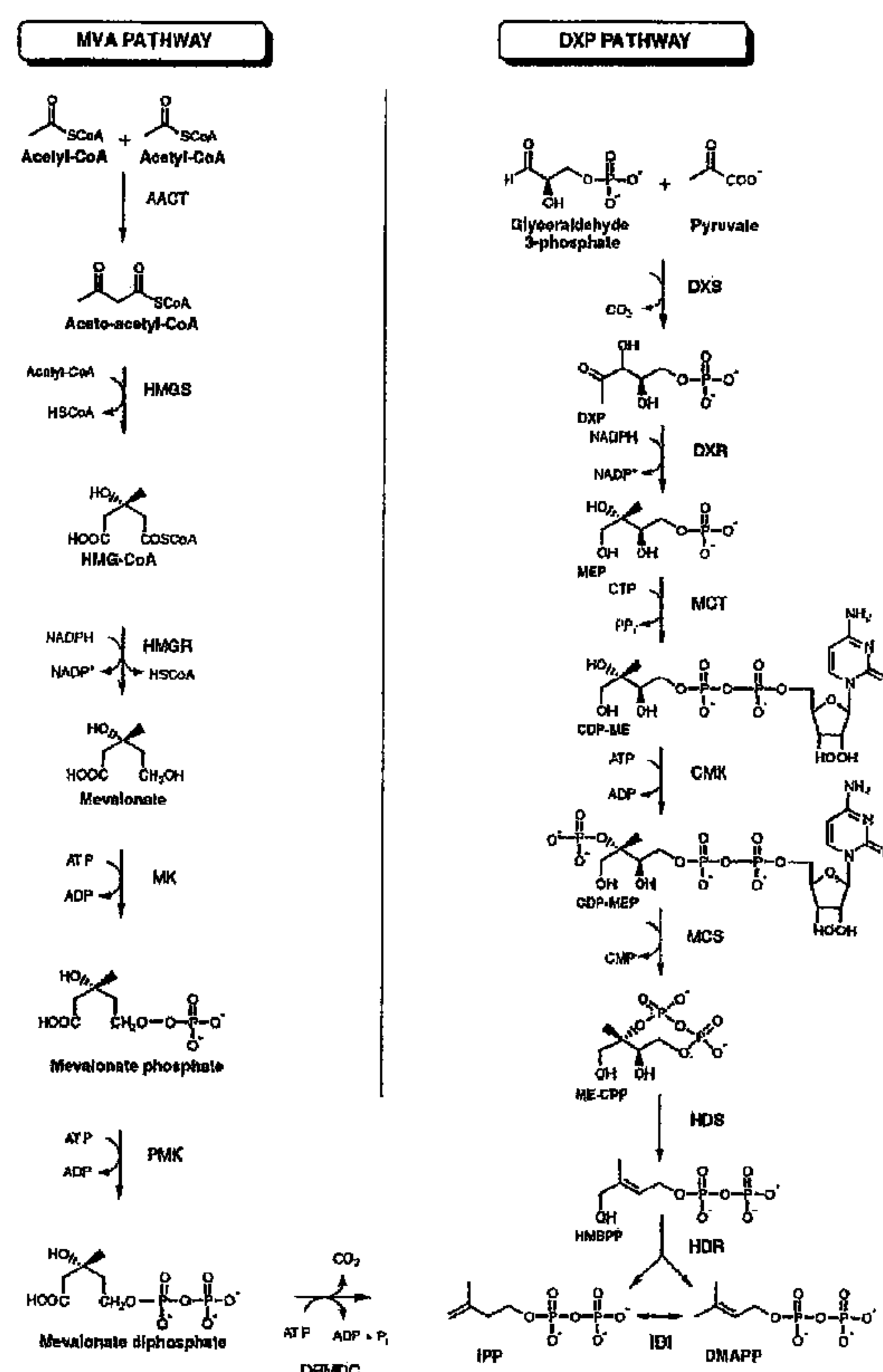


(86) Date de dépôt PCT/PCT Filing Date: 2008/12/15  
(87) Date publication PCT/PCT Publication Date: 2009/06/18  
(85) Entrée phase nationale/National Entry: 2010/06/11  
(86) N° demande PCT/PCT Application No.: US 2008/086869  
(87) N° publication PCT/PCT Publication No.: 2009/076676  
(30) Priorité/Priority: 2007/12/13 (US61/013,574)

(51) Cl.Int./Int.Cl. *C12N 1/21* (2006.01),  
*C12N 1/15* (2006.01), *C12N 1/19* (2006.01),  
*C12N 1/20* (2006.01), *C12N 15/00* (2006.01),  
*C12N 15/52* (2006.01), *C12N 9/00* (2006.01),  
*C12P 5/02* (2006.01)  
(71) Demandeurs/Applicants:  
DANISCO US INC., US;  
THE GOODYEAR TIRE & RUBBER COMPANY, US  
(72) Inventeurs/Inventors:  
CERVIN, MARGUERITE, US;  
WHITED, GREGORY M., US;  
CHOTANI, GOPAL K., US;  
VALLE, FERNANDO, US; ...

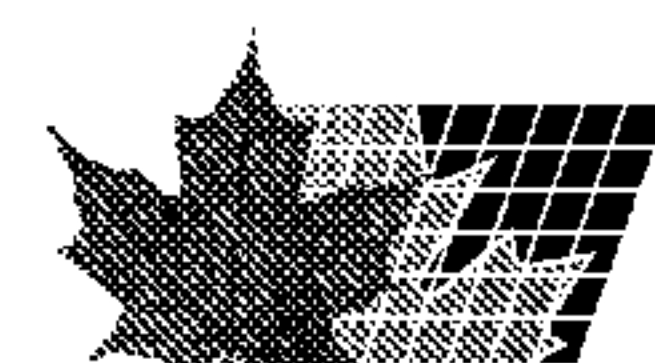
(54) Titre : COMPOSITIONS ET METHODES DE PRODUCTION D'ISOPRENE  
(54) Title: COMPOSITIONS AND METHODS FOR PRODUCING ISOPRENE

Figure 19A



(57) Abrégé/Abstract:

The invention features methods for producing isoprene from cultured cells. The invention also provides compositions that include these cultured cells.





(72) **Inventeurs(suite)/Inventors(continued):** FIORESI, CAROL, US; SANFORD, KARL J., US; MCAULIFFE, JOSEPH C., US; FEHER, FRANK J., US; PUHALA, AARON S., US; MIASNIKOV, ANDREI, US; ALDOR, ILANA S., US

(74) **Agent:** BERESKIN & PARR LLP/S.E.N.C.R.L.,S.R.L.



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

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
18 June 2009 (18.06.2009)

PCT

(10) International Publication Number  
**WO 2009/076676 A2**

## (51) International Patent Classification:

*C12P 5/02* (2006.01) *C12N 1/21* (2006.01)  
*C12N 1/20* (2006.01)

## (21) International Application Number:

PCT/US2008/086869

## (22) International Filing Date:

15 December 2008 (15.12.2008)

## (25) Filing Language:

English

## (26) Publication Language:

English

## (30) Priority Data:

61/013,574 13 December 2007 (13.12.2007) US

## (71) Applicants (for all designated States except US):

**GENENCOR INTERNATIONAL, INC.** [US/US]; 925  
Page Mill Road, Palo Alto, CA 94304-1013 (US). **THE**  
**GOODYEAR TIRE & RUBBER COMPANY** [US/US];  
1144 East Market Street, Akron, OH 44316-0001 (US).

## (72) Inventors; and

(75) Inventors/Applicants (for US only): **CERVIN, Mar-**  
**guerite** [CA/US]; 404 Palomar Drive, Redwood City, CA

94062 (US). **WHITED, Gregory, M.** [US/US]; 304 South  
Road, Belmont, CA 94002 (US). **CHOTANI, Gopal, K.**  
[IN/US]; 10166 Carol Lee Drive, Cupertino, CA 95014  
(US). **VALLE, Fernando** [MX/US]; 232 Bancroft Road,  
Burlingame, CA 94010 (US). **FIORESI, Carol** [US/US];  
408 Northumberland Avenue, Redwood City, CA 94061  
(US). **SANFORD, Karl, J.** [US/US]; 10434 Alicia Court,  
Cupertino, CA 95014 (US). **MCAULIFFE, Joseph, C.**  
[AU/US]; 663 Madrone Avenue, Sunnyvale, CA 94085  
(US). **FEHER, Frank, J.** [US/US]; 4305 Sierra Drive,  
Copley, OH 44321 (US). **PUHALA, Aaron, S.** [US/US];  
923 Catlin Court, Kent, OH 44242 (US). **MIASNIKOV,**  
**Andrei** [FI/US]; 1769 Latham Street, #A, Mountain View,  
CA 94041 (US). **ALDOR, Ilana, S.** [CA/US]; 220 College  
Avenue, Apt. D, Palo Alto, CA 94306 (US).

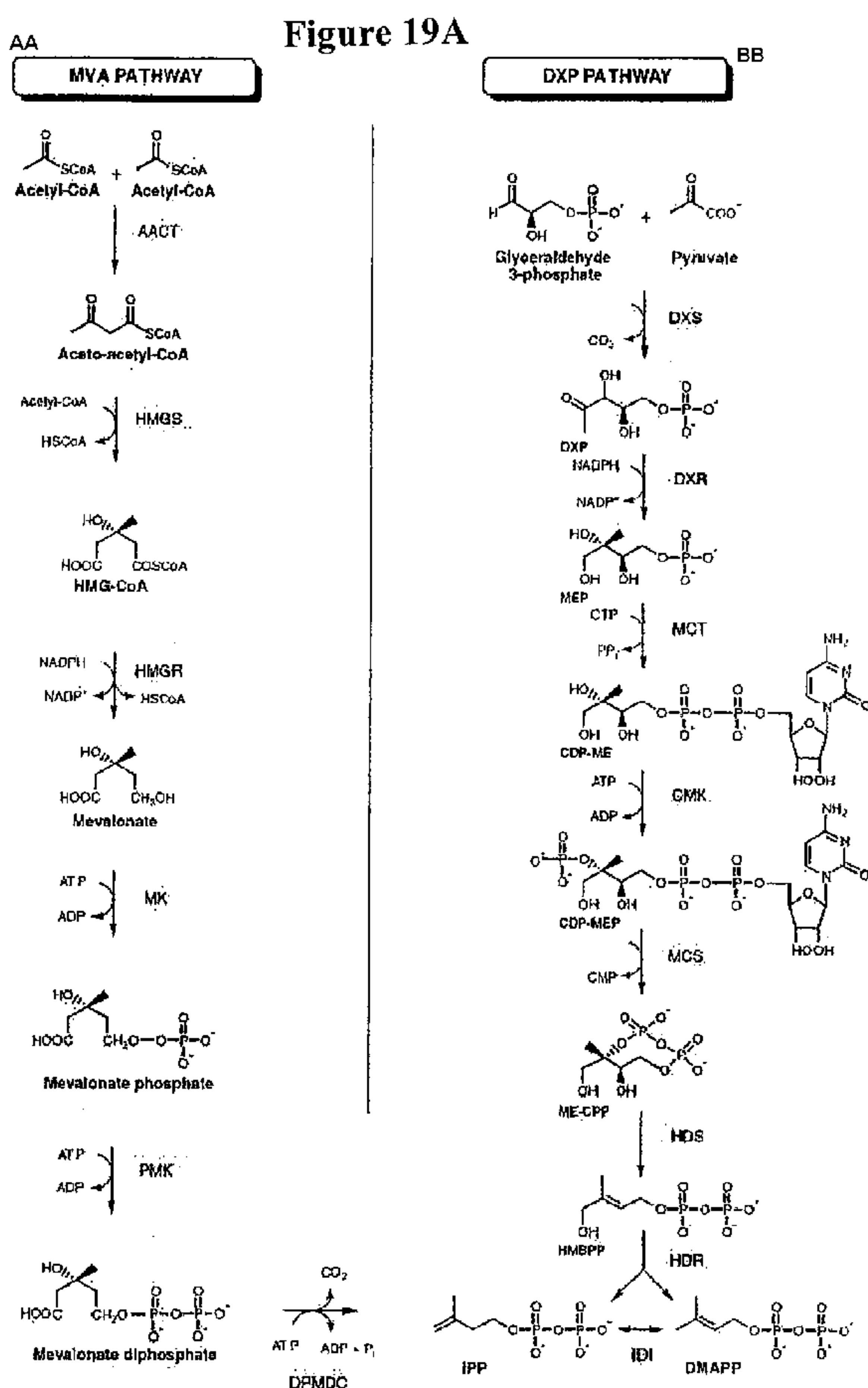
(74) Agent: **YEH, Jenny, J.**; Morrison & Foerster LLP, 755  
Page Mill Road, Palo Alto, CA 94304-1018 (US).

(81) Designated States (unless otherwise indicated, for every  
kind of national protection available): AE, AG, AL, AM,  
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA,

[Continued on next page]

## (54) Title: COMPOSITIONS AND METHODS FOR PRODUCING ISOPRENE

(57) Abstract: The invention features methods for producing  
isoprene from cultured cells. The invention also provides com-  
positions that include these cultured cells.





WO 2009/076676 A2



CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH,

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



## COMPOSITIONS AND METHODS FOR PRODUCING ISOPRENE

### CROSS-REFERENCE TO RELATED APPLICATIONS

**[0001]** This application claims priority benefit of United States Provisional patent application No. 61/013,574, filed on December 13, 2007. The content of that application is hereby incorporated by reference herein in its entirety.

### FIELD OF THE INVENTION

**[0002]** The present invention relates generally to methods for producing isoprene from cultured cells and compositions that include these cultured cells.

### BACKGROUND OF THE INVENTION

**[0003]** Isoprene (2-methyl-1,3-butadiene) is the critical starting material for a variety of synthetic polymers, most notably synthetic rubbers. Isoprene is naturally produced by a variety of microbial, plant, and animal species. In particular, two pathways have been identified for the biosynthesis of isoprene: the mevalonate (MVA) pathway and the non-mevalonate (DXP) pathway (Figures 19A and 19B). However, the yield of isoprene from naturally-occurring organisms is commercially unattractive. About 800,000 tons per year of *cis*-polyisoprene are produced from the polymerization of isoprene; most of this polyisoprene is used in the tire and rubber industry. Isoprene is also copolymerized for use as a synthetic elastomer in other products such as footwear, mechanical products, medical products, sporting goods, and latex.

**[0004]** Currently, the tire and rubber industry is based on the use of natural and synthetic rubber. Natural rubber is obtained from the milky juice of rubber trees or plants found in the rainforests of Africa. Synthetic rubber is based primarily on butadiene polymers. For these polymers, butadiene is obtained as a co-product from ethylene and propylene manufacture.

**[0005]** While isoprene can be obtained by fractionating petroleum, the purification of this material is expensive and time-consuming. Petroleum cracking of the C5 stream of hydrocarbons produces only about 15% isoprene. Thus, more economical methods for producing isoprene are needed. In particular, methods that produce isoprene at rates, titers, and



purity that are sufficient to meet the demands of a robust commercial process are desirable. Also desired are systems for producing isoprene from inexpensive starting materials.

#### BRIEF SUMMARY OF THE INVENTION

**[0006]** In one aspect, the invention features cells in culture that produce isoprene. In some embodiments, the invention provides cells in culture that produce greater than about 400 nmole of isoprene/gram of cells for the wet weight of the cells/hour (nmole/g<sub>wcm</sub>/hr) of isoprene. In some embodiments, the cells have a heterologous nucleic acid that (i) encodes an isoprene synthase polypeptide and (ii) is operably linked to a promoter. In some embodiments, the cells are cultured in a culture medium that includes a carbon source, such as, but not limited to, a carbohydrate (*e.g.*, xylose or glucose), acetate, glycerol, glycerine, dihydroxyacetone, one-carbon source, oil, animal fat, animal oil, fatty acid, lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, renewable carbon source (*e.g.*, a hydrolyzed biomass carbon source), polypeptide (*e.g.*, a microbial or plant protein or peptide), yeast extract, component from a yeast extract, or any combination of two or more of the foregoing.

**[0007]** In some embodiments, the invention provides cells in culture that have an average volumetric productivity of isoprene greater than about 0.1 mg/L<sub>broth</sub>/hr. In some embodiments, the invention provides cells in culture that have a peak volumetric productivity of isoprene greater than about 0.5 mg/L<sub>broth</sub>/hr. In some embodiments, the cells have a heterologous nucleic acid that (i) encodes an isoprene synthase polypeptide and (ii) is operably linked to a promoter. In some embodiments, the cells are cultured in a culture medium that includes a carbon source, such as, but not limited to, a carbohydrate (*e.g.*, xylose or glucose), acetate, glycerol, glycerine, dihydroxyacetone, one-carbon source, oil, animal fat, animal oil, fatty acid, lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, renewable carbon source (*e.g.*, a hydrolyzed biomass carbon source), polypeptide (*e.g.*, a microbial or plant protein or peptide), yeast extract, component from a yeast extract, or any combination of two or more of the foregoing.

**[0008]** In some embodiments, the invention provides cells in culture that convert more than about 0.002% of the carbon in a cell culture medium into isoprene. In some embodiments, the cells have a heterologous nucleic acid that (i) encodes an isoprene synthase polypeptide and (ii)



is operably linked to a promoter. In some embodiments, the cells are cultured in a culture medium that includes a carbon source, such as, but not limited to, a carbohydrate (*e.g.*, xylose or glucose), acetate, glycerol, glycerine, dihydroxyacetone, one-carbon source, oil, animal fat, animal oil, fatty acid, lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, renewable carbon source (*e.g.*, a hydrolyzed biomass carbon source), polypeptide (*e.g.*, a microbial or plant protein or peptide), yeast extract, component from a yeast extract, or any combination of two or more of the foregoing.

**[0009]** In some embodiments, the invention provides cells in culture that comprise a heterologous nucleic acid encoding an isoprene synthase polypeptide. In some embodiments, the cells have a heterologous nucleic acid that (i) encodes an isoprene synthase polypeptide and (ii) is operably linked to a promoter. In some embodiments, the cells are cultured in a culture medium that includes a carbon source, such as, but not limited to, a carbohydrate (*e.g.*, xylose or glucose), acetate, glycerol, glycerine, dihydroxyacetone, one-carbon source, oil, animal fat, animal oil, fatty acid, lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, renewable carbon source (*e.g.*, a hydrolyzed biomass carbon source), polypeptide (*e.g.*, a microbial or plant protein or peptide), yeast extract, component from a yeast extract, or any combination of two or more of the foregoing.

**[0010]** In one aspect, the invention features methods of producing isoprene, such as methods of using any of the cells described herein to produce isoprene. In some embodiments, the method involves culturing cells under conditions sufficient to produce greater than about 400 nmole/g<sub>wcm</sub>/hr of isoprene. In some embodiments, the method also includes recovering isoprene produced by the cells. In some embodiments, the method includes purifying isoprene produced by the cells. In some embodiments, the method includes polymerizing the isoprene. In some embodiments, the cells have a heterologous nucleic acid that (i) encodes an isoprene synthase polypeptide and (ii) is operably linked to a promoter. In some embodiments, the cells are cultured in a culture medium that includes a carbon source, such as, but not limited to, a carbohydrate (*e.g.*, xylose or glucose), acetate, glycerol, glycerine, dihydroxyacetone, one-carbon source, oil, animal fat, animal oil, fatty acid, lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, renewable carbon source (*e.g.*, a hydrolyzed biomass carbon source),



polypeptide (*e.g.*, a microbial or plant protein or peptide), yeast extract, component from a yeast extract, or any combination of two or more of the foregoing.

**[0011]** In one aspect, the invention features methods of producing isoprene, such as methods of using any of the cells described herein to produce isoprene. In some embodiments, the method involves culturing cells under conditions resulting in an average volumetric productivity of isoprene greater than about 0.1 mg/L<sub>broth</sub>/hr. In some embodiments, the method involves culturing cells under conditions resulting in a peak volumetric productivity of isoprene greater than about 0.5 mg/L<sub>broth</sub>/hr. In some embodiments, the method also includes recovering isoprene produced by the cells. In some embodiments, the method includes purifying isoprene produced by the cells. In some embodiments, the method includes polymerizing the isoprene. In some embodiments, the cells have a heterologous nucleic acid that (i) encodes an isoprene synthase polypeptide and (ii) is operably linked to a promoter. In some embodiments, the cells are cultured in a culture medium that includes a carbon source, such as, but not limited to, a carbohydrate (*e.g.*, xylose or glucose), acetate, glycerol, glycerine, dihydroxyacetone, one-carbon source, oil, animal fat, animal oil, fatty acid, lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, renewable carbon source (*e.g.*, a hydrolyzed biomass carbon source), polypeptide (*e.g.*, a microbial or plant protein or peptide), yeast extract, component from a yeast extract, or any combination of two or more of the foregoing.

**[0012]** In some embodiments, the method includes culturing cells under conditions sufficient to convert more than about 0.002% of the carbon in a cell culture medium into isoprene. In some embodiments, the method also includes recovering isoprene produced by the cells. In some embodiments, the method includes purifying isoprene produced by the cells. In some embodiments, the method includes polymerizing the isoprene. In some embodiments, the cells have a heterologous nucleic acid that (i) encodes an isoprene synthase polypeptide and (ii) is operably linked to a promoter. In some embodiments, the cells are cultured in a culture medium that includes a carbon source, such as, but not limited to, a carbohydrate (*e.g.*, xylose or glucose), acetate, glycerol, glycerine, dihydroxyacetone, one-carbon source, oil, animal fat, animal oil, fatty acid, lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, renewable carbon source (*e.g.*, a hydrolyzed biomass carbon source), polypeptide (*e.g.*, a



microbial or plant protein or peptide), yeast extract, component from a yeast extract, or any combination of two or more of the foregoing.

**[0013]** In some embodiments of any of the aspects of the invention, the cells in culture produce isoprene at greater than or about 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, 10,000, 12,500, 20,000, 30,000, 40,000, 50,000, 75,000, 100,000, 125,000, 150,000, 188,000, or more nmole/g<sub>wcm</sub>/hr isoprene. In some embodiments, the cells in culture have an average volumetric productivity of isoprene at greater than or about 0.1, 1.0, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1100, 1200, 1300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,100, 2,200, 2,300, 2,400, 2,500, 2,600, 2,700, 2,800, 2,900, 3,000, 3,100, 3,200, 3,300, 3,400, 3,500, or more mg of isoprene/L of broth/hr (mg/L<sub>broth</sub>/hr, wherein the volume of broth includes the volume of the cells and the cell medium). In some embodiments, the cells in culture have a peak volumetric productivity of isoprene at greater than or about 0.5, 1.0, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1100, 1200, 1300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,100, 2,200, 2,300, 2,400, 2,500, 2,600, 2,700, 2,800, 2,900, 3,000, 3,100, 3,200, 3,300, 3,400, 3,500, 3,750, 4,000, 4,250, 4,500, 4,750, 5,000, 5,250, 5,500, 5,750, 6,000, 6,250, 6,500, 6,750, 7,000, 7,250, 7,500, 7,750, 8,000, 8,250, 8,500, 8,750, 9,000, 9,250, 9,500, 9,750, 10,000, 12,500, 15,000, or more mg of isoprene/L of broth/hr (mg/L<sub>broth</sub>/hr, wherein the volume of broth includes the volume of the cells and the cell medium). In some embodiments of any of the aspects of the invention, the cells in culture convert greater than or about 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.12, 0.14, 0.16, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.4, 1.6, 2.0, 2.2, 2.4, 2.6, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0, 12.0, 13.0, 14.0, 15.0, 16.0, 17.0, 18.0, 19.0, 20.0, 21.0, 22.0, 23.0, 23.2, 23.4, 23.6, 23.8, 24.0, 25.0, 30.0, 31.0, 32.0, 33.0, 35.0, 37.5, 40.0, 45.0, 47.5, 50.0, 55.0, 60.0, 65.0, 70.0, 75.0, 80.0, 85.0, 90.0 molar %, or more of the carbon in the cell culture medium into isoprene. In some embodiments of any of the aspects of the invention, the cells in culture produce isoprene at greater than or about 1, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, 10,000, 100,000, or more ng of isoprene/gram of cells for the wet weight of the cells /hr (ng/g<sub>wcm</sub>/h). In some embodiments of any of the aspects of the invention, the cells in culture produce a cumulative titer (total amount) of isoprene at greater than or about 1, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000,



2,500, 3,000, 4,000, 5,000, 10,000, 50,000, 100,000, or more mg of isoprene/L of broth (mg/L<sub>broth</sub>, wherein the volume of broth includes the volume of the cells and the cell medium). Other exemplary rates of isoprene production and total amounts of isoprene production are disclosed herein.

**[0014]** In some embodiments of any of the aspects of the invention, the cells further comprise a heterologous nucleic acid encoding an IDI polypeptide. In some embodiments of any of the aspects of the invention, the cells further comprise an insertion of a copy of an endogenous nucleic acid encoding an IDI polypeptide. In some embodiments of any of the aspects of the invention, the cells further comprise a heterologous nucleic acid encoding a DXS polypeptide. In some embodiments of any of the aspects of the invention, the cells further comprise an insertion of a copy of an endogenous nucleic acid encoding a DXS polypeptide. In some embodiments of any of the aspects of the invention, the cells further comprise one or more nucleic acids encoding an IDI polypeptide and a DXS polypeptide. In some embodiments of any of the aspects of the invention, one nucleic acid encodes the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide. In some embodiments of any of the aspects of the invention, one vector encodes the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide. In some embodiments, the vector comprises a selective marker, such as an antibiotic resistance nucleic acid.

**[0015]** In some embodiments of any of the aspects of the invention, the heterologous isoprene synthase nucleic acid is operably linked to a T7 promoter, such as a T7 promoter contained in a medium or high copy plasmid. In some embodiments of any of the aspects of the invention, the heterologous isoprene synthase nucleic acid is operably linked to a Trc promoter, such as a Trc promoter contained in a medium or high copy plasmid. In some embodiments of any of the aspects of the invention, the heterologous isoprene synthase nucleic acid is operably linked to a Lac promoter, such as a Lac promoter contained in a low copy plasmid. In some embodiments of any of the aspects of the invention, the heterologous isoprene synthase nucleic acid is operably linked to an endogenous promoter, such as an endogenous alkaline serine protease promoter. In some embodiments, the heterologous isoprene synthase nucleic acid integrates into a chromosome of the cells without a selective marker.



[0016] In some embodiments of any of the aspects of the invention, at least a portion of the cells maintain the heterologous isoprene synthase nucleic acid for at least or about 5, 10, 20, 40, 50, 60, 65, or more cell divisions in a continuous culture (such as a continuous culture without dilution). In some embodiments of any of the aspects of the invention, the nucleic acid comprising the isoprene synthase, IDI, or DXS nucleic acid also comprises a selective marker, such as an antibiotic resistance nucleic acid.

[0017] In some embodiments of any of the aspects of the invention, the cells further comprise a heterologous nucleic acid encoding an MVA pathway polypeptide (such as an MVA pathway polypeptide from *Saccharomyces cerevisi*, *Methanosarcina mazei*, or *Enterococcus faecalis*). In some embodiments of any of the aspects of the invention, the cells further comprise an insertion of a copy of an endogenous nucleic acid encoding an MVA pathway polypeptide (such as an MVA pathway polypeptide from *Saccharomyces cerevisia*, *Methanosarcina mazei*, or *Enterococcus faecalis*). In some embodiments of any of the aspects of the invention, the cells comprise an isoprene synthase, DXS, and MVA pathway nucleic acid. In some embodiments of any of the aspects of the invention, the cells comprise an isoprene synthase nucleic acid, a DXS nucleic acid, an IDI nucleic acid, and a MVA pathway nucleic (in addition to the IDI nucleic acid).

[0018] In some embodiments of any of the aspects of the invention, the isoprene synthase polypeptide is a naturally-occurring polypeptide from a plant such as *Pueraria* (e.g., *Pueraria montana*) or *Populus* (e.g., *Populus tremuloides*, *Populus alba* (*P. alba*), *Populus nigra*, *Populus trichocarpa*, or the hybrid, *Populus alba* x *Populus tremula*).

[0019] In some embodiments of any of the aspects of the invention, the cells are bacterial cells, such as gram-positive bacterial cells (e.g., *Bacillus* cells such as *Bacillus subtilis* cells or *Streptomyces* cells such as *Streptomyces lividans*, *Streptomyces coelicolor*, *Streptomyces coelicolor*, *Streptomyces albus*, or *Streptomyces griseus* cells). In some embodiments of any of the aspects of the invention, the cells are gram-negative bacterial cells (e.g., *Escherichia* cells such as *Escherichia coli* cells or *Pantoea* cells such as *Pantoea citrea* cells). In some embodiments, the *E. coli* cells are *E. coli FadR atoC* mutant cells. In some embodiments, the *E. coli* cells express (such as constitutively express) *ybhE* (also known as *pgl*). In some embodiments of any of the aspects of the invention, the cells are fungal, cells such as



filamentous fungal cells (e.g., *Trichoderma* cells such as *Trichoderma reesei* cells or *Aspergillus* cells such as *Aspergillus oryzae* and *Aspergillus niger*) or yeast cells (e.g., *Yarrowia* cells such as *Yarrowia lipolytica* cells).

[0020] In some embodiments of any of the aspects of the invention, the microbial polypeptide carbon source includes one or more polypeptides from yeast or bacteria. In some embodiments of any of the aspects of the invention, the plant polypeptide carbon source includes one or more polypeptides from soy, corn, canola, jatropha, palm, peanut, sunflower, coconut, mustard, rapeseed, cottonseed, palm kernel, olive, safflower, sesame, or linseed.

[0021] In one aspect, the invention provides a tire comprising polyisoprene. In some embodiments, the polyisoprene is produced by (i) polymerizing isoprene from any of the compositions or methods described herein or (ii) polymerizing isoprene recovered from any of the compositions or methods described herein. In some embodiments, the polyisoprene comprises *cis*-1,4-polyisoprene.

[0022] In one aspect, the invention features a product (such as a tire) produced by any of the compositions or methods of the invention.

[0023] It is to be understood that one, some, or all of the properties of the various embodiments described herein may be combined to form other embodiments of the present invention. These and other aspects of the invention will become apparent to one of skill in the art.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0024] Figure 1 is the nucleotide sequence of a kudzu isoprene synthase gene codon-optimized for expression in *E. coli* (SEQ ID NO:1). The atg start codon is in italics, the stop codon is in bold and the added *Pst*I site is underlined.

[0025] Figure 2 is a map of pTrcKudzu.

[0026] Figure 3 is the nucleotide sequence of pTrcKudzu (SEQ ID NO:2). The RBS is underlined, the kudzu isoprene synthase start codon is in bold capital letters and the stop codon is in bold, capital, italics letters. The vector backbone is pTrcHis2B.



- [0027] Figure 4 is a map of pETNHisKudzu.
- [0028] Figure 5 is the nucleotide sequence of pETNHisKudzu (SEQ ID NO:5).
- [0029] Figure 6 is a map of pCL-lac-Kudzu.
- [0030] Figure 7 is the nucleotide sequence of pCL-lac-Kudzu (SEQ ID NO:7).
- [0031] Figure 8A is a graph showing the production of isoprene in *E. coli* BL21 cells with no vector.
- [0032] Figure 8B is a graph showing the production of isoprene in *E. coli* BL21 cells with pCL-lac-Kudzu
- [0033] Figure 8C is a graph showing the production of isoprene in *E. coli* BL21 cells with pTrcKudzu.
- [0034] Figure 8D is a graph showing the production of isoprene in *E. coli* BL21 cells with pETN-HisKudzu.
- [0035] Figure 9A is a graph showing OD over time of fermentation of *E. coli* BL21/pTrcKudzu in a 14 liter fed batch fermentation.
- [0036] Figure 9B is a graph showing isoprene production over time of fermentation of *E. coli* BL21/pTrcKudzu in a 14 liter fed batch fermentation.
- [0037] Figure 10A is a graph showing the production of isoprene in *Pantaea citrea*. Control cells without recombinant kudzu isoprene synthase. Grey diamonds represent isoprene synthesis, black squares represent OD<sub>600</sub>.
- [0038] Figure 10B is a graph showing the production of isoprene in *Pantaea citrea* expressing pCL-lac Kudzu. Grey diamonds represent isoprene synthesis, black squares represent OD<sub>600</sub>.
- [0039] Figure 10C is a graph showing the production of isoprene in *Pantaea citrea* expressing pTrcKudzu. Grey diamonds represent isoprene synthesis, black squares represent OD<sub>600</sub>.



[0040] Figure 11 is a graph showing the production of isoprene in *Bacillus subtilis* expressing recombinant isoprene synthase. BG3594comK is a *B. subtilis* strain without plasmid (native isoprene production). CF443-BG3594comK is a *B. subtilis* strain with pBSKudzu (recombinant isoprene production). IS on the y-axis indicates isoprene.

[0041] Figure 12 is the nucleotide sequence of pBS Kudzu #2 (SEQ ID NO:57).

[0042] Figure 13 is the nucleotide sequence of kudzu isoprene synthase codon-optimized for expression in *Yarrowia* (SEQ ID NO:8).

[0043] Figure 14 is a map of pTrex3g comprising a kudzu isoprene synthase gene codon-optimized for expression in *Yarrowia*.

[0044] Figure 15 is the nucleotide sequence of vector pSPZ1(MAP29Spb) (SEQ ID NO:11).

[0045] Figure 16 is the nucleotide sequence of the synthetic kudzu (*Pueraria montana*) isoprene gene codon-optimized for expression in *Yarrowia* (SEQ ID NO:12).

[0046] Figure 17 is the nucleotide sequence of the synthetic hybrid poplar (*Populus alba* x *Populus tremula*) isoprene synthase gene (SEQ ID NO:13). The ATG start codon is in bold and the stop codon is underlined.

[0047] Figure 18A shows a schematic outlining construction of vectors pYLA 1, pYL1 and pYL2.

[0048] Figure 18B shows a schematic outlining construction of the vector pYLA(POP1).

[0049] Figure 18C shows a schematic outlining construction of the vector pYLA(KZ1)

[0050] Figure 18D shows a schematic outlining construction of the vector pYLI(KZ1)

[0051] Figure 18E shows a schematic outlining construction of the vector pYLI(MAP29)

[0052] Figure 18F shows a schematic outlining construction of the vector pYLA(MAP29)

[0053] Figure 19A shows the MVA and DXP metabolic pathways for isoprene (based on F. Bouvier *et al.*, Progress in Lipid Res. 44: 357-429, 2005). The following description includes



alternative names for each polypeptide in the pathways and a reference that discloses an assay for measuring the activity of the indicated polypeptide (each of these references are each hereby incorporated by reference in their entireties, particularly with respect to assays for polypeptide activity for polypeptides in the MVA and DXP pathways). **Mevalonate Pathway:** **AACT**; Acetyl-CoA acetyltransferase, MvaE, EC 2.3.1.9. Assay: *J. Bacteriol.*, 184: 2116–2122, 2002; **HMGS**; Hydroxymethylglutaryl-CoA synthase, MvaS, EC 2.3.3.10. Assay: *J. Bacteriol.*, 184: 4065–4070, 2002; **HMGR**; 3-Hydroxy-3-methylglutaryl-CoA reductase, MvaE, EC 1.1.1.34. Assay: *J. Bacteriol.*, 184: 2116–2122, 2002; **MVK**; Mevalonate kinase, ERG12, EC 2.7.1.36. Assay: *Curr Genet* 19:9-14, 1991. **PMK**; Phosphomevalonate kinase, ERG8, EC 2.7.4.2, Assay: *Mol Cell Biol.*, 11:620–631, 1991; **DPMDC**; Diphosphomevalonate decarboxylase, MVD1, EC 4.1.1.33. Assay: *Biochemistry*, 33:13355-13362, 1994; **IDI**; Isopentenyl-diphosphate delta-isomerase, IDI1, EC 5.3.3.2. Assay: *J. Biol. Chem.* 264:19169-19175, 1989. **DXP Pathway:** **DXS**; 1-Deoxyxylulose-5-phosphate synthase, dxs, EC 2.2.1.7. Assay: *PNAS*, 94:12857-62, 1997; **DXR**; 1-Deoxy-D-xylulose 5-phosphate reductoisomerase, dxr, EC 2.2.1.7. Assay: *Eur. J. Biochem.* 269:4446–4457, 2002; **MCT**; 4-Diphosphocytidyl-2C-methyl-D-erythritol synthase, IspD, EC 2.7.7.60. Assay: *PNAS*, 97: 6451–6456, 2000; **CMK**; 4-Diphosphocytidyl-2-C-methyl-D-erythritol kinase, IspE, EC 2.7.1.148. Assay: *PNAS*, 97:1062-1067, 2000; **MCS**; 2C-Methyl-D-erythritol 2,4-cyclodiphosphate synthase, IspF, EC 4.6.1.12. Assay: *PNAS*, 96:11758-11763, 1999; **HDS**; 1-Hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase, ispG, EC 1.17.4.3. Assay: *J. Org. Chem.*, 70:9168 -9174, 2005; **HDR**; 1-Hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase, IspH, EC 1.17.1.2. Assay: *JACS*, 126:12847-12855, 2004.

**[0054]** Figure 19B illustrates the classical and modified MVA pathways. 1, acetyl-CoA acetyltransferase (**AACT**); 2, HMG-CoA synthase (**HMGS**); 3, HMG-CoA reductase (**HMGR**); 4, mevalonate kinase (**MVK**); 5, phosphomevalonate kinase (**PMK**); 6, diphosphomevalonate decarboxylase (**MVD** or **DPMDC**); 7, isopentenyl diphosphate isomerase (**IDI**); 8, phosphomevalonate decarboxylase (**PMDC**); 9, isopentenyl phosphate kinase (**IPK**). The classical MVA pathway proceeds from reaction 1 through reaction 7 via reactions 5 and 6, while a modified MVA pathway goes through reactions 8 and 9. P and PP in the structural formula are phosphate and pyrophosphate, respectively. This figure was taken from Koga and Morii, *Microbiology and Mol. Biology Reviews*, 71:97-120, 2007, which is incorporated by reference in



its entirety, particularly with respect to nucleic acids and polypeptides of the modified MVA pathway. The modified MVA pathway is present, for example, in some archaeal organisms, such as *Methanosarcina mazei*.

[0055] Figure 20 shows graphs representing results of the GC-MS analysis of isoprene production by recombinant *Y. lipolytica* strains without (left) or with (right) a kudzu isoprene synthase gene. The arrows indicate the elution time of the authentic isoprene standard.

[0056] Figure 21 is a map of pTrcKudzu yIDI DXS Kan.

[0057] Figure 22 is the nucleotide sequence of pTrcKudzu yIDI DXS Kan (SEQ ID NO:20).

[0058] Figure 23A is a graph showing production of isoprene from glucose in BL21/pTrcKudzukan. Time 0 is the time of induction with IPTG (400  $\mu$ mol). The x-axis is time after induction; the y-axis is OD<sub>600</sub> and the y2-axis is total productivity of isoprene ( $\mu$ g/L headspace or specific productivity ( $\mu$ g/L headspace/OD)). Diamonds represent OD<sub>600</sub>, circles represent total isoprene productivity ( $\mu$ g/L) and squares represent specific productivity of isoprene ( $\mu$ g/L/OD).

[0059] Figure 23B is a graph showing production of isoprene from glucose in BL21/pTrcKudzu yIDI kan. Time 0 is the time of induction with IPTG (400  $\mu$ mol). The x-axis is time after induction; the y-axis is OD<sub>600</sub> and the y2-axis is total productivity of isoprene ( $\mu$ g/L headspace or specific productivity ( $\mu$ g/L headspace/OD)). Diamonds represent OD<sub>600</sub>, circles represent total isoprene productivity ( $\mu$ g/L) and squares represent specific productivity of isoprene ( $\mu$ g/L/OD).

[0060] Figure 23C is a graph showing production of isoprene from glucose in BL21/pTrcKudzu DXS kan. Time 0 is the time of induction with IPTG (400  $\mu$ mol). The x-axis is time after induction; the y-axis is OD<sub>600</sub> and the y2-axis is total productivity of isoprene ( $\mu$ g/L headspace or specific productivity ( $\mu$ g/L headspace/OD)). Diamonds represent OD<sub>600</sub>, circles represent total isoprene productivity ( $\mu$ g/L) and squares represent specific productivity of isoprene ( $\mu$ g/L/OD).

[0061] Figure 23D is a graph showing production of isoprene from glucose in BL21/pTrcKudzu yIDI DXS kan. Time 0 is the time of induction with IPTG (400  $\mu$ mol). The



x-axis is time after induction; the y-axis is OD<sub>600</sub> and the y2-axis is total productivity of isoprene (μg/L headspace or specific productivity (μg/L headspace/OD)). Diamonds represent OD<sub>600</sub>, circles represent total isoprene productivity (μg/L) and squares represent specific productivity of isoprene (μg/L/OD).

**[0062]** Figure 23E is a graph showing production of isoprene from glucose in BL21/pCL PtrcKudzu. Time 0 is the time of induction with IPTG (400 μmol). The x-axis is time after induction; the y-axis is OD<sub>600</sub> and the y2-axis is total productivity of isoprene (μg/L headspace or specific productivity (μg/L headspace/OD)). Diamonds represent OD<sub>600</sub>, circles represent total isoprene productivity (μg/L) and squares represent specific productivity of isoprene (μg/L/OD).

**[0063]** Figure 23F is a graph showing production of isoprene from glucose in BL21/pCL PtrcKudzu yIDI. Time 0 is the time of induction with IPTG (400 μmol). The x-axis is time after induction; the y-axis is OD<sub>600</sub> and the y2-axis is total productivity of isoprene (μg/L headspace or specific productivity (μg/L headspace/OD)). Diamonds represent OD<sub>600</sub>, circles represent total isoprene productivity (μg/L) and squares represent specific productivity of isoprene (μg/L/OD).

**[0064]** Figure 23G is a graph showing production of isoprene from glucose in BL21/pCL PtrcKudzu DXS. Time 0 is the time of induction with IPTG (400 μmol). The x-axis is time after induction; the y-axis is OD<sub>600</sub> and the y2-axis is total productivity of isoprene (μg/L headspace or specific productivity (μg/L headspace/OD)). Diamonds represent OD<sub>600</sub>, circles represent total isoprene productivity (μg/L) and squares represent specific productivity of isoprene (μg/L/OD).

**[0065]** Figure 23H is a graph showing production of isoprene from glucose in BL21/pTrcKudzuIDIDXkan. The arrow indicates the time of induction with IPTG (400 μmol). The x-axis is time after induction; the y-axis is OD<sub>600</sub> and the y2-axis is total productivity of isoprene (μg/L headspace or specific productivity (μg/L headspace/OD)). Black diamonds represent OD<sub>600</sub>, black triangles represent isoprene productivity (μg/L) and white squares represent specific productivity of isoprene (μg/L/OD).

**[0066]** Figure 24 is a map of pTrcKKDyIkIS kan.

**[0067]** Figure 25 is a nucleotide sequence of pTrcKKDyIkIS kan (SEQ ID NO:33).



[0068] Figure 26 is a map of pCL PtrcUpperPathway.

[0069] Figure 27 is a nucleotide sequence of pCL PtrcUpperPathway (SEQ ID NO:46).

[0070] Figure 28 shows a map of the cassette containing the lower MVA pathway and yeast *idi* for integration into the *B. subtilis* chromosome at the *nprE* locus. *nprE* upstream/downstream indicates 1 kb each of sequence from the *nprE* locus for integration. *aprE* promoter (alkaline serine protease promoter) indicates the promoter (-35, -10, +1 transcription start site, RBS) of the *aprE* gene. MVK1 indicates the yeast mevalonate kinase gene. RBS-PMK indicates the yeast phosphomevalonate kinase gene with a *Bacillus* RBS upstream of the start site. RBS-MPD indicates the yeast diphosphomevalonate decarboxylase gene with a *Bacillus* RBS upstream of the start site. RBS-IDI indicates the yeast *idi* gene with a *Bacillus* RBS upstream of the start site. Terminator indicates the terminator alkaline serine protease transcription terminator from *B. amyloliquefaciens*. SpecR indicates the spectinomycin resistance marker. “*nprE* upstream repeat for amp.” indicates a direct repeat of the upstream region used for amplification.

[0071] Figure 29 is a nucleotide sequence of cassette containing the lower MVA pathway and yeast *idi* for integration into the *B. subtilis* chromosome at the *nprE* locus (SEQ ID NO:47).

[0072] Figure 30 is a map of p9796-poplar.

[0073] Figure 31 is a nucleotide sequence of p9796-poplar (SEQ ID NO:48).

[0074] Figure 32 is a map of pTrcPoplar.

[0075] Figure 33 is a nucleotide sequence of pTrcPoplar (SEQ ID NO:49).

[0076] Figure 34 is a map of pTrcKudzu yIDI Kan.

[0077] Figure 35 is a nucleotide sequence of pTrcKudzu yIDI Kan (SEQ ID NO:50).

[0078] Figure 36 is a map of pTrcKudzuDXS Kan.

[0079] Figure 37 is a nucleotide sequence of pTrcKudzuDXS Kan (SEQ ID NO:51).

[0080] Figure 38 is a map of pCL PtrcKudzu.



- [0081] Figure 39 is a nucleotide sequence of pCL PtrcKudzu (SEQ ID NO:52).
- [0082] Figure 40 is a map of pCL PtrcKudzu A3.
- [0083] Figure 41 is a nucleotide sequence of pCL PtrcKudzu A3 (SEQ ID NO:53).
- [0084] Figure 42 is a map of pCL PtrcKudzu yIDI.
- [0085] Figure 43 is a nucleotide sequence of pCL PtrcKudzu yIDI (SEQ ID NO:54).
- [0086] Figure 44 is a map of pCL PtrcKudzu DXS.
- [0087] Figure 45 is a nucleotide sequence of pCL PtrcKudzu DXS (SEQ ID NO:55).
- [0088] Figure 46 shows graphs representing isoprene production from biomass feedstocks. Panel A shows isoprene production from corn stover, Panel B shows isoprene production from bagasse, Panel C shows isoprene production from softwood pulp, Panel D shows isoprene production from glucose, and Panel E shows isoprene production from cells with no additional feedstock. Grey squares represent OD<sub>600</sub> measurements of the cultures at the indicated times post-inoculation and black triangles represent isoprene production at the indicated times post-inoculation.
- [0089] Figure 47A shows a graph representing isoprene production by BL21 ( $\lambda$ DE3) pTrcKudzu yIDI DXS (kan) in a culture with no glucose added. Squares represent OD<sub>600</sub>, and triangles represent isoprene produced ( $\mu$ g/ml).
- [0090] Figure 47B shows a graph representing isoprene production from 1% glucose feedstock invert sugar by BL21 ( $\lambda$ DE3) pTrcKudzu yIDI DXS (kan). Squares represent OD<sub>600</sub>, and triangles represent isoprene produced ( $\mu$ g/ml).
- [0091] Figure 47C shows a graph representing isoprene production from 1% invert sugar feedstock by BL21 ( $\lambda$ DE3) pTrcKudzu yIDI DXS (kan). Squares represent OD<sub>600</sub>, and triangles represent isoprene produced ( $\mu$ g/ml).



[0092] Figure 47D shows a graph representing isoprene production from 1% AFEX corn stover feedstock by BL21 ( $\lambda$ DE3) pTrcKudzu yIDI DXS (kan). Squares represent OD<sub>600</sub>, and triangles represent isoprene produced ( $\mu$ g/ml).

[0093] Figure 48 shows graphs demonstrating the effect of yeast extract of isoprene production. Panel A shows the time course of optical density within fermentors fed with varying amounts of yeast extract. Panel B shows the time course of isoprene titer within fermentors fed with varying amounts of yeast extract. The titer is defined as the amount of isoprene produced per liter of fermentation broth. Panel C shows the effect of yeast extract on isoprene production in *E. coli* grown in fed-batch culture.

[0094] Figure 49 shows graphs demonstrating isoprene production from a 500 L bioreactor with *E. coli* cells containing the pTrcKudzu and yIDI and DXS plasmid. Panel A shows the time course of optical density within the 500-L bioreactor fed with glucose and yeast extract. Panel B shows the time course of isoprene titer within the 500-L bioreactor fed with glucose and yeast extract. The titer is defined as the amount of isoprene produced per liter of fermentation broth. Panel C shows the time course of total isoprene produced from the 500-L bioreactor fed with glucose and yeast extract.

[0095] Figure 50 is a map of pJMupperpathway2.

[0096] Figure 51 is the nucleotide sequence of pJMupperpathway2 (SEQ ID NO:56).

[0097] Figure 52 is a map of pBS Kudzu #2.

[0098] Figure 53A is a graph showing growth during fermentation time of *Bacillus* expressing recombinant kudzu isoprene synthase in 14 liter fed batch fermentation. Black diamonds represent a control strain (BG3594comK) without recombinant isoprene synthase (native isoprene production) and grey triangles represent *Bacillus* with pBSKudzu (recombinant isoprene production).

[0099] Figure 53B is a graph showing isoprene production during fermentation time of *Bacillus* expressing recombinant kudzu isoprene synthase in 14 liter fed batch fermentation. Black diamonds represent a control strain (BG3594comK) without recombinant isoprene



synthase (native isoprene production) and grey triangles represent *Bacillus* with pBSKudzu (recombinant isoprene production).

[0100] Figures 54 is a map of plasmid pET24 P.alba HGS.

[0101] Figures 55A and 55B are the nucleotide sequence of plasmid pET24 P.alba HGS (SEQ ID NO:87).

[0102] Figure 56 is a schematic diagram showing restriction sites used for endonuclease digestion to construct plasmid EWL230 and compatible cohesive ends between BspHI and NcoI sites.

[0103] Figure 57 is a map of plasmid EWL230.

[0104] Figures 58A and 58B are the nucleotide sequence of plasmid EWL230 (SEQ ID NO:88).

[0105] Figure 59 is a schematic diagram showing restriction sites used for endonuclease digestion to construct plasmid EWL244 and compatible cohesive ends between NsiI and PstI sites.

[0106] Figure 60 is a map of EWL244.

[0107] Figures 61A and 61B are the nucleotide sequence of plasmid EWL244 (SEQ ID NO:89).

[0108] Figure 62 is a map of plasmids MCM484-487.

[0109] Figures 63A-63C are the nucleotide sequence of plasmid MCM484 (SEQ ID NO:90).

[0110] Figures 64A-64C are the nucleotide sequence of plasmid MCM485 (SEQ ID NO:91).

[0111] Figures 65A-65C are the nucleotide sequence of plasmid MCM486 (SEQ ID NO:92).

[0112] Figures 66A-66C are the nucleotide sequence of plasmid MCM487 (SEQ ID NO:93).



**[0113]** Figures 67A-67D are graphs of isoprene production by *E. coli* strain (EWL256) expressing genes from the MVA pathway and grown in fed-batch culture at the 15-L scale without yeast extract feeding. Figure 67A shows the time course of optical density within the 15-L bioreactor fed with glucose. Figure 67B shows the time course of isoprene titer within the 15-L bioreactor fed with glucose. The titer is defined as the amount of isoprene produced per liter of fermentation broth. Figure 67C shows the time course of total isoprene produced from the 15-L bioreactor fed with glucose. Figure 67D shows the total carbon dioxide evolution rate (TCER), or metabolic activity profile, within the 15-L bioreactor fed with glucose.

**[0114]** Figures 68A-68E are graphs of isoprene production by *E. coli* strain (EWL256) expressing genes from the MVA pathway and grown in fed-batch culture at the 15-L scale with yeast extract feeding. Figure 68A shows the time course of optical density within the 15-L bioreactor fed with glucose. Figure 68B shows the time course of isoprene titer within the 15-L bioreactor fed with glucose. The titer is defined as the amount of isoprene produced per liter of fermentation broth. Figure 68C shows the time course of total isoprene produced from the 15-L bioreactor fed with glucose. Figure 68D shows the volumetric productivity within the 15-L bioreactor fed with glucose. An average value of 1.1 g/L/hr was maintained for a 40-hour period (23-63 hours) with yeast extract feeding. Figure 68E shows the carbon dioxide evolution rate (CER), or metabolic activity profile, within the 15-L bioreactor fed with glucose.

**[0115]** Figures 69A-69D shows production of isoprene from different carbon sources via the MVA (pathway). Figure 69A shows growth of *E. coli* EWL256, which contains both the MVA pathway and isoprene synthase, on either glucose, biomass hydrolysate, glycerol, or acetate as the only carbon source. The different carbon sources were added to a concentration of 1% in the media. A negative control with no added carbon source was included. Growth was measured as optical density at 600 nM. Figure 69B shows specific productivity of isoprene from *E. coli* EWL256 containing both the MVA pathway and isoprene synthase when grown on either glucose, biomass hydrolysate, glycerol, or acetate as only carbon source. The different carbon sources were added to a concentration of 1% in the media. A negative control with no added carbon source was included. Samples were taken 190 minutes, 255 minutes and 317 minutes after inoculation and isoprene produced by the bacteria was measured using GC-MS. Figure 69C shows growth of *E. coli* EWL256 on either glucose or xylose as the only carbon source.



The different carbon sources were added to a concentration of 1% in the media. A negative control with no added carbon source was included. Growth was measured as optical density at 600nm. Figure 69D shows specific productivity of isoprene from *E. coli* EWL256 when grown on either glucose or xylose as only carbon source. The carbon sources were added to a concentration of 1% in the media. A negative control with no added carbon source was included. Samples were taken 260 minutes, 322 minutes and 383 minutes after inoculation and isoprene produced by the bacteria was measured using GC-MS.

**[0116]** Figures 70A and 70B show the production of isoprene by *E. coli* strains from glucose and from fatty acid, respectively. For Figure 70A, eleven colonies from the transformation of WW4 with pMCM118, the plasmid bearing the lower mevalonic acid pathway, were picked to verify the presence of the lower pathway. Cell from the colonies were cultured in TM3 medium containing 0.1% yeast extract and 2% glucose. Aliquots of induced culture were assayed for isoprene production after 4 hours of induction. All colonies showed the production of isoprene. The inducer IPTG had a strong growth inhibitory effect as was evident from the 3 to 4.6-fold reduced cell density in going from 50 to 900 uM concentration of the inducer (data not shown). The graph shows that higher induction, yields a higher specific titer of isoprene. For Figure 70B, the production culture was inoculated from a washed overnight culture at 1 to 10 dilution. The culture was grown for several hours and induced with 50 uM IPTG. The left bar shows isoprene assay results four hours after induction followed by a one hour isoprene accumulation assay. The middle bar shows the one hour normalized value for the same culture with the same induction period but analyzed by a 12 hour isoprene accumulation assay. The right bar shows the value for a one hour isoprene accumulation assay of the culture that was induced for 13 hours.

**[0117]** Figure 71 is a map of the *E. coli-Streptomyces* shuttle vector pUWL201PW (6400 bp) used for cloning isoprene synthase from Kudzu. Tsr, thiostrepton resistance gene. Picture is taken from Doumith *et al.*, *Mol. Gen. Genet.* 264: 477-485, 2000.

**[0118]** Figure 72 shows isoprene formation by *Streptomyces albus* wild type strain ("wt") and strains harboring plasmid pUWL201PW (negative control) or pUWL201\_iso (encoding isoprene synthase from Kudzu).



[0119] Figure 73A is a map of the *M. mazei* archaeal Lower Pathway operon.

[0120] Figures 73B and 73C are the nucleotide sequence of the *M. mazei* archaeal lower Pathway operon (SEQ ID NO:113).

[0121] Figure 74A is a map of MCM376 - MVK from *M. mazei* archaeal Lowerin pET200D.

[0122] Figures 74B and 74C are the nucleotide sequence of MCM376 - MVK from *M. mazei* archaeal Lowerin pET200D (SEQ ID NO:114).

[0123] Figures 75A-75D show growth and specific productivity of isoprene production for EWL256 compared to RM11608-2. Growth (OD550) is represented by the white diamonds; specific productivity of isoprene is represented by the solid bars. The x-axis is time (hours) post-induction with either 200 (Figures 75A and 75B) or 400 (Figures 75C and 75D)  $\mu$ M IPTG. Y-1 axis is productivity of isoprene ( $\mu$ g/L/OD/hr) and Y-2 is arbitrary units of optical density at a wavelength of 550. These values for the OD550 must be multiplied by 6.66 to obtain the actual OD of the culture.

[0124] Figure 76 is a map of plasmid pBBRCMPGI1.5-pgl.

[0125] Figures 77A and 77B are the nucleotide sequence of plasmid pBBRCMPGI1.5-pgl (SEQ ID NO:122).

[0126] Figures 78A-78F are graphs of isoprene production by *E. coli* strain expressing *M. mazei* mevalonate kinase, *P. alba* isoprene synthase, and *pgl* (RHM11608-2), and grown in fed-batch culture at the 15-L scale. Figure 78A shows the time course of optical density within the 15-L bioreactor fed with glucose. Figure 78B shows the time course of isoprene titer within the 15-L bioreactor fed with glucose. The titer is defined as the amount of isoprene produced per liter of fermentation broth. Method for calculating isoprene: cumulative isoprene produced in 59 hrs, g/Fermentor volume at 59 hrs, L [=] g/L broth. Figure 78C also shows the time course of isoprene titer within the 15-L bioreactor fed with glucose. Method for calculating isoprene:  $\int$ (Instantaneous isoprene production rate, g/L/hr)dt from t = 0 to 59 hours [=] g/L broth. Figure 78D shows the time course of total isoprene produced from the 15-L bioreactor fed with glucose. Figure 78E shows volumetric productivity within the 15-L bioreactor fed with glucose. Figure



78F shows carbon dioxide evolution rate (CER), or metabolic activity profile, within the 15-L bioreactor fed with glucose.

[0127] Figure 79A is a map of plasmid pJ201:19813.

[0128] Figures 79B and 79C are the nucleotide sequence of pJ201:19813 (SEQ ID NO:123).

#### DETAILED DESCRIPTION OF THE INVENTION

[0129] The invention features compositions and methods for the production of increased amounts of isoprene. In particular, these compositions and methods increase the rate of isoprene production and increase the total amount of isoprene that is produced. For example, cell culture systems that generate  $4.8 \times 10^4$  nmole/ $g_{wcm}$ /hr of isoprene have been produced (Table 1). The efficiency of these systems is demonstrated by the conversion of  $\sim 23.6$  molar % yield (10.7 weight % yield) of the carbon that the cells consume from a cell culture medium into isoprene (% carbon yield). As shown in the Examples and Table 2, approximately 60.5 g of isoprene per liter of broth was generated. Isoprene was produced at a peak specific rate of  $1.88 \times 10^5$  nmol/OD/hr ( $1.88 \times 10^5$  nmole/ $g_{wcm}$ /hr). If desired, even greater amounts of isoprene can be obtained using other conditions, such as those described herein. In some embodiments, a renewable carbon source is used for the production of isoprene. The compositions and methods of the present invention are desirable because they allow high isoprene yield per cell, high carbon yield, high isoprene purity, high productivity, low energy usage, low production cost and investment, and minimal side reactions. This efficient, large scale, biosynthetic process for isoprene production provides an isoprene source for synthetic isoprene-based rubber and provides a desirable, low-cost alternative to using natural rubber.

[0130] As discussed further below, the amount of isoprene produced by cells can be greatly increased by introducing a heterologous nucleic acid encoding an isoprene synthase polypeptide (e.g., a plant isoprene synthase polypeptide) into the cells. Isoprene synthase polypeptides convert dimethylallyl diphosphate (DMAPP) into isoprene. As shown in the Examples, a heterologous *Pueraria Montana* (kudzu) or *Populus alba* (Poplar) isoprene synthase polypeptide was expressed in a variety of host cells, such as *Escherichia coli*, *Pantoea citrea*, *Bacillus subtilis*, *Yarrowia lipolytica*, and *Trichoderma reesei*. As also shown in the Examples, a heterologous *Methanosarcina mazei* (*M. mazei*) mevalonate kinase (MVK) was expressed in



host cells such as *Escherichia coli* to increase isoprene production. All of these cells produced more isoprene than the corresponding cells without the heterologous isoprene synthase polypeptide. As illustrated in Tables 1 and 2, large amounts of isoprene are produced using the methods described herein. For example, *B. subtilis* cells with a heterologous isoprene synthase nucleic acid produced approximately 10-fold more isoprene in a 14 liter fermentor than the corresponding control *B. subtilis* cells without the heterologous nucleic acid (Table 2). The production of 60.5 g of isoprene per liter of broth (mg/L, wherein the volume of broth includes both the volume of the cell medium and the volume of the cells) by *E. coli* and 30 mg/L by *B. subtilis* in fermentors indicates that significant amounts of isoprene can be generated (Table 2). If desired, isoprene can be produced on an even larger scale or other conditions described herein can be used to further increase the amount of isoprene. The vectors listed in Tables 1 and 2 and the experimental conditions are described in further detail below and in the Examples section.

**Table 1: Exemplary yields of isoprene from a shake flask using the cell cultures and methods of the invention.** The assay for measuring isoprene production is described in Example I, part II. For this assay, a sample was removed at one or more time points from the shake flask and cultured for 30 minutes. The amount of isoprene produced in this sample was then measured. The headspace concentration and specific rate of isoprene production are listed in Table 1 and described further herein.

| Strain                                                  | Isoprene Production in a Headspace vial*                |                                                                                                                   |
|---------------------------------------------------------|---------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
|                                                         | Headspace concentration<br>$\mu\text{g/L}_{\text{gas}}$ | Specific Rate<br>$\mu\text{g/L}_{\text{broth}}/\text{hr}/\text{OD}$<br>( $\text{nmol/g}_{\text{wcm}}/\text{hr}$ ) |
| <i>E. coli</i> BL21/ pTrcKudzu IS                       | 1.40                                                    | 53.2<br>(781.2)                                                                                                   |
| <i>E. coli</i> BL21/ Pcl DXS yidi Kudzu IS              | 7.61                                                    | 289.1<br>( $4.25 \times 10^3$ )                                                                                   |
| <i>E. coli</i> BL21/MCM127 with kudzu IS and entire MVA | 23.0                                                    | 874.1<br>( $1.28 \times 10^4$ )                                                                                   |



| pathway                                                                                               |            |                                                   |
|-------------------------------------------------------------------------------------------------------|------------|---------------------------------------------------|
| <i>E. coli</i> BL21/ Pet N-HisKudzu IS                                                                | 1.49       | 56.6<br>(831.1)                                   |
| <i>Pantoea citrea</i> /pTrcKudzu IS                                                                   | 0.66       | 25.1<br>(368.6)                                   |
| <i>E. coli</i> w/ <i>Poplar</i> IS<br>[Miller (2001)]                                                 | -          | 5.6<br>(82.2)                                     |
| <i>Bacillus licheniformis</i> Fall<br>US 5849970                                                      | -          | 4.2<br>(61.4)                                     |
| <i>Yarrowia lipolytica</i> with<br>kudzu isoprene synthase                                            | ~0.05 µg/L | ~2<br>(~30)                                       |
| <i>Trichoderma reesei</i> with<br>kudzu isoprene synthase                                             | ~0.05 µg/L | ~2<br>(~30)                                       |
| <i>E. coli</i> BL21/<br>pTrcKKD <sub>y</sub> I <sub>k</sub> IS with kudzu<br>IS and lower MVA pathway | 85.9       | 3.2 x 10 <sup>3</sup><br>(4.8 x 10 <sup>4</sup> ) |

\*Normalized to 1 mL of 1 OD<sub>600</sub>, cultured for 1 hour in a sealed headspace vial with a liquid to headspace volume ratio of 1:19.



**Table 2: Exemplary yields of isoprene in a fermentor using the cell cultures and methods of the invention.** The assay for measuring isoprene production is described in Example I, part II. For this assay, a sample of the off-gas of the fermentor was taken and analyzed for the amount of isoprene. The peak headspace concentration (which is the highest headspace concentration during the fermentation), titer (which is the cumulative, total amount of isoprene produced per liter of broth), and peak specific rate of isoprene production (which is the highest specific rate during the fermentation) are listed in Table 2 and described further herein.

| Strain                                                      | Isoprene Production in Fermentors                                         |                                                  |                                                                                                                                      |
|-------------------------------------------------------------|---------------------------------------------------------------------------|--------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------|
|                                                             | Peak Headspace concentration**<br>( $\mu\text{g}/\text{L}_{\text{gas}}$ ) | Titer<br>( $\text{mg}/\text{L}_{\text{broth}}$ ) | Peak Specific rate<br>$\mu\text{g}/\text{L}_{\text{broth}}/\text{hr}/\text{OD}$<br>( $\text{nmol}/\text{g}_{\text{wcm}}/\text{hr}$ ) |
| <i>E. coli</i> BL21 /pTrcKudzu with Kudzu IS                | 52                                                                        | 41.2                                             | 37<br>(543.3)                                                                                                                        |
| <i>E. coli</i> FM5/pTrcKudzu IS                             | 3                                                                         | 3.5                                              | 21.4<br>(308.1)                                                                                                                      |
| <i>E. coli</i> BL21/ triple strain (DXS, yidi, IS)          | 285                                                                       | 300                                              | 240<br>( $3.52 \times 10^3$ )                                                                                                        |
| <i>E. coli</i> FM5/ triple strain (DXS, yidi, IS)           | 50.8                                                                      | 29                                               | 180.8<br>( $2.65 \times 10^3$ )                                                                                                      |
| <i>E. coli</i> /MCM127 with Kudzu IS and entire MVA pathway | 1094                                                                      | 250                                              | 875<br>( $1.28 \times 10^4$ )                                                                                                        |
| <i>Bacillus subtilis</i> wild-type                          | 1.5                                                                       | 2.5                                              | 0.8<br>(11.7)                                                                                                                        |



|                                                                                                                  |                    |                         |                                              |
|------------------------------------------------------------------------------------------------------------------|--------------------|-------------------------|----------------------------------------------|
| <i>Bacillus pBS Kudzu IS</i>                                                                                     | 16.6               | ~30<br>(over 100 hours) | 5<br>(73.4)                                  |
| <i>Bacillus Marburg 6051 [Wagner and Fall (1999)]</i>                                                            | 2.04               | 0.61                    | 24.5<br>(359.8)                              |
| <i>Bacillus Marburg 6051 Fall US 5849970</i>                                                                     | 0.7                | 0.15                    | 6.8<br>(100)                                 |
| <i>E. coli</i><br>BL21/pCLPtrcUpper<br>Pathway and<br>gil.2KKDyI and<br>pTrcAlba-mMVK                            | $2.03 \times 10^4$ | $3.22 \times 10^4$      | $5.9 \times 10^3$<br>( $8.66 \times 10^4$ )  |
| <i>E. coli</i><br>BL21/pCLPtrcUpper<br>Pathway and<br>gil.2KKDyI and<br>pTrcAlba-mMVK<br>plus<br>pBBRCMPGI1.5pgl | $3.22 \times 10^4$ | $6.05 \times 10^4$      | $1.28 \times 10^4$<br>( $1.88 \times 10^5$ ) |

\*\*Normalized to an off-gas flow rate of 1 vvm (1 volume off-gas per 1 L<sub>broth</sub> per minute).

[0131] Additionally, isoprene production by cells that contain a heterologous isoprene synthase nucleic acid can be enhanced by increasing the amount of a 1-deoxy-D-xylulose-5-phosphate synthase (DXS) polypeptide and/or an isopentenyl diphosphate isomerase (IDI) polypeptide expressed by the cells. For example, a DXS nucleic acid and/or an IDI nucleic acid can be introduced into the cells. The DXS nucleic acid may be a heterologous nucleic acid or a duplicate copy of an endogenous nucleic acid. Similarly, the IDI nucleic acid may be a



heterologous nucleic acid or a duplicate copy of an endogenous nucleic acid. In some embodiments, the amount of DXS and/or IDI polypeptide is increased by replacing the endogenous DXS and/or IDI promoters or regulatory regions with other promoters and/or regulatory regions that result in greater transcription of the DXS and/or IDI nucleic acids. In some embodiments, the cells contain both a heterologous nucleic acid encoding an isoprene synthase polypeptide (*e.g.*, a plant isoprene synthase nucleic acid) and a duplicate copy of an endogenous nucleic acid encoding an isoprene synthase polypeptide.

**[0132]** The encoded DXS and IDI polypeptides are part of the DXP pathway for the biosynthesis of isoprene (Figure 19A). DXS polypeptides convert pyruvate and D-glyceraldehyde-3-phosphate into 1-deoxy-D-xylulose-5-phosphate. While not intending to be bound by any particular theory, it is believed that increasing the amount of DXS polypeptide increases the flow of carbon through the DXP pathway, leading to greater isoprene production. IDI polypeptides catalyze the interconversion of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). While not intending to be bound by any particular theory, it is believed that increasing the amount of IDI polypeptide in cells increases the amount of IPP that is converted into DMAPP, which in turn is converted into isoprene.

**[0133]** For example, fermentation of *E. coli* cells with a kudzu isoprene synthase, *S. cerevisia* IDI, and *E. coli* DXS nucleic acids was used to produce isoprene. The levels of isoprene varied from 50 to 300 µg/L over a time period of 15 hours (Example 7, part VII). As another example, fermentation of *E. coli* with *M. mazei* mevalonate kinase (MVK), *P. alba* isoprene synthase, the upper MVA pathway, and the integrated lower MVA pathway was used to produce isoprene. The levels of isoprene varied from 32 to 35.6 g/L over a time period of 67 hours (Example 10, part III).

**[0134]** In yet another example, fermentation of *E. coli* with *M. mazei* mevalonate kinase (MVK), *P. alba* isoprene synthase, *pgl* over-expression (RHM111608-2), the upper MVA pathway, and the integrated lower MVA pathway were used to produce isoprene. The levels of isoprene vary from 33.2 g/L to 40.0 g/L over a time period of 40 hours or 48.6 g/L to 60.5 g/L over a time period of 59 hours (Example 13, part (ii)).



[0135] In some embodiments, the presence of heterologous or extra endogenous isoprene synthase, IDI, and DXS nucleic acids causes cells to grow more reproducibly or remain viable for longer compared to the corresponding cell with only one or two of these heterologous or extra endogenous nucleic acids. For example, cells containing heterologous isoprene synthase, IDI, and DXS nucleic acids grew better than cells with only heterologous isoprene synthase and DXS nucleic acids or with only a heterologous isoprene synthase nucleic acid. Also, heterologous isoprene synthase, IDI, and DXS nucleic acids were successfully operably linked to a strong promoter on a high copy plasmid that was maintained by *E. coli* cells, suggesting that large amounts of these polypeptides could be expressed in the cells without causing an excessive amount of toxicity to the cells. While not intending to be bound to a particular theory, it is believed that the presence of heterologous or extra endogenous isoprene synthase and IDI nucleic acids may reduce the amount of one or more potentially toxic intermediates that would otherwise accumulate if only a heterologous or extra endogenous DXS nucleic acid was present in the cells.

[0136] In some embodiments, the production of isoprene by cells that contain a heterologous isoprene synthase nucleic acid is augmented by increasing the amount of a MVA polypeptide expressed by the cells (Figures 19A and 19B). Exemplary MVA pathway polypeptides include any of the following polypeptides: acetyl-CoA acetyltransferase (AA-CoA thiolase) polypeptides, 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase) polypeptides, 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase) polypeptides, mevalonate kinase (MVK) polypeptides, phosphomevalonate kinase (PMK) polypeptides, diphosphomevalonate decarboxylase (MVD) polypeptides, phosphomevalonate decarboxylase (PMDC) polypeptides, isopentenyl phosphate kinase (IPK) polypeptides, IDI polypeptides, and polypeptides (*e.g.*, fusion polypeptides) having an activity of two or more MVA pathway polypeptides. For example, one or more MVA pathway nucleic acids can be introduced into the cells. In some embodiments, the cells contain the upper MVA pathway, which includes AA-CoA thiolase, HMG-CoA synthase, and HMG-CoA reductase nucleic acids. In some embodiments, the cells contain the lower MVA pathway, which includes MVK, PMK, MVD, and IDI nucleic acids. In some embodiments, the cells contain the entire MVA pathway, which includes AA-CoA thiolase, HMG-CoA synthase, HMG-CoA reductase, MVK, PMK, MVD, and IDI nucleic acids. In some embodiments, the cells contain an entire MVA pathway that includes



AA-CoA thiolase, HMG-CoA synthase, HMG-CoA reductase, MVK, PMDC, IPK, and IDI nucleic acids. The MVA pathway nucleic acids may be heterologous nucleic acids or duplicate copies of endogenous nucleic acids. In some embodiments, the amount of one or more MVA pathway polypeptides is increased by replacing the endogenous promoters or regulatory regions for the MVA pathway nucleic acids with other promoters and/or regulatory regions that result in greater transcription of the MVA pathway nucleic acids. In some embodiments, the cells contain both a heterologous nucleic acid encoding an isoprene synthase polypeptide (*e.g.*, a plant isoprene synthase nucleic acid) and a duplicate copy of an endogenous nucleic acid encoding an isoprene synthase polypeptide.

[0137] For example, *E. coli* cells containing a nucleic acid encoding a kudzu isoprene synthase polypeptide and nucleic acids encoding *Saccharomyces cerevisia* MVK, PMK, MVD, and IDI polypeptides generated isoprene at a rate of  $6.67 \times 10^4$  nmol/L<sub>broth</sub>/OD<sub>600</sub>/hr (*see* Example 8). Additionally, a 14 liter fermentation of *E. coli* cells with nucleic acids encoding *Enterococcus faecalis* AA-CoA thiolase, HMG-CoA synthase, and HMG-CoA reductase polypeptides produced 22 grams of mevalonic acid (an intermediate of the MVA pathway). A shake flask of these cells produced 2-4 grams of mevalonic acid per liter. These results indicate that heterologous MVA pathways nucleic acids are active in *E. coli*. *E. coli* cells that contain nucleic acids for both the upper MVA pathway and the lower MVA pathway as well as a kudzu isoprene synthase (strain MCM 127) produced significantly more isoprene (874 µg/L<sub>broth</sub>/hr/OD) compared to *E. coli* cells with nucleic acids for only the lower MVA pathway and the kudzu isoprene synthase (strain MCM 131) (*see* Table 3 and Example 8, part VIII).

[0138] As another example, *E. coli* cells containing a nucleic acid encoding a *P. alba* isoprene synthase polypeptide and a nucleic acid encoding *M. mazei* MVK polypeptide generated 320.6 g (at a peak specific rate of  $9.54 \times 10^4$  nmol/L<sub>broth</sub>/OD<sub>600</sub>/hr (*i.e.*  $9.5 \times 10^{-5}$  mol/L<sub>broth</sub>/OD<sub>600</sub>/hr)) of isoprene during a 67 hour fermentation in the absence of yeast extract feeding or 395.5 g (at a peak specific rate of  $8.66 \times 10^4$  nmol/L<sub>broth</sub>/OD<sub>600</sub>/hr) during a 68 hour fermentation in the presence of yeast extract feeding (*see* Example 10).

[0139] In some embodiments, at least a portion of the cells maintain the heterologous isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acid for at least about 5, 10, 20, 50, 75, 100, 200, 300, or more cell divisions in a continuous culture (such as a continuous culture without



dilution). In some embodiments of any of the aspects of the invention, the nucleic acid comprising the heterologous or duplicate copy of an endogenous isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acid also comprises a selective marker, such as a kanamycin, ampicillin, carbenicillin, gentamicin, hygromycin, phleomycin, bleomycin, neomycin, or chloramphenicol antibiotic resistance nucleic acid.

[0140] As indicated in Example 7, part VI, the amount of isoprene produced can be further increased by adding yeast extract to the cell culture medium using *E. coli* cells with kudzu isoprene synthase, *S. cerevisia* IDI, and *E. coli* DXS nucleic acids to produce isoprene. In particular, the amount of isoprene produced was linearly proportional to the amount of yeast extract in the cell medium for the concentrations tested (Figure 48C). Additionally, approximately 0.11 grams of isoprene per liter of broth was produced from a cell medium with yeast extract and glucose (Example 7, part VIII). Increasing the amount of yeast extract in the presence of glucose resulted in more isoprene being produced than increasing the amount of glucose in the presence of yeast extract. Also, increasing the amount of yeast extract allowed the cells to produce a high level of isoprene for a longer length of time and improved the health of the cells.

[0141] Isoprene production was also demonstrated using three types of hydrolyzed biomass (bagasse, corn stover, and soft wood pulp) as the carbon source (Figures 46A-C and Figures 69A and 69B). *E. coli* cells with kudzu isoprene synthase, *S. cerevisia* IDI, and *E. coli* DXS nucleic acids produced as much isoprene from these hydrolyzed biomass carbon sources as from the equivalent amount of glucose (*e.g.*, 1% glucose, w/v). *E. coli* cells expressing *P. alba* isoprene synthase and the MVA pathway produced isoprene at a higher initial growth rate from ammonia fiber expansion (AFEX) pretreated corn stover than from the equivalent amount of glucose. (Figures 69A and 69B). If desired, any other biomass carbon source can be used in the compositions and methods of the invention. Biomass carbon sources are desirable because they are cheaper than many conventional cell mediums, thereby facilitating the economical production of isoprene.

[0142] Additionally, invert sugar was shown to function as a carbon source for the generation of isoprene (Figure 47D).



[0143] Additionally, xylose, acetate, and glycerol were also shown to function as a carbon source for the generation of isoprene (Figures 69A-69D). For example, *E. coli* cells with *P. alba* isoprene synthase and the MVA pathway grown on acetate as the only carbon source had a specific productivity of isoprene about twice as high as during growth on glucose (Example 10, Part IV; Figures 69A and 69B).

[0144] In some embodiments, an oil is included in the cell medium. For example, *B. subtilis* cells containing a kudzu isoprene synthase nucleic acid produced isoprene when cultured in a cell medium containing an oil and a source of glucose (Example 4, part III). As another example, *E. coli* fadR atoC mutant cells containing the upper and lower MVA pathway plus kudzu isoprene synthase produced isoprene when cultured in a cell medium containing palm oil and a source of glucose (Example 12, part II). In some embodiments, more than one oil (such as 2, 3, 4, 5, or more oils) is included in the cell medium. While not intending to be bound to any particular theory, it is believed that (i) the oil may increase the amount of carbon in the cells that is available for conversion to isoprene, (ii) the oil may increase the amount of acetyl-CoA in the cells, thereby increasing the carbon flow through the MVA pathway, and/or (iii) the oil may provide extra nutrients to the cells, which is desirable since a lot of the carbon in the cells is converted to isoprene rather than other products. In some embodiments, cells that are cultured in a cell medium containing oil naturally use the MVA pathway to produce isoprene or are genetically modified to contain nucleic acids for the entire MVA pathway. In some embodiments, the oil is partially or completely hydrolyzed before being added to the cell culture medium to facilitate the use of the oil by the host cells.

#### *Exemplary Polypeptides and Nucleic Acids*

[0145] Various isoprene synthase, DXS, IDI, and/or MVA pathway polypeptides and nucleic acids can be used in the compositions and methods of the invention.

[0146] As used herein, "polypeptides" includes polypeptides, proteins, peptides, fragments of polypeptides, and fusion polypeptides that include part or all of a first polypeptide (*e.g.*, an isoprene synthase, DXS, IDI, or MVA pathway polypeptide) and part or all of a second polypeptide (*e.g.*, a peptide that facilitates purification or detection of the fusion polypeptide, such as a His-tag). In some embodiments, the fusion polypeptide has an activity of two or more



MVA pathway polypeptides (such as AA-CoA thiolase and HMG-CoA reductase polypeptides). In some embodiments, the polypeptide is a naturally-occurring polypeptide (such as the polypeptide encoded by an *Enterococcus faecalis mvaE* nucleic acid) that has an activity of two or more MVA pathway polypeptides.

[0147] In various embodiments, a polypeptide has at least or about 50, 100, 150, 175, 200, 250, 300, 350, 400, or more amino acids. In some embodiments, the polypeptide fragment contains at least or about 25, 50, 75, 100, 150, 200, 300, or more contiguous amino acids from a full-length polypeptide and has at least or about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 100% of an activity of a corresponding full-length polypeptide. In particular embodiments, the polypeptide includes a segment of or the entire amino acid sequence of any naturally-occurring isoprene synthase, DXS, IDI, or MVA pathway polypeptide. In some embodiments, the polypeptide has one or more mutations compared to the sequence of a wild-type (*i.e.*, a sequence occurring in nature) isoprene synthase, DXS, IDI, or MVA pathway polypeptide.

[0148] In some embodiments, the polypeptide is an isolated polypeptide. As used herein, an “isolated polypeptide” is not part of a library of polypeptides, such as a library of 2, 5, 10, 20, 50 or more different polypeptides and is separated from at least one component with which it occurs in nature. An isolated polypeptide can be obtained, for example, by expression of a recombinant nucleic acid encoding the polypeptide.

[0149] In some embodiments, the polypeptide is a heterologous polypeptide. By “heterologous polypeptide” is meant a polypeptide whose amino acid sequence is not identical to that of another polypeptide naturally expressed in the same host cell.

[0150] As used herein, a “nucleic acid” refers to two or more deoxyribonucleotides and/or ribonucleotides in either single or double-stranded form. In some embodiments, the nucleic acid is a recombinant nucleic acid. By “recombinant nucleic acid” means a nucleic acid of interest that is free of one or more nucleic acids (*e.g.*, genes) which, in the genome occurring in nature of the organism from which the nucleic acid of interest is derived, flank the nucleic acid of interest. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a



prokaryote or eukaryote, or which exists as a separate molecule (*e.g.*, a cDNA, a genomic DNA fragment, or a cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. In some embodiments, an isoprene synthase, DXS, IDI, or MVA pathway nucleic acid is operably linked to another nucleic acid encoding all or a portion of another polypeptide such that the recombinant nucleic acid encodes a fusion polypeptide that includes an isoprene synthase, DXS, IDI, or MVA pathway polypeptide and all or part of another polypeptide (*e.g.*, a peptide that facilitates purification or detection of the fusion polypeptide, such as a His-tag). In some embodiments, part or all of a recombinant nucleic acid is chemically synthesized.

**[0151]** In some embodiments, the nucleic acid is a heterologous nucleic acid. By “heterologous nucleic acid” is meant a nucleic acid whose nucleic acid sequence is not identical to that of another nucleic acid naturally found in the same host cell.

**[0152]** In particular embodiments, the nucleic acid includes a segment of or the entire nucleic acid sequence of any naturally-occurring isoprene synthase, DXS, IDI, or MVA pathway nucleic acid. In some embodiments, the nucleic acid includes at least or about 50, 100, 150, 200, 300, 400, 500, 600, 700, 800, or more contiguous nucleotides from a naturally-occurring isoprene synthase nucleic acid DXS, IDI, or MVA pathway nucleic acid. In some embodiments, the nucleic acid has one or more mutations compared to the sequence of a wild-type (*i.e.*, a sequence occurring in nature) isoprene synthase, DXS, IDI, or MVA pathway nucleic acid. In some embodiments, the nucleic acid has one or more mutations (*e.g.*, a silent mutation) that increase the transcription or translation of isoprene synthase, DXS, IDI, or MVA pathway nucleic acid. In some embodiments, the nucleic acid is a degenerate variant of any nucleic acid encoding an isoprene synthase, DXS, IDI, or MVA pathway polypeptide.

**[0153]** “Codon degeneracy” refers to divergence in the genetic code permitting variation of the nucleotide sequence without affecting the amino acid sequence of an encoded polypeptide. The skilled artisan is well aware of the “codon-bias” exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a nucleic acid for improved expression in a host cell, it is desirable in some embodiments to design the nucleic acid such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.



[0154] The accession numbers of exemplary isoprene synthase, DXS, IDI, and/or MVA pathway polypeptides and nucleic acids are listed in Appendix 1 (the accession numbers of Appendix 1 and their corresponding sequences are herein incorporated by reference in their entireties, particularly with respect to the amino acid and nucleic acid sequences of isoprene synthase, DXS, IDI, and/or MVA pathway polypeptides and nucleic acids). The Kegg database also contains the amino acid and nucleic acid sequences of numerous exemplary isoprene synthase, DXS, IDI, and/or MVA pathway polypeptides and nucleic acids (*see*, for example, the world-wide web at “genome.jp/kegg/pathway/map/map00100.html” and the sequences therein, which are each hereby incorporated by reference in their entireties, particularly with respect to the amino acid and nucleic acid sequences of isoprene synthase, DXS, IDI, and/or MVA pathway polypeptides and nucleic acids). In some embodiments, one or more of the isoprene synthase, DXS, IDI, and/or MVA pathway polypeptides and/or nucleic acids have a sequence identical to a sequence publicly available on December 12, 2007 or December 11, 2008, such as any of the sequences that correspond to any of the accession numbers in Appendix 1 or any of the sequences present in the Kegg database. Additional exemplary isoprene synthase, DXS, IDI, and/or MVA pathway polypeptides and nucleic acids are described further below.

*Exemplary Isoprene Synthase Polypeptides and Nucleic Acids*

[0155] As noted above, isoprene synthase polypeptides convert dimethylallyl diphosphate (DMAPP) into isoprene. Exemplary isoprene synthase polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of an isoprene synthase polypeptide. Standard methods can be used to determine whether a polypeptide has isoprene synthase polypeptide activity by measuring the ability of the polypeptide to convert DMAPP into isoprene *in vitro*, in a cell extract, or *in vivo*. In an exemplary assay, cell extracts are prepared by growing a strain (*e.g.*, the *E. coli*/pTrcKudzu strain described herein) in the shake flask method as described in Example 1. After induction is complete, approximately 10 mLs of cells are pelleted by centrifugation at 7000 x *g* for 10 minutes and resuspended in 5 ml of PEB without glycerol. The cells are lysed using a French Pressure cell using standard procedures. Alternatively the cells are treated with lysozyme (Ready-Lyse lysozyme solution; EpiCentre) after a freeze/thaw at -80C.



**[0156]** Isoprene synthase polypeptide activity in the cell extract can be measured, for example, as described in Silver *et al.*, *J. Biol. Chem.* 270:13010-13016, 1995 and references therein, which are each hereby incorporated by reference in their entireties, particularly with respect to assays for isoprene synthase polypeptide activity. DMAPP (Sigma) is evaporated to dryness under a stream of nitrogen and rehydrated to a concentration of 100 mM in 100 mM potassium phosphate buffer pH 8.2 and stored at -20 °C. To perform the assay, a solution of 5 µl of 1M MgCl<sub>2</sub>, 1 mM (250 µg/ml) DMAPP, 65 µl of Plant Extract Buffer (PEB) (50 mM Tris-HCl, pH 8.0, 20 mM MgCl<sub>2</sub>, 5% glycerol, and 2 mM DTT) is added to 25 µl of cell extract in a 20 ml Headspace vial with a metal screw cap and teflon coated silicon septum (Agilent Technologies) and cultured at 37 °C for 15 minutes with shaking. The reaction is quenched by adding 200 µl of 250 mM EDTA and quantified by GC/MS as described in Example 1, part II.

**[0157]** Exemplary isoprene synthase nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an isoprene synthase polypeptide. Exemplary isoprene synthase polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein.

**[0158]** In some embodiments, the isoprene synthase polypeptide or nucleic acid is from the family Fabaceae, such as the Faboideae subfamily. In some embodiments, the isoprene synthase polypeptide or nucleic acid is a polypeptide or nucleic acid from *Pueraria montana* (kudzu) (Sharkey *et al.*, *Plant Physiology* 137: 700-712, 2005), *Pueraria lobata*, poplar (such as *Populus alba*, *Populus nigra*, *Populus trichocarpa*, *Populus alba x tremula* (CAC35696), or *Populus alba*) (Sasaki *et al.*, *FEBS Letters* 579(11): 2514-2518, 2005; Miller *et al.*, *Planta* 213: 483-487, 2001), aspen (such as *Populus tremuloides*) (Silver *et al.*, *JBC* 270(22): 13010-1316, 1995), or English Oak (*Quercus robur*) (Zimmer *et al.*, WO 98/02550), which are each hereby incorporated by reference in their entireties, particularly with respect to isoprene synthase nucleic acids and the expression of isoprene synthase polypeptides. Suitable isoprene synthases include, but are not limited to, those identified by Genbank Accession Nos. AY341431, AY316691, AY279379, AJ457070, and AY182241, which are each hereby incorporated by reference in their entireties, particularly with respect to sequences of isoprene synthase nucleic



acids and polypeptides. In some embodiments, the isoprene synthase polypeptide or nucleic acid is not a naturally-occurring polypeptide or nucleic acid from *Quercus robur* (*i.e.*, the isoprene synthase polypeptide or nucleic acid is an isoprene synthase polypeptide or nucleic acid other than a naturally-occurring polypeptide or nucleic acid from *Quercus robur*). In some embodiments, the isoprene synthase nucleic acid or polypeptide is a naturally-occurring polypeptide or nucleic acid from poplar. In some embodiments, the isoprene synthase nucleic acid or polypeptide is not a naturally-occurring polypeptide or nucleic acid from poplar.

#### *Exemplary DXS Polypeptides and Nucleic Acids*

[0159] As noted above, 1-deoxy-D-xylulose-5-phosphate synthase (DXS) polypeptides convert pyruvate and D-glyceraldehyde-3-phosphate into 1-deoxy-D-xylulose-5-phosphate. Exemplary DXS polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of a DXS polypeptide. Standard methods (such as those described herein) can be used to determine whether a polypeptide has DXS polypeptide activity by measuring the ability of the polypeptide to convert pyruvate and D-glyceraldehyde-3-phosphate into 1-deoxy-D-xylulose-5-phosphate *in vitro*, in a cell extract, or *in vivo*. Exemplary DXS nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of a DXS polypeptide. Exemplary DXS polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein.

#### *Exemplary IDI Polypeptides and Nucleic Acids*

[0160] Isopentenyl diphosphate isomerase polypeptides (isopentenyl-diphosphate delta-isomerase or IDI) catalyses the interconversion of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) (*e.g.*, converting IPP into DMAPP and/or converting DMAPP into IPP). Exemplary IDI polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of an IDI polypeptide. Standard methods (such as those described herein) can be used to determine whether a polypeptide has IDI polypeptide activity by measuring the ability of the polypeptide to interconvert IPP and DMAPP *in vitro*, in a cell extract, or *in vivo*. Exemplary IDI nucleic acids



include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an IDI polypeptide. Exemplary IDI polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein.

*Exemplary MVA Pathway Polypeptides and Nucleic Acids*

**[0161]** Exemplary MVA pathway polypeptides include acetyl-CoA acetyltransferase (AA-CoA thiolase) polypeptides, 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase) polypeptides, 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase) polypeptides, mevalonate kinase (MVK) polypeptides, phosphomevalonate kinase (PMK) polypeptides, diphosphomevalonate decarboxylase (MVD) polypeptides, phosphomevalonate decarboxylase (PMDC) polypeptides, isopentenyl phosphate kinase (IPK) polypeptides, IDI polypeptides, and polypeptides (*e.g.*, fusion polypeptides) having an activity of two or more MVA pathway polypeptides. In particular, MVA pathway polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of an MVA pathway polypeptide. Exemplary MVA pathway nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an MVA pathway polypeptide. Exemplary MVA pathway polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein.

**[0162]** In particular, acetyl-CoA acetyltransferase polypeptides (AA-CoA thiolase or AACT) convert two molecules of acetyl-CoA into acetoacetyl-CoA. Standard methods (such as those described herein) can be used to determine whether a polypeptide has AA-CoA thiolase polypeptide activity by measuring the ability of the polypeptide to convert two molecules of acetyl-CoA into acetoacetyl-CoA *in vitro*, in a cell extract, or *in vivo*.

**[0163]** 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase or HMGS) polypeptides convert acetoacetyl-CoA into 3-hydroxy-3-methylglutaryl-CoA. Standard methods (such as those described herein) can be used to determine whether a polypeptide has HMG-CoA



synthase polypeptide activity by measuring the ability of the polypeptide to convert acetoacetyl-CoA into 3-hydroxy-3-methylglutaryl-CoA *in vitro*, in a cell extract, or *in vivo*.

**[0164]** 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase or HMGR) polypeptides convert 3-hydroxy-3-methylglutaryl-CoA into mevalonate. Standard methods (such as those described herein) can be used to determine whether a polypeptide has HMG-CoA reductase polypeptide activity by measuring the ability of the polypeptide to convert 3-hydroxy-3-methylglutaryl-CoA into mevalonate *in vitro*, in a cell extract, or *in vivo*.

**[0165]** Mevalonate kinase (MVK) polypeptide phosphorylates mevalonate to form mevalonate-5-phosphate. Standard methods (such as those described herein) can be used to determine whether a polypeptide has MVK polypeptide activity by measuring the ability of the polypeptide to convert mevalonate into mevalonate-5-phosphate *in vitro*, in a cell extract, or *in vivo*.

**[0166]** Phosphomevalonate kinase (PMK) polypeptides phosphorylates mevalonate-5-phosphate to form mevalonate-5-diphosphate. Standard methods (such as those described herein) can be used to determine whether a polypeptide has PMK polypeptide activity by measuring the ability of the polypeptide to convert mevalonate-5-phosphate into mevalonate-5-diphosphate *in vitro*, in a cell extract, or *in vivo*.

**[0167]** Diphosphomevalonate decarboxylase (MVD or DPMDC) polypeptides convert mevalonate-5-diphosphate into isopentenyl diphosphate (IPP). Standard methods (such as those described herein) can be used to determine whether a polypeptide has MVD polypeptide activity by measuring the ability of the polypeptide to convert mevalonate-5-diphosphate into IPP *in vitro*, in a cell extract, or *in vivo*.

**[0168]** Phosphomevalonate decarboxylase (PMDC) polypeptides convert mevalonate-5-phosphate into isopentenyl phosphate (IP). Standard methods (such as those described herein) can be used to determine whether a polypeptide has PMDC polypeptide activity by measuring the ability of the polypeptide to convert mevalonate-5-phosphate into IP *in vitro*, in a cell extract, or *in vivo*.



[0169] Isopentenyl phosphate kinase (IPK) polypeptides phosphorylate isopentyl phosphate (IP) to form isopentenyl diphosphate (IPP). Standard methods (such as those described herein) can be used to determine whether a polypeptide has IPK polypeptide activity by measuring the ability of the polypeptide to convert IP into IPP *in vitro*, in a cell extract, or *in vivo*.

[0170] Exemplary IDI polypeptides and nucleic acids are described above.

*Exemplary Methods for Isolating Nucleic Acids*

[0171] Isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acids can be isolated using standard methods. Methods of obtaining desired nucleic acids from a source organism of interest (such as a bacterial genome) are common and well known in the art of molecular biology (*see*, for example, WO 2004/033646 and references cited therein, which are each hereby incorporated by reference in their entireties, particularly with respect to the isolation of nucleic acids of interest). For example, if the sequence of the nucleic acid is known (such as any of the known nucleic acids described herein), suitable genomic libraries may be created by restriction endonuclease digestion and may be screened with probes complementary to the desired nucleic acid sequence. Once the sequence is isolated, the DNA may be amplified using standard primer directed amplification methods such as polymerase chain reaction (PCR) (U.S. Patent No. 4,683,202, which is incorporated by reference in its entirety, particularly with respect to PCR methods) to obtain amounts of DNA suitable for transformation using appropriate vectors.

[0172] Alternatively, isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acids (such as any isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acids with a known nucleic acid sequence) can be chemically synthesized using standard methods.

[0173] Additional isoprene synthase, DXS, IDI, or MVA pathway polypeptides and nucleic acids which may be suitable for use in the compositions and methods described herein can be identified using standard methods. For example, cosmid libraries of the chromosomal DNA of organisms known to produce isoprene naturally can be constructed in organisms such as *E. coli*, and then screened for isoprene production. In particular, cosmid libraries may be created where large segments of genomic DNA (35-45 kb) are packaged into vectors and used to transform appropriate hosts. Cosmid vectors are unique in being able to accommodate large quantities of DNA. Generally cosmid vectors have at least one copy of the cos DNA sequence which is



needed for packaging and subsequent circularization of the heterologous DNA. In addition to the cos sequence, these vectors also contain an origin of replication such as ColEI and drug resistance markers such as a nucleic acid resistant to ampicillin or neomycin. Methods of using cosmid vectors for the transformation of suitable bacterial hosts are well described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2<sup>nd</sup> ed., Cold Spring Harbor, 1989, which is hereby incorporated by reference in its entirety, particularly with respect to transformation methods.

[0174] Typically to clone cosmids, heterologous DNA is isolated using the appropriate restriction endonucleases and ligated adjacent to the cos region of the cosmid vector using the appropriate ligases. Cosmid vectors containing the linearized heterologous DNA are then reacted with a DNA packaging vehicle such as bacteriophage. During the packaging process, the cos sites are cleaved and the heterologous DNA is packaged into the head portion of the bacterial viral particle. These particles are then used to transfect suitable host cells such as *E. coli*. Once injected into the cell, the heterologous DNA circularizes under the influence of the cos sticky ends. In this manner, large segments of heterologous DNA can be introduced and expressed in host cells.

[0175] Additional methods for obtaining isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acids include screening a metagenomic library by assay (such as the headspace assay described herein) or by PCR using primers directed against nucleotides encoding for a length of conserved amino acids (for example, at least 3 conserved amino acids). Conserved amino acids can be identified by aligning amino acid sequences of known isoprene synthase, DXS, IDI, and/or MVA pathway polypeptides. Conserved amino acids for isoprene synthase polypeptides can be identified based on aligned sequences of known isoprene synthase polypeptides. An organism found to produce isoprene naturally can be subjected to standard protein purification methods (which are well known in the art) and the resulting purified polypeptide can be sequenced using standard methods. Other methods are found in the literature (*see*, for example, Julsing *et al.*, *Applied. Microbiol. Biotechnol.* 75: 1377-84, 2007; Withers *et al.*, *Appl Environ Microbiol.* 73(19):6277-83, 2007, which are each hereby incorporated by reference in their entireties, particularly with respect to identification of nucleic acids involved in the synthesis of isoprene).



[0176] Additionally, standard sequence alignment and/or structure prediction programs can be used to identify additional DXS, IDI, or MVA pathway polypeptides and nucleic acids based on the similarity of their primary and/or predicted polypeptide secondary structure with that of known DXS, IDI, or MVA pathway polypeptides and nucleic acids. Standard databases such as the swissprot-trembl database (world-wide web at “expasy.org”, Swiss Institute of Bioinformatics Swiss-Prot group CMU - 1 rue Michel Servet CH-1211 Geneva 4, Switzerland) can also be used to identify isoprene synthase, DXS, IDI, or MVA pathway polypeptides and nucleic acids. The secondary and/or tertiary structure of an isoprene synthase, DXS, IDI, or MVA pathway polypeptide can be predicted using the default settings of standard structure prediction programs, such as PredictProtein (630 West, 168 Street, BB217, New York, N.Y. 10032, USA). Alternatively, the actual secondary and/or tertiary structure of an isoprene synthase, DXS, IDI, or MVA pathway polypeptide can be determined using standard methods. Additional isoprene synthase, DXS, IDI, or MVA pathway nucleic acids can also be identified by hybridization to probes generated from known isoprene synthase, DXS, IDI, or MVA pathway nucleic acids.

#### *Exemplary Promoters and Vectors*

[0177] Any of the isoprene synthase, DXS, IDI, or MVA pathway nucleic acid described herein can be included in one or more vectors. Accordingly, the invention also features vectors with one more nucleic acids encoding any of the isoprene synthase, DXS, IDI, or MVA pathway polypeptides that are described herein. As used herein, a “vector” means a construct that is capable of delivering, and desirably expressing one or more nucleic acids of interest in a host cell. Examples of vectors include, but are not limited to, plasmids, viral vectors, DNA or RNA expression vectors, cosmids, and phage vectors. In some embodiments, the vector contains a nucleic acid under the control of an expression control sequence.

[0178] As used herein, an “expression control sequence” means a nucleic acid sequence that directs transcription of a nucleic acid of interest. An expression control sequence can be a promoter, such as a constitutive or an inducible promoter, or an enhancer. An “inducible promoter” is a promoter that is active under environmental or developmental regulation. The expression control sequence is operably linked to the nucleic acid segment to be transcribed.



[0179] In some embodiments, the vector contains a selective marker. The term “selective marker” refers to a nucleic acid capable of expression in a host cell that allows for ease of selection of those host cells containing an introduced nucleic acid or vector. Examples of selectable markers include, but are not limited to, antibiotic resistance nucleic acids (*e.g.*, kanamycin, ampicillin, carbenicillin, gentamicin, hygromycin, phleomycin, bleomycin, neomycin, or chloramphenicol) and/or nucleic acids that confer a metabolic advantage, such as a nutritional advantage on the host cell. Exemplary nutritional selective markers include those markers known in the art as *amdS*, *argB*, and *pyr4*. Markers useful in vector systems for transformation of *Trichoderma* are known in the art (*see, e.g.*, Finkelstein, Chapter 6 in *Biotechnology of Filamentous Fungi*, Finkelstein *et al.*, Eds. Butterworth-Heinemann, Boston, MA, Chap. 6., 1992; and Kinghorn *et al.*, *Applied Molecular Genetics of Filamentous Fungi*, Blackie Academic and Professional, Chapman and Hall, London, 1992, which are each hereby incorporated by reference in their entireties, particularly with respect to selective markers). In some embodiments, the selective marker is the *amdS* nucleic acid, which encodes the enzyme acetamidase, allowing transformed cells to grow on acetamide as a nitrogen source. The use of an *A. nidulans amdS* nucleic acid as a selective marker is described in Kelley *et al.*, *EMBO J.* 4:475 – 479, 1985 and Penttila *et al.*, *Gene* 61:155-164, 1987 (which are each hereby incorporated by reference in their entireties, particularly with respect to selective markers). In some embodiments, an isoprene synthase, DXS, IDI, or MVA pathway nucleic acid integrates into a chromosome of the cells without a selective marker.

[0180] Suitable vectors are those which are compatible with the host cell employed. Suitable vectors can be derived, for example, from a bacterium, a virus (such as bacteriophage T7 or a M-13 derived phage), a cosmid, a yeast, or a plant. Protocols for obtaining and using such vectors are known to those in the art (*see, for example*, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2<sup>nd</sup> ed., Cold Spring Harbor, 1989, which is hereby incorporated by reference in its entirety, particularly with respect to the use of vectors).

[0181] Promoters are well known in the art. Any promoter that functions in the host cell can be used for expression of an isoprene synthase, DXS, IDI, or MVA pathway nucleic acid in the host cell. Initiation control regions or promoters, which are useful to drive expression of isoprene synthase, DXS, IDI, or MVA pathway nucleic acids in various host cells are numerous



and familiar to those skilled in the art (*see*, for example, WO 2004/033646 and references cited therein, which are each hereby incorporated by reference in their entireties, particularly with respect to vectors for the expression of nucleic acids of interest). Virtually any promoter capable of driving these nucleic acids is suitable for the present invention including, but not limited to, CYC1, HIS3, GAL1, GAL10, ADH1, PGK, PHO5, GAPDH, ADCI, TRP1, URA3, LEU2, ENO, and TPI (useful for expression in *Saccharomyces*); AOX1 (useful for expression in *Pichia*); and lac, trp,  $\lambda P_L$ ,  $\lambda P_R$ , T7, tac, and trc (useful for expression in *E. coli*).

[0182] In some embodiments, a glucose isomerase promoter is used (*see*, for example, U.S. Patent No. 7,132,527 and references cited therein, which are each hereby incorporated by reference in their entireties, particularly with respect promoters and plasmid systems for expressing polypeptides of interest). Reported glucose isomerase promoter mutants can be used to vary the level of expression of the polypeptide encoded by a nucleic acid operably linked to the glucose isomerase promoter (U.S. Patent No. 7,132,527). In various embodiments, the glucose isomerase promoter is contained in a low, medium, or high copy plasmid (U.S. Patent No. 7,132,527).

[0183] In various embodiments, an isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acid is contained in a low copy plasmid (*e.g.*, a plasmid that is maintained at about 1 to about 4 copies per cell), medium copy plasmid (*e.g.*, a plasmid that is maintained at about 10 to about 15 copies per cell), or high copy plasmid (*e.g.*, a plasmid that is maintained at about 50 or more copies per cell). In some embodiments, the heterologous or extra endogenous isoprene synthase, DXS, IDI, or MVA pathway nucleic acid is operably linked to a T7 promoter. In some embodiments, the heterologous or extra endogenous isoprene synthase, DXS, IDI, or MVA pathway nucleic acid operably linked to a T7 promoter is contained in a medium or high copy plasmid. In some embodiments, the heterologous or extra endogenous isoprene synthase, DXS, IDI, or MVA pathway nucleic acid is operably linked to a Trc promoter. In some embodiments, the heterologous or extra endogenous isoprene synthase, DXS, IDI, or MVA pathway nucleic acid operably linked to a Trc promoter is contained in a medium or high copy plasmid. In some embodiments, the heterologous or extra endogenous isoprene synthase, DXS, IDI, or MVA pathway nucleic acid is operably linked to a Lac promoter. In some embodiments, the heterologous or extra endogenous isoprene synthase, DXS, IDI, or MVA pathway nucleic acid



operably linked to a Lac promoter is contained in a low copy plasmid. In some embodiments, the heterologous or extra endogenous isoprene synthase, DXS, IDI, or MVA pathway nucleic acid is operably linked to an endogenous promoter, such as an endogenous *Escherichia*, *Pantoea*, *Bacillus*, *Yarrowia*, *Streptomyces*, or *Trichoderma* promoter or an endogenous alkaline serine protease, isoprene synthase, DXS, IDI, or MVA pathway promoter. In some embodiments, the heterologous or extra endogenous isoprene synthase, DXS, IDI, or MVA pathway nucleic acid operably linked to an endogenous promoter is contained in a high copy plasmid. In some embodiments, the vector is a replicating plasmid that does not integrate into a chromosome in the cells. In some embodiments, part or all of the vector integrates into a chromosome in the cells.

[0184] In some embodiments, the vector is any vector which when introduced into a fungal host cell is integrated into the host cell genome and is replicated. Reference is made to the Fungal Genetics Stock Center Catalogue of Strains (FGSC, the world-wide web at “fgsc.net” and the references cited therein, which are each hereby incorporated by reference in their entireties, particularly with respect to vectors) for a list of vectors. Additional examples of suitable expression and/or integration vectors are provided in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., Cold Spring Harbor, 1989, Current Protocols in Molecular Biology (F. M. Ausubel *et al.* (eds) 1987, Supplement 30, section 7.7.18); van den Hondel *et al.* in Bennett and Lasure (Eds.) More Gene Manipulations in Fungi, Academic Press pp. 396-428, 1991; and U.S. Patent No. 5,874,276, which are each hereby incorporated by reference in their entireties, particularly with respect to vectors. Particularly useful vectors include pFB6, pBR322, PUC18, pUC100, and pENTR/D.

[0185] In some embodiments, an isoprene synthase, DXS, IDI, or MVA pathway nucleic acid is operably linked to a suitable promoter that shows transcriptional activity in a fungal host cell. The promoter may be derived from one or more nucleic acids encoding a polypeptide that is either endogenous or heterologous to the host cell. In some embodiments, the promoter is useful in a *Trichoderma* host. Suitable non-limiting examples of promoters include *cbh1*, *cbh2*, *egl1*, *egl2*, *pepA*, *hfb1*, *hfb2*, *xyn1*, and *amy*. In some embodiments, the promoter is one that is native to the host cell. For example, in some embodiments when *T. reesei* is the host, the promoter is a native *T. reesei* promoter. In some embodiments, the promoter is *T. reesei cbh1*, which is an



inducible promoter and has been deposited in GenBank under Accession No. D86235, which is incorporated by reference in its entirety, particularly with respect to promoters. In some embodiments, the promoter is one that is heterologous to the fungal host cell. Other examples of useful promoters include promoters from the genes of *A. awamori* and *A. niger* glucoamylase (glaA) (Nunberg *et al.*, *Mol. Cell Biol.* 4:2306-2315, 1984 and Boel *et al.*, *EMBO J.* 3:1581-1585, 1984, which are each hereby incorporated by reference in their entireties, particularly with respect to promoters); *Aspergillus niger* alpha amylases, *Aspergillus oryzae* TAKA amylase, *T. reesei xln1*, and the *T. reesei cellobiohydrolase 1* (EP 137280, which is incorporated by reference in its entirety, particularly with respect to promoters).

[0186] In some embodiments, the expression vector also includes a termination sequence. Termination control regions may also be derived from various genes native to the host cell. In some embodiments, the termination sequence and the promoter sequence are derived from the same source. In another embodiment, the termination sequence is endogenous to the host cell. A particularly suitable terminator sequence is *cbh1* derived from a *Trichoderma* strain (such as *T. reesei*). Other useful fungal terminators include the terminator from an *A. niger* or *A. awamori* glucoamylase nucleic acid (Nunberg *et al.*, *Mol. Cell Biol.* 4:2306-2315, 1984 and Boel *et al.*, *EMBO J.* 3:1581-1585, 1984; which are each hereby incorporated by reference in their entireties, particularly with respect to fungal terminators). Optionally, a termination site may be included. For effective expression of the polypeptides, DNA encoding the polypeptide are linked operably through initiation codons to selected expression control regions such that expression results in the formation of the appropriate messenger RNA.

[0187] In some embodiments, the promoter, coding, region, and terminator all originate from the isoprene synthase, DXS, IDI, or MVA pathway nucleic acid to be expressed. In some embodiments, the coding region for an isoprene synthase, DXS, IDI, or MVA pathway nucleic acid is inserted into a general-purpose expression vector such that it is under the transcriptional control of the expression construct promoter and terminator sequences. In some embodiments, genes or part thereof are inserted downstream of the strong *cbh1* promoter.

[0188] An isoprene synthase, DXS, IDI, or MVA pathway nucleic acid can be incorporated into a vector, such as an expression vector, using standard techniques (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, 1982, which is hereby



incorporated by reference in its entirety, particularly with respect to the screening of appropriate DNA sequences and the construction of vectors). Methods used to ligate the DNA construct comprising a nucleic acid of interest (such as an isoprene synthase, DXS, IDI, or MVA pathway nucleic acid), a promoter, a terminator, and other sequences and to insert them into a suitable vector are well known in the art. For example, restriction enzymes can be used to cleave the isoprene synthase, DXS, IDI, or MVA pathway nucleic acid and the vector. Then, the compatible ends of the cleaved isoprene synthase, DXS, IDI, or MVA pathway nucleic acid and the cleaved vector can be ligated. Linking is generally accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide linkers are used in accordance with conventional practice (*see*, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., Cold Spring Harbor, 1989, and Bennett and Lasure, More Gene Manipulations in Fungi, Academic Press, San Diego, pp 70–76, 1991, which are each hereby incorporated by reference in their entireties, particularly with respect to oligonucleotide linkers). Additionally, vectors can be constructed using known recombination techniques (*e.g.*, Invitrogen Life Technologies, Gateway Technology).

**[0189]** In some embodiments, it may be desirable to over-express isoprene synthase, DXS, IDI, or MVA pathway nucleic acids at levels far higher than currently found in naturally-occurring cells. This result may be accomplished by the selective cloning of the nucleic acids encoding those polypeptides into multicopy plasmids or placing those nucleic acids under a strong inducible or constitutive promoter. Methods for over-expressing desired polypeptides are common and well known in the art of molecular biology and examples may be found in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., Cold Spring Harbor, 1989, which is hereby incorporated by reference in its entirety, particularly with respect to cloning techniques.

**[0190]** The following resources include descriptions of additional general methodology useful in accordance with the invention: Kreigler, Gene Transfer and Expression; A Laboratory Manual, 1990 and Ausubel *et al.*, Eds. Current Protocols in Molecular Biology, 1994, which are each hereby incorporated by reference in their entireties, particularly with respect to molecular biology and cloning techniques.

#### *Exemplary Source Organisms*



[0191] Isoprene synthase, DXS, IDI, or MVA pathway nucleic acids (and their encoded polypeptides) can be obtained from any organism that naturally contains isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acids. As noted above, isoprene is formed naturally by a variety of organisms, such as bacteria, yeast, plants, and animals. Organisms contain the MVA pathway, DXP pathway, or both the MVA and DXP pathways for producing isoprene (Figure 19). Thus, DXS nucleic acids can be obtained, *e.g.*, from any organism that contains the DXP pathway or contains both the MVA and DXP pathways. IDI and isoprene synthase nucleic acids can be obtained, *e.g.*, from any organism that contains the MVA pathway, DXP pathway, or both the MVA and DXP pathways. MVA pathway nucleic acids can be obtained, *e.g.*, from any organism that contains the MVA pathway or contains both the MVA and DXP pathways.

[0192] In some embodiments, the nucleic acid sequence of the isoprene synthase, DXS, IDI, or MVA pathway nucleic acid is identical to the sequence of a nucleic acid that is produced by any of the following organisms in nature. In some embodiments, the amino acid sequence of the isoprene synthase, DXS, IDI, or MVA pathway polypeptide is identical to the sequence of a polypeptide that is produced by any of the following organisms in nature. In some embodiments, the isoprene synthase, DXS, IDI, or MVA pathway nucleic acid or polypeptide is a mutant nucleic acid or polypeptide derived from any of the organisms described herein. As used herein, “derived from” refers to the source of the nucleic acid or polypeptide into which one or more mutations is introduced. For example, a polypeptide that is “derived from a plant polypeptide” refers to polypeptide of interest that results from introducing one or more mutations into the sequence of a wild-type (*i.e.*, a sequence occurring in nature) plant polypeptide.

[0193] In some embodiments, the source organism is a fungus, examples of which are species of *Aspergillus* such as *A. oryzae* and *A. niger*, species of *Saccharomyces* such as *S. cerevisiae*, species of *Schizosaccharomyces* such as *S. pombe*, and species of *Trichoderma* such as *T. reesei*. In some embodiments, the source organism is a filamentous fungal cell. The term “filamentous fungi” refers to all filamentous forms of the subdivision Eumycotina (*see*, Alexopoulos, C. J. (1962), *Introductory Mycology*, Wiley, New York). These fungi are characterized by a vegetative mycelium with a cell wall composed of chitin, cellulose, and other complex polysaccharides. The filamentous fungi are morphologically, physiologically, and genetically



distinct from yeasts. Vegetative growth by filamentous fungi is by hyphal elongation and carbon catabolism is obligatory aerobic. The filamentous fungal parent cell may be a cell of a species of, but not limited to, *Trichoderma*, (e.g., *Trichoderma reesei*, the asexual morph of *Hypocrea jecorina*, previously classified as *T. longibrachiatum*, *Trichoderma viride*, *Trichoderma koningii*, *Trichoderma harzianum*) (Sheir-Neirs *et al.*, Appl. Microbiol. Biotechnol 20: 46-53, 1984; ATCC No. 56765 and ATCC No. 26921); *Penicillium sp.*, *Humicola sp.* (e.g., *H. insolens*, *H. lanuginosa*, or *H. grisea*); *Chrysosporium sp.* (e.g., *C. lucknowense*), *Gliocladium sp.*, *Aspergillus sp.* (e.g., *A. oryzae*, *A. niger*, *A. sojae*, *A. japonicus*, *A. nidulans*, or *A. awamori*) (Ward *et al.*, Appl. Microbiol. Biotechnol. 39: 7380743, 1993 and Goedegebuur *et al.*, Genet 41: 89-98, 2002), *Fusarium sp.*, (e.g., *F. roseum*, *F. gramineum*, *F. cerealis*, *F. oxysporum*, or *F. venenatum*), *Neurospora sp.*, (e.g., *N. crassa*), *Hypocrea sp.*, *Mucor sp.*, (e.g., *M. miehei*), *Rhizopus sp.* and *Emericella sp.* (see also, Innis *et al.*, Sci. 228: 21-26, 1985). The term “*Trichoderma*” or “*Trichoderma sp.*” or “*Trichoderma spp.*” refer to any fungal genus previously or currently classified as *Trichoderma*.

[0194] In some embodiments, the fungus is *A. nidulans*, *A. awamori*, *A. oryzae*, *A. aculeatus*, *A. niger*, *A. japonicus*, *T. reesei*, *T. viride*, *F. oxysporum*, or *F. solani*. *Aspergillus* strains are disclosed in Ward *et al.*, Appl. Microbiol. Biotechnol. 39:738-743, 1993 and Goedegebuur *et al.*, Curr Gene 41:89-98, 2002, which are each hereby incorporated by reference in their entireties, particularly with respect to fungi. In particular embodiments, the fungus is a strain of *Trichoderma*, such as a strain of *T. reesei*. Strains of *T. reesei* are known and non-limiting examples include ATCC No. 13631, ATCC No. 26921, ATCC No. 56764, ATCC No. 56765, ATCC No. 56767, and NRRL 15709, which are each hereby incorporated by reference in their entireties, particularly with respect to strains of *T. reesei*. In some embodiments, the host strain is a derivative of RL-P37. RL-P37 is disclosed in Sheir-Neiss *et al.*, Appl. Microbiol. Biotechnology 20:46–53, 1984, which is hereby incorporated by reference in its entirety, particularly with respect to strains of *T. reesei*.

[0195] In some embodiments, the source organism is a yeast, such as *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Pichia sp.*, or *Candida sp.*

[0196] In some embodiments, the source organism is a bacterium, such as strains of *Bacillus* such as *B. licheniformis* or *B. subtilis*, strains of *Pantoea* such as *P. citrea*, strains of



*Pseudomonas* such as *P. alcaligenes*, strains of *Streptomyces* such as *S. albus*, *S. lividans*, or *S. rubiginosus*, or strains of *Escherichia* such as *E. coli*.

[0197] As used herein, “the genus *Bacillus*” includes all species within the genus “*Bacillus*,” as known to those of skill in the art, including but not limited to *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. lautus*, and *B. thuringiensis*. It is recognized that the genus *Bacillus* continues to undergo taxonomical reorganization. Thus, it is intended that the genus include species that have been reclassified, including but not limited to such organisms as *B. stearothermophilus*, which is now named “*Geobacillus stearothermophilus*.” The production of resistant endospores in the presence of oxygen is considered the defining feature of the genus *Bacillus*, although this characteristic also applies to the recently named *Alicyclobacillus*, *Amphibacillus*, *Aneurinibacillus*, *Anoxybacillus*, *Brevibacillus*, *Filobacillus*, *Gracilibacillus*, *Halobacillus*, *Paenibacillus*, *Salibacillus*, *Thermobacillus*, *Ureibacillus*, and *Virgibacillus*.

[0198] In some embodiments, the source organism is a gram-positive bacterium. Non-limiting examples include strains of *Streptomyces* (e.g., *S. albus*, *S. lividans*, *S. coelicolor*, or *S. griseus*) and *Bacillus*. In some embodiments, the source organism is a gram-negative bacterium, such as *E. coli* or *Pseudomonas sp.*

[0199] In some embodiments, the source organism is a plant, such as a plant from the family Fabaceae, such as the Faboideae subfamily. In some embodiments, the source organism is kudzu, poplar (such as *Populus alba x tremula* CAC35696 or *Populus alba*) (Sasaki *et al.*, *FEBS Letters* 579(11): 2514-2518, 2005), aspen (such as *Populus tremuloides*), or *Quercus robur*.

[0200] In some embodiments, the source organism is an algae, such as a green algae, red algae, glaucophytes, chlorarachniophytes, euglenids, chromista, or dinoflagellates.

[0201] In some embodiments, the source organism is a cyanobacteria, such as cyanobacteria classified into any of the following groups based on morphology: *Chroococcales*, *Pleurocapsales*, *Oscillatoriales*, *Nostocales*, or *Stigonematales*.

*Exemplary Host Cells*



**[0202]** A variety of host cells can be used to express isoprene synthase, DXS, IDI, and/or MVA pathway polypeptides and to produce isoprene in the methods of the claimed invention. Exemplary host cells include cells from any of the organisms listed in the prior section under the heading “*Exemplary Source Organisms.*” The host cell may be a cell that naturally produces isoprene or a cell that does not naturally produce isoprene. In some embodiments, the host cell naturally produces isoprene using the DXP pathway, and an isoprene synthase, DXS, and/or IDI nucleic acid is added to enhance production of isoprene using this pathway. In some embodiments, the host cell naturally produces isoprene using the MVA pathway, and an isoprene synthase and/or one or more MVA pathway nucleic acids are added to enhance production of isoprene using this pathway. In some embodiments, the host cell naturally produces isoprene using the DXP pathway and one or more MVA pathway nucleic acids are added to produce isoprene using part or all of the MVA pathway as well as the DXP pathway. In some embodiments, the host cell naturally produces isoprene using both the DXP and MVA pathways and one or more isoprene synthase, DXS, IDI, or MVA pathway nucleic acids are added to enhance production of isoprene by one or both of these pathways.

#### *Exemplary Transformation Methods*

**[0203]** Isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acids or vectors containing them can be inserted into a host cell (*e.g.*, a plant cell, a fungal cell, a yeast cell, or a bacterial cell described herein) using standard techniques for expression of the encoded isoprene synthase, DXS, IDI, and/or MVA pathway polypeptide. Introduction of a DNA construct or vector into a host cell can be performed using techniques such as transformation, electroporation, nuclear microinjection, transduction, transfection (*e.g.*, lipofection mediated or DEAE-Dextrin mediated transfection or transfection using a recombinant phage virus), incubation with calcium phosphate DNA precipitate, high velocity bombardment with DNA-coated microprojectiles, and protoplast fusion. General transformation techniques are known in the art (*see, e.g.*, Current Protocols in Molecular Biology (F. M. Ausubel *et al.* (eds) Chapter 9, 1987; Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., Cold Spring Harbor, 1989; and Campbell *et al.*, *Curr. Genet.* 16:53-56, 1989, which are each hereby incorporated by reference in their entireties, particularly with respect to transformation methods). The expression of heterologous polypeptide in *Trichoderma* is described in U.S. Patent No. 6,022,725; U.S. Patent No.



6,268,328; U.S. Patent No. 7,262,041; WO 2005/001036; Harkki *et al.*; *Enzyme Microb. Technol.* 13:227-233, 1991; Harkki *et al.*, *Bio Technol.* 7:596-603, 1989; EP 244,234; EP 215,594; and Nevalainen *et al.*, "The Molecular Biology of *Trichoderma* and its Application to the Expression of Both Homologous and Heterologous Genes," in *Molecular Industrial Mycology*, Eds. Leong and Berka, Marcel Dekker Inc., NY pp. 129 – 148, 1992, which are each hereby incorporated by reference in their entireties, particularly with respect to transformation and expression methods). Reference is also made to Cao *et al.*, (*Sci.* 9:991–1001, 2000; EP 238023; and Yelton *et al.*, *Proceedings. Natl. Acad. Sci. USA* 81:1470-1474, 1984 (which are each hereby incorporated by reference in their entireties, particularly with respect to transformation methods) for transformation of *Aspergillus* strains. The introduced nucleic acids may be integrated into chromosomal DNA or maintained as extrachromosomal replicating sequences.

**[0204]** Any method known in the art may be used to select transformants. In one non-limiting example, stable transformants including an *amdS* marker are distinguished from unstable transformants by their faster growth rate and the formation of circular colonies with a smooth, rather than ragged outline on solid culture medium containing acetamide. Additionally, in some cases a further test of stability is conducted by growing the transformants on a solid non-selective medium (*e.g.*, a medium that lacks acetamide), harvesting spores from this culture medium, and determining the percentage of these spores which subsequently germinate and grow on selective medium containing acetamide.

**[0205]** In some embodiments, fungal cells are transformed by a process involving protoplast formation and transformation of the protoplasts followed by regeneration of the cell wall in a known manner. In one specific embodiment, the preparation of *Trichoderma sp.* for transformation involves the preparation of protoplasts from fungal mycelia (*see*, Campbell *et al.*, *Curr. Genet.* 16:53-56, 1989, which is incorporated by reference in its entirety, particularly with respect to transformation methods). In some embodiments, the mycelia are obtained from germinated vegetative spores. The mycelia are treated with an enzyme that digests the cell wall resulting in protoplasts. The protoplasts are then protected by the presence of an osmotic stabilizer in the suspending medium. These stabilizers include sorbitol, mannitol, potassium chloride, magnesium sulfate, and the like. Usually the concentration of these stabilizers varies



between 0.8 M and 1.2 M. It is desirable to use about a 1.2 M solution of sorbitol in the suspension medium.

**[0206]** Uptake of DNA into the host *Trichoderma sp.* strain is dependent upon the calcium ion concentration. Generally, between about 10 mM CaCl<sub>2</sub> and 50 mM CaCl<sub>2</sub> is used in an uptake solution. In addition to the calcium ion in the uptake solution, other compounds generally included are a buffering system such as TE buffer (10 mM Tris, pH 7.4; 1 mM EDTA) or 10 mM MOPS, pH 6.0 buffer (morpholinepropanesulfonic acid) and polyethylene glycol (PEG). While not intending to be bound to any particular theory, it is believed that the polyethylene glycol acts to fuse the cell membranes, thus permitting the contents of the medium to be delivered into the cytoplasm of the *Trichoderma sp.* strain and the plasmid DNA to be transferred to the nucleus. This fusion frequently leaves multiple copies of the plasmid DNA integrated into the host chromosome.

**[0207]** Usually a suspension containing the *Trichoderma sp.* protoplasts or cells that have been subjected to a permeability treatment at a density of 10<sup>5</sup> to 10<sup>7</sup>/mL (such as 2 x 10<sup>6</sup>/mL) are used in the transformation. A volume of 100 µL of these protoplasts or cells in an appropriate solution (e.g., 1.2 M sorbitol and 50 mM CaCl<sub>2</sub>) are mixed with the desired DNA. Generally, a high concentration of PEG is added to the uptake solution. From 0.1 to 1 volume of 25% PEG 4000 can be added to the protoplast suspension. In some embodiments, about 0.25 volumes are added to the protoplast suspension. Additives such as dimethyl sulfoxide, heparin, spermidine, potassium chloride, and the like may also be added to the uptake solution and aid in transformation. Similar procedures are available for other fungal host cells (*see, e.g.,* U.S. Patent Nos. 6,022,725 and 6,268,328, which are each hereby incorporated by reference in their entireties, particularly with respect to transformation methods).

**[0208]** Generally, the mixture is then cultured at approximately 0°C for a period of between 10 to 30 minutes. Additional PEG is then added to the mixture to further enhance the uptake of the desired nucleic acid sequence. The 25% PEG 4000 is generally added in volumes of 5 to 15 times the volume of the transformation mixture; however, greater and lesser volumes may be suitable. The 25% PEG 4000 is desirably about 10 times the volume of the transformation mixture. After the PEG is added, the transformation mixture is then cultured either at room temperature or on ice before the addition of a sorbitol and CaCl<sub>2</sub> solution. The protoplast



suspension is then further added to molten aliquots of a growth medium. When the growth medium includes a growth selection (*e.g.*, acetamide or an antibiotic) it permits the growth of transformants only.

**[0209]** The transformation of bacterial cells may be performed according to conventional methods, *e.g.*, as described in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1982, which is hereby incorporated by reference in its entirety, particularly with respect to transformation methods.

#### *Exemplary Cell Culture Media*

**[0210]** The invention also includes a cell or a population of cells in culture that produce isoprene. By “cells in culture” is meant two or more cells in a solution (*e.g.*, a cell medium) that allows the cells to undergo one or more cell divisions. “Cells in culture” do not include plant cells that are part of a living, multicellular plant containing cells that have differentiated into plant tissues. In various embodiments, the cell culture includes at least or about 10, 20, 50, 100, 200, 500, 1,000, 5,000, 10,000 or more cells.

**[0211]** Any carbon source can be used to cultivate the host cells. The term “carbon source” refers to one or more carbon-containing compounds capable of being metabolized by a host cell or organism. For example, the cell medium used to cultivate the host cells may include any carbon source suitable for maintaining the viability or growing the host cells.

**[0212]** In some embodiments, the carbon source is a carbohydrate (such as monosaccharide, disaccharide, oligosaccharide, or polysaccharids), invert sugar (*e.g.*, enzymatically treated sucrose syrup), glycerol, glycerine (*e.g.*, a glycerine byproduct of a biodiesel or soap-making process), dihydroxyacetone, one-carbon source, oil (*e.g.*, a plant or vegetable oil such as corn, palm, or soybean oil), acetate, animal fat, animal oil, fatty acid (*e.g.*, a saturated fatty acid, unsaturated fatty acid, or polyunsaturated fatty acid), lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, polypeptide (*e.g.*, a microbial or plant protein or peptide), renewable carbon source (*e.g.*, a biomass carbon source such as a hydrolyzed biomass carbon source), yeast extract, component from a yeast extract, polymer, acid, alcohol, aldehyde, ketone, amino acid, succinate, lactate, acetate, ethanol, or any combination of two or more of the foregoing. In some embodiments, the carbon source is a product of



photosynthesis, including, but not limited to, glucose. In some embodiment, the carbohydrate is xylose or glucose.

**[0213]** Exemplary monosaccharides include glucose and fructose; exemplary oligosaccharides include lactose and sucrose, and exemplary polysaccharides include starch and cellulose. Exemplary carbohydrates include C6 sugars (*e.g.*, fructose, mannose, galactose, or glucose) and C5 sugars (*e.g.*, xylose or arabinose). In some embodiments, the cell medium includes a carbohydrate as well as a carbon source other than a carbohydrate (*e.g.*, glycerol, glycerine, dihydroxyacetone, one-carbon source, oil, animal fat, animal oil, fatty acid, lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, renewable carbon source, or a component from a yeast extract). In some embodiments, the cell medium includes a carbohydrate as well as a polypeptide (*e.g.*, a microbial or plant protein or peptide). In some embodiments, the microbial polypeptide is a polypeptide from yeast or bacteria. In some embodiments, the plant polypeptide is a polypeptide from soy, corn, canola, jatropha, palm, peanut, sunflower, coconut, mustard, rapeseed, cottonseed, palm kernel, olive, safflower, sesame, or linseed.

**[0214]** In some embodiments, the concentration of the carbohydrate is at least or about 5 grams per liter of broth (g/L, wherein the volume of broth includes both the volume of the cell medium and the volume of the cells), such as at least or about 10, 15, 20, 30, 40, 50, 60, 80, 100, 150, 200, 300, 400, or more g/L. In some embodiments, the concentration of the carbohydrate is between about 50 and about 400 g/L, such as between about 100 and about 360 g/L, between about 120 and about 360 g/L, or between about 200 and about 300 g/L. In some embodiments, this concentration of carbohydrate includes the total amount of carbohydrate that is added before and/or during the culturing of the host cells.

**[0215]** Exemplary lipids are any substance containing one or more fatty acids that are C4 and above fatty acids that are saturated, unsaturated, or branched.

**[0216]** Exemplary oils are lipids that are liquid at room temperature. In some embodiments, the lipid contains one or more C4 or above fatty acids (*e.g.*, contains one or more saturated, unsaturated, or branched fatty acid with four or more carbons). In some embodiments, the oil is obtained from soy, corn, canola, jatropha, palm, peanut, sunflower, coconut, mustard, rapeseed,



cottonseed, palm kernel, olive, safflower, sesame, linseed, oleagineous microbial cells, Chinese tallow, or any combination of two or more of the foregoing.

**[0217]** Exemplary fatty acids include compounds of the formula  $\text{RCOOH}$ , where “R” is a hydrocarbon. Exemplary unsaturated fatty acids include compounds where “R” includes at least one carbon-carbon double bond. Exemplary unsaturated fatty acids include, but are not limited to, oleic acid, vaccenic acid, linoleic acid, palmitelaidic acid, and arachidonic acid. Exemplary polyunsaturated fatty acids include compounds where “R” includes a plurality of carbon-carbon double bonds. Exemplary saturated fatty acids include compounds where “R” is a saturated aliphatic group. In some embodiments, the carbon source includes one or more  $\text{C}_{12}$ - $\text{C}_{22}$  fatty acids, such as a  $\text{C}_{12}$  saturated fatty acid, a  $\text{C}_{14}$  saturated fatty acid, a  $\text{C}_{16}$  saturated fatty acid, a  $\text{C}_{18}$  saturated fatty acid, a  $\text{C}_{20}$  saturated fatty acid, or a  $\text{C}_{22}$  saturated fatty acid. In an exemplary embodiment, the fatty acid is palmitic acid. In some embodiments, the carbon source is a salt of a fatty acid (*e.g.*, an unsaturated fatty acid), a derivative of a fatty acid (*e.g.*, an unsaturated fatty acid), or a salt of a derivative of fatty acid (*e.g.*, an unsaturated fatty acid). Suitable salts include, but are not limited to, lithium salts, potassium salts, sodium salts, and the like. Di- and triglycerols are fatty acid esters of glycerol.

**[0218]** In some embodiments, the concentration of the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride is at least or about 1 gram per liter of broth (g/L, wherein the volume of broth includes both the volume of the cell medium and the volume of the cells), such as at least or about 5, 10, 15, 20, 30, 40, 50, 60, 80, 100, 150, 200, 300, 400, or more g/L. In some embodiments, the concentration of the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride is between about 10 and about 400 g/L, such as between about 25 and about 300 g/L, between about 60 and about 180 g/L, or between about 75 and about 150 g/L. In some embodiments, the concentration includes the total amount of the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride that is added before and/or during the culturing of the host cells. In some embodiments, the carbon source includes both (i) a lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride and (ii) a carbohydrate, such as glucose. In some embodiments, the ratio of the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride to the carbohydrate is about 1:1 on a carbon basis (*i.e.*, one carbon in the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride per carbohydrate carbon). In particular



embodiments, the amount of the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride is between about 60 and 180 g/L, and the amount of the carbohydrate is between about 120 and 360 g/L.

**[0219]** Exemplary microbial polypeptide carbon sources include one or more polypeptides from yeast or bacteria. Exemplary plant polypeptide carbon sources include one or more polypeptides from soy, corn, canola, jatropha, palm, peanut, sunflower, coconut, mustard, rapeseed, cottonseed, palm kernel, olive, safflower, sesame, or linseed.

**[0220]** Exemplary renewable carbon sources include cheese whey permeate, cornsteep liquor, sugar beet molasses, barley malt, and components from any of the foregoing. Exemplary renewable carbon sources also include acetate, glucose, hexose, pentose and xylose present in biomass, such as corn, switchgrass, sugar cane, cell waste of fermentation processes, and protein by-product from the milling of soy, corn, or wheat. In some embodiments, the biomass carbon source is a lignocellulosic, hemicellulosic, or cellulosic material such as, but are not limited to, a grass, wheat, wheat straw, bagasse, sugar cane bagasse, soft wood pulp, corn, corn cob or husk, corn kernel, fiber from corn kernels, corn stover, switch grass, rice hull product, or a by-product from wet or dry milling of grains (*e.g.*, corn, sorghum, rye, triticate, barley, wheat, and/or distillers grains). Exemplary cellulosic materials include wood, paper and pulp waste, herbaceous plants, and fruit pulp. In some embodiments, the carbon source includes any plant part, such as stems, grains, roots, or tubers. In some embodiments, all or part of any of the following plants are used as a carbon source: corn, wheat, rye, sorghum, triticate, rice, millet, barley, cassava, legumes, such as beans and peas, potatoes, sweet potatoes, bananas, sugarcane, and/or tapioca. In some embodiments, the carbon source is a biomass hydrolysate, such as a biomass hydrolysate that includes both xylose and glucose or that includes both sucrose and glucose.

**[0221]** In some embodiments, the renewable carbon source (such as biomass) is pretreated before it is added to the cell culture medium. In some embodiments, the pretreatment includes enzymatic pretreatment, chemical pretreatment, or a combination of both enzymatic and chemical pretreatment (*see, for example, Farzaneh et al., Bioresource Technology* 96 (18): 2014-2018, 2005; U.S. Patent No. 6,176,176; U.S. Patent No. 6,106,888; which are each hereby incorporated by reference in their entireties, particularly with respect to the pretreatment of renewable carbon



sources). In some embodiments, the renewable carbon source is partially or completely hydrolyzed before it is added to the cell culture medium.

**[0222]** In some embodiments, the renewable carbon source (such as corn stover) undergoes ammonia fiber expansion (AFEX) pretreatment before it is added to the cell culture medium (*see*, for example, Farzaneh *et al.*, *Bioresource Technology* 96 (18): 2014-2018, 2005). During AFEX pretreatment, a renewable carbon source is treated with liquid anhydrous ammonia at moderate temperatures (such as about 60 to about 100 °C) and high pressure (such as about 250 to about 300 psi) for about 5 minutes. Then, the pressure is rapidly released. In this process, the combined chemical and physical effects of lignin solubilization, hemicellulose hydrolysis, cellulose decrystallization, and increased surface area enables near complete enzymatic conversion of cellulose and hemicellulose to fermentable sugars. AFEX pretreatment has the advantage that nearly all of the ammonia can be recovered and reused, while the remaining serves as nitrogen source for microbes in downstream processes. Also, a wash stream is not required for AFEX pretreatment. Thus, dry matter recovery following the AFEX treatment is essentially 100%. AFEX is basically a dry to dry process. The treated renewable carbon source is stable for long periods and can be fed at very high solid loadings in enzymatic hydrolysis or fermentation processes. Cellulose and hemicellulose are well preserved in the AFEX process, with little or no degradation. There is no need for neutralization prior to the enzymatic hydrolysis of a renewable carbon source that has undergone AFEX pretreatment. Enzymatic hydrolysis of AFEX-treated carbon sources produces clean sugar streams for subsequent fermentation use.

**[0223]** In some embodiments, the concentration of the carbon source (*e.g.*, a renewable carbon source) is equivalent to at least or about 0.1, 0.5, 1, 1.5, 2, 3, 4, 5, 10, 15, 20, 30, 40, or 50% glucose (w/v). The equivalent amount of glucose can be determined by using standard HPLC methods with glucose as a reference to measure the amount of glucose generated from the carbon source. In some embodiments, the concentration of the carbon source (*e.g.*, a renewable carbon source) is equivalent to between about 0.1 and about 20% glucose, such as between about 0.1 and about 10% glucose, between about 0.5 and about 10% glucose, between about 1 and about 10% glucose, between about 1 and about 5% glucose, or between about 1 and about 2% glucose.



**[0224]** In some embodiments, the carbon source includes yeast extract or one or more components of yeast extract. In some embodiments, the concentration of yeast extract is at least 1 gram of yeast extract per liter of broth (g/L, wherein the volume of broth includes both the volume of the cell medium and the volume of the cells), such as at least or about 5, 10, 15, 20, 30, 40, 50, 60, 80, 100, 150, 200, 300, or more g/L. In some embodiments, the concentration of yeast extract is between about 1 and about 300 g/L, such as between about 1 and about 200 g/L, between about 5 and about 200 g/L, between about 5 and about 100 g/L, or between about 5 and about 60 g/L. In some embodiments, the concentration includes the total amount of yeast extract that is added before and/or during the culturing of the host cells. In some embodiments, the carbon source includes both yeast extract (or one or more components thereof) and another carbon source, such as glucose. In some embodiments, the ratio of yeast extract to the other carbon source is about 1:5, about 1:10, or about 1:20 (w/w).

**[0225]** Additionally the carbon source may also be one-carbon substrates such as carbon dioxide, or methanol. Glycerol production from single carbon sources (*e.g.*, methanol, formaldehyde, or formate) has been reported in methylotrophic yeasts (Yamada *et al.*, *Agric. Biol. Chem.*, 53(2) 541-543, 1989, which is hereby incorporated by reference in its entirety, particularly with respect to carbon sources) and in bacteria (Hunter *et al.*, *Biochemistry*, 24, 4148-4155, 1985, which is hereby incorporated by reference in its entirety, particularly with respect to carbon sources). These organisms can assimilate single carbon compounds, ranging in oxidation state from methane to formate, and produce glycerol. The pathway of carbon assimilation can be through ribulose monophosphate, through serine, or through xylulose-momophosphate (Gottschalk, *Bacterial Metabolism*, Second Edition, Springer-Verlag: New York, 1986, which is hereby incorporated by reference in its entirety, particularly with respect to carbon sources). The ribulose monophosphate pathway involves the condensation of formate with ribulose-5-phosphate to form a six carbon sugar that becomes fructose and eventually the three carbon product glyceraldehyde-3-phosphate. Likewise, the serine pathway assimilates the one-carbon compound into the glycolytic pathway via methylenetetrahydrofolate.

**[0226]** In addition to one and two carbon substrates, methylotrophic organisms are also known to utilize a number of other carbon containing compounds such as methylamine, glucosamine and a variety of amino acids for metabolic activity. For example, methylotrophic yeast are



known to utilize the carbon from methylamine to form trehalose or glycerol (Bellion *et al.*, *Microb. Growth Cl Compd.*, [Int. Symp.], 7<sup>th</sup> ed., 415-32. Editors: Murrell *et al.*, Publisher: Intercept, Andover, UK, 1993, which is hereby incorporated by reference in its entirety, particularly with respect to carbon sources). Similarly, various species of *Candida* metabolize alanine or oleic acid (Sulter *et al.*, *Arch. Microbiol.* 153(5), 485-9, 1990, which is hereby incorporated by reference in its entirety, particularly with respect to carbon sources).

**[0227]** In some embodiments, cells are cultured in a standard medium containing physiological salts and nutrients (*see, e.g.*, Pourquie, J. *et al.*, *Biochemistry and Genetics of Cellulose Degradation*, eds. Aubert *et al.*, Academic Press, pp. 71-86, 1988 and Ilmen *et al.*, *Appl. Environ. Microbiol.* 63:1298-1306, 1997, which are each hereby incorporated by reference in their entireties, particularly with respect to cell media). Exemplary growth media are common commercially prepared media such as Luria Bertani (LB) broth, Sabouraud Dextrose (SD) broth, or Yeast medium (YM) broth. Other defined or synthetic growth media may also be used, and the appropriate medium for growth of particular host cells are known by someone skilled in the art of microbiology or fermentation science.

**[0228]** In addition to an appropriate carbon source, the cell medium desirably contains suitable minerals, salts, cofactors, buffers, and other components known to those skilled in the art suitable for the growth of the cultures or the enhancement of isoprene production (*see, for example*, WO 2004/033646 and references cited therein and WO 96/35796 and references cited therein, which are each hereby incorporated by reference in their entireties, particularly with respect cell medias and cell culture conditions). In some embodiments where an isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acid is under the control of an inducible promoter, the inducing agent (*e.g.*, a sugar, metal salt or antimicrobial), is desirably added to the medium at a concentration effective to induce expression of an isoprene synthase, DXS, IDI, and/or MVA pathway polypeptide. In some embodiments, cell medium has an antibiotic (such as kanamycin) that corresponds to the antibiotic resistance nucleic acid (such as a kanamycin resistance nucleic acid) on a vector that has one or more DXS, IDI, or MVA pathway nucleic acids.

#### *Exemplary Cell Culture Conditions*



**[0229]** Materials and methods suitable for the maintenance and growth of bacterial cultures are well known in the art. Exemplary techniques may be found in Manual of Methods for General Bacteriology Gerhardt *et al.*, eds), American Society for Microbiology, Washington, D.C. (1994) or Brock in Biotechnology: A Textbook of Industrial Microbiology, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA, which are each hereby incorporated by reference in their entirety, particularly with respect to cell culture techniques. In some embodiments, the cells are cultured in a culture medium under conditions permitting the expression of one or more isoprene synthase, DXS, IDI, or MVA pathway polypeptides encoded by a nucleic acid inserted into the host cells.

**[0230]** Standard cell culture conditions can be used to culture the cells (*see*, for example, WO 2004/033646 and references cited therein, which are each hereby incorporated by reference in their entirety, particularly with respect to cell culture and fermentation conditions). Cells are grown and maintained at an appropriate temperature, gas mixture, and pH (such as at about 20 to about 37°C, at about 6% to about 84% CO<sub>2</sub>, and at a pH between about 5 to about 9). In some embodiments, cells are grown at 35 °C in an appropriate cell medium. In some embodiments, *e.g.*, cultures are cultured at approximately 28 °C in appropriate medium in shake cultures or fermentors until desired amount of isoprene production is achieved. In some embodiments, the pH ranges for fermentation are between about pH 5.0 to about pH 9.0 (such as about pH 6.0 to about pH 8.0 or about 6.5 to about 7.0). Reactions may be performed under aerobic, anoxic, or anaerobic conditions based on the requirements of the host cells. Exemplary culture conditions for a given filamentous fungus are known in the art and may be found in the scientific literature and/or from the source of the fungi such as the American Type Culture Collection and Fungal Genetics Stock Center.

**[0231]** In various embodiments, the cells are grown using any known mode of fermentation, such as batch, fed-batch, or continuous processes. In some embodiments, a batch method of fermentation is used. Classical batch fermentation is a closed system where the composition of the media is set at the beginning of the fermentation and is not subject to artificial alterations during the fermentation. Thus, at the beginning of the fermentation the cell medium is inoculated with the desired host cells and fermentation is permitted to occur adding nothing to the system. Typically, however, “batch” fermentation is batch with respect to the addition of



carbon source and attempts are often made at controlling factors such as pH and oxygen concentration. In batch systems, the metabolite and biomass compositions of the system change constantly until the time the fermentation is stopped. Within batch cultures, cells moderate through a static lag phase to a high growth log phase and finally to a stationary phase where growth rate is diminished or halted. In some embodiments, cells in log phase are responsible for the bulk of the isoprene production. In some embodiments, cells in stationary phase produce isoprene.

[0232] In some embodiments, a variation on the standard batch system is used, such as the Fed-Batch system. Fed-Batch fermentation processes comprise a typical batch system with the exception that the carbon source is added in increments as the fermentation progresses. Fed-Batch systems are useful when catabolite repression is apt to inhibit the metabolism of the cells and where it is desirable to have limited amounts of carbon source in the cell medium. Fed-batch fermentations may be performed with the carbon source (*e.g.*, glucose) in a limited or excess amount. Measurement of the actual carbon source concentration in Fed-Batch systems is difficult and is therefore estimated on the basis of the changes of measurable factors such as pH, dissolved oxygen, and the partial pressure of waste gases such as CO<sub>2</sub>. Batch and Fed-Batch fermentations are common and well known in the art and examples may be found in Brock, *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition (1989) Sinauer Associates, Inc., which is hereby incorporated by reference in its entirety, particularly with respect to cell culture and fermentation conditions.

[0233] In some embodiments, continuous fermentation methods are used. Continuous fermentation is an open system where a defined fermentation medium is added continuously to a bioreactor and an equal amount of conditioned medium is removed simultaneously for processing. Continuous fermentation generally maintains the cultures at a constant high density where cells are primarily in log phase growth.

[0234] Continuous fermentation allows for the modulation of one factor or any number of factors that affect cell growth or isoprene production. For example, one method maintains a limiting nutrient such as the carbon source or nitrogen level at a fixed rate and allows all other parameters to moderate. In other systems, a number of factors affecting growth can be altered continuously while the cell concentration (*e.g.*, the concentration measured by media turbidity)



is kept constant. Continuous systems strive to maintain steady state growth conditions. Thus, the cell loss due to media being drawn off is balanced against the cell growth rate in the fermentation. Methods of modulating nutrients and growth factors for continuous fermentation processes as well as techniques for maximizing the rate of product formation are well known in the art of industrial microbiology and a variety of methods are detailed by Brock, *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition (1989) Sinauer Associates, Inc., which is hereby incorporated by reference in its entirety, particularly with respect to cell culture and fermentation conditions.

**[0235]** In some embodiments, cells are immobilized on a substrate as whole cell catalysts and subjected to fermentation conditions for isoprene production.

**[0236]** In some embodiments, bottles of liquid culture are placed in shakers in order to introduce oxygen to the liquid and maintain the uniformity of the culture. In some embodiments, an incubator is used to control the temperature, humidity, shake speed, and/or other conditions in which a culture is grown. The simplest incubators are insulated boxes with an adjustable heater, typically going up to ~65 °C. More elaborate incubators can also include the ability to lower the temperature (via refrigeration), or the ability to control humidity or CO<sub>2</sub> levels. Most incubators include a timer; some can also be programmed to cycle through different temperatures, humidity levels, *etc.* Incubators can vary in size from tabletop to units the size of small rooms.

**[0237]** If desired, a portion or all of the cell medium can be changed to replenish nutrients and/or avoid the build up of potentially harmful metabolic byproducts and dead cells. In the case of suspension cultures, cells can be separated from the media by centrifuging or filtering the suspension culture and then resuspending the cells in fresh media. In the case of adherent cultures, the media can be removed directly by aspiration and replaced. In some embodiments, the cell medium allows at least a portion of the cells to divide for at least or about 5, 10, 20, 40, 50, 60, 65, or more cell divisions in a continuous culture (such as a continuous culture without dilution).

**[0238]** In some embodiments, a constitutive or leaky promoter (such as a Trc promoter) is used and a compound (such as IPTG) is not added to induce expression of the isoprene synthase,



DXS, IDI, or MVA pathway nucleic acid(s) operably linked to the promoter. In some embodiments, a compound (such as IPTG) is added to induce expression of the isoprene synthase, DXS, IDI, or MVA pathway nucleic acid(s) operably linked to the promoter.

#### *Exemplary Production of Isoprene*

[0239] In some embodiments, the cells are cultured in a culture medium under conditions permitting the production of isoprene by the cells. By “peak absolute productivity” is meant the maximum absolute amount of isoprene in the off-gas during the culturing of cells for a particular period of time (*e.g.*, the culturing of cells during a particular fermentation run). By “peak absolute productivity time point” is meant the time point during a fermentation run when the absolute amount of isoprene in the off-gas is at a maximum during the culturing of cells for a particular period of time (*e.g.*, the culturing of cells during a particular fermentation run). In some embodiments, the isoprene amount is measured at the peak absolute productivity time point. In some embodiments, the peak absolute productivity for the cells is about any of the isoprene amounts disclosed herein.

[0240] By “peak specific productivity” is meant the maximum amount of isoprene produced per cell during the culturing of cells for a particular period of time (*e.g.*, the culturing of cells during a particular fermentation run). By “peak specific productivity time point” is meant the time point during the culturing of cells for a particular period of time (*e.g.*, the culturing of cells during a particular fermentation run) when the amount of isoprene produced per cell is at a maximum. The peak specific productivity is determined by dividing the total productivity by the amount of cells, as determined by optical density at 600nm (OD<sub>600</sub>). In some embodiments, the isoprene amount is measured at the peak specific productivity time point. In some embodiments, the peak specific productivity for the cells is about any of the isoprene amounts per cell disclosed herein.

[0241] By “peak volumetric productivity” is meant the maximum amount of isoprene produced per volume of broth (including the volume of the cells and the cell medium) during the culturing of cells for a particular period of time (*e.g.*, the culturing of cells during a particular fermentation run). By “peak specific volumetric productivity time point” is meant the time point during the culturing of cells for a particular period of time (*e.g.*, the culturing of cells during a



particular fermentation run) when the amount of isoprene produced per volume of broth is at a maximum. The peak specific volumetric productivity is determined by dividing the total productivity by the volume of broth and amount of time. In some embodiments, the isoprene amount is measured at the peak specific volumetric productivity time point. In some embodiments, the peak specific volumetric productivity for the cells is about any of the isoprene amounts per volume per time disclosed herein.

[0242] By “peak concentration” is meant the maximum amount of isoprene produced during the culturing of cells for a particular period of time (*e.g.*, the culturing of cells during a particular fermentation run). By “peak concentration time point” is meant the time point during the culturing of cells for a particular period of time (*e.g.*, the culturing of cells during a particular fermentation run) when the amount of isoprene produced per cell is at a maximum. In some embodiments, the isoprene amount is measured at the peak concentration time point. In some embodiments, the peak concentration for the cells is about any of the isoprene amounts disclosed herein.

[0243] By “average volumetric productivity” is meant the average amount of isoprene produced per volume of broth (including the volume of the cells and the cell medium) during the culturing of cells for a particular period of time (*e.g.*, the culturing of cells during a particular fermentation run). The average volumetric productivity is determined by dividing the total productivity by the volume of broth and amount of time. In some embodiments, the average specific volumetric productivity for the cells is about any of the isoprene amounts per volume per time disclosed herein.

[0244] By “cumulative total productivity” is meant the cumulative, total amount of isoprene produced during the culturing of cells for a particular period of time (*e.g.*, the culturing of cells during a particular fermentation run). In some embodiments, the cumulative, total amount of isoprene is measured. In some embodiments, the cumulative total productivity for the cells is about any of the isoprene amounts disclosed herein.

[0245] In some embodiments, the cells in culture produce isoprene at greater than or about 1, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, 10,000, 12,500, 20,000, 30,000, 40,000, 50,000, 75,000,



100,000, 125,000, 150,000, 188,000, or more nmole of isoprene/gram of cells for the wet weight of the cells/hour (nmole/g<sub>wcm</sub>/hr). In some embodiments, the amount of isoprene is between about 2 to about 200,000 nmole/g<sub>wcm</sub>/hr, such as between about 2 to about 100 nmole/g<sub>wcm</sub>/hr, about 100 to about 500 nmole/g<sub>wcm</sub>/hr, about 150 to about 500 nmole/g<sub>wcm</sub>/hr, about 500 to about 1,000 nmole/g<sub>wcm</sub>/hr, about 1,000 to about 2,000 nmole/g<sub>wcm</sub>/hr, about 2,000 to about 5,000 nmole/g<sub>wcm</sub>/hr, about 5,000 to about 10,000 nmole/g<sub>wcm</sub>/hr, about 10,000 to about 50,000 nmole/g<sub>wcm</sub>/hr, about 50,000 to about 100,000 nmole/g<sub>wcm</sub>/hr, about 100,000 to about 150,000 nmole/g<sub>wcm</sub>/hr, or about 150,000 to about 200,000 nmole/g<sub>wcm</sub>/hr. In some embodiments, the amount of isoprene is between about 20 to about 200,000 nmole/g<sub>wcm</sub>/hr, about 100 to about 5,000 nmole/g<sub>wcm</sub>/hr, about 200 to about 2,000 nmole/g<sub>wcm</sub>/hr, about 200 to about 1,000 nmole/g<sub>wcm</sub>/hr, about 300 to about 1,000 nmole/g<sub>wcm</sub>/hr, about 400 to about 1,000 nmole/g<sub>wcm</sub>/hr, about 1,000 to about 5,000 nmole/g<sub>wcm</sub>/hr, about 2,000 to about 20,000 nmole/g<sub>wcm</sub>/hr, about 5,000 to about 50,000 nmole/g<sub>wcm</sub>/hr, about 10,000 to about 100,000 nmole/g<sub>wcm</sub>/hr, about 20,000 to about 150,000 nmole/g<sub>wcm</sub>/hr, or about 20,000 to about 200,000 nmole/g<sub>wcm</sub>/hr.

**[0246]** The amount of isoprene in units of nmole/g<sub>wcm</sub>/hr can be measured as disclosed in U.S. Patent No. 5,849,970, which is hereby incorporated by reference in its entirety, particularly with respect to the measurement of isoprene production. For example, two mL of headspace (*e.g.*, headspace from a culture such as 2 mL of culture cultured in sealed vials at 32 °C with shaking at 200 rpm for approximately 3 hours) are analyzed for isoprene using a standard gas chromatography system, such as a system operated isothermally (85 °C) with an n-octane/porasil C column (Alltech Associates, Inc., Deerfield, Ill.) and coupled to a RGD2 mercuric oxide reduction gas detector (Trace Analytical, Menlo Park, CA) (*see*, for example, Greenberg *et al*, *Atmos. Environ.* 27A: 2689-2692, 1993; Silver *et al.*, *Plant Physiol.* 97:1588-1591, 1991, which are each hereby incorporated by reference in their entireties, particularly with respect to the measurement of isoprene production). The gas chromatography area units are converted to nmol isoprene via a standard isoprene concentration calibration curve. In some embodiments, the value for the grams of cells for the wet weight of the cells is calculated by obtaining the A<sub>600</sub> value for a sample of the cell culture, and then converting the A<sub>600</sub> value to grams of cells based on a calibration curve of wet weights for cell cultures with a known A<sub>600</sub> value. In some embodiments, the grams of the cells is estimated by assuming that one liter of broth (including



cell medium and cells) with an  $A_{600}$  value of 1 has a wet cell weight of 1 gram. The value is also divided by the number of hours the culture has been incubating for, such as three hours.

**[0247]** In some embodiments, the cells in culture produce isoprene at greater than or about 1, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, 10,000, 100,000, or more ng of isoprene/gram of cells for the wet weight of the cells/hr ( $\text{ng/g}_{\text{wcm}}/\text{h}$ ). In some embodiments, the amount of isoprene is between about 2 to about 5,000  $\text{ng/g}_{\text{wcm}}/\text{h}$ , such as between about 2 to about 100  $\text{ng/g}_{\text{wcm}}/\text{h}$ , about 100 to about 500  $\text{ng/g}_{\text{wcm}}/\text{h}$ , about 500 to about 1,000  $\text{ng/g}_{\text{wcm}}/\text{h}$ , about 1,000 to about 2,000  $\text{ng/g}_{\text{wcm}}/\text{h}$ , or about 2,000 to about 5,000  $\text{ng/g}_{\text{wcm}}/\text{h}$ . In some embodiments, the amount of isoprene is between about 20 to about 5,000  $\text{ng/g}_{\text{wcm}}/\text{h}$ , about 100 to about 5,000  $\text{ng/g}_{\text{wcm}}/\text{h}$ , about 200 to about 2,000  $\text{ng/g}_{\text{wcm}}/\text{h}$ , about 200 to about 1,000  $\text{ng/g}_{\text{wcm}}/\text{h}$ , about 300 to about 1,000  $\text{ng/g}_{\text{wcm}}/\text{h}$ , or about 400 to about 1,000  $\text{ng/g}_{\text{wcm}}/\text{h}$ . The amount of isoprene in  $\text{ng/g}_{\text{wcm}}/\text{h}$  can be calculated by multiplying the value for isoprene production in the units of  $\text{nmole/g}_{\text{wcm}}/\text{hr}$  discussed above by 68.1 (as described in Equation 5 below).

**[0248]** In some embodiments, the cells in culture produce a cumulative titer (total amount) of isoprene at greater than or about 1, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, 10,000, 50,000, 100,000, or more mg of isoprene/L of broth ( $\text{mg/L}_{\text{broth}}$ , wherein the volume of broth includes the volume of the cells and the cell medium). In some embodiments, the amount of isoprene is between about 2 to about 5,000  $\text{mg/L}_{\text{broth}}$ , such as between about 2 to about 100  $\text{mg/L}_{\text{broth}}$ , about 100 to about 500  $\text{mg/L}_{\text{broth}}$ , about 500 to about 1,000  $\text{mg/L}_{\text{broth}}$ , about 1,000 to about 2,000  $\text{mg/L}_{\text{broth}}$ , or about 2,000 to about 5,000  $\text{mg/L}_{\text{broth}}$ . In some embodiments, the amount of isoprene is between about 20 to about 5,000  $\text{mg/L}_{\text{broth}}$ , about 100 to about 5,000  $\text{mg/L}_{\text{broth}}$ , about 200 to about 2,000  $\text{mg/L}_{\text{broth}}$ , about 200 to about 1,000  $\text{mg/L}_{\text{broth}}$ , about 300 to about 1,000  $\text{mg/L}_{\text{broth}}$ , or about 400 to about 1,000  $\text{mg/L}_{\text{broth}}$ .

**[0249]** The specific productivity of isoprene in mg of isoprene/L of headspace from shake flask or similar cultures can be measured by taking a 1 ml sample from the cell culture at an  $\text{OD}_{600}$  value of approximately 1.0, putting it in a 20 mL vial, incubating for 30 minutes, and then measuring the amount of isoprene in the headspace (as described, for example, in Example I, part II). If the  $\text{OD}_{600}$  value is not 1.0, then the measurement can be normalized to an  $\text{OD}_{600}$  value



of 1.0 by dividing by the OD<sub>600</sub> value. The value of mg isoprene/L headspace can be converted to mg/L<sub>broth</sub>/hr/OD<sub>600</sub> of culture broth by multiplying by a factor of 38. The value in units of mg/L<sub>broth</sub>/hr/OD<sub>600</sub> can be multiplied by the number of hours and the OD<sub>600</sub> value to obtain the cumulative titer in units of mg of isoprene/L of broth.

**[0250]** In some embodiments, the cells in culture have an average volumetric productivity of isoprene at greater than or about 0.1, 1.0, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1100, 1200, 1300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,100, 2,200, 2,300, 2,400, 2,500, 2,600, 2,700, 2,800, 2,900, 3,000, 3,100, 3,200, 3,300, 3,400, 3,500, or more mg of isoprene/L of broth/hr (mg/L<sub>broth</sub>/hr, wherein the volume of broth includes the volume of the cells and the cell medium). In some embodiments, the average volumetric productivity of isoprene is between about 0.1 to about 3,500 mg/L<sub>broth</sub>/hr, such as between about 0.1 to about 100 mg/L<sub>broth</sub>/hr, about 100 to about 500 mg/L<sub>broth</sub>/hr, about 500 to about 1,000 mg/L<sub>broth</sub>/hr, about 1,000 to about 1,500 mg/L<sub>broth</sub>/hr, about 1,500 to about 2,000 mg/L<sub>broth</sub>/hr, about 2,000 to about 2,500 mg/L<sub>broth</sub>/hr, about 2,500 to about 3,000 mg/L<sub>broth</sub>/hr, or about 3,000 to about 3,500 mg/L<sub>broth</sub>/hr. In some embodiments, the average volumetric productivity of isoprene is between about 10 to about 3,500 mg/L<sub>broth</sub>/hr, about 100 to about 3,500 mg/L<sub>broth</sub>/hr, about 200 to about 1,000 mg/L<sub>broth</sub>/hr, about 200 to about 1,500 mg/L<sub>broth</sub>/hr, about 1,000 to about 3,000 mg/L<sub>broth</sub>/hr, or about 1,500 to about 3,000 mg/L<sub>broth</sub>/hr.

**[0251]** In some embodiments, the cells in culture have a peak volumetric productivity of isoprene at greater than or about 0.5, 1.0, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1100, 1200, 1300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,100, 2,200, 2,300, 2,400, 2,500, 2,600, 2,700, 2,800, 2,900, 3,000, 3,100, 3,200, 3,300, 3,400, 3,500, 3,750, 4,000, 4,250, 4,500, 4,750, 5,000, 5,250, 5,500, 5,750, 6,000, 6,250, 6,500, 6,750, 7,000, 7,250, 7,500, 7,750, 8,000, 8,250, 8,500, 8,750, 9,000, 9,250, 9,500, 9,750, 10,000, 12,500, 15,000, or more mg of isoprene/L of broth/hr (mg/L<sub>broth</sub>/hr, wherein the volume of broth includes the volume of the cells and the cell medium). In some embodiments, the peak volumetric productivity of isoprene is between about 0.5 to about 15,000 mg/L<sub>broth</sub>/hr, such as between about 0.5 to about 10 mg/L<sub>broth</sub>/hr, about 1.0 to about 100 mg/L<sub>broth</sub>/hr, about 100 to about 500 mg/L<sub>broth</sub>/hr, about 500 to about 1,000 mg/L<sub>broth</sub>/hr, about 1,000 to about 1,500 mg/L<sub>broth</sub>/hr, about 1,500 to about 2,000 mg/L<sub>broth</sub>/hr, about 2,000 to about 2,500 mg/L<sub>broth</sub>/hr, about 2,500 to



about 3,000 mg/L<sub>broth</sub>/hr, about 3,000 to about 3,500 mg/L<sub>broth</sub>/hr, about 3,500 to about 5,000 mg/L<sub>broth</sub>/hr, about 5,000 to about 7,500 mg/L<sub>broth</sub>/hr, about 7,500 to about 10,000 mg/L<sub>broth</sub>/hr, about 10,000 to about 12,500 mg/L<sub>broth</sub>/h, or about 12,500 to about 15,000 mg/L<sub>broth</sub>/hr. In some embodiments, the peak volumetric productivity of isoprene is between about 10 to about 15,000 mg/L<sub>broth</sub>/hr, about 100 to about 2,500 mg/L<sub>broth</sub>/hr, about 1,000 to about 5,000 mg/L<sub>broth</sub>/hr, about 2,500 to about 7,500 mg/L<sub>broth</sub>/hr, about 5,000 to about 10,000 mg/L<sub>broth</sub>/hr, about 7,500 to about 12,500 mg/L<sub>broth</sub>/hr, or about 10,000 to about 15,000 mg/L<sub>broth</sub>/hr.

**[0252]** The instantaneous isoprene production rate in mg/L<sub>broth</sub>/hr in a fermentor can be measured by taking a sample of the fermentor off-gas, analyzing it for the amount of isoprene (in units such as mg of isoprene per L<sub>gas</sub>) as described, for example, in Example I, part II and multiplying this value by the rate at which off-gas is passed through each liter of broth (*e.g.*, at 1 vvm (volume of air/volume of broth/minute) this is 60 L<sub>gas</sub> per hour). Thus, an off-gas level of 1 mg/L<sub>gas</sub> corresponds to an instantaneous production rate of 60 mg/L<sub>broth</sub>/hr at air flow of 1 vvm. If desired, the value in the units mg/L<sub>broth</sub>/hr can be divided by the OD<sub>600</sub> value to obtain the specific rate in units of mg/L<sub>broth</sub>/hr/OD. The average value of mg isoprene/L<sub>gas</sub> can be converted to the total product productivity (grams of isoprene per liter of fermentation broth, mg/L<sub>broth</sub>) by multiplying this average off-gas isoprene concentration by the total amount of off-gas sparged per liter of fermentation broth during the fermentation. Thus, an average off-gas isoprene concentration of 0.5 mg/L<sub>broth</sub>/hr over 10 hours at 1 vvm corresponds to a total product concentration of 300 mg isoprene/L<sub>broth</sub>.

**[0253]** In some embodiments, the cells in culture convert greater than or about 0.0015, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.12, 0.14, 0.16, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.4, 1.6, 2.0, 2.2, 2.4, 2.6, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0, 12.0, 13.0, 14.0, 15.0, 16.0, 17.0, 18.0, 19.0, 20.0, 21.0, 22.0, 23.0, 23.2, 23.4, 23.6, 23.8, 24.0, 25.0, 30.0, 31.0, 32.0, 33.0, 35.0, 37.5, 40.0, 45.0, 47.5, 50.0, 55.0, 60.0, 65.0, 70.0, 75.0, 80.0, 85.0, or 90.0 molar % of the carbon in the cell culture medium into isoprene. In some embodiments, the percent conversion of carbon into isoprene is between about 0.002 to about 90.0 molar %, such as about 0.002 to about 0.005%, about 0.005 to about 0.01%, about 0.01 to about 0.05%, about 0.05 to about 0.15%, 0.15 to about 0.2%, about 0.2 to about 0.3%, about 0.3 to about 0.5%, about 0.5 to about 0.8%, about 0.8 to about 1.0%, about 1.0 to about 1.6%, about 1.6 to about 3.0%, about 3.0 to



about 5.0%, about 5.0 to about 8.0%, about 8.0 to about 10.0%, about 10.0 to about 15.0%, about 15.0 to about 20.0%, about 20.0 to about 25.0%, about 25.0% to 30.0%, about 30.0% to 35.0%, about 35.0% to 40.0%, about 45.0% to 50.0%, about 50.0% to 55.0%, about 55.0% to 60.0%, about 60.0% to 65.0%, about 65.0% to 70.0%, about 75.0% to 80.0%, about 80.0% to 85.0%, or about 85.0% to 90.0%. In some embodiments, the percent conversion of carbon into isoprene is between about 0.002 to about 0.4 molar %, 0.002 to about 0.16 molar %, 0.04 to about 0.16 molar %, about 0.005 to about 0.3 molar %, about 0.01 to about 0.3 molar %, about 0.05 to about 0.3 molar %, about 0.1 to 0.3 molar %, about 0.3 to about 1.0 molar %, about 1.0 to about 5.0 molar %, about 2 to about 5.0 molar %, about 5.0 to about 10.0 molar %, about 7 to about 10.0 molar %, about 10.0 to about 20.0 molar %, about 12 to about 20.0 molar %, about 16 to about 20.0 molar %, about 18 to about 20.0 molar %, about 18 to 23.2 molar %, about 18 to 23.6 molar %, about 18 to about 23.8 molar %, about 18 to about 24.0 molar %, about 18 to about 25.0 molar %, about 20 to about 30.0 molar %, about 30 to about 40.0 molar %, about 30 to about 50.0 molar %, about 30 to about 60.0 molar %, about 30 to about 70.0 molar %, about 30 to about 80.0 molar %, or about about 30 to about 90.0 molar %

**[0254]** The percent conversion of carbon into isoprene (also referred to as “% carbon yield”) can be measured by dividing the moles carbon in the isoprene produced by the moles carbon in the carbon source (such as the moles of carbon in batched and fed glucose and yeast extract). This number is multiplied by 100% to give a percentage value (as indicated in Equation 1).

**[0255]** Equation 1

$$\% \text{ Carbon Yield} = (\text{moles carbon in isoprene produced}) / (\text{moles carbon in carbon source}) * 100$$

**[0256]** For this calculation, yeast extract can be assumed to contain 50% w/w carbon. As an example, for the 500 liter described in Example 7, part VIII, the percent conversion of carbon into isoprene can be calculated as shown in Equation 2.

**[0257]** Equation 2

$$\% \text{ Carbon Yield} = (39.1 \text{ g isoprene} * 1/68.1 \text{ mol/g} * 5 \text{ C/mol}) / [(181221 \text{ g glucose} * 1/180 \text{ mol/g} * 6 \text{ C/mol}) + (17780 \text{ g yeast extract} * 0.5 * 1/12 \text{ mol/g})] * 100 = 0.042\%$$



[0258] For the two 500 liter fermentations described herein (Example 7, parts VII and VIII), the percent conversion of carbon into isoprene was between 0.04-0.06%. A 0.11-0.16% carbon yield has been achieved using 14 liter systems as described herein.

[0259] One skilled in the art can readily convert the rates of isoprene production or amount of isoprene produced into any other units. Exemplary equations are listed below for interconverting between units.

Units for Rate of Isoprene production (total and specific)

[0260] Equation 3

$$1 \text{ g isoprene/L}_{\text{broth}}/\text{hr} = 14.7 \text{ mmol isoprene/L}_{\text{broth}}/\text{hr} \text{ (total volumetric rate)}$$

[0261] Equation 4

$$1 \text{ nmol isoprene /g}_{\text{wcm}}/\text{hr} = 1 \text{ nmol isoprene /L}_{\text{broth}}/\text{hr}/\text{OD}_{600} \text{ (This conversion assumes that one liter of broth with an OD}_{600} \text{ value of 1 has a wet cell weight of 1 gram.)}$$

[0262] Equation 5

$$1 \text{ nmol isoprene/g}_{\text{wcm}}/\text{hr} = 68.1 \text{ ng isoprene/g}_{\text{wcm}}/\text{hr} \text{ (given the molecular weight of isoprene)}$$

[0263] Equation 6

$$1 \text{ nmol isoprene/L}_{\text{gas}} \text{ O}_2/\text{hr} = 90 \text{ nmol isoprene/L}_{\text{broth}}/\text{hr} \text{ (at an O}_2 \text{ flow rate of 90 L/hr per L of culture broth)}$$

[0264] Equation 7

$$1 \text{ ug isoprene/L}_{\text{gas}} \text{ isoprene in off-gas} = 60 \text{ ug isoprene/L}_{\text{broth}}/\text{hr} \text{ at a flow rate of 60 L}_{\text{gas}} \text{ per L}_{\text{broth}} \text{ (1 vvm)}$$

Units for Titer (total and specific)

[0265] Equation 8



1 nmol isoprene/mg cell protein = 150 nmol isoprene/L<sub>broth</sub>/OD<sub>600</sub> (This conversion assumes that one liter of broth with an OD<sub>600</sub> value of 1 has a total cell protein of approximately 150 mg) (specific productivity)

[0266] Equation 9

1 g isoprene/L<sub>broth</sub> = 14.7 mmol isoprene/L<sub>broth</sub> (total titer)

[0267] If desired, Equation 10 can be used to convert any of the units that include the wet weight of the cells into the corresponding units that include the dry weight of the cells.

[0268] Equation 10

Dry weight of cells = (wet weight of cells)/3.3

[0269] In some embodiments encompassed by the invention, a cell comprising a heterologous nucleic acid encoding an isoprene synthase polypeptide produces an amount of isoprene that is at least or about 2-fold, 3-fold, 5-fold, 10-fold, 25-fold, 50-fold, 100-fold, 150-fold, 200-fold, 400-fold, or greater than the amount of isoprene produced from a corresponding cell grown under essentially the same conditions without the heterologous nucleic acid encoding the isoprene synthase polypeptide.

[0270] In some embodiments encompassed by the invention, a cell comprising a heterologous nucleic acid encoding an isoprene synthase polypeptide and one or more heterologous nucleic acids encoding a DXS, IDI, and/or MVA pathway polypeptide produces an amount of isoprene that is at least or about 2-fold, 3-fold, 5-fold, 10-fold, 25-fold, 50-fold, 100-fold, 150-fold, 200-fold, 400-fold, or greater than the amount of isoprene produced from a corresponding cell grown under essentially the same conditions without the heterologous nucleic acids.

#### *Exemplary Isoprene Purification Methods*

[0271] In some embodiments, any of the methods described herein further include recovering the isoprene. For example, the isoprene produced using the compositions and methods of the invention can be recovered using standard techniques. such as gas stripping, fractionation, adsorption/desorption, pervaporation, thermal or vacuum desorption of isoprene from a solid



phase, or extraction of isoprene immobilized or absorbed to a solid phase with a solvent (see, for example, U.S. Patent Nos. 4,703,007 and 4,570,029, which are each hereby incorporated by reference in their entireties, particularly with respect to isoprene recovery and purification methods). In some embodiments, the recovery of isoprene involves the isolation of isoprene in a liquid form (such as a neat solution of isoprene or a solution of isoprene in a solvent). Gas stripping involves the removal of isoprene vapor from the fermentation off-gas stream in a continuous manner. Such removal can be achieved in several different ways including, but not limited to, adsorption to a solid phase, partition into a liquid phase, or direct condensation. In some embodiments, membrane enrichment of a dilute isoprene vapor stream above the dew point of the vapor resulting in the condensation of liquid isoprene.

[0272] The recovery of isoprene may involve one step or multiple steps. In some embodiments, the removal of isoprene vapor from the fermentation off-gas and the conversion of isoprene to a liquid phase are performed simultaneously. For example, isoprene can be directly condensed from the off-gas stream to form a liquid. In some embodiments, the removal of isoprene vapor from the fermentation off-gas and the conversion of isoprene to a liquid phase are performed sequentially. For example, isoprene may be adsorbed to a solid phase and then extracted from the solid phase with a solvent.

[0273] In some embodiments, any of the methods described herein further include purifying the isoprene. For example, the isoprene produced using the compositions and methods of the invention can be purified using standard techniques. Purification refers to a process through which isoprene is separated from one or more components that are present when the isoprene is produced. In some embodiments, the isoprene is obtained as a substantially pure liquid. Examples of purification methods include (i) distillation from a solution in a liquid extractant and (ii) chromatography. As used herein, "purified isoprene" means isoprene that has been separated from one or more components that are present when the isoprene is produced. In some embodiments, the isoprene is at least about 20%, by weight, free from other components that are present when the isoprene is produced. In various embodiments, the isoprene is at least or about 25%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 95%, or 99%, by weight, pure. Purity can be assayed by any appropriate method, *e.g.*, by column chromatography, HPLC analysis, or GC-MS analysis.



[0274] In some embodiments, any of the methods described herein further include polymerizing the isoprene. For example, standard methods can be used to polymerize the purified isoprene to form *cis*-polyisoprene or other down stream products using standard methods.

[0275] Additional methods and compositions are described in United States Provisional patent application No. 61/097,186, filed on September 15, 2008, United States Provisional patent application No. 61/097,189, filed on September 15, 2008, and United States Provisional patent application No. 61/097,163, filed on September 15, 2008, all of which are incorporated by reference in their entireties, particular with respect to compositions and methods for producing isoprene.

## EXAMPLES

[0276] The examples, which are intended to be purely exemplary of the invention and should therefore not be considered to limit the invention in any way, also describe and detail aspects and embodiments of the invention discussed above. Unless indicated otherwise, temperature is in degrees Centigrade and pressure is at or near atmospheric. The foregoing examples and detailed description are offered by way of illustration and not by way of limitation. All publications, patent applications, and patents cited in this specification are herein incorporated by reference as if each individual publication, patent application, or patent were specifically and individually indicated to be incorporated by reference. In particular, all publications cited herein are expressly incorporated herein by reference for the purpose of describing and disclosing compositions and methodologies which might be used in connection with the invention. Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

### **Example 1: Production of isoprene in *E. coli* expressing recombinant kudzu isoprene synthase**

#### **I. Construction of vectors for expression of the kudzu isoprene synthase in *E. coli***



[0277] The protein sequence for the kudzu (*Pueraria montana*) isoprene synthase gene (IspS) was obtained from GenBank (AAQ84170). A kudzu isoprene synthase gene, optimized for *E. coli* codon usage, was purchased from DNA2.0 (SEQ ID NO:1). The isoprene synthase gene was removed from the supplied plasmid by restriction endonuclease digestion with *Bsp*LU11I/*Pst*I, gel-purified, and ligated into pTrcHis2B (Invitrogen) that had been digested with *Nco*I/*Pst*I. The construct was designed such that the stop codon in the isoprene synthase gene 5' to the *Pst*I site. As a result, when the construct was expressed the His-Tag is not attached to the isoprene synthase protein. The resulting plasmid, pTrcKudzu, was verified by sequencing (Figures 2 and 3).

[0278] The isoprene synthase gene was also cloned into pET16b (Novagen). In this case, the isoprene synthase gene was inserted into pET16b such that the recombinant isoprene synthase protein contained the N-terminal His tag. The isoprene synthase gene was amplified from pTrcKudzu by PCR using the primer set pET-His-Kudzu-2F: 5'-CGTGAGATCATATGTGTGCGACCTCTTCTCAATTTAC (SEQ ID NO:3) and pET-His-Kudzu-R: 5'-CGGTCGACGGATCCCTGCAGTTAGACATACATCAGCTG (SEQ ID NO:4). These primers added an *Nde*I site at the 5'-end and a *Bam*H1 site at the 3' end of the gene respectively. The plasmid pTrcKudzu, described above, was used as template DNA, Herculanase polymerase (Stratagene) was used according to manufacture's directions, and primers were added at a concentration of 10 pMols. The PCR was carried out in a total volume of 25 µl. The PCR product was digested with *Nde*I/*Bam*H1 and cloned into pET16b digested with the same enzymes. The ligation mix was transformed into *E. coli* Top10 (Invitrogen) and the correct clone selected by sequencing. The resulting plasmid, in which the kudzu isoprene synthase gene was expressed from the T7 promoter, was designated pETNHISKudzu (Figures 4 and 5).

[0279] The kudzu isoprene synthase gene was also cloned into the low copy number plasmid pCL1920. Primers were used to amplify the kudzu isoprene synthase gene from pTrcKudzu described above. The forward primer added a *Hind*III site and an *E. coli* consensus RBS to the 5' end. The *Pst*I cloning site was already present in pTrcKudzu just 3' of the stop codon so the reverse primer was constructed such that the final PCR product includes the *Pst*I site. The sequences of the primers were: *Hind*III-rbs-Kudzu F: 5'-



CATATGAAAGCTTGTATCGATTAAATAAGGAGGAATAAACC (SEQ ID NO:6) and BamH1-Kudzu R:

[0280] 5'- CGGTCGACGGATCCCTGCAGTTAGACATACATCAGCTG (SEQ ID NO:4). The PCR product was amplified using Herculase polymerase with primers at a concentration of 10 pmol and with 1 ng of template DNA (pTrcKudzu). The amplification protocol included 30 cycles of (95° C for 1 minute, 60° C for 1 minute, 72° C for 2 minutes). The product was digested with *HindIII* and *PstI* and ligated into pCL1920 which had also been digested with *HindIII* and *PstI*. The ligation mix was transformed into *E. coli* Top10. Several transformants were checked by sequencing. The resulting plasmid was designated pCL-lac-Kudzu (Figures 6 and 7).

## II. Determination of isoprene production

[0281] For the shake flask cultures, one ml of a culture was transferred from shake flasks to 20 ml CTC headspace vials (Agilent vial cat# 5188 2753; cap cat# 5188 2759). The cap was screwed on tightly and the vials incubated at the equivalent temperature with shaking at 250 rpm. After 30 minutes the vials were removed from the incubator and analyzed as described below (see Table 1 for some experimental values from this assay).

[0282] In cases where isoprene production in fermentors was determined, samples were taken from the off-gas of the fermentor and analyzed directly as described below (see Table 2 for some experimental values from this assay).

[0283] The analysis was performed using an Agilent 6890 GC/MS system interfaced with a CTC Analytics (Switzerland) CombiPAL autosampler operating in headspace mode. An Agilent HP-5MS GC/MS column (30 m x 0.25 mm; 0.25 µm film thickness) was used for separation of analytes. The sampler was set up to inject 500 µL of headspace gas. The GC/MS method utilized helium as the carrier gas at a flow of 1 ml/minutes. The injection port was held at 250° C with a split ratio of 50:1. The oven temperature was held at 37° C for the 2 minute duration of the analysis. The Agilent 5793N mass selective detector was run in single ion monitoring (SIM) mode on m/z 67. The detector was switched off from 1.4 to 1.7 minutes to allow the elution of permanent gases. Under these conditions isoprene (2-methyl-1,3-butadiene) was observed to



elute at 1.78 minutes. A calibration table was used to quantify the absolute amount of isoprene and was found to be linear from 1  $\mu\text{g/L}$  to 200  $\mu\text{g/L}$ . The limit of detection was estimated to be 50 to 100 ng/L using this method.

### III. Production of isoprene in shake flasks containing *E. coli* cells expressing recombinant isoprene synthase

[0284] The vectors described above were introduced to *E. coli* strain BL21 (Novagen) to produce strains BL21/ptrcKudzu, BL21/pCL-lac-Kudzu and BL21/pETHisKudzu. The strains were spread for isolation onto LA (Luria agar) and carbenicillin (50  $\mu\text{g/ml}$ ) and incubated overnight at 37° C. Single colonies were inoculated into 250 ml baffled shake flasks containing 20 ml Luria Bertani broth (LB) and carbenicillin (100  $\mu\text{g/ml}$ ). Cultures were grown overnight at 20° C with shaking at 200 rpm. The OD<sub>600</sub> of the overnight cultures were measured and the cultures were diluted into a 250 ml baffled shake flask containing 30 ml MagicMedia (Invitrogen) and carbenicillin (100  $\mu\text{g/ml}$ ) to an OD<sub>600</sub> ~ 0.05. The culture was incubated at 30° C with shaking at 200 rpm. When the OD<sub>600</sub> ~ 0.5 - 0.8, 400  $\mu\text{M}$  IPTG was added and the cells were incubated for a further 6 hours at 30° C with shaking at 200 rpm. At 0, 2, 4 and 6 hours after induction with IPTG, 1 ml aliquots of the cultures were collected, the OD<sub>600</sub> was determined and the amount of isoprene produced was measured as described above. Results are shown in Figure 8.

### IV. Production of Isoprene from BL21/ptrcKudzu in 14 liter fermentation

[0285] Large scale production of isoprene from *E. coli* containing the recombinant kudzu isoprene synthase gene was determined from a fed-batch culture. The recipe for the fermentation media (TM2) per liter of fermentation medium was as follows: K<sub>2</sub>HPO<sub>4</sub> 13.6 g, KH<sub>2</sub>PO<sub>4</sub> 13.6 g, MgSO<sub>4</sub> \* 7H<sub>2</sub>O 2 g, citric acid monohydrate 2 g, ferric ammonium citrate 0.3 g, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 3.2 g, yeast extract 5 g, 1000X Modified Trace Metal Solution 1 ml. All of the components were added together and dissolved in diH<sub>2</sub>O. The pH was adjusted to 6.8 with potassium hydroxide (KOH) and q.s. to volume. The final product was filter sterilized with 0.22  $\mu$  filter (only, do not autoclave). The recipe for 1000X Modified Trace Metal Solution was as



follows: Citric Acids \* H<sub>2</sub>O 40 g, MnSO<sub>4</sub> \* H<sub>2</sub>O 30 g, NaCl 10 g, FeSO<sub>4</sub> \* 7H<sub>2</sub>O 1 g, CoCl<sub>2</sub> \* 6H<sub>2</sub>O 1 g, ZnSO \* 7H<sub>2</sub>O 1 g, CuSO<sub>4</sub> \* 5H<sub>2</sub>O 100 mg, H<sub>3</sub>BO<sub>3</sub> 100 mg, NaMoO<sub>4</sub> \* 2H<sub>2</sub>O 100 mg. Each component was dissolved one at a time in diH<sub>2</sub>O, pH to 3.0 with HCl/NaOH, then q.s. to volume and filter sterilized with a 0.22 µ filter.

[0286] This experiment was carried out in 14 L bioreactor to monitor isoprene formation from glucose at the desired fermentation, pH 6.7 and temperature 34° C. An inoculum of *E. coli* strain BL21/ptrcKudzu taken from a frozen vial was prepared in soytone-yeast extract-glucose medium. After the inoculum grew to OD<sub>550</sub> = 0.6, two 600 ml flasks were centrifuged and the contents resuspended in 70 ml supernatant to transfer the cell pellet (70 ml of OD 3.1 material) to the bioreactor. At various times after inoculation, samples were removed and the amount of isoprene produced was determined as described above. Results are shown in Figure 9.

#### **Example 2: Production of isoprene in *E. coli* expressing recombinant poplar isoprene synthase**

[0287] The protein sequence for the poplar (*Populus alba* x *Populus tremula*) isoprene synthase (Schnitzler, J-P, *et al.* (2005) *Planta* 222:777-786) was obtained from GenBank (CAC35696). A gene, codon optimized for *E. coli*, was purchased from DNA2.0 (p9796-poplar, Figures 30 and 31). The isoprene synthase gene was removed from the supplied plasmid by restriction endonuclease digestion with *Bsp*LU11I /*Pst*I, gel-purified, and ligated into pTrcHis2B that had been digested with *Nco*I/*Pst*I. The construct is cloned such that the stop codon in the insert is before the *Pst*I site, which results in a construct in which the His-Tag is not attached to the isoprene synthase protein. The resulting plasmid pTrcPoplar (Figures 32 and 33), was verified by sequencing.

#### **Example 3: Production of isoprene in *Pantoea citrea* expressing recombinant kudzu isoprene synthase**

[0288] The pTrcKudzu and pCL-lac Kudzu plasmids described in Example 1 were electroporated into *P. citrea* (U.S. Pat. No. 7,241,587). Transformants were selected on LA



containing carbenicillin (200 µg/ml) or spectinomycin (50 µg/ml) respectively. Production of isoprene from shake flasks and determination of the amount of isoprene produced was performed as described in Example 1 for *E. coli* strains expressing recombinant kudzu isoprene synthase. Results are shown in Figure 10.

**Example 4: Production of isoprene in *Bacillus subtilis* expressing recombinant kudzu isoprene synthase**

I. Construction of a *B. subtilis* replicating plasmid for the expression of kudzu isoprene synthase

[0289] The kudzu isoprene synthase gene was expressed in *Bacillus subtilis aprEnprE Pxyl-comK* strain (BG3594comK) using a replicating plasmid (pBS19 with a chloramphenicol resistance cassette) under control of the *aprE* promoter. The isoprene synthase gene, the *aprE* promoter and the transcription terminator were amplified separately and fused using PCR. The construct was then cloned into pBS19 and transformed into *B. subtilis*.

a) Amplification of the *aprE* promoter

[0290] The *aprE* promoter was amplified from chromosomal DNA from *Bacillus subtilis* using the following primers:

CF 797 (+) Start *aprE* promoter MfeI

5'- GACATCAATTGCTCCATTTTCTTCTGCTATC (SEQ ID NO:58)

CF 07-43 (-) Fuse *aprE* promoter to Kudzu *ispS*

5'- ATTGAGAAGAGGTCGCACACACTCTTTACCCTCTCCTTTTA (SEQ ID NO:59)

b) Amplification of the isoprene synthase gene

[0291] The kudzu isoprene synthase gene was amplified from plasmid pTrcKudzu (SEQ ID NO:2). The gene had been codon optimized for *E. coli* and synthesized by DNA 2.0. The following primers were used:

CF 07-42 (+) Fuse the *aprE* promoter to kudzu isoprene synthase gene (GTG start codon)



5'- TAAAAGGAGAGGGTAAAGAGTGTGTGCGACCTCTTCTCAAT (SEQ ID NO:60)

CF 07-45 (-) Fuse the 3' end of kudzu isoprene synthase gene to the terminator

5'- CCAAGGCCGGTTTTTTTTTAGACATACATCAGCTGGTTAATC (SEQ ID NO:61)

c) Amplification of the transcription terminator

**[0292]** The terminator from the alkaline serine protease of *Bacillus amyliquefaciens* was amplified from a previously sequenced plasmid pJHPms382 using the following primers:

CF 07-44 (+) Fuse the 3' end of kudzu isoprene synthase to the terminator

5'- GATTAACCAGCTGATGTATGTCTAAAAAAAACCGGCCTTGG (SEQ ID NO:62)

CF 07-46 (-) End of *B. amyliquefaciens* terminator (BamHI)

5'- GACATGACGGATCCGATTACGAATGCCGTCTC (SEQ ID NO:63)

**[0293]** The kudzu fragment was fused to the terminator fragment using PCR with the following primers:

CF 07-42 (+) Fuse the *aprE* promoter to kudzu isoprene synthase gene (GTG start codon)

5'- TAAAAGGAGAGGGTAAAGAGTGTGTGCGACCTCTTCTCAAT (SEQ ID NO:61)

CF 07-46 (-) End of *B. amyliquefaciens* terminator (BamHI)

5'- GACATGACGGATCCGATTACGAATGCCGTCTC (SEQ ID NO:63)

**[0294]** The kudzu-terminator fragment was fused to the promoter fragment using PCR with the following primers:

CF 797 (+) Start *aprE* promoter MfeI

5'- GACATCAATTGCTCCATTTTCTTCTGCTATC (SEQ ID NO:64)

CF 07-46 (-) End of *B. amyliquefaciens* terminator (BamHI)

5'- GACATGACGGATCCGATTACGAATGCCGTCTC (SEQ ID NO:63)



**[0295]** The fusion PCR fragment was purified using a Qiagen kit and digested with the restriction enzymes *MfeI* and *BamHI*. This digested DNA fragment was gel purified using a Qiagen kit and ligated to a vector known as pBS19, which had been digested with *EcoRI* and *BamHI* and gel purified.

**[0296]** The ligation mix was transformed into *E. coli* Top 10 cells and colonies were selected on LA and 50 carbenicillin plates. A total of six colonies were chosen and grown overnight in LB and 50 carbenicillin and then plasmids were isolated using a Qiagen kit. The plasmids were digested with *EcoRI* and *BamHI* to check for inserts and three of the correct plasmids were sent in for sequencing with the following primers:

CF 149 (+) *EcoRI* start of *aprE* promoter

5'- GACATGAATTCCTCCATTTTCTTCTGC (SEQ ID NO:65)

CF 847 (+) Sequence in pXX 049 (end of *aprE* promoter)

5'- AGGAGAGGGTAAAGAGTGAG (SEQ ID NO:66)

CF 07-45 (-) Fuse the 3' end of kudzu isoprene synthase to the terminator

5'- CCAAGGCCGGTTTTTTTTTAGACATACATCAGCTGGTTAATC (SEQ ID NO:61)

CF 07-48 (+) Sequencing primer for kudzu isoprene synthase

5'- CTTTTCATCACCCACCTGAAG (SEQ ID NO:67)

CF 07-49 (+) Sequencing in kudzu isoprene synthase

5'- GGCGAAATGGTCCAACAACAAAATTATC (SEQ ID NO:68)

**[0297]** The plasmid designated pBS Kudzu #2 (Figures 52 and 12) was correct by sequencing and was transformed into BG 3594 comK, a *Bacillus subtilis* host strain. Selection was done on LA and 5 chloramphenicol plates. A transformant was chosen and struck to single colonies on LA and 5 chloramphenicol, then grown in LB and 5 chloramphenicol until it reached an OD<sub>600</sub>



of 1.5. It was stored frozen in a vial at -80° C in the presence of glycerol. The resulting strain was designated CF 443.

## II. Production of isoprene in shake flasks containing *B. subtilis* cells expressing recombinant isoprene synthase

**[0298]** Overnight cultures were inoculated with a single colony of CF 443 from a LA and Chloramphenicol (Cm, 25 µg/ml). Cultures were grown in LB and Cm at 37° C with shaking at 200 rpm. These overnight cultures (1 ml) were used to inoculate 250 ml baffled shake flasks containing 25 ml Grants II media and chloramphenicol at a final concentration of 25 µg/ml. Grants II Media recipe was 10 g soytone, 3 ml 1M K<sub>2</sub>HPO<sub>4</sub>, 75 g glucose, 3.6 g urea, 100 ml 10X MOPS, q.s. to 1 L with H<sub>2</sub>O, pH 7.2; 10X MOPS recipe was 83.72 g MOPS, 7.17 g tricine, 12 g KOH pellets, 10 ml 0.276M K<sub>2</sub>SO<sub>4</sub> solution, 10 ml 0.528M MgCl<sub>2</sub> solution, 29.22 g NaCl, 100 ml 100X micronutrients, q.s. to 1 L with H<sub>2</sub>O; and 100X micronutrients recipe was 1.47 g CaCl<sub>2</sub>\*2H<sub>2</sub>O, 0.4 g FeSO<sub>4</sub>\*7H<sub>2</sub>O, 0.1 g MnSO<sub>4</sub>\*H<sub>2</sub>O, 0.1 g ZnSO<sub>4</sub>\*H<sub>2</sub>O, 0.05 g CuCl<sub>2</sub>\*2H<sub>2</sub>O, 0.1 g CoCl<sub>2</sub>\*6H<sub>2</sub>O, 0.1 g Na<sub>2</sub>MoO<sub>4</sub>\*2H<sub>2</sub>O, q.s. to 1 L with H<sub>2</sub>O. Shake flasks were incubated at 37° C and samples were taken at 18, 24, and 44 hours. At 18 hours the headspaces of CF443 and the control strain were sampled. This represented 18 hours of accumulation of isoprene. The amount of isoprene was determined by gas chromatography as described in Example 1. Production of isoprene was enhanced significantly by expressing recombinant isoprene synthase (Figure 11).

## III. Production of isoprene by CF443 in 14 L fermentation

**[0299]** Large scale production of isoprene from *B. subtilis* containing the recombinant kudzu isoprene synthase gene on a replication plasmid was determined from a fed-batch culture. Bacillus strain CF 443, expressing a kudzu isoprene synthase gene, or control stain which does not express a kudzu isoprene synthase gene were cultivated by conventional fed-batch fermentation in a nutrient medium containing soy meal (Cargill), sodium and potassium phosphate, magnesium sulfate and a solution of citric acid, ferric chloride and manganese chloride. Prior to fermentation the media is macerated for 90 minutes using a mixture of enzymes including cellulases, hemicellulases and pectinases (see, WO95/04134). 14-L batch fermentations are fed with 60% wt/wt glucose (Cargill DE99 dextrose, ADM Versadex greens or



Danisco invert sugar) and 99% wt/wt oil (Western Family soy oil, where the 99% wt/wt is the concentration of oil before it was added to the cell culture medium). Feed was started when glucose in the batch was non-detectable. The feed rate was ramped over several hours and was adjusted to add oil on an equal carbon basis. The pH was controlled at 6.8 – 7.4 using 28% w/v ammonium hydroxide. In case of foaming, antifoam agent was added to the media. The fermentation temperature was controlled at 37°C and the fermentation culture was agitated at 750 rpm. Various other parameters such as pH, DO%, airflow, and pressure were monitored throughout the entire process. The DO% is maintained above 20. Samples were taken over the time course of 36 hours and analyzed for cell growth (OD<sub>550</sub>) and isoprene production. Results of these experiments are presented in Figures 53A and 53B.

#### IV. Integration of the kudzu isoprene synthase (*ispS*) in *B. subtilis*.

**[0300]** The kudzu isoprene synthase gene was cloned in an integrating plasmid (pJH101-cmpR) under the control of the *aprE* promoter. Under the conditions tested, no isoprene was detected.

#### **Example 5: Production of isoprene in *Trichoderma***

##### I. Construction of vectors for expression of the kudzu isoprene synthase in *Trichoderma reesei*

**[0301]** The *Yarrowia lipolytica* codon-optimized kudzu IS gene was synthesized by DNA 2.0 (SEQ ID NO:8) (Figure 13). This plasmid served as the template for the following PCR amplification reaction: 1 µl plasmid template (20 ng/ul), 1 µl Primer EL-945 (10 uM) 5'-GCTTATGGATCCTCTAGACTATTACACGTACATCAATTGG (SEQ ID NO:9), 1 µl Primer EL-965 (10uM) 5'-CACCATGTGTGCAACCTCCTCCCAGTTTAC (SEQ ID NO:10), 1 µl dNTP (10mM), 5 µl 10x PfuUltra II Fusion HS DNA Polymerase Buffer, 1 µl PfuUltra II Fusion HS DNA Polymerase, 40 µl water in a total reaction volume of 50 µl. The forward primer contained an additional 4 nucleotides at the 5'-end that did not correspond to the *Y. lipolytica* codon-optimized kudzu isoprene synthase gene, but was required for cloning into the pENTR/D-TOPO vector. The reverse primer contained an additional 21 nucleotides at the 5'-end that did not correspond to the *Y. lipolytica* codon-optimized kudzu isoprene synthase gene, but were inserted for cloning into other vector backbones. Using the MJ Research PTC-200



Thermocycler, the PCR reaction was performed as follows: 95° C for 2 minutes (first cycle only), 95° C for 30 seconds, 55° C for 30 seconds, 72° C for 30 seconds (repeat for 27 cycles), 72° C for 1 minute after the last cycle. The PCR product was analyzed on a 1.2% E-gel to confirm successful amplification of the *Y. lipolytica* codon-optimized kudzu isoprene synthase gene.

[0302] The PCR product was then cloned using the TOPO pENTR/D-TOPO Cloning Kit following manufacturer's protocol: 1 µl PCR reaction, 1 µl Salt solution, 1 µl TOPO pENTR/D-TOPO vector and 3 µl water in a total reaction volume of 6 µl. The reaction was incubated at room temperature for 5 minutes. One microliter of TOPO reaction was transformed into TOP10 chemically competent *E. coli* cells. The transformants were selected on LA and 50 µg/ml kanamycin plates. Several colonies were picked and each was inoculated into a 5 ml tube containing LB and 50 µg/ml kanamycin and the cultures grown overnight at 37° C with shaking at 200 rpm. Plasmids were isolated from the overnight culture tubes using QIAprep Spin Miniprep Kit, following manufacturer's protocol. Several plasmids were sequenced to verify that the DNA sequence was correct.

[0303] A single pENTR/D-TOPO plasmid, encoding a *Y. lipolytica* codon-optimized kudzu isoprene synthase gene, was used for Gateway Cloning into a custom-made pTrex3g vector. Construction of pTrex3g is described in WO 2005/001036 A2. The reaction was performed following manufacturer's protocol for the Gateway LR Clonase II Enzyme Mix Kit (Invitrogen): 1 µl *Y. lipolytica* codon-optimized kudzu isoprene synthase gene pENTR/D-TOPO donor vector, 1 µl pTrex3g destination vector, 6 µl TE buffer, pH 8.0 in a total reaction volume of 8 µl. The reaction was incubated at room temperature for 1 hour and then 1 µl proteinase K solution was added and the incubation continued at 37° C for 10 minutes. Then 1 µl of reaction was transformed into TOP10 chemically competent *E. coli* cells. The transformants were selected on LA and 50 µg/ml carbenicillin plates. Several colonies were picked and each was inoculated into a 5 ml tube containing LB and 50 µg/ml carbenicillin and the cultures were grown overnight at 37° C with shaking at 200 rpm. Plasmids were isolated from the overnight culture tubes using QIAprep Spin Miniprep Kit (Qiagen, Inc.), following manufacturer's protocol. Several plasmids were sequenced to verify that the DNA sequence was correct.



**[0304]** Biolistic transformation of *Y. lipolytica* codon-optimized kudzu isoprene synthase pTrex3g plasmid (Figure 14) into a quad delete *Trichoderma reesei* strain was performed using the Biolistic PDS-1000/HE Particle Delivery System (see WO 2005/001036 A2). Isolation of stable transformants and shake flask evaluation was performed using protocol listed in Example 11 of patent publication WO 2005/001036 A2.

## II. Production of isoprene in recombinant strains of *T. reesei*

**[0305]** One ml of 15 and 36 hour old cultures of isoprene synthase transformants described above were transferred to head space vials. The vials were sealed and incubated for 5 hours at 30° C. Head space gas was measured and isoprene was identified by the method described in Example 1. Two of the transformants showed traces of isoprene. The amount of isoprene could be increased by a 14 hour incubation. The two positive samples showed isoprene at levels of about 0.5 µg/L for the 14 hour incubation. The untransformed control showed no detectable levels of isoprene. This experiment shows that *T. reesei* is capable of producing isoprene from endogenous precursor when supplied with an exogenous isoprene synthase.

## **Example 6: Production of isoprene in *Yarrowia***

### I. Construction of vectors for expression of the kudzu isoprene synthase in *Yarrowia lipolytica*.

**[0306]** The starting point for the construction of vectors for the expression of the kudzu isoprene synthase gene in *Yarrowia lipolytica* was the vector pSPZ1(MAP29Spb). The complete sequence of this vector (SEQ ID No:11) is shown in Figure 15.

**[0307]** The following fragments were amplified by PCR using chromosomal DNA of a *Y. lipolytica* strain GICC 120285 as the template: a promotorless form of the URA3 gene, a fragment of 18S ribosomal RNA gene, a transcription terminator of the *Y. lipolytica* XPR2 gene and two DNA fragments containing the promoters of XPR2 and ICL1 genes. The following PCR primers were used:

ICL1 3



5'- GGTGAATTCAGTCTACTGGGGATTCCCAAATCTATATATACTGCAGGTGAC (SEQ ID NO:69)

ICL1 5

5'- GCAGGTGGGAAACTATGCACTCC (SEQ ID NO:70)

XPR 3

5'- CCTGAATTCTGTTGGATTGGAGGATTGGATAGTGGG (SEQ ID NO:71)

XPR 5

5'- GGTGTCGACGTACGGTCGAGCTTATTGACC (SEQ ID NO:72)

XPRT3

5'- GGTGGGCCCCGCATTTTGCCACCTACAAGCCAG (SEQ ID NO:73)

XPRT 5

5'- GGTGAATTCTAGAGGATCCCAACGCTGTTGCCTACAACGG (SEQ ID NO:74)

Y18S3

5'- GGTGCGGCCGCTGTCTGGACCTGGTGAGTTTCCCCG (SEQ ID NO:75)

Y18S 5

5'- GGTGGGCCCATTAAATCAGTTATCGTTTATTTGATAG (SEQ ID NO:76)

YURA3

5'- GGTGACCAGCAAGTCCATGGGTGGTTTGATCATGG (SEQ ID NO:77)

YURA 50

5'- GGTGCGGCCGCCTTTGGAGTACGACTCCAACCTATG (SEQ ID NO:78)

YURA 51

5'- GCGGCCGCAGACTAAATTTATTTTCAGTCTCC (SEQ ID NO:79)



[0308] For PCR amplification the PfuUltraII polymerase (Stratagene), supplier-provided buffer and dNTPs, 2.5  $\mu$ M primers and the indicated template DNA were used as per the manufacturer's instructions. The amplification was done using the following cycle: 95° C for 1 min; 34x (95° C for 30 sec; 55° C for 30 sec; 72° C for 3 min) and 10 min at 72° C followed by a 4° C incubation.

[0309] Synthetic DNA molecules encoding the kudzu isoprene synthase gene, codon-optimized for expression in *Yarrowia*, was obtained from DNA 2.0 (Figure 16; SEQ ID NO:12). Full detail of the construction scheme of the plasmids pYLA(KZ1) and pYLI(KZ1) carrying the synthetic kudzu isoprene synthase gene under control of XPR2 and ICL1 promoters respectively is presented in Figure 18. Control plasmids in which a mating factor gene (MAP29) is inserted in place of an isoprene synthase gene were also constructed (Figure 18E and 18F).

[0310] A similar cloning procedure can be used to express a poplar (*Populus alba* x *Populus tremula*) isoprene synthase gene. The sequence of the poplar isoprene is described in Miller B. *et al.* (2001) *Planta* 213, 483-487 and shown in Figure 17 (SEQ ID NO:13). A construction scheme for the generation the plasmids pYLA(POP1) and pYLI(POP1) carrying synthetic poplar isoprene synthase gene under control of XPR2 and ICL1 promoters respectively is presented in Figure 18A and B.

## II. Production of isoprene by recombinant strains of *Y. lipolytica*.

[0311] Vectors pYLA(KZ1), pYLI(KZ1), pYLA(MAP29) and pYLI(MAP29) were digested with *Sac*II and used to transform the strain *Y. lipolytica* CLIB 122 by a standard lithium acetate/polyethylene glycol procedure to uridine prototrophy. Briefly, the yeast cells grown in YEPD (1% yeast extract, 2% peptone, 2% glucose) overnight, were collected by centrifugation (4000 rpm, 10 min), washed once with sterile water and suspended in 0.1 M lithium acetate, pH 6.0. Two hundred  $\mu$ l aliquots of the cell suspension were mixed with linearized plasmid DNA solution (10-20  $\mu$ g), incubated for 10 minutes at room temperature and mixed with 1 ml of 50% PEG 4000 in the same buffer. The suspensions were further incubated for 1 hour at room temperature followed by a 2 minutes heat shock at 42° C. Cells were then plated on SC his leu plates (0.67% yeast nitrogen base, 2% glucose, 100 mg/L each of leucine and histidine). Transformants appeared after 3-4 days of incubation at 30° C.



**[0312]** Three isolates from the pYLA(KZ1) transformation, three isolates from the pYLI(KZ1) transformation, two isolates from the pYLA(MAP29) transformation and two isolates from the pYLI(MAP29) transformation were grown for 24 hours in YEP7 medium (1% yeast extract, 2% peptone, pH 7.0) at 30° C with shaking. Cells from 10 ml of culture were collected by centrifugation, resuspended in 3 ml of fresh YEP7 and placed into 15 ml screw cap vials. The vials were incubated overnight at room temperature with gentle (60 rpm) shaking. Isoprene content in the headspace of these vials was analyzed by gas chromatography using mass-spectrometric detector as described in Example 1. All transformants obtained with pYLA(KZ1) and pYLI(KZ1) produced readily detectable amounts of isoprene (0.5 µg/L to 1 µg/L, Figure 20). No isoprene was detected in the headspace of the control strains carrying phytase gene instead of an isoprene synthase gene.

**Example 7: Production of isoprene in *E. coli* expressing kudzu isoprene synthase and *idi*, or *dxs*, or *idi* and *dxs***

I. Construction of vectors encoding kudzu isoprene synthase and *idi*, or *dxs*, or *idi* and *dxs* for the production of isoprene in *E. coli*

i) Construction of pTrcKudzuKan

**[0313]** The *bla* gene of pTrcKudzu (described in Example 1) was replaced with the gene conferring kanamycin resistance. To remove the *bla* gene, pTrcKudzu was digested with *Bsp*HI, treated with Shrimp Alkaline Phosphatase (SAP), heat killed at 65° C, then end-filled with Klenow fragment and dNTPs. The 5 kbp large fragment was purified from an agarose gel and ligated to the *kan*<sup>r</sup> gene which had been PCR amplified from pCR-Blunt-II-TOPO using primers MCM22 5'- GATCAAGCTTAACCGGAATTGCCAGCTG (SEQ ID NO:14) and MCM23 5'- GATCCGATCGTCAGAAGAACTCGTCAAGAAGGC (SEQ ID NO:15), digested with *Hind*III and *Pvu*I, and end-filled. A transformant carrying a plasmid conferring kanamycin resistance (pTrcKudzuKan) was selected on LA containing kanamycin 50 µg/ml.

ii) Construction of pTrcKudzu yIDI Kan



**[0314]** pTrcKudzuKan was digested with *Pst*I, treated with SAP, heat killed and gel purified. It was ligated to a PCR product encoding *idi* from *S. cerevisiae* with a synthetic RBS. The primers for PCR were NsiI-YIDI 1 F 5'-CATCAATGCATCGCCCTTAGGAGGTAAAAAAAATGAC (SEQ ID NO:16) and PstI-YIDI 1 R 5'-CCTTCTGCAGGACGCGTTGTTATAGC (SEQ ID NO:17); and the template was *S. cerevisiae* genomic DNA. The PCR product was digested with *Nsi*I and *Pst*I and gel purified prior to ligation. The ligation mixture was transformed into chemically competent TOP10 cells and selected on LA containing 50 µg/ml kanamycin. Several transformants were isolated and sequenced and the resulting plasmid was called pTrcKudzu-yIDI(kan) (Figures 34 and 35).

### iii) Construction of pTrcKudzu DXS Kan

**[0315]** Plasmid pTrcKudzuKan was digested with *Pst*I, treated with SAP, heat killed and gel purified. It was ligated to a PCR product encoding *dxs* from *E. coli* with a synthetic RBS. The primers for PCR were MCM13 5'-GATCATGCATTCGCCCTTAGGAGGTAAAAAACATGAGTTTTGATATTGCCAAATAC CCG (SEQ ID NO:18) and MCM14 5'-CATGCTGCAGTTATGCCAGCCAGGCCTTGAT (SEQ ID NO:19); and the template was *E. coli* genomic DNA. The PCR product was digested with *Nsi*I and *Pst*I and gel purified prior to ligation. The resulting transformation reaction was transformed into TOP10 cells and selected on LA with kanamycin 50 µg/ml. Several transformants were isolated and sequenced and the resulting plasmid was called pTrcKudzu-DXS(kan) (Figures 36 and 37).

### iv) Construction of pTrcKudzu-yIDI-dxs (kan)

**[0316]** pTrcKudzu-yIDI(kan) was digested with *Pst*I, treated with SAP, heat killed and gel purified. It was ligated to a PCR product encoding *E. coli dxs* with a synthetic RBS (primers MCM13 5'-GATCATGCATTCGCCCTTAGGAGGTAAAAAACATGAGTTTTGATATTGCCAAATAC CCG (SEQ ID NO:18) and MCM14 5'-CATGCTGCAGTTATGCCAGCCAGGCCTTGAT (SEQ ID NO:19); template TOP10 cells) which had been digested with *Nsi*I and *Pst*I and gel purified. The final plasmid was called pTrcKudzu-yIDI-dxs (kan) (Figures 21 and 22).



## v) Construction of pCL PtrcKudzu

**[0317]** A fragment of DNA containing the promoter, structural gene and terminator from Example 1 above was digested from pTrcKudzu using *SspI* and gel purified. It was ligated to pCL1920 which had been digested with *PvuII*, treated with SAP and heat killed. The resulting ligation mixture was transformed into TOP10 cells and selected in LA containing spectinomycin 50 µg/ml. Several clones were isolated and sequenced and two were selected. pCL PtrcKudzu and pCL PtrcKudzu (A3) have the insert in opposite orientations (Figures 38-41).

## vi) Construction of pCL PtrcKudzu yIDI

**[0318]** The *NsiI-PstI* digested, gel purified, IDI PCR amplicon from (ii) above was ligated into pCL PtrcKudzu which had been digested with *PstI*, treated with SAP, and heat killed. The ligation mixture was transformed into TOP10 cells and selected in LA containing spectinomycin 50 µg/ml. Several clones were isolated and sequenced and the resulting plasmid is called pCL PtrcKudzu yIDI (Figures 42 and 43).

## vii) Construction of pCL PtrcKudzu DXS

**[0319]** The *NsiI-PstI* digested, gel purified, DXS PCR amplicon from (iii) above was ligated into pCL PtrcKudzu (A3) which had been digested with *PstI*, treated with SAP, and heat killed. The ligation mixture was transformed into TOP10 cells and selected in LA containing spectinomycin 50 µg/ml. Several clones were isolated and sequenced and the resulting plasmid is called pCL PtrcKudzu DXS (Figures 44 and 45).

## II. Measurement of isoprene in headspace from cultures expressing kudzu isoprene synthase, *idi*, and/or *dxs* at different copy numbers.

**[0320]** Cultures of *E. coli* BL21(λDE3) previously transformed with plasmids pTrcKudzu(kan) (A), pTrcKudzu-yIDI kan (B), pTrcKudzu-DXS kan (C), pTrcKudzu-yIDI-DXS kan (D) were grown in LB kanamycin 50 µg/mL. Cultures of pCL PtrcKudzu (E), pCL PtrcKudzu, pCL PtrcKudzu-yIDI (F) and pCL PtrcKudzu-DXS (G) were grown in LB spectinomycin 50 µg/mL. Cultures were induced with 400 µM IPTG at time 0 (OD<sub>600</sub>



approximately 0.5) and samples taken for isoprene headspace measurement (see Example 1). Results are shown in Figure 23A-23G.

**[0321]** Plasmid pTrcKudzu-yIDI-dxs (kan) was introduced into *E. coli* strain BL21 by transformation. The resulting strain BL21/pTrc Kudzu IDI DXS was grown overnight in LB containing kanamycin (50 µg/ml) at 20° C and used to inoculate shake flasks of TM3 (13.6 g K<sub>2</sub>PO<sub>4</sub>, 13.6 g KH<sub>2</sub>PO<sub>4</sub>, 2.0 g MgSO<sub>4</sub>\*7H<sub>2</sub>O), 2.0 g citric acid monohydrate, 0.3 g ferric ammonium citrate, 3.2 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.2 g yeast extract, 1.0 ml 1000x Modified Trace Metal Solution, adjusted to pH 6.8 and q.s. to H<sub>2</sub>O, and filter sterilized) containing 1% glucose. Flasks were incubated at 30° C until an OD<sub>600</sub> of 0.8 was reached, and then induced with 400 µM IPTG. Samples were taken at various times after induction and the amount of isoprene in the head space was measured as described in Example 1. Results are shown in Figure 23H.

### III. Production of isoprene from biomass in *E. coli*/pTrcKudzu yIDI DXS

**[0322]** The strain BL21 pTrcKudzuIDIDXS was tested for the ability to generate isoprene from three types of biomass; bagasse, corn stover and soft wood pulp with glucose as a control. Hydrolysates of the biomass were prepared by enzymatic hydrolysis (Brown, L and Torget, R., 1996, NREL standard assay method Lap-009 “Enzymatic Saccharification of Lignocellulosic Biomass”) and used at a dilution based upon glucose equivalents. In this example, glucose equivalents were equal to 1% glucose. A single colony from a plate freshly transformed cells of BL21 (DE3) pTrcKudzu yIDI DXS (kan) was used to inoculate 5 ml of LB plus kanamycin (50 µg/ml). The culture was incubated overnight at 25° C with shaking. The following day the overnight culture was diluted to an OD<sub>600</sub> of 0.05 in 25 ml of TM3 and 0.2% YE and 1% feedstock. The feedstock was corn stover, bagasse, or softwood pulp. Glucose was used as a positive control and no glucose was used as a negative control. Cultures were incubated at 30° C with shaking at 180 rpm. The culture was monitored for OD<sub>600</sub> and when it reached an OD<sub>600</sub> of ~0.8, cultures were analyzed at 1 and 3 hours for isoprene production as described in Example 1. Cultures are not induced. All cultures containing added feedstock produce isoprene equivalent to those of the glucose positive control. Experiments were done in duplicate and are shown in Figure 46.

### IV. Production of isoprene from invert sugar in *E. coli*/pTrcKudzuIDIDXS



**[0323]** A single colony from a plate freshly transformed cells of BL21 ( $\lambda$ DE3)/pTrcKudzu yIDI DXS (kan) was used to inoculate 5 mL of LB and kanamycin (50  $\mu$ g/ml). The culture was incubated overnight at 25° C with shaking. The following day the overnight culture was diluted to an OD<sub>600</sub> of 0.05 in 25 ml of TM3 and 0.2% YE and 1% feedstock. Feedstock was glucose, inverted glucose or corn stover. The invert sugar feedstock (Danisco Invert Sugar) was prepared by enzymatically treating sucrose syrup. AFEX corn stover was prepared as described below (Part V). The cells were grown at 30° C and the first sample was measured when the cultures reached an OD<sub>600</sub> ~0.8-1.0 (0 hour). The cultures were analyzed for growth as measured by OD<sub>600</sub> and for isoprene production as in Example 1 at 0, 1 and 3 hours. Results are shown in Figure 47.

#### V. Preparation of hydrolysate from AFEX pretreated corn stover

**[0324]** AFEX pretreated corn stover was obtained from Michigan Biotechnology Institute. The pretreatment conditions were 60% moisture, 1:1 ammonia loading, and 90 °C for 30 minutes, then air dried. The moisture content in the AFEX pretreated corn stover was 21.27%. The contents of glucan and xylan in the AFEX pretreated corn stover were 31.7% and 19.1% (dry basis), respectively. The saccharification process was as follows; 20 g of AFEX pretreated corn stover was added into a 500 ml flask with 5 ml of 1 M sodium citrate buffer pH 4.8, 2.25 ml of Accellerase 1000, 0.1 ml of Grindamyl H121 (Danisco xylanase product from *Aspergillus niger* for bread-making industry), and 72.65 ml of DI water. The flask was put in an orbital shaker and incubated at 50° C for 96 hours. One sample was taken from the shaker and analyzed using HPLC. The hydrolysate contained 38.5 g/l of glucose, 21.8 g/l of xylose, and 10.3 g/l of oligomers of glucose and/or xylose.

#### VI. The effect of yeast extract on isoprene production in *E. coli* grown in fed-batch culture

**[0325]** Fermentation was performed at the 14-L scale as previously described with *E. coli* cells containing the pTrcKudzu yIDI DXS plasmid described above. Yeast extract (Bio Springer, Montreal, Quebec, Canada) was fed at an exponential rate. The total amount of yeast extract delivered to the fermentor was varied between 70-830 g during the 40 hour fermentation. Optical density of the fermentation broth was measured at a wavelength of 550 nm. The final optical density within the fermentors was proportional to the amount of yeast extract added



(Figure 48A). The isoprene level in the off-gas from the fermentor was determined as previously described. The isoprene titer increased over the course of the fermentation (Figure 48B). The amount of isoprene produced was linearly proportional to the amount of fed yeast extract (Figure 48C).

#### VII. Production of isoprene in 500 L fermentation of pTrcKudzu DXS yIDI

**[0326]** A 500 liter fermentation of *E. coli* cells with a kudzu isoprene synthase, *S. cerevisiae* IDI, and *E. coli* DXS nucleic acids (*E. coli* BL21 ( $\lambda$ DE3) pTrc Kudzu dxs yidi) was used to produce isoprene. The levels of isoprene varied from 50 to 300  $\mu$ g/L over a time period of 15 hours. On the basis of the average isoprene concentrations, the average flow through the device and the extent of isoprene breakthrough, the amount of isoprene collected was calculated to be approximately 17 g.

#### VIII. Production of isoprene in 500 L fermentation of *E. coli* grown in fed-batch culture

Medium Recipe (per liter fermentation medium):

**[0327]**  $K_2HPO_4$  7.5 g,  $MgSO_4 \cdot 7H_2O$  2 g, citric acid monohydrate 2 g, ferric ammonium citrate 0.3 g, yeast extract 0.5 g, 1000X Modified Trace Metal Solution 1 ml. All of the components were added together and dissolved in  $diH_2O$ . This solution was autoclaved. The pH was adjusted to 7.0 with ammonium gas ( $NH_3$ ) and q.s. to volume. Glucose 10 g, thiamine \* HCl 0.1 g, and antibiotic were added after sterilization and pH adjustment.

**[0328]** 1000X Modified Trace Metal Solution:

**[0329]** Citric Acids \*  $H_2O$  40 g,  $MnSO_4 \cdot H_2O$  30 g, NaCl 10 g,  $FeSO_4 \cdot 7H_2O$  1 g,  $CoCl_2 \cdot 6H_2O$  1 g,  $ZnSO_4 \cdot 7H_2O$  1 g,  $CuSO_4 \cdot 5H_2O$  100 mg,  $H_3BO_3$  100 mg,  $NaMoO_4 \cdot 2H_2O$  100 mg. Each component is dissolved one at a time in DI  $H_2O$ , pH to 3.0 with HCl/NaOH, then q.s. to volume and filter sterilized with 0.22 micron filter.

**[0330]** Fermentation was performed in a 500-L bioreactor with *E. coli* cells containing the pTrcKudzu yIDI DXS plasmid. This experiment was carried out to monitor isoprene formation from glucose and yeast extract at the desired fermentation pH 7.0 and temperature 30° C. An



inoculum of *E. coli* strain taken from a frozen vial was prepared in soytone-yeast extract-glucose medium. After the inoculum grew to OD 0.15, measured at 550 nm, 20 ml was used to inoculate a bioreactor containing 2.5-L soytone-yeast extract-glucose medium. The 2.5-L bioreactor was grown at 30° C to OD 1.0 and 2.0-L was transferred to the 500-L bioreactor.

[0331] Yeast extract (Bio Springer, Montreal, Quebec, Canada) and glucose were fed at exponential rates. The total amount of glucose and yeast extract delivered to the bioreactor during the 50 hour fermentation was 181.2 kg and 17.6 kg, respectively. The optical density within the bioreactor over time is shown in Figure 49A. The isoprene level in the off-gas from the bioreactor was determined as previously described. The isoprene titer increased over the course of the fermentation (Figure 49B). The total amount of isoprene produced during the 50 hour fermentation was 55.1 g and the time course of production is shown in Figure 49C.

#### **Example 8: Production of isoprene in *E. coli* expressing kudzu isoprene synthase and recombinant mevalonic acid pathway genes**

##### **I. Cloning the lower MVA pathway**

[0332] The strategy for cloning the lower mevalonic pathway was as follows. Four genes of the mevalonic acid biosynthesis pathway; mevalonate kinase (MVK), phosphomevalonate kinase (PMK), diphosphomevalonate decarboxylase (MVD) and isopentenyl diphosphate isomerase genes were amplified by PCR from *S. cerevisiae* chromosomal DNA and cloned individually into the pCR BluntII TOPO plasmid (Invitrogen). In some cases, the *idi* gene was amplified from *E. coli* chromosomal DNA. The primers were designed such that an *E. coli* consensus RBS (AGGAGGT (SEQ ID NO:80) or AAGGAGG (SEQ ID NO:81)) was inserted at the 5' end, 8 bp upstream of the start codon and a *Pst*I site was added at the 3' end. The genes were then cloned one by one into the pTrcHis2B vector until the entire pathway was assembled.

[0333] Chromosomal DNA from *S. cerevisiae* S288C was obtained from ATCC (ATCC 204508D). The MVK gene was amplified from the chromosome of *S. cerevisiae* using primers MVKF (5'-AGGAGGTAAAAAACATGTCATTACCGTTCTTAACTTCTGC, SEQ ID NO:21) and MVK-PstI-R (5'-ATGGCTGCAGGCCTATCGCAAATTAGCTTATGAAGTCCATGGTAAATTCGTG, SEQ ID NO:22) using PfuTurbo as per manufacturer's instructions. The correct sized PCR product (1370 bp) was identified by electrophoresis through a 1.2% E-gel (Invitrogen) and cloned into



pZeroBLUNT TOPO. The resulting plasmid was designated pMVK1. The plasmid pMVK1 was digested with *SacI* and *TaqI* restriction endonucleases and the fragment was gel purified and ligated into pTrcHis2B digested with *SacI* and *BstBI*. The resulting plasmid was named pTrcMVK1.

[0334] The second gene in the mevalonic acid biosynthesis pathway, PMK, was amplified by PCR using primers: PstI-PMK1 R (5'-GAATTCGCCCTTCTGCAGCTACC, SEQ ID NO:23) and BsiHKA I-PMK1 F (5'-CGACTGGTGCACCCTTAAGGAGGAAAAAACATGTCAG, SEQ ID NO:24). The PCR reaction was performed using Pfu Turbo polymerase (Stratagene) as per manufacturer's instructions. The correct sized product (1387 bp) was digested with *PstI* and *BsiHKA I* and ligated into pTrcMVK1 digested with *PstI*. The resulting plasmid was named pTrcKK. The MVD and the *idi* genes were cloned in the same manner. PCR was carried out using the primer pairs PstI-MVD 1 R (5'-GTGCTGGAATTCGCCCTTCTGCAGC, SEQ ID NO:25) and NsiI-MVD 1 F (5'-GTAGATGCATGCAGAATTCGCCCTTAAGGAGG, SEQ ID NO:26) to amplify the MVD gene and PstI-YIDI 1 R (5'-CCTTCTGCAGGACGCGTTGTTATAGC, SEQ ID NO:27) and NsiI-YIDI 1 F (5'-CATCAATGCATCGCCCTTAGGAGGTAAAAAAAATGAC, SEQ ID NO:28) to amplify the yIDI gene. In some cases the IPP isomerase gene, *idi* from *E. coli* was used. To amplify *idi* from *E. coli* chromosomal DNA, the following primer set was used: PstI-CIDI 1 R (5'-GTGTGATGGATATCTGCAGAATTCG, SEQ ID NO:29) and NsiI-CIDI 1 F (5'-CATCAATGCATCGCCCTTAGGAGGTAAAAAACATG, SEQ ID NO:30). Template DNA was chromosomal DNA isolated by standard methods from *E. coli* FM5 (WO 96/35796 and WO 2004/033646, which are each hereby incorporated by reference in their entireties, particularly with respect to isolation of nucleic acids). The final plasmids were named pKKDIy for the construct encoding the yeast *idi* gene or pKKDIc for the construct encoding the *E. coli* *idi* gene. The plasmids were transformed into *E. coli* hosts BL21 for subsequent analysis. In some cases the isoprene synthase from kudzu was cloned into pKKDIy yielding plasmid pKKDIyIS.

[0335] The lower MVA pathway was also cloned into pTrc containing a kanamycin antibiotic resistance marker. The plasmid pTrcKKDIy was digested with restriction endonucleases *ApaI* and *PstI*, the 5930 bp fragment was separated on a 1.2% agarose E-gel and purified using the Qiagen Gel Purification kit according to the manufacturer's instructions. The plasmid



pTrcKudzuKan, described in Example 7, was digested with restriction endonucleases *Apa*I and *Pst*I, and the 3338 bp fragment containing the vector was purified from a 1.2% E-gel using the Qiagen Gel Purification kit. The 3338 bp vector fragment and the 5930 bp lower MVA pathway fragment were ligated using the Roche Quick Ligation kit. The ligation mix was transformed into *E. coli* TOP10 cells and transformants were grown at 37° C overnight with selection on LA containing kanamycin (50 µg/ml). The transformants were verified by restriction enzyme digestion and one was frozen as a stock. The plasmid was designated pTrcKanKKDIy.

## II. Cloning a kudzu isoprene synthase gene into pTrcKanKKDIy

[0336] The kudzu isoprene synthase gene was amplified by PCR from pTrcKudzu, described in Example 1, using primers MCM50 5'-

GATCATGCATTTCGCCCTTAGGAGGTAAAAAACATGTGTGCGACCTCTTCTCAATTT  
ACT (SEQ ID NO:31) and MCM53 5'-

CGGTCGACGGATCCCTGCAGTTAGACATACATCAGCTG (SEQ ID NO:32). The resulting PCR fragment was cloned into pCR2.1 and transformed into *E. coli* TOP10. This fragment contains the coding sequence for kudzu isoprene synthase and an upstream region containing a RBS from *E. coli*. Transformants were incubated overnight at 37° C with selection on LA containing carbenicillin (50 µg/ml). The correct insertion of the fragment was verified by sequencing and this strain was designated MCM93.

[0337] The plasmid from strain MCM93 was digested with restriction endonucleases *Nsi*I and *Pst*I to liberate a 1724 bp insert containing the RBS and kudzu isoprene synthase. The 1724 bp fragment was separated on a 1.2% agarose E-gel and purified using the Qiagen Gel Purification kit according to the manufacturer's instructions. Plasmid pTrcKanKKDIy was digested with the restriction endonuclease *Pst*I, treated with SAP for 30 minutes at 37° C and purified using the Qiagen PCR cleanup kit. The plasmid and kudzu isoprene synthase encoding DNA fragment were ligated using the Roche Quick Ligation kit. The ligation mix was transformed into *E. coli* TOP10 cells and transformants were grown overnight at 37° C with selection on LA containing Kanamycin at 50 µg/ml. The correct transformant was verified by restriction digestion and the plasmid was designated pTrcKKDyIkISKan (Figures 24 and 25). This plasmid was transformed into BL21(λDE3) cells (Invitrogen).



III. Isoprene production from mevalonate in *E. coli* expressing the recombinant lower mevalonate pathway and isoprene synthase from kudzu.

[0338] Strain BL21/pTrcKKDyIkISKan was cultured in MOPS medium (Neidhardt *et al.*, (1974) *J. Bacteriology* 119:736-747) adjusted to pH 7.1 and supplemented with 0.5% glucose and 0.5% mevalonic acid. A control culture was also set up using identical conditions but without the addition of 0.5% mevalonic acid. The culture was started from an overnight seed culture with a 1% inoculum and induced with 500  $\mu$ M IPTG when the culture had reached an OD<sub>600</sub> of 0.3 to 0.5. The cultures were grown at 30° C with shaking at 250 rpm. The production of isoprene was analyzed 3 hours after induction by using the head space assay described in Example 1. Maximum production of isoprene was  $6.67 \times 10^4$  nmol/L<sub>broth</sub>/OD<sub>600</sub>/hr where L<sub>broth</sub> is the volume of broth and includes both the volume of the cell medium and the volume of the cells. The control culture not supplemented with mevalonic acid did not produce measurable isoprene.

IV. Cloning the upper MVA pathway

[0339] The upper mevalonate biosynthetic pathway, comprising two genes encoding three enzymatic activities, was cloned from *Enterococcus faecalis*. The *mvaE* gene encodes a protein with the enzymatic activities of both acetyl-CoA acetyltransferase and 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase, the first and third proteins in the pathway, and the *mvaS* gene encodes second enzyme in the pathway, HMG-CoA synthase. The *mvaE* gene was amplified from *E. faecalis* genomic DNA (ATCC 700802D-5) with an *E. coli* ribosome binding site and a spacer in front using the following primers:

CF 07-60 (+) Start of *mvaE* w/ RBS + ATG start codon *SacI*

5' - GAGACATGAGCTCAGGAGGTAAAAAACATGAAAACAGTAGTTATTATTG (SEQ ID NO:34)

CF 07-62 (-) Fuse *mvaE* to *mvaS* with RBS in between

5' - TTTATCAATCCCAATTGTCATGTTTTTTTACCTCCTTTATTGTTTTCTTAAATC (SEQ ID NO:35)



**[0340]** The *mvaS* gene was amplified from *E. faecalis* genomic DNA (ATCC 700802D-5) with a RBS and spacer from *E. coli* in front using the following primers:

CF 07-61 (+) Fuse *mvaE* to *mvaS* with RBS in between

5' -

GATTTAAGAAAACAATAAAGGAGGTAAAAAACATGACAATTGGGATTGATAAA  
(SEQ ID NO:36)

CF 07-102 (-) End of *mvaS* gene *Bg*III

5' -GACATGACATAGATCTTTAGTTTCGATAAGAACGAACGGT (SEQ ID NO:37)

**[0341]** The PCR fragments were fused together with PCR using the following primers:

CF 07-60 (+) Start of *mvaE* w/ RBS + ATG start codon *Sac*I

5' -GAGACATGAGCTCAGGAGGTAAAAAACATGAAACAGTAGTTATTATTG (SEQ ID NO:34)

CF 07-102 (-) End of *mvaS* gene *Bg*III

5'-GACATGACATAGATCTTTAGTTTCGATAAGAACGAACGGT (SEQ ID NO:37)

**[0342]** The fusion PCR fragment was purified using a Qiagen kit and digested with the restriction enzymes *Sac*I and *Bg*III. This digested DNA fragment was gel purified using a Qiagen kit and ligated into the commercially available vector pTrcHis2A, which had been digested with *Sac*I and *Bg*III and gel purified.

**[0343]** The ligation mix was transformed into *E. coli* Top 10 cells and colonies were selected on LA and 50 µg/ml carbenicillin plates. A total of six colonies were chosen and grown overnight in LB and 50 µg/ml carbenicillin and plasmids were isolated using a Qiagen kit. The plasmids were digested with *Sac*I and *Bg*III to check for inserts and one correct plasmid was sequenced with the following primers:

CF 07-58 (+) Start of *mvaE* gene

5' – ATGAAACAGTAGTTATTATTGATGC (SEQ ID NO:38)

CF 07-59 (-) End of *mvaE* gene

5' – ATGTTATTGTTTTCTTAAATCATTTAAAATAGC (SEQ ID NO:39)

CF 07-82 (+) Start of *mvaS* gene

5' – ATGACAATTGGGATTGATAAAATTAG (SEQ ID NO:40)



CF 07-83 (-) End of *mvaS* gene

5' – TTAGTTTCGATAAGAACGAACGGT (SEQ ID NO:41)

CF 07-86 (+) Sequence in *mvaE*

5' – GAAATAGCCCCATTAGAAGTATC (SEQ ID NO:42)

CF 07-87 (+) Sequence in *mvaE*

5' – TTGCCAATCATATGATTGAAAATC (SEQ ID NO:43)

CF 07-88 (+) Sequence in *mvaE*

5' – GCTATGCTTCATTAGATCCTTATCG (SEQ ID NO:44)

CF 07-89 (+) Sequence *mvaS*

5' – GAAACCTACATCCAATCTTTTGCCC (SEQ ID NO:45)

**[0344]** The plasmid called pTrcHis2AUpperPathway#1 was correct by sequencing and was transformed into the commercially available *E. coli* strain BL21. Selection was done on LA and 50 µg/ml carbenicillin. Two transformants were chosen and grown in LB and 50 µg/ml carbenicillin until they reached an OD<sub>600</sub> of 1.5. Both strains were frozen in a vial at -80° C in the presence of glycerol. Strains were designated CF 449 for pTrcHis2AUpperPathway#1 in BL21, isolate #1 and CF 450 for pTrcHis2AUpperPathway#1 in BL21, isolate #2. Both clones were found to behave identically when analyzed.

#### V. Cloning of UpperMVA Pathway into pCL1920

**[0345]** The plasmid pTrcHis2AUpperPathway was digested with the restriction endonuclease *SspI* to release a fragment containing pTrc-*mvaE-mvaS*-(His tag)-terminator. In this fragment, the his-tag was not translated. This blunt ended 4.5 kbp fragment was purified from a 1.2% E-gel using the Qiagen Gel Purification kit. A dephosphorylated, blunt ended 4.2 kbp fragment from pCL1920 was prepared by digesting the vector with the restriction endonuclease *PvuII*, treating with SAP and gel purifying from a 1.2% E-gel using the Qiagen Gel Purification kit. The two fragments were ligated using the Roche Quick Ligation Kit and transformed into TOP10 chemically competent cells. Transformants were selected on LA containing spectinomycin (50 µg/ml). A correct colony was identified by screening for the presence of the insert by PCR. The plasmid was designated pCL PtrcUpperPathway (Figures 26 and 27).



#### VI. Strains expressing the combined Upper and Lower Mevalonic Acid Pathways

**[0346]** To obtain a strain with a complete mevalonic acid pathway plus kudzu isoprene synthase, plasmids pTrcKKDyIkISkan and pCLpTrcUpperPathway were both transformed into BL21( $\lambda$ DE3) competent cells (Invitrogen) and transformants were selected on LA containing kanamycin (50  $\mu$ g/ml) and Spectinomycin (50  $\mu$ g/ml). The transformants were checked by plasmid prep to ensure that both plasmids were retained in the host. The strain was designated MCM127.

#### VII. Production of mevalonic acid from glucose in *E. coli*/pUpperpathway

**[0347]** Single colonies of the BL21/pTrcHis2A-*mvaE/mvaS* or FM5/pTrcHis2A-*mvaE/mvaS* are inoculated into LB and carbenicillin (100  $\mu$ g/ml) and are grown overnight at 37° C with shaking at 200 rpm. These cultures were diluted into 50 ml medium in 250 ml baffled flasks to an OD<sub>600</sub> of 0.1. The medium was TM3, 1 or 2% glucose, carbenicillin (100  $\mu$ g/ml) or TM3, 1% glucose, hydrolyzed soy oil, and carbenicillin (100  $\mu$ g/ml) or TM3 and biomass (prepared bagasse, corn stover or switchgrass). Cultures were grown at 30° C with shaking at 200 rpm for approximately 2-3 hours until an OD<sub>600</sub> of 0.4 was reached. At this point the expression from the *mvaE mvaS* construct was induced by the addition of IPTG (400  $\mu$ M). Cultures were incubated for a further 20 or 40 hours with samples taken at 2 hour intervals to 6 hour post induction and then at 24, 36 and 48 hours as needed. Sampling was done by removing 1 ml of culture, measuring the OD<sub>600</sub>, pelleting the cells in a microfuge, removing the supernatant and analyzing it for mevalonic acid.

**[0348]** A 14 liter fermentation of *E. coli* cells with nucleic acids encoding *Enterococcus faecalis* AA-CoA thiolase, HMG-CoA synthase, and HMG-CoA reductase polypeptides produced 22 grams of mevalonic acid with TM3 medium and 2% glucose as the cell medium. A shake flask of these cells produced 2-4 grams of mevalonic acid per liter with LB medium and 1% glucose as the cell culture medium. The production of mevalonic acid in these strains indicated that the MVA pathway was functional in *E. coli*.

#### VIII. Production of isoprene from *E. coli* BL21 containing the upper and lower MVA pathway plus kudzu isoprene synthase.



**[0349]** The following strains were created by transforming in various combinations of plasmids containing the upper and lower MVA pathway and the kudzu isoprene synthase gene as described above and the plasmids containing the idi, dxs, and dxr and isoprene synthase genes described in Example 7. The host cells used were chemically competent BL21( $\lambda$ DE3) and the transformations were done by standard methods. Transformants were selected on L agar containing kanamycin (50  $\mu$ g/ml) or kanamycin plus spectinomycin (both at a concentration of 50  $\mu$ g/ml). Plates were grown at 37° C. The resulting strains were designated as follows:

Grown on Kanamycin plus Spectinomycin (50  $\mu$ g/ml each)

MCM127 - pCL Upper MVA and pTrcKKDyIkIS (kan) in BL21( $\lambda$ DE3)

MCM131 - pCL1920 and pTrcKKDyIkIS (kan) in BL21( $\lambda$ DE3)

MCM125 - pCL Upper MVA and pTrcHis2B (kan) in BL21( $\lambda$ DE3)

Grown on Kanamycin (50  $\mu$ g/ml)

MCM64 - pTrcKudzu yIDI DXS (kan) in BL21( $\lambda$ DE3)

MCM50 - pTrcKudzu (kan) in BL21( $\lambda$ DE3)

MCM123 - pTrcKudzu yIDI DXS DXR (kan) in BL21( $\lambda$ DE3)

**[0350]** The above strains were streaked from freezer stocks to LA and appropriate antibiotic and grown overnight at 37° C. A single colony from each plate was used to inoculate shake flasks (25 ml LB and the appropriate antibiotic). The flasks were incubated at 22° C overnight with shaking at 200 rpm. The next morning the flasks were transferred to a 37° C incubator and grown for a further 4.5 hours with shaking at 200 rpm. The 25 ml cultures were centrifuged to pellet the cells and the cells were resuspended in 5 ml LB and the appropriate antibiotic. The cultures were then diluted into 25 ml LB, % glucose, and the appropriate antibiotic to an OD<sub>600</sub> of 0.1. Two flasks for each strain were set up, one set for induction with IPTG (800  $\mu$ M) the second set was not induced. The cultures were incubated at 37° C with shaking at 250 rpm. One set of the cultures were induced after 1.50 hours (immediately following sampling time point 1). At each sampling time point, the OD<sub>600</sub> was measured and the amount of isoprene determined as described in Example 1. Results are presented in Table 3. The amount of isoprene made is presented as the amount at the peak production for the particular strain.



**Table 3. Production of isoprene in *E. coli* strains**

| Strain | Isoprene ( $\mu\text{g/L}_{\text{broth}}/\text{hr}/\text{OD}$ ) |
|--------|-----------------------------------------------------------------|
| MCM50  | 23.8                                                            |
| MCM64  | 289                                                             |
| MCM125 | ND                                                              |
| MCM131 | Trace                                                           |
| MCM127 | 874                                                             |

ND: not detected

Trace: peak present but not integrable.

#### IX. Analysis of mevalonic acid

**[0351]** Mevalonolactone (1.0 g, 7.7 mmol) (CAS# 503-48-0) was supplied from Sigma-Aldrich (WI, USA) as a syrup that was dissolved in water (7.7 mL) and was treated with potassium hydroxide (7.7 mmol) in order to generate the potassium salt of mevalonic acid. The conversion to mevalonic acid was confirmed by  $^1\text{H}$  NMR analysis. Samples for HPLC analysis were prepared by centrifugation at 14,000 rpm for 5 minutes to remove cells, followed by the addition of a 300  $\mu\text{L}$  aliquot of supernatant to 900  $\mu\text{L}$  of  $\text{H}_2\text{O}$ . Perchloric acid (36  $\mu\text{L}$  of a 70% solution) was then added followed by mixing and cooling on ice for 5 minutes. The samples were then centrifuged again (14,000 rpm for 5 min) and the supernatant transferred to HPLC. Mevalonic acid standards (20, 10, 5, 1 and 0.5 g/L) were prepared in the same fashion. Analysis of mevalonic acid (20  $\mu\text{L}$  injection volume) was performed by HPLC using a BioRad Aminex 87-H+ column (300 mm by 7.0 mm) eluted with 5 mM sulfuric acid at 0.6 mL/min with refractive index (RI) detection. Under these conditions mevalonic acid eluted as the lactone form at 18.5 minutes.

#### **Example 9. Construction of the upper and lower MVA pathway for integration into *Bacillus subtilis***

##### **I. Construction of the Upper MVA pathway in *Bacillus subtilis***



[0352] The upper pathway from *Enterococcus faecalis* is integrated into *B. subtilis* under control of the *aprE* promoter. The upper pathway consists of two genes; *mvaE*, which encodes for AACT and HMGR, and *mvaS*, which encodes for HMGS. The two genes are fused together with a stop codon in between, an RBS site in front of *mvaS*, and are under the control of the *aprE* promoter. A terminator is situated after the *mvaE* gene. The chloramphenicol resistance marker is cloned after the *mvaE* gene and the construct is integrated at the *aprE* locus by double cross over using flanking regions of homology.

[0353] Four DNA fragments are amplified by PCR such that they contain overhangs that will allow them to be fused together by a PCR reaction. PCR amplifications are carried out using Herculase polymerase according to manufacturer's instructions.

#### 1: PaprE

CF 07-134 (+) Start of *aprE* promoter PstI

5'- GACATCTGCAGCTCCATTTTCTTCTGC (SEQ ID NO:82)

CF 07-94 (-) Fuse PaprE to *mvaE*

5'- CAATAATAACTACTGTTTTCACTCTTTACCCTCTCCTTTTAA (SEQ ID NO:83)

Template: *Bacillus subtilis* chromosomal DNA

#### 2: *mvaE*

CF 07-93 (+) fuse *mvaE* to the *aprE* promoter (GTG start codon)

5'- TTAAAAGGAGAGGGTAAAGAGTGAAAACAGTAGTTATTATTG (SEQ ID NO:84)

CF 07-62 (-) Fuse *mvaE* to *mvaS* with RBS in between

5'- TTTATCAATCCCAATTGTCATGTTTTTTTACCTCCTTTATTGTTTTCTTAAATC  
(SEQ ID NO:35)

Template: *Enterococcus faecalis* chromosomal DNA (from ATCC)

#### 3. *mvaS*

CF 07-61 (+) Fuse *mvaE* to *mvaS* with RBS in between



5'-

GATTTAAGAAAACAATAAAGGAGGTAAAAAACATGACAATTGGGATTGATAAA  
(SEQ ID NO:36)

CF 07-124 (-) Fuse the end of *mvaS* to the terminator

5'- CGGGGCCAAGGCCGGTTTTTTTAGTTTCGATAAGAACGAACGGT (SEQ ID  
NO:85)

Template: *Enterococcus faecalis* chromosomal DNA

4. *B. amyliquefaciens* alkaline serine protease terminator

CF 07-123 (+) Fuse the end of *mvaS* to the terminator

5'- ACCGTTTCGTTCTTATCGAAACTAAAAAAACCGGCCTTGGCCCCG (SEQ ID  
NO:86)

CF 07-46 (-) End of *B. amyliquefaciens* terminator BamHI

5'- GACATGACGGATCCGATTACGAATGCCGTCTC (SEQ ID NO:63)

Template: *Bacillus amyliquefaciens* chromosomal DNA

#### PCR Fusion Reactions

5. Fuse *mvaE* to *mvaS*

CF 07-93 (+) fuse *mvaE* to the *aprE* promoter (GTG start codon)

5'- TTAAAAGGAGAGGGTAAAGAGTGAAAACAGTAGTTATTATTG (SEQ ID NO:84)

CF 07-124 (-) Fuse the end of *mvaS* to the terminator

5'- CGGGGCCAAGGCCGGTTTTTTTAGTTTCGATAAGAACGAACGGT (SEQ ID  
NO:85)

Template: #2 and 3 from above

6. Fuse *mvaE-mvaS* to *aprE* promoter



CF 07-134 (+) Start of *aprE* promoter PstI

5'- GACATCTGCAGCTCCATTTTCTTCTGC (SEQ ID NO:82)

CF 07-124 (-) Fuse the end of *mvaS* to the terminator

5'- CGGGGCCAAGGCCGGTTTTTTTAGTTTCGATAAGAACGAACGGT (SEQ ID NO:85)

Template #1 and #4 from above

7. Fuse *PapE-mvaE-mvaS* to terminator

CF 07-134 (+) Start of *aprE* promoter PstI

5'- GACATCTGCAGCTCCATTTTCTTCTGC (SEQ ID NO:82)

CF 07-46 (-) End of *B. amyliquefaciens* terminator BamHI

5'- GACATGACGGATCCGATTACGAATGCCGTCTC (SEQ ID NO:63)

Template: #4 and #6

**[0354]** The product is digested with restriction endonucleases *PstI/BamHI* and ligated to pJM102 (Perego, M. 1993. Integrational vectors for genetic manipulation in *Bacillus subtilis*, p. 615–624. In A. L. Sonenshein, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis* and other gram-positive bacteria: biochemistry, physiology, and molecular genetics. American Society for Microbiology, Washington, D.C.) which is digested with *PstI/BamHI*. The ligation is transformed into *E.coli* TOP 10 chemically competent cells and transformants are selected on LA containing carbenicillin (50 µg/ml). The correct plasmid is identified by sequencing and is designated pJMUpperpathway2 (Figures 50 and 51). Purified plasmid DNA is transformed into *Bacillus subtilis aprEnprE Pxyl-comK* and transformants are selected on L agar containing chloramphenicol (5 µg/ml). A correct colony is selected and is plated sequentially on L agar containing chloramphenicol 10, 15 and 25 µg/ml to amplify the number of copies of the cassette containing the upper pathway.



[0355] The resulting strain is tested for mevalonic acid production by growing in LB containing 1% glucose and 1%. Cultures are analyzed by GC for the production of mevalonic acid.

[0356] This strain is used subsequently as a host for the integration of the lower mevalonic acid pathway.

[0357] The following primers are used to sequence the various constructs above.

Sequencing primers:

CF 07-134 (+) Start of *aprE* promoter PstI

5'- GACATCTGCAGCTCCATTTTCTTCTGC (SEQ ID NO:82)

CF 07-58 (+) Start of *mvaE* gene

5'- ATGAAAACAGTAGTTATTATTGATGC (SEQ ID NO:38)

CF 07-59 (-) End of *mvaE* gene

5'- ATGTTATTGTTTTCTTAAATCATTTAAAATAGC (SEQ ID NO:39)

CF 07-82 (+) Start of *mvaS* gene

5'- ATGACAATTGGGATTGATAAAATTAG (SEQ ID NO:40)

CF 07-83 (-) End of *mvaS* gene

5'- TTAGTTTCGATAAGAACGAACGGT (SEQ ID NO:41)

CF 07-86 (+) Sequence in *mvaE*

5'- GAAATAGCCCCATTAGAAGTATC (SEQ ID NO:42)

CF 07-87 (+) Sequence in *mvaE*

5'- TTGCCAATCATATGATTGAAAATC (SEQ ID NO:43)

CF 07-88 (+) Sequence in *mvaE*

5'- GCTATGCTTCATTAGATCCTTATCG (SEQ ID NO:44)



CF 07-89 (+) Sequence *mvaS*

5'- GAAACCTACATCCAATCTTTTGCCC (SEQ ID NO:45)

[0358] Transformants are selected on LA containing chloramphenicol at a concentration of 5 µg/ml. One colony is confirmed to have the correct integration by sequencing and is plated on LA containing increasing concentrations of chloramphenicol over several days, to a final level of 25 µg/ml. This results in amplification of the cassette containing the genes of interest. The resulting strain is designated CF 455: pJMupperpathway#1 X *Bacillus subtilis aprEnprE Pxyl comK* (amplified to grow on LA containing chloramphenicol 25 µg/ml).

## II. Construction of the Lower MVA pathway in *Bacillus subtilis*

[0359] The lower MVA pathway, consisting of the genes *mvk1*, *pmk*, *mpd* and *idi* are combined in a cassette consisting of flanking DNA regions from the *nprE* region of the *B. subtilis* chromosome (site of integration), the *aprE* promoter, and the spectinomycin resistance marker (see Figures 28 and 29). This cassette is synthesized by DNA2.0 and is integrated into the chromosome of *B. subtilis* containing the upper MVA pathway integrated at the *aprE* locus. The kudzu isoprene synthase gene is expressed from the replicating plasmid described in Example 4 and is transformed into the strain with both upper and lower pathways integrated.

## **Example 10: Production of isoprene in *E. coli* expressing *M. mazei* mevalonate kinase and *P. alba* isoprene synthase**

### I. Construction of vectors and strains encoding *M. mazei* mevalonate kinase (MVK) and *P. alba* isoprene synthase

#### (i) Construction of strain EWL201 (BL21, Cm-GI1.2-KKDyI)

[0360] *E. coli* BL21 (Novagen brand, EMD Biosciences, Inc.) was a recipient strain, transduced with MCM331 P1 lysate (lysate prepared according to the method described in Ausubel, *et al.*, *Current Protocols in Molecular Biology*. John Wiley and Sons, Inc.).

Transductants were selected for by spreading cells onto L Agar and 20 µg/µl chloramphenicol. The plates were incubated overnight at 30°C. Analysis of transductants showed no colonies on



control plates (water + cells control plate for reversion and water and P1 lysate control plate for lysate contamination).

**[0361]** Four transductants were picked and used to inoculate 5 mL L Broth and 20 µg/µl chloramphenicol. The cultures were grown overnight at 30°C with shaking at 200 rpm. To make genomic DNA preps of each transductant for PCR analysis, 1.5mL of overnight cell culture were centrifuged. The cell pellet was resuspended with 400µl Resuspension Buffer (20mM Tris, 1mM EDTA, 50mM NaCl, pH 7.5) and 4µl RNase, DNase-free (Roche) was added. The tubes were incubated at 37°C for 30 minutes followed by the addition of 4µl 10% SDS and 4µl of 10mg/ml Proteinase K stock solution (Sigma-Aldrich). The tubes were incubated at 37°C for 1 hour. The cell lysate was transferred into 2ml Phase Lock Light Gel tubes (Eppendorf) and 200µl each of saturated phenol pH7.9 (Ambion Inc.) and chloroform were added. The tubes were mixed well and microcentrifuged for 5 minutes. A second extraction was done with 400µl chloroform and the aqueous layer was transferred to a new eppendorf tube. The genomic DNA was precipitated by the addition of 1ml of 100% ethanol and centrifugation for 5 minutes. The genomic DNA pellet was washed with 1ml 70% ethanol. The ethanol was removed and the genomic DNA pellet was allowed to air dry briefly. The genomic DNA pellet was resuspended with 200µl TE.

**[0362]** Using Pfu Ultra II DNA polymerase (Stratagene) and 200ng/µl of genomic DNA as template, 2 different sets of PCR reaction tubes were prepared according to manufacturer's protocol. For set 1, primers MCM130 and GB Cm-Rev (Table 4) were used to ensure transductants were successfully integrated into the attTn7 locus. PCR parameters for set 1 were 95°C for 2 minutes (first cycle only), 95°C for 25 seconds, 55°C for 25 seconds, 72°C for 25 seconds (repeat steps 2-4 for 28 cycles), 72°C for 1 minute. For set 2, primers MVD For and MVD Rev (Table 4) were used to ensure that the gi1.2-KKDyI operon integrated properly. PCR parameters for set 2 were 95°C for 2 minutes (first cycle only), 95°C for 25 seconds, 55°C for 25 seconds, 72°C for 10 seconds (repeat steps 2-4 for 28 cycles), 72°C for 1 minute. Analysis of PCR amplicons on a 1.2% E-gel (Invitrogen Corp.) showed that all 4 transductant clones were correct (picked one and designated as strain EWL201).

ii) Construction of Strain EWL204 (BL21, loopout-GI1.2-KKDyI)



**[0363]** The chloramphenicol marker was looped out of strain EWL201 using plasmid pCP20 as described by Datsenko and Wanner (2000) (Datsenko *et al.*, *Proc Natl. Acad. Sci USA* 97:6640-6645, 2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. (Datsenko *et al.*, *PNAS*, 97: 6640-6645, 2000). EWL201 cells were grown in L Broth to midlog phase and then washed three times in ice-cold, sterile water. An aliquot of 50µl of cell suspension was mixed with 1µl of pCP20 and the cell suspension mixture was electroporated in a 2mm cuvette (Invitrogen Corp.) at 2.5 Volts and 25µF using a Gene Pulser Electroporator (Bio-Rad Inc.). 1ml of LB was immediately added to the cells, then transferred to a 14ml polypropylene tube (Sarstedt) with a metal cap. Cells were allowed to recover by growing for 1 hour at 30°C. Transformants were selected on L Agar and 20µg/µl chloramphenicol and 50µg/µl carbenicillin and incubated at 30°C overnight. The next day, a single clone was grown in 10ml L Broth and 50µg/µl carbenicillin at 30°C until early log phase. The temperature of the growing culture was then shifted to 42°C for 2 hours. Serial dilutions were made, the cells were then spread onto LA plates (no antibiotic selection), and incubated overnight at 30°C. The next day, 20 colonies were picked and patched onto L Agar (no antibiotics) and LA and 20µg/µl chloramphenicol plates. Plates were then incubated overnight at 30°C. Cells able to grow on LA plates, but not LA and 20µg/µl chloramphenicol plates, were deemed to have the chloramphenicol marker looped out (picked one and designated as strain EWL204).

iii) Construction of plasmid pEWL230 (pTrc *P. alba*)

**[0364]** Generation of a synthetic gene encoding *Populus alba* isoprene synthase (*P. alba* HGS) was outsourced to DNA2.0 Inc. (Menlo Park, CA) based on their codon optimization method for *E. coli* expression. The synthetic gene was custom cloned into plasmid pET24a (Novagen brand, EMD Biosciences, Inc.) and delivered lyophilized (Figures 54, 55A and 55B).

**[0365]** A PCR reaction was performed to amplify the *P. alba* isoprene synthase (*P. alba* HGS) gene using pET24 *P. alba* HGS as the template, primers MCM182 and MCM192, and Herculase II Fusion DNA polymerase (Stratagene) according to manufacturer's protocol. PCR conditions were as follows: 95°C for 2 minutes (first cycle only), 95°C for 25 seconds, 55°C for 20 seconds, 72°C for 1 minute, repeat for 25 cycles, with final extension at 72°C for 3 minutes.



The *P. alba* isoprene synthase PCR product was purified using QIAquick PCR Purification Kit (Qiagen Inc.).

[0366] *P. alba* isoprene synthase PCR product was then digested in a 20µl reaction containing 1µl *Bsp*HI endonuclease (New England Biolabs) with 2µl 10X NEB Buffer 4. The reaction was incubated for 2 hours at 37°C. The digested PCR fragment was then purified using the QIAquick PCR Purification Kit. A secondary restriction digest was performed in a 20µl reaction containing 1µl *Pst*I endonuclease (Roche) with 2µl 10X Buffer H. The reaction was incubated for 2 hours at 37°C. The digested PCR fragment was then purified using the QIAquick PCR Purification Kit. Plasmid pTrcHis2B (Invitrogen Corp.) was digested in a 20µl reaction containing 1µl *Nco*I endonuclease (Roche), 1µl *Pst*I endonuclease, and 2µl 10X Buffer H. The reaction was incubated for 2 hours at 37°C. The digested pTrcHis2B vector was gel purified using a 1.2% E-gel (Invitrogen Corp.) and extracted using the QIAquick Gel Extraction Kit (Qiagen) (Figure 56). Using the compatible cohesive ends of *Bsp*HI and *Nco*I sites, a 20µl ligation reaction was prepared containing 5µl *P. alba* isoprene synthase insert, 2µl pTrc vector, 1µl T4 DNA ligase (New England Biolabs), 2µl 10X ligase buffer, and 10µl ddH<sub>2</sub>O. The ligation mixture was incubated at room temperature for 40 minutes. The ligation mixture was desalted by floating a 0.025µm nitrocellulose membrane filter (Millipore) in a petri dish of ddH<sub>2</sub>O and applying the ligation mixture gently on top of the nitrocellulose membrane filter for 30 minutes at room temperature. MCM446 cells (See section II) were grown in LB to midlog phase and then washed three times in ice-cold, sterile water. An aliquot of 50µl of cell suspension was mixed with 5µl of desalted pTrc *P.alba* HGS ligation mix. The cell suspension mixture was electroporated in a 2mm cuvette at 2.5 Volts and 25µF using a Gene Pulser Electroporator. 1ml of LB is immediately added to the cells, then transferred to a 14ml polypropylene tube (Sarstedt) with a metal cap. Cells were allowed to recover by growing for 2 hour at 30°C. Transformants were selected on L Agar and 50µg/µl carbenicillin and 10mM mevalonic acid and incubated at 30°C. The next day, 6 transformants were picked and grown in 5ml L Broth and 50µg/µl carbenicillin tubes overnight at 30°C. Plasmid preps were performed on the overnight cultures using QIAquick Spin Miniprep Kit (Qiagen). Due to the use of BL21 cells for propagating plasmids, a modification of washing the spin columns with PB Buffer 5X and PE Buffer 3X was incorporated to the standard manufacturer's protocol for achieving high



quality plasmid DNA. Plasmids were digested with *Pst*I in a 20µl reaction to ensure the correct sized linear fragment. All 6 plasmids were the correct size and shipped to Quintara Biosciences (Berkeley, CA) for sequencing with primers MCM65, MCM66, EL1000 (Table 4). DNA sequencing results showed all 6 plasmids were correct. Picked one and designated plasmid as EWL230 (Figures 57, 58A and 58B).

iv) Construction of plasmid pEWL244 (pTrc *P. alba*-mMVK)

**[0367]** A PCR reaction was performed to amplify the *Methanosarcina mazei* (*M. mazei*) MVK gene using MCM376 as the template (see section v), primers MCM165 and MCM177 (see Table 4), and Pfu Ultra II Fusion DNA polymerase (Stratagene) according to manufacturer's protocol. PCR conditions were as follows: 95°C for 2 minutes (first cycle only), 95°C for 25 seconds, 55°C for 25 seconds, 72°C for 18 seconds, repeat for 28 cycles, with final extension at 72°C for 1 minute. The *M. mazei* MVK PCR product was purified using QIAquick PCR Purification Kit (Qiagen Inc.,)

**[0368]** The *M. mazei* MVK PCR product was then digested in a 40µl reaction containing 8µl PCR product, 2µl *Pme*I endonuclease (New England Biolabs), 4µl 10X NEB Buffer 4, 4µl 10X NEB BSA, and 22µl of ddH<sub>2</sub>O. The reaction was incubated for 3 hours at 37°C. The digested PCR fragment was then purified using the QIAquick PCR Purification Kit. A secondary restriction digest was performed in a 47µl reaction containing 2µl *Nsi*I endonuclease (Roche), 4.7µl 10X Buffer H, and 40µl of *Pme*I digested *M. mazei* MVK fragment. The reaction was incubated for 3 hours at 37°C. The digested PCR fragment was then gel purified using a 1.2% E-gel and extracted using the QIAquick Gel Extraction Kit. Plasmid EWL230 was digested in a 40µl reaction containing 10µl plasmid, 2µl *Pme*I endonuclease, 4µl 10X NEB Buffer 4, 4µl 10X NEB BSA, and 20µl of ddH<sub>2</sub>O. The reaction was incubated for 3 hours at 37°C. The digested PCR fragment was then purified using the QIAquick PCR Purification Kit. A secondary restriction digest was performed in a 47µl reaction containing 2µl *Pst*I endonuclease, 4.7µl 10X Buffer H, and 40µl of *Pme*I digested EWL230 linear fragment. The reaction was incubated for 3 hours at 37°C. The digested PCR fragment was then gel purified using a 1.2% E-gel and extracted using the QIAquick Gel Extraction Kit (Figure 59). Using the compatible cohesive ends of *Nsi*I and *Pst*I sites, a 20µl ligation reaction was prepared containing 8µl *M.*



*mazei* MVK insert, 3µl EWL230 plasmid, 1µl T4 DNA ligase, 2µl 10X ligase buffer, and 6µl ddH<sub>2</sub>O. The ligation mixture was incubated overnight at 16°C. The next day, the ligation mixture was desalted by floating a 0.025µm nitrocellulose membrane filter in a petri dish of ddH<sub>2</sub>O and applying the ligation mixture gently on top of the nitrocellulose membrane filter for 30 minutes at room temperature. MCM446 cells were grown in LB to midlog phase and then washed three times in ice-cold, sterile water. An aliquot of 50µl of cell suspension was mixed with 5µl of desalted pTrc P.alba-mMVK ligation mix. The cell suspension mixture was electroporated in a 2mm cuvette at 2.5 Volts and 25uFd using a Gene Pulser Electroporator. 1ml of LB is immediately added to the cells, then the cells are transferred to a 14ml polypropylene tube with a metal cap. Cells were allowed to recover by growing for 2 hour at 30°C.

Transformants were selected on LA and 50µg/µl carbenicillin and 5mM mevalonic acid plates and incubated at 30°C. The next day, 6 transformants were picked and grown in 5ml LB and 50µg/µl carbenicillin tubes overnight at 30°C. Plasmid preps were performed on the overnight cultures using QIAquick Spin Miniprep Kit. Due to the use of BL21 cells for propagating plasmids, a modification of washing the spin columns with PB Buffer 5X and PE Buffer 3X was incorporated to the standard manufacturer's protocol for achieving high quality plasmid DNA. Plasmids were digested with PstI in a 20µl reaction to ensure the correct sized linear fragment. Three of the 6 plasmids were the correct size and shipped to Quintara Biosciences for sequencing with primers MCM65, MCM66, EL1000, EL1003, and EL1006 (Table 4). DNA sequencing results showed all 3 plasmids were correct. Picked one and designated plasmid as EWL244 (Figures 60 and 61A-B).

v) Construction of plasmid MCM376 - MVK from *M. mazei* archaeal Lower in pET200D.

**[0369]** The MVK ORF from the *M. mazei* archaeal Lower Pathway operon (Figures 73A-C) was PCR amplified using primers MCM161 and MCM162 (Table 4) using the Invitrogen Platinum HiFi PCR mix. 45 uL of PCR mix was combined with 1 uL template, 1 uL of each primer at 10 uM, and 2 uL water. The reaction was cycled as follows: 94 °C for 2:00 minutes; 30 cycles of 94 °C for 0:30 minutes, 55 °C for 0:30 minutes and 68 °C for 1:15 minutes; and then 72 °C for 7:00 minutes, and 4 °C until cool. 3 uL of this PCR reaction was ligated to Invitrogen pET200D plasmid according to the manufacturer's protocol. 3 uL of this ligation was introduced into Invitrogen TOP10 cells, and transformants were selected on LA/kan50. A



plasmid from a transformant was isolated and the insert sequenced, resulting in MCM376 (Figures 74A-C).

vi) Construction of strain EWL251 (BL21(DE3), Cm-GI1.2-KKDyI, pTrc P.alba-mMVK)

**[0370]** MCM331 cells (which contain chromosomal construct gi1.2KKDyI encoding *S. cerevisiae* mevalonate kinase, mevalonate phosphate kinase, mevalonate pyrophosphate decarboxylase, and IPP isomerase) were grown in LB to midlog phase and then washed three times in ice-cold, sterile water. Mixed 50µl of cell suspension with 1µl of plasmid EWL244. The cell suspension mixture was electroporated in a 2mm cuvette at 2.5 Volts and 25uFd using a Gene Pulser Electroporator. 1ml of LB is immediately added to the cells, then the cells were transferred to a 14ml polypropylene tube with a metal cap. Cells were allowed to recover by growing for 2 hours at 30°C. Transformants were selected on LA and 50µg/µl carbenicillin and 5mM mevalonic acid plates and incubated at 37°C. One colony was selected and designated as strain EWL251.

vii) Construction of strain EWL256 (BL21(DE3), Cm-GI1.2-KKDyI, pTrc P.alba-mMVK, pCL Upper MVA)

**[0371]** EWL251 cells were grown in LB to midlog phase and then washed three times in ice-cold, sterile water. Mixed 50µl of cell suspension with 1µl of plasmid MCM82 (which is pCL PtrcUpperPathway encoding *E. faecalis* mvaE and mvaS). The cell suspension mixture was electroporated in a 2mm cuvette at 2.5 Volts and 25uFd using a Gene Pulser Electroporator. 1ml of LB was immediately added to the cells, then transferred to a 14ml polypropylene tube with a metal cap. Cells were allowed to recover by growing for 2 hour at 30°C. Transformants were selected on LA and 50µg/µl carbenicillin and 50µg/µl spectinomycin plates and incubated at 37°C. Picked one colony and designated as strain EWL256.

Table 4: Primer Sequences

| Primer name | Primer sequence                      |
|-------------|--------------------------------------|
| MCM130      | ACCAATTGCACCCGGCAGA (SEQ ID NO:94)   |
| GB Cm Rev   | GCTAAAGCGCATGCTCCAGAC (SEQ ID NO:95) |
| MVD         | GACTGGCCTCAGATGAAAGC (SEQ ID NO:96)  |



|            |                                                                                       |
|------------|---------------------------------------------------------------------------------------|
| For        |                                                                                       |
| MVD<br>Rev | CAAACATGTGGCATGGAAAG (SEQ ID NO:97)                                                   |
| MCM182     | GGGCCCCGTTTAAACTTTAACTAGACTCTGCAGTTAGCGTTCAAACGGCAGAA<br>(SEQ ID NO:98)               |
| MCM192     | CGCATGCATGTCATGAGATGTAGCGTGTCCACCGAAAA (SEQ ID NO:99)                                 |
| MCM65      | ACAATTTACACAGGAAACAGC (SEQ ID NO:100)                                                 |
| MCM66      | CCAGGCAAATTCTGTTTTATCAG (SEQ ID NO:101)                                               |
| EL1000     | GCACTGTCTTTCCGTCTGCTGC (SEQ ID NO:102)                                                |
| MCM165     | GCGAACGATGCATAAAGGAGGTAAAAAACATGGTATCCTGTTCTGCGCCGGG<br>TAAGATTTACCTG (SEQ ID NO:103) |
| MCM177     | GGGCCCCGTTTAAACTTTAACTAGACTTTAATCTACTTTCAGACCTTGC (SEQ ID<br>NO:104)                  |
| EL1003     | GATAGTAACGGCTGCGCTGCTACC (SEQ ID NO:105)                                              |
| EL1006     | GACAGCTTATCATCGACTGCACG (SEQ ID NO:106)                                               |
| MCM161     | CACCATGGTATCCTGTTCTGCG (SEQ ID NO:107)                                                |
| MCM162     | TTAATCTACTTTCAGACCTTGC (SEQ ID NO:108)                                                |

## II. Construction of MCM442-449: BL21 and BL21(DE3) with FRT-cmR-FRT-gil.x-mKKDyI

### i) Construction of Template for Recombination

**[0372]** FRT-based recombination cassettes, and plasmids for Red/ET-mediated integration and antibiotic marker loopout were obtained from Gene Bridges GmbH (Germany). Procedures using these materials were carried out according to Gene Bridges protocols. Primers MCM193 and MCM195 were used to amplify the resistance cassette from the FRT-gb2-Cm-FRT template using Stratagene Hercules II Fusion kit according to the manufacturer's protocol. The 50uL reaction was cycled as follows: 95°C, 2 minutes; (95°C, 20 seconds, 55°C, 20 seconds, 72°C, 1 minute) x 5, (95°C, 20 seconds, 60°C, 20 seconds, 72°C, 1 minute) x 25; 72°C, 3 minutes; 4°C until cool. The amplicon was purified by a Qiagen PCR column according to the manufacturer's protocol and eluted in 30uL EB (Elution Buffer). DNA was digested with *NdeI* and *PciI* in a 20uL reaction with 1x Roche H buffer and 0.5uL BSA. Plasmid MCM376 was digested in a 10uL reaction containing 1uL each of *NdeI*, *NcoI*, and Roche H buffer. Reactions proceeded overnight at 37°C, and then cut DNA was purified on Qiagen PCR columns and eluted in 30uL EB. The PCR product was ligated into MCM376 in a reaction containing 1uL vector, 3uL PCR



product, 1uL Roche Quick Ligase Buffer 2, 5uL Buffer1, and 1uL Ligase. The reaction proceeded at room temperature for 3 hours and then 5uL was transformed into Invitrogen TOP10 cells according to the manufacturer's protocol. Transformants were selected on L agar (LA) and chloramphenicol (10 ug/mL) at 37°C overnight.

[0373] Transformant colonies were patched onto LA containing chloramphenicol (30 ug/mL) and kanamycin (50 ug/mL) for storage and sent to Quintara (Berkeley, CA) for sequencing. Four clones, one each with the four different nucleotides at the "N" in primer MCM195, were found to have the correct sequence for the inserted promoter. Clones were grown in 5mL LB containing chloramphenicol (30 ug/mL) and kanamycin (50 ug/mL) and used for the preparation of plasmid DNA. This plasmid was retransformed into TOP10 cells and strains were frozen as:

Table 5: MCM484-487

|        |                                                         |
|--------|---------------------------------------------------------|
| MCM484 | cmR-gi1.6-MVK(mazei) in pET (clone A1-3, variable nt A) |
| MCM485 | cmR-gi1.0-MVK(mazei) in pET (clone B4-6, variable nt C) |
| MCM486 | cmR-gi1.2-MVK(mazei) in pET (clone C1-5, variable nt G) |
| MCM487 | cmR-gi1.5-MVK(mazei) in pET (clone C3-3, variable nt T) |

ii) Creation of Recombination Target Strains MCM349 and MCM441

[0374] The chloramphenicol resistance (cmR) marker was looped out of strain MCM331 using plasmid pGB706 (GeneBridges) according to Manufacturer's instructions. MCM331 cells were grown to mid-log in LB and washed three times in iced, sterile water. A 1uL aliquot of pGB706 DNA was added to 50uL of cell suspension and this mixture was electroporated in a 2mm cuvette at 2.5 volts, 25uF followed immediately by recovery in 500uL LB for one hour at 30°C. Transformants were selected on LB containing tetracycline (5 ug/mL) at 30°C. The following day, a clone was grown up at 30°C in LB containing tetracycline (5 ug/mL) until visibly turbid (OD<sub>600</sub>~0.5-0.8). This culture was streaked onto LB and grown overnight at 37°C. A clone that was unable to grow on LB containing chloramphenicol (10 ug/mL) or LB containing tetracycline (5 ug/mL) was frozen as MCM348. Plasmid MCM356 (pRedET carbencillin; GeneBridges) was electroporated in as described above and transformants were selected on LB containing carbencillin (50 ug/mL) at 30°C. A clone was grown in LB carbencillin (50 ug/mL) at 30°C and frozen as MCM349.



[0375] Strain MCM441 was created by electrotransforming plasmid MCM356 into EWL204 as above.

iii) Recombination of FRT-cmR-FRT-gil.x-mMVK into MCM349 and MCM441

[0376] Plasmids MCM484-487 were used as template for PCR amplification with primers MCM120 and MCM196 and Herculase II Fusion kit, according to the manufacturer's protocol. Three reactions per template were carried out, with 0, 1, or 3uL DMSO. The 50uL reactions were cycled as follows: 95°C, 2 minutes; (95°C, 20 seconds; 55°C 20 seconds; 72°C, 1.5 minutes) for five cycles; (95°C, 20 seconds; 60°C 20 seconds; 72°C, 1.5 minutes) for 25 cycles; 72°C for 3 minutes; 4°C, overnight.] The three reactions from a given template were pooled and purified on Qiagen PCR columns and eluted with 30 uL EB at 60°C. 5uL DNA was digested with 1uL *DpnI* in 1x Roche Buffer A for 3 hours at 37°C. This DNA was then microdialyzed against excess water for 30 minutes.

[0377] Strains were grown in 5mL LB containing carbenicillin (50 ug/mL) from fresh streaks at 30C to an OD600 of ~0.5. 40mM L-arabinose was added and cultures were incubated at 37C for 1.5 hours. Cells were harvested and electroporated with 3uL dialyzed amplicons above, and then recovered in 500uL SOC at 37C for 1.5-3 hours. Transformants were selected on LA plates containing chloramphenicol (5 ug/mL) at 37°C.

[0378] Kanamycin sensitive clones were screened by PCR for insertion of the amplicon. PCR products from positive clones were sequenced to verify the sequence of inserted DNA. Amplicons were consistent with the FRT-gil.2-yKKDyI at attTn7 in MCM441 and 348 being replaced by FRT-cmR-FRT-gil.x-mKKDyI (The yK and mK designations refer to the mevalonate kinase from *Saccharomyces cerevisiae* and *Methanosarcina mazei* respectively).

Table 6A: The following strains were grown in LB containing chloramphenicol (5 ug/mL) and frozen.

| Strain ID | Name                                      | Parent | Recombination Amplicon Template |
|-----------|-------------------------------------------|--------|---------------------------------|
| MCM442    | BL21(DE3) cmR-gil.6mKKDyI A1, clone37 (A) | MCM349 | MCM484                          |



|        |                                            |        |        |
|--------|--------------------------------------------|--------|--------|
| MCM443 | BL21(DE3) cmR-gi1.0mKKDyI B4, clone27 (C ) | MCM349 | MCM485 |
| MCM444 | BL21(DE3) cmR-gi1.2mKKDyI C1, clone16 (G)  | MCM349 | MCM486 |
| MCM445 | BL21(DE3) cmR-gi1.5mKKDyI C3, clone7 (T)   | MCM349 | MCM487 |
| MCM446 | BL21 cmR-gi1.6mKKDyI A1-3 (A)              | MCM441 | MCM484 |
| MCM447 | BL21 cmR-gi1.0mKKDyI B4-6 (C )             | MCM441 | MCM485 |
| MCM448 | BL21 cmR-gi1.2mKKDyI C1-5 (G)              | MCM441 | MCM486 |
| MCM449 | BL21 cmR-gi1.5mKKDyI C3-3 (T)              | MCM441 | MCM487 |

Table 6B: Primers

|        |                                                                                                                                                          |
|--------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| MCM120 | AAAGTAGCCGAAGATGACGGTTTGTACATGGAGTTGGCAGGAT<br>GTTTGATTAAAAGCAATTAACCCTCACTAAAGGGCGG (SEQ ID<br>NO:109)                                                  |
| MCM193 | GATATACATATGAATTAACCCTCACTAAAGG (SEQ ID NO:110)                                                                                                          |
| MCM195 | GCATGCATGACATGTTTTTTTACCTCCTTTGTTATCCGCTCACAAT<br>TAGTGGTTGAATTATTTGCTCAGGATGTGGCATNGTCAAGGGCG<br>CGGCCGCGATCTAATACGACTCACTATAGGGCTCG (SEQ ID<br>NO:111) |
| MCM196 | AGGCTCTCAACTCTGACATGTTTTTTTCCTCCTTAAGGGTGCAGG<br>CCTATCGCAAATTAGCTTAATCTACTTTCAGACCTTGCTCGG<br>(SEQ ID NO:112)                                           |

III. The effect of yeast extract on isoprene production in *E. coli* expressing genes from the mevalonic acid pathway and grown in fed-batch culture at the 15-L scale

[0379] Medium Recipe (per liter fermentation medium):

[0380]  $K_2HPO_4$  7.5 g,  $MgSO_4 \cdot 7H_2O$  2 g, citric acid monohydrate 2 g, ferric ammonium citrate 0.3 g, yeast extract 0.5 g, 1000X Modified Trace Metal Solution 1 ml. All of the components were added together and dissolved in  $diH_2O$ . This solution was autoclaved. The pH was adjusted to 7.0 with ammonium hydroxide (30%) and q.s. to volume. Glucose 10 g, thiamine \* HCl 0.1 g, and antibiotics were added after sterilization and pH adjustment.

[0381] 1000X Modified Trace Metal Solution:



**[0382]** Citric Acids \* H<sub>2</sub>O 40 g, MnSO<sub>4</sub> \* H<sub>2</sub>O 30 g, NaCl 10 g, FeSO<sub>4</sub> \* 7H<sub>2</sub>O 1 g, CoCl<sub>2</sub> \* 6H<sub>2</sub>O 1 g, ZnSO<sub>4</sub> \* 7H<sub>2</sub>O 1 g, CuSO<sub>4</sub> \* 5H<sub>2</sub>O 100 mg, H<sub>3</sub>BO<sub>3</sub> 100 mg, NaMoO<sub>4</sub> \* 2H<sub>2</sub>O 100 mg. Each component is dissolved one at a time in Di H<sub>2</sub>O, pH to 3.0 with HCl/NaOH, then q.s. to volume and filter sterilized with a 0.22 micron filter.

**[0383]** Fermentation was performed in a 15-L bioreactor with BL21 (DE3) *E. coli* cells containing the upper mevalonic acid (MVA) pathway (pCL Upper), the integrated lower MVA pathway (gil.2KKDyI), and high expression of mevalonate kinase from *M. mazei* and isoprene synthase from *P. alba* (pTrcAlba-mMVK). This experiment was carried out to monitor isoprene formation from glucose at the desired fermentation pH 7.0 and temperature 30°C. A frozen vial of the *E. coli* strain was thawed and inoculated into tryptone-yeast extract medium. After the inoculum grew to OD 1.0, measured at 550 nm, 500 mL was used to inoculate a 15-L bioreactor bringing the initial volume to 5-L.

i) Production of isoprene in *E.coli* cells (EL256) grown in fed-batch culture without yeast extract feeding

**[0384]** Glucose was fed at an exponential rate until cells reached the stationary phase. After this time the glucose feed was decreased to meet metabolic demands. The total amount of glucose delivered to the bioreactor during the 67 hour fermentation was 3.9 kg. Induction was achieved by adding isopropyl-beta-D-1-thiogalactopyranoside (IPTG). The IPTG concentration was brought to 102 uM when the optical density at 550 nm (OD<sub>550</sub>) reached a value of 9. The IPTG concentration was raised to 192 uM when OD<sub>550</sub> reached 140. The OD<sub>550</sub> profile within the bioreactor over time is shown in Figure 67A. The isoprene level in the off gas from the bioreactor was determined using a Hiden mass spectrometer. The isoprene titer increased over the course of the fermentation to a final value of 35.6 g/L (Figure 67B). The total amount of isoprene produced during the 67 hour fermentation was 320.6 g and the time course of production is shown in Figure 67C. The metabolic activity profile, as measured by TCER, is shown in Figure 67D. The molar yield of utilized carbon that went into producing isoprene during fermentation was 17.9%. The weight percent yield of isoprene from glucose was 8.1%.

**[0385]** Production of isoprene in *E.coli* cells (EL256) grown in fed-batch culture with yeast extract feeding



[0386] Glucose was fed at an exponential rate until cells reached the stationary phase. After this time the glucose feed was decreased to meet metabolic demands. The total amount of glucose delivered to the bioreactor during the 68 hour fermentation was 7.1 kg. A total of 1.06 kg of yeast extract was also fed during the fermentation. Induction was achieved by adding IPTG. The IPTG concentration was brought to 208  $\mu$ M when the optical density at 550 nm ( $OD_{550}$ ) reached a value of 7. The IPTG concentration was raised to 193  $\mu$ M when  $OD_{550}$  reached 180. The  $OD_{550}$  profile within the bioreactor over time is shown in Figure 68A. The isoprene level in the off gas from the bioreactor was determined using a Hiden mass spectrometer. The isoprene titer increased over the course of the fermentation to a maximum value of 32.2 g/L (Figure 68B). The total amount of isoprene produced during the 68 hour fermentation was 395.5 g and the time course of production is shown in Figure 68C. The time course of volumetric productivity is shown in Figure 68D and shows that an average rate of 1.1 g/L/hr was maintained for between 23 and 63 hours. The metabolic activity profile, as measured by CER, is shown in Figure 68E. The molar yield of utilized carbon that went into producing isoprene during fermentation was 10.3%. The weight percent yield of isoprene from glucose was 5.2%.

#### IV. Production of isoprene from different carbon sources in *E. coli* harboring the mevalonic acid (MVA) pathway and isoprene synthase (EWL256)

[0387] Media Recipe (per liter fermentation media):

[0388]  $K_2HPO_4$  13.6 g,  $KH_2PO_4$  13.6g,  $MgSO_4 \cdot 7H_2O$  2 g, citric acid monohydrate 2 g, ferric ammonium citrate 0.3 g,  $(NH_4)_2SO_4$  3.2 g, yeast extract 0.2 g, 1000X Modified Trace Metal Solution 1 ml. All of the components were dissolved sequentially in  $diH_2O$ . The pH was adjusted to 6.8 with ammonium hydroxide (30%) and brought to volume. Media was filter sterilized with a 0.22 micron filter. Carbon source was added to a final concentration of 1%. Required antibiotics were added after sterilization and pH adjustment.

[0389] 1000X Trace Metal Solution (per liter fermentation media):

[0390] Citric Acids  $\cdot H_2O$  40 g,  $MnSO_4 \cdot H_2O$  30 g, NaCl 10 g,  $FeSO_4 \cdot 7H_2O$  1 g,  $CoCl_2 \cdot 6H_2O$  1 g,  $ZnSO_4 \cdot 7H_2O$  1 g,  $CuSO_4 \cdot 5H_2O$  100 mg,  $H_3BO_3$  100 mg,  $NaMoO_4 \cdot 2H_2O$  100



mg. Each component was dissolved one at a time in diH<sub>2</sub>O, pH to 3.0 with HCl/NaOH, and then brought to volume and filter sterilized with a 0.22 micron filter.

i) Preparation of AFEX biomass hydrolysate

**[0391]** AFEX pretreated corn stover was hydrolyzed to prepare biomass hydrolysate containing both xylose, glucose and acetate.

**[0392]** AFEX pretreated corn stover, received from Michigan Biotechnology Institute, was used. The pretreatment conditions were, 60% moisture, 1:1 ammonia loading, and 90 °C for 30 minutes, then air dried. The moisture content in the AFEX pretreated corn stover was 21.27%. Content of glucan and xylan in the AFEX pretreated corn stover were 31.7% and 19.1% (dry basis) respectively. The enzyme used was accellerase 1000, Grindamyl H121 (Danisco xylanase product from *Aspergillus niger* for bread-making industry).

**[0393]** For saccharification, 20 g of AFEX pretreated corn stover was added into a 500 ml flask, together with 5 ml of 1 M pH 4.8 sodium citrate buffer, 2.25 ml of Accellerase 1000, 0.1 ml of Grindamyl H121, and 72.65 ml of DI water. The flask was put in an orbital shaker, and incubated at 50°C for 96 hours.

**[0394]** For analysis, one sample was taken from the shaker, and analyzed using HPLC. The hydrolysate contained 37.2 g/l of glucose and 24.3 g/L of xylose, and 7.6 g/L of oligomers of glucose and/or xylose. Additionally, the hydrolysate also contains 1.17 g/L acetate.

ii) Experimental procedure

**[0395]** An inoculum of the *E. coli* strain EWL256 containing the MVA pathway and isoprene synthase was taken from a frozen vial and streaked onto an LB broth agar plate containing spectinomycin (50 ug/mL) and carbinicillin (50 ug/mL) and incubated at 30°C overnight. A single colony was inoculated into TM3 media containing glucose, xylose, glycerol, acetate or biomass as only carbon source and grown overnight at 30°C. Cells grow on acetate reached a significantly lower optical density. Cells grown on glucose, glycerol, biomass hydrolysate or acetate were diluted into 20 mL of TM3 media containing the respective carbon sources to reach an optical density of between 0.1 measured at 600nm. A negative control not containing any



carbon source was prepared from the glucose overnight culture. A separate experiment was performed with glucose and xylose, where the cultures were diluted to an optical density of 0.05. All culture conditions (except for acetate and glycerol) were tested in duplicates and the presented results are averaged between these cultures. Production of isoprene was induced with 200 $\mu$ M IPTG from the beginning of the experiment. The flasks were incubated at 30°C in an orbital shaker (200 rpm) and growth was followed by measuring optical density. After the glucose fed cultures had reached an optical density of approximately 0.4, samples were analyzed for isoprene production from all the tested carbon sources every hour for three hours. Samples of 100  $\mu$ L were transferred in duplicates to 2 mL glass vials, sealed and incubated for 30 min at 30°C. The bacteria were then heat killed by incubation at 80°C for 8 minutes. The amount of produced isoprene was measured using GC-MS and specific productivity ( $\mu$ g/L\*hr) was calculated.

### iii) Results

**[0396]** Significant production of isoprene could be demonstrated during growth on all the tested carbon sources. These carbon sources are examples of common alcohols, organic acids, sugars containing 5 or 6 carbon units (C5 or C6), and biomass hydrolysate.

**[0397]** The initial growth rate on biomass hydrolysate was comparable to the growth rate on glucose (Figure 69A). The initial specific productivity during growth on biomass hydrolysate was significantly higher than during growth on glucose. This demonstrates that biomass hydrolysate can be used as an efficient source of carbon for the production of isoprene. The specific productivity declined after 255 minutes of growth on biomass hydrolysate (Figure 69B). The bacteria had a slower growth rate with xylose as only carbon source when compared to glucose (Figure 69C), but a significant specific isoprene productivity was measured (Figure 69D). This shows that both C5 and C6 sugars can be utilized for the production of isoprene via the mevalonate acid pathway.

**[0398]** Surprisingly, bacteria grown on acetate as the only carbon source had a specific productivity of isoprene approximately twice as high as during growth on glucose (Figure 69A). The bacteria grew slower on acetate when compared to glucose (Figure 69B), but the performed



experiment demonstrates that acetate can also be used as a carbon source for the production of isoprene. Acetate was also present in the biomass hydrolysate as measured by HPLC.

[0399] The bacteria grew well with glycerol as only carbon source (Figure 69A) and significant production of isoprene was demonstrated (Figure 69B). This shows that common alcohols may also be used as carbon sources for production of isoprene via the mevalonate acid pathway.

**Example 11: Expression of isoprene-synthase from plant in *Streptomyces* sp.**

[0400] The gene for isoprene synthase Kudzu was obtained from plasmid pJ201:19813. Plasmid pJ201:19813 encodes isoprene synthase from *Pueraria lobata* (Kudzu plant) and was codon-optimized for *Pseudomonas fluorescens*, *Pseudomonas putida*, *Rhodopseudomonas palustris* and *Corynebacterium* (Figures 79A-79C (SEQ ID NO:123)). Digestion of plasmid pJ201:19813 with restriction enzymes NdeI and BamHI liberated gene iso19813 that was ligated into the *Streptomyces-E.coli* shuttle vector pUWL201PW (Doumith *et al.*, *Mol. Gen. Genet.* 264: 477-485, 2000; Figure 71) to generate pUWL201\_iso. Successful cloning was verified by restriction analysis of pUWL201\_iso. Expression of isoprene synthase iso19813 was under control of the *erm*-promoter which allows for constitutive expression in *Streptomyces* species, but not for expression in *E. coli*.

[0401] PUWL201PW (no insert) and pUWL201\_iso were introduced in *Streptomyces albus* J1074 (Sanchez *et al.*, *Chem. Biol.* 9:519-531, 2002) by transformation of protoplasts as described by Hopwood *et al.*, *The John innes foundation, Norwich*, 1985.

[0402] A 200 µl aliquot of protoplast suspensions was transformed with 1.9 µg pUWL201PW or 2.9 µg pUWL201\_iso. After incubation overnight at 28°C on non-selective R5-agarplates, positive transformants were selected by further incubation for 4 days in R3-overlay agar containing thiostrepton (250 µg/ml). Thiostrepton resistant transformants were examined for presence of the pUWL-plasmids by plasmid preparation using Plasmid Mini Kit (Qiagen). Prepared plasmid DNA was reintroduced in *E. coli* DH5α to generate sufficient amounts of plasmid DNA to be analyzed by restriction analysis. Positive transformants were selected on ampicillin-containing L-agar plates and insert analysis was done by digestion of plasmid DNA



with *NdeI* and *BamHI* endonucleases. Isoprene synthase was identified as a 1.7 kb fragment in positive pUWL201 iso clones while in the control strains (pUWL201PW) no such fragment was observed.

**[0403]** Wild type strain and transformants of *S. albus* containing control plasmid pUWL201PW or isoprene synthase encoding pUWL201\_iso were analyzed for isoprene formation. Strains were cultivated in duplicate on solid media (tryptic soy broth agar, TSB; 2.5 ml) in presence or absence of thiostrepton (200 µg/ml) and incubated for 4 days at 28°C in sealed head-space vials (total volume 20 ml). 500 µl head-space samples (end point measurements) were analyzed by GC-MS in SIM-mode and isoprene was identified according to reference retention times and molecular masses (67 m/z). Isoprene present in head-space samples was quantified by previously generated calibration curves. While wild-type *S. albus* and control strains harboring pUWL201PW produced isoprene in concentrations slightly higher than the detection limit (0.04 – 0.07 ppm), *S. albus* harboring pUWL201\_iso produced isoprene in at least tenfold excess compared to controls (0.75 ppm; Figure 72). The results demonstrate successful expression of plant-derived isoprene synthase in a prokaryotic organism of the *Actinomycetes* group.

**Example 12: Production of isoprene or mevalonate from fatty acid or palm oil in *E. coli* *fadR atoC* LS5218 containing the upper or upper and lower Mevalonic Acid pathway plus kudzu isoprene synthase.**

**[0404]** *Escherichia coli fadR atoC* strain LS5218 (#6966) was obtained from the Coli Genetic Stock Center. FadR encodes a transcription repressor that negatively regulates expression of the genes encoding fatty acid degradation enzymes (Campbell *et al.*, *J. Bacteriol.* 183: 5982-5990, 2001). AtoC is a response regulator in a two-component regulatory system with AtoS, regulates acetolactate metabolism. The *fadR atoC* strain allows constitutive expression of the fatty acid degradation genes and incorporates long chain fatty acids into long-chain-length polyhydroxyalkanoates. When palm oil is used as a carbon source for either mevalonate or isoprene production, the palm oil was converted to glycerol plus fatty acid. Methods for this are well known in the art, and it can be done either enzymatically by incubation with a lipase (for



example Porcine pancreatic lipase, *Candida rugosa* lipase, or other similar lipases) or chemically by saponification with a base such as sodium hydroxide.

i) *E. coli fadR atoC* strain expressing the upper Mevalonic Acid pathway

**[0405]** Strain WW4 was created by electroporating pCLPtrcUpperPathway into LS5218 using standard methods (Sambrooke *et al.*, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., Cold Spring Harbor, 1989). Incorporation of the plasmid was demonstrated by the production of mevalonic acid (MVA) when cells were cultured in TM3 medium supplemented with either C12 fatty acid (FA) or palm oil as the carbon source. To demonstrate production of MVA by WW4 from fatty acid, cells from an overnight culture were diluted 1 to 100 into 5 mL of modified TM3 medium (TM3 without yeast extract) supplemented with 0.25% C12 FA (Sigma cat # L9755). The first sign of MVA production (24 mg/L) was apparent after overnight incubation at 30°C of the IPTG induced culture. Production increased over three days with the final level of 194 mg/L of MVA produced. To demonstrate production of MVA by WW4 from oil, cells from an overnight culture were diluted 1 to 100 into modified TM3 medium supplemented with 200 mg of digested palm oil per 5 mL of TM3 medium. The first sign of MVA production (50 mg/L) was apparent after overnight incubation of the IPTG induced culture at 30°C. Production increased over three days with a final level of 500 mg/L of MVA produced.

ii) *E. coli fadR atoC* strain expressing the upper and lower MVA pathway plus kudzu isoprene synthase

**[0406]** *Escherichia coli* strain WW4 (LS5218 fadR atoC pCLPtrcUpperPathway) was transformed with pMCM118 [pTrcKKDyIkIS] to yield WW10. The incorporation of the plasmid was demonstrated by evidence of production of isoprene when the strain was cultured in TM3 and glucose and induced with IPTG (100, 300, or 900 uM). The strain was relatively sensitive to IPTG and showed a significant growth defect even at 100 uM IPTG. These results are shown in Figure 70A.

**[0407]** To test isoprene production from dodecanoic acid, WW10 was cultured overnight in L broth containing spectinomycin (50 ug/ml), and kanamycin (50 ug/ml) at 37C with shaking at 200 rpm. The cells were washed with modified TM3 medium by centrifugation and resuspension in their original culture volume with this medium. The washed and resuspended



cells from this starter culture were diluted 1 to 100 and 1 to 10 into 5 mL of modified TM3 medium containing 0.125% C12 Fatty Acid (Sigma cat # L9755).

**[0408]** To demonstrate production of mevalonate from palm oil, the oil was predigested with lipase at 37°C and 250 rpm for several days to release the fatty acids (evidence of hydrolysis was judged by the foam formed when tubes were shaken).

**[0409]** In addition, a culture was set up by diluting the washed cells at 1 to 10 into modified TM3 medium contained in test tubes with palm oil. A further tube was set up by the addition of 0.125% C12FA to the remainder (2.5 mL) of the washed cells without further dilution (bioconversion). After 3.75 hours of growth at 30°C with shaking at 250 rpm all of the cultures were induced by the addition of 50 uM IPTG. Incubation was continued for 4 hours after which time 200 uL of each of the cultures was assayed for isoprene accumulation with a modified head space assay (1 hour accumulation at 30°C with shaking at 500 rpm). An additional isoprene assay was conducted by a 12 hour incubation of the assay glass block prior to GCMS analysis. Incubation of the induced cultures was continued overnight and 200 uL aliquots were again assayed for isoprene production (1 hour, 30 deg, 500 rpm Shel-Lab shaker) the following morning. Analysis of these cultures showed the production of significant levels of isoprene. The highest levels of isoprene were observed in the culture which was seeded at 1/10 dilution from the overnight starter culture after it had been incubated and induced overnight. This result suggests that this culture continued to grow and increase in cell density. These results are shown in Figure 70B. Cell density could not be measured directly because the fatty acid suspension had a turbid appearance. Cell density of this culture was therefore determined by plating an aliquot of the culture and showed  $8 \times 10^7$  colony forming units. This corresponds approximately to an OD<sub>600</sub> of 0.1. Nevertheless, this culture provided significant isoprene production; no isoprene is observed for similar strains without the pathway described in this example.

**Example 13: Improvement of isoprene production by constitutive expression of *ybhE* in *E. coli*.**

**[0410]** This example shows production of isoprene in a strain constitutively expressing *ybhE* (pgl) compared to a control strain with wild type *ybhE*. The gene *ybhE* (pgl) encodes a 6-phosphogluconolactonase that suppresses posttranslational gluconylation of heterologously expressed proteins and improves product solubility and yield while also improving biomass



yield and flux through the pentose phosphate pathway (Aon *et al. Applied and Environmental Microbiology*, 74(4): 950-958, 2008).

**[0411]** The BL21 strain of *E. coli* producing isoprene (EWL256) was constructed with constitutive expression of the *ybhE* gene on a replicating plasmid pBBR1MCS5(Gentamycin) (obtained from Dr. K. Peterson, Louisiana State University).

**[0412]** FRT-based recombination cassettes, and plasmids for Red/ET-mediated integration and antibiotic marker loopout were obtained from Gene Bridges GmbH (Germany). Procedures using these materials were carried out according to Gene Bridges protocols. Primers Pgl-F and PglGI1.5-R were used to amplify the resistance cassette from the FRT-gb2-Cm-FRT template using Stratagene Herculase II Fusion kit according to the manufacturer's protocol. The PCR reaction (50 uL final volume) contained: 5 uL buffer, 1 uL template DNA (FRT-gb2-Cm-F from Gene Bridges), 10 pmols of each primer, and 1.5 uL 25mM dNTP mix, made to 50 uL with dH<sub>2</sub>O. The reaction was cycled as follows: 1 x 2 minutes, 95°C then 30 cycles of (30 seconds at 95°C; 30 seconds at 63°C; 3 minutes at 72°C).

**[0413]** The resulting PCR product was purified using the QiaQuick PCR purification kit (Qiagen) and electroporated into electrocompetent MG1655 cells harboring the pRed-ET recombinase-containing plasmid as follows. Cells were prepared by growing in 5 mLs of L broth to and OD<sub>600</sub>~0.6 at 30°C. The cells were induced for recombinase expression by the addition of 4% arabinose and allowed to grow for 30 minutes at 30°C followed by 30 minutes of growth at 37°C. An aliquot of 1.5 mLs of the cells was washed 3-4 times in ice cold dH<sub>2</sub>O. The final cell pellet was resuspended in 40 uL of ice cold dH<sub>2</sub>O and 2-5 uL of the PCR product was added. The electroporation was carried out in 1-mm gap cuvettes, at 1.3 kV in a Gene Pulser Electroporator (Bio-Rad Inc.). Cells were recovered for 1-2 hours at 30°C and plated on L agar containing chloramphenicol (5 ug/mL). Five transformants were analyzed by PCR and sequencing using primers flanking the integration site (2 primer sets: pgl and 49 rev and 3' EcoRV-pglstop; Bottom Pgb2 and Top GB's CMP (946)). A correct transformant was selected and this strain was designated MG1655 GI1.5-pgl::CMP.

**[0414]** The chromosomal DNA of MG1655 GI1.5-pgl::CMP was used as template to generate a PCR fragment containing the FRT-CMP-FRT-GI1.5 - *ybhE* construct. This construct was



cloned into pBBR1MCS5(Gentamycin) as follows. The fragment, here on referred to as CMP-GI1.5-pgl, was amplified using the 5' primer Pglconfirm-F and 3' primer 3' EcoRV-pglstop. The resulting fragment was cloned using the Invitrogen TOPO-Blunt cloning kit into the plasmid vector pCR-Blunt II-TOPO as suggested from the manufacturer. The *NsiI* fragment harboring the CMP-GI1.5-pgl fragment was cloned into the *PstI* site of pBBR1MCS5(Gentamycin). A 20µl ligation reaction was prepared containing 5µl CMP-GI1.5-pgl insert, 2µl pBBR1MCS5(Gentamycin) vector, 1µl T4 DNA ligase (New England Biolabs), 2µl 10X ligase buffer, and 10µl ddH<sub>2</sub>O. The ligation mixture was incubated at room temperature for 40 minutes then 2-4 uL were electroporated into electrocompetent Top10 cells (Invitrogen) using the parameters disclosed above. Transformants were selected on L agar containing 10 ug/ml chloramphenicol and 5 ug/ml Gentamycin. The sequence of the selected clone was determined using a number of the primers described above as well as with the in-house T3 and Reverse primers provided by Sequetech, CA. This plasmid was designated pBBRCMPGI1.5-pgl (Figures 77A-B and SEQ ID NO:122).

**[0415]** Plasmid pBBRCMPGI1.5-pgl was electroporated into EWL256, as described above in Example 10 and transformants were plated on L agar containing Chloramphenicol (10 ug/mL), Gentamycin (5 ug/mL), spectinomycin (50 ug/mL), and carbenicillin (50 ug/mL). One transformant was selected and designated RM11608-2.

Primers:

Pgl-F

5'-  
ACCGCCAAAAGCGACTAATTTTAGCTGTTACAGTCAGTTGAATTAACCCTCACTAAA  
GGGCGGCCGC-3' (SEQ ID NO:115)

PglGI1.5-R

5'-  
GCTGGCGATATAAACTGTTTGCTTCATGAATGCTCCTTTGGGTTACCTCCGGGAAAC  
GCGGTTGATTTGTTTAGTGGTTGAATTATTTGCTCAGGATGTGGCATAGTCAAGGGC  
GTGACGGCTCGCTAATACGACTCACTATAGGGCTCGAG-3' (SEQ ID NO:116)

3' EcoRV-pglstop:

5'-CTT GAT ATC TTA GTG TGC GTT AAC CAC CAC (SEQ ID NO:117)



pgl +49 rev: CGTGAATTTGCTGGCTCTCAG (SEQ ID NO:118)

Bottom Pgb2: GGTTTAGTTCCTCACCTTGTC (SEQ ID NO:119)

Top GB's CMP (946): ACTGAAACGTTTTTCATCGCTC (SEQ ID NO:120)

Pglconfirm-F

5'-ACCGCCAAAAGCGACTAATTTTAGCT-3' (SEQ ID NO:121)

i) Small scale analysis

**[0416]** Media Recipe (per liter fermentation media):

**[0417]** K<sub>2</sub>HPO<sub>4</sub> 13.6 g, KH<sub>2</sub>PO<sub>4</sub> 13.6 g, MgSO<sub>4</sub> \* 7H<sub>2</sub>O 2 g, citric acid monohydrate 2 g, ferric ammonium citrate 0.3 g, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 3.2 g, yeast extract 1 g, 1000X Trace Metals Solution 1 ml. All of the components were added together and dissolved in diH<sub>2</sub>O. The pH was adjusted to 6.8 with ammonium hydroxide (30%) and brought to volume. Media was filter-sterilized with a 0.22 micron filter. Glucose 5.0 g and antibiotics were added after sterilization and pH adjustment.

**[0418]** 1000X Trace Metal Solution (per liter fermentation media):

**[0419]** Citric Acid \* H<sub>2</sub>O 40g, MnSO<sub>4</sub> \* H<sub>2</sub>O 30g, NaCl 10g, FeSO<sub>4</sub> \* 7H<sub>2</sub>O 1g, CoCl<sub>2</sub> \* 6H<sub>2</sub>O 1g, ZnSO<sub>4</sub> \* 7H<sub>2</sub>O 1g, CuSO<sub>4</sub> \* 5H<sub>2</sub>O 100mg, H<sub>3</sub>BO<sub>3</sub> 100mg, NaMoO<sub>4</sub> \* 2H<sub>2</sub>O 100mg. Each component is dissolved one at a time in diH<sub>2</sub>O. The pH is adjusted to 3.0 with HCl/NaOH, and then the solution is brought to volume and filter-sterilized with a 0.22 micron filter.

a) Experimental procedure

**[0420]** Isoprene production was analyzed by growing the strains in a Cellerator<sup>TM</sup> from MicroReactor Technologies, Inc. The working volume in each of the 24 wells was 4.5 mL. The temperature was maintained at 30°C, the pH setpoint was 7.0, the oxygen flow setpoint was 20 sccm and the agitation rate was 800 rpm. An inoculum of *E. coli* strain taken from a frozen vial was streaked onto an LB broth agar plate (with antibiotics) and incubated at 30°C. A single



colony was inoculated into media with antibiotics and grown overnight. The bacteria were diluted into 4.5 mL of media with antibiotics to reach an optical density of 0.05 measured at 550 nm.

[0421] Off-gas analysis of isoprene was performed using a gas chromatograph-mass spectrometer (GC-MS) (Agilent) headspace assay. Sample preparation was as follows: 100  $\mu$ L of whole broth was placed in a sealed GC vial and incubated at 30°C for a fixed time of 30 minutes. Following a heat kill step, consisting of incubation at 70°C for 5 minutes, the sample was loaded on the GC.

[0422] Optical density (OD) at a wavelength of 550 nm was obtained using a microplate reader (Spectramax) during the course of the run. Specific productivity was obtained by dividing the isoprene concentration ( $\mu$ g/L) by the OD reading and the time (hour).

[0423] The two strains EWL256 and RM11608-2 were assessed at 200 and 400  $\mu$ M IPTG induction levels. Samples were analyzed for isoprene production and cell growth (OD550) at 1, 2.5, 4.75, and 8 hours post-induction. Samples were done in duplicate.

#### b) Results

[0424] The experiment demonstrated that at 2 different concentrations of IPTG the strain expressing the *ybhE* (*pgl*) had a dramatic 2-3 fold increase in specific productivity of isoprene compared to the control strain.

ii) Isoprene fermentation from *E. coli* expressing *M. mazei* mevalonate kinase, *P. alba* isoprene synthase, and *pgl* over-expression (RHM111608-2) and grown in fed-batch culture at the 15-L scale

[0425] Medium Recipe (per liter fermentation medium)

[0426]  $K_2HPO_4$  7.5 g,  $MgSO_4 \cdot 7H_2O$  2 g, citric acid monohydrate 2 g, ferric ammonium citrate 0.3 g, yeast extract 0.5 g, 1000X Modified Trace Metal Solution 1 ml. All of the components were added together and dissolved in  $diH_2O$ . This solution was autoclaved. The pH was adjusted to 7.0 with ammonium hydroxide (30%) and q.s. to volume. Glucose 10 g, thiamine \* HCl 0.1 g, and antibiotics were added after sterilization and pH adjustment.



[0427] 1000X Modified trace Metal Solution:

[0428] Citric Acids \* H<sub>2</sub>O 40 g, MnSO<sub>4</sub> \* H<sub>2</sub>O 30 g, NaCl 10 g, FeSO<sub>4</sub> \* 7H<sub>2</sub>O 1 g, CoCl<sub>2</sub> \* 6H<sub>2</sub>O 1 g, ZnSO<sub>4</sub> \* 7H<sub>2</sub>O 1 g, CuSO<sub>4</sub> \* 5H<sub>2</sub>O 100 mg, H<sub>3</sub>BO<sub>3</sub> 100 mg, NaMoO<sub>4</sub> \* 2H<sub>2</sub>O 100 mg. Each component is dissolved one at a time in Di H<sub>2</sub>O, pH to 3.0 with HCl/NaOH, then q.s. to volume and filter sterilized with a 0.22 micron filter

[0429] Fermentation was performed in a 15-L bioreactor with BL21 (DE3) *E. coli* cells containing the upper mevalonic acid (MVA) pathway (pCL Upper), the integrated lower MVA pathway (gi1.2KKDyI), high expression of mevalonate kinase from *M. mazei* and isoprene synthase from *P. alba* (pTrcAlba-mMVK), and high expression of *pgl* (pBBR-pgl). This experiment was carried out to monitor isoprene formation from glucose at the desired fermentation pH 7.0 and temperature 34°C. A frozen vial of the *E. coli* strain was thawed and inoculated into tryptone-yeast extract medium. After the inoculum grew to OD 1.0, measured at 550 nm, 500 mL was used to inoculate a 15-L bioreactor bringing the initial volume to 5-L.

[0430] Glucose was fed at an exponential rate until cells reached the stationary phase. After this time the glucose feed was decreased to meet metabolic demands. The total amount of glucose delivered to the bioreactor during the 40 hour (59 hour) fermentation was 3.1 kg (4.2 kg at 59 hour). Induction was achieved by adding IPTG. The IPTG concentration was brought to 110 uM when the optical density at 550 nm (OD<sub>550</sub>) reached a value of 4. The IPTG concentration was raised to 192 uM when OD<sub>550</sub> reached 150. The OD<sub>550</sub> profile within the bioreactor over time is shown in Figure 78A. The isoprene level in the off gas from the bioreactor was determined using a Hiden mass spectrometer. The isoprene titer increased over the course of the fermentation to a maximum value of 33.2 g/L at 40 hours (48.6 g/L at 59 hours) (Figure 78B). The isoprene titer increased over the course of the fermentation to a maximum value of 40.0 g/L at 40 hours (60.5 g/L at 59 hours) (Figure 78C). The total amount of isoprene produced during the 40-hour (59-hour) fermentation was 281.3 g (451.0 g at 59 hours) and the time course of production is shown in Figure 78D. The time course of volumetric productivity is shown in Figure 78E and shows that an average rate of 1.0 g/L/hr was maintained between 0 and 40 hours (1.4 g/L/hour between 19 and 59 hour). The metabolic activity profile, as measured by CER, is shown in Figure 78F. The molar yield of utilized carbon that went into producing



isoprene during fermentation was 19.6% at 40 hours (23.6% at 59 hours). The weight percent yield of isoprene from glucose was 8.9% at 40 hours (10.7% at 59 hours).

**[0431]** Unless defined otherwise, the meanings of all technical and scientific terms used herein are those commonly understood by one of skill in the art to which this invention belongs. Singleton, *et al.*, Dictionary of Microbiology and Molecular Biology, 2nd ed., John Wiley and Sons, New York (1994), and Hale & Marham, The Harper Collins Dictionary of Biology, Harper Perennial, N.Y. (1991) provide one of skill with a general dictionary of many of the terms used in this invention. It is to be understood that this invention is not limited to the particular methodology, protocols, and reagents described, as these may vary. One of skill in the art will also appreciate that any methods and materials similar or equivalent to those described herein can also be used to practice or test the invention.

**[0432]** The headings provided herein are not limitations of the various aspects or embodiments of the invention which can be had by reference to the specification as a whole.

**[0433]** For use herein, unless clearly indicated otherwise, use of the terms “a”, “an,” and the like refers to one or more.

**[0434]** Reference to “about” a value or parameter herein includes (and describes) embodiments that are directed to that value or parameter per se. For example, description referring to “about X” includes description of “X.” Numeric ranges are inclusive of the numbers defining the range.

**[0435]** It is understood that aspects and embodiments of the invention described herein include “comprising,” “consisting,” and “consisting essentially of” aspects and embodiments.



## Appendix 1

### Exemplary 1-deoxy-D-xylulose-5-phosphate synthase nucleic acids and polypeptides

ATH: AT3G21500(DXPS1) AT4G15560(CLA1) AT5G11380(DXPS3)

OSA: 4338768 4340090 4342614

CME: CMF089C

PFA: MAL13P1.186

TAN: TA20470

TPV: TP01\_0516

ECO: b0420(dxS)

ECJ: JW0410(dxS)

ECE: Z0523(dxS)

ECS: ECs0474

ECC: c0531(dxS)

ECI: UTI89\_C0443(dxS)

ECP: ECP\_0479

ECV: APECO1\_1590(dxS)

ECW: EcE24377A\_0451(dxS)

ECX: EcHS\_A0491

STY: STY0461(dxS)

STT: t2441(dxS)

SPT: SPA2301(dxS)

SEC: SC0463(dxS)

STM: STM0422(dxS)

YPE: YPO3177(dxS)

YPK: y1008(dxS)

YPM: YP\_0754(dxS)

YPA: YPA\_2671

YPN: YPN\_0911

YPP: YPDSF\_2812



YPS: YPTB0939(dxs)  
YPI: YpsIP31758\_3112(dxs)  
SFL: SF0357(dxs)  
SFX: S0365(dxs)  
SFV: SFV\_0385(dxs)  
SSN: SSON\_0397(dxs)  
SBO: SBO\_0314(dxs)  
SDY: SDY\_0310(dxs)  
ECA: ECA1131(dxs)  
PLU: plu3887(dxs)  
BUC: BU464(dxs)  
BAS: BUsg448(dxs)  
WBR: WGLp144(dxs)  
SGL: SG0656  
KPN: KPN\_00372(dxs)  
BFL: Bfl238(dxs)  
BPN: BPEN\_244(dxs)  
HIN: HI1439(dxs)  
HIT: NTHI1691(dxs)  
HIP: CGSHiEE\_04795  
HIQ: CGSHiGG\_01080  
HDU: HD0441(dxs)  
HSO: HS\_0905(dxs)  
PMU: PM0532(dxs)  
MSU: MS1059(dxs)  
APL: APL\_0207(dxs)  
XFA: XF2249  
XFT: PD1293(dxs)  
XCC: XCC2434(dxs)  
XCB: XC\_1678  
XCV: XCV2764(dxs)  
XAC: XAC2565(dxs)



XOO: XOO2017(dxS)  
XOM: XOO\_1900(XOO1900)  
VCH: VC0889  
VVU: VV1\_0315  
VVY: VV0868  
VPA: VP0686  
VFI: VF0711  
PPR: PBPRA0805  
PAE: PA4044(dxS)  
PAU: PA14\_11550(dxS)  
PAP: PSPA7\_1057(dxS)  
PPU: PP\_0527(dxS)  
PST: PSPTO\_0698(dxS)  
PSB: Psyr\_0604  
PSP: PSPPH\_0599(dxS)  
PFL: PFL\_5510(dxS)  
PFO: Pfl\_5007  
PEN: PSEEN0600(dxS)  
PMY: Pmen\_3844  
PAR: Psyc\_0221(dxS)  
PCR: Pcryo\_0245  
ACI: ACIAD3247(dxS)  
SON: SO\_1525(dxS)  
SDN: Sden\_2571  
SFR: Sfri\_2790  
SAZ: Sama\_2436  
SBL: Sbal\_1357  
SLO: Shew\_2771  
SHE: Shewmr4\_2731  
SHM: Shewmr7\_2804  
SHN: Shewana3\_2901  
SHW: Sputw3181\_2831



ILO: IL2138(dxS)  
CPS: CPS\_1088(dxS)  
PHA: PSHAa2366(dxS)  
PAT: Patl\_1319  
SDE: Sde\_3381  
PIN: Ping\_2240  
MAQ: Maqu\_2438  
MCA: MCA0817(dxS)  
FTU: FTT1018c(dxS)  
FTF: FTF1018c(dxS)  
FTW: FTW\_0925(dxS)  
FTL: FTL\_1072  
FTH: FTH\_1047(dxS)  
FTA: FTA\_1131(dxS)  
FTN: FTN\_0896(dxS)  
NOC: Noc\_1743  
AEH: Mlg\_1381  
HCH: HCH\_05866(dxS)  
CSA: Csal\_0099  
ABO: ABO\_2166(dxS)  
AHA: AHA\_3321(dxS)  
BCI: BCI\_0275(dxS)  
RMA: Rmag\_0386  
VOK: COSY\_0360(dxS)  
NME: NMB1867  
NMA: NMA0589(dxS)  
NMC: NMC0352(dxS)  
NGO: NGO0036  
CVI: CV\_2692(dxS)  
RSO: RSc2221(dxS)  
REU: Reut\_A0882  
REH: H16\_A2732(dxS)



RME: Rmet\_2615  
BMA: BMAA0330(dxs)  
BMV: BMASAVP1\_1512(dxs)  
BML: BMA10299\_1706(dxs)  
BMN: BMA10247\_A0364(dxs)  
BXE: Bxe\_B2827  
BUR: Bcep18194\_B2211  
BCN: Bcen\_4486  
BCH: Bcen2424\_3879  
BAM: Bamb\_3250  
BPS: BPSS1762(dxs)  
BPM: BURPS1710b\_A0842(dxs)  
BPL: BURPS1106A\_A2392(dxs)  
BPD: BURPS668\_A2534(dxs)  
BTE: BTH\_II0614(dxs)  
BPE: BP2798(dxs)  
BPA: BPP2464(dxs)  
BBR: BB1912(dxs)  
RFR: Rfer\_2875  
POL: Bpro\_1747  
PNA: Pnap\_1501  
AJS: Ajs\_1038  
MPT: Mpe\_A2631  
HAR: HEAR0279(dxs)  
MMS: mma\_0331  
NEU: NE1161(dxs)  
NET: Neut\_1501  
NMU: Nmul\_A0236  
EBA: ebA4439(dxs)  
AZO: azo1198(dxs)  
DAR: Daro\_3061  
TBD: Tbd\_0879



MFA: Mfla\_2133  
HPY: HP0354(dxS)  
HPJ: jhp0328(dxS)  
HPA: HPAG1\_0349  
HHE: HH0608(dxS)  
HAC: Hac\_0968(dxS)  
WSU: WS1996  
TDN: Tmden\_0475  
CJE: Cj0321(dxS)  
CJR: CJE0366(dxS)  
CJJ: CJJ81176\_0343(dxS)  
CJU: C8J\_0298(dxS)  
CJD: JJD26997\_1642(dxS)  
CFF: CFF8240\_0264(dxS)  
CCV: CCV52592\_1671(dxS) CCV52592\_1722  
CHA: CHAB381\_1297(dxS)  
CCO: CCC13826\_1594(dxS)  
ABU: Abu\_2139(dxS)  
NIS: NIS\_0391(dxS)  
SUN: SUN\_2055(dxS)  
GSU: GSU0686(dxS-1) GSU1764(dxS-2)  
GME: Gmet\_1934 Gmet\_2822  
PCA: Pcar\_1667  
PPD: Ppro\_1191 Ppro\_2403  
DVU: DVU1350(dxS)  
DVL: Dvul\_1718  
DDE: Dde\_2200  
LIP: LI0408(dsX)  
DPS: DP2700  
ADE: Adeh\_1097  
MXA: MXAN\_4643(dxS)  
SAT: SYN\_02456



SFU: Sfum\_1418  
PUB: SAR11\_0611(dxs)  
MLO: mlr7474  
MES: Meso\_0735  
SME: SMc00972(dxs)  
ATU: Atu0745(dxs)  
ATC: AGR\_C\_1351  
RET: RHE\_CH00913(dxs)  
RLE: RL0973(dxs)  
BME: BMEI1498  
BMF: BAB1\_0462(dxs)  
BMS: BR0436(dxs)  
BMB: BruAb1\_0458(dxs)  
BOV: BOV\_0443(dxs)  
BJA: bl12651(dxs)  
BRA: BRADO2161(dxs)  
BBT: BBta\_2479(dxs)  
RPA: RPA0952(dxs)  
RPB: RPB\_4460  
RPC: RPC\_1149  
RPD: RPD\_4305  
RPE: RPE\_1067  
NWI: Nwi\_0633  
NHA: Nham\_0778  
BHE: BH04350(dxs)  
BQU: BQ03540(dxs)  
BBK: BARBAKC583\_0400(dxs)  
CCR: CC\_2068  
SIL: SPO0247(dxs)  
SIT: TM1040\_2920  
RSP: RSP\_0254(dxsA) RSP\_1134(dxs)  
JAN: Jann\_0088 Jann\_0170



RDE: RD1\_0101(dxs) RD1\_0548(dxs)  
MMR: Mmar10\_0849  
HNE: HNE\_1838(dxs)  
ZMO: ZMO1234(dxs) ZMO1598(dxs)  
NAR: Saro\_0161  
SAL: Sala\_2354  
ELI: ELI\_12520  
GOX: GOX0252  
GBE: GbCGDNIH1\_0221 GbCGDNIH1\_2404  
RRU: Rru\_A0054 Rru\_A2619  
MAG: amb2904  
MGM: Mmc1\_1048  
SUS: Acid\_1783  
BSU: BG11715(dxs)  
BHA: BH2779  
BAN: BA4400(dxs)  
BAR: GBAA4400(dxs)  
BAA: BA\_4853  
BAT: BAS4081  
BCE: BC4176(dxs)  
BCA: BCE\_4249(dxs)  
BCZ: BCZK3930(dxs)  
BTK: BT9727\_3919(dxs)  
BTL: BALH\_3785(dxs)  
BLI: BL01523(dxs)  
BLD: BLi02598(dxs)  
BCL: ABC2462(dxs)  
BAY: RBAM\_022600  
BPU: BPUM\_2159  
GKA: GK2392  
GTN: GTNG\_2322  
LMO: lmo1365(tktB)



LMF: LMOF2365\_1382(dxS)  
LIN: lin1402(tktB)  
LWE: lwe1380(tktB)  
LLA: L108911(dxSA) L123365(dxSB)  
LLC: LACR\_1572 LACR\_1843  
LLM: llmg\_0749(dxSB)  
SAK: SAK\_0263  
LPL: lp\_2610(dxS)  
LJO: LJ0406  
LAC: LBA0356  
LSL: LSL\_0209(dxS)  
LGA: LGAS\_0350  
STH: STH1842  
CAC: CAC2077 CA\_P0106(dxS)  
CPE: CPE1819  
CPF: CPF\_2073(dxS)  
CPR: CPR\_1787(dxS)  
CTC: CTC01575  
CNO: NT01CX\_1983  
CTH: Cthe\_0828  
CDF: CD1207(dxS)  
CBO: CBO1881(dxS)  
CBA: CLB\_1818(dxS)  
CBH: CLC\_1825(dxS)  
CBF: CLI\_1945(dxS)  
CKL: CKL\_1231(dxS)  
CHY: CHY\_1985(dxS)  
DSY: DSY2348  
DRM: Dred\_1078  
PTH: PTH\_1196(dxS)  
SWO: Swol\_0582  
CSC: Csac\_1853



TTE: TTE1298(dxs)  
MTA: Moth\_1511  
MPE: MYPE730  
MGA: MGA\_1268(dxs)  
MTU: Rv2682c(dxs1) Rv3379c(dxs2)  
MTC: MT2756(dxs)  
MBO: Mb2701c(dxs1) Mb3413c(dxs2)  
MLE: ML1038(dxs)  
MPA: MAP2803c(dxs)  
MAV: MAV\_3577(dxs)  
MSM: MSMEG\_2776(dxs)  
MMC: Mmcs\_2208  
CGL: NCgl1827(cgl1902)  
CGB: cg2083(dxs)  
CEF: CE1796  
CDI: DIP1397(dxs)  
CJK: jk1078(dxs)  
NFA: nfa37410(dxs)  
RHA: RHA1\_ro06843  
SCO: SCO6013(SC1C3.01) SCO6768(SC6A5.17)  
SMA: SAV1646(dxs1) SAV2244(dxs2)  
TWH: TWT484  
TWS: TW280(Dxs)  
LXX: Lxx10450(dxs)  
CMI: CMM\_1660(dxsA)  
AAU: AAur\_1790(dxs)  
PAC: PPA1062  
TFU: Tfu\_1917  
FRA: Francci3\_1326  
FAL: FRAAL2088(dxs)  
ACE: Acel\_1393  
SEN: SACE\_1815(dxs) SACE\_4351



BLO: BL1132(dxS)  
BAD: BAD\_0513(dxS)  
FNU: FN1208 FN1464  
RBA: RB2143(dxS)  
CTR: CT331(dxS)  
CTA: CTA\_0359(dxS)  
CMU: TC0608  
CPN: CPn1060(tktB\_2)  
CPA: CP0790  
CPJ: CPj1060(tktB\_2)  
CPT: CpB1102  
CCA: CCA00304(dxS)  
CAB: CAB301(dxS)  
CFE: CF0699(dxS)  
PCU: pc0619(dxS)  
TPA: TP0824  
TDE: TDE1910(dxS)  
LIL: LA3285(dxS)  
LIC: LIC10863(dxS)  
LBJ: LBJ\_0917(dxS)  
LBL: LBL\_0932(dxS)  
SYN: sll1945(dxS)  
SYW: SYNW1292(Dxs)  
SYC: syc1087\_c(dxS)  
SYF: Synpcc7942\_0430  
SYD: Syncc9605\_1430  
SYE: Syncc9902\_1069  
SYG: sync\_1410(dxS)  
SYR: SynRCC307\_1390(dxS)  
SYX: SynWH7803\_1223(dxS)  
CYA: CYA\_1701(dxS)  
CYB: CYB\_1983(dxS)



TEL: tll0623  
GVI: gll0194  
ANA: alr0599  
AVA: Ava\_4532  
PMA: Pro0928(dxS)  
PMM: PMM0907(Dxs)  
PMT: PMT0685(dxS)  
PMN: PMN2A\_0300  
PMI: PMT9312\_0893  
PMB: A9601\_09541(dxS)  
PMC: P9515\_09901(dxS)  
PMF: P9303\_15371(dxS)  
PMG: P9301\_09521(dxS)  
PMH: P9215\_09851  
PMJ: P9211\_08521  
PME: NATL1\_09721(dxS)  
TER: Tery\_3042  
BTH: BT\_1403 BT\_4099  
BFR: BF0873 BF4306  
BFS: BF0796(dxS) BF4114  
PGI: PG2217(dxS)  
CHU: CHU\_3643(dxS)  
GFO: GFO\_3470(dxS)  
FPS: FP0279(dxS)  
CTE: CT0337(dxS)  
CPH: Cpha266\_0671  
PVI: Cvib\_0498  
PLT: Plut\_0450  
DET: DET0745(dxS)  
DEH: cbdb\_A720(dxS)  
DRA: DR\_1475  
DGE: Dgeo\_0994



TTH: TTC1614

TTJ: TTHA0006

AAE: aq\_881

TMA: TM1770

PMO: Pmob\_1001



**Exemplary acetyl-CoA-acetyltransferase nucleic acids and polypeptides**

HSA: 38(ACAT1) 39(ACAT2)

PTR: 451528(ACAT1)

MCC: 707653(ACAT1) 708750(ACAT2)

MMU: 110446(Acat1) 110460(Acat2)

RNO: 25014(Acat1)

CFA: 484063(ACAT2) 489421(ACAT1)

GGA: 418968(ACAT1) 421587(RCJMB04\_34i5)

XLA: 379569(MGC69098) 414622(MGC81403) 414639(MGC81256)

444457(MGC83664)

XTR: 394562(acat2)

DRE: 30643(acat2)

SPU: 759502(LOC759502)

DME: Dmel\_CG10932 Dmel\_CG9149

CEL: T02G5.4 T02G5.7 T02G5.8(kat-1)

ATH: AT5G48230(ACAT2/EMB1276)

OSA: 4326136 4346520

CME: CMA042C CME087C

SCE: YPL028W(ERG10)

AGO: AGOS\_ADR165C

PIC: PICST\_31707(ERG10)

CAL: CaO19.1591(erg10)

CGR: CAGL0L12364g

SPO: SPBC215.09c

MGR: MGG\_01755 MGG\_13499

ANI: AN1409.2

AFM: AFUA\_6G14200 AFUA\_8G04000

AOR: AO090103000012 AO090103000406

CNE: CNC05280

UMA: UM03571.1

DDI: DDB\_0231621



PFA: PF14\_0484  
TET: TTHERM\_00091590 TTHERM\_00277470 TTHERM\_00926980  
TCR: 511003.60  
ECO: b2224(atoB)  
ECJ: JW2218(atoB) JW5453(yqeF)  
ECE: Z4164(yqeF)  
ECS: ECs3701  
ECC: c2767(atoB) c3441(yqeF)  
ECI: UTI89\_C2506(atoB) UTI89\_C3247(yqeF)  
ECP: ECP\_2268 ECP\_2857  
ECV: APECO1\_3662(yqeF) APECO1\_4335(atoB) APECO1\_43352(atoB)  
ECX: EcHS\_A2365  
STY: STY3164(yqeF)  
STT: t2929(yqeF)  
SPT: SPA2886(yqeF)  
SEC: SC2958(yqeF)  
STM: STM3019(yqeF)  
SFL: SF2854(yqeF)  
SFX: S3052(yqeF)  
SFV: SFV\_2922(yqeF)  
SSN: SSON\_2283(atoB) SSON\_3004(yqeF)  
SBO: SBO\_2736(yqeF)  
ECA: ECA1282(atoB)  
ENT: Ent638\_3299  
SPE: Spro\_0592  
HIT: NTHI0932(atoB)  
XCC: XCC1297(atoB)  
XCB: XC\_2943  
XCV: XCV1401(thlA)  
XAC: XAC1348(atoB)  
XOO: XOO1881(atoB)  
XOM: XOO\_1778(XOO1778)



VCH: VCA0690  
VCO: VC0395\_0630  
VVU: VV2\_0494 VV2\_0741  
VVY: VVA1043 VVA1210  
VPA: VPA0620 VPA1123 VPA1204  
PPR: PBPRB1112 PBPRB1840  
PAE: PA2001(atoB) PA2553 PA3454 PA3589 PA3925  
PAU: PA14\_38630(atoB)  
PPU: PP\_2051(atoB) PP\_2215(fadAx) PP\_3754 PP\_4636  
PPF: Pput\_2009 Pput\_2403 Pput\_3523 Pput\_4498  
PST: PSPTO\_0957(phbA-1) PSPTO\_3164(phbA-2)  
PSB: Psyr\_0824 Psyr\_3031  
PSP: PSPPH\_0850(phbA1) PSPPH\_2209(phbA2)  
PFL: PFL\_1478(atoB-2) PFL\_2321 PFL\_3066 PFL\_4330(atoB-2) PFL\_5283  
PFO: Pfl\_1269 Pfl\_1739 Pfl\_2074 Pfl\_2868  
PEN: PSEEN3197 PSEEN3547(fadAx) PSEEN4635(phbA)  
PMY: Pmen\_1138 Pmen\_2036 Pmen\_3597 Pmen\_3662 Pmen\_3820  
PAR: Psyc\_0252 Psyc\_1169  
PCR: Pcryo\_0278 Pcryo\_1236 Pcryo\_1260  
PRW: PsycPRwf\_2011  
ACI: ACIAD0694 ACIAD1612 ACIAD2516(atoB)  
SON: SO\_1677(atoB)  
SDN: Sden\_1943  
SFR: Sfri\_1338 Sfri\_2063  
SAZ: Sama\_1375  
SBL: Sbal\_1495  
SBM: Shew185\_1489  
SBN: Sbal195\_1525  
SLO: Shew\_1667 Shew\_2858  
SPC: Sputcn32\_1397  
SSE: Ssed\_1473 Ssed\_3533  
SPL: Spea\_2783



SHE: Shewmr4\_2597  
SHM: Shewmr7\_2664  
SHN: Shewana3\_2771  
SHW: Sputw3181\_2704  
ILO: IL0872  
CPS: CPS\_1605 CPS\_2626  
PHA: PSHAa0908 PSHAa1454(atoB) PSHAa1586(atoB)  
PAT: Patl\_2923  
SDE: Sde\_3149  
PIN: Ping\_0659 Ping\_2401  
MAQ: Maqu\_2117 Maqu\_2489 Maqu\_2696 Maqu\_3162  
CBU: CBU\_0974  
LPN: lpg1825(atoB)  
LPF: lpl1789  
LPP: lpp1788  
NOC: Noc\_1891  
AEH: Mlg\_0688 Mlg\_2706  
HHA: Hhal\_1685  
HCH: HCH\_05299  
CSA: Csal\_0301 Csal\_3068  
ABO: ABO\_0648(fadAx)  
MMW: Mmwyl1\_0073 Mmwyl1\_3021 Mmwyl1\_3053 Mmwyl1\_3097 Mmwyl1\_4182  
AHA: AHA\_2143(atoB)  
CVI: CV\_2088(atoB) CV\_2790(phaA)  
RSO: RSc0276(atoB) RSc1632(phbA) RSc1637(bktB) RSc1761(RS02948)  
REU: Reut\_A0138 Reut\_A1348 Reut\_A1353 Reut\_B4561 Reut\_B4738  
Reut\_B5587 Reut\_C5943 Reut\_C6062  
REH: H16\_A0170 H16\_A0867 H16\_A0868 H16\_A0872 H16\_A1297  
H16\_A1438(phaA) H16\_A1445(bktB) H16\_A1528 H16\_A1713 H16\_A1720  
H16\_A1887 H16\_A2148 H16\_B0380 H16\_B0381 H16\_B0406 H16\_B0662  
H16\_B0668 H16\_B0759 H16\_B1369 H16\_B1771  
RME: Rmet\_0106 Rmet\_1357 Rmet\_1362 Rmet\_5156



BMA: BMA1316 BMA1321(phbA) BMA1436  
 BMV: BMASAVP1\_A1805(bktB) BMASAVP1\_A1810(phbA)  
 BML: BMA10299\_A0086(phbA) BMA10299\_A0091  
 BMN: BMA10247\_1076(bktB) BMA10247\_1081(phbA)  
 BXE: Bxe\_A2273 Bxe\_A2335 Bxe\_A2342 Bxe\_A4255 Bxe\_B0377 Bxe\_B0739  
       Bxe\_C0332 Bxe\_C0574 Bxe\_C0915  
 BVI: Bcep1808\_0519 Bcep1808\_1717 Bcep1808\_2877 Bcep1808\_3594  
       Bcep1808\_4015 Bcep1808\_5507 Bcep1808\_5644  
 BUR: Bcep18194\_A3629 Bcep18194\_A5080 Bcep18194\_A5091  
       Bcep18194\_A6102 Bcep18194\_B0263 Bcep18194\_B1439  
       Bcep18194\_C6652 Bcep18194\_C6802 Bcep18194\_C6874  
       Bcep18194\_C7118 Bcep18194\_C7151 Bcep18194\_C7332  
 BCN: Bcen\_1553 Bcen\_1599 Bcen\_2158 Bcen\_2563 Bcen\_2998 Bcen\_6289  
 BCH: Bcen2424\_0542 Bcen2424\_1790 Bcen2424\_2772 Bcen2424\_5368  
       Bcen2424\_6232 Bcen2424\_6276  
 BAM: Bamb\_0447 Bamb\_1728 Bamb\_2824 Bamb\_4717 Bamb\_5771 Bamb\_5969  
 BPS: BPSL1426 BPSL1535(phbA) BPSL1540  
 BPM: BURPS1710b\_2325(bktB) BURPS1710b\_2330(phbA)  
       BURPS1710b\_2453(atoB-2)  
 BPL: BURPS1106A\_2197(bktB) BURPS1106A\_2202(phbA)  
 BPD: BURPS668\_2160(bktB) BURPS668\_2165(phbA)  
 BTE: BTH\_I2144 BTH\_I2256 BTH\_I2261  
 PNU: Pnuc\_0927  
 BPE: BP0447 BP0668 BP2059  
 BPA: BPP0608 BPP1744 BPP3805 BPP4216 BPP4361  
 BBR: BB0614 BB3364 BB4250 BB4804 BB4947  
 RFR: Rfer\_0272 Rfer\_1000 Rfer\_1871 Rfer\_2273 Rfer\_2561 Rfer\_2594  
       Rfer\_3839  
 POL: Bpro\_1577 Bpro\_2140 Bpro\_3113 Bpro\_4187  
 PNA: Pnap\_0060 Pnap\_0458 Pnap\_0867 Pnap\_1159 Pnap\_2136 Pnap\_2804  
 AAV: Aave\_0031 Aave\_2478 Aave\_3944 Aave\_4368  
 AJS: Ajs\_0014 Ajs\_0124 Ajs\_1931 Ajs\_2073 Ajs\_2317 Ajs\_3548



Ajs\_3738 Ajs\_3776  
VEI: Veis\_1331 Veis\_3818 Veis\_4193  
DAC: Daci\_0025 Daci\_0192 Daci\_3601 Daci\_5988  
MPT: Mpe\_A1536 Mpe\_A1776 Mpe\_A1869 Mpe\_A3367  
HAR: HEAR0577(phbA)  
MMS: mma\_0555  
NEU: NE2262(bktB)  
NET: Neut\_0610  
EBA: ebA5202 p2A409(tioL)  
AZO: azo0464(fadA1) azo0469(fadA2) azo2172(thlA)  
DAR: Daro\_0098 Daro\_3022  
HPA: HPAG1\_0675  
HAC: Hac\_0958(atoB)  
GME: Gmet\_1719 Gmet\_2074 Gmet\_2213 Gmet\_2268 Gmet\_3302  
GUR: Gura\_3043  
BBA: Bd0404(atoB) Bd2095  
DOL: Dole\_0671 Dole\_1778 Dole\_2160 Dole\_2187  
ADE: Adeh\_0062 Adeh\_2365  
AFW: Anae109\_0064 Anae109\_1504  
MXA: MXAN\_3791  
SAT: SYN\_02642  
SFU: Sfum\_2280 Sfum\_3582  
RPR: RP737  
RCO: RC1134 RC1135  
RFE: RF\_0163(paaJ)  
RBE: RBE\_0139(paaJ)  
RAK: A1C\_05820  
RBO: A1I\_07215  
RCM: A1E\_04760  
PUB: SAR11\_0428(thlA)  
MLO: mlr3847  
MES: Meso\_3374



PLA: Plav\_1573 Plav\_2783  
SME: SMa1450 SMc03879(phbA)  
SMD: Smed\_0499 Smed\_3117 Smed\_5094 Smed\_5096  
ATU: Atu2769(atoB) Atu3475  
ATC: AGR\_C\_5022(phbA) AGR\_L\_2713  
RET: RHE\_CH04018(phbAch) RHE\_PC00068(ypc00040) RHE\_PF00014(phbAf)  
RLE: RL4621(phaA) pRL100301 pRL120369  
BME: BMEI0274 BMEII0817  
BMF: BAB1\_1783(phbA-1) BAB2\_0790(phbA-2)  
BMS: BR1772(phbA-1) BRA0448(phbA-2)  
BMB: BruAb1\_1756(phbA-1) BruAb2\_0774(phbA-2)  
BOV: BOV\_1707(phbA-1)  
OAN: Oant\_1130 Oant\_3107 Oant\_3718 Oant\_4020  
BJA: bll0226(atoB) bll3949 bll7400 bll7819 blr3724(phbA)  
BRA: BRADO0562(phbA) BRADO0983(pimB) BRADO3110 BRADO3134(atoB)  
BBT: BBta\_3558 BBta\_3575(atoB) BBta\_5147(pimB) BBta\_7072(pimB)  
BBta\_7614(phbA)  
RPA: RPA0513(pcaF) RPA0531 RPA3715(pimB)  
RPB: RPB\_0509 RPB\_0525 RPB\_1748  
RPC: RPC\_0504 RPC\_0636 RPC\_0641 RPC\_0832 RPC\_1050 RPC\_2005  
RPC\_2194 RPC\_2228  
RPD: RPD\_0306 RPD\_0320 RPD\_3105 RPD\_3306  
RPE: RPE\_0168 RPE\_0248 RPE\_3827  
NWI: Nwi\_3060  
XAU: Xaut\_3108 Xaut\_4665  
CCR: CC\_0510 CC\_0894 CC\_3462  
SIL: SPO0142(bktB) SPO0326(phbA) SPO0773 SPO3408  
SIT: TM1040\_0067 TM1040\_2790 TM1040\_3026 TM1040\_3735  
RSP: RSP\_0745 RSP\_1354 RSP\_3184  
RSH: Rsph17029\_0022 Rsph17029\_2401 Rsph17029\_3179 Rsph17029\_3921  
RSQ: Rsph17025\_0012 Rsph17025\_2466 Rsph17025\_2833  
JAN: Jann\_0262 Jann\_0493 Jann\_4050



RDE: RD1\_0025 RD1\_0201(bktB) RD1\_3394(phbA)  
PDE: Pden\_2026 Pden\_2663 Pden\_2870 Pden\_2907 Pden\_4811 Pden\_5022  
DSH: Dshi\_0074 Dshi\_3066 Dshi\_3331  
MMR: Mmar10\_0697  
HNE: HNE\_2706 HNE\_3065 HNE\_3133  
NAR: Saro\_0809 Saro\_1069 Saro\_1222 Saro\_2306 Saro\_2349  
SAL: Sala\_0781 Sala\_1244 Sala\_2896 Sala\_3158  
SWI: Swit\_0632 Swit\_0752 Swit\_2893 Swit\_3602 Swit\_4887 Swit\_5019  
Swit\_5309  
ELI: ELI\_01475 ELI\_06705 ELI\_12035  
GBE: GbCGDNIH1\_0447  
ACR: Acry\_1847 Acry\_2256  
RRU: Rru\_A0274 Rru\_A1380 Rru\_A1469 Rru\_A1946 Rru\_A3387  
MAG: amb0842  
MGM: Mmc1\_1165  
ABA: Acid345\_3239  
BSU: BG11319(mmgA) BG13063(yhfS)  
BHA: BH1997 BH2029 BH3801(mmgA)  
BAN: BA3687 BA4240 BA5589  
BAR: GBAA3687 GBAA4240 GBAA5589  
BAA: BA\_0445 BA\_4172 BA\_4700  
BAT: BAS3418 BAS3932 BAS5193  
BCE: BC3627 BC4023 BC5344  
BCA: BCE\_3646 BCE\_4076 BCE\_5475  
BCZ: BCZK3329(mmgA) BCZK3780(thl) BCZK5044(atoB)  
BCY: Bcer98\_2722 Bcer98\_3865  
BTK: BT9727\_3379(mmgA) BT9727\_3765(thl) BT9727\_5028(atoB)  
BTL: BALH\_3262(mmgA) BALH\_3642(fadA) BALH\_4843(atoB)  
BLI: BL03925(mmgA)  
BLD: BLi03968(mmgA)  
BCL: ABC0345 ABC2989 ABC3617 ABC3891(mmgA)  
BAY: RBAM\_022450



BPUM: BPUM\_2374(yhfS) BPUM\_2941 BPUM\_3373  
OIH: OB0676 OB0689 OB2632 OB3013  
GKA: GK1658 GK3397  
SAU: SA0342 SA0534(vraB)  
SAV: SAV0354 SAV0576(vraB)  
SAM: MW0330 MW0531(vraB)  
SAR: SAR0351(thl) SAR0581  
SAS: SAS0330 SAS0534  
SAC: SACOL0426 SACOL0622(atoB)  
SAB: SAB0304(thl) SAB0526  
SAA: SAUSA300\_0355 SAUSA300\_0560(vraB)  
SAO: SAOUHSC\_00336 SAOUHSC\_00558  
SAJ: SaurJH9\_0402  
SAH: SaurJH1\_0412  
SEP: SE0346 SE2384  
SER: SERP0032 SERP0220  
SHA: SH0510(mvaC) SH2417  
SSP: SSP0325 SSP2145  
LMO: lmo1414  
LMF: LMOF2365\_1433  
LIN: lin1453  
LWE: lwe1431  
LLA: L11745(thiL) L25946(fadA)  
LLC: LACR\_1665 LACR\_1956  
LLM: llmg\_0930(thiL)  
SPY: SPy\_0140 SPy\_1637(atoB)  
SPZ: M5005\_Spy\_0119 M5005\_Spy\_0432 M5005\_Spy\_1344(atoB)  
SPM: spyM18\_0136 spyM18\_1645(atoB)  
SPG: SpyM3\_0108 SpyM3\_1378(atoB)  
SPS: SPs0110 SPs0484  
SPH: MGAS10270\_Spy0121 MGAS10270\_Spy0433 MGAS10270\_Spy1461(atoB)  
SPI: MGAS10750\_Spy0124 MGAS10750\_Spy0452 MGAS10750\_Spy1453(atoB)



SPI: MGAS2096\_Spy0123 MGAS2096\_Spy0451 MGAS2096\_Spy1365(atoB)  
SPK: MGAS9429\_Spy0121 MGAS9429\_Spy0431 MGAS9429\_Spy1339(atoB)  
SPF: SpyM50447(atoB2)  
SPA: M6\_Spy0166 M6\_Spy0466 M6\_Spy1390  
SPB: M28\_Spy0117 M28\_Spy0420 M28\_Spy1385(atoB)  
SAK: SAK\_0568  
LJO: LJ1609  
LAC: LBA0626(thiL)  
LSA: LSA1486  
LDB: Ldb0879  
LBU: LBUL\_0804  
LBR: LVIS\_2218  
LCA: LSEI\_1787  
LGA: LGAS\_1374  
LRE: Lreu\_0052  
EFA: EF1364  
OOE: OEOE\_0529  
STH: STH2913 STH725 STH804  
CAC: CAC2873 CA\_P0078(thiL)  
CPE: CPE2195(atoB)  
CPF: CPF\_2460  
CPR: CPR\_2170  
CTC: CTC00312  
CNO: NT01CX\_0538 NT01CX\_0603  
CDF: CD1059(thlA1) CD2676(thlA2)  
CBO: CBO3200(thl)  
CBE: Cbei\_0411 Cbei\_3630  
CKL: CKL\_3696(thlA1) CKL\_3697(thlA2) CKL\_3698(thlA3)  
AMT: Amet\_4630  
AOE: Clos\_0084 Clos\_0258  
CHY: CHY\_1288 CHY\_1355(atoB) CHY\_1604 CHY\_1738  
DSY: DSY0632 DSY0639 DSY1567 DSY1710 DSY2402 DSY3302



DRM: Dred\_0400 Dred\_1491 Dred\_1784 Dred\_1892  
SWO: Swol\_0308 Swol\_0675 Swol\_0789 Swol\_1486 Swol\_1934 Swol\_2051  
TTE: TTE0549(paaJ)  
MTA: Moth\_1260  
MTU: Rv1135A Rv1323(fadA4) Rv3546(fadA5)  
MTC: MT1365(phbA)  
MBO: Mb1167 Mb1358(fadA4) Mb3576(fadA5) Mb3586c(fadA6)  
MBB: BCG\_1197 BCG\_1385(fadA4) BCG\_3610(fadA5) BCG\_3620c(fadA6)  
MLE: ML1158(fadA4)  
MPA: MAP2407c(fadA3) MAP2436c(fadA4)  
MAV: MAV\_1544 MAV\_1573 MAV\_1863 MAV\_5081  
MSM: MSMEG\_2224 MSMEG\_4920  
MUL: MUL\_0357  
MVA: Mvan\_1976 Mvan\_1988 Mvan\_4305 Mvan\_4677 Mvan\_4891  
MGI: Mflv\_1347 Mflv\_1484 Mflv\_2040 Mflv\_2340 Mflv\_4356 Mflv\_4368  
MMC: Mmcs\_1758 Mmcs\_1769 Mmcs\_3796 Mmcs\_3864  
MKM: Mkms\_0251 Mkms\_1540 Mkms\_1805 Mkms\_1816 Mkms\_2836 Mkms\_3159  
Mkms\_3286 Mkms\_3869 Mkms\_3938 Mkms\_4227 Mkms\_4411 Mkms\_4580  
Mkms\_4724 Mkms\_4764 Mkms\_4776  
MJL: Mjls\_0231 Mjls\_1739 Mjls\_1750 Mjls\_2819 Mjls\_3119 Mjls\_3235  
Mjls\_3800 Mjls\_3850 Mjls\_4110 Mjls\_4383 Mjls\_4705 Mjls\_4876  
Mjls\_5018 Mjls\_5063 Mjls\_5075  
CGL: NCgl2309(cgl2392)  
CGB: cg2625(pcaF)  
CEF: CE0731 CE2295  
CJK: jk1543(fadA3)  
NFA: nfa10750(fadA4)  
RHA: RHA1\_ro01455 RHA1\_ro01623 RHA1\_ro01876 RHA1\_ro02517(catF)  
RHA1\_ro03022 RHA1\_ro03024 RHA1\_ro03391 RHA1\_ro03892  
RHA1\_ro04599 RHA1\_ro05257 RHA1\_ro08871  
SCO: SCO5399(SC8F4.03)  
SMA: SAV1384(fadA5) SAV2856(fadA1)



ART: Arth\_1160 Arth\_2986 Arth\_3268 Arth\_4073  
NCA: Noca\_1371 Noca\_1797 Noca\_1828 Noca\_2764 Noca\_4142  
TFU: Tfu\_1520 Tfu\_2394  
FRA: Francci3\_3687  
FRE: Franean1\_1044 Franean1\_2711 Franean1\_2726 Franean1\_3929  
Franean1\_4037 Franean1\_4577  
FAL: FRAAL2514 FRAAL2618 FRAAL5910(atoB)  
ACE: Acel\_0626 Acel\_0672  
SEN: SACE\_1192(mmgA) SACE\_2736(fadA6) SACE\_4011(catF)  
SACE\_6236(fadA4)  
STP: Strop\_3610  
SAQ: Sare\_1316 Sare\_3991  
RXY: Rxyl\_1582 Rxyl\_1842 Rxyl\_2389 Rxyl\_2530  
FNU: FN0495  
BGA: BG0110(fadA)  
BAF: BAPKO\_0110(fadA)  
LIL: LA0457(thiL1) LA0828(thiL2) LA4139(fadA)  
LIC: LIC10396(phbA)  
LBJ: LBJ\_2862(paaJ-4)  
LBL: LBL\_0209(paaJ-4)  
SYN: slr1993(phaA)  
SRU: SRU\_1211(atoB) SRU\_1547  
CHU: CHU\_1910(atoB)  
GFO: GFO\_1507(atoB)  
FJO: Fjoh\_4612  
FPS: FP0770 FP1586 FP1725  
RRS: RoseRS\_3911 RoseRS\_4348  
RCA: Rcas\_0702 Rcas\_3206  
HAU: Haur\_0522  
DRA: DR\_1072 DR\_1428 DR\_1960 DR\_2480 DR\_A0053  
DGE: Dgeo\_0755 Dgeo\_1305 Dgeo\_1441 Dgeo\_1883  
TTH: TTC0191 TTC0330



TTJ: TTHA0559

TME: Tmel\_1134

FNO: Fnod\_0314

PMO: Pmob\_0515

HMA: rrnAC0896(acaB3) rrnAC2815(aca2) rrnAC3497(yqeF)  
rrnB0240(aca1) rrnB0242(acaB2) rrnB0309(acaB1)

TAC: Ta0582

TVO: TVN0649

PTO: PTO1505

APE: APE\_2108

SSO: SSO2377(acaB-4)

STO: ST0514

SAI: Saci\_0963 Saci\_1361(acaB1)

MSE: Msed\_0656

PAI: PAE1220

PIS: Pisl\_0029 Pisl\_1301

PCL: Pcal\_0781

PAS: Pars\_0309 Pars\_1071

CMA: Cmaq\_1941



**Exemplary HMG-CoA synthase nucleic acids and polypeptides**

HSA: 3157(HMGCS1) 3158(HMGCS2)  
PTR: 457169(HMGCS2) 461892(HMGCS1)  
MCC: 702553(HMGCS1) 713541(HMGCS2)  
MMU: 15360(Hmgcs2) 208715(Hmgcs1)  
RNO: 24450(Hmgcs2) 29637(Hmgcs1)  
CFA: 479344(HMGCS1) 607923(HMGCS2)  
BTA: 407767(HMGCS1)  
SSC: 397673(CH242-38B5.1)  
GGA: 396379(HMGCS1)  
XLA: 380091(hmgcs1) 447204(MGC80816)  
DRE: 394060(hmgcs1)  
SPU: 578259(LOC578259)  
DME: Dmel\_CG4311(Hmgs)  
CEL: F25B4.6  
ATH: AT4G11820(BAP1)  
OSA: 4331418 4347614  
CME: CMM189C  
SCE: YML126C(ERG13)  
AGO: AGOS\_ADL356C  
PIC: PICST\_83020  
CAL: CaO19\_7312(CaO19.7312)  
CGR: CAGL0H04081g  
SPO: SPAC4F8.14c(hcs)  
MGR: MGG\_01026  
ANI: AN4923.2  
AFM: AFUA\_3G10660 AFUA\_8G07210  
AOR: AO090003000611 AO090010000487  
CNE: CNC05080 CNG02670  
UMA: UM05362.1  
ECU: ECU10\_0510



DDI: DDBDRAFT\_0217522 DDB\_0219924(hgsA)  
TET: TTHERM\_00691190  
TBR: Tb927.8.6110  
YPE: YPO1457  
YPK: y2712(pksG)  
YPM: YP\_1349(pksG)  
YPA: YPA\_0750  
YPN: YPN\_2521  
YPP: YPDSF\_1517  
YPS: YPTB1475  
CBD: COXBU7E912\_1931  
TCX: Tcr\_1719  
DNO: DNO\_0799  
BMA: BMAA1212  
BPS: BPSS1002  
BPM: BURPS1710b\_A2613  
BPL: BURPS1106A\_A1384  
BPD: BURPS668\_A1470  
BTE: BTH\_II1670  
MXA: MXAN\_3948(tac) MXAN\_4267(mvaS)  
BSU: BG10926(pksG)  
OIH: OB2248  
SAU: SA2334(mvaS)  
SAV: SAV2546(mvaS)  
SAM: MW2467(mvaS)  
SAR: SAR2626(mvaS)  
SAS: SAS2432  
SAC: SACOL2561  
SAB: SAB2420(mvaS)  
SAA: SAUSA300\_2484  
SAO: SAOUHSC\_02860  
SAJ: SaurJH9\_2569



SAH: SaurJH1\_2622  
SEP: SE2110  
SER: SERP2122  
SHA: SH0508(mvaS)  
SSP: SSP0324  
LMO: lmo1415  
LMF: LMOF2365\_1434(mvaS)  
LIN: lin1454  
LWE: lwe1432(mvaS)  
LLA: L13187(hmcM)  
LLC: LACR\_1666  
LLM: llmg\_0929(hmcM)  
SPY: SPy\_0881(mvaS.2)  
SPZ: M5005\_Spy\_0687(mvaS.1)  
SPM: spyM18\_0942(mvaS2)  
SPG: SpyM3\_0600(mvaS.2)  
SPS: SPs1253  
SPH: MGAS10270\_Spy0745(mvaS1)  
SPI: MGAS10750\_Spy0779(mvaS1)  
SPJ: MGAS2096\_Spy0759(mvaS1)  
SPK: MGAS9429\_Spy0743(mvaS1)  
SPF: SpyM51121(mvaS)  
SPA: M6\_Spy0704  
SPB: M28\_Spy0667(mvaS.1)  
SPN: SP\_1727  
SPR: spr1571(mvaS)  
SPD: SPD\_1537(mvaS)  
SAG: SAG1316  
SAN: gbs1386  
SAK: SAK\_1347  
SMU: SMU.943c  
STC: str0577(mvaS)



STL: stu0577(mvaS)  
STE: STER\_0621  
SSA: SSA\_0338(mvaS)  
SSU: SSU05\_1641  
SSV: SSU98\_1652  
SGO: SGO\_0244  
LPL: lp\_2067(mvaS)  
LJO: LJ1607  
LAC: LBA0628(hmcS)  
LSA: LSA1484(mvaS)  
LSL: LSL\_0526  
LDB: Ldb0881(mvaS)  
LBU: LBUL\_0806  
LBR: LVIS\_1363  
LCA: LSEI\_1785  
LGA: LGAS\_1372  
LRE: Lreu\_0676  
PPE: PEPE\_0868  
EFA: EF1363  
OOE: OEOE\_0968  
LME: LEUM\_1184  
NFA: nfa22120  
SEN: SACE\_4570(pksG)  
BBU: BB0683  
BGA: BG0706  
BAF: BAPKO\_0727  
FJO: Fjoh\_0678  
HAL: VNG1615G(mvaB)  
HMA: rrnAC1740(mvaS)  
HWA: HQ2868A(mvaB)  
NPH: NP2608A(mvaB\_1) NP4836A(mvaB\_2)



**Exemplary hydroxymethylglutaryl-CoA reductase nucleic acids and polypeptides**

HSA: 3156(HMGCR)  
PTR: 471516(HMGCR)  
MCC: 705479(HMGCR)  
MMU: 15357(Hmgcr)  
RNO: 25675(Hmgcr)  
CFA: 479182(HMGCR)  
BTA: 407159(HMGCR)  
GGA: 395145(RCJMB04\_14m24)  
SPU: 373355(LOC373355)  
DME: Dmel\_CG10367(Hmgcr)  
CEL: F08F8.2  
OSA: 4347443  
SCE: YLR450W(HMG2) YML075C(HMG1)  
AGO: AGOS\_AER152W  
CGR: CAGL0L11506g  
SPO: SPCC162.09c(hmg1)  
ANI: AN3817.2  
AFM: AFUA\_1G11230 AFUA\_2G03700  
AOR: AO090103000311 AO090120000217  
CNE: CNF04830  
UMA: UM03014.1  
ECU: ECU10\_1720  
DDI: DDB\_0191125(hmgA) DDB\_0215357(hmgB)  
TBR: Tb927.6.4540  
TCR: 506831.40 509167.20  
LMA: LmjF30.3190  
VCH: VCA0723  
VCO: VC0395\_0662  
VVU: VV2\_0117  
VYV: VVA0625



VPA: VPA0968  
VFI: VFA0841  
PAT: Patl\_0427  
CBU: CBU\_0030 CBU\_0610  
CBD: COXBU7E912\_0151 COXBU7E912\_0622(hmgA)  
TCX: Tcr\_1717  
DNO: DNO\_0797  
CVI: CV\_1806  
SUS: Acid\_5728 Acid\_6132  
SAU: SA2333(mvaA)  
SAV: SAV2545(mvaA)  
SAM: MW2466(mvaA)  
SAB: SAB2419c(mvaA)  
SEP: SE2109  
LWE: lwe0819(mvaA)  
LLA: L10433(mvaA)  
LLC: LACR\_1664  
LLM: llmg\_0931(mvaA)  
SPY: SPy\_0880(mvaS.1)  
SPM: spyM18\_0941(mvaS1)  
SPG: SpyM3\_0599(mvaS.1)  
SPS: SPs1254  
SPH: MGAS10270\_Spy0744  
SPI: MGAS10750\_Spy0778  
SPJ: MGAS2096\_Spy0758  
SPK: MGAS9429\_Spy0742  
SPA: M6\_Spy0703  
SPN: SP\_1726  
SAG: SAG1317  
SAN: gbs1387  
STC: str0576(mvaA)  
STL: stu0576(mvaA)



STE: STER\_0620  
SSA: SSA\_0337(mvaA)  
LPL: lp\_0447(mvaA)  
LJO: LJ1608  
LSL: LSL\_0224  
LBR: LVIS\_0450  
LGA: LGAS\_1373  
EFA: EF1364  
NFA: nfa22110  
BGA: BG0708(mvaA)  
SRU: SRU\_2422  
FPS: FP2341  
MMP: MMP0087(hmgA)  
MMQ: MmarC5\_1589  
MAC: MA3073(hmgA)  
MBA: Mbar\_A1972  
MMA: MM\_0335  
MBU: Mbur\_1098  
MHU: Mhun\_3004  
MEM: Memar\_2365  
MBN: Mboo\_0137  
MTH: MTH562  
MST: Msp\_0584(hmgA)  
MSI: Msm\_0227  
MKA: MK0355(HMG1)  
AFU: AF1736(mvaA)  
HAL: VNG1875G(mvaA)  
HMA: rrnAC3412(mvaA)  
HWA: HQ3215A(hmgR)  
NPH: NP0368A(mvaA\_2) NP2422A(mvaA\_1)  
TAC: Ta0406m  
TVO: TVN1168



PTO: PTO1143  
PAB: PAB2106(mvaA)  
PFU: PF1848  
TKO: TK0914  
RCI: RCIX1027(hmgA) RCIX376(hmgA)  
APE: APE\_1869  
IHO: Igni\_0476  
HBU: Hbut\_1531  
SSO: SSO0531  
STO: ST1352  
SAI: Saci\_1359  
PAI: PAE2182  
PIS: Pisl\_0814  
PCL: Pcal\_1085  
PAS: Pars\_0796



**Exemplary mevalonate kinase nucleic acids and polypeptides**

HSA: 4598(MVK)  
MCC: 707645(MVK)  
MMU: 17855(Mvk)  
RNO: 81727(Mvk)  
CFA: 486309(MVK)  
BTA: 505792(MVK)  
GGA: 768555(MVK)  
DRE: 492477(zgc:103473)  
SPU: 585785(LOC585785)  
DME: Dmel\_CG33671  
OSA: 4348331  
SCE: YMR208W(ERG12)  
AGO: AGOS\_AER335W  
PIC: PICST\_40742(ERG12)  
CGR: CAGL0F03861g  
SPO: SPAC13G6.11c  
MGR: MGG\_06946  
ANI: AN3869.2  
AFM: AFUA\_4G07780  
AOR: AO090023000793  
CNE: CNK01740  
ECU: ECU09\_1780  
DDI: DDBDRAFT\_0168621  
TET: TTHERM\_00637680  
TBR: Tb927.4.4070  
TCR: 436521.9 509237.10  
LMA: LmjF31.0560  
CBU: CBU\_0608 CBU\_0609  
CBD: COXBU7E912\_0620(mvk)  
LPN: lpg2039



LPF: lpl2017  
LPP: lpp2022  
BBA: Bd1027(lmbP) Bd1630(mvk)  
MXA: MXAN\_5019(mvk)  
OIH: OB0225  
SAU: SA0547(mvaK1)  
SAV: SAV0590(mvaK1)  
SAM: MW0545(mvaK1)  
SAR: SAR0596(mvaK1)  
SAS: SAS0549  
SAC: SACOL0636(mvk)  
SAB: SAB0540(mvaK1)  
SAA: SAUSA300\_0572(mvk)  
SAO: SAOUHSC\_00577  
SEP: SE0361  
SER: SERP0238(mvk)  
SHA: SH2402(mvaK1)  
SSP: SSP2122  
LMO: lmo0010  
LMF: LMOF2365\_0011  
LIN: lin0010  
LWE: lwe0011(mvk)  
LLA: L7866(yeaG)  
LLC: LACR\_0454  
LLM: llmg\_0425(mvk)  
SPY: SPy\_0876(mvaK1)  
SPZ: M5005\_Spy\_0682(mvaK1)  
SPM: spyM18\_0937(mvaK1)  
SPG: SpyM3\_0595(mvaK1)  
SPS: SPs1258  
SPH: MGAS10270\_Spy0740(mvaK1)  
SPI: MGAS10750\_Spy0774(mvaK1)



SPJ: MGAS2096\_Spy0753(mvaK1)  
SPK: MGAS9429\_Spy0737(mvaK1)  
SPF: SpyM51126(mvaK1)  
SPA: M6\_Spy0699  
SPB: M28\_Spy0662(mvaK1)  
SPN: SP\_0381  
SPR: spr0338(mvk)  
SPD: SPD\_0346(mvk)  
SAG: SAG1326  
SAN: gbs1396  
SAK: SAK\_1357(mvk)  
SMU: SMU.181  
STC: str0559(mvaK1)  
STL: stu0559(mvaK1)  
STE: STER\_0598  
SSA: SSA\_0333(mvaK1)  
SSU: SSU05\_0289  
SSV: SSU98\_0285  
SGO: SGO\_0239(mvk)  
LPL: lp\_1735(mvaK1)  
LJO: LJ1205  
LAC: LBA1167(mvaK)  
LSA: LSA0908(mvaK1)  
LSL: LSL\_0685(eRG)  
LDB: Ldb0999(mvk)  
LBU: LBUL\_0906  
LBR: LVIS\_0858  
LCA: LSEI\_1491  
LGA: LGAS\_1033  
LRE: Lreu\_0915  
PPE: PEPE\_0927  
EFA: EF0904(mvk)



OOE: OEOE\_1100  
LME: LEUM\_1385  
NFA: nfa22070  
BGA: BG0711  
BAF: BAPKO\_0732  
FPS: FP0313  
MMP: MMP1335  
MAE: Maeo\_0775  
MAC: MA0602(mvk)  
MBA: Mbar\_A1421  
MMA: MM\_1762  
MBU: Mbur\_2395  
MHU: Mhun\_2890  
MEM: Memar\_1812  
MBN: Mboo\_2213  
MST: Msp\_0858(mvk)  
MSI: Msm\_1439  
MKA: MK0993(ERG12)  
HAL: VNG1145G(mvk)  
HMA: rrnAC0077(mvk)  
HWA: HQ2925A(mvk)  
NPH: NP2850A(mvk)  
PTO: PTO1352  
PHO: PH1625  
PAB: PAB0372(mvk)  
PFU: PF1637(mvk)  
TKO: TK1474  
RCI: LRC399(mvk)  
APE: APE\_2439  
HBU: Hbut\_0877  
SSO: SSO0383  
STO: ST2185



SAI: Saci\_2365(mvk)

MSE: Msed\_1602

PAI: PAE3108

PIS: Pisl\_0467

PCL: Pcal\_1835



**Exemplary phosphomevalonate kinase nucleic acids and polypeptides**

HSA: 10654(PMVK)  
PTR: 457350(PMVK)  
MCC: 717014(PMVK)  
MMU: 68603(Pmvk)  
CFA: 612251(PMVK)  
BTA: 513533(PMVK)  
DME: Dmel\_CG10268  
ATH: AT1G31910  
OSA: 4332275  
SCE: YMR220W(ERG8)  
AGO: AGOS\_AER354W  
PIC: PICST\_52257(ERG8)  
CGR: CAGL0F03993g  
SPO: SPAC343.01c  
MGR: MGG\_05812  
ANI: AN2311.2  
AFM: AFUA\_5G10680  
AOR: AO090010000471  
CNE: CNM00100  
UMA: UM00760.1  
DDI: DDBDRAFT\_0184512  
TBR: Tb09.160.3690  
TCR: 507913.20 508277.140  
LMA: LmjF15.1460  
MXA: MXAN\_5017  
OIH: OB0227  
SAU: SA0549(mvaK2)  
SAV: SAV0592(mvaK2)  
SAM: MW0547(mvaK2)  
SAR: SAR0598(mvaK2)



SAS: SAS0551  
SAC: SACOL0638  
SAB: SAB0542(mvaK2)  
SAA: SAUSA300\_0574  
SAO: SAOUHSC\_00579  
SAJ: SaurJH9\_0615  
SEP: SE0363  
SER: SERP0240  
SHA: SH2400(mvaK2)  
SSP: SSP2120  
LMO: lmo0012  
LMF: LMOF2365\_0013  
LIN: lin0012  
LWE: lwe0013  
LLA: L10014(yebA)  
LLC: LACR\_0456  
LLM: llmg\_0427  
SPY: SPy\_0878(mvaK2)  
SPZ: M5005\_Spy\_0684(mvaK2)  
SPM: spyM18\_0939  
SPG: SpyM3\_0597(mvaK2)  
SPS: SPs1256  
SPH: MGAS10270\_Spy0742(mvaK2)  
SPI: MGAS10750\_Spy0776(mvaK2)  
SPJ: MGAS2096\_Spy0755(mvaK2)  
SPK: MGAS9429\_Spy0739(mvaK2)  
SPF: SpyM51124(mvaK2)  
SPA: M6\_Spy0701  
SPB: M28\_Spy0664(mvaK2)  
SPN: SP\_0383  
SPR: spr0340(mvaK2)  
SPD: SPD\_0348(mvaK2)



SAG: SAG1324  
SAN: gbs1394  
SAK: SAK\_1355  
SMU: SMU.938  
STC: str0561(mvaK2)  
STL: stu0561(mvaK2)  
STE: STER\_0600  
SSA: SSA\_0335(mvaK2)  
SSU: SSU05\_0291  
SSV: SSU98\_0287  
SGO: SGO\_0241  
LPL: lp\_1733(mvaK2)  
LJO: LJ1207  
LAC: LBA1169  
LSA: LSA0906(mvaK2)  
LSL: LSL\_0683  
LDB: Ldb0997(mvaK)  
LBU: LBUL\_0904  
LBR: LVIS\_0860  
LCA: LSEI\_1092  
LGA: LGAS\_1035  
LRE: Lreu\_0913  
PPE: PEPE\_0925  
EFA: EF0902  
NFA: nfa22090  
BGA: BG0710  
BAF: BAPKO\_0731  
NPH: NP2852A  
SSO: SSO2988  
STO: ST0978  
SAI: Saci\_1244



**Exemplary diphosphomevalonate decarboxylase nucleic acids and polypeptides**

HSA: 4597(MVD)  
PTR: 468069(MVD)  
MCC: 696865(MVD)  
MMU: 192156(Mvd)  
RNO: 81726(Mvd)  
CFA: 489663(MVD)  
GGA: 425359(MVD)  
DME: Dmel\_CG8239  
SCE: YNR043W(MVD1)  
AGO: AGOS\_AGL232C  
PIC: PICST\_90752  
CGR: CAGL0C03630g  
SPO: SPAC24C9.03  
MGR: MGG\_09750  
ANI: AN4414.2  
AFM: AFUA\_4G07130  
AOR: AO090023000862  
CNE: CNL04950  
UMA: UM05179.1  
DDI: DDBDRAFT\_0218058  
TET: TTHERM\_00849200  
TBR: Tb10.05.0010 Tb10.61.2745  
TCR: 507993.330 511281.40  
LMA: LmjF18.0020  
CBU: CBU\_0607(mvaD)  
CBD: COXBU7E912\_0619(mvaD)  
LPN: lpg2040  
LPF: lpl2018  
LPP: lpp2023



TCX: Tcr\_1734  
DNO: DNO\_0504(mvaD)  
BBA: Bd1629  
MXA: MXAN\_5018(mvaD)  
OIH: OB0226  
SAU: SA0548(mvaD)  
SAV: SAV0591(mvaD)  
SAM: MW0546(mvaD)  
SAR: SAR0597(mvaD)  
SAS: SAS0550  
SAC: SACOL0637(mvaD)  
SAB: SAB0541(mvaD)  
SAA: SAUSA300\_0573(mvaD)  
SAO: SAOUHSC\_00578  
SAJ: SaurJH9\_0614  
SAH: SaurJH1\_0629  
SEP: SE0362  
SER: SERP0239(mvaD)  
SHA: SH2401(mvaD)  
SSP: SSP2121  
LMO: lmo0011  
LMF: LMOF2365\_0012(mvaD)  
LIN: lin0011  
LWE: lwe0012(mvaD)  
LLA: L9089(yeaH)  
LLC: LACR\_0455  
LLM: llmg\_0426(mvaD)  
SPY: SPy\_0877(mvaD)  
SPZ: M5005\_Spy\_0683(mvaD)  
SPM: spyM18\_0938(mvd)  
SPG: SpyM3\_0596(mvaD)  
SPS: SPs1257



SPH: MGAS10270\_Spy0741(mvaD)  
SPI: MGAS10750\_Spy0775(mvaD)  
SPJ: MGAS2096\_Spy0754(mvaD)  
SPK: MGAS9429\_Spy0738(mvaD)  
SPF: SpyM51125(mvaD)  
SPA: M6\_Spy0700  
SPB: M28\_Spy0663(mvaD)  
SPN: SP\_0382  
SPR: spr0339(mvd1)  
SPD: SPD\_0347(mvaD)  
SAG: SAG1325(mvaD)  
SAN: gbs1395  
SAK: SAK\_1356(mvaD)  
SMU: SMU.937  
STC: str0560(mvaD)  
STL: stu0560(mvaD)  
STE: STER\_0599  
SSA: SSA\_0334(mvaD)  
SSU: SSU05\_0290  
SSV: SSU98\_0286  
SGO: SGO\_0240(mvaD)  
LPL: lp\_1734(mvaD)  
LJO: LJ1206  
LAC: LBA1168(mvaD)  
LSA: LSA0907(mvaD)  
LSL: LSL\_0684  
LDB: Ldb0998(mvaD)  
LBU: LBUL\_0905  
LBR: LVIS\_0859  
LCA: LSEI\_1492  
LGA: LGAS\_1034  
LRE: Lreu\_0914



PPE: PEPE\_0926  
EFA: EF0903(mvaD)  
LME: LEUM\_1386  
NFA: nfa22080  
BBU: BB0686  
BGA: BG0709  
BAF: BAPKO\_0730  
GFO: GFO\_3632  
FPS: FP0310(mvaD)  
HAU: Haur\_1612  
HAL: VNG0593G(dmd)  
HMA: rrnAC1489(dmd)  
HWA: HQ1525A(mvaD)  
NPH: NP1580A(mvaD)  
PTO: PTO0478 PTO1356  
SSO: SSO2989  
STO: ST0977  
SAI: Saci\_1245(mvd)  
MSE: Msed\_1576



**Exemplary isopentenyl phosphate kinases (IPK) nucleic acids and polypeptides**

*Methanobacterium thermoautotrophicum* gi|2621082

*Methanococcus jannaschii* DSM 2661 gi|1590842 ;

*Methanocaldococcus jannaschii* gi|1590842

*Methanothermobacter thermautotrophicus* gi|2621082

*Picrophilus torridus* DSM9790 (IG-57) gi|48477569

*Pyrococcus abyssi* gi|14520758

*Pyrococcus horikoshii* OT3 gi|3258052

*Archaeoglobus fulgidus* DSM4304 gi|2648231



**Exemplary isopentenyl-diphosphate Delta-isomerase (IDI) nucleic acids and polypeptides**

HSA: 3422(IDI1) 91734(IDI2)  
PTR: 450262(IDI2) 450263(IDI1)  
MCC: 710052(LOC710052) 721730(LOC721730)  
MMU: 319554(Idi1)  
RNO: 89784(Idi1)  
GGA: 420459(IDI1)  
XLA: 494671(LOC494671)  
XTR: 496783(idi2)  
SPU: 586184(LOC586184)  
CEL: K06H7.9(idi-1)  
ATH: AT3G02780(IPP2)  
OSA: 4338791 4343523  
CME: CMB062C  
SCE: YPL117C(IDI1)  
AGO: AGOS\_ADL268C  
PIC: PICST\_68990(IDI1)  
CGR: CAGL0J06952g  
SPO: SPBC106.15(idi1)  
ANI: AN0579.2  
AFM: AFUA\_6G11160  
AOR: AO090023000500  
CNE: CNA02550  
UMA: UM04838.1  
ECU: ECU02\_0230  
DDI: DDB\_0191342(ipi)  
TET: TTHERM\_00237280 TTHERM\_00438860  
TBR: Tb09.211.0700  
TCR: 408799.19 510431.10  
LMA: LmjF35.5330  
EHI: 46.t00025



ECO: b2889(idi)  
ECJ: JW2857(idi)  
ECE: Z4227  
ECS: ECs3761  
ECC: c3467  
ECI: UTI89\_C3274  
ECP: ECP\_2882  
ECV: APECO1\_3638  
ECW: EcE24377A\_3215(idi)  
ECX: EcHS\_A3048  
STY: STY3195  
STT: t2957  
SPT: SPA2907(idi)  
SEC: SC2979(idi)  
STM: STM3039(idi)  
SFL: SF2875(idi)  
SFX: S3074  
SFV: SFV\_2937  
SSN: SSON\_3042 SSON\_3489(yhfK)  
SBO: SBO\_3103  
SDY: SDY\_3193  
ECA: ECA2789  
PLU: plu3987  
ENT: Ent638\_3307  
SPE: Spro\_2201  
VPA: VPA0278  
VFI: VF0403  
PPR: PBPRA0469(mvaD)  
PEN: PSEEN4850  
CBU: CBU\_0607(mvaD)  
CBD: COXBU7E912\_0619(mvaD)  
LPN: lpg2051



LPF: lpl2029  
LPP: lpp2034  
TCX: Tcr\_1718  
HHA: Hhal\_1623  
DNO: DNO\_0798  
EBA: ebA5678 p2A143  
DVU: DVU1679(idi)  
DDE: Dde\_1991  
LIP: LI1134  
BBA: Bd1626  
AFW: Anae109\_4082  
MXA: MXAN\_5021(fni)  
RPR: RP452  
RTY: RT0439(idi)  
RCO: RC0744  
RFE: RF\_0785(fni)  
RBE: RBE\_0731(fni)  
RAK: A1C\_04190  
RBO: A1I\_04755  
RCM: A1E\_02555  
RRI: A1G\_04195  
MLO: mlr6371  
RET: RHE\_PD00245(ypd00046)  
XAU: Xaut\_4134  
SIL: SPO0131  
SIT: TM1040\_3442  
RSP: RSP\_0276  
RSH: Rsph17029\_1919  
RSQ: Rsph17025\_1019  
JAN: Jann\_0168  
RDE: RD1\_0147(idi)  
DSH: Dshi\_3527



BSU: BG11440(ypgA)  
BAN: BA1520  
BAR: GBAA1520  
BAA: BA\_2041  
BAT: BAS1409  
BCE: BC1499  
BCA: BCE\_1626  
BCZ: BCZK1380(fni)  
BCY: Bcer98\_1222  
BTK: BT9727\_1381(fni)  
BTL: BALH\_1354  
BLI: BL02217(fni)  
BLD: BLi02426  
BAY: RBAM\_021020(fni)  
BPU: BPUM\_2020(fni)  
OIH: OB0537  
SAU: SA2136(fni)  
SAV: SAV2346(fni)  
SAM: MW2267(fni)  
SAR: SAR2431(fni)  
SAS: SAS2237  
SAC: SACOL2341(fni)  
SAB: SAB2225c(fni)  
SAA: SAUSA300\_2292(fni)  
SAO: SAOUHSC\_02623  
SEP: SE1925  
SER: SERP1937(fni-2)  
SHA: SH0712(fni)  
SSP: SSP0556  
LMO: lmo1383  
LMF: LMOF2365\_1402(fni)  
LIN: lin1420



LWE: lwe1399(fni)  
LLA: L11083(yebB)  
LLC: LACR\_0457  
LLM: llmg\_0428(fni)  
SPY: SPy\_0879  
SPZ: M5005\_Spy\_0685  
SPM: spyM18\_0940  
SPG: SpyM3\_0598  
SPS: SPs1255  
SPH: MGAS10270\_Spy0743  
SPI: MGAS10750\_Spy0777  
SPJ: MGAS2096\_Spy0756  
SPK: MGAS9429\_Spy0740  
SPF: SpyM51123(fni)  
SPA: M6\_Spy0702  
SPB: M28\_Spy0665  
SPN: SP\_0384  
SPR: spr0341(fni)  
SPD: SPD\_0349(fni)  
SAG: SAG1323  
SAN: gbs1393  
SAK: SAK\_1354(fni)  
SMU: SMU.939  
STC: str0562(idi)  
STL: stu0562(idi)  
STE: STER\_0601  
SSA: SSA\_0336  
SGO: SGO\_0242  
LPL: lp\_1732(idi1)  
LJO: LJ1208  
LAC: LBA1171  
LSA: LSA0905(idi)



LSL: LSL\_0682  
LDB: Ldb0996(fni)  
LBU: LBUL\_0903  
LBR: LVIS\_0861  
LCA: LSEI\_1493  
LGA: LGAS\_1036  
LRE: Lreu\_0912  
EFA: EF0901  
OOE: OEOE\_1103  
STH: STH1674  
CBE: Cbei\_3081  
DRM: Dred\_0474  
SWO: Swol\_1341  
MTA: Moth\_1328  
MTU: Rv1745c(idi)  
MTC: MT1787(idi)  
MBO: Mb1774c(idi)  
MBB: BCG\_1784c(idi)  
MPA: MAP3079c  
MAV: MAV\_3894(fni)  
MSM: MSMEG\_1057(fni) MSMEG\_2337(fni)  
MUL: MUL\_0380(idi2)  
MVA: Mvan\_1582 Mvan\_2176  
MGI: Mflv\_1842 Mflv\_4187  
MMC: Mmcs\_1954  
MKM: Mkms\_2000  
MJL: Mjls\_1934  
CGL: NCgl2223(cgl2305)  
CGB: cg2531(idi)  
CEF: CE2207  
CDI: DIP1730(idi)  
NFA: nfa19790 nfa22100



RHA: RHA1\_ro00239  
SCO: SCO6750(SC5F2A.33c)  
SMA: SAV1663(idi)  
LXX: Lxx23810(idi)  
CMI: CMM\_2889(idiA)  
AAU: AAur\_0321(idi)  
PAC: PPA2115  
FRA: Francci3\_4188  
FRE: Franean1\_5570  
FAL: FRAAL6504(idi)  
KRA: Krad\_3991  
SEN: SACE\_2627(idiB\_2) SACE\_5210(idi)  
STP: Strop\_4438  
SAQ: Sare\_4564 Sare\_4928  
RXY: Rxyl\_0400  
BBU: BB0684  
BGA: BG0707  
SYN: sll1556  
SYC: syc2161\_c  
SYF: Synpcc7942\_1933  
CYA: CYA\_2395(fni)  
CYB: CYB\_2691(fni)  
TEL: tll1403  
ANA: all4591  
AVA: Ava\_2461 Ava\_B0346  
TER: Tery\_1589  
SRU: SRU\_1900(idi)  
CHU: CHU\_0674(idi)  
GFO: GFO\_2363(idi)  
FJO: Fjoh\_0269  
FPS: FP1792(idi)  
CTE: CT0257



CCH: Cag\_1445  
CPH: Cpha266\_0385  
PVI: Cvib\_1545  
PLT: Plut\_1764  
RRS: RoseRS\_2437  
RCA: Rcas\_2215  
HAU: Haur\_4687  
DRA: DR\_1087  
DGE: Dgeo\_1381  
TTH: TT\_P0067  
TTJ: TTHB110  
MJA: MJ0862  
MMP: MMP0043  
MMQ: MmarC5\_1637  
MMX: MmarC6\_0906  
MMZ: MmarC7\_1040  
MAE: Maeo\_1184  
MVN: Mevan\_1058  
MAC: MA0604(idi)  
MBA: Mbar\_A1419  
MMA: MM\_1764  
MBU: Mbur\_2397  
MTP: Mthe\_0474  
MHU: Mhun\_2888  
MLA: Mlab\_1665  
MEM: Memar\_1814  
MBN: Mboo\_2211  
MTH: MTH48  
MST: Msp\_0856(fni)  
MSI: Msm\_1441  
MKA: MK0776(lldD)  
AFU: AF2287



HAL: VNG1818G(idi) VNG6081G(crt\_1) VNG6445G(crt\_2) VNG7060 VNG7149

HMA: rrnAC3484(idi)

HWA: HQ2772A(idiA) HQ2847A(idiB)

NPH: NP0360A(idiB\_1) NP4826A(idiA) NP5124A(idiB\_2)

TAC: Ta0102

TVO: TVN0179

PTO: PTO0496

PHO: PH1202

PAB: PAB1662

PFU: PF0856

TKO: TK1470

RCI: LRC397(fni)

APE: APE\_1765.1

SMR: Smar\_0822

IHO: Igni\_0804

HBU: Hbut\_0539

SSO: SSO0063

STO: ST2059

SAI: Saci\_0091

MSE: Msed\_2136

PAI: PAE0801

PIS: Pisl\_1093

PCL: Pcal\_0017

PAS: Pars\_0051

TPE: Tpen\_0272



**Exemplary isoprene synthase nucleic acids and polypeptides**

Genbank Accession Nos.

AY341431

AY316691

AY279379

AJ457070

AY182241



## CLAIMS

1. Cells in culture that produce greater than about 400 nmole/g<sub>wcm</sub>/hr of isoprene.
2. Cells in culture that have an average volumetric productivity of isoprene greater than about 0.1 mg/L<sub>broth</sub>/hr.
3. Cells in culture that convert more than about 0.002 molar percent of the carbon that the cells consume from a cell culture medium into isoprene.
4. Cells in culture comprising a heterologous nucleic acid encoding an isoprene synthase polypeptide, wherein the cells (i) produce greater than about 400 nmole/g<sub>wcm</sub>/hr of isoprene, (ii) convert more than about 0.002 molar percent of the carbon that the cells consume from a cell culture medium into isoprene, or (iii) have an average volumetric productivity of isoprene greater than about 0.1 mg/L<sub>broth</sub>/hr.
5. Cells in culture comprising a heterologous nucleic acid encoding an isoprene synthase polypeptide, wherein the heterologous nucleic acid is operably linked to a promoter, and wherein the production of isoprene by the cells is greater than about 400 nmole/g<sub>wcm</sub>/hr.
6. Cells in culture comprising a heterologous nucleic acid encoding an isoprene synthase polypeptide, wherein the heterologous nucleic acid is operably linked to a promoter, and wherein the cells have an average volumetric productivity of isoprene greater than about 0.1 mg/L<sub>broth</sub>/hr.
7. Cells in culture comprising a heterologous nucleic acid encoding an isoprene synthase polypeptide, wherein the heterologous nucleic acid is operably linked to a promoter, and wherein more than about 0.002 molar percent of the carbon that the cells consume from a cell culture medium is converted into isoprene.
8. The cells of any one of claims 1, 4, and 5, wherein the cells produce greater than about  $4.0 \times 10^3$  nmole/g<sub>wcm</sub>/hr of isoprene.



9. The cells of claim 8, wherein the cells produce greater than about  $4.0 \times 10^4$  nmole/g<sub>wcm</sub>/hr of isoprene.
10. The cells of claim 9, wherein the cells produce greater than about  $1.5 \times 10^5$  nmole/g<sub>wcm</sub>/hr of isoprene.
11. The cells of any one of claims 1, 4, and 5, wherein the cells produce between about 400 nmole/g<sub>wcm</sub>/hr to about  $2.0 \times 10^5$  nmole/g<sub>wcm</sub>/hr of isoprene.
12. The cells of any one of claims 2, 4, and 6, wherein the cells have an average volumetric productivity of isoprene greater than about 100 mg/L<sub>broth</sub>/hr.
13. The cells of claim 12, wherein the cells have an average volumetric productivity of isoprene greater than about 1,000 mg/L<sub>broth</sub>/hr.
14. The cells of claim 13, wherein the cells have an average volumetric productivity of isoprene greater than about 1,500 mg/L<sub>broth</sub>/hr.
15. The cells of any one of claims 2, 4, and 6, wherein the cells have an average volumetric productivity of isoprene between about 0.1 to about 1,500 mg/L<sub>broth</sub>/hr
16. The cells of any one of claims 3, 4, and 7, wherein the cells convert more than about 0.05 molar percent of the carbon into isoprene.
17. The cells of claim 16, wherein the cells convert more than about 1 molar percent of the carbon into isoprene.
18. The cells of claim 17, wherein the cells convert more than about 10 molar percent of the carbon into isoprene.
19. The cells of claim 18, wherein the cells convert more than about 20 molar percent of the carbon into isoprene.



20. The cells of any one of claims 3, 4, and 7, wherein the cells convert between about 0.002 to about 25 molar percent of the carbon into isoprene.
21. The cells of any one of claims 4-7, wherein the isoprene synthase polypeptide is a plant isoprene synthase polypeptide.
22. The cells of any one of claims 1-21, further comprising a heterologous nucleic acid encoding an IDI polypeptide.
23. The cells of any one of claims 1-22, further comprising an insertion of a copy of an endogenous nucleic acid encoding an IDI polypeptide.
24. The cells of any one of claims 1-23, further comprising a heterologous nucleic acid encoding a DXS polypeptide.
25. The cells of any one of claims 1-24, further comprising an insertion of a copy of an endogenous nucleic acid encoding a DXS polypeptide.
26. The cells of any one of claims 1-25, further comprising one or more nucleic acids encoding an IDI polypeptide and a DXS polypeptide.
27. The cells of claim 26, wherein one nucleic acid encodes the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide.
28. The cells of claim 27, wherein one plasmid encodes the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide.
29. The cells of any one of claims 1-28, further comprising a heterologous nucleic acid encoding an MVA pathway polypeptide.



30. The cells of claim 29, wherein the MVA pathway polypeptide is a mevalonate kinase (MVK).
31. The cells of claim 30, wherein the MVK is a polypeptide from *Methanosarcina*.
32. The cells of claim 31, wherein the MVK is a polypeptide from *Methanosarcina mazei*.
33. The cells of any one of claims 4-32, wherein the isoprene synthase polypeptide is a naturally-occurring polypeptide from *Pueraria*.
34. The cells of claim 33, wherein the isoprene synthase polypeptide is a naturally-occurring polypeptide from *Pueraria montana*.
35. The cells of any one of claims 4-32, wherein the isoprene synthase polypeptide is a naturally-occurring polypeptide from *Populus*.
36. The cells of claim 35, wherein the isoprene synthase polypeptide is a naturally-occurring polypeptide from *Populus alba*.
37. The cells of any one of claims 35 and 36, further comprising a heterologous nucleic acid encoding an MVA pathway polypeptide.
38. The cells of claim 37, wherein the MVK is a polypeptide from *Methanosarcina*.
39. The cells of claim 38, wherein the MVK is a polypeptide from *Methanosarcina mazei*.
40. The cells of any one of claims 1-39, wherein the cells are bacterial cells.
41. The cells of claim 40, wherein the cells are gram-positive bacterial cells.
42. The cells of claim 41, wherein the cells are *Bacillus* cells.



43. The cells of claim 42, wherein the cells are *Bacillus subtilis* cells.
44. The cells of claim 40, wherein the cells are gram-negative bacterial cells.
45. The cells of claim 44, wherein the cells are *Escherichia* or *Pantoea* cells.
46. The cells of claim 45, wherein the cells are *Escherichia coli* or *Pantoea citrea* cells.
47. The cells of any one of claims 1-39, wherein the cells are fungal cells.
48. The cells of claim 47, wherein the cells are *Trichoderma* cells.
49. The cells of claim 48, wherein the cells are *Trichoderma reesei* cells.
50. The cells of claim 47, wherein the cells are yeast cells.
51. The cells of claim 50, wherein the cells are *Yarrowia* cells.
52. The cells of claim 51, wherein the cells are *Yarrowia lipolytica* cells.
53. A method for producing isoprene, the method comprising culturing cells under conditions sufficient to produce greater than about 400 nmole/g<sub>wcm</sub>/hr of isoprene.
54. A method for producing isoprene, the method comprising culturing cells under conditions resulting in an average volumetric productivity of isoprene greater than about 0.1 mg/L<sub>broth</sub>/hr.
55. A method for producing isoprene, the method comprising culturing cells under conditions sufficient to convert more than about 0.002 molar percent of the carbon that the cells consume from a cell culture medium into isoprene.



56. A method for producing isoprene, the method comprising culturing cells comprising a heterologous nucleic acid encoding an isoprene synthase polypeptide, wherein the heterologous nucleic acid is operably linked to a promoter, and wherein production of isoprene by the cells is greater than about 400 nmole/g<sub>wcm</sub>/hr.

57. A method for producing isoprene, the method comprising culturing cells comprising a heterologous nucleic acid encoding an isoprene synthase polypeptide, wherein the heterologous nucleic acid is operably linked to a promoter, and wherein the cells have an average volumetric productivity of isoprene greater than about 0.1 mg/L<sub>broth</sub>/hr.

58. A method for producing isoprene, the method comprising culturing cells comprising a heterologous nucleic acid encoding an isoprene synthase polypeptide, wherein the heterologous nucleic acid is operably linked to a promoter, and wherein more than about 0.002 molar percent of the carbon that the cells consume from a cell culture medium is converted into isoprene.

59. A method of producing isoprene, the method comprising  
(a) culturing cells comprising a heterologous nucleic acid encoding an isoprene synthase polypeptide under suitable culture conditions for the production of isoprene, wherein the cells (i) produce greater than about 400 nmole/g<sub>wcm</sub>/hr of isoprene, (ii) convert more than about 0.002 molar percent of the carbon that the cells consume from a cell culture medium into isoprene, or (iii) have an average volumetric productivity of isoprene greater than about 0.1 mg/L<sub>broth</sub>/hr of isoprene, and  
(b) producing isoprene.

60. The method of any one of claims 53, 56, and 59, wherein the cells produce greater than about  $4.0 \times 10^3$  nmole/g<sub>wcm</sub>/hr of isoprene.

61. The method of claim 60, wherein the cells produce greater than about  $4.0 \times 10^4$  nmole/g<sub>wcm</sub>/hr of isoprene.

62. The method of claim 61, wherein the cells produce greater than about  $1.5 \times 10^5$  nmole/g<sub>wcm</sub>/hr of isoprene.



63. The method of any one of claims 53, 56, and 59 wherein the cells produce between about 400 nmole/g<sub>wcm</sub>/hr to about  $2.0 \times 10^5$  nmole/g<sub>wcm</sub>/hr of isoprene.
64. The method of any one of claims 54, 57, and 59, wherein the cells have an average volumetric productivity of isoprene greater than about 100 mg/L<sub>broth</sub>/hr.
65. The method of claim 64, wherein the cells have an average volumetric productivity of isoprene greater than about 1,000 mg/L<sub>broth</sub>/hr.
66. The method of claim 65, wherein the cells have an average volumetric productivity of isoprene greater than about 1,500 mg/L<sub>broth</sub>/hr.
67. The cells of any one of claims 54, 57, and 59, wherein the cells have an average volumetric productivity of isoprene between about 0.1 to about 1,500 mg/L<sub>broth</sub>/hr
68. The method of any one of claims 55, 58, and 59, wherein the cells convert more than about 0.05 molar percent of the carbon into isoprene.
69. The method of claim 68, wherein the cells convert more than about 1 molar percent of the carbon into isoprene.
70. The method of claim 69, wherein the cells convert more than about 10 molar percent of the carbon into isoprene.
71. The method of claim 70, wherein the cells convert more than about 20 molar percent of the carbon into isoprene.
72. The method of any one of claims 55, 58, and 59, wherein the cells convert between about 0.002 to about 25 molar percent of the carbon into isoprene.



73. The method of any one of claims 53-72, further comprising recovering the isoprene produced by the cells.
74. The method of any one of claims 56-73, wherein the isoprene synthase polypeptide is a plant isoprene synthase polypeptide.
75. The method of any one of claims 53-74, wherein the cells further comprise a heterologous nucleic acid encoding an IDI polypeptide.
76. The method of any one of claims 53-75, wherein the cells further comprise an insertion of a copy of an endogenous nucleic acid encoding an IDI polypeptide.
77. The method of any one of claims 53-76, wherein the cells further comprise a heterologous nucleic acid encoding a DXS polypeptide.
78. The method of any one of claims 53-77, wherein the cells further comprise an insertion of a copy of an endogenous nucleic acid encoding a DXS polypeptide.
79. The method of any one of claims 53-78, wherein the cells further comprise one or more nucleic acids encoding an IDI polypeptide and a DXS polypeptide.
80. The method of claim 79, wherein one nucleic acid encodes the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide.
81. The method of claim 80, wherein one plasmid encodes the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide.
82. The method of any one of claims 53-81, wherein the cells further comprise a heterologous nucleic acid encoding an MVA pathway polypeptide.
83. The cells of claim 82, wherein the MVA pathway polypeptide is a mevalonate kinase (MVK).



84. The cells of claim 83, wherein the MVK is a polypeptide from *Methanosarcina*.
85. The cells of claim 84, wherein the MVK is a polypeptide from *Methanosarcina mazei*.
86. The method of any one of claims 56-85, wherein the isoprene synthase polypeptide is a naturally-occurring polypeptide from *Pueraria*.
87. The method of claim 86, wherein the isoprene synthase polypeptide is a naturally-occurring polypeptide from *Pueraria montana*.
88. The method of any one of claims 56-85, wherein the isoprene synthase polypeptide is a naturally-occurring polypeptide from *Populus*.
89. The method of claim 88, wherein the isoprene synthase polypeptide is a naturally-occurring polypeptide from *Populus alba*.
90. The method of any one of claims 88 and 89, further comprising a heterologous nucleic acid encoding an MVA pathway polypeptide.
91. The cells of claim 90, wherein the MVK is a polypeptide from *Methanosarcina*.
92. The cells of claim 91, wherein the MVK is a polypeptide from *Methanosarcina mazei*.
93. The method of any one of claims 53-92, wherein the cells are bacterial cells.
94. The method of claim 93, wherein the cells are gram-positive bacterial cells.
95. The method of claim 94, wherein the cells are *Bacillus* cells.
96. The method of claim 95, wherein the cells are *Bacillus subtilis* cells.
97. The method of claim 93, wherein the cells are gram-negative bacterial cells.



98. The method of claim 97, wherein the cells are *Escherichia* or *Pantoea* cells.
99. The method of claim 98, wherein the cells are *Escherichia coli* or *Pantoea citrea* cells.
100. The method of any one of claims 53-92 wherein the cells are fungal cells.
101. The method of claim 100, wherein the cells are *Trichoderma* cells.
102. The method of claim 100, wherein the cells are *Trichoderma reesei* cells.
103. The method of claim 100, wherein the cells are yeast cells.
104. The method of claim 103, wherein the cells are *Yarrowia* cells.
105. The method of claim 104, wherein the cells are *Yarrowia lipolytica* cells.



**Figure 1**

1-

atgtgtgcgacctcttctcaattttactcagattaccgagcataattcccgtcgttccgcaaact  
atcagccaaacctgtggaatttcgaattcctgcaatccctgggagaacgacctgaaagtggaaaa  
gctggaggagaaaagcgaccaaaactggaggaagaagttcgctgcatgatcaaccgtgtagacacc  
cagccgctgtccctgctggagctgatcgacgatgtgcagcgccctgggtctgacctacaaatttg  
aaaaagacatcattaaagccctggaaaacatcgtactgctggacgaaaaacaaaaagaacaaatc  
tgacctgcacgcaaccgctctgtctttccgtctgctgcgtcagcacggtttcgagggtttctcag  
gatgttttttgagcgtttcaaggataaagaaggtggtttcagcggtgaactgaaaggtgacgtcc  
aaggcctgctgagcctgtatgaagcgtcttacctgggttttcgagggtgagaacctgctggagga  
ggcgctaccttttccatcacccacctgaagaacaacctgaaagaaggcattaataccaaggtt  
gcagaacaagtgagccacgccttggaaactgccatatcaccagcgtctgcaccgtctggaggcac  
gttggttctctggataaatacgaaccgaaagaaccgcatcaccagctgctgctggagctggcgaa  
gctggatttttaacatggtacagacctgcaccagaaagagctgcaagatctgtcccgtggtgg  
accgagatgggacctggctagcaaaactggattttgtacgcgaccgcctgatggaagtttatttct  
gggcactgggtatggcgccagacccgcagtttggtgaatgtcgcaaagctgttactaaaatgtt  
tggtctggtgacgatcatcgatgacgtgtatgacgtttatggcactctggacgaactgcaactg  
ttcaccgatgctgtagagcgtgggacgttaacgctattaacacctgccggactatatgaaac  
tgtgtttcctggcactgtacaacaccgttaacgacacgtcctattctattctgaaagagaaagg  
tcataacaacctgtcctatctgacgaaaagctggcgtgaactgtgcaaagcctttctgcaagag  
gcgaaatggtccaacaacaaaattatcccggcttttctccaagtacctggaaaacgccagcgttt  
cctcctccggtgtagcgctgctggcgccgtcttacttttccgtatgccagcagcaggaagacat  
ctccgaccacgcgctgcgttccctgaccgacttccatggtctggtgcgttctagctgcgttatac  
ttccgcctgtgcaacgatctggccacctctgcggcggagctggaacgtggcgagactaccaatt  
ctatcattagctacatgcacgaaaacgatggtaccagcgaggaacaggccccgcgaagaactgcg  
taaactgatcgacgccgaatggaaaaagatgaatcgtgaacgcgttagcgactccacctgctg  
cctaaagcgttcatggaaatcgacgttaacatggcacgtgtttcccactgcacctaccagtatg  
gcgatggtctgggtcgcccagactacgcgactgaaaaccgcatcaaactgctgctgattgacct  
tttcccgattaaccagctgatgtatgtc

taactgcag

(SEQ ID NO:1)



2/167

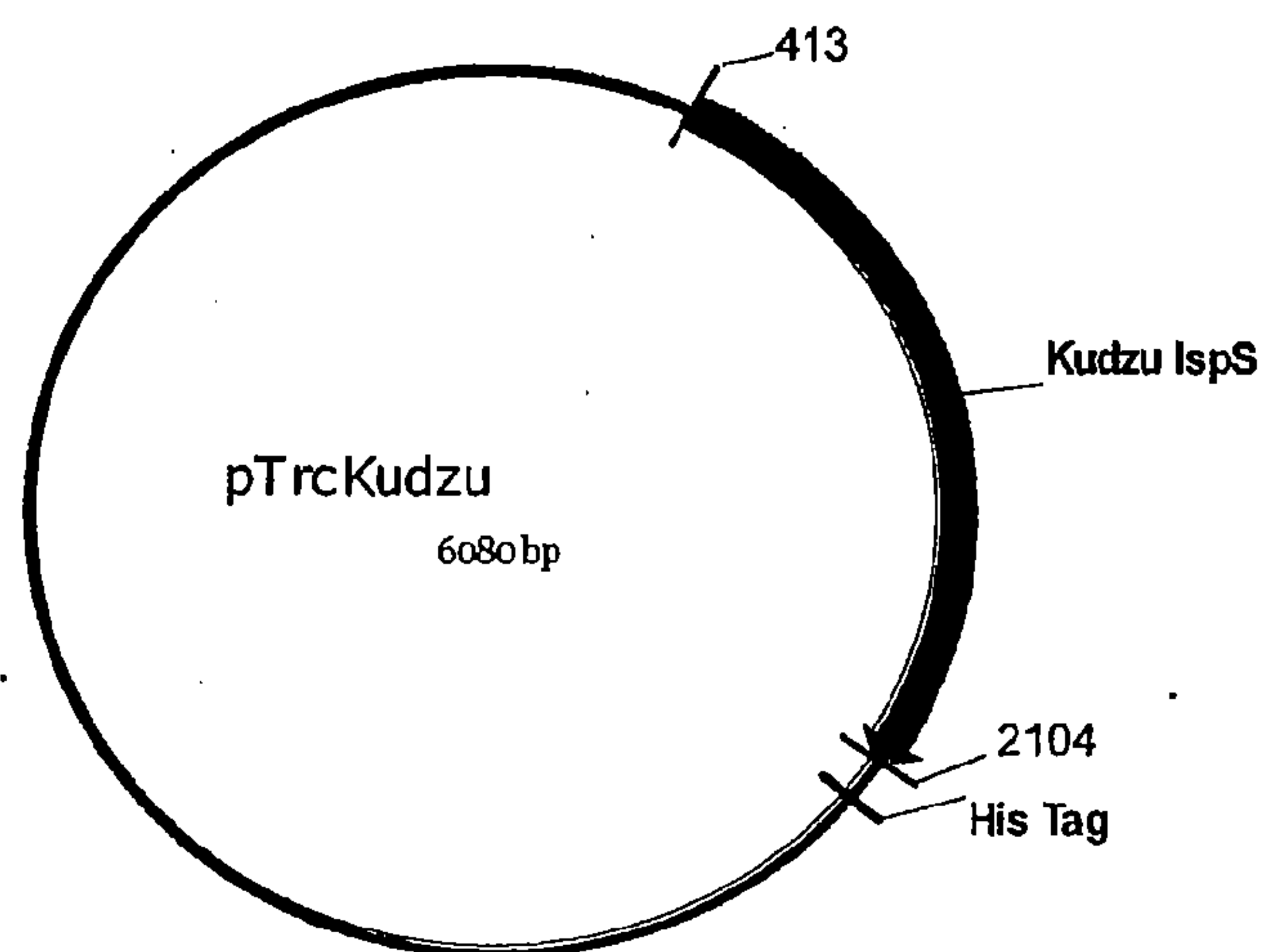
**Figure 2**



**Figure 3A**

1-

gtttgacagcttatcatcgactgcacggtgcaccaatgcttctggcgtcaggcagccatcggaa  
gctgtggtatggctgtgcaggtcgtaaatcactgcataattcgtgtcgctcaaggcgcactccc  
gttctggataatgttttttgcgccgacatcataacggttctggcaaataattctgaaatgagctg  
ttgacaattaatcatccggctcgtataatgtgtggaattgtgagcggataacaatttcacacag  
gaaacagcgcgcgtgagaaaaagcgaagcggcactgctctttaacaatttatcagacaatctgt  
gtgggcactcgaccggaattatcgattaactttattattaaaaattaaagaggtatatattaat  
gtatcgattaaataaggaggaataaacc**ATG**tgtgcgacctcttctcaatttactcagattacc  
gagcataattcccgtcgttccgcaaactatcagccaaacctgtggaatttcgaattcctgcaat  
ccctggagaacgacctgaaagtggaaaagctggaggagaaagcgaccaaactggaggaagaagt  
tcgctgcatgatcaaccgtgtagacaccagccgctgtccctgctggagctgatcgacgatgtg  
cagcgcctgggtctgacctacaaatttgaaaaagacatcattaaagccctggaaaacatcgtac  
tgctggacgaaaaacaaaaagaacaaatctgacctgcacgcaaccgctctgtctttccgtctgct  
gcgtcagcacgggtttcgagggtttctcaggatgtttttgagcgtttcaaggataaagaagggtggt  
ttcagcgggtgaactgaaagggtgacgtccaaggcctgctgagcctgtatgaagcgtcttacctgg  
gttttcgaggggtgagaacctgctggaggaggcgcgtaccttttccatcacccacctgaagaacaa  
cctgaaagaaggcattaataccaagggttgacagaacaagtgagccacgcccctggaaactgccatat  
caccagcgtctgcaccgtctggaggcacgttggttcctggataaatacgaaccgaaagaaccgc  
atcaccagctgctgctggagctggcgaagctggatttttaacatggtacagaccctgcaccagaa  
agagctgcaagatctgtcccgtggtggaccgagatgggcctggctagcaaactggatttttgta  
cgcgaccgcctgatggaagtttattttctgggcactgggtatggcgccagaccgcagtttggtg  
aatgtcgcaaagctgttactaaaatgtttggtctggtgacgatcatcgatgacgtgtatgacgt  
ttatggcactctggacgaactgcaactgttcaccgatgctgtagagcgtgggacgttaacgct  
attaacaccctgccggactatatgaaactgtgtttcctggcactgtacaacaccgttaacgaca  
cgtcctattctattctgaaagagaaaggtcataacaacctgtcctatctgacgaaaagctggcg  
tgaactgtgcaaagcctttctgcaagaggcgaaatgggtccaacaacaaaattatcccggctttc  
tccaagtacctggaaaacgccagcgtttcctcctccggtgtagcgtgctggcgccgtcttact  
tttccgtatgccagcagcaggaagacatctccgaccacgcgctgcgttccctgaccgacttcca  
tggtctggtgcgttctagctgcgttatcttccgcctgtgcaacgatctggccacctctgcggcg  
gagctggaacgtggcgagactaccaattctatcattagctacatgcacgaaaacgatggtacca  
gcgaggaacaggcccgcgaagaactgcgtaaactgatcgacgccgaatggaaaaagatgaatcg  
tgaacgcgttagcgactccacctgctgcctaaagcgttcatggaaatcgcagttaacatggca  
cgtgtttcccaactgcacctaccagtatggcgatgggtctgggtcgccagactacgcgactgaaa  
accgcatcaaactgctgctgattgaccctttcccgattaaccagctgatgtatgtc**TAA**ctgca  
gctggtaccatatgggaattcgaagctttctagaacaaaaactcatctcagaagaggatctgaa  
tagcgccgtcgaccatcatcatcatcattgagtttaaacggtctccagcttggtgttttg  
gcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggctctgataaa  
acagaatttgcttggcggcagtagcgcggtggtcccacctgaccccatgccgaactcagaagtg  
aaacgccgtagcgccgatggtagtgtgggtctcccatgcgagagtagggaactgccaggcat  
caaataaaacgaaaggctcagtcgaaagactgggcctttcggttttatctgttggttggtga  
acgctctcctgagtaggacaaatccgcggggagcggatttgaaacgttgcggaagcaacggcccg



**Figure 3B**

aggggtggcgggcaggacgccccgccataaactgccaggcatcaaattaagcagaaggccatcctg  
acggatggccttttttgcgtttctacaaactctttttgtttatttttctaaatacattcaaatat  
gtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatg  
agtattcaacattttccgtgtcgccttattcccttttttgcggcattttgccttcctgtttttg  
ctcaccagaaacgctgggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggtta  
catcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttcca  
atgatgagcacttttaagttctgctatgtggcgcggtattatcccgtgttgacgcccgggcaag  
agcaactcggtcgcccgcatacactattctcagaatgacttggttgagtactcaccagtcacaga  
aaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgat  
aacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgc  
acaacatgggggatcatgtaactcgccttgatcgttggaaccggagctgaatgaagccatacc  
aaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaact  
ggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttg  
caggaccacttctgcgctcggcccttccggctggctgggtttattgctgataaatctggagccgg  
tgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtat  
gttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagatag  
gtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatacttttagattga  
tttaaaacttcattttttaatttaaaaggatctaggtgaagatccttttttgataatctcatgacc  
aaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggat  
cttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctacc  
agcgggtgggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagc  
agagcgcagataccaaataactgtccttctagtgtagccgtagttaggccaccacttcaagaact  
ctgtagcaccgcctacatacctcgtctctgctaactcgttaccagtggctgctgccagtgggcga  
taagtcgtgtcttacccgggttggaactcaagacgatagttaccggataaggcgcagcggtcgggc  
tgaacgggggggttcgtgcacacagcccagcttgagcgaacgacctaaccgaactgagatacc  
tacagcgtgagctatgagaaagcgcacgcttcccgaaggggagaaaggcggacaggtatccggt  
aagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggatctt  
tatagtcctgtcgggtttcgcacactctgacttgagcgtcgattttttgtgatgctcgtcagggg  
ggcggagcctatggaaaaacgccagcaacgcggcctttttacgggttccctggccttttgctggcc  
ttttgctcacatgttctttcctgcggttatccctgattctgtggataaccgtattaccgccttt  
gagtgagctgataccgctcgcgcagccgaacgaccgagcgcagcagtcagtgagcaggaag  
cggaagagcgcctgatgcggtatttttctccttacgcatctgtgcggtatttcacaccgcatatg  
gtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatcgc  
tacgtgactgggtcatggctgcgccccgacaccccgccaacacccgctgacgcgccctgacgggc  
ttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcag  
aggtttttaccgctcatcaccgaaacgcgcgagggcagcagatcaattcgcgcgcgaaggcgaagc  
ggcatgcattttacgttgacaccatcgaatgggtgcaaaacctttcgcgggtatggcatgatagcgc  
ccggaagagagtcaattcaggggtgggtgaatgtgaaaccagtaacggttatacagatgtcgcagagt  
atgccggtgtctcttatcagaccgtttcccgcgtgggtgaaccaggccagccacgtttctgcgaa  
aacgcgggaaaaagtggaaagcggcgatggcggagctgaattacattcccaaccgcgtggcaca  
caactggcggggcaaacagtcgttgctgattggcgttgccacctccagtcctggccctgcacgcgc  
cgtcgcaaatgtgcgcgcgattaaatctcgcgcgcatcaactgggtgccagcgtgggtgggtgc  
gatggtagaacgaagcggcgctcgaagcctgtaaagcggcggtgcacaatcttctcgcgcaacgc  
gtcagtgggctgatcattaactatccgctggatgaccaggatgccattgctgtggaagctgcct



**Figure 3C**

Gcactaatgttccggcggttatttcttgatgtctctgaccagacacccatcaacagtattattt  
ctcccatgaagacggtacgcgactgggcggtggagcatctgggtcgcatcattgggtcaccagcaa  
atcgcgctgttagcggggcccatcattagttctgtctcggcgcgctctgcgtctgggtggctggc  
ataaatatctcactcgcaatcaaattcagccgatagcgggaacgggaaggcgactggagtgcc  
atgtccggttttcaacaaaccatgcaaattgctgaatgagggcatcgttcccactgcgatgctg  
ggttgccaacgatcagatggcgctgggcgcaatgcgcgccattaccgagtcggggctgcgcgtt  
ggtgcggatatctcggtagtgggatacgacgataccgaagacagctcatgttatatcccgc  
cgtcaaccaccatcaaacaggattttcgctgctggggcaaaccagcgtggaccgcttgctg  
caactctctcagggccagggcgtgaagggcaatcagctgttgcccgtctcactggtgaaa  
agaaaaaccaccctggcgc ccaatacgcaaaccgcctctccccgcgcgcttggccgatt  
cattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacg  
caattaatgtgagtttagcgcgaattgatctg (SEQ ID NO:2)



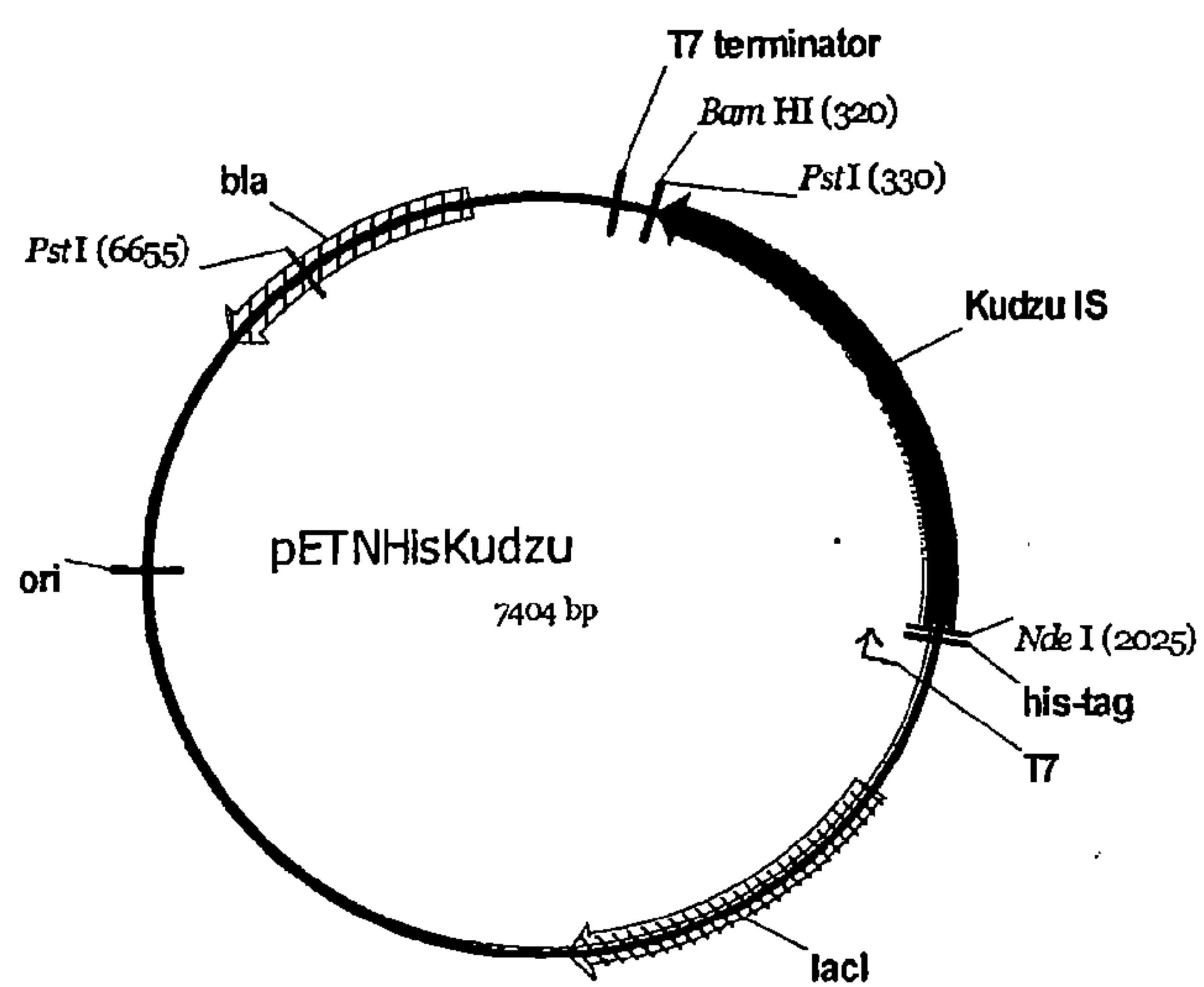
**Figure 4**



**Figure 5A**

1-

ttctcatggttgacagcttatcatcgataagctttaatgcggtagtttatcacagttaaattgc  
taacgcagtcagggcaccgtgtatgaaatctaacaatgcgctcatcgctcatcctcggcaccgtca  
ccctggatgctgtaggcataggcttggttatgccggtactgccgggcctcttgccgggatatccg  
gatatagttcctcctttcagcaaaaaacccctcaagaccgcttagaggccccaaggggttatg  
ctagttattgctcagcgggtggcagcagccaactcagcttcctttcgggctttgttagcagccgg  
atccctgcagttagacatacatcagctggttaatcgggaaaggggtcaatcagcagcagtttgat  
gcgggttttcagtcgcgtagtctgggcgacccagaccatcgccatactggtaggtgcagtgaggaa  
acacgtgccatgttaactgcgatttccatgaacgcttaggcagcaggggtggagtcgctaacgc  
gttcacgattcatctttttccattcggcgctcgatcagtttacgcagttcttcgcgggcctgttc  
ctcgctggtaccatcggttttcgtgcatgtagctaataatgtagaattggtagtctcgccacgttcc  
agctccgcccgcagaggtggccagatcggtgcacaggcggaagataacgcagctagaacgcacca  
gaccatggaagtcggtcagggaaacgcagcgcgtgggtcggagatgtcttcctgctgctggcatac  
ggaaaagtaagacggcgccagcagcgtacaccggaggaggaaacgcgtggcggtttccagggtac  
ttggagaaagccgggataattttgttggtggaccatttcgcctcttgcaaaaggctttgcaca  
gttcacgccagcttttcgtcagataggacaggttggttatgacctttctctttcagaatagaata  
ggacgtgtcggttaacgggtgttgtagcagtgccaggaaacacagtttcatatagtcgggcaggggtg  
ttaatagcggttaacgtcccagcgcctctacagcatcggtgaacagttgcagttcgtccagagtg  
cataaacgtcatacagtcacgatgatcgctcaccagaccacacatttttagtaacagctttgcg  
acattcaccaaactgcgggtctggcgccatacccagtgcccagaaataaacttccatcaggcg  
tcgcgtacaaaatccagtttgctagccaggcccatctcggtccaccagcgggacagatcttgca  
gctctttctggtgcaggggtctgtaccatgttaaaatccagcttcgccagctccagcagcagctg  
gtgatgcgggttctttcgggttcgtatttatccaggaaaccaacgtgcctccagacgggtgcagacgc  
tggtgatatggcagttccaggggcggtgggtcacttggtctgcaaccttggtattaatgccttctt  
tcaggttggtcttcaggtgggtgatggaaaaggtagcgccctcctccagcaggttctcaccctc  
gaaacccaggtaagacgcttcatacagggtcagcagggccttgagcgtcacctttcagttcacccg  
ctgaaaccaccttctttatccttgaaacgctcaaaaacatcctgagaaacctcgaaaccgtgct  
gacgcagcagacggaaagacagagcggttgcggtgcaggtcagatttggttctttttgttttcgtc  
cagcagtacgatgttttccagggttttaatgatgtctttttcaaatttgtaggtcagaccagg  
cgctgcacatcgctcgatcagctccagcagggacagcgggtgggtgtctacacgggtgatcatgc  
agcgaacttcttcctccagtttggtcgctttctcctccagcttttccactttcaggtcgttctc  
cagggttgtaggaattcgaaattccacaggtttgggtgatagtttgcggaacgcaggggaatta  
tgctcggtaatctgagtaaatgagaagaggtcgcacacatatgacgaccttcgatatggccgc  
tgctgtgatgatgatgatgatgatgatggcccatgggtatatctccttcttaaaagttaa  
acaaaattatttctagaggggaattgttatccgctcacaattcccctatagtgagtcgtattaa  
tttcgcgggatcgagatctcgatcctctacgccggacgcacatcggtggccggcatcacccggcgcca  
caggtgcgggttgctggcgccctatatcgccgacatcacccgatggggaagatcgggctcgccactt  
cgggctcatgagcgttggttcggcggtgggtatgggtggcagggccccgtggccgggggactgttg  
ggcgccatctccttgcatgcaccattccttgccggcggtgctcaacggcctcaacctactac  
tgggctgcttctaatgcaggagtcgcataaggagagcgctcgagatcccggacaccatcgaat  
ggcgcaaaacctttcgcggtatggcatgatagcggccggaagagagtcattcaggggtggtgaa  
tgtgaaaccagtaacgttatagatgtcgagagtatgccggtgtctcttatcagaccgtttcc  
cgcggtggtgaaccaggccagccacgtttctgcgaaaacgcgggaaaaagtggaagcggcgatgg  
cggagctgaattacattcccaaccgcgtggcacaacaactggcgggcaaacagtcgttgctgat  
tggcggttgccacctccagtcctggccctgcacgcgcgtcgcaaattgtcgcggcgattaaatct



**Figure 5B**

cgcgccgatcaactgggtgccagcgtgggtgggtgtcgatggtagaacgaagcggcgtcgaagcct  
gtaaagcggcgggtgcacaatcttctcgcgcaacgcgtcagtggtgatcattaactatccgct  
ggatgaccaggatgccattgctgtggaagctgcctgcactaatgttccggcgttattttcttgat  
gtctctgaccagacacccatcaacagtattattttctcccatgaagacgggtacgcgactgggcg  
tggagcatctggtcgcattgggtcaccagcaaatacgcgctgttagcggggccatttaagttctgt  
ctcggcgcgtctgcgtctggctggctggcataaatatctcactcgcaatcaaattcagccgata  
gcggaacgggaaggcgactggagtgccatgtccggttttcaacaaacctatgcaaatactgaatg  
agggcatcgttcccactgcgatgctgggttgccaacgatcagatggcgcgtgggcgcaatgcgcgc  
cattaccgagtcggggtgcgcgttgggtgcggatatctcggtagtgggatacgacgataccgaa  
gacagctcatgttatatcccggcgttaaccacctcaaacaggattttcgccctgctggggcaaa  
ccagcgtggaccgcttgctgcaactctctcagggccaggcgggtgaagggaatcagctgttgcc  
cgtctcactgggtgaaaagaaaaaccacctggcgcccaataacgcaaaccgcctctccccgcgcg  
ttggccgattcattaatgcagctggcacgacagggtttcccgactggaaagcgggcagtgagcgc  
aacgcaattaatgtaagttagctcactcattaggcacccgggatctcgaccgatgcccttgagag  
ccttcaaccacagtcagctccttccgggtgggcgcggggcatgactatcgtcgcgcgacttatgac  
tgtcttctttatcatgcaactcgtaggacagggtgccggcagcgcctctgggtcattttcggcgag  
gaccgctttcgtcggagcgcgacgatgatcgccctgtcgttgcggtatttcggaatcttgacg  
ccctcgtcgaagccttcgtcactgggtcccgcaccaaacgtttcggcgagaagcaggccattat  
cgccggcatggcggccgacgcgcgtgggttacgtcttgctggcggttcgcgacgcgaggctggatg  
gccttccccattatgattcttctcgttccggcgggcatcgggatgcccgcttgacaggccatgc  
tgtccaggcaggtagatgacgaccatcagggaacagcttcaaggatcgtcgcgcgctcttaccag  
cctaacttcgatcactggaccgctgatcgtcacggcgatttatgccgcctcggcgagcacatgg  
aacgggttggtcatggattgtaggcgccgcctataccttgtctgcctccccgcgttgctgcgcg  
gtgcatggagccggggccacctcgacctgaatggaagccggcgggcacctcgctaacggattcacc  
actccaagaattggagccaatcaattcttgcgagaaactgtgaatgcgcaaaccaacccttggc  
agaacatatccatcgcgtccgccatctccagcagccgcacgcggcgcatctcgggcagcgttgg  
gtcctggccacgggtgcgcgatgatcgtgctcctgtcgttgaggaccggctaggctggcgggggt  
tgcttactgggttagcagaatgaatcacgatacgcgagcgaacgtgaagcgactgctgctgca  
aaacgtctgcgacctgagcaacaacatgaatgggtcttcgggtttccgtgtttcgtaaagtctgga  
aacgcggaagtcagcgcctgcaccattatgttccggatctgcatcgcaggatgctgctggcta  
ccctgtggaacacctacatctgtattaacgaagcgtggcattgaccctgagtatttttctct  
gggtcccgcgcgatccataccgccagttgtttacctcacaacgttccagtaaccggggcatgttc  
atcatcagtaaccggtatcgtgagcatcctctctcgtttcatcggtatcattacccccatgaac  
agaaatcccccttacacggaggcatcagtgaccaaacaggaaaaaacgccttaacatggccc  
gctttatcagaagccagacattaacgcttctggagaaactcaacgagctggacgcggatgaaca  
ggcagacatctgtgaatcgttcacgaccacgctgatgagctttaccgcagctgctcgcgcgt  
ttcgggtgatgacgggtgaaaacctctgacacatgcagctcccggagacgggtcacagcttgtctgt  
aagcggatgccgggagcagacaagcccgtcagggcgcgctcagcgggtgttggcgggtgtcgggg  
cgacgcatgaccagtcacgtagcgatagcggagtgtatactggcttaactatgcggcatcag  
agcagattgtactgagagtgcaccatatatgcgggtgtgaaataccgcacagatgcgttaaggaga  
aaataccgcacagggcgtcttccgcttccctcgtcactgactcgtcgcgtcggtcgttcggc  
tgcggcgagcgggtatcagctcactcaaaggcggtaatacgggttatccacagaatcaggggataa  
cgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaaccgtaaaaaggccgcgttg  
ctggcggtttttccatagggtccgccccctgacgagcatcacaataatcgacgctcaagtcaga



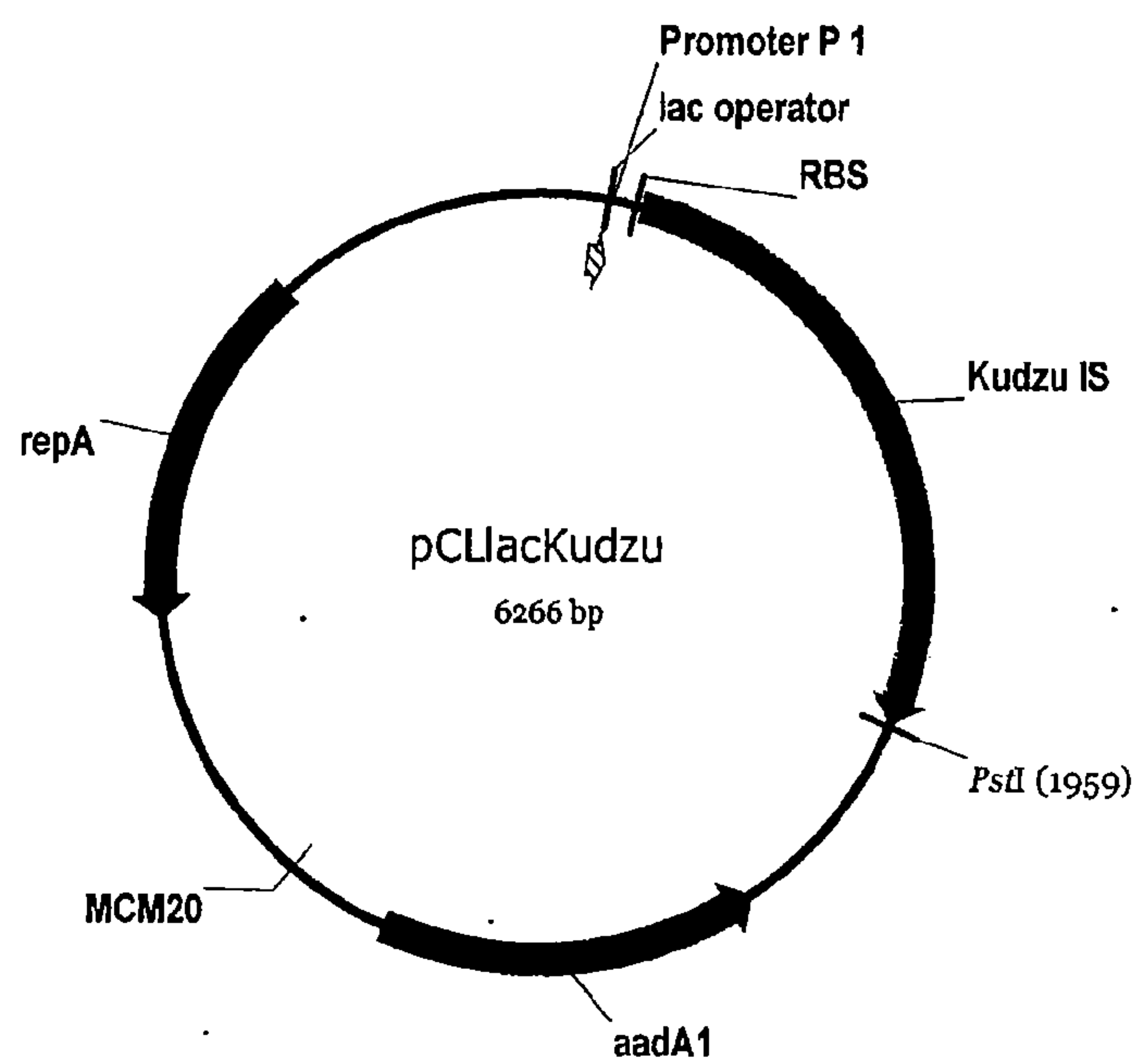
**Figure 5C**

ggtggcgaaaacccgacaggactataaagataaccaggcggtttccccctggaagctccctcgtgcg  
ctctcctggttccgacctgcccgttaccggatacctgtccgcctttctcccttcgggaagcgtg  
gcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgttcgctccaagctgg  
gctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttga  
gtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcaga  
gcgagggtatgtaggcgggtgctacagagttcttgaagtgggtggcctaactacggctacactagaa  
ggacagtattttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctc  
ttgatccggcaaacaaccaccgctggtagcgggtgggtttttttgtttgcaagcagcagattacg  
cgcagaaaaaaaggatctcaagaagatcctttgatctttttctacgggggtctgacgctcagtga  
acgaaaactcacgttaagggatttttggtcatgagattatcaaaaaggatcttcacctagatcct  
tttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttgggtctgacagt  
taccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgtttcatccatagttg  
cctgactccccgctcgtgtagataactacgatacgggaggggttaccatctggccccagtgctgc  
aatgataccgcgagacccacgctcaccgggtccagatttatcagcaataaaccagccagccgga  
agggccgagcgcagaagtgggtcctgcaactttatccgcctccatccagtctattaattgttgcc  
gggaagctagagtaagtagttcgccagttaatagtttgcgcaacggttggttgccattgctgcagg  
catcgtgggtgtcacgctcgtcgttttggtatggcttcattcagctccgggttcccaacgatcaagg  
cgagttacatgatcccccatggttgcaaaaaagcgggttagctccttcgggtcctccgatcgttg  
tcagaagtaagttggccgcagtggttatcactcatggttatggcagcactgcataattctcttac  
tgtcatgccatccgtaagatgctttttctgtgactgggtgagtactcaaccaagtcattctgagaa  
tagtgatgoggcgaccgagttgctcttgcccggcgtcaacacgggataataccgcgccacata  
gcagaactttaaaagtgtcatcattggaaaacggttcttcggggcgaaaactctcaaggatctt  
accgctggtgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatctttt  
actttcaccagcggtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataa  
gggcgacacggaaatggtgaatactcatactcttcctttttcaatattattgaagcatttatca  
gggttattgtctcatgagcggatacatatttgatgtatttagaaaaataaacaatataggggtt  
ccgcgcacattttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaa  
cctataaaaaataggcgtatcacgaggccctttcgtcttcaagaa

(SEQ ID NO:5)



10/167

**Figure 6**



11/167

**Figure 7A**

1-

cccgctcttactgtcgggaattcgcggttgcccgattcattaatgcagctggcacgacagggtttcc  
cgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcacccc  
caggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttc  
acacaggaaacagctatgaccatgattacgccaaagcttgtatcgattaaataaggaggaataaa  
ccatgtgtgcgacctcttctcaatttactcagattaccgagcataattcccgctcgttccgcaaa  
ctatcagccaaacctgtggaatttcgaattcctgcaatccctggagaacgacctgaaagtggaa  
aagctggaggagaaagcgaccaaaactggagggaagaagttcgctgcatgatcaaccgtgtagaca  
cccagccgctgtccctgctggagctgatcgacgatgtgcagcgcctgggtctgacctacaaatt  
tgaaaaagacatcattaagccctggaaaacatcgactgctggacgaaaacaaaaagaacaaa  
tctgacctgcacgcaaccgctctgtctttccgtctgctgcgtcagcacgggtttcgagggtttctc  
aggatgtttttgagcgtttcaaggataaagaagggtggtttcagcgggtgaactgaaagggtgacgt  
ccaaggcctgctgagcctgtatgaagcgtcttacctgggtttcgagggtgagaacctgctggag  
gaggcgcgtaccttttccatcacccacctgaagaacaacctgaaagaaggcattaataccaagg  
ttgcagaacaagtgcagccacgacctggaactgccatatcaccagcgtctgcacctctggaggc  
acgttgggttccctggataaatacgaaccgaaagaaccgcatcaccagcgtctgctggagctggcg  
aagctggatttttaacatggtacagacctgcaccagaaagagctgcaagatctgtcccgctgggt  
ggaccgagatgggcctggctagcaaaactggattttgtacgcgaccgacctgatggaagtttattt  
ctgggcactgggtatggcgccagacccgcagtttggtgaatgtcgcaaagctgttactaaaatg  
tttgggtctggtgacgatcatcgatgacgtgtatgacgtttatggcactctggacgaactgcaac  
tgttcaccgatgctgtagagcgtgggacgttaacgctattaacacctgccggactatatgaa  
actgtgtttcctggcactgtacaacaccttaacgacacgtcctattctattctgaaagagaaa  
ggtcataacaacctgtcctatctgacgaaaagctggcgtgaactgtgcaaagcctttctgcaag  
aggcgaaatgggtccaacaacaaaattatcccggctttctccaagtacctggaaaacgccagcgt  
ttcctcctccggtgtagcgtgctggcgccgctcttacttttccgtatgccagcagcaggaagac  
atctccgaccacgcgctgcgttccctgaccgacttccatgggtctggtgcgttctagctgcgtta  
tcttccgcctgtgcaacgatctggccacctctgcggcgagctggaacgtggcgagactaccaa  
ttctatcattagctacatgcacgaaaacgatgggtaccagcgaggaacaggcccgcgaagaactg  
cgtaaaactgatcgacgccgaatggaaaaagatgaatcgtgaacgcgttagcgactccacctgc  
tgccataaagcgttcatggaaatcgacgttaacatggcacgtgtttccactgcacctaccagta  
tggcgatggtctgggtcgcccagactacgcgactgaaaaccgcatcaaactgctgctgattgac  
cctttcccgattaaccagctgatgtatgtctaactgcaggctcactctagaggatccccgggtta  
ccgagctcgaattcactggccgctcgttttacaacgtcgtgactgggaaaacctggcggttacc  
aacttaatcgcccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcac  
cgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgccctgatgcggtattttctc  
cttacgcatctgtgcggtatttcacaccgcatatgggtgcaactctcagtacaatctgctctgatg  
ccgcatagtttaagccagccccgacacccgccaaacacccgctgacgagcttagtaaaagccctcgc  
tagatttttaatgcggatggttgcgattacttcgccaactattgcgataacaagaaaaagccagcc  
tttcatgatataatctcccaatttgtgtagggcttattatgcacgcttaaaaaataataaaagcag  
acttgacctgatagtttggctgtgagcaattatgtgcttagtgcatctaacgcttgagttaagc  
cgcgccgcgaagcggcgctgggcttgaacgaattgttagacattatgtccgactaccttggtga  
tctcgccctttcacgtagtggaacaaattcttccaactgatctgcgcgcgaggccaagcgatcttc  
ttcttgtccaagataagcctgtctagcttcaagtatgacgggctgatactgggcccggcaggcgc  
tccattgccagtcggcagcgacatccttcggcgcgattttgccgggttactgcgctgtaccaa  
tgccgggacaacgtaagcactacatttcgctcatcgccagcccagtcgggcccgcgagttccatag



12/167

**Figure 7B**

cgttaagggtttcatttagcgcctcaaatagatcctgttcaggaaccggatcaaagagttcctcc  
gccgctggacctaccaaggcaacgctatgttctcttgcttttgtcagcaagatagccagatcaa  
tgtcgatcgtggctggctcgaagatacctgcaagaatgtcattgcgctgccattctccaaattg  
cagttcgcgcttagctggataacgccacggaatgatgtcgtcgtgcacaacaatgggtgacttct  
acagcgcgggagaatctcgtctctccaggggaagccgaagtttccaaaagggtcgttgatcaaag  
ctcgcgcggttggttcatcaagccttacgggtcacccgtaaccagcaaatacaatatactgtgtgg  
cttcaggccgcatccactgaggagccgtacaaatgtacggccagcaacgctcggttcgagatgg  
cgctcgatgacgccaactacctctgatagttgagtcgatacttcggcgatcacccgcttccctca  
tgatgttttaactttgttttagggcgactgccctgctgcgtaacatcgttgctgctccataacat  
caaacatcgacccacggcgtaacgcgcttgctgcttggtatgcccagggcatagactgtacccca  
aaaaaacagtcataacaagccatgaaaaccgcccactgcgcggttaccaccgctgcgttcgggtca  
aggttctggaccagttgcgtgagcgcatacgcctacttgcttacagcttacgaaccgaacaggc  
ttatgtccactgggttcgtgccttcatccgtttccacgggtgtgcgtcacccggcaaccttgggc  
agcagcgaagtcgaggcatttctgtccttggttggaacgagcgcgaagggttccggctctccacgc  
atcgtcaggcattggcgcccttgctgttcttctacggcaagggtgctgtgcacggatctgccttg  
gcttcaggagatcggaagacctcgcccgctcgcggcgcttgccgggtggtgctgaccccgatgaa  
gtgggttcgcatectcggttttctggaaggcgagcatcgtttgttcgcccagcttctgtatggaa  
ggggcatgcggatcagtgagggtttgcaactgcgggtcaaggatctggatttcgatcacggcac  
gatcatcgtgcgggagggcaagggtccaaggatcgggccttgatgttacccgagagcttggca  
cccagcctgcgcgagcaggggaattaattcccacgggttttgctgcccgcaaacgggctgttct  
gggtgttgctagtttggttatcagaatcgcagatccgggttcagccggtttgccggctgaaagcgc  
tatttcttccagaattgccatgattttttcccccacgggagggcgctcactgggtcccgtgttgctg  
gcagctttgatcgataagcagcatcgccctgtttcaggctgtctatgtgtgactgttgagctgt  
aacaagttgtctcaggtgttcaatttcatgttctagttgctttgttttactggtttcacctgtt  
ctattagggtgttacatgctgttcatctgtttacattgtcgatctgttcatgggtgaacagctttga  
atgcacccaaaaactcgtaaaagctctgatgtatctatcttttttacaccgttttcatctgtgca  
tatggacagttttccctttgatatgtaacgggtgaacagttgttctacttttggttggttagtctt  
gatgcttcactgatagatacaagagccataagaacctcagatccttccgtatttagccagtatg  
ttctctagtgtggttcggtgtttttgcgtgagccatgagaacgaaccattgagatcatacttac  
tttgcatgtcactcaaaaattttgcctcaaaaactgggtgagctgaatttttgacagttaaagcatc  
gtgtagtgtttttcttagtccgttatgtaggttaggaatctgatgtaatgggtgttggtattttg  
tcaccattcatttttatctggttggttctcaagttcgggttacgagatccatttgtctatctagtt  
caacttggaataatcaacgtatcagtcgggcggcctcgcttatcaaccaccaatttcatattgct  
gtaagtgtttaaatctttacttattgggtttcaaaaaccattgggttaagccttttaaaactcatgg  
tagttattttcaagcattaacatgaacttaaattcatcaaggctaattctctatatttgccctgt  
gagttttcttttggttagttcttttaataaccactcataaatcctcatagagtatttggtttc  
aaaagacttaacatgttccagattatattttatgaatttttttaactggaaaagataaggcaat  
atctcttcaactaaaaactaattctaatttttcgcttgagaacttgccatagtttggtccactgga  
aaatctcaaagcctttaaccaaaggattcctgatttccacagttctcgtcatcagctctctgggt  
tgcttttagctaatacaccataagcattttccctactgatgttcatcatctgagcgtattgggtta  
taagtgaacgataccgtccgttctttcccttgtaggggtttcaatcgtggggttgagtagtgcca  
cacagcataaaattagcttgggttcatgctccgttaagtcatagcgactaatcgctagttcatt  
tgctttgaaaacaactaattcagacatacatctcaattgggtctaggtgatttttaatactatac  
caattgagatgggctagtcaatgataattactagtccttttcccttgagttgtgggtatctgta



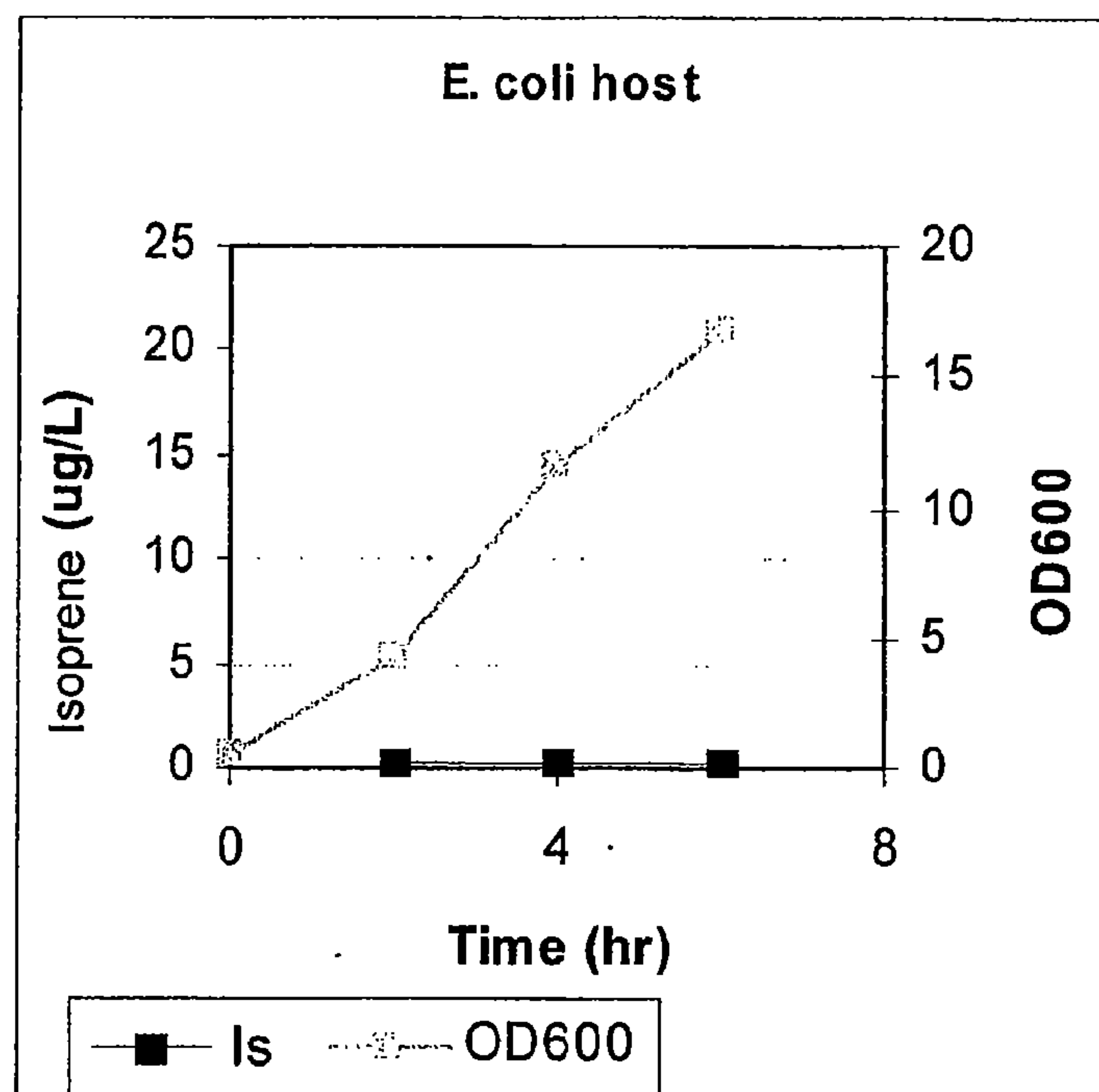
13/167

**Figure 7C**

Aattctgctagacctttgctggaaaacttgtaaattctgctagaccctctgtaaattccgctag  
acctttgtgtgtttttttttgtttatattcaagtggttataatttatagaataaagaaagaataa  
aaaaagataaaaagaatagatcccagccctgtgtataactcactacttttagtcagttccgcagt  
attacaaaaggatgtcgcaaacgctgtttgctcctctacaaaacagaccttaaaaccctaaagg  
cttaagtagcacccctcgcaagctcggggcaaatacgctgaatattccttttgtctccgaccatcag  
gcacctgagtcgctgtcttttttcgtgacattcagttcgctgcgctcacggctctggcagtgat  
gggggtaaatggcactacaggcgcccttttatggattcatgcaaggaaactaccataatacaag  
aaaagcccgtcacgggcttctcagggcggttttatggcgggtctgctatgtggtgctatctgact  
ttttgctgttcagcagttcctgcctctgattttccagtctgaccacttcggattatcccgtga  
caggtcattcagactggctaatagcacccagtaaggcagcggtatcatcaacaggctta  
(SEQ ID NO:7)

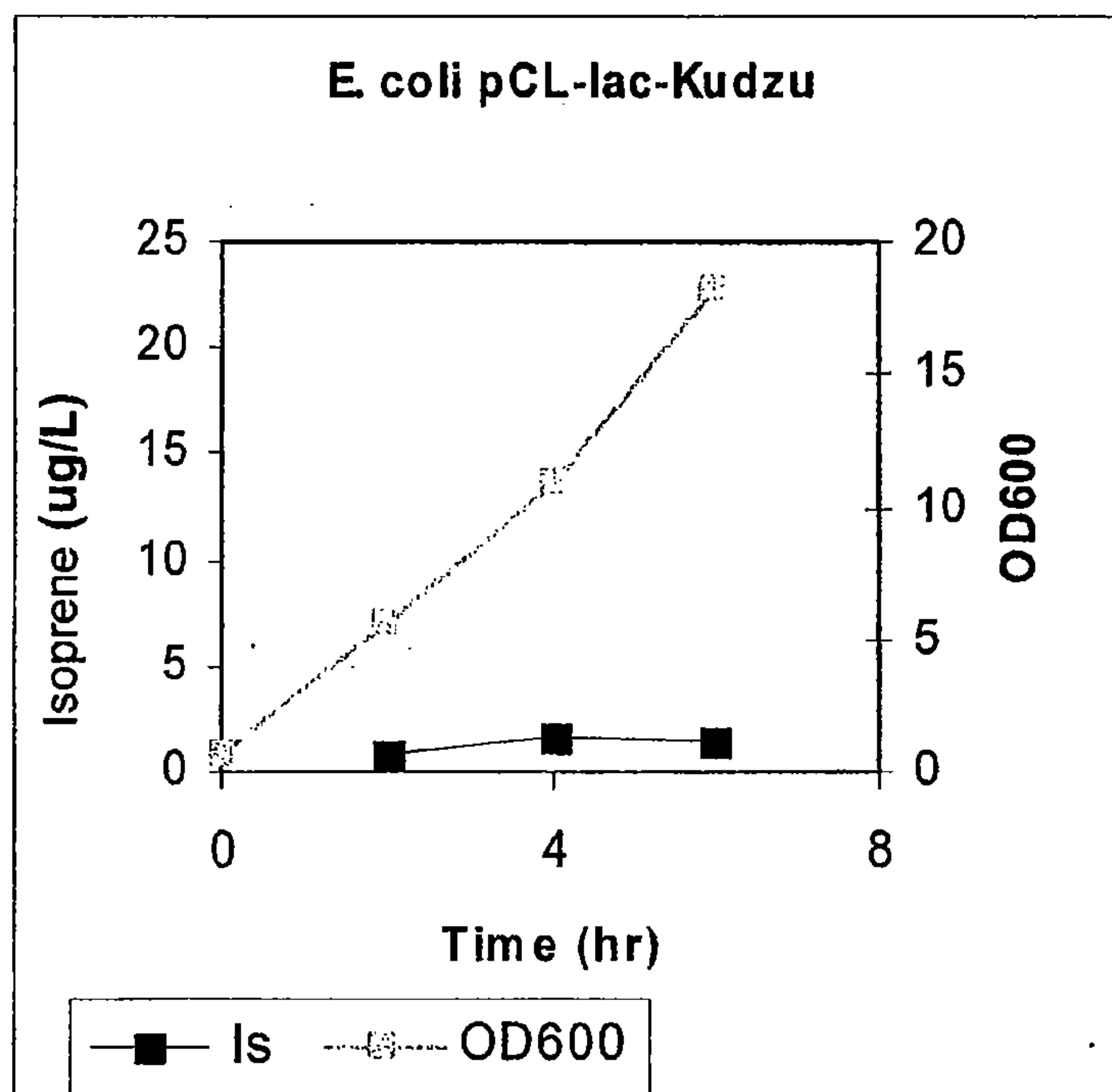


14/167

**Figure 8A**

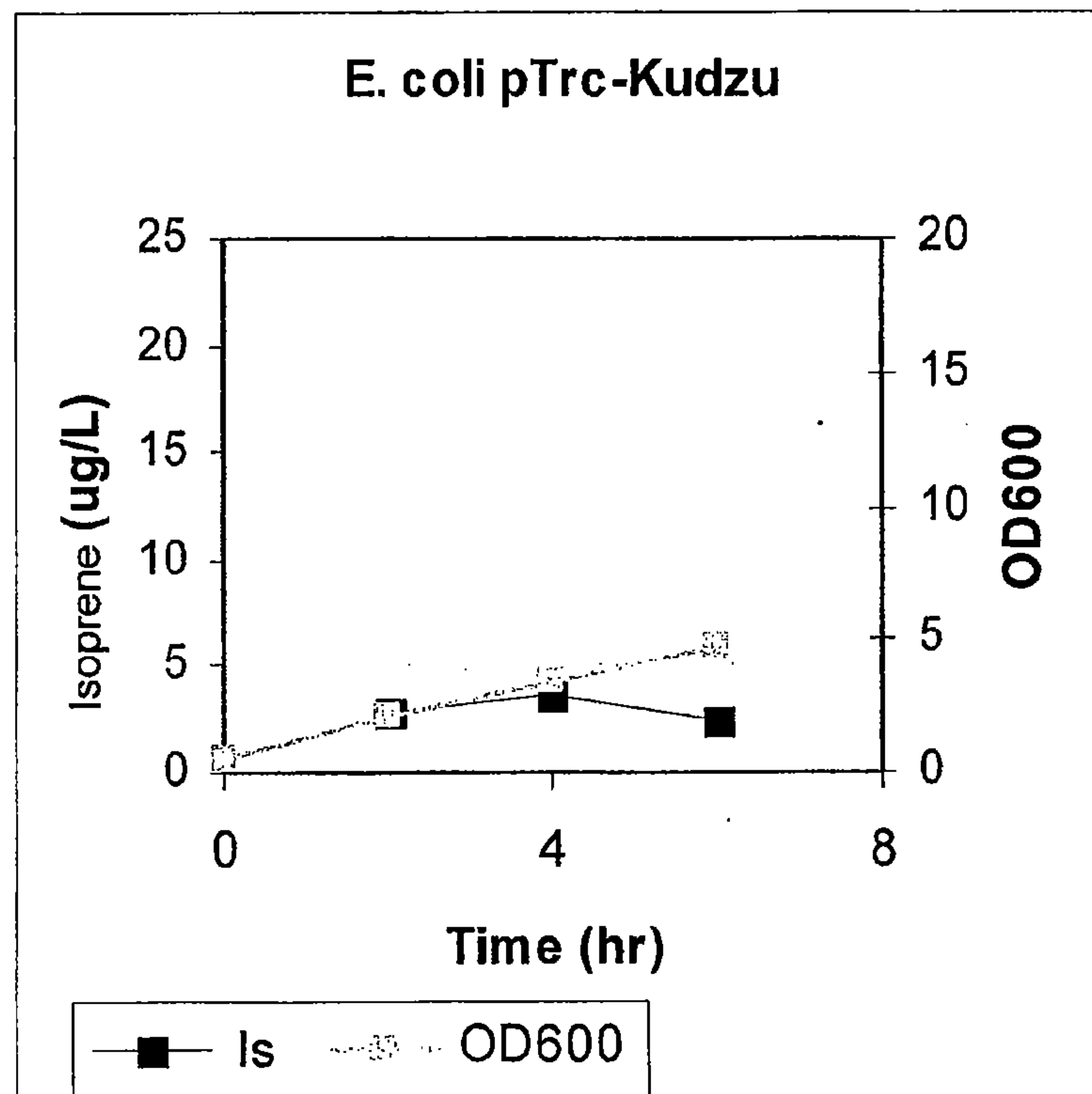


15/167

**Figure 8B**

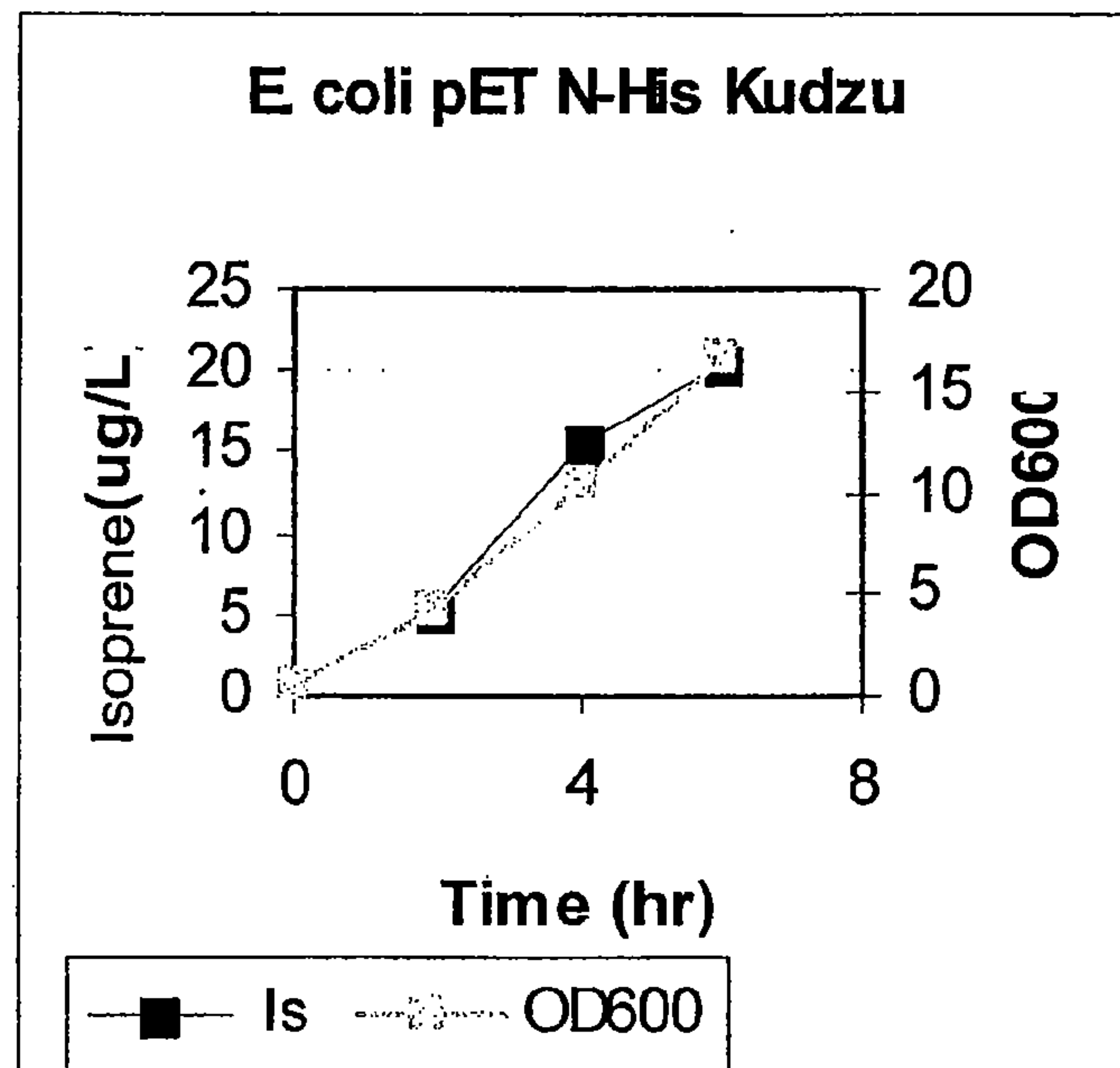


16/167

**Figure 8C**



17/167

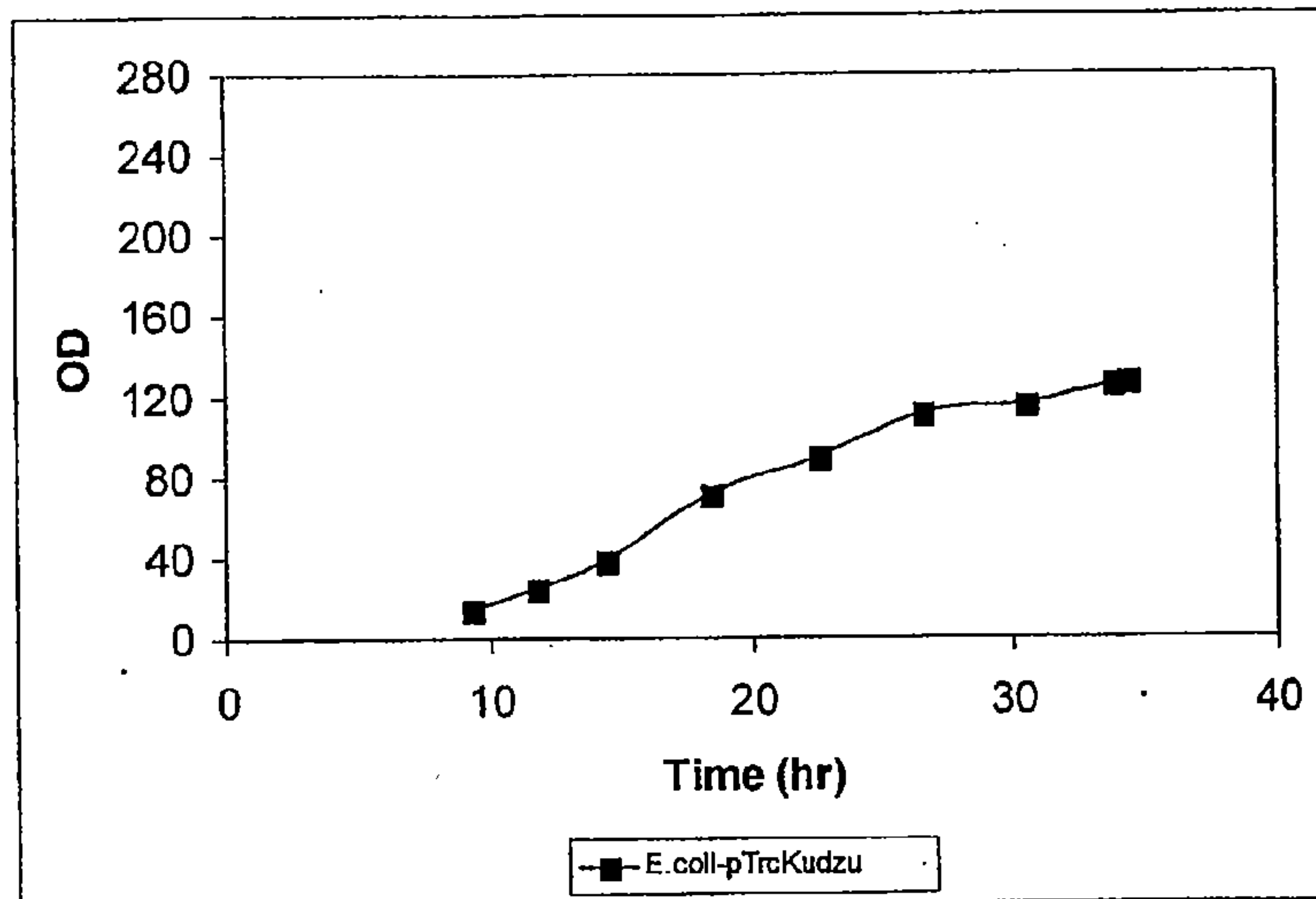
**Figure 8D**



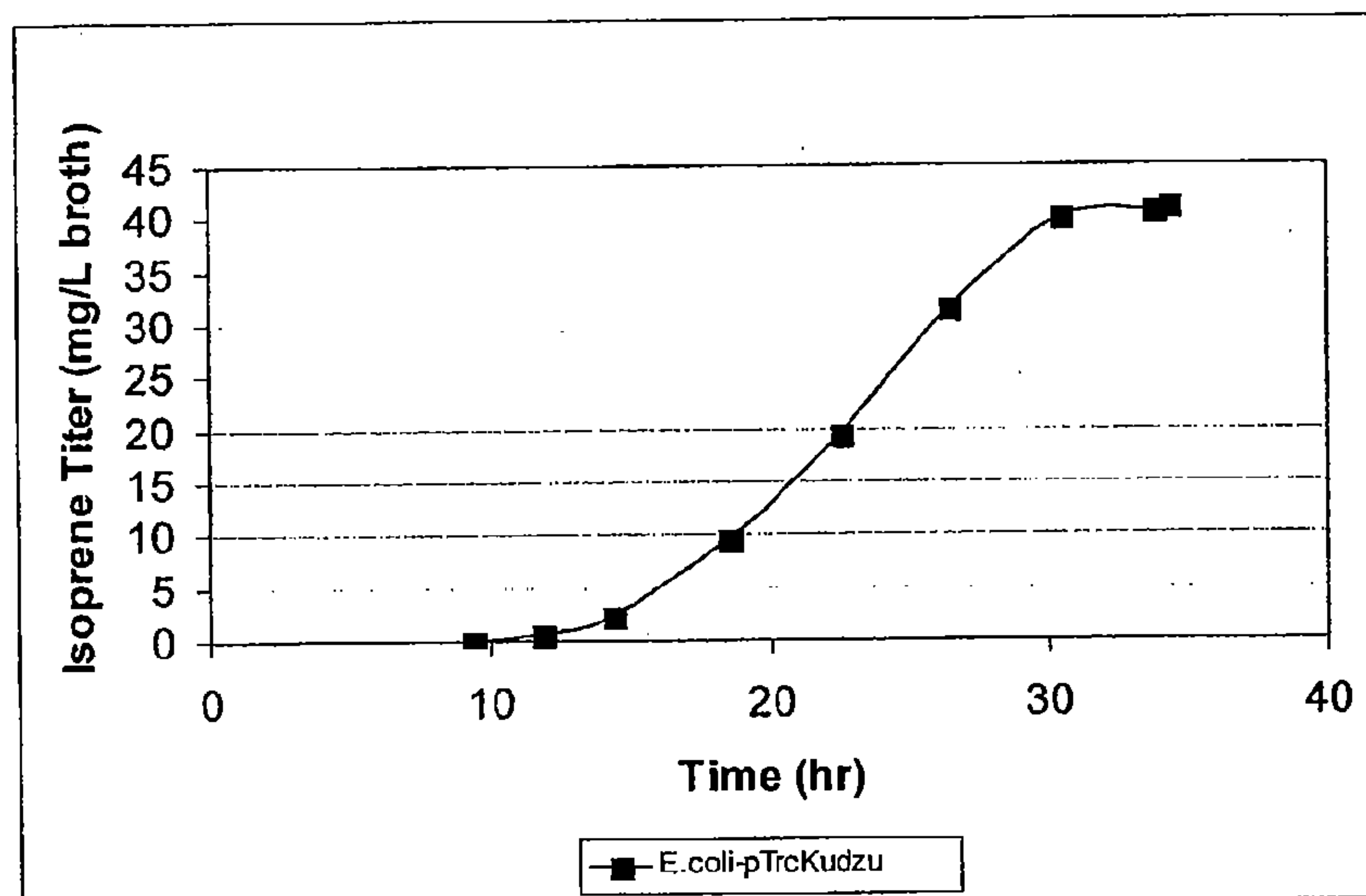
18/167

**Figure 9**

A.



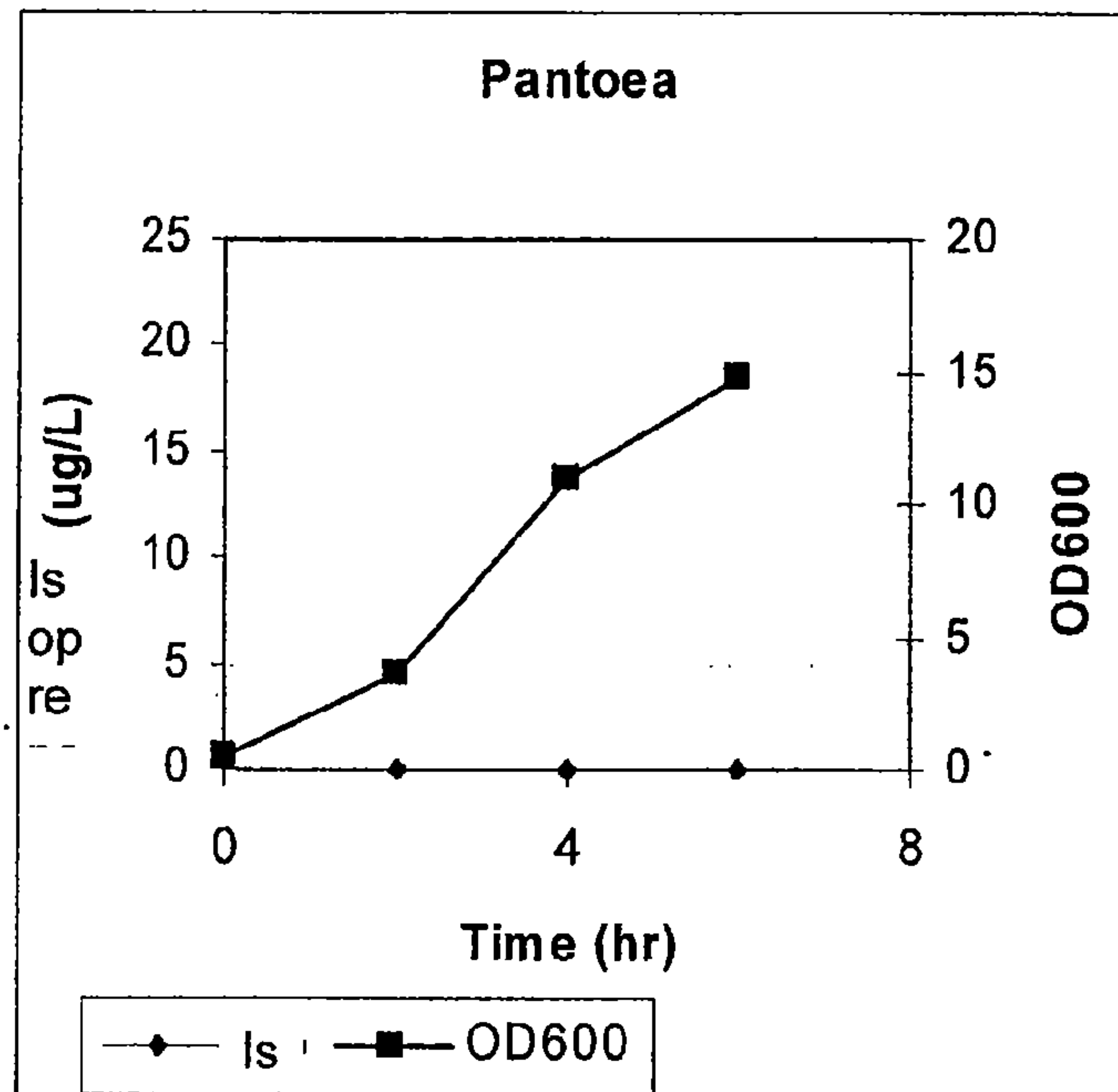
B.





19/167

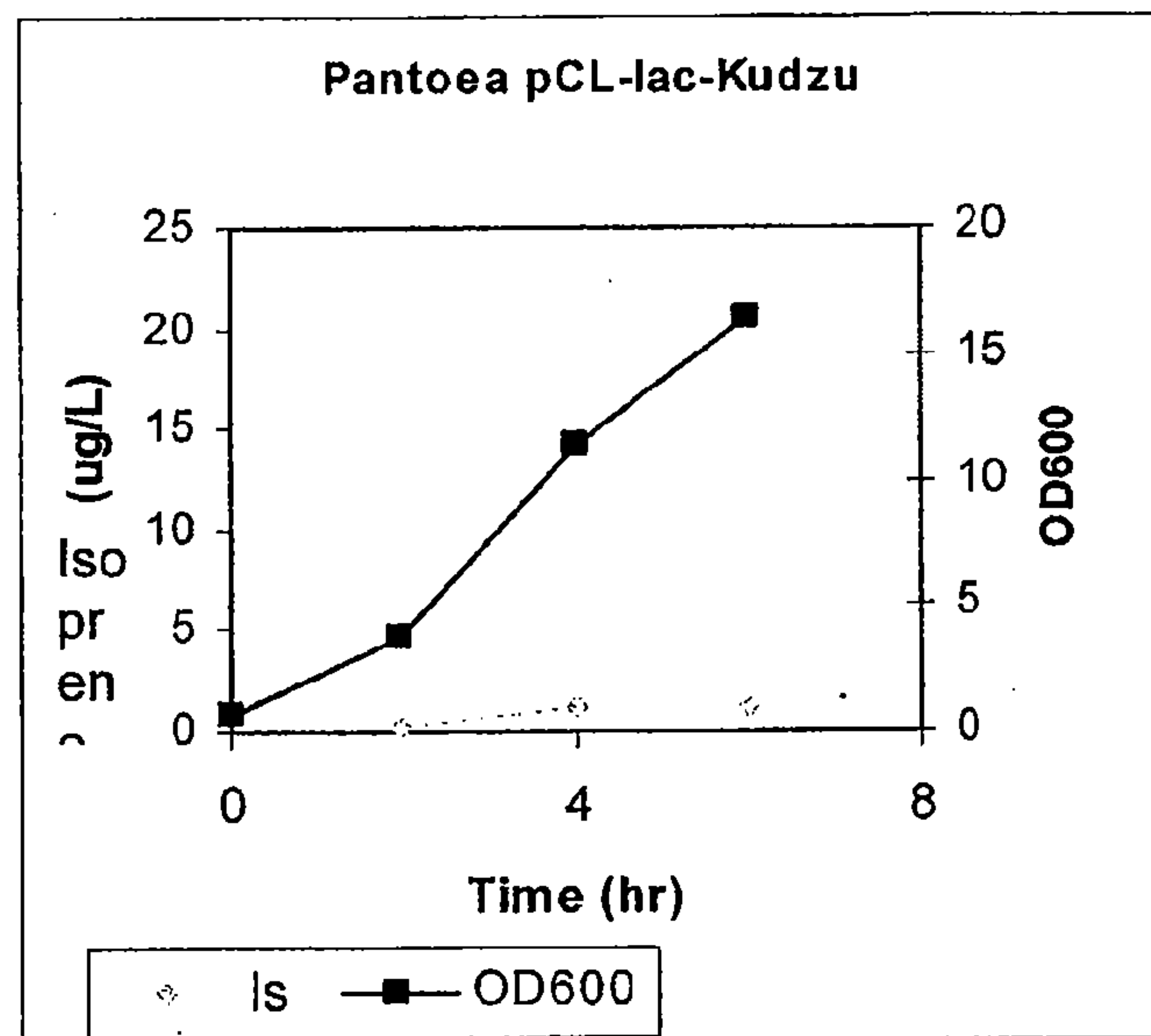
Figure 10A





20/167

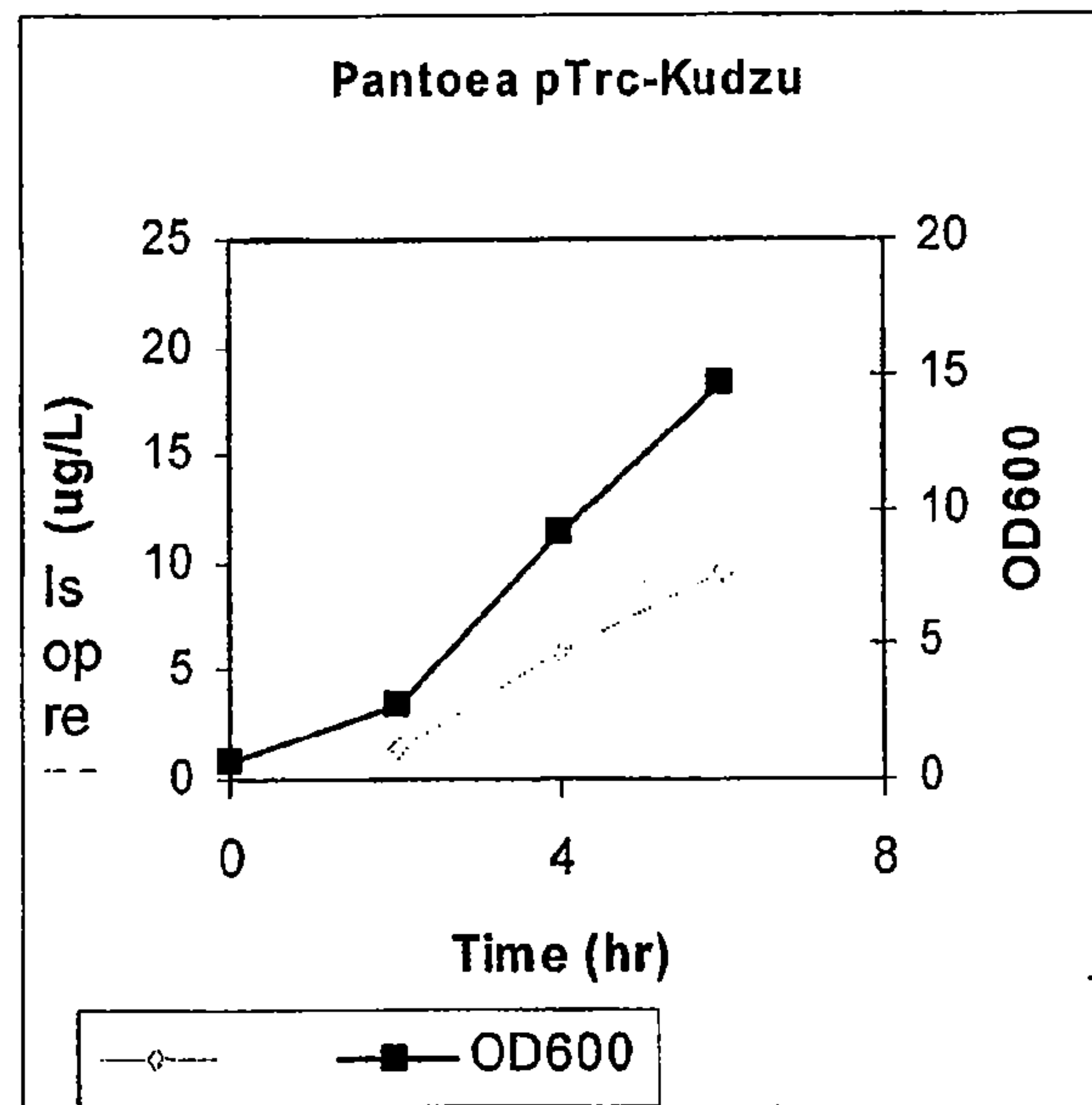
Figure 10B





21/167

Figure 10C





22/167

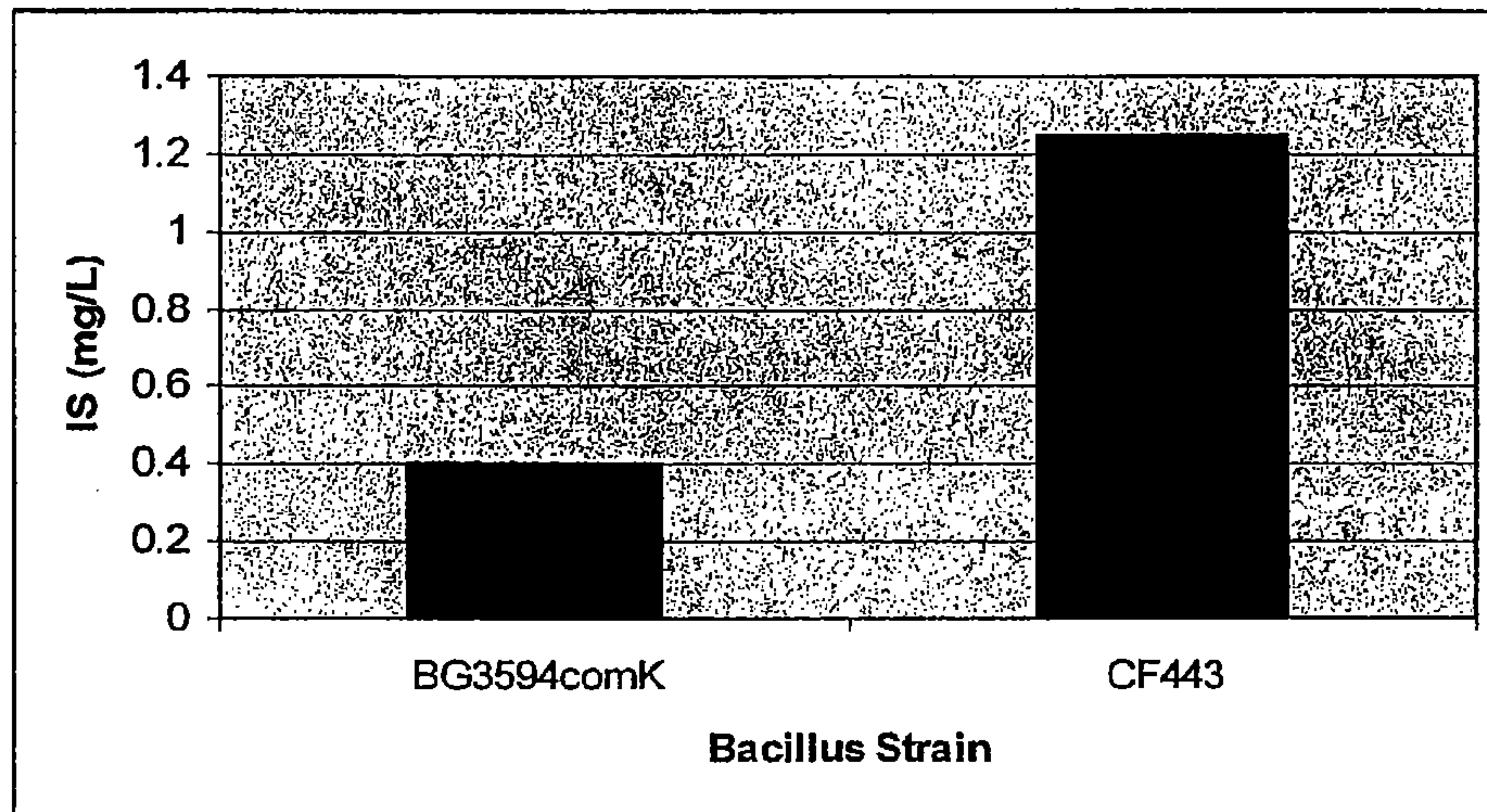
**Figure 11**



**Figure 12A**

1-

gaattgctccatttttcttctgctatcaaaataacagactcgtgattttccaaacgagctttcaa  
aaaagcctctgccccttgcaaatacgatgcctgtctataaaattcccgatattgggttaaacagc  
ggcgcaatggcgccgcacatctgatgtctttgcttggcgaatgttcacatcttatttcttcctccct  
ctcaataattttttcattctatcccttttctgtaaagtattttttcagaatacttttatcatc  
atgctttgaaaaaatatcacgataatatccattgttctcacggaagcacacgcaggtcatttga  
acgaattttttcgacaggaatttgccgggactcaggagcatttaacctaaaaaagcatgacatt  
tcagcataatgaacatttactcatgtctattttcggttcttttctgtatgaaaatagttatttgcg  
agtctctacggaaatagcgagagatgatatacctaataagagataaaatcatctcaaaaaaatg  
ggtctactaaaaatattattccatctattacaataaattcacagaatagtcttttaagtaagtct  
actctgaatttttttaaaaggagagggtaaaagagtgtgtgcgacctcttctcaatttactcaga  
ttaccgagcataattcccgctcgttccgcaaactatcagccaaacctgtggaatttcgaattcct  
gcaatccctggagaacgacctgaaagtggaaaagctggaggagaaagcgaccaactggaggaa  
gaagttcgctgcatgatcaaccgtgtagacacccagccgctgtccctgctggagctgatcgacg  
atgtgcagcgccctgggtctgacctacaaatttgaaaaagacatcattaaagccctggaaaacat  
cgtactgctggacgaaaacaaaaagaacaaatctgacctgcacgcaaccgctctgtctttccgt  
ctgctgcgtcagcacggttttcgaggtttctcaggatgtttttgagcggtttcaaggataaagaag  
gtggttttcagcggtgaactgaaaggtgacgtccaaggcctgctgagcctgtatgaagcgtctta  
cctgggttttcgaggggtgagaacctgctggaggaggcggtaccttttccatcacccacctgaag  
aacaacctgaaagaaggcatttaataccaagggtgcagaacaagtgcagccacgcccctggaactgc  
catatcacccagcgtctgcacctctggaggcacgttggttccctggataaatacgaaccgaaaga  
accgcatcacccagctgctgctggagctggcgaagctggatttttaacatggtacagacctgcac  
cagaaagagctgcaagatctgtcccgtggtggaccgagatgggcctggctagcaaaactggatt  
ttgtacgcgaccgcctgatggaagtttatttctgggcactgggtatggcgccagaccgcagtt  
tggtgaatgtcgcaaagctgttactaaaatgtttggtctggtgacgatcatcgatgacgtgtat  
gacgtttatggcactctggacgaactgcaactgttcaccgatgctgttagagcgctgggacgtta  
acgctattaacacctgcccggactatatgaaactgtgtttcctggcactgtacaacacctgtaa  
cgacacgtcctatttctatttctgaaagagaaaggtcataacaacctgtcctatctgacgaaaagc  
tggcgtgaactgtgcaaagcctttctgcaagaggcgaaatgggtccaacaacaaaattatcccg  
ctttctccaagtacctggaaaacgccagcgtttcctcctccggtgtagcgctgctggcgccgtc  
ttacttttccgtatgccagcagcaggaagacatctccgaccacgcgctgcgttccctgaccgac  
ttccatgggtctggtgcgttctagctgcgttatcttccgcctgtgcaacgatctggccacctctg  
cggcgagctggaacgtggcgagactaccaattctatcattagctacatgcacgaaaacgatgg  
taccagcgaggaacaggcccgcgaagaactgcgttaaactgatcgacgcgcaatggaaaaagatg  
aatcgtgaacgcgttagcgactccacctgctgcctaaagcgttcattggaaatcgaggttaaca  
tggcacgtgtttccactgcacctaccagtatggcgatggtctgggtcgcccagactacgcgac  
tgaaaacccgcatcaaactgctgctgattgaccttttcccgatttaaccagctgatgtatgtctaa  
aaaaaacccgaccttgccccgcgggttttttattatttttcttccctccgcatgttcaatccgct  
ccataatcgacggatggctccctctgaaaatttttaacgagaaacggcggttgacctggctcag  
tcccgtaacggccaagtcctgaaacgtctcaatcgccgcttcccggtttccggctcagctcaatg  
ccgtaacggctcggcggtgttttctgataccgggagacggcattcgtaatcggatcctctagag  
tcgacctgcaggcatgcaagctttgcctcgcgcttccggtgatgacggtgaaaacctctgaca  
catgcagctcccgagacgggtcacagcttgtctgtaagcggatgccgggagcagacaagccgct  
cagggcgcgctcagcggtgttggtggcggtgtcggggagcagccatgacctcagctagcgata



24/167

**Figure 12B**

gcggagtgtataactggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatg  
cgggtgtgaaataaccgcacagatgcgtaaggagaaaaataaccgcatcaggcgctcttccgcttcct  
cgctcactgactcgctgcgctcggtcggttcggctgcggcgagcggtatcagctcactcaaaggc  
ggtaatacgggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccag  
caaaaggccaggaaccgtaaaaaaggccgcttgctggcggtttttccataggctccgccccctg  
acgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagata  
ccaggcggtttccccctggaagctccctcgctgcgctctctgttccgaccctgccgcttacccgga  
tacctgtccgcctttctcccttcgggaagcggtggcgcttttctcaatgctcacgctgtagggtatc  
tcagttcggtgtagggtcggttcgctccaagctgggctgtgtgcacgaaccccccggttcagccga  
ccgctgcgccttatccggtaactatcgctcttgagtccaacccggtaagacacgacttatcgcca  
ctggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttct  
tgaagtgggtggcctaactacgggtacactagaaggacagtatttgggtatctgcgctctgctgaa  
gccagttaccttcggaaaaagagttggtagctcttgatccggcaaacacaccacgctggtagc  
ggtaggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctt  
tgatcttttctacgggggtctgacgctcagtggaacgaaaactcacgttaagggatttttggtcat  
gagattatcaaaaaggatcgaaagtcgggttcagaaaaagaaggatatggatctggagctgtaata  
taaaaaaccttcttcaactaacgggggcaggttagtgacattagaaaaccgactgtaaaaagtaca  
gtcggcattatctcatattataaaaagccagtcattaggcctatctgacaattcctgaatagagt  
tcataaacaatcctgcatgataaccatcacaaacagaatgatgtacctgtaaagatagcggtaa  
atataattgaattacctttattaatgaattttcctgctgtaataatgggtagaaggtaattacta  
ttattattgatattttaagttaaacccagtaaatgaagtcctatggaataatagaaagagaaaaag  
cattttcaggtataggtgttttgggaaacaatttaaaagaaccattatatttctctacatcaga  
aaggtataaatcataaaactccttgaagtcattctttacaggagtccaaataaccagagaatgtt  
ttagatacaccatcaaaaattgtataaagtggtcttaacttatcccaataacctaactctccgt  
cgctattgtaaccagttctaaaagctgtatttgagtttatcaccttgctcactaagaaaataaa  
tgcagggtataaatttatatccttcttggttttatgtttcggtataaaacactaatatcaatttct  
gtggttataactaaaagtcgtttgttggttcaaataatgattaaatatctcttttctcttccaat  
tgtctaaatcaatttttattaaagttcatttgatagcctcctaaatttttatctaaagtgaatt  
taggaggcttacttgctgctttcttcattagaatcaatccttttttaaagtcattattactgt  
aacataaatatataatttttaaaaatatccacttttatccaattttcggtttgttgaactaatgggt  
gcttttagttgaagaataaagaccacattaaaaaatgtggtccttttggtgttttttttaaaggattt  
gagcgtacgcgaaaaatccttttctttctttcttatcttgataataagggttaactattgccggt  
tgtccattcatggctgaactctgcttcctctgttgacatgacacacatcatctcaatatccgaa  
tagggcccatcagctctgacgaccaagagagccataaacaccaatagccttaacatcatcccat  
atttatccaatattcggttccttaatttcattgaacaatcttcattctttctctctagtcattat  
tattgggtccattcactattctcattcccttttcagataatttttagatttgcttttctaaataag  
aatatttggagagcaccgttcttattcagctattaataactcgtcttcctaagcatccttcaat  
ccttttaataacaattatagcatctaattctcaacaaactggcccggtttgttgaactactcttt  
aataaaataaattttccggttcccaattccacattgcaataatagaaaatccatcttcatcggt  
ttttcgctcatcatctgtatgaatcaaatcgcttcttctgtgtcatcaaggtttaattttttat  
gtatttcttttaacaaaccaccataggagattaaccttttacgggtgtaaaccttcctccaaatc  
agacaaacggtttcaaattcttttcttcatcatcggtcataaaatccgtatcctttacaggatat  
tttgcagtttcgtcaattgccgattgtatatccgatttatatttttttcggtcgaatcattt  
gaacttttacatttggtatcatagtctaatttcattgcctttttccaaaattgaatccattgttt



25/167

**Figure 12C**

ttgattcacgtagtttttctgttatttctaaaataagttgggtccacacataccattacatgcatg  
tgctgattataagaattatctttattattttattgtcacatccgttgacgcataaaaccaacaa  
gatttttattaattttttttatatattgcatcattcggcgaaatccttgagccatatctgtcaaact  
cttatttaattcttcgccatcataaacatttttaactgttaatgtgagaaacaaccaacgaact  
gttggcttttgtttaataacttcagcaacaaccttttgtgactgaatgccatgtttcattgctc  
tcctccagttgcacattggacaaagcctggatttgcaaaaccacactcgataccactttctttc  
gcctgtttcacgattttgtttatactctaataatttcagcacaatcttttactctttcagccttt  
ttaaattcaagaatatgcagaagttcaaagtaatcaacattagcgaattttcttttctctccatg  
gtctcacttttccactttttgtcttgtccactaaaacccttgatttttcatctgaataaatgct  
actattaggacacataatattaaaagaaacccccatctatttagttatttgttttagtcacttat  
aactttaacagatgggggtttttctgtgcaaccaattttaagggttttcaatactttaaaacaca  
tacataccaacacttcaacgcacctttcagcaactaaaataaaaatgacgttatttctatatgt  
atcaagataagaaagaacaagttcaaaaccatcaaaaaaagacaccttttcagggtgcttttttt  
attttataaaactcattccctgatctcgacttcgttctttttttacctctcgggttatgagttagt  
tcaaattcgttcttttttaggttctaaatcgtgtttttcttgggaattgtgctgttttatccttta  
ccttgtctacaaaccccttaaaaacgttttttaaggccttttaagccgtctgtacgttccttaag

(SEQ ID NO:57)



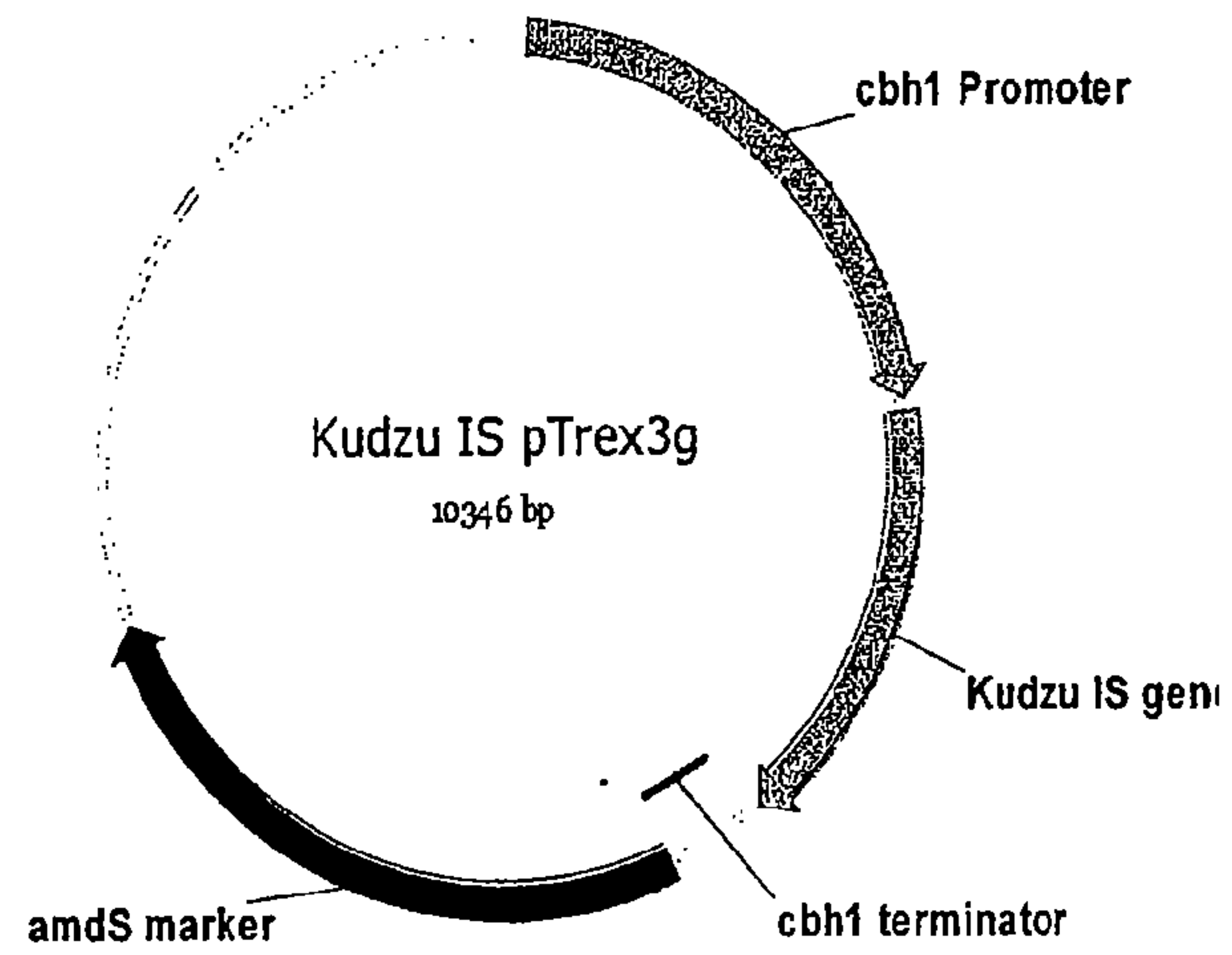
**Figure 13**

ATGTGTGCAACCTCCTCCCAGTTTACTCAGATTACCGAGCATAATTCTCGACGATCTGCTAACT  
ACCGAGCCGAACCTTTGGAACCTTTGAGTTTCTCCAGTCTCTCGAAAATGACCTGAAGGTGGAAAA  
GCTCGAGGAGAAGGCGACCAAACCTCGAGGAGGAGGTGCGATGTATGATCAACAGAGTTGACACC  
CAACCCCTGTCTTTGCTGGAGCTGATCGACGATGTGCAGCGGTTGGGTTTGACTTATAAATTCG  
AGAAGGACATTATCAAGGCACTGGAGAACATTGTGCTCCTCGACGAGAACAAAGAACAAGTC  
TGATCTTCACGCTACCGCTCTCTCTTTCCGACTTCTTCGACAACACGGCTTCGAGGTGTCGCAG  
GACGTCTTCGAGAGATTTAAGGACAAGGAGGGGAGGATTTAGCGGCGAGCTGAAGGGAGACGTTT  
AGGGTCTTCTCTCCTTGACGAGGCGTCCTACCTGGGATTCGAGGGGAGAGAACCTCCTGGAGGA  
AGCTCGTACATTTTCCATCACTCACCTTAAGAATAACCTTAAGGAGGGAATTAACACCAAGGTG  
GCCGAGCAGGTTTCTCACGCCCTGGAGCTCCCCTACCACCAACGGCTCCATAGACTGGAGGCTC  
GTTGGTTCCTGGACAAATATGAGCCAAAGGAGCCTCATCATCAGTTGCTGTTGGAGTTGGCCAA  
GCTGGACTTCAATATGGTTCAGACGCTGCACCAAAGGAGTTGCAGGACCTGTCTCGATGGTGG  
ACCGAGATGGGATTGGCCTCGAAGCTGGATTTTGTCCGTGACCGACTTATGGAGGTCTATTTTT  
GGGCCCTTGGAATGGCGCCTGACCCCCAGTTCCGGAGAGTGCCGGAAGGCGGTGACGAAGATGTT  
CGGTCTTGTGACTATCATCGACGACGTCTACGATGTCTACGGCACACTCGACGAGTTGCAGCTG  
TTCCTGACGCCGTGAGCGATGGGATGTGAACGCCATTAATACTCTCCCTGACTATATGAAGC  
TGTGCTTCCTGGCTCTGTACAACACTGTCAACGATACCTCGTACTCTATCCTCAAGGAGAAGGG  
ACACAACAATCTCTCCTACTTGACCAAATCCTGGCGAGAACTGTGCAAGGCTTTTCTGCAGGAG  
GCTAAATGGTCCAATAACAAGATCATTCCTGCTTTTTCTAAATACCTGGAAAATGCCTCGGTGT  
CGAGCTCTGGCGTCGCCCTTCTGGCCCCCTTCTACTTCTCCGTCTGCCAGCAGCAGGAGGATAT  
TTCCGATCATGCTCTTAGATCGCTGACCGATTTTCACGGCCTCGTGCGATCTTCCTGCGTGATT  
TTTCGGTTGTGTAATGACCTTGCGACCTCTGCTGCTGAGCTGGAACGAGGCGAGACTACAAATT  
CCATTATTTCTTACATGCACGAAAACGATGGAACATCTGAAGAACAGGCTAGAGAGGAACTGCG  
AAAGTTGATCGACGCCGAGTGGAAGAAGATGAACAGAGAGCGGGTGTCCGACTCTACCCTGCTT  
CCCAAGGCCTTCATGGAGATCGCCGTGAACATGGCTCGAGTTTCCCATTGTACTTACCAGTACG  
GTGACGGCCTGGGTGCTCCGGACTACGCTACAGAGAACCGAATCAAGCTGCTGCTCATCGACCC  
CTTCCCTATCAACCAATTGATGTACGTGTAA

(SEQ ID NO:8)



27/167

**Figure 14**



28/167

**Figure 15A**

```
1  TCGACCGGTG AGAAGAACAG CATCGGGACA AGGGAAGGAA GAACAAAGAC AAAGAAAACA
61 AAAGAAAGCA ATTGAAAACA AAACAAAACA ATTTTCATTC CTTCTCTTAT CATTCCCTTTT
121 CTTTTCTTTT CTCTCATTC ACGCACTCCA TCGTATCCGT ATTCCTCTTA TTTTTTCTCT
181 TTCTCTATAT CCATTTCTTT CTCTCTAGGT GTGTCTCTC TCTCTCTTCA ATTTCTCTAC
241 TCCGCATTCC AACGCATCCT TCCCCCAACC TCCCATTTC TCCTTACGGC CCGATAGCGA
301 TCGTCTTTCC CTCGCTATCA CTCGCTACCG GCCCCTCCTC TGCACCGTAA CCTCCTACGT
361 ATTTACCATA TCATAAAGTT TTTTCCGACG CTTATCGCTG ACCCCTGTC GCCCTCCTAT
421 TGGCTTCCGG ATTATCTTCT TGTCCATAAG GTGATCCATG CTTCTCTGAG ATTCCCGAAA
481 TGTGTCCACT TTGGCGGGGA ATCATTCCAT CCACTTCTTT CTCTCTCGCT TTCCTCATTC
541 GGCCTCCTCC TTCCGCGTCT CATTGGTCTT CCGCTCCGTT TTTGCTTTGC CGATGTTACT
601 TGGGGAGAGG TGGGATAATC CTTTCGCAAA AACTCGGTTT GACGCCTCC ATGGTATAAA
661 TAGTGGGTGG TGGACAGGTG CCTTCGCTTT TCTTTAAGCA AGAGAATCCC ATTGTCTTGA
721 CTATCACGAA TTCACATACA TTATGAAGAT CACCGCTGTC ATTGCCCTTT TATTCTCACT
781 TGCTGCTGCC TCACCTATTC CAGTTGCCGA TCCTGGTGTG GTTTCAGTTA GCAAGTCATA
841 TGCTGATTTT CTTCTGTTT ACCAAAGTTG GAACACTTTT GCTAATCCTG ATAGACCCAA
901 CCTTAAGAAG AGAAATGATA CACCTGCAAG TGGATATCAA GTTGAAAAAG TCGTAATTTT
961 GTCACGTCAC GGTGTTAGGG CCCCTACAAA AATGACTCAA ACCATGCGTG ATGTCACTCC
1021 TAATACATGG CCAGAATGGC CCGTTAAATT AGGATATATT ACACCAAGAG GTGAACACTT
1081 GATATCACTT ATGGGCGGTT TTTACCGTCA AAAATTCCAG CAACAAGGAA TCCTTTCTCA
1141 GGGCTCCTGT CCTACTCCTA ACTCCATATA TGTCTGGGCT GACGTGATC AGCGTACTTT
1201 AAAAAGTGGT GAAGCATTC CTGCTGGTTT GGCACCACAA TGTGGCTTGA CAATTCATCA
1261 CCAACAAAAT CTTGAGAAAG CTGATCCTCT TTTTCATCCC GTTAAAGCTG GAACCTGCTC
1321 TATGGATAAA ACTCAAGTTC AACAAGCTGT TGAGAAGGAG GCACAACTC CTATAGATAA
1381 TTTGAATCAA CATTACATCC CCTTTTTCAG TTTAATGAAT ACAACATTAA ATTTTAGTAC
1441 TTCTGCCTGG TGCCAAAAC ACTCTGCTGA TAAATCCTGT GACCTAGGTT TATCCATGCC
1501 TTCTAAATTG TCCATAAAAG ATAATGGTAA CAAGGTCGCA TTGGATGGAG CTATTGGTCT
1561 ATCCTCTACT TTGGCCGAGA TTTTCTTCT TGAATATGCT CAAGGCATGC CTCAAGCTGC
1621 TTGGGGTAAC ATCCACTCAG AGCAAGAGTG GGCTTCCTTG CTAAAGTTGC ATAATGTTCA
1681 ATTCGATTTG ATGGCCCGAA CACCTTATAT TGCTCGACAT AACGGTACTC CTTTATTGCA
1741 AGCTATATCA AATGCCCTTA ATCCCAACGC CACTGAATCA AAACCTCCAG ATATTTTACC
1801 TGATAACAAA ATATTGTTCA TTGCAGGTCA TGACACAAAT ATTGCTAATA TAGCCGGCAT
1861 GTTAAATATG CGTTGGACAT TACCAGGTCA ACCAGATAAT ACTCCTCCAG GTGGTGGCCT
1921 AGTATTTGAA CGTCTTGCTG ATAAAAGTGG AAAACAATAT GTTTCTGTAT CTATGGTTTA
1981 TCAAACACTA GAACAACTTC GATCACAGAC TCCCCTTTCT CTAAATCAGC CTGCCGGATC
2041 TGTTCAACTT AAAATTCCAG GTTGCAATGA TCAAACAGCC GAGGGTTACT GTCCTCTTTC
2101 CACTTTTACA AGAGTTGTTT CCAATCTGT TGAACCTGGA TGCCAACCTC AATAATGAGG
2161 ATCCAAGTAA GGAATGAGA ATGTGATCCA CTTTAAATTC CTAATGAATA CATGCCTATA
2221 GTTCTTTTCT TTTGTTCTTT ATGTCGTTTT TCGATGGTAC GGCCGTTGTC AATCTCAGTT
2281 TGTGTGCTTG GTTGAGCTT GGTTCAAAT CTGTTTCTCT CATGAATCTT TTACCATTTT
2341 ACCACACGTT TATACCATTC TCTCATAGAA TCTTCATCAA ACCATCTCGG GGTAGAGTG
2401 GAAAGAAAGT CTTGTTCTTT TATTTCTTTT TTTCCATCTT CAAGGCTTTT CTTTTCTTCC
2461 TCCTCCTCGT TCATCTTGAG GTTTGACGTG TCTGTTTGA ATTTTGAGCT GTTGAGCAT
2521 CTTATTTTTT GTTTTGCGAA AACGAAGCGC TTTACTCTCT TCATCAGTTG GACGATTGTA
2581 CCTTTGAAAA CCAACTACTT TTGCATGTTT TGTATAGAAA TCAATGATAT TAGAATCCCA
2641 TCCTTTAATT TCTTTCAAAG TAGTTGAGCT ATAGTTAAGT GTAAGGGCCC TACTGCGAAA
2701 GCATTTGCCA AGGATGTTTT CATTAAATCA GAACGAAAGT TAGGGGATCG AAGACGATCA
2761 GATACCGTCG TAGTCTTAAC CATAAACTAT GCCGACTAGG GATCGGGCAA TGTTCATTT
2821 ATCGACTTGC TCGGCACCTT ACGAGAAATC AAAGTCTTTG GGTTCCGGGG GGAGTATGGT
2881 CGCAAGGCTG AAACCTAAAG GAATTGACGG AAGGGCACCA CAATGGAGTG GAGCCTGCGG
2941 CTTAATTTGA CTCAACACGG GGAACTCAC CAGGTCCAGA CATAGTAAGG ATTGACAGAT
3001 TGAGAGCTCT TTCTTGATTC TATGGGTGGT GGTGCATGGC CGTTCTTAGT TGGTGGAGTG
3061 ATTTGTCTGC TTAATTGCGA TAACGAACGA GACCTTAACC TGCTAAATAG CTGGATCAGC
3121 CATTTTGGCT GATCATTAGC TTCTTAGAGG GACTATTGGC ATAAAGCCAA TGGAGTTTG
3181 AGGCAATAAC AGGTCTGTGA TGCCCTTAGA TGTTCTGGGC CGCACGCGCG CTACACTGAC
3241 GGAGCCAACG AGTTGAAAAA AATCTTTTGA TTTTTTATCC TTGGCCGGAA GGTCTGGGTA
3301 ATCTTGTTAA ACTCCGTCGT GCTGGGGATA GAGCATTGCA ATTATTGCGG CCGCTCCTCA
3361 ATTCGATGTT GCAGATTTTA CAAGTTTTTA AAATGTATTT CATTATTACT TTTTATATGC
3421 CTAATAAAAA AGCCATAGTT TAATCTATAG ATAACTTTTT TTCCAGTGCA CTAACGGACG
```



29/167

**Figure 15B**

```

3481 TTACATTCCC ATACAAAAC TCGTAGTTAA AGCTAAGGAA AAGTTAATAT CATGTTAATT
3541 AAATACGCTA TTTACAATAA GACATTGAAC TCATTTATAT CGTTGAATAT GAATAACCAA
3601 TTTCAGCGAA TTTTAAACAA ACATCGTTCA CCTCGTTTAA GGATATCTTG TGTATGGGGT
3661 GTTGACTTGC TTTATCGAAT AATTACCGTA CCTGTAATTG GCTTGCTGGA TATAGCGGTA
3721 GTCTAATATC TAGCAAAAAT CTTTTGGGTG AAAAGGCTTG CAATTTACAG ACACCGAACT
3781 ATTTGTCATT TTTTAATAAG GAAGTTTTCC ATAAATTCCT GTAATTCCTG GTTGATCTAA
3841 TTGAAAAGAG TAGTTTTGCA TCACGATGAG GAGGGCTTTT GTAGAAAGAA ATACGAACGA
3901 AACGAAAATC AGCGTTGCCA TCGCTTTGGA CAAAGCTCCC TTACCTGAAG AGTCGAATTT
3961 TATTGATGAA CTTATAACTT CCAAGCATGC AAACCAAAG GGAGAACAAG TAATCCAAGT
4021 AGACACGGGA ATTGGATTCT TGGATCACAT GTATCATGCA CTGGCTAAAC ATGCAGGCTG
4081 GAGCTTACGA CTTTACTCAA GAGGTGATTT AATCATCGAT GATCATCACA CTGCAGAAGA
4141 TACTGCTATT GCACTTGGTA TTGCATTCAA GCAGGCTATG GGTAACCTTG CCGGCGTTAA
4201 AAGATTTGGA CATGCTTATT GTCCACTTGA CGAAGCTCTT TCTAGAAGCG TAGTTGACTT
4261 GTCGGGACGG CCCTATGCTG TTATCGATTT GGGATTAAAG CGTGAAAAGG TTGGGGAATT
4321 GTCCTGTGAA ATGATCCCTC ACTTACTATA TTCCTTTTCG GTAGCAGCTG GAATTACTTT
4381 GCATGTTACC TGCTTATATG GTAGTAATGA CCATCATCGT GCTGAAAGCG CTTTTAAATC
4441 TCTGGCTGTT GCCATGCGCG CGGCTACTAG TCTTACTGGA AGTTCTGAAG TCCCAAGCAC
4501 GAAGGGAGTG TTGTAAAGAT GAATTGGATT ATGTCAGGAA AAGAACGACA ATTTTGCATC
4561 CAAATTGTCT AAATTTTAGA GTTGCTTGAA AACAATAGAA CCTTACTTGC TTTATAATTA
4621 CGTTAATTAG AAGCGTTATC TCGTGAAGGA ATATAGTACG TAGCCGTATA AATTGAATTG
4681 AATGTTACG TTATAGAATA GAGACACTTT GCTGTTCAAT GCGTCGTCAC TTACCATACT
4741 CACTTTATTA TACGACTTTA AGTATAAACT CCGCGGTTAT GGTAAAATTA ATGATGCACA
4801 AACGTCCGAT TCCATATGGG TACACTACAA TTAAATACTT TTAAGCTGAT CCCCCACACA
4861 CCATAGCTTC AAAATGTTTC TACTCCTTTT TTAATCTTCC AGATTTTCTC GGAATCCGCG
4921 CATCGCCGTA CCACTTCAAA ACACCCAAGC ACAGCATACT AAATTTTCCC TCTTTCTTCC
4981 TCTAGGGTGT CGTTAATTAC CCGTACTAAA GGTTTGGAAA AGAAAAAGA GACCGCCTCG
5041 TTTCTTTTTC TTCGTCGAAA AAGGCAATAA AAATTTTAT CACGTTTCTT TTTCTTGAAA
5101 TTTTTTTTTT TAGTTTTTTT CTCTTTCAGT GACCTCCATT GATATTTAAG TTAATAAACG
5161 GTCTTCAATT TCTCAAGTTT CAGTTTCATT TTTCTTGTTT TATTACAAC TTTTTTACTT
5221 CTTGTTCAAT AGAAAGAAAG CATAGCAATC TAATCTAAGG GCGGTGTTGA CAATTAATCA
5281 TCGGCATAGT ATATCGGCAT AGTATAATAC GACAAGGTGA GGAATAAAC CATGGCCAAG
5341 TTGACCAGTG CCGTTCGGT GCTCACCAGG CGCGACGTCG CCGGAGCGGT CGAGTTCTGG
5401 ACCGACCGGC TCGGGTTCTC CCGGGAATTC GTGGAGGACG ACTTCGCCGG TGTGGTCCGG
5461 GACGACGTGA CCCTGTTTCA CAGCGCGGTC CAGGACCAGG TGGTGCCGGA CAACACCCTG
5521 GCCTGGGTGT GGGTGCGCGG CCTGGACGAG CTGTACGCCG AGTGGTCGGA GGTGCTGTCC
5581 ACGAACTTCC GGGACGCCTC CGGGCCGGCC ATGACCGAGA TCGGCGAGCA GCCGTGGGGG
5641 CGGGAGTTCT CCCTGCGCGA CCCGGCCGGC AACTGCGTGC ACTTCGTGGC CGAGGAGCAG
5701 GACTGACACG TCCGACGGCG GCCACGGGT CCCAGGCCTC GGAGATCCGT CCCCCTTTTC
5761 CTTTGTGAT ATCATGTAAT TAGTTATGTC ACGCTTACAT TCACGCCCTC CCCCCACATC
5821 CGCTCTAACC GAAAAGGAAG GAGTTAGACA ACCTGAAGTC TAGGTCCCTA TTTATTTTTT
5881 TATAGTTATG TTAGTATTAA GAACGTTATT TATATTTCAA ATTTTCTTT TTTTTCTGTA
5941 CAGACGCGAG CTTCCAGTA AATGTGCCAT CTCGTAGGCA GAAAACGGTT CCCCCTAGG
6001 GTCTCTCTCT TGGCCTCCTT TCTAGGTCGG GCTGATTGCT CTTGAAGCTC TCTAGGGGGG
6061 CTCACACCAT AGGCAGATAA CGTTCCCCAC CGGCTCGCCT CGTAAGCGCA CAAGGACTGC
6121 TCCCAAAGAT CCTAGGCGGG ATTTTGCCGA TTTCCGCCCTA AAGGAACCGG AACACGTAGA
6181 AAGCCAGTCC GCAGAAACGG TGCTGACCCC GGATGAATGT CAGCTACTGG GCTATCTGGA
6241 CAAGGGAAAA CGCAAGCGCA AAGAGAAAGC AGGTAGCTTG CAGTGGGCTT ACATGGCGAT
6301 AGCTAGACTG GCGGTTTTTA TGGACAGCAA GCGAACCAGA ATTGCCAGCT GGGGCGCCCT
6361 CTGGTAAGGT TGGGAAGCCC TGCAAAGTAA ACTGGATGGC TTTCTTGCCG CCAAGGATCT
6421 GATGGCGCAG GGGATCAAGA TCTGATCAAG AGACAGGATG AGGATCGTTT CGCATGATTG
6481 AACAAGATGG ATTGCACGCA GGTCTCTCCG CCGCTTGGGT GGAGAGGCTA TTCGGCTATG
6541 ACTGGGCACA ACAGACAATC GGCTGCTCTG ATGCCGCCGT GTTCCGGCTG TCAGCGCAGG
6601 GCGGCCCGGT TCTTTTGTG AAGACCGACC TGTCCGGTGC CCTGAATGAA CTGCAGGACG
6661 AGGCAGCGCG GCTATCGTGG CTGGCCACGA CGGGCGTTCC TTGCGCAGCT GTGCTCGACG
6721 TTGTCACTGA AGCGGGAAGG GACTGGCTGC TATTGGGCGA AGTGCCGGGG CAGGATCTCC
6781 TGTCATCTCG CCTTGCTCCT GCCGAGAAAG TATCCATCAT GGCTGATGCA ATGCGGCGGC
6841 TGCATACGCT TGATCCGGCT ACCTGCCCAT TCGACCACCA AGCGAAACAT CGCATCGAGC
6901 GAGCACGTAC TCGGATGGAA GCCGGTCTTG TCGATCAGGA TGATCTGGAC GAAGAGCATC
6961 AGGGGCTCGC GCCAGCCGAA CTGTTCCGCA GGCTCAAGGC GCGCATGCCC GACGGCGAGG

```



30/167

**Figure 15C**

```
7021 ATCTCGTCGT GATCCATGGC GATGCCTGCT TGCCGAATAT CATGGTGGAA AATGGCCGCT
7081 TTTCTGGATT CAACGACTGT GGCCGGCTGG GTGTGGCGGA CCGCTATCAG GACATAGCGT
7141 TGGATACCCG TGATATTGCT GAAGAGCTTG GCGGCGAATG GGCTGACCGC TTCCTCGTGC
7201 TTTACGGTAT CGCCGCTCCC GATTCGCAGC GCATCGCCTT CTATCGCCTT CTTGACGAGT
7261 TCTTCTGAAT TGAAAAAGGT ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAACT
7321 TCATTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT
7381 CCCTTAACGT GAGTTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC
7441 TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT
7501 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAAGTGG
7561 CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC
7621 TGCTGCCAGT GCGGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA
7681 TAAGGCGCAG CGGTCGGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC
7741 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCATTGA GAAAGCGCCA CGCTTCCCGA
7801 AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG
7861 GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG
7921 ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG
8041 CAACGCGGCC TTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTTCC
8101 TCGGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC
8161 TCGCCGCAGC CGAACGACCG AGCGCAGCGA G
```

(SEQ ID NO:11)



31/167

**Figure 16**

```
1 GAATTCAAAA CAAAATGTGT GCAACCTCCT CCCAGTTTAC TCAGATTACC GAGCATAATT
61 CTCGACGATC TGCTAACTAC CAGCCGAACC TTTGGAACTT TGAGTTTCTC CAGTCTCTCG
121 AAAATGACCT GAAGGTGGAA AAGCTCGAGG AGAAGGCGAC CAAACTCGAG GAGGAGGTGC
181 GATGTATGAT CAACAGAGTT GACACCCAAC CCCTGTCTTT GCTGGAGCTG ATCGACGATG
241 TGCAGCGGTT GGGTTTGACT TATAAATTCG AGAAGGACAT TATCAAGGCA CTGGAGAACA
301 TTGTGCTCCT CGACGAGAAC AAGAAGAACA AGTCTGATCT TCACGCTACC GCTCTCTCTT
361 TCCGACTTCT TCGACAACAC GGCTTCGAGG TGTCGCAGGA CGTCTTCGAG AGATTTAAGG
421 ACAAGGAGGG AGGATTTAGC GGCGAGCTGA AGGGAGACGT TCAGGGTCTT CTCTCCTTGT
481 ACGAGGCGTC CTACCTGGGA TTCGAGGGAG AGAACCTCCT GGAGGAAGCT CGTACATTTT
541 CCATCACTCA CCTTAAGAAT AACCTTAAGG AGGGAATTAA CACCAAGGTG GCCGAGCAGG
601 FTTCTCACGC CCTGGAGCTC CCCTACCACC AACGGCTCCA TAGACTGGAG GCTCGTTGGT
661 TCCTGGACAA ATATGAGCCA AAGGAGCCTC ATCATCAGTT GCTGTTGGAG TTGGCCAAGC
721 TGGACTTCAA TATGGTTCAG ACGCTGCACC AAAAGGAGTT GCAGGACCTG TCTCGATGGT
781 GGACCGAGAT GGGATTGGCC TCGAAGCTGG ATTTTGTCCG TGACCGACTT ATGGAGGTCT
841 ATTTTTGGGC CCTTGGAATG GCGCCTGACC CCCAGTTCCG AGAGTGCCGG AAGGCGGTGA
901 CGAAGATGTT CGGTCTTGTG ACTATCATCG ACGACGTCTA CGATGTCTAC GGCACACTCG
961 ACGAGTTGCA GCTGTTCACT GACGCCGTCG AGCGATGGGA TGTGAACGCC ATTAATACTC
1021 TCCCTGACTA TATGAAGCTG TGCTTCCTGG CTCTGTACAA CACTGTCAAC GATACCTCGT
1081 ACTCTATCCT CAAGGAGAAG GGACACAACA ATCTCTCCTA CTTGACCAA TCCTGGCGAG
1141 AACTGTGCAA GGCTTTTCTG CAGGAGGCTA AATGGTCCAA TAACAAGATC ATTCCTGCTT
1201 TTTCTAAATA CCTGGAAAAT GCCTCGGTGT CGAGCTCTGG CGTCGCCCTT CTGGCCCCCT
1261 CCTACTTCTC CGTCTGCCAG CAGCAGGAGG ATATTTCCGA TCATGCTCTT AGATCGCTGA
1321 CCGATTTTCA CGGCCTCGTG CGATCTTCCT GCGTGATTTT TCGGTTGTGT AATGACCTTG
1381 CGACCTCTGC TGCTGAGCTG GAACGAGGCG AACTACAAA TTCCATTATT TCTTACATGC
1441 ACGAAAACGA TGGAAATCTT GAAGAACAGG CTAGAGAGGA ACTGCGAAAG TTGATCGACG
1501 CCGAGTGGAA GAAGATGAAC AGAGAGCGGG TGTCCGACTC TACCCTGCTT CCAAGGCCT
1561 TCATGGAGAT CGCCGTGAAC ATGGCTCGAG TTTCCCATTG TACTTACCAG TACGGTGACG
1621 GCCTGSGTCG TCCGGACTAC GCTACAGAGA ACCGAATCAA GCTGCTGCTC ATCGACCCCT
1681 TCCCTATCAA CCAATTGATG TACGTGTAAT AGTCTAGAGG ATCC
```

(SEQ ID NO:12)



32/167

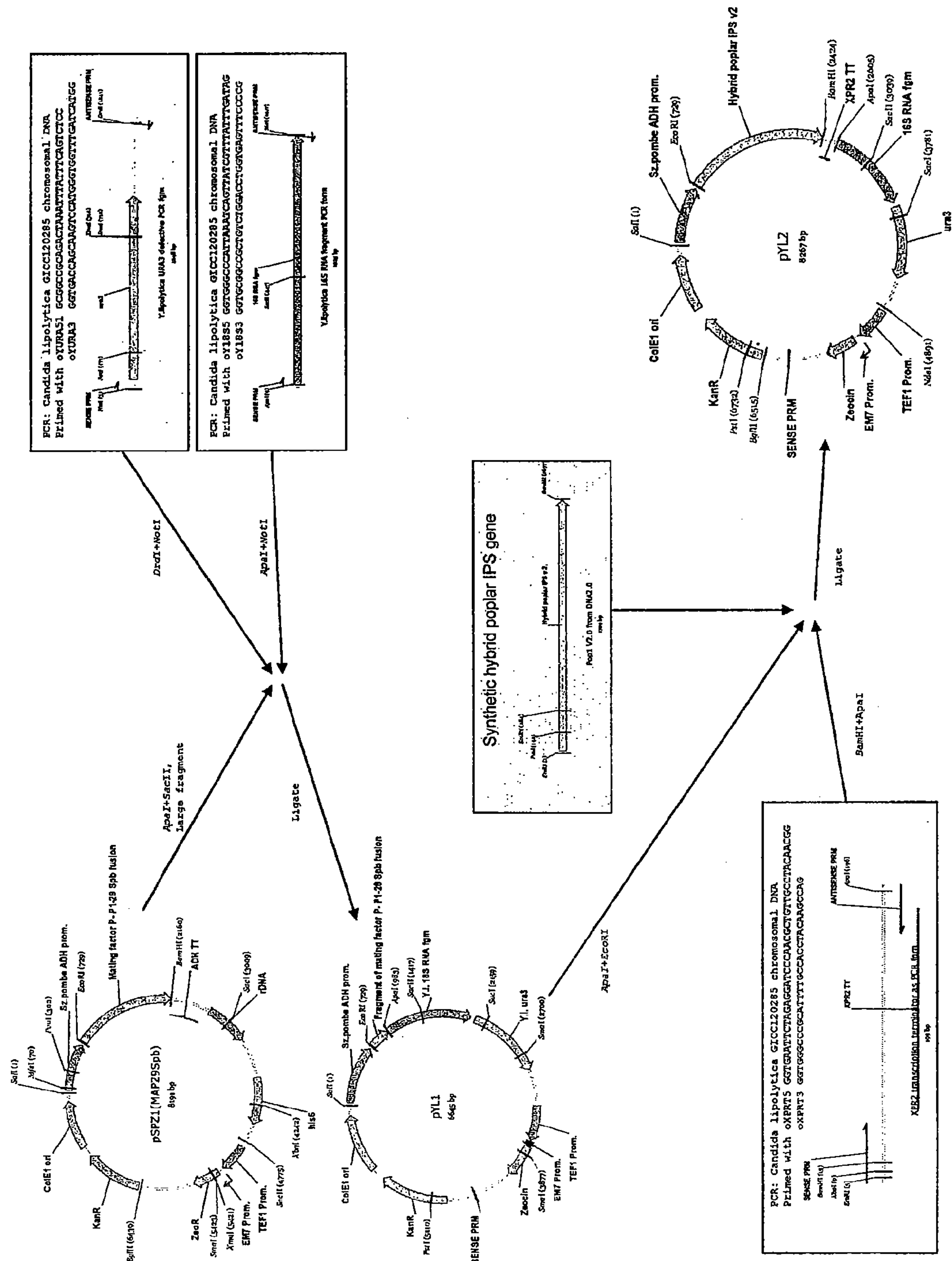
**Figure 17**

```
1 GAATTCAACA AAAATGTGCT CTGTTTCCAC TGAGAACGTG TCCTTTACTG AGACTGAGAC
61 TGAAGCACGT AGAAGCGCCA ACTACGAACC CAACTCCTGG GATTATGACT TTCTGCTGTC
121 TTCTGACACC GACGAGTCGA TCGAGGTTTA TAAGGATAAG GCCAAGAAAC TTGAGGCCGA
181 GGTCAGACGA GAGATTAACA ACGAGAAGGC CGAGTTCCTG ACCCTTCTTG AGCTGATCGA
241 CAACGTTCAA CGACTTGGTC TTGGTTACCG TTTCGAATCC GATATCCGAC GTGCATTGGA
301 TCGATTTGTC TCGTCCGGAG GTTTCGATGG TGTGACTAAG ACGTCGCTGC ACGCCACAGC
361 TCTTTCCTTC AGACTGTTGC GGCAGCATGG ATTTGAGGTT TCCCAGGAAG CCTTTTCTGG
421 TTTCAAGGAT CAGAACGGAA ACTTTTTTGA GAATCTCAAG GAGGACACCA AGGCCATCCT
481 GTCGTGTAT GAGGCCTCGT TCCTGGCTCT TGAGGGCGAG AATATTCTGG ATGAGGCTCG
541 GGTTCGCT ATTCGCACC TGAAGGAGTT GTCGGAGGAA AAGATCGGAA AGGAACTGGC
601 CGAGCAGGTC AACCATGCAC TTGAACCTCC CCTGCATCGA CGTACCCAGC GACTGGAGGC
661 CGTGTGGAGC ATCGAGGCGT ACAGAAAAA GGAGGATGCT AATCAGGTTT TGCTCGAACT
721 CGCTATCCTC GACTATAACA TGATTAGAG CGTGTACCAG CGTGAATTGC GAGAGACAAG
781 CCGGTGGTGG CGACGGGTGG GACTGGCCAC GAAGCTCCAC TTTGCTAAAG ATCGATTGAT
841 TGAGTCGTTT TACTGGGCAG TGGGTGTGGC CTTTGAGCCT CAGTACTCCG ACTGCCGAAA
901 CTCCGTTGCA AAGATGTTTT CTTTGTTCAC TATCATCGAC GACATCTACG ATGTTTACGG
961 CACTCTCGAT GAACTCGAAC TCTTCACGGA CGCTGTCGAG CGATGGGATG TGAATGCCAT
1021 TAATGATCTG CCAGATTATA TGAAGTTGTG TTTCTTGGCG CTCTACAACA CAATTAATGA
1081 AATTGCCTAC GACAACCTCA AGGACAAGGG AGAGAACATT CTGCCCTACC TTAATAAAGC
1141 CTGGGCCGAC CTGTGTAACG CCTTTTTGCA GGAAGCCAAG TGGCTCTATA ACAAATCTAC
1201 TCCTACATTT GATGACTACT TCGGCAACGC TTGGAAGTCT TCCAGCGGCC CTCTCCAGTT
1261 GATCTTCGCT TACTTTGCAG TGGTCCAGAA CATCAAGAAA GAGGAGATTG AGAACCTCCA
1321 GAAGTATCAC GACATCATCT CCCGACCTTC GCACATCTTT CGACTGTGCA ATGACCTTGC
1381 CTCCGCATCC GCTGAGATTG CCCGAGGAGA AACAGCCAAT TCTGTGTCGT GTTACATGCG
1441 TACAAAGGGC ATCTCCGAGG AGCTGGCTAC CGAGTCTGTG ATGAACCTGA TCGATGAAAC
1501 CTGTAAGAAG ATGAACAAAG AGAACTGGG CGGTTCTCTG TTCGCCAAAC CATTTGTTGA
1561 AACCGCGATC AATCTGGCTC GTCAGTCTCA TTGTACTTAC CATAACGGTG ACGCGCACAC
1621 TTCGCCGAC GAATTGACCC GTAAGCGTGT GCTTTCGGTG ATTACCGAGC CGATCCTGCC
1681 GTTCGAAAGA TAATAGGATC C
```

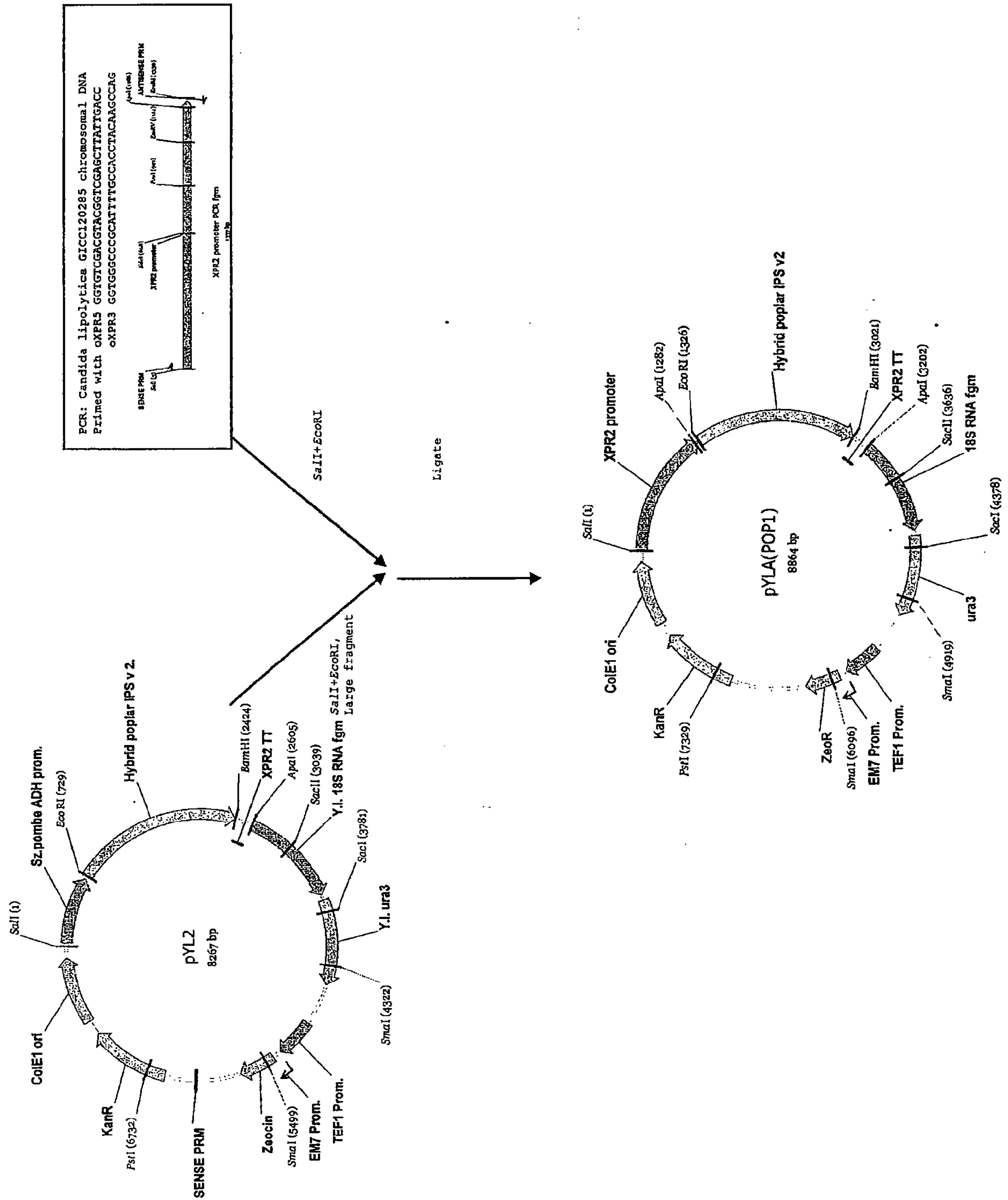
(SEQ ID NO:13)



### Figure 18A

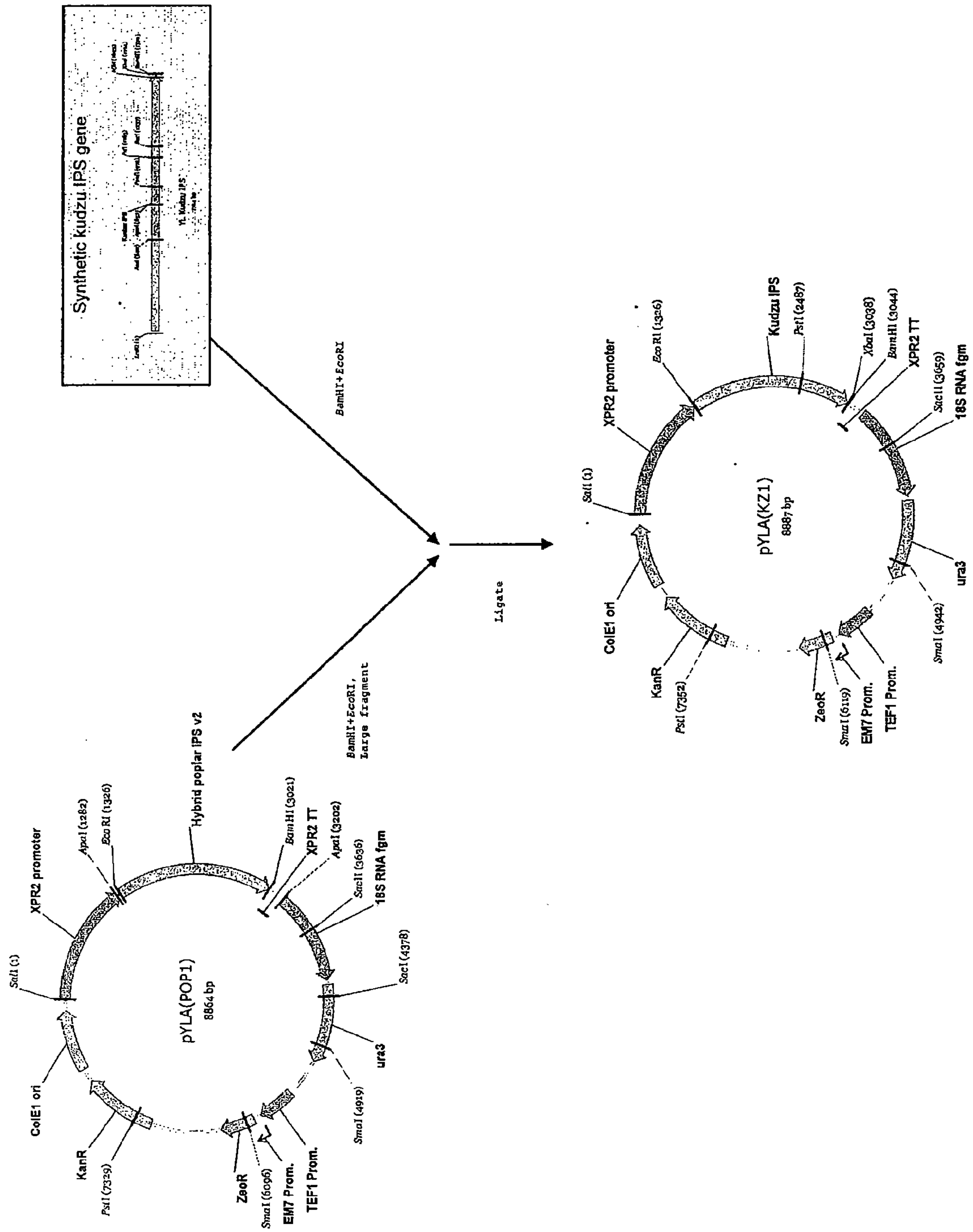




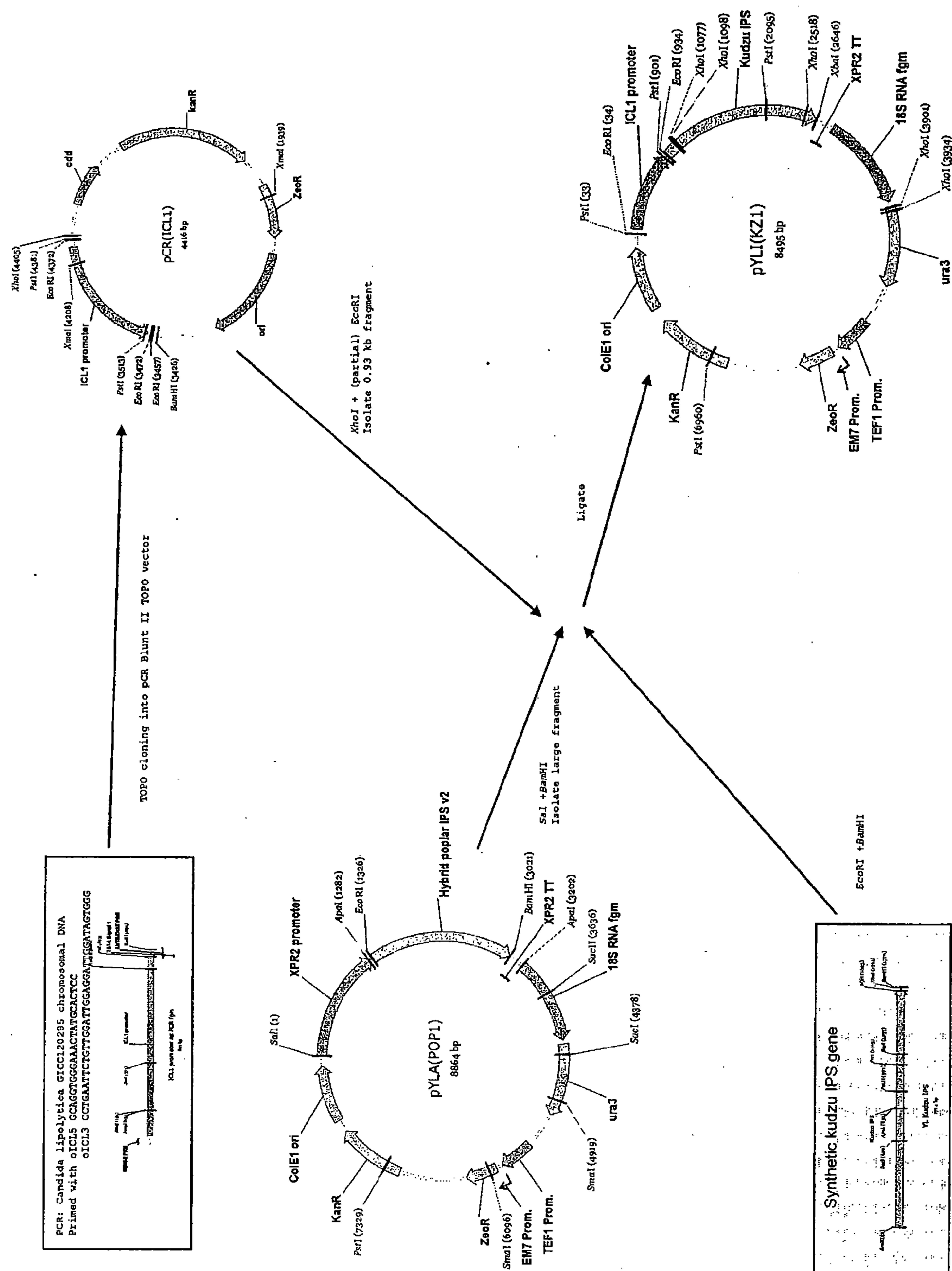




### Figure 18C



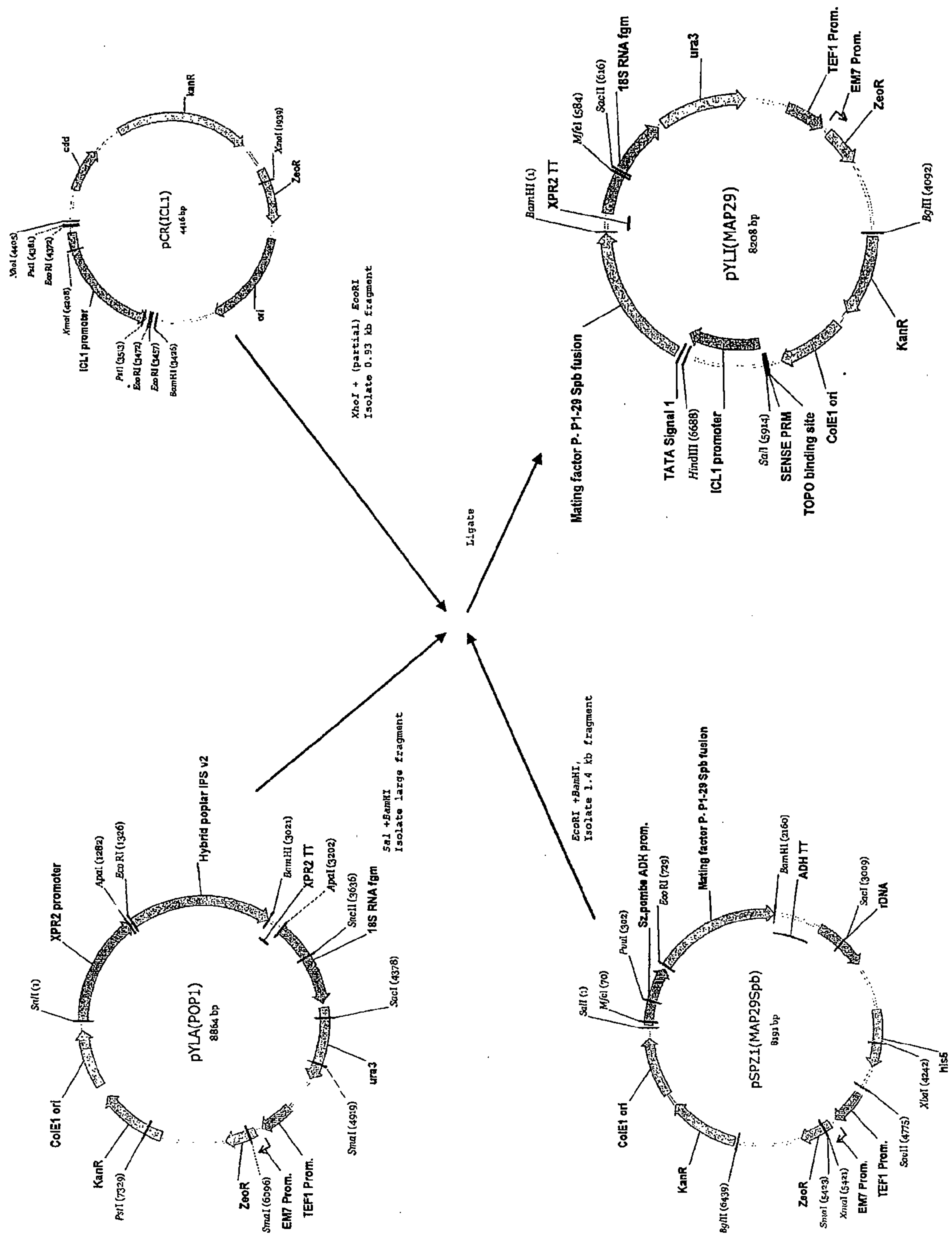




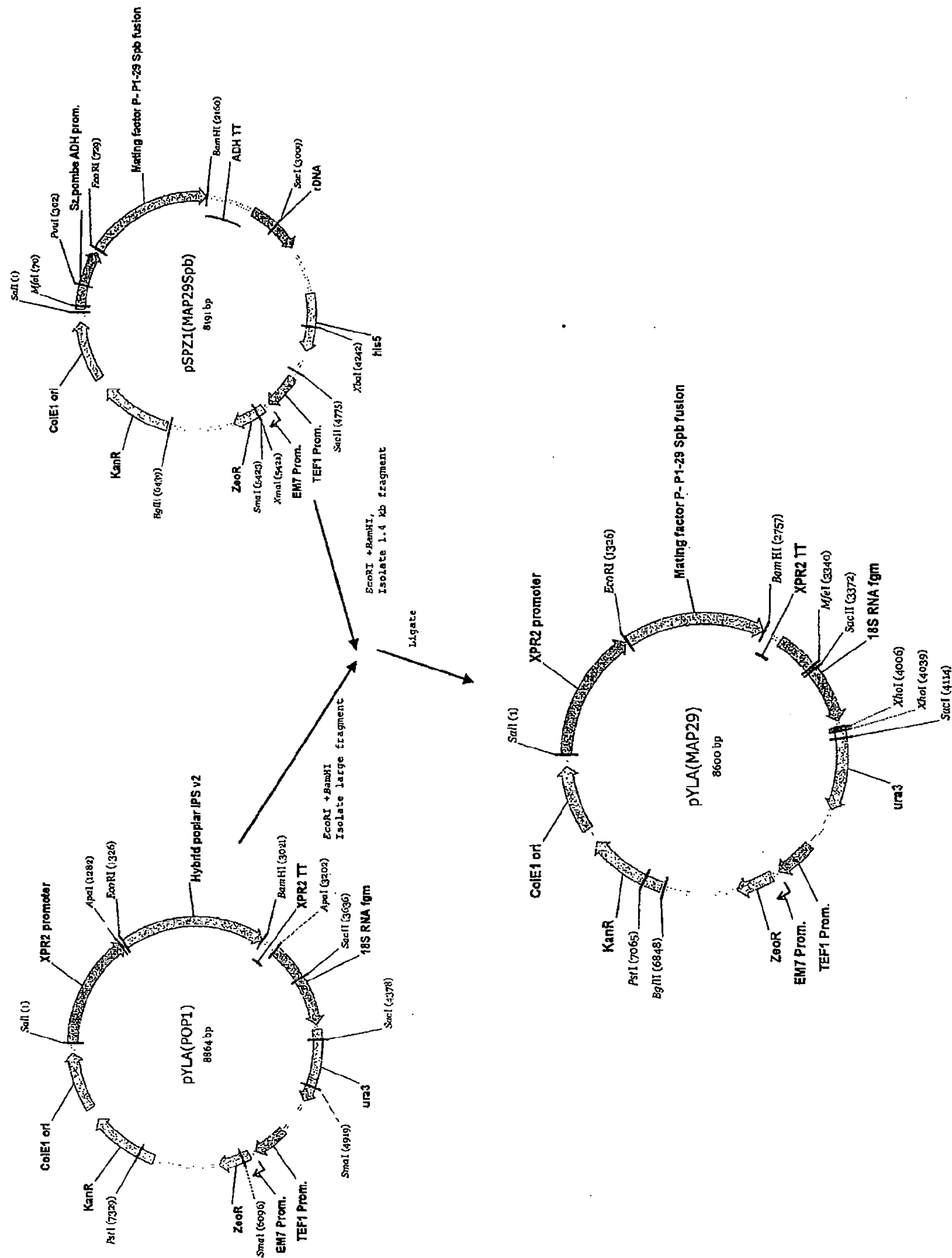


37/167

Figure 18E



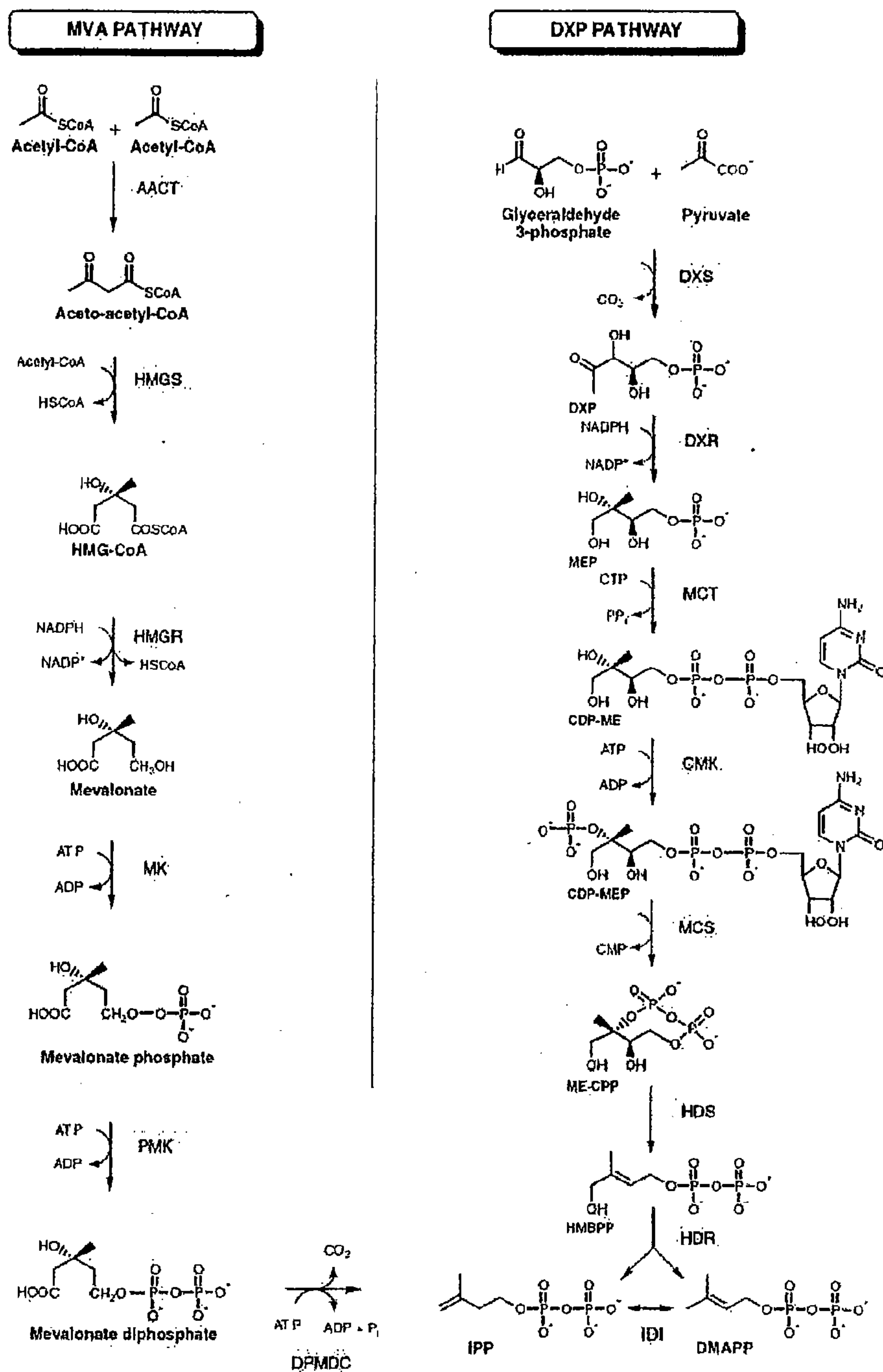






39/167

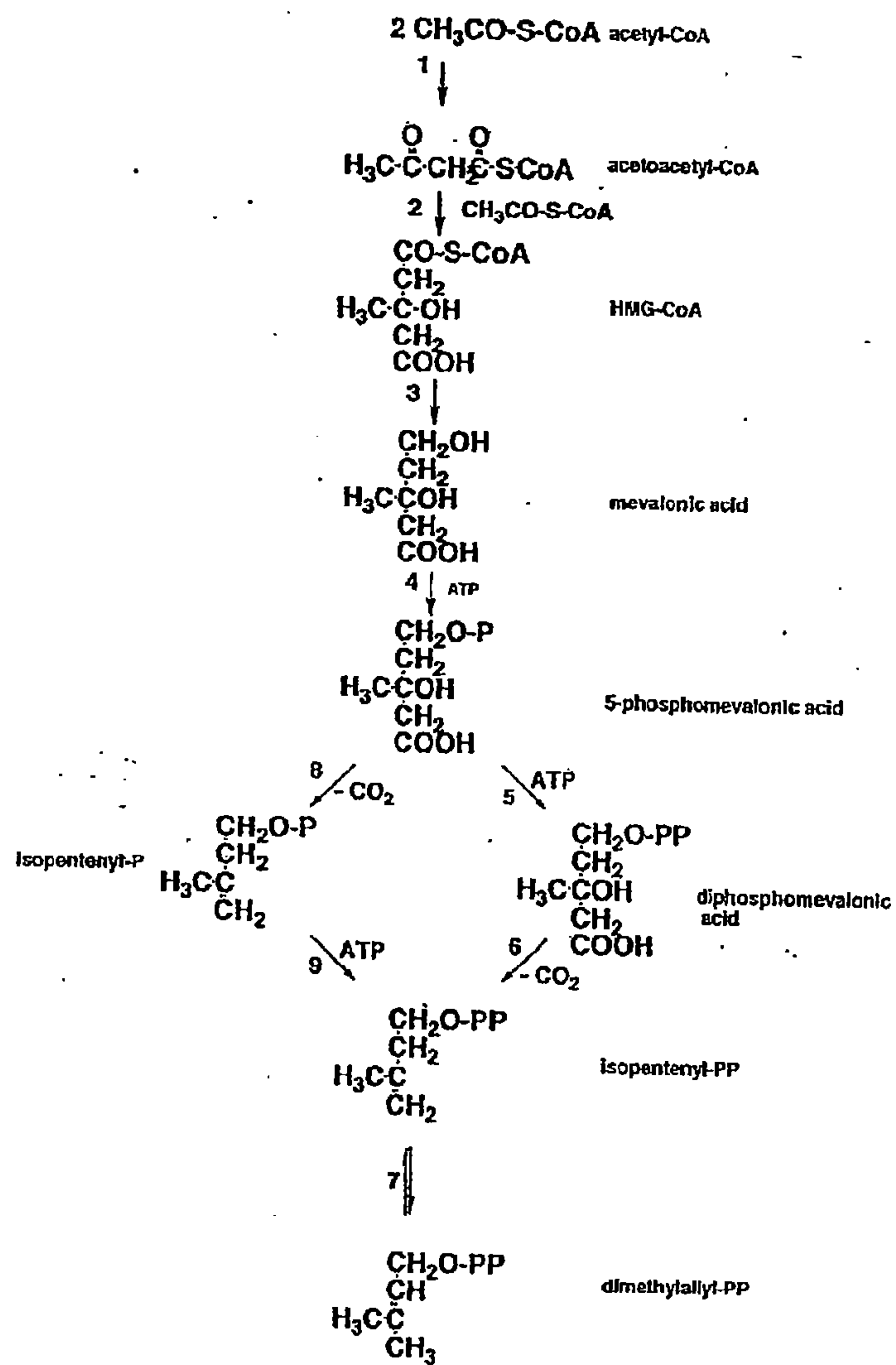
Figure 19A





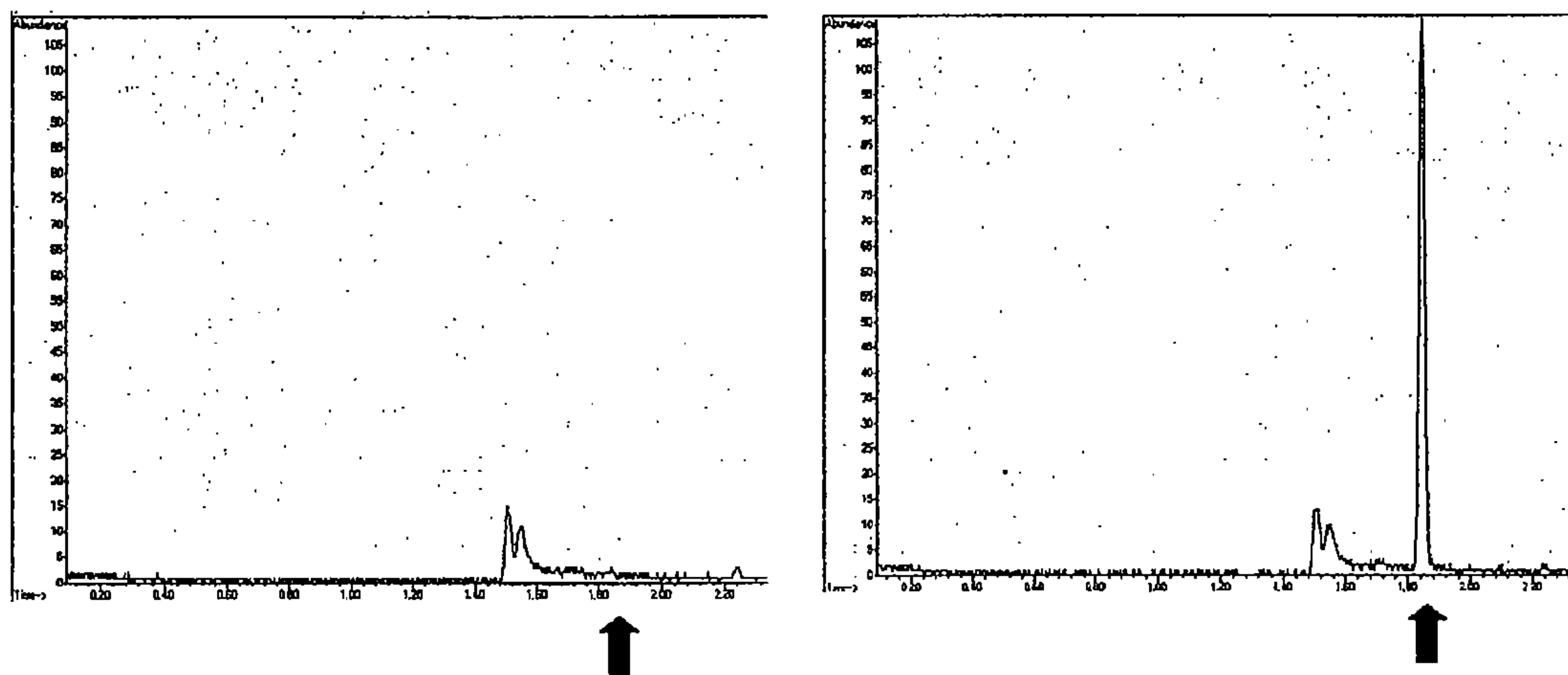
40/167

Figure 19B





41/167

**Figure 20***Y. lipolytica* CLIB122:: pYLA(MAP29)*Y. lipolytica* CLIB122:: pYLA(KZ1)



42/167

Figure 21

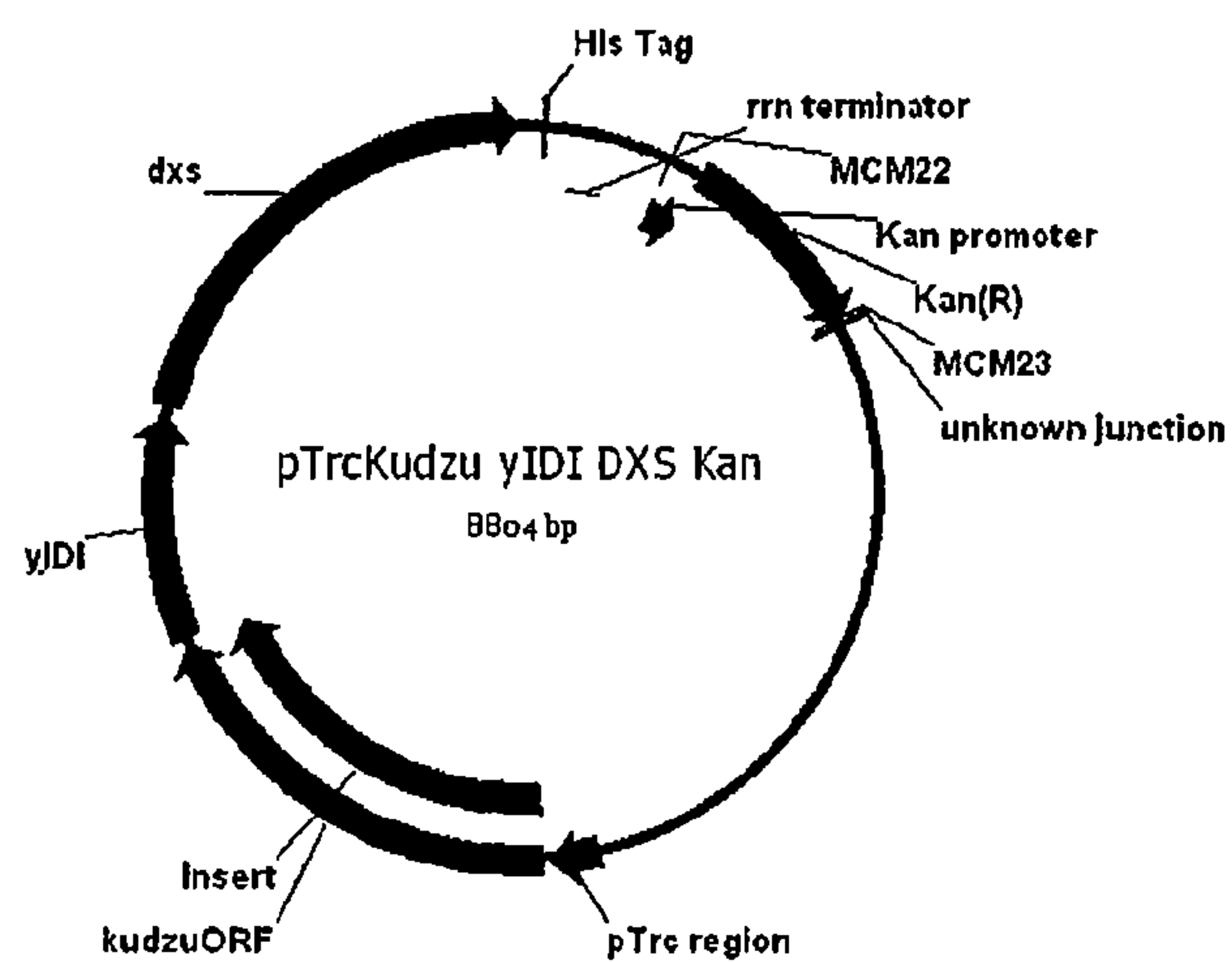




**Figure 22A**

1-

gctggtaccatatgggaattcgaagctttctagaacaaaaactcatctcagaagaggatctgaa  
tagcgccgctcgaccatcatcatcatcattgagtttaaaccggtctccagcttggctgttttg  
gcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggctctgataaa  
acagaatttgcctggcggcagtagcgcggtgggtcccacctgaccccatgccgaactcagaagt  
aaacgcgctagcgccgatggtagtgtgggggtctcccatgcgagagtagggaaactgccaggcat  
caaataaaacgaaaggctcagtcgaaagactgggcctttcggttttatctgttggttgctcggtga  
acgctctcctgagtaggacaaatccgcccgggagcggatttgaacggttgcaagcaacggcccg  
aggggtggcggggcaggacgcccgcataaaactgccaggcatcaaattaagcagaaggccatcctg  
acggatggcctttttgctgtttctacaaactctttttgtttatttttctaaatacattcaaata  
gtatccgcttaaccggaattgccagctggggcgccctctggtaagggttggaagccctgcaaag  
taaactggatggctttctcgccgccaaggatctgatggcgaggggatcaagctctgatcaaga  
gacaggatgaggatcgtttcgcatgattgaacaagatggattgcacgcagggttctccggccgct  
tgggtggagaggctattcggctatgactgggcacaacagacaatcggtctgctctgatgccgccc  
tggtccggctgtcagcgagggggcgcccggttctttttgtcaagaccgaacctgtccggtgccct  
gaatgaactgcaagacgaggcagcgcggtatcggtggctggccacgacgggcttcccttgccga  
gctgtgctcgacggtgtcactgaagcgggaagggaactgggtgctattggggcgaagtgccggggc  
aggatctcctgtcatctcaccttgctcctgccgagaaagtatccatcatggctgatgcaatgcg  
gcggctgcatacgcttgatccggctacctgccattcgaccaccaagcgaaacatcgcatcgag  
cgagcacgtactcggtggaagccggtcttgtcgatcaggatgatctggacgaagagcatcagg  
ggctcgcgccagccgaactgttcgccaggctcaaggcgagcatgcccgacggcgaggatctcgt  
cgtgacccatggcgatgcctgcttgccgaatatcatgggtggaaaatggccgcttttctggattc  
atcgactgtggccggctgggtgtggcggaaccgctatcaggacatagcggttggtacctcgat  
ttgctgaagagcttggcggcgaatgggtgaccgcttctcgtgctttacggtatcgccgctcc  
cgattcgacgagcatcgcttctatcgcttcttgacgagttcttctgacatgacaaaaatccc  
ttaacgtgagttttcggttccactgagcgctcagaccccgtagaaaagatcaaaggatcttcttga  
gatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtgg  
tttggtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgca  
gataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagca  
ccgcctacatacctcgctctgctaactctgttaccagtggctgctgccagtgggcgataagtctg  
gtcttaccgggttggaactcaagacgatatgttaccggataaggcgagcggtcgggctgaacggg  
gggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgt  
gagctatgagaaagcgccacgcttcccgaagggagaaaggcgagcaggatatccggttaagcggca  
gggtcggaaacaggagagcgacagggagcttccaggggggaaacgcctgggtatctttatagtcc  
tgtegggtttcgccacctctgacttgagcgctcgatttttgatgctcgtcagggggggcgagc  
ctatggaaaaacgcccagcaacgcggcctttttacgggttccctggccttttgctggccttttgctc  
acatgttctttcctgcttatccctgattctgtggataaccgtattaccgcctttgagtgagc  
tgataccgctcgccgcagccgaacgaccgagcgagcagcagtcagtgagcgaggaagcgggaagag  
cgctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatgggtgactc  
tcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatcgctacgtgac  
tgggtcatggctgcgccccgacacccgccaacacccgctgacgcgcccctgacgggcttgtctgc  
tcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagagggttttc  
accgtcatcaccgaaacgcgcgaggcagcagatcaattcgcgcgcaaggcggaagcggcatgca  
tttacgttgacaccatcgaatgggtgcaaacctttcgcggtatggcatgatagcgcccgggaaga



**Figure 22B**

gagtcaattcaggggtggtgaatgtgaaaccagtaacggttatacagatgtcgcagagtatgccggt  
gtctcttatcagaccggtttcccgcggtggtgaaccaggccagccacggtttctgcgaaaacgcggg  
aaaaagtggaaagcggcgatggcgagctgaattacattcccaaccgcgtggcacaacaactggc  
gggcaaacagtcggttgctgattggcggtgccacctccagtcctggccctgcacgcgccgtcgcaa  
attgtcgcggcgattaaatctcgcgcgcgatcaactgggtgccagcgtgggtggtgctgatggttag  
aacgaagcggcgctcgaagcctgtaaagcggcggtgcacaatcttctcgcgcgaacgcgtcagtgg  
gctgatcattaactatccgctggatgaccaggatgccattgctgtggaagctgcctgcactaat  
gttccggcggttatttcttgatgtctctgaccagacacccatcaacagttatttttctcccatg  
aagacgggtacgcgactgggcgtggagcatctggtcgcattgggtcaccagcaaatacgcgtgtt  
agcggggcccatthaagttctgtctcggcgcgctctgcgtctggctggctggcataaatactcact  
cgcaatcaaattcagccgatagcgggaacgggaaggcgactggagtgccatgtccggttttcaac  
aaaccatgcaaatactgaatgagggcatcgttcccactgcgatgctggttgccaacgatcagat  
ggcgctggggcgcaatgcgcgccattaccgagtcggggctgcgcgttggtgcggtatatctcggtta  
gtgggatacgcgataccgaagacagctcatgttatatcccgccgtcaaccaccatcaaacagg  
attttcgcctgctggggcaaaccagcgtggaccgcttgctgcaactctctcagggccaggcggt  
gaagggaatcagctggtgcccgtctcactggtgaaaagaaaaaccaccctggcgcccaatacgc  
caaaccgcctctccccgcgcgttgggcgattcattaatgcagctggcacgacagggtttcccgac  
tggaagcggggcagtgagcgcaacgcaattaatgtgagtttagcgcgaattgatctggttttgaca  
gcttatcatcgactgcacgggtgcaccaatgcttctggcgctcaggcagccatcggaagctgtggt  
atggctgtgcaggtcgtaaatcactgcataattcgtgtcgcctcaaggcgactcccgttctgga  
taatgttttttgcgccgacatcataacgggttctggcaaataattctgaaatgagctggttgacaat  
taatcatccggctcgtataatgtgtggaattgtgagcggataacaatttcacacaggaaacagc  
gccgctgagaaaaagcgaagcggcactgctctttaacaatttatcagacaatctgtgtgggcac  
tcgaccggaattatcgattaactttattattaaaaattaaagaggtatatattaatgtatcgat  
taaataaggaggaataaaccatgtgtgcgacctcttctcaatttactcagattaccgagcataa  
ttcccgtcgttccgcaaactatcagccaaacctgtggaatttcgaattcctgcaatccctggag  
aacgacctgaaagtggaaaagctggaggagaaagcgaccaaactggagggaagaagtctcgtgca  
tgatcaaccgtgtagacaccagccgctgtccctgctggagctgatcgacgatgtgcagcgct  
gggtctgacctacaaatttgaaaaagacatcattaagccctggaaaacatcgactgctggac  
gaaaacaaaaagaacaaatctgacctgcacgcaaccgctctgtctttccgtctgctgcgtcagc  
acggtttcgagggtttctcaggatgtttttgagcgtttcaaggataaagaagggtggtttcagcgg  
tgaactgaaagggtgacgtccaaggcctgctgagcctgtatgaagcgtcttacctgggtttcgag  
ggtgagaacctgctggaggaggcgctaccttttccatcacccacctgaagaacaacctgaaag  
aaggcattaataccaagggtgcagaacaagtgcagccacgcctggaaactgccatatcaccagcg  
tctgcaccgtctggaggcacggttggttccctggataaatacgaaccgaaagaacgcgatcaccag  
ctgctgctggagctggcgaagctggatttttaacatggtacagacctgcaccagaaagagctgc  
aagatctgtcccgtggtggaccgagatgggcctggctagcaaactggattttgtacgcgaccg  
cctgatggaagttttatttctgggcactgggtatggcgccagacccgcagtttggtgaatgtcgc  
aaagctgttactaaaatgtttggtctggtgacgatcatcgatgacgtgtatgacgtttatggca  
ctctggacgaactgcaactgttcaccgatgctgtagagcgctgggacgttaacgctattaacac  
cctgccggactatatgaaactgtgtttcctggcactgtacaacaccgttaacgacacgtcctat  
tctattctgaaagagaaagggtcataacaacctgtcctatctgacgaaaagctggcggtgaactgt  
gcaaagcctttctgcaagaggcgaaatgggtccaacaacaaaattatcccggctttctccaagta  
cctggaaaacgccagcgtttcctcctccggtgtagcgtgctggcgccgtcttacttttccgta



**Figure 22C**

tgccagcagcaggaagacatctccgaccacgcgctgcgttccctgaccgacttccatgggtctgg  
tgcgttctagctgcgttatcttccgcctgtgcaacgatctggccacctctgcggcggagctgga  
acgtggcgcgagactaccaattctatcattagctacatgcacgaaaacgatgggtaccagcgcaggaa  
caggcccgcgaagaactgcgtaaactgatcgacgcgcaatggaaaaagatgaatcgtgaacgcg  
ttagcgactccacctgctgcctaaagcgttcatggaaatcgagtttaacatggcacgtgtttc  
ccactgcacctaccagtatggcgatgggtctgggtcgcccagactacgcgactgaaaaccgcac  
aaactgctgctgattgaccttttcccgattaccagctgatgtatgtctaactgcacgcctt  
aggaggtaaaaaaaaaatgactgccgacaacaatagtatgccccatgggtgcagtatctagttacg  
ccaaattagtgcaaaaccaaacacctgaagacatttttgggaagagtttccctgaaattattccatt  
acaacaaagacctaatacccgatctagtgcacgtcaaatagcgaaagcggagaaacatgtttt  
tctgggtcatgatgaggagcaaattaagttaataatgaaaattgtattgttttggattgggacg  
ataatgctatttgggtgccggtaccaagaaagtttgtcattttaatggaaaatattgaaaagggttt  
actacatcgtgcatttctccgtcttttattttcaatgaacaagggtgaattacttttacaacaaaga  
gccactgaaaaaataactttccctgatctttggactaacacatgctgctctcatccactatgta  
ttgatgacgaattagggtttgaagggtgaagctagacgataagattaaggggcgctattactgcggc  
ggtgagaaaactagatcatgaattagggtattccagaagatgaaactaagacaaggggtaagttt  
cacttttttaaacagaatccattacatggcaccaagcaatgaacctgggggtgaacatgaaattg  
attacatcctattttataagatcaacgcctaaagaaaacttgactgtcaacccaaacgtcaatga  
agttagagacttcaaatagggtttcaccaaatgatttgaaaactatgtttgctgacccaagttac  
aagtttacgccttgggtttaagattatttgcgagaattacttattcaactgggtgggagcaattag  
atgacctttctgaagtggaaaatgacaggcaaattcatagaatgctataacaacgcgtcctgca  
ttcgcccttaggaggtaaaaaaacatgagttttgatattgccaaatacccgacctggcactgg  
tcgactccacccaggagttacgactggtgccgaaagagagtttaccgaaactctgcgacgaact  
gcgcgcgtattttactcgacagcgtgagccgttccagcgggcacttcgcctccgggctgggcacg  
gtcgaactgaccgtggcgctgcactatgtctacaacaccccggttgaccaattgatttgggatg  
tggggcatcaggcttatccgcataaaattttgaccggacgcgcgcgacaaaatcggcaccatccg  
tcagaaaggcgggtctgcacctgttcccgtggcgcgcggaagcgaatatgacgtattaagcgtc  
gggcattcatcaacctccatcagtgccggaattgggtattgcgggttgctgccgaaaaagaaggca  
aaaatcgccgcacctgtctgtgtcattggcgatggcgcgattaccgcaggcatggcggtttgaagc  
gatgaatcacgcgggcgatatccgtcctgatatgctgggtgattctcaacgacaatgaaatgtcg  
atttccgaaaatgtcggcgcgctcaacaacctctggcacagctgctttccggtaagctttact  
cttcactgcgcgaaggcgggaaaaaagttttctctggcggtgccgccaattaaagagctgctcaa  
acgcaccgaagaacatattaaaggcatggtagtgctggcacgttggtttgaagagctgggcttt  
aactacatcgccccgggtggacgggtcacgatgtgctggggcttatcaccacgctaaagaacatgc  
gcgacctgaaaggcccgagttcctgcataatcatgacaaaaaagggtcgtgggttatgaaccggc  
agaaaaagacccgatcactttccacgcctgcctaaatttgatccctccagcgggttggttgccg  
aaaagtagcggcggtttgccgagctattcaaaaatctttggcgactgggttggtgcgaaacggcag  
cgaaagacaacaagctgatggcgattactccggcgatgcgtgaagggtccggcatgggtcgagtt  
ttcacgtaaatcccggtatcgctacttcgacgtggcaattgccgagcaacacgcgggtgaccttt  
gctgcgggtctggcgattgggtgggtacaaaccattgtcgcgatttactccactttcctgcaac  
gcgcctatgatcagggtgctgcacgtggcgattcaaaagcttccgggtcctgttcgccatcga  
ccgcgcgggcattgttggtgctgacggtcaaaccatcagggtgcttttgatctctcttacctg  
cgctgcataccggaaatgggtcattatgaccccgagcgatgaaaacgaatgtcgccagatgctct  
ataccggctatcactataacgatggcccgctcagcgggtgcgctaccgcgtggcaacgcggtcgg



46/167

**Figure 22D**

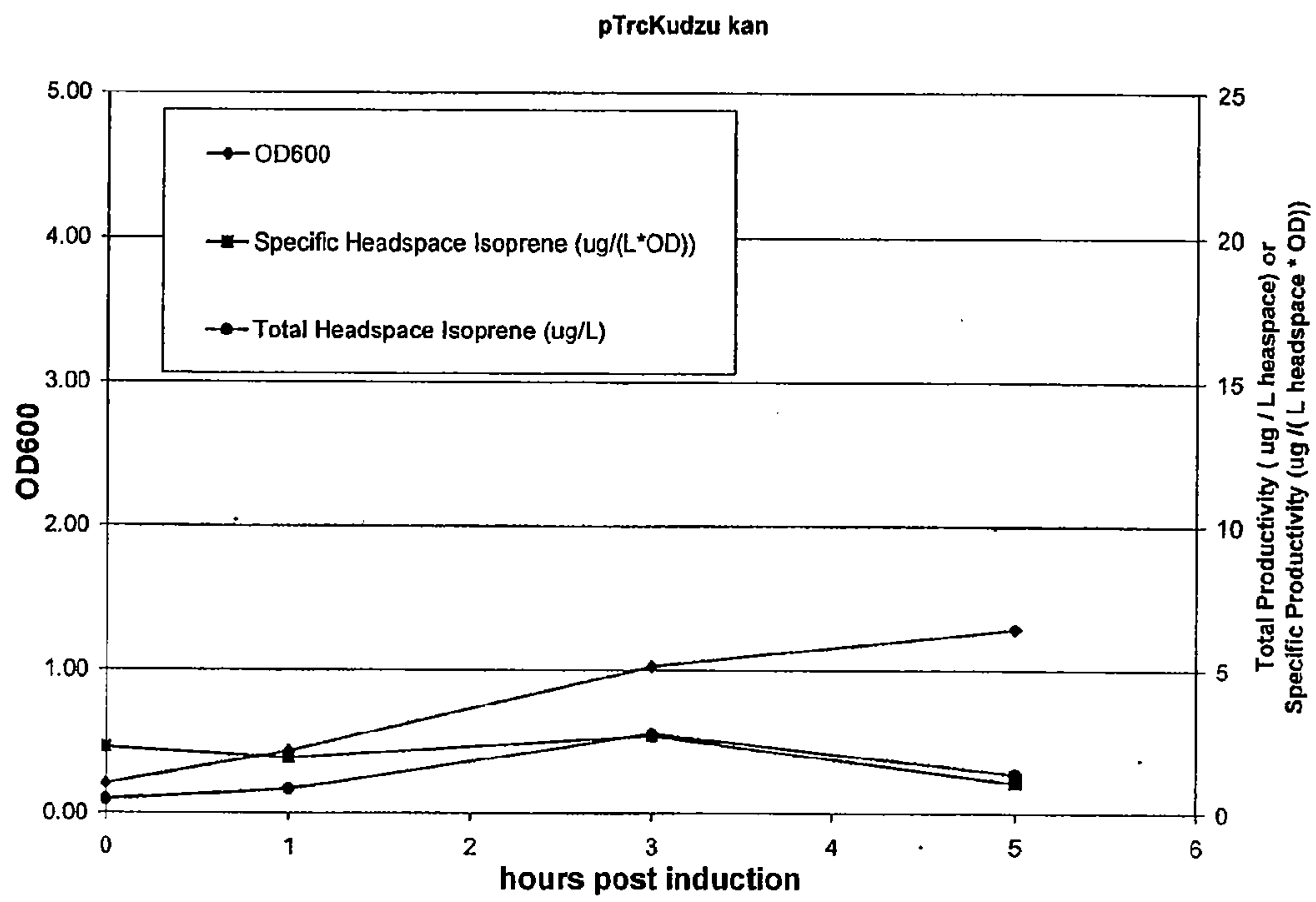
cgtggaactgacgccgctggaaaaactaccaattggcaaaggcattgtgaagcgtcgtggcgag  
aaactggcgatccttaacttttggtacgctgatgccagaagcggcgaaagtcgccgaatcgctga  
acgccacgctggtcgatatgcgttttgtgaaaccgcttgatgaagcgttaattctggaaatggc  
cgccagccatgaagcgtggtcaccgtagaagaaaacgccattatgggcggcgagcagcggc  
gtgaacgaagtgctgatggcccatcgtaaaccagtaccgctgctgaacattggcctgccggact  
tctttattccgcaagggaactcaggaagaaatgcgcgccgaactcggcctcgatgccgctggtat  
ggaagccaaaatcaaggcctggctggcataactgca

(SEQ ID NO:20)



47/167

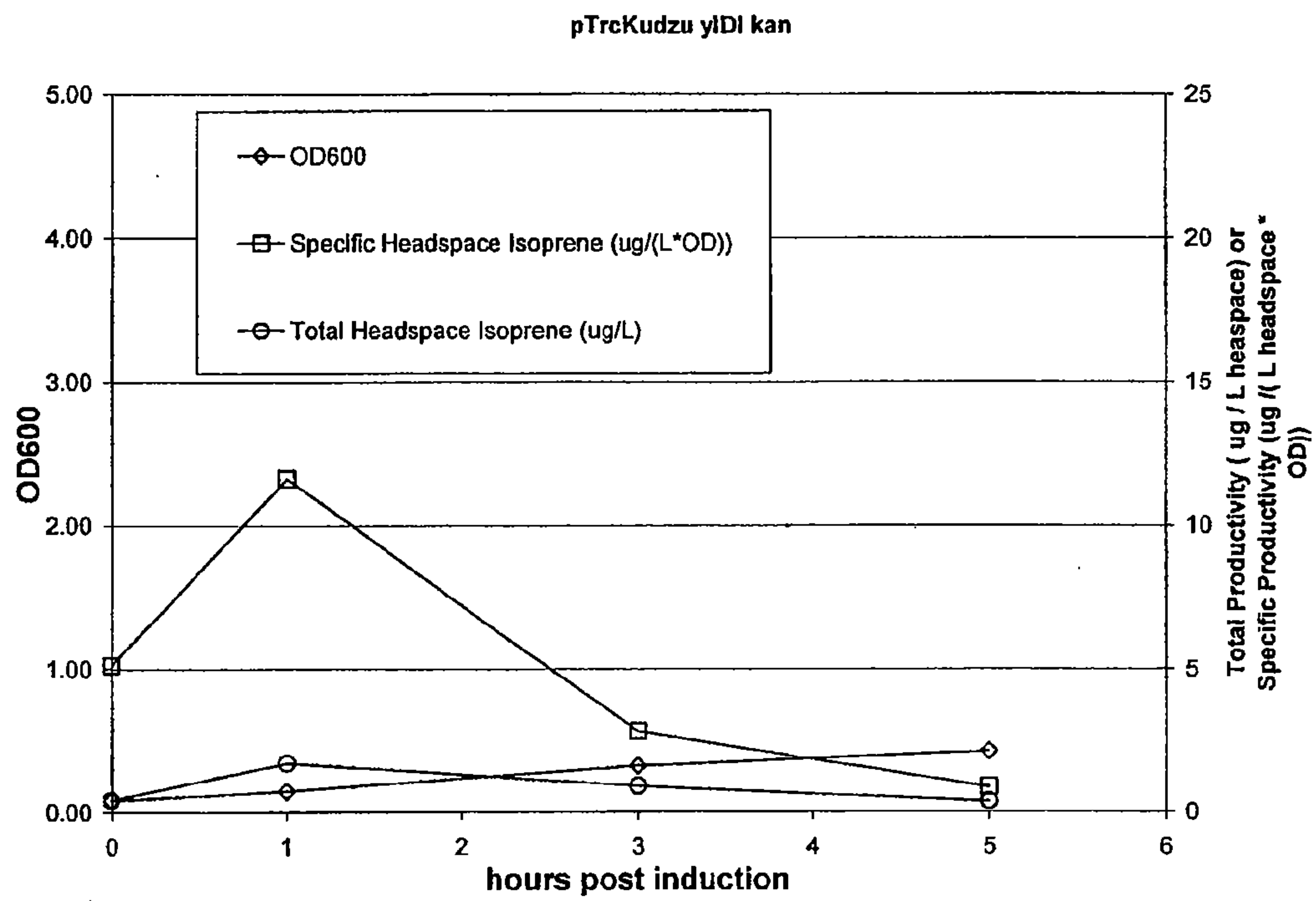
Figure 23A





48/167

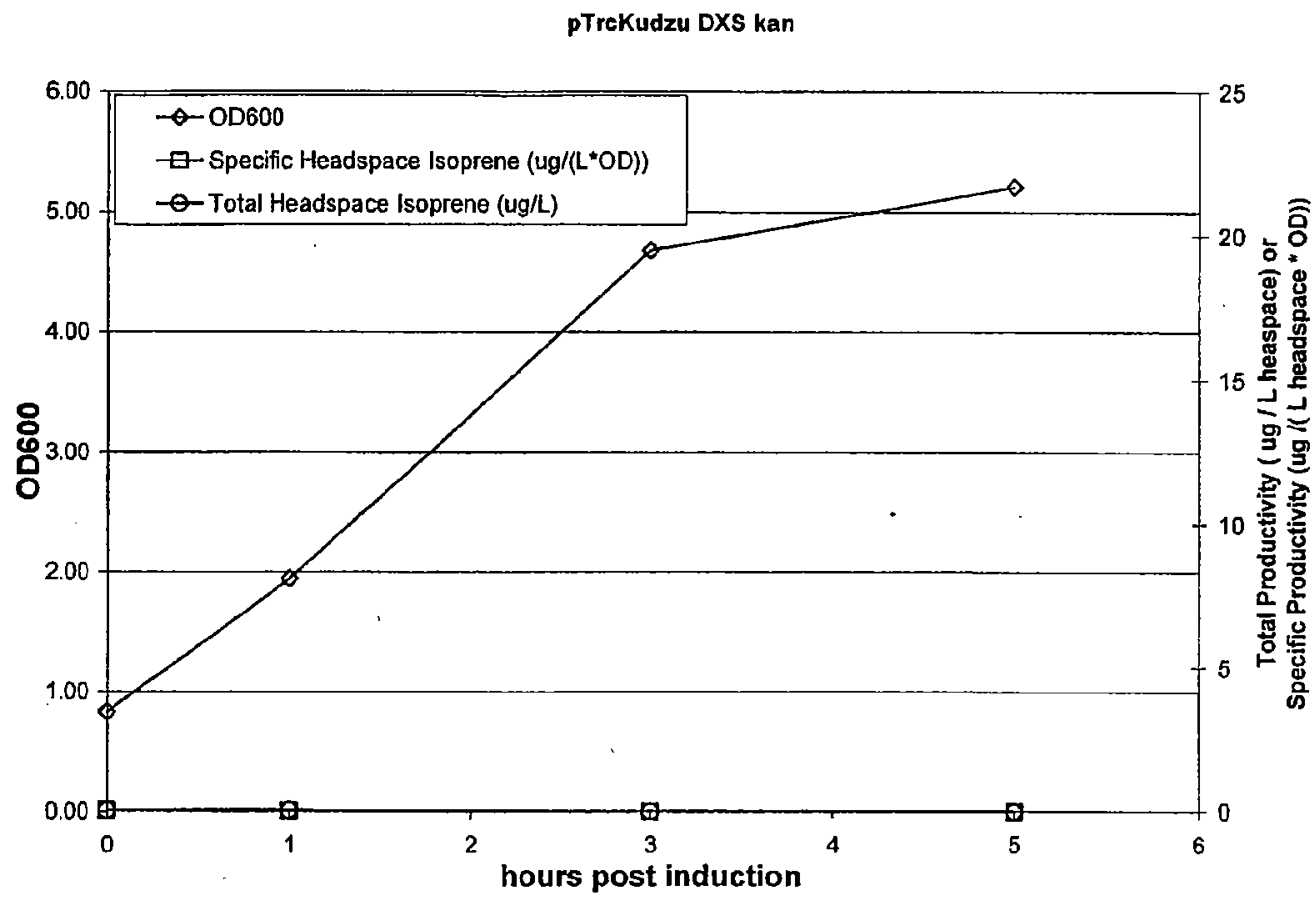
Figure 23B





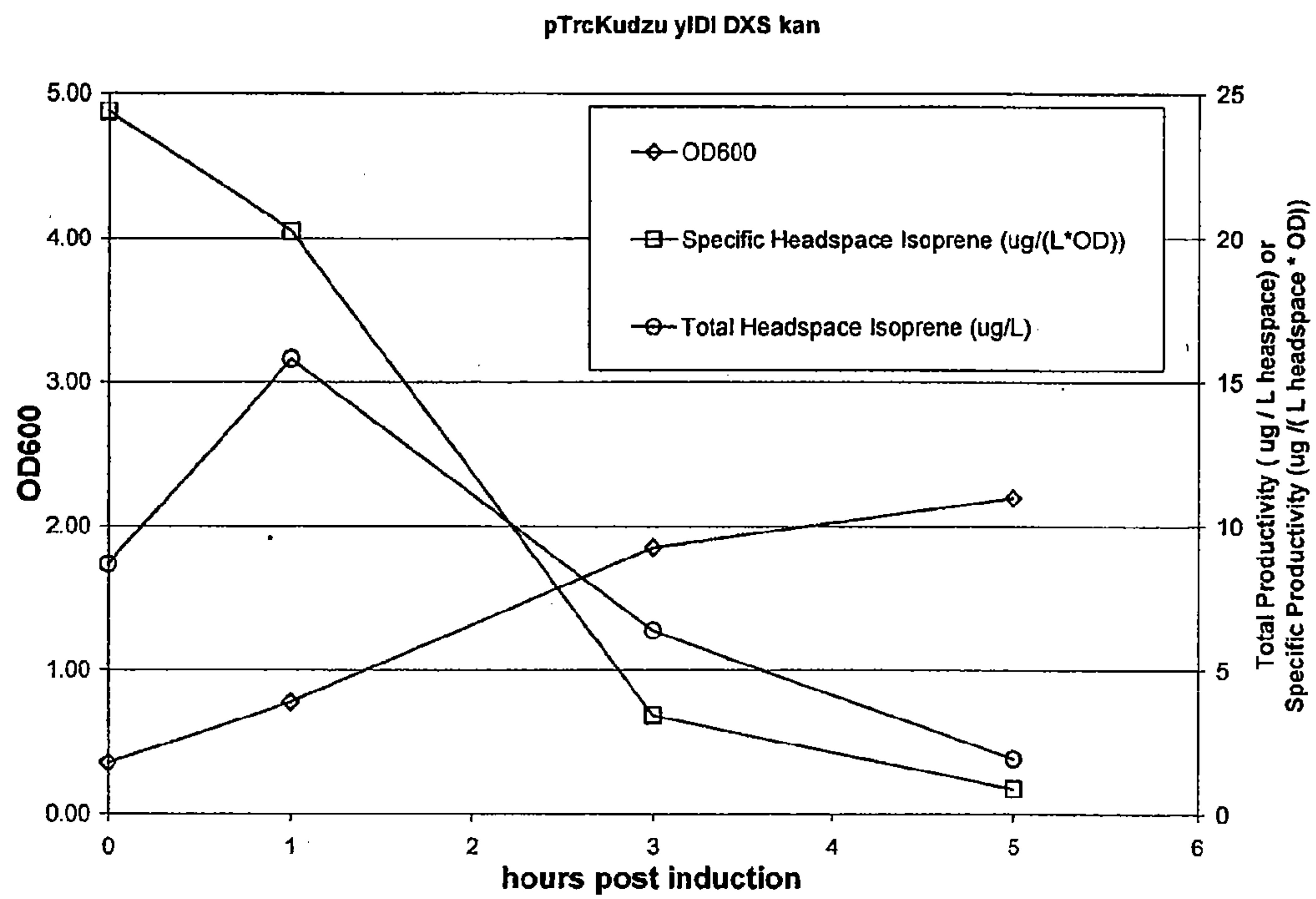
49/167

Figure 23C





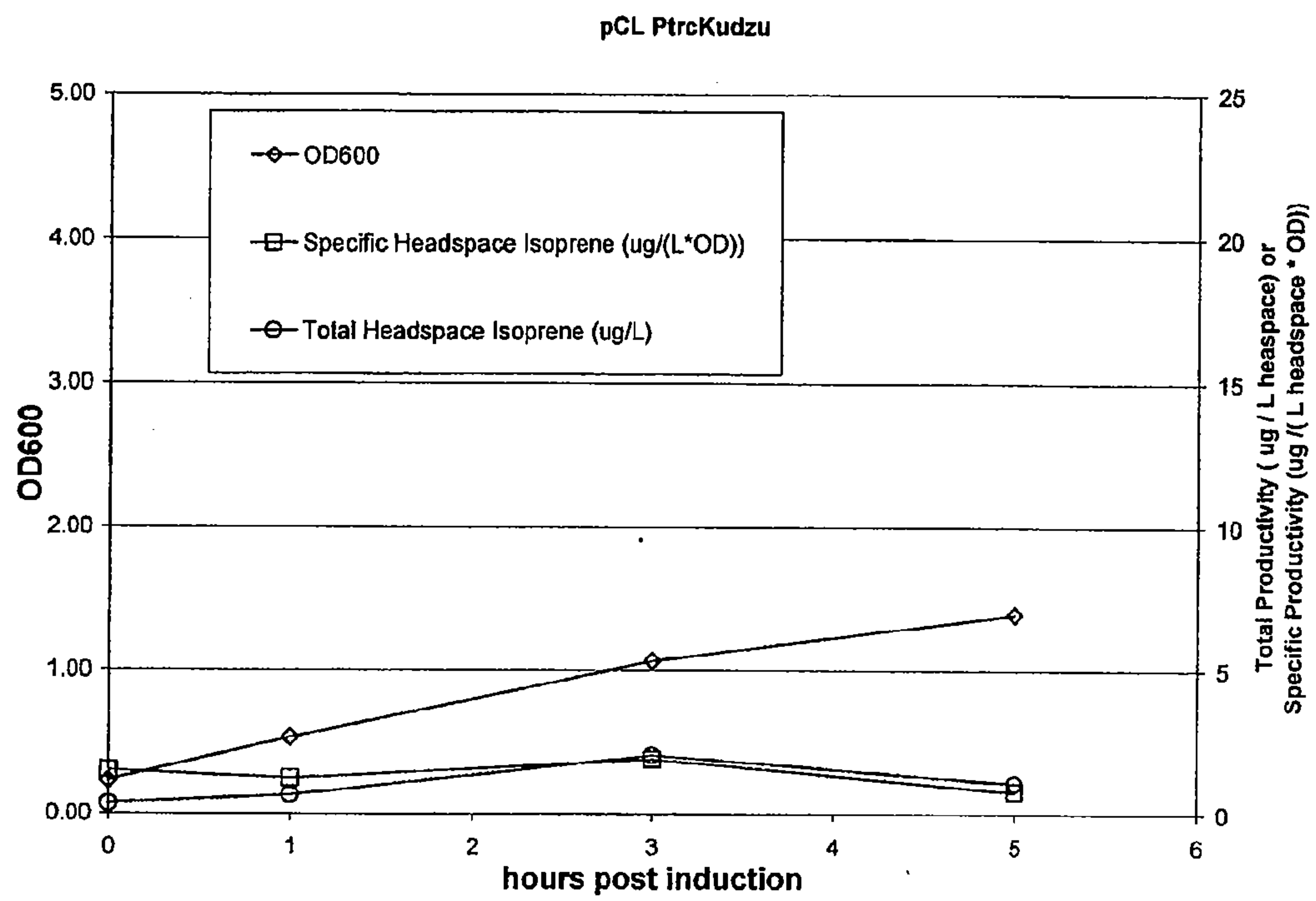
50/167

**Figure 23D**



51/167

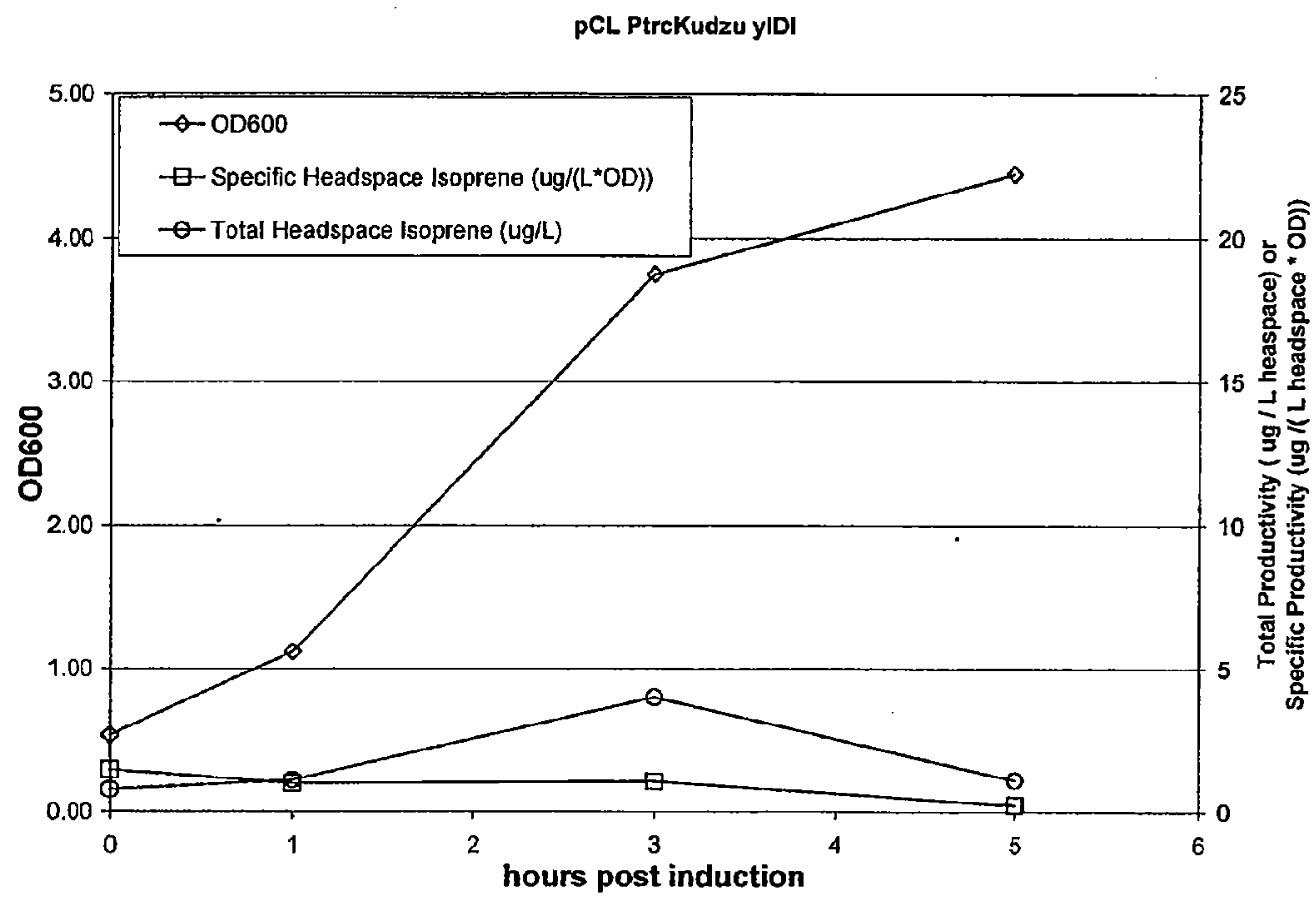
Figure 23E





52/167

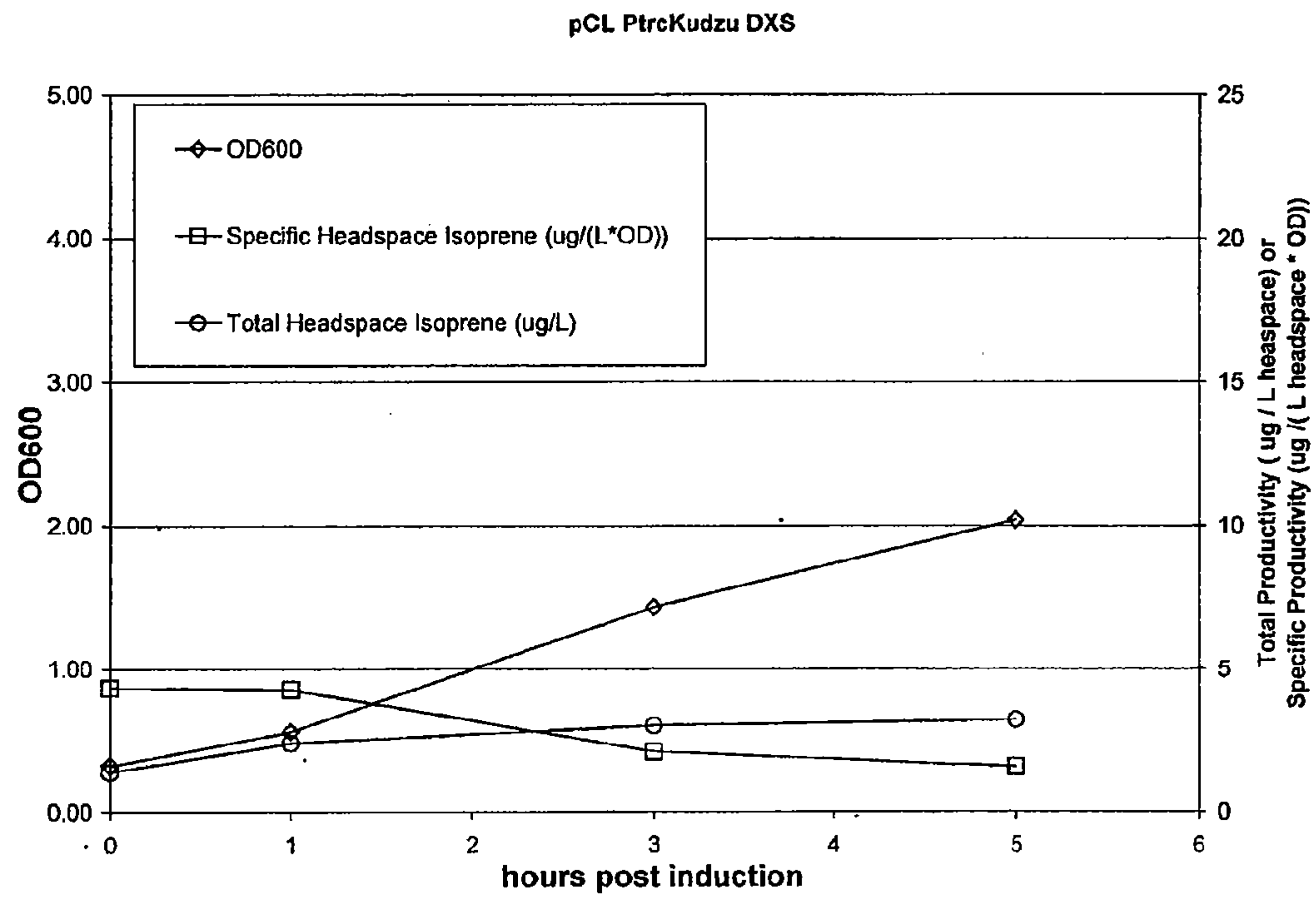
Figure 23F





53/167

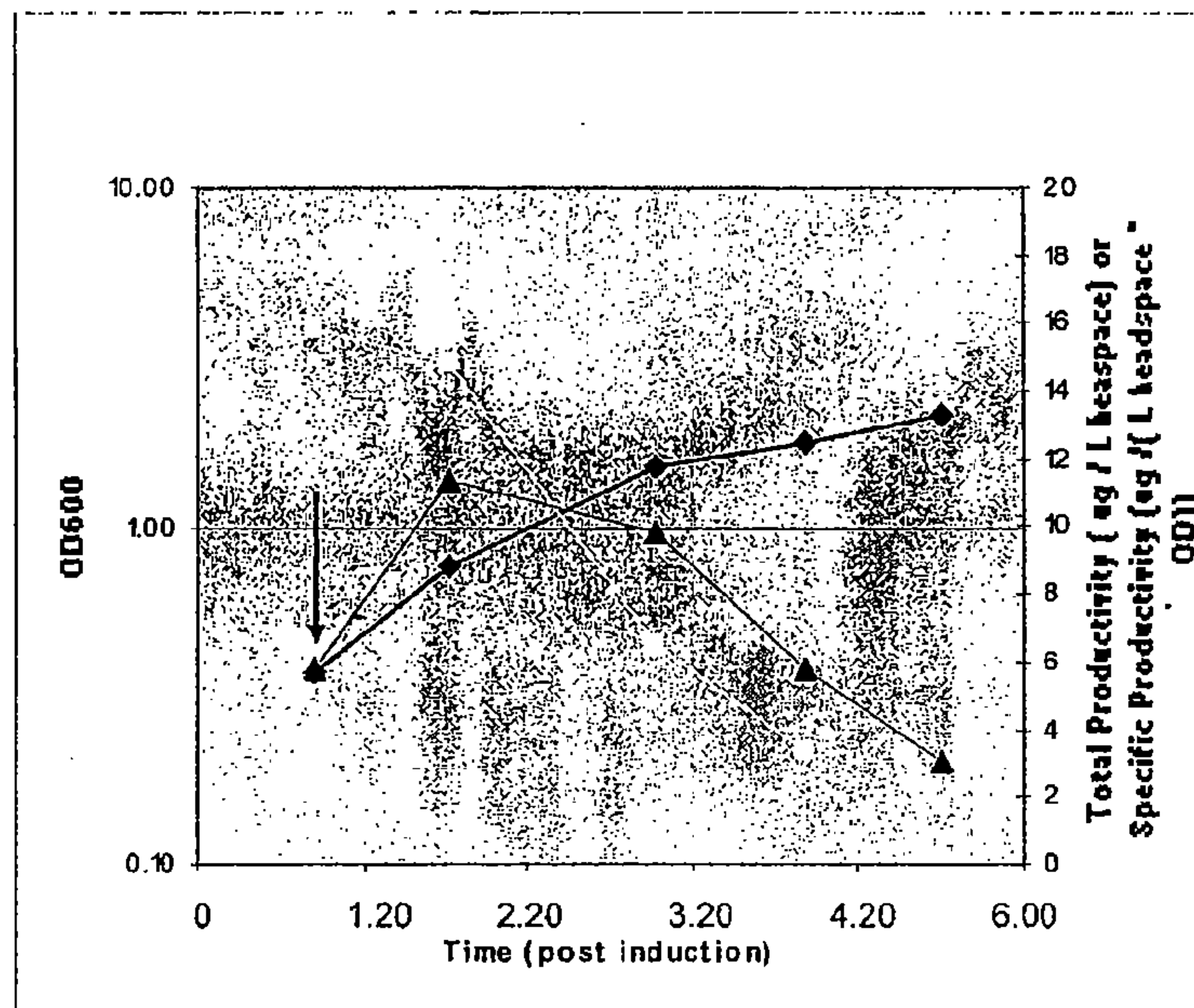
Figure 23G





54/167

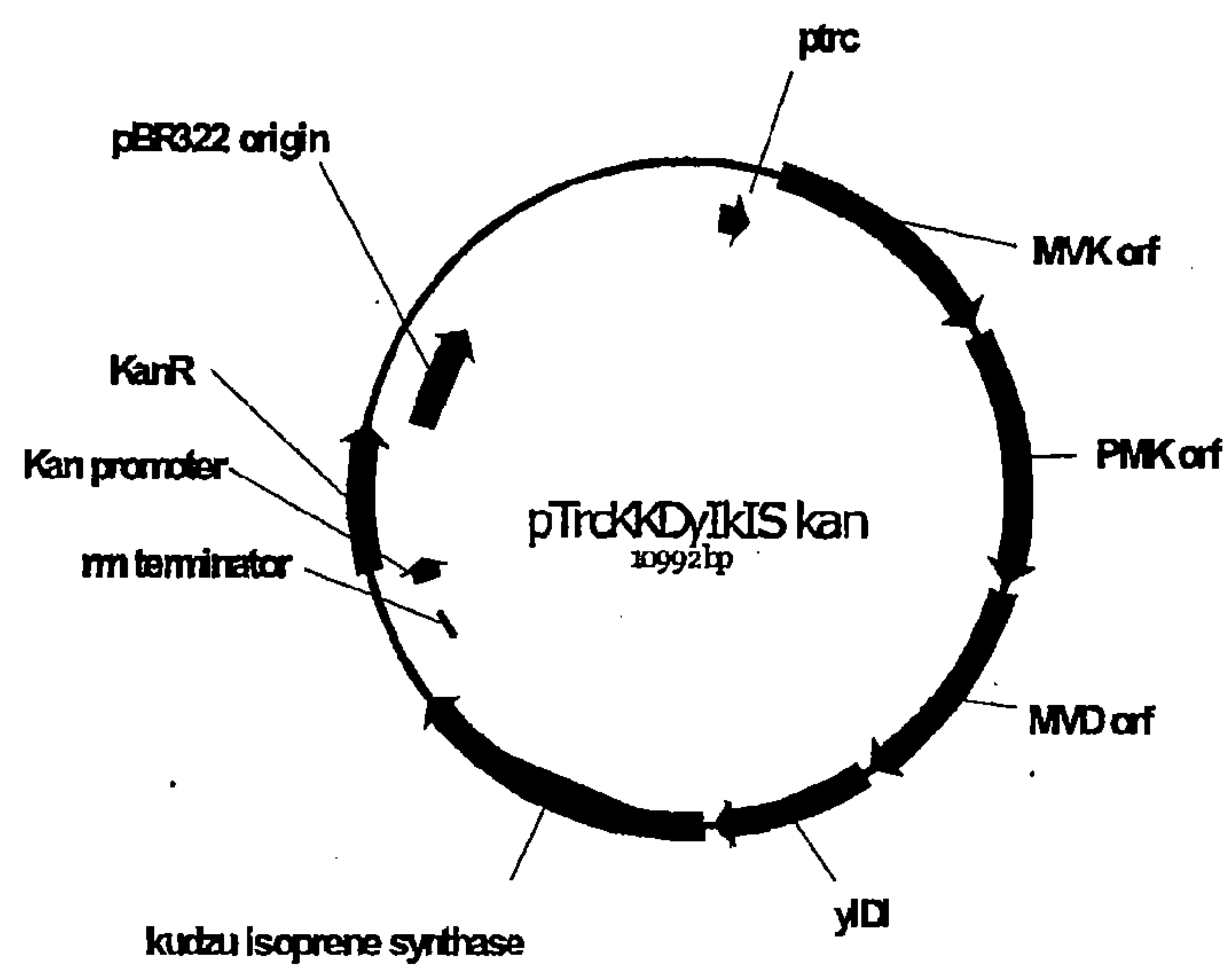
Figure 23H





55/167

Figure 24





56/167

**Figure 25A**

5' -

gtttgacagcttatcatcgactgcacggtgcaccaatgcttctggcgtcaggcagccatcggaa  
gctgtggtatggctgtgcaggtcgtaaactactgcataattcgtgtcgctcaaggcgcactccc  
gttctggataatgttttttgcgccgacatcataacggttctggcaaataattctgaaatgagctg  
ttgacaattaatcatccggctcgtataatgtgtggaattgtgagcggataacaatttcacacag  
gaaacagcgcgcgtgagaaaaagcgaagcggcactgctctttaacaatttatcagacaatctgt  
gtgggcactcgaccggaattatcgattaactttattattaaaaattaaagaggtatatattaat  
gtatcgattaaataaggaggaataaaccatggatccgagctcggatccactagtaacggccgcc  
agtgtgctggaattcgcccttaggaggtaaaaaaacatgtcattaccgttcttaacttctgcac  
cgggaaagggttattatttttgggtgaacactctgctgtgtacaacaagcctgccgtcgctgctag  
tgtgtctgcgttgagaacctacctgctaataagcagagtcattctgcaccagatactattgaattg  
gacttcccggacattagctttaatcataagtgggtccatcaatgatttcaatgccatcacaggagg  
atcaagtaaaactcccaaaaattggccaagggtcaacaagccaagcagatggcttgtctcaggaact  
cgtagtcttttggatccgttgtagtcaactatccgaatccttccactaccatgcagcgttt  
tgtttcctgtatatgtttgtttgcctatgcccccatgccagaataattaagttttctttaaggt  
ctactttacccatcgggtgctgggttgggctcaagcgcctctatttctgtatcactggccttagc  
tatggcctacttggggggggttaataggatctaatagacttggaaaagctgtcagaaaacgataag  
catatagtgaatcaatgggccttcataaggtgaaaagtgtattcacgggtaccccttcaggaatag  
ataacgctgtggccacttatggtaatgccctgctatttgaaaaagactcacataatggaacaat  
aaacacaaaacaattttaagttcttagatgatttcccagccattccaatgatcctaacctatact  
agaattccaagggtctacaaaagatcttggtgctcgcgttcgtgtgttggtcaccgagaaatttc  
ctgaagttatgaagccaattctagatgccatgggtgaatgtgccctacaaggcttagagatcat  
gactaagttaagtaaatgtaaaggcaccgatgacgaggctgtagaaactaataatgaactgtat  
gaacaactattggaattgataagaataaatcatggactgcttgtctcaatcgggtgtttctcatc  
ctggattagaacttattaaaaatctgagcgatgatttgagaattgggtccacaaaacttaccgg  
tgctgggtggcggcggttgctctttgactttgttacgaagagacattactcaagagcaaattgac  
agcttcaaaaagaaattgcaagatgatttttagttacgagacatttgaaacagacttgggtggga  
ctggctgctgtttgttaagcgcaaaaaatttgataaagatcttaaaatcaaataccctagtatt  
ccaattatttgaaaataaaaactaccacaaagcaacaattgacgatctattattgccaggaaac  
acgaatttaccatggacttcataagctaatttgcgataggcctgcacccttaaggaggaaaaaa  
acatgtcagagttgagagccttcagtgcgccagggaaagcgttactagctggtggatatttagt  
tttagatacaaaaatatgaagcatttgtagtcggattatcggcaagaatgcatgctgtagcccat  
ccttacgggttcattgcaagggtctgataagtttgaaagtgcgtgtgaaaagttaaacaatttaaag  
atggggagtggctgtaccatataagtcctaaaagtggcttcattcctgtttcgataggcggatc  
taagaaccctttcattgaaaaagttatcgctaacgtatttagctactttaaacctaacatggac  
gactactgcaatagaaacttggttcgttattgatattttctctgatgatgcctaccattctcagg  
aggatagcgttaccgaacatcgtggcaacagaagattgagttttcattcgcacagaattgaaga  
agttcccaaaaacaggggtgggctcctcggcaggttttagtcacagttttaactacagctttggcc  
tccttttttgtatcggacctggaaaataatgtagacaaatatagagaagttattcataatttag  
cacaagttgctcattgtcaagctcagggtaaaattggaagcgggtttgatgtagcggcggcagc  
atatggatctatcagatatagaagattcccaccgcattaatctctaatttgccagatattgga  
agtgtacttacggcagtaaaactggcgcattttggttgatgaagaagactggaatattacgatta  
aaagtaaccatttaccttcgggattaactttatggatgggcgatattaagaatggttcagaaac  
agtaaaactggtccagaaggtaaaaaattggtatgattcgcataatgccagaagcttgaaaata



57/167

**Figure 25B**

tatacagaactcgatcatgcaaattctagattttatggatggactatctaaactagatcgcttac  
acgagactcatgacgattacagcgatcagatatttgagtctcttgagaggaatgactgtacctg  
tcaaaagtatcctgaaatcacagaagtttagagatgcagttgccacaatttagacgttcctttaga  
aaaataactaaagaatctggtgccgatatcgaaacctcccgtaaaaactagcttattggatgatt  
gccagaçcttaaaaaggagttcttacttgcttaataacctggtgctggtggttatgacgccattgc  
agtgattactaagcaagatggtgatcttagggctcaaaccgctaatagacaaaagatttttctaag  
gttcaatggctggatgtaactcaggctgactgggggtgttaggaaagaaaaagatccggaaactt  
atcttgataaataacttaaggtagctgcatgcagaattcgcccttaaggaggaaaaaaaatga  
ccgttttacacagcatccgttaccgcacccgtcaacatcgcaacccttaagtattgggggaaaag  
ggacacgaagttgaatctgcccaccaattcgctccatatcagtgactttatcgcaagatgacctc  
agaacgttgacctctgcggtactgcacctgagtttgaacgcgacactttgtggttaaatggag  
aaccacacagcatcgacaatgaaagaactcaaaattgtctgctgcgcgacctacgccaatgaagaa  
ggaaatggaatcgaaggacgcctcattgcccacattatctcaatggaaactccacattgtctcc  
gaaaataactttcctacagcagctgggttttagcttcctccgctgctggctttgctgcattggctct  
ctgcaattgctaagttataccaattaccacagtcacttcagaaatatctagaatagcaagaaa  
ggggtctggttcagcttgtagatcgttgtttgccggatacgtggcctgggaaatgggaaaagct  
gaagatggtcatgattccatggcagtaaaaatcgacagacagctctgactggcctcagatgaaag  
cttggtgctcctagttgtcagcgatattaaaaaggatgtgagttccactcagggtatgcaattgac  
cgtggcaacctccgaactatttaaaagaagaattgaacatgtcgtaccaaaagagatttgaagtc  
atgcgtaaagccattggtgaaaaagatttgcgcacctttgcaaaggaaacaatgatggattcca  
actctttccatgccacatgtttggtactctttccctccaatattctacatgaatgacacttccaa  
gcgtatcatcagttggtgccacaccattaatcagttttacggagaaacaatcgttgcatacacg  
tttgatgcaggtccaaatgctgtgtgtgtactacttagctgaaaatgagtcgaaactctttgcat  
ttatctataaattgtttggtctctgttcctggatgggacaagaaatttactactgagcagcttga  
ggctttcaaccatcaatttgaatcatctaactttactgcacgtgaattggatcttgagttgcaa  
aaggatggtgccagagtgatttttaactcaagtcgggttcaggccccacaagaaacaaacgaatctt  
tgattgacgcaaagactggtctaccaaaaggaataagatcaattcgctgcatcgcccttaggagg  
taaaaaaaaatgactgccgacaacaatagtatgccccatggtgcagtatctagttacgccaaat  
tagtgcaaaaaccaaacacctgaagacatttttggaagagtttcctgaaattattccattacaaca  
aagacctaataccgatctagtgcagcgtcaaatgacgaaagcggagaaacatgtttttctggt  
catgatgaggagcaaattaagttaataatgaaaattgtattgttttggtattgggacgataatg  
ctattggtgccggtaccaagaaagtttgctcatttaattgaaaataattgaaaagggtttactaca  
tcgtgcattctccgtctttattttcaatgaacaagggtgaattacttttacaacaaagagccact  
gaaaaaataactttccctgatctttggactaacacatgctgctctcatccactatgtattgatg  
acgaattagggttgagggttaagctagacgataagattaagggcgctattactgcggcggtgag  
aaaactagatcatgaattagggtattccagaagatgaaactaagacaaggggtaagtttcacttt  
ttaaacagaatccattacatggcaccaagcaatgaaccatgggggtgaacatgaattgattaca  
tcctattttataagatcaacgctaaagaaaacttgactgtcaacccaaacgtcaatgaagttag  
agacttcaaatgggtttcaccaaattgatttgaaaactatgtttgctgacccaagttacaagttt  
acgccttggtttaagattatttgcgagaattacttattcaactggtgggagcaattagatgacc  
tttctgaagtggaaaatgacaggcaaattcatagaatgctataacaacgcgtcctgcattcgcc  
cttaggaggtaaaaaaacatgtgtgcgacctcttctcaatttactcagattaccgagcataatt  
cccgctcggtccgcaaaactatcagccaaacctgtggaatttcgaattcctgcaatccctggagaa  
cgacctgaaagtggaaaagctggaggagaaagcgaccaaaactggaggaagaagttcgctgcatg



**Figure 25C**

atcaaccgtgtagacacccagccgctgtccctgctggagctgatcgacgatgtgcagcgcctgg  
gtctgacctacaaatTTGAAAAAGACATCATTAAAGCCCTGGAAAACATCGTACTGCTGGACGA  
AAACAAAAAGAACAAATCTGACCTGCACGCAACCGCTCTGTCTTTCGGTCTGCTGCGTCAGCAC  
GGTTTCGAGGTTTCTCAGGATGTTTTTGAGCGTTTCAAGGATAAAGAAGGTGGTTTCAGCGGTG  
AACTGAAAGGTGACGTCCAAGGCCTGCTGAGCCTGTATGAAGCGTCTTACCTGGGTTTCGAGGG  
TGAGAACCTGCTGGAGGAGGCGCGTACCTTTTCCATCACCCACCTGAAGAACAACCTGAAAGAA  
GGCATTAAATACCAAGGTTGCAGAACAAGTGAGCCACGCCCTGGAAGTGCATATCACCGCGTC  
TGCAACCGTCTGGAGGCACGTTGGTTCCTGGATAAATACGAACCGAAAGAACCGCATCACCGAGCT  
GCTGCTGGAGCTGGCGAAGCTGGATTTTAAACATGGTACAGACCCTGCACCAGAAAGAGCTGCAA  
GATCTGTCCCGCTGGTGGACCGAGATGGGCCTGGCTAGCAAACCTGGATTTTGTACGCGACCGCC  
TGATGGAAGTTTATTTCTGGGCACTGGGTATGGCGCCAGACCCGCGAGTTTGGTGAATGTCGCAA  
AGCTGTTACTAAAATGTTTGGTCTGGTGACGATCATCGATGACGTGTATGACGTTTATGGCACT  
CTGGACGAAGTGAAGTGTTCACCGATGCTGTAGAGCGCTGGGACGTTAACGCTATTAAACACC  
TGCCGGACTATATGAAACTGTGTTTCTGGCACTGTACAACACCGTTAACGACACGTCCTATT  
TATTCTGAAAGAGAAAGGTCATAACAACCTGTCCTATCTGACGAAAAGCTGGCGTGAAGTGTGC  
AAAGCCTTTCTGCAAGAGGCGAAATGGTCCAACAACAAAATTATCCCGGCTTTCTCCAAGTACC  
TGGAACACGCCAGCGTTTCTCTCTCGGTGTAGCGCTGCTGGCGCCGTCTTACTTTTCCGTATG  
CCAGCAGCAGGAAGACATCTCCGACCACGCGCTGCGTTCCTGACCGACTTCCATGGTCTGGTG  
CGTTCTAGCTGCGTTATCTTCCGCCTGTGCAACGATCTGGCCACCTCTGCGGCGGAGCTGGAAC  
GTGGCGAGACTACCAATTCTATCATTAGCTACATGCACGAAAACGATGGTACCAGCGAGGAACA  
GGCCCGCGAAGAAGTGCCTAAACTGATCGACGCCGAATGGAAAAAGATGAATCGTGAACGCGTT  
AGCGACTCCACCCTGCTGCCTAAAGCGTTCTATGGAAATCGCAGTTAACATGGCACGTGTTTCCC  
ACTGCACCTACCAGTATGGCGATGGTCTGGGTGCGCCAGACTACGCGACTGAAAACCGCATCAA  
ACTGCTGCTGATTGACCCTTTCCCGATTAACCAGCTGATGTATGTCTAACTGCAGCTGGTACCA  
TATGGGAATTCGAAGCTTTCTAGAACAAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTC  
GACCATCATCATCATCATTGAGTTTAAACGGTCTCCAGCTTGGCTGTTTTGGCGGATGAGA  
GAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCGAGAAGCGGTCTGATAAACAGAATTTG  
CCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAAGTGAAGTGAACGCGCGTA  
GCGCCGATGGTAGTGTGGGTCTCCCCATGCGAGAGTAGGGAAGTGCAGGCATCAAATAAAAC  
GAAAGGCTCAGTCGAAAGACTGGGCCTTTCTGTTTTATCTGTTGTTTGTGCGGTGAACGCTCTCT  
GAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGG  
GCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCC  
TTTTTGCGTTTTCTACAAACTCTTTTTGTTTTATTTTTCTAAATACATTCAAATATGTATCCGCTT  
AACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTGGGAAGCCCTGCAAAGTAAACTGGAT  
GGCTTTCTCGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGATGA  
GGATCGTTTTCGATGATTGAACAAGATGGATTGCACGCGAGGTTCTCCGGCCGCTTGGGTGGAGA  
GGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCT  
GTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCTGAATGAAGT  
CAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCG  
ACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCAGGGGCGAGGATCTCT  
GTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCAT  
ACGCTTGATCCGGCTACCTGCCATTGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTA  
CTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCC  
AGCCGAAGTGTTCGCCAGGCTCAAGGCGAGCATGCCCGACGGCGAGGATCTCGTCGTGACCCAT



59/167

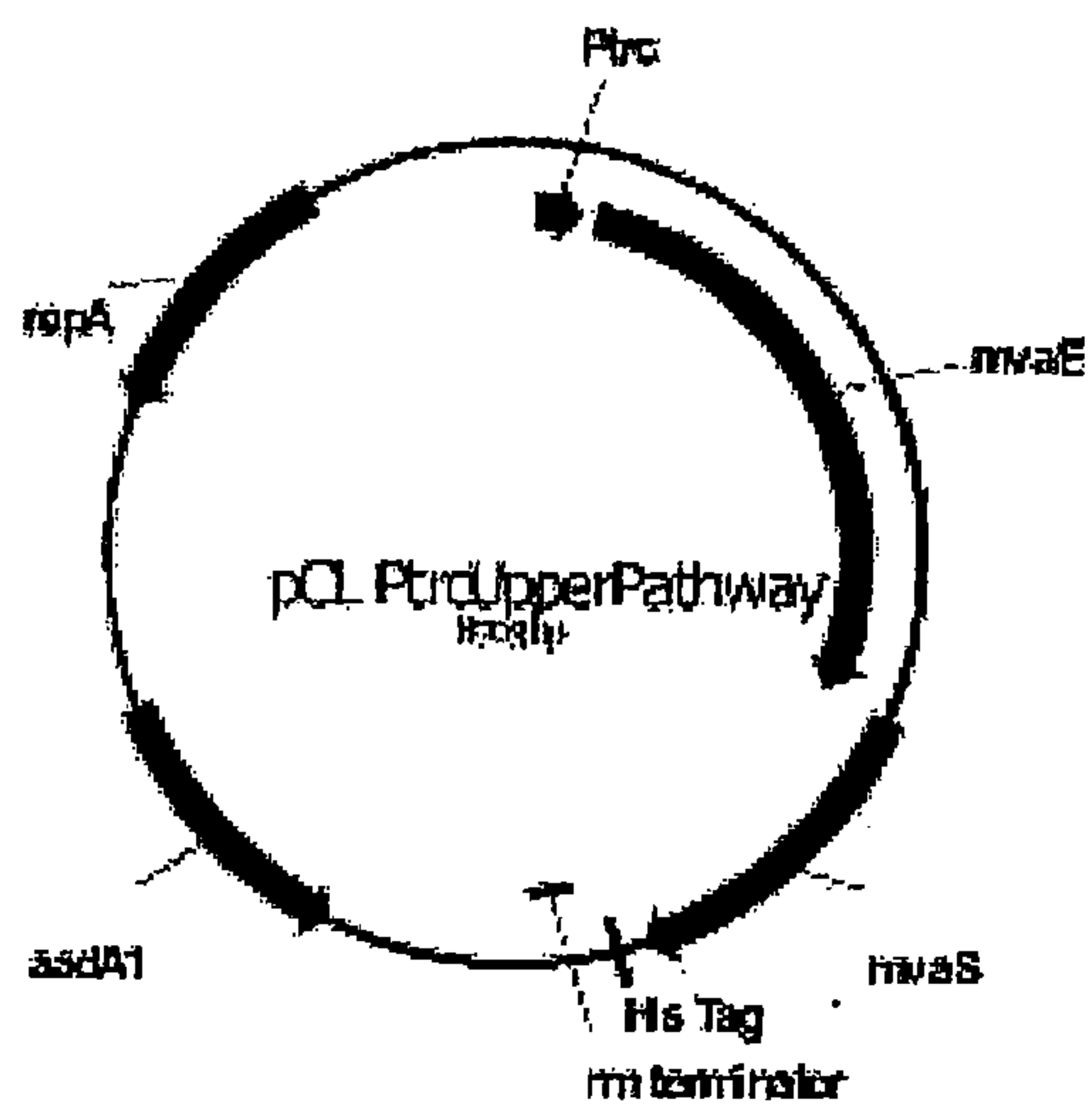
**Figure 25D**

ggcgatgcctgcttgccgaatatcatggtggaaaatggccgctttttctggattcatcgactgtg  
gccggctgggtgtggcggaccgctatcaggacatagcgttggctaccgctgatattgctgaaga  
gcttggcggcgaatgggctgaccgcttcctcgtgctttacgggtatcgccgctcccgattcgcag  
cgcatcgcttctatcgcttcttgacgagttcttctgacgcatgacccaaaatcccttaacgtg  
agttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttt  
ttttctgcgcgtaatctgctgcttgcaacaaaaaaaccaccgctaccagcgggtggtttggttg  
ccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaa  
atactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctac  
atacctcgctctgctaactctgttaccagtgggtgctgccagtggcgataagtcgtgtcttacc  
gggttggactcaagacgatagttaccggataaggcgcagcggctcgggctgaacgggggggttcgt  
gcacacagcccagcttggagcgaacgacctaaccgaactgagatacctacagcgtgagctatg  
agaaagcgcacgcttcccgaaggagaaaggcggacaggtatccggtaagcggcagggtcggga  
acaggagagcgcacgagggagcttccagggggaaacgcctgggtatctttatagtcctgtcgggt  
ttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggaggagcctatggaa  
aaacgccagcaacgcggcctttttacgggttcctggccttttgccttttgcctcacatgttc  
tttctgcttctatccctgattctgtggataaccgtattaccgcctttgagtgagctgataccg  
ctcgccgcagccgaacgaccgagcgcagcagtcagtgagcaggaagcgggaagagcgcctgat  
gcggtatttttctccttacgcatctgtgcggtatttcacaccgcatatgggtgcactctcagtaca  
atctgctctgatgccgcatagttaagccagtatacactccgctatcgctacgtgactgggtcat  
ggctgcgccccgacacccgccaacacccgctgacgcgcctgacgggcttgtctgctcccggca  
tccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagagggttttcaccgctcat  
caccgaaacgcgcgaggcagcagatcaattcgcgcgcgaaggcgaagcggcatgcatttacgtt  
gacaccatcgaatgggtgcaaaacctttcgcggtatggcatgatagcgcgccggaagagagtcaat  
tcaggggtggtgaatgtgaaaccagtaacgttatacgatgtcgcagagtatgccggtgtctctta  
tcagaccgtttcccgcgtggtgaaccaggccagccacgtttctgcgaaaacgcgggaaaaagt  
gaagcggcgatggcggagctgaattacattcccaaccgcgtggcacaacaactggcggggcaaac  
agtcggttgcctgattggcgttgccacctccagtctggccctgcacgcgcctgcgcaaattgtcgc  
ggcgattaaatctcgcgcgcgatcaactgggtgccagcgtgggtggtgtcgatggtagaacgaagc  
ggcgtcgaagcctgtaaagcggcgggtgcacaatcttctcgcgcaacgcgtcagtgggctgatca  
ttaactatccgctggatgaccaggatgccattgctgtggaagctgcctgcactaatgttccggc  
gttattttcttgatgtctctgaccagacacccatcaacagttatttttctcccatgaagacgggt  
acgcgactgggcgtggagcatctggtcgcattgggtcaccagcaaatcgcgctgttagcgggccc  
cattaagtctgtctcggcgcgtctgcgtctggctggctggcataaatatctcactcgcaatca  
aattcagccgatagcgggaacgggaaggcgactggagtgccatgtccgggttttcaacaaaccatg  
caaatgctgaatgagggcatcgcttcccactgcgatgctgggttgccaacgatcagatggcgcctgg  
gcgcaatgcgcgccattaccgagtcggggctgcgcgttgggtgcggatatctcggtagtgggata  
cgacgataccgaagacagctcatgttatatcccggcgtcaaccaccatcaaacaggattttcgc  
ctgctggggcaaacagcgtggaccgcttgcctgcaactctctcagggccaggcgggtgaagggca  
atcagctgttgccgctctcactgggtgaaaagaaaaaccacctggcgcccaatacgcaaacgc  
ctctccccgcgcgttggccgattcattaatgcagctggcagcagaggtttcccgactggaaagc  
gggcagtgagcgcgaacgcaattaatgtgagttagcgcgaattgatctg  
(SEQ ID NO:33)



60/167

Figure 26





61/167

**Figure 27A**

5' -

cccgtcttactgtcgggaattcgcggttgccgattcattaatgcagattctgaaatgagctggt  
gacaattaatcatccggctcgtataatgtgtggaattgtgagcggataacaatttcacacagga  
aacagcgccgctgagaaaaagcgaagcggcactgctctttaacaatttatcagacaatctgtgt  
gggcactcgaccggaattatcgattaactttattattaaaaattaaagaggtatatattaatgt  
atcgattaaataaggaggaataaaccatggatccgagctcaggaggtaaaaaaacatgaaaaca  
gtagttattattgatgcattacgaacaccaattggaaaatataaaggcagcttaagtcaagtaa  
gtgccgtagacttaggaacacatgttacaacacaacttttaaaaagacattccactattttctga  
agaaattgatcaagtaatctttggaaatgttttacaagctggaaatggccaaaatcccgcacga  
caaatagcaataaacagcggtttgtctcatgaaattcccgcaatgacggttaatgaggtctgcg  
gatcaggaatgaaggccgttattttggcgaaacaattgattcaattaggagaagcgggaagtttt  
aattgctggcgggattgagaatatgtcccaagcacctaaattacaacgtttttaattacgaaaca  
gaaagctacgatgcgcctttttctagtatgatgtatgatggattaacgggatgccttttagtggtc  
aggcaatgggcttaactgctgaaaatgtggccgaaaagtatcatgtaactagagaagagcaaga  
tcaatttttctgtacattcacaattaaaagcagctcaagcacaagcagaagggatattcgctgac  
gaaatagccccattagaagtatcaggaacgcttgtggagaaagatgaagggattcgccttaatt  
cgagcgttgagaagctaggaacgcttaaaacagtttttaagaagacgggtactgtaacagcagg  
gaatgcatcaaccattaatgatggggcttctgctttgattattgcttcacaagaatatgccgaa  
gcacacgggtcttccttattttagctattattcgagacagtgtggaagtcgggtattgatccagcct  
atatgggaatttcgccgattaaagccattcaaaaactgttagcgcgcaatcaacttactacgga  
agaaattgatctgtatgaaatcaacgaagcatttgcagcaacttcaatcgtgggtccaaagagaa  
ctggcttttaccagaggaaaaggtcaacatttatgggtggcgggtatttcattaggtcatgcgattg  
gtgccacaggtgctcgtttattaacgagtttaagttatcaattaaatcaaaaagaaaagaaata  
tggagtggcttctttatgtatcggcggtggcttaggactcgctatgctactagagagacctcag  
caaaaaaaaaaacagccgatttttatcaaatgagtcctgaggaacgcctggcttctcttcttaatg  
aaggccagatttctgctgatacaaaaaaagaatttgaaaatacggctttatcttcgcagattgc  
caatcatatgattgaaaatcaaatcagtgaacacagaagtgccgatgggcggttggttacattta  
acagtggacgaaactgattattttggtaccaatggcgacagaagagccctcagttattgcggtt  
tgagtaatggtgcaaaaatagcacaaggattttaaacagtgaaatcaacaacgcttaatgcgtgg  
acaaatcggttttttacgatgttgagatcccagatcattgattgataaactacaagtaagagaa  
gcggaagtttttcaacaagcagagtttaagttatccatctatcgttaaacggggcgggcggttaa  
gagatttgcaatatcgtaacttttgatgaatcatttgtatctgtcgaacttttttagtagatgttaa  
ggatgcaatgggggcaaatatcgttaacgctatgttggaagggtgtggccgagttgttccgtgaa  
tggtttgcgagcaaaaagattttattcagtatttttaagtaattatgccacggagtcggttggtta  
cgatgaaaacggctattccagtttcacgtttaagtaaggggagcaatggccgggaaattgctga  
aaaaattgttttagcttcacgctatgcttcatttagatccttatcgggcagtcacgcataaaaa  
ggaatcatgaatggcattgaagctgtagtttttagctacaggaaatgatacacgcgctgttagcg  
cttcttgctcatgcttttgcggtgaaggaaggctcgctaccaaggcttgactagttggacgctgga  
tggcgaacaactaattggtgaaatttcagttccgcttgcttttagccacgggttgccggtgccaca  
aaagtcttacctaaatctcaagcagctgctgatttgttagcagtgacggatgcaaaagaactaa  
gtcgagtagtagcggctgttggtttggcacaataatttagcggcggttacgggccttagtctctga  
aggaattcaaaaaggacacatggctctacaagcacgttcttttagcgatgacggtcggagctact  
ggtaaagaagttgaggcagtcgctcaacaattaaaacgtcaaaaaacgatgaaccaagaccgag  
ccatggctatttttaaatgatttaagaaaacaataaaggaggtaaaaaaacatgacaattgggat



62/167

**Figure 27B**

tgataaaattagtttttttgtgcccccttattatattgatatgacggcactggctgaagccaga  
aatgtagaccctggaaaatttcatattgggtattgggcaagaccaaattggcggtgaacccaatca  
gccaagatattgtgacatttgcagccaatgccgcagaagcgatcttgaccaagaagataaaga  
ggccattgatatgggtgattgtcgggactgagtccagtatcgatgagtcaaaagcggccgcagtt  
gtcttacatcgtttaattggggattcaacctttcgctcgctctttcgaaatcaaggaagcttggt  
acggagcaacagcaggcttacagtttagctaagaatcacgtagccttacatccagataaaaaagt  
cttggctcgtagcggcagatattgcaaaatatggcttaaattctggcggtgagcctacacaagga  
gctggggcggttgcaatgtagttgctagtgaaccgcgcattttggctttaaaagaggataatg  
tgatgctgacgcaagatatctatgacttttggcgctccaacaggccaccggtatcctatggctga  
tggtcctttgtcaaacgaaacctacatccaatcttttgcccaagtctgggatgaacataaaaaa  
cgaaccggtcttgattttgcagattatgatgcttttagcgttccatattccttacacaaaaatgg  
gcaaaaaagccttatttagcaaaaatctccgaccaaaactgaagcagaacaggaacgaatttttagc  
ccgttatgaagaaagtatcgctctatagtcgctcgctaggaacttggtatacgggttcactttat  
ctgggactcatttcccttttagaaaaatgcaacgactttaaccgcaggcaatcaaattggtttat  
tcagttatggttctggtgctgctgctgaatttttactggtgaattagtagctggttatcaaaa  
tcatttacaaaaagaaactcatttagcactgctggataatcggacagaactttctatcgctgaa  
tatgaagccatgtttgcagaaacttttagacacagacattgatcaaacgttagaagatgaattaa  
aatatagtattttctgctattaataataaccgttcgttcttatcgaaactaagagatctgcagctg  
gtaccatattgggaattcgaagcttgggcccgaacaaaaactcatctcagaagaggatctgaata  
gcgccgtcgaccatcatcatcatcattgagtttaaaccggtctccagcttggctgttttggc  
ggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggctctgataaaac  
agaatttgctggcggcagtagcgcggtggtcccacctgaccccatgccgaactcagaagtga  
acgccgtagcgcgatggtagtggtgggtctcccatgcgagagtagggaactgccaggcatca  
aataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttggttgctcggtgaac  
gctctcctgagtaggacaaatccgccgggagcggatttgaacggttgcaagcaacggcccggag  
ggtggcgggcaggacgcccgcataaaactgccaggcatcaaattaagcagaaggccatcctgac  
ggatggcctttttgcgtttctacaaactctttttgtttatttttctaaatacattcaaatatgt  
atccgctcatgagacaataaccctgataaatgcttcaataatctggcgtaatagcgaagaggcc  
cgcaccgatcgcccttcccaacagttgcgcgacctgaatggcgaatggcgacctgatgcggtatt  
ttctccttacgcatctgtgcggtattttcacaccgcatatggtgcactctcagtacaatctgctc  
tgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgagcttagtaaagcc  
ctcgctagatttttaattgcggatggtgcgattacttcgccaactattgcgataacaagaaaaagc  
cagcctttcatgatatatctcccaatttgtgttagggcttattatgcacgcttaaaaaataaaa  
agcagacttgacctgatagtttggctgtgagcaattatgtgcttagtgcatctaacgcttgagt  
taagccgcgcgcgaagcggcgctcggcttgaacgaattgtagacattatttgccgactacctt  
ggtgatctcgcttttcacgtagtggaacaaattcttccaactgatctgcgcgcgaggccaagcga  
tcttcttcttgtccaagataagcctgtctagcttcaagtatgacgggctgatactgggcccggca  
ggcgctccattgcccagtcggcagcgacatccttcggcgcgattttgccgggttactgcgctgta  
ccaaatgcgggacaacgtaagcactacatttcgctcatcgccagcccagtcgggcccgcgagttc  
catagcggttaaggtttcathtagcgctcaaatagatcctgttcaggaaccggatcaaagagtt  
cctccgcgcgtggacctaccaaggcaacgctatgttctcttgcttttgtcagcaagatagccag  
atcaatgtcgatcgtggctggctcgaagataacctgcaagaatgtcattgcgctgccattctcca  
aattgcagttcgcgcttagctggataacgccacggaatgatgtcgtcgctgcacaacaatggtga  
cttctacagcgcggagaatctcgctctctccaggggaagccgaagtttccaaaaggctcggtgat



63/167

**Figure 27C**

caaagctcgcgcggtgtttcatcaagccttacgggtcacccgtaaccagcaaatcaatatcactg  
tgtggcttcaggccgcatccactgaggagccgtacaaatgtacggccagcaacgtcgggttcga  
gatggcgctcgatgacgccaactacctctgatagttgagtcgatacttcggcgatcacccgttc  
cctcatgatgtttaactttgttttagggcgactgccctgctgcgtaacatcgttgctgctccat  
aacatcaaacatcgacccacggcgtaacgcgcttgctgcttggtatgcccagggcatagactgta  
ccccaaaaaacagtcataacaagccatgaaaaccgcccactgcgcggttaccaccgctgcgttc  
ggtcaagggttctggaccagttgcgtgagcgcatagctacttgcatcagcttacgaaccgaa  
caggcttatgtccactgggttcgtgccttcacccgttccacgggtgtgcgtcacccggcaacct  
tgggcagcagcgaagtcgaggcatttctgtcctggctggcgaacgagcgcgaagggttcgggtctc  
cacgcacgtcaggcattggcgcccttgctgttcttctacggcaagggtgctgtgcacggatctg  
ccctggcttcaggagatcggaagacctcgccgctcgccggcgttgccgggtggtgctgacccggg  
atgaagtgggttcgcatcctcgggttttctggaaggcgagcatcgtttgttcgccagcttctgta  
tggaacggggcatgcggatcagtgagggtttgcaactgcgggtcaaggatctggatttcgatcac  
ggcacgatcatcgtgcgggagggcaagggtccaaggatcgggccttgatgttacccgagagct  
tggcaccagcctgcgcgagcaggggaattaattcccaagggttttgctgcccgcaaacgggct  
gttctggtgttgctagtttgttatcagaatcgagatccggcttcagccggtttgcgggtgaa  
agcgctatttcttccagaattgcatgatttttccccacgggagggcgtcactgggtcccgtgt  
tgtcggcagctttgattcgataagcagcatcgccgttttcaggctgtctatgtgtgactgttga  
gctgtaacaagttgtctcaggtgttcaatttcatgttctagttgctttgttttactggtttcac  
ctgttctattaggtgttacatgctgttcatctgttacattgtcgtatctgttcatggtgaacagc  
tttgaatgcacaaaaactcgtaaaagctctgatgtatctatcttttttacaccggttttcatct  
gtgcataatggacagttttccctttgatatgtaacgggtgaacagttgttctacttttgtttgtta  
gtcttgatgcttcactgatagatacaagagccataagaacctcagatccttccgtatttagcca  
gtatgttctctagtggtgttcgttgtttttgcgtgagccatgagaacgaaccattgagatcata  
cttactttgcatgtcactcaaaaattttgcctcaaaaactgggtgagctgaatttttgcagttaaa  
gcatcgtgtagtggttttcttagtccgttatgtaggttaggaatctgatgtaatgggtgttggtta  
ttttgtcaccattcatttttatctggttggttctcaagttcgggttaacgagatccatttgtctatc  
tagttcaacttggaataatcaacgtatcagtcgggcggcctcgcttatcaaccaccaatttcata  
ttgctgtaagtgttttaaatctttacttattgggtttcaaaaccattgggttaagccttttaact  
catggtagttattttcaagcattaacatgaacttaaatcatcaaggctaattctctatatttgc  
cttgtagttttcttttgtgttagttcttttaataaccactcataaatcctcatagagtatttg  
ttttcaaaagacttaacatgttccagattatattttatgaatttttttaactggaaaagataag  
gcaatatctcttcaactaaaaactaatttctaatttttcgcttgagaacttggcatagtttgtcca  
ctggaaaatctcaaagcctttaaccaaaaggattcctgatttccacagttctcgtcatcagctct  
ctgggttgcttttagctaatacaccataagcattttccctactgatgttcatcatctgagcgtatt  
gggtataagtgaacgataaccgtccgttctttccttgtaggggttttcaatcgtgggggttagtag  
tgccacacagcataaaaattagcttggtttcatgctccgttaagtcatacgactaatcgctagt  
tcatttgctttgaaaacaactaattcagacatacatctcaattgggtctaggtgattttaatcac  
tataccaattgagatgggctagtcaatgataattactagtccttttcccttgagttgtgggtat  
ctgtaaattctgctagacctttgctggaaaacttgtaaattctgctagacctctgtaaattcc  
gctagacctttgtgtgtttttttgttttatattcaagtgggtataatttatagaataaagaaag  
aataaaaaaagataaaaagaatagatcccagccctgtgtataactcactacttttagtcagttcc  
gcagtattacaaaaggatgtcgaaacgctgtttgctcctctacaaaacagaccttaaaacct  
aaaggcttaagtagcaccctcgcaagctcgggcaaatcgctgaatattccttttgtctccgacc



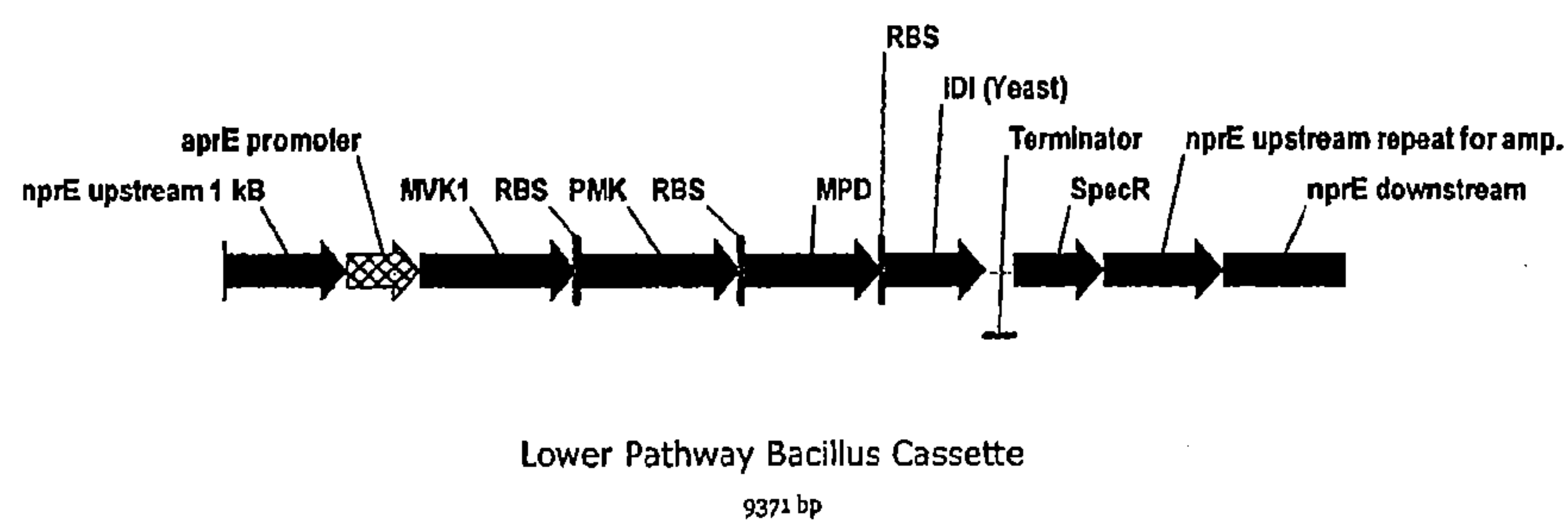
64/167

**Figure 27D**

Atcagggcacctgagtcgctgtcttttttcgtgacattcagttcgctgcgctcacggctctggcag  
tgaatgggggtaaatggcactacagggcgctttttatggattcatgcaaggaaactacccataat  
acaagaaaagcccgtcacgggcttctcagggcgttttatggcgggtctgctatgtgggtgctatc  
tgacttttttgctgttcagcagttcctgccctctgattttccagtctgaccacttcggattatcc  
cgtgacaggtcattcagactggctaatagcacccagtaaggcagcggtatcatcaacaggctta  
(SEQ ID NO:46)



65/167

**Figure 28**



66/167

**Figure 29A**

5' -

tgtaaccttttgctttcaaatagagtagaaataatgcacatccatgtttgtatcgtgcaaataaag  
tgtttcatccgtaggaaaaaatgacttttagtatctgttccgctttttctgatgaaatgtgctcc  
ccgacaaaattgaatgaatcatggacatttgctggctttgatacagcgaaagcagccgttccta  
tgttatatatcggatttaacagcaggacaaaaaacaccatgacagccatcgtcacccacttatt  
cacacgcacataaaccttttctgacttttggaacagatgatagctcatcaaaaatcccgccatt  
gccaaataaatcgtatatggcattactgcaccataatcttttgagatttgattgggatatggcg  
caagcagcaagacaagcagtcgataatcagcgtataaaataagcctagtaagatcttatccgt  
tctccaatacagcttgaaaaaactacattcaacgcaatgggaagagtgatgatgaaaaacaga  
aacacgaatgcaatcggctccatcccatccgggtattccttccaatacgaagaaactaaaaa  
tcatttgtacgatcggcaaacctgacaacagcaaggctcgaacgtataaaacttaccctttccgcc  
atgatcacgcggcatcagcatatagtgaagagccgtcagcagcacatatccgtataacaaaaa  
tgacagcagcggcagcagttcttttccgctcctctcttaagtaagcgtggtgaagtttggtgatt  
gcacctggtgaataagttcaacagacactcccgccagcagcacatccgcaatataacaccgc  
caagaacattgtgctgctgccggtttattttgggatgatgcaccaaagatataagcccgccaga  
acaacaattgaccattgaatcagcaggggtgctttgtctgcttaataaaaaataacgttcgaat  
gcaatacataatgactgaataactccaacacgaacaacaactccattttcttctgctatcaaaa  
taacagactcgtgattttccaaacgagctttcaaaaaagcctctgccccttgcaaatcggatgc  
ctgtctataaaattcccgatatttggttaaacagcggcgcaatggcgggccgcatctgatgtcttt  
gcttggcgaatgttcattcttatttcttccctctcaataattttttcattctatcccttttc  
tgtaaagtttatttttccagaatacttttatcatcatgctttgaaaaaatatcacgataatatcc  
attgttctcacggaagcacacgcaggtcatttgaacgaattttttcgacaggaatttgccggga  
ctcaggagcatttaacctaaaaaagcatgacatttcagcataatgaacatttactcatgtctat  
tttcggttcttttctgtatgaaaatagttatttcgagctctctacggaaatagcgagagatgatat  
acctaaatagagataaaatcatctcaaaaaaatgggtctactaaaatattattccatctattac  
aataaattcacagaatagtcttttaagtaagtctactctgaatttttttaaaaggagagggtaa  
agagtgtcattaccgttcttaacttctgcaccgggaaagggttattatttttgggtgaacactctg  
ctgtgtacaacaagcctgccgtcgtgctagtgtgtctgcgttgagaacctacctgctaataag  
cgagtcatctgcaccagatactattgaattggacttcccgacatttagctttaatcataagtg  
tccatcaatgatttcaatgccatcacccaggatcaagtaaactcccaaaaattggccaaggctc  
aacaagccaccgatggcttgtctcaggaactcgttagtcttttggtatccgttggttagctcaact  
atccgaatccttccactaccatgcagcgttttggttctctgtatatgtttgtttgcctatgcccc  
catgccaagaatattaagttttctttaaggtctactttacccatccgtgctgggttgggtctcaa  
ggcctctattttctgtatcactggccttagctatggcctacttgggggggttaataggatctaa  
tgacttggaaaagctgtcagaaaacgataagcatatagtgaatcaatgggccttcataagtgaa  
aagtgtattcacggtacccttcaggaatagataacgctgtggccacttatggtaatgccctgc  
tatttgaaaaagactcacataatggaacaataaacacaaacaattttaagttcttagatgattt  
cccagccattccaatgatcctaactataactagaattccaagggtctacaaaagatcttggtgct  
cgcgcttcgtgtgttgggtcacccagaaaatttctgaagttatgaagccaattctagatgccatgg  
gtgaatgtgcctacaaggcttagagatcatgactaagtttaagtaaatgtaaaggcaccgatga  
cgaggctgtagaaactaataatgaactgtatgaacaactattggaattgataagaataaatcat  
ggactgcttgtctcaatcgggtgtttctcatcctggattagaacttattaaaaatctgagcgatg  
atttgagaattgggtccacaaaacttaccgggtgctgggtggcgggcgggttgcctttgactttgtt  
acgaagagacattactcaagagcaaatgacagcttcaaaaagaaattgcaagatgattttagt



67/167

**Figure 29B**

tacgagacattttgaaacagacttgggtgggactggctgctgtttgttaagcgcaaaaaatttga  
ataaagatcttaaaatcaaataccctagtagtattccaattattttgaaaataaaaactaccacaaagca  
acaaattgacgatctattattgccaggaaacacgaattttaccatggacttcataaaaggagagg  
gtgtcagagttgagagccttcagtgtcccagggaagcggttactagctgggtggatatttagttt  
tagatacaaaaatatgaagcattttgtagtcgggattatcggcaagaatgcatgctgtagcccatcc  
ttacgggttcattgcaagggctctgataagttttgaagtgcgtgtgaaaagtaaacattttaagat  
ggggagtggctgtaccatataagtcctaaaagtggttcattcctgtttcgataggcggatcta  
agaaccctttcattgaaaaagttatcgctaactgtatttagctactttaaacctaactggacga  
ctactgcaatagaaaacttggttcgttattgatattttctctgatgatgcctaccattctcaggag  
gatagcgttaccgaacatcgtggcaacagaagattgagttttcattcgcacagaattgaagaag  
ttcccaaaacagggctgggctcctcggcagggttttagtcacagtttttaactacagctttggcctc  
cttttttgtatcggacctggaaaataatgtagacaaatatagagaagttattcataatttagca  
caagttgctcattgtcaagctcagggtaaaaattggaagcgggtttgatgtagcggcggcagcat  
atggatctatcagatatagaagattcccacccgcattaatctctaatttgccagatattggaag  
tgctacttacggcagtaaaactggcgcattttggttgatgaagaagactggaatattacgattaaa  
agtaaccatttaccttcgggattaactttatggatgggogatatataagaatggttcagaaacag  
taaaactgggtccagaaggtaaaaaattggtatgattcgcataatgccagaaagcttgaaaatata  
tacagaactcgatcatgcaaattctagattttatggatggactatctaaactagatcgcttacac  
gagactcatgacgattacagcgatcagatatttgagtctcttgagaggaatgactgtacctgtc  
aaaagtatcctgaaatcacagaagttagagatgcagttgccacaattagacgttccttttagaaa  
aataactaaagaatctggtgccgatatcgaacctcccgtacaaactagcttattggatgattgc  
cagaccttaaaaggagttcttacttgcttaataacctgggtgctgggtggttatgacgccattgcag  
tgattactaagcaagatggtgatcttagggctcaaaccgctaatagacaaaagattttctaaagg  
tcaatggctggatgtaactcaggctgactggggtggttaggaaagaaaaagatccggaaacttat  
cttgataaataaaaggagaggggtgacctttacacagcatccgttaccgcacccgtcaacatcg  
caacccttaagtattgggggaaaaggacacgaagttgaatctgccaccaattcgtccatctc  
agtgactttatcgcaagatgacctcagaacggtgacctctcgggctactgcacctgagtttgaa  
cgcgacactttgtggttaaattggagaaccacacagcatcgacaatgaaagaactcaaaattgtc  
tgcgcgacctacgccaattaagaaaggaaatggaatcgaaggacgcctcattgccacattatc  
tcaatggaaactccacattgtctccgaaaataactttcctacagcagctgggttagcttcctcc  
gctgctggctttgctgcattgggtctctgcaattgctaagttataccaattaccacagtcaactt  
cagaaatatctagaatagcaagaaaggggtctggttcagcttgtagatcggttggttgccggata  
cgtggcctgggaaatgggaaaagctgaagatgggtcatgattccatggcagtagacaaatcgacagac  
agctctgactggcctcagatgaaagcttggtgtcctagttgtcagcgatattaaaaaggatgtga  
gttccactcaggggtatgcaattgacctggcaacctccgaactatttaaaagaagaattgaaca  
tgtcgtaccaaagagattttgaagtcatgcgttaaagccattggtgaaaaagatttcgccaccttt  
gcaaaggaaacaatgatggattccaactctttccatgccacatggtttggactctttccctccaa  
tattctacatgaatgacacttccaagcgtatcatcagttggtgccacaccattaatcagtttta  
cggagaaacaatcggttgcatcacggtttgatgcagggtccaaatgctgtgtgtactacttagct  
gaaaatgagtcgaaactctttgcatttatctataaattggttggtctgttccctggatgggaca  
agaaatttactactgagcagcttgaggctttcaaccatcaatttgaaatcatctaactttactgc  
acgtgaattggatcttgagttgcaaaaggatggtgccagagtgatttttaactcaagtcgggttca  
ggcccaagaagaacaaacgaatctttgattgacgcaaagactggtctaccaaaaggaataaaagg  
agaggggtgactgccgacaacaatagtagtgcctccatgggtgcagtagttagttacgccaatttagt



**Figure 29C**

gcaaaaccaaacacctgaagacatTTTTggaagagtttcctgaaattattccattacaacaaaga  
cctaatacccgatctagtgagacgtcaaattgacgaaagcggagaaacatgtttttctgggtcatg  
atgaggagcaaattaagttaattgaatgaaaattgtattgttttggattgggacgataatgctat  
tgggtgccggtaccaagaaagtttgtcatttaattgaaaatattgaaaagggtttactacatcgt  
gcattctccgtctttatTTTTcaatgaacaagggtgaattacttttacaacaaagagccactgaaa  
aaataactttccctgatctttggactaacacatgctgctctcatccactatgtattgatgacga  
attagggtttgaagggttaagctagacgataagattaaggggcgctattactgcggcggtgagaaaa  
ctagatcatgaattagggtattccagaagatgaaactaagacaaggggtaagtttcactttttaa  
acagaatccattacatggcaccaagcaatgaaccatgggggtgaacatgaaattgattacatcct  
atTTTtataagatcaacgctaaagaaaacttgactgtcaacccaaacgtcaatgaagttagagac  
ttcaaattgggtttcaccaaattgatttgaaaactatgtttgctgacccaagttacaagtttacgc  
cttgggtttaagattatttgcgagaattacttattcaactgggtgggagcaattagatgacctttc  
tgaagtggaaaatgacaggcaattcatagaatgctataaaaaaacggccttggccccgccg  
gttttttatttatttttcttcccgcatgttcaatccgctccataatcgacggatggctccctc  
tgaaaatttttaacgagaaacggcggttgacccggctcagtcctcgtaacggccaagtccctgaaa  
cgtctcaatcgccgcttcccggtttccgggtcagctcaatgocgtaacgggtcggcggtgtttcc  
tgataccggggagacggcattcgtaatTTgaatacatacgaacaaattaataaagtgaaaaaat  
acttcggaaacattttaaaaaataaccttattgggtacttacatgtttggatcaggagttgagagt  
ggactaaaaccaaattagtgatcttgacttttttagtcgctcgtatctgaaccattgacagatcaaa  
gtaaagaaatacttatacaaaaaatttagacctatttcaaaaaaaataggagataaaagcaactt  
acgatatattgaattaacaattattattcagcaagaaatgggtaccgtggaatcatcctcccaaa  
caagaattttatttatggagaatgggttacaagagctttatgaacaaggatacattcctcagaagg  
aattaaattcagatttaaccataatgctttaccaagcaaaacgaaaaataaaagaatatcgg  
aaattatgacttagaggaattactacctgatattccattttctgatgtgagaagagccattatg  
gattcgtcagaggaattaatagataattatcaggatgatgaaaccaactctatattaactttat  
gccgtatgatttttaactatggacacgggtaaaatcataccaaaagatattgcgggaaatgcagt  
ggctgaatcttctccattagaacatagggagagaattttgttagcagttcgtagttatcttggga  
gagaatattgaatggactaatgaaaatgtaaatttaactataaactattttaataaacagattaa  
aaaaattataatgtaaccttttgctttcaaattgagtagaaataatgcacatccatgtttgtatcg  
tgcaataaaagtgtttcatccgtaggaaaaaatgacttttagtatctgttccgctttttctgatg  
aatgtgctccccgacaaaattgaatgaatcatggacatttgctggctttgatacagcgaaagc  
agccgttcctatgttatatatcggatttaacagcaggacaaaaaacacatgacagccatcgtc  
accacttattcacacgcacataaacctttcctgacttttggaacagatgatagctcatcaaaa  
atcccgccattgccaataaatcgtatatggcattactgcaccataatcttttgagatttgatt  
gggatatggcgcaagcagcaagacaagcagtcggataatcagcgtataaaaataagcctagtaag  
atcttatccgttctccaatacagcttgaaaaacactacattcaacgcaatgggaagagtgatga  
tgaaaaacagaaacacgaatgcaatcgggtccatcccatccgggtattccttccaataacgaaaa  
gaaactaaaaatcatttgtagatcggcaaaactgacaacagcaagggtcgaacgtataaaactta  
ccctttccgcatgatcacgcccagcagcatatagtgaaaagccgtcagcagcacatatccgt  
atacaaaaaaatgcagcagcggcagcagttcttttccgtcctctcttaagtaagcgctggtgaa  
gtttgttgattgcacctgggtgaataagttcaacagacactcccgccagcagcacaatccgcaat  
ataacacccgccaagaacattgtgcgctgccggtttatTTTgggatgatgcacaaaagatata  
agcccgccagaacaacaattgaccattgaatcagcaggggtgctttgtctgcttaataataaaata  
acgttcgaaatgcaatacataatgactgaataactccaacacgaacaacaaaagtgcgcatttt



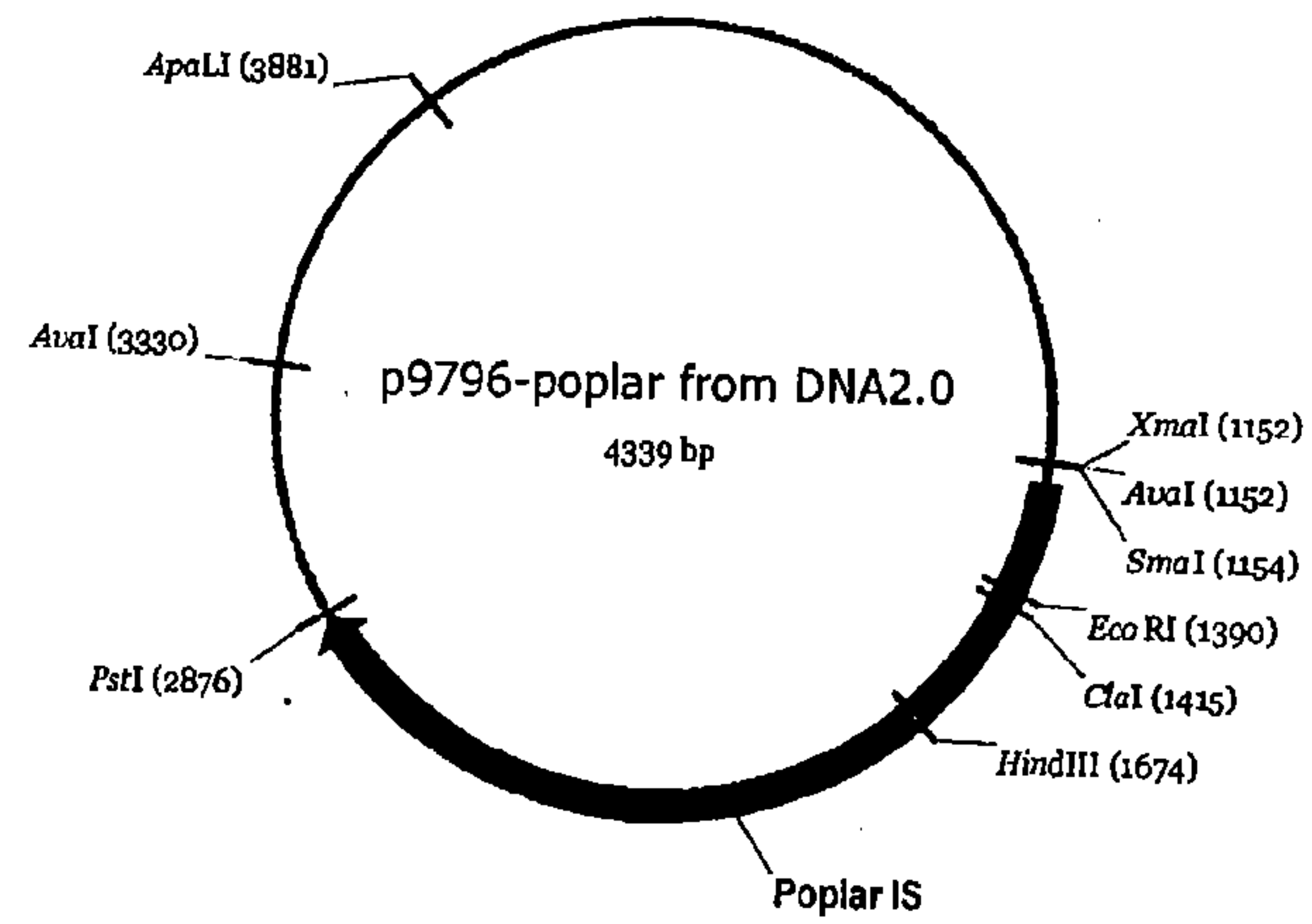
69/167

**Figure 29D**

Ataaaagctaattgattcagttccacataattgatagacgaattctgtctacagggtcacgtggctat  
gtgaaggatcgcgcggtccagtttaagagcaaaaacattgacaaaaaaattttatgtctaaaat  
ttactattaatatatttgtatgtataataagattctcctggccagggggaatcttattttttgtg  
gaggatcatttcatgaggaaaaatgagttccagcttaacgtctctaatctcagcttttgcccgtg  
catatcacagccgatatgacacacctcttatttttgatgattttatcgcaaaagatctcattaa  
cgaaaaagagttttatcgacatcagtaaaaaatatgattcaagaaatatcgtttttcaacaaagag  
atcgccgaacgtcttcaaaatgatcctgaaaaaatattaaaatgggttgacaaaatccagctgt  
ctccaacgcccctagcacgtgcttcttattgtgaaaaagtcttgacacgaattaatcctggg  
ggcaaaacagtatgtcattcttggagcgggactggatactttctgctttcggcatccagaatta  
gaaaacagcttacagggttttcgaggttgatcatccggccacacagcaattgaaaaaaaaataagc  
tgaaggatgcaaatctgacaattccgggtcatcttcattttgttcctatggatttcaccaaac  
gttttcgtatgatcctctcttagatgaaggatttaaaaacacaaaaacattcttcagccttctc  
ggagtgtcttattatgtaacacgggaagaaaaatgcaagcttgatcagcaattttattttctcatg  
tcccgcctggaagctctattgtttttgattatgaggacgaaacactttttacagcaaaaggac  
gtcgaatcgagttgaacatatggtgaagatggctgccgcaagcggggaaccgatgaaatcatgt  
ttcacttatcaagagattgaacatctg  
(SEQ ID NO:47)



70/167

**Figure 30**



71/167

**Figure 31A**

5' -

tagaaaaactcatcgagcatcaaatagaaactgcaatttattcatatcaggattatcaataccat  
atTTTTgaaaaagccgttttctgtaatgaaggagaaaaactcacccgaggcagttccataggatggc  
aagatcctgggtatcggtctgcgattccgactcgtccaacatcaatacaacctattaatttcccc  
tcgtcaaaaataaggttatcaagtgagaaatcaccatgagtgacgactgaatccggtgagaatg  
gcaaaagtttatgcatttctttccagacttggtcaacaggccagccattacgctcgtcatcaaa  
atcactcgcacatcaaccaaaccgttattcattcgtgattgcgccctgagcggaggcgaaatacgcga  
tcgctgtttaaaggacaattacaaacaggaatcgagtgaaccggcgaggaacactgccagcg  
catcaacaatatTTTcacctgaatcaggatattcttctaatacctggaacgctgtttttccggg  
gatcgcagtggtgagtaaccatgcacatcaggagtacggataaaatgcttgatggtcggaagt  
ggcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgctac  
ctttgccatgtttcagaaacaactctggcgcatcgggcttcccatacaagcgatagattgtcgc  
acctgattgcccgcacattatcgcgagcccatttatacccatataaatcagcatccatgttgga  
tttaatcgcggcctcgcagctttcccggtgaatatggctcatattcttctttttcaatattatt  
gaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataa  
acaaataggggtcagtggttacaaccaattaaccaattctgaacattatcgcgagcccatttata  
cctgaatatggctcataacacccttggttgccctggcggcagtagcgcggtgggtcccacctgac  
cccatgccgaactcagaagtgaacgcgcgtagcgcgcgatggtagtggtgggactcccatgcca  
gagtagggaaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgcc  
cgggctaattaggggggtgtcgcccttttagtcgctgaacatgtgctctgtttctaccgagaacgt  
ttccttcactgagacggaaaccgaggcacgtcgtagcgcgaactacgagccgaatagctgggac  
tacgatttctgctgtcttccgatactgacgaatctattgaggtgtacaaagacaaagcaaaga  
aactggaggctgaagtgcgcgcgcaaatcaaacgagaaagctgaattcctgactctgctgga  
gctgatcgataacgtacagcgcctgggtctgggttaccgcttcgaatctgatatccgtcgcgca  
ctggatcgtttcgtaagcagcggcggtttcgatggcgtgaccaaaccgagcctgcacgctaccg  
cgctgtccttccgtctgctgctcagcacggcttcgaagtcttcaggaagcattctccgggtt  
caaagatcaaaacggtaacttcctggaaaacctgaaagaagacactaaggcgatcctgagcctg  
tatgaggcaagctttctggccctggagggtgagaacatcctggatgaggcgcgcgctattcgcca  
tctcccatctgaaagagctgtctgaagagaaaaatcggttaaggaaactggcagagcaggttaatca  
cgcaactggaactgccgctgcacgtcgtaccagcgtctggaggcggtttgggtccatcgaagcg  
taccgcaaaaaggaggatgctaaccaggttctgctggaactggccatcctggactacaacatga  
tccagtcggtttaccagcgtgatctgcgtgaaacctcccggttggtggcgccgtgtgggcctggc  
gaccaaactgcacttcgctaaggaccgcctgattgagtccttttactgggcagtcggcggttgcg  
ttcgaacctcagtattctgactgccgtaacagcgttgcgaaaatgttcagcttcgttactatta  
tcgacgacatctacgacgtttacggtaactctggacgagctggaactgtttaccgacgctgtcga  
acgttgggatgttaacgccatcaacgatctgcctgactacatgaaactgtgcttcctggcactg  
tataacacgatcaacgaaattgcatacgacaacctgaaagacaaagggtgaaaacatcctgccgt  
acctgactaaagcgtgggcggatctgtgtaacgcttttctgcaagaagcgaaatggctgtataa  
caaatccactccgacctttgacgattatttcggcaatgcctggaaatccagctctggcccgtg  
caactgatcttcgcttattttgcggttggtccaaaacatcaaaaaggaggaaattgaaaacctgc  
aaaaataccacgatatcattagccgtccttctcatatctttcgccctgtgcaacgacctggcaag  
cgcgctccgcagagatcgacgctggcgaaaccgctaactctgtttcctgctacatgcgcaccaag  
ggcatttccgaagagctggcaaccgagagcgtaatgaatctgatcgacgaaacctgtaagaaaa  
tgaacaaagaaaaactgggtgggtccctgttcgctaaaccggttcgtagagactgctattaacct



72/167

**Figure 31B**

ggcacgtcagagccactgcacctaccacaatggtgacgcacatactagcccggatgaactgact  
cgtaaacgtgtactgtctgttatcacgaaccgattctgccgttcgaacgttaactgcagcgtc  
aatcgaaagggcgacacaaaatttattctaaatgcataataaatactgataacatcttatagtt  
tgtattatattttgtattatcgttgacatgtataattttgatatacaaaaactgattttcccttt  
attattttcgagattttattttcttaattctctttaacaaactagaaatattgtatatatacaaaa  
atcataaataatagatgaatagtttaattataggtgttcatcaatcgaaaaagcaacgtatctt  
atttaaagtgcgttgctttttctcatttataagggttaaataattctcatatatcaagcaaagt  
gacaggcgcccttaaatattctgacaaatgctctttccctaaactcccccataaaaaaacccg  
ccgaagcgggtttttacgttatatttgcggttaacgattactcgttatcagaaccgcccaggggg  
cccagacttaagactggccgtcggttttacaacacagaaagagttttagaaacgcaaaaaggcc  
atccgtcaggggccttctgcttagtttgatgcctggcagttccctactctcgccctccgcttcc  
tcgctcactgactcgctgcgctcggtcggtcggtcggtcggttatcagctcactcaaagg  
cggtaatacgggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaaggcca  
gcaaaaggccaggaaccgtaaaaaggccgcgttgctggcggtttttccataggctccgccccct  
gacgagcatcacaaaaatcgacgctcaagtgcagaggtggcgaaacccgacaggactataaagat  
accaggcggtttccccctggaagctccctcgctgcgctctcctgttccgaccctgccgcttaccgg  
atacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtat  
ctcagttcgggtgtaggtcggttcgctccaagctgggctgtgtgcacgaaccccccggttcagccc  
accgctgcgccttatccggtaactatcgctcttgagtccaacccggtaagacacgacttatcgcc  
actggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttc  
ttgaagtgggtgggctaactacggctacactagaagaacagtatattgggtatctgcgctctgctga  
agccagttaccttcggaaaaagagttggtagctcttgatccggcaacaaaccaccgctggtag  
cgggtgggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcct  
ttgatcttttctacggggtctgaagctcagtggaacgacgcgcgtaactcacgttaagggtat  
tttgggtcatgagcttgcgccgtcccgtcaagtcagcgtaaatgctctgcttt

(SEQ ID NO:48)



73/167

**Figure 32**

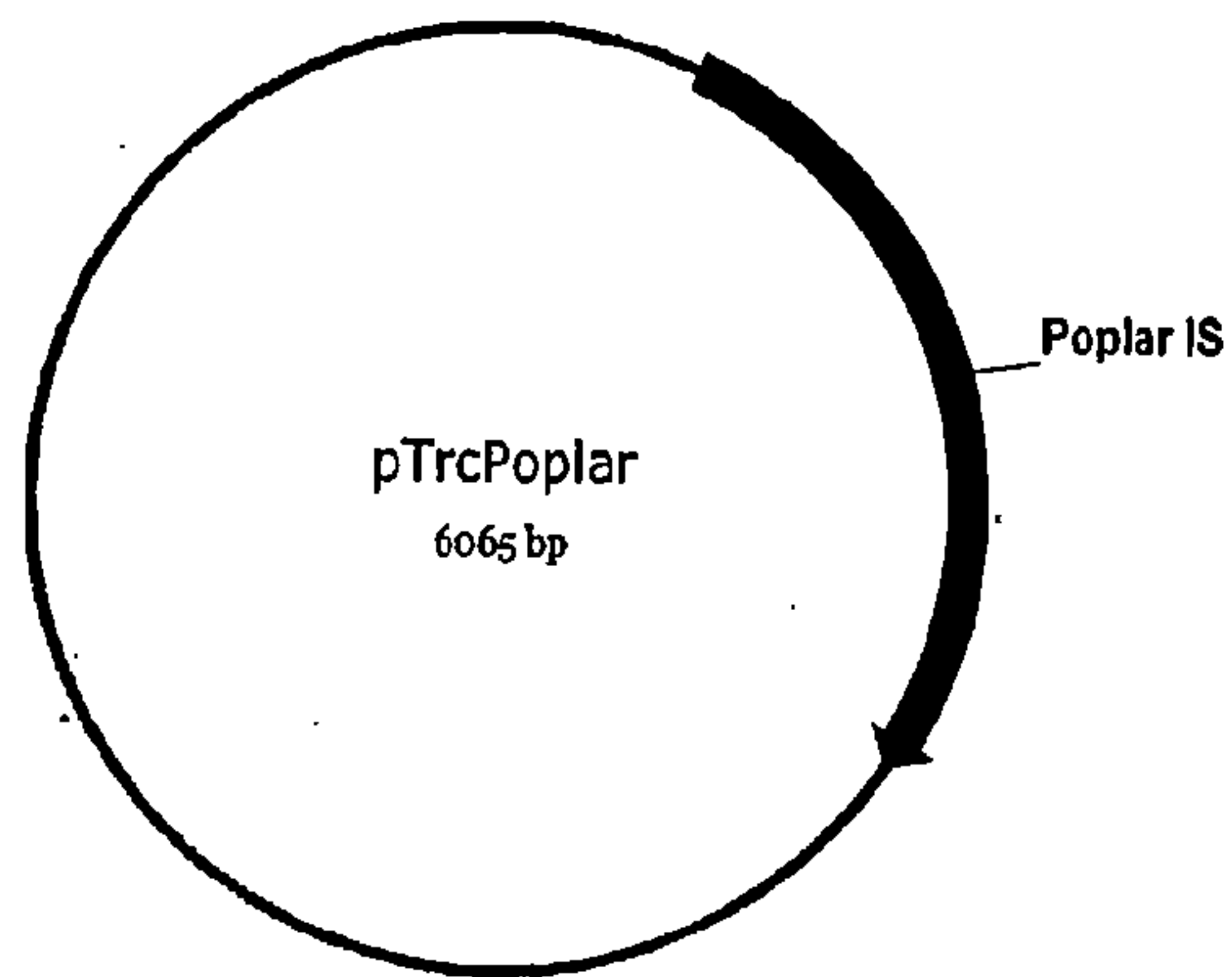




**Figure 33A**

5' -

gtttgacagcttatcatcgactgcacgggtgcaccaatgcttctggcgctcaggcagccatcggaa  
gctgtggtatggctgtgcaggctcgtaaactactgcataattcgtgtcgctcaaggcgactccc  
gttctggataatgttttttgcgcccacatcataacgggttctggcaaataattctgaaatgagctg  
ttgacaattaatcatccggctcgtataatgtgtggaattgtgagcggataacaatttcacacag  
gaaacagcgccgctgagaaaaagcgaagcggcactgctctttaacaatttatcagacaatctgt  
gtggggcactcgaccggaattatcgattaactttattattaaaaattaaagaggtatatattaat  
gtatcgattaaataaggaggaataaaccatgtgctctgttttctaccgagaacgtttccttact  
gagacggaaaccgaggcacgtcgtagcgcgaactacgagccgaatagctgggactacgatttcc  
tgctgtcttccgatactgacgaatctattgaggtgtacaaagacaaagcaaagaaactggaggc  
tgaagtgcgcccgcgaaattaacaacgagaaagctgaattcctgactctgctggagctgatcgat  
aacgtacagcgccctgggtctgggttacggcttcgaatctgatatccgtcgcgcactggatcggt  
tcgtaagcagcggcggtttcgatggcgtgaccaaaccgagcctgcacgctaccgcgctgtcctt  
ccgtctgctgcgtcagcacgggttcgaagtttctcaggaagcattctccggtttcaaagatcaa  
aacggtaacttcctggaaaacctgaaagaagacactaaggcgatcctgagcctgtatgaggcaa  
gctttctggccctggaggggtgagaacatcctggatgaggcgcgctatttcgccatctcccatct  
gaaagagctgtctgaagagaaaatcggttaaggaaactggcagagcagggttaatcacgcactggaa  
ctgccgctgcatcgctcgtaaccagcgtctggaggcggtttgggtccatcgaagcgtaccgcaaaa  
aggaggatgctaaccagggttctgctggaactggccatcctggactacaacatgatccagtccgt  
ttaccagecgtgatctgcgtgaaacctcccggttggtggcgccgtgtgggcctggcgaccaaactg  
cacttcgctaaggaccgcctgattgagtccttttactgggcagtcggcggttgcggttcgaacctc  
agtattctgactgccgtaacagcgttgcgaaaatgttcagcttcggttactattatcgacgacat  
ctacgacgtttacgggtactctggacgagctggaactgtttaccgacgctgtcgaacgttgggat  
gttaacgccatcaacgatctgcctgactacatgaaactgtgcttcctggcactgtataacacga  
tcaacgaaattgcatacgacaacctgaaagacaaagggtgaaaacatcctgccgtacctgactaa  
agcgtgggcggatctgtgtaacgcttttctgcaagaagcgaaatggctgtataacaaatccact  
ccgacctttgacgattatctcggaatgcctggaaatccagctctggcccgtgcaactgatct  
tcgcttattttgcggttggtccaaaacatcaaaaaggaggaaattgaaaacctgcaaaaatacca  
cgatatcattagccgtccttctcatatcttttcgctgtgcaacgacctggcaagcgcgctccgca  
gagatcgacgtggcgaaacctgtaactctgtttcctgctacatgcgcaccaagggcatttccg  
aagagctggcaaccgagagcgtaatgaatctgatcgacgaaacctgtaagaaaatgaacaaaga  
aaaactgggtggctccctgttcgctaaaccttcgtagagactgctattaacctggcacgtcag  
agccactgcacctaccacaatgggtgacgcacatactagcccggatgaactgactcgtaaacgtg  
tactgtctgttatcaccgaaccgattctgccgttcgaacgttaactgcagctggtaccatatgg  
gaattcgaagctttctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgacca  
tcatcatcatcattgagtttaaacgggtctccagcttggtgttttggcggtgagagaaga  
ttttcagcctgatacagattaaatcagaacgcagaagcgggtctgataaaacagaatttgcctgg  
cggcagtagcgcgggtgggtcccacctgaccccatgccgaactcagaagtgaacgccgtagcgc  
gatggtagtgtgggtctccccatgcgagagtaggggaactgccaggcatcaaataaaacgaaag  
gctcagtcgaaagactgggcctttcgttttatctgttggttgctcggtgaacgctctcctgagta  
ggacaaatccgccgggagcggatttgaacgttgcaagcaacggcccgagggtggcgggcagg  
acggccgccataaactgccaggcatcaaattaagcagaaggccatcctgacggatggccttttt  
gcgtttctacaaactctttttgtttatttttctaaatacattcaaatatgtatccgctcatgag  
acaataacctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttc



75/167

**Figure 33B**

cgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcaccagaaacgc  
tgggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatct  
caacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcactttt  
aaagttctgctatgtggcgcggtattatcccgtgttgacgcggggcaagagcaactcggtcgcc  
gcatacactattctcagaatgacttgggttgagtactcaccagtcacagaaaagcatcttacgga  
tggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaac  
ttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgacacaacatgggggatc  
atgtaactcgcccttgatcggttggaaccggagctgaatgaagccataccaaacgacgagcgtga  
caccacgatgcctgtagcaatggcaacaacggttgcgcaaaactattaactggcgaactacttact  
ctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgc  
gctcggcccttccggctggctgggtttattgctgataaatctggagccggtgagcgtgggtctcg  
cggtatcattgcagcactggggccagatggtaagccctcccgatcgtagttatctacacgacg  
gggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgatta  
agcattggtaactgtcagaccaagtgttactcatatatacttttagattgatttaaaacttcattt  
ttaatttaaaaggatctaggtgaagatcctttttgataatctcatgacaaaaatcccttaacgt  
gagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctt  
tttttctgcgcgtaatctgctgcttgcaaaaaaaaccacgcctaccagcgggtgggtttgttt  
gccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacca  
aatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgccta  
catacctcgctctgctaactctgttaccagtggctgctgccagtggcgataagtcgtgtcttac  
cgggttggaactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggggttcg  
tgacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctat  
gagaaagcggccacgcttcccgaaggagaaaaggcggacaggtatccggtaagcggcagggtcgg  
aacaggagagcgcacgaggagcttccaggggggaaacgcctgggtatctttatagtcctgtcggg  
tttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcagggggggcggagcctatgga  
aaaacgccagcaacgcggcctttttacgggttcctggccttttgcctggccttttgcctcacatgtt  
ctttcctgcgttatccctgattctgttgataaccgtattaccgcctttgagtgagctgatacc  
gctcggcgcagccgaacgaccgagcgcagcagtcagtgagcaggaagcgggaagagcgcctga  
tgcggtatttttctccttacgcacatctgtgcggtatttcacaccgcataatggtgcactctcagtac  
aatctgctctgatgccgcatagttaagccagtatacactccgctatcgctacgtgactgggtca  
tggctgcgccccgacacccgccaacacccgctgacgcgccttgacgggcttgtctgctcccggc  
atccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagagggttttcaccgtca  
tcaccgaaacgcgcgaggcagcagatcaattcgcgcgcggaaggcgaagcggcatgcatttacgt  
tgacaccatcgaatgggtgcaaaacctttcgcggtatggcatgatagcgcgccggaagagagtc  
ttcaggggtggatgaatgtgaaaccagtaacggttatacagatgtcgcagagtatgccggtgtctctt  
atcagaccgtttcccgcgtgggtgaaccaggccagccacggtttctgcgaaaacgcgggaaaaagt  
ggaagcggcgatggcggagctgaattacattcccaaccgcgtggcacaacaactggcgggcaaa  
cagtcgttgctgattggcggttgccacctccagtctggccctgcacgcgcgctcgcaaatgtcgc  
cggcgattaaatctcgcgccgatcaactgggtgccagcgtgggtgggtgctgatggtagaacgaag  
cggcgtcgaagcctgtaaagcggcggtgcacaatcttctcgcgcaacgcgtcagtgggctgatc  
attaactatccgctggatgaccaggatgccattgctgtggaagctgcctgcactaatgttccgg  
cgttattttcttgatgtctctgaccagacacccatcaacagtattattttctcccatgaagacgg  
tacgcgactgggcgtggagcatctggtcgcattgggtcaccagcaaatcgcgctgttagcgggc  
ccattaagttctgtctcggcgcgctctgcgtctggctggctggcataaatatctcactcgcaatc



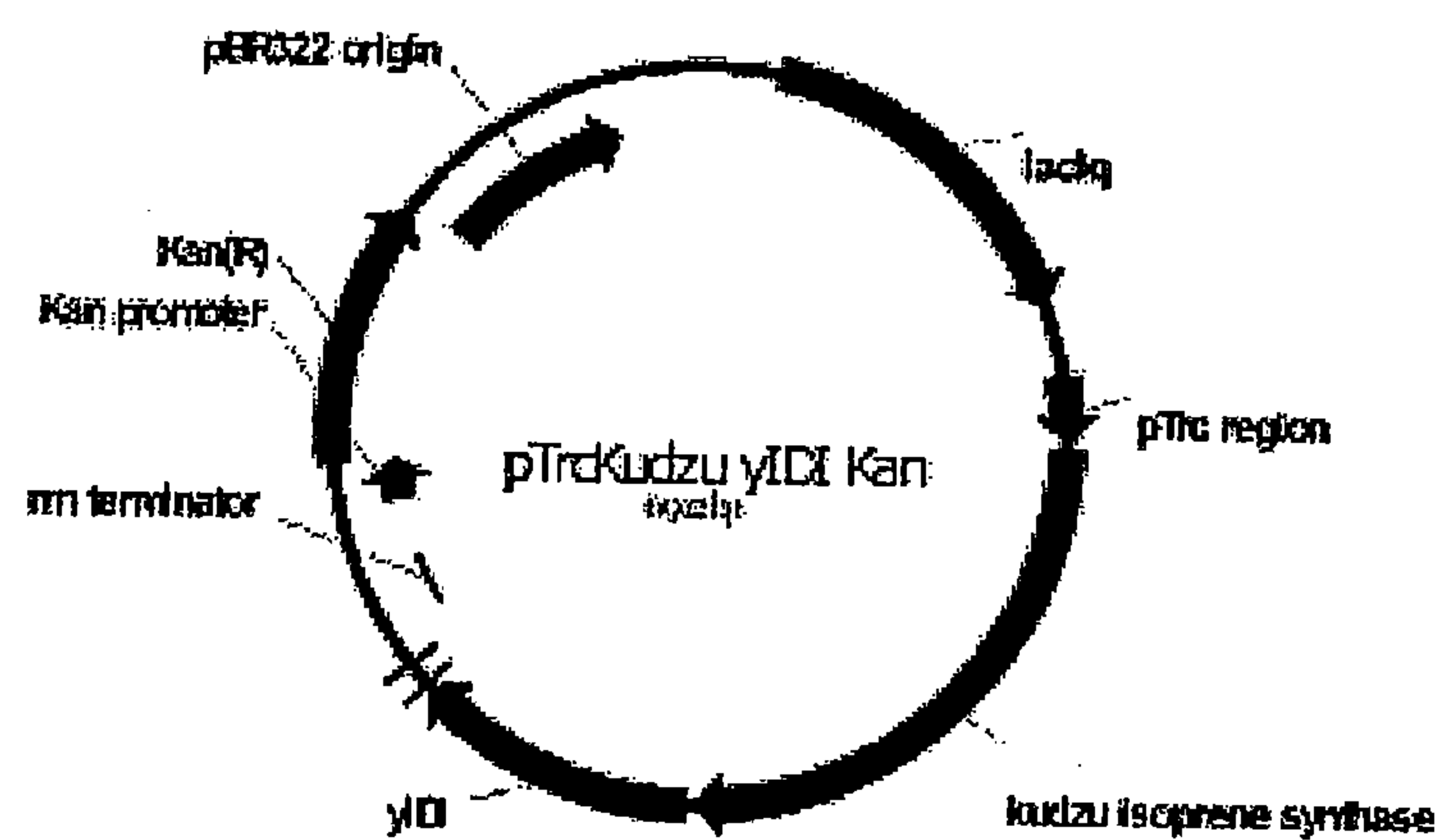
76/167

**Figure 33C**

aaattcagccgatagcggaacgggaaggcgactggagtgccatgtccggttttcaacaaaccat  
gcaaatgctgaatgagggcatcggtcccactgcatgctggttgccaacgatcagatggcgctg  
ggcgcaatgcgcgccattaccgagtcgggctgcgcggttggtgcggatatctcggtagtgggat  
acgacgataccgaagacagctcatgttatatcccgccgtcaaccaccatcaaacaggattttcg  
cctgctggggcaaaccagcgtggaccgcttgctgcaactctctcagggccaggcgggtgaagggc  
aatcagctggtgcccgtctcactggtgaaaagaaaaaccaccctggcgccaatacgcaaaccg  
cctctccccgcgcggttgcccgattcattaatgcagctggcacgacaggtttcccgactggaaag  
cgggcagtgagcgcaacgcaattaatgtgagttagcgcggaattgatctg  
(SEQ ID NO:49)



77/167

**Figure 34**



78/167

**Figure 35A**

5' -

ttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcag  
aggttttcaccgtcatcaccgaaacgcgcgaggcagcagatcaattcgcgcgcgaaggcgaagc  
ggcatgcattttacgttgacaccatcgaatgggtgcaaaacctttcgcggtatggcatgatagcgc  
ccggaagagagtcaattcaggggtggtgaatgtgaaaccagtaacggttatacgatgtcgcagagt  
atgccggtgtctcttatcagaccgtttcccgcgtggtgaaccaggccagccacgtttctgcgaa  
aacgcgggaaaaaagtgggaagcggcgatggcggagctgaattacattcccaaccgcgtggcaca  
caactggcggggcaaacagtcgttgctgattggcgttgccacctccagtctggccctgcacgcgc  
cgtcgcaaatgtgcgggcgattaaatctcgcgcgcgatcaactgggtgccagcgtgggtggtgtc  
gatggtagaacgaagcggcgctcgaagcctgtaaagcggcggtgcacaatcttctcgcgcgaacgc  
gtcagtgggctgatcattaactatccgctggatgaccaggatgccattgctgtggaagctgcct  
gcactaatgttccggcggttatttcttgatgtctctgaccagacacccatcaacagtattatttt  
ctcccatgaagacgggtacgcgactgggcgtggagcatctggctcgcatgtgggtcaccagcaa  
atcgcgtgttagcggggccattaaagtctgtctcggcgcgctctgcgtctggctggctggcata  
aatatctcactcgcaatcaaattcagccgatagcggaaacgggaaggcgactggagtgccatgtccg  
gtttcaacaaccatgcaaatgctgaatgagggcatcgttcccactgcgatgctgggttgccaac  
gatcagatggcgctgggcgcaatgcgcgccattaccgagtcggggctgcgcgttggtgcggata  
tctcggtagtgggatacgcgataccgaagacagctcatgttatatcccggcggtcaaccaccat  
caaacaggattttcgcctgctggggcaaacaccagcgtggaccgcttgctgcaactctctcagggc  
caggcgggtgaagggaatcagctgttgcccgctctcactggtgaaaagaaaaaccaccctggcgc  
ccaatacgcaaacccgctctccccgcgcgttgcccgattcattaatgcagctggcacgcaggt  
ttcccgactggaaagcgggcagtgagcgcgaacgcaattaatgtgagttagcgcgaattgatctg  
gtttgacagcttatcatcgactgcacggtgcaccaatgcttctggcgctcaggcagccatcgga  
gctgtgggtatggctgtgcaggtcgtaaatacactgcataattcgtgtcgtcaaggcgcactccc  
gttctggataatgttttttgccgcgacatcataacggttctggcaaatattctgaaatgagctg  
ttgacaattaatcatccggctcgtataatgtgtggaattgtgagcggataacaattttcacacag  
gaaacagcgcgcgtgagaaaaagcgaagcggcactgctctttaacaatttatcagacaatctgt  
gtgggcactcgaccggaattatcgattaactttattattaaaaattaaagaggatatatattaat  
gtatcgattaaataaggaggaataaaccatgtgtgcgacctcttctcaatttactcagattacc  
gagcataattcccgtcgttccgcaaacctatcagccaaacctgtggaatttcgaattcctgcaat  
ccctgggagaacgcactgaaagtggaaaagctggaggagaaagcgaccaaactggaggaagaagt  
tcgctgcatgatcaaccgtgtagacacccagccgctgtccctgctggagctgatcgacgatgtg  
cagcgcctgggtctgacctacaaatttgaaaaagacatcattaagccctggaaaacatcgtag  
tgctggacgaaaaacaaaaagaacaaatctgacctgcacgcaaccgctctgtctttccgtctgct  
gcgtcagcacgggtttcgagggtttctcaggatgtttttgagcgtttcaaggataaagaagggtgt  
ttcagcgggtgaactgaaagggtgacgtccaaggcctgctgagcctgtatgaagcgtcttacctgg  
gtttcgagggtgagaacctgctggaggaggcgcgtaccttttccatcaccacctgaagaacaa  
cctgaaagaaggcattaataccaagggttgagaacaagtgagccacgccctggaaactgccatat  
caccagcgtctgcaccgtctggaggcacgttggttccctggataaatacgaaccgaaagaaccgc  
atcaccagctgctgctggagctggcgaagctggatttttaacatggtacagaccctgcaccagaa  
agagctgcaagatctgtcccgtggtggaccgagatgggcctggctagcaaacctggatttttgta  
cgcgaccgcctgatggaagtttatttctgggcactgggtatggcgccagaccgcagtttggtg  
aatgtcgcaaacgtgttactaaaatgtttgggtctggtgacgatcatcgatgacgtgtatgacgt



79/167

**Figure 35B**

ttatggcactctggacgaactgcaactgttcacccgatgctgtagagcgctgggacgttaacgct  
attaacaccctgccggactatatgaaactgtgtttcctggcactgtacaacaccgttaacgaca  
cgtcctattctatttctgaaagagaaagggtcataacaacctgtcctatctgacgaaaagctggcg  
tgaactgtgcaaagccttttctgcaagaggcgaaatgggtccaacaacaaaattatcccggctttc  
tccaagtacctggaaaacgccagcggtttcctcctccggtgtagcgctgctggcgccgtcttact  
tttccgtatgccagcagcaggaagacatctccgaccacgcgctgcgttccctgaccgacttcca  
tgggtctgggtgcgttctagctgcgttatcttccgcctgtgcaacgatctggccacctctgcggcg  
gagctggaacgtggcgagactaccaattctatcattagctacatgcacgaaaacgatggtagca  
gcgaggaacaggcccgcggaagaactgcgtaaactgatcgacgccgaatggaaaaagatgaatcg  
tgaacgcgttagcgactccaccctgctgcctaaagcggttcattggaaatcgcagttaacatggca  
cgtgtttcccactgcacctaccagtatggcgatgggtctgggtcgcccagactacgcgactgaaa  
accgcatcaaactgctgctgattgaccctttcccgattaaccagctgatgtatgtctaactgca  
tcgcccttaggaggtataaaaaaaatgactgccgacaacaatagtagtgcctccatgggtgcagtatc  
tagttacgccaaattagtgcaaaaccaaacaacctgaagacatttttgggaagagtttccctgaaatt  
attccattacaacaaagacctaatacccgatctagtgcagcgtcaaatgacgaaagcggagaaa  
catgtttttctgggtcatgatgaggagcaaatgaagttaatgaatgaaaattgtattgttttggga  
ttggggacgataatgctattgggtgccggtaccaagaaagtttgtcattttaatggaaaatattgaa  
aagggtttactacatcgtgcattctccgtctttattttcaatgaacaaggtgaattacttttac  
aacaagagccactgaaaaaataactttccctgatctttggactaacacatgctgctctcatcc  
actatgtattgatgacgaattagggtttgaagggtgaagctagacgataagattaaggggcgctatt  
actgcggcggtgagaaaactagatcatgaattagggtattccagaagatgaaactaagacaaggg  
gtaagtttctacttttttaaacagaatccattacatggcaccaagcaatgaaccatgggggtgaaca  
tgaaattgattacatcctattttataagatcaacgctaaagaaaacttgactgtcaacccaaac  
gtcaatgaagtttagagacttcaaattgggtttcaccaaattgatttgaaaactatgttttgctgacc  
caagttacaagtttacgccttgggtttaagattatttgcgagaattacttattcaactgggtggga  
gcaattagatgacctttctgaagtggaaaatgacaggcaaatcatagaatgctataacaacgc  
gtcctgcagctggtaccatatgggaattcgaagctttctagaacaaaaactcatctcagaagag  
gatctgaatagcgccgctcgaccatcatcatcatcattgagtttaaacgggtctccagcttgg  
ctgtttttggcggtatgagagaagattttcagcctgatacagattaaatcagaacgcgagaagcgg  
ctgataaaacagaatttgcctggcggcagtagcgcggtgggtcccacctgaccccatgccgaact  
cagaagtgaacgcgtagcgccgatggtagtgggggtctcccatgagagagtaggggaactg  
ccaggcatcaaataaaaacgaaagggtcagtcgaaagactgggcctttcgtttttatctgttgttt  
gtcgggtgaacgctctcctgagtaggacaaatccgcccgggagcggtttgaacggttgcaagcaa  
cggcccgggaggggtggcgggcaggacgcccgcataaaactgccaggcatcaaattaagcagaagg  
ccatcctgacggatggcctttttgcgtttctacaaactctttttgtttatttttctaaatacat  
tcaaatatgtatccgcttaaccggaattgccagctggggcgccctctggtaagggttggaagcc  
ctgcaaagtaaaactggatggctttctcgccgccaaggatctgatggcgaggggatcaagctct  
gatcaagagacaggatgaggatcgtttgcgatgattgaacaagatggattgcacgcaggttctc  
cggccgcttgggtggagaggctattcggctatgactgggcacaaacagacaatcggctgctctga  
tgccgcgctgttccggctgtcagcgcagggggcgcccggttctttttgtcaagaccgacctgtcc  
gggtgccctgaatgaactgcaagacgaggcagcgcggctatcgtggctggccacgacggggcggtc  
cttgccgagctgtgctcgacgttgtcactgaagcgggaagggaactggctgctattgggcgaagt  
gccggggcaggatctcctgtcatctcaccttgcctcctgccgagaaagtatccatcatggctgat



80/167

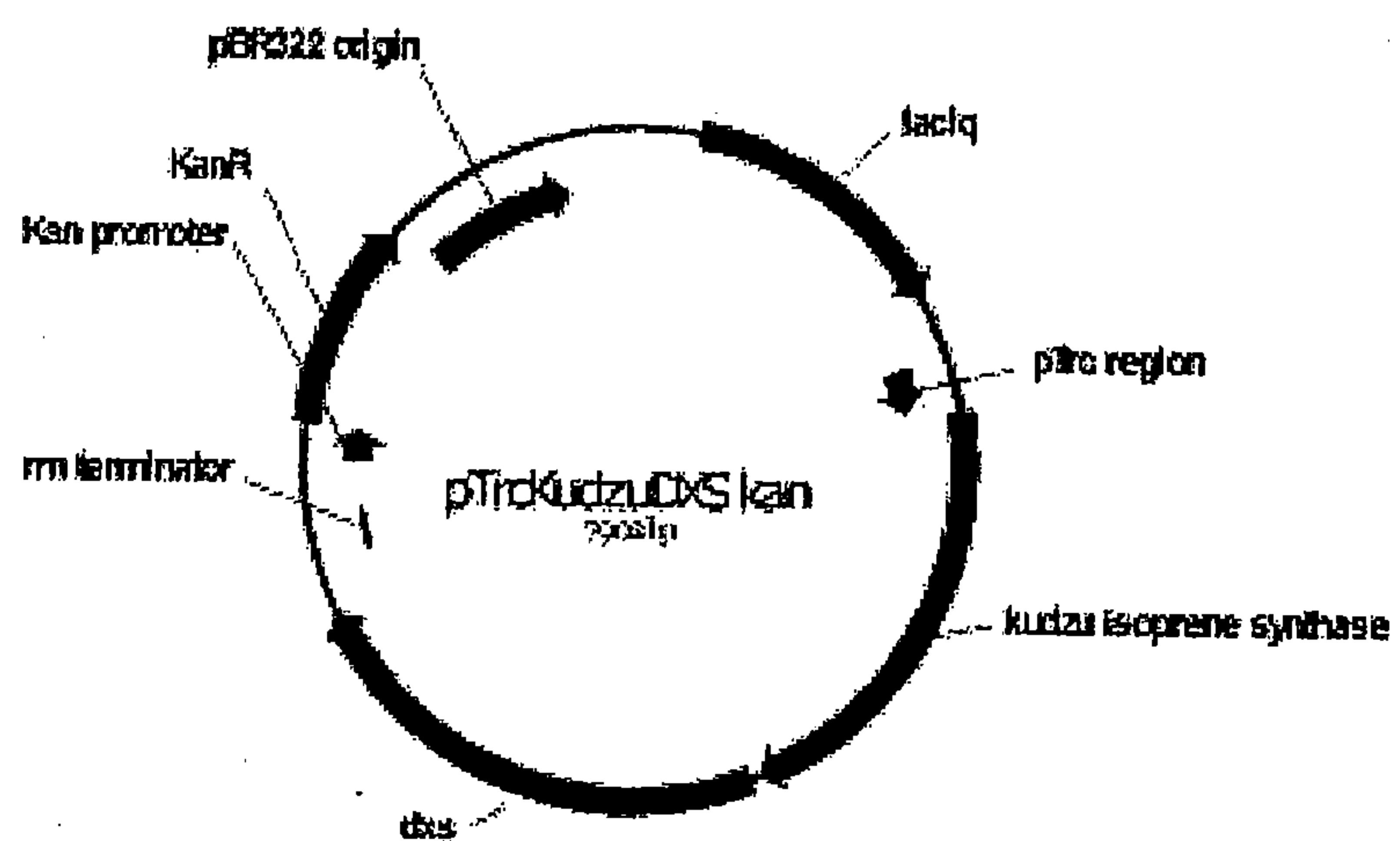
**Figure 35C**

gcaatgcgggcggctgcatacgccttgatccggctacctgcccattcgaccaccaagcgaaacatc  
gcatcgagcgagcacgtactcggatggaagccggctcttgatcgatcaggatgatctggacgaaga  
gcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgagcatgcccgcggcgag  
gatctcgtcgtgacccatggcgatgcctgcttgccgaatatcatggtggaaaatggccgctttt  
ctggattcatcgactgtggccggctgggtgtggcggaccgctatcaggacatagcgttggctac  
ccgtgatattgctgaagagcttggcggcgaatgggctgaccgcttcctcgtgctttacggtatc  
gccgctcccgatcgcagcgcacgccttctatcgcccttcttgacgagttcttctgacatgacc  
aaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggat  
cttcttgagatcctttttttctgcgcgtaatatgctgcttgcaaacaaaaaaccaccgctacc  
agcgggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagc  
agagcgcagataccaaataactgtccttctagtgtagccgtagttaggccaccacttcaagaact  
ctgtagcaccgcctacatacctcgctctgctaatactgttaccagtggctgctgccagtggcga  
taagtctgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggc  
tgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacc  
tacagcgtgagctatgagaaagcgccacgcttcccgaaggagaaaggcggacaggtatccggt  
aagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggatatctt  
tatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcagggg  
ggcggagcctatggaaaaacgccagcaacgcggcctttttacgggttcctggccttttgcctggcc  
ttttgctcacatgttctttcctgcgttatccctgattctgtggataaccgtattaccgccttt  
gagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaag  
cggaagagcgcctgatgcggtattttctccttacgcacatctgtgcggtatttcacaccgcataatg  
gtgcaactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatcgc  
tacgtgactgggtcatggctgcgccccgacacccgcccaacacccgctgacgcgccttgacgggc  
(SEQ ID NO:50)



81/167

Figure 36





82/167

**Figure 37A**

5'-

ttgtctgctcccggtacatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcag  
aggttttcaccgtcatcaccgaaacgcgcgaggcagcagatcaattcgcgcgcgaaggcgaagc  
ggcatgcatttacgttgacaccatcgaatgggtgcaaaacctttcgcggtatggcatgatagcgc  
ccggaagagagtcaattcaggggtggtgaatgtgaaaccagtaaacgtttatacgtatgtcgcagagt  
atgccggtgtctcttatcagaccgtttcccgcggtggtgaaccaggccagccacgtttctgcgaa  
aacgcgggaaaaagtggaaagcggcgatggcggagctgaattacattcccaaccgcgtggcaca  
caactggcgggcaaacagtcgttgctgattggcggttgccacctccagtctggccctgcacgcgc  
cgtcgcaaatgtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtggtggtgtc  
gatggtagaacgaagcggcgtcgaagcctgtaaagcggcggtgcacaatcttctcgcgcaacgc  
gtcagtgggctgatcattaactatccgctggatgaccaggatgccattgctgtggaagctgcct  
gcactaatgttccggcggttatttcttgatgtctctgaccagacacccatcaacagtattattt  
ctcccatgaagacggtaacgcgactgggcggtgagcatctgggtcgcatattgggtcaccagcaa  
atcgcgtgttagcgggcccattaaagtctgtctcggcgcgctctgcgtctggctggctggcata  
aatatctcactcgcaatcaaattcagccgatagcggaaacgggaaggcgactggagtgccatgtccg  
gtttcaacaaccatgcaaatgctgaatgagggcatcgttccactgcgatgctggttgccaac  
gatcagatggcgctgggcgcaatgcgcgccattaccgagtcggggctgcgcgttggtgcggata  
tctcggtagtgggatacgcagataccgaagacagctcatgttatatcccgcgctcaaccacat  
caaacaggattttcgcctgctggggcaaaccagcgtggaccgcttgctgcaactctctcagggc  
caggcggtgaagggaatcagctgttgccgctctcactggtgaaaagaaaaaccacctggcgc  
ccaatacgcaaacgcctctccccgcgcgttgcccgattcattaatgcagctggcagcagagt  
ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagcgcgaattgatctg  
gtttgacagcttatcatcgactgcacggtgcaccaatgcttctggcgctcaggcagccatcggaa  
gctgtggtatggctgtgcaggctcgtaaatcactgcataattcgtgtcgtcgaaggcgactccc  
gttctggataatgttttttgcgccgacatcataacgggttctggcaaatattctgaaatgagctg  
ttgacaattaatcatccggctcgtataatgtgtggaattgtgagcggataacaatttcacacag  
gaaacagcgcgcgtgagaaaaagcgaagcggcactgctctttaacaatttatcagacaatctgt  
gtgggcactcgaccggaattatcgattaactttattattaaaaattaaagaggtatatattaat  
gtatcgattaaataaggaggaataaaccatgtgtgcgacctcttctcaatttactcagattacc  
gagcataattcccgctcgttccgcaaacctatcagccaaacctgtggaatttcgaattcctgcaat  
ccctggagaacgcacctgaaagtggaaaagctggaggagaaagcgaccaaacctggaggaagaagt  
tcgctgcatgatcaaccgtgtagacacccagccgctgtccctgctggagctgatcgacgatgtg  
cagcgctgggtctgacctacaaatttgaaaaagacatcattaaagccctggaaaacatcgtac  
tgctggacgaaaaacaaaaagaacaaatctgacctgcacgcaaccgctctgtctttccgtctgct  
gcgtcagcacggtttcgagggtttctcaggatgtttttgagcggttcaaggataaagaagggtggt  
ttcagcgggtgaactgaaagggtgacgtccaaggcctgctgagcctgtatgaagcgtcttacctgg  
gtttcgagggtgagaacctgctggaggaggcgcgctacctttccatcacccacctgaagaacaa  
cctgaaagaaggcattaataccaagggtgcagaacaagtgagccacgcctggaaactgccatat  
caccagcgtctgcaccgtctggaggcacgttggttcctggataaatacgaaccgaaagaaccgc  
atcaccagctgctgctggagctggcgaagctggatttttaacatggtacagacctgcaccagaa  
agagctgcaagatctgtcccgctgggtggaccgagatgggcctggctagcaaacctggattttgta  
cgcgaccgcctgatggaagttttatttctgggcactgggtatggcgccagaccgcagtttggtg  
aatgtcgcaagctgttactaaaatgtttgggtctggtgacgatcatcgatgacgtgtatgacgt  
ttatggcactctggacgaactgcaactgttcaccgatgctgtagagcgtgggacgttaacgct



83/167

**Figure 37B**

attaacaccctgccggactatatgaaactgtgtttcctggcactgtacaacaccgttaacgaca  
cgtcctatttctatttctgaaagagaaagggtcataacaacctgtcctatctgacgaaaagctggcg  
tgaactgtgcaaagcctttctgcaagaggcgaaatgggtccaacaacaaaattatcccggctttc  
tccaagtacctggaaaacgccagcgtttcctcctcoggtgtagcgctgctggcgccgtcttact  
tttccgtatgccagcagcaggaagacatctccgaccacgcgctgcgttccctgaccgacttcca  
tggtctgggtgcgttctagctgcgttatcttccgcctgtgcaacgatctggccacctctgcggcg  
gagctggaacgtggcgagactaccaattctatcattagctacatgcacgaaaacgatggtagca  
gcgaggaacaggcccgcgagaactgcgtaaactgatcgacgccgaatggaaaaagatgaatcg  
tgaacgcgttagcgactccacctgctgcctaaagcgttcatggaaatcgagttaacatggca  
cgtgtttcccaactgcacctaccagtatggcgatgggtctgggtcgcccagactacgcgactgaaa  
accgcatcaaactgctgctgattgaccttttcccgattaccagctgatgtatgtctaactgca  
ttcgcccttaggaggtataaaaaaacatgagttttgatattgccaaataccgcacctggcactgg  
tcgactccaccaggagttacgactgttgccgaaagagagtttaccgaaactctgcgacgaact  
gcgcgcgtatttactcgacagcgtgagccgttccagcgggcacttcgcctccgggctgggcacg  
gtcgaaactgaccgtggcgctgcactatgtctacaacaccccggttgaccaattgatttgggatg  
tggggcatcaggcttatccgcataaaattttgaccggacgccgcgacaaaatcggcaccatccg  
tcagaaaggcggtctgcacccgttcccgtggcgcgcggaagcgaatatgacgtattaagcgtc  
gggcattcatcaacctccatcagtgcgggaattgggtattgcgggtgctgccgaaaaagaaggca  
aaaatcgccgcaccgtctgtgtcattggcgatggcgcgattaccgcaggcatggcggtttgaagc  
gatgaatcacgcggggcgatatccgtcctgatatgctgggtgattctcaacgacaatgaaatgtcg  
atttccgaaaatgtcggcgcgctcaacaacctctggcacagctgctttccggtaagctttact  
cttcaactgcgcgaaggcgggaaaaaagttttctctggcggtgccgccaattaaagagctgctcaa  
acgcaccgaagaacatattaaaggcatggtagtgctggcacgttggttgaagagctgggcttt  
aactacatcggcccggtggacgggtcacgatgtgctggggcttatcaccacgctaaagaacatgc  
gcgacctgaaaggcccgagttcctgcataatcatgacaaaaaaaggctcgtgggttatgaaccggc  
agaaaaagaccgatcactttccacgcctgcctaaatttgatccctccagcgggttggttgccg  
aaaagtagcggcgggttgccgagctattcaaaaatctttggcgactgggtgtgcaaacggcgag  
cgaaagacaacaagctgatggcgattactccggcgatgcgtgaagggtccggcatggtcgagtt  
ttcacgtaaatcccggatcgctacttcgacgtggcaattgccgagcaaacacgcgggtgaccttt  
gctgcgggtctggcgattgggtgggtacaaaccattgtcgcgatttactccactttcctgcaac  
gcgcctatgatcagggtgctgcatgacgtggcgattcaaaagcttccgggtcctgttcgccatcga  
ccgcgcgggcattgttggtgctgacgggtcaaaccatcagggtgcttttgatctctcttacctg  
cgctgcataaccggaaatgggtcattatgaccccgagcgatgaaaacgaatgtcgccagatgctct  
ataccggctatcactataacgatggcccggtcagcgggtgcgctaccgcggtggcaacgcgggtcgg  
cgtggaaactgacgccgctggaaaaactaccaattggcaaaggcattgtgaagcgtcgtggcgag  
aaactggcgatccttaactttggtacgctgatgccagaagcggcgaaagtgcgccgaatcgctga  
acgccacgctggtcgatatgcgttttgtgaaaccgcttgatgaagcgttaattctggaaatggc  
cgccagccatgaagcgtggtcaccgtagaagaaaacgccattatgggcggcgagggcagcggc  
gtgaacgaagtgtgatggcccatcgtaaacagtagcccggtgctgaacattggcctgccggact  
tctttattccgcaaggaaactcaggaagaaatgcgcgccgaactcggcctcgatgccgctggat  
ggaagccaaaatcaaggcctggctggcataactgcagctggtagcatatgggaattcgaagctt  
tctagaacaaaaactcatctcagaagaggatctgaatagcgcgctcgaccatcatcatcatcat  
cattgagtttaaacgggtctccagcttgggtgttttgccggatgagagaagattttcagcctgat  
acagattaaatcagaacgcgagaagcgggtctgataaaacagaatttgccctggcggcagtagcgcg



84/167

**Figure 37C**

gtggtccacactgaccccatgccgaactcagaagtgaacgcgctagcgccgatggtagtggtg  
ggtctcccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaag  
actgggcctttcgttttatctgttggttggtcggtgaacgctctcctgagtaggacaaatccgcc  
gggagcggatTTgaacgTTgcgaagcaacggcccgagggtggcgggcaggacgcccgccataa  
actgccaggcatcaaattaagcagaaggccatcctgacggatggcctTTTTgCGTTTTctacaaa  
ctctTTTTgtttatttttctaaatacattcaaatatgtatccgcttaaccggaattgccagctg  
ggcgccctctggtaagggtgggaagccctgcaaagtaaactggatggcctttctcgccgccaag  
gatctgatggcgaggggatcaagctctgatcaagagacaggatgaggatcgtttcgcatgatt  
gaacaagatggattgcacgcagggttctccggccgcttgggtggagaggctattcggtctatgact  
gggcacaacagacaatcggtctctgatgccgccgtgttccggctgtcagcgagggggcgccc  
ggttctTTTTgtcaagaccgacctgtccgggtgcctgaatgaactgcaagacgaggcagcgcg  
ctatcgTggctggccacgacgggcttcccttgcgagctgtgctcgacgttgtcactgaagcgg  
gaagggaactggctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcc  
tgccgagaaaagtatccatcatggctgatgcaatgcggcggtgcatacgcttgatccggctacc  
tgcccatctgaccaccaagcgaaacatcgcatcgagcgagcacgtactcggttggaagccggctc  
ttgtcgatcaggatgatctggacgaagagcatcaggggctcgcgccagccgaactgttcgccag  
gctcaaggcgagcatgcccgacggcgaggatctcgTcgTgacctatggcgatgcctgcttgccg  
aatatcatggTggaaaatggccgcttttctggattcatcgactgtggccggctgggtgtggcg  
accgctatcaggacatagcgTtggtaccctgtgatattgctgaagagcttggcggcgaatgggc  
tgaccgcttctcgTgctttacggTatcgccgctcccgattcgagcgcatcgcttctatcg  
cttcttgacgagTtcttctgacgcatgacaaaaatcccttaacgtgagTTTTcgTtccactgag  
cgTcagaccccgtagaaaaagatcaaaggatcttcttgagatcctTTTTtctgCGcgtaatctg  
ctgcttgcaaaacaaaaaaaccaccgctaccagcggtggTttgtttgccggatcaagagctacca  
actctTTTTccgaaggtaactggcttcagcagagcgagataccaaatactgtccttctagtgt  
agccgtagTtaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaat  
cctgttaccagTggctgctgccagTggcgataagtcgTgtcttaccgggttgactcaagacga  
tagTtaccggataaggcgagcggtcgggctgaacggggggTtcgTgcacacagcccagcttgg  
agcgaacgacctacaccgaactgagatacctacagcgTgagctatgagaaagcgccacgcttcc  
cgaaggggagaaaggcgagacaggtatccggtaagcggcagggtcggaacaggagagcgacgagg  
gagcttccagggggaaacgcctggTatctttatagTcctgtcgggTttcgccacctctgacttg  
agcgTcgatTTTTgtgatgctcgTcagggggggcgagcctatggaaaaacgccagcaacgcggc  
ctTTTTacggTtccTggcctTTTTgctggcctTTTTgctcacatgttcttctcTgcttatccct  
gatTctgtggataaccgtattaccgcctttgagTgagctgataccgctcgccgagccgaacga  
ccgagcgagcgagTcagTgagcgaggaagcggaagagcgccctgatgcggTattttctccttac  
gcatctgtgCGgtatttccacaccgcatatggTgcactctcagTacaatctgctctgatgccgca  
tagTtaagccagTatacactccgctatcgctacgtgactgggtcatggctgcgccccgacaccc  
gccaacacccgctgacgcgcctgacgggc

(SEQ ID NO:51)



85/167

Figure 38

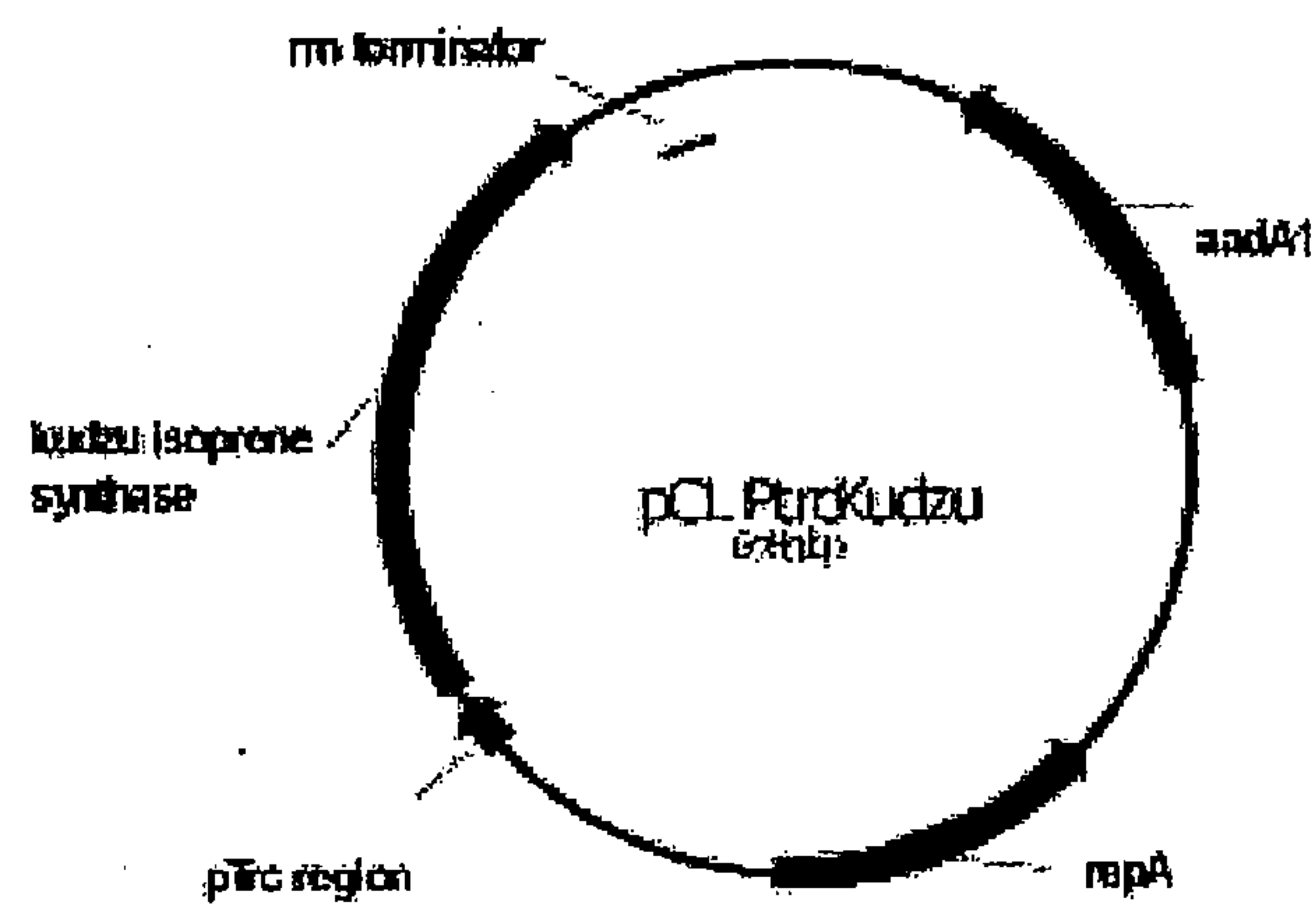




**Figure 39A**

5' -

ctggcgtaatagcgaagaggcccgacccgatcgcccttcccaacagttgcgagcctgaatggc  
gaatggcgccotgatgcggtatTTTTctccttacgcacatctgtgcggtatTTTcacaccgcatatggt  
gcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacc  
cgctgacgagcttagtaaagccctcgctagattTTTaatgcggatggtgcgattacttcgccaac  
tattgcgataacaagaaaaagccagcctttcatgatatactcccaatttgtgtagggcttatt  
atgcacgccttaaaaaataataaaagcagacttgacctgatagtttggctgtgagcaattatgtgc  
ttagtgcatctaacgcttgagttaagccgcgcgcgaagcggcgctcggttgaacgaattgtta  
gacattatTTTgocgactaccttggtgatctcgcccttcacgtagtggacaaattcttccaactg  
atctgcgcgcgaggccaagcgatcttcttcttgtccaagataagcctgtctagcttcaagtatg  
acgggctgatactgggcccggcaggcgctccattgccagtcggcagcgacatccttcggcgcgga  
TTTTgcccgttactgcgctgtaccaaattgcgggacaacgtaagcactacatttcgctcatcgcc  
agcccagtcggggcggcgagttccatagcgttaaggTTTTcatttagcgccctcaaatagatcctgt  
tcaggaaccggatcaaagagttcctccgcgcgtggacctaccaaggcaacgctatggtctcttg  
cttttgtcagcaagatagccagatcaatgtcgatcgtggctggctcgaagatacctgcaagaat  
gtcattgcgctgccattctccaaattgcagttcgcgcttagctggataacgccacggaatgatg  
tcgtcgtgcacaacaatggtgacttctacagcgcggagaatctcgctctctccagggggaagccg  
aagTTTTccaaaaggtcggtgatcaaagctcgccgcggttgtttcatcaagccttacggtcaccgt  
aaccagcaaatcaatatcactgtgtggcttcaggccgcatccactgcggagccgtacaaatgt  
acggccagcaacgtcggttcgagatggcgctcgatgacgccaactacctctgatagttgagtcg  
atacttcggcgatcacccgcttccctcatgatgtttaactTTTgttttagggcgactgcctgtctg  
cgtaacatcggttctgctccataacatcaaacatcgacccacggcgtaacgcgcttctgctgcttg  
gatgcccagggcatagactgtacccccaaaaaacagtcataacaagccatgaaaaccgcccactg  
cgccgttaccaccgctgcgttcggtcaagggttctggaccagttgcgtgagcgcatacgctactt  
gcattacagcttacgaaccgaacaggcttatgtccactgggttcgctgccttcacccgtttccac  
ggtgtgcgtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcga  
acgagcgcaaggTTTTcggtctccacgcacatcgtcaggcattggcgcccttgctgttctctacgg  
caagggtgctgtgcacggatctgccctggcttcaggagatcggaagacctcgcccgctcgccggcg  
ttgccgggtggtgctgaccccgatgaagtgggttcgcacctcggtTTTTctggaaggcgagcatc  
gtttgttcgcccagcttctgtatggaacgggcatgcggatcagtgagggttgcactgcgggt  
caaggatctggatttcgatcacggcacgatcatcgtcggggaggggcaagggtccaaggatcgg  
gccttgatgttacccgagagcttggcaccagcctgcgcgagcaggggaattaattcccacggg  
TTTTgctgcccgcaaaccgggtgttctgggtgttgctagtttgttatcagaatcgagatccggc  
ttcagccggtttgccggctgaaagcgctatTTTctccagaattgccatgattTTTTccccacgg  
gaggcgtcactggctcccgtgttgtcggcagctttgattcgataagcagcatcgctgtttcag  
gctgtctatgtgtgactgttgagctgtaacaagttgtctcagggtgttcaatttcattgttctagt  
tgctttgttttactgggtttcacctgttctattagggtgttacatgctgttcatctgttacattgt  
cgatctgttcatggtgaacagctttgaatgcacaaaaactcgtaaaagctctgatgtatctat  
cttttttacaccgTTTTcatctgtgcataatggacagttttccctttgatatgtaacgggtgaaca  
gttgttctactTTTTgtttgttagtcttgatgcttactgatagatacaagagccataagaacct  
cagatccttccgtatTTtagccagtatgttctctagtgtgggttcggttgtttttgcgtgagccatg  
agaacgaaccattgagatcatacttactttgcatgtcactcaaaaattttgcctcaaaactggt  
gagctgaattTTTTgcagttaaagcatcgtagtgtttttcttagtccgttatgtaggtaggaa  
tctgatgtaatgggtgttggtatTTTTgtcaccattcatttttatctggttgttctcaagttcgg



87/167

**Figure 39B**

ttacgagatccatttgtctatctagttcaacttggaaaatcaacgtatcagtcggggcggcctcg  
cttatcaaccaccaattttcatattgctgtaagtgttttaaatctttacttattggtttcaaaacc  
cattgggttaagcctttttaaactcatggtagttattttcaagcattaacatgaacttaaattcat  
caaggctaattctctatatttgccttgtgagttttcttttgtgttagttcttttaataaccactc  
ataaatcctcatagagtatttgttttcaaaagacttaacatgttccagattatattttatgaat  
ttttttaactggaaaagataaggcaatatctcttcaactaaaaactaatttctaatttttcgcttg  
agaacttggcatagtttgtccactggaaaatctcaaagcctttaaccaaggattcctgatttc  
cacagttctcgtcatcagctctctgggttgctttagctaatacaccataagcattttccctactg  
atgttcatcatctgagcgtattgggttataagtgaacgataaccgtccgttctttccttgtaggggt  
tttcaatcgtgggggttgagtagtgccacacagcataaaaattagcttgggtttcatgctccgttaa  
gtcatagcgaactaatcgctagttcatttgcctttgaaaacaactaattcagacatacatctcaat  
tggtctaggtgatttttaatacactataccaattgagatgggctagtcaatgataattactagtcc  
ttttcctttgagttgtgggtatctgtaaattctgctagacctttgctggaaaacttgtaaattc  
tgctagacctctgtaaattccgctagacctttgtgtgtttttttgttttatattcaagtgggt  
ataatttatagaataaaagaaaagataaaaaaagataaaaagaatagatcccagccctgtgtata  
actcactacttttagtcagttccgcagttattacaaaaggatgtcgcaaacgctgtttgctcctct  
acaaaacagaccttaaaaccctaaaggcttaagtagcacctcgcgaagctcgggcaaatoctg  
aatattccttttgtctccgacctcaggcacctgagtcgctgtctttttcgtgacattcagttc  
gctgcgctcacggctctggcagtgaaatgggggttaaattggcactacaggcgccttttatggattc  
atgcaaggaaaactaccataatacaagaaaagcccgctcacgggcttctcagggcgttttatggc  
gggtctgctatgtgggtgctatctgactttttgctgttcagcagttcctgcccctctgattttcca  
gtctgaccacttcggattatcccgtgacagggtcattcagactggctaattgcaccagtaaggca  
gcggtatcatcaacaggcttaccgctcttactgtcgggaattcgcggttgccgattcattaatg  
cagattctgaaatgagctgttgacaattaatcatccggctcgtataatgtgtggaattgtgagc  
ggataacaatttcacacaggaaacagcgcgctgagaaaaagcgaagcggcactgctctttaac  
aatttatcagacaatctgtgtgggcactcgaccggaattatcgattaaactttattattaaaaat  
taaagagggtatatattaatgtatcgattaaataaggaggaataaaccatgtgtgctgacctcttc  
tcaatttactcagattaccgagcataattcccgtcgttccgcaaaactatcagccaaacctgtgg  
aatttcgaattcctgcaatccctggagaacgacctgaaagtggaaaagctggaggagaaagcga  
ccaaactggaggaagaagttcgtgcatgatcaaccgtgtagacaccagccgctgtccctgct  
ggagctgatcgacgatgtgcagcgcctgggtctgacctacaaatttgaaaaagacatcatta  
gcccctggaaaacatcgtactgctggacgaaaaacaaaagaacaaatctgacctgcacgcaaccg  
ctctgtctttccgtctgctgctgcgtcagcacgggtttcgagggtttctcaggatgtttttgagcgttt  
caaggataaagaaggtgggtttcagcgggtgaactgaaaggtgacgtccaaggcctgctgagcctg  
tatgaagcgtcttacctgggtttcagagggtgagaacctgctggaggaggcgcgtaccttttcca  
tcaccacacctgaagaacaacctgaaagaaggcattaataccaagggttcagacaagtgagcca  
cgccctggaactgccatatcaccagcgtctgcaccgtctggaggcacgttggttcctggataaa  
tacgaaccgaaagaaccgcatcaccagctgctgctggagctggcgaagctggatttttaacatgg  
tacagacctgcaccagaaagagctgcaagatctgtcccgtgggtggaccgagatgggcctggc  
tagcaaaactggattttgtacgcgaccgcctgatggaagtattttctgggcactgggtatggcg  
ccagacctgcagtttgggtgaatgtcgcaaagctgttactaaaatgtttgggtctggtgacgatca  
tcgatgacgtgtatgacgttttatggcactctggacgaactgcaactgttcaccgatgctgtaga  
gcgctgggacgttaacgctattaacaccctgccggactatatgaaactgtgtttcctggcactg  
tacaacacctgtaacgacacgtcctattctatttctgaaagagaaaggtcataacaacctgtcct



88/167

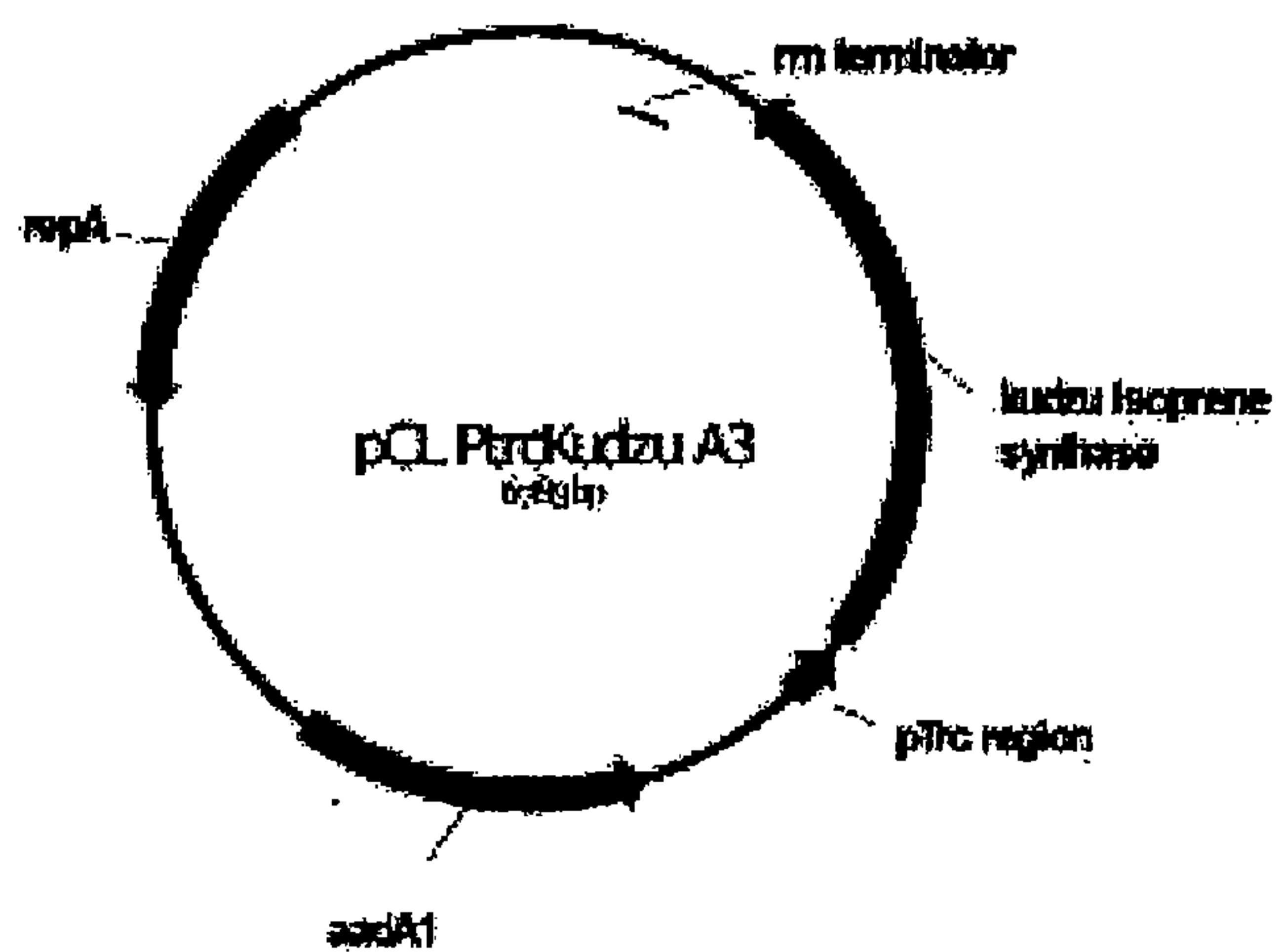
**Figure 39C**

atctgacgaaaagctggcgtgaactgtgcaaagcctttctgcaagaggcgaaatgggtccaacaa  
caaaattatcccggcttttctccaagtacctggaaaacgccagcgtttcctcctccgggtgtagcg  
ctgctggcgccgtcttacttttccgtatgccagcagcaggaagacatctccgaccacgcgctgc  
gttccctgaccgaacttccatgggtctgggtgcggttctagctgcggttatcttccgcctgtgcaacga  
tctggccacctctgcggcggagctggaacgtggcgagactaccaattctatcattagctacatg  
cacgaaaacgatgggtaccagcaggaacaggcccgcgaagaactgcgtaaactgatcgacgccg  
aatggaaaaagatgaatcgtgaacgcggttagcgactccaccctgctgcctaaagcggttcatgga  
aatcgagtttaacatggcacgtgtttcccactgcacctaccagtatggcgatgggtctgggtcgc  
ccagactacgcgactgaaaaccgcatcaaactgctgctgattgaccctttcccgatataaccagc  
tgatgtatgtctaactgcagctggtaccatatgggaattcgaagctttctagaacaaaaactca  
tctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtttaaacggt  
ctccagcttggctgttttggcggatgagagaagattttcagcctgatacagattaaatcagaac  
gcagaagcgggtctgataaaaacagaatttgcctggcggcagtagcgcggtgggtcccacctgaccc  
catgccgaactcagaagtgaacgccgtagcgccgatggtagtgtgggggtctccccatgacgaga  
gtaggggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgtttt  
atctgttggttgctcgggtgaacgctctcctgagtaggacaaatccgccgggagcggatttgaacg  
ttgcgaagcaacggcccggagggtggcgggcaggacgcccgcataaactgccaggcatcaaat  
taagcagaaggccatcctgacggatggcctttttgcgtttctacaaactctttttgtttat  
tctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataat  
(SEQ ID NO:52)



89/167

Figure 40





90/167

**Figure 41A**

5' -

cccgcttactgtcgggaattcgcggttgcccgattcattaatgcagattattgaagcatttatc  
agggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataacaaaaagagtt  
ttagaaaacgcaaaaaggccatccgtcaggatggccttctgcttaatttgatgcctggcagttt  
atggcggggtcctgcccgccaccctccggggcgttgcttcgcaacggtcaaattccgctcccgg  
cggatttgctcactcaggagagcgttcaccgacaaacaacagataaaacgaaaggcccagtc  
ttcgactgagccttttcgtttttatttgatgcctggcagttccctactctcgcatggggagacccc  
acactaccatcggcgctacggcggtttcacttctgagttcggcatggggtcagggtgggaccaccg  
cgctactgccgccaggcaaatctgttttatcagaccgcttctgcgttctgatttaattctgtat  
caggctgaaaatcttctctcatccgcaaaaacagccaagctggagaccgtttaaactcaatgat  
gatgatgatgatggtcgacggcgctattcagatcctcttctgagatgagtttttgttctagaaa  
gcttcgaattcccatatgggtaccagctgcagtttagacatacatcagctgggttaatcgggaaagg  
gtcaatcagcagcagtttgatgcgggttttcagtcgcgtagctctggggcagccagaccatcgcca  
tactggtaggtgcagtgaggaaacacgtgccatgttaactgcgatttccatgaacgctttaggca  
gcagggtggagtcgctaacgcgttcacgattcatctttttccattcggcgctcgatcagtttacg  
cagttcttcgcggggcctgttcctcgctgggtaccatcggttttcgtgcattgtagctaatgatagaa  
ttggtagtctcgccacgttccagctccgccgcagaggtggccagatcggttgacaggcggaaga  
taacgcagctagaacgcaccagaccatggaagtcggtcagggaacgcagcgcgtggtcggagat  
gtcttcctgctgctggcatacggaaaagtaagacggcgccagcagcgtacaccggaggaggaa  
acgctggcggttttccaggtaacttgagaaagccgggataattttgttggttggaccatttcgcct  
cttgcagaaaggctttgcacagttcacgccagcttttctgcagataggacagggttggtatgacc  
tttctctttcagaatagaataggacgtgtcgttaacgggtgttgtagcagtgccaggaaacacagt  
ttcatatagtcgggcagggtgttaatagcgttaacgtcccagcgtctacagcatcggtgaaca  
gttgtagtctcgccagagtgccataaacgtcatacacgtcatcgatgatcgtcaccagacccaaa  
catttttagtaacagcttttgcgacattcaccaaaactgcgggtctggcgccatacccagtgcccag  
aaataaaacttccatcaggcgggtcgcgtaaaaaatccagtttgctagccaggcccatctcggtcc  
accagcgggacagatcttgcagctctttctggtgcagggtctgtaccatgttaaaatccagctt  
cgccagctccagcagcagctgggtgatgcgggttctttcggttcgtattttatccagggaaccaacgt  
gcctccagacgggtgcagacgctgggtgatatggcagttccaggggcgtgggtcacttggtctgcaa  
ccttggtattaatgccttctttcagggtgttcttcagggtgggtgatggaaaaggtagcgccctc  
ctccagcaggttctcaccctcgaaacccaggtaagacgcttcatacaggctcagcaggccttg  
acgtcacctttcagttcacccgtgaaaccaccttctttatccttgaaacgctcaaaaacatcct  
gagaaacctcgaaaccgtgctgacgcagcagacggaaagacagagcgggttgctgcagggtcaga  
tttggttctttttgttttcgtccagcagtagatgttttccagggtttaaattgatgtctttttca  
aatttgtaggtcagaccaggcgctgcacatcgctcgatcagctccagcagggaacagcggctggg  
tgtctacacgggtgatcatgcagcgaacttcttcctccagtttggtcgctttctcctccagctt  
ttccactttcaggctcgttctccagggttgatgcaggaaattcgaaattccacagggtttggctgatag  
tttgcggaacgacgggaattatgctcggtaatctgagtaaattgagaagaggctgcacacatgg  
tttattcctccttatttaattcgatacattaatatatacctctttaatttttaataataaagtta  
atcgataattccggtcgagtgcccacacagattgtctgataaattgttaaaagagcagtgccgct  
tcgctttttctcagcggcgctgtttcctgtgtgaaattgttatccgctcacaattccacacatt  
atacagagccggatgattaattgtcaacagctcatttcagaatctggcgtaatatagcgaaggcc  
cgccaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgccctgatgcggatt  
ttctccttacgcatctgtgcgggtatttcacaccgcatatgggtgcactctcagtacaattctgctc



91/167

**Figure 41B**

tgatgccgcatagttaagccagccccgacacccgcccaacacccgctgacgagcttagtaaagcc  
ctcgctagattttaatgcggtatgttgcgattacttcgccaactattgcgataacaagaaaaagc  
cagcctttcatgatatatctcccaatttgtgtagggcttattatgcacgcttaaaaataataaa  
agcagacttgacctgatagtttggctgtgagcaattatgtgcttagtgcatctaacgcttgagt  
taagccgcgcgcggaagcggcgctcggttgaaacgaattgttagacattatttgccgactacctt  
ggatgatctcgctttcacgtagtggacaaattcttccaactgatctgcgcgcgaggccaagcga  
tcttcttcttgtccaagataagcctgtctagcttcaagtatgacgggctgatactgggcccggca  
ggcgctccattgcccagtcggcagcgacatccttcggcgcgattttgcccgggttactgcgctgta  
ccaaatgcgggacaacgtaagcactacatttcgctcatcgccagcccagtcggggcggcgagttc  
catagcggttaaggttttcatttagcgcctcaaataagatcctgttcaggaaccggatcaaagagtt  
cctccgcgcgctggacctaccaaggcaacgctatgttctcttgcgttttgtcagcaagatagccag  
atcaatgtcgatcgtggctggctcgaagatacctgcaagaatgtcattgcgctgccattctcca  
aattgcagttcgcgcttagctggataacgccacggaatgatgtcgtcgtgcacaacaatgggtga  
cttctacagcgcggagaatctcgctctctccaggggaagccgaagtttccaaaagggtcggtgat  
caaagctcgccgcgttggttcatcaagccttacgggtcacgtaaccagcaaatcaatatcactg  
tgtggcttcaggccgcatccactgcggagccgtacaaatgtacggccagcaacgctcggttcga  
gatggcgctcgatgacgccaactacctctgatagttgagtcgatacttcggcgatcacccgcttc  
cctcatgatgtttaactttgttttagggcgactgcctgctgcgtaacatcggttgctgctccat  
aacatcaaacatcgacccacggcgtaacgcgcttgctgcttggtatgcccgaggcatagactgta  
ccccaaaaaacagtcataacaagccatgaaaaccgccactgcgcggttaccaccgctgcgttc  
ggtaagggttctggaccagttgcgtgagcgcatacgcctacttgcatcacagcttacgaaccgaa  
caggcttatgtccactgggttcgtgccttcacccgtttccacgggtgtgcgtcacccggcaacct  
tgggcagcagcgaagtcgaggcatttctgtcctggctggcgaacgagcgcgaaggtttcggtctc  
cacgcacgtcaggcattggcgcccttgctgttcttctacggcaaggtgctgtgcacggatctg  
ccctggcttcaggagatcgggaagacctcgccgctcgcgccgcttgccgggtggtgctgaccccg  
atgaagtgggttcgcatcctcggttttctggaaggcgagcatcggttggttcgcccagcttctgta  
tggaacgggcatcgggatcagtgagggtttgcaactgcgggtcaaggatctggatttcgatcac  
ggcacgatcatcgtgcgggagggcaagggctccaaggatcgggccttgatgttacccgagagct  
tggcaccacagcctgcgcgagcaggggaattaattccacgggttttgctgcccgcgaacgggct  
gttctgggtgttgctagtttggttatcagaatcgagatccggcttcagccgggtttgccggctgaa  
agcgctatttcttccagaattgccatgatttttcccccagggagggcgtcactggctcccgtgt  
tgtcggcagctttgatctgataagcagcatcgctgtttcaggctgtctatgtgtgactgttga  
gctgtaacaagttgtctcagggtgttcaatttcatgttctagttgctttgttttactgggtttcac  
ctgttctattaggtgttacatgctgttcatctgttacattgtcgcgactgttcatgggtgaacagc  
tttgaatgcacaaaaactcgtaaaagctctgatgtatctatcttttttacaccggttttcatct  
gtgcatatggacagttttccctttgatatgtaacgggtgaacagttgttctacttttgtttgtta  
gtcttgatgcttactgatagatacaagagccataagaacctcagatccttccgtatttagcca  
gtatgttctctagtgtgggttcgttggttttgcgtgagccatgagaacgaaccattgagatcata  
cttactttgcatgtcactcaaaaattttgcctcaaaaactgggtgagctgaatttttgcagttaaa  
gcatcgtgtagtggttttcttagtccgttatgtaggttaggaatctgatgtaatgggtgttggtga  
ttttgtcaccattcatttttatctggttggttctcaagttcggttacgagatccatttgtctatc  
tagttcaacttggaaaatcaacgtatcagtcggggcgccctcgcttatcaaccaccaatttcata  
ttgctgtaagtgtttaaatctttacttattgggtttcaaaaccattgggttaagccttttaaact  
catggtagttattttcaagcattaacatgaacttaaattcatcaaggctaattctctatatttgc



92/167

**Figure 41C**

cttgtgagttttcttttgtgttagttcttttaataaccactcataaatcctcatagagtatttg  
ttttcaaaagacttaacatgttccagattatattttatgaatttttttaactggaaaagataag  
gcaatatctcttcaactaaaaactaattctaatttttcgcttgagaacttggcatagtttgtcca  
ctggaaaatctcaaagcctttaaccaaaaggattcctgatttccacagttctcgtcatcagctct  
ctgggttgcttttagctaatacaccataagcattttccctactgatgttcatcatctgagcgtatt  
ggttataagtgaacgataaccgtccgttctttccttgtaggggttttcaatcgtgggggttgagtag  
tgccacacagcataaaaattagcttggtttcatgctccgttaagtcatagcgactaatcgctagt  
tcatttgctttgaaaacaactaattcagacatacatctcaattgggtctaggtgattttaatcac  
tataccaattgagatgggctagtcaatgataattactagtccttttcctttgagttgtgggtat  
ctgtaaattctgctagaccctttgctggaaaacttgtaaattctgctagaccctctgtaaattcc  
gctagaccctttgtgtgttttttttgtttatatattcaagtgggtataatttatagaataaagaaag  
aataaaaaaagataaaaaagaatagatcccagccctgtgtataactcactacttttagtcagttcc  
gcagtattacaaaaggatgtcgcaaacgctgtttgctcctctacaaaacagacccttaaaaccct  
aaaggcttaagtagcaccctcgcaagctcgggcaaatacgctgaatattccttttgtctccgacc  
atcaggcacctgagtcgctgtctttttcgtgacattcagttcgtcgctcacggctctggcag  
tgaatgggggtaaatggcactacaggcgcccttttatggattcatgcaaggaaactaccataat  
acaagaaaagcccgtcacgggcttctcaggggcgttttatggcggtctgctatgtgggtgctatc  
tgactttttgctgttcagcagttcctgccctctgattttccagtcctgaccacttcggattatcc  
cgtgacaggtcattcagactggctaatagcacccagtaaggcagcggtatcatcaacaggctta  
(SEQ ID NO:53)



93/167

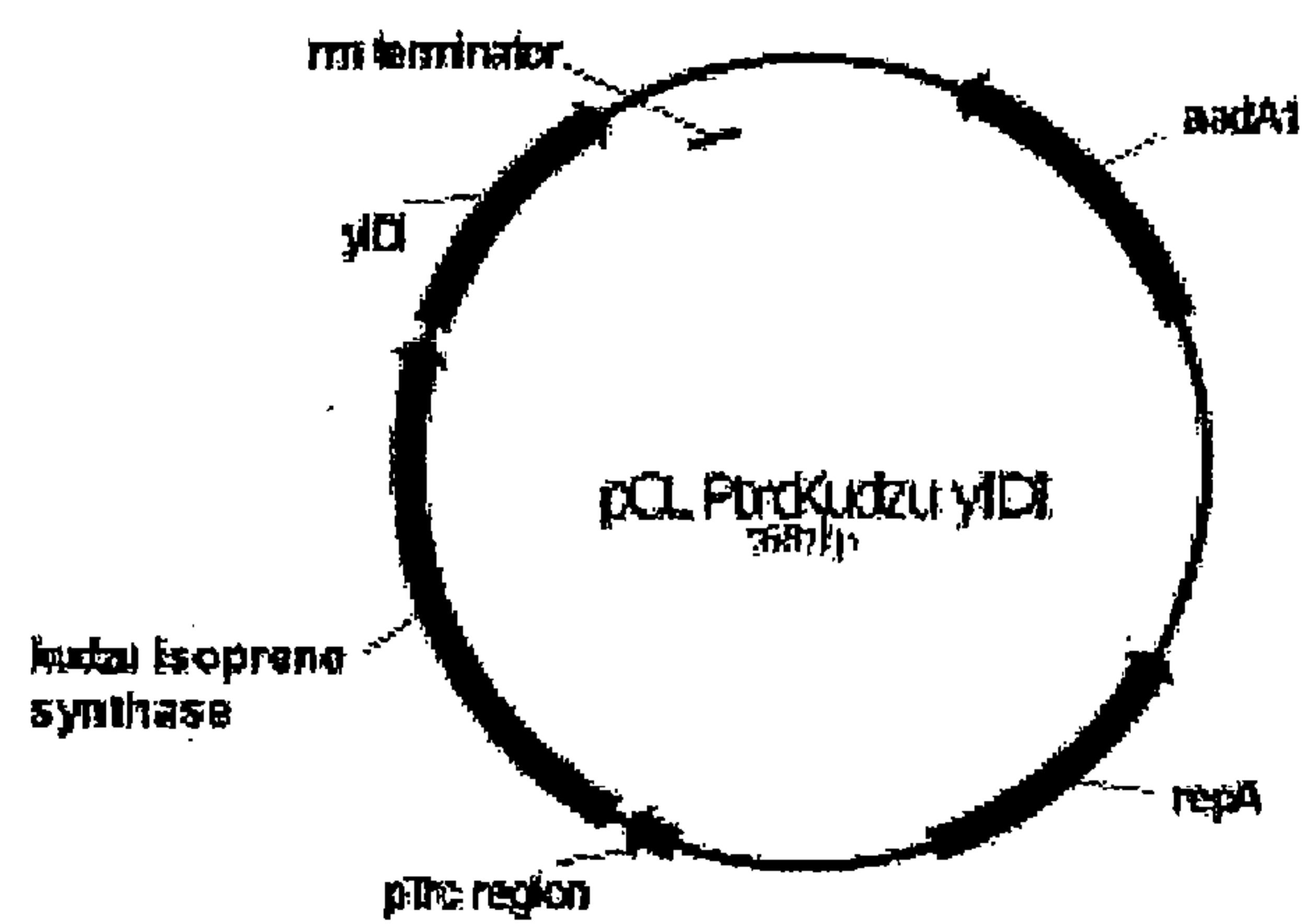
**Figure 42**



**Figure 43A**

5' -

ctggcgtaatagcgaagaggcccgacccgatcgcccttcccaacagttgcgagcctgaatggc  
gaatggcgccctgatgcggtatTTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATATGGT  
gcactctcagtacaatctgctctgatgccgcatagttaagccagcccgacacccgccaacacc  
cgctgacgagcttagtaaagccctcgctagatTTTAAATGCGGATGTTGCGATTACTTCGCCAAC  
tattgcgataacaagaaaaagccagcctttcatgatatatctcccaatttGTGTAGGGCTTATT  
atgcacgcttaaaaaataataaaaagcagacttgacctgatagtttggctgtgagcaattatgtgc  
ttagtgcatctaacgcttgagttaagccgcgcgcgaagcggcgctcggccttgaacgaattgtta  
gacattatTTTgccgactaccttgggtgatctcgccctttcacgtagtggaacaaattcttccaactg  
atctgcgcgcgaggccaagcgatcttcttcttgtccaagataagcctgtctagcttcaagtatg  
acgggctgatactgggcccggcaggcgctccattgcccgagtcggcagcgacatccttcggcgcgga  
TTTTGCCGGTTACTGCGCTGTACCAAATGCGGGACAACGTAAGCACTACATTTGCTCATCGCC  
agcccagtcggggcggcgagttccatagcgttaaggTTTTcatttagcgcctcaaatagatcctgt  
tcaggaaccggatcaaagagttcctccgcgcgtggacctaccaaggcaacgctatgttctcttg  
cttttgtcagcaagatagccagatcaatgtcgatcgtggctggctcgaagatacctgcaagaat  
gtcattgcgctgccattctccaaattgcagttcgcgcttagctggataacgccacggaatgatg  
tcgtcgtgcacaacaatgggtgacttctacagcgcggagaatctcgctctctccagggggaagccg  
aagtttccaaaaggTCGTTGATCAAAGCTCGCCGCGTTGTTTcatcaagccttacggtcaccgt  
aaccagcaaatacatatcactgtgtggcttcaggccgccatccactgcggagccgtacaaatgt  
acggccagcaacgtcggttcgagatggcgctcgatgacgccaaactacctctgatagttgagtcg  
atacttcggcgatcacccgcttccctcatgatgtttaactTTTgTTTTagggcgactgccctgctg  
cgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgcttgctgcttg  
gatgcccgaggcatagactgtacccccaaaaaacagtcataacaagccatgaaaaccgccactg  
cgccgttaccaccgctgcggttcgggtcaagggttctggaccagttgcgtgagcgcatacgcctactt  
gcattacagcttacgaaccgaacaggccttatgtccactgggttcgtgccttcatccgtttccac  
gggtgtgcgtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcga  
acgagcgcgaaggTTTcggctctccacgcacatcgtcaggcattggcgcccttgctgttcttctacgg  
caagggtgctgtgcacggatctgccctggcttcaggagatcggaagacctcggccgctcgcgggcgc  
ttgccggtgggtgctgaccccgatgaagtgggttcgcacctcgggttttctggaaggcgagcatc  
gtttgttcgcccagcttctgtatggaacgggcatgcggatcagtgagggtttgcaactgcgggt  
caaggatctggatttcgatcacggcacgatcatcgtgcgggaggggcaagggtccaaggatcgg  
gccttgatgttacccgagagcttggcaccacagcctgcgcgagcaggggaattaattcccacggg  
TTTTgctgcccgcaaacgggctgttctgggtgttgctagtttgttatcagaatcgcagatccggc  
ttcagccgggtttgccggctgaaagcgcctatTTTcttccagaattgccatgattttttccccacgg  
gaggcgtcactggctcccgtgttgctggcagccttgattcgataagcagcatcgccgtgtttcag  
gctgtctatgtgtgactgttgagctgtaacaagttgtctcagggtgttcaatttcatgttctagt  
tgctttgttttactggtttcacctgttctattagggtgttacatgctgttcatctgtttacattgt  
cgatctgttcatgggtgaacagctttgaatgcacaaaaactcgtaaaagctctgatgtatctat  
cttttttacaccgTTTTcatctgtgcataatggacagttttccctttgatatgtaacgggtgaaca  
gttggttctactTTTgtttgttagtcttgatgcttcaactgatagatacaagagccataagaacct  
cagatccttccgtatTTtagccagtatgttctctagtgtgggttcgttgTTTTtgcgtagccatg  
agaacgaaccattgagatcatacttactttgcatgtcactcaaaaattttgcctcaaaaactggg  
gagctgaatttttgagttaaagcatcgtgtagtgTTTTtcttagtccgttatgtaggtaggaa  
tctgatgtaatgggtgttggtatTTTTgtcaccattcatttttatctgggtgttctcaagttcgg



**Figure 43B**

ttacgagatccatttgtctatctagttcaacttggaaaatcaacgtatcagtcgggcggcctcg  
cttatcaaccaccaattttcatattgctgtaagtgtttaaatctttacttattggtttcaaaacc  
cattgggttaagccttttaaaactcatggtagttattttcaagcattaacatgaacttaaattcat  
caaggctaattctctatatttgccttgtgagttttcttttgtgtagttcttttaataaccactc  
ataaatcctcatagagtatttgttttcaaaagacttaacatgttccagattatattttatgaat  
ttttttaactggaaaagataaggcaatatctcttcaactaaaaactaattctaatttttcgcttg  
agaacttggcatagtttgtccactggaaaatctcaaagcctttaaccaaggattcctgatttc  
cacagttctcgtcatcagctctctgggttgcctttagctaatacaccataagcattttccctactg  
atgttcatcatctgagcgtattgggtataagtgaacgataccgtccgttctttccttgtagggg  
tttcaatcgtggggttgagtagtgccacacagcataaaattagcttgggttcatgctccgttaa  
gtcatagcgactaatcgctagttcatttgccttgaaaacaactaattcagacatacatctcaat  
tggcttaggtgatttttaatacactataccaattgagatgggctagtcaatgataattactagtcc  
ttttcctttgagttgtgggtatctgtaaattctgctagaccttgcctggaaaacttgtaaattc  
tgctagacctctgtaaattccgctagaccttgcctgtgtgtttttttgtttatattcaagtgggt  
ataatttatagaataaagaaagaataaaaaaagataaaaaagaatagatcccagccctgtgtata  
actcactacttttagtcagttccgcagttattacaaaaggatgtcgcaaacgctgtttgctcctct  
acaaaacagaccttaaaaccctaaaggcttaagtagcacctcgcaagctcgggcaaatacgctg  
aatattccttttgcctccgacctcagggcacctgagtcgctgtctttttcgtgacattcagttc  
gctgcgctcacggctctggcagtgaaatgggggtaaatggcactacaggcgcccttttatggattc  
atgcaaggaaactaccataatacaagaaaagcccgtcacgggcttctcaggggcggttttatggc  
gggtctgctatgtggtgctatctgactttttgctgttcagcagttcctgccctctgattttcca  
gtctgaccacttcggattatcccgtgacagggtcattcagactggctaatagcacccagtaaggca  
gcggtatcatcaacaggcttaccgctcttactgtcgggaattcgcgcttgggcgattcattaatg  
cagattctgaaatgagctgttgacaattaatcatccggctcgtataatgtgtggaattgtgagc  
ggataacaatttcacacaggaaacagcgccgctgagaaaaagcgaagcggcactgctctttaac  
aatttatcagacaatctgtgtgggcactcgaccggaattatcgattaactttattattaaaaat  
taaagagggtatatattaatgtatcgattaaataaggaggaataaaccatgtgtgcgacctcttc  
tcaatttactcagattaccgagcataattcccgtcgttccgcaaacctatcagccaaacctgtgg  
aatttcgaattcctgcaatccctggagaacgacctgaaagtggaaaagctggaggagaaagcga  
ccaaactggaggaagaagttcgctgcatgatcaaccgtgtagacacccagccgctgtccctgct  
ggagctgatcgacgatgtgcagcgccctgggtctgacctacaaatttgaaaaagacatcattaaa  
gccctggaaaacatcgtaactgctggacgaaaacaaaaagaacaaatctgacctgcacgcaaccg  
ctctgtctttccgtctgctgctcagcacgggtttcgagggttctcaggatgtttttgagcggtt  
caaggataaagaaggtgggttcagcggtgaactgaaaggtgacgtccaaggcctgctgagcctg  
tatgaagcgtcttacctgggtttcgagggtgagaacctgctggaggaggcgctaccttttcca  
tcacccacctgaagaacaacctgaaagaaggcattaataccaagggttgagaacaagtgaacca  
cgccctggaaactgccatatcaccagcgtctgcaccgtctggaggcacgttgggtcctggataaa  
tacgaaccgaaagaaccgcatcaccagctgctgctggagctggcgaagctggatttttaacatgg  
tacagacctgcaccagaaagagctgcaagatctgtcccgtggtggaccgagatgggcctggc  
tagcaaacctggattttgtacgcgaccgcctgatggaagtttatttctgggcactgggtatggcg  
ccagacctgcagtttgggtgaatgtcgcaagctgttactaaaatgtttgggtctgggtgacgatca  
tcgatgacgtgtatgacgtttatggcactctggacgaactgcaactgttcaccgatgctgtaga  
gcgctgggacgttaacgctattaacacctgcccggactatatgaaactgtgtttcctggcactg  
tacaacacctgtaacgacacgtcctattctattctgaaagagaaaggtcataacaacctgtcct



**Figure 43C**

atctgacgaaaagctggcgtgaactgtgcaaagcctttctgcaagaggcgaaatgggtccaacaa  
caaaattatcccggctttctccaagtacctggaaaacgccagcgtttcctcctccgggtgtagcg  
ctgctggcgccgtcttacttttccgtatgccagcagcaggaagacatctccgaccacgcgctgc  
gttccctgaccgacttccatgggtctgggtgcgttctagctgcgttatcttccgcctgtgcaacga  
tctggccacctctgcggcgaggctggaacgtggcgagactaccaattctatcattagctacatg  
cacgaaaacgatgggtaccagcgaggaacaggccccgcgaagaactgcgtaaactgatcgacgccg  
aatggaaaaagatgaatcgtgaacgcgttagcgactccaccctgctgcctaaagcggttcatgga  
aatcgagttaacatggcacgtgtttcccactgcacctaccagtatggcgatgggtctgggtcgc  
ccagactacgcgactgaaaaccgcatcaaactgctgctgattgaccctttcccgattaaccagc  
tgatgtatgtctaactgcatcgcccttaggaggtataaaaaaaatgactgccgacaacaatagta  
tgccccatgggtgcagtatctagttacgccaattagtgcaaaaccaaacacctgaagacatttt  
ggaagagtttcctgaaattattccattacaacaaagacctaatacccgatctagtgagacgtca  
aatgacgaaagcggagaaacatgtttttctgggtcatgatgaggagcaaattaagttaatgaatg  
aaaattgtattgttttggattgggacgataatgctattgggtgccggtaccaagaaagtttgtca  
tttaattggaaaatattgaaaagggtttactacatcgtgcattctccgtctttattttcaatgaa  
caaggtgaattacttttacaacaaagagccactgaaaaaataactttccctgatctttggacta  
acacatgctgctctcatccactatgtattgatgacgaattaggtttgaagggttaagctagacga  
taagattaagggcgctattactgcggcggtgagaaaactagatcatgaattaggtattccagaa  
gatgaaactaagacaaggggtaagtttcactttttaacagaatccattacatggcaccaagca  
atgaaccatgggggtgaacatgaaattgattacatcctattttataagatcaacgctaaagaaaa  
cttgactgtcaacccaaacgtcaatgaagttagagacttcaaattgggtttcaccaaattgatttg  
aaaactatgtttgctgacccaagttacaagtttacgccttgggtttaagattatttgcgagaatt  
acttattcaactgggtgggagcaattagatgacctttctgaagtggaaaatgacaggcaaatca  
tagaatgctataacgacgcgtcctgcagctggtaccatatgggaattcgaagctttctagaacg  
aaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtt  
taaacgggtctccagcttggctgttttggcggtatgagagaagattttcagcctgatacagattaa  
atcagaacgcagaagcgggtctgataaaacagaatttgcctggcggcagtagcgcggtgggtccca  
cctgaccccatgccgaactcagaagtgaacgcgctagcgccgatggtagtgtgggggtctcccc  
atgcgagagtaggggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcct  
ttcgttttatctgttgtttgtcggtgaacgctctcctgagtaggacaaatccgcccgggagcgga  
tttgaacgttgccaagcaacggccccggagggtggcgggcaggacgcccgcataaaactgccagg  
catcaaattaagcagaaggccatcctgacggatggcctttttgcgtttctacaaactctttttg  
tttattttttctaaatacattcaaataatgtatccgctcatgagacaataaccctgataaatgctt  
caataat

(SEQ ID NO:54)



97/167

Figure 44

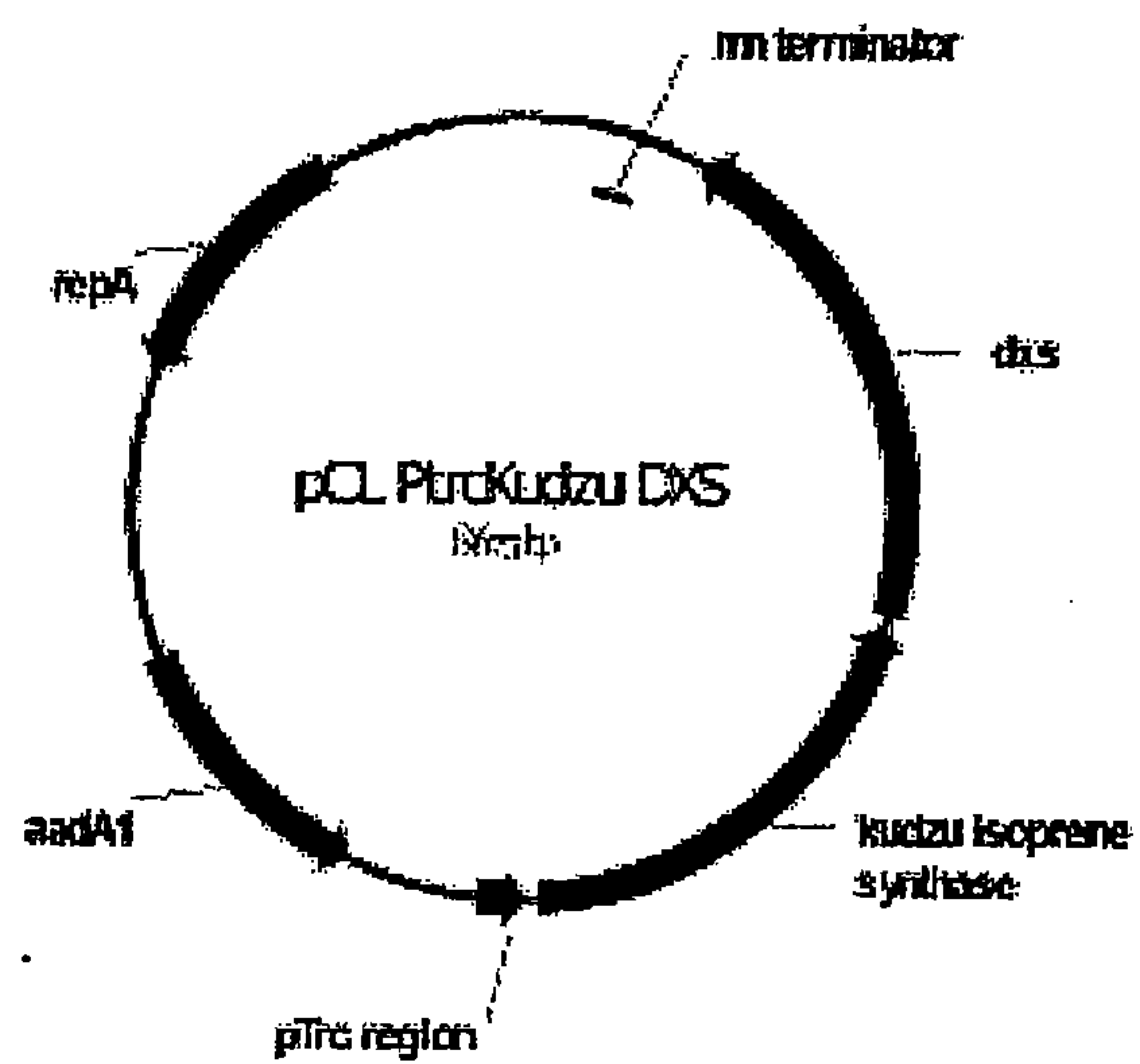




**Figure 45A**

5' -

cccgctcttactgtcgggaattcgcggttgcccgattcattaatgcagattattgaagcatttatc  
agggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaataaaaaagagtt  
tgtagaaacgcaaaaaggccatccgtcaggatggccttctgcttaatttgatgcctggcagttt  
atggcggggtcctgcccggccaccctccggggcgttgcttcgcaacggttcaaataccgctcccgg  
cggatttgtcctactcaggagagcgttcaccgacaaacaacagataaaacgaaaggcccagttct  
ttcgactgagcctttcgtttttatttgatgcctggcagttccctactctcgcatggggagacccc  
acactaccatcggcgctacggcgtttcacttctgagttcggcatgggggtcaggtgggaccaccg  
cgctactgccgccaggcaaatctgttttatcagaccgcttctgcgttctgatttaactctgtat  
caggctgaaaatcttctctcatccgccaaaacagccaagctggagaccgtttaactcaatgat  
gatgatgatgatggtcgacggcgctattcagatcctctctgagatgagttttgttctagaaa  
gcttcgaattcccatatgggtaccagctgcagttatgccagccaggccttgattttggcttccat  
accagcggcatcgaggccgagttcggcgcgcatttcttctgagttccttgccgaataaagaag  
tccggcaggccaatgttcagcacgggtactggtttacgatggggccatcagcacttcgttcacgc  
cgctgcctgcgcccgcataatggcggttttcttctacgggtgaccagcgcttcatggctggcggc  
catttccagaattaacgcttcatcaagcgggtttcacaaaacgcatatcgaccagcggtggcggtc  
agcgattcggcgactttcggcgcttctggcatcagcgtaccaaagttaaggatcgccagtttct  
cgccacgacgcttcacaatgcctttgccaatggtagttttccagcggcggtcagttccacgcc  
gaccggttgccacgcgggttagcgcacogctgacggggccatcgttatagtgatagccggtatag  
agcatctggcgacattcgttttctcgtcgggggtcataatgaccatttccgggtatgcagcgca  
ggtaagagagatcaaaagcaccctgatgggtttgaccgtcagcaccaacaatgcccgcgcgggtc  
gatggcgaacaggaccggaagcttttgaatcgccacgtcatgcagcacctgatcataggcgcgt  
tgcaggaaagtggagtaaatcgcgacaatgggtttgtacccaaccaatcgccagaccgcgagcaa  
aggtcaccgcggtgttgctcggcaattgccacgtcgaagtagcgatccgggaatttacgtgaaaa  
ctcgaccatgcgggaaccttcacgcacatcgccggagtaatcgccatcagcttgttgtcttctcgt  
gccgtttcgcacaaccagtcgccaagatttttgaatagctcggcaaacgcgcgtacttttcg  
gcaaacaccgctggagggtatcaaatttaggcacggcggtggaaagtgatcgggtctttttctgc  
cggttcataaccacgaccttttttgggtcatgatatgcaggaaactgcgggcctttcaggtcgcgc  
atgttcttttagcgtgggtgataagccccagcacatcgtgaccgtccaccggggccgatgtagttaa  
agcccagctcttcaaacaacgtgccaggcactaccatgcctttaatatgttcttcgggtgcgttt  
gagcagctctttaattggcggcacgccagagaaaaactttttcccgccttcgcgcagtgaaag  
taaagcttaccggaagcagctgtgccagatgggtgttgagcgcgcgcacattttcggaatcg  
acatttcattgtcgttgagaatcaccagcatatcaggacggatatcgcccgcgtgattcatcgc  
ttcaaacgccatgcctgcggtaatcgcgccatcgccaatgacacagacgggtgcggcgatttttg  
ccttcttttttcggcagcaaccgcaataccaattccggcactgatggagggtgatgaatgccga  
cgcttaatacgtcatattcgttttcgccgcgccacgggaacgggtgcagaccgcctttctgacg  
gatgggtgccgattttgtcgcggcgctccgggtcaaaattttatgcggataagcctgatgcccaca  
tcccaaatcaattggtcaaacgggggtgttgtagacatagtgacgcgccacgggtcagttcgaccg  
tgcccagcccggaggcgaagtgcccgctggaacgggtcacgctgtcgagtaaatagcggcgag  
ttcgtcgcagagtttcggtaaaactctctttcggcaacagtcgtaactcctgggtggagtcgacc  
agtgccagggtcgggtattttggcaatatcaaaactcatgtttttttacctcctaagggcgaaatg  
cagttagacatacatcagctgggttaatcgggaaagggtcaatcagcagcagtttgatgcgggtt  
tcagtcgcgtagtctgggcgacccagaccatcgccatactggtaggtgcagtgggaaacacgtg  
ccatgttaactgcgatttccatgaacgccttttaggcagcagggtggagtcgctaacgcgttcacg



**Figure 45B**

attcatctttttccattcggcgtcgatcagtttacgcagttcttcgcgggcctgttcctcgctg  
gtaccatcgtttttcgtgcatgtagctaataatgatagaattggtagtctcgccacggttccagctccg  
ccgcagaggtggccagatcgttgcacaggcgggaagataacgcagctagaacgcaccagaccatg  
gaagtcggtcagggaacgcagcgcgtggtcggagatgtcttcctgctgctggcatacggaaaag  
taagacggcgccagcagcgtacaccggaggaggaaacgctggcgttttccagggtacttggaga  
aagccgggataattttgttgttggaccatttcgcctcttgacagaaaggctttgcacagttcacg  
ccagcttttcgtcagataggacaggttggttatgacctttctctttcagaatagaataggacgtg  
tcgttaacgggtgttgtagcagtgccaggaaacacagtttcatatagtcgggcagggtgttaatag  
cgtaaacgtcccagcgcctctacagcatcggtgaacagttgcagttcgtccagagtgccataaac  
gtcatacacgtcatcgatgatcgtcaccagaccaaacatttttagtaaacagctttgcgacattca  
ccaaactgcgggtctggcgccatacccagtgcccagaaataaaacttccatcaggcggtcgcgta  
caaaatccagtttgctagccaggcccatctcggtccaccagcgggacagatcttgacagctcttt  
ctgggtgcagggtctgtaccatgttaaaatccagcttcgccagctccagcagcagctgggtgatgc  
ggttcttttcgggttcgtatttatccaggaaccaacgtgcctccagacgggtgcagacgctgggtgat  
atggcagttccagggcgtgggtcacttggttctgcaaccttggtattaatgccttctttcaggtt  
gttcttcaggtgggtgatggaaaaggtagcgcctcctccagcaggttctcaccctcgaaacc  
aggtaagacgcttcatacaggctcagcaggccttggaagtcacctttcagttcacccgctgaaac  
caccttctttatccttgaaacgctcaaaaacatcctgagaaacctcgaaaccgtgctgacgcag  
cagacggaaagacagagcgggtgcgtgcaggtcagatttggttctttttgttttcgtccagcag  
acgatgttttccagggttttaataatgatgtctttttcaaatttgtaggtcagaccaggcgtgca  
catcgctcgatcagctccagcaggacagcggctgggtgtctacacgggtgatcatgcagcgaac  
ttcttctccagtttggtcgttttctcctccagcttttccactttcaggtcgttctccagggtat  
tgcaggaattcgaaattccacaggtttggctgatagtttgcggaacgacgggaattatgctcgg  
taatctgagtaaattgagaagaggtagcacacatggtttatctcctcttatttaatacgatacat  
taatatatacctctttaatttttaataataaaagttaatcgataattccggtcgagtgcccacac  
agattgtctgataaattgttaaaagagcagtgccgcttcgctttttctcagcggcgctgtttcct  
gtgtgaaattgttatccgctcacaattccacacattatacagagccggatgattaattgtcaaca  
gctcatttcagaatctggcgtaataagcgaagaggcccgccacccgatcgcccttcccaacagttgc  
gcagcctgaatggcgaatggcgctgatgcggtattttctccttacgcacatctgtgcggtatttc  
acaccgcataatgggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccga  
caccgcgaacacccgctgacgagcttagtaaaagccctcgctagattttaatgcggatgttgcg  
attacttcgccaactattgcgataacaagaaaaagccagcctttcatgatatactctcccaattt  
gtgtaggggttattatgcacgcttaaaaaataataaaagcagacttgacctgatagtttggtgt  
gagcaattatgtgcttagtgcatctaacgcttgagtttaagccgcgcgcgaagcggcgctcggct  
tgaacgaattgtagacattatttgccgactaccttggtgatctcgcctttcacgtagtggaca  
aattcttccaactgatctgcgcgcgaggccaagcagatcttcttcttgccaagataagcctgtc  
tagcttcaagtatgacgggtgatactgggcccggcaggcgtccattgccagtcggcagcgac  
atccttcggcgcgattttgccggttactgcgctgtaccaaatgcgggacaacgtaagcactaca  
tttcgctcatcgccagcccagtcgggcccgcaggttccatagcgttaagggtttcatttagcgcct  
caaatagatcctgttcaggaaccggatcaaagagttcctccgcccgtggacctaccaaggcaac  
gctatgttctcttgcttttggtcagcaagatagccagatcaatgtcgatcgtggctggctcgaag  
atacctgcaagaatgtcattgcgctgccatttctccaaattgcagttcgcgcttagctggataac  
gccacggaatgatgtcgctcgtgcacaacaatgggtgacttctacagcgcggagaatctcgctctc  
tccaggggaagccgaagtttccaaaagggtcggttgatcaaagctcgccgcggttggtttcatcaagc



100/167

**Figure 45C**

cttacgggtcaccgtaaccagcaaatcaatatcactgtgtggcttcaggccgcatccactgcgg  
agccgtacaaatgtacggccagcaacgtcggttcgagatggcgctcgatgacgccaactacctc  
tgatagttgagtcgatacttcggcgatcaccgcttccctcatgatgtttaactttgttttaggg  
cgactgccctgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaac  
gcgcttgctgcttggtatgcccgaggcatagactgtacccccaaaaaacagtacataacaagccat  
gaaaaccgcccactgcgcccgttaccaccgctgcgttcgggtcaaggttctggaccagttgcgtgag  
cgcatacgctacttgcattacagcttacgaaccgaacaggcttatgtccactgggttcgtgcct  
tcatecggtttccacgggtgtgcgtcaccggaacacctggggcagcagcgaagtcgaggcatttct  
gtcctgggtggcgaaacgagcgcgaaggtttcggtctccacgcacatcgtcaggcattggcgcccttg  
ctgttcttctacggcaaggtgctgtgcacggatctgccctggcttcaggagatcggaagacctc  
ggccgtcgcggcgcttgccgggtggtgctgaccccggtatgaagtgggttcgcacatcctcggttttct  
ggaaggcgagcatcggttggttcgcccagcttctgtatggaacgggcatgcggatcagtgagggt  
ttgcaactgcgggtcaaggatctggatttcgatcacggcacgatcatcgctgcgggagggcaagg  
gctccaaggatcgggcccttgatgttaaccgagagcttggcaccagcctgcgcgagcaggggaa  
ttaattcccacgggttttgctgcccgcaaacgggctgttctgggtgttgctagtttggttatcaga  
atcgcagatccggcttcagccgggttgccggctgaaagcgctatttcttccagaattgccatga  
ttttttcccccacgggagggcgtcactggctcccgtgttgctcggcagctttgattcgataagcagc  
atcgccctgtttcaggctgtctatgtgtgactgttgagctgtaacaagttgtctcagggtgttcaa  
tttcatgttctagttgctttgttttactgggtttcacctgttctattaggtgttacatgctgttc  
atctgttacattgtcgatctgttcatgggtgaacagctttgaatgcacccaaaaactcgtaaaagc  
tctgatgtatctatcttttttacaccgttttcatctgtgcataatggacagttttccctttgata  
tgtaacgggtgaacagttgttctacttttggttggttagtcttgatgcttcactgatagatacaag  
agccataagaacctcagatccttccgtatttagccagtatgttctctagtgtgggttcgttggtt  
ttgcgtgagccatgagaacgaaccattgagatcatacttactttgcatgtcactcaaaaatttt  
gcctcaaaaactgggtgagctgaatttttgacagttaaagcatcgtagtgtttttcttagtccgt  
tatgtaggttaggaatctgatgtaatgggtgttggtattttgctcaccattcatttttatctggtt  
gttctcaagttcggttacgagatccatttgctctatctagttcaacttggaataatcaacgtatca  
gtcggggcgccctcgcttatcaaccaccaatttcatattgctgtaagtgtttaaatctttactta  
ttgggtttcaaaaaccattgggttaagccttttaaaactcatggtagttattttcaagcattaacat  
gaacttaaaattcatcaaggctaattctctatatttgccctgtgagttttcttttggttagttct  
tttaataaccactcataaatcctcatagagtatttggtttcaaaaagacttaacatgttccagat  
tatattttatgaatttttttaactggaaaagataaggcaatatctcttcaactaaaaactaatc  
taatttttgccttgagaacttggcatagtttggtccactggaaaatctcaaagccttttaaccaa  
ggattcctgattttccacagttctcgtcatcagctctctgggttgcttttagctaatacaccataag  
cattttccctactgatgttcatcatctgagcgtattgggttataagtgaacgataccgtccgttc  
tttcccttgtagggttttcaatcggtgggttgagtagtgccacacagcataaaaattagcttggtt  
tcatgctccgttaagtcatagcgactaatcgctagttcatttgctttgaaaacaactaattcag  
acatacatctcaattgggtctaggtgatttttaatcactataccaattgagatgggctagtcaatg  
ataattactagtccttttccctttgagttgtgggtatctgtaaattctgctagacctttgctgga  
aaacttgtaaattctgctagacctctgtaaattccgctagacctttgtgtgtttttttgttt  
atattcaagtgggttataatttatagaataaagaaagaataaaaaaagataaaaaagaatagatcc  
cagccctgtgtataactcactacttttagtcagttccgcagctattacaaaaggatgtcgcaaacg  
ctgtttgctcctctacaaaacagaccttaaaaccctaaaggcttaagtagcaccctcgcaagct  
cgggcaaatacgctgaatattccttttgctcctcgaccatcaggcacctgagtcgctgtcttttc



101/167

**Figure 45D**

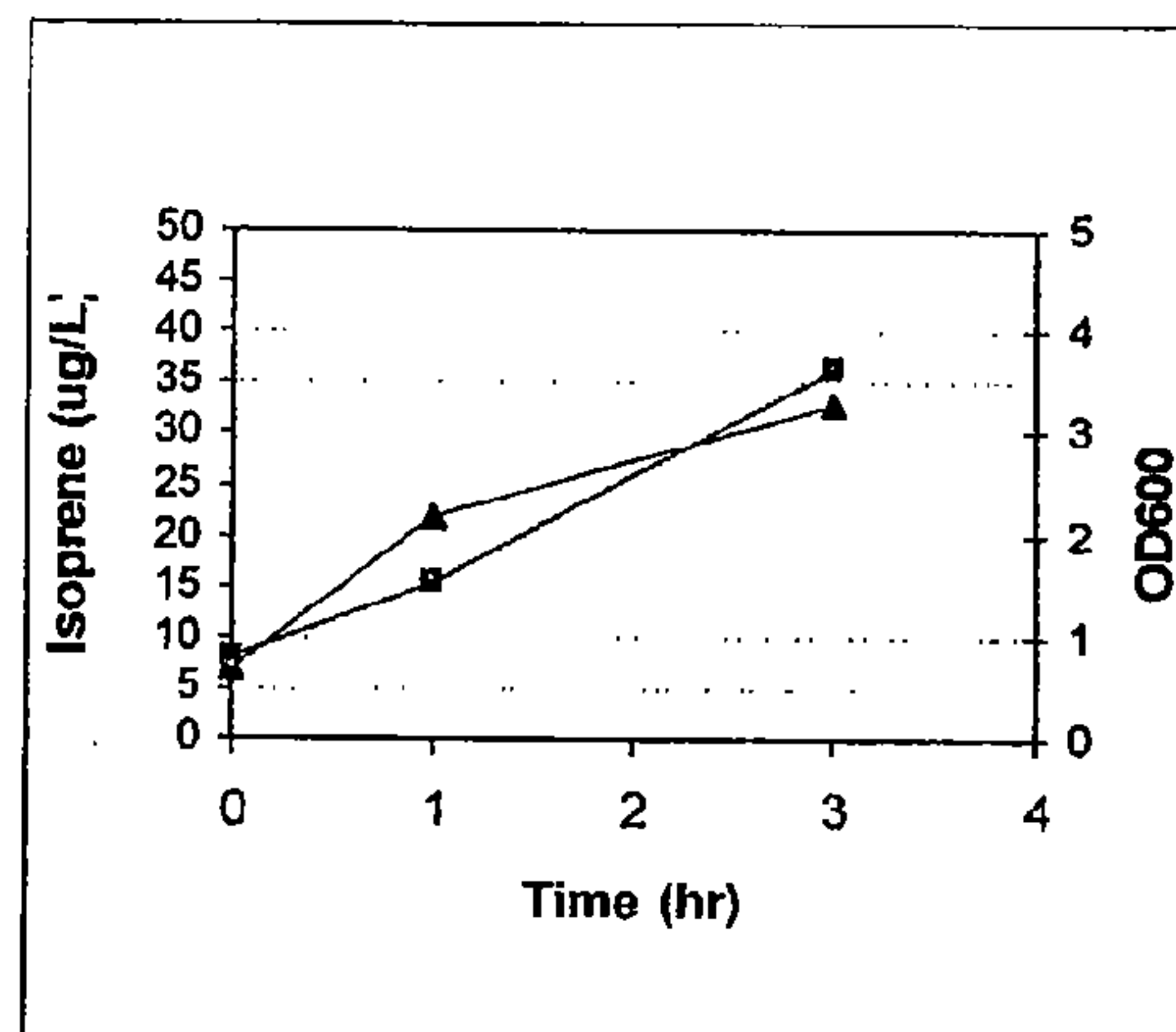
Gtgacattcagttcgctgcgctcacggctctggcagtgaatgggggtaaattggcactacagggc  
ccttttatggattcatgcaaggaaactacccataatacaagaaaagcccgtcacgggcttctca  
gggcgttttatggcgggtctgctatgtgggtgctatctgactttttgctgttcagcagttcctgc  
cctctgattttccagtctgaccacttcggattatcccgtgacaggtcattcagactggctaata  
cacccagtaaggcagcggtatcatcaacaggctta  
(SEQ ID NO:55)



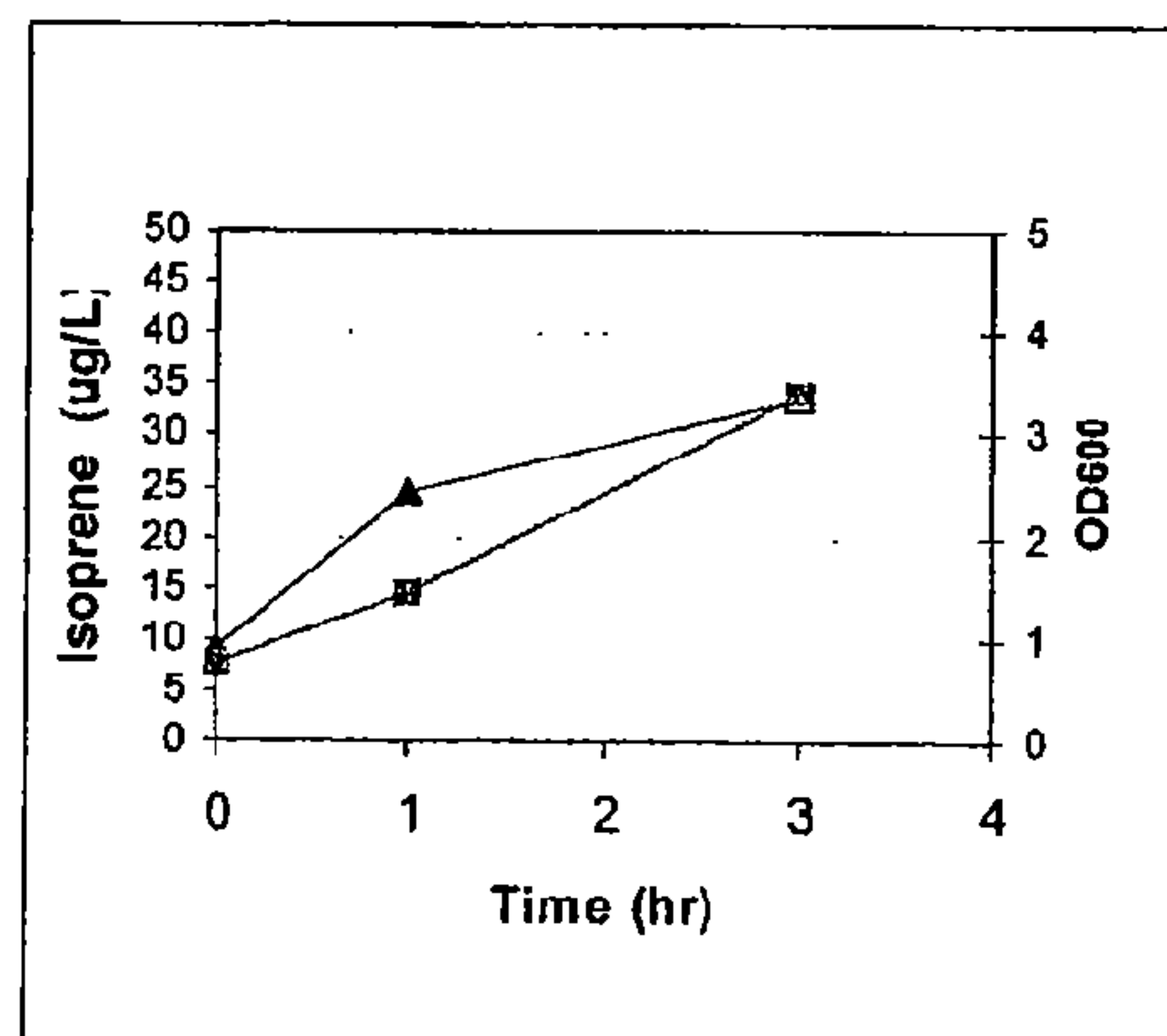
102/167

**Figure 46A**

A.

**Figure 46B**

B.

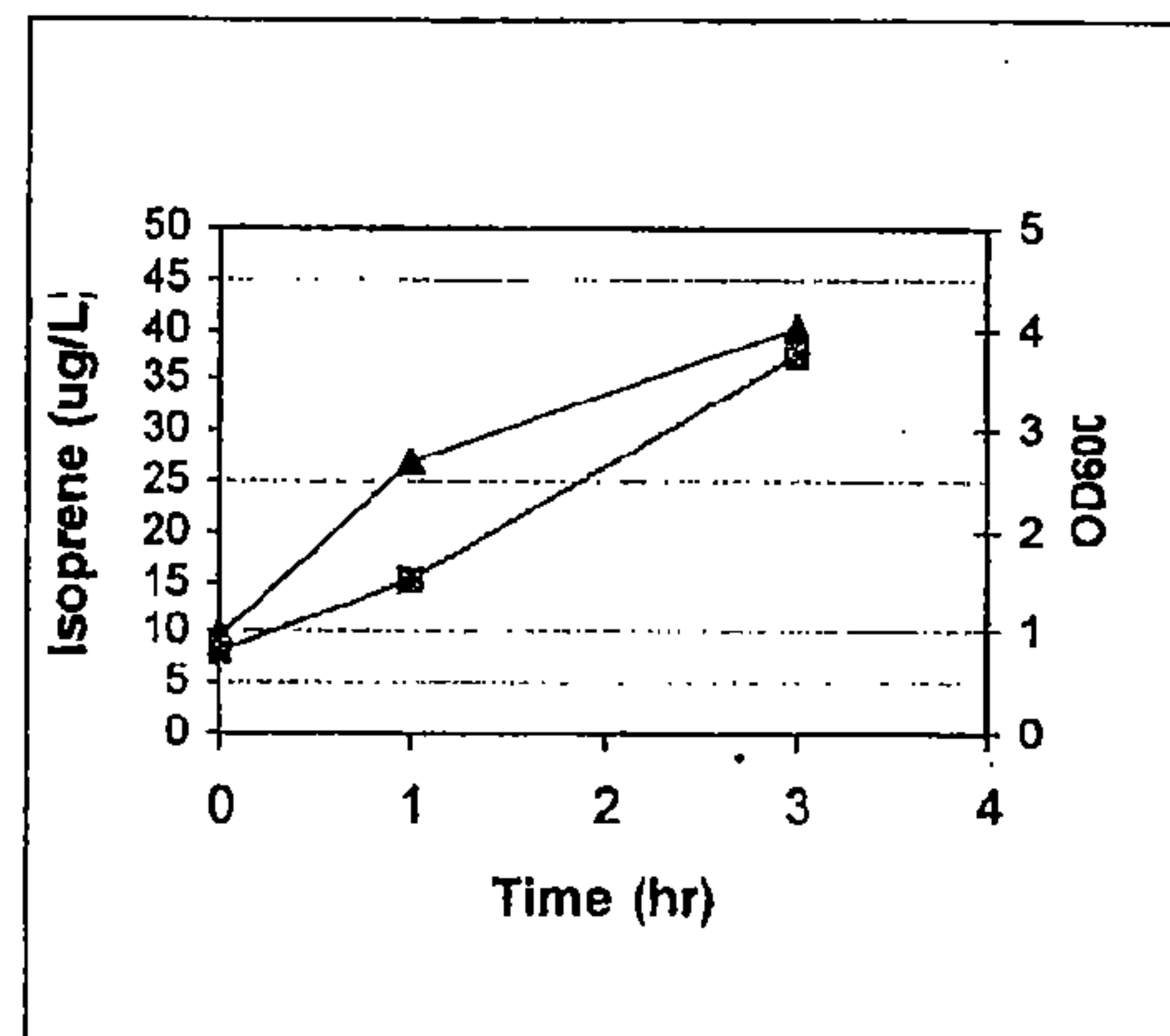




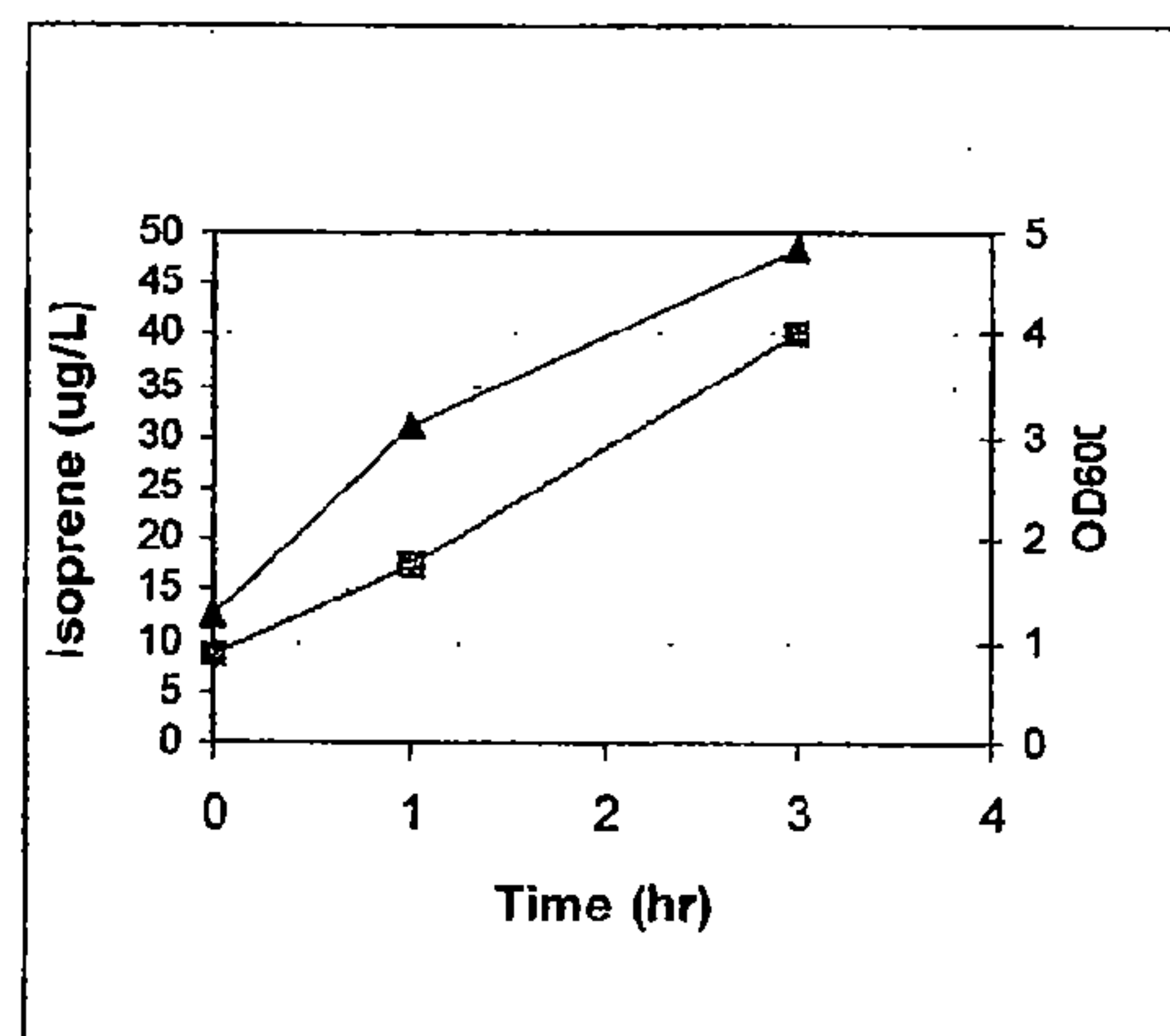
103/167

**Figure 46C**

C.

**Figure 46D**

D.

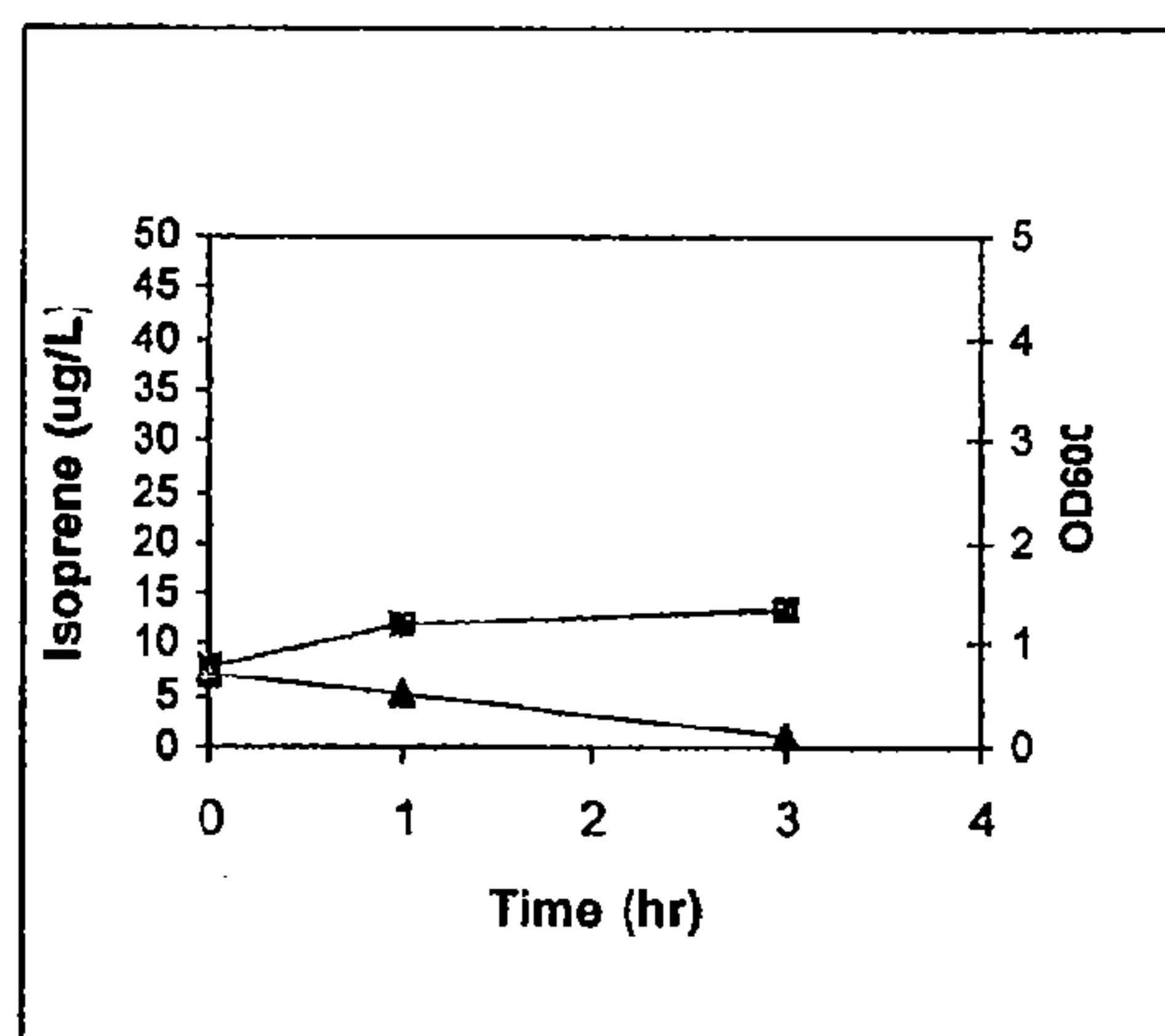




104/167

Figure 46E

E.





105/167

Figure 47A

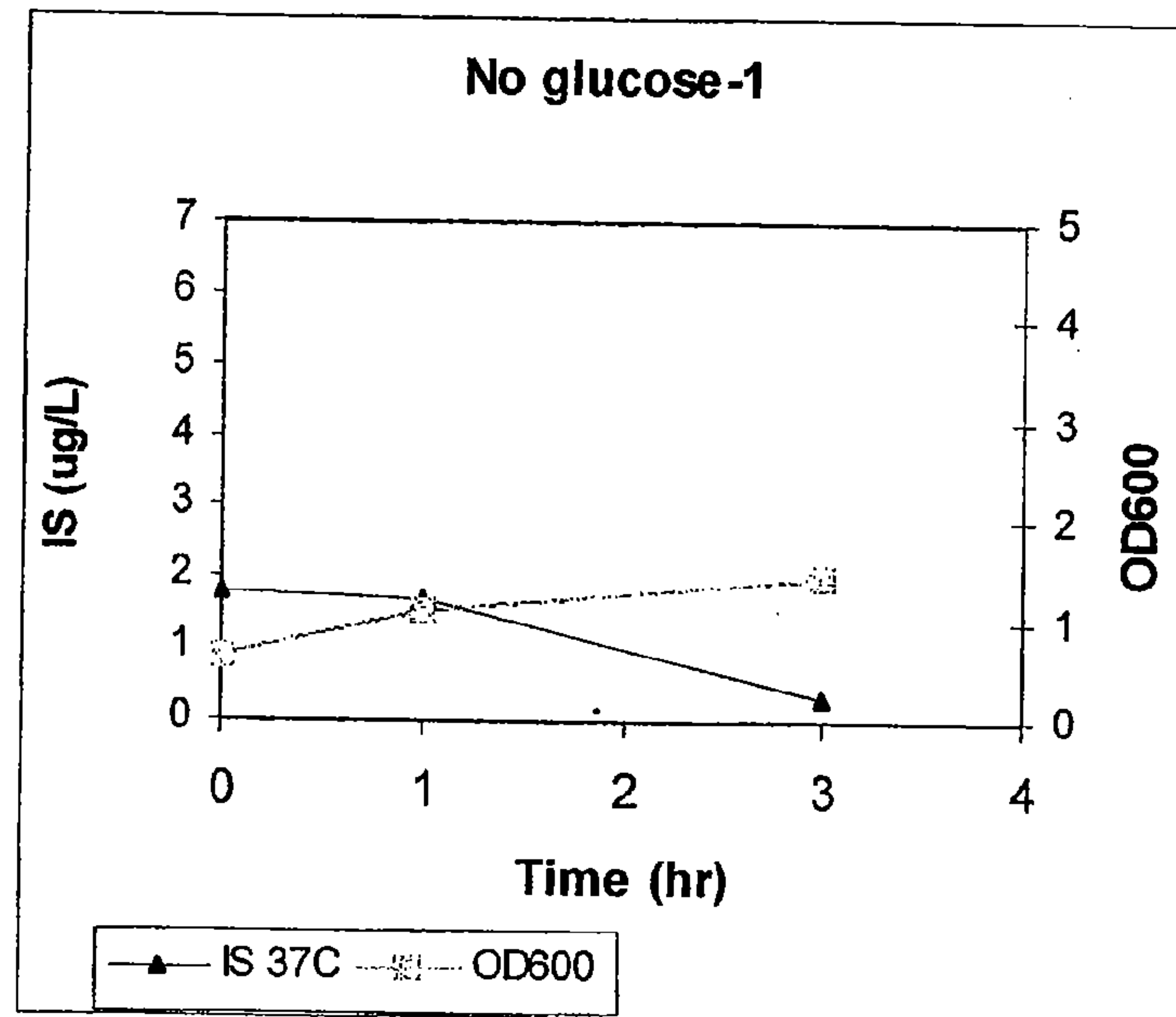
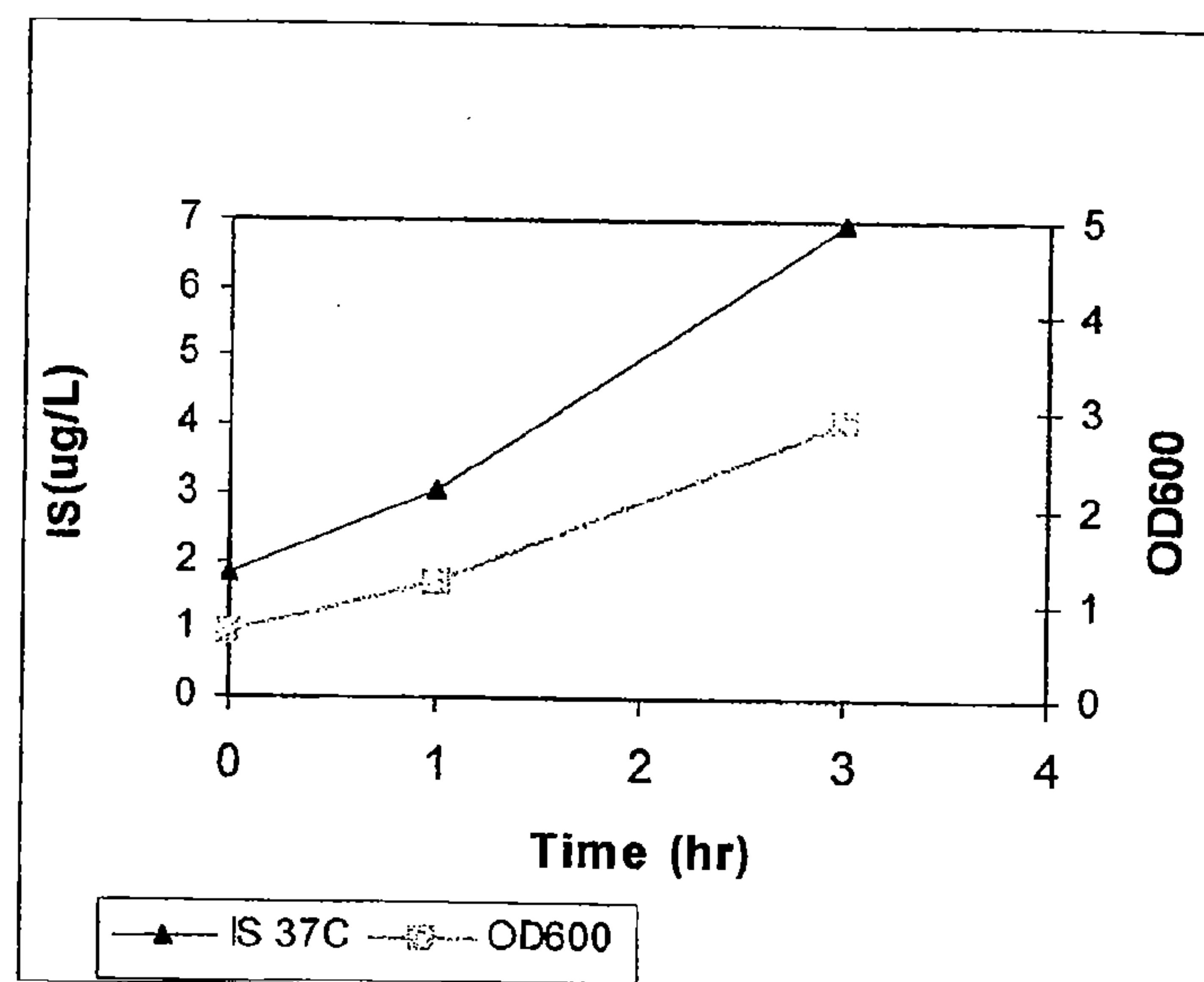
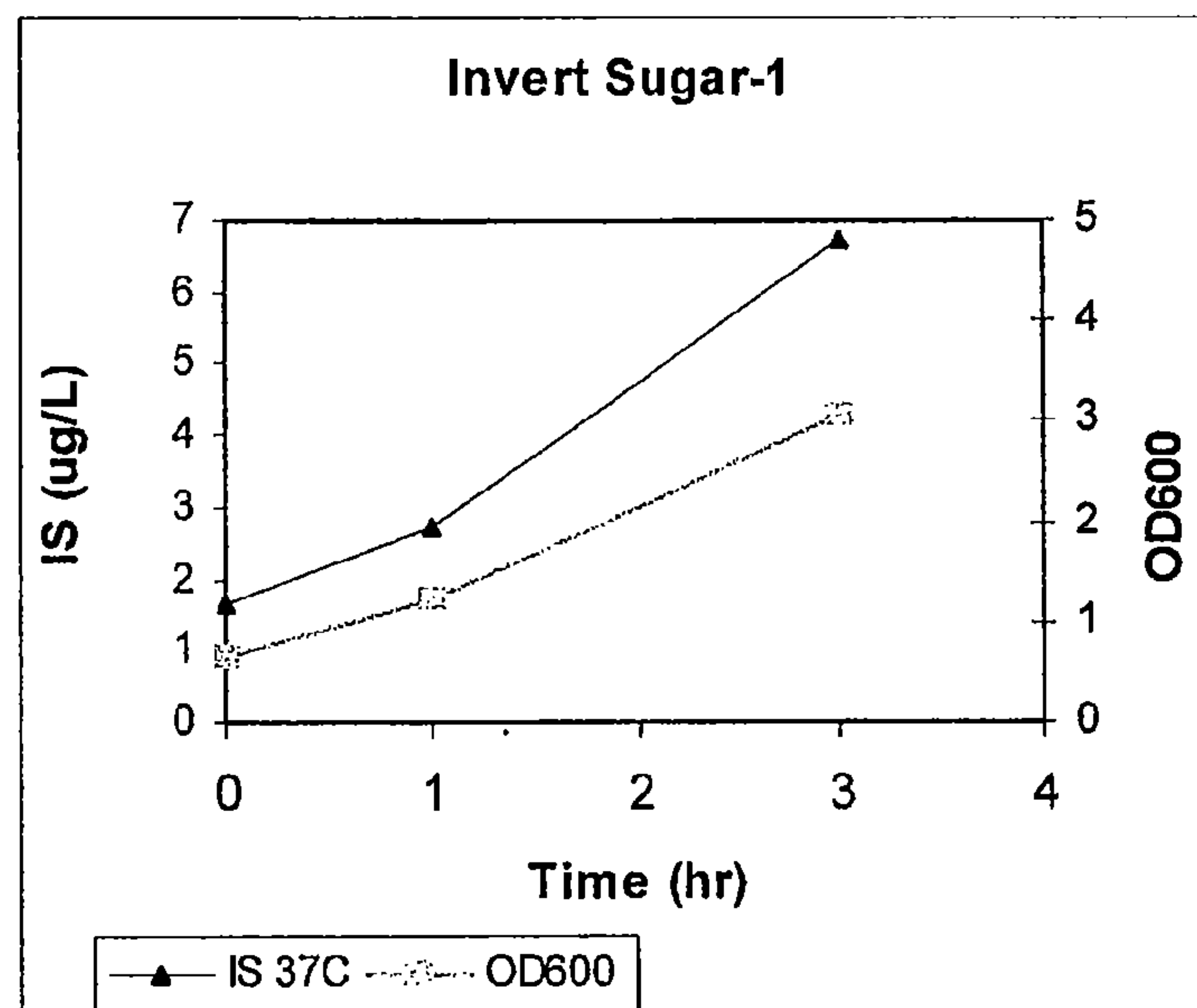
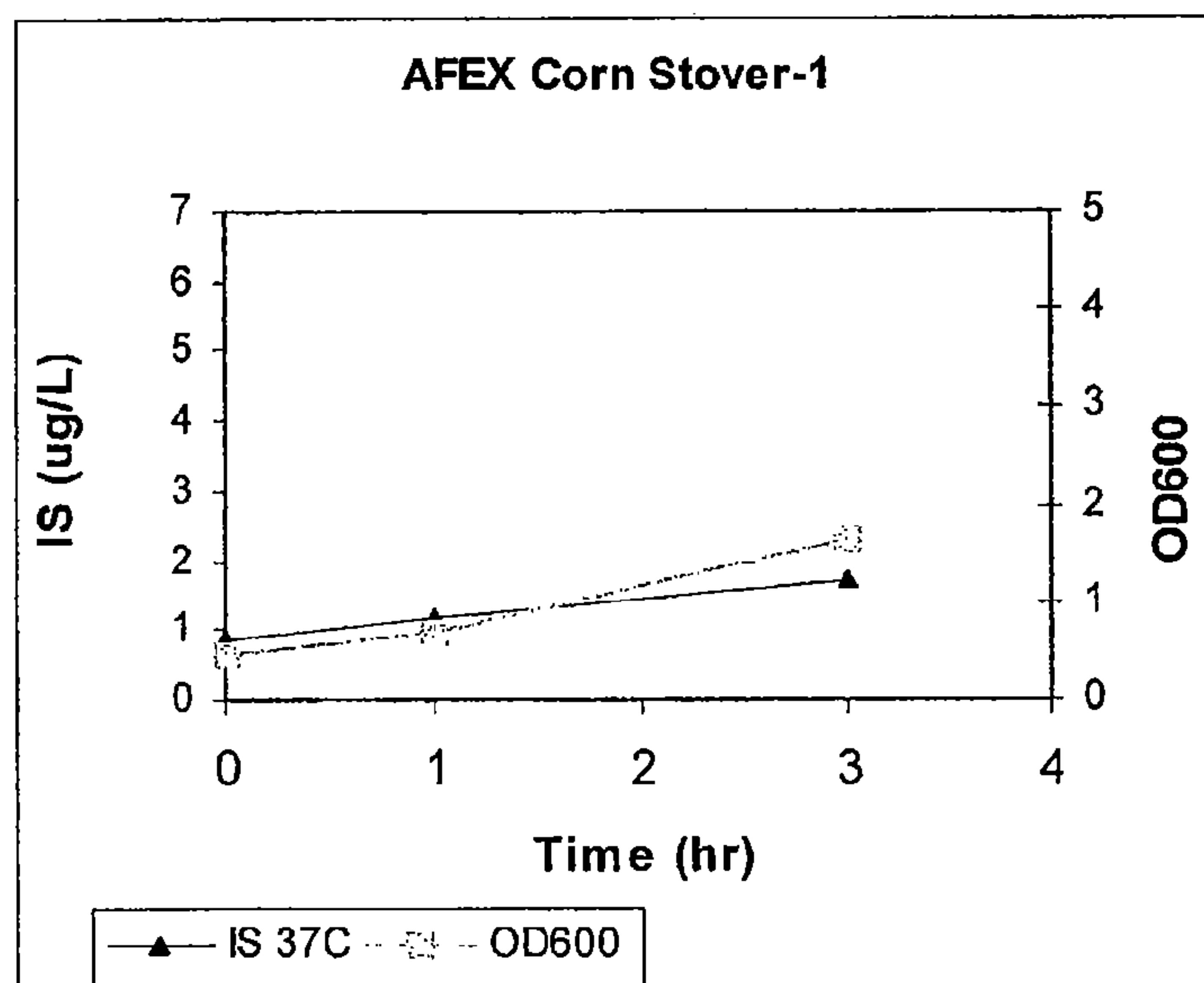


Figure 47B



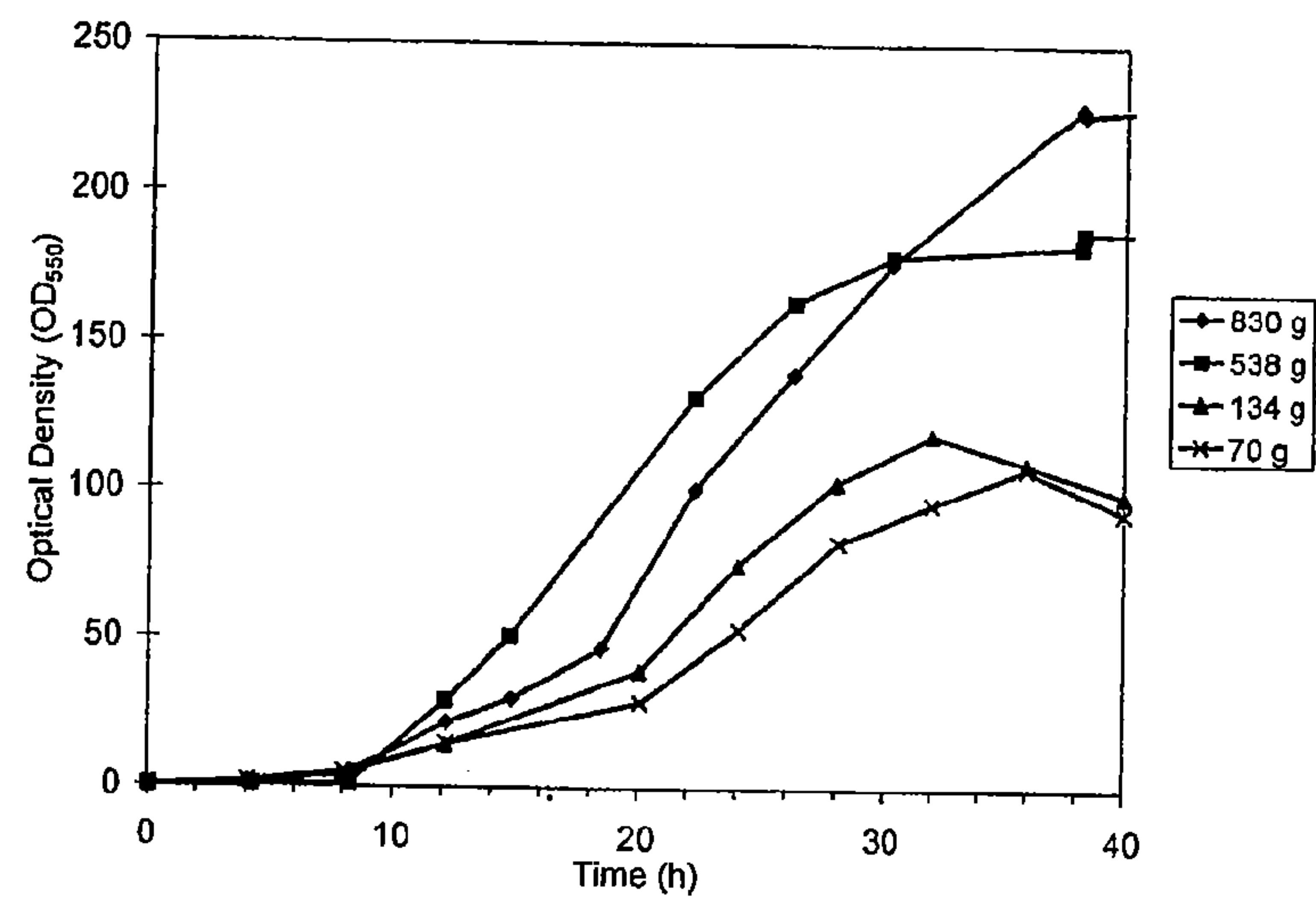
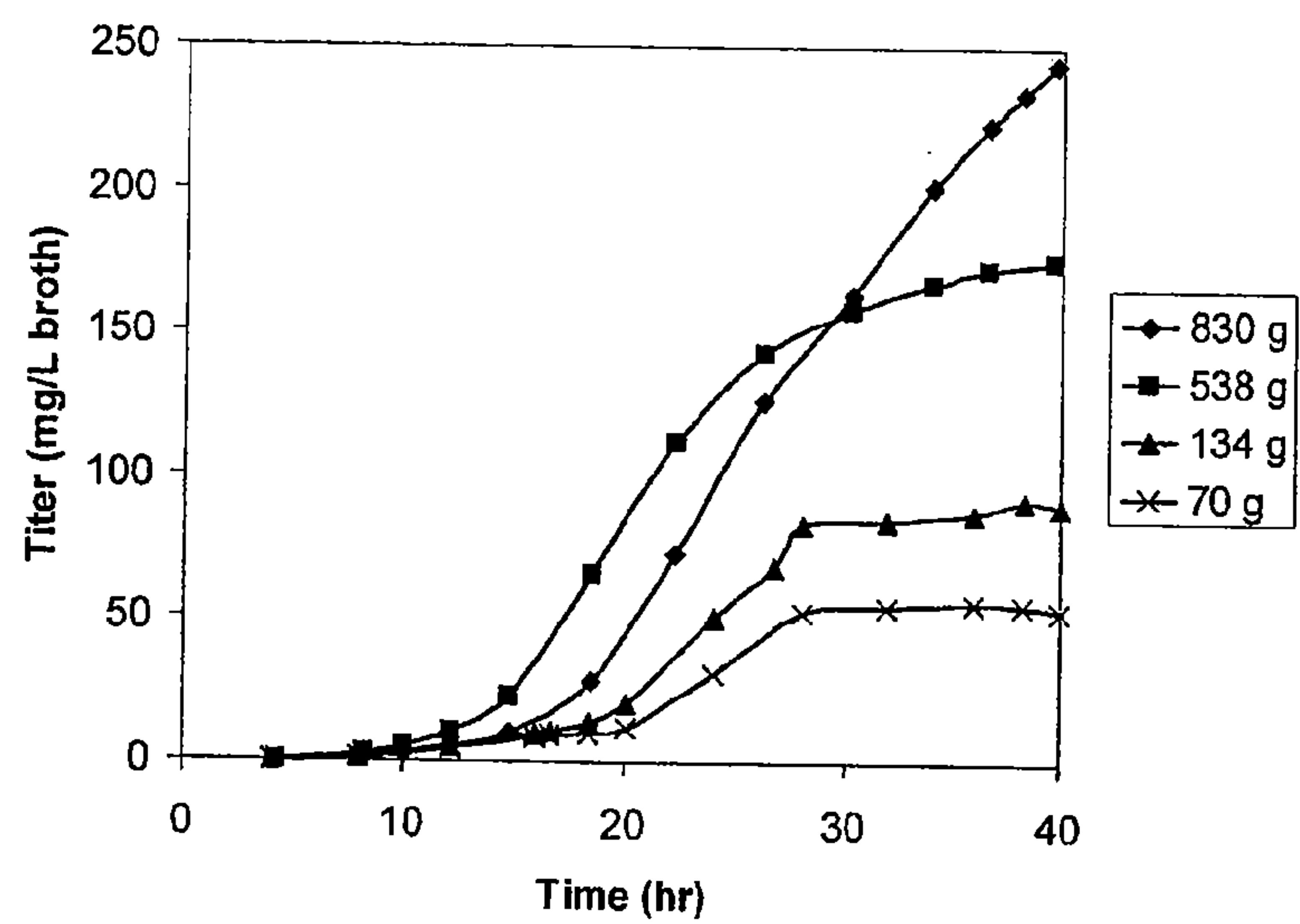


106/167

**Figure 47C****Figure 47D**

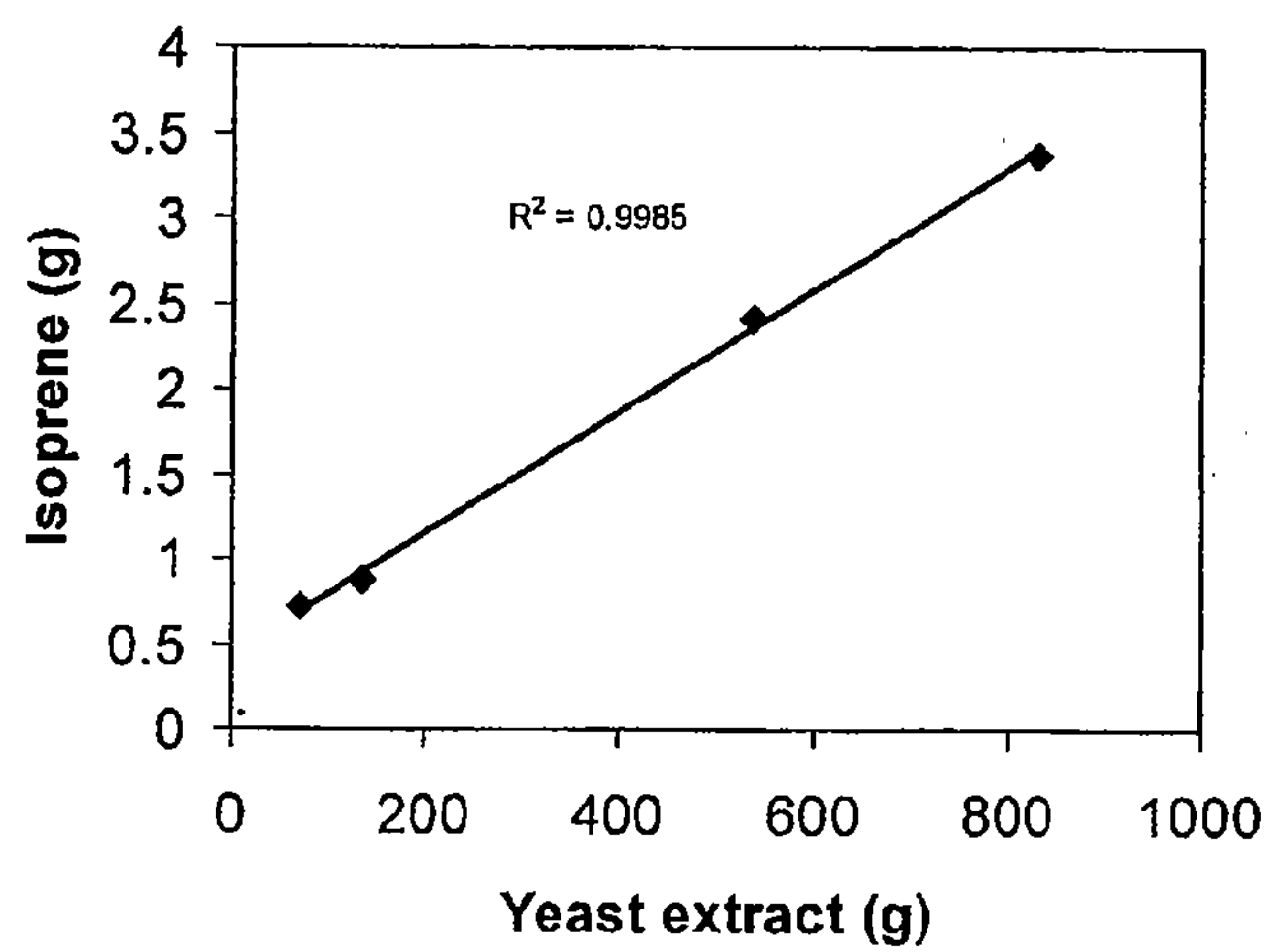


107/167

**Figure 48A****Figure 48B**

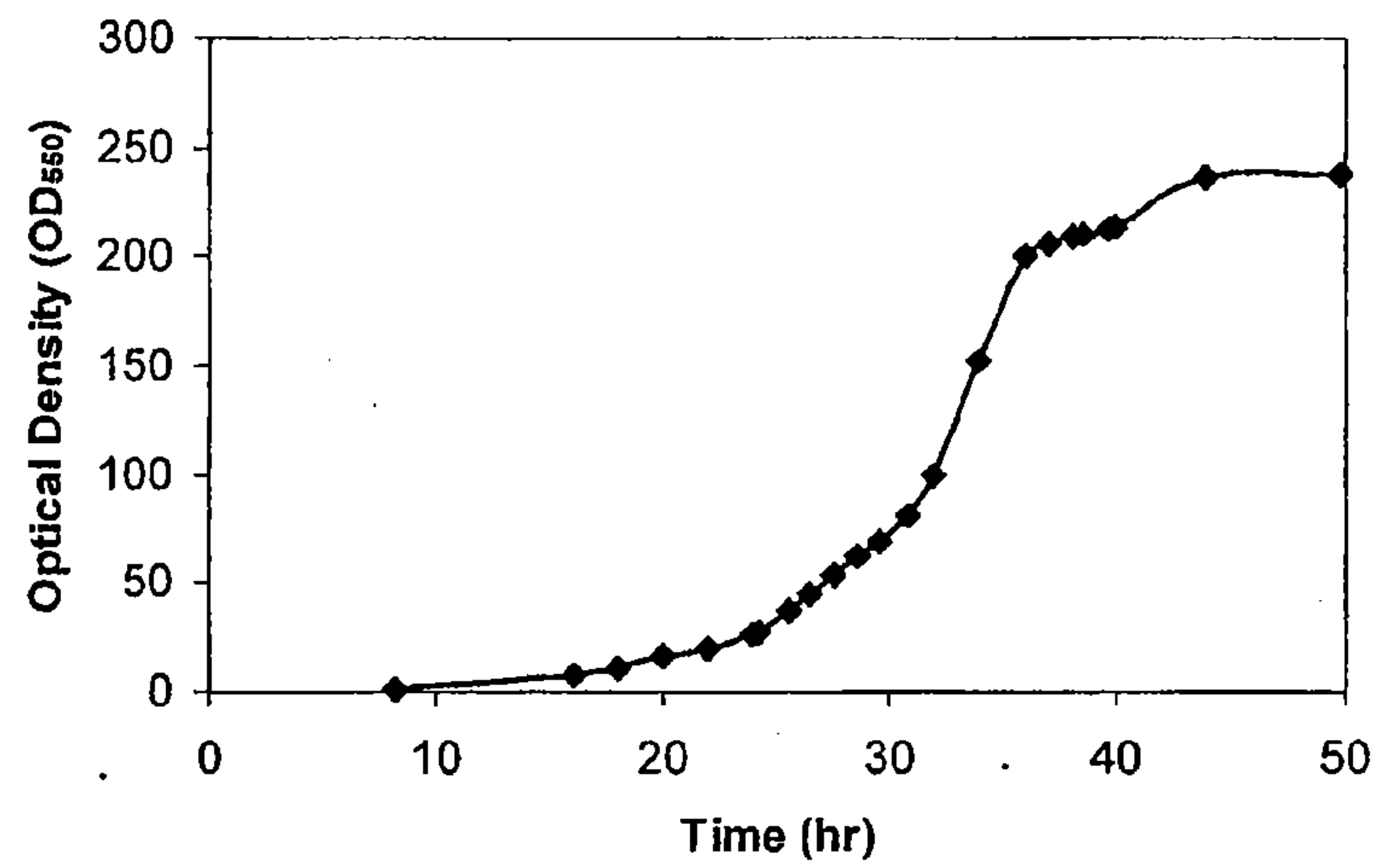
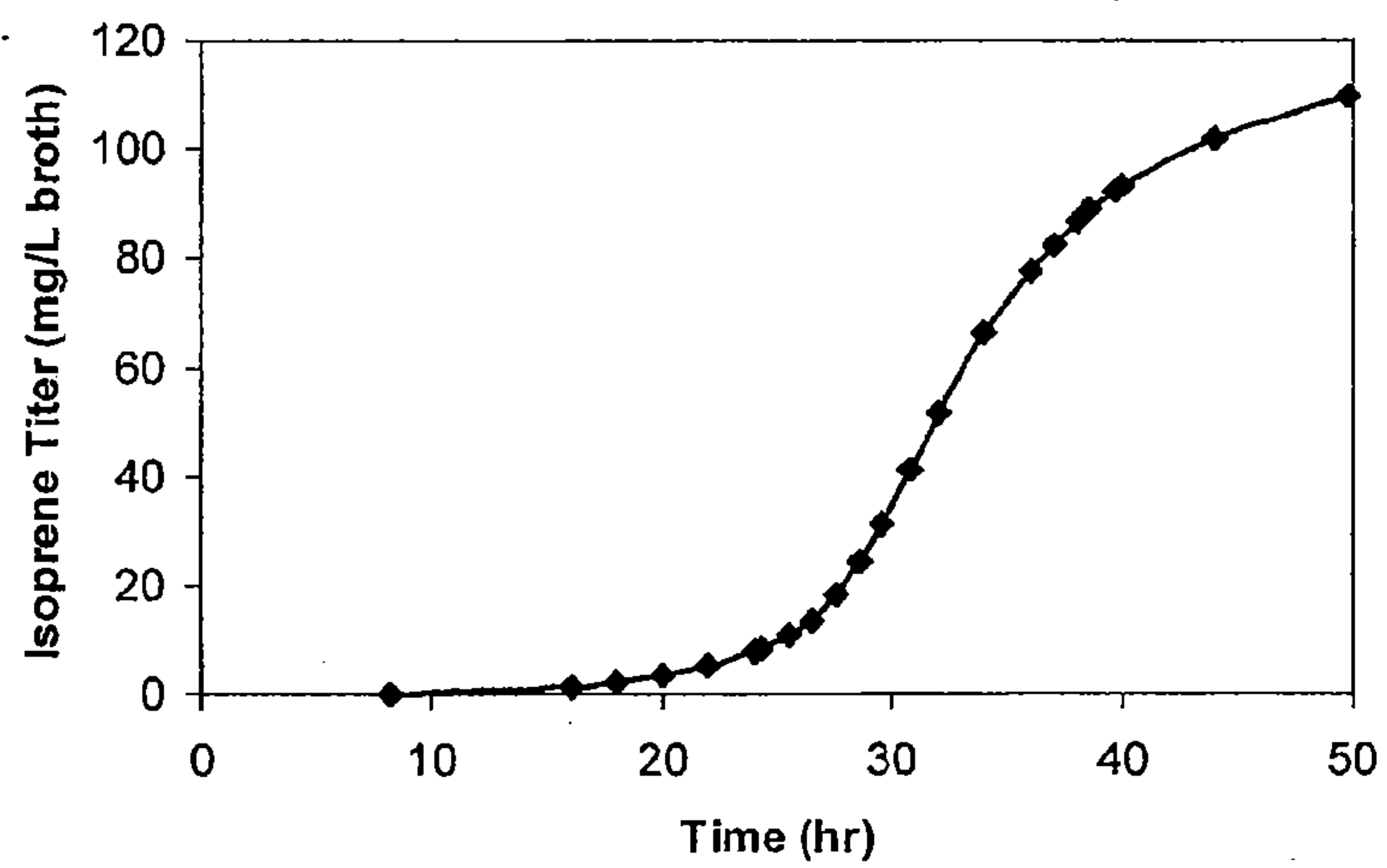


108/167

**Figure 48C**

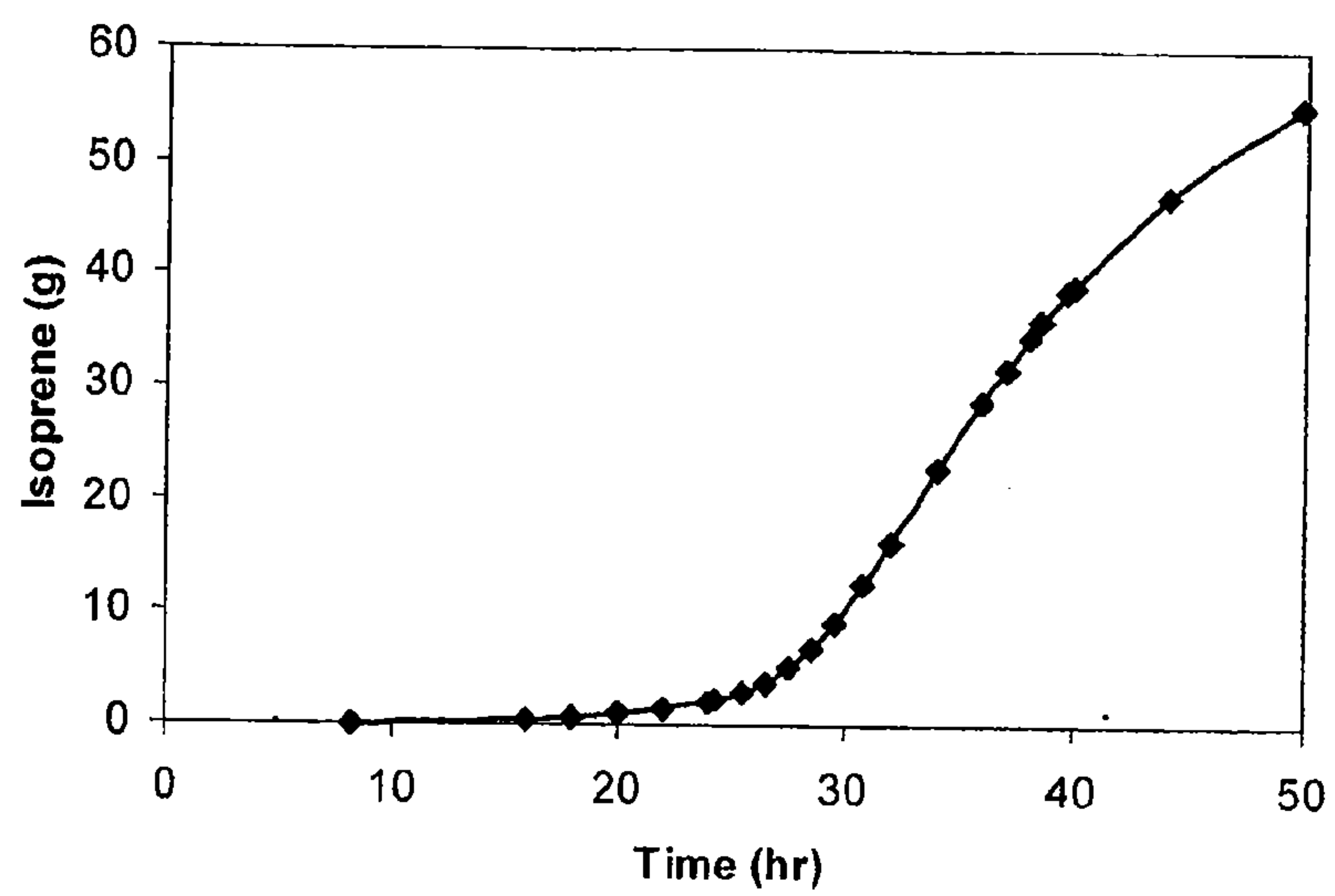


109/167

**Figure 49A****Figure 49B**

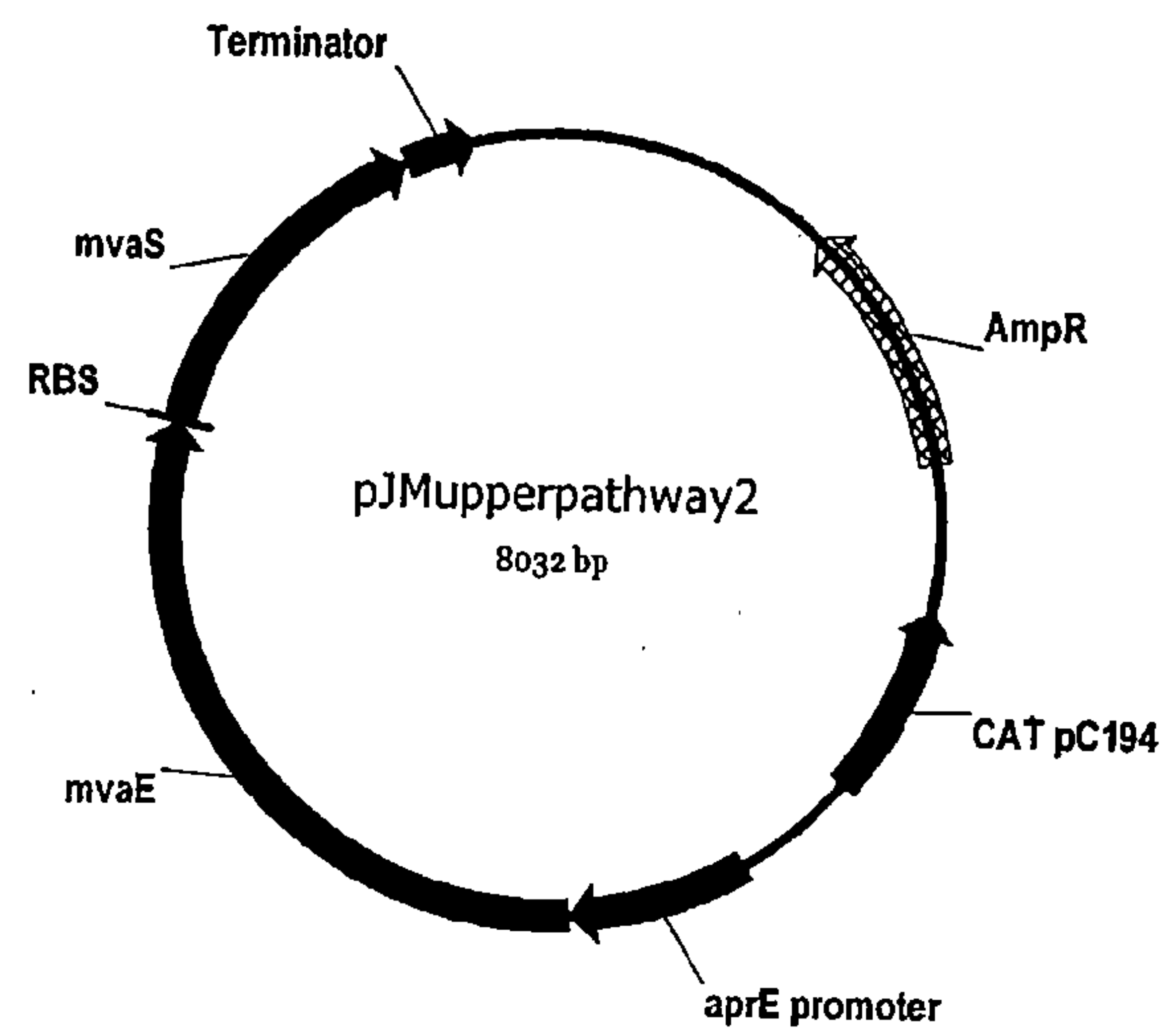


110/167

**Figure 49C**



111/167

**Figure 50**



112/167

**Figure 51A**

5' -

tcgctgcgctcggctcgttcggctgcggcgagcgggtatcagctcactcaaaggcggtaatacgggt  
tatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccag  
gaaccgtaaaaaggccgcgttgctggcggtttttccataggctccgccccctgacgagcatcac  
aaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataaccaggcgtttc  
cccctggaagctccctcgtgcgctctcctgttccgacctgocgcttaccggatacctgtccgc  
ctttctcccttcgggaagcgtggcgcttttctcatagctcacgctgtaggtatctcagttcgggtg  
taggtcgttcgctccaagctgggctgtgtgcacgaaccccccggttcagcccgaccgctgcgcct  
tatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagc  
cactggtaacaggattagcagagcgaggtatgtaggcgggtgctacagagttcttgaagtgggtg  
cctaactacggctacactagaagaacagtatcttggtatctgcgctctgctgaagccagttacct  
tcggaaaaagagttggttagctcttgatccggcaaaacaaaccacgctggtagcgggtggtttttt  
tgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttct  
acgggggtctgacgctcagtggaacgaaaactcacgttaagggatttttggtcatgagattatcaa  
aaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatata  
tgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgt  
ctatttcgttcatccatagttgcctgactccccgctcgtgtagataactacgatacgggagggt  
taccatctggccccagtgctgcaatgataccgcgagacccacgctcacccggctccagatttate  
agcaataaaccagccagccggaagggccgagcgcagaagtggctcctgcaactttatccgctcc  
atccagtcctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgca  
acgttggtgccattgctacaggcatcgtgggtgtcacgctcgtcgttttggtatggcttcattcag  
ctccggttcccaacgatcaaggcgagttacatgatcccccatgttggtgcaaaaaagcgggttagc  
tccttcgggtcctccgatcgttggtcagaagtaagttggccgcagtggtatcactcatgggttatgg  
cagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtagta  
ctcaaccaagtcattctgagaatagtgatgcgggcagccaggttgctcttgcccggcgtcaata  
cgggataataccgcgccacatagcagaactttaaaagtgtcatcattggaaaacggttcttcgg  
ggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaaccactcgtgcacc  
caactgatcttcagcatcttttactttcaccagcggttctgggtgagcaaaaacaggaaggcaa  
aatgccgcaaaaaagggaataagggcgacacggaaatgttgaaatactcatactcttcctttttc  
aatattattgaagcatttatcaggggttattgtctcatgagcggatacatatttgaaatgtattta  
gaaaaataaacaatataggggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaa  
accattattatcatgacattaacctataaaaaataggcgtatcacgagggccctttcgtctcgcgc  
gtttcgggtgatgacggtgaaaacctctgacacatgcagctcccggagacgggtcacagcttgctc  
gtaagcggatgccgggagcagacaagcccgtcagggcgcgctcagcgggtgttgggcgggtgtcgg  
ggctggcttaactatgcggcatcagagcagattgtactgagagtgaccatagatctggagctg  
taatataaaaaccttcttcaactaacggggcaggttagtgacattagaaaaccgactgtaaaaa  
gtacagtcggcattatctcatattataaaagccagtcattaggcctatctgacaattcctgaat  
agagttcataaacaatcctgcatgataaccatcacaacagaatgatgtacctgtaaagatagc  
ggtaaatatattgaattacctttattaatgaattttcctgctgtaataatgggtagaaggtaat  
tactattattattgatatttaagttaaaccagtaaatgaagtccatggaataatagaaagaga  
aaaagcattttcaggtataggtgttttgggaaacaatttccccgaaccattatatttctctaca  
tcagaaagggtataaatcataaaactctttgaagtcattctttacaggagtcacaataaccagaga  
atgttttagatacaccatcaaaaattgtataaagtggctctaacttatcccaataacctaactc  
tccgtcgtctattgtaaccagttctaaaagctgtatttgagtttatcacccttgctcactaagaaa  
ataaatgcagggtataaatttatatccttcttggttttatgtttc



113/167

**Figure 51B**

ggtataaaacactaatatcaatttctgtggttataactaaaagtcgtttggttggttcaaataatg  
attaaatatctcttttctcttccaattgtctaaatcaattttattaaagttcatttgatatgcc  
tcctaaatttttatctaaagtgaatttaggaggcttacttgtctgctttcttcattagaatcaa  
tcctttttttaaagtcfaatattactgtaacataaatatattttaaaaatatcccactttatc  
caattttcgtttggtgaactaatgggtgctttagttgaagaataaaagacctatgcggtgtgaa  
ataccgcacagatgcgtaaggagaaaataccgcacatcaggcgccattcgccattcagggtgcgca  
actgttgggaaggcgatcggtgcgggcctcttcgctattacgccagctggcgaaagggggatg  
tgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaaacgacg  
gccagtgccaagcttgcattgcctgcactccattttcttctgctatcaaaataacagactcgtga  
ttttccaaacgagctttcaaaaaagcctctgccccttgcaaatcggtatgcctgtctataaaatt  
cccgatattggttaaacagcgggcgcaatggcgggccgcattctgatgtctttgcttggcgaaatgtt  
catcttattttcttctcctcctctcaataattttttcattctatcccttttctgtaaagtttattt  
ttcagaatacttttatcatcatgctttgaaaaaatatcacgataatattcattgttctcacgga  
agcacacgcaggtcatttgaacgaattttttcgacagggaatttgccgggactcaggagcattta  
acctaaaaaagcatgacatttcagcataatgaacatttactcatgtctattttcgcttcttttct  
gtatgaaaatagttattttcgagctctctacggaaatagcgagagatgatatacctaaatagagat  
aaaatcatctcaaaaaaatgggtctactaaaatattattccatctattacaataaattcacaga  
atagctcttttaagtaagtctactctgaatttttttaaaggagagggtaaaagagtgaaaacagt  
agttattattgatgcattacgaacaccaattggaaaatataaaaggcagcttaagtcaagtaagt  
gccgtagacttaggaacacatgttacaacacaactttttaaagacattccactatttctgaag  
aaattgatcaagtaatctttggaaatgttttacaagctggaaatggccaaaatcccgcacgaca  
aatagcaataaacagcggtttgtctcatgaaattcccgcattgacgggttaatgaggtctgcgga  
tcaggaatgaaggccgttattttggcgaaacaattgattcaattaggagaagcggaagttttaa  
ttgctggcggttggagaatatgtcccaagcacctaaattacaacgtttttaattacgaaacaga  
aagctacgatgcgcctttttctagtatgatgatgatggattaacggatgcctttagtggtcag  
gcaatgggcttaactgctgaaaatgtggccgaaaagtatcatgtaactagagaagagcaagatc  
aattttctgtacattcacaattaaaagcagctcaagcacagcagaagggatattcgctgacga  
aatagccccattagaagtatcaggaacgcttgtggagaaagatgaagggttcgccctaattcg  
agcgttgagaagctaggaacgcttaaaacagttttttaaagaagacgggtactgtaacagcagga  
atgcatcaaccattaatgatggggcttctgctttgattattgcttcacaagaatatgccgaagc  
acacgggtcttcttatttagctattattcgagacagtgtggaagtcggtattgatccagcctat  
atgggaatttcgccgattaaagccattcaaaaactgttagcgcgcaatcaacttactacggaag  
aaattgatctgtatgaaatcaacgaagcatttgcagcaacttcaatcgtgggtccaaagagaact  
ggctttaccagaggaaaagggtcaacatttatgggtggcggtattttcattaggtcatgcgattgggt  
gccacaggtgctcgtttattaacgagtttaagttatcaattaaatcaaaaagaaaagaaatatg  
gagtggcttctttatgtatcggcgggtggcttaggactcgctatgctactagagagacctcagca  
aaaaaaaaaacagccgatttttatcaaatgagtcctgaggaacgcctggcttctcttcttaatgaa  
ggccagatttctgctgatacaaaaaaagaatttgaaaatacgggttttatcttcgcagattgcca  
atcatatgattgaaaatcaaatcagtgaacacagaagtgccgatgggcgttggcttacatttaac  
agtggacgaaactgattattttggtaccaatggcgacagaagagccctcagttattgcggctttg  
agtaatggtgcaaaaatagcacaaggattttaaacagtgatcaacaacgcttaatgcgtggac  
aaatcgttttttacgatgttgcagatcccagctcattgattgataaactacaagtaagagaagc  
ggaagtttttcaacaagcagagtttaagttatccatctatcgttaaacggggcgggcgttaaga  
gatttgcaatatcgctacttttgatgaatcatttgtatctgtcgacttttttagtagatgttaagg  
atgcaatgggggcaaatatcgtttaacgctatggttgaagggtgtg



114/167

**Figure 51C**

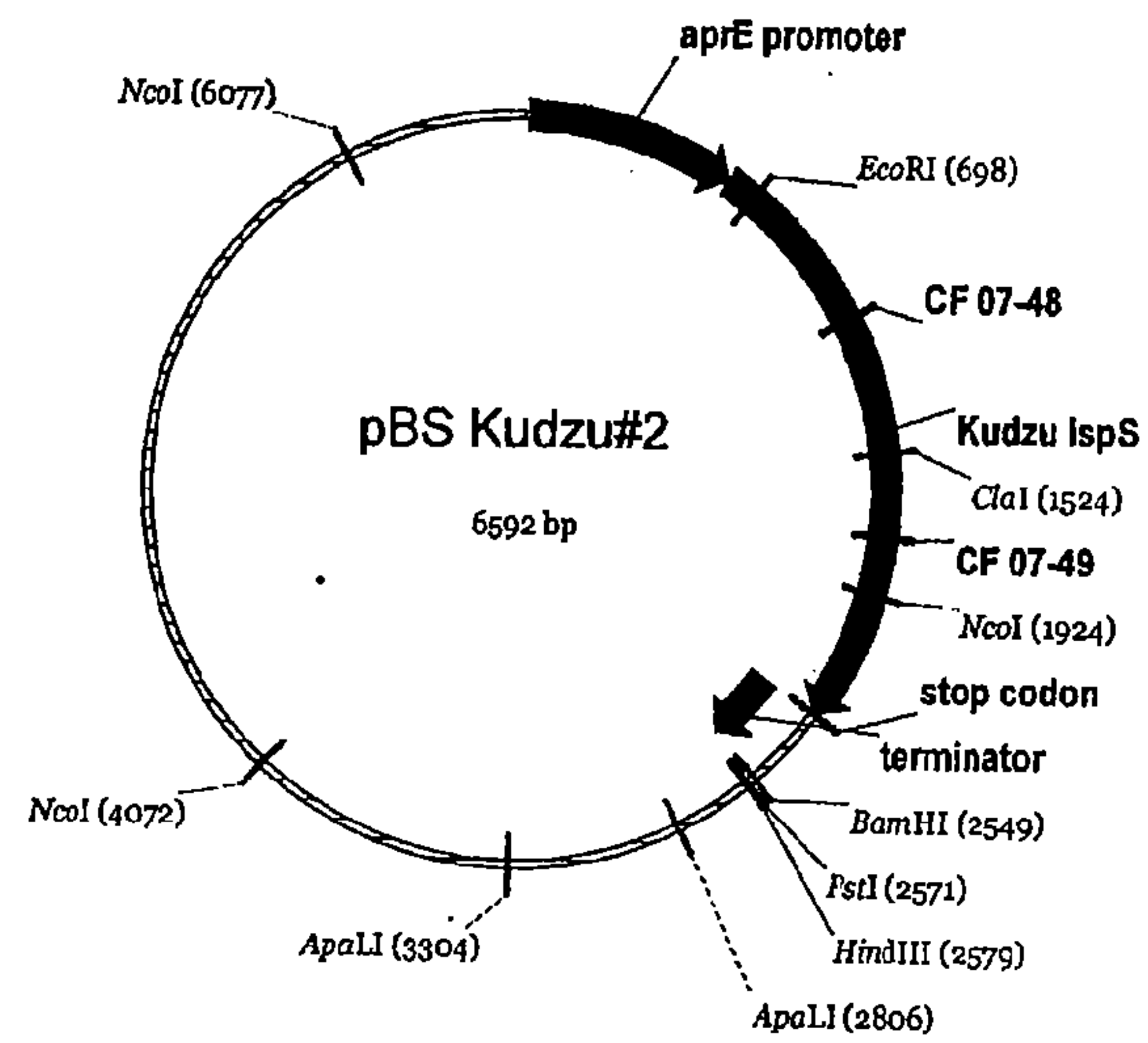
gccgagttgttccgtgaatggtttgcggagcaaaagattttattcagtattttaagtaattatg  
ccacggagtcggttggtacgatgaaaacggctattccagtttcacgtttaagtaaggggagcaa  
tggccgggaaattgctgaaaaaattgttttagcttcacgctatgcttcattagatccttatcgg  
gcagtcacgcataacaaaggaatcatgaatggcattgaagctgtagtttttagctacaggaaatg  
atacacgcgctggttagcgcttcttgtcatgcttttgcggtgaaggaaggctcgctaccaaggctt  
gactagttggacgctggatggcgaacaactaattggtgaaatttcagttccgcttgctttagcc  
acggttggcgggtgccacaaaagtcttacctaataatctcaagcagctgctgatttggttagcagtga  
cggatgcaaaagaactaagtcgagtagtagcggctgttggtttggcacaaaatttagcggcggtt  
acgggccttagtctctgaaggaattcaaaaaggacacatggctctacaagcacgttcttttagcg  
atgacggtcggagctactggtaaagaagttgaggcagtcgctcaacaattaaaacgtcaaaaaa  
cgatgaaccaagaccgagccatggctatttttaaatgatttaagaaaacaataaaaaggagaggggt  
gacaattgggattgataaaattagttttttgtgcccccttattatattgatatgacggcactg  
gctgaagccagaaatgtagaccctggaaaatttcatttgggtattgggcaagaccaaattggcgg  
tgaaccaatcagccaagatattgtgacatttgcagccaatgccgcagaagcgatcttgaccaa  
agaagataaagaggccattgatatggtgattgtcgggactgagtcagtatcgatgagtcaaaa  
gcggccgcagttgtcttacatcgtttaattggggattcaacctttcgctcgctctttcgaaatca  
aggaagcttggttacggagcaacagcaggcttacagtttagctaagaatcacgtagccttacatcc  
agataaaaaagtcttggtcgtagcggcagatattgcaaaatatggcttaaatcttggcgggtgag  
cctacacaaggagctggggcggttgcaatggttagttgctagtgaaccgcgcatttttggctttaa  
aagaggataatgtgatgctgacgcaagatatctatgacttttggcgtccaacaggccaccgta  
tcctatggtcgatggtcctttgtcaaacgaaacctacatccaatcttttgcccaagtctgggat  
gaacataaaaaacgaaccggtcttgattttgcagattatgatgcttttagcgttccatattcctt  
acacaaaaatgggcaaaaaagccttattagcaaaaatctccgaccaaactgaagcagaacagga  
acgaatttttagcccggttatgaagaaagtatcgtctatagtcgctcgctaggaacttgatatacg  
ggttcacttttatctgggactcatttcccttttagaaaaatgcaacgactttaaccgcaggcaatc  
aaattggttttattcagttatggttctggtgctgctgctgaatttttactggtgaattagtagc  
tggttatcaaaatcatttacaaaaagaaactcatttagcactgctggataatcggacagaactt  
tctatcgctgaatatgaagccatgtttgcagaaacttttagacacagacattgatcaaacgttag  
aagatgaattaaaatatagtatcttctgctattaataataaccgttcgttcttatcgaaactaaa  
aaaaccggccttggccccgcgggttttttattatttttcttccctccgcatgttcaatccgctcc  
ataatcgacggatggctccctctgaaaattttaacgagaaacggcgggttgacccggctcagtc  
ccgtaacggccaagtcctgaaacgtctcaatcgccgcttcccggtttccggtcagctcaatgcc  
gtaacggtcggcggcggttttctgataccgggagacggcattcgtaatcgggatccccgggtac  
cgagctcgaattcgtaatcatgtcatagctgtttcctgtgtgaaattgttatccgctcacaatt  
ccacacaacatacagagccggaagcataaagtgtaaagcctgggggtgcctaataagtgagtgagctaac  
tcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgca  
ttaatgaatcggccaacgcgcggggagaggcggtttgcgtattggggcgctcttccgcttcctcg  
ctcactgac

(SEQ ID NO:56)



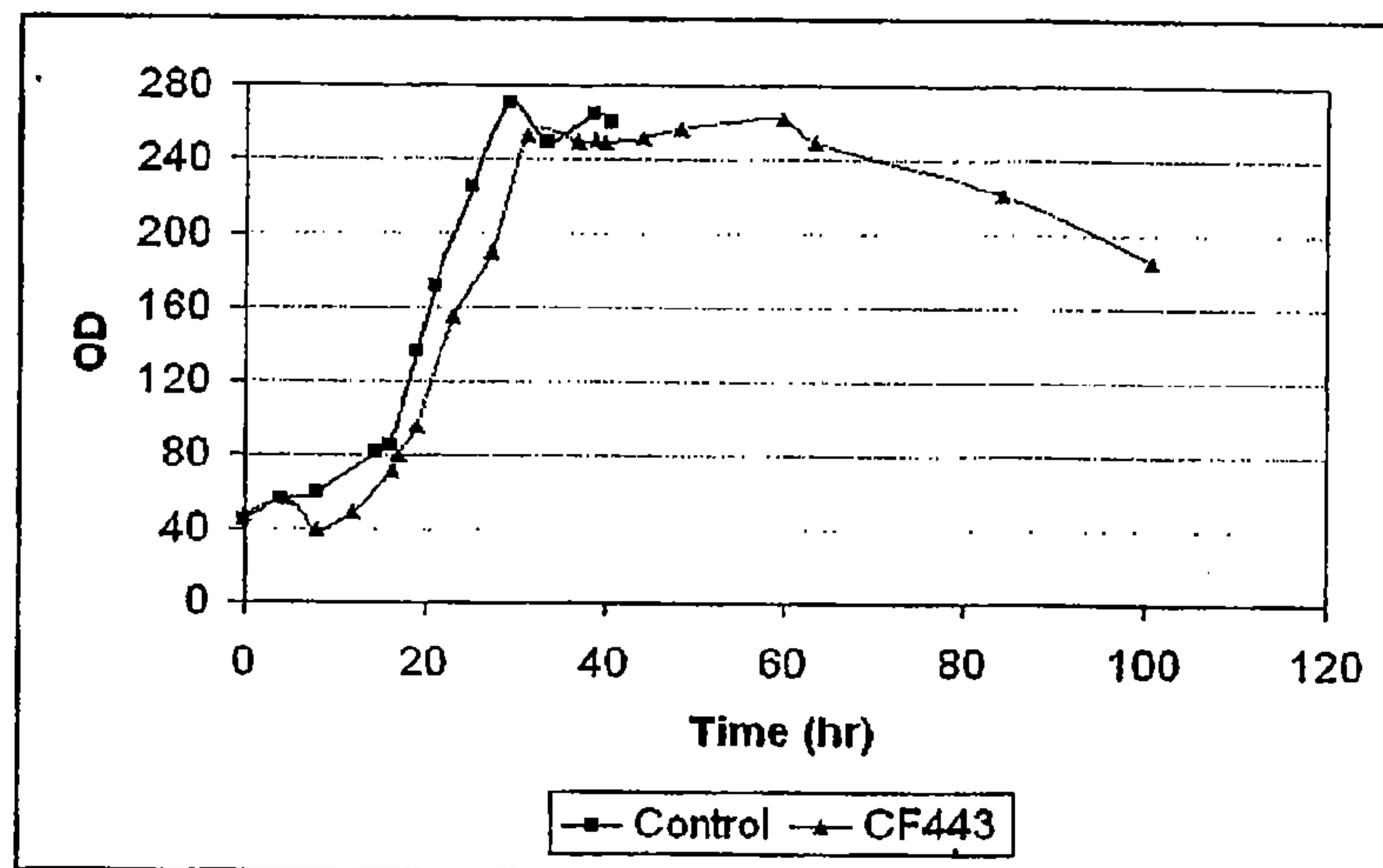
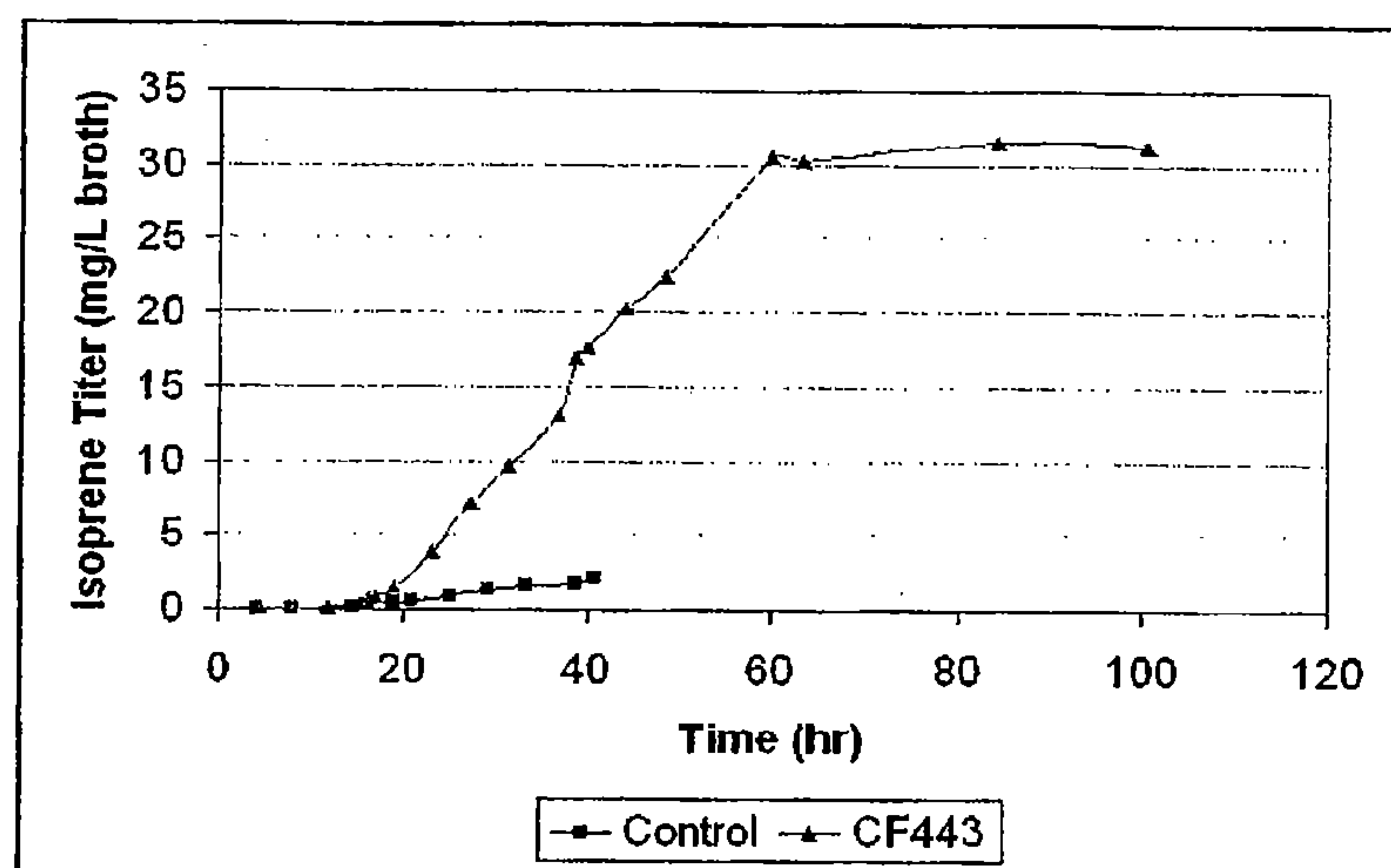
115/167

Figure 52



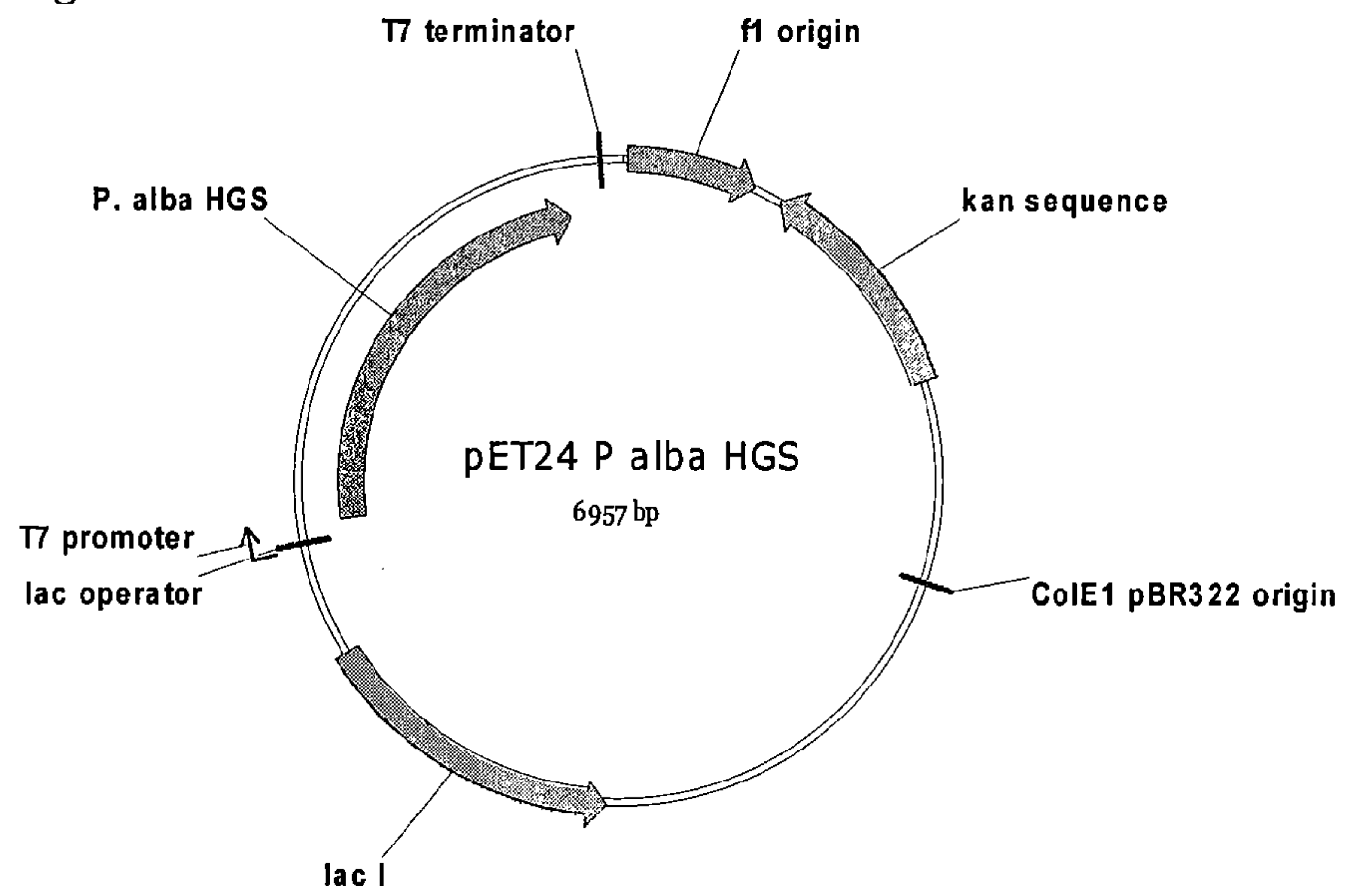


116/167

**Figure 53A****Figure 53B**



117/167

**Figure 54**



118/167

Figure 55A

1-

tggcgaatgggacgcgccctgtacggcgccattaagcgcggcggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccc  
tagcgcccgcctcttctgcttttcccttcccttctcgcacgttcgcccgttcccccgtcaagctctaaatcgggggctcccttaggggtcc  
gatttagtgcttacggcacctcgacccccaaaaacttgattaggggtgatgggtcacgtagtgggccatcgccctgatagacgggttttgcgcc  
tttgacgttggagtcacgttcttaatagtggactcttgttccaaaactggaacaacactcaaccctatctcggtctattcttttgatttataagggat  
tttgccgatttcggcctattgggttaaaaaatgagctgatttaacaaaaatttaacgcgaatttaacaaaatattaacgtttacaatttcaggtggca  
cttttcggggaaatgtgcgcggaacccctatttggttatttttctaaatacattcaataatgtatccgctcatgaattaattcttagaaaaactcatcg  
agcatcaaatgaaactgcaatttattcatatcaggattatcaataccatattttgaaaaagccgttctgtaataaggagaaaaactcaccgag  
gcagttccataggtatggcaagatcctggatcggctcgcgattccgactcgtccaacatcaatacaacctatttaatttcccctcgtcaaaaataa  
ggttatcaagtgagaaatcaccatgagtgcgactgaatccgggtgagaatggcaaaagtattatgcatttcttccagacttggtcaacaggcca  
gccattacgctcgtcatcaaaatcactcgcacatcaaccaaaccgttattcattcgtgattgcgcctgagcgcgagacgaaatacgcgacgctgtta  
aaaggacaattacaaacaggaatcgaatgaaccggcgaggaacactgccagcgcacatcaacaatattttcacctgaatcaggatattcttc  
taatacctggaatgctgttttcccggggatcgcagtggtgagtaacatgcacatcaggagtagcgataaaatgcttgatggtcggaagag  
gcataaattccgctcagccagtttagtctgaccatctcatctgtaacatcattggcaacgctaccttgccatgtttcagaaacaactctggcgcat  
cgggcttccatacaatcgatagattgtcgcacctgattgcccagacattatcgcgagccatttatacccatataaatcagcatccatgttgga  
ttaatcgcggcctagagcaagacgttcccgttgaaatggctcataacacccctgtattactgtttatgtaagcagacagttttattgttcatga  
ccaaaatcccttaacgtgagtttctgtccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttctgcgcgtaa  
tctgctgcttgcaacaaaaaaaccaccgctaccagcgggtggtttgttgcggatcaagagctaccaactcttttccgaaggtaactggcctt  
cagcagagcgcagataccaaatactgtccttctagtgtagccgttagtgaggccaccacttcaagaactctgtagcaccgcctacatacctcgc  
tctgctaactctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgca  
gcggctcgggctgaacggggggttcgtgcacacagcccagcttgagcgaacgacacacgaactgagatacctacagcgtgagctat  
gagaaagcgcacgcttcccgaaggagaaaggcgagaggtatccggtaagcggcagggtcggaaacaggagagcgcacgagggga  
gcttccagggggaaacgcctggtatctttatagtcctgtcgggttcgccacctctgacttgagcgtcgattttgtgatgctcgtcagggggg  
cggagcctatggaaaaacgccagcaacgcggccttttacgggtcctggccttttctggtccttttgcacatgttcttctcgttatccct  
gattctgtggataaccgtattaccgccttgagtgcgctgataccgctcgcgcagccgaacgaccgagcgcagcagtgactgagcgcgag  
gaagcgggaagagcgcctgatgcggtattttctccttacgcacatctgtgcggtatttcacaccgcataatgtgtgcactctcagtacaatctgctct  
gatgccgcatagttaagccagtatacactccgctatcgtacgtgactgggtcatggtgcgccccgacacccgccaacacccgctgacgc  
gccctgacgggcttgtctgtccccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagagggtttcacctgcatc  
accgaaacgcgcgaggcagctgcggttaaagctcatcagcgtggctcgtgaagcgattcacagatgtctgcctgttcacccgctcagctcg  
ttgagtttctccagaagcgttaatgtctggcttctgataaagcgggcatgttaaggcggtttttctgtttggtcactgatgcctccgtgtaag  
ggggatttctgttcatgggggtaatgataccgatgaaacgagagaggatgtcacgatacgggttactgatgatgaacatgcccgttactg  
gaacgttgtgagggtaaacaactggcggtatggatgcggcgggaccagagaaaaatcactcagggtcaatgccagcgttcgttaataca  
gatgtaggtgttccacagggtagccagcagcatcctgcgatgcagatccggaacataatggtgcaggcgctgacttccgcgtttccagac  
tttacgaaacacggaaaccgaagaccattcatgttgtgtcaggtcgcagacgttttcagcagcagtcgcttcacgttcgctcgcgtatcg  
gtgattcattctgctaaccagtaaggcaaccccgccagccttagccgggtcctcaacgacagggagcacgatcatgcgcacccgtggggcc  
gccatgccggcgataatggcctgcttctcgcgaaacgtttgggtggcgggaccagtgacgaaggcttgagcgcaggcggtgcaagattccg  
aataccgcaagcgcagagccgatcctcgtcgcgtccagcgaagcggtcctcgcgaaaatgacccagagcgtgcccggcacctgtc  
ctacgagttgcatgataaagaagacagtcataagtgcggcgacgatagtcaccccgcgcccaccggaaggagctgactgggttgaag  
gctctcaagggcacggctcagatcccgtgcctaatagtgagtaactacattaattgcgttgcgctcactgcccgtttccagtcggga  
aacctgtcgtgccagctgcattaatgaatcgccaacgcgcggggagagcggtttgcgtattgggcgccagggtggttttttccacca  
gtgagacgggcaacagctgattgcccttcaccgcctggcctgagagagttgcagcaagcggtccacgctggtttgccccagcaggcga  
aaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcgggtatcgtctatcccactaccgagatatccgcaccaacgcgc  
agccccgactcggtaatggcgcgcattgcgccagcgcctatctgatcgttggaaccagcatcgcagtggaacgatgccctcattcagc  
atttgcagtggtttgtgaaaaccggacatggcactccagtcgccttcccgttccgctatcggtgaatttgattgcgagtgagatattatgccag



119/167

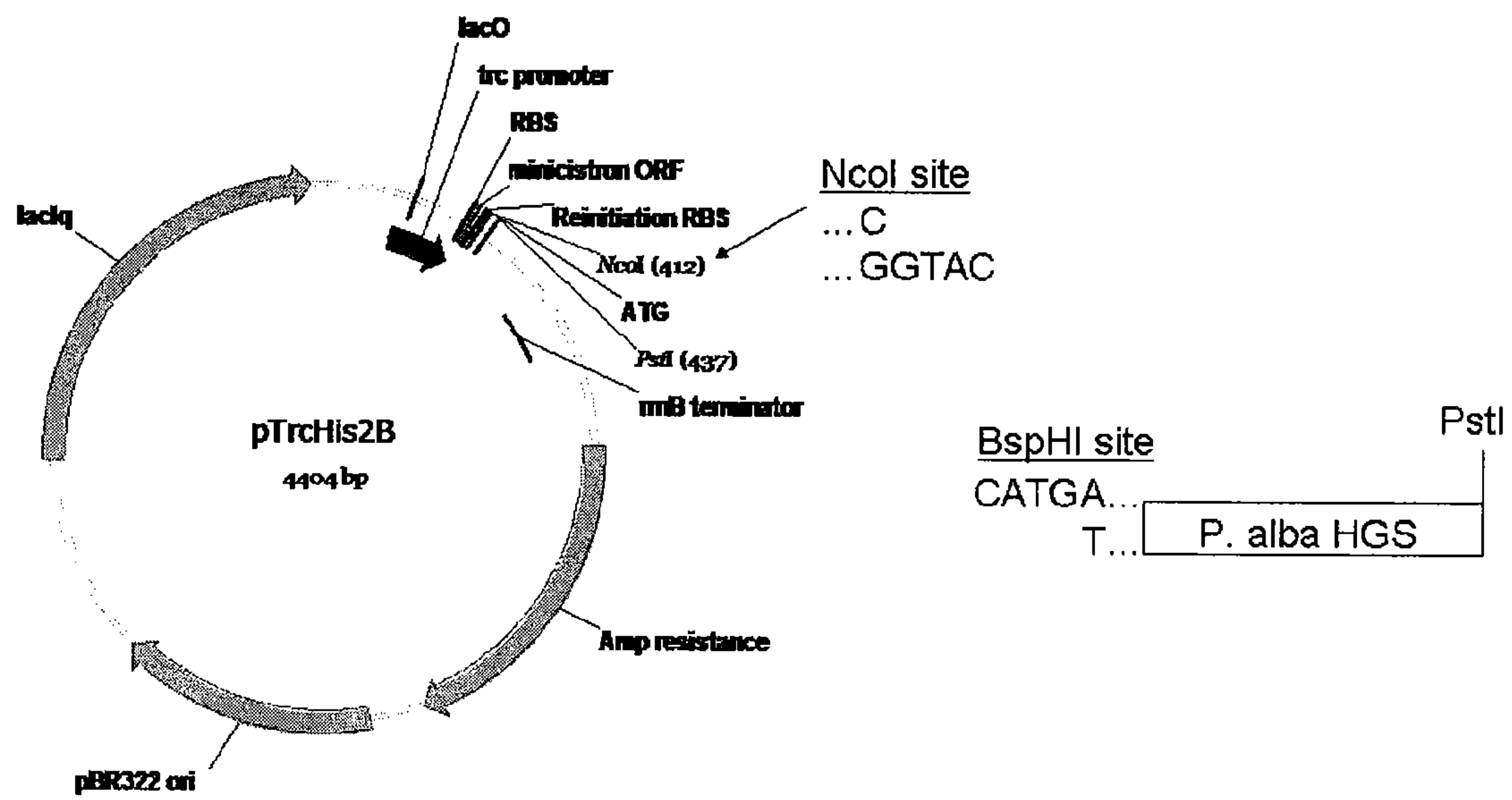
**Figure 55B**

ccagccagacgcagacgcgccgagacagaacttaatgggcccgctaacagcgcgatttgctggtgacccaatgcgaccagatgctccac  
gcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatgggtgtctggtcagagacatcaagaaataacgccggaacattagtc  
aggcagctccacagcaatggcatcctggatccagcggatagttaatgatcagcccactgacgcgttgccgcgagaagattgtgcaccgc  
cgctttacaggcttcgacgccgcttcgttctaccatcgacaccaccacgctggcaccagttgatcggcgcgagatttaacgccgcgacaa  
tttgcgacggcgctgcagggccagactggaggtggcaacgccaatcagcaacgactgtttgcccgccagttgttgccacgcggttgg  
gaatgtaattcagctccgccatcgccgcttccacttttcccgcgttttcgcagaaacgtggctggcctggttcaccacgcgggaaacggtct  
gataagagacaccggcatactctgcgacatcgataacgttactggtttcacattcaccacctgaattgactctcttcgggcgctatcatgc  
cataccgcgaaagggtttgcgccattcgatgggtgtccgggatctcgacgctctcccttatgcgactcctgcattaggaagcagcccagtagta  
ggttgaggccggtgagcaccgccgccgaaggaatggtgatgcaaggagatggcgcccaacagtcccccgccacggggcctgcca  
ccatacccacgccgaaacaagcgtcatgagcccgaagtggcgagcccgatcttcccatcggtgatgtcggcgatataggcgccagca  
accgcacctgtggcgccggtgatgccggccacgatgcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgactacta  
taggggaattgtgagcgggataacaattcccctctagaaataattttgtttaactttaagaaggagatatacatatgcgtttagcgtgtccaccg  
aaaatgtgtcttccaccgaaactgaaaccgaagctcgtcttgcgaactacgaacctaacagctgggactatgattacctgctgtcctccga  
cacggacgagtcctcgaagtatacaaagacaaaagcgaaaaagctggaagccgaagttcgtcgcgagattaataacgaaaaagcagaat  
ttctgaccctgctggaactgattgacaacgtccagcgctgggctgggttaccgtttcgagtctgatataccgtgggtgcgctggatcgcttcgt  
ttcctccggcggttcgatgcggtaaccaagacttccctgcacgggtacggcactgtctttccgtctgctgcgtcaacacgggtttgaggttctc  
aggaagcgttcagcggcttcaaagacaaaacggcaacttccctggagaacctgaaggaagatatcaaagctatcctgagcctgtacgagg  
ccagcttccctggctctggaaggcgaaaacatcctggacgaggcggaaggttttcgcaatctctcatctgaaagaactgtctgaagaaaagatc  
ggtaaagagctggcagaacaggtgaacctgcactggaactgccactgcacgccgtactcagcgtctggaagcagtatggtctatcgag  
gcctaccgtaaaaaggaggacgcgaatcaggttctgctggagctggcaattctggattacaacatgatccagtctgtataccagcgtgatct  
gcgtgaaacgtcccgttgggtggcgctggtgtgggtctggcgaccaaactgcactttgctcgtgaccgcctgattgagagcttctactgggccc  
tggtgttagcatcgaaccgcaatactccgactgccgtaactccgtcgcaaaaatgtttcttctgtaaccattatcgacgatatctacgatgat  
acggcacccctggacgaactggagctgttactgatgcagttgagcgttgggacgtaaacgccatcaacgacctgccggattacatgaaact  
gtgcttctggctctgtataacactattaacgaaatcgctacgacaacctgaaagataaagggtgagaacatcctgccgtatctgaccaaagc  
ctgggctgacctgtgcaacgcttccctgcaagaagccaagtggctgtacaacaaatctactccgacctttgacgactacttcggcaacgcatg  
gaaatcctctcttggcccgtgcaactgggtgttcgcttacttcgtgtcgtgcagaacattaaaaaggaagagatcgaaaacctgcaaaaata  
ccatgacaccatctctcgtcttcccatatcttccgtctgtgcaatgacctggctagcgcgtctgcggaaattgcgcgtggtgaaaccgcaaat  
agcgtttctgttacatgcgcactaaaggtatctccgaagaactggctaccgaaagcgtgatgaatctgatcgatgaaacctggaaaaagatg  
aacaaggaaaaactgggtggttagcctgttcgcgaaaccgttcgtggaaaccgcgatcaacctggcacgtcaatctcactgcacttatac  
cggcgacgcgcatacctctccggatgagctgacccgaaacgcgttctgtctgtaatactgaaccgattctgccgtttgaacgctaaggat  
ccgaattcgagctccgtcgacaagcttgcggccgactcgagcaccaccaccaccactgagatccggctgctaacaagcccgaaa  
ggaagctgagttggctgtgccaccgtgagcaataactagcataaccccttggggcctctaaacgggtcttgagggtttttgctgaaag  
gaggaactatatccggat (SEQ ID NO:87)



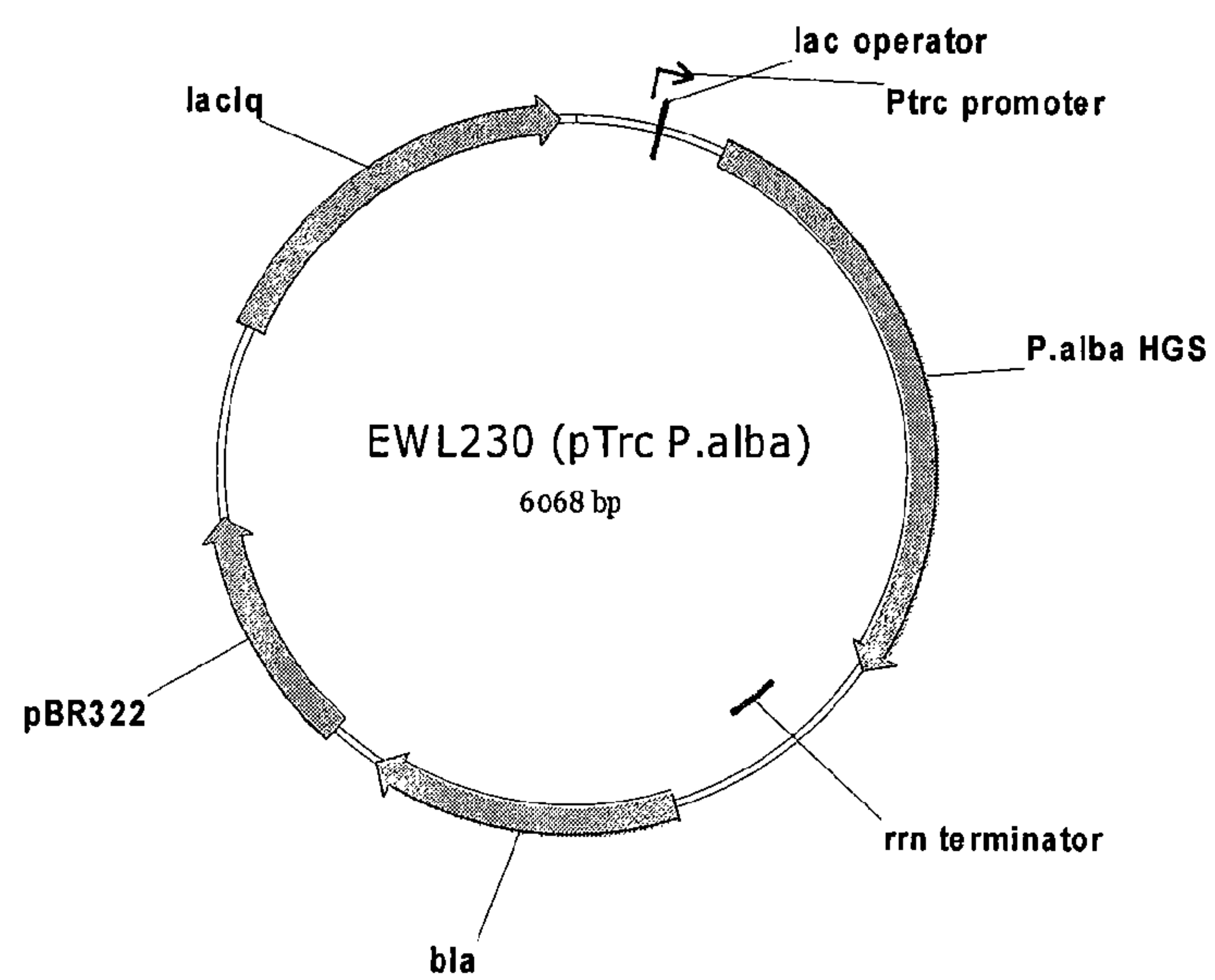
120/167

Figure 56





121/167

**Figure 57**



122/167

**Figure 58A**

1-

gtttgacagcttatcatcgactgcacgggtgcaccaatgcttctggcgtcaggcagccatcggaagctgtggtatggctgtgcaggtcgtaa  
cactgcataattcgtgtcgctcaaggcgcactcccgttctggataatgtttttgcgccgacatcataacggttctggcaaatattctgaaatgag  
ctgttgacaattaatcatccggctcgtataatgtgtggaattgtgagcggataacaatttcacacaggaaacagcgccgctgagaaaaagcg  
aagcggcactgctcttaacaatttatcagacaatctgtgtgggcactcgaccggaattatcgattaactttattattaaaaattaaagaggatat  
attaatgtatcgattaaataaggaggaataaaccatgagatgtagcgtgtccaccgaaaatgtgtcttcaccgaaactgaaaccgaagctcg  
tcgttctgcgaactacgaacctaacagctgggactatgattacctgctgtcctccgacacggacgagtcctcgaagtatacaagacaaag  
cgaaaaagctggaagccgaagttcgtcgcgagattaataacgaaaaagcagaatttctgacctgctggaactgattgacaacgtccagcg  
cctgggcctgggttaccgttctgagctcgtatccgtgggtgcgctggatcgcttcgttctcctccggcggcttcgatcggttaaccaagacttcc  
ctgcacgggtacggcactgtcttccgtctgctgcgtcaacacgggtttgaggtttctcaggaagcgttcagcggcttcaaagacccaaaacggc  
aacttctggagaacctgaaggaagatatcaaagctatcctgagcctgtacgagggcagcttctggctctggaaggcgaaaacatcctgg  
acgaggcggaaggttttcgcaatctctcatctgaaagaactgtctgaagaaaagatcggtaaagagctggcagaacaggtgaacctgcact  
ggaaactgccactgcacgccgtactcagcgtctggaagcagtatggtctatcgaggcctaccgtaaaaaggaggacgcgaatcaggttctg  
ctggagctggcaattctggattacaacatgatccagtctgtataaccagcgtgatctgcgtgaaacgtcccgttggtggcgtcgtgtgggtctg  
gcgaccaaaactgcactttgctcgtgaccgcctgattgagagcttctactgggccgtgggtgtagcattcgaaccgcaataactccgactgccg  
taactccgtcgcataaatgtttcttctgtaaccattatcgacgatatctacgatgtatacggcaccctggacgaactggagctgttactgatgc  
agttgagcgttgggacgtaaacgccatcaacgacctgccgattacatgaaactgtgcttctggctctgtataacactattaacgaaatcgcc  
tacgacaacctgaaagataaaggtgagaacatcctgccgtatctgaccaaagcctgggctgacctgtgcaacgcttctcctgcaagaagcca  
agtggctgtacaacaaatctactccgaccttgacgactacttcggcaacgcatggaaatcctcttctggcccgtgcaactggtgttcgctta  
cttcgctgtcgtgcagaacattaaaaaggaagagatcgaaaacctgcaaaaataccatgacaccatctctcgtccttcccatatcttccgtctg  
tgcaatgacctggctagcgcgtctgcggaaattgcgcgtggtgaaaccgcaaatagcgttcttgtttacatgcgcactaaaggtagtctccgaa  
gaactggctaccgaaagcgtgatgaatctgatcgatgaaacctggaaaaagatgaacaaggaaaaactgggtggtagcctgttcgcgaaa  
ccgttcgtggaaccgcgatcaacctggcacgtcaatctcactgcacttatcataacggcgacgcgcataacctctccgatgagctgacct  
gcaaacgcgttctgtctgtaactactgaaccgattctgccgttgaacgctaactgcagctggtaccatatgggaattcgaagcttctagaac  
aaaaactcatctcagaagaggatctgaatagcgcctgcaccatcatcatcatcattgagtttaaaccggtctccagcttggctgttttggcg  
gatgagagaagatttccagcctgatacagattaaatcagaacgcagaagcggctctgataaaacagaatttgcctggcggcagtagcgcggt  
ggtcccatcctgaccccatgccgaactcagaagtgaacgccgtagcgcgatggtagtgtgggtctcccatgcgagagtaggggaact  
gccaggcatcaataaaaacgaaaggctcagtcgaaagactgggccttctgtttatctgttgttgcgtgtaacgctctcctgagtaggacaa  
atccgcggggagcggatttgaacgttgcgaagcaacggcccgagggtggcgggcaggacggccgataaactgccaggcatcaaat  
taagcagaaggccatcctgacggatggccttttgcgtttctacaaactcttttgtttattttctaaatacattcaaatatgtatccgctcatgaga  
caataacctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgccttattcccttttttgcggcattttg  
ccttctgttttgcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtggttacctgaactggatctc  
aacagcggtaagatccttgagagtttgcggccgaagaacgtttccaatgatgagcacttttaaagtctgtatgtggcgcggtattatcccg  
tgttgacggcggaagagcaactcggctgcgcgatacactatttccagaatgacttgggtgagtactaccagtcacagaaaagcatcttac  
ggatggcatgacagtaagagaattatgcagtgtgccataacctgagtgataacactgcggccaacttacttctgacaacgatcggagga  
ccgaaggagctaacgcgtttttgcacaacatgggggatcatgtaactgccttgatcgttgggaaccggagctgaatgaagccataccaaa  
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaaactattaactggcgaactacttacttagcttcccggcaac  
aattaatagactggatggaggggataaagtgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggag  
ccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggttaagccctcccgtatcgtatctacacgacggggagtc  
ggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggttaactgtcagaccaagtttactcatatata  
cttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgcataatctcatgaccaaatacccttaacgtgagtttctgt  
ccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttctgcgcgtaatctgctgcttgcacacaaaaaacca  
ccgctaccagcgggtggttgttgcggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagataccaaataact  
gtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgctaactctgttaccagtggctgc



123/167

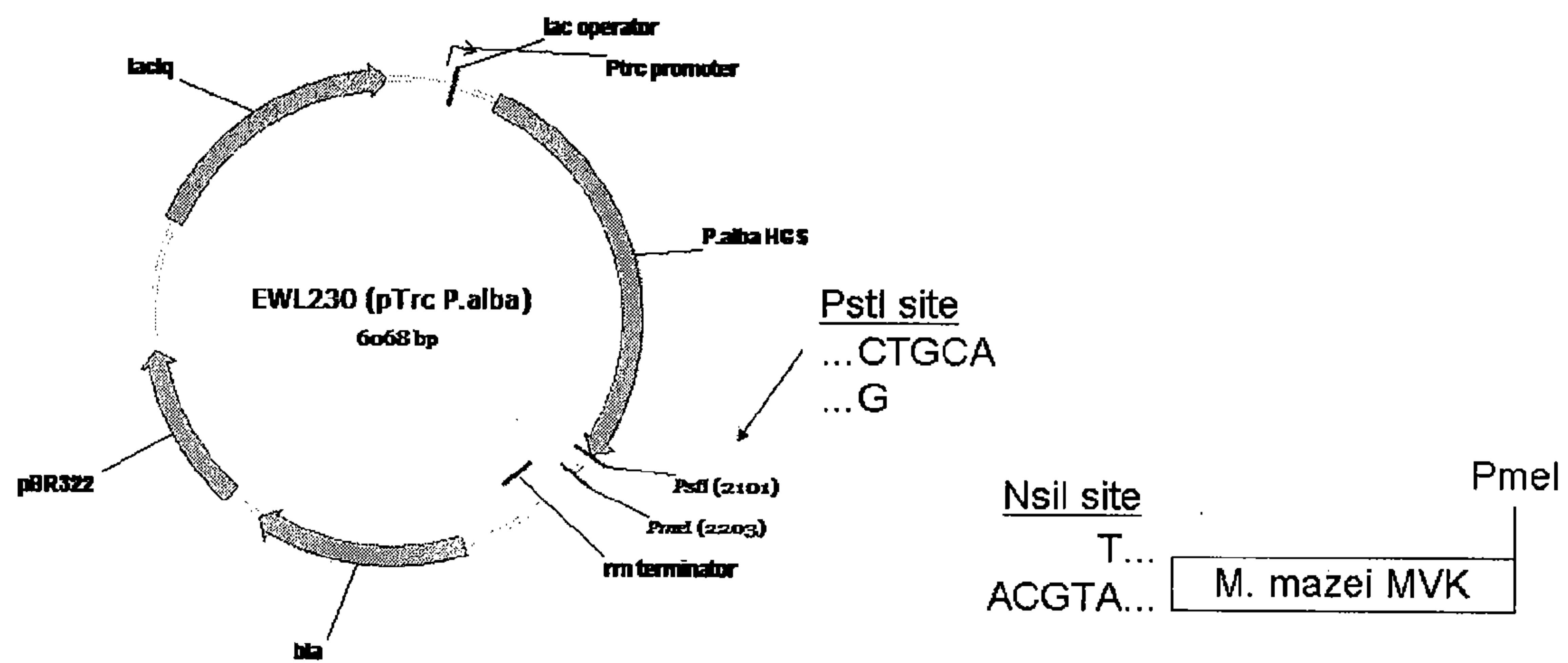
**Figure 58B**

tgccagtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggctcgggctgaacgggggggttc  
gtgcacacagcccagcttgagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaag  
ggagaaaggcggacaggtatccggttaagcggcagggctcgaacaggagagcgcacgagggagcttccagggggaaacgcctggtat  
ctttatagtcctgtcgggttcgccacctctgacttgagcgtcgatttttgatgctcgtcagggggcgagcctatggaaaaacgccagca  
acgcggcctttttacggttcctggccttttgcctggccttttgcacatgttcttctgcgttateccctgattctgtggataaccgtattaccgct  
ttgagtgagctgataccgctcggcgagccgaacgaccgagcgcagcagtcagtgagcaggaagcgggaagagcgcctgatgcggt  
atcttccttacgcatctgtcgggtatttcacaccgcataatggcgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacact  
ccgctatcgctacgtgactgggtcatggctgcgccccgacaccgccaacaccgctgacgcgccctgacgggcttgctgtctcccggca  
tccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttaccgctatcaccgaaacgcgcgaggcagcagatca  
attcgcgcgcgaaggcgaagcggcatgcatttacgttgacaccatcgatgggtgcaaaacctttcgcggatggcatgatagcgccggaa  
gagagtcaattcaggggtggtgaatgtgaaaccagtaacgttatacagatgtcgcagagtatgccggtgtctcttatacagaccgttcccgcgtg  
gtgaaccaggccagccacgtttctgcgaaaacgcgggaaaaagtggaaagcggcgatggcggagctgaattacattcccaaccgctggc  
acaacaactggcggggcaaacagtcgttgctgattggcgttgccacctccagctcggccctgcacgcgccgtcgcaattgtcgcggcgatt  
aaatctcgcgccgatcaactgggtgccagcgtggtggtgtcgatggtagaacgaagcggcgctgaagcctgtaaagcggcggtgcacaa  
tcttctcgcgcaacgcgtcagtggtgctgataactatccgctggatgaccaggatgccattgctgtggaagctgcctgcactaatgttccg  
gcgttatttcttgatgtctctgaccagacacccatcaacagtatttttctccatgaagacgggtacgcgactgggcgtggagcatctggtcgc  
attgggtcaccagcaaactcgcgtgttagcggggccattaagtctgtctcggcgcgctcgtcgtctggctggctggcataaatatctcactcg  
caatcaaattcagccgatagcggaaacgggaaggcgactggagtgccatgtccggttttaacaaacctgcaaatgctgaatgagggcac  
gttcccactgcgatgctggttgccaacgatcagatggcgtgggcgcaatgcgcgccattaccgagtcggggctgcgcgttggtgcggat  
atctcggtagtgggatacgacgataccgaagacagctcatgttatatcccgcctcaaccacctcaaacaggattttcgcctgctggggca  
aaccagcgtggaccgcttgctgcaactctcagggccaggcgggtgaagggaatcagctgttgcctcactggtgaaaagaaaaac  
cacctggcgcccaatacgaaccgcctctcccgcgcggttgccgattcattaatgcagctggcagcacaggttcccgactggaaagc  
gggcagtgagcgaacgcaattaatgtgagttagcgcgaattgatctg (SEQ ID NO:88)



124/167

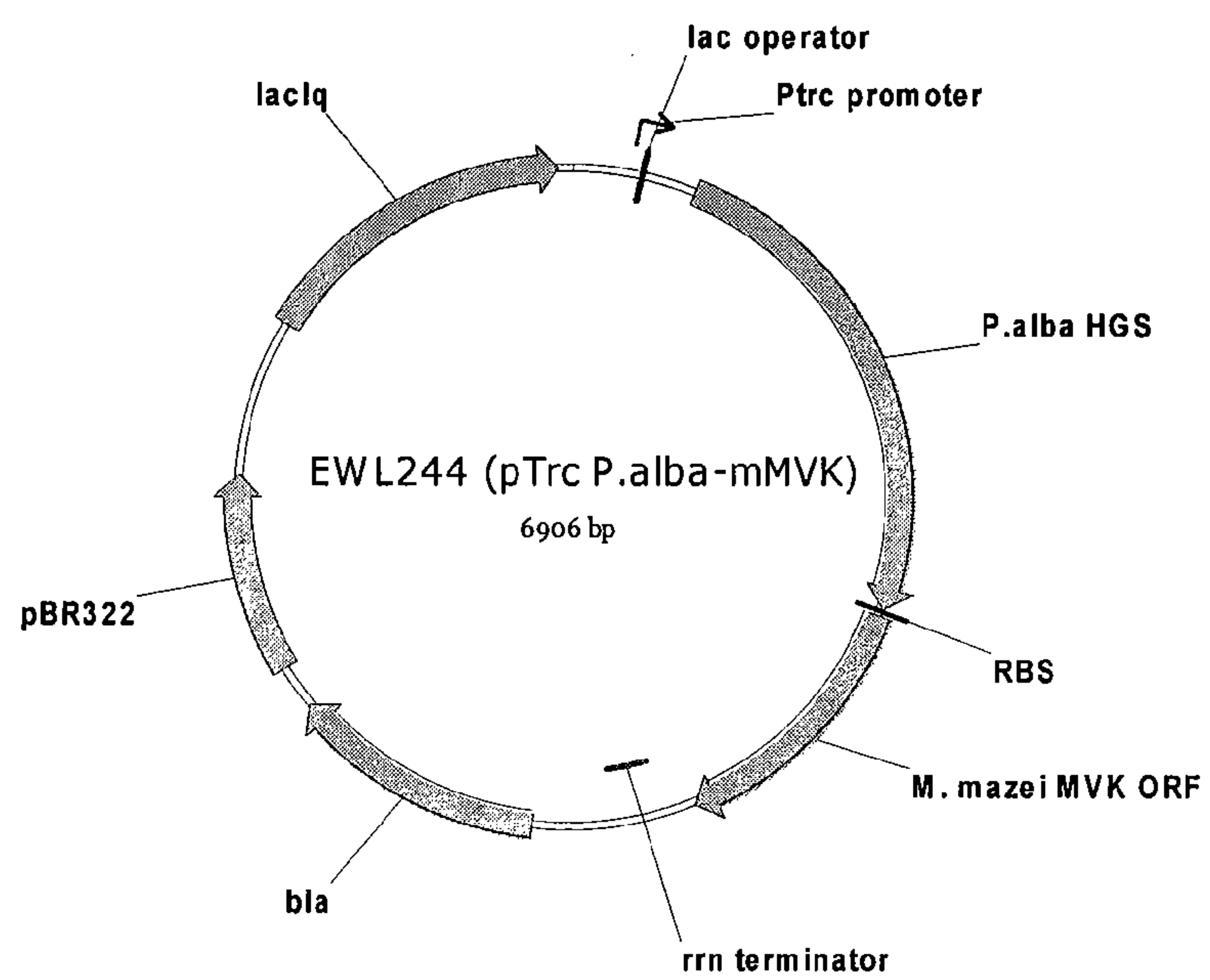
Figure 59





125/167

Figure 60





126/167

Figure 61A

1-

gtttgacagcttatcatcgactgcacgggtgcaccaatgcttctggcgtcaggcagccatcggaagctgtggtatggctgtgcaggtcgtaaat  
cactgcataattcgtgtcgctcaaggcgcactcccgttctggataatgtttttgcgcccacatcataacgggtctggcaaatattctgaaatgag  
ctgttgacaattaatcatccggctcgtataatgtgtggaattgtgagcggataacaatttcacacaggaaacagcgcgcgtgagaaaaagcg  
aagcggcactgctctttaacaatttatcagacaatctgtgtgggcactcgaccggaattatcgattaactttattattaaaaattaaagaggtatat  
attaatgtatcgattaaataaggaggaataaaccatgagatgtagcgtgtccaccgaaaatgtgtctttcaccgaaactgaaaccgaagctcg  
tcgttctgcgaactacgaacctaacagctgggactatgattacctgctgtcctccgacacggacgagtcctcgaagtatacaaaagacaaaag  
cgaaaaagctggaagccgaagttcgtcgcgagattaataacgaaaaagcagaatttctgacctgctggaactgattgacaacgtccagcg  
cctgggcctgggttaccgtttcagtgctgataatccgtgggtgcgctggatcgcttcgtttcctccggcggcttcgatgcggttaaccaagacttcc  
ctgcacgggtacggcactgtctttccgtctgctgcgtcaacacgggttttgagggtttctcaggaagcgttcagcggcttcaaagacccaaaacggc  
aacttctggagaacctgaaggaagatatcaaagctatcctgagcctgtacgaggccagcttctggctctggaaggcgaaaacatcctgg  
acgagggcgaaggttttcgcaatctctcatctgaaagaactgtctgaagaaaagatcggtaaagagctggcagaacaggtgaaccatgcact  
ggaaactgccactgcacgcctgactcagcgtctggaagcagtatggtctatcgaggccctaccgtaaaaaggaggacgcgaatcaggttctg  
ctggagctggcaattctggattacaacatgatccagtctgtataaccagcgtgatctgcgtgaaacgtcccgttgggtggcgtcgtgtgggtctg  
gcgaccaaactgcactttgctcgtgaccgcctgattgagagcttctactgggccgtgggtgtagcattcgaaccgcaataactccgactgccg  
taactccgtcgcataaatgttttcttcgtaaccattatcgacgatatctacgatgtatacggcaccctggacgaactggagctgtttactgatgc  
agttgagcgttgggacgtaaacgccatcaacgacctgccgattacatgaaactgtgctttctggctctgtataacactattaacgaaatcgcc  
tacgacaacctgaaagataaagggtgagaacatcctgccgtatctgaccaaagcctgggctgacctgtgcaacgctttcctgcaagaagcca  
agtggctgtacaacaaatctactccgacctttgacgactacttcggcaacgcagtggaatcctcttctggcccgtgcaactggtgttcgctta  
cttcgctgtcgtgcagaacattaaaaaggaagagatcgaaaacctgcaaaaataccatgacaccatctctcgtccttcccatatcttccgtctg  
tgcaatgacctggctagcgcgtctgcggaaattgcgctggtgaaaccgcaaatagcgtttctgttacatgcgcactaaaggatctccgaa  
gaactggctaccgaaagcgtgatgaatctgatcgatgaaacctggaaaaagatgaacaaggaaaaactgggtggtagcctgttcgcgaaa  
ccgttcgtggaaccgcgatcaacctggcacgtcaatctcactgcacttatcataacggcgacgcgcatacctctccggatgagctgaccc  
gcaaacgcgttctgtctgtaactactgaaccgattctgccgtttgaacgctaactgcataaaggagggtaaaaaaacatggtatcctgttctgcg  
ccgggtgaagatttacctgttcgggtgaacacgccgtagtttatggcgaaactgcaattgcgtgtgcggtggaactgcgtaccctgttcgcgcg  
gaactcaatgactctatcactattcagagccagatcgccgcacccggtctggatttcgaaaagcacccttatgtgtctgcggttaattgagaaa  
atgcgcaaatctattcctattaacgggtgttttctgaccgtcgattccgacatcccgggtgggctccgggtctgggttagcagcgcagccgttactat  
cgcgctctattggtgcgctgaacgagctgttcggctttggcctcagcctgcaagaaatcgctaaactggggccacgaaatcgaaattaaagtac  
agggtgccgcgtccccaaccgatacgtatgtttctaccttcggcgggcgtgggttaccatcccggaaacgtcgcaaacgaaactccggactg  
cggcattgtgattggcgataccggcggtttctcctccaccaaagagttagtagctaacgtacgtcagctgcgcgaaagctacccggattgat  
cgaaccgctgatgacctctattggcaaaatctctcgtatcggcgaaacactggttctgtctggcgactacgcacatccatcgccgcctgatgaa  
cgtcaaccagggtctcctggacgccctgggcgttaacatcttagaactgagccagctgatctattccgctcgtgcggcaggtgcgtttggcg  
ctaaaatcacgggcgtggtggcggtggtgtatggttgcgctgaccgctccggaaaaatgcaaccaagtggcagaagcggtagcaggc  
gctggcggtaaagtactatcactaaaccgaccgagcaaggtctgaaagtagattaaagtctagttaaagtttaaacggtctccagcttggct  
gttttggcggtatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggctctgataaaacagaatttgcctggcggcagt  
agcgcggtgggtcccacctgaccccatgccgaactcagaagtgaacgccgtagcgcgatggtagtgtgggtctccccatgcgagagt  
agggaactgccaggcatacaataaaacgaaaggctcagtcgaaagactgggctttcgtttatctgttgtttgtcgggtgaacgctctcctgag  
taggacaaatccgccgggagcggatttgaacgttgcgaagcaacggcccgagggtggcgggcaggacgcccgcataaaactgccag  
gcatcaaattaagcagaaggccatcctgacggatggccttttgcgtttctacaaactcttttgtttattttctaaatacattcaataatglatccg  
ctcatgagacaataacctgataaatgttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgccttattccctttttg  
cggcattttgccttctgttttgcacccagaaacgctggtgaaagttaaagatgctgaagatcagttgggtgcacgagtggttacctga  
actggatctcaacagcggtaagatccttgagagttttcgcgccgaagaacgtttccaatgatgagcacttttaagttctgctatgtggcgcg  
gtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttgggtgagtactaccagtcacagaa  
aagcatcttacggatggcatgacagtaagagaattatgcagtgtgccataacctgagtgataacactgcggccaacttacttctgacaacg



127/167

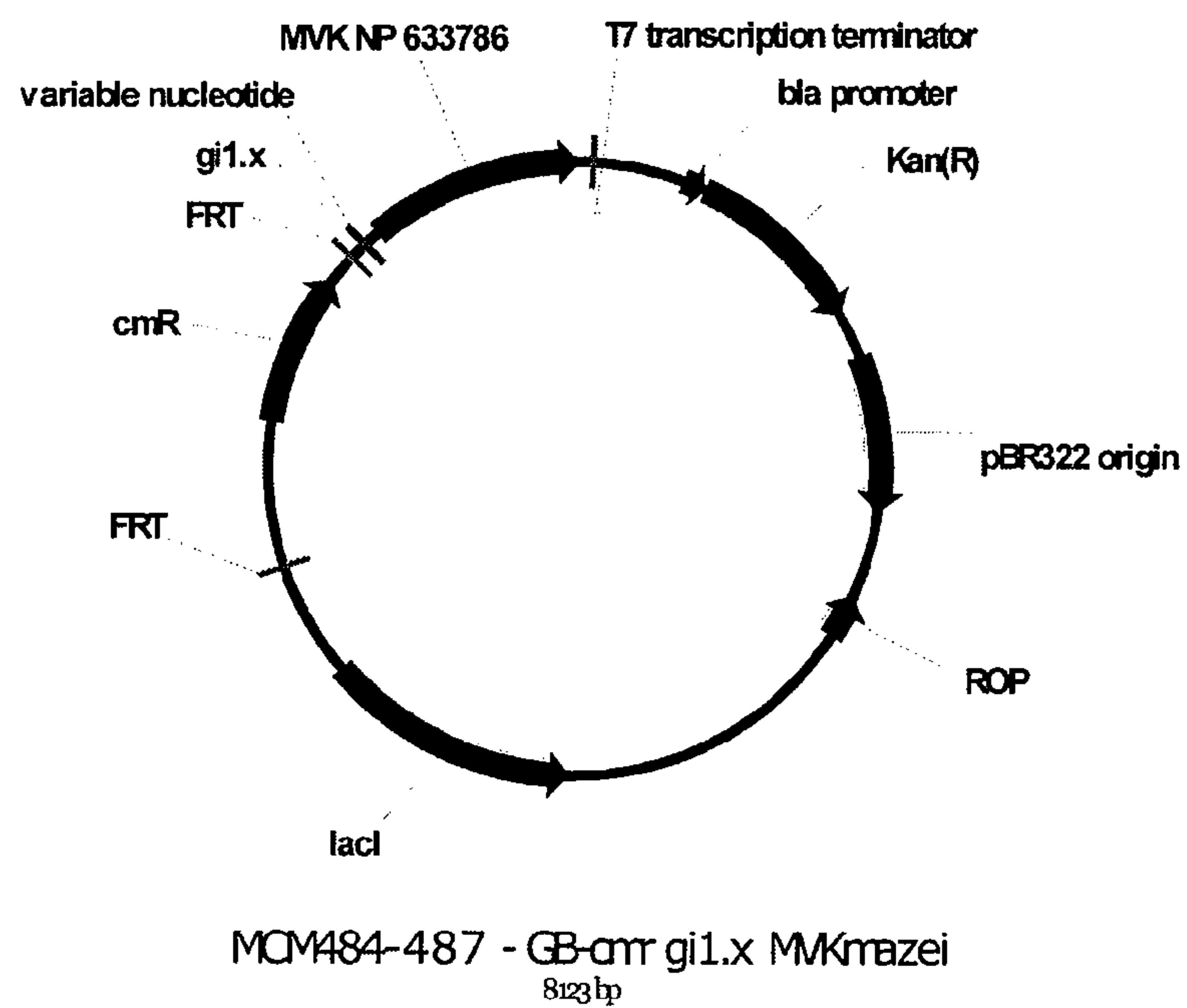
**Figure 61B**

atcggaggaccgaaggagctaaccgctttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaag  
ccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttacttagct  
tcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggccctccggctggctggttattgctgat  
aaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctccgctatcgtagtattctacacgac  
ggggagtcaggcaactatggatgaacgaaatagacagatcgtgagataggtgcctcactgattaagcattggtaactgtcagaccaagttt  
actcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatcctttttgataatctcatgacaaaaatcccttaacgt  
gagtttctgtccactgagcgtcagacccccgtagaaaagatcaaaggatcttcttgagatcctttttctgcgcgtaatctgctgcttgcaaca  
aaaaaaccaccgctaccagcgggtggttggttgcccggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagat  
accaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgctaactctgttac  
cagtggctgctgccagtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggctcgggctgaa  
cgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccac  
gcttcccgaaggagaaaaggcggacaggtatccggtaagcggcagggctcggaaacaggagagcgcacgagggagcttccagggggaa  
acgcctggtatctttatagtcctgtcgggttccgccacctctgacttgagcgtcgattttgtgatgctcgtcagggggcgaggcctatggaaa  
aacgccagcaacgcggcctttttacggttccctggccttttgctggccttttctcacatgttcttctcgttatccctgattctgttgataacc  
gtattaccgcctttgagtgagctgataccgctcgcgcagccgaacgaccgagcgcagcagtcagtgagcaggaagcgggaagagcg  
cctgatgcggtattttctccttacgcctctgtgcggtatttcacaccgcataatggtgcactctcagtacaatctgctctgatgccgcatagttaag  
ccagtatacactccgctatcgtacgtgactgggtcatggctgcgccccgacaccgcccaacaccgctgacgcgcctgacgggcttgt  
ctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtttcaccgtcatcaccgaaacgcgcga  
ggcagcagatcaattcgcgcgcgaaggcgaagcggcatgcatttacgttgacaccatcgatggtgcaaaacctttcgcggtatggcatga  
tagcggccggaagagagtcaattcagggtggtgaatgtgaaaccagtaacgttatacgtatgctgcagagtatgccggtgtctcttatcagac  
cgtttcccgcgtggtgaaccaggccagccacgtttctgcgaaaacgcgggaaaaagtggaaagcggcgatggcggagctgaattacattcc  
caaccgcgtggcacaacaactggcggggcaaacagtcgttgctgattggcgttgccacctccagtctggccctgcacgcgcctgcgcaatt  
gtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtggtggtgtcgatggtagaacgaagcggcgtcgaagcctgtaaagcg  
gcggtgcacaatcttctcgcgcaacgcgtcagtgggctgataactatccgctggatgaccaggatgccattgctgtggaagctgcctg  
cactaatgttccggcggtattttcttgatgtctctgaccagacacccatcaacagtattattttctcccatgaagacgggtacgcgactgggcgtgg  
agcatctggtcgcattgggtcaccagcaaatcgcgctgttagcggggccattaagtctgtctcggcgcgtctgcgtctggctggctggcata  
aatactcactcgcgaatcaaatcagccgatagcgggaacgggaaggcgactggagtgccatgtccgggtttcaacaaacctgcaaatgctg  
aatgagggcatcgttcccactgcgatgctggttgccaacgatcagatggcgctggggcgcaatgcgcgccattaccgagtcgggctgcgc  
gttggtgcggatctcggtagtgggatacagcagataccgaagacagctcatgttatatcccggcgtcaaccacatcaaacaggattttcgc  
ctgctggggcaaacaccagcgtggaccgcttgcgaactctctcagggccaggcgggtgaagggaatcagctgttggcgtctcactggtg  
aaaagaaaaaccacctggcgcccaatacgaacaccgcctctccccgcgcttggccgattcattaatgcagctggcacgacagggttccc  
gactggaaagcgggcagtgagcgaacgcaattaatgtgagttagcgcgaattgatctg (SEQ ID NO:89)



128/167

Figure 62





129/167

Figure 63A

1-

aaggcgagctcaacgatccggctgctaacaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataacc  
ccttggggcctctaaacgggtcttgaggagttttttgcgaaaggaggaaactatatccggatatcccgaagaggcccgagcaccggcat  
aaccaagcctatgcctacagcatccagggtgacgggtgccgaggatgacgatgagcgcattgttagatttcatacacggctgctgactgcgtt  
agcaatttaactgtgataaactaccgcattaaagcttatcgatgataagctgtcaaacatgagaatttaattcttgaagacgaaaggcctcgtg  
atacgccatttttatagggttaattgtcatgataataatggtttcttagacgtcagggtggcacttttcggggaaatgtgcgcggaacccctatttgtt  
attttctaaatacattcaaatatgtatccgctcatgagacaataaacctgataaatgcttcaataatattgaaaaaggaagagtatgattgaacaa  
gatggattgcacgcagggttccggccgcttgggtggagaggctattcggctatgactgggcacaactgacaatcggtgctctgatgccg  
ccgtgttccggctgtcagcgcagggggcggcgttcttttgcgaagaccgacctgtccgggtgccctgaatgaactgcaggacgaggcagc  
gcggctatcgtggctggccacgacgggcgttcccttgcgcagctgtgctcgacgttgcactgaagcgggaagggaactggctgctattggg  
cgaagtgccggggcaggatctcctgtcatctcacctgtcctgccgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgc  
ttgatccggctacctgcccattcgaccaccaagcgaacatcgcatcgagcgggcacgtactcggatggaagccggtcttgcgatcagga  
tgatctggacgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcgatgccgacggcgaggatctcgtc  
gtgacacatggcgatgcctgcttgcgaatatcatggtggaaaatggccgcttttctggattcatcgactgtggccggctgggtgtggcgga  
ccgctatcaggacatagcgttggctacccgtgatatgtcgaagagcttggcggcgaaatgggctgaccgcttccctcgtgctttacgggtatcgc  
cgctcccgaattcgcagcgcacgccttctatgccttcttgacgagtcttctgagcgggactctggggttcgaaatgaccgaccaagcgac  
gcctaactgtcagaccaagttactcatatatactttagattgatttaaaacttcattttaattaaaaggatctaggtgaagatccttttgataatc  
catgacaaaaatcccttaacgtgagtttctgtccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttctgcg  
cgtaatctgctgcttgcacacaaaaaaccaccgctaccagcgggtggttgcgggatcaagagctaccaactcttttccgaaggtaact  
ggcttcagcagagcgcagatacacaactgtccttctagtgtagccgtagttaggccaccactcaagaactctgtagcaccgcctacatac  
ctcgtctgctaactctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataag  
gcgcagcggctcgggtgaacggggggttctgtcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtg  
agctatgagaaagcgcacgcttcccgaaggagaaaggcggacaggtatccggtgaagcggcagggtcggaaacaggagagcgcacg  
agggagcttccagggggaaacgcctggtatctttatagtcctgtcgggttccgccacctctgacttgagcgtcgattttgtgatgctcgtcagg  
ggggcggagcctatggaaaaacgccagcaacgcggccttttaccggttcttggccttttgcctacatgttcttctcgtgctt  
cccctgattctgtggataaccgtattaccgcctttgagttagctgataccgctcggcgagccgaacgaccgagcgcagcagtcagttag  
cgaggaagcgggaagagcgcctgatgcgggtatttctccttacgcactctgtgcgggtatttcacaccgcaatgggtgcactctcagtacaatctgct  
ctgatccgcgatagtttaagccagtatacactccgctatcgtacgtgactgggtcatggctgcgccccgacacccgccaacaccgctgac  
gcgccctgacgggcttgtctgtcctccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtttaccgctc  
atcaccgaaacgcgcgagggcagctgcggtaaaagctcatcagcgtggtcgtgaagcgattcacagatgtctgcctgttcatccgcgtccagc  
tcgttgagtttctcagaagcgttaattgtctggttctgataaagcgggcatgtaaggggcgggttttctggttggtcactgatgcctccgtgt  
aagggggatttctgttcatgggggttaattgataccgatgaaacgagagaggatgtcacgatacgggttactgatgatgaacatgccgggtta  
ctggaacgttgtgagggttaacaactggcgggtatggatgcggcgggaccagagaaaaatcactcagggtcaatgccagcgttctgtaatt  
acagatgtaggtgttccacagggttagccagcagcatcctgcgatgcagatccggaacataatgggtgcagggcgtgacttccgcgtttcca  
gactttacgaaacacggaaaccgaagaccattcatgttgttgcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgctcgcgta  
tcggtgattcattctgctaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggcc  
aggaccaacgctgcccagatgcgccgcgtgcggctgctggagatggcggacgcgatggatatgttctgccaagggttggttgcgcat  
tcacagttctccgcaagaattgattggctccaattcttggagtggatccgtttagcgagggtgccgccggcttccattcaggctcaggtggc  
ccggctccatgcaccgcgacgcaacgcggggaggcagacaaggatataggcgggcgcctacaatccatgccaacccgttccatgtgctc  
gccgagggcgataaatcccggtgacgatcagcgggtccaatgatcgaagttaggttggttaagagccgcgagcagatccttgaagctgtccc  
tgatggctgctcatctacctgctggacagcatggcctgcaacgcgggcaccccgatgccggcggaagcgagaagaatcataatggggaa  
ggccatccagcctcgcgtcgcgaacgccagcaagacgtagcccagcgcgtcggccgccatgccggcgataatggcctgcttctcgcg  
aaacgttgggtggcgggaccagtgacgaaggctgagcagggcgtgcaagattccgaataccgcaagcgacaggccgatcatcgtcgc  
gctccagcgaaagcggctcctcgcggaatgaccagagcgtgcccggcacctgtcctacgagttgcatgataaagaagacagtcataag



130/167

**Figure 63B**

tgcggcgacgatagtcatgccccgcgccaccggaaggagctgactgggttgaaggctctcaagggcacatcggtcgagatccccggtgcct  
aatgagtgagctaacttacatlaattgcgttgcgctcactgcccgccttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggcc  
acgcgcggggagagggcggttgcgtattgggcgccagggtggtttttccttaccagtgagacgggcaacagctgattgcccttcaccgcc  
tggccctgagagagttgcagcaagcgggtccacgctggtttgccccagcaggcgaaaatcctgtttgatggttggttaacggcgggatataac  
atgagctgtcttcggtatcgtcgtatcccactaccgagatatccgcaccaacgcgcagccccggactcggtaatggcgcgcatcgcgccag  
cgccatctgatcgttggcaaccagcatcgagtggaacgatgcctcattcagcatttgcattggtttgtgaaaaccggacatggcactcca  
gtcgccttcccgttccgctatcggctgaatttgattgcgagtgagatatttatgccagccagccagacgcagacgcgccgagacagaactta  
atggggcccgtaacagcgcgatttgcgttggtgaccaatgcgaccagatgtccacgcccagtcgcgtaccgtcttcatgggagaaaaataat  
actgttgatgggtgtctggtcagagacatcaagaataacgccggaacattagtgcaggcagcttccacagcaatggcatcctggtcatcca  
gcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgccgctttacaggcttcgacgccgcttcgttctaccatc  
gacaccaccacgctggcaccacgttgatcggcgcgagatttaatgcgcgcgacaatttgcgacggcgcggtgcagggccagactggaggt  
ggcaacgccaatcagcaacgactgtttgcccgccagttgttgtgccacgcggttgggaatgtaattcagctccgccatcgccgcttccacttt  
tcccgcgttttcgcagaaacgtggctggcctggttcaccacgcgggaaacggctcgataagagacaccggcatactctgcgacatcgtata  
acgttactggtttcacattcaccacctgaattgactctcttccgggcgctatcatgccataccgcgaaaggttttgcgccattcgtatggtgtcc  
gggatctcgacgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgccgccgcaaggaat  
ggtgcatgcaaggagatggcgcccaacagtcccccgccacggggcctgccaccataccacgccgaaacaagcgctcatgagccccg  
aagtggcgagcccgatcttccccatcggtagtgcggcgatataggcgccagcaaccgcacctgtggcgccggtgatgccggccacgat  
gcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgaactcactataggggaattgtgagcggataacaattccccctctaga  
aataatttgtttaactttaagaaggagatatacatatgaattaacctcactaaagggcgggcggaagtccctattctctagaaagtataggaa  
cttcattctaccgggtaggggaggcgcttttcccaaggcagctcggagcatgcgctttagcagccccgctgggcacttggcgctacacaagt  
ggcctctggcctcgcacacattccacatccaccggtaggcgccaaccggctccgttcttgggtggcccccttcgcgccaccttccactcctcc  
cctagtcaggaagttcccccccgccccgcagctcgcgtcgtgcaggacgtgacaaatggaagtagcacgtctcactagtctcgtgcagatg  
gacagcaccgctgagcaatggaagcgggtaggcccttgggggcagcggccaatagcagcttctccttcgcttcttgggctcagaggctg  
ggaagggggtgggtccggggggcggggtcagggggcggggtcagggggcggggcggggcgccgaaggctcctccggaggccccggcattct  
gcacgcttcaaaagcgcacgtctgccgcgtgttctcctcttctcatctccgggccttccgacctgcagcagcacgtgttgacaattaatcat  
cggcatagtatatcggcatagtataatacgaaggtgaggaactaaacctatggagaaaaaaatcactggatataccaccgttgatatatcc  
caatggcatcgtaaagaacattttgaggcattcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggcctttttaag  
accgtaaagaaaaataagcacaagttttatccggcctttattcacattcttggccgcctgatgaatgtcatccggaattccgtatggcaatgaa  
agacgggtgagctggtgatatgggatagtgttcacccttgttacaccgctttccatgagcaaaactgaaacgttttcacgtctctggagtgaatacc  
acgacgatttccggcagtttctacacatatattcgaagatgtggcgtgttacgggtgaaaacctggcctatttccctaaggggttattgagaat  
atgttttctgtctcagccaatccctgggtgagtttcaccagttttgatttaaacgtggccaatatggacaacttcttcgccccggttttcaccatgg  
gcaaatattatacgaaggcgacaaggtgctgatgccgctggcgattcaggttcacatgcgcttgggtgatggcttccatgtcggcagaatgc  
ttaatgaattacaacagtagtgcgatgagtggcagggcggggcgtaagcgggactctgggggtcgaataaagaccgaccaagcgacgtct  
gagagctccctggcgaattcggtagcaataaaagagctttattttcatgatctgtgtgttgggttttgtgtgcggcgcggaagttcctattctctag  
aaagtataggaaacttctcagaccctatagttagtcgtattagatcgcggccgcgcccttgacaatgccacatcctgagcaataattcaacc  
actaattgtgagcggataacaaaggaggtaaaaaaacatggtatcctgttctgcgccgggtaagatttacctgttcggtgaacacgccgtagt  
ttatggcgaaactgcaattgcgtgtgcgggtggaactgcgtacccgtgttcgcgcggaactcaatgactctatcactattcagagccagatcgg  
ccgcaccggctctggatttcgaaaagcacccttatgtgtctgcggtaattgagaaaaatgcgcaaatctattcctattaacgggtgttttcttgaccgt  
cgattccgacatcccgggtgggctccggctgtgggtagcagcgcagccgttactatcgcgtctattggtgcgctgaacgagctgttcggccttg  
gcctcagcctgcaagaaatcgctaaactgggccacgaaatcgaaattaaagtacaggggtgccgcgtccccaaccgatacgtatgtttctacc  
ttcggcgggcgtggttaccatcccggaaactgcgcaaaactgaaaactccggactgcggcattgtgattggcgataaccggcggttttctcctccac  
caaagagttagtagctaacgtacgtcagctgcgcgaaagctacccggatttgatgaaccgctgatgaccttattggcaaaatctctcgtat  
cggcgaaacaactggttctgtctggcgactacgcattccatcgccgcctgatgaacgtcaaccagggtctcctggacgccttgggcgttaac  
atcttagaactgagccagctgatctattccgctcgtgcggcaggtgcgtttggcgctaaaaatcacggggcgctggcgggcggtggcgtgatggt



131/167

**Figure 63C**

tgcgctgaccgctccggaaaaatgcaaccaagtggcagaagcggtagcaggcgctggcggtaaagtgactatcactaaaccgaccgag  
caaggtctgaaagtagattaa (SEQ ID NO:90)



132/167

Figure 64A

1-

aaggcgagctcaacgatccggctgctaacaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataacc  
ccttggggcctctaaacgggtcttgaggagtttttgctgaaaggagggaactatatccggatatcccgcaagaggccggcagtagccgcat  
aaccaagcctatgcctacagcatccagggtgacgggtgccgaggatgacgatgagcgcatgttagatttcatacacgggtgcctgactgcgtt  
agcaatttaactgtgataaactaccgcattaaagcttatcgatgataagctgtcaaacatgagaatttaattctgaagacgaaagggcctcgtg  
atacgcctattttataggtaatgtcatgataataatggttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtt  
attttctaaatacattcaaatatgtatccgctcatgagacaataacccgtataaatgcttcaataatattgaaaaggaagagtatgattgaacaa  
gatggattgcacgcaggttctccggccgcttgggtggagaggctattcggctatgactgggcacaactgacaatcggtgctctgatgccg  
ccgtgttccggctgtcagcgcagggggcgcccggttctttgtcaagaccgacctgtccgggtgccctgaatgaactgcaggacgaggcagc  
gcggctatcgtggctggccacgacgggcgttcttgcgcagctgtgctcgacgttgcactgaagcgggaagggaactggctgctattggg  
cgaagtgccggggcaggatctcctgtcatctcacctgtcctgccgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgc  
ttgatccggctacctgcccattcgaccaccaagcgaaacatcgcatcgagcgggcacgtactcggatggaagccggctctgtcgatcagga  
tgatctggacgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcgcatgcccagggcgaggatctcgtc  
gtgacacatggcgatgcctgcttgcgaatatcatggtggaaaatggccgctttctggattcatcgactgtggccggctgggtgtggcgga  
ccgctatcaggacatagcgttggctacccgtgatattgctgaagagcttggcggcgaaatgggctgaccgcttcctcgtgctttacgggtatcgc  
cgctcccgatctgcagcgcacgccttctatgccttcttgacgagttcttctgagcgggactctggggttcgaaatgaccgaccaagcgac  
gcctaactgtcagaccaagttactcatatatactttagattgatttaaaacttcattttaatttaaaaggatctaggtgaagatccttttgataatct  
catgacaaaaatcccttaacgtgagtttctgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttctgcg  
cgtaatctgctgcttgcacacaaaaaaccaccgctaccagcgggtggttgggttgcgggatcaagagctaccaactcttttccgaaggtaact  
ggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccactcaagaactctgtagcaccgcctacatac  
ctcgtctgtctaactcgttaccagtggtgctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataag  
gcgcagcggctcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtg  
agctatgagaaagcgccacgcttcccgaaggggagaaaggcggacaggtatccggtaagcggcagggctcggaacaggagagcgcacg  
agggagcttcagggggaaacgcctggtatctttatagtcctgtcgggttccgccacctctgacttgagcgtcgattttgtgatgctcgtcagg  
ggggcggagcctatggaaaaacgccagcaacgcggcctttttacgggtccttggccttttgcctcactgttcttctcgtgctt  
ccccgtattctgtggataaccgtattaccgctttgagttagctgataccgctcggcgagccgaacgaccgagcgcagcagtcagtgag  
cgaggaagcgggaagagcgcctgatgcgggtattttctccttacgcacatctgtcgggtatttcacaccgcaatgggtgcactctcagtacaatctgct  
ctgatgccgcatagtttaagccagttatacactccgctatcgtacgtgactgggtcatggctgcgccccgacaccgccaacacccgctgac  
gcgcccgtacgggcttctgtcctccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagagggtttaccgctc  
atcaccgaaacgcgcgagggcagctgcggtaaaagctcatcagcgtggtcgtgaagcgattcacagatgtctgctgttcatccgcgtccagc  
tcgttgagtttctccagaagcgttaatgtctggttctgataaagcgggccatgttaaggggcgtttttctggttgggtcactgatgcctccgtgt  
aagggggatttctgttcatgggggtaatgataccgatgaaacgagagaggatgtcacgatacgggttactgatgatgaacatgcccggtta  
ctggaacgttgtgagggtaaacaactggcgggtatggatgcggcgggaccagagaaaaatcactcagggtcaatgccagcgttcgttaat  
acagatgtaggtgttccacagggttagccagcagcatcctgcgatgcagatccggaacataatgggtgcagggcgctgacttcgcgtttcca  
gactttacgaaacacggaaaccgaagaccattcatgttgttgcaggtcgcagacgttttcagcagcagtcgcttcacgttcgctcgcgta  
tcgggtgattcattctgctaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggcc  
aggaccaacgctgcccagatgcgccgcgtgcggctgctggagatggcggacgcgatggatatgttctgccaagggttggttgcgcat  
tcacagttctccgcaagaattgattggctccaattcttggagtggatccgttagcgaggtgccgcccggcttccattcaggtcgaggtggc  
ccggctccatgcaccgcgacgcaacgcggggaggcagacaaggtatagggcgggcctacaatccatgccaaaccgttccatgtgctc  
gccgaggcggcataaatcgccgtgacgatcagcgggtccaatgatcgaagttaggttggttaagagccgcgagcgatccttgaagctgtccc  
tgatggctgctcatctacctgcctggacagcatggcctgcaacgcgggcaccccgatgccggcggaagcgagaagaatcataatggggaa  
ggccatccagcctcgcgtcgcgaacgccagcaagacgtagcccagcgcgtcggccgccatgccggcgataatggcctgcttctcgcgg  
aaacgttgggtggcgggaccagtgacgaaggcttgagcgagggcgtgcaagattccgaataccgcaagcgacaggccgatcatcgtcgc  
gctccagcgaaagcggctcctcgcggaaaatgaccagagcgtgcgggcacctgtcctacgagttgcatgataaagaagacagtcataag



133/167

**Figure 64B**

tgcggcgacgatagtcatgccccgcgcccaccggaaggagctgactgggttgaaggctctcaaggggcatcggtcgagatcccgggtgcct  
aatgagtgagctaacttacattaattgcgttgcgctcactgcccgtttccagtcgggaaacctgtcgtgccagctgcattaatgaatcgcca  
acgcgcggggagaggcggttgcgtattggcgccagggtggttttttccaccagtgcagacgggcaacagctgattgcccttcaccgcc  
tggccctgagagagttgcagcaagcggtccacgctggtttgccccagcaggcgaaaatcctgtttgatggttggttaacggcgggatataac  
atgagctgtcttcggtatcgtcgtatcccactaccgagatatccgcaccaacgcgcagcccggactcggtaatggcgcgcatcgccccag  
cgccatctgatcggttgcaaccagcatcgagtggaacgatgccctcattcagcatttgcattggtttgttgaaccgggacatggcactcca  
gtcgccctcccgttcgctatcggtgaatttgattgcgagtgagatatccagccagccagacgcagacgcgcggagacagaactta  
atgggcccgtaacagcgcgatttgctggtgaccaatgcgaccagatgtccacgcccagtcgcgtaccgttctatgggagaaaataat  
actgttgatgggtgtctggtcagagacatcaagaaataacgccggaacattagtgcaggcagcttccacagcaatggcactcctggtcatcca  
gcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgcccgtttacaggcttcgacgcgcgttcgttctaccatc  
gacaccaccacgctggcaccagttgatcggcgcgagatttaacgccgcgacaatttgcgacggcgcggtgcagggccagactggaggt  
ggcaacgccaatcagcaacgactgtttgcccggcagttgttgccacgcggttgggaatgtaattcagctccgccatcgccgttccacttt  
tcccgcgttttcgagaaacgtggctggcctgggtcaccacgcgggaaacgggtctgataagagacaccggcactctgcgacatcgata  
acgttactgggttcacattcaccacctgaattgactctcttcggggcgctatcatgccataccgcgaaagggtttgcgccattcgatggtgtcc  
gggatctcgacgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgcccgcgcaaggaat  
ggtgcatgcaaggagatggcgcccaacagtcccccggccacggggcctgccaccataccacgcgcgaaacaagcgctcatgagccg  
aagtggcgagcccgatcttccccatcggtgatgtcggcgatataggcgccagcaaccgcacctgtggcgccggtgatgccggccacgat  
gcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgaactactatagggggaattgtgagcggataacaattcccctctaga  
aataattttgttaactttaagaaggagatatacatatgaattaaccctcactaaaggggcgccgcgaagttcctattctctagaaagtataggaa  
cttcattctaccgggtaggggaggcgcttttcccaaggcagctctggagcatgcgctttagcagccccgctgggcacttggcgctacacaagt  
ggcctctggcctcgacacattccacatccaccggtaggcgccaaccggctccgttcttgggtggcccccttcgcgccaccttccactcctcc  
cctagtcaggaagttcccccccgccccgcagctcgcgctgctgcaggacgtgacaaatggaagtagcacgtctcactagtctcgtgcagatg  
gacagcaccgctgagcaatggaagcgggtaggcctttggggcagcggccaatagcagcttgccttcgcttctggttcagaggctg  
ggaaggggtgggtccggggggcgggctcagggggcgggctcagggggcgggcgggcgcccgaaggtcctccggaggccccgcattct  
gcacgcttcaaaagcgacgtctgcgcgctgttctcctctcctcatctccgggccttcgacctgcagcagcacgtgttgacaattaatcat  
cggcatagtatatcggcatagtataatacgaaggtgaggaactaaaccatggagaaaaaaatcactggatataccaccgttgatatatcc  
caatggcatcgtaaagaacattttgaggcatttcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggcctttttaaag  
accgtaaagaaaaataagcacaagttttatccggcctttattcacattcttggccgctgatgaatgctcatccggaattccgtatggcaatgaa  
agacgggtgagctggtgatatgggatatgttcaccctgttacaccggtttccatgagcaaaactgaaacgttttcctcgtctggtgagtaatacc  
acgacgatttccggcagtttctacacatatattcgcaagatgtggcggttacggtgaaaacctggcctatttccctaaagggtttattgagaat  
atgttttctcgtctcagccaatccctgggtgagtttaccagtttgatttaaactggccaatatggacaacttcttcgccccggtttaccatgg  
gcaaatattatacgaaggcgacaaggtgctgatgccgctggcgattcaggttcacatgccgtttgtgatggcttccatgtcggcagaatgc  
ttaatgaattacaacagtactgcgatgagtggcagggcgggcgtaagcgggactctgggggtcgaataaagaccgaccaagcgacgtct  
gagagctccctggcgaattcggtaccaataaaagagctttattttcatgatctgtgtgttgggtttgtgtgcggcgcggaagttcctattctctag  
aaagtataggaacttctcgagccctatagttagtcgtattagatcgcgccgcgccttgacctgccacatcctgagcaataattcaacc  
actaattgtgagcggataacaaaggaggtaaaaaacatggtatcctgttctgcgcgggtaagatttacctgttcggtgaacacgccgtagt  
ttatggcgaaactgcaattgcgtgtgcggtggaactgcgtacccgtgttcgcgcggaactcaatgactctatcactattcagagccagatcg  
ccgcaccggctctggatttcgaaaagcacccttatgtgtctgcggttaattgagaaaatgcgcaaatctattcctattaacgggtgtttcttgaccgt  
cgattccgacatcccgggtgggtcctcggtctgggttagcagcgcagccgttactatcgctctattggtgcgctgaacgagctgttcggctttg  
gcctcagcctgcaagaaatcgctaaactgggccacgaaatcgaaattaaagtacagggtgcgcgtccccaaccgatacgtatgtttctacc  
ttcggcgggcgtggttaccatcccggaacgtcgcaaaactgaaaactccggactgcggcattgtgattggcgataccggcggttttctcctccac  
caaagagtagtagtaacgtacgtcagctgcgcgaaagctacccggatttgatcgaaccgctgatgacctctattggcaaaatctctcgtat  
cggcgaacaactggttctgtctggcgactacgcacatccatcgccgcctgatgaacgtcaaccagggtctcctggacgccctgggcgttaac  
atcttagaactgagccagctgatctattccgctcgtgcggcaggtgcgtttggcgctaaaatcacgggcgctggcgggcggtggtgtatggt



134/167

**Figure 64C**

tgcgctgaccgctccggaaaaatgcaaccaagtggcagaagcggtagcaggcgctggcggtaaagtgactatcactaaaccgaccgag  
caaggtctgaaagtagattaa (SEQ ID NO:91)



135/167

Figure 65A

1-

aagggcgagctcaacgatccggctgctaacaagccccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataacc  
ccttggggcctctaaacgggtcttgaggagtttttgctgaaaggaggaactatatccggatatcccgaagaggcccgagctaccggcat  
aaccaagcctatgcctacagcatccagggtgacgggtgccgaggatgacgatgagcgcattgttagatttcatacacgggtgcctgactgcgtt  
agcaatttaactgtgataaactaccgcattaaagcttatcgatgataagctgtcaaacatgagaattaattcttgaagacgaaagggcctcgtg  
atacgctatttttatagggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtt  
atttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgattgaacaa  
gatggattgcacgcaggttctccggcgcgttgggtggagaggctattcggctatgactgggcacaactgacaatcggctgctctgatgccg  
ccgtgttccggctgtcagcgcagggggcggcggttcttttgcagaccgacctgtccgggtgccctgaatgaactgcaggacgaggcagc  
gcggctatcgtggctggccacgacgggcgttccctgcgcagctgtgctcgacgtgtcactgaagcgggaagggactggctgctattggg  
cgaagtgcgggggcaggatctcctgtcatctcacctgtcctgccgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgc  
ttgatccggctacctgcccattcgaccaccaagcgaacatcgcacgagcgggcacgtactcggatggaagccggctcttgcgatcagga  
tgatctggacgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcgatgcccgacggcgaggatctcgtc  
gtgacacatggcgatgcctgcttgcgaatatcatggtggaaaaatggcgcttttctggattcatcactgtggccggctgggtgtggcgga  
ccgctatcaggacatagcgttggctacccgtgatattgtgaagagcttggcggcgaatgggctgaccgcttctcgtgctttacgggtatcgc  
cgctcccgatcgcagcgcacgccttctatgccttcttgacgagttcttctgagcgggactctgggggttcgaaatgaccgaccaagcgac  
gcctaactgtcagaccaagttaactcatatatacttttagattgatttaaaacttcattttaatttaaaaggatctaggtgaagatccttttgataatc  
catgaccaaatacccttaacgtgagtttctgctccactgagcgtcagaccccgtagaaaagatcaaaaggatcttcttgagatcctttttctgcg  
cgtaatctgctgcttgcacacaaaaaaccaccgctaccagcgggtggttgggttccggatcaagagctaccaactcttttccgaaggtaact  
ggcttcagcagagcgcagatacacaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatac  
ctcgtctgctaactcgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataag  
gcgcagcggctcgggctgaacgggggggtcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtg  
agctatgagaaagcgccacgcttcccgaaggagaaaggcggacaggtatccggtgaagcggcagggctcggaaacaggagagcgcacg  
agggagcttccagggggaaacgcctgggtatctttatagtcctgtcgggttcgccacctctgacttgagcgtcgattttgtgatgctcgtcagg  
ggggcgagcctatggaaaaacgccagcaacgcggcctttttacgggtccttggccttttgcctacatgttcttctcgtgctt  
cccctgattctgttgataaccgtattaccgcctttgagttagctgataccgctcggcgagccgaacgaccgagcgcagcgagtcagttag  
cgaggaagcgggaagagcgcctgatgcgggtatttctccttacgcatctgtgcgggtatttcacaccgcaatgggtgcactctcagtacaatctgct  
ctgatccgcatagttaagccagtatacactccgctatcgtactgactgggtcatggctgcgccccgacacccgccaacacccgctgac  
gcgcctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtttaccgctc  
atcaccgaaacgcgcgagggcagctgcggtaaagctcatcagcgtggctgtgaagcgattcacagatgtctgcctgttcatccgcgtccagc  
tcgttgagtttctccagaagcgttaatgtctggcttctgataaagcggggccatgtaaggggcggttttctcgttgggtcactgatgcctccgtgt  
aaggggggatttctgttcatgggggttaatgataccgatgaaacgagagaggatgtcacgatacgggttactgatgatgaacatgcccggtta  
ctggaacgttgtgagggtaaacaaactggcggtatggatgcggcgggaccagagaaaaatcactcaggggtcaatgccagcgttctgtaat  
acagatgtaggtgttccacagggtagccagcagcatcctgcgatgcagatccggaacataatgggtgcagggcgctgacttccgcgtttcca  
gactttacgaaacacggaaaccgaagaccattcatgttgttgcctcaggtcgcagacgtttgcagcagcagtcgcttcacgttcgctcgcgta  
tcgggtgattcattctgtaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggcc  
aggaccaacgctgcccagatgcggcgctgcgggtgctggagatggcgagcgcgatggatatgttctgccaagggttgggttgcgcgt  
tcacagttctccgcaagaattgattggctccaattcttggagtgggtgaatccgttagcgagggtgccggcgttccattcaggtcgaggtggc  
ccggctccatgcaccgcgacgcaacgcggggagggcagacaagggtataggggcgggcgcctacaatccatgccaacccgttccatgtgctc  
gccgaggcgggcataaatcgccgtgacgatcagcgggtccaatgatcgaagttaggttggttaagagccgcgagcgcgatccttgaagctgtccc  
tgatggctgctcatctacctgcctggacagcatggcctgcaacgcgggcatcccgatgccggcggaagcgagaagaatcataatggggaa  
ggccatccagcctcgcgtcgcgaacgccagcaagacgtagcccagcgcgtcggccgcatgccggcgataatggcctgcttctcgcgg  
aaacgtttgggtggcgggaccagtgacgaaggcttgagcgagggcggtgcaagattccgaataaccgcaagcgacaggccgatcatcgtcgc  
gctccagcgaaagcgggtcctcgcggaaaatgaccagagcgtgcgggcacctgtcctacgagttgcatgataaagaagacagtcataag



**Figure 65B**

tgcggcgacgatatgcatgccccgcgcccaccggaaggagctgactgggttgaaggctctcaagggcacggtcgagatcccgggtgcct  
aatgagtgagctaacttacattaattgcgttgcgctcactgcccgtttccagtcgggaaacctgtcgtgccagctgcattaatgaatcgcca  
acgcgcggggagaggcggttgcgtattggcgccagggtggtttttttaccagtgcagacgggcaacagctgattgcccttcaccgcc  
tgccctgagagagttgcagcaagcgggtccacgctggttgcgccagcaggcgaaaacctgttgatggtggttaacggcgggatataac  
atgagctgtcttcggtatcgtcgtatcccactaccgagatatccgcaccaacgcgcagcccggactcggtaatggcgcgcatcgcgccag  
cgccatcgtatcgttggcaaccagcagcgcagtggtgggaacgatgccctcattcagcatttgcattggttggtaaaccgggacatggcactcca  
gtcgccttcccgttccgctatcggtgaatttgattgcgagtgcagatatccagccagccagacgcagacgcgcggagacagaactta  
atgggcccgcataacagcgcgatttgctggtgaccaatgcgaccagatgtccacgcccagtcgcgtaccgtctcatgggagaaaataat  
actgttgatgggtgtctggtcagagacatcaagaaataacgccggaacattagtgcaggcagcttccacagcaatggcatcctggtcatcca  
gcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgcccgtttacaggcttcgacgcgcgttcgttctaccatc  
gacaccaccacgctggcaccagtgatcggcgcgagatttaacgccgcgacaatttgcgacggcgcggtgcaggggccagactggaggt  
ggcaacgccaatcagcaacgactgttgcgccagttgtgtgccacgcggttgggaatgtaattcagctccgccatcgccgttccacttt  
tcccgcgttttcgcagaaacgtggctggcctggttcaccacgcgggaaacgggtctgataagagacaccggcatactctgcgacatcgtata  
acgttactgggttcacattcaccacctgaattgactcttccgggcgctatcatgccataaccgcgaaagggtttgcgccattcgtatggtgtcc  
gggatctcgacgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgcccgcgcaaggat  
ggtgcatgcaaggagatggcgcccaacagtcccccgccacggggcctgccaccataccacgcggaacaagcgctcatgagccg  
aagtggcgagcccgatcttcccacggtgatgtcgcgatataggcgccagcaaccgcacctgtggcgccggtgatgccggccacgat  
gcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgaactcactataggggaattgtgagcgggataacaattcccctctaga  
aataattttgttaacttaagaaggagatatacatatgaattaaccctcactaaaggcgccgcgaagttcctattctctagaaagtataggaa  
cttcattctaccgggtaggggaggcgcttttcccaaggcagctctggagcatgcgttttagcagccccgctgggcaacttggcgctacacaagt  
ggcctctggcctcgacacattccacatccaccggtaggcgccaaccggctccgttcttgggtggcccccttcgcgccaccttccactcctcc  
cctagtcaggaagttccccccgcgccgagctcgcgctcgtgcaggacgtgacaaatggaagtagcacgtctcactagtctcgtgcagatg  
gacagcaccgctgagcaatggaagcgggtaggcctttggggcagcggccaatagcagcttgccttcgcttctgggctcagaggctg  
ggaaggggtgggtccgggggcggggtcagggggcggggtcagggggcggggcggcgcccgaaggtcctccggaggcccggcattct  
gcacgctcaaaagcgcacgtctgccgcgtgttctcctctcctcatctccgggcttccgacctgcagcagcacgtgttgacaattaatcat  
cggcatagtatacggcatagtataatacgaaggtgaggaactaaaccatggagaaaaaaatcactggatataccaccgttgatatacc  
caatggcatcgtaaagaacatttgaggcattcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggccttttaag  
accgtaaagaaaaataagcacaagtttatccggccttattcacattcttgcgcctgatgaatgctcatccggaattccgtatggcaatgaa  
agacgggtgagctggtgatatgggatatgttcaccctgttacaccgtttccatgagcaaaactgaaacgtttcatcgctctggagtgaaatcc  
acgacgatttccggcagtttctacacatatattcgcaagatgtggcggttacgggtgaaaacctggcctatttccctaaagggttattgagaat  
atgttttctgctcagccaatccctgggtgagtttaccagtttgattaaacgtggccaatatggacaacttcttcgccccggtttaccatgg  
gcaaatattatacgaaggcgacaaggtgctgatgccgctggcgattcaggttcattcatgccgttgtgatggcttccatgtcggcagaatgc  
ttaatgaattacaacagtaactgcgatgagtgagggcgggggcgtaagcgggactctgggggtcgaataaagaccgaccaagcgacgtct  
gagagctccctggcgaattcgggtaccaataaaagagcttattttcatgatctgtgtgttgggtttgtgtgcggcgcggaagttcctattcttag  
aaagtataggaacttctcgagccctatagtgcgtgattagatcgcgccgcgccccttgacgatgccacatcctgagcaataattcaacc  
actaattgtgagcgggataacaaaggaggtaaaaaacatggtatcctgttctgcgccgggtaagattacgttgcgtgaacacgccgtagt  
ttatggcgaaactgcaattgcgtgtgcggtggaactgcgtaccctgttgcgcgggaactcaatgactctatcactattcagagccagatcg  
ccgcaccgggtctggatttcgaaaagcacccttatgtgtcgcggttaattgagaaaatgcgcaaatctattcctattaacgggtgtttcttgaccgt  
cgattccgacatcccgggtgggtcctcggtctgggttagcagcgcagccgttactatcgctctattggtgcgctgaacgagctgttcggctttg  
gcctcagcctgcaagaaatcgtaaaactgggccacgaaatcgaaattaaagtacagggtgccgcgtccccaaccgatacgtatgtttctacc  
ttcggcgggcgtggttaccatcccggaaactgcgcaaaactgaaaactccggactgcggcattgtgattggcgataccggcggttttctcctccac  
caaagagttagtagctaactgtacgtcagctgcgcgaaagctaccgggattgatcgaaccgctgatgacctctattggcaaaatctctcgtat  
cggcgaaacaaactggttctgtctggcgactacgcattccatcgccgcctgatgaacgtcaaccagggtctcctggacgccctgggcgttaac  
atcttagaactgagccagctgatctattccgctcgtgcggcaggtgcgttggcgctaaaatcacgggcgctggcgggcggtggctgtatggt



137/167

**Figure 65C**

tgcgctgaccgctccggaaaaatgcaaccaagtggcagaagcggtagcaggcgctggcggtaaagtgactatcactaaaccgaccgag  
caaggctgaaagtagattaa (SEQ ID NO:92)



**Figure 66A**

1-

aagggcgagctcaacgatccggctgctaacaaagcccgaaggaagctgagttggctgctgccaccgctgagcaataactagcataacc  
ccttggggcctctaaacgggtcttgaggagtttttctgtaaaggaggaactatatccggatatcccgaagaggcccggcagtaccggcat  
aaccaagcctatgcctacagcatccagggtgacggtgccgaggatgacgatgagcgcattgttagatttcatacacggtgacctgactgcgt  
agcaatttaactgtgataaactaccgcattaaagcttatcgatgataagctgtcaaacatgagaatttaattcttgaagacgaaagggcctcgtg  
atacgctattttataggtaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtt  
attttctaaatacattcaaataatgtatccgctcatgagacaataaccctgataaatgtctcaataatattgaaaaaggaagagtatgattgaacaa  
gatggattgcacgcaggttctccggccgcttgggtggagaggctattcggctatgactgggcacaactgacaatcggctgctctgatgccg  
ccgtgttccggctgtcagcgcaggggcccgggtcttttctcaagaccgacctgtccgggtgccctgaatgaactgcaggacgaggcagc  
gcggctatcgtggctggccacgacgggcgttcccttgcgcagctgtgctcgacgttgcactgaagcgggaagggaactggctgctattggg  
cgaagtgcgggggcaggatctcctgtcatctcaccttgcctcctgccgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgc  
ttgatccggctacctgcccattcgaccaccaagcgaaacatcgcatcgagcgggcacgtactcggatggaagccgggtcttctcgatcagga  
tgatctggacgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcgatcccgcagggcgaggatctcgtc  
gtgacacatggcgatgcctgcttgcgaatatcatgggtggaaaatggccgcttttctggattcatcgactgtggccggctgggtgtggcgga  
ccgctatcaggacatagcgttggctacccgtgatattgtgaagagcttggcggcgaatgggctgaccgcttctcgtgctttacgggtatgc  
cgctcccgaattcgacgcgcacgccttctatcgcttcttgacgagttcttctgagcgggactctggggttcgaaatgaccgaccaagcgac  
gcctaactgtcagaccaagtttactcatatatacttttagattgatttaaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatct  
catgacaaaaatccettaacgtgagtttctgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcg  
cgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcgggtggttgttgcgggatcaagagctaccaactcttttccgaaggtaact  
ggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatac  
ctcgtctgctaactcgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataag  
gcgcagcggctcgggctgaacgggggggtcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtg  
agctatgagaaagcgcacgcttcccgaaggggagaaaggcggacaggtatccggtaagcggcagggctcgggaacaggagagcgcacg  
agggagcttccagggggaaacgcctggatctttatagtcctgtcgggttccgccacctctgacttgagcgtcgattttgtgatgctcgtcagg  
ggggcggagcctatgaaaaacgccagcaacgcggcctttttacgggttcccttggccttttgccttctgctcacatgttcttctcgtt  
ccctgattctgttgataaccgtattaccgcctttgagttagctgataccgctcggcgagccgaacgaccgagcgcagcgagtcagttag  
cgaggaagcgggaagagcgcctgatgcggtattttctcttacgcactgtgcegggtatttcacaccgcaatggtgcactctcagtacaatctgct  
ctgatccgcgatagttaagccagtatacactccgctatcgctacgtgactgggtcatggctgcgccccgacaccgcgaacacccgctgac  
gcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagagggtttaccgct  
atcaccgaaacgcgcgaggcagctgcggtaaagctcatcagcgtggctcgtgaagcgattcacagatgtctgctgttcatccgcgtccagc  
tcgttgagtttctccagaagcgttaatgtctggcttctgataaagcgggccatgttaaggggcggtttttcctgttgggtcactgatgcctccgtgt  
aaggggggatttctgttcatgggggtaatgataccgatgaacgagagaggatgtcacgatacgggttactgatgatgaacatgcccggtta  
ctggaacgttgtgagggtaaacaactggcgggtatggatgcggcgggaccagagaaaaatcactcaggggtcaatgccagcgttctgta  
acagatgtaggtgttccacagggttagccagcagcatcctgcgatgcagatccggaacataatggtgcagggcgctgacttccgcgtttcca  
gactttacgaaacacggaaaccgaagaccattcatgttgttgcaggtcgcagacgtttgcagcagcagtcgcttccgctcgtcgcgta  
tcgggtgattcattctgtaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggcc  
aggaccaacgctgcccagatgcgccgcgtgcggctgctggagatggcggacgcgatggatatgttctgccaaagggttgggttgcgcat  
tcacagttctccgcaagaattgattggctccaattcttggagtggatccgttagcgaggtgccgcccgggttccattcagggtcaggtggc  
ccgggtccatgcaccgcgacgcaacgcggggaggcagacaaggtatagggcgggcgcctacaatccatgccaacccgttccatgtgctc  
gccgaggcgggcataaatcgccgtgacgatcagcgggtccaatgatcgaagttaggctggtaagagccgcgagcgatccttgaagctgtccc  
tgatggctgctcatctacctgcctggacagcatggcctgcaacgcgggcatcccgatgccgccggaagcgagaagaatcataatggggaa  
ggccatccagcctcgcgtcgcgaacgccagcaagacgtagcccagcgcgtcggccgccatgccggcgataatggcctgcttctcgcg  
aaacgttgggtggcgggaccagtgacgaaggcttgagcgagggcggtgcaagattccgaataccgcaagcgacaggccgatcatcgtcgc  
gctccagcgaaagcgggtcctcgcggaaaatgaccagagcgtgcccggcacctgtcctacgagttgcatgataaagaagacagtcataag



139/167

**Figure 66B**

tgcggcgacgatagtcatgccccgcgcccaccggaaggagctgactgggttgaaggctctcaagggcacggtcgagatcccgggtgcct  
aatgagttagctaaacttacattaattgcgttgcgctcactgcccgtttccagtcgggaaacctgtcgtgccagctgcattaatgaatcgcca  
acgcgcggggagaggcggttgcgtattgggcgccagggtggtttttttaccagttagacgggcaacagctgattgcccttcaccgcc  
tggccctgagagagttgcagcaagcggtccacgctggttggcccagcagggcgaaaatcctgttgatggtggttaacggcgggatataac  
atgagctgtcttcggtatcgtcgtatcccactaccgagatatccgcaccaacgcgcagcccggactcggtaatggcgcgcatcgcgccag  
cgccatctgatcgttggcaaccagcagcagtggtggaacgatgccctcattcagcatttgcattggttggtaaaaccgggacatggcactcca  
gtcgccttcccgttccgctatcggtgaatttgattgcgagttagatatttatgccagccagccagacgcagacgcgcggagacagaactta  
atgggcccgcctaacagcgcgatttgctggtgacctaatgcgaccagatgtccacgcccagtcgcgtaccgtctcatgggagaaaataat  
actgttgatgggtgtcgtgtagagacatcaagaaataacgccggaacattagtgcaggcagcttcacagcaatggcatcctggtcatcca  
gcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgccgcttacaggcttcgacgcgcgttcgttctaccatc  
gacaccaccagctggcaccagttgatcggcgcgagatttaatcgccgcgacaatttgcgacggcgcggtgcagggccagactggaggt  
ggcaacgccaatcagcaacgactgtttgcccgccagttgttgcacgcggttgggaatgtaattcagctccgccatcgccgcttcacttt  
ttcccgcgttttcgagaaacgtggctggcctggttcaccacgcgggaaacggctcgataagagacaccggcatactctgcgacatcgtata  
acgttactgggttcacattcaccacctgaattgactctcttcgggcgctatcatgccataaccgcgaaagggttttgcgccattcgatggtgtcc  
gggatctcgacgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttagggcggttagcaccgccgcccgaaggat  
ggtgcatgcaaggagatggcgcccaacagtcccccggccacggggcctgccaccataccacgccgaaacaagcgctcatgagccg  
aagtggcgagcccgatcttccccatcggtgatgtcgcgatataggcgccagcaaccgcacctgtggcgccggtgatgccggccacgat  
gcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgaactcactataggggaattgtgagcgggataacaattcccctctaga  
aataattttgttaactttaagaaggagatatacatatgaattaacctcactaaaggcgccgcgaagttcctattctctagaaagtataggaa  
cttcattctaccgggttaggggagggcgcttttcccaaggcagctctggagcatgcgttttagcagccccgctgggcacttggcgctacacaagt  
ggcctctggcctgcacacattccacatccaccggtaggcgccaaccggctccgttcttgggtggcccccttcgcgccaccttcactctcc  
cctagttaggaagttcccccccgccccgcagctcgcgtcgtgcaggacgtgacaaatggaagtagcacgtctcactagtctcgtgcagatg  
gacagcaccgctgagcaatggaagcgggtaggcctttggggcagcggccaatagcagcttgccttcgctttctgggctcagaggctg  
ggaaggggtgggtccggggggcgggctcagggggcgggctcagggggcgggcgggcgcccgaaggtcctccggaggccccggcattct  
gcacgctcaaaaagcgcacgtctgccgcgctgttctcctctcctcatctccgggcttgcagctgcagcagcagctgttgacaattaatcat  
cggcatagtatcggcatagtataatacgaaggtgaggaactaaaccatggagaaaaaatacactggatataccaccgttgatataatcc  
caatggcatcgtaaagaacattttgaggcattcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggccttttaag  
accgtaaagaaaaataagcacaagttttatccggcctttattcacattcttgcgcctgatgaatgctcatccggaattccgtatggcaatgaa  
agacgggtgagctggtgatatgggtagtgttcacccttgttacaccggtttccatgagcaaaactgaaacgtttcatcgctctggagtgaatacc  
acgacgatttccggcagtttctacacatatattcgcaagatgtggcggttacggtgaaaacctggcctatttccctaaagggtttattgagaat  
atgttttctcgtcagccaateccctgggtgagtttaccagttttgatttaaactggccaatatggacaacttcttcgccccgtttaccatgg  
gcaaatattatacgaaggcgacaaggtgctgatgccgctggcgattcaggttcatcatgccgttgtgatggcttccatgtcggcagaatgc  
ttaatgaattacaacagtactgcgatgagtgagggcgggggcgtaagcgggactctgggggtcgaataaagaccgaccaagcgacgtct  
gagagctccctggcgaaatcggtaccaataaaagagctttatttcatgatctgtgtgttgggtttgtgtgcggcgcggaagttcctattcttag  
aaagtataggaacttctcagccctatagttagtcgtattagatcgcgggcgcgcccttgactatgccacatcctgagcaataattcaacc  
actaattgtgagcgggataacaaaggaggtaaaaaacatggtatcctgttctgcgccgggtaagatttacctgttcggtgaacacgccgtagt  
ttatggcgaaactgcaattgcgtgtcggtggaactgcgtacccgtgttcgcgcgggaactcaatgactctatcactattcagagccagatcg  
ccgcaccggtctggatttcgaaaagcacccttatgtgtctgcggttaattgagaaaatgcgcaaatctattcctattaacgggtgtttcttgaccgt  
cgattccgacatcccgggtgggctccggtctgggtagcagcgcagccgttactatcgctctattggtgcgctgaacgagctgttcggctttg  
gcctcagcctgcaagaaatcgctaaactgggccacgaaatcgaaattaaagtacagggtgccgcgtccccaaccgatacgtatgtttctacc  
ttcgggcggtggttaccatcccgggaacgtcgcaactgaaaactccggactgcggcattgtgattggcgataccggcggttttctcctccac  
caaagagttagtagctaactgacgtcagctgcgcgaaagctacccggatttgatcgaaccgctgatgacctctattggcaaaatctctcgtat  
cggcgaaactggttctgtctggcgactacgcacatccatcgccgcctgatgaacgtcaaccagggtctcctggacgccctggcggttaac  
atcttagaactgagccagctgatctattccgctcgtgcggcaggtgcgtttggcgctaaaatcacgggcgctggcgggcggtggctgtatggt



140/167

**Figure 66C**

tgcgctgaccgctccggaaaaatgcaaccaagtggcagaagcggtagcaggcgctggcggtaaagtgactatcactaaaccgaccgag  
caaggtctgaaagtagattaa (SEQ ID NO:93)



141/167

Figure 67A

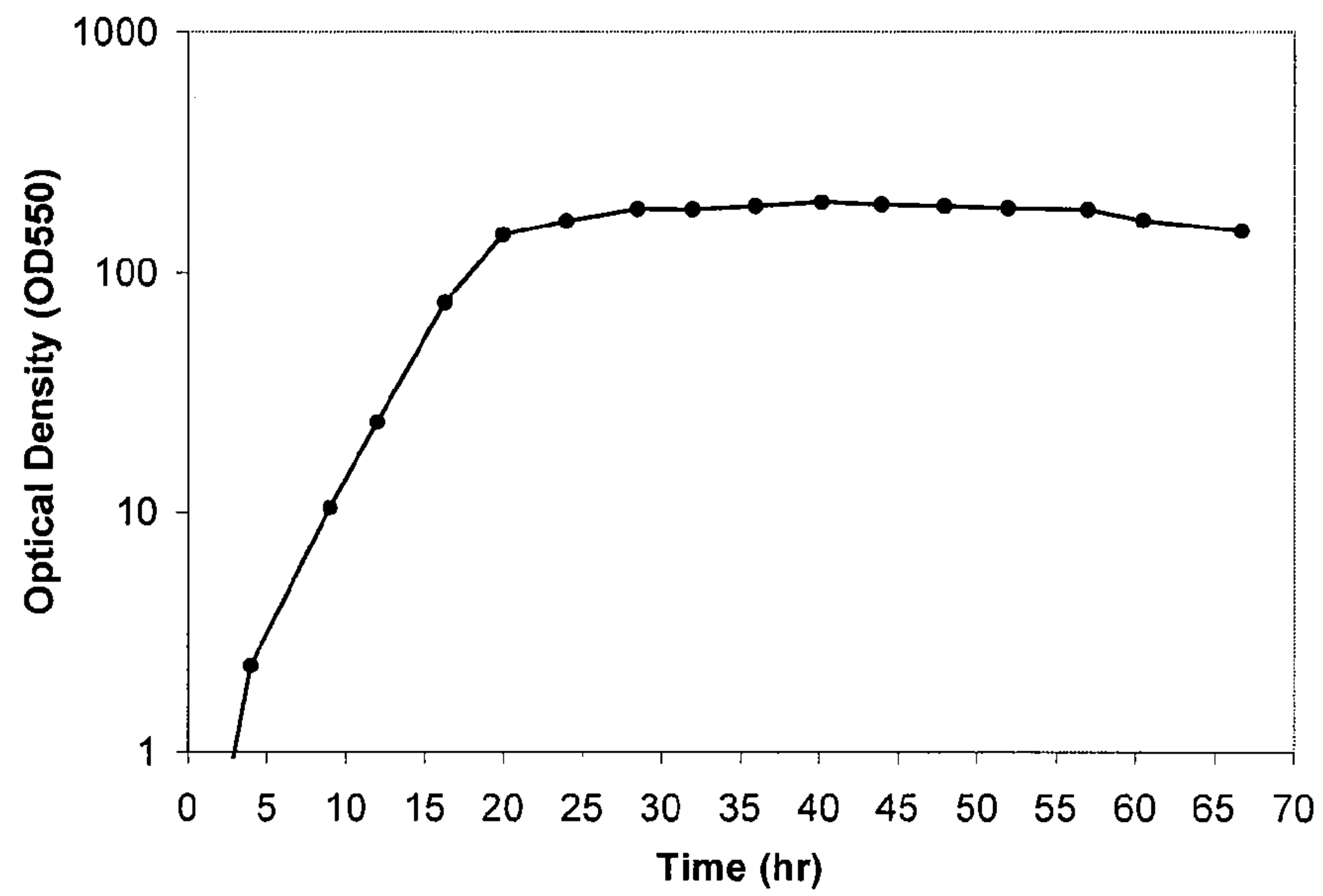
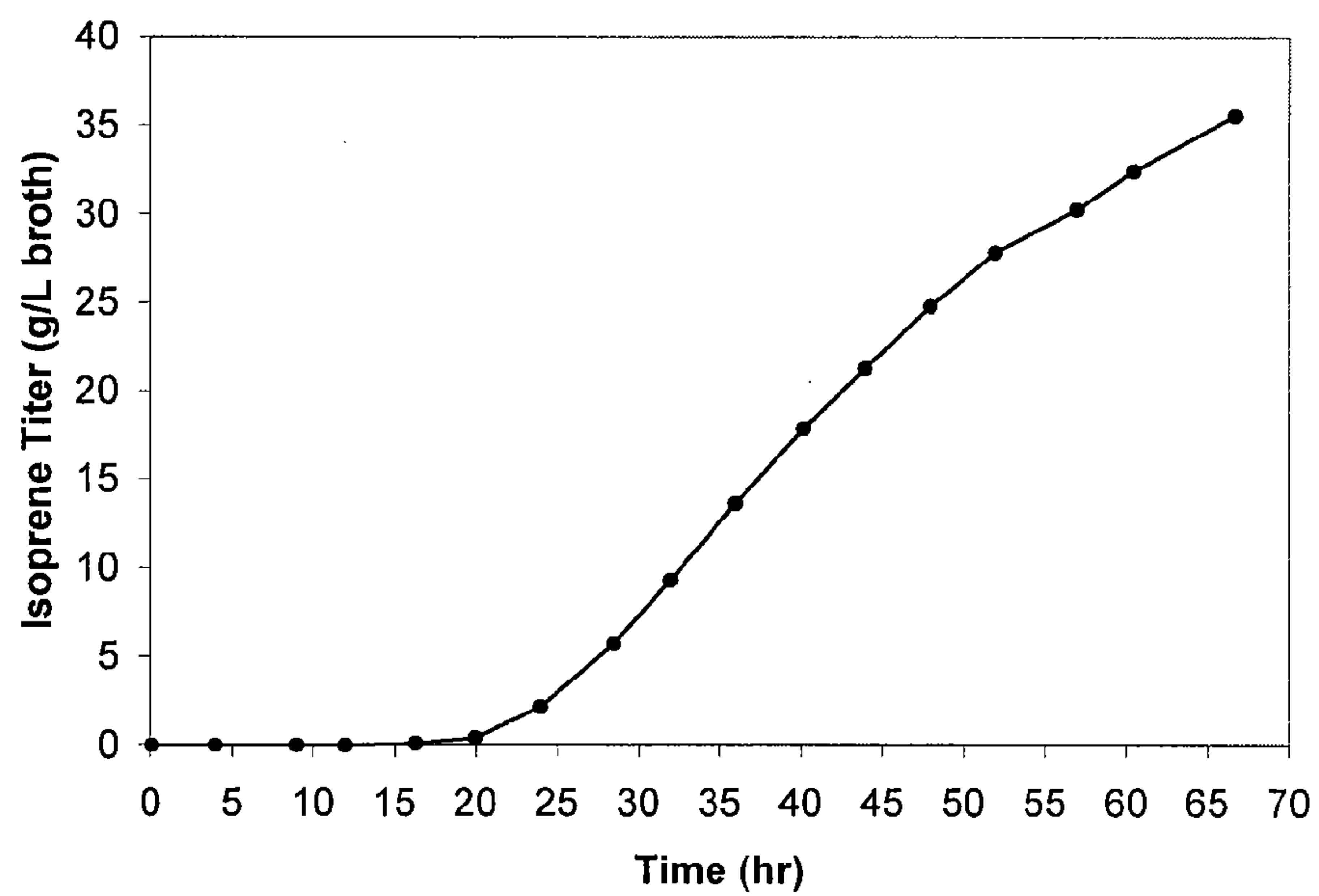


Figure 67B





142/167

Figure 67C

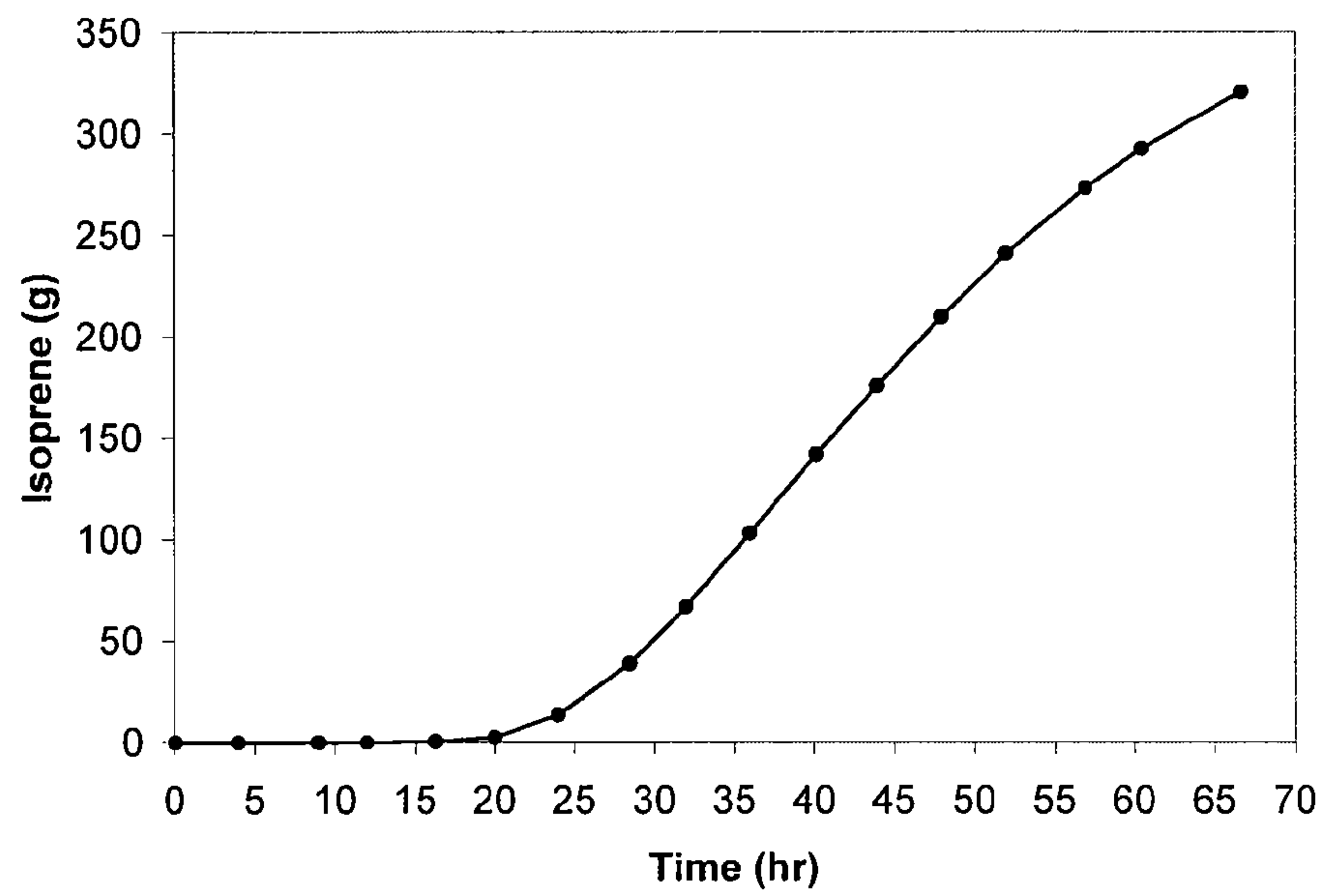
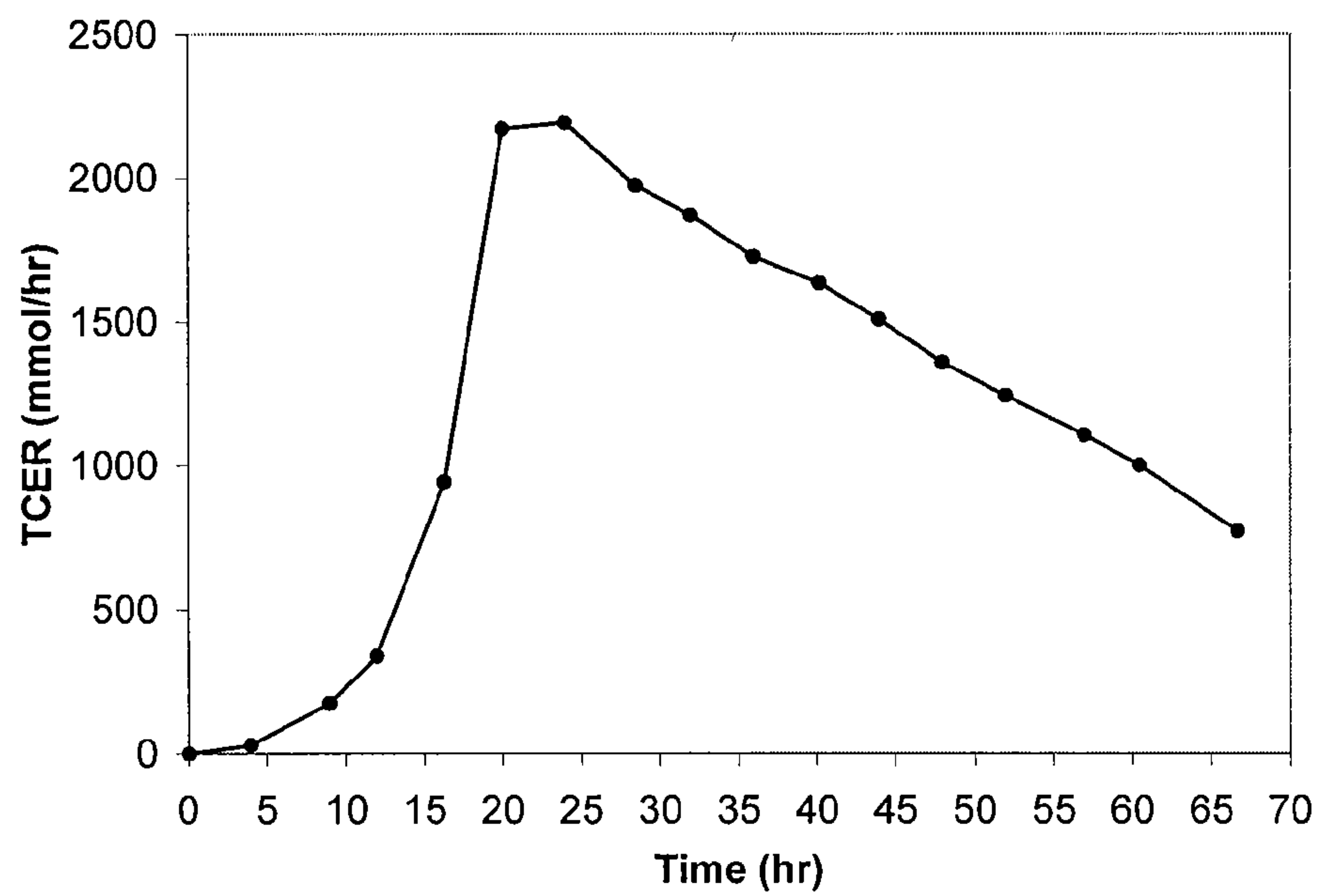


Figure 67D





143/167

Figure 68A

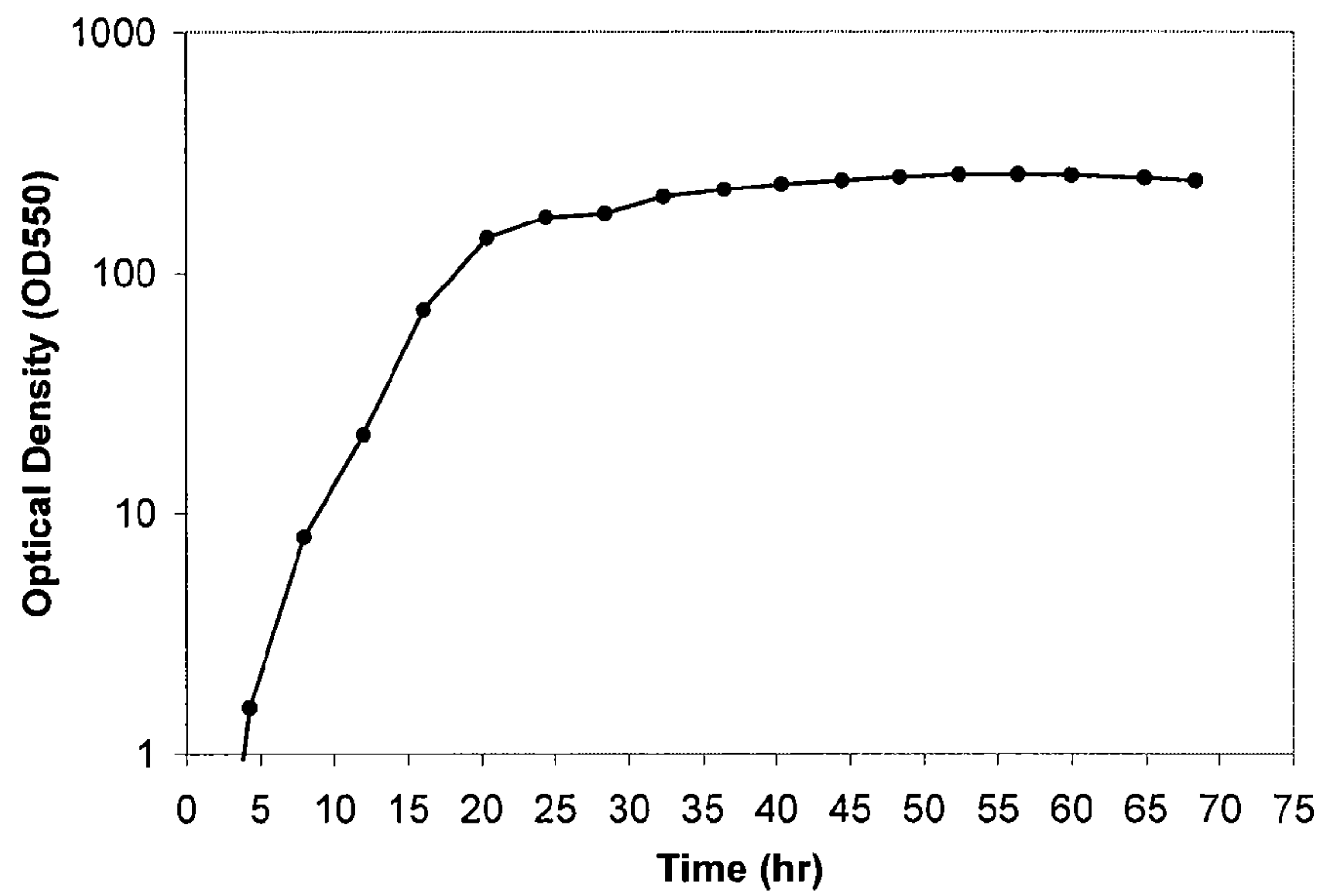
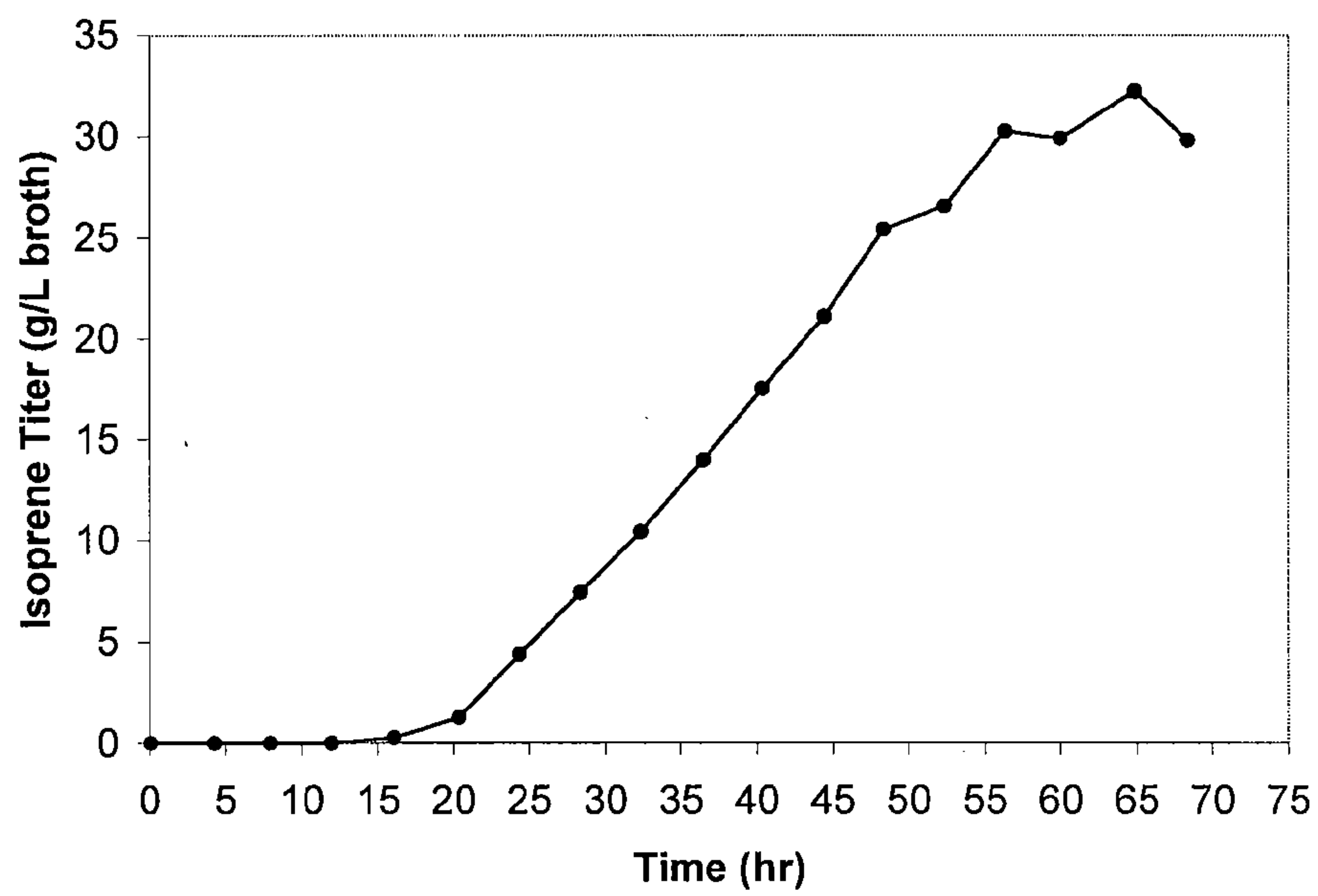


Figure 68B





144/167

Figure 68C

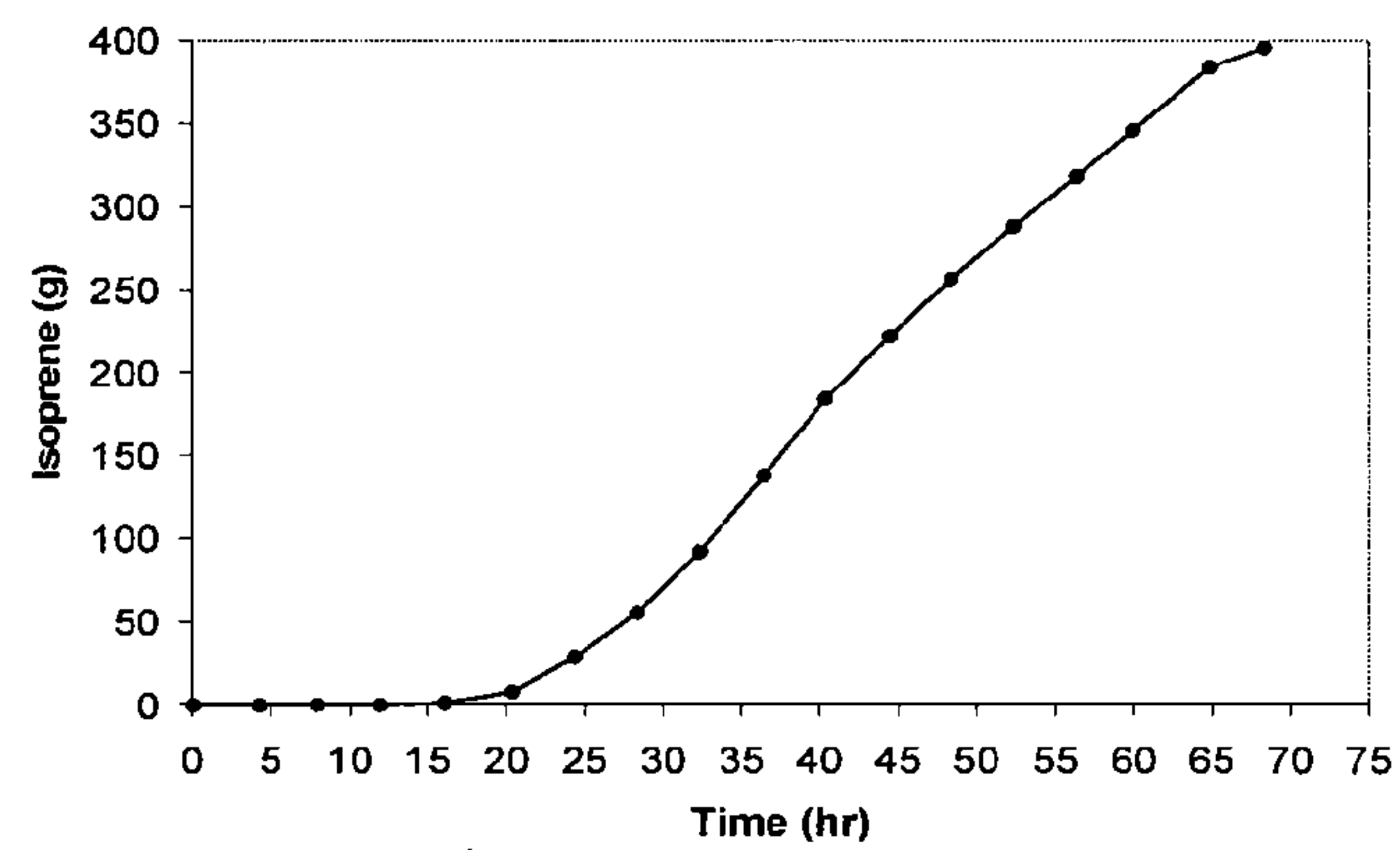


Figure 68D

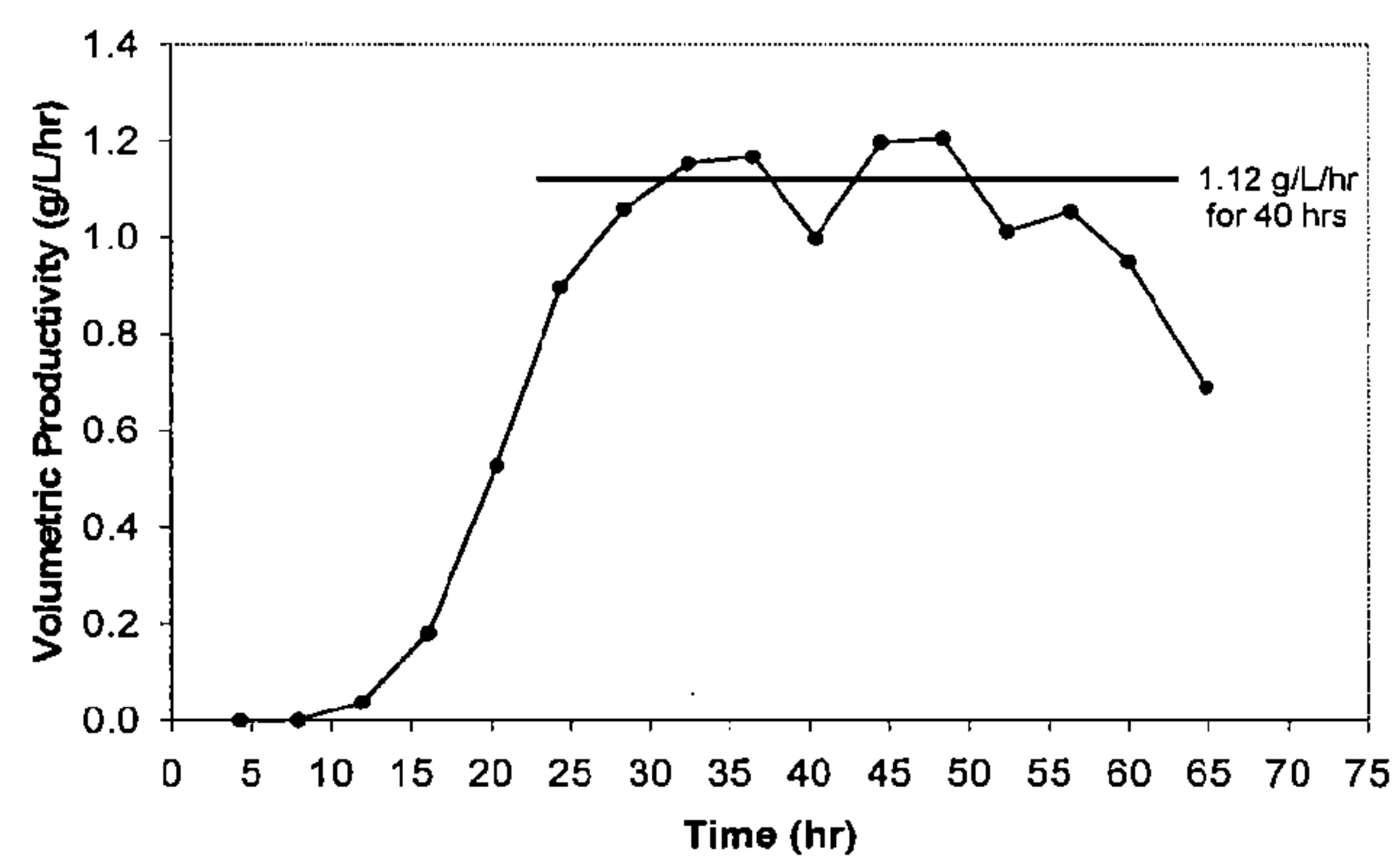
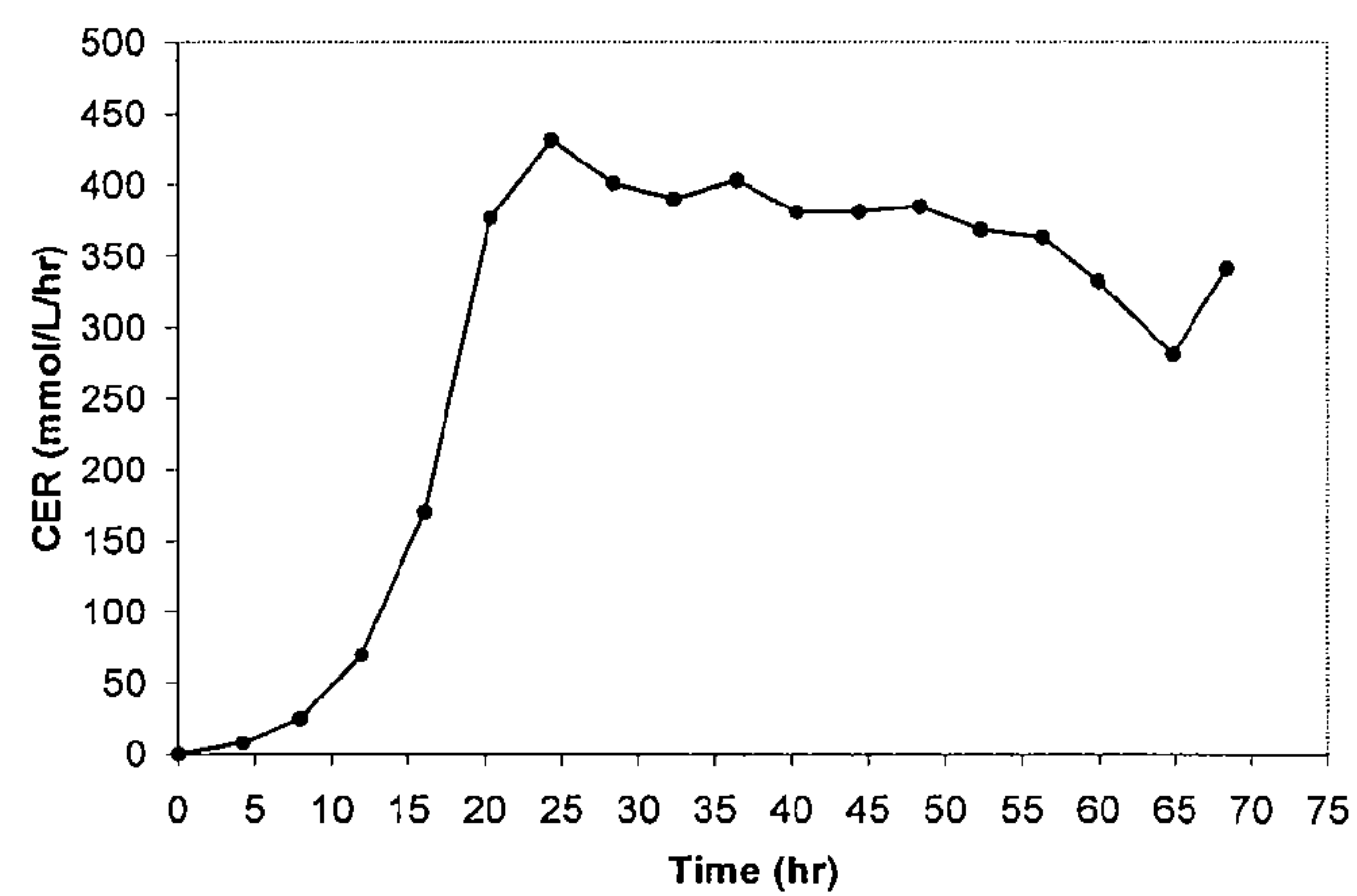


Figure 68E





145/167

Figure 69A

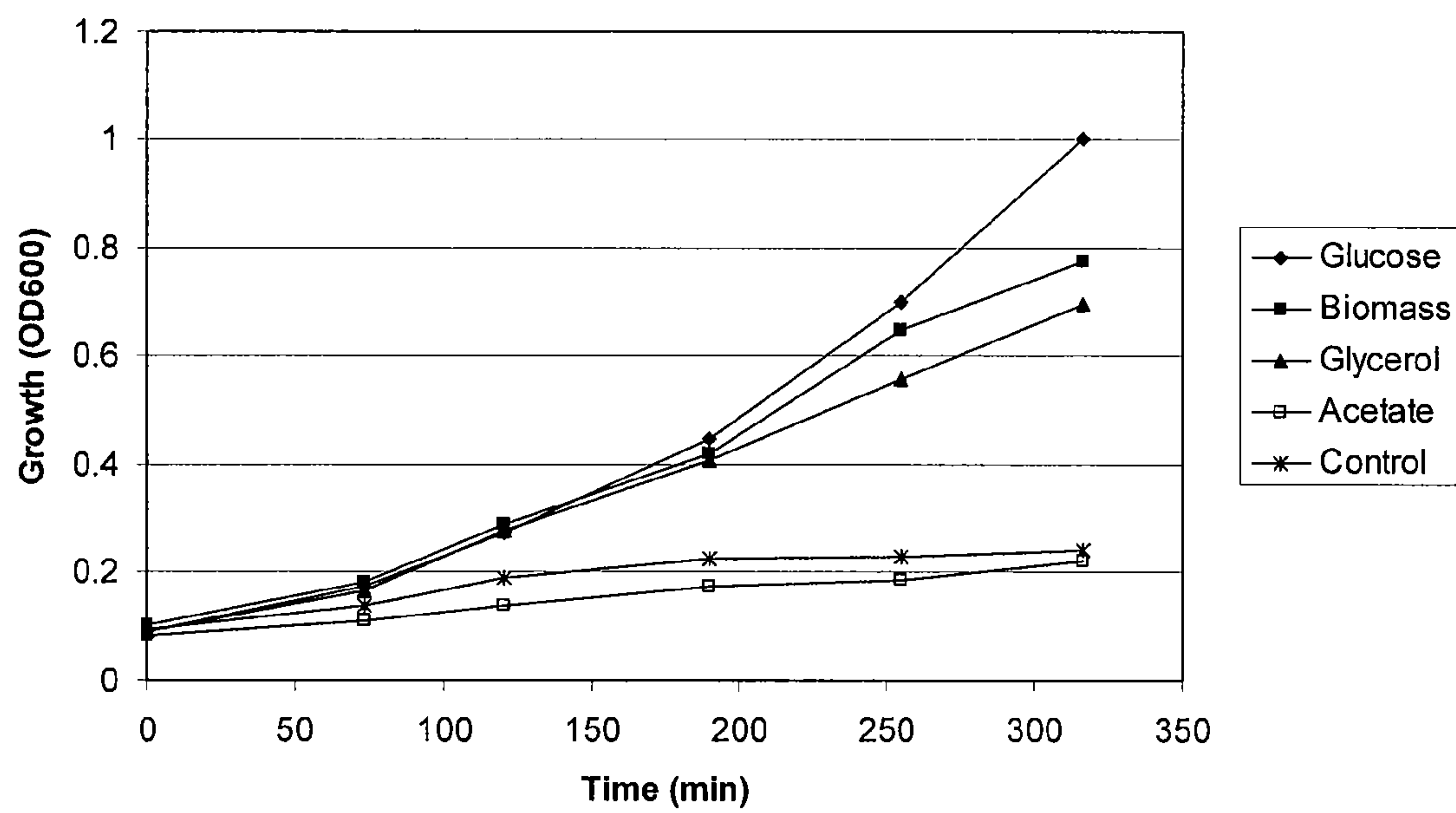
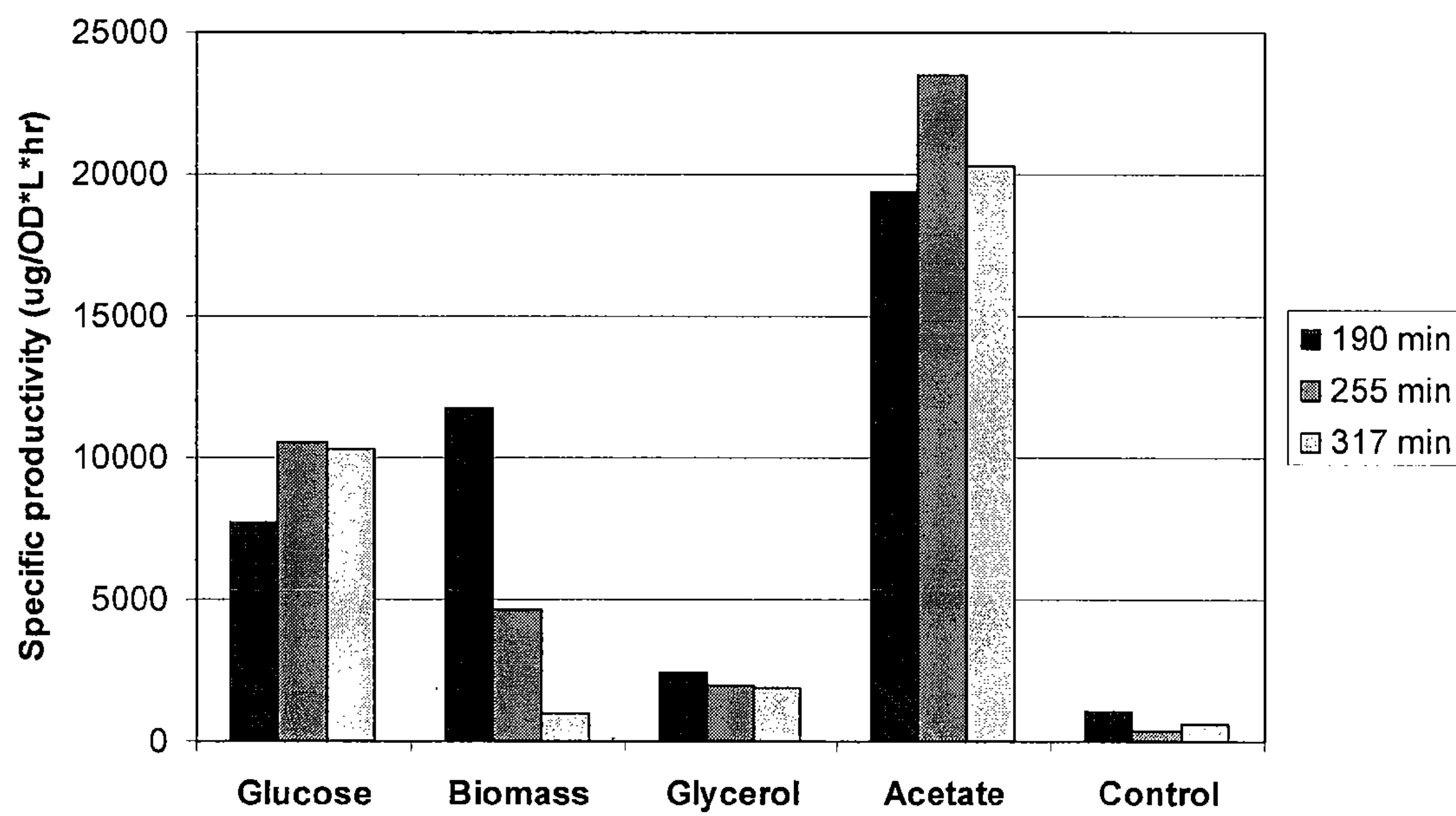


Figure 69B





146/167

Figure 69C

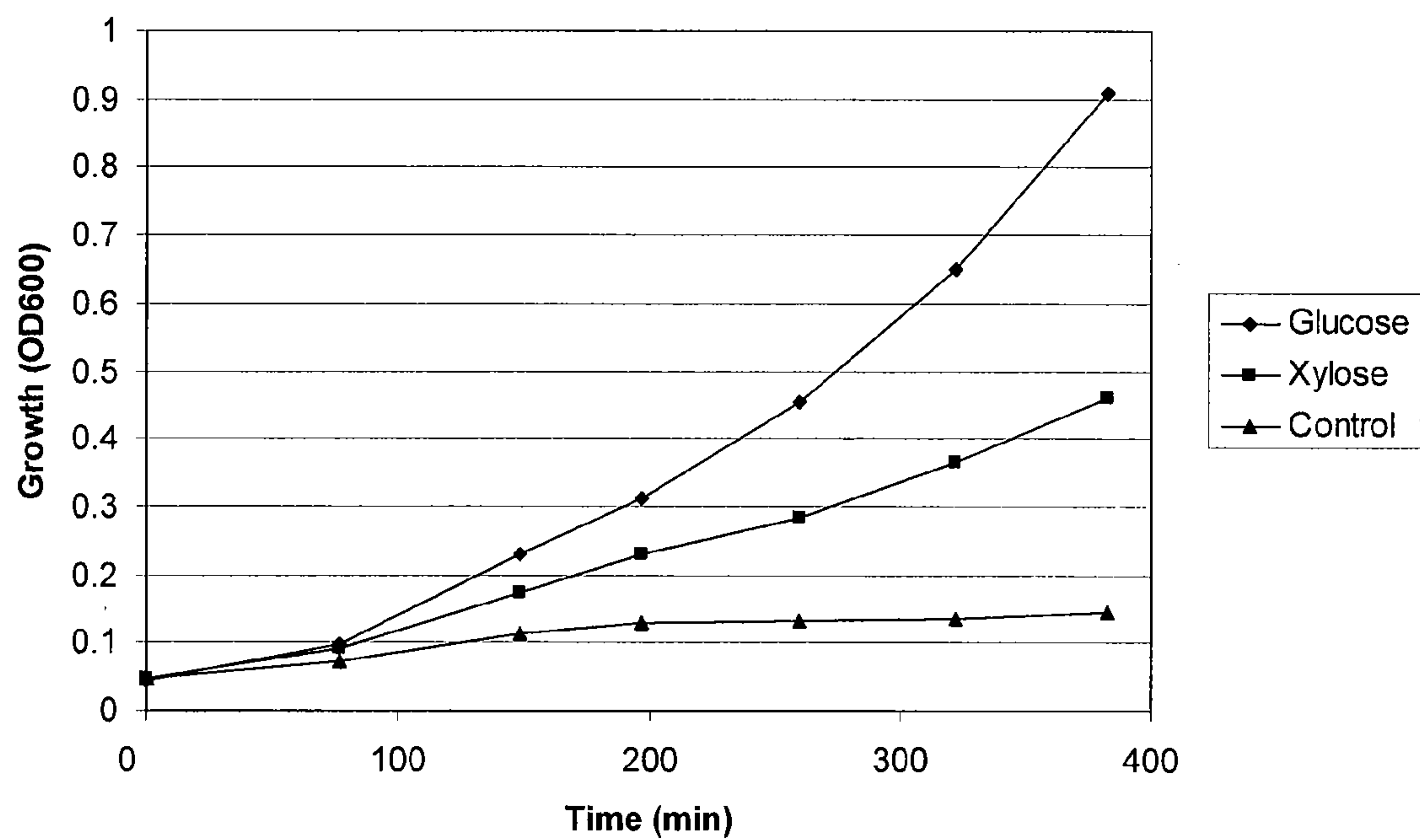
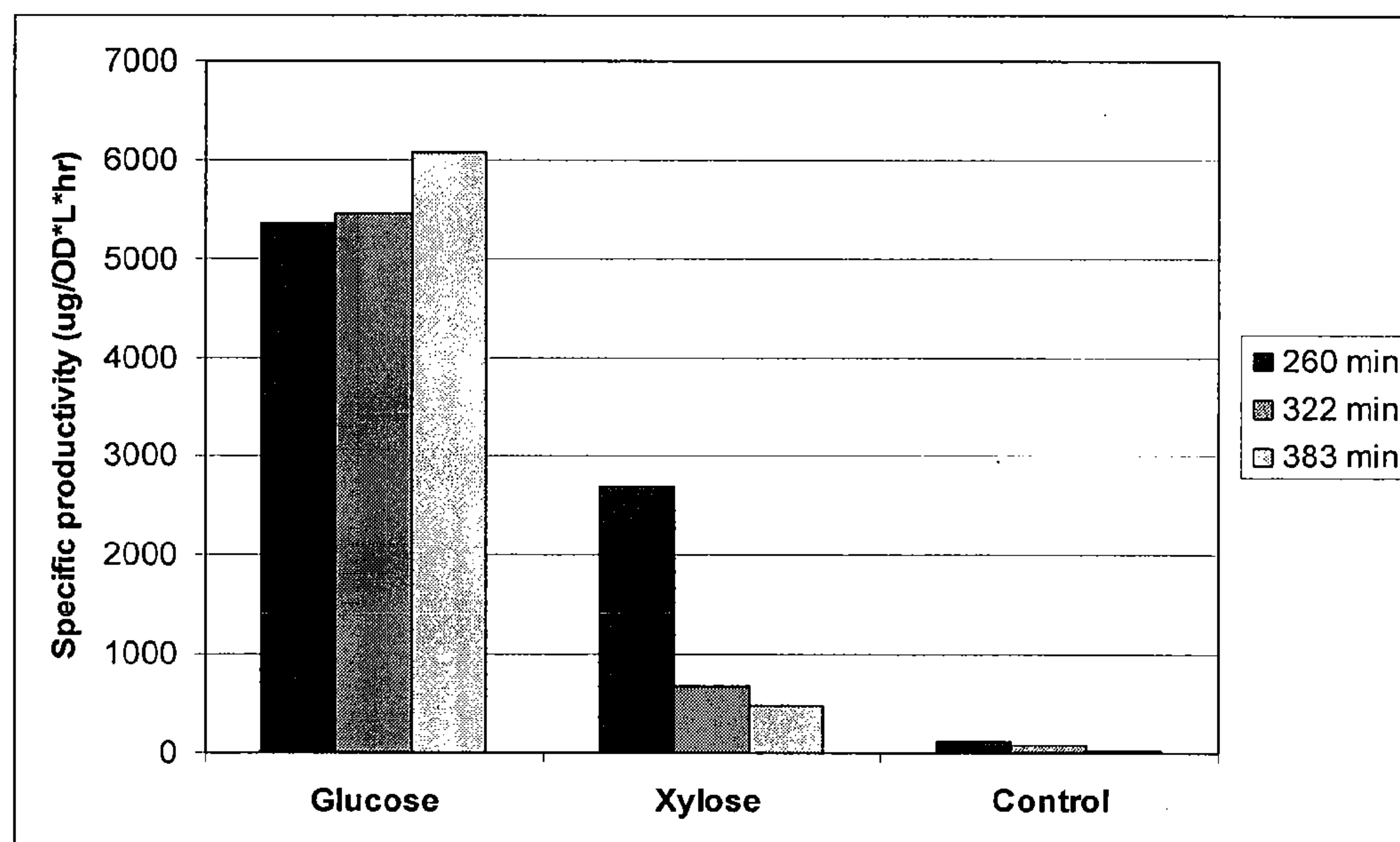


Figure 69D





147/167

Figure 70A

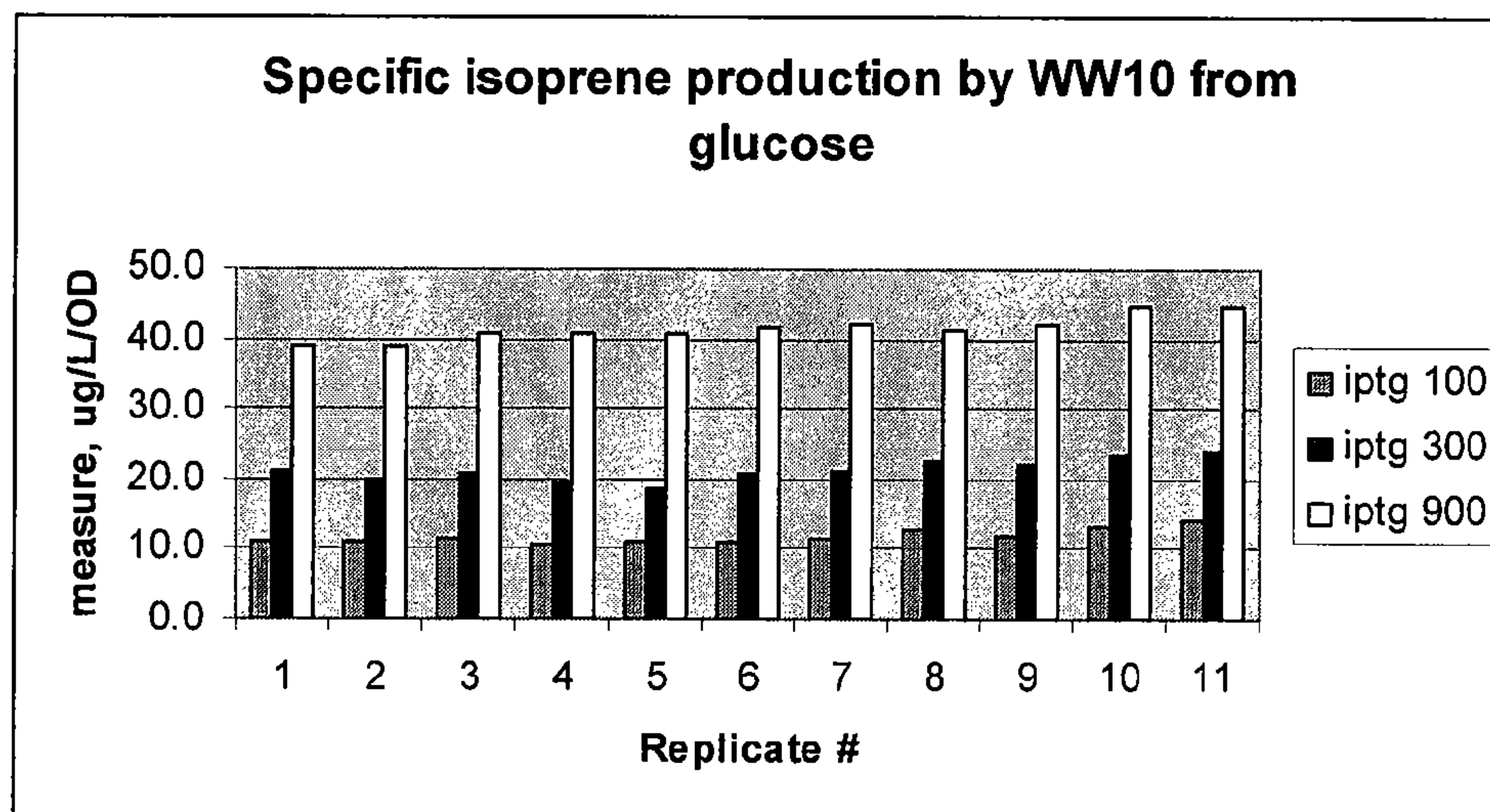
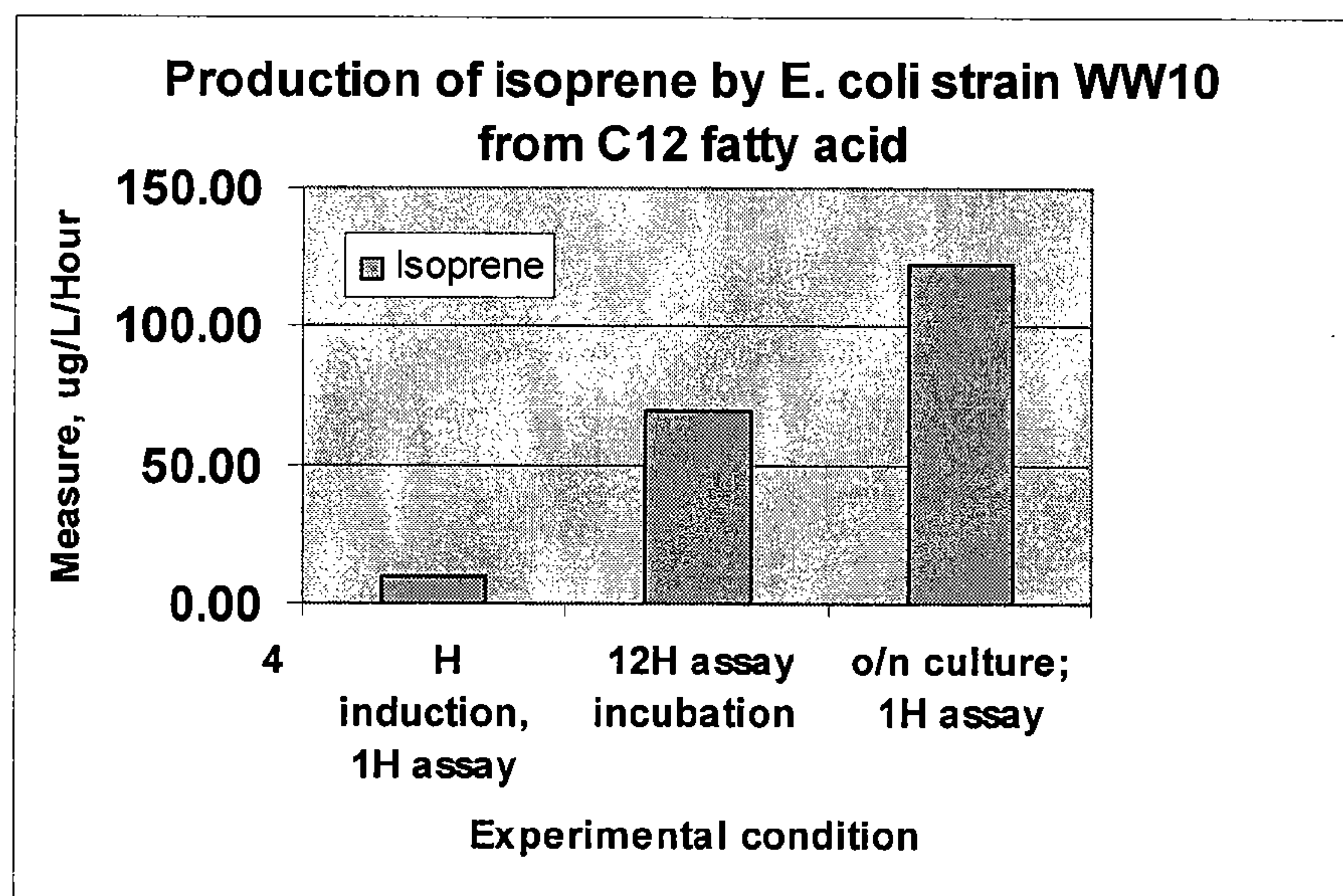


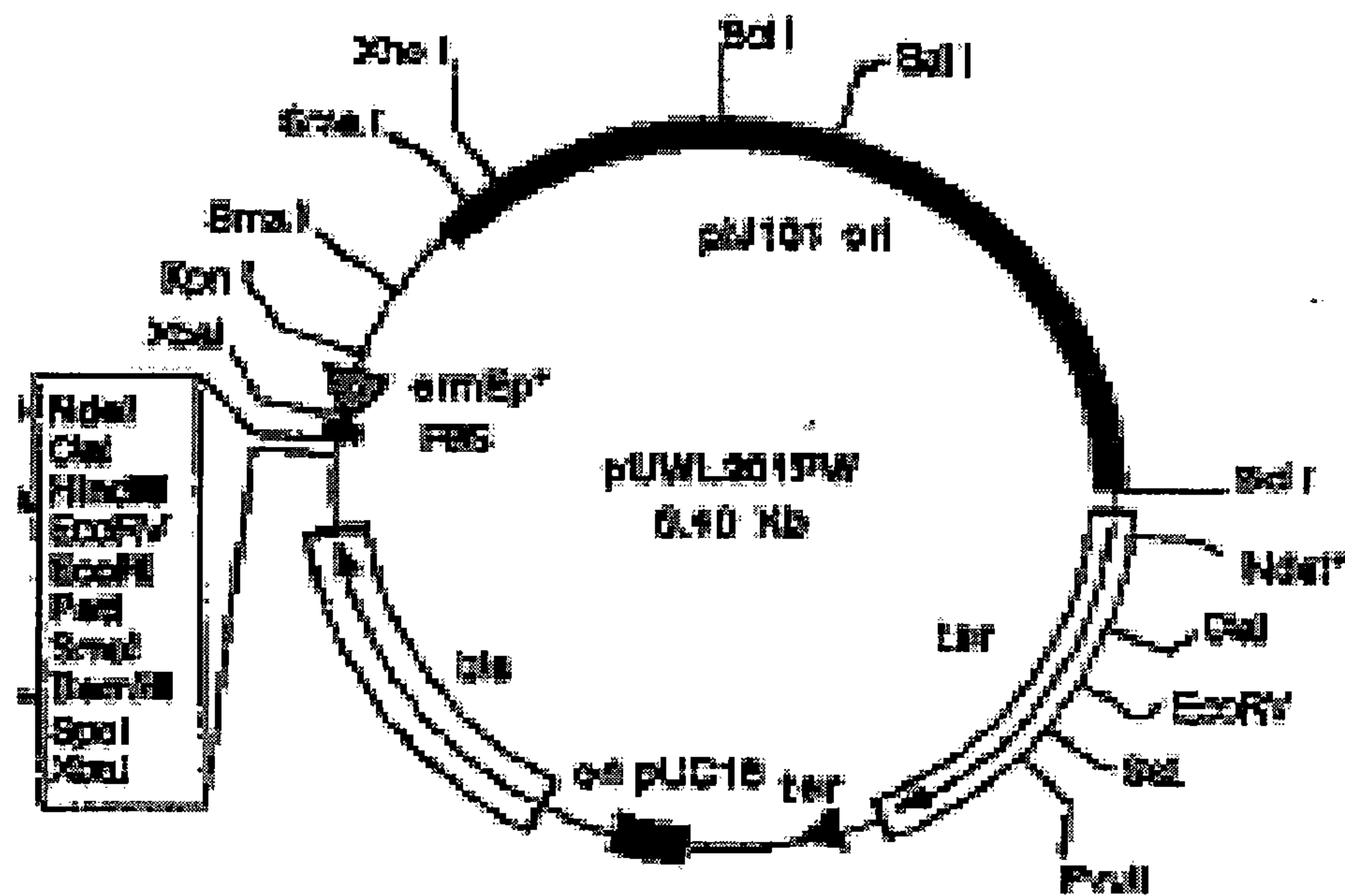
Figure 70B





148/167

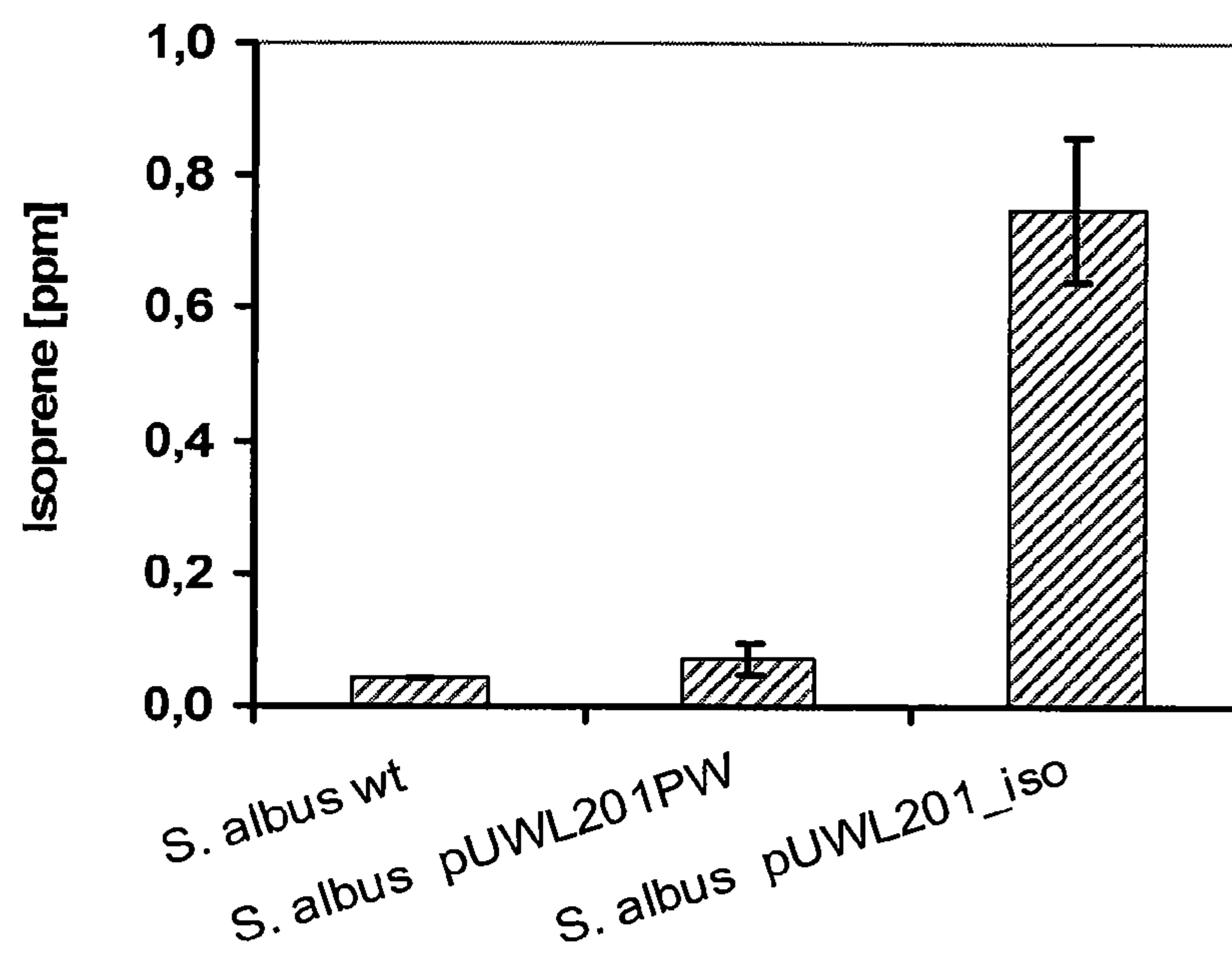
Figure 71





149/167

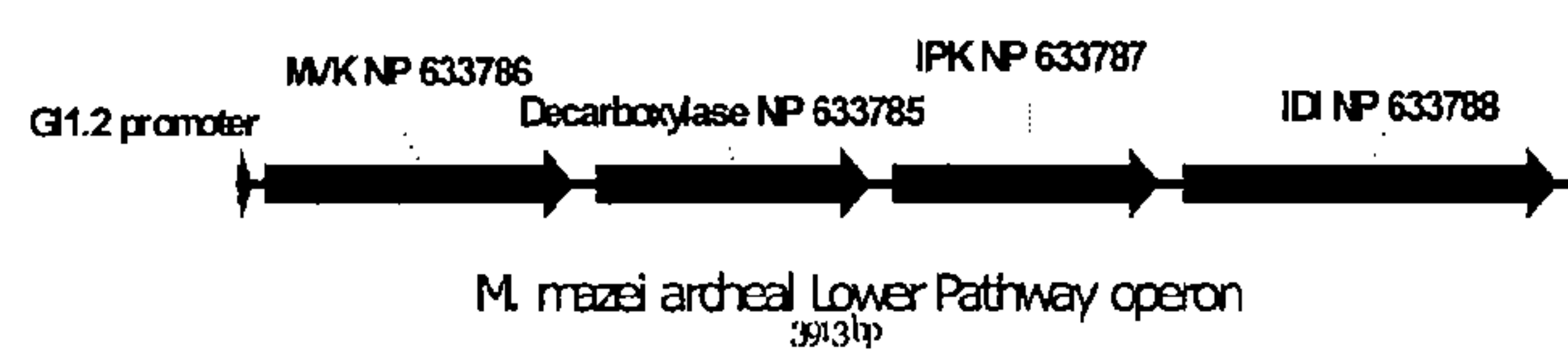
Figure 72





150/167

Figure 73A





151/167

**Figure 73B**

gcggccgcgcccttgacgatgccacatcctgagcaaataattcaaccactaattgtgagcggataacacaaggaggaaacagc  
catggtatcctgttctgcgcccggtaagatttacctgttcgggtgaacacgccgtagtttatggcgaaactgcaattgcgtgtgcggt  
ggaactgcgtaccctgttcgcgcgggaactcaatgactctatcactattcagagccagatcgcccgccaccggctctggatttcgaa  
aagcacccttatgtgtctgcggttaattgagaaaaatgcgcaaatctattcctattaacgggtgtttcttgaccgtcgattccgacatccc  
ggtgggctccggctctgggttagcagcgcagccgttactatcgctctattgggtgcgctgaacgagctgttcggctttggcctcagc  
ctgcaagaaatcgctaaactgggccacgaaatcgaaattaaagtacagggtgccgcgtccccaaccgalacgtatgtttctacct  
tcggcggcgtgggttaccatcccggaaactgcgcaaatgaaaactccggactgcggcattgtgattggcgataccggcggtttctc  
ctccaccaaaagagtttagtagctaactgtacgtcagctgcgcgaaagctacccggatttgatcgaaccgtgatgacctctattggc  
aaaatctctcgtatcgcggaacaactggttctgtctggcgactacgcatccatcgcccgctgatgaacgtcaaccagggtctcct  
ggacgccctgggcggttaacatcttagaactgagccagctgatctattccgctcgtgcggcaggtgcgtttggcgctaaaaatcacg  
ggcgctggcgggcggtgggtgtatggttgcgctgaccgctccggaaaaatgcaaccaagtggcagaagcggttagcaggcgct  
ggcggtaaagtgactatcactaaaccgaccgagcaaggtctgaaagtagattaagccttgacttaataagctgcttatttcgccctta  
tggtacctagtaggaggaaaaaaacatggaaatgcgtcaaccggctgtcgcaggtcaattctaccactgcgttgcgagaacct  
ggaaaacgaactgaaacgctgcttcgaaggcctggagatccgcgaacaagaagtgcgtggcgagctctgcccgcacgccggt  
tatatgtactctggcaaagttgcggcgacgtctatgccactctgccggaagctgatacctacgtaattctcgcccgaaccacac  
cggctacggtagccctgtctctgtgagccgtgaaacttgaagacccccgttgggcaatatcgaatgttgacctggaactggcgga  
cggcttcctgggttccatcgtagatgcggatgaactcggtcacaaatacgaacactctatcgaagttcagctgccgtttctgcaata  
ccgtttgaacgcgatttcaaaattctgccaatctgcatgggtatgcaagacgaagaaaccgcggtcgaagtaggtaacctgctg  
gcggatctgatcagcgagtcggtaaacgtgctgtgatcctgcgaagctctgattcaccactatgagacggctgaacgtgcc  
aagaaatcgattccgaagttattgattctatcctgaactttgacatctctggcatgtacgacgcctgtatcgccgtaacgcctctgtt  
gcggttacggcccgatcaccgctatgctgacggcaagcaaaaagctggcggtctcgtgcgactttgctgaaatacgcaaac  
agcggtgacgtgtccggtgataaagacgctgtggtgggtacgccgccatcctcgttgagtaagctgattaaaggtgaacagat  
aggatttcgtcatggatcctacaaggaggaaaaaaacatgaatgcttctaatgaaccgggtgattctgaaactgggtggctctgcta  
ttaccgacaaaggtgcctacgaaggcgtagttaaggaagctgatttgcgcacatgcacaggaagtttagcggtttccgtggca  
agatgatcgtgggtcatggtgctggttagcttcggccatacgtacgcgaagaaatacggcctggaccgtaccttcgaccagagg  
gcgcaattgttactcatgaatctgttaaaaagctgcctccaaagttaggtgctctgaatagcttcggcgtgcgtgctatcgcg  
tgcatcctatggactgcgcagtatgccgtaacggctgatacgaacgatgtatctggactccatcaagttaatgctggaaaaaggt  
ctggtgccggttctgcacggcgacgtcgcaatggatattgaactgggcacttgatcctgtccggtgatcaaatcgttccttacctg  
gccaaagaactgggtatctcccgcctcggcctgggcagcgcagaggatgggtgtgctggatatggagggcaaacctgtaccgg  
aaatcaccaccagaaactttcgaagagttccgccactgcatcggtggttctggttctactgatgtaaccggtggcatgctgggcaa  
gtgctggaacttctggaattgagcaaaaaattcttccattactagctacatttcaacgctggtaaagcagacaacatctaccgcttct  
gaatggtgagtcacatcggcactcgcatcagcccggaacagcgtgtttaagctagttattaacctaaatgctctaaaccagttatga  
gctctacaaggaggaaaaaaacatgattaacactaccagccgccgcaaaattgaacacctgaaactctgcgcagaatccccgg  
ttgaagcgcgtcaggatatctgccggctttgaagacgttactctgatccaccgcgctttaccggagctgaacatggatgaactgga  
cctcagcgttgatttccgtgggttaaaccgcatcaaaagcggcgttccgtgattgcgtctatcaggggtgggtcaccagataccatcccgg  
ttaacgctgcgctggcagctgctgctgaggagctgggtgttggcatcggcgttgggtctcagcgcgcggccattgatgatccga  
gccaggaagacagcttccgtgtagtgcgtgatgaagccccagatgcgtttgttatggcaacgtcggcgagcacagatccgtc  
agtatggtgttgaaggtgttgaaaaactgatcgaatgattgacgcagatgccttggcaatccacctgaactttctgcaagaagcg  
gtccaaccggaaggtgaccgcgacgcgaccggttgcctggacatgattaccgaaatttgctctcagattaaaactccggtaatcg  
tgaaagaaaccggtgcaggcattagccgtgaagatgcgattctgttccagaaagctggcgtgagcgcgaatcgacgttggcggc  
gcgggcggcacctcctgggctggcgctgaggtctaccgtgctaaagaaagccgtgactctgtttagcgagcggttaggtgagct  
gttttgggatttcggcattccgacggtagcttctctgattgaatcccgcgtttccttgcgctgatcgaaccggcggtatccgtaac  
ggtctggacattgctaaaagcattgctctcggcgcaagcgtgccagcggcgtctgcccgttcgttgggtccgtccctggagggc  
aaagaatccgttgatcgtgtgctgagctgcatgctggaagaatttaaagcagcaatgttttgcggttgcggcaacatcaaaga



152/167

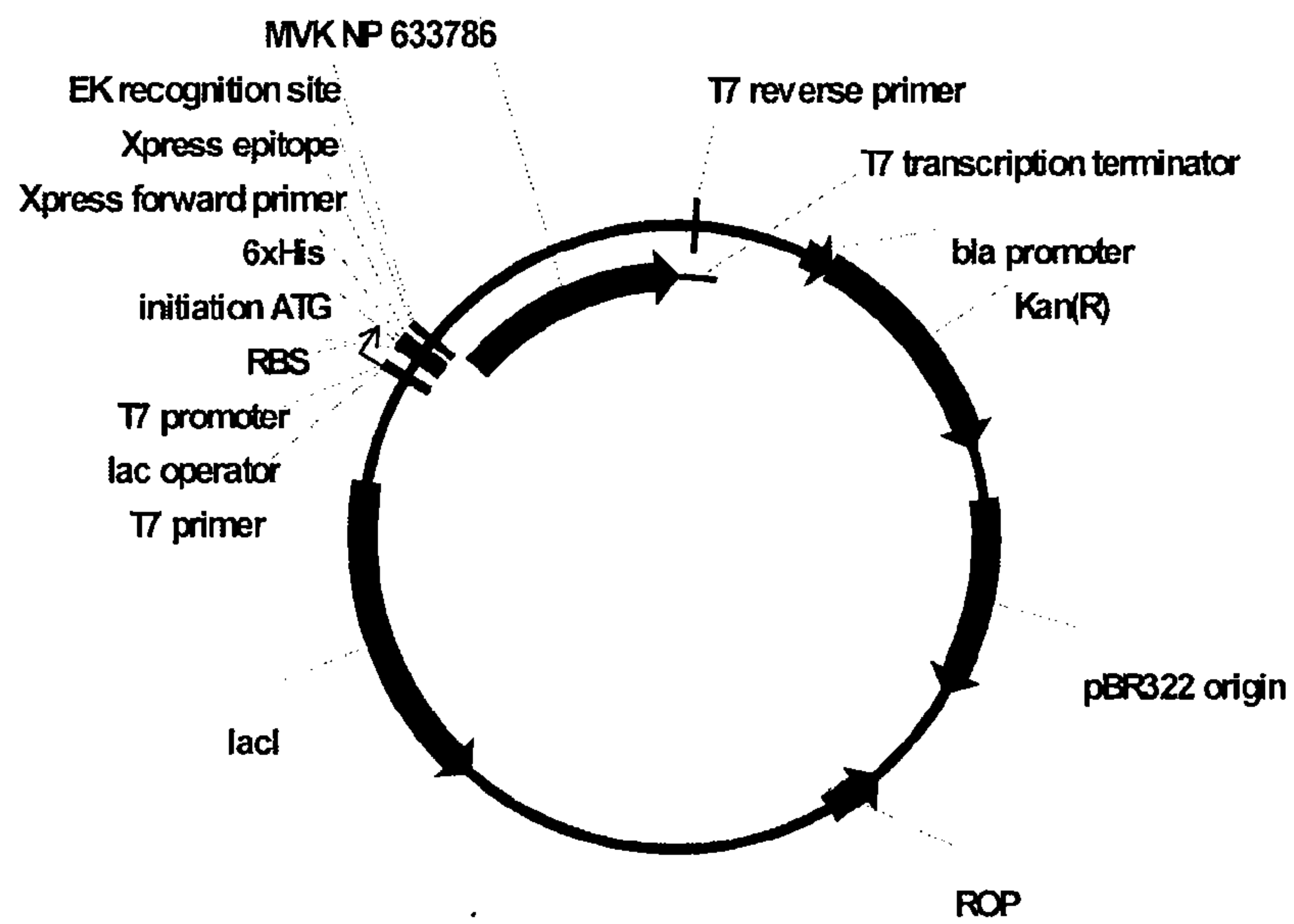
**Figure 73C**

cctgcacaactctccagtagtggttaactgggtggacccgcgaatacctggagcagcgcggtttaacgttaaggacctctccctg  
ccgggcaacgctctgtaagcttcaacgcgtctacaaataaaaaaggcacgtcagatgacgtgcctttttctgtctaga  
(SEQ ID NO:113)



153/167

Figure 74A



MCM376 - MMK from *M. mazei* archeal Lowerin pET200D  
6647 bp



154/167

**Figure 74B**

aaggcgagctcaacgatccggctgctaacaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactag  
cataacccttggggcctctaaacgggtcttgaggagtttttctgaaaggaggaactatatccggatatcccgcaagaggccc  
ggcagtaccggcataaccaagcctatgcctacagcatccagggtgacgggtgccgaggatgacgatgagcgcatgttagattc  
atacacgggtgcctgactgcgttagcaatttaactgtgataaactaccgcattaaagcttatcgatgataagctgtcaaactgagaa  
ttaattctgaagacgaaagggcctcgtgatacgcctattttataggtaatgtcatgataataatggtttcttagacgtcaggtggca  
cttttcggggaaatgtgcgcggaacccctattgtttattttctaaatacattcaaatatgtatccgctcatgagacaataaccctgat  
aatgcttcaataatattgaaaaaggaagagtattgaacaagatggattgcacgcaggttctccggccgcttgggtggagag  
gctattcggctatgactgggcacaactgacaatcggctgctctgatccgcccgtgttccggctgtcagcgcagggggcgcccgtt  
tcttttgtcaagaccgacctgtccgggtgcctgaatgaactgcaggacgaggcagcgcggctatcgtggctggccacgacgg  
gcgttccttgcgcagctgtgctcagcttgcactgaagcgggaagggaactggctgctattgggcgaagtgcggggcgaggat  
ctcctgtcatctcaccttgcctgccgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgttgatccggctac  
ctgcccattcgaccaccaagcgaaacatcgcatcgagcggggcacgtactcggatggaagccggctcttgatcaggatgatct  
ggacgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcgatcccgcaggcgaggatctcgt  
cgtgacacatggcgatgcctgcttgcgaataatcatgggtggaaaatggccgctttctggattcatcgactgtggccggctgggtg  
tgccggaccgctatcaggacatagcgttggctaccctgatattgtgaagagcttggcggcgaatgggctgaccgcttctcgt  
gctttacgggtatcgcgctcccgatcgcagcgcategccttctatcgccttcttgacgagttcttctgagcgggactctggggtc  
gaaatgaccgaccaagcgacgcctaactgtcagaccaagttaactcatatatacttttagattgatttaaaacttcatttttaatttaaaa  
ggatctaggtgaagatccttttgataatctcatgacaaaatcccttaacgtgagtttctgcttccactgagcgtcagaccccgtaga  
aaagatcaaaggatcttcttgagatcctttttctgcgcgtaactgtctgcttgcacacaaaaaaccaccgctaccagcggtggtt  
tgcttgcgggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagataaccaatactgtccttctagtgt  
agccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgctaatectgttaccagtggctgctgcc  
agtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggctcgggctgaacggggg  
gttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccac  
gcttcccgaaggagaaaaggcggacaggtatccggtgaagcggcagggctcggaaacaggagagcgcacgagggagcttcca  
gggggaaacgcctggtatctttatagtcctgtcgggttccgacactctgacttgagcgtcgattttgtgatgctcgtcagggggg  
cggagcctatggaaaaacgccagcaacgcggccttttacgggttcttggccttttgccttcttgcacatgttcttctcgtt  
atcccctgattctgtggataaccgtattaccgcctttgagtgagctgataaccgctcgcgcagccgaacgaccgagcgcagcga  
gtcagtgagcaggaagcgggaagagcgcctgatgcggtattttctccttacgcactctgtcgggtatttcacaccgcaatggtgca  
ctctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatcgtactgactgggtcatggctgcgcccc  
gacacccgccaacaccgctgacgcgcctgacgggcttgtctgctccggcatccgcttacagacaagctgtgaccgtctcc  
gggagctgcatgtgtcagaggtttcaccgtcatcaccgaaacgcgcgaggcagctgcggtaaagctcatcagcgtggtcgtg  
aagcgattcacagatgtctgctgttcacccgcgtccagctcgttgagtttctccagaagcgttaatgtctggcttctgataaagcgg  
gccatgttaaggggcggttttctggttgcactgatgcctccgtgtaagggggatttctgttcatgggggtaatgataccgatga  
aacgagagaggatgtcacgatacgggttactgatgatgaacatgcccggttactggaacgttgtgagggtaaacaactggcg  
gtatggatgcggcgggaccagagaaaaatcactcagggtcaatgccagcgttcgttaatacagatgtaggtgttccacagggt  
agccagcagcatcctgcgatgcagatccggaacataatggtgcagggcgctgacttccgcgttccagactttacgaaacacgg  
aaaccgaagaccattcatgttgttgcaggtcgcagacgtttgcagcagcagctcgttcacgttcgctcgcgtatcggtgattca  
ttctgtaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggccagg  
accaacgctgcccagatgcgcccgtgcgggtgctggagatggcggacgcgatggatatgttctgccaagggttgggttgc  
gcattcacagttctccgcaagaattgattggctccaattcttgagtggtgaatccgttagcgaggtgccgcccgttccattcagg  
tcgaggtggccccggctccatgcaccgcgacgcaacgcggggaggcagacaaggatataggcgggcgccctacaatccatgcc  
aaccggttccatgtgctcgcgagggcgccataaatcgccgtgacgatcagcgggtccaatgatcgaagttaggctggtaagagc  
cgcgagcgatecttgaagctgtccctgatggctgctcatctacctgcctggacagcatggcctgcaacgcgggcatcccgatgcc  
gccggaagcgagaagaatcataatgggggaaggccatccagcctcgcgtcgcgaacgccagcaagacgtagcccagcgcgt



155/167

**Figure 74C**

cgcccgccatgccggcgataatggcctgcttctcgccgaaacgtttggtggcgggaccagtgcgaaggcttgagcgagggc  
gtgcaagattccgaataccgcaagcgacaggccgatcatcgctgcgctccagcgaaagcggctcctcgccgaaaatgaccag  
agcgctgccggcacctgtcctacgagttgcatgataaagaagacagtcataagtgcggcgacgatagtcatgccccgcgcca  
ccggaaggagctgactgggtgaaggctctcaagggcacggtcgagatcccggcgctaatgagtgagctaacttacattaatt  
gcgttgcgctcactgcccgtttccagtcgggaaacctgtcgtgccagctgcattaatgaatcgcccaacgcgcggggagagg  
cggtttgcgtattgggcgccagggtgggtttttctttaccagtgagacgggcaacagctgattgcccttcaccgctggccctga  
gagagttgcagcaagcgggtccacgctgggttgcgccagcaggcgaaaatcctgtttgatggtgggtaacggcgggatataacat  
gagctgtcttcggtatcgctgatacccactaccgagatatccgcaccaacgcgcagcccggactcggtaatggcgcgcatgtgc  
cccagcgccatctgatcggtggcaaccagcatcgagtggaacgatgccctcattcagcatttgcattggttggaaaaccgg  
acatggcactccagtcgccttcccgttccgctatcggtgaatttgattgcgagtgagatatttatgccagccagccagacgcaga  
cgcgccgagacagaacttaatgggcccgtaacagcgcgatttgctggtgacccaatgcgaccagatgctccacgcccagtcg  
cgtaccgtcttcatgggagaaaataactgttgatgggtgtctggtcagagacatcaagaaataacgccggaacattagtgcag  
gcagcttccacagcaatggcatcctggatccagcggatagttatgatcagcccactgacgcgttgcgcgagaagattgtgc  
accgccgtttacaggcttcgacgccgcttcgttctaccatcgacaccaccagctggcaccagttgatcggcgcgagatttaa  
tcgccgcgacaatttgcgacggcgctgcagggccagactggaggtggcaacgccaatcagcaacgactgtttgcccgccag  
ttgttgcacgcggttgggaatgtaattcagctccgccatcgccgcttccacttttcccgcgttttcgcagaaacgtggctggc  
ctggttaccacgcgggaaacgggtctgataagagacaccggcactctgcgacatcgtataacgttactggtttcacattacca  
ccctgaattgactcttccgggcgtatcatgccataccgcgaaaggttttgcgccattcgatggtgtccgggatctcgacgctct  
cccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgcccgcaaggaaatggtgcatg  
caaggagatggcgcccaacagtcccccggccacggggcctgccaccatacccacgccgaaacaagcgcctcatgagcccga  
agtggcgagcccgatcttcccacggtgatgtcggcgatataggcgccagcaaccgcacctgtggcgccgggtgatccggc  
cacgatgcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgaactactataggggaattgtgagcggataac  
aattcccctctagaataattttgtttaactttaagaaggagatatacatatcggggggttctcatcatcatcatcatggtatggta  
gcatgactggtggacagcaaatgggtcgggatctgtacgacgatgacgataaggatcatcccttaccatggtatcctgttctgc  
gccgggtaagatttacctgttcggtgaacacgccgtagtttatggcgaaactgcaattgcgtgtgcggtggaactgcgtaccgt  
gttcgcgcggaactcaatgactctatcactattcagagccagatcgggccgaccggcttgatttcgaaaagcacccttatgtgtc  
tgcggttaattgagaaaatgcgcaaatctattcctattaacgggtgtttcttgaccgtcgattccgacatcccgggtgggctccggtctg  
ggtagcagcgcagccgttactatcgctctattggtgcgctgaacgagctgttcggctttggcctcagcctgcaagaaatcgcta  
aactggggccacgaaatcgaattaaagtacagggtgccgcgtcccaaccgatacgtatgtttctaccttcggcggcgtggttac  
catcccggaacgtcgcaactgaaaactccggactgcggcattgtgattggcgataccggcggtttctcctccaccaaagagtta  
gtagctaacgtacgtcagctgcgcgaaagctacccggattgatcgaaccgctgatgacctctattggcaaaatctctcgtatcg  
cgaacaactggttctgtctggcgactacgcattccatcgcccgctgatgaacgtcaaccagggtctcctggacgccctgggcgt  
taacatcttagaactgagccagctgatctattccgctcgtgcggcaggtgcgtttggcgctaaaatcacgggcgctggcgggcggt  
ggctgtatggttgcgctgaccgctccggaaaaatgcaaccaagtggcagaagcggttagcaggcgctggcggttaaagtactat  
cactaaaccgaccgagcaaggtctgaaagtagattaa  
(SEQ ID NO:114)



156/167

Figure 75A

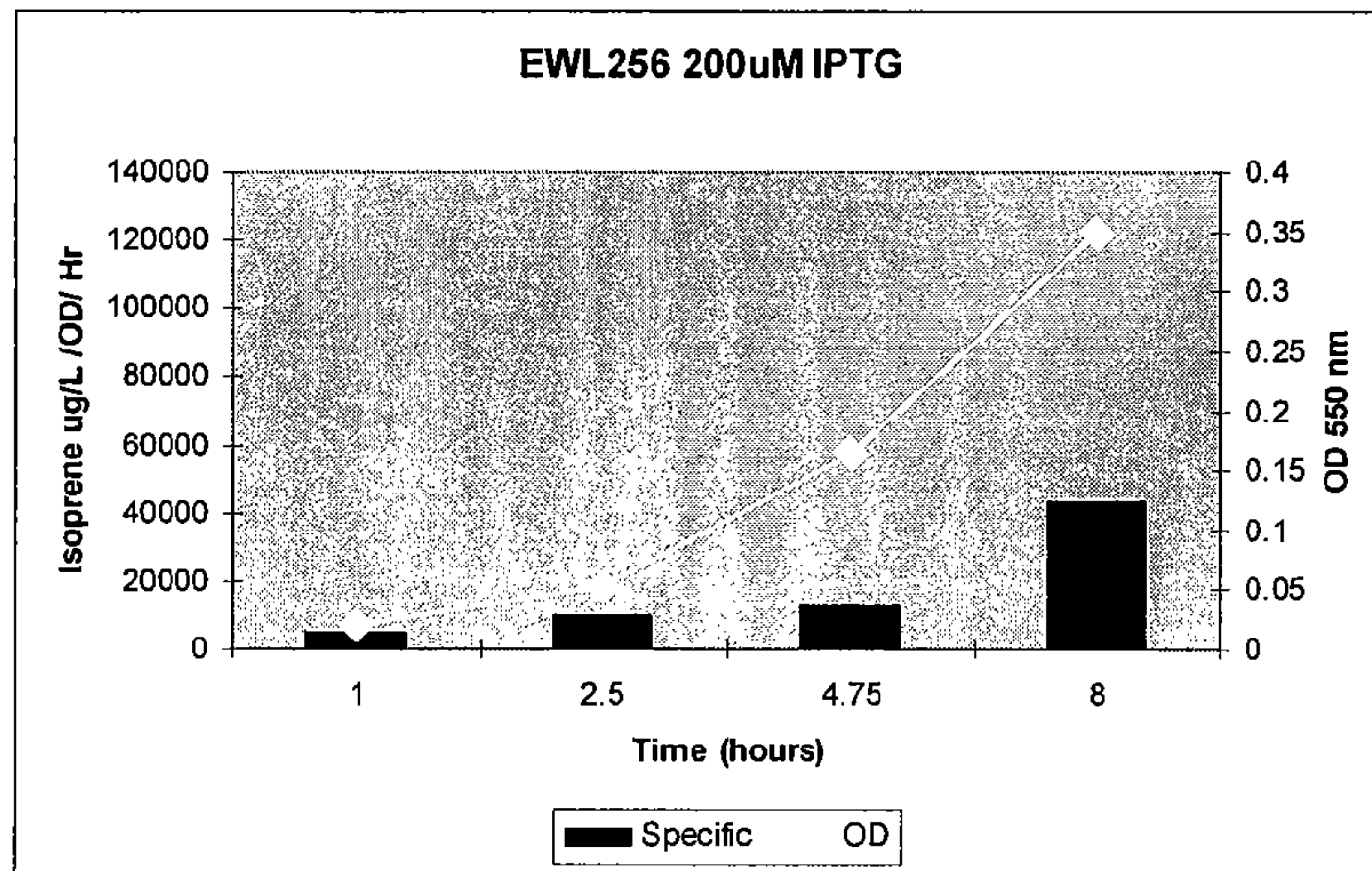
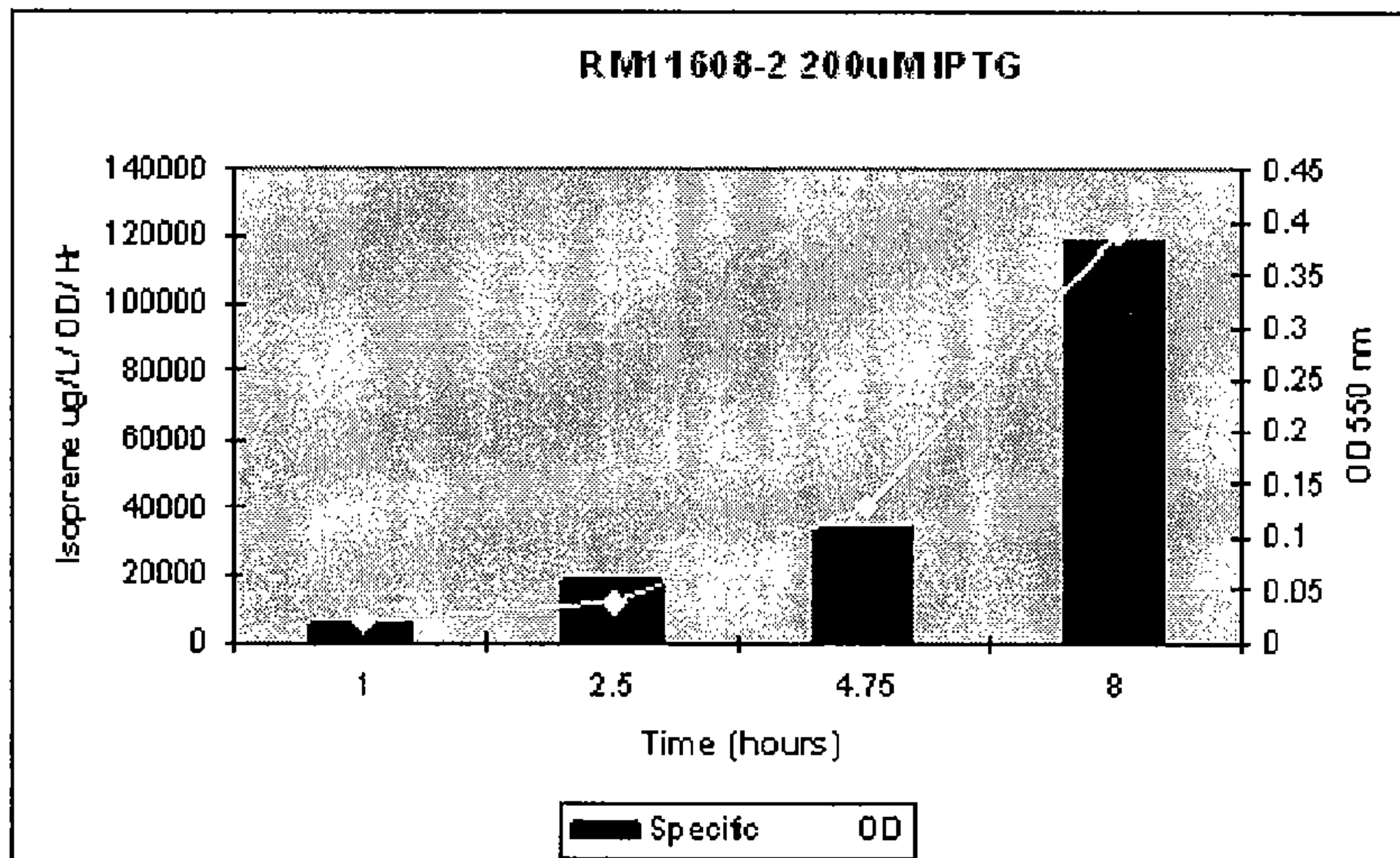


Figure 75B





157/167

Figure 75C

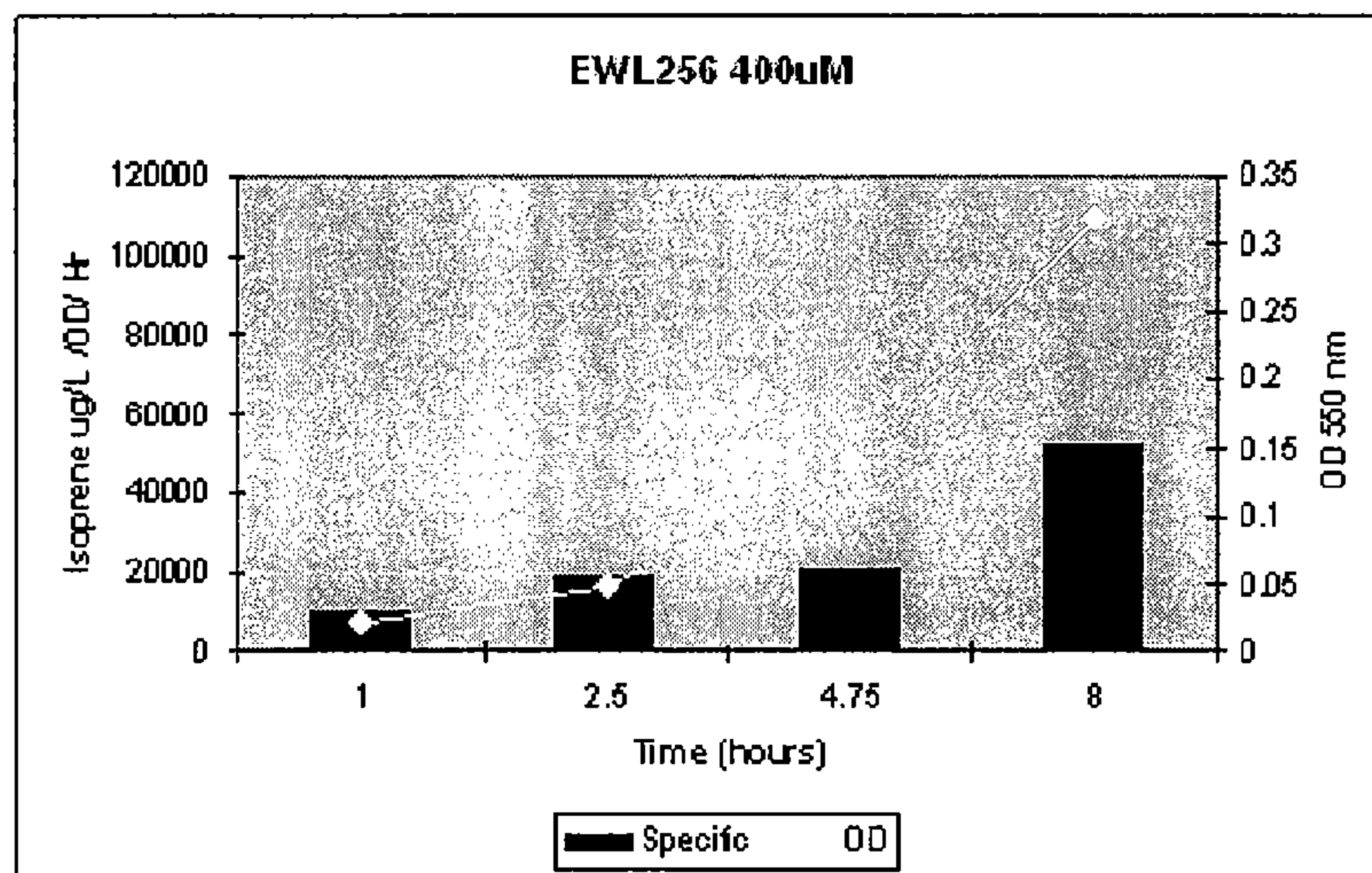
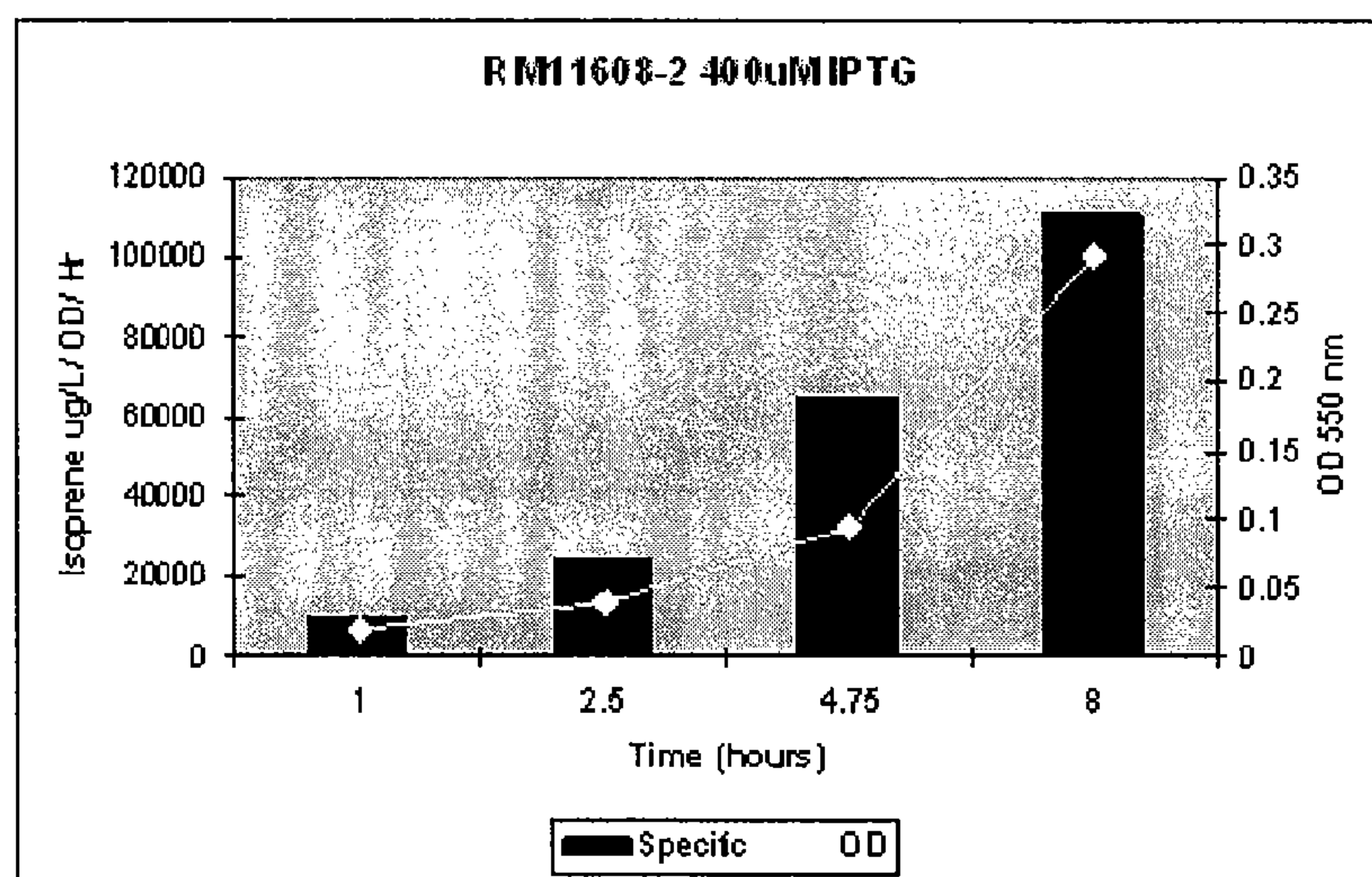


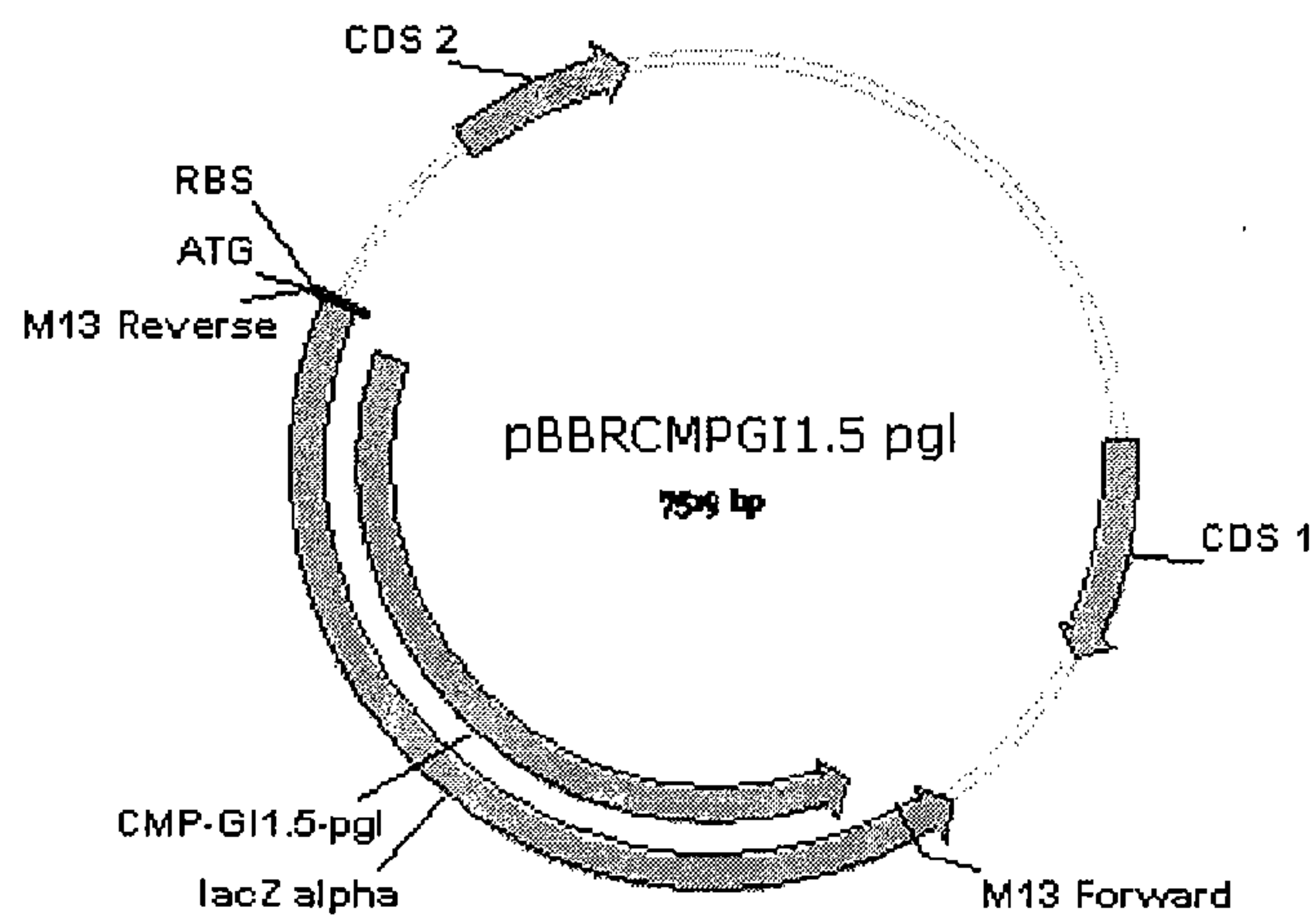
Figure 75D





158/167

Figure 76



CDS 2: Gentamycin resistance gene; CDS1: E. coli replication protein.



159/167

Figure 77A

1-

ctcgggcccgtctcttgggcttgatcgcccttcttgcgcattcacgcgtcctgcggcggcctgtagggcaggctcataccctgccgaacc  
gcttttgcagccggctcgccacggcttccggcgtctcaacgcgtttgagattcccagcttttcggccaatccctgcgggtcataggcgcgt  
ggctcgaccgcttgcgggctgatggtgacgtggcccaactggtggcgcctccagggcctcgtagaacgcctgaatgcgcgtgtgacgtgc  
cttgcctgcctcgatgccccgttgcagccctagatcgccacagcggccgcaaactggtctggctgcgggtcatctgcgtttgttgcga  
tgaactccttggccgacagcctgccgtcctgcgtcagcggcaccacgaacgcggctcatgtgcgggctggttcgtcacgggtggatgctggc  
cgtaacgatgcgacccgccccgtacttgcgcgcagccacttgcgccttctcgaagaacgcgcctgctgttcttggctggccgacttcca  
ccattccgggctggccgtcatgacgtactgcaccgccaacacagcgtccttgcgcgccttctcgtgcagcaactcgcgcagtcggcccatc  
gcttcacgggtgctgctggccgcccagtgctcgttcttggcgtcctgctggcgtcagcgttgggcgtctcgcgcctcgcggtaggcgtgctt  
gagactggccgccacgttgcctatttgcagccttcttgcacgcgtatgcgcctatgcgcctatgcgccttcccttttgggtgtccaacc  
ggctcgacgggggcagcgcgaaggcgggtgcctccggcggggccactcaatgcttgagtatactcactagactttgcttcgcaaagtcgtgac  
cgctacggcgggtgcggcgcctacgggcttgcctcgggcttcgcctgcgcgtcgtgcgcctccttgcagcccgtggatatgtg  
gacgatggccgcgagcggccaccggctggctcgttcgctcggccgtggacaacctgctggacaagctgatggacaggctgcgcct  
gcccacgagcttgaccacagggttgcctaccggctaccagccttcgaccacataccaccggctccaactgcgcggcctgcggcctt  
gccccatcaattttttaatttcttggggaaaagcctccggcctgcggcctgcgcgttcgcttgcgggttgacaccaagtgaaggcgg  
gtcaaggctcgcgcagcgcaccgcgcagcggcttggccttgacgcgcctggaacgacccaagcctatgcgagtgggggcagtcgaagg  
cgaagcccgcgcctgcccccgagcctcagcggcggcagtgccgggggttccaagggggcagcgcacacttgggcaaggccgaag  
gcccgcgagtcgatcaacaagccccggagggggccacttttgcgggagggggagccgcgcgaaggcgtgggggaaccccgaggg  
gtgccccttcttgggcaccaaagaactagatatagggcgaaatgcgaaagacttaaaaatcaacaacttaaaaaaggggggtacgcaacag  
ctcattgcggcaccccccgcaatagctcattgcgtagggttaaagaaaatctgtaattgactgccacttttacgcaacgcataattgttgcgcg  
tgccgaaaagtgcagctgattgcgcgtggtgccgaaccgtgcggcacccctaccgcatggagataagcatggccacgcagtcagaga  
aatcggcattcaagccaagaacaagccccgtcactgggtgcaaacggaacgcaaagcgcgtgaggcgtggccgggcttattgcgagg  
aaaccacggcggcaatgctgctgcacacctcgtggcgcagatgggccaccagaacgcgcgtggtggcagccagaagacacttccaa  
gctcatcggacgttcttgcggacggtccaatacgcagtcaggacttgggtggccgagcgtggtatcctcgtcgtgaagctcaacggcccc  
ggcaccgtgtcggcctacgtgggtcaatgaccgcgtggcgtggggccagccccgcgaccagttgcgcctgtcgggtgtcagtcgcgcgt  
gggtggttgatcacgacgaccaggacgaatcgtgttggggcatggcgacctgcgcgcacccgacctgtatccgggcgagcagcaac  
taccgaccggccccggcgaggagccgcccagccagccccgcatccgggcagtggaaccagacctgccagccttgaccgaaacggag  
gaatgggaacggcgcgggcagcagcgcctgccgatccccgatgagccgtgtttctggacgatggcgagccgttgagccgcccagac  
gggtcacgctgcgcgcgggtagcacttgggttgcgcagcaacccgtaagtgcgctgttccagactatcggtgtagccgcctcgcgc  
ctataccttgtctgcctccccgcgttgcgtcgcgggtgcagtgagccggggccacctgcacctgaatggaagccggcggcacctcgtaacg  
gattcaccgttttatcaggctctgggaggcagaataaatgatcatatcgtcaattattacctccacggggagagcctgagcaaactggcctca  
ggcatttgagaagcacacgggtcacactgctccggtagtcaataaacgggtaaaccagcaatagacataagcggctatttaacgacctgcc  
ctgaaccgacgaccgggtcgaatttgccttgcgaatttctgccattcatccgcttattatcacttattcaggcgtagcaccaggcgttaaggga  
ccaataactgccttaaaaaaattacgccccgcctgccactcagcagtcggcctattggttaaaaaatgagctgatttaacaaaaattaac  
gcgaattttaacaaaatattaacgcttacaatttccattcaggtgcgcaactgttgggaaggcgatcgggtgcgggcctcttcgct  
attacgccagctggcgaaagggggatgtgctgcaaggcgattaagtgggtaacgccagggtttccagtcacgacgttgtaaacgacg  
gccagtgagcgcgcgtaatacactcactatagggcgaaattggagctccaccgcgggtggcggcgcctctagaactagtgatccccgg  
gctgcgtgcgcgagcggccagtgatgatactgcgaattcgccttcttgatatcttagtgtgcttaaccaccacccacattggtc  
cctgcccagccgcatagcggccttttcatgcagtagccccgtcgcgcaacaattcgtataccgagatgtggtgagatttttggccggcg  
caatcagatacttgcgcgtgatcaacattgaagccgcgcggctgggtttccgttggtggaagccttcttactcaacacgctgccatctc  
cgaaacgctgaaaacggtaacaggctggcgggtacggtgcagggctataaatggcgacctccgggggtgatgaatatcagccgccc  
aacgggtgtcggagaagtttccggcatcataatccagcgtctggacacattcgatattaccgtgcggatcttccagtcacacatccactga  
gctgttaactcattgacgcaatacgcataattgttcgttggatggaataccatatgacgcgggcccggcccttcaacgggtggtcacttccgca  
gggtcctgcgcacgagatgaccatcagcgtgaccgtaaacaggcaaatgcgatcctgtttaatgccggaaccacagcgtacggttgt



160/167

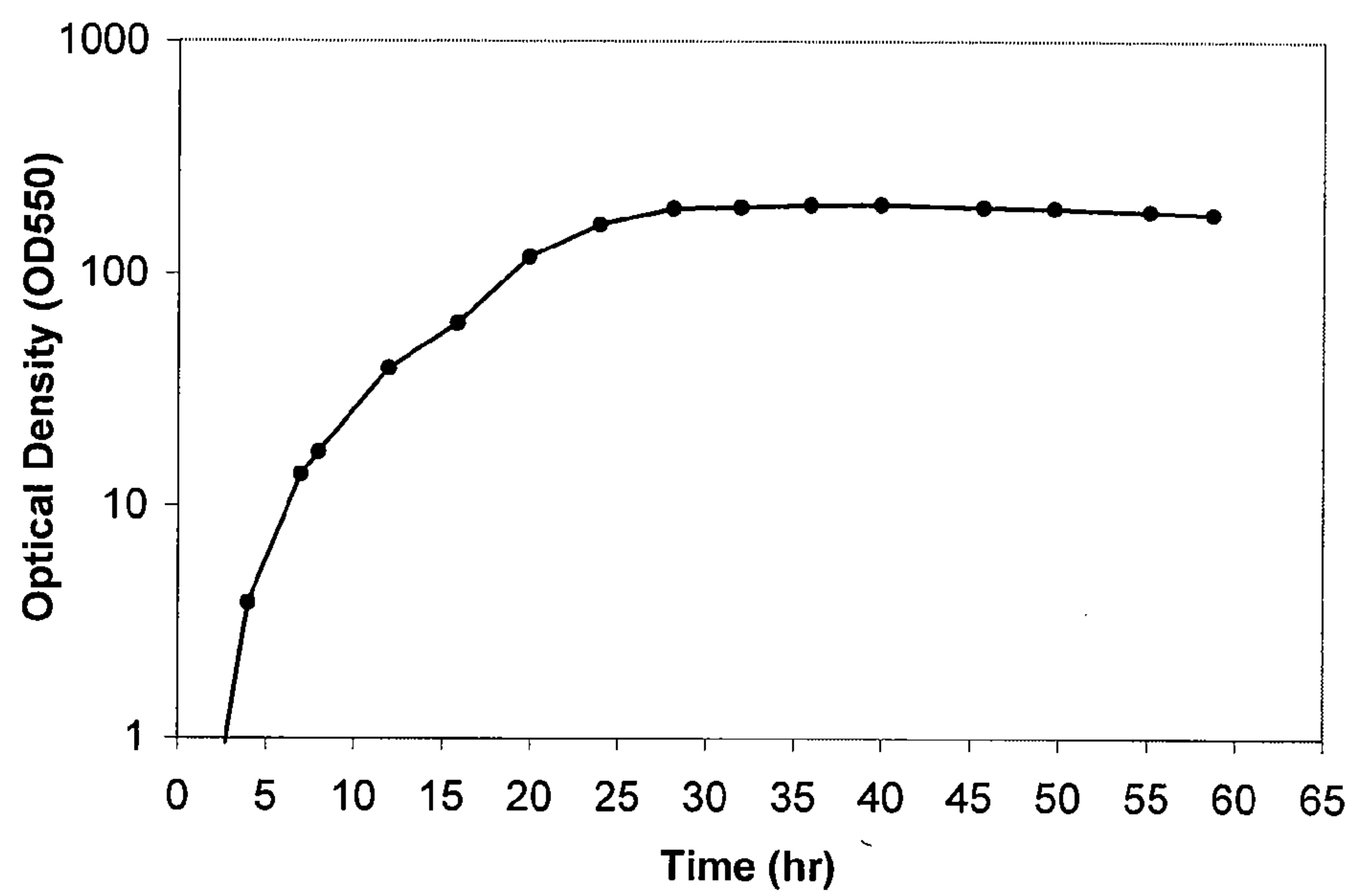
**Figure 77B**

ccggtgagatattggcggaatggcaaccgtccagcccctcgaccacatcgacgacgcccactggcaggccatcttcagacgcgttacgc  
tcacgttaccgcgcatgtgaagaacctacaaagacaaaactgcccctggtgatcggtggaaatatgcgtcggactaccgggcagcgcagactc  
tgccgcaaaggtcagtgcccatcgctcggggcgatacgatacgccaggacgcgaaactcagggcgaaaccaacatagagataacgt  
ttgtccgggctgaccaccatcggtgcacctgccccggcacatcgacaacctgtgtcagcgtcagtgccgcttcattgattcagattccagac  
gtgaatttgctggctctcagggctggcgatataaactgtttgcttcattgaatgctcctttgggttacctccgggaaacgcggttgatttggttagtg  
gtgaattatttgctcaggaatgtggcatagtcaggggcgtgacggctcgtaatacaactcactatagggtcgaggaagtccctatactttcta  
gagaataggaactccgcgcgcacacaaaaaccaacacacagatcatgaaaataaagctcttttattggtaccgaattcgccaggagct  
ctcagacgtcgttggtcggtctttattcgaaccccagagtcccgttacgccccgcccctgccactcatcgacgtactgttgtaattcattaagc  
attctgccgacatggaagccatcacaaacggcatgatgaacctgaatcgccagcggcatcagcaccttgctgccttgctgataatatttgccc  
atggtgaaaacggggggcgaagaagttgtccatattggccacgtttaaatcaaaactggtgaaactcaccagggttggtgagacgaaaa  
acatattctcaataaaccttttagggaaataggccaggtttcacgtaacacgccacatcttgcaatatatgtgtagaaactgccggaaatc  
gtcgtggtattcactccagagcgatgaaaacgttcagtttgcctcatggaaaacggtgtaacaagggtgaacactatcccatatcaccagctc  
accgtctttcattgccatacggaattccggatgagcattcatcaggcgggcaagaatgtgaataaaggccggataaaacttgcttattttct  
ttacggtctttaaaaaggccgtaatatccagctgaacggtctggttataggtagcactgactgaaatgcctcaaatgttctttacgat  
gccattgggatatacaacggtggtatatccagtattttttctccatggttagttcctcacttgctgattatactatgccgatactatgccg  
atgattaattgtcaacacgtgctgctgcaggtcgaaaggcccgagatgaggaagaggagaacagcgcggcagacgtgcgctttgaag  
cgtgcagaatgccgggctccggaggaccttcggggcgcccgccccgccccgagcccgccccgagcccgccccggaccaccctt  
cccagcctctgagcccagaaagcgaaggagcaaagctgctattggccgctgccccaaaggcctaccgcttcattgctcagcgggtgctg  
tccatctgcacgagactagtgcagcgtgctacttccattgtcacgtcctgcacgacgcgagctgcggggcgggggggaactcctgacta  
ggggaggagtggaaaggtggcgcgaaggggccaccaagaacggagccggttggcgctaccggtggatgtggaatgtgtgcgaggc  
cagaggccacttggtgtagcgccaagtcccagcggggctgctaaagcgcatgctccagactgccttgggaaaagcgctcccctaccg  
gtagaatgaagttcctatactttctagagaataggaacttcgcggccgccctttagttaggggttaattcaactgactgtaacagctaaaattagt  
cgcttttggcggaaggggcgaattccagcacactggcgggcgttactagtggatccgagctcggtaccaagcttgatgcaggaattcgatat  
caagcttatcgataccgtcgacctcgagggggggggcccggtaccagctttgttccctttagttaggggttaattgcgcgcttggcgtaatcatg  
gtcatagctgttccctgtgtgaaattgttatccgctcacattccacacaacatacgagccggaagcataaagtgtaaagcctgggggtgcctaa  
tgagtgcgtaactcacattaattgcgttgcgctcactgcccgtttccagtcgggaaacctgtcgtgccagctgcattaatgaatcgcccaac  
gcgcgggggagaggcggttgcgtattggcgcatgataaaaactgttgtaattcattaagcattctgccgacatggaagccatcacaaacg  
gcatgatgaacctgaatcgccagcggcatcagcaccttgctgccttgctgataatatttgccatggacgcacaccgtggaaacggatgaa  
ggcacgaaccagttgacataagcctgttcggttcgtaaaactgtaattgcaagtagcgtatgcgtcacgcaactggtccagaaccttgaccg  
aacgcagcgggtggtaacggcgagtggtgggtttcatggcttggtatgactgtttttgtacagtctatgcctcgggcatccaagcagcaagc  
gcgttacgccgtgggtcgatgtttgatgttatggagcagcaacgatgttacgcagcagcaacgatgttacgcagcagggcagtcgccctaa  
aacaaggttaggtggctcaagtatgggcatcattcgacatgtaggctcgccctgaccaagtcaaatccatgcgggctgctcttgatctttc  
ggctgtgagttcggagacgtagccacctactcccaacatcagccggactccgattacctcggaacttgctccgtagtaagacattcatgc  
gcttgctgccttcgaccaagaagcggttggtggcgctctcgcggttacgttctgccaggttgagcagccgcgtagtgcagatctatatctat  
gatctcgagctctccggcgagcaccggaggcagggcattgccaccgcgtcatcaatctcctcaagcatgaggccaacgcgcttggtgct  
tatgtgatctacgtgcaagcagattacggtgacgatcccgagtggtctctatacaaaagtgggcatacggggaagaagtgcactttgat  
atcgaccaagtaccgccacctaacaattcgttcaagccgagatcggttcccggccgagggttggttggttaattgtcacaacgccgcc  
aggtggcacttttcggggaaatgtgcgcgcccgcgttctgctggcgctgggcctgtttctggcgctggacttcccgtgttccgtcagcag  
cttttcgccacggccttgatgatcgcgggcgcccttggtctgcataatccgattcaacggccccaggcggtccagaacgggcttcaggcg  
tcccgaaggt (SEQ ID NO:122).



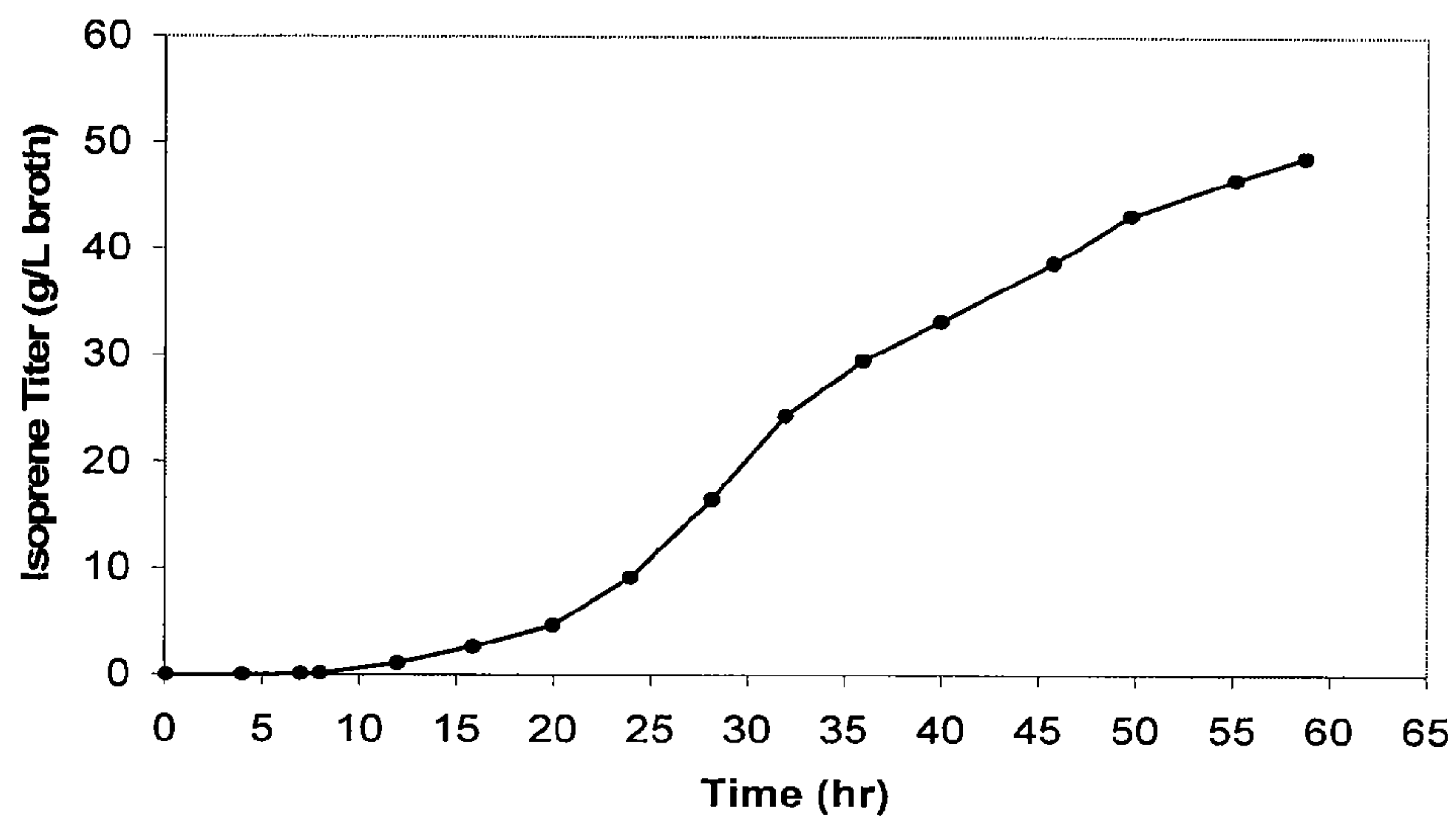
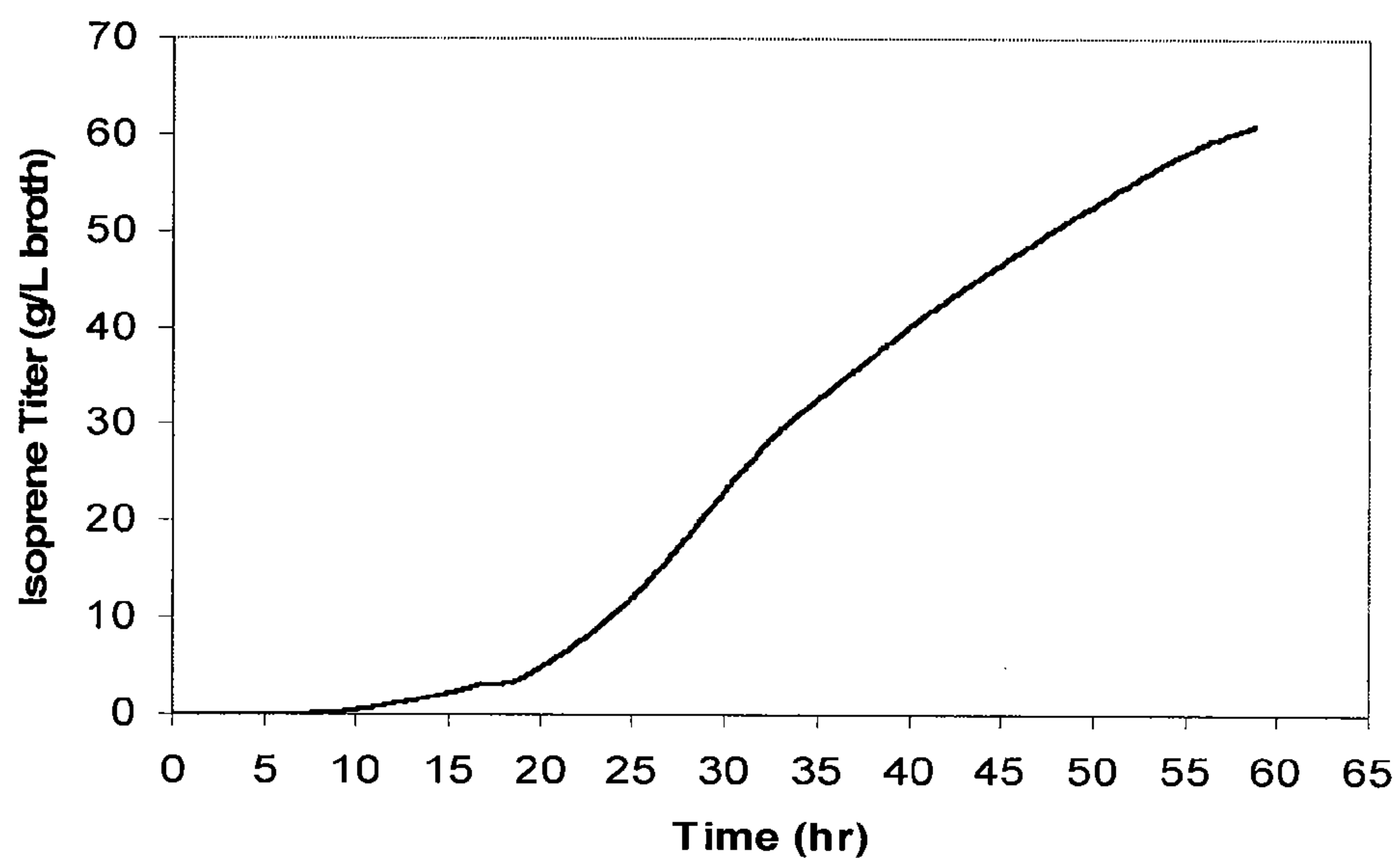
161/167

Figure 78A





162/167

**Figure 78B****Titer 78C**



163/167

Figure 78D

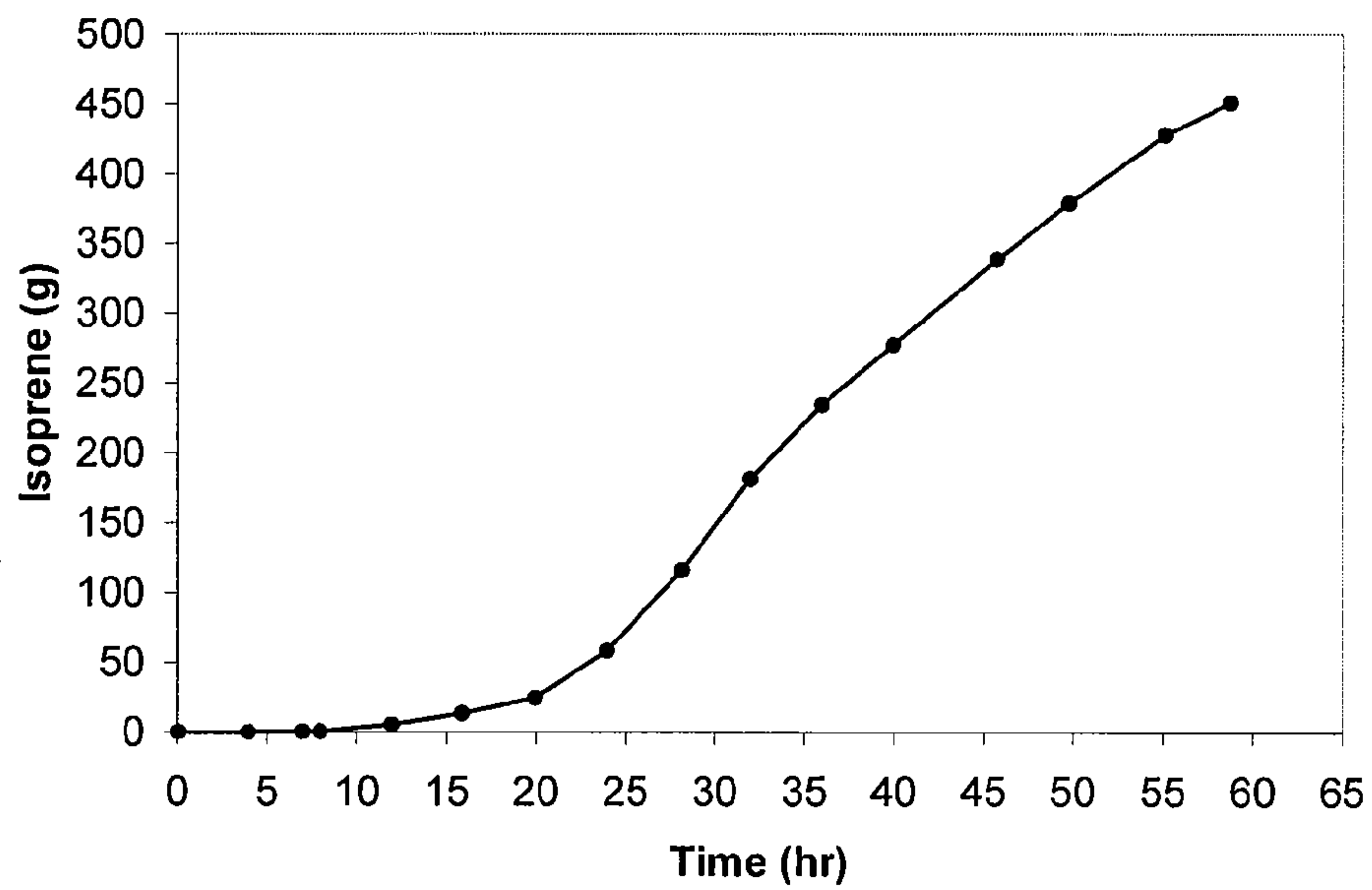
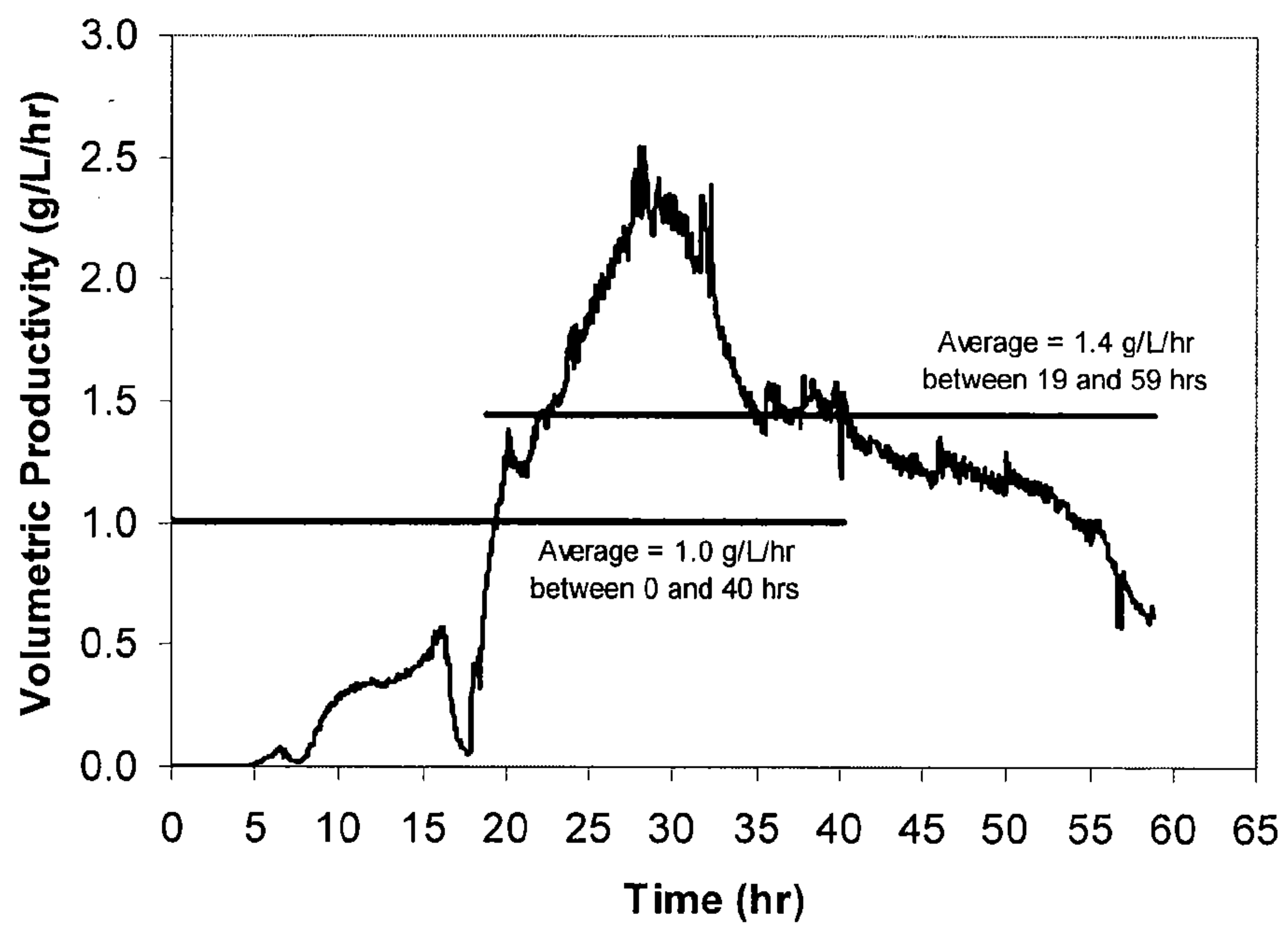


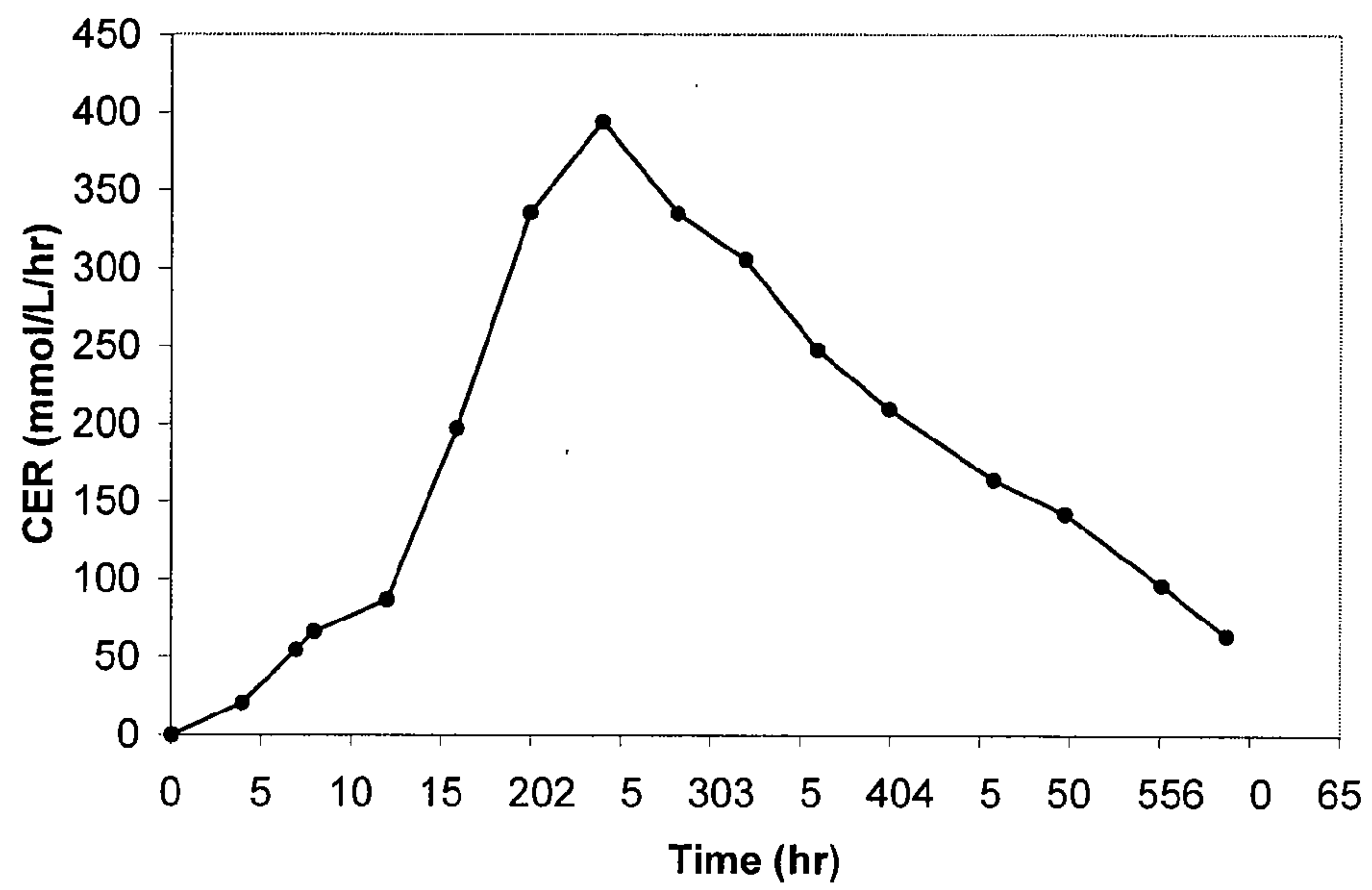
Figure 78E





164/167

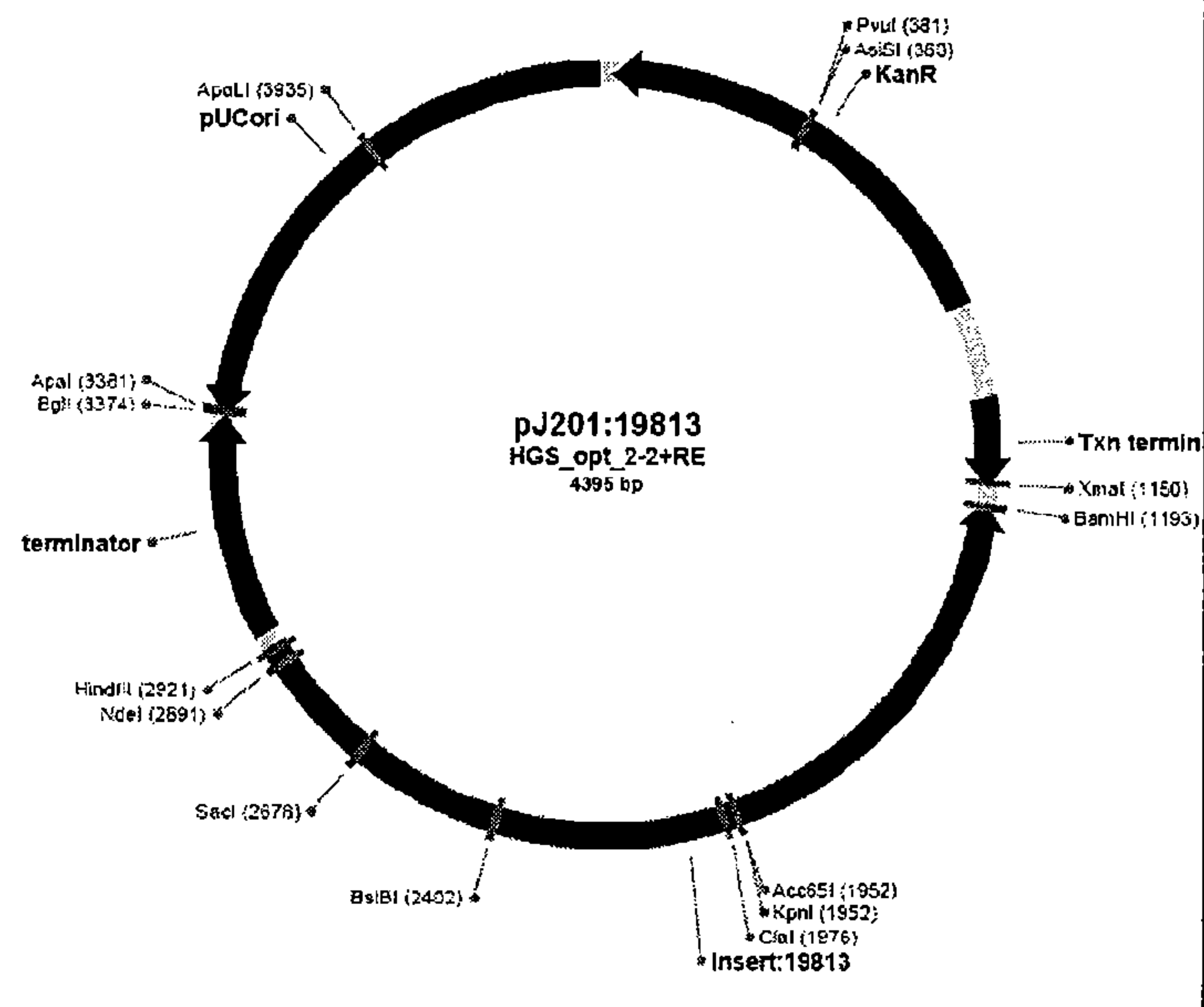
Figure 78F





165/167

Figure 79A





166/167

Figure 79B

5'-

tagaaaaactcatcgagcatcaaataaaactgcaatttattcatatcaggattatcaataccatattttgaaaaagccgtttctgtaatgaagga  
gaaaactcaccgaggcagttccataggatggcaagatcctggatcgggtcgcgattccgactcgtccaacatcaatacaacctattaattcc  
cctcgtcaaaaataaggttatcaagtgagaaatcaccatgagtgacgactgaatccgggtgagaatggcaaaagtttatgcatttcttccagac  
ttgttcaacaggccagccattacgctcgtcatcaaaatcactcgcacatcaacaaaccgttattcatcgtgattgcgcctgagcgaggcgaaat  
acgcgatcgtgtttaaaggacaattacaacaggaatcgagtgaaccggcgaggaacactgccagcgcatcaacaatatttcacctg  
aatcaggatatttcttaataacctggaacgctgttttccggggatcgagtggtgagtaaccatgcatcatcaggagtagcgataaaatgctt  
gatggtcggaagtggcataaattccgctcagccagtttagtctgacctctcatctgtaacatcattggcaacgctaccttggcatgtttcagaa  
acaactctggcgcatcgggcttccatacaagcgatagattgtcgacactgattgcccacattatcgcgagcccatttatacccatataaatc  
agcatccatgttggaatttaacgcggcctcgacgtttccggttgaatatggctcatatttctcttttcaatattattgaagcatttatcagggttat  
tgtctcatgagcggatacatattgaatgtatttagaaaaataaacaatagggggtcagtggttacaaccaattaaccaattctgaacattatcgcg  
agcccatttataacctgaatatggctcataacacccctgtttgcctggcggcagtagcgcggtgtgtcccacctgaccccatgccgaactcaga  
agtgaacgcccgtagcgccgatggtagtgtggggactccccatcgagagtagggaaactgccaggcatacaataaaacgaaaggctcag  
tcgaaagactgggaccttcgcccgggctaattaggggggtgtcgcccttcgattgacgggtacgggacctcacacgtacatcagctggttgat  
ggggaacgggtcgatgagcagcagcttgatgcgggttcggtggcgtaatccggggcgccagcccgtcccatattggttaggtgcagtg  
gctcacgcgggccatgttcacggcgatctccatgaacgccttcggcagcaggggtgctgtccgacacgcgctcgcggttcattttcttccact  
cggcgctcgatcagcttgccgagctcttcgcgggcctgttctcgtcgtcgctcgttcgtgcatgtagctgatgatgctgttggtggtttcg  
ccgcttcgagttccgcccaggtcgccagatcgttcacagccgaaagatcacgcaggacgagcgccaccaggccgtggaagtcgg  
tcaggagcgaggggcggtggtccgagatatcttctgctgtgagcagaccgagaagtagctcgggcgccagcagcgcgaccccgctgga  
ggacacgctggcggttctcaggtactgtctgaaggcggggatgatctgttattgtccacttggcttcttcaggaaggccttgacagttcg  
cgccagcttttggtcagatagctcaggttattgtggcccttctcctcaggatggagtaggacgtgtcgttcacgggtgtgtacagggccagga  
agcacagcttcataatagtcgggcagcgtgttgatggcggtcacgtcccagcgttccaccgcgtcggtgaagagctgcagttcgtccagggtta  
ccgtacacgtcatagacgtcatcgataatggtgaccagaccgaacatcttggtgacggccttgcggcattcgccgaactgcgggtccggcg  
ccatgccagcgcccagaagtacacttccatcaggcggtcccgcacgaaatccagcttgctggcgaggcccatctcggtccaccaccgg  
ctcagggtctcgagctcttttgggtgcaggggtctggaccatgttgaatcgagtttgccagttccagcagcagctggtgatgcggctccttgg  
gttcgtactgtccagaaaccaccgcgctccagggcggtgcagggcgttgatgatacggcagctccagcgcggtgggacacctgctcgcca  
ccttcgtgttgatccccctcttgaggtgttcttcagatgggtgatgctgaaggtacgggcctcctcagcagatttctgccttcgaaaccgaga  
tagctggcctcgtacagggtcagcagggcctgcacgtcaccttcagttccccggagaagcccccttcttgccttgaagcgctcgaacac  
gtcctggctcacctcaaagccatgctgccgcagcagggcggaagctcagggcggtcgcggtgcagatcgctttgttcttcttattctcgtccag  
caggacgatgttctcagcgcttgatgatatcttctcaaatgttaggtcaggcccaggcgctgcacgtcgtcgatgagctccagcaggt  
caggggctgggtgtccaccgggtgatcatgaacgcacctctctcctcagcttggtggccttctcttcagcttctccaccttcaggtcgtttt  
ccaggctctgcaggaactcgaagtccacaggttgggctggtagtgcggaccgacggctattatgctcggtgatctgggtgaactggctg  
ctggtggcgacatatgtatatctcttcttaaagttaaacaagcttaagatgttcagcgacaagggcgacacaaaatttattctaaatgcataat  
aaatactgataacatcttatagtttgtattatatttgtattatcgttgacatgtataattttagatatcaaaaactgatttcccttattatttcgagattta  
tttcttaattctctttaaacaactagaaatattgtatatacaaaaaatcataaataatagatgaatagtttaattataggtgttcatcaatcgaaaaag  
caacgtatcttatttaaagtgcgttgctttttctcatttataagggttaaataatttctatatacaagcaaaagtgcagggcgcccttaaattctga  
caaatgctctttccctaaactcccccataaaaaaacccgccgaagcgggttttacgttatttgcggattaacgattactcgttatcagaaccg  
cccagggggcccgagcttaagactggcgctggtttacaacacagaaagagttttagaaacgcaaaaaggccatccgctcaggggccttct  
gcttagttgatgcctggcagttccctactctcgcttccgcttctcgtcactgactcgtcgctcggtcgttcggctcgggcgagcggtat  
cagctcactcaaaggcggttaatacgggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggc  
caggaaccgtaaaaaggcggtgtgctggcggttttccataggtccgccccctgacgagcatcaaaaaatcgacgctcaagtcagag  
gtggcgaaaccgcagagactataaagataaccaggcggtttccccctggaagctccctcgtcgctctctgttccgacctgccgcttaccg  
gatacctgtccgcttctccttcgggaagcgtggcgcttttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaag  
ctgggctgtgtgcacgaacccccgttcagccgaccgctgcgcttattccggtaactatcgtcttagtccaacccggtgaagacacgactt



167/167

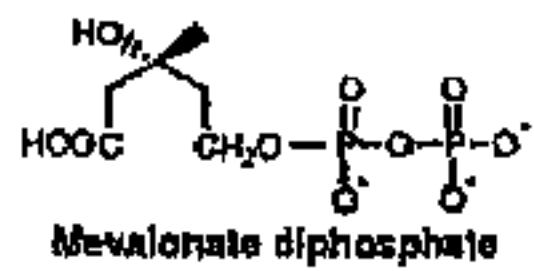
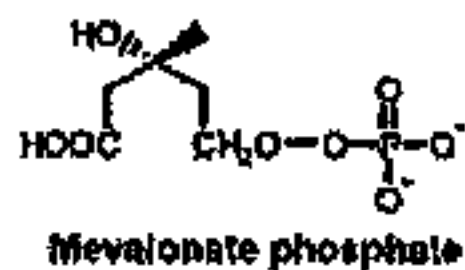
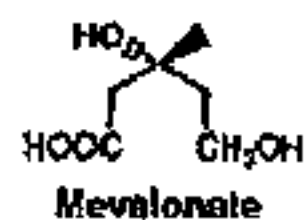
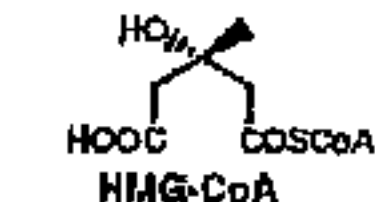
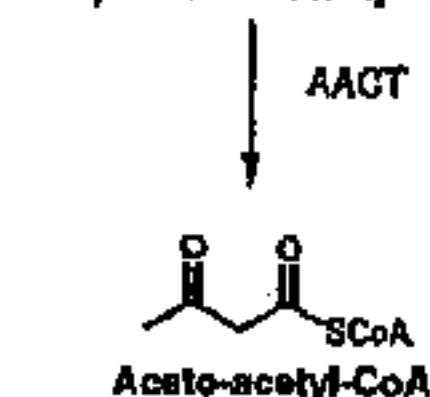
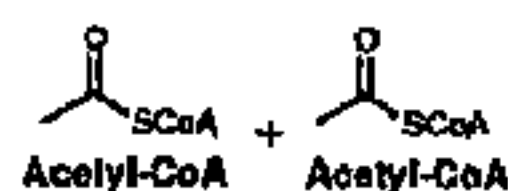
**Figure 79C**

atgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtgggtgggctaactac  
ggctacactagaagaacagtatttggatatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaca  
accaccgctggtagcgggtgggtttttgtttgcaagcagcagattacgcgcagaaaaaaggatctcaagaagatcctttgatctttctacgg  
ggtctgacgctcagtggaaacgacgcgcgcgtaactcacgttaagggattttggatcatgagcttgcgccgtcccgtaagtcagcgtaatgct  
ctgcttt (SEQ ID NO:123)



Figure 19A

**MVA PATHWAY**



**DXP PATHWAY**

