

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 May 2010 (14.05.2010)

PCT

(10) International Publication Number
WO 2010/052335 A1

(51) International Patent Classification:

C07K 14/61 (2006.01) **C12N 1/20** (2006.01)
C12N 1/16 (2006.01)

(21) International Application Number:

PCT/EP2009/064903

(22) International Filing Date:

10 November 2009 (10.11.2009)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

08168750.1 10 November 2008 (10.11.2008) EP

(71) Applicant (for all designated States except US): **NOVO NORDISK A/S** [DK/DK]; Novo Allé, DK-2880 Bagsværd (DK).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **WÖLDIKE, Helle Fabricius** [DK/DK]; Novo Nordisk A/s, DK-2760 Måløv (DK). **JIANHUI, Deng** [CN/CN]; Novo Nordisk R/d Center China, 1st Fl, Tower B no. 29, Life Science Park Road, Changping District, Beijing 102206 (CN). **XIN, Zhao** [CN/CN]; Novo Nordisk R/d Center China, 1st Fl, Tower B no. 29, Life Science Park Road, Changping District, Beijing 102206 (CN). **YUN, Liu** [CN/CN]; Novo Nordisk R/d Center China, 1st Fl, Tower B no. 29, Life Science Park Road, Changping District, Beijing 102206 (CN). **JING, Su** [CN/CN]; Novo Nordisk R/d Center China, 1st Fl, Tower B no. 29, Life Science Park Road, Changping District, Beijing 102206 (CN).

(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, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, 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, PE, 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.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, 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, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: E. COLI BL21 STRAIN LACKING A FUNCTIONAL GROUP II CAPSULAR GENE CLUSTER

(57) Abstract: The present invention relates to a novel non-pathogenic Escherichia coli (E. coli) B BL21 strain comprising a deletion of group II capsular gene cluster, and the use thereof for the production of peptides.



WO 2010/052335 A1

E. COLI BL21 STRAIN LACKING A FUNCTIONAL GROUP II CAPSULAR GENE CLUSTER**FIELD OF THE INVENTION**

The present invention relates to the manipulation of *Escherichia coli* (*E. coli*) genome, to an improved non-pathogenic *E. coli* B strain, and its use thereof for the production of peptides and proteins.

BACKGROUND OF THE INVENTION

Bacteria have been used to produce a wide range of commercial products. The use of genetically engineered bacteria like *E. coli* as host cells for the production of biological reagents such as peptides and nucleic acids in the laboratory and industry, is well-known in the art. Bacteria in their natural environment are exposed to many conditions that are not normally experienced in standard industrial or laboratory growth, and thus their genomes carry a large number of condition-dependent, stress-induced genes or otherwise non-essential genes which may not be needed in industrial or laboratory use of the organisms, or may even be unwanted.

Bacterial pathogens contain virulence factors such as capsular polysaccharides at the cell surface, which mediate the interactions between the bacterium and its immediate environment and have a number of functions directly related to their pathogenicity. More than 80 serologically and chemically distinct types of polysaccharide capsules have been described in *E. coli* and termed K antigens. The polysaccharide capsules have been classified into 4 groups, of which the group II capsular polysaccharides are similar to those of *Haemophilus influenzae* and *Neisseria meningitidis*. The group II capsular gene clusters contain most of the capsular types associated with invasive disease and have a conserved genetic organisation consisting of three functional regions. Regions 1 and 3 are conserved among group II gene clusters and encode peptides necessary for the transport of the polysaccharide from its site of synthesis to the cell surface, while region 2 is serotype specific and encodes the enzymes responsible for biosynthesis and polymerization of the individual monosaccharides that comprise the particular polysaccharide. Region 1 comprises 6 genes *kpsFEDUCS*, organised in a single transcriptional unit. Region 2 is serotype specific and differs among group 2 antigens. Region 3 contains 2 genes *kpsM* and *kpsT* organised in a single transcriptional unit [Andreishcheva et al. Gene (2006) 384: 113-119].

The *E. coli* strains K-12 and B are not encapsulated and are commonly used strains for the production of peptides. The *E. coli* K-12 genome has been fully sequenced and does not contain any capsular genes. It has been found that *E. coli* B BL21(DE3), which is a pre-

ferred strain for high-yield expression of recombinant peptides, contains a cluster of genes coding for group II capsular polysaccharides and it is thought that an ancestor of *E. coli* B must have acquired these capsular genes by horizontal transfer. The K antigens located in the capsule of *E. coli* are often associated with a pathogenic phenotype in strains harbouring the functional form of the genes. It has been shown that in *E. coli* B BL21, the region 2 of group II capsular gene cluster has become inactive by an insertion element (IS1), preventing the production of the antigen, while region 1 and 3 supporting export and presentation of wildtype capsular antigen, are still intact [Andreishcheva et al. Gene (2006) 384: 113-119]. However, there is a theoretical possibility of re-establishing a functional group II capsular gene cluster, e.g. if the IS1 in region 2 is lost and thereby the group II cluster may be re-established. Alternatively, if the IS1 mutated region is exchanged by an exogenous wildtype region 2, the group II cluster may also be re-established.

The present invention provides an *E. coli* B BL21 (DE3) strain with permanent knock out (KO) of the group II capsular gene cluster, bringing this strain into the same safety class as other commercially important strains such as *E. coli* K-12.

SUMMARY OF THE INVENTION

The present invention relates to an *E. coli* B BL21 strain having an irreversibility inactivated group II capsular gene cluster.

In one embodiment *E. coli* B BL21 strain is characterised by having a total or partial deletion of the group II capsular gene cluster.

In one embodiment *E. coli* B BL21 strain is characterised by having total deletion of the group II capsular gene cluster.

In one embodiment *E. coli* B BL21 strain is characterised by having a partial deletion of the group II capsular gene cluster.

In a further embodiment, the *E. coli* B BL21 strain is characterised by having one or more point mutations in the group II capsular gene cluster.

In another embodiment, the *E. coli* strain is characterised by having at least 5% of group II capsular gene cluster permanently inactivated by deletion.

In another embodiment, the *E. coli* strain is characterised by having at least 25% of group II capsular gene cluster permanently inactivated by deletion.

In another embodiment, the *E. coli* strain is characterised by having at least 40% of group II capsular gene cluster permanently inactivated by deletion.

In another embodiment, the *E. coli* strain is characterised by having at least 55% of group II capsular gene cluster permanently inactivated by deletion.

In another embodiment, the *E. coli* strain is characterised by having from 75% to 100% of the group II capsular gene cluster permanently inactivated by deletion.

In another embodiment, the *E. coli* strain is characterised by having at least 85% of group II capsular gene cluster permanently inactivated by deletion.

5 In another embodiment, the *E. coli* strain is characterised by having at least 95% of group II capsular gene cluster permanently inactivated by deletion.

In one embodiment, 30% of the group II capsular gene cluster is deleted.

In another embodiment, 50% of the group II capsular gene cluster is deleted.

In another embodiment, 70% of the group II capsular gene cluster is deleted.

10 In another embodiment, 80% of the group II capsular gene cluster is deleted.

In another embodiment, 90% of the group II capsular gene cluster is deleted.

In one embodiment, the *E. coli* strain is *E. coli* B BL21(DE3) and in a further embodiment the *E. coli* strain is *E. coli* B BL21(DE3) Δ *dadX* Δ *alr*.

The present invention also contemplates a method for producing an *E. coli* B strain
15 wherein the group II capsular gene cluster is deleted by a method comprising the steps of:

- a. replacing of group II capsular gene cluster by a selection marker,
- b. eliminating the selection marker and
- c. confirming the deletion by PCR analysis.

In one embodiment, the selection marker is an antibiotic resistance gene.

20 In another embodiment, the antibiotic resistance gene is selected from the group consisting of chloramphenicol, tetracyclin, ampicillin, kanamycin, vancomycin and erythromycin.

In yet another embodiment, the antibiotic resistance gene is chloramphenicol.

The invention is also related to a method for producing a peptide by introducing an
25 expression vector comprising a DNA sequence encoding a peptide in the *E. coli* strain of the present invention.

In one embodiment the present invention also contemplates a method for the production of a peptide comprising culturing an *E. coli* B strain according to the present invention, in a suitable culture medium under condition for the expression of DNA encoding the
30 said peptide. The peptide to be expressed by the expression vector may be any peptide which can be expressed in *E. coli*. Representative examples of such peptides are hGH, glucagon-like peptides, interleukins, insulin analogues, wildtype adiponectin, FGF-21, trypsin, aprotinin, amylin and leptin and analogs thereof, as well as enzymes such as sialidase, transglutaminase (tGase), HRV (Human Rhino Virus) 3C protease, Tobacco Etch Virus
35 (TEV) protease and variants thereof.

In one embodiment, the *E. coli* strain of the present invention comprises an expression vector comprising a DNA sequence encoding a peptide.

In one embodiment, the peptide to be produced is hGH.

In another embodiment, the peptide to be produced is FGF-21.

5 In another embodiment, the peptide to be produced is amylin.

In another embodiment, the peptide to be produced is leptin.

In another embodiment, the peptide to be produced is sialidase.

DESCRIPTION OF THE DRAWINGS

Fig.1. shows the genetic organisation of the *E. coli* B BL21(DE3) capsular gene cluster with non-coding flanking regions. Open reading frames (ORFs) of the B BL21(DE3) capsular gene cluster (large arrows); direction of transcription (arrowheads): region 1 genes (grey arrows); region 2 ORFs (black arrows); region 3 genes (white arrows); IS1 element (white box). [Andreishcheva et al. Gene (2006) 384: 113-119]. Small arrows indicate the position of the non-coding left arm (LA) and the right arm (RA).

15 Fig.2. is a schematic representation of the workflow of the group II capsular gene cluster knock out method: a) replacement of group II capsular gene cluster by an antibiotic resistance gene and b) elimination of the antibiotic resistance gene.

Fig.3. shows PCR screening of knock out candidates. The right size of 1686bp (LA+part of cat, A1-A3 in upper panel, B1-B3 in lower panel) and 1578bp (RA+part of cat, A1-A3 in upper panel, B1-B3 in lower panel) amplified from knockout candidates (but not its parental strain) (BL or D) indicate a successful knock out of group II gene cluster. For the negative control, only the parental strain has a band with size around 3kb but not group II knock out strain. BL standard for BL21(DE3), D standard for BL21(DE3) Δ dadX Δ alr strain. A1-A3 are different colony numbers of BL21(DE3) knock out candidates and B1-B3 are knock out candidates of BL21(DE3) Δ dadX Δ alr strain.

Fig.4. is a comparison of the growth of host *E. coli* B strains in shake flask. The symbols used refer to the following strains: ●, BL21(DE3); □, BL21(DE3) Δ (kpsM-kpsF); —, BL21(DE3) Δ dadX Δ alr; Δ , BL21(DE3) Δ dadX Δ alr Δ (kpsM-kpsF).

Fig.5. shows the fermentation growth curve for the different *E. coli* B strains. The symbols used refer to the following strains: ♦, BL21(DE3); □, BL21(DE3) Δ (kpsM-kpsF); —, BL21(DE3)/pNNC1-g; ○, BL21(DE3) Δ (kpsM-kpsF)/pNNC1-g.

Fig.6. shows a SDS-polyacrylamide gel electrophoresis (PAGE) of hGH expression at different induction times: a) expression in *E. coli* B BL21(DE3)/pNNC1-g and b) expression in *E. coli* B BL21(DE3) Δ (kpsM-kpsF)/pNNC1-g. In both a) and b), lanes 1-3 show different

concentrations of hGH: 5µg, 2.5µg and 1.25µg, respectively. Lane 4 shows expression of hGH before induction and lanes 5-14 show the expression of hGH at different induction times from 1-10 hours, every hour. All the samples were normalized to the same loading OD.

Fig.7. shows the growth of the different *E. coli* B strains before and after induction of hGH expression. The symbols used refer to the following strains: ♦: BL21(DE3)Δ*dadX*Δ*alr*/pNNC1-1h; □: BL21Δ*dadX*Δ*alr*Δ(*kpsM-kpsF*)/pNNC1-1h; Δ, BL21(DE3)Δ*dadX* Δ*alr*/pNNC1-1h induced; —, BL21(DE3)Δ*dadX*Δ*alr*Δ(*kpsM-kpsF*)/ pNNC1-1h induced.

Fig.8. shows the expression of hGH analyzed by SDS-PAGE at different induction times: a) *E. coli* B BL21(DE3)Δ*dadX*Δ*alr* /pNNC1-1h; b) *E. coli* B BL21(DE3)Δ*dadX*Δ*alr*Δ(*kpsM-kpsF*) /pNNC1-1h. In both a) and b), lanes 1-3 show expression of hGH at different concentrations, 5µg, 2.5µg and 1.25µg, respectively. Lane 4 shows expression of hGH before induction and lanes 5-13 show the expression of hGH at induction times from 1-9 hours, every hour. All samples were normalized to the same loading OD.

Fig.9. shows the growth curve of *E. coli* B BL21(DE3)Δ*dadX*Δ*alr*Δ(*kpsM-kpsF*) /pNNC1a.23 in a high cell-density fermentation in FDM1-2 medium. Expression of hGH mutant, MEAE-hGH(L101C), was induced by adding IPTG to a final concentration of 0.2mM when the cell density reached 70 OD₆₀₀ and continued to cultivate at 25°C. The final cell density reached approximately 100 OD₆₀₀ after 8 hours' induction.

Fig.10. shows expression of MEAE-hGH(L101C) during high-cell density fermentation. Lane 1: marker; lanes 2-4: purified hGH as the standard at concentrations of 5, 2.5, 1.25µg, respectively; lanes 5-13: samples taken at different time intervals after being induced for 0,1,2,3,4,5,6,7 and 8 hours, respectively. Arrow indicates the target protein, MEAE-hGH(L101C).

Fig.11. shows SDS-PAGE analysis of fermentation samples from expression of *Arthrobacter ureafaciens* sialidase in *E. coli* B BL21(DE3)Δ(*kpsM-kpsF*). Lane 1: marker; lane 2: pre-induction, homogenate; lane 3: post-induction, homogenate; lane 4: final sample, homogenate; lane 5: pre-induction, insoluble fraction; lane 6: post-induction, insoluble fraction; lane 7: final sample, insoluble fraction; lane 8: pre-induction, soluble fraction; lane 9: post-induction, soluble fraction; lane 10: final sample, soluble fraction; lanes 11-12: sialidase standard (150 and 300 mg/l). All samples were 10x diluted.

Fig.12. shows expression of FGF21, amylin and leptin all from *H.sapiens* in *E. coli* B BL21 DE3 and in *E. coli* B BL21 DE3Δ(*kpsM-kpsF*). Lane 1: marker; lanes 3-8: expression in BL21 DE3; lanes 9-14: expression in BL21 DE3 Δ(*kpsM-kpsF*); lanes 3 and 9: uninduced; Lanes 4 and 10: induced FGF21; lanes 5 and 11: uninduced; lanes 6 and 12: induced

amylin; lanes 7 and 13: uninduced; lanes 8 and 14: induced leptin.

DESCRIPTION OF THE INVENTION

The present invention relates to an *E. coli* B BL21 strain with an irreversibly inactivated group II cluster providing a permanently safe strain for the production of peptides.

5 The impairment or inactivation of the group II capsular gene cluster of the wildtype or any other strain of *E. coli* B BL21(DE3) may be accomplished by any suitable technique including deletion of the whole cluster or a part thereof or by one or more point mutations in the gene cluster.

10 The *E. coli* strain of the present invention has the advantage of providing a more robust and high yielding process for production of peptides, compared to commonly industrially used strains, such as *E. coli* K-12.

15 The *E. coli* B BL21(DE3) used in the present invention may be a double knockout (2xKO) as described in WO2008/025744. This *E. coli* B BL21(DE3) strain has a deletion of 2 genes *alr* and *dadX* which both code for D,L-alanine racemase. This double knock out strain cannot grow without either the addition to the growth medium of D-alanine or without the presence of one of the racemase genes on a plasmid, which can supply the necessary D-alanine. This 2xKO D-alanine auxotrophic *E. coli* strain, referred to herein as *E. coli* B BL21(DE3) $\Delta alr\Delta dadX$, enables a non-antibiotic selection procedure for maintenance of plasmids in the cells.

20 In one embodiment of the present invention, one or more genes within the group II capsular gene cluster of the wildtype *E. coli* B BL21(DE3) or *E. coli* B BL21(DE3) $\Delta alr\Delta dadX$ strain are deleted. The *E. coli* strain B BL21 (DE3) is described by Studier and Moffat, [J.Mol.Biol. (1986) 189:113-130] and Studier et al. [Methods in Enzymology (1990) 185:60-89] and is available from Stratagene and Novagen.

25 In one embodiment of the present invention, one or more genes within the group II capsular gene cluster of the wildtype *E. coli* B BL21(DE3) are deleted. In another embodiment of the present invention, one or more genes within the group II capsular gene cluster of *E. coli* B BL21(DE3) $\Delta alr\Delta dadX$ are deleted.

30 In another embodiment, the gene or genes of choice for deletion within the group II capsular gene cluster are those responsible for potential pathogenicity. Deletion of the entire 16kb group II capsular gene cluster provides the prevention of the potential re-establishment of the pathogenic capsular phenotype. Figure 1 shows the genes involved in group II capsular gene cluster of capsular strains including region 1 genes *kpsFEDUCS*, region 2 ORF1-4

and region 3 *kpsM* and *kpsT*. Any one of these genes or combinations thereof may be deleted.

The total deletion of the group II capsular gene cluster, means deletion of the 3 regions, ie., $\Delta(kpsM-kpsF)$. The *E. coli* B strains obtained are *E. coli* B BL21(DE3) $\Delta(kpsM-kpsF)$ and *E. coli* B BL21(DE3) $\Delta alr\Delta dadX \Delta(kpsM-kpsF)$.

In one embodiment, the present invention contemplates any modification sufficient to inactivate the group II gene cluster, apart from the insertion element (IS1).

Bacteria are haploid organisms, i.e. they have only one chromosome, so recombination usually occurs only when new DNA is picked up from the environment, for example, through transformation. Furthermore, bacteria are not readily transformable with linear DNA, e.g., because of intracellular exonucleases that degrade linear DNA. However, recombination-proficient mutants lacking exonucleases of the recombination complex, are transformable with linear DNA. Recombinases are enzymes that catalyze natural recombination reactions and it has been shown that the λ Red shows an enhanced recombination rate when compared with other recombination systems. The method for deletion of the group II capsular gene cluster of the strain of the present invention may be based on the Red-mediated recombination technology developed by Datsenko and Wanner [(2000) PNAS 97(12): 6640-6645] or other conventional methodologies.

Figure 2 shows a schematic view of how to delete the 16kb group II capsular gene cluster of a strain of the present invention. In one embodiment, the first step in the method is the replacement of the 16 kb group II capsular gene cluster with a selection marker that is generated by PCR by using primers with nucleotide homology extensions or regions, RA and LA (see Fig. 1). This is accomplished by Red-mediated recombination in these flanking homology extensions. After selection, the second step in this method is the elimination of the selection marker by using a helper plasmid expressing the FLP recombinase, which acts on the directly repeated FRT (FLP recognition target) sites flanking the selection marker. The Red and FLP helper plasmids are cured by growth at 37°C, since they are temperature sensitive replicons. The third step includes the confirmation of the deletion by PCR analysis.

Bacterial transformation may be referred to as a stable genetic change brought about by taking up DNA and competence refers to the state of being able to take up exogenous DNA. Some bacteria are naturally capable of taking up DNA under laboratory conditions and such species carry sets of genes specifying machinery for bringing DNA across the cell's membrane or membranes, while others have to be induced by laboratory procedures in which cells are passively made permeable to DNA, using conditions that do not normally occur in nature. Chilling cells in the presence of divalent cations such as Ca^{2+}

(in CaCl_2) prepares the cell walls to become permeable to plasmid DNA. Cells are incubated on ice with the DNA and then briefly heat shocked (e.g. 42 °C for 30–120 seconds), which causes the DNA to enter the cell and is a well-known method in the art [Sambrook et al., A Laboratory Manual (1989) CSH]. Electroporation is another way to make holes in bacterial (and other) cells, by briefly shocking them with an electric field in the kV range. Plasmid DNA can enter the cell through these holes. This method is amenable to use with large plasmid DNA. Natural membrane-repair mechanisms will rapidly close these holes after the shock. In order to persist and be stably maintained in the cell, a plasmid DNA molecule must contain an origin of replication, which allows it to be replicated in the cell independently of the chromosome.

When employing any of the gene inactivation methodologies described herein, any targeted inactivation of part or the whole group II gene cluster is detected by a selection marker so that the modified *E. coli* B BL21 strain clone can be identified and isolated.

The selection marker may be any suitable marker gene such as an antibiotic resistance gene or an auxotrophic marker.

Suitable antibiotic resistance genes are chloramphenicol, tetracyclin, ampicillin, kanamycin, vancomycin and erythromycin or any other representatives of the β -lactams, the aminoglycosides, the glycopeptides or the macrolide antibiotics.

The term “cloning” as used herein entails making multiple identical copies of a sequence of DNA. The target DNA sequence is then inserted into a cloning vector. Because this vector originates from a self-replicating virus, plasmid, or higher organism cell when the appropriate size DNA is inserted the “target and vector DNA fragments are then ligated” and create a recombinant DNA molecule. The recombinant DNA molecules are then put into a bacteria strain (usually *E. coli*) which produces several identical copies after transformation.

Transformation is the DNA uptake mechanism possessed by bacteria. However, only one recombinant DNA molecule with a given replication origin can be established within a single bacteria cell, so each clone harbors just one DNA insert.

The terms “genome” or “genetic composition” as used herein refer to the entire genetic matter, both structural and functional, which is inheritable and/or otherwise transferable, and is contained within an identified cell or microorganism. The genome (genetic composition) of a microorganism is thus understood to include both chromosomal DNA and extrachromosomal DNA (e.g. plasmids, R factors and the like).

The term “about” as used herein means in reasonable vicinity of the stated numerical value, such as plus or minus 10%.

The term "knockout" as used herein, means a genetically engineered bacteria strain in which one or more genes have been deleted or inactivated (have been "knocked out" of the organism). Knockout organisms are normally used as models for studying the role of genes which have been sequenced but have unknown functions. Knockout techniques may be used to remove undesired and/or potentially pathogenic genes from organisms of interest.

The term "gene cluster" as used herein means a set of two or more genes which are genetically and functionally coupled. Because populations from a common ancestor tend to possess varieties of the same gene clusters, they are useful for tracing back recent evolutionary history.

The term "parental strain" may be any bacterial strain from which the reduced genome strain is derived. Representative examples of parent strains according to the present invention include, but are not limited to, *E. coli* strains such as B or C, or a strain with a genome sequence substantially identical thereto.

The terms "level" or "amount" of peptide produced in a given bacterial host cell, as used herein may refer to a quantitative measure of peptide present or detectable, in accordance with the method of production selected, in the cell culture. For example, if the method of peptide production chosen results in an accumulation of peptide within the cell itself, then the level or amount of peptide produced will be that quantitative measure of peptide present with the cells themselves. If the method of peptide production involves secretion of the peptide, then the level of peptide production will be that quantitative measure of peptide present in the growth medium.

The term "yield" of peptide, as used herein refers to a quantitative amount of an essentially pure peptide, which peptide is retrieved and/or purified essentially free from peptides naturally associated with the host strain. In ascertaining a quantitative "yield" or "amount" of peptide produced by a bacterial host cell, such as *E. coli* B BL21 strains, conventional peptide-specific methods such as immunoassay, high performance liquid chromatography (HPLC), enzymatic activity, spectrophotometric techniques, SDS-PAGE and the like can be employed.

The term "expression control sequence" used herein means a promoter or array of transcription factor binding sites that direct transcription of a nucleic acid operatively linked thereto.

The term "free end" used in reference to a double-stranded nucleic acid may mean a linear nucleic acid with blunt free ends or sticky free ends, or a combination thereof.

The term "gene" used herein refers to a nucleic acid (e.g., DNA or RNA) that comprises a nucleic acid sequence encoding a polypeptide or precursor thereto. The polypeptide can be encoded by a full length coding sequence or by any portion of the coding sequence so long as the desired activity or functional properties (e.g., enzymatic activity, ligand binding, signal transduction, antigenicity etc.) of the full-length or fragment are retained. The term also encompasses the sequences located adjacent to the coding region on both the 5' and 3' ends that contribute to the gene being transcribed into a full-length mRNA. The term "gene" encompasses both cDNA and genomic forms of a gene. A genomic form or clone of a gene may contain the coding region interrupted with non-coding sequences termed (e.g., introns).

The terms "introduce" or "introduced", when used in reference to adding a nucleic acid to a strain, means that the nucleic acid may be integrated into the chromosome of the strain or contained on a vector, such as a plasmid, in the strain.

The terms "open reading frame" or "ORF" as used herein refers to a genetic sequence that does not contain stop codons.

The terms "peptide", "polypeptide" or "protein" are used interchangeably herein and mean a peptide, polypeptide and protein, whether native or recombinant, as well as fragments, derivatives, homologues, variants and fusions thereof.

The growth of group II knock out strains was evaluated in shake flasks and in high density fermentors and compared with their parental strains. The experiments showed that group II knock out does not affect the growth of the cells in either strain background. Expression vector harbouring several peptides were transformed into group II deleted strains for expression tests in a high density fermentor.

In one embodiment, the peptides for production in the *E. coli* B strains of the present invention may be selected from the group consisting of, but not limited to hGH, glucagon-like peptides, interleukins, insulin analogs, wildtype adiponectin, FGF-21, trypsin, aprotinin, amylin and leptin and analogs thereof. Enzymes such as sialidase, transglutaminase (tGase), HRV (Human Rhino Virus) 3C protease, Tobacco Etch Virus (TEV) protease and variants thereof, and any other peptide or peptide for which an expression system is operable in *E. coli*.

The present invention also relates to a method of preparing human peptides as mentioned above. In general, a cloned wildtype peptide nucleic acid sequence is modified to encode the desired protein. This modified sequence is then inserted into an expression vector, which is in turn transformed or transfected into host cells. Operable expression systems typically include the conventional gene expression components set forth by Studier and Mof-

fat [J.Mol.Biol. (1986) 189:113-130] and Studier et al. [Methods in Enzymology (1990) 185:60-89].

The peptides to be produced in the host cells can be synthesized as fusion peptides, secreted into the media or into the periplasmic space, produced in free form in the cytoplasm, or accumulated in intracellular bodies such as in inclusion bodies. Purification of the peptides similarly involves the use of aforementioned conventional, publicly available methodologies which are adapted, again by conventional techniques, to the peptide of interest.

The present invention is suitable for the production of any type of peptide.

In one embodiment, the peptide selected for production in the group II knock out *E. coli* B BL21 host cells is human growth hormone (hGH) and in one embodiment the human growth hormone (hGH) is produced from a precursor peptide as described in EP217814, which describes how to make hGH from an hGH precursor.

In a further embodiment the growth hormone is MEAE(SEQ ID 16)-hGH(L101C), which is a precursor of an hGH analogue, wherein the leucine at position 101 in the hGH molecule is substituted by cysteine.

In another embodiment the glucagon-like peptide is GLP-1, GLP-1 and/or derivatives thereof.

The term "glucagon-like peptide" as used herein means the glucagon family of peptides, exendins and analogues thereof. The glucagon family of peptides are encoded by the preproglucagon gene and encompasses three small peptides with a high degree of homology, i.e. glucagon (1-29), GLP-1 (1-37) and GLP-2 (1-33). Exendins are peptides expressed in lizards and like GLP-1, are insulinotropic. Examples of exendins are exendin-3 and exendin-4.

The term "glucagon-like peptide analogue" as used herein refers to a modified glucagon-like peptide wherein one or more amino acid residues of the glucagon-like peptide have been substituted by other amino acid residues and/or wherein one or more amino acid residues have been deleted from the glucagon-like peptide and/or wherein one or more amino acid residues have been added to the glucagon-like peptide. Such addition or deletion of amino acid residues can take place at any position of the peptide.

EXPERIMENTAL PART

For the preparation of glycerol stocks of the strains, a single colony was transferred into a 12ml test tube containing 2ml of LB medium with the appropriate antibiotics and grown at 37°C, with shaking at 220rpm for 2 hours.

For the preparation of fermentation inoculae, the cells from a glycerol stock were

streaked on LB plates with and without addition of ampicilin. The overnight cells were washed with LB medium and inoculated into 500ml shake flasks containing 100ml each of fermentation complex medium with appropriate antibiotic concentration and grown at 37°C with shaking at 220rpm for 3-4 hours. The cells were then inoculated into the reactor to a concentration of approximately 0.5 OD₆₀₀.

Fermentations were conducted under aerobic conditions in a 5l bioreactor with an initial volume of fermentation complex medium of 2.5l, with appropriate antibiotics. At time zero, the reactor was inoculated to an OD₆₀₀ of approximately 0.5. The pH was maintained at 7.0 with 5N NH₃ and H₂SO₄ and a constant air flow rate of 6ml/min. was sparged through the culture during the fermentation period. The temperature was maintained at 37°C. The agitation was cascade-controlled by maintaining dissolved oxygen level up to 20% of saturated O₂.

A fed-batch experiment was performed to compare the growth and peptide expression of group II deleted strain with its parental strain in a high density fermentor. The fermentor was charged with complex medium supplemented with 27g/l of glycerol. At time zero, the fermentor was inoculated to an OD₆₀₀ of 0.5. Glycerol and yeast extract were added continuously in increasing steps as basic glycerol was consumed, which was indicated by an increase in dissolved oxygen. The samples for determination of OD₆₀₀ and protein expression were collected from the fermentor and analyzed by spectrophotometer and by SDS-PAGE.

Induction of protein expression in bacteria is well known in the art. In the present invention the induction of protein expression is made by addition of isopropyl-β-D-thiogalactopyranoside (IPTG).

Plasmids and strains used in certain embodiments of the invention are listed in Tables 1 and 2, respectively. The construction of plasmids was performed using standard molecular biology techniques as described in Sambrook et al. [A Laboratory Manual (1989) CSH].

Table 1. Plasmids used in the examples

Plasmid	Genotype	Origin
pKD46	Amp ^R	<i>E. coli</i> Genetic Stock Center
pCP20	Amp ^R Cm ^R	<i>E. coli</i> Genetic Stock Center
pEZ-T	Cloning vector Amp ^R	GenStar
pGEM-T	Cloning vector Amp ^R	Promega
pNNC1-g (pET11d-MEAE-hGH)	Expression vector Amp ^R expressing MEAE-hGH	Based on pET 11d from Novagen
pNNC1-1h (pET11d-MEAE-hGH, <i>alr</i> (<i>alr</i> promoter and coding sequence))	Expression vector <i>alr</i> Expressing MEAE-hGH and D,L-alanine racemase	Based on pET 11d from Novagen

Table 2. Strains used in the examples

Strain	Genotype	Reference
BL21(DE3)	F-ompT hsdS _B (rB-mB-)gal dcm (DE3)	Novagen
BL21(DE3) Δ <i>alr</i> Δ <i>dadx</i>	F-ompT hsdS _B (rB-mB-)gal dcm (DE3) Δ <i>alr</i> Δ <i>dadx</i>	WO2008/025744
BL21(DE3) Δ (<i>kpsM-kpsF</i>):: <i>cat</i>	F-ompThsdS _B (rB-mB-)gal dcm(DE3) Δ (<i>kpsM-kpsF</i>):: <i>cat</i>	Present invention
BL21(DE3) Δ <i>dadX</i> Δ <i>alr</i> Δ (<i>kpsM-kpsF</i>):: <i>cat</i>	F-ompThsdS _B (rB-mB-)gal dcm(DE3) Δ <i>alr</i> Δ <i>dadx</i> Δ (<i>kpsM-kpsF</i>):: <i>cat</i>	Present invention
BL21(DE3) Δ (<i>kpsM-kpsF</i>)	F- ompT hsdS _B (rB-mB-)gal dcm(DE3) Δ (<i>kpsM-kpsF</i>)	Present invention
BL21(DE3) Δ <i>dadX</i> Δ <i>alr</i> Δ (<i>kpsM-kpsF</i>)	F-ompThsdS _B (rB-mB-)gal dcm(DE3) Δ <i>alr</i> Δ <i>dadx</i> Δ (<i>kpsM-kpsF</i>)	Present invention

Table 3. List of primers used to carry out the invention

PRIMER NAME	CONSENSUS 5' TO 3' (DNA DOUBLE STRANDED)
LA_BL_GII_F	SacII+LA TCCTTACCGCGGGTCAAACCCGCTTACCGGGC (SEQ ID 1)
LA_BL_GII_R	EcoRI+LA down CTACCGGAATTCTTGATGATGTGATCCTAATCTC (SEQ ID 2)
RA1_BL_GII_F	BamHI+RA1 TTCGGATCCATCAGCTCCTTTGCACGGAA (SEQ ID 3)
RA1_BL_GII_R	SacI+RA1 down GAACGAGCTCACAACCTCCGAATCATCAATA (SEQ ID 4)
RA_SEQ_R1	GCACCACTAATTCTTAAGAACCCGCCACAAG (SEQ ID 5)
LA_SEQ_F1	ACTGGAATTTGGACGGGGGTAGGTATTGCC (SEQ ID 6)
LA_SEQ_F2	TTGATGTCAGTGCTGCTGAGAAGCCTGGGATG (SEQ ID 7)
Cat_SEQ_F	CGACGATTTCCGGCAGTTTCTAC (SEQ ID 8)
CAT_ID_R	GCCAGGTTTTACCGTAACACGCC (SEQ ID 9)
RA6_BL_GII_F	BamHI+RA6 CCTGTGGATCCAGAACGTTGTTG (SEQ ID 10)
RA6_BL_GII_R	SacI+RA6 down GAACGAGCTCATGTTGATAAAAGTGAAGTC (SEQ ID 11)

EXAMPLES**Example 1: Construction of replacement cassette**

5 The *E. coli* B BL21(DE3) strain was used as a source of DNA for PCR amplification. Oligonucleotide primers (LA_BL_GII_F/LA_BL_GII_R) (Table 3) were synthesized, targeting a 1001bp region upstream of the 5' end of *KpsM* containing SacII/EcoRI cloning site. This region was amplified from genomic DNA as left arm (LA). For the right arm (RA), oligonucleotide primers (RA1_BL_GII_F/RA1_BL_GII_R) (Table 3) were synthesized, targeting a

10 1108bp region upstream of the 5' end of *KpsF* containing BamHI/SacI cloning site. PCR products of the left arm and the right arm were generated. The DNA amplification reaction was performed in a thermal cycler, using a program that included the following parameters: denaturation at 94°C for 50s, annealing at 57°C for 60s and primer extension at 72°C for 60s, for a total of 30 cycles. The PCR amplification product with correct band size was observed

15 by gel electrophoresis on a 1% agarose gel, from which 50µl of the amplified product were extracted from the gel and used to construct a replacement fragment. A list of oligonucleotide sequences is shown in Table 3.

The amplified left arm was gel-purified and ligated into plasmid pEZ-T, resulting in plasmid pEZ-T-LA. The restriction enzymes EcoRI/BamHI digested the FRT-cat-FRT frag-

ment which was ligated into EcoRI/BamHI site of pEZ-T-LA, resulting plasmid pEZ-T-LA-FRT-cat-FRT. The amplified right arm was ligated into plasmid pGEM-T and was digested with BamHI/SacI. The BamHI/SacI-digested DNA fragment, which carries an ampicillin resistance gene was ligated into plasmid pEZ-T-LA-FRT-cat-FRT. The chloramphenicol resistance gene provides a selectable marker for identifying the replacement fragment. The resulting plasmid pEZ-T-LA-FRT-cat-FRT-RA, was used to produce a knockout BL21(DE3) strain.

Example 2: Production of a knock out BL21(DE3) strain

E. coli BL21(DE3) was grown in LB medium overnight at 37°C and the heat shock competent cells of BL21(DE3) were prepared by standard methods. The plasmid pKD46 was transformed into BL21(DE3) competent cells and the transformed cells were grown to 0.4 OD_{600nm} in LB broth supplemented with 100µg/ml ampicillin at 30°C. The expression of Red recombinase was induced by L -arabinose (final concentration 1mmol/l) for one hour.

For group II gene cluster knock out, 500ng of replacement fragment amplified from pEZ-T-LA-FRT-cat-FRT-RA were transferred into BL21(DE3) competent cells by electroporation. The electrotransformation competent cells were prepared by standard methods. Recovery of the cells was carried out at 30°C for 1.5 hours. The cells were plated on LB agar supplemented with 8µg/ml chloramphenicol. Chloramphenicol-resistant colonies of BL21(DE3)Δ*kpsM-kpsF*::*cat* were formed after overnight incubation at 30°C.

To ensure complete knock out of the group II gene cluster at the right position of the chromosome, oligonucleotide primers were designed with homology to the outside of the left and the right arms of the group II gene cluster. Lacking the flanking sequences of *E. coli* BL21(DE3) outside the 18kb, the published sequence of *E. coli* 536 (Accession no. CP000247.1) was used for design of the PCR screening primers, assuming a similar location of group II cluster in the 2 genomes, based on high homology in the near flanking regions. Three chloramphenicol-resistant colonies were tested for gene replacement.

Phenotype test of chloramphenicol-resistant colonies was performed by growth in LB+Cm (25µg/ml) for at least two generations. All of the chloramphenicol-resistant colonies could grow very well in LB+Cm (25µg/ml) medium. Chloramphenicol-resistant D-alanine auxotrophic strains have also been tested in the same process with supplementation of D-alanine.

A successful replacement was confirmed by bands of 1686bp and 1578bp.

Oligonucleotide primers (LA_SEQ_F1/Cat_ID_R) were used to amplify a band containing left arm and part of *cat* gene (size 1686bp), and oligonucleotide primers (RA_SEQ_R1/Cat_SEQ_F) were used to amplify a band containing right arm and part of the *cat* gene (size 1578bp). For three chloramphenicol-resistant colonies right size of the bands indicated that the 16kb group II gene cluster had been replaced by the *cat* gene (Fig.3). A negative control of 3000 bp band which contained the right arm including *kpsF* and part of *kpsE*, was created and should appear only in the control strain (before replacement) but not in the chloramphenicol-resistant strain. Oligonucleotide primers (RA6_BL_GII_F/RA6_BL_GII_R) (Table 3) were used to amplify *kpsF* and *kpsE*. The PCR results showed that only the control strain had this band, not the chloramphenicol-resistant strain.

The same elimination process was applied to the D-alanine auxotrophic strain BL21(DE3) Δ *dadX* Δ *alr*, which needs a supplement of D-alanine during growth in LB medium.

The chloramphenicol-resistant strains *E. coli* B BL21(DE3) Δ (*kpsM-kpsF*)::*cat* and BL21(DE3) Δ *dadX* Δ *alr* Δ (*kpsM-kpsF*)::*cat* were confirmed by their phenotype and PCR screening, confirming that the group II capsular gene cluster had been replaced by the *cat* gene.

Example 3: Elimination of drug-resistant marker and characterization of group II knock out strain

E. coli BL21(DE3) Δ (*kpsM-kpsF*)::*cat* and BL21(DE3) Δ *dadX* Δ *alr*(*kpsM-kpsF*)::*cat* cells from frozen stock were grown in LB medium overnight at 37°C with 25µg/ml chloramphenicol or 25µg/ml chloramphenicol, plus 100mM D-alanine for the second strain. The electroporation competent cells of the two knock out strains were prepared by standard methods. The plasmid pCP20 (Cm^R, Amp^R) was transformed into the two knock out competent strains. The elimination of the drug-resistant marker was carried out by growing the transformed cells in SOC broth at 30°C for 1 hour with gentle shaking. The single colonies of ampicillin resistant candidate clones on the plate were selected after overnight culture at 30°C. The candidate clones were grown in LB medium only or LB medium supplemented with D-alanine at 37°C for three hours, to eliminate the plasmid pCP20. Five single colonies, which appeared on LB medium only or LB medium supplemented with D-alanine, were further characterized by PCR screening.

Phenotype testing of *cat* elimination strains was performed by growing the single colonies in LB, LB+Cm (8µg/ml) and LB+Amp. All the *cat* elimination colonies could grow only on LB medium. For *cat* eliminated D-alanine auxotrophic strains, all five colonies could

grow only in LB medium with supplementary D-alanine. Thus, the *cat* gene had been eliminated.

A successful elimination of the *cat* gene should be confirmed by a 2381bp band. Oligonucleotide primers (LA_SEQ_F2/RA_SEQ_R1) were used to amplify a band containing left arm and right arm. For all the single colonies, the appearance of a 2381bp band indicated that the *cat* gene had been eliminated.

The knock out strains *E. coli* B BL21(DE3) $\Delta(kpsM-kpsF)$ and B BL21(DE3) $\Delta dadX \Delta alr \Delta(kpsM-kpsF)$ were confirmed by their phenotype and PCR screening. Glycerol stocks of both strains were prepared for storage.

Example 4: Growth of the knock out host strains in shake flask

E. coli B BL21(DE3) and B BL21(DE3) $\Delta(kpsM-kpsF)$ cells from frozen stocks were grown on LB plate overnight at 37°C. Appropriate amounts of cells washed out from overnight plates were inoculated into 250ml baffled shake flasks containing 25ml LB medium to a concentration of approximately 0.05 OD₆₀₀. The cells were grown at 37°C with shaking at 220rpm for 8 hours. A sample of each strain was taken every ½ hour to compare the growth of the cells at OD₆₀₀.

The same working flow was applied to *E. coli* B BL21(DE3) $\Delta dadX \Delta alr$ and B BL21(DE3) $\Delta dadX \Delta alr \Delta(kpsM-kpsF)$ supplemented with D-alanine for comparison of growth in shake flask.

The results in Fig.4 show that growth of BL21 $\Delta(kpsM-kpsF)$ over time was identical to its parental strain BL21(DE3) in shake flask using LB medium. The same conclusion could be drawn when comparing BL21(DE3) $\Delta dadX \Delta alr \Delta(kpsM-kpsF)$ with BL21(DE3) $\Delta dadX \Delta alr$.

Example 5: Comparison of growth and MEAE-hGH expression between host strain BL21(DE3) and BL21(DE3) $\Delta kpsM-kpsF$ in high density fermentation

E. coli B BL21(DE3) $\Delta(kpsM-kpsF)$ cells from frozen stock were grown on LB plates. Heat shock competent cells of the knock out strain were prepared by standard methods. The plasmid pNNC1-g (pET11d-MEAE-hGH(Amp^R)) containing an expression cassette of hGH was transformed into B BL21(DE3) $\Delta(kpsM-kpsF)$ competent cells, from which a glycerol stock of transformed cells was prepared.

High density fermentations were conducted under aerobic conditions as described above.

Figure 5 shows that the growth of BL21 $\Delta(kpsM-kpsF)$ over time was identical to its

parental strain BL21(DE3) in high density fermentor in complex medium. The same conclusion could be drawn when comparing BL21(DE3) $\Delta(kpsM-kpsF)$ /pNNC1-g with BL21(DE3)/pNNC1-g. The cell density reached approximately 140 OD₆₀₀ in the uninduced state. Expression of hGH was carried out by adding IPTG to a concentration of 1mM when the growth of cells reached 80 OD₆₀₀. After 10 hours of IPTG induction at 30°C, the cell density reached approximately 100 OD₆₀₀.

E. coli B BL21(DE3)/pNNC1-g and B BL21(DE3) $\Delta(kpsM-kpsF)$ /pNNC1-g are efficient in the production of hGH as seen on SDS-PAGE. The increase in hGH expression level follows increasing induction time from 1 to 10 hours shown on the SDS-PAGE gel (Fig.6).

The ability of the two strains to produce hGH was identical, i.e., no significant difference was observed in the expression of hGH by host strain B BL21(DE3) $\Delta(kpsM-kpsF)$ when compared to B BL21(DE3), harbouring the gene II cluster.

Example 6: Comparison of growth and MEAE-hGH expression between host strain

BL21(DE3) $\Delta alr\Delta dadX$ and BL21(DE3) $\Delta alr\Delta dadX\Delta(kpsM-kpsF)$ in high density fermentation

E. coli B BL21(DE3) $\Delta dadX\Delta alr\Delta(kpsM-kpsF)$ cells from frozen stock were grown in LB medium supplemented with D-alanine. Heat shock competent cells of the knock out strain were prepared by standard methods. The pNNC1-1h (pET11d-MEAE-hGH and *alrP-alrCD* (Amp^R)) containing a D-alanine complementary gene was digested using AatII/PsiI, following by blunt end and self-ligation of the plasmid, resulting in a *bla*- plasmid for selection of the *alr* gene on the plasmid. The plasmid pNNC1-1h (pET11d-MEAE-hGH, *alrP-alrCD*, *bla*-) was transformed into BL21(DE3) $\Delta dadX\Delta alr\Delta(kpsM-kpsF)$ competent cells. Colonies appearing overnight on LB plate without supplementary D-alanine were isolated for the preparation of glycerol stocks.

High density fermentations were conducted as described in Example 5.

Figure 7 shows that the growth of *E. coli* B BL21(DE3) $\Delta dadX\Delta alr\Delta(kpsM-kpsF)$ /pNNC1-1h over time was identical to its parental strain BL21(DE3) $\Delta dadX\Delta alr$ /pNNC1-1h in a high density fermentor in FCM1 medium. The same conclusion could be drawn under induction of hGH when comparing B BL21(DE3) $\Delta dadX\Delta alr$ /pNNC1-1h with BL21(DE3) $\Delta dadX\Delta alr\Delta(kpsM-kpsF)$ /pNNC1-1h. The cell density reached approximately 120 OD₆₀₀ in the uninduced state. Expression of hGH was carried out by adding IPTG to a concentration of 1mM when the growth of cells reached 80 OD₆₀₀. After 9 hours of IPTG induction at 30°C, the cell density reached approximately 100 OD₆₀₀.

E. coli B BL21(DE3) $\Delta dadX\Delta alr$ /pNNC1-1h and B BL21(DE3) $\Delta dadX\Delta alr\Delta(kpsM-$

kpsF) /pNNC1-1h were efficient in production of hGH as can be seen from the results by SDS-PAGE (Fig.8). Increase in hGH expression level follows increasing induction time. Both strains had identical ability to produce hGH.

5 Example 7: Evaluation of cell growth and expression of MEAE-hGH(L101C) using host strain *E. coli* B BL21(DE3) Δ *alr* Δ *dadX* Δ (*kpsM-kpsF*) in high-cell density fermentation

E. coli B BL21(DE3) Δ *dadX* Δ *alr* Δ (*kpsM-kpsF*) cells from frozen stock were grown in LB medium supplemented with D-alanine. Competent cells of this triple knock-out strain for transformation using heat-shock method were prepared by standard methods.

10 To delete the β -lactamase gene coding for ampicillin-resistance, the plasmid pNNC1a.23 (pET11d-MEAE-hGH(L101C) encoding MEAE-hGH(L101C) and containing *alrP*-*alrCD* for complementing the D-alanine-deficiency of the host cell was digested using *Aat*II/*Psi*I, followed by end-blunting and self-ligation for removal of the β -lactamase gene. The resulting plasmid, pNNC1a.23 (pET11d-MEAE-hGH, *alr*, *bla*⁻), was transformed into the host
15 cell, BL21(DE3) Δ *dadX* Δ *alr* Δ (*kpsM-kpsF*), by heat-shock method. A single colony grown up from LB plate without supplementation with D-alanine was selected to prepare the glycerol stocks for storage at -80°C.

Two batches of high-cell density fermentation were conducted as described in Example 5. Cell growth was monitored during the process and the results are shown in Fig.9.

20 Expression of MEAE-hGH(L101C) by strain *E. coli* B BL21(DE3) Δ *dadX* Δ *alr* Δ (*kpsM-kpsF*) harbouring plasmid pNNC1a.23 was evaluated in high-cell density fermentation. The result showed that the cell density at harvest can reach to about 100 measured at OD_{600nm}, and the expression level of the target protein, MEAE-hGH(L101C), reached about 7g/l after 8 hours' induction, as shown in Fig.10.

25

Example 8: Expression of *Arthrobacter ureafaciens* sialidase in *E. coli* BL21(DE3) Δ (*kpsM-kpsF*)

Arthrobacter ureafaciens sialidase (AUS, 53.4 kDa theoretical mass) was produced in *Escherichia coli* BL21(DE3) Δ (*kpsM-kpsF*) with a C-terminal His-tag. The *Arthrobacter ure-*
30 *afaciens* sialidase gene, with a His6-tag in the C-terminus, was cloned in expression vector pET-24a (Novagen) under control of the T7 promoter. In order to produce the sialidase enzyme from *Arthrobacter ureafaciens* a synthetic codon optimized gene encoding the following protein sequence was generated (SEQ ID 12):

MAPTPPNSPTLPPGSFSETNLAADRRTAANFFYRIPALTYLGNDVVLAAWDGRPGS
35 AADAPNPNSIVQRRSTDGGKTWGPVQVIAAGHVADASGPYRGYSYSDPSYIYDAEANKVFAFF

VYSKDQGFSGSQFGNDDADRNVISSAVIESSDAGVTWSQPRLLTSVTKPGTSKTNPAAGDV
 RSNFASSGEGIQLKYGPHKGRLIQQYAGDVRQADGSNKIQAYSVYSDDHGVTVHKGANVG
 DRMDENKTVELSDGRVLLNSRDNANRGYRKVAVSTDGGATYGPVSQDTELPDPANNGAIA
 RMFPNAAQGSADAKKLIFTNANSKTGRENVSARVSCDDGETWPGVVRTIRSGFSAYSTVTRL
 5 ADGKFGVLYEGNYTDNMPFATFDDAWLNYVCAPLAVPAVNIAPSATQEVPTVTNQEATTL
 SGATATVYTPSGWSATTVPVPDVAPGASVTVTVALTAPADASGPRSLNAAFTTADGRVSQF
 TFTATTPVAPQVGLTITGSALEHHHHHH

The synthetic gene was introduced into the pET-24a expression vector as a XbaI-XhoI fragment to produce the AUS expression vector pLLC028. The resulting vector was
 10 transformed into competent BL21(DE3) $\Delta(kpsM-kpsF)$ and transformants were preserved in glycerol at -80°C.

One cryotube of the frozen initial cell clone (ICC) was used to inoculate two yeast extract media (yeast extract 5 g/l, glucose 10 g/l, agar 20 g/l) agar flask. The agar medium contained kanamycin for selection. Incubation was at 30°C for 25 hours. After incubation, the
 15 cells were washed from the agar surface using sterile 0.9% NaCl solution. A 10 l fermentation was inoculated using pre-culture from two agar flasks. The volume of the obtained inoculum was 75 ml, and the OD_{600nm} was 4.3, resulting in a start OD_{600nm} in the fermentor of approximately 0.03. The fermentation was carried out in a 20 l stainless steel vessel with a start working volume of 10 l. The fermentation starts with a 8 hour long batch phase on a complex
 20 medium containing 40 g/l yeast extract, 15 g/l glycerol (87%), salts and trace metals, followed by glycerol feeding at a predetermined profile. During the first 17 hours, fermentation temperature was 37°C. After that, the temperature was lowered to 30°C. At 18 hours elapsed fermentation time, the inducer IPTG was added at 0.1 mM in order to induce protein expression. The duration of induction can be up to 22 hours. The recovery process consists of cen-
 25 trifugation, high pressure homogenisation and crossflow microfiltration using Sartoclon Slice Cassettes with a pore size of 0.2 µm.

The AUS fermentation was analysed using SDS-PAGE and an enzymatic activity assay. SDS-PAGE analysis revealed that the expressed sialidase appears both in the soluble and pellet fraction after bead-milling. Figure 11 shows a typical fermentation result from a
 30 10L batch run at 30°C induction temperature and 0.1 mM IPTG.

Example 9: Expression of human fibroblast growth factor 21 (FGF21) in BL21 DE3 and BL21 DE3 $\Delta(kpsM-kpsF)$

The DNA and amino acid sequences for human FGF21 are disclosed by e.g., Nishimura *et al* in Biochim. Biophys. Acta 1492(1):203-206 (2000). The sequences are also
 35

available from public databases (Accession nos. EMBL:AB021975 and UNIPROT:Q9NSA1, respectively).

The native polypeptide is synthesised with a signal peptide (in italics below) of 28 amino acids for secretion, while the mature FGF21 polypeptide consists of the remaining 181 amino acids (SEQ ID 13):

*MDSDETGF**EH**SGLWVSVLAGLLLGACQAH**PI**DSSPLLQFGGQVRQRYLYTDDA*
*QQTEAHLEI**REDGTVGGAADQS**PESLLQLKALKPGVIQILGVKTSRFLCQRPD**GALYGSLHF*
*DPEACSFRELLLEDGYNVYQSEAHGLPLHLPGNKSPHRDPAPRGP**ARFLPLPGLPPA-*
*LPEPPGILAPQPPDV**GSSDPLSMVGPS* QGRSPSYAS

The mature FGF21 polypeptide was cloned and expressed as an intracellular protein in *E. coli*, without the signal peptide, but with an added N-terminal methionine. In particular, a 550 bp coding region including at the 3'-end the ATG codon for Met, as well as Nde1 and BamH1 restriction sites at the 3'- and 5'-ends, respectively, were inserted into the expression vector pET 11c (Novagen) in Nde1-BamH1 under control of the phage T7 promoter, and transformed into *E. coli* B BL21(DE3) and into *E. coli* B BL21 DE3Δ(*kpsM-kpsF*). The cells were grown in LB amp 100 ug/ml to OD₄₅₀ 0.5, and expression was induced with 0.3 mM IPTG for 4 hours at 37°C. Crude extracts of cells were made by sonication for analysis of FGF21 expression.

The Coomassie stained SDS-PAGE gel in Fig. 12 shows equally successful expression of FGF21 in *E. coli* B BL21 DE3 and *E. coli* B BL21 DE3Δ(*kpsM-kpsF*). Although the calculated molecular weight (MW) of the thus expressed FGF21 (Met-FGF21) is 19.5 kD, it migrated on the gel as a 25 kD protein, which is likely due to the high content of prolines, delaying the movement of the protein.

The current example shows that the yield for FGF21 is not affected by the genetic manipulation of the host strain.

Example 10: Expression of human amylin in BL21 DE3 and BL21 DE3Δ(*kpsM-kpsF*)

The sequence of the 37 aminoacid amyloid peptide amylin was published by Cooper *et al.* in Proc. Natl. Acad. Sci. (84) : 8628-8632, (1987). The gene was cloned and fused it N-terminally to a ribosomal protein RL23 from *Thermotoga maritima*. The 146 amino acids fusion protein has the following sequence (SEQ ID 14):

KQEKLSLHDVLRPIITEKALILREQRKYVFEVNPLANKNLVKEAVEKLFNVKVEKVN
 LNMKPKPKRRGIFEGKTRSWKKAVVTLKEGYTIKELEGEHSSSSDDDDKKCNTAT-
 CATQRLA-NFLVHSSNNFGAILSSTNVGSNTY

The fusion gene was cloned as described in example 9 as a 450 bp Nde1- BamH1 fragment into pET 11c under control of the T7 promoter, and the plasmid was transformed into BL21 DE3 and BL21 DE3 $\Delta(kpsM-kpsF)$.

Expression was performed as described in example 9, and the result is shown in Fig.12, again demonstrating equally high yield expression in the two *E. coli* B hosts BL21 DE3 and BL21 DE3 $\Delta(kpsM-kpsF)$.

Example 11: Expression of human leptin in BL21 DE3 and BL21 DE3 $\Delta(kpsM-kpsF)$

The sequence of the 146 amino acids long human leptin was published by Masuzaki *et al.* in Diabetes (44): 855-858 (1995). The protein is secreted in humans with the addition of a 21 amino acids signal peptide. The mature leptin was expressed intracellularly in *E. coli* has the following sequence (SEQ ID 15):

MVPIQKVQDDTKTLIKTIVTRINDISHTQSVSSKQKVTGLDFIPGLHPILTLSKMDQTL
AVYQQILTSMPSRNVIQISNDLENLRDLLHVLAFSKSCHLPWASGLETLDLGGVLEASGYST
EVVALSRLQGSLQDMLWQLDLSPGC

The leptin gene was cloned as described in example 9 as a 450 bp Nde1-BamH1 fragment into pET11c under control of the T7 promoter, and the plasmid was transformed into BL21 DE3 and BL21 DE3 $\Delta(kpsM-kpsF)$.

Expression was performed as described in example 9, and the result is shown in Fig.12, again illustrating equally high yield expression in the two *E. coli* hosts BL21 DE3 and BL21 DE3 $\Delta(kpsM-kpsF)$.

It is concluded that the elimination of group II capsular gene cluster does not affect the growth of the cells in any growth conditions. It can also be concluded that group II capsular gene cluster elimination does not affect protein expression. Therefore, a more safe strain with complete removal of group II capsular gene cluster is hereby produced for using *E. coli* B BL21(DE3) as commercial host strain for therapeutic protein production. The operation of deleting group II capsular genes in *E. coli* B BL21(DE3) brings it into the same safety category as *E. coli* K12, since the group II gene cluster is the only significant difference between the two strains.

CLAIMS

1. An *E. coli* B BL21 strain characterised by having an irreversibly inactivated group II capsular gene cluster.
2. An *E. coli* B BL21 strain according to claim 1, characterised by having a total or partial
5 deletion of group II capsular gene cluster.
3. An *E. coli* B BL21 strain according to claim 1, characterised by having a total deletion of group II capsular gene cluster.
4. An *E. coli* B BL21 strain according to any of the preceding claims, wherein the *E. coli* is *E. coli* B BL21(DE3).
- 10 5. An *E. coli* strain according to any of the claims 1-3, wherein the *E. coli* is *E. coli* B BL21(DE3) Δ *dadX* Δ *alr*.
6. An *E. coli* B BL21 strain according to any of the preceding claims, wherein the strain further comprises an expression vector with an inserted DNA sequence encoding a peptide.
7. An *E. coli* B BL21 strain according to any of the preceding claims, wherein the peptide is
15 selected from the group consisting of hGH, glucagon-like peptides, interleukins, insulin analogs, wildtype adiponectin, FGF-21, trypsin, aprotinin, amylin and leptin and analogs thereof and enzymes such as sialidase, transglutaminase (tGase), HRV (Human Rhino Virus) 3C protease, Tobacco Etch Virus (TEV) protease and variants thereof.
8. A method for producing an *E. coli* B strain according to any of the preceding claims,
20 wherein the group II capsular gene cluster is deleted by a method comprising the steps of:
 - a. replacing of group II capsular gene cluster by a selection marker,
 - b. eliminating the selection marker and
 - c. confirming the deletion by PCR analysis.
- 25 9. A method for production of an *E. coli* B strain according to claim 8, where the selection marker is an antibiotic resistance gene.
10. A method for production of an *E. coli* B strain according to claim 9, where the antibiotic resistance gene is selected from the group consisting of chloramphenicol, tetracyclin, ampicillin, kanamycin, vancomycin and erythromycin.
- 30 11. A method for production of an *E. coli* B strain according to claim 9, where the antibiotic is chloramphenicol.
12. A method for the production of a peptide comprising culturing in a suitable culture medium, an *E. coli* B strain according to any of claims 1-7 followed by isolation and purification of the expressed peptide by well-known technique.
- 35 13. A method according to claim 12 wherein the peptide is hGH.

14. A method according to claim 12 wherein the peptide is FGF-21.
15. A method according to claim 12 wherein the peptide is amylin.

1/7

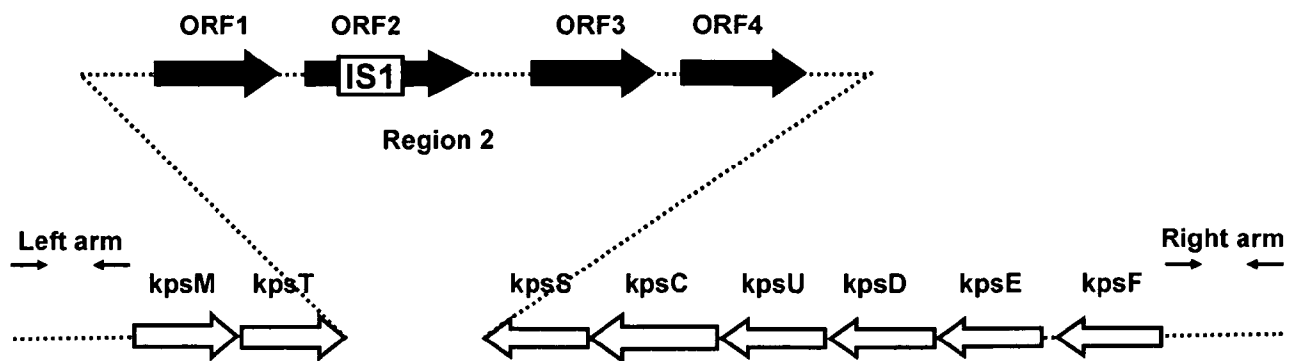


Fig. 1/12

Fermentation growth curve

□, BL21(DE3) Δ dadX Δ alr Δ (kpsM-kpsF)/pNNC1a.23

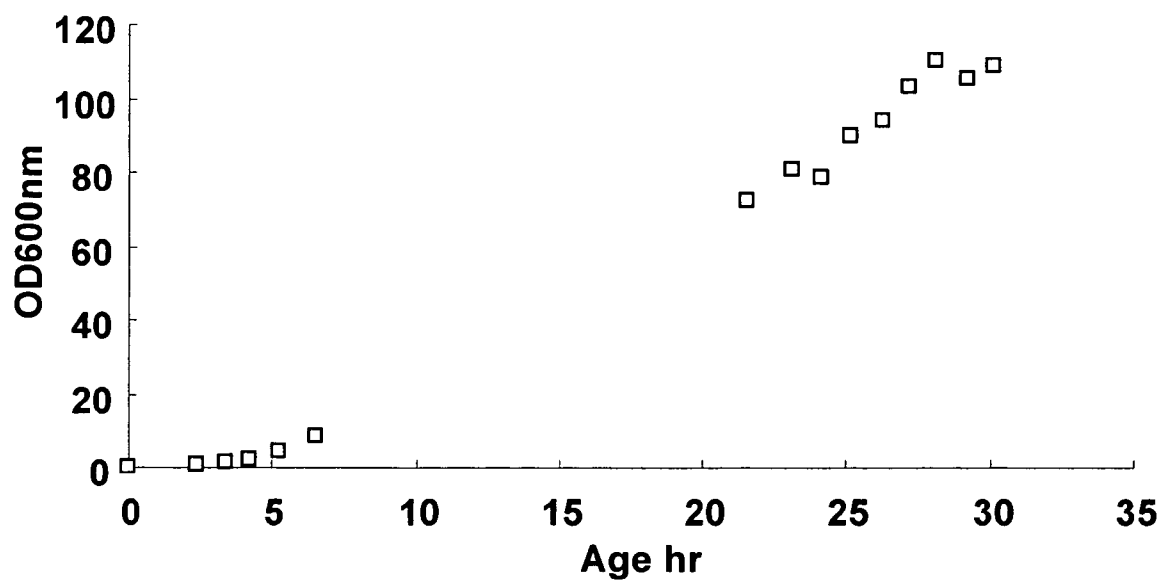
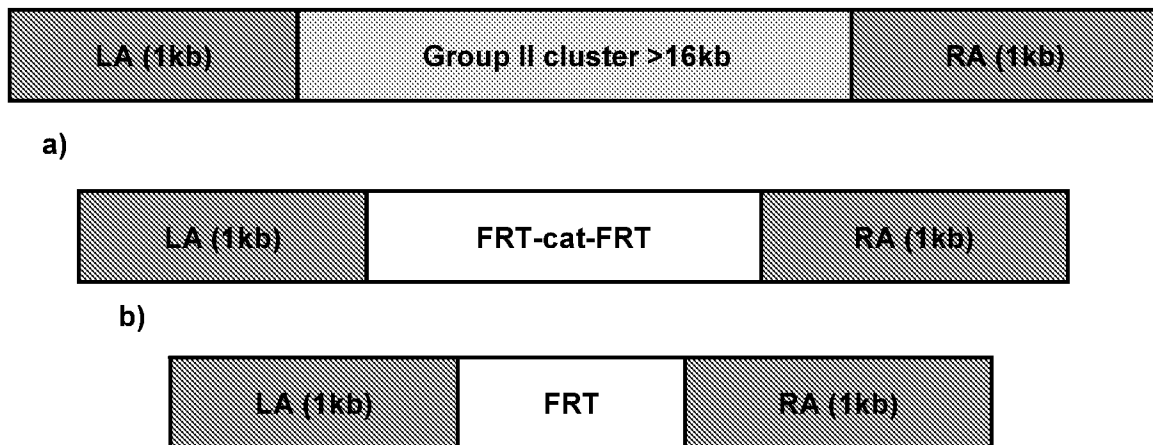


Fig. 9/12

**Fig. 2/12**

3/7

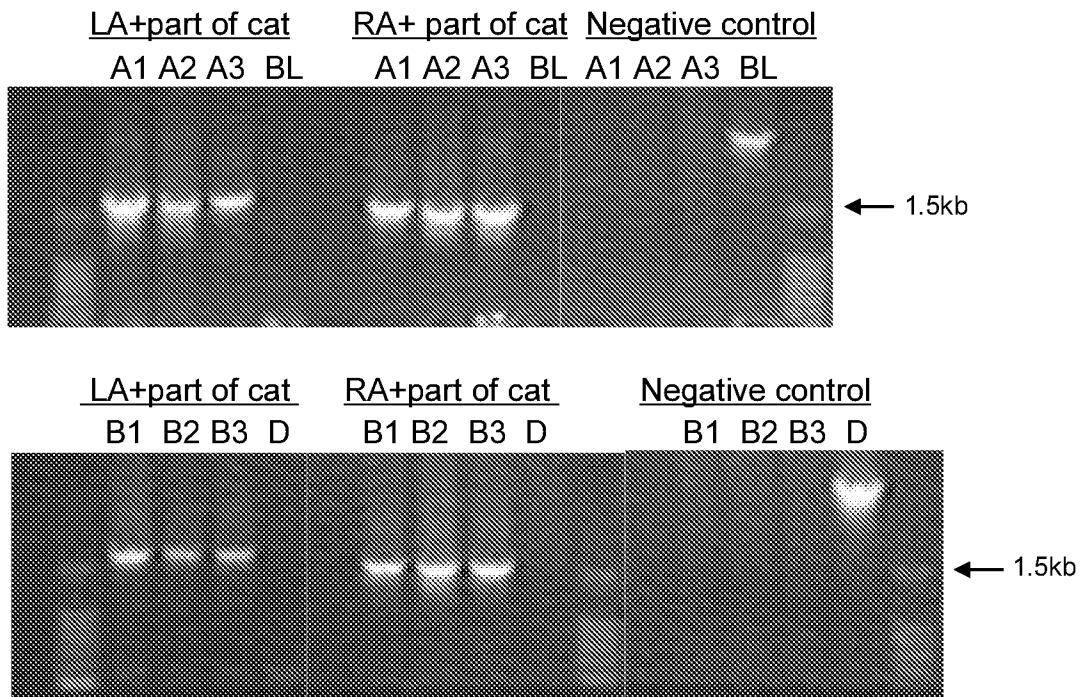


Fig. 3/12

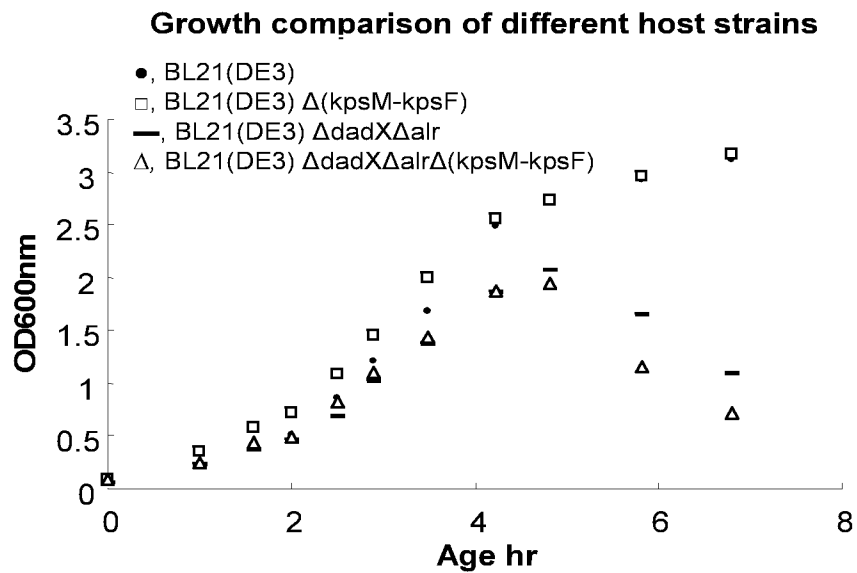


Fig. 4/12

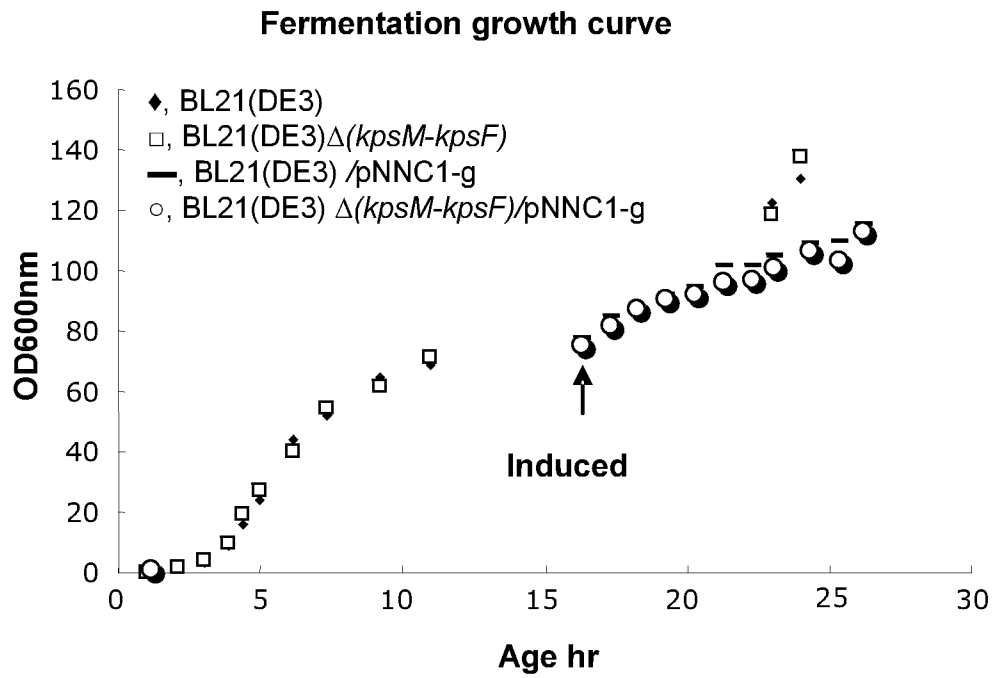


Fig.5/12

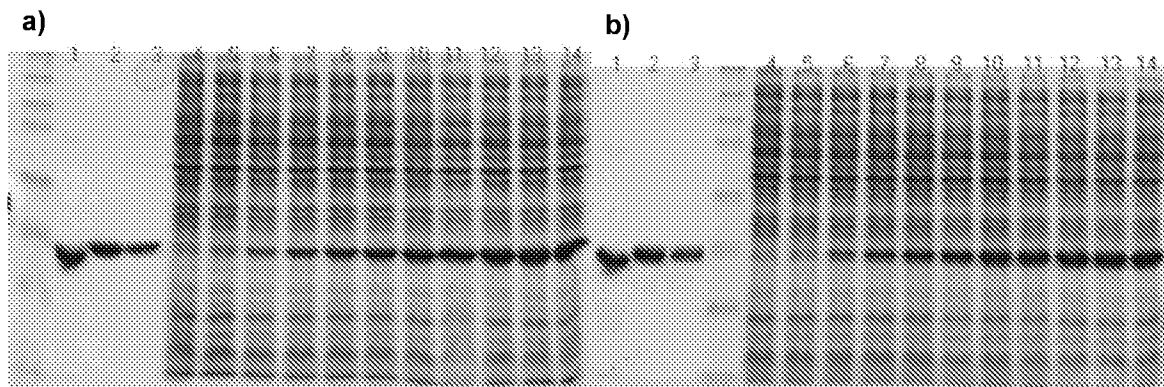


Fig. 6/12

5/7

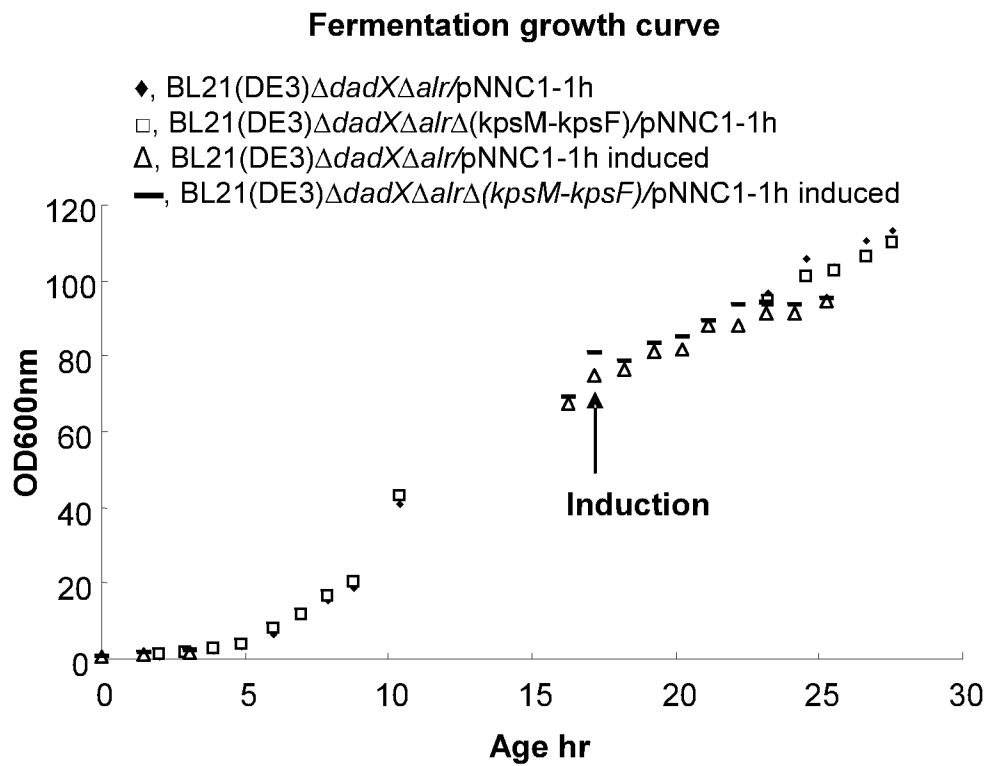


Fig.7/12

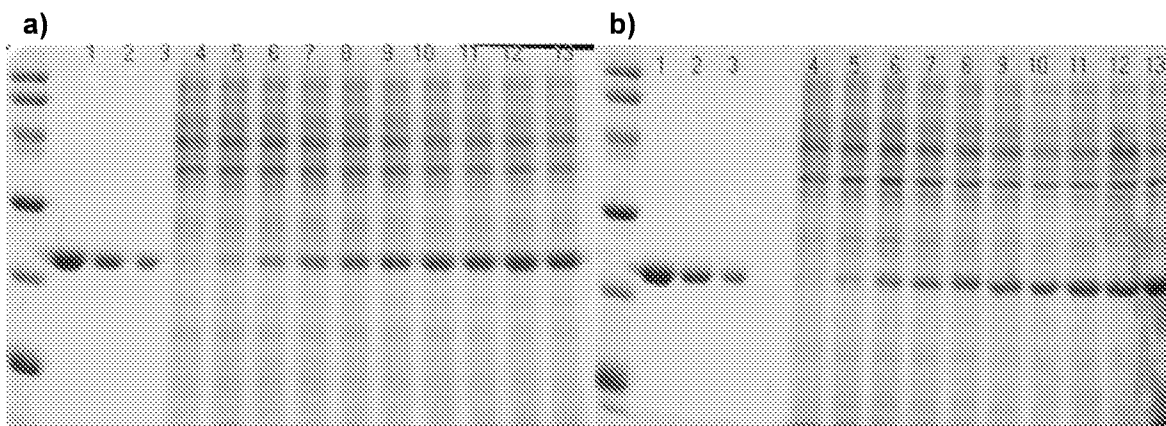


Fig. 8/12

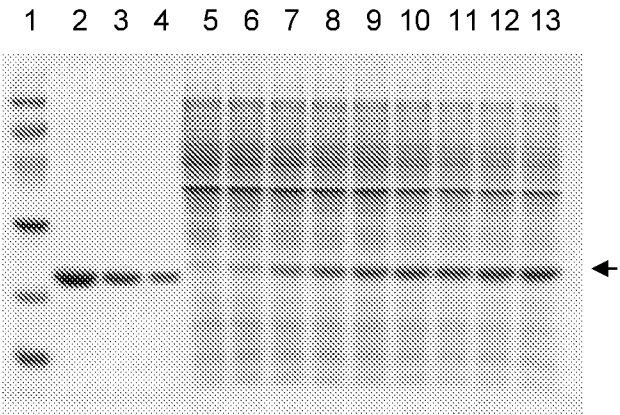


Fig. 10/12

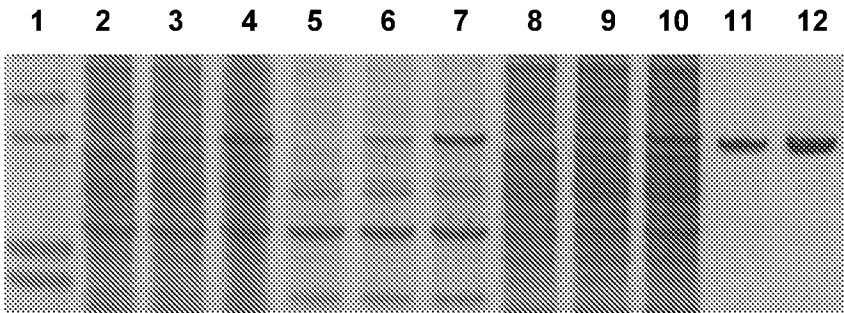


Fig. 11/12

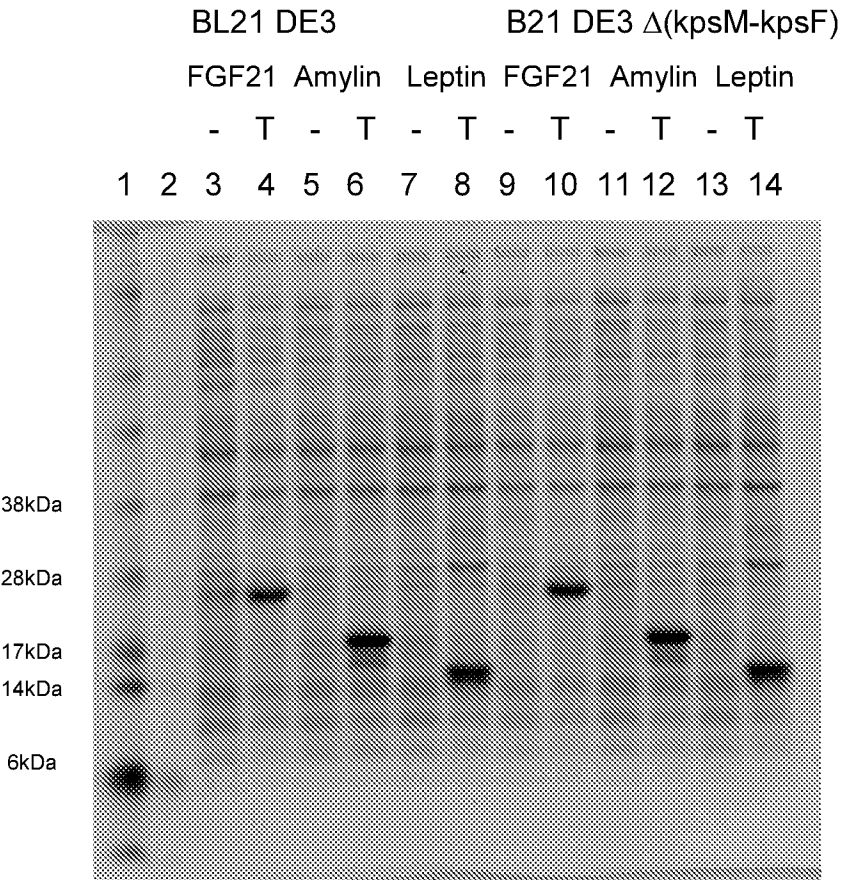


Fig. 12/12

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2009/064903

A. CLASSIFICATION OF SUBJECT MATTER INV. C07K14/61 C12N1/16 C12N1/20		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C07K C12N C12R		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, BIOSIS, EMBASE, FSTA		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ANDREISHCHEVA ET AL: "Escherichia coli BL21(DE3) chromosome contains a group II capsular gene cluster" GENE, ELSEVIER, AMSTERDAM, NL, vol. 384, 24 November 2006 (2006-11-24), pages 113-119, XP005764907 ISSN: 0378-1119 cited in the application	1
Y	page 116, column 1, paragraph 3 page 116, column 2 - page 117, column 1, paragraph 1 page 113 - page 119 page 118, column 2, paragraph 1 ----- -/--	2-15
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents :		
A document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *&* document member of the same patent family		
Date of the actual completion of the international search 24 March 2010	Date of mailing of the international search report 07/04/2010	
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer Bonello, Steve	

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2009/064903

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MUEHLDOERFER I ET AL: "GENETIC ASPECTS OF ESCHERICHIA COLI VIRULENCE" MICROBIAL PATHOGENESIS, ACADEMIC PRESS LIMITED, NEW YORK, NY, US, vol. 16, 1 March 1994 (1994-03-01), pages 171-181, XP000983208 ISSN: 0882-4010 page 176, column 2 -----	2-7
Y	WO 97/49416 A (VIRGINIA TECH INTELL PROP [US]) 31 December 1997 (1997-12-31) page 2, line 32 - page 3, line 2 claims 1,4 -----	2-7
Y	DATSENKO KIRILL A ET AL: "One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products". PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE, WASHINGTON, DC.; US, vol. 97, no. 12, 6 June 2000 (2000-06-06), pages 6640-6645, XP002210218 ISSN: 0027-8424 page 6640, column 2, paragraph 3 page 6640, column 2, paragraph 4 page 6645 the whole document -----	8-11
Y	WO 02/14468 A (PHAGE BIOTECHNOLOGY CORP [US]) 21 February 2002 (2002-02-21) page 15, line 30 - page 16, line 16 -----	12-14
Y	LOPEST D H J ET AL: "Amyloidogenicity and cytotoxicity of recombinant mature human islet amyloid polypeptide (rhIAPP)" JOURNAL OF BIOLOGICAL CHEMISTRY 20041008 AMERICAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BIOLOGY INC. US, vol. 279, no. 41, 8 October 2004 (2004-10-08), pages 42803-42810, XP002574331 page 42803, column 1 - page 42804, column 2 page 42805; figures 1,2 page 42810, column 1, paragraph 3 -----	11,15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2009/064903

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9749416	A	31-12-1997	AU 734254 B2	07-06-2001
			AU 3390497 A	14-01-1998
			BR 9710986 A	11-12-2001
			CA 2258168 A1	31-12-1997
			CN 1223586 A	21-07-1999
			CZ 9804189 A3	14-07-1999
			EP 0964689 A1	22-12-1999
			HK 1020004 A1	20-06-2003
			JP 2001522347 T	13-11-2001
			KR 20000022276 A	25-04-2000
			NO 986099 A	13-01-1999
			NZ 333387 A	23-06-2000
			PL 330863 A1	07-06-1999
			RU 2234941 C2	27-08-2004
			SK 177198 A3	13-03-2000
			TR 9802716 T2	21-06-2001
			UA 70287 C2	15-10-2004
			US 6326001 B1	04-12-2001
			US 6086894 A	11-07-2000
			ZA 9705586 A	23-03-1998
WO 0214468	A	21-02-2002	AT 358135 T	15-04-2007
			AU 8491401 A	25-02-2002
			AU 2001284914 B2	09-02-2006
			CA 2419203 A1	21-02-2002
			DE 60127561 T2	13-12-2007
			EP 1309604 A2	14-05-2003
			ES 2284677 T3	16-11-2007
			JP 2004528005 T	16-09-2004