SCALABLE FERMENTATION PROCESS

Inventors: Marcel EMMERLING, Schlieren (CH); Frank Hennecke, Dietikon (CH); Holger Pfrunder, Zurich (CH); Martin Rhiel, Bonstetten (CH); Philipp Steiner, Schlieren (CH)

Assignee: Cytos Biotechnology AG, Zurich-Schlieren (CH)

Filed: Dec. 22, 2011

Related U.S. Application Data
Continuation of application No. 11/921,023, filed on Nov. 26, 2007, now abandoned, filed as application No. PCT/EP2006/062628 on May 24, 2006.

ABSTRACT
This invention provides a robust fermentation process for the expression of a capsid protein of a bacteriophage which is forming a VLP by self-assembly, wherein the process is scalable to a commercial production scale and wherein the expression rate of the capsid protein is controlled to obtain improved yield of soluble capsid protein. This is achieved by combining the advantages of fed-batch culture and of lactose induced expression systems with specific process parameters providing improved repression of the promoter during the growth phase and high plasmid retention throughout the process.
Figure 1

Fermentation QT0602, Induction with Cofeed
(20% Glycerol + 20% Lactose)
SCALABLE FERMENTATION PROCESS

FIELD OF THE INVENTION

[0001] This invention is related to the field of protein expression and fermentation technology. A process for the efficient expression of recombinant bacteriophage capsid protein in a bacterial host is described. The process leads to high yield of recombinant capsid protein which is capable of forming a virus-like particle (VLP) by self-assembly. Furthermore, the process is scalable from laboratory scale to fermenter volumes larger than 50 litres.

BACKGROUND OF THE INVENTION

[0002] Recent vaccination strategies make use of viruses or virus-like particles (VLPs) to enhance the immune response towards antigens. For example, WO02/056905 demonstrates the utility of VLPs as a carrier to present antigens linked thereto in a highly ordered repetitive array. Such antigen arrays can cause a strong immune response, in particular antibody responses, against the linked antigen and are even capable of breaking the immune system’s inherent tolerance towards self antigens. Such antigen arrays are therefore useful in the production of vaccines for the treatment of infectious diseases and allergies as well as for the efficient induction of self-specific immune responses, e.g. for the treatment of cancer, rheumatoid arthritis and various other diseases.

[0003] As indicated in WO02/056905 capsid proteins of bacteriophages are particularly suited as antigen carrier. They have been shown to efficiently self-assemble into VLPs upon expression in a bacterial host (Kasilein et al. 1983, Gene 23:245-254; Kozlovskaya et al. 1986, Dokl. Akad. Nauk SSSR 287:452-455). Moreover, capsid proteins of bacteriophages such as derived from fr (Pushk et al. 1993, Protein Engineering 6(8):883-891), Qβ (Kozlovskaya et al. 1993, Gene 137:133-137; Ciliens et al. 2000, FEBS Letters 2471:1-4; Vasiljeva et al. 1998, FEBS Letters 431:7-11) and MS-2 (WO92/13081; Mastico et al. 1993, Journal of General Virology 74:541-548; Heald et al. 2000, Vaccine 18:251-258) have been produced in bacterial hosts using inducible promoters such as the trp promoter or a trpT fusion (in the case of fr and Qβ) or the tae promoter using IPTG as inducer substance (in the case of MS-2). The use of inducible promoters is beneficial, to avoid possible toxic effects of the recombinant capsid protein and the metabolic burden of protein expression which both might reduce the growth of the bacterial expression host and, ultimately, the yield of expressed protein.

[0004] However, the expression systems used so far for the expression of capsid proteins of bacteriophages have been applied in small scale fermentations, i.e. in laboratory scale and small batch cultures with volumes of typically clearly below 1 litre. An scale up of these systems comprising volumes of 50 litre and more is expected to diminish in a great extent the respective capsid protein yield due to increased promoter leakage and/or lowered plasmid retention.

[0005] A further problem associated with commercially desired high-level expression and rapid accumulation of recombinant capsid proteins of bacteriophages is the formation of incorrectly folded protein species and the formation of so called inclusion bodies, i.e. protein aggregates, which are insoluble and which may hamper further downstream processes. Thus, for bacteriophage MS-2 coat protein the formation of protein aggregates and of protein species which lost their ability to self-assemble to VLPs have been reported when the protein was expressed under the control of the strong T7 promoter after IPTG induction using the pET expression system (Peabody & Al-Bitar 2001, Nucleic Acid Research 29(22):e113).

[0006] High expression rates of the recombinant capsid protein may therefore have a negative impact on the yield of correctly assembled VLPs. The production of VLP-based vaccines in a commercial scale requires, therefore, the establishment of an efficient, and in particular scalable fermentation process for the expression of recombinant capsid protein of bacteriophages leading to a product of constant quality and purity having the capability of self-assembling into VLPs, whereby the formation of insoluble fractions of the capsid protein is minimized or avoided.

[0007] Therefore, it is an object of the present invention to provide a process for expression of a recombinant capsid protein of a bacteriophage which avoids or minimizes the disadvantage or disadvantages of the prior art processes, and in particular, which is scalable to a commercial scale and still leading to a product of constant quality and purity and the capability of self-assemble to VLPs, and wherein the formation of insoluble fraction of the capsid protein is minimized or avoided.

SUMMARY OF THE INVENTION

[0008] The invention relates to a process for expression of a recombinant capsid protein of a bacteriophage, or a mutant or fragment thereof being capable of forming a VLP by self-assembly, said process comprising the steps of: a) introducing an expression plasmid into a bacterial host, wherein said expression plasmid comprises an expression construct, wherein said expression construct comprises (i) a first nucleotide sequence encoding said recombinant capsid protein, or mutant or fragment thereof, and (ii) a promoter being inducible by lactose; b) cultivating said bacterial host in a medium comprising a major carbon source; wherein said cultivating is performed in batch culture and under conditions under which said promoter is repressed by lacI, wherein said lacI is overexpressed by said bacterial host; c) feeding said batch culture with said major carbon source; and d) inducing said promoter with an inducer, wherein preferably said feeding of said batch culture with said major carbon source is continued.

[0009] This invention provides a robust fermentation process for the expression of a capsid protein of a bacteriophage which is forming a VLP by self-assembly, wherein the process is scalable to a commercial production scale and wherein the expression rate of the capsid protein leads to improved yield of soluble capsid protein. This is, in particular, achieved by improved repression of the promoter during the growth phase and high plasmid retention throughout the process. The expression system further avoids formation of insoluble protein aggregates by limiting the maximum expression rate occurring during the production phase.

[0010] In a preferred embodiment said bacteriophage is a RNA bacteriophage. More preferably, said RNA bacteriophage is selected from the group consisting of: a) bacteriophage Qβ; b) bacteriophage AP205; c) bacteriophage fr; d) bacteriophage G1A; e) bacteriophage SP; f) bacteriophage MS2; g) bacteriophage M11; h) bacteriophage MX1; i) bacteriophage NL95; j) bacteriophage f2; k) bacteriophage PF7 and l) bacteriophage R17. Preferably, said RNA bacteriophage is Qβ. More preferably said recombinant capsid protein comprises or alternatively consists of a nucleic acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID
NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, and SEQ ID NO:11. Still more preferably said recombinant capsid protein comprises SEQ ID NO:5, most preferably said recombinant capsid protein consists of SEQ ID NO:5.

[0011] In a further preferred embodiment said recombinant capsid protein comprises or alternatively consists of an amino acid sequence selected from the group consisting of SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:14. More preferably said recombinant capsid protein comprises SEQ ID NO:12, most preferably said recombinant capsid protein consists of SEQ ID NO:12.

[0012] In another embodiment of the present invention, said expression construct comprises a first stop codon, and wherein said first stop codon is TAA, and wherein preferably said TAA is located directly 3' of said first nucleotide sequence.

[0013] In a further embodiment said expression construct comprises a first stop codon and a second stop codon, wherein said first stop codon is located directly 3' of said first nucleotide sequence and wherein said second stop codon is located directly 3' of said first stop codon, and wherein at least one of said first or second stop codon is TAA.

[0014] In a further embodiment said expression construct comprises a first nucleotide sequence and a second nucleotide sequence, wherein said first nucleotide sequence is encoding a recombinant capsid protein, preferably QP CP or a mutant or fragment thereof, and wherein said second nucleotide sequence is encoding any other protein, preferably the QP A1 protein or a mutant or fragment thereof, and wherein said first and said second nucleotide sequence are separated by exactly one sequence stretch comprising at least one TAA stop codon. In a preferred embodiment said expression construct comprises or alternatively consists of the nucleotide sequence of SEQ ID NO:6.

[0015] In a further embodiment said expression plasmid comprises or, more preferably, consists of the nucleotide sequence of SEQ ID NO:1.

[0016] In one embodiment of the invention said promoter is selected from the group consisting of the a) tac promoter; b) trc promoter; c) t'c promoter; d) lac promoter; e) lacUV5 promoter; f) Psyn promoter; g) lpp' promoter; h) lpp-lac promoter; i) T7-lac promoter; j) 13-lac promoter; k) T5-lac promoter; and l) a promoter having at least 50% sequence homology to SEQ ID NO:2. In a preferred embodiment said promoter has at least 50%, 60%, 70%, 80%, 90, or 95%, preferably 98 to 100%, most preferably 99% sequence homology to SEQ ID NO:2. In a further preferred embodiment said promoter is selected from the group consisting of tac promoter, trc promoter and tac promoter. Even more preferably said promoter is the tac promoter. Most preferably said promoter comprises or alternatively consists of the nucleotide sequence of SEQ ID NO:2.

[0017] In one embodiment said major carbon source is glucose or glycerol, preferably glycerol.

[0018] In one embodiment said feeding of said batch culture is performed with a flow rate, wherein said flow rate increases with an exponential coefficient and wherein preferably said exponential coefficient μ is below μmax.

[0019] In a further embodiment said inducing of said promoter is performed by co-feeding said batch culture with said inducer, preferably lactose and said major carbon source, preferably glycerol, at a constant flow rate.

[0020] In a further embodiment said inducing of said promoter is performed by co-feeding said batch culture with said inducer, preferably lactose and said major carbon source, preferably glycerol, at an increasing flow rate.

[0021] In a further embodiment said inducer is lactose, wherein preferably said lactose and said major carbon source are co-fed to said batch culture in a ratio of about 2:1 to 1:4 (w/w).

[0022] In a further embodiment said inducer is IPTG wherein preferably the concentration of said IPTG said medium is 0.001 to 5 mM, preferably 0.001 to 1 mM, more preferably 0.005 to 1 mM, still more preferably 0.005 to 0.5 mM. In a very preferred embodiment said concentration of IPTG is about 0.01 mM, most preferably 0.01 mM.

[0023] In one embodiment said lacI is overexpressed by said bacterial host, wherein said overexpression is caused by lacI or lacQ1, preferably by lacI. In one embodiment said bacterial host comprises said lacI gene or said lacQ1 gene, preferably said lacI gene on its chromosome. In a further preferred embodiment said bacterial host comprises said lacI gene or said lacQ1 gene, preferably said lacI gene on a plasmid, preferably on a high copy number plasmid. In a further preferred embodiment said bacterial host comprises said lacI gene or said lacQ1 gene, preferably said lacI gene on said expression plasmid.

[0024] In one embodiment said bacterial host is selected from the group consisting of the strains E. coli RB791, E. coli DH120 and E. coli Y1088. Preferably said bacterial host is E. coli RB791.

[0025] In one embodiment said bacterial host comprises β-galactosidase activity.

[0026] In one embodiment said cultivating and said feeding of said batch culture and said inducing of said promoter is performed at a temperature which is below the optimal growth temperature of said bacterial host. Preferably said temperature is between 25° C. and 35° C., more preferably between 25 and 33° C., even more preferably between 27 and 32° C., still more preferably between 28 and 31° C. Even more preferably said temperature is about 30° C., most preferably said temperature is 30° C.

[0027] In one embodiment said cultivating and said feeding of said batch culture is performed at a temperature which is below the optimal growth temperature of said bacterial host, wherein preferably said temperature is between 25° C. and 35° C., more preferably between 25 and 33° C., even more preferably between 27 and 32° C., still more preferably between 28 and 31° C., even more preferably said temperature is about 30° C., most preferably said temperature is 30° C., and said inducing of said promoter is performed at the optimal growth temperature of the bacterial host, preferably at about 37° C.

[0028] In one embodiment said cultivating and said feeding of said batch culture and said inducing of said promoter is performed in the absence of an antibiotic.

[0029] In a specific embodiment said expression plasmid comprises or alternatively consists of the nucleotide sequence of SEQ ID NO:1, said major carbon source is glycerol, said feeding of said batch culture is performed with a flow rate, wherein said flow rate increases with an exponential coefficient μi, and wherein said exponential coefficient μi is below μmax, said inducing of said promoter by co-feeding said batch culture is performed with a constant flow rate, wherein lactose and glycerol are co-fed to the batch culture in a ratio of about 2.1 to about 1.4 (w/w), preferably about 1:1 to about 1:4 (w/w), most preferably about 1:3 (w/w), and wherein said cultivating and feeding of said batch culture and said inducing of said promoter is performed at a temperature between 27 and 32° C., preferably about 30° C., most preferably 30° C.
In a further specific embodiment said expression plasmid comprises or alternatively consists of the nucleotide sequence of SEQ ID NO: 30, said major carbon source is glycerol, said feeding of said batch culture is performed with a flow rate, wherein said flow rate increases with an exponential coefficient μ, and wherein said exponential coefficient μ is below μ<sub>max</sub>, said inducing of said promoter by co-feeding said batch culture is performed with a constant flow rate, wherein lactose and said major carbon source are co-fed to the batch culture in a ratio of about 2:1 to about 1:4 (w/w), preferably about 1:1 to about 1:4 (w/w), most preferably about 1:3 (w/w), and wherein said cultivating and feeding of said batch culture and said inducing of said promoter is performed at a temperature between 27 and 32° C., preferably about 30° C., most preferably 30° C.

**DESCRIPTION OF THE FIGURES**

**FIG. 1:** Fermentation profile with pTac-nSD-Qbmut (SEQ ID NO:1) in Rb791 in 2 l culture. Co-feeding during production phase was performed with medium containing 20% glycerol and 20% lactose. Shown are glycerol concentration [g/l] (circles), lactose concentration [g/l] (triangles); β-Gal activity [U/ml*OD<sub>1</sub>l] (squares) and OD600 (diamonds) plotted against the process time [h].

**DETAILED DESCRIPTION OF THE INVENTION**

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs.

“about”: within the meaning of the present application the expression about shall have the meaning of ±10%. For example about 100 shall mean 90 to 110.

“promoter which is inducible by lactose”: as used herein refers to a promoter which comprises regulatory elements of the lac operon. Such promoters are repressed by lac and can be induced by lactose or the synthetic inducer IPTG. The skilled person is aware that induction of a promoter by lactose requires β-galactosidase activity in the bacterial host.

“located directly 3′”: a nucleotide sequence N2 which is located directly 3′ of another nucleotide sequence N1 refers to a continuous sequence having the conformation 5′-N1-N2-3′ wherein N1 and N2 are directly connected and not separated by additional sequence elements.

“sequence stretch”: as used herein the term “sequence stretch” refers to a continuous nucleotide sequence which consists of less than 50, preferably less than 20, more preferably less than 10, even more preferably less than 5 nucleotides. In a further preferred embodiment the sequence stretch comprises or alternatively consists of at least one, preferably one, TAA stop codon. In another embodiment the sequence stretch comprises or alternatively consists of at least one, preferably one, TAA and at least one, preferably one, TGA stop codon. In further preferred embodiment the sequence stretch comprises or alternatively consists of SEQ ID NO: 32.

“bacterial host”: as used herein the term “bacterial host” refers to a bacterial organism which is hosting or capable of hosting an expression plasmid of the invention, wherein “hosting” involves the replication of the expression plasmid and maintenance of the expression plasmid during cell division.

“culture”: in the context of the instant invention a “culture” comprises a bacterial host in a medium (“bacterial culture”), wherein typically said medium is supporting the growth of said bacterial host.

“batch culture” as used herein relates to a culture, i.e. a bacterial host in a medium, wherein said culture constitutes a closed system, i.e. typically and preferably no addition or removal of medium takes place during the cultivation time. Therefore, in contrast to a continuous culture, typically and preferably the density of the bacterial host in the batch culture continuously increases with progressing cultivation time. Batch culture does not exclude the addition of compounds required for the control of the process, such as, for example, inducer, oxygen, and alkali or acid to control the pH.

“feed batch culture”: as used herein is a culture which is supplied with additional medium comprising a substrate, preferably the major carbon source of the bacterial host (feed or co-feed medium). In the context of the application this process is referred to by the terms “feeding said batch culture” (medium comprises the major carbon source) and “co-feeding said batch culture” (medium comprises the major carbon source and the inducer, preferably lactose). Typically and preferably, no removal of medium except for analytical purposes takes place during cultivation time of a feed batch culture.

“Preculture”: a culture, preferably a batch culture, which is used to produce the inoculum for a culture of a larger volume, e.g. the culture in which the recombinant capsid protein is produced (production culture). A preculture can be performed in two or more steps, wherein a second preculture is inoculated with a first preculture etc. to produce a sufficiently large inoculum for the production culture. The first and/or subsequent precultures may comprise an antibiotic to improve plasmid stability.

“substrate”: as used herein refers to, a compound in the culture medium which contributes to the carbon and energy supply of the bacterial host. The terms “substrate” therefore encompasses any compound contained in the medium contributing to the carbon supply of the bacterial host. Typical substrates for bacteria are sugar, starch, glycerol, acetate and any other organic compound which can be metabolized by bacteria. Therefore, the term “substrate” includes the major carbon source but also, for example, lactose.

“Major carbon source” as Used herein refers to the compound in the culture medium which contributes most to the carbon and energy supply of the bacterial host during the growth phase. The major carbon source thus is the major substrate of the bacterial host. The major carbon source is typically a sugar such as sucrose or glucose, or glycerol, and preferably glucose or glycerol. Though lactose could in principal act as a major carbon source for a bacterial host, in the context of the instant invention the term “major carbon source” typically and preferably does not include lactose.

**PHASES OF THE PROCESS OF THE INVENTION**

The process of the invention is characterised by different phases which refer to different physiological conditions of the bacterial host with respect to its growth and the repression induction status of the expression construct.
with said inducing said promoter with an inducer. The growth phase can be further divided in a “batch phase” and a “feed phase”. Said batch phase is initiated by said cultivating said bacterial host in a medium. The batch phase comprised a “lag phase” during which the bacterial host is not yet growing or growing with a non-exponential rate, typically and preferably a linear rate. The growth phase further comprises an “exponential growth phase” which directly follows the lag phase. No feeding of said culture takes place during the batch phase, thus the exponential growth phase is terminated by the consumption of the substrate by the bacterial host. The growth phase further comprises a “feed phase” which is directly following the batch phase and which is initiated by said feeding of said batch culture with said major carbon source. The feed phase is characterised by a growth rate of the bacterial host which is directly dependent on the flow rate of the feed medium containing the major carbon source.

**0046** “production phase”: The growth phase is followed by the production phase which is initiated by said inducing said promoter with an inducer, wherein typically and preferably said feeding of mid batch culture with said major carbon source is continued.

**0047** “Conditions under which the promoter is repressed”: it is to be understood that the repression of a promoter is an equilibrium of formation and dissociation of the repressor-operator complex and that even stringently repressed promoters may show a certain expression rate also in the absence of their inducer. Therefore, as used within the Application the term “conditions under which the promoter is repressed” relates to conditions, wherein at the end of the growth phase, i.e. directly before the addition of inducer to the culture, the recombinant capsid protein is expressed to a level which does not exceed a concentration in the medium of 200 mg/l, preferably 150 mg/l, more preferably 100 mg/l, as determined by the HPLC method of Example 17. Most preferably, the concentration of the recombinant protein is below the detection level of said method.

**0048** “Inducer”: within the meaning of the in invention the term “inducer” relates to any substance which directly or indirectly interacts with an inducible promoter and thereby facilitates expression from said promoter; for example, inducers of “a promoter inducible by lactose”, such as the lac or tac promoter, are IPTG, lactose and allo lactose.

**0049** “Coat protein”/“capsid protein”: The term “coat protein” and the interchangeably used term “capsid protein” within this application, refers to a viral protein, preferably a subunit of a natural capsid of a virus, preferably of a RNA bacteriophage, which is capable of being incorporated into a virus capsid or a VLP. For example, the specific gene product of the coat protein gene of RNA bacteriophage QB is referred to as “QB CP”, whereas the “coat proteins” or “capsid proteins” of bacteriophage QB comprise the “QB CP” as well as the A1 protein.

**0050** “Recombinant capsid protein”: A capsid protein which is synthesised by a recombinant host cell.

**0051** “Polypeptide”: As used herein, the term “polypeptide” refers to a polymer composed of amino acid residues, generally natural amino acid residues, linked together through peptide bonds. Although a polypeptide may not necessarily be limited in size, the term polypeptide is often used in conjunction with peptide of a size of about ten to about 50 amino acids.

**0052** “Protein”: As used herein, the term protein refers to a polypeptide generally of a size of above 20, more particularly of above 50 amino acid residues. Proteins generally have a defined three dimensional structure although they do not necessarily need to, and are often referred to as folded, in opposition to peptides and polypeptides which often do not possess a defined three-dimensional structure, but rather can adopt a large number of different conformations, and are referred to as unfolded.

**0053** “Recombinant host cell”: As used herein, the term “recombinant host cell” refers to a host cell into which one or more nucleic acid molecules of the invention have been introduced.

**0054** “Recombinant VLP”: The term “recombinant VLP”, as used herein, refers to a VLP that is obtained by a process which comprises at least one step of recombinant DNA technology. The term “VLP recombinantly produced”, as used herein, refers to a VLP that is obtained by a process which comprises at least one step of recombinant DNA technology. Thus, the terms “recombinant VLP” and “VLP recombinantly produced” are interchangeably used herein and should have the identical meaning.

**0055** “RNA-bacteriophage”: As used herein, the term “RNA-bacteriophage” refers to RNA viruses infecting bacteria, preferably to single-stranded positive-sense RNA viruses infecting bacteria.

**0056** “Virus-like particle (VLP)”: as used herein, the term “virus-like particle” refers to a structure resembling a virus particle or it refers to a non-replicative or non-infec tious, preferably a non-replicative and non-infectious virus particle, or it refers to a non-replicative or non-infectious, preferably a non-replicative and non-infectious structure resembling a virus particle, preferably a capsid of a virus. The term “non-replicative”, as used herein, refers to being incapable of replicating the genome comprised by the VLP. The term “non-infectious”, as used herein, refers to being incapable of entering the host cell. Preferably a virus-like particle in accordance with the invention is non-replicative and/or non-infectious since it lacks all or part of the viral genome or genome function. Typically a virus-like particle lacks all or part of the replicative and infectious components of the viral genome. A virus-like particle in accordance with the invention may contain nucleic acid distinct from their genome. A typical and preferred embodiment of a virus-like particle in accordance with the present invention is a viral capsid such as the viral capsid of the corresponding virus, bacteriophage, preferably RNA-phages. The terms “viral capsid” or “capsid”, refer to a macromolecular assembly composed of viral protein subunits. Typically, there are 60, 120, 180, 240, 300, 360 and more than 360 viral protein subunits. Typically and preferably, the interactions of these subunits lead to the formation of viral capsid or viral-capsid like structure with an inherent repetitive organisation, wherein said structure is, typically, spherical or tubular. For example, the capsids of RNA bacteriophages or HBcAg have a spherical form of icosahedral symmetry.

**0057** “Virus-like particle of a RNA bacteriophage”: As used herein, the term “virus-like particle of a RNA bacteriophage” refers to a virus-like particle comprising, or preferably consisting essentially of or consisting of coat proteins, mutants or fragments thereof, of a RNA bacteriophage. In addition, virus-like particle of a RNA bacteriophage resembling the structure of a RNA bacteriophage, being non-replicative and/or non-infectious, and lacking at least the gene or
genes encoding for the replication machinery of the RNA bacteriophage, and typically also lacking the gene or genes encoding the protein or proteins responsible for viral attachment to or entry into the host. Preferred VLPs derived from RNA bacteriophages exhibit icosahedral symmetry and consist of 180 subunits. A preferred method to render a virus-like particle of a RNA bacteriophage non replicative and/or non-infectious is by genetic manipulation.

[0050] one, a, or an. When the terms “one,” “a,” or “an” are used in this disclosure, they mean “at least one” or “one or more,” unless otherwise indicated.

[0059] “Sequence identity”: The amino acid sequence identity of polypeptides can be determined conventionally using known computer programs such as the Bestfit program. When using Bestfit or any other sequence alignment program, preferably using Bestfit, to determine whether a particular sequence is, for instance, 95% identical to a reference amino acid sequence, the parameters are set such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed. This aforementioned method in determining the percentage of identity between polypeptides is applicable to all proteins, polypeptides or a fragment thereof disclosed in this invention.

[0060] “Sequence homology”: The homology of nucleotide sequences can for example be determined by the program blast which is an implementation of the BLAST algorithm, preferably using the default settings of the software.

[0061] “Fragment of a protein”, in particular fragment of a recombinant protein or recombinant coat protein, as used herein, is defined as a polypeptide, which is of at least 70%, preferably at least 80%, more preferably at least 90%, even more preferably at least 95% the length of the wild-type recombinant protein, or coat protein, respectively and which preferably retains the capability of forming VLP. Preferably the fragment is obtained by at least one internal deletion, at least one truncation or at least one combination thereof. Further preferably the fragment is obtained by at least one internal deletion, at least one truncation or at least one combination thereof. Further preferably the fragment is obtained by at least one truncation or at least one combination thereof. Most preferably the fragment is obtained by exactly one internal deletion, exactly one truncation or by a combination thereof.

[0062] The term “fragment of a recombinant protein” or “fragment of a coat protein” shall further encompass polypeptide, which has at least 80%, preferably 90%, even more preferably 95% amino acid sequence identity with the “fragment of a recombinant protein” or “fragment of a coat protein”, respectively, as defined above and which is preferably capable of assembling into a virus-like particle.

[0063] The term “mutant recombinant protein” or the term “mutant of a recombinant protein” as interchangeably used in this invention, or the term “mutant coat protein” or the term “mutant of a coat protein”, as interchangeably used in this invention, refers to a polypeptide having an amino acid sequence derived from the wild type recombinant protein, or coat protein, respectively, wherein the amino acid sequence is at least 80%, preferably at least 85%, 90%, 95%, 97%, or 99% identical to the wild type sequence and preferably retains the ability to assemble into a VLP.

[0064] The invention is related to an efficient fermentation process for the production of a VLP of a bacteriophage. The process is improved with respect to yield of the VLP and can be scaled up to a commercial production scale. The process encompasses the expression of recombinant capsid protein of bacteriophages in a bacterial host under conditions which allow the capsid protein to self-assemble into VLPs spontaneously.

[0065] Specific examples of VLPs which can be produced by the process of the invention are VLPs of bacteriophages, preferably RNA bacteriophages. In one preferred embodiment of the invention, the virus-like particle of the invention comprises, consists essentially of, or alternatively consists of, recombinant coat proteins, mutants or fragments thereof, of RNA phage. Preferably, the RNA-phage is selected from the group consisting of a) bacteriophage Qβ; b) bacteriophage R17; c) bacteriophage fr; d) bacteriophage GA; e) bacteriophage SP; f) bacteriophage MS2; g) bacteriophage M11; h) bacteriophage MX1; i) bacteriophage NL95; k) bacteriophage f2; l) bacteriophage PP7 and m) bacteriophage AP205.

[0066] In one preferred embodiment of the invention VLPs are produced comprising coat protein, mutants or fragments thereof, of RNA bacteriophages, wherein the coat protein has an amino acid sequence selected from the group consisting of: (a) SEQ ID NO:5 referring to Qβ CP; (b) a mixture of SEQ ID NO:5 and SEQ ID NO:15 (Qβ A1 protein); (c) SEQ ID NO:16 (R17 capsid protein); (d) SEQ ID NO:17 (fr capsid protein); (e) SEQ ID NO:18 (GA capsid protein); (f) SEQ ID NO:19 (SP capsid protein); (g) a mixture of SEQ ID NO:19 and SEQ ID NO:20; (h) SEQ ID NO:21 (MS2 capsid protein); (i) SEQ ID NO:22 (M11 capsid protein); (j) SEQ ID NO:23 (MX1 capsid protein); (k) SEQ ID NO:24 (NL95 capsid protein); (l) SEQ ID NO:25 (f2 capsid protein); (m) SEQ ID NO:26 (PP7 capsid protein); and (n) SEQ ID NO:12 (AP205 capsid protein).

[0067] Upon expression in E. coli, the N-terminal methionine of Qβ coat protein is usually removed (Stoll, E. et al., J. Biol. Chem. 252:990-993 (1977)). VLP composed of Qβ coat proteins where the N-terminal methionine has not been removed, or VLPs comprising a mixture of Qβ coat proteins where the N-terminal methionine is either cleaved or present are also within the scope of the present invention.

[0068] In one preferred embodiment of the invention, the VLP is a mosaical VLP comprising or alternatively consisting of more than one amino acid sequence, preferably two amino acid sequences, of coat proteins, mutants or fragments thereof, of a RNA bacteriophage.

[0069] In one very preferred embodiment, the VLP comprises or alternatively consists of two different coat proteins of a RNA bacteriophage, said two coat proteins have an amino acid sequence of SEQ ID NO: 5 and SEQ ID NO:15, or of SEQ ID NO:19 and SEQ ID NO:20.

[0070] In preferred embodiments of the present invention, the produced VLP comprises, or alternatively consists essentially of, or alternatively consists of recombinant coat proteins, mutants or fragments thereof, of the RNA-bacteriophage Qβ, fr, AP205 or GA.

[0071] In one preferred embodiment, the VLP is a VLP of RNA-phage Qβ. The capsid or virus-like particle of Qβ shows an icosahedral phage-like capsid structure with a diameter of 25 nm and T=3 quasi symmetry. The capsid contains 180 copies of the coat protein, which are linked in covalent pentamers and hexamers by disulfide bridges (Golmohammadi, R. et al., Structure 4:543-5554 (1996)).
[0072] Preferred virus-like particles of RNA bacteriophages, in particular of Qβ and frn in accordance of this invention are disclosed in WO 02/056905, the disclosure of which is herewith incorporated by reference in its entirety. Particular Example 18 of WO 02/056905 gave detailed description of preparation of VLP particles from Qβ.

[0073] In another preferred embodiment, the VLP is a VLP of RNA bacteriophage AP205. Assembly-competent mutant forms of AP205 VLPs, including AP205 coat protein with the substitution of proline at amino acid 5 to threonine, may also be used in the practice of the invention and leads to other preferred embodiments of the invention. WO 2004/007538 describes, in particular in Example 1 and Example 2, how to obtain VLP comprising AP205 coat proteins, and hereby in particular the expression and the purification thereto. WO 2004/007538 is incorporated herein by way of reference.

[0074] In one preferred embodiment, the VLP comprises or consists of a mutant coat protein of a virus, preferably a RNA bacteriophage, wherein the mutant coat protein has been modified by removal of at least one lysine residue by way of substitution and/or by way of deletion. In another preferred embodiment, the VLP of the invention comprises or consists of a mutant coat protein of a virus, preferably a RNA bacteriophage, wherein the modified coat protein has been modified by addition of at least one lysine residue by way of substitution and/or by way of insertion. The deletion, substitution or addition of at least one lysine residue allows varying the degree of coupling with an antigen.

[0075] VLPs or capsids of Qβ coat protein display a defined number of lysine residues on their surface, with a defined topology with three lysine residues pointing towards the interior of the capsid and interacting with the RNA, and four other lysine residues exposed to the exterior of the capsid.

[0076] Qβ mutants, of which exposed lysine residues are replaced by arginines are also encompassed by the present invention. Preferably these mutant coat proteins comprise or alternatively consist of an amino acid sequence selected from the group of a) Qβ-240 (SEQ ID NO:7, Lys13→Arg); b) Qβ-243 (SEQ ID NO:8, Asn10→Lys); c) Qβ-250 (SEQ ID NO:9, Lys2→Arg); d) Qβ-251 (SEQ ID NO:10, Lys16→Arg); and e) Qβ-259 (SEQ ID NO:11, Lys2→Arg, Lys16→Arg). The construction, expression and purification of the above indicated Qβ mutant coat proteins, mutant Qβ coat proteins and VLPs and capsids, respectively, are described in WO2005/056905. In particular, this is hereby referred to Example 18 of above mentioned application.

[0077] In a further preferred embodiment the recombinant capsid protein is a capsid protein of bacteriophage AP205 having the amino acid sequence depicted in SEQ ID NO:12 or a mutation thereof, which is capable of forming a VLP, for example the proteins AP205P5T (SEQ ID NO:13) or AP205N14D (SEQ ID NO:14).

[0078] In a very preferred embodiment said recombinant capsid protein is composed of the 133 amino acid coat protein C of E. coli RNA bacteriophage Qβ comprising or preferably consisting of the amino acid sequence depicted in SEQ ID NO:5, wherein preferably said recombinant capsid protein is capable of forming a VLP by self-assembly.

[0079] In one embodiment, the expression construct comprises a first stop codon and a second stop codon, wherein said first stop codon is located directly 3’ of said first nucleotide sequence and wherein said second stop codon is located directly 3’ of said first stop codon, and wherein at least one of said first or second stop codon is TAA. For example, plasmid pTac-nSDAP205 (SEQ ID NO:30) comprises the naturally occurring TAA stop codon as a first stop codon and an additional TGA stop codon directly 3’ of the first stop codon.

[0080] In a preferred embodiment the expression construct comprises a first nucleotide sequence and a second nucleotide sequence, wherein said first nucleotide sequence is encoding a recombinant capsid protein, preferably Qβ CP, or a mutant or fragment thereof, most preferably SEQ ID NO:5, and wherein said second nucleotide sequence is encoding any other protein, preferably the Qβ A1 protein or a mutant or fragment thereof, most preferably SEQ ID NO:15, and wherein said first and said second nucleotide sequence are separated by exactly one sequence stretch comprising at least one TAA stop codon. In one embodiment said TAA stop codon is generated by replacing the naturally occurring stop codon, preferably TGA by the sequence TAA. Alternatively and more preferably said TAA stop codon is generated by replacing the naturally occurring stop codon, preferably TGA by the sequence TAAATGA (SEQ ID NO:32).

[0081] For example, the region of Qβ gene C corresponds to the NCBI GenBank Acc. No. M99039 (nucleotides 46-1062). Gene C contains a first nucleotide sequence encoding the 133-amino acid Qβ coat protein (SEQ ID NO:5) and a second nucleotide sequence encoding the 329-amino acid read through protein A1 (SEQ ID NO:15). Nucleotides 1-399 of SEQ ID NO:6 (nucleotides 46-444 of NCBI GenBank Acc. No. M99039) correspond to said first nucleotide sequence encoding the 133-amino acid Qβ CP. Nucleotides 400 to 402 of SEQ ID NO:6 correspond to the strong TAA stop codon and nucleotides 403 to 405 of SEQ ID NO:6 to the leaky TGA stop codon, which is followed by said second nucleotide sequence (Qβ A1). Surprisingly, it was found that the presence of the nucleotide sequence relating to A1 in the expression construct results in higher RNA stability and, thus, in improved yield of Qβ CP and VLP as compared to a construct wherein the A1 sequence is deleted.

[0082] The expression of a recombinant protein can significantly reduce the growth rate of the bacterial host due to toxic effects of the accumulating protein and due to the metabolic burden caused by the protein synthesis. In particular cell lysis and low plasmid retention may occur. Inducible promoters provide the possibility to separate the growth phase from the production phase of a fermentation process. Inducible promoters are repressed by a repressor molecule during the growth phase of the bacterial host and are induced by exposing the bacterial host to inductive conditions during the production phase. Inducible promoters therefore allow the bacterial host to grow fast, preferably exponentially during the growth phase and to reach high cell densities. Thus, inducible promoters provide for high yield of the expression product at the end of the production phase. Therefore, the usage of inducible promoters for the expression of recombinant protein is preferred.

[0083] A well known example for an inducible promoter is the lac promoter which forms part of the lac operon and which can be induced by addition of lactose or the strong synthetic inducer isopropylthio-β-D-galactosid (IPTG) to the growth medium of the bacterial host Donavan et al. 2000 (Can. J. Microbiol 46:532-541) report on an improved process for the expression of a monoclonal antibody fragment under the control of the lac promoter. Further examples of inducible promoters are provided in table 1 of Makrides 1996 (Microbiological Reviews, p. 512-538).
A typical drawback of expression systems based on inducible promoters is the “leakiness” of the promoter, meaning that the promoter is only insufficiently repressed and causes a certain expression rate of the recombinant protein during the growth phase. This typically leads to a reduced cell density or to plasmid instability, and, as a consequence, to reduced yield of the recombinant protein (Makrides 1996; Microbiological Reviews, p. 512-538). An example of a promoter which is prone to insufficient repression is the VIB promoter which is repressed under high oxygen conditions and induced upon oxygen depletion.

For the purpose of the invention promoters are preferred which are stringently repressed. In one embodiment the promoter is repressed by the repressor lac. Examples of such promoters are disclosed in Makrides 1996 (Microbiol. Rev. 60:512-538), Goldstein & Doi 1995 (Biotechnology Annual Review 1:105-128), Hannig & Makrides 1998 (TIBTECH 16:54-60) and Stevens 2000 (Structures 8, R177-R185). In a preferred embodiment the promoter is inducible by lactose, more preferably it is selected from the group consisting of lac, lacUV5, tac, trc, P_promoter, lpp-lac, T7-lac, T3-lac, and T5-lac. Especially preferred for the purpose of the invention is the tac promoter (SEQ ID NO:2) or a mutation or variant thereof. Within the scope of the invention the mutants or truncated or deleted variants of the tac promoter having a sequence homology with SEQ ID NO:2 which is at least 50%, 60%, 70%, 80, 90, or 95%, preferably 98 to 100%, most preferably 99%. Wherein the promoter strength of such mutated truncated or deleted variant is comparable to that of the promoter of SEQ ID NO:2. The skilled person will be able to determine the promoter strength of a given sequence by comparative expression studies using standard methods. In a specific embodiment of the invention the promoter driving the expression of the recombinant capsid protein comprises or alternatively consists of SEQ ID NO:2. The tac promoter is a fusion product of the -10 region of the lacUV5 promoter and the -35 region of the trp promoter and combines the high transcription efficiency of tip with the regulatory elements of the lac promoter (de Boer et al. 1983, PNAS 80:21-25; Aman et al. 1983 Gene 25:167-178). It provides for sufficiently high expression rates and high protein yield while avoiding the formation of insoluble or incorrectly folded recombinant protein which may occur with stronger promoters, such as the T7 promoter. The tic and the tic promoter are mutated versions of the tac promoter (Brosius et al 1985, The Journal of Biological Chemistry 260 (6):3539-3541). In a further preferred embodiment the promoter is selected from the group consisting of tic, trc and tac.

For the construction of an expression construct for the purpose of the invention the promoter is operably linked to said first nucleotide sequence encoding the recombinant capsid protein via a ribosome binding site (Shine-Dalgarno sequence, SD), typically comprising an ATG start codon at its 3' end. Suitable Shine-Dalgarno sequences for the purpose of the invention are well known in the art (Dalbge et al. 1988, DNA 7 (6):399-405; Ringquist et al. 1992, Mol. Micr. 6:1219-1229). In one embodiment of the invention the expression construct comprises the SD sequence of Dalbge et al. 1988 (DNA 7 (6):399-405) which is depicted in SEQ ID NO:4. In another, preferred, embodiment the expression construct comprises a Shine-Dalgarno sequence of Ringquist et al. 1992 (Mol. Micr. 6:1219-1229, SEQ ID NO:3, nSD). Surprisingly, it was found that SEQ ID NO:3 is particularly suited for the purpose of the invention because it results in improved expression levels and improved yield of recombinant capsid protein. SEQ ID NO:3 is especially suited to enhance the expression of AP205 capsid protein. In a preferred embodiment of the invention the expression construct comprises a Shine-Dalgarno sequence selected from the group consisting of SEQ ID NO:3 and SEQ ID NO:4, preferably said Shine-Dalgarno sequence is SEQ ID NO:3.

Transcriptional terminators are functional elements of expression constructs. The skilled person will be able to choose a suitable terminator sequence form a wide range of sources. In a preferred embodiment of the invention said expression construct comprises a terminator sequence, wherein preferably said terminator sequence is operably linked to said first nucleotide sequence, wherein further preferably said terminator sequence is the rRNB terminator sequence, most preferably SEQ ID NO:28.

For the purpose of plasmid selection the skilled person will typically use an antibiotic resistance marker gene. Examples of antibiotic resistance genes which are widely used in the art and which are suitable for the purpose of the invention are resistance genes against the antibiotics ampicillin, tetracyclin and kanamycin. The use of kanamycin as a selective agent in the frame of a process for the production of a VLP is generally preferred because of the lower allergenic potential of kanamycin as compared to alternative antibiotics and because of the lower safety concerns resulting thereof for the use of the VLP as a vaccine. Furthermore, kanamycin provides better plasmid retention as compared to alternative antibiotics such as ampicillin. The kanamycin 3'-phosphotransferase gene (SEQ ID NO:29) which is derived from the transposon Tn903 is therefore a particularly useful selectable marker gene.

The addition of antibiotics to the medium is generally undesirable in a commercial production process for cost and safety reasons. In the context of the invention antibiotics, preferably kanamycin, are typically and preferably used for the selection of the expression strain. Media used in the production process are essentially free of antibiotics, in particular kanamycin. However, addition of an antibiotic to precultures used to produce the inoculum for the production culture can improve plasmid retention throughout the process (Example 10).

The skilled person will create expression plasmids comprising expression constructs which are useful for the production of VLPs of bacteriophages by combining the genetic elements described above applying standard methods of molecular biology. Particularly useful expression plasmids for the purpose of the invention are pLac-nSDQb-mut (SEQ ID NO:1) for the production of Qβ VLP and pLac-nSDAP205 (SEQ ID NO:30) for the production of AP205 VLP. The construction of these specific expression plasmids is described in detail in the Examples section.

The expression plasmids are transformed to a bacterial expression host by any method known in the art, preferably by electroporation. Individual clones of the host comprising the expression plasmid are selected for maximal expression of the recombinant capsid protein by SDS-PAGE after cell lysis. Selected clones of the expression host comprising the expression plasmid can be stored as frozen glycerol cultures.
[0092] Said bacterial host can be chosen from any bacterial strain capable of replicating and maintaining said expression plasmid during cell division. Preferred bacterial hosts are *Escherichia coli* strains having the specific features described in the following sections.

[0093] The repression of the promoter is improved by overexpression of the repressor by the bacterial host. In one embodiment said cultivating of said bacterial host is performed in batch culture and under conditions under which said promoter is repressed by lacI. In a preferred embodiment the gene causing overexpression of said lacI in said bacterial host is located on a plasmid, preferably on said expression plasmid. Alternatively, said gene is located on a separate plasmid contained in said bacterial host, wherein said separate plasmid preferably is a high copy number plasmid. Alternatively, and most preferably said gene is located on the chromosome of said bacterial host.

[0094] One example of a gene causing overexpression of lacI is lacIβ (Menzella et al. 2005, Biotechnology and Bioengineering 82 (7):809-817) which is a single CG to TA change at –35 of the promoter region of lacI which causes a 10 fold increase in LacI expression. A further example is lacIβ (Glascocck & Weickert 1998, Gene 223 (1-2):221-231). Improved repression of the promoter during the growth phase results in improved plasmid retention and higher cell density and, ultimately, in improved protein yield. For example, bacterial strains comprising the lacIβ gene overexpress the lacI repressor molecule and therefore prevent the formation of the recombinant protein during the growth phase more efficiently than strains comprising the wildtype gene. In a preferred embodiment the gene causing overexpression of said lacI is lacIβ or lacIβ, preferably lacIβ. In a specifically preferred embodiment said bacterial host comprises the lacIβ gene on its chromosome.

[0095] In one embodiment said inducing of said promoter is performed with an inducer, wherein said inducer is preferably selected from IPTG and lactose, most preferably said inducer is lactose. Upon exposure of the bacterial host to an inducer, the repressor is inactivated and the promoter becomes active. Addition of the strong inducer IPTG to the culture medium results in an immediate increase of the expression rate of the recombinant protein to a high level because IPTG directly enters the cells by diffusion and initiates the active repressor lacI. Inactivated lacI repressor molecules dissociates from the operator and allow high level transcription from the promoter. IPTG is not metabolized by the cell and the transcription continues with high rates until other metabolic parameters become limiting.

[0096] As mentioned before, high expression rates may lead to the formation of insoluble recombinant protein which is not capable of forming a VLP by self-assembly. Induction of protein expression with high concentrations of IPTG is particularly prone to the formation of insoluble protein. Therefore, induction of the promoter is preferably achieved by the addition of IPTG in concentrations which are below the concentration which causes the expression to occur at its maximum rate (Kopetzki et al. 1989, Mol Gen Genet 216: 149-155).

[0097] In a preferred embodiment said inducing of said promoter is performed with IPTG, wherein the concentration of said IPTG in said medium is about 0.001 to 5 mM, preferably 0.001 to 1 mM, more preferably 0.005 to 1 mM, still more preferably 0.005 to 0.5 mM. In a specifically preferred embodiment the concentration of said IPTG is about 0.01 mM, most preferably 0.01 mM.

[0098] Alternatively, induction of the promoter is achieved by the addition of lactose. Induction of recombinant protein expression with lactose requires that the bacterial host is capable of taking up lactose from the medium, e.g. by Lac permease and that it comprises β-galactosidase activity. The Lac permease dependent uptake of lactose into the cells follows a slower kinetic than the uptake of IPTG by diffusion. Furthermore, lactose does not directly interact with the lac operon but is converted by β-galactosidase to allolactose (1-6-O-β-galactopyranosyl-D-glucose) which is the actual inducer of the promoter. Induction of recombinant protein expression by the addition of lactose is advantageous because it avoids the immediate increase of the expression rate to a maximum level upon addition of the inducer and, thus, it reduces the risk of the formation of insoluble protein.

[0099] Allolactose is metabolised by the bacterial host during the production phase and contributes carbon and metabolic energy to the bacterial metabolism. This may further contribute to improved protein yield as compared to induction with IPTG. Furthermore, induction by lactose allows to a certain extend the control of the expression rate of the recombinant protein during the production phase via the lactose concentration in the medium. Induction by lactose is further preferred in a pharmaceutical production process because IPTG is expensive and is believed to be toxic. Its removal needs to be demonstrated at the end of a the production process.

[0100] In a preferred embodiment said inducing of said promoter is performed by the addition of lactose to said batch culture, wherein preferably said bacterial host is capable of taking up lactose from the medium and wherein further preferably said bacterial host comprises β-galactosidase activity. Such bacteria strains can, for example, be obtained from strain collections such as ATCC (http://www.atcc.org). In a preferred embodiment, said bacterial host is an *E. coli* strain, preferably an *E. coli* strain selected from the group consisting of RB791, DH20, Y1088, W3110 and MG1655. Most preferably, said bacterial host is *E. coli* RB791. In a still more preferred embodiment said promoter is the tac promoter or mutant or variant thereof and said bacterial host is an *E. coli* strain which further comprises a gene causing overexpression of a repressor of the tac promoter, wherein said gene preferably is lacIβ. The pH of the culture medium is controlled in a range between 6.5 and 8.0. At an appropriate pH, the desired protein is expressed in the bacterial host, allowing the isolation and purification of the desired protein. The preferred pH range is 7.0 to 7.5. In a specific embodiment, the cultivation medium is buffered at a pH of 7.5. In another embodiment, the cultivation medium is not buffered. In a still more preferred embodiment the cultivation medium is buffered at a pH of 7.0 to 7.5. In yet another embodiment, the cultivation medium is not buffered. The cultivation medium is preferably not buffered.

[0101] The process of the invention comprises a growth phase, wherein said growth phase comprises a batch and a feed phase, wherein said growth phase and simultaneously said batch phase are initiated by said cultivating said bacterial host and wherein said feed phase is initiated by said feeding of said batch culture with said major carbon source.
The oxidative capacity of bacteria cells is limited and high concentrations of the substrate may cause the formation of reduced products like acetate, which may lead to undesired acidification of the medium and to reduced growth of the bacteria. Therefore, the bacterial host is grown in a fed-batch culture on a minimum medium with a limited quantity of substrate. In one embodiment said cultivating of said bacterial host is performed in a medium comprising said major carbon source, wherein said medium preferably is a minimal medium, preferably a chemically defined minimal medium. Most preferably said medium is R27 medium as described in Example 5.

At the end of the batch phase, when the substrate contained in the medium is almost exhausted, medium containing the major carbon source (feed medium) is fed to said batch culture at the same rate as the desired growth rate of the bacterial host, i.e., the growth rate of the bacterial host is limited by the feed rate of the substrate. It is understood by the skilled person that the decisive parameter is the actual mass flow of the substrate, preferably the major carbon source, and other nutrients required to maintain growth. Since in practice a constant composition of the feed medium can be assumed, the flow rate refers to the volume flow of the medium. The same consideration applies to the co-feed medium (see below).

Therefore, in one embodiment said feeding of said batch culture with said major carbon source is performed with a flow rate, wherein said flow rate is limiting the growth rate of said bacterial host.

During the feed phase the growth rate can be freely selected in a wide range up to the maximum growth rate ($\mu_{\text{max}}$) if no inhibition occurs. The actual value of $\mu_{\text{max}}$ is highly dependent on the bacterial strain, the expression construct and the growth conditions. The skilled person will understand that the determination of $\mu_{\text{max}}$ is performed under conditions under which the promoter is repressed.

For a given experimental setup, $\mu$ can be determined from the growth curve of the culture by plotting biomass concentration (x) as determined by OD$_{660}$ or cell wet weight (CWW) against the cultivation time and determining the exponential growth coefficient $\mu$ based on the equation $x=x_0e^{\mu t}$. The actual value of $\mu_{\text{max}}$ is determined as the growth rate $\mu$ of an exponentially growing batch culture in the beginning of the batch phase when no substrate limitation occurs, i.e., without supply of additional medium by feeding. The growth rate $\mu$ can be determined by computing the ratio of the difference between natural logarithm of the total biomass $X$ measured at time $t_1$ and natural logarithm of the total biomass $X$ measured at time $t_0$ to the time difference ($t_1-t_0$); $\mu=\frac{(\ln X_1-\ln X_0)}{(t_1-t_0)}$.

Feed-batch culture allows the maintenance of a constant growth rate ($\mu$). In a preferred embodiment the substrate, preferably the major carbon source, is fed during the feed phase according to the exponential increase of the biomass (x). If during the feed phase the substrate is supplied at the same rate it is consumed, the culture is in a quasi steady state, analogous to the cultivation in a continuous culture. Because biomass formation and substrate consumption are correlated over the substrate-referred yield coefficient $Y_{sx}$ (biomass [g]/substrate [g]), the substrate quantity (s) per unit time (t) to be supplied is calculated according to the formula $ds/dt=\mu Y_{sx} x_{0,\text{tot}} t^{\mu}$, wherein $x_{0,\text{tot}}$ is the total biomass at feed start.

Therefore, in a preferred embodiment said feeding of said batch culture with said major carbon source is performed with a flow rate, wherein said flow rate increases with an exponential coefficient $\mu$, and wherein preferably said exponential coefficient $\mu$ is below $\mu_{\text{max}}$. Thus, the growth rate of said bacterial host during the feed phase is set to a value which is below $\mu_{\text{max}}$. In a preferred embodiment said exponential coefficient $\mu$ is about 30% to 70%, most preferably about 50% of $\mu_{\text{max}}$. In a specific embodiment of the invention $\mu$ is set to an absolute value of 0.15 to 0.45 h$^{-1}$, more preferably 0.25 to 0.35 h$^{-1}$, most preferably $\mu$ is 0.3 h$^{-1}$, provided that the set up of the process is such that these values are below $\mu_{\text{max}}$.

Bacteria are able to utilise a wide range of different substrates. For the purpose of the invention, preferred major carbon sources are glucose and glycerol, preferably glycerol. Although the maximum specific growth rate ($\mu_{\text{max}}$) of the expression host which can be achieved may be higher with glucose than with glycerol, glycerol causes less acetate formation and provides higher biomass yield per substrate ($Y_{sx}$) and, ultimately, higher yield of the recombinant protein. Furthermore, the handling of the liquid substrate glycerol is easier than that of solid carbon sources like glucose which need to be dissolved in a separate process step.

As mentioned before, plasmid retention, i.e. the maintenance of the expression plasmid in the bacterial host during the fermentation process, is essential for optimal yield of the recombinant protein. Plasmid retention can be assessed by spreading bacteria cells on a solid medium to form single colonies and testing individual colonies for their antibiotic resistance. For example, a plasmid retention of 100% means that 100 out of 100 tested colonies comprise the specific antibiotic resistance which conferred by the expression plasmid. For the purpose of the invention plasmid retention at the end of the fermentation process is more than 80%, preferably more than 90%, more preferably more than 95%, even more preferably more than 97% and most preferably 100%.

The optimal growth temperature of a bacterial strain is the temperature at which it reaches its highest maximal growth rate ($\mu_{\text{max}}$). Under otherwise not limiting conditions for most E. coli strains this temperature is about 37° C. However, growth of the bacterial strain comprising the expression construct at the optimal growth temperature and in the absence of a selective antibiotic may favour the loss of the expression plasmid, whereas plasmid retention is generally improved when the expression strain is grown at lower temperature. Although the maximum growth rate of the expression strain is lower when the strain is grown at temperatures below its optimal growth temperature as compared to growth at the optimal growth temperature, the yield of recombinant protein may be equal or even better at the lower temperature due to improved plasmid retention.

In one embodiment of the invention, said cultivating of said bacterial host and/or said feeding of said batch culture with said major carbon source and/or said inducing said promoter with an inducer is therefore performed at a temperature below the optimal growth temperature of said bacterial host. In a preferred embodiment said temperature is between 20 and 37° C, preferably between 23 and 35° C, more preferably between 25 and 33° C, even more preferably between 27 and 32° C, still more preferably between 28 and 31° C. Still more preferably said temperature is about 30° C, most preferably said temperature is 30° C.
The process of the invention comprises a production phase, wherein said production phase is initiated by said inducing said promoter with an inducer. The time point for the initiation of said production phase can be determined based on cultivation time and/or growth parameters.

The growth of the bacterial host during the fermentation process can be assessed by determining the optical density at 600 nm (OD_{600}), the cell wet weight (CWW [g/l]) and the cell dry weight (CDW [g/l]). These parameters can be used to define the optimal time point for the start of the production phase by addition of the inducer, preferably lactose, to the medium. It is apparent for the skilled person that on one hand higher CWW at the beginning of the production phase can be achieved by an extended feed phase and may lead to improved yield of the recombinant protein but that on the other hand over-aged cultures may show insufficient protein expression. The optimal time point for the beginning of the production phase, which is initiated by said inducing of said promoter with an inducer, therefore needs to be determined for the specific production conditions. For example, for expression of Q8 CP in E. coli RB791 in a total volume of 2 l, induction is started after ca. 14 h, when OD_{600} has reached about 40 to 60. Surprisingly, similar parameters were found for the same process in a 50 l scale, where induction start is also after ca. 14 h when OD_{600} has reached about 50.

Therefore, in one embodiment of the invention, said inducing of said promoter with said inducer is performed 10 h to 16 h after the beginning of said growth phase, preferably after 12 h to 15 h, more preferably after 13 h to 15 h, most preferably after about 14 h, wherein preferably said inducing of said promoter with said inducer is performed when the OD_{600} has reached about 40 to 60, preferably about 50.

In a further embodiment, said inducing of said promoter with said inducer is performed after an extended feed phase, wherein preferably said inducing of said promoter with said inducer is performed 14 h to 20 h after the beginning of said cultivating of said bacterial host in a medium, preferably after 15 h to 18 h, more preferably after 16 h to 17 h, most preferably after about 16.5 h, wherein preferably said inducing of said promoter with said inducer is performed when the OD_{600} has reached about 80 to 90, preferably about 85.

In one embodiment of the invention said inducing of said promoter with said inducer is performed when the OD_{600} reached a value of 25 to 60, preferably 25 to 55, more preferably 30 to 50, most preferably 30 to 40. In a specifically preferred embodiment said inducing of said promoter with said inducer is performed when the OD_{600} is 35.

In another embodiment of the invention said inducing of said promoter with said inducer is performed after an extended feed phase, when the OD_{600} reached a value of 60 to 120, preferably 70 to 110, more preferably 80 to 100, most preferably 80 to 90. In a specifically preferred embodiment the induction is started after an extended feed phase when OD_{600} is about 85, preferably 85.

Induction with IPTG: In one embodiment of the invention said inducing of said promoter with an inducer is achieved by the addition of IPTG, wherein preferably said feeding of the culture with the major carbon source is continued with a constant or an increasing flow rate of said major carbon source exponentially increasing flow rate of the major carbon source.

Induction with Lactose: As described above, the induction of protein expression can alternatively be achieved by the addition of lactose to the culture medium. In one embodiment of the invention, at the beginning of the production phase the exponential feed of the substrate is interrupted and the culture is supplied with a constant flow of induction medium containing 100 to 300 g/l, preferably 100 g/l lactose as the sole carbon source (lactose feed medium). Preferably, the constant flow rate of lactose equals approximately the flow rate of the substrate at the end of the feed phase.

In a preferred embodiment of the invention said inducing of said promoter with an inducer is achieved by the addition of lactose, wherein preferably said lactose is fed to said batch culture in a continuous flow during and wherein preferably said feeding of said batch culture with said major carbon source is not continued.

Upon addition of lactose to the culture, the β-galactosidase activity increases, lactose is converted to allolactose which induces the tac promoter and the expression of the recombinant capsid is initiated. In parallel, allolactose is further metabolised and contributes to the energy supply for the bacterial host. The equilibrium of the feeding rate of the induction medium and the lactose consumption by the cells thus determines the expression rate. The enzymatic reactions involved in this cascade allow to control the process in such a way that the formation of inclusion bodies is minimised. The progress of induction process can be monitored by determining the β-galactosidase activity in the culture, e.g. by a β-Gal Assay Kit (Invitrogen, K1455-01).

In a more preferred embodiment of the invention said inducing of said promoter with an inducer is achieved by the addition of lactose, wherein preferably said lactose is fed to said batch culture in a continuous flow during and wherein preferably said feeding of said batch culture with said major carbon source is continued.

Discontinuous Addition of Inducer: Said inducer can be added to the culture discontinuously by a single addition at the beginning of the production phase or by a few subsequent additions during the production phase. Discontinuous addition of the inducer, especially by a single addition is particularly suited when the inducer is IPTG since IPTG is not metabolized by the bacterial host. Therefore, typically and preferably no replacement of metabolised IPTG is necessary during the production phase. In one embodiment said inducing of said promoter with an inducer is performed by the addition of said inducer, preferably IPTG or lactose, most preferably IPTG, to said medium, wherein said inducer is added to about its final concentration at once by a single addition at the beginning of the production phase, wherein preferably said feeding of said batch culture with said major carbon source is continued. In a preferred embodiment said inducing of said promoter with an inducer is performed by the addition of IPTG to said medium, wherein said IPTG is added to about its final concentration at once by a single addition, wherein preferably said feeding of said batch culture with said major carbon source is continued. Alternatively, said inducing of said promoter with an inducer is performed by the addition of said inducer, preferably IPTG or lactose, most preferably lactose, to said medium, wherein said addition is performed in several steps, preferably in 1 to 5, more prefer-
ably in 2 to 4, most preferably in 3 steps during the production phase, wherein preferably said feeding of said batch culture with said major carbon source is continued.

[0125] Continuous Addition (Feeding) of Inducer: Preferably, said inducer is added to the medium in a continuous flow, preferably throughout the production phase. The continuous addition of the inducer is particularly suited for lactose, since lactose is metabolized by the bacterial host and therefore a continuous addition of lactose during the production phase allows to maintain a lactose concentration in the medium which allows for efficient induction of the promoter. In a preferred embodiment, said inducing of said promoter with an inducer is performed by feeding said batch culture with said inducer, wherein preferably said inducer is IPTG or lactose, most preferably lactose, and wherein said feeding is performed in a continuous flow, wherein further preferably said feeding is performed throughout the production phase.

[0126] Co-Feeding of Inducer and Major Carbon Source: The expression of the recombinant protein is an energy demanding process. To prevent yield loss which might be caused by the excessive consumption of the inducer by the bacterial host and low expression rates resulting thereof, the culture can be additionally supplemented with substrate, preferably the major carbon source, during the production phase, wherein the flow rate of inducer and/or the major carbon source is constant or increasing, preferably constant. When during the production phase the culture is supplemented with substrate at an increasing flow rate, the flow rate is preferably increasing with an exponential rate.

[0127] Co-Feeding with Constant Flow Rate: In a preferred embodiment said inducing of said promoter with an inducer is performed by co-feeding said batch culture with said inducer and said major carbon source, wherein said inducer is preferably IPTG or lactose, most preferably lactose, and wherein said major carbon source is glucose or glycerol, preferably glycerol, wherein said inducer, preferably lactose and said major carbon source, preferably glycerol are co-fed to said batch culture at a flow rate wherein said flow rate is preferably about constant. In a further preferred embodiment said flow rate is chosen to allow feeding of said major carbon source to said batch culture at about the same rate as at the end of the growth phase. In a still further preferred embodiment said inducer, preferably lactose, and said major carbon source, preferably glycerol, are contained in the same medium (co-feed medium). In a further preferred embodiment said co-feed medium is fed to said batch culture with a flow rate, wherein said flow rate is preferably about constant, and wherein further preferably said flow rate is chosen to allow feeding of said major carbon source to said batch culture at about the same rate as at the end of the growth phase. In a very preferred embodiment said inducer is lactose and said major carbon source is glycerol, wherein said lactose and said glycerol are co-fed to said batch culture in a ratio of about 2:1 to 1:4 (w/w).

[0128] In a further preferred embodiment of the invention lactose and said major carbon source, preferably glycerol, are co-fed to said batch culture in a ratio of 0.1 to 1.0 (w/w), preferably about 2:1 to about 1:2 (w/w), more preferably about 1:1 to 1.3 (w/w), most preferably the ratio is about 1:3 (w/w). In a preferred embodiment the ratio of lactose and the major carbon source, preferably glycerol, is 1:1 (w/w). In another preferred embodiment the ratio of lactose and the major carbon source, preferably glycerol, is 1:3 (w/w).

[0129] Co-Feeding with Increasing Flow Rate: Alternatively, said inducing of said promoter with an inducer is performed by co-feeding said batch culture with said inducer and said major carbon source, wherein said inducer is preferably IPTG or lactose, most preferably lactose, and wherein said major carbon source is glucose or glycerol, preferably glycerol, wherein said inducer, preferably lactose and said major carbon source, preferably glycerol are co-fed to said batch culture at a flow rate, wherein said flow rate is increasing, wherein said flow rate may increase with a linear or with an exponential characteristic, wherein preferably the initial flow rate is chosen to allow feeding of said major carbon source to said batch culture at about the same rate as at the end of the growth phase.

[0130] Further alternatively said inducing of said promoter with an inducer is performed by co-feeding said batch culture with said inducer and said major carbon source, wherein said inducer is preferably IPTG or lactose, most preferably lactose, and wherein said major carbon source is glucose or glycerol, preferably glycerol, wherein said inducer, preferably lactose is fed to said batch culture at a first flow rate, and wherein said major carbon source, preferably glycerol is fed to said batch culture at a second flow rate, wherein said first flow rate is constant or increasing, preferably constant, and wherein said second flow rate is constant or increasing, preferably increasing, wherein preferably the initial value of said second flow rate is chosen to allow feeding of said major carbon source to said batch culture at about the same rate as at the end of the growth phase. In a very preferred embodiment said inducer is lactose and said major carbon source is glycerol, wherein said lactose and said glycerol are co-fed to said batch culture in a ratio of about 2:1 to 1:4 (w/w).

[0131] The growth of the bacterial host as determined by CDW, CW and OD increases during the production phase at a growth rate which is lower than that during the growth phase and which is decreasing with the process time. In a further embodiment of the invention, said inducing said promoter with an inducer is performed by co-feeding said inducer, preferably lactose and said major carbon source, preferably glycerol, to said batch culture with an increasing flow rate, preferably with a flow rate wherein the incremental increase of the flow rate is adapted to the actual growth rate of the culture. In a further preferred embodiment said inducer, preferably lactose, and said major carbon source, preferably glycerol, are contained in the same medium (co-feed medium), wherein preferably the ratio between lactose and glycerol in said medium (co-feed medium) ranges from about 0:1 to 1:0 (w/w), preferably about 2:1 to about 1:4 (w/w), more preferably about 1:1 to 1:3 (w/w), most preferably the ratio is about 1:3 (w/w). In a preferred embodiment the ratio of lactose and the major carbon source, preferably glycerol, is 1:1 (w/w). In an another preferred embodiment the ratio of lactose and the major carbon source, preferably glycerol, is 1:3 (w/w).

[0132] In one embodiment of the invention said inducing of said promoter with an inducer is performed by co-feeding said inducer, preferably lactose and said major carbon source, preferably glycerol to said batch culture, wherein said inducer, preferably lactose and said major carbon source, preferably glycerol are contained in separate media which are separately fed to said culture.
At the end of the production phase the cells are harvested by centrifugation. Typically, cells are harvested about 5 h after induction start, when a final OD_{600} of 90 to 130 is reached. Further extension of the production phase leads to higher OD_{600} and CWW values and therefore to further improved yield of the expression construct.

Harvested cells may be suspended in a storage buffer and stored at ~80°C for further processing.

The total protein content of the cells is determined after cell lysis by SDS PAGE or LDS PAGE and comparison with a protein standard. The content of soluble protein is determined by HPLC. The identity of the expressed capsid protein is determined by western blotting. The concentration of assembled VLPs can be analysed by size exclusion chromatography (Example 18). VLP can preparatively be purified from lysed cells by chromatographic methods.

Scale-up of the process of the invention to large volumes is possible with only minor adaptations. The invention encompasses culture volumes in the range of 100 ml up to 6000 l. Preferred culture volumes are 40 to 100 l, most preferably about 50 l. It is apparent for the skilled person that larger culture volumes in particular require larger volumes of the preculture which is used for inoculation. For example, a preculture may be performed in two or more steps with increasing preculture volume. To ensure plasmid retention in large culture volumes, the precultures which are used as inoculum may contain an antibiotic to maintain selection pressure. The skilled person is aware that plasmid retention can further be improved by reducing the number of generations which is necessary to reach the desired final cell density. Therefore, it is advantageous to inoculate the precultures and the batch cultures with high cell densities. In a preferred embodiment the initial OD_{600} of the preculture is 0.1 to 0.4, preferably about 0.3.

In one embodiment, prior to said cultivation step, said process further comprises the step of introducing said bacterial host into a medium, wherein said introducing is performed with an inoculum, wherein said inoculum is produced in a preculture process comprising the step of growing said bacterial host in a medium comprising an antibiotic, preferably kanamycin. More preferably, said preculture process comprises the steps of growing said bacterial host in a first medium comprising an antibiotic, preferably kanamycin, and diluting said first medium comprising the bacterial host with a second medium to an OD_{600} of 0.1 to 0.4, preferably about 0.3, wherein said second medium is essentially free of an antibiotic, and further cultivating said bacterial host.

Furthermore, it is apparent for the skilled person, that the fermentation process of the invention is an aerobic process which requires adequate oxygen supply of the bacteria in the culture. The oxygen demand of the bacterial host is, inter alia, increasing with increasing cell density and increasing growth rate. Depending on the total volume and the oxygen demand of the bacterial host, oxygen can, for example, be supplied by stirring and/or by aeration with air. Alternatively, oxygen can also be supplied by aeration with pure oxygen or a mixture of pure oxygen with any other gas, preferably air, wherein pure oxygen refers to the technically pure gas as commonly available for technical purposes. A further possibility of supplying oxygen to the bacterial host is increasing the oxygen partial pressure in the medium by increasing the pressure in the fermenter.

In a preferred embodiment of the invention, said cultivating said bacterial host and/or said feeding of said batch culture and/or said inducing of said promoter with an inducer is performed under conditions, wherein said bacterial host is supplied with oxygen, preferably by aeration with air, most preferably by aeration with air in a constant flow, wherein preferably said oxygen is supplied throughout the entire process, most preferably throughout the lag-, growth- and production phase, and wherein further preferably the partial pressure of oxygen is monitored in the culture medium and wherein the bacterial host is alternatively or additionally supplied with oxygen by aeration with pure oxygen, preferably when the partial pressure of oxygen in the medium (pO₂) is below a certain threshold. In a specifically preferred embodiment said threshold of pO₂ is in the range of 0% to 60%, preferably 10% to 50%, more preferably 20% to 45% most preferably said threshold is about 40%.

Oxygen supply, preferably by aeration with air and/or pure oxygen to maintain the preferred pO₂ as described above, is routinely applied in the process of the invention, preferably for culture volumes of 2 l and more. Aeration with oxygen in the described manner is especially preferred in the scaled-up process, most preferably at 40 to 100 l and above.

Therefore, one embodiment of the invention is a process for expression of a recombinant capsid protein of a bacteriophage or a mutant or fragment thereof being capable of forming a VLP by self-assembly, said process comprising the steps of: a) introducing an expression plasmid into a bacterial host, wherein said expression plasmid comprises an expression construct, wherein said expression construct comprises (i) a first nucleotide sequence encoding said recombinant capsid protein, or mutant or fragment thereof, and (ii) a promoter being inducible by lactose; b) cultivating said bacterial host in a medium comprising a major carbon source; wherein said cultivating is performed in batch culture and under conditions which said promoter is repressed by lactose wherein said plasmid is overexpressed by said bacterial host; c) feeding said batch culture with said major carbon source; and d) inducing said promoter with an inducer, wherein said feeding of said batch culture with said major carbon source is continued; wherein throughout steps b) to d) of said process oxygen is supplied to said bacterial host by a pO₂ in said medium of at least about 10% to 50%, preferably about 40%; and wherein further preferably said oxygen is supplied by aeration with air, pure oxygen, or a mixture of both, preferably by a mixture of air and pure oxygen.

EXAMPLES

Example 1

Cloning Strategy for the Expression Plasmid pTace-nSD-Qb-mut (SEQ ID NO:1)

The coat protein-encoding gene (C) of E. coli RNA bacteriophage QB is amplified from plasmid pSDQb-mut (SEQ ID NO:33). The plasmid contains the sequence of gene C coding for the 133-aa QB coat protein (CP) and the 329-aa read through protein (A). To prevent read-through, nucleotides 445-450 according to NCBI GenBank Acc. No. M99030 TGAACA (SEQ ID NO:31) are replaced by the sequence TAATGA (SEQ ID NO:32).

The coat protein-encoding gene C from plasmid pSDQb-mut is amplified by PCR. Oligonucleotide Qb-FOIR3/2 (SEQ ID NO:34) with an internal EcoRI site and a synthetic Shine-Dalgarno (SD, SEQ ID NO:4) sequence anneals to the 5' end of the Qb CP gene. Oligonucleotide Qbhang-REV2/2 (SEQ ID NO:35) contains an internal HindIII site and primes to the 3' end of the noncoding region of gene C. The 1054 by amplified PCR fragment includes nucleotides 46-1062 of NCBI GenBank Acc. No. M99039 (except
the nucleotide changes described above) and the synthetic SD sequence. The PCR fragment is digested with the restriction enzymes HindIII EcoRI and the resulting 1036 bp fragment is inserted into the HindIII/EcoRI restriction sites of a modified pKK223-3 vector (Pharmacia, NCBI GenBank Acc. No.: M77749, SEQ ID NO:27). In this modified pKK223-3 vector the ampicillin resistance gene is replaced with the kanamycin resistance gene of vector pUC4K (Pharmacia, NCBI GenBank Acc. No.: X06404, SEQ ID NO:37).

Example 1

Expression of Qβ CP Under Control of the tac Promoter and nSD

[0147] The E. coli strain RB791 was transformed with plasmids pTac-nSD-Qb-mut (SEQ ID NO:1). The clone was grown in shake flasks. Each flask contained 100 ml of R40 medium (main culture medium, Hyperr 7455, glycerol, see Example 5) with kanamycin (25 μg/ml) and was inoculated with overnight cultures at a start OD₆₀₀ of 0.3. The shake flasks were incubated for 4 h (OD₆₀₀ between 4 and 5) at 30°C and an agitation of 220 rpm. The induction was carried out with 0.5% of lactose for 4 h. Protein production was determined by SDS-PAGE. The gel showed a strong protein band which was identified as Qβ CP.

Example 4

Expression of AP205 CP Under Control of the tac Promoter and SD vs. nSD

[0148] 9 clones of pTac-nSDAP205 (SEQ ID NO:30) and 6 clones of pTac-SDAP205 were screened in shake flasks. pTac-SDAP205 (SEQ ID NO:40) is identical to pTac-nSDAP205 but comprises the Shine-Dalgarno sequence of SEQ ID NO:4 instead of that of SEQ ID NO:3. Each flask contained 50 ml of R40 medium (main culture medium, Hyperr 7455, glycerol, see Example 5) with kanamycin (25 μg/ml) and was inoculated with overnight cultures at a start OD₆₀₀ of 0.3 for pTac-nSDAP205 or 0.4 for pTac-SDAP205. The shake flasks were incubated for 4 h at 30°C and an agitation of 220 rpm. The induction was carried out with 0.5% of lactose. Protein production was determined by SDS-PAGE. For all tested clones expression of AP205 CP was significantly stronger from pTac-nSDAP205 than from pTac-SDAP205.

Example 5

Composition of Culture Media

[0150] Culture media were composed as described in Table 1.

### TABLE 1

<table>
<thead>
<tr>
<th>Component</th>
<th>Main Medium + Hyperr R27</th>
<th>Main Medium + Hyperr + Glycerol R40</th>
<th>Feed Medium + 50% Glycerol R41</th>
<th>Induction Medium + 20% Lactose R42</th>
<th>Main Medium + Bacto YET + Glycerol R43</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na₂HPO₄·2H₂O</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>K₂HPO₄</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>K₃HPO₄</td>
<td>5.2</td>
<td>5.2</td>
<td>5.2</td>
<td>5.2</td>
<td>5.2</td>
</tr>
<tr>
<td>Citrate</td>
<td>3.86</td>
<td>3.86</td>
<td>3.86</td>
<td>3.86</td>
<td>3.86</td>
</tr>
</tbody>
</table>
TABLE 1-continued Composition of Culture media. Concentrations in [g/L]

<table>
<thead>
<tr>
<th>Component</th>
<th>Main Medium + Bacto YE + Glycerol R43</th>
<th>Main Medium + Bacto YE + Glycerol R42</th>
<th>Feed Medium + 50% Glycerol + 20% Lactose R41</th>
<th>Induction Medium + 20% Hypep + Glycerol + 20% Lactose R40</th>
<th>Main Medium + Hypep + Glycerol R27</th>
</tr>
</thead>
<tbody>
<tr>
<td>(NH₄)₂SO₄</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Vit B₁</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>CaCl₂ 2H₂O</td>
<td>0.0147</td>
<td>0.0147</td>
<td>0.0147</td>
<td>0.0147</td>
<td>0.0147</td>
</tr>
<tr>
<td>MgSO₄ 7H₂O</td>
<td>0.5</td>
<td>0.5</td>
<td>9</td>
<td>9</td>
<td>0.5</td>
</tr>
<tr>
<td>FeCl₃ 6H₂O</td>
<td>0.054</td>
<td>0.054</td>
<td>0.054</td>
<td>0.054</td>
<td>0.054</td>
</tr>
<tr>
<td>CoCl₂ 6H₂O</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
<tr>
<td>MnCl₂ 4H₂O</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>CuCl₂ 2H₂O</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
</tr>
<tr>
<td>H₃BO₃</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>Na₂MoO₄ 2H₂O</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
<tr>
<td>ZnCl₂COO₂ 2H₂O</td>
<td>0.0026</td>
<td>0.0026</td>
<td>0.0026</td>
<td>0.0026</td>
<td>0.0026</td>
</tr>
<tr>
<td>Glucose</td>
<td>—</td>
<td>5</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>Glycerol</td>
<td>—</td>
<td>—</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>Lactose anhydrous</td>
<td>—</td>
<td>—</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>HyPep</td>
<td>5</td>
<td>5</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Bacto Yeast Extract</td>
<td>5</td>
<td>5</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

Example 6
Expression of Qβ CP in a Fed-batch Process (2 L Scale)

The fermentation process was performed in a bioreactor (Applikon 5 L, fish bottom equipped with 2 disc stirrer (6 cm), baffles (3x16 cm), pH-, pO₂-, and temperature control, and fermenter software BioXpert Version 2.22)

5 mL cryo culture of RB1791 transformed with plasmids pEac-nSD-Qβ-mut were inoculated in 100 mL Erlenmeyer flasks containing 50 mL medium R40 (25 µg/mL kanamycin) and cultivated for 14 h at 30°C and 220 RPM over night. After 14 h an OD₃₀₀ value of 6.0 was reached. For batch fermentation, 2 L of medium (R40) were pumped into the bioreactor. In Table 2 the cultivation parameters are listed.

TABLE 2

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Set point</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stirrer speed</td>
<td>1000</td>
<td>[rpm]</td>
</tr>
<tr>
<td>Air supply</td>
<td>2.5</td>
<td>[L/min]</td>
</tr>
<tr>
<td>O₂-supply, maximal</td>
<td>2</td>
<td>[L/min]</td>
</tr>
<tr>
<td>Temperature</td>
<td>30</td>
<td>[°C]</td>
</tr>
<tr>
<td>O₂-saturation</td>
<td>&gt;40</td>
<td>[%]</td>
</tr>
<tr>
<td>pH</td>
<td>6.8</td>
<td></td>
</tr>
</tbody>
</table>

The Qβ content in percent of the total protein content was calculated as follows, assuming that 50% of the E. coli biomass is protein:

\[
\text{Biomass (g/L)}/2=\text{total protein (g/L)}
\]

\[
\text{Qβ (g/L)}=\text{total protein (g/L)}\cdot100=\text{Qβ/total protein [%].}
\]

In the fed-batch mode, which followed the batch mode, a feeding phase was added. In the feeding phase substrate is supplied to the cells in the reactor according to a defined profile. The feed profile depends on the selected growth rate µ, the yield coefficient biomass to glycerol (Yₓₓ glycerol), the volume (VF), and the concentration of substrate in the feed (cf), substrate concentration. The feed was calculated using the following equation:

\[
\text{Feed equation:} \quad \text{mf}=(\mu\cdot Y_{x/x g}+m)\cdot V_F \cdot C_F^{\text{cf}}
\]

\[
\text{pump}=(mf/cf+b)/a
\]

For the determination of the calibration parameters a and b, a pump calibration was carried out. In addition, the feed tube with feed bottle was clamped into the feed pump and the pump was run with 7, 14 and 21% pump performance. The pumped feed volume per time was noted. In a resulting diagram of the relation of pump performance [%] to pumped feed solution [ml/h], the slope (a) and the Y-axis section (b) was determined. On the bioreactor the parameters in Table 3 were set for fed-batch cultivation.
TABLE 3 Parameters for fed-batch cultivation in bioreactor.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Set point</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stirrer speed</td>
<td>1000</td>
<td>[rpm]</td>
</tr>
<tr>
<td>Air supply</td>
<td>2.5</td>
<td>[L/min]</td>
</tr>
<tr>
<td>$O_2$ supply, maximal</td>
<td>2</td>
<td>[L/min]</td>
</tr>
<tr>
<td>Temperature</td>
<td>30</td>
<td>['C.']</td>
</tr>
<tr>
<td>$O_2$ saturation</td>
<td>&gt;40</td>
<td>[-]</td>
</tr>
<tr>
<td>pH</td>
<td>6.8</td>
<td>[-]</td>
</tr>
</tbody>
</table>

After reaching a process time of approximately 7h (end of batch) the feed pump was turned on automatically. After further 7h cultivation, when the $OD_{600}$ reached 55-60, the feed medium (for biomass propagation) was exchanged with the medium induction R42 (for biomass propagation and induction). After 5h feeding of R42 was stopped and the culture was harvested by centrifugation.

Analysis of Process Parameters:

The following process parameters were routine analysed: The $pO_2$, pH, temperature and stirrer speed were measured online throughout the process. The optical density was measured offline at 600nm. The determination of the $\beta$-galactosidase activity was performed using a $\beta$-Gal Assay Kit (Invitrogen, cat. no. K1455-01). The activity was specified as units per mL $OD_{600}=1.0$. It is defined as the quantity of Ortho-Nitrophenyl-$\beta$-D-Galactopyranosid (ONPG) in nmol, which is hydrolysed per minute and mL bacteria suspension ($OD_{600}=1.0$). The accumulated product was analysed by SDS-PAGE, the total protein content (soluble and insoluble protein) was determined and using HPLC analysis, the soluble fraction was measured. Cell disruption of E. coli was performed in lysis buffer (50 mM glucose, 25 mM tris/HCl (pH8), 15 mM EDTA (pH8.0) with and ultrasonic homogeniser (Bandelin Sonoplus, HD2070), 250 $\mu$L bacteria suspension with an $OD_{600}$ of 50 were centrifuged with 14000 RPM for 10 min. The pellet was resuspended in 250 $\mu$L lysis buffer (vortex) and placed at room temperature for 5 min. Afterwards, the cells were disrupted for 20 s with ultrasonic at 10% device performance (cells on ice) and then the cell suspension was centrifuged at 14000 RPM, 10 min. The supernatant (soluble protein) was then analysed by SDS-PAGE and HPLC.

Samples before induction and at end of production (after 5h induction) were taken from the bioreactor for analysis of Qb65 formation analyzed by SDS-PAGE standardized to $OD_{50}$. At the end of cultivation, 1.9L of the culture was harvested. After centrifugation, the following cell pellets were obtained in three independent reactor runs: 1.) End $OD_{600}$ of 84: 194 g CWW; 2.) End $OD_{600}$ of 88: 200 g CWW; 3.) End $OD_{600}$ of 86: 201 g CWW.

The plasmid retention in run 1 and 2 was 100% at induction start and 100% at harvest. Based on comparison with a Qb CP standard on SDS-PAGE the yield was roughly estimated to be about 5 g/l Qb65 CP. HPLC analysis revealed a concentration of about 6 g/l Qb VLP.

Example 7 Selection of Carbon Source and Bacterial Strain

Glucose and glycerol as carbon sources were compared. In order to test the growth behavior of each of the strains DH20 and RB791 on these carbon sources, shake flask experiments were conducted with medium containing glucose (R27) and medium containing glycerol (R40). Both media were supplemented with 25 mg/ml kanamycin. Each culture was started with an initial $OD_{600}$ of 0.3. Induction was performed by adding 0.5% lactose. The maximum specific growth rates ($\mu_{max}$) and the yield coefficients ($Y_{x_\sigma}$) were determined and are listed in Table 4. RB791 grew faster on both, glucose and glycerol. In addition, the resulting yield coefficients were higher. Although glucose allowed higher maximum specific growth rates ($\mu_{max}$), the yield coefficients ($Y_{x_\sigma}$) was higher for glycerol.

Table 4 Maximum specific growth rates and the yield coefficients of the cultivation experiments with RB791 and DH20 on glucose and glycerol.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Carbon source</th>
<th>Value after 6h</th>
<th>Max. specific.</th>
<th>Yield coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$OD_{600}$</td>
<td>$\mu_{max}$</td>
<td>$Y_{x_\sigma}$ biomass [g/g]</td>
</tr>
<tr>
<td>RB791</td>
<td>glucose</td>
<td>6.24</td>
<td>0.44</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>glycerol</td>
<td>4.04</td>
<td>0.21</td>
<td>0.62</td>
</tr>
<tr>
<td>DH20</td>
<td>glucose</td>
<td>2.52</td>
<td>0.42</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>glycerol</td>
<td>2.82</td>
<td>0.25</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Determination of Optimal Temperature

The influence of temperature on product formation was investigated. Two shake flask cultures were inoculated and incubated at 30°C and 220 rpm. After an $OD_{600}$ of 5 was reached, the cultures were induced with lactose. Subsequently, one culture was continued to be incubated at 37°C and the other culture at 23°C. Results of the SDS-PAGE revealed that expression levels at 4 and 5h after induction are higher in the culture induced at 37°C. Induction of the cultures for 19h showed a higher Qb level in the cultures induced at 23°C.

Example 9 Induction by Co-Feed of Lactose and Glycerol

A feed solution of 20% glycerol and 20% lactose was composed (R42) and applied to fermentation as described in Example 6 at induction start. FIG. 1 provides an overview over relevant process parameters throughout the entire process time. Expression was induced at 13.5 h at an $OD_{600}$ of about 55. Upon induction, the feed pump rate was set to constant. Glycerol did not accumulate with feeding. Lactose accumulated to 4 g/l and then it started to diminish. The $\beta$-galactosidase activity rose to 10 U/ml and decreased thereafter. Compared with the previous fermentation runs a.) lactose applied as a single lactose pulse at induction start, no feeding; b.) continuous lactose feed without glycerol, the activity was with 7 U/ml higher and the maximum activity was already reached after 2 h as compared to 4 h in runs a.) and b.).
Example 10

Plasmid Retention

[0175] The effect of the following operating conditions on the plasmid retention was tested in the process described in Example 6: 1) Preculture starting volume, 2) Kanamycin in the preculture, 3) Growth and/or induction at 37°C vs. 30°C. The results are summarised in Table 5. Precultures were started with volumes of 5 µl out of the cell bank vial. Inoculation of a small volume allowed growth of a preculture over-night. The preculture for QT0103_F8 contained 25 mg/L kanamycin, whereas the preculture for QT0103_F7 did not contain any kanamycin. Both fermentations were operated at 30°C and induced for 5 h. Judging from the plasmid retentions before and after 5 h induction, supplementing the preculture with kanamycin has a positive effect on plasmid retention. Plasmid retention remained at 98% before and after 5 h induction. In contrast, plasmid retentions reached only values of 80% when kanamycin was omitted from the preculture. For a subsequent run, QT0203_F7, the preculture was also started with 5 µl and grown in kanamycin containing medium. The resulting fermentation in the bioreactor was operated at 37°C from the beginning. Operation at 37°C had a detrimental effect on the plasmid stability. While the plasmid retention was at 99% before induction, it dropped to 0% after 5 h induction. In order to test whether a shorter preculture and thus, less generations, would improve the plasmid retention after 5 h induction, a set of precultures were started with 300 µl volume from a thawed cell bank vial and grown in kanamycin free medium. Two fermenters were operated at 30°C, for the whole run. An additional two fermenters were operated first at 30°C, for cell growth and then switched to 37°C, for the production phase. The resulting plasmid stabilities were all at 100% before and 5 h after induction.

Example 11

Variation in Time Point of Induction

[0176] In a process essentially as described in Example 6 the exponential feed profile was programmed to start 7 h after the inoculation of the bioreactor. Under standard conditions, the scheduled time for induction was at 14 h process time. In order to test the effect of variations in the time point of induction on the final cell densities, one culture was induced at 13.5 h (resulting in 6.5 h of exponential feed) and another culture at 14.5 h (resulting in 7.5 h of exponential feed). One culture induced at the regular 14 h time point served as a control (7 h of exponential feed). Results are summarised in Table 6. Cell density increased with increasing length of feeding. Judged from a linear regression analysis of the available data points for final CWW, a linear relationship appears to exist (r²=0.92).

TABLE 5

Summary of plasmid retention before and 5 h after induction obtained under different operating conditions in terms of generations in the preculture, with and without kanamycin in the preculture, and growth and/or induction at 37°C.

<table>
<thead>
<tr>
<th>Bioreactor run</th>
<th>Preculture Culture Volume [µl]</th>
<th>Kanamycin [mg/L]</th>
<th>Starting Plasmid retention [%] before 5 h</th>
<th>Induc. Plasmid retention [%]</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>QT0103_F8</td>
<td>5</td>
<td>25</td>
<td>98</td>
<td>98</td>
<td>whole process at 30°C,</td>
</tr>
<tr>
<td>QT0103_F7</td>
<td>5</td>
<td>no</td>
<td>80</td>
<td>80</td>
<td>whole process at 30°C,</td>
</tr>
<tr>
<td>QT0203_F7</td>
<td>5</td>
<td>25</td>
<td>99</td>
<td>0</td>
<td>Bioreactor run at 37°C,</td>
</tr>
<tr>
<td>QT0603_F7</td>
<td>300</td>
<td>no</td>
<td>100</td>
<td>100</td>
<td>whole process at 30°C,</td>
</tr>
<tr>
<td>QT0703_F8</td>
<td>300</td>
<td>no</td>
<td>100</td>
<td>100</td>
<td>Induction at 37°C, rest of the process at 30°C,</td>
</tr>
<tr>
<td>QT0803_F9</td>
<td>300</td>
<td>no</td>
<td>100</td>
<td>100</td>
<td>whole process at 30°C,</td>
</tr>
<tr>
<td>QT0903_F10</td>
<td>300</td>
<td>no</td>
<td>100</td>
<td>100</td>
<td>Induction at 37°C, rest of the process at 30°C,</td>
</tr>
</tbody>
</table>

TABLE 6

Variations in time point of induction: effect on final cell density in terms of OD_600 and CWW.

<table>
<thead>
<tr>
<th>Reactor</th>
<th>Process Time</th>
<th>Duration of Exp. Feed Phase [h]</th>
<th>Final OD_600 [g/L]</th>
<th>Final CWW [g/L]</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2</td>
<td>13 h 32 min</td>
<td>6.5</td>
<td>83.4</td>
<td>116.5</td>
</tr>
<tr>
<td>F1</td>
<td>14 h 02 min</td>
<td>7.0</td>
<td>82.4</td>
<td>122.5</td>
</tr>
<tr>
<td>F3</td>
<td>14 h 29 min</td>
<td>7.5</td>
<td>100.4</td>
<td>141.1</td>
</tr>
</tbody>
</table>

Dec. 20, 2012
Example 12

Variation in Time Point of Harvest

[0177] Harvest of the culture in a process essentially as described in Example 6 is performed manually. Under standard conditions, the scheduled time for harvest was at 19 h process time. The operation “Harvest” involves the manual ending of the bioreactor operations. In order to test the effect of variations in the time point of harvest on the final cell densities, one culture was harvested at 18.8 h (resulting in 4.8 h of induction) and another culture at 19.5 h (resulting in 5.5 h of induction). One culture harvested at the regular 19 h time point served as a control (5 h of induction). Results are summarized in Table 7. Cell density increased with increasing length of induction because the cells are still growing while induced.

### Table 7

<table>
<thead>
<tr>
<th>Reactor</th>
<th>Process Time Point of Harvest</th>
<th>Length of Induction [h]</th>
<th>Final OD&lt;sub&gt;600&lt;/sub&gt;</th>
<th>Final CWW [g/L]</th>
</tr>
</thead>
<tbody>
<tr>
<td>F5</td>
<td>18 h 50 min</td>
<td>4.8</td>
<td>91.4</td>
<td>122.4</td>
</tr>
<tr>
<td>F4</td>
<td>19 h 00 min</td>
<td>5.0</td>
<td>92.2</td>
<td>127.5</td>
</tr>
<tr>
<td>F6</td>
<td>19 h 30 min</td>
<td>5.5</td>
<td>96.0</td>
<td>132.4</td>
</tr>
</tbody>
</table>

Example 13

Effect of Temperature

[0178] The effect of fermentation temperature in a process essentially as described in Example 6 was investigated by running 6 fermentations at 5 different temperature setpoints. Results are summarized in Table 8. Final cell densities were sensitive to the fermentation temperature with an optimum at a temperature of 30°C.

### Table 8

<table>
<thead>
<tr>
<th>Reactor</th>
<th>Temperature [°C]</th>
<th>Final OD&lt;sub&gt;600&lt;/sub&gt;</th>
<th>Final CWW [g/L]</th>
</tr>
</thead>
<tbody>
<tr>
<td>F5</td>
<td>25.0</td>
<td>37.8</td>
<td>62</td>
</tr>
<tr>
<td>F4</td>
<td>27.5</td>
<td>80.0</td>
<td>117</td>
</tr>
<tr>
<td>F3</td>
<td>30.0</td>
<td>92.8</td>
<td>123</td>
</tr>
<tr>
<td>F4</td>
<td>30.0</td>
<td>92.4</td>
<td>125</td>
</tr>
<tr>
<td>F5</td>
<td>32.5</td>
<td>85.0</td>
<td>111</td>
</tr>
<tr>
<td>F6</td>
<td>35.0</td>
<td>79.6</td>
<td>107</td>
</tr>
</tbody>
</table>

Example 14

Sealed-Up Fermentation (50 L)

[0179] The process described in Example 6 was scaled up to a volume of 50 L in order to achieve scale-up capability from the 2 L working volume bioreactor system to a larger volume. Key process parameters for the scaled-up process are summarized in Table 9.

### Table 9

<table>
<thead>
<tr>
<th>Process Parameters of in 50 L bioreactor.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Preculture 1</td>
</tr>
<tr>
<td>Preculture 2</td>
</tr>
<tr>
<td>Inoculation of Bioreactor</td>
</tr>
<tr>
<td>Induction Start</td>
</tr>
<tr>
<td>End of Culture</td>
</tr>
</tbody>
</table>

*Relative to the time of bioreactor inoculation.

Example 15

Effect of Extended Exponential Feed

[0181] The exponential feeding phase for fermentations performed according to Examples 6 or 14 was 7 h. After this time the cells reached a density for induction, which increased during induction to the targeted maximum OD<sub>600</sub> of around 100 to 130 as final cell density. Final OD<sub>600</sub>, final CWW, final CDW, plasmid retention at induction start and harvest and Qβ concentration at the end of culture are determined for reactor runs performed as described in Examples 6 and 14, as in Example 14, wherein the exponential feeding phase is extended to a duration up to 11 h, preferably to 10 h.

Example 16

Effect of Increased Feeding During Production

[0182] Example 9 demonstrates that the glycerol does not accumulate during production phase, indicating that production might be limited by the feeding rate of induction medium. Effect of extended feeding rate of induction medium on final OD<sub>600</sub>, final CWW, final CDW, plasmid retention at induction start and harvest and Qβ concentration at the end of culture is determined in reactor runs as described in Example.
6 and 14, preferably as in Example 14, wherein the feeding rate during production is increased. Alternatively or additionally, the ratio between lactose and glycerol in the feed medium shifted towards a higher glycerol and a lower lactose concentration.

Example 17

**HPLC Analysis of Qβ CP**

**[0183]** Qβ CP was measured with an HPLC system as follows: A sample containing Qβ CP was diluted appropriately in 1x reaction buffer (50 mM tris(hydroxymethyl)aminomethane buffer pH 8.0) containing 10 mM 1.4-Dithio-DL-threitol and incubated for 15 min at 50°C in a thermomixer. After incubation the sample was centrifuged and the supernatant was stored at 2°C to 10°C until HPLC analysis. 10 to 100 µL of the sample were injected.

**[0184]** Qβ was quantified with a regression curve of known Qβ standards regressed to the HPLC peak area detected at 215 nm after elution from a C2 reversed phase column, 300 A, 5 µm, 4.6x150 mm, Vydac Inc., Hesperia, USA (Cat. No. 214TP5415) thermally equilibrated at 50°C. The flow rate through the system was 1 ml/min consisting of mobile phase A (0.12% trifluoroacetic acid in water) and mobile phase B (0.12% trifluoroacetic acid in acetonitrile) with the following gradient of phase B: 0 to 2 min constant at 40%, 2 to 8 min linear increase to 50%, 8 to 10 min constant at 50%, 10 to 10.1 min linear decrease to 40%, and 10.1 to 12 min constant at 40%.

Example 18

**Determination of Qβ VI.P by Analytical Size Exclusion Chromatography**

**[0185]** Analysis of Qβ particles by analytical size exclusion chromatography was performed using a Tskgel G5000 PW XL7 column (10 µm, 7.8x300 mm, Tosoh Biosep; Cat.-No. 08023) equilibrated in phosphate buffered saline (20 mM Na2HPO4/NaH2PO4, 150 mM NaCl pH 7.2). Elution was performed by an isocratic gradient for 20 min at 0.8 ml/min in phosphate buffered saline. The Qbeta concentration was determined from a regression curve of known Qβ standards regressed to the HPLC peak area detected at 260 nm.

Example 19

**Effect of Extended Exponential Feed**

**[0186]** The exponential feeding phase for fermentations performed according to Examples 6 or 14 was 7 h. After this time the cells reached a density for induction, which increased during induction to the targeted maximum OD600 of around 100 to 130 as final cell density. Final OD600, final CWW, plasmid retention before induction and at harvest and Qβ concentration at the end of culture were determined for reactor runs performed as described in Examples 6 and 14, preferably as in Example 14, wherein the exponential feeding phase was extended to a duration up to 12 h. In addition, the concentration of glycerol and lactose in the induction feed were changed to 300 g/L and 100 g/L respectively. The results are summarized in Table 10.

---

### TABLE 10

<table>
<thead>
<tr>
<th>Duration</th>
<th>OD600</th>
<th>CWW [g/L]</th>
<th>Plasmid Retention [%]</th>
<th>Peak Oxygen [mW]</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>86</td>
<td>122</td>
<td>100</td>
<td>99</td>
</tr>
<tr>
<td>8</td>
<td>112</td>
<td>184</td>
<td>99</td>
<td>98</td>
</tr>
<tr>
<td>9</td>
<td>136</td>
<td>217</td>
<td>100</td>
<td>98</td>
</tr>
<tr>
<td>10</td>
<td>164</td>
<td>228</td>
<td>99</td>
<td>98</td>
</tr>
<tr>
<td>11</td>
<td>200</td>
<td>262</td>
<td>100</td>
<td>97</td>
</tr>
<tr>
<td>12</td>
<td>90</td>
<td>186</td>
<td>99</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>90</td>
<td>186</td>
<td>99</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>90</td>
<td>186</td>
<td>99</td>
<td>100</td>
</tr>
</tbody>
</table>

**[0187]** According to LDS-PAGE analysis, the specific Qbeta concentration of all cultivations except for the cultivation with 12 h exponential feeding was the same. An optimum regarding absolute Qbeta yield and oxygen consumption was found for 9.5 h exponential feeding. Therefore, the process is preferably run with 9.5 h exponential feeding phase.

Example 20

**Scaled Up Fermentation (50 L)**

**[0188]** The process described in Example 6 and with 9.5 h exponential feeding phase with 300 g/L glycerol and 100 g/L lactose as described in Example 19 was scaled up to a volume of 50 L in order to evaluate scale-up capability from the 2 L working volume bioreactor system to a larger volume. Key process parameters for the scaled up process are summarized in Table 11.

---

### TABLE 11

<table>
<thead>
<tr>
<th>Culture Step</th>
<th>Description</th>
<th>Time [h]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preiculture</td>
<td>200 µL from cell bank vial were transferred into 800 ml preculture medium contained in 3000 mL shake flask and cultured for 18 h (2 flask)</td>
<td>-18*</td>
</tr>
<tr>
<td>Inculation of Bioreactor</td>
<td>Pooled total volume (approx. 1.6 L) was transferred into the 50 L Bioreactor, Initial volume: 35 L</td>
<td>0</td>
</tr>
<tr>
<td>Induction Start</td>
<td>The exponential feeding profile was switched to constant and feed was switched to induction feed</td>
<td>16.5</td>
</tr>
<tr>
<td>End of Culture</td>
<td>Culture was completed after 5 h of induction</td>
<td>21.5</td>
</tr>
</tbody>
</table>

*Relative to the time of bioreactor inoculation.

**[0189]** It was necessary to change the preicluture procedure in order to inoculate the larger reactor with approximately the same cell density. The cultures in the 50 L bioreactors were performed with the time profile optimised for the 2 L scale as described in Example 19. The final cell wet weight for six cultures was 188 g/L ± 29. Plasmid retention was 97.3% ± 1.4 at the end of culture. The concentration of Qβ CP protein in the medium at the end of culture was determined by C₄ reversed phase HPLC (Example 17) to 10.8 g/L ± 0.3. The total amount of Qβ CP was 540 g for one 50 L run. The crude extract of approximately two times concentrated biomass was
analysed for Qb CP and Qb VLP (Example 18). The concentration of Qb CP was 19.1 g/L±0.4 (C4 reversed phase HPLC), the concentration of Qb VLP was 18.8 g/L±1.1. Therefore, the VLP-yield of the fermentation process is estimated to approximately 9-11 g/l fermentation broth at the time of harvest.
| aaactgtctt | gtttacattaa | acagtaatac | aagggagtgtt | atgagccata | ttcaacggga | 1800 |
| aacgtcttgcc | tcggagcgcg | gattaatttc | caaactggat | gtctgattat | atgggtataa | 1860 |
| atgggctcgg | gataatctgg | ggcaacaggt | tcggacatac | tatcgagggt | atggaagaag | 1920 |
| cgatgccgca | gcgggtcttc | tgaacacagg | caagggcaga | gtctgcaatac | atgcatacga | 1980 |
| tggatgtggc | gtaccaactt | gcgtgacggga | atttatgctc | cttcgagacca | tcaagcactt | 2040 |
| tacccgtactc | ctgtattgatg | caaggtctct | cccccgtgcc | atcccagggga | aacaacatt | 2100 |
| caaggtattta | gaagaatatac | ctggatccgg | tgaaatatatt | gtgtagtccg | tgccagtgtt | 2160 |
| cctgcgctgcg | ttcgatccga | tctctgctgt | taatgcctct | ctaaaccgcc | atgccgatat | 2220 |
| ttggctctgct | cagggcgaat | caagtaatga | taacggttgg | gttgatgaca | gtgatattttga | 2280 |
| tgcacagagc | aatgggctg | ctgctgaaaca | agttggagaa | gaatacgtata | agctttttggc | 2340 |
| atcttcgacc | gatttcgtcc | tcagaatctgg | tctcttcc | cttggataacc | tattttttttgc | 2400 |
| cagagagaggga | ttaataggtt | gttatgattg | ttgaaagattg | ggatcgcag | aagcatacoca | 2460 |
| ggtctacctg | atctggataa | gcgcctcgcg | tgagttttttt | cttctacactt | aagaaaggttt | 2520 |
| ttggcataaaa | ttagctttagg | ataactcctga | tgtgaataaa | ttgcatgttc | atttttagt | 2580 |
| cagatggttt | tttctcaaoct | gtgaaccaagt | ttaactcatat | gtaacttgtag | ttgatatttga | 2640 |
| actcactttt | taatattaaaa | ggtatctaggt | gaagatccctt | tttgataacta | taactgacaaa | 2700 |
| aatccctattc | gcgttagtttt | ctcctacaact | agcgtccagac | cccgtagaaa | agcatcaaggg | 2760 |
| atccttcgta | gatcccttttt | ttctgtgctt | aatacagctg | tcgtcaaca | aaaaaaaca | 2820 |
| gctacacggc | gttggcggttg | tggcagctac | aagctcgsca | cccctttttttttc | cgaagtttcag | 2880 |
| tggcaaccggc | agagccagga | taacaaatatc | tgtcctctctt | gccgtagctg | actgagggc | 2940 |
| ccactcgaag | aacgttggagc | cacocctccat | atccctgcgt | ctcgtaatccc | tgtatcaggt | 3000 |
| ggtctgctgct | agtggagata | agtggctgct | tcaggggggt | gccctactag | gatcaaaaca | 3060 |
| ggaataagcgc | cgctgctgctg | gcggataaaggggg | ggtcggctgc | cacacgccca | gctggagagc | 3120 |
| aaacgaotacg | aacgaaagta | gataccatac | ggtctgatat | ttgaaaagagc | caagctcc | 3180 |
| cgaaggggga | aagggacagca | ggtacctcgct | aagggcaggg | gtcggaaacag | gagagcgcagc | 3240 |
| gagagacgag | cccagggggga | aagcomtgcgta | tttttatagt | ctgctcgggtt | tggcacactt | 3300 |
| ctgagcactc | gggtgctttt | tcgtgatcgc | gtcggaggggg | gcggagcattc | gaaaaaaca | 3360 |
| cagcaacagcg | gctttttaaaggggg | gcgtccttggc | cttttgtcgt | cttttttcact | gtcgggttct | 3420 |
| tctcctcgtta | tcccctggtttc | cgggtatatc | cgtgagttatac | cttctgatcc | cacaacctcc | 3480 |
| cgctgagggc | aagccagcag | cccagcagcag | cgagcctagc | aggagggagc | gcgaagagcg | 3540 |
| cctgtagggc | cttttttcctc | ctgtgattttc | ttcagcactt | cgggagctgct | ggctggatac | 3600 |
| ttcacatca | atctgctctg | atgcctgata | gttgagaacag | tatcactctcc | gctagctgtc | 3660 |
| cggtatggggg | ccgtctcgcag | gcocgcaacc | cccaaacaac | cgctgcaagc | gcccgtgagg | 3720 |
| ggttggctgc | tcggcagtactc | cgctttactac | cagctgtga | cgcgtcctcgg | ggtgtagcag | 3780 |
| tgtctcaggtg | ttttttcttct | cttggtgatc | aaccagagaa | gcggcagggag | aagttgcgtaa | 3840 |
| gcggcggagc | gagaagactt | aacaagatgtc | gcggtagcct | ccgtgctcagc | ctcctgtgcgt | 3900 |
| ttcacagacs | ggtagtaatctc | ctgctgcttct | aataaagcggg | ccsatcttaag | gggggttttt | 3960 |
| ttcgggttgctt | ctgctgctg | ctgctgctg | ggggagtttc | ttctgatgagc | ggtagatga | 4020 |
-continued

cgatatgaac gagagaggt gctcagcata cgggtgtaag atgatgaaca tgcgctggta 4080
cgagattt gtaggggtaa agaaccgccg ttgtaggatc ggacggagat cagagaaatc 4140
aaactcaagc aataaggagt attgtaata acaatagttg ggatcggaca ggttagccag 4200
cagcctacgy cctgcggaag cccggaatata tgggtggcgg ggctggacctt cccgtttccc 4260
agacttttaag aacaaagga aacagagacc accatatgttg ttgctcaggtt gcgagcaagt 4320
ttgcaagacc agcgctccca cgttggctcg cgcattccgt attcattctg caacacgcta 4380
aggcaacccc gcgacgcatt cgggttgtctc caacagcagg gaaacgatcat gcggacccgg 4440
agggcagagc caagcctgcc gctggtagcgc cgggttggcgg tgctggagatt gagggagcgcg 4500
atggatatgtt tagtttgcag gcagttggtcc tctcttgtaa aatgcttgag 4560
gtcctcaattc ttgagggtgt gaaactcgtta gcgagggcgc acgggtttccc attcaggtcgt 4620
agttgctgac gctctccgca cccggagcga cggggaggg gacgacaaggt ttagggcgg 4680
cggctcaccat ccacgtaacc cgtagcgtcgc tgccgatggac ggccgcataaa atgcgcctga 4740
cgatccaggg tccatgtgcg gaagttggag ttgactaagggc gcggatccgga 4800
ctgccctggt gtcggacact aclctttcggt gcgacggaata cccccggcgg 4860
tggcggcga agggcgggca gtcctgtacg agcctggcg gcggagaactgg 4920
ccggaacagc gactcctgcc cggatggcc cccgagcgag gcacccctaa ggtcgtctggc 4980
cgaacagcttt gcggataagg gcagcgaggt gcagcttgac gggcttggtgc aaggtctccga 5040
atatcgccag ggacagccag acctccgaa gctcgaacctgc atcgggctgcc 5100
tgcgggctg cggctggcgac atcttcgctcc ggctgcaccg caagatcccgag 5160
gtgcggcgcg catagctctgg cccggcgcgg gccgacagga gtgtcctcgg tttgacggtc 5220
tcaacggccggct tcccctcttt gcacatccctc cttggagcag ccacgcaagttg 5280
gttgagcggcgac cgtcctccag cagggagcag cggaggaaggta cggaggaagctg 5340
cccagaaagcc cccgacgacag ggctggccgag cctccgacac gccgaacacag ggctctcgatgt 5400
ggcgcggagtg cggacgacag ctctccgattc cggccgatgg ccggacgaccag 5460
ccgacacgttg ggcgcggaga atggcggctcag cggctgatggc ggcgccggtcg 5520
atacgctgca cggctgacca atggctctggc cgctggacag cccgagggag cttggatt 5579

<210> SEQ ID NO 2
<211> LENGTH: 245
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: promoter sequence - chemically synthesized
<400> SEQUENCE: 2

cgacactgacg gtcgacaacatt gtttctggcg tcaggcegccgt atcggaagctgtggttaggc 60
tgagcaagtc gtaaatcctag gtaatcctag tcgaaggtccgt ggtgagttgccg 120
aatggtttttt gcggctcagacatcagggtt cggccaaaattctcagtaaagagctgtggagc 180
aatatcgccag ccaggcttcag atgtctggga atggtagggc gacaatcctag tccagcagga 240
aacag
245

<210> SEQ ID NO 3
<211> LENGTH: 19
-continued

<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Shine-Dalgarno Sequence - chemically synthesized

<400> SEQUENCE: 3

gaagggagaa asaaaaatg

<210> SEQ ID NO 4
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Shine-Dalgarno Sequence - chemically synthesized

<400> SEQUENCE: 4

agaggttaaa aacgagatg

<210> SEQ ID NO 5
<211> LENGTH: 133
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage Q-beta

<400> SEQUENCE: 5

Met Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly
  1   5   10   15
Lys Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly
   20  25   30
Val Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Lys Arg
  35   40   45
Val Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys
  50  55   60
Val Gln Val Lys Ile Gln Asp Pro Thr Ala Cys Thr Ala Asn Gly Ser
  65   70  75   80
Cys Asp Pro Ser Val Thr Arg Gln Ala Tyr Ala Asp Val Thr Phe Ser
   85  90  95
Phe Thr Gln Tyr Ser Thr Asp Glu Glu Arg Ala Phe Val Arg Thr Glu
 100 105 110
Leu Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln
 115 120 125
Leu Asn Pro Ala Tyr
 130

<210> SEQ ID NO 6
<211> LENGTH: 1017
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Expression construct - chemically synthesized

<400> SEQUENCE: 6

atggcaaata tagaaactgt tactttaggt aacatcgagga aagatgagaa aacaactctgt
  60
gttccctacct gcgttggggt aatacccact aacgacgtg cctcgctttc acaagcgggt
 120
gcagttctct gcgttggaag gcgtgtacac gtttggttat ctcagccttc tggcaatcg
 180
aagaaactca aagtcagatc taagatcccc aacgagcctc gttgcaatcg aacaggttct
 240
-continued

tgagacctc cggacctgg cacggtggat cgtgacgctg ccccttctgcct ccacgccagct
300
agacctgatg agggagctgg tttgggtggc gcagagctgat cgtgagctgct
360
cctaggctgg atctgctgtag ttcagctgag acacgcttgct cattgctggct
420
ggtgctgctg ggtgcaaaaa ccagaagctg attacgctgac caagattgctgc accagccgca
490
gggacaggtg agtatacttg cccctttgaca atttggctccc tagaggaggt ttagagctgt
540
cctactaaga accagccccgct gctactttat atatgctttg aacctccagcc tggcgaatatt
600
gagtggccct tcaaaagatct ttgagcagta aacaagttgct gcagatttgga tttcgggttt
660
agtataacca cgtctccgctg ttggtggctgc aacgattata ttgacccgta tgcagctatt
720
cctgtctactg atcaggtcctgt gcctgatcac cagattgata ttgcagaggg caagaaaccct
780
gggtcgtttgg gcccttgtcga gcagttcttc tattcctaag tgcataaatgc tttgctttct
840
cctaggcata ttgagcagta ctacagcgag gcctcgtttag ttggtttttg ggcagatcga
900
tccaggtgtg gcctcagatcc gttgccccc aactagtggt ttaaagcttt atgtcttatt
960
cagagccttg tattgctttc tctggtcttag taactaaggga tgaatgccat gcttaag
1017

<210> SEQ ID NO 7
<211> LENGTH: 132
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage Q-beta

<400> SEQUENCE: 7

Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Arg Asp Gly Lys
1 5 10 15
Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val
20 25 30
Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Lys Arg Val
35 40 45
Thr Val Ser Val Ser Gln Pro Ser Arg Asn Gly Lys Asn Tyr Lys Val
50 55 60
Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys
65 70 75 80
Asp Pro Ser Val Thr Arg Gln Lys Tyr Ala Asp Val Thr Phe Ser Phe
85 90 95
Thr Gln Tyr Ser Thr Asp Glu Arg Ala Phe Val Arg Thr Glu Leu
100 105 110
Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu
115 120 125
Asn Pro Ala Tyr
130

<210> SEQ ID NO 8
<211> LENGTH: 132
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage Q-beta

<400> SEQUENCE: 8

Ala Lys Leu Glu Thr Val Thr Leu Gly Lys Ile Gly Lys Arg Asp Gly Lys
1 5 10 15
Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val
20 25 30
Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Lys Arg Val
<table>
<thead>
<tr>
<th>35</th>
<th>40</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys Val</td>
<td>Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys</td>
<td>Asp Pro Ser Val Thr Arg Gln Lys Tyr Ala Asp Val Thr Phe Ser Phe</td>
</tr>
<tr>
<td>50</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys</td>
<td>Asp Pro Ser Val Thr Arg Gln Lys Tyr Ala Asp Val Thr Phe Ser Phe</td>
<td>Thr Gln Tyr Ser Thr Asp Glu Glu Arg Ala Phe Val Arg Thr Glu Leu</td>
</tr>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>Asp Pro Ser Val Thr Arg Gln Lys Tyr Ala Asp Val Thr Phe Ser Phe</td>
<td>Thr Gln Tyr Ser Thr Asp Glu Glu Arg Ala Phe Val Arg Thr Glu Leu</td>
<td>Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu</td>
</tr>
<tr>
<td>85</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu</td>
<td>Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu</td>
<td>Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu</td>
</tr>
<tr>
<td>115</td>
<td>120</td>
<td>125</td>
</tr>
<tr>
<td>Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu</td>
<td>Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu</td>
<td>Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu</td>
</tr>
<tr>
<td>130</td>
<td>135</td>
<td>140</td>
</tr>
</tbody>
</table>

**SEQ ID NO 9**
**LENGTH:** 132
**TYPE:** PRT
**ORGANISM:** Bacteriophage Q-beta

**SEQUENCE:**

```
Ala Arg Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Arg Asp Gly Lys 15
Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val 20
Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Arg Val 35
Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys Val 50
Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys 65
Asp Pro Ser Val Thr Arg Gln Lys Tyr Ala Asp Val Thr Phe Ser Phe 85
Thr Ala Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val 100
Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu
```

**SEQ ID NO 10**
**LENGTH:** 132
**TYPE:** PRT
**ORGANISM:** Bacteriophage Q-beta

**SEQUENCE:**

```
Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly Arg 15
Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val 20
Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Arg Val 35
Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys Val 50
Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys
```

**SEQ ID NO 11**
**LENGTH:** 132
**TYPE:** PRT
**ORGANISM:** Bacteriophage Q-beta

**SEQUENCE:**

```
Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly Arg 15
Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val 20
Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Arg Val 35
Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys Val 50
Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys
```

**SEQ ID NO 12**
**LENGTH:** 132
**TYPE:** PRT
**ORGANISM:** Bacteriophage Q-beta

**SEQUENCE:**

```
Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly Arg 15
Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val 20
Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Arg Val 35
Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys Val 50
Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys
```

**SEQ ID NO 13**
**LENGTH:** 132
**TYPE:** PRT
**ORGANISM:** Bacteriophage Q-beta

**SEQUENCE:**

```
Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly Arg 15
Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val 20
Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Arg Val 35
Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys Val 50
Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys
```

**SEQ ID NO 14**
**LENGTH:** 132
**TYPE:** PRT
**ORGANISM:** Bacteriophage Q-beta

**SEQUENCE:**

```
Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly Arg 15
Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val 20
Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Arg Val 35
Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys Val 50
Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys
```

**SEQ ID NO 15**
**LENGTH:** 132
**TYPE:** PRT
**ORGANISM:** Bacteriophage Q-beta

**SEQUENCE:**

```
Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly Arg 15
Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val 20
Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Arg Val 35
Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys Val 50
Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys
```

**SEQ ID NO 16**
**LENGTH:** 132
**TYPE:** PRT
**ORGANISM:** Bacteriophage Q-beta

**SEQUENCE:**

```
Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly Arg 15
Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly Val 20
Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Arg Val 35
Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys Val 50
Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser Cys
```
### continued

<table>
<thead>
<tr>
<th>65</th>
<th>70</th>
<th>75</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp Pro Ser Val Thr Arg Gln Lys Tyr Ala Asp Val Thr Phe Ser Phe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>90</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Thr Gln Tyr Ser Thr Asp Glu Arg Ala Phe Val Arg Thr Glu Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>120</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>Asp Pro Ala Tyr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SEQ ID NO 11**

**LENGTH: 132**

**TYPE: PRT**

**ORGANISM: Bacteriophage Q-beta**

**SEQUENCE: 11**

<table>
<thead>
<tr>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala Arg Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly Arg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Lys Arg Val</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys Val</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Thr Val Ser Thr Val Ser Gln Pro Asp Val Thr Ala Cys Thr Ala Gly Ser Cys</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td>Thr Gln Tyr Ser Thr Asp Glu Arg Ala Phe Val Arg Thr Glu Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>90</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>120</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>Asp Pro Ala Tyr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SEQ ID NO 12**

**LENGTH: 131**

**TYPE: PRT**

**ORGANISM: Bacteriophage AP205**

**SEQUENCE: 12**

<table>
<thead>
<tr>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met Ala Asn Lys Pro Met Gln Pro Ile Thr Ser Thr Ala Asn Lys Ile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Leu Arg Glu Val Lys Val Gly Ala Glu Leu Asn Asn Val Ser</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Gly Gln Tyr Val Ser Val Tyr Lys Arg Pro Ala Pro Lys Pro Glu Gly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Cys Ala Asp Ala Cys Val Ile Met Pro Asn Glu Asn Gln Ser Ile Arg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td>Thr Val Ile Ser Gly Ser Ala Glu Asn Leu Ala Thr Leu Lys Ala Glu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>90</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Trp Glu Thr His Lys Arg Asn Val Asp Thr Leu Phe Ala Ser Gly Asn</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ala Gly Leu Gly Phe Leu Asp Pro Thr Ala Ala Ile Val Ser Ser Asp  

Thr Thr Ala

110

<210> SEQ ID NO 13
<211> LENGTH: 131
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage AP205

<400> SEQUENCE: 13
Met Ala Asn Lys Thr Met Gln Pro Ile Thr Ser Thr Ala Asn Lys Ile  
1   5   10   15
Val Trp Ser Asp Pro Thr Arg Leu Ser Thr Thr Phe Ser Ala Ser Leu  
20  25  30
Leu Arg Gln Arg Val Lys Val Gly Ile Ala Glu Asn Asn Val Ser  
35  40  45
Gly Glu Tyr Val Ser Val Tyr Lys Arg Pro Ala Pro Lys Pro Glu Gly  
50  55  60
Cys Ala Asp Ala Cys Val Ile Met Pro Asn Glu Asn Gln Ser Ile Arg  
65  70  75  80
Thr Val Ile Ser Gly Ser Ala Glu Asn Leu Ala Thr Leu Lys Ala Glu  
85  90  95
Trp Glu Thr His Lys Arg Asn Val Asp Thr Leu Phe Ala Ser Gly Asn  
100 105 110
Ala Gly Leu Gly Phe Leu Asp Pro Thr Ala Ala Ile Val Ser Ser Asp  
115 120 125
Thr Thr Ala

130

<210> SEQ ID NO 14
<211> LENGTH: 131
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage AP205

<400> SEQUENCE: 14
Met Ala Asn Lys Pro Met Gln Pro Ile Thr Ser Thr Ala Asn Lys Ile  
1   5   10   15
Val Trp Ser Asp Pro Thr Arg Leu Ser Thr Thr Phe Ser Ala Ser Leu  
20  25  30
Leu Arg Gln Arg Val Lys Val Gly Ile Ala Glu Asn Asn Val Ser  
35  40  45
Gly Glu Tyr Val Ser Val Tyr Lys Arg Pro Ala Pro Lys Pro Glu Gly  
50  55  60
Cys Ala Asp Ala Cys Val Ile Met Pro Asn Glu Asn Gln Ser Ile Arg  
65  70  75  80
Thr Val Ile Ser Gly Ser Ala Glu Asn Leu Ala Thr Leu Lys Ala Glu  
85  90  95
Trp Glu Thr His Lys Arg Asn Val Asp Thr Leu Phe Ala Ser Gly Asn  
100 105 110
Ala Gly Leu Gly Phe Leu Asp Pro Thr Ala Ala Ile Val Ser Ser Asp  
115 120 125
Thr Thr Ala
<210> SEQ ID NO 15
<211> LENGTH: 329
<212> TYPE: DNA
<213> ORGANISM: Bacteriophage Q-beta

<400> SEQUENCE: 15

Met Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly
1 5 10 15
Lys Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly
20 25 30
Val Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Lys Arg
35 40 45
Val Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys
50 55 60
Val Gln Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser
65 70 75 80
Cys Asp Pro Ser Val Thr Arg Gln Ala Tyr Ala Asp Val Thr Phe Ser
85 90 95
Phe Thr Gln Tyr Ser Thr Asp Glu Glu Arg Ala Phe Val Arg Thr Glu
100 105 110
Leu Ala Ala Leu Leu Ala Ser Leu Leu Ile Asp Ala Ile Asp Gln
115 120 125
Leu Asn Pro Ala Tyr Trp Thr Leu Ile Ala Gly Gly Gly Gly Ser Gly
130 135 140
Ser Lys Pro Asp Pro Val Ile Pro Asp Pro Ile Asp Pro Pro Pro Pro
145 150 155 160
Gly Thr Gly Lys Tyr Thr Cys Pro Phe Ala Ile Trp Ser Leu Gly Glu
165 170 175
Val Tyr Glu Pro Pro Thr Lys Asn Arg Pro Trp Pro Ile Tyr Asn Ala
180 185 190
Val Gln Leu Gln Pro Arg Glu Phe Asp Val Ala Leu Lys Asp Leu Leu
195 200 205
Gly Asn Thr Lys Trp Arg Asp Trp Asp Arg Ser Arg Tyr Thr Thr
210 215 220
Phe Arg Gly Cys Arg Gly Asn Gly Tyr Ile Asp Leu Asp Ala Thr Tyr
225 230 235 240
Leu Ala Thr Asp Gln Ala Met Arg Asp Gln Lys Tyr Asp Ile Arg Glu
245 250 255
Gly Lys Lys Pro Gly Ala Phe Gly Asn Ile Glu Glu Phe Ile Tyr Leu
260 265
Lys Ser Ile Asn Ala Tyr Cys Ser Leu Ser Asp Ile Ala Ala Tyr His
275 280 285
Ala Asp Gly Val Ile Val Gly Phe Trp Arg Asp Pro Ser Ser Gly Gly
290 295 300
Ala Ile Pro Phe Asp Phe Thr Lys Phe Asp Lys Thr Lys Cys Pro Ile
305 310 315 320
Gln Ala Val Ile Val Val Pro Arg Ala
325

<210> SEQ ID NO 16
<211> LENGTH: 129
**Continued**

```plaintext
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage R17
<400> SEQUENCE: 16

Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asn Asp Gly Gly Thr Gly
1  5 10 15
Asn Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Val Ala Glu Trp
20 25 30
Ile Ser Ser Asn Ser Arg Ser Gin Ala Tyr Lys Val Thr Cys Ser Val
35 40 45
Arg Gin Ser Ser Ala Gin Arg Tyr Thr Ile Lys Val Glu Val
50 55 60
Pro Lys Val Ala Thr Gin Thr Val Gly Val Glu Leu Pro Val Ala
60 70 75 80
Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile Pro Ile Phe Ala
90 95 100
Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met Gin Gly Leu Leu
105 110
Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala Asn Ser Gly Ile
115 120 125
Tyr

<210> SEQ ID NO 17
<211> LENGTH: 130
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage fr
<400> SEQUENCE: 17

Met Ala Ser Asn Phe Glu Glu Phe Val Leu Val Asp Asn Gly Gly Thr
1  5 10 15
Gly Asp Val Lys Val Ala Pro Ser Asn Phe Ala Asn Gly Val Ala Glu
20 25 30
Trp Ile Ser Ser Asn Ser Arg Ser Gin Ala Tyr Lys Val Thr Cys Ser
35 40 45
Val Arg Gin Ser Ser Ala Asn Arg Arg Lys Tyr Thr Val Lys Val Glu
50 55 60
Val Pro Lys Val Ala Thr Gin Val Gin Gly Gly Glu Val Leu Pro Val
65 70 75 80
Ala Ala Trp Arg Ser Tyr Met Asn Met Glu Leu Thr Ile Pro Val Phe
85 90 95
Ala Thr Asn Asp Asp Cys Ala Leu Ile Val Lys Ala Leu Gin Gly Thr
100 105 110
Phe Lys Thr Gly Asn Pro Ile Ala Thr Ala Ile Ala Ala Asn Ser Gly
115 120 125
Ile Tyr
130

<210> SEQ ID NO 18
<211> LENGTH: 130
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage GA
<400> SEQUENCE: 18

Met Ala Thr Leu Arg Ser Phe Val Leu Val Asp Asn Gly Gly Thr Gly
1  5 10 15
```
Asn Val Thr Val Val Pro Val Ser Asn Ala Asn Gly Val Ala Glu Trp
20 25 30
Leu Ser Asn Asn Ser Arg Ser Gln Ala Tyr Arg Val Thr Ala Ser Tyr
35 40 45
Arg Ala Ser Gly Ala Asp Lys Arg Lys Tyr Ala Ile Lys Leu Glu Val
50 55 60
Pro Lys Ile Val Thr Gln Val Val Asn Gly Val Glu Leu Pro Gly Ser
65 70 75 80
 Ala Trp Lys Ala Tyr Ala Ser Ile Asp Leu Thr Ile Pro Ile Phe Ala
85 90 95
 Ala Thr Asp Arg Val Thr Ile Ser Lys Ser Leu Ala Gly Leu Phe
100 105 110
Lys Val Gly Asn Pro Ile Ala Glu Ala Ile Ser Ser Gln Ser Gly Phe
115 120 125
Tyr Ala
130

<210> SEQ ID NO 19
<211> LENGTH: 132
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage SP

<400> SEQUENCE: 19
Met Ala Lys Leu Asn Gln Val Thr Leu Ser Lys Ile Gly Lys Asn Gly
1 5 10 15
 Asp Gln Thr Leu Thr Leu Thr Pro Arg Gly Val Asn Pro Thr Asn Gly
20 25 30
 Val Ala Ser Leu Ser Glu Ala Gly Ala Val Pro Ala Leu Glu Lys Arg
35 40 45
 Val Thr Val Ser Val Ala Gln Pro Ser Arg Asn Arg Lys Asn Phe Lys
50 55 60
 Val Gln Ile Lys Leu Gln Asn Pro Thr Ala Cys Thr Arg Asp Ala Cys
65 70 75 80
 Asp Pro Ser Val Thr Arg Ser Ala Phe Ala Asp Val Thr Leu Ser Phe
85 90 95
 Thr Ser Tyr Ser Thr Asp Glu Glu Arg Ala Leu Ile Arg Thr Glu Leu
100 105 110
 Ala Ala Leu Leu Ala Asp Pro Leu Ile Val Asp Ala Ile Asp Asn Leu
115 120 125
Asn Pro Ala Tyr
130

<210> SEQ ID NO 20
<211> LENGTH: 329
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage SP

<400> SEQUENCE: 20
Ala Lys Leu Asn Gln Val Thr Leu Ser Lys Ile Gly Lys Asn Gly Asp
1 5 10 15
 Gln Thr Leu Thr Leu Thr Pro Arg Gly Val Asn Pro Thr Asn Gly Val
20 25 30
 Ala Ser Leu Ser Glu Ala Gly Ala Val Pro Ala Leu Glu Lys Arg Val
<table>
<thead>
<tr>
<th>Thr</th>
<th>Val</th>
<th>Ser</th>
<th>Val</th>
<th>Ala</th>
<th>Gln</th>
<th>Pro</th>
<th>Ser</th>
<th>Arg</th>
<th>Asn</th>
<th>Arg</th>
<th>Lys</th>
<th>Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Ile</td>
<td>Lys</td>
<td>Leu</td>
<td>Gln</td>
<td>Asn</td>
<td>Pro</td>
<td>Thr</td>
<td>Ala</td>
<td>Cys</td>
<td>Thr</td>
<td>Arg</td>
<td>Asp</td>
</tr>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ser</td>
<td>Val</td>
<td>Thr</td>
<td>Arg</td>
<td>Ser</td>
<td>Ala</td>
<td>Phe</td>
<td>Ala</td>
<td>Asp</td>
<td>Val</td>
<td>Thr</td>
<td>Leu</td>
</tr>
<tr>
<td>95</td>
<td>90</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Tyr</td>
<td>Ser</td>
<td>Thr</td>
<td>Asp</td>
<td>Glu</td>
<td>Glu</td>
<td>Arg</td>
<td>Ala</td>
<td>Leu</td>
<td>Ile</td>
<td>Arg</td>
<td>Thr</td>
</tr>
<tr>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Leu</td>
<td>Leu</td>
<td>Ala</td>
<td>Asp</td>
<td>Pro</td>
<td>Leu</td>
<td>Ile</td>
<td>Val</td>
<td>Asp</td>
<td>Ala</td>
<td>Ile</td>
<td>Asp</td>
</tr>
<tr>
<td>115</td>
<td>120</td>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Pro</td>
<td>Ser</td>
<td>Tyr</td>
<td>Trp</td>
<td>Ala</td>
<td>Leu</td>
<td>Val</td>
<td>Ala</td>
<td>Ser</td>
<td>Ser</td>
<td>Gly</td>
<td>Gly</td>
</tr>
<tr>
<td>130</td>
<td>135</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Pro</td>
<td>Ser</td>
<td>Asp</td>
<td>Pro</td>
<td>Asp</td>
<td>Val</td>
<td>Pro</td>
<td>Val</td>
<td>Val</td>
<td>Pro</td>
<td>Asp</td>
<td>Val</td>
</tr>
<tr>
<td>145</td>
<td>150</td>
<td>155</td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Gly</td>
<td>Thr</td>
<td>Gly</td>
<td>Arg</td>
<td>Tyr</td>
<td>Lys</td>
<td>Cys</td>
<td>Pro</td>
<td>Phe</td>
<td>Ala</td>
<td>Cys</td>
<td>Tyr</td>
</tr>
<tr>
<td>165</td>
<td>170</td>
<td>175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Ile</td>
<td>Tyr</td>
<td>Glu</td>
<td>Val</td>
<td>Gly</td>
<td>Gly</td>
<td>Ser</td>
<td>Pro</td>
<td>Ile</td>
<td>Tyr</td>
<td>Glu</td>
<td>Arg</td>
</tr>
<tr>
<td>180</td>
<td>185</td>
<td>190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Asp</td>
<td>Glu</td>
<td>Val</td>
<td>Ser</td>
<td>Thr</td>
<td>Phe</td>
<td>Asp</td>
<td>Tyr</td>
<td>Ala</td>
<td>Leu</td>
<td>Glu</td>
<td>Asp</td>
</tr>
<tr>
<td>195</td>
<td>200</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Asn</td>
<td>Thr</td>
<td>Asn</td>
<td>Trp</td>
<td>Arg</td>
<td>Asn</td>
<td>Trp</td>
<td>Asp</td>
<td>Gln</td>
<td>Arg</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>210</td>
<td>215</td>
<td>220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Ala</td>
<td>Asn</td>
<td>Arg</td>
<td>Arg</td>
<td>Arg</td>
<td>Gly</td>
<td>Asn</td>
<td>Gly</td>
<td>Tyr</td>
<td>Ile</td>
<td>Asp</td>
<td>Leu</td>
</tr>
<tr>
<td>225</td>
<td>230</td>
<td>235</td>
<td>240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Met</td>
<td>Gln</td>
<td>Ser</td>
<td>Asp</td>
<td>Phe</td>
<td>Val</td>
<td>Leu</td>
<td>Ser</td>
<td>Gly</td>
<td>Arg</td>
</tr>
<tr>
<td>245</td>
<td>250</td>
<td>255</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Arg</td>
<td>Lys</td>
<td>Val</td>
<td>Lys</td>
<td>Phe</td>
<td>Pro</td>
<td>Gly</td>
<td>Ala</td>
<td>Phe</td>
<td>Gly</td>
<td>Ser</td>
<td>Ile</td>
</tr>
<tr>
<td>260</td>
<td>265</td>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Asn</td>
<td>Ile</td>
<td>Gln</td>
<td>Gly</td>
<td>Asp</td>
<td>Ala</td>
<td>Trp</td>
<td>Leu</td>
<td>Asp</td>
<td>Leu</td>
<td>Ser</td>
<td>Glu</td>
</tr>
<tr>
<td>275</td>
<td>280</td>
<td>285</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Arg</td>
<td>Ser</td>
<td>Tyr</td>
<td>Gly</td>
<td>Met</td>
<td>Val</td>
<td>Ile</td>
<td>Gly</td>
<td>Phe</td>
<td>Trp</td>
<td>Thr</td>
<td>Asp</td>
</tr>
<tr>
<td>290</td>
<td>295</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Gln</td>
<td>Leu</td>
<td>Pro</td>
<td>Thr</td>
<td>Asp</td>
<td>Phe</td>
<td>Thr</td>
<td>Gln</td>
<td>Phe</td>
<td>Asn</td>
<td>Ser</td>
<td>Ala</td>
</tr>
<tr>
<td>305</td>
<td>310</td>
<td>315</td>
<td>320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Gln</td>
<td>Thr</td>
<td>Val</td>
<td>Ile</td>
<td>Ile</td>
<td>Ile</td>
<td>Pro</td>
<td>Ser</td>
<td>325</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SEQ ID NO 21**
**LENGTH: 130**
**ORGANISM: Bacteriophage MS2**

**SEQUENCE:**

```
Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
 1  5  10  15
Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Val Ala Glu
 20  25  30
Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
 35  40  45
Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
 50  55  60
Val Pro Lys Val Ala Thr Gln Thr Val Gly Gly Val Glu Leu Pro Val
```
 Ala Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile Pro Ile Phe  
   85                        90                        95  
 Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met Gln Gly Leu  
   100      105                        110  
 Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala Asn Ser Gly  
   115                        120                        125  
 Ile Tyr  
   130  

<210> SEQ ID NO 22  
<211> LENGTH: 133  
<212> TYPE: PRT  
<213> ORGANISM: Bacteriophage M11  

<400> SEQUENCE: 22  
Met Ala Lys Leu Gln Ala Ile Thr Leu Ser Gly Ile Gly Lys Gly  
   1 5 10 15  
Asp Val Thr Leu Asp Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly  
   20 25 30  
Val Ala Ala Leu Ser Glu Ala Val Pro Ala Leu Glu Lys Arg  
   35 40 45  
Val Thr Ile Ser Val Ser Glu Pro Ser Arg Arg Lys Asn Tyr Lys  
   50 55 60  
Val Gln Val Lys Ile Gln Asn Pro Thr Ser Cys Thr Ala Ser Gly Thr  
   65 70 75 80  
Cys Asp Pro Ser Val Thr Arg Ser Ala Tyr Ser Asp Val Thr Phe Ser  
   85 90 95  
Phe Thr Gln Tyr Ser Thr Val Glu Glu Arg Ala Leu Val Arg Thr Glu  
   100 105 110  
Leu Gln Ala Leu Leu Ala Asp Pro Met Leu Val Asn Ala Ile Asp Asn  
   115 120 125  
Leu Asn Pro Ala Tyr  
   130  

<210> SEQ ID NO 23  
<211> LENGTH: 133  
<212> TYPE: PRT  
<213> ORGANISM: Bacteriophage MX1  

<400> SEQUENCE: 23  
Met Ala Lys Leu Gln Ala Ile Thr Leu Ser Gly Ile Gly Lys Asn Gly  
   1 5 10 15  
Asp Val Thr Leu Asn Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly  
   20 25 30  
Val Ala Ala Leu Ser Glu Ala Val Pro Ala Leu Glu Lys Arg  
   35 40 45  
Val Thr Ile Ser Val Ser Glu Pro Ser Arg Arg Lys Asn Tyr Lys  
   50 55 60  
Val Gln Val Lys Ile Gln Asn Pro Thr Ser Cys Thr Ala Ser Gly Thr  
   65 70 75 80  
Cys Asp Pro Ser Val Thr Arg Ser Ala Tyr Ser Asp Val Thr Phe Ser  
   85 90 95  
Phe Thr Gln Tyr Ser Thr Val Glu Glu Arg Ala Leu Val Arg Thr Glu  
   100 105 110  
Leu Gln Ala Leu Leu Ala Asp Pro Met Leu Val Asn Ala Ile Asp Asn  
   115 120 125  
Leu Asn Pro Ala Tyr  
   130  

Leu Lys Ala Leu Leu Ala Asp Pro Met Leu Ile Asp Ala Ile Asp Asn
115 120 125
Leu Asn Pro Ala Tyr
130
<210> SEQ ID NO: 24
<211> LENGTH: 330
<212> TYPE: PRT
<213> ORGANISM: Bacteriophage NL95
<400> SEQUENCE: 24
Met Ala Lys Leu Asn Lys Val Thr Leu Thr Gly Ile Gly Lys Ala Gly
1 5 10 15
Asn Gln Thr Leu Thr Leu Thr Pro Arg Gly Val Asn Pro Thr Asn Gly
20 25 30
Val Ala Ser Leu Ser Glu Ala Gly Ala Val Pro Ala Leu Glu Lys Arg
35 40 45
Val Thr Val Ser Val Ala Gln Ser Pro Ser Arg Asn Arg Lys Asn Tyr Lys
50 55 60
Val Gln Ile Lys Leu Gln Asn Pro Thr Ala Cys Thr Lys Asp Ala Cys
65 70 75 80
Asp Pro Ser Val Thr Arg Ser Gly Ser Arg Asp Val Thr Leu Ser Phe
85 90 95
Thr Ser Tyr Ser Thr Glu Arg Glu Ala Leu Ile Arg Thr Glu Leu
100 105 110
Ala Ala Leu Leu Lys Asp Leu Ile Val Asp Ala Ile Asp Asn Leu
115 120 125
Asn Pro Ala Tyr Trp Ala Ala Leu Leu Ala Ala Ser Pro Gly Gly Gly
130 135 140
Asn Asn Pro Tyr Pro Gly Val Pro Asp Ser Pro Asn Val Lys Pro Pro
145 150 154 160
Gly Gly Thr Gly Thr Tyr Arg Arg Gly Thr Val Asp Pro Phe Ala Cys Tyr Arg Arg Gly
165 170 175
Glu Leu Ile Thr Glu Ala Lys Asp Asp Lys Ala Cys Ala Leu Tyr Ala Cys
180 185 190
Gly Ser Glu Ala Leu Val Glu Phe Gly Thr Tyr Leu Glu Asp Phe Leu
195 200 205
Gly Asn Glu Phe Trp Arg Asn Trp Asp Gly Arg Leu Ser Lys Tyr Asp
210 215 220
Ile Glu Thr His Arg Arg Cys Arg Gly Asn Gly Tyr Val Asp Leu Asp
225 230 235 240
 Ala Ser Val Met Gln Ser Asp Glu Tyr Val Leu Ser Gly Ala Tyr Asp
245 250 255
Val Val Lys Met Gln Pro Pro Gly Thr Phe Asp Ser Pro Arg Tyr Tyr
260 265 270
Leu His Leu Met Asp Gly Ile Tyr Val Asp Leu Ala Glu Val Thr Ala
275 280 285
Tyr Arg Ser Tyr Gly Met Val Ile Gly Phe Trp Thr Asp Ser Lys Ser
290 295 300
Pro Gln Leu Pro Thr Asp Phe Thr Arg Phe Asn Arg His Asn Cys Pro
305 310 315 320
**SEQ ID NO**: 25  
**LENGTH**: 129  
**TYPE**: PRT  
**ORGANISM**: Bacteriophage f2

**SEQUENCE**: 25

```
Ala Ser Asn Phe Thr Glu Phe Val Leu Val Asn Asp Gly Gly Thr Gly
```

<210> SEQ ID NO 25  
<211> LENGTH: 129  
<212> TYPE: PRT  
<213> ORGANISM: Bacteriophage f2

**SEQUENCE**: 26

```
Met Ser Lys Thr Ile Val Leu Ser Val Gly Glu Ala Thr Arg Thr Leu
```

<210> SEQ ID NO 26  
<211> LENGTH: 128  
<212> TYPE: PRT  
<213> ORGANISM: Bacteriophage f2

**SEQUENCE**: 26

```
Thr Glu Ile Gln Ser Thr Ala Asp Arg Glu Ile Phe Glu Glu Lys Val
```

<210> SEQ ID NO 27  
<211> LENGTH: 4586  
<212> TYPE: DNA  
<213> ORGANISM: artificial sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Cloning Vector - chemically synthesized
<400> SEQUENCE: 27

```
ttctgtttcc tgggtgaat tggtatcgc ccacattcgg acatattata acagctgcgt  
attaaattgc aacagcatt ttcagatat ttgccgaaac cgttattagc tgggccaaa  
acacatattc cagaacagc cgtagccttg aggcaagcag gattagcatt gattaccgc  
cgagcagccc atacagctgc ttcgctgagc cagagcatac ggtgccacgt  
gcattgtgata aggcccgtcc ctctccgccc ggaagcactg tggcggcact cagccgccc  
acaggtcggc tggctacgcc ctatatcgc gacataccgc atggggcagg tgggtgctgc  
cacttgccgt gcctgacgcc tggcagcggc gtcggcggc ccgtggcggc  
ggagcttggg cgcagcttcg ccctctctgg cggcgcgctg cgtcgggagc  
cacacttcg cacttgccgt ctctctctgc acatcgcgac ataaagggg ggtcgagcgc  
ccctggctga ttttgggcgg caaatgctgc ctgtggtggc gcgatgcgt gcggccggc  
ttcggtgttct cggagtccgc gcaacccgct tggccggttg  
acacgacttc cgtgatcgc gacgtccgct ctcggtgctg tcgtgccggtc  
gcaatgttgc acggttggtc gcggagaagg gcgtgggagc gtagctgcg tcgtggcgc  
tgttgctggt cgggtgggag gttggttagc atggcctgc ttcgctgcgc ccctcaaagc  
tttttgcg  
gtccctgctgg cacttcagag tctgtctcgc gctgtctgctg  
catcctagc gacactgcgc  
cggggccct ctcggtgc gggcggcgcc cggcgtcctgc gcgtggtcgc ggtctgttgc  
gttgctggt ccgggccccg gggcgtggct gcgtggtggt ccggggccgcct  
gttggcgttgc gcgtggtggt ccggggccgcct  
gttggcgttgc gcgtggtggt ccggggccgcct  
gttggcgttgc gcgtggtggt ccggggccgcct  
```
-continued

tacgccatac ggcgtctttc cgtacctctg ctcactgaact cgctggtacg ggtggttccgg 2280
cgtgcggcgc ccgtataagc tcaactaaag gcctaaatcgc gtttatccagc agaatcagg 2340
gataaagcgcc aaagaacact tgtgacaaaaa gcgccagc4aa aggccaggaag cggaaaaaag 2400
ggcgcggctgc ttgcttcttg ccctatgtcgc cgcccccttct gcggcgctac gaaaaaactga 2460
gctgtaagcg cagagggtcgc aaaccgagca ggaacttatgc gataccaggg gtctccccctt 2520
ggtagcctcc tccgtgcttc tccgtgctcc acacctgcgc gctacggata ctgcctgccc 2580
	ttttctccc gggaagctgt ggcgtctttc ctagctcaac gtcttraggtg tcttctggtg 2640
gtgcgttccg ctctgctcaca gcgggtgtgt tgtgacggac ccccgctctccg gcggcgctccg 2700
tgcggcttcat cggattctgc ccctacgag taccctgcg gcgggcttcag ttagctggca 2760
tcggcgacg ccacagcaaa cagggattaca agagcgaggt aatagggcgc tgtcigccag 2820
tcttgcaagt gggctgccat tcaagcgtaa actagaagga cgttaattgg tattgctgct 2880
tcggctgacg catttacctt cggaaagaagag gtttctatag ctgtgctaggg ccacaaaaaacc 2940
acccctgcgta gggctgttct ttttttctgg gccctagccgga tataccgcat cccaaaaaggg 3000
	ttctcaagac atctctctgt cttttttctgg gggctctag ctcagctgga cggaaaaactca 3060
cgtagcttagg tttttntcttag cagattatata aaaggactatctct ctaacgtagg ccccttttaaat 3120
taaaattgaa gtttttaactta aatcataatg atatagtggta aaacttggctc tcagcgttcc 3180
cactgtaata tccgtgtgcct acctatccca gcctactgtcg ttaccctgctat tccctgactg 3240
gctgtccttc cccttctctgt gataatcctg ataaggcagg gtttccatoo tggccagctg 3300
gctgcttgctg caccctgcgga caccagccatc ccgcgtttagg attacccagc aataacccg 3360
ccacgctggaa gcggccgactgc cagutaagta cctggaaactt tcaagcgctt catactgcct gcctgctct 3420
ataaattgtg gcggccgactgc tagattatac gattgccagt tttaattgtt ggcctgactg 3480
gttgccttgg ctctacctgct gtccgtctgc cggagtctgtg tgggtatggcc ttcacctcctg 3540
tccctgtctgc ccacgataag gggatccatacg tgcctccccag tggtagccgaa aaaagcggctg 3600
agctctccgc gtcagtctcc gataactgtgta agataaggtgc cgcagctgctt actaatcagctg 3660
gttattcgcc cactctggctt ccctcttctct tctctccagt caggtgcataa cggtaattgttg 3720
aatgtaagttg atccataacca gttcattctga cagttgacta tcggccgactg gcctgcttctg 3780
tgcggctctgg ccacaaactgc taataaggctc cccactcgca gaaacctataat aatgcgtctcag 3840
atggaacac tccctctctg gggaaaaac tcaaggtgctt tcaaacggctt agatcctggg 3900
tgggtgcctgc ccacgctgcctg cttccagctc ctctctctctc ccaagactgcgtg 3960
tcgggagagcg cttccttctgg gcctcagata gctggctata aacgtaataa aagggtaaag 4020
aatagttgaa gcggctcagct tggctagtctgc ctctctctctc ctaacgctactgc gatgtctttatg 4080

tgctctcagc gcgggatact atttgatagt attgaatat gcctacaaat aaatgttgttc 4140

ggcgtcgatgc gggcgaacat ggttctgcttg cttcctctgc aatctccttc acctgggctgc 4200
tgcggctgatt ccgtctctcg gcag-aggccagc ccaacgataac ccaacagactgc 4260
cagtctcttc gcagctgttct ctctctctctg cggagagctg ccaggccagctc cccagactgac acaaaagggc 4320
ccactcttcc gcagcctctt ttcgctctttc ggatgcttgt tcaaggtcag 4380
ggggcagccc caccacatac cggctgctctgc ggtctccccag ttcctgcttc gggatgctttg 4440

cagtctggtgac caccgctgct gctgcctgag cgaattcttgc ttctttctcag ggctgtctggc 4500
gttcgatcct aacgcatgac agcagcttca cctccggcgg cagagccttg 4560

gctcagcgtc gacggtgcct cggga 4586

<210> SEQ ID NO 28
<211> LENGTH: 425
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Terminator Sequence - chemically synthesized

<400> SEQUENCE: 28

cgttttggc ggatgagaga agattttcag cctgatcag attaattcag aacgcagaa 60
cgtctgata aacaaagatt atgggtgtcc ggtgcaggtg tggggtcacc cttgacccat 120
gcgcaacct cca cctgctagca cctggtgagt tggggtcct cccatcgcag 180
agtggagac gctgcaagct ccataatatt cgaagcctca gtgaaggagc tgggctttc 240
gttttacttg tgggttgtcg tggagactc tcctgtaag gacaaatctcg cgggagccg 300
atttgacgtc tggcagacac cggcggcaggg gtaggggagg cccgcaacgt 360
cagcgtcatc aataaggcgc aagcgcaccc tcacggtgac ctttcctgcag 420
cctt 425

<210> SEQ ID NO 29
<211> LENGTH: 816
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Resistance gene - chemically synthesized

<400> SEQUENCE: 29

ttagaaaaac tcatcgcagca tcaatgtgaa cttgcaattta ttcatatcag gattatcaat 60
aacatatattt tgaagagggc gtatcttgtaa tgaagggagaa aatooacgga gggagtctca 120
tagagcgtacg gatctctttg atcggcgtgc gattgacgat cttgcaacact caatcacaac 180
ttataattt ccctcttcaaa aataaggatt atcaagtaag aatooacgccg tagtaagcag 240
tgaacctgtg gctttgcgaa aagctttgtc ctttcccctg cgaactttgc caacaggcga 300
gcattacgc tcctgcttcaaa aatooacgct atcaaccacaa ccgtttactca ttcggtgattg 360
cgcttgactc gagaagaaat cgcagacgcgt gtaaaaaggc caatttaaacc cgggaactca 420
atgcacccgg cgcagcagcct ctcgacgacgc atcaaccacaatt ttttaccttg aatgggata 480
ttttttaaat atccttggagt cttgttcccgg gaggattcga gttgtaggta accatgatcc 540
ataagggatt ggatgtgtgt gcagggaggg ataaattcgg tcagcaggt 600
tagtagtacgt atccatcgct taaatacatt cgggacgctct cttttgctac gttgcaaaa 660
cacatcggcc gcatcggtct tccccataaa tctgtaagatt gttgagccttg attttcccag 720
attgtcggga gcccacatat taccatatctct taccacattt tttttcagtt ttatcgcgg 780
cctcgcagca gacgtttccc gttgaaattg gctcat 816
-continued

<400> SEQUENCE: 30

ggcgccgagt gcctactatac acgggataat tctggctgcct caagggcgcac tccgggtcttg

60
gcatttctt tggggcgag cagcaataac gcccggcggag tctggcgtgca atatggtggtat

gacctttgc tggagtttgc gtgagtttgc ggcggactaa atctgcaat

180
ggcggccgg agctcgggat gcgggttgggg ctggctggttag gcggggctggt

240
catacgggct ctaaagtcgc ggacgagcgtg taggggtgccct tctggcgtgca

gcagcgcct ctgggtgtct cgtggcggtt cggggtttgt ggcggctggtgg

360
aggatggtc ctgggggttc tctggcgtgca cgggggtttgt ggcggctggtgg

420
tggggcggca gacatcggg caacgtcttt gcggggtttgt ggcggctggtgg

480
tggcgggg cggggtttgt ggcggctggtgg

540
tggcgggg cggggtttgt ggcggctggtgg

600
tggggtttg tggggcgtg cgggggggttt ggcggctggtgg

660
tggggtttg tggggcgtg cgggggggttt ggcggctggtgg

720
tggggcggc cggggtttgt ggcggctggtgg

780
tggggcggc cggggtttgt ggcggctggtgg

840
tggggcggc cggggtttgt ggcggctggtgg

900
tggggcggc cggggtttgt ggcggctggtgg

960
tggggcggc cggggtttgt ggcggctggtgg

1020
tggggcggc cggggtttgt ggcggctggtgg

1080
tggggcggc cggggtttgt ggcggctggtgg

1140
tggggcggc cggggtttgt ggcggctggtgg

1200
tggggcggc cggggtttgt ggcggctggtgg

1260
tggggcggc cggggtttgt ggcggctggtgg

1320
tggggcggc cggggtttgt ggcggctggtgg

1380
tggggcggc cggggtttgt ggcggctggtgg

1440
tggggcggc cggggtttgt ggcggctggtgg

1500
tggggcggc cggggtttgt ggcggctggtgg

1560
tggggcggc cggggtttgt ggcggctggtgg

1620
tggggcggc cggggtttgt ggcggctggtgg

1680
tggggcggc cggggtttgt ggcggctggtgg

1740
tggggcggc cggggtttgt ggcggctggtgg

1800
tggggcggc cggggtttgt ggcggctggtgg

1860
tggggcggc cggggtttgt ggcggctggtgg

1920
tggggcggc cggggtttgt ggcggctggtgg

1980
tggggcggc cggggtttgt ggcggctggtgg

2040
tggggcggc cggggtttgt ggcggctggtgg

2100
tggggcggc cggggtttgt ggcggctggtgg

2160
tggggcggc cggggtttgt ggcggctggtgg

2220
tggttgcccgt atcaagagct acctactcct tttcgaaggg tacactggttc cagcagagcc 2280
cagatataac atactgctct tctactgttag cggtagctag gcacacacct caagactctt 2340
gttgacacgt ctatacaact gctgtcggta atccgtgtac cagtggctgc tgcagcggtc 2400
gttcagtctgt ccctcagcgg gttggactac aagagactgt taccagataac ggctcgcccg 2460
tcgggtgtgaa gggggggtgct gtagacacag cccagcttgg agogaagcgac ctacaggca 2520
tgtagaatac taccagcctga gctatagaga aagccacacc gttccgaaggg gagaagggcg 2580
gaccaggtac eggtagccag cagggtcggga acagagagcc gcagcggagga gtcoccaaggg 2640
ggaaagctct ggtactctttata gtactcgttgc gggttccgct acctctgtac tggagtgtacga 2700
ctgctctgtgat gctgtcagcc gggggggtgct gtagacacag cccagcttgg agogaagcgac 2760
ttcagcttcc tgcctgttctgtgttcgctt gttcacaatgg ttttctctgc gttatcccct 2820
gattttgtag atacgcatgt taacgccttt gctgtagctg atacgctcgt cggacagccgc 2880
agcagcgacc ggcagcagtc agrgagcgcag gaagcgaggg gacgcttgtag ggggtaatttt 2940
cctcttacag atcgtgctggttttttact gcatagtagc gacacacttc ctaaactatgc 3000
tctgatgccc ctatgtaaga ccagttata ctcgcgtatc gtaagctgac tggtcagtgg 3060
catcggccgg caacccgacca aacccgctgct agccgtcctgct ggtccctggcctggcttccc 3120
catcggccgg caacccgacca aacccgctgct agccgtcctgct ggtccctggcctggcttccc 3180
cytcataacg gaaaccgacta cggcagctgc gagaacgctt catcagctgtg tgcagcggtc 3240
attaacagat gtagccgctt tctcctgcttg ccagctgtag ggtgcctctgc agaagctgta 3300
atgtgtcgtg ttcgctaaga cggccgctttg ttttctttgtggcagcggtc 3360
atgtgtcgtg ttcgtaaga gttgcggtcat ttttgcggtag gttgcggtcat 3420
agtcctgcct gtagagggga ttttgcggtc ttttgcggtag gttgcggtcat 3480
gttgagtcct cagttataac atcagctgtg actcagctgtg aaacctcctg gttgtcagcggtc 3540
agcctggcct aactgcttgc ctgctttatcc ccagccttgtg ccagccacct atcagctgtg 3600
agtcctgctt gtcgtagacat ggtcagcgtt gttgcagctgc ccagctgtag gttgcggtcat 3660
agcctggcct aactgcttgc ctgctttatcc ccagccttgtg ccagccacct atcagctgtg 3720
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 3780
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 3840
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 3900
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 3960
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 4020
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 4080
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 4140
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 4200
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 4260
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 4320
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 4380
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 4440
tcagctgctg tctcagctgtg ggtgtgactac tctgtaaagcc cccgggcaacg 4500
...-continued

cggcactgt cttacagtt gcatgataaa gaagacagtc ataagtgcgg cgaagatagt 4560
cagcggccgg gcccacccga aagagctgac tgggttgaaag gctotcaaggg gcatcggctg 4620
aagcetccct ttatgcaat gctgcccattg ggaagaagcc aagtataggt tgaagcggtt 4680
ggacgccggc ggccagagta atgggtctgtg caagggagtg gcggcccaaca gctcccggcc 4740
cacgagggtct gcacacatc ccagacgagaa acaaggcgtc atgagccccg aagcggcagc 4800
cgatctctcc ccatcyggta ttgtgcccatt atagggcaca gcaaccgcac ctgcggtgcc 4860
ggtgtagtcgg gcacacagtc gtccaggtga gaggatcgg gttatctgcg tgcacggtgc 4920
accgaatgcgt ctggctcgctg ccagacatcg ggacgctggtg tact 4963

<210> SEQ ID NO 31
<211> LENGTH: 6
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Stop Codon - chemically synthesized

<400> SEQUENCE: 31
tgaaca

<210> SEQ ID NO 32
<211> LENGTH: 6
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Stop Codon - chemically synthesized

<400> SEQUENCE: 32
taatga

<210> SEQ ID NO 33
<211> LENGTH: 4625
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Plasmid - chemically synthesized

<400> SEQUENCE: 33
tcgccggttt cgggtagac ggtgaacacc tctgacact gcagctccgc gagaagttca 60
cagcgtgctt gtctaggggt gcgggaagca gcacagcccg tcagggcgct tcaggggtglt 120
cttgccgggtt cgggcttcgg cttaactatg cggccatcaga gcaagttgta ctgagatgc 180
acccatcgct ggttgaaata cgccacagat gcgtagggag aaaaatccgc atccaggccg 240
attccgcatt cagctgccgc aacctgtgag ggggccccgt cggcggggcc tctccgtat 300
tacgcagct gcagccggag ggtatgtgtgc caagggatt aagttggctga aagccagggt 360
tttccccgtcc acaaggctgt aacaagcagg ccaaggtgatt cagagttgat tttatctctt 420
aggcataag cagcaggaac gactatcagct gttgtaatag gcacattattg cttataacac 480
cttggtagct cagctgcttt ggcaccacca ctgctgattg ggcgcacaaa ggcaactac 540
acgaccctgg cgtagctggc ccagatcagc ttaagagcgc atataagctat ttagactta 600
agataaatga atcgcatcat gttacccaga gccaagggt cttggccctgc ggaatattca 660
tcacttcgt cagcagcagtc ctgctagta gcaagataag ccctccatct gcagcctataa 720
ccatggccac gcgaagccgc gagaaggtgt taactaagcc gagaacccca atcagcgcac 780
tttgatattg ccaaaagatct ttggagggca acaataaatg cggaggcttg gatgccaaca 840
gcattataga tagccgacgct cgtgctctta gtagagggct ccgaaacccc ccggccctaa 900
caatattcga aggacagact ataacttcatt gtcctctggee gggatctaat cggggtgatcc 960
ggataacccg gactgggttt tggacgctcg gggccacccc ccggaaaccc cgagcagcgat tcctaatatt 1020
gctggttccagcgatcgatag ccggagagct agcagcagct ccggagagag 1090
tctgtaacga ccaagactcg tctctactcg gataataact gctgaagagaa aaagggactc 1140
tcgccatagc cctggcgagct aacgagttgc ccaaaagaga cagtctgagct gaaagcttgc 1200
gggtccttga caacttcctgt gacttccgag ttcctagcgat cggagcagcagct cggagcagcag 1260
gaaagcttgg ccgcttttgag cggagcagcagct ccggagagag 1320
cgctgtattgc gatctacccg aacgagttgc cggagcagcagct ccggagagag 1380
gtgaacatc aagtaacagtc cctaaccttt gctacgtttt ttaccttcccttt 1440
tatatacttc aacgagttgc ccaagagagct ccggagagag 1500
ttaaatacttc aacgagttgc ccaagagagct ccggagagag 1560
cattatgaa cccgctggggc ggcgggcgc cggagcagcagct ccggagagag 1620
tctctctgcc tcatacgggt cggagcagcagct ccggagagag 1680
taaaggtcgt tttcgggtgg gacttccctgc gcggagcagcagct ccggagagag 1740
gcggagcagcagct ccggagagag 1800
agcttggctgg gcggcgtgctgg gcggagcagcagct ccggagagag 1860
cgcttctgcttg ccgctttttg gcggagcagcagct ccggagagag 1920
gtggagcagcagct ccggagagag 1980
tctctctgcc ttcgagagcttg gcggagcagcagct ccggagagag 2040
gggtgttggc aacgccctgg ccggagagag 2100
tggagagcttg gcggagcagcagct ccggagagag 2160
attagaggg cggagtgtgg ccggagagag 2220
ggctctcttcc gcggagcagcagct ccggagagag 2280
aaaagatgttcgccggcgc ccaacccaag cgggagtgttg ccggagagag 2340
gtggtgctg gcggagcagcagct ccggagagag 2400
tctctctgcc gcggagcagcagct ccggagagag 2460
tattcttgtctgg gcggagcagcagct ccggagagag 2520
ccggagcagcagct ccggagagag 2580
atttctgctg ccggagcagcagct ccggagagag 2640
acctactacggcc cccggagcagcagct ccggagagag 2700
cgctctcttcc gcggagcagcagct ccggagagag 2760
gcttggctgg gcggagcagcagct ccggagagag 2820
gttcagcagcagct ccggagagag 2880
gttcagcagcagct ccggagagag 2940
ggttcagcagcagct ccggagagag 3000
ggttcagcagcagct ccggagagag 3060
cttaoctgtaa tcgcatactgtaa aagatgcatt tctgtgactg tctagtgaggg 3120
cggagtctg ccctggaag aaggtggtgc tgactctac ccggctcgtc ggtccccatc 3180
atccagccag aagttgaggg aagcoaggttt gatggagacct ttggtgttagg tgggaccgatt 3240
gggtcttttt aacctttttct ttgctcagga aacgcttctg tctgctgggaat gattgcgtgt 3300
cgatcctcct acgcagcaa aagtctgccatt tattctacac aagccgcgctc gcttcaagtc 3360
agcgttaatgc tcggcggatg tttaacacaa ttaaccacttt tagattgaga aaactcatcg 3420
agcatcaaat gasaactgcaaa ttttctcata tcggagttatc ctaatacataa tttttggaaa 3480
agcgcgttctt gltattgaggg aggaaaactcct ccggctcagct tcctagctgg ggaggactcc 3540
tcgtaagct ccggcttctc gagctctgaoa actataatac aaactattaa tttttctcct 3600
tccaaacattt ggtctacacag tggaaagagc cacataaggta cagactgtaat ccggctgagaat 3660
ggcgaaaacct tattgtacttc ttctcctagct ttctcctagc gcacagcatt aagctatgcct 3720
tcacaatcac ggcagcttc gaacacgtaa ttcattctgttt attagcctgc aggacagcga 3780
aattgaggttg gcgggctga aagagaaatta cctggatcag gattttcttc tattacgcttg 3840
aacacggtca gcgcagcataa aatattttta cctggatcag gattttcttc tattacgcttg 3900
aatggtgttt tcgggctggat cgacaggtttg atagcactag cattatactg agtaagggta 3960
aalagctcgg ggtctccagag aggataata ctaacgctagc acgtttgacct cccacatcag 4020
tcggactcag ctcgctctg ggcgcgtcctg cccgtttcgt cacacacgct ggacagcctg 4080
cggccctcct gattcttggct ccagggctcct ctcagcccag tggccgctg 4140
ctttccctcctt catatctaccttc aatcagagc tctctacagc gcgcagcttg 4200
tccggcttct aatttgaatc tccagtctttt ggttaactgttt tattgtaagc agaaggttttt 4260
attgctcaag atatattcttt atattctttag ctaatgggca actagagagct ctagagcctct 4320
aagctttgct ctctcccccct cctattatga aagatattag cagattttg agtctagcag 4380
gcgaatactac ttaggtattt ggttaaagaat agaattcctt ggtttggcag cccattttcct 4440
cgaaaaagtt cacttgaagt ctaagagcct attatatctc tggacagattttcct 4500
agcggtatca cggccggctg ttgctc 4525

<210> SEQ ID NO 34
<211> LENGTH: 56
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Sequence - chemically synthesised
<400> SEQUENCE: 34

gccgcgcaat tcggcaggtta aaaaaacagtg gcaaagattag agacagttac tttagg 56

<210> SEQ ID NO 35
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Sequence - chemically synthesised
<400> SEQUENCE: 35

gcatgcaagc ttagcatgc aatttcactct teg 33
-continued

<210> SEQ ID NO 36
<211> LENGTH: 57
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Sequence - chemically synthesized

<400> SEQUENCE: 36

gcgcgcgaa ttcaggaggg aaaaaaaat gagcaaaatta gacactgtaa cttagg  57

<210> SEQ ID NO 37
<211> LENGTH: 3914
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Cloning Vector - chemically synthesized

<400> SEQUENCE: 37

tgcgcgctgttt cgggtcagac ggtgcacacct gcagctcccg gacacggctca  60
cagcccgtct gtcagcggc gctgccccag ccaacgcgct tcagcggctgct  120
ttggcggcgt ttgctgctgg cttactactc gcgctgcagc gcaagctggct ctgagctgctc  180
accatatgc ggtcagacat cgcagccagc cggaagggag cactgacagc  240
attctgcaat cagccgcctg cgggcttgcc aagggctgac gcgcggccgcgtgct  300
tacgccgtgc gcgcccaaggg gggatcagctg caaggtggtgtg gacacgggtg  360
atctccatgc cgcagcctgct taaacagcagc cccgagctgc gctgagctgctc  420
agggaggggag gcgcgtggtc gtcgcagctgc tcagaaaggt cgctgctgact cataccaggg  480
tctagctgcc ccatcatccgc gcggagaaggt gcgggagcgc gcgttagctg gatgttgaggt  540
ctaggtgac caggtgctga ttttgaacct ttgctctgggc acgcaaggtc ctgctgctgc  600
gggagggag gcgcgtggtc gtcgcagctgc cctctcagct gcacaaaggt gcgttatcagc aaccaaacgg  660
cgtcagctgc aagtcgtcagc ggtgcttata cagaattaac cagcctgact  720
tgcagagact ccatagcacagt gatttattt gtagatttaa tatacattaga  780
tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
ttgtaaatac tggccatagt gcgtttcctgt gttgaattgt tattccgctca caattcaca 1740
cacatcgy gcgggaagca taaagttgaa agcctggtggt gcctaatagag tgaactaact 1800
cacatttaatt ggttctgctg caactgcgcgg tttcccagtg gggaaacctgt ggtggcagctg 1860
gcattatgta atccgtagaac gcgaggagggag aggcggttgct cgatttgagc gctctcggc 1920
tctgctgcto atcgactgct gcggctggtg cgggagaggg cggggaccga ctaggtgta 1980
cctaaagggc gtatactcgt tagctcagcata atccagggagt aacgcagggag aagccagctg 2040
agcggaaaaa cagccaaaggg cagggaaaccc taaaagagggcg cgttctgctg cgtttttcaca 2100	
taggtctcgcc cccccctgcag aacctgcaca ataaacctgc tcaaagtgctg gttggctgaga 2160
cgcacaggg tataaaggtt accagacctt tccccctgag aagctccctc tgcggtgcc 2220
ttcgcctacc ttgcggtgcat gctttccgctt gttggtggtc tggcagcttctc ttcggtgtgg 2280
gttttttaat gcgctgcttg ctattttgctg tagctttccgttc gttttccaaag 2340
ggggctgttgc cagcagccccc cctgggtggtc cgagcctggc gctttctctc tctccctc 2400
ttcggcgtct cagccgcatgt gactgcctgtg taaactacgtg 2460
gattacgga gcagaggtgctt taggctgtgtg tctagaggtct ttggaagttt ggccttaacta 2520
cggctacact aggaagacagc tattttgtat ctgctgtcttg cgtggtcagcg ttacctctgg 2580
aaaaagagtg ggttagctcttg gatccgggaa acaacacaac gcgtttggctg cggggtttttt 2640
tgttgcagcg cagcagccga aacagatacct caacagagtc tttttctgtct 2700
tttcaggggg gtcgggcggc aagctcggag ttaggggatttt cttgctctag 2760
atttttaaaa aggttctcatt ccctactcct ttttattttta aataggtttt ttaaaactaat 2820
cattaatata tagatatatttt caatgtctag caagttacctca tggagcagcc 2880
tacccacgcat atcagctcat ttgctttcat atagcgtgcct gactctcctca cttccggt 2940
aactaogata cggggaggtt cactacgctgg cccacagtct gttaagttgg cagcgaacta 3000
ccgctacgcc cgctgctgactaat gcagacgcg ggcggggagt gggagccgagc 3060
aagttgctct gcacaatttt ccggctcctag cactctttctt aatgtgtgcccc gggaagatag 3120
agttaagttc gctgcagcaatatagttgctc caaacgtgttt gcacacgctc caagcagctg 3180
ggtgctcaag gcctgttttg gtttgtccat ttactggtcct gcgttcacaa gataagggcg 3240
agttactagt cccctccttg tggcgaaaaa aagggctacc aactctttgg tcttggtgctt 3300
tgctgagaagt ggtgtgcgtc cagttgtcct atccagttct atggcagcct tggcatactc 3360
tcttaacctgc atcgctgctct tttggtctgt gcgtgactag ctaaccgaact 3420
attctggagaa taggtgactg gcggacgagg tttgtctgctg cgggctgcaaa taacaggataa 3480
tacgccggtc ctaagggaaa ctttaaaggg gcgaacaaaat cgcgtgcac 3540
aatacctca aggattttcc gcctgtggag atcgctgctgc atgtaaccc cttcgtcgc 3600
cacattgata tcagctatttt caaagttgctt gggggagaaa aacatggag 3660
gcmaaagtcc gcmaaagaag gaatgaaggg gcacagggaaa tgggtgactt ctatctcttc 3720
ccttttccaa tatttgacg gacatttca gcgctttgt gcataactgct 3780
tgaattttg ataagaaaata aucacatagc ggtttgccag atcttttcacc gcgaagttgc 3840
actgcaacgtc tagaaaccact tatttcatc gacattacaa ttaaaccata gggggtttc 3900
gggcccttc gctc 3914
<210> SEQ ID NO: 38
<211> LENGTH: 57
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer sequence - chemically synthesized
<400> SEQUENCE: 38

ggcgcgcag tctaaggagga aaaaaaaaaa ggcacaatac gcatacgac ccgac 57

<210> SEQ ID NO: 39
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Sequence - chemically synthesized
<400> SEQUENCE: 39

gcagtgaac gtcataacgc aagrtaacgc gacgtaacg tagc 44

<210> SEQ ID NO: 40
<211> LENGTH: 4962
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Expression Construct - chemically synthesized
<400> SEQUENCE: 40

ggcgcgcag gtcgaactac atcgcataac tcctgctact cagaccgcac tcgggttctgg 60
gatatttatt gttgctggac gccataacgc gtatctctact gatctactcgtt 120
gacaattact cagcgcttcg tatataatctg aatattgtta ggcgtttaac gagttactca aatcttactca 180
ggaaacagca tccggggcgt tattaattgcat ggctataactgtat aactctatctttcttgctggctgctttataaaccat cagagtttat caacggtcttgctttcttc cctttcctttactcttc ttcctttccttctctcttctctcttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttc
tggggcaatac aggtgcaacga atctatcgat tggatgggaa gcggcatgcg ccagagttgtg 1320
cttgaaacaa tcggcaaggag acgcgtgcca agtatgtatg agatgagatgt tgtaggctaa 1380
aetgtggcat gccagattatg cctcttcgga ccataaagca ttttcctcggt aetatggatg 1440
atcgatggt ggtaacctca cgcgtcccccgg gaaaaaacgc atccagagta ttgagaatt 1500
atcctgtaac aggtgaaatat agttgatgtg ogcgtggcagt ggctctggcgc cgggtcactt 1560
cyattctggtttgtatctgt ctttttaaca gcgggtggtg atttcgctct gctctagggc 1620
aatcagcatgc gataaagggat tcggggatgtg ctagtgatc cttgtacagag ctgtaatgct 1680
ggctctgtgga acagcttcgg aagaaatagc atatccctg gccatctcctt gcggagtccg 1740
tgctcaacca tggaggatat tcgcttgata acccttattt tcagaggggg aaaaatattg 1800
gttggattgag tcgggaagcc gcgagatagc cagacgaca cagagatcttc gcgatcttcat 1860
ggaaagtcgct gggtggattt ttcgctcttg acaagaaacc gttcttttcgaa aataatgga 1920
tgataaacctgc gataagcctc tcgtctatgt gcctgagcag ttttttctaaa 1980
cgcgtgcaacc gttctaccctc atgattgatc ttaaacaatct ttttaatatta 2040
aaaggtcctca gcgcaagactt ctttttctaat atctctagac caaaatcctc taagctgatt 2100
tttctgctttg tagctgctgca gacccctgtag aaagacataa aaggctctttc tcagatctt 2160
ttttttctcg cgtctatcgc tcgctggaaaa ccaaaaaaac acggctaccc gcggggggttt 2220
tggtgcgga ctgagggctc ccaaccttttc ttcgagaagtt gcactggcttc agcagagcgc 2280
agataacaaacct ttcgcgtcctc ctatggtagc cggctgtgag ctcaccaaccta aagagacctg 2340
tagcagcctc tccatctccct gctgtctatc gcttggatgtgc gggagtggcgc 2400
ataagctgtag tttgtaagggg ttggctctccaa gcgcggaatt acctgataag gcggagcggcg 2460
cgggtgcagggggctggcg tgcagacaggc cgcctgtagga gcggagcacc ttcaccgaaac 2520
tgagatacct cagcgtgctg gcttgagaaa cgccagctgt tcgcagaggg agaaaagggg 2580
acaggttacc ggtgaagggc aggggtccgaag cgggagacgg cacgaggggt gccagccaggg 2640
gaaagcgctgt gttttctttct gtttggccctt tcctgctgtt gcctgggtgt 2700
aaggctctgctttgcagggcag aacgggggctgtagttcct cggctttcctt 2760
tgcggtctct tcagcagggg ggggggagggg tttgaaacaaa gcggacgcaac gggtgcctttt 2820
atctgcgcgtct gcagcgagtct ccgtgtttgtccttc acacgatatg agtgcagttc tgcagcagc 2880
cggacgagcc cagcgcgtcata gcggagcagc aacgcgactag cggcgttgcat 2940
tccagatcagc tcggcctttc gcctgatgtg cctctctcagc acaattcgcgct 3000
tcgtgagcgcct tcaagccggt cagccctgactc gtggagtactc ggctctggggc 3060
tgagcgcgcctc cagcgcgtcctgc cggcgcggacg cgcgtcgcgtg gcgtgcgtggtc 3120
atcagcgttc atagccgcttgc gcggagcggc atgtgctcaag gtttcttgcct 3180
gtacccagcc aacagcgctg cggcgcctggc tgaagctcca tcaagctggtg ctgtgaagcga 3240
ttcctagctgc tcggcttgcgtt ccctgctgctg cggcggtctg aggcttttcca gaagctttaa 3300
tgcagctggct gcagatgtgctt tcagctggtaag gggtaaagtt tcttttctttc tgggaatctg 3360
tggcctgcttgc aagggggttctgctcttcag gtttttcaagc 3420
gatcgcctgct atgcgagggct gtaagatgtaga aagcgagagag 3480
taaacacactgg gcggatcgga tgcgggaggg cggagagaaata ctcattcaag gcgtagccc 3540
1. A process for expression of a recombinant capsid protein of a RNA bacteriophage being capable of forming a VLP by self-assembly, said process comprising the steps of:
   a.) introducing an expression plasmid into a bacterial host, wherein said expression plasmid comprises an expression construct, wherein said expression construct comprises (i) a first nucleotide sequence encoding said recombinant capsid protein, or mutant or fragment thereof, and (ii) a promoter being inducible by lactose;
   b.) cultivating said bacterial host in a medium comprising a major carbon source, wherein said cultivating is performed in batch culture and under conditions which said promoter is repressed by lactulose, wherein said lacI is overexpressed by said bacterial host;
   c.) feeding said batch culture with said major carbon source; and
   d.) inducing said promoter with an inducer, wherein said feeding of said batch culture with said major carbon source is continued.

2. (canceled)

3. The process of claim 1, wherein said bacteriophage is selected from the group consisting of:
   a.) bacteriophage QB;
   b.) bacteriophage AP205;
   c.) bacteriophage fr;
   d.) bacteriophage GA;
   e.) bacteriophage SP;
   f.) bacteriophage MS2;
   g.) bacteriophage M11;
   h.) bacteriophage MX1;
   i.) bacteriophage NL195;
   j.) bacteriophage f2;
   k.) bacteriophage PP7 and
   l.) bacteriophage R17.

4. The process of claim 1, wherein said RNA bacteriophage is QB.

5. The process of claim 1, wherein said recombinant capsid protein has the amino acid sequence of SEQ ID NO: 5.

6. The process of claim 1, wherein said expression construct comprises a first stop codon and a second stop codon, wherein said first stop codon is located directly 3' of said first stop codon.
nucleotide sequence and wherein said second stop codon is located directly 3' of said first stop codon, and wherein at least one of said first or second stop codon is TAA.

7. The process of claim 1, wherein said expression construct comprises a first nucleotide sequence and a second nucleotide sequence, wherein said first nucleotide sequence is encoding Qβ CP, or a mutant or fragment thereof; and wherein said second nucleotide sequence is encoding the Qβ A1 protein or a mutant or fragment thereof; and wherein said first and second nucleotide sequence are separated by exactly one sequence stretch comprising at least one TAA stop codon.

8. The process of claim 1, wherein said expression construct comprises or alternatively consists of the nucleotide sequence of SEQ ID NO:6.

9. The process of claim 1, wherein said expression plasmid comprises or preferably consists of the nucleotide sequence of SEQ ID NO:1.

10. (canceled)

11. The process of claim 1, wherein said promoter is selected from the group consisting of the

a.) tac promoter;

b.) trc promoter;

c.) lac promoter;

d.) lac promoter;

e.) lacUV5 promoter;

f.) P_{aro} promoter;

g.) lpp^+ promoter;

h.) lpp-lac promoter;

i.) T7-lac promoter;

j.) T3-lac promoter;

k.) T5-lac promoter; and

l.) a promoter having at least 50% sequence homology to SEQ ID NO:2.

12. The process of claim 1, wherein said promoter comprises the nucleotide sequence of SEQ ID NO:2.

13. The process of claim 1, wherein said major carbon source is glycerol.

14. The process of claim 1, wherein said feeding of said batch culture is performed with a flow rate, wherein said flow rate increases with an exponential coefficient μ and wherein preferably said exponential coefficient μ is below μ_{max}.

15. The process of claim 1, wherein said inducing of said promoter is performed by continuous feeding said batch culture with said inducer and said major carbon source at a constant flow rate.

16. The process of claim 1, wherein said inducing of said promoter is performed by continuous feeding said batch culture with said inducer and said major carbon source at an increasing flow rate.

17. The process of claim 15, wherein said inducer is lactose and wherein preferably said lactose and said major carbon source are co-fed to said batch culture in a ratio of about 2:1 to 1:4 (w/w).

18. The process of claim 1, wherein said inducer is IPTG and wherein preferably the concentration of said IPTG in said medium is 0.001 to 5 mM.

19. (canceled)

20. The process of claim 1, wherein lactose is over-expressed by said bacterial host, wherein said overexpression is caused by lacI or lacQ, preferably by lacI.

21. (canceled)

22. (canceled)

23. The process of claim 1, wherein said inducer is lactose and wherein said bacterial host comprises β-galactosidase activity.

24. The process of claim 1, wherein said cultivating and said feeding of said batch culture and said inducing of said promoter is performed at a temperature which is below the optimal growth temperature of said bacterial host.

25. The process of claim 9 wherein:

a.) said major carbon source is glycerol;

b.) said feeding of said batch culture is performed with a flow rate, wherein said flow rate increases with an exponential coefficient μ, and wherein preferably said exponential coefficient μ is below μ_{max};

c.) said inducer is lactose;

d.) and said lactose and said major carbon source are co-fed to said batch culture in a ratio of 2:1 to 1:4 (w/w), preferably 1:1 to 1:3 (w/w), most preferably 1:3 (w/w);

e.) said bacterial host is *E. coli* RB791; and

f.) said cultivating and feeding of said batch culture and said inducing of said promoter is performed at a temperature of about 30°C.

26. The process of claim 1 wherein:

a.) said expression plasmid comprises or preferably consists of the nucleotide sequence of SEQ ID NO:30;

b.) said major carbon source is glycerol;

c.) said feeding of said batch culture is performed with a flow rate, wherein said flow rate increases with an exponential coefficient μ, and wherein preferably said exponential coefficient μ is below μ_{max};

d.) said inducer is lactose;

e.) said lactose and said major carbon source are co-fed to the batch culture in a ratio of 2:1 to 1:4 (w/w), preferably 1:1 to 1:3 (w/w), most preferably 1:3 (w/w);

f.) said bacterial host is *E. coli* RB791; and

g.) said cultivating and feeding of said batch culture and said inducing of said promoter is performed at a temperature of about 30°C.

27. The process of claim 1, wherein throughout steps b.) to d.) of said process oxygen is supplied to said bacterial host by a pO2, in the medium of at least about 40%.

28. The process of claim 16, wherein said inducer is lactose and wherein preferably said lactose and said major carbon source are co-fed to said batch culture in a ratio of about 2:1 to 1:4 (w/w).

* * * * *