



(11)

EP 3 134 426 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
28.10.2020 Bulletin 2020/44

(21) Application number: **15738474.4**

(22) Date of filing: **21.04.2015**

(51) Int Cl.:
C07K 14/005 ^(2006.01) **C12N 15/09** ^(2006.01)

(86) International application number:
PCT/IB2015/052915

(87) International publication number:
WO 2015/162560 (29.10.2015 Gazette 2015/43)

(54) A METHOD OF OBTAINING A POLYEPITOPIC PROTEIN AS WELL AS A DNA VECTOR FOR EMBODYING THIS METHOD

VERFAHREN ZUR HERSTELLUNG POLYEPITOPISCHER PROTEINE UND DNA VEKTOREN ZUR DURCHFÜHRUNG DES VERFAHRENS

PROCEDE POUR PREPARER DES PROTEINES POLYEPITOPES ET DES VECTEURS ADNS POUR LEUR PRODUCTION

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **21.04.2014 PL 40795014**

(43) Date of publication of application:
01.03.2017 Bulletin 2017/09

(73) Proprietor: **Bioventures Institute Sp. Z O.O. 60-141 Poznan (PL)**

(72) Inventors:

- **SKOWRON, Piotr**
PL-80-303 Gdansk (PL)
- **ZYLICZ-STACHULA, Agnieszka**
PL-80-306 Gdansk (PL)
- **ZOLNIERKIEWICZ, Olga**
PL-80-306 Gdansk (PL)
- **SKOWRON, Malgorzata**
PL-81-601 Gdynia (PL)
- **JANUS, Lukasz**
PL-60-694 Poznan (PL)
- **JEZEWSKA-FRACKOWIAK, Joanna**
PL-80-331 Gdansk (PL)
- **KREFFT, Daria**
PL-84-240 Reda (PL)
- **NIDZWORSKI, Dawid**
PL-80-807 Gdansk (PL)
- **SZEMIAKO, Kasjan**
PL-80-287 Gdansk (PL)

- **MACIEJEWSKA, Natalia**
PL-13-206 Plosnica (PL)
- **NOWAK, Marta**
PL-84-240 Reda (PL)
- **SZYMANSKA, Aneta**
PL-80-283 Gdansk (PL)

(74) Representative: **Witek, Rafal**
WTS Patent Attorneys
Witek, Sniezko & Partners
ul. Weigla 12
53-114 Wroclaw (PL)

(56) References cited:

- **GUENTHER STEPHAN ET AL: "A novel method for efficient amplification of whole hepatitis B virus genomes permits rapid functional analysis and reveals deletion mutants in immunosuppressed patients", JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 69, no. 9, 1 January 1995 (1995-01-01), pages 5437-5444, XP002359956, ISSN: 0022-538X**
- **LEE J H ET AL: "Sequential amplification of cloned DNA as tandem multimers using class-IIS restriction enzymes", GENETIC ANALYSIS: BIOMOLECULAR ENGINEERING, ELSEVIER SCIENCE PUBLISHING, US, vol. 13, no. 6, 1 December 1996 (1996-12-01), pages 139-145, XP004017262, ISSN: 1050-3862, DOI: 10.1016/S1050-3862(96)00164-7**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 3 134 426 B1

- Eric R. Geertsma ET AL: "A Versatile and Efficient High-Throughput Cloning Tool for Structural Biology", Biochemistry, vol. 50, no. 15, 19 April 2011 (2011-04-19), pages 3272-3278, XP055080307, ISSN: 0006-2960, DOI: 10.1021/bi200178z

Description

[0001] The subject of the present invention is a method of obtaining a polyepitopic protein as well as a DNA vector for embodying this method. The proteins obtained according to the present invention protein may find a number of uses, and in particular for the production of improved vaccines.

[0002] Shen, S-H P.N.A.Sci USA 81:4627-4631 (1984) discloses a method of obtaining of a multimeric proinsulin (up to 7 copies). The multimers were obtained using the consecutive concatenation of DNA lengths using synthetic DNA fragments cut with the IIS endonuclease - SfaNI, alkaline phosphatase, polynucleotide kinase and DNA ligase. This method, however, did not make it possible to produce a multiply repeated DNA sequence in one reaction, because each time only a single additional monomer copy could be added.

Lennick et al. Gene 61:103-112 (1987) discloses multimeric gene encoding 8 copies of a peptide hormone - atrial natriuretic peptide. The method described does not use a vector that would make it possible to generate multiple copies of sequences encoding a peptide, and the resulting multimeric gene produced *in vitro* was finally cloned into a commonly used expression vector. This method also has not made it possible to easily multiply the DNA sequence in a single reaction, because each time a only single additional monomer copy could be added.

Kim, S. C. and Szybalski, W., Gene 71:1-8 (1988) discloses a method of directionally amplifying a cloned DNA fragment using a IIS restriction endonuclease - BspMI, the pSK3 DNA vector as well as DNA ligase. 30 monomer copies were obtained in the concatemer. However, the use of BspMI is made difficult in practice, because it does not digest the substrate DNA completely. In effect, it's not possible to maintain ORF continuity, and the amplification introduces additional DNA sequences into the amplified concatemer. The method described does not make it possible to repeat the amplification cycle which significantly limits the possibility of obtaining the desired number of repeating DNA segments. The vector disclosed was not an expression vector.

Lee, J. H. et al., Genetic Analysis: Biomolecular Engineering 13:139-145 (1996) discloses a method of directionally amplifying a cloned DNA fragment using the IIS restriction endonucleases BspMI and BbsI, the pBBS1 DNA vector as well as DNA ligase. This method, based on the autoligation of 4-nucleotide (nt) sticky DNA ends makes it impossible to maintain ORF continuity, since it made it possible to repeat the amplification cycle in order to achieve the desired number of copies of the amplified DNA segment. The vector used was not an expression vector. The method was used to amplify a short antibacterial peptide gene - mogainin (108 copies).

Lee, J. H. et al., Protein Expression and purification 12:53-60 (1998) discloses another variant of a method based on the previously described vector, pBBS1. The vector used was not an expression vector. The authors used the BbsI enzyme to generate 4-nt sticky ends, seriously impeding ORF continuity. The document describes the production of no more than 6 copies of the DNA monomer DNA in the amplified sequence. The method was used for amplifying the mogainin antibacterial peptide gene as well as bufferin II.

Wang, Y.-Q and Cai, J.Y. Appl Biochem Biotechnol. 141:203-13 (2007) discloses another variant of multimerizing genes encoding antibiotic peptides using the autoligation of synthetic DNA fragments, containing asymmetric sticky ends in the presence of 2 DNA adapters. The adapters contained sequences recognised by the restriction endonucleases Sail and EcoRI. The procedure used no vector, the amplification is hard to control and requires the addition of consecutive portions of the synthetic DNA monomer during the reaction. The disclosed method of polymeric gene construction, designed for the production of antibacterial peptides, resulted in 8 monomer copies in a polymeric protein.

Lee et al, Genetic Analysis: Biomolecular Engineering 13 (1996) 139-145, discloses sequential amplification of cloned DNA as tandem multimers using class-IIS restriction enzymes (see: Title, Abstract, and Figure 1). However, obtained DNA multimers are not encoding polyantigenic multimers in form of a single ORF, which would be expressed as a single polypeptide.

Greetsma and Dutzler, Biochemistry 2011, 3272-3278, discloses a cloning vector in which a DNA insert is introduced that encodes a polypeptide. The introduction of the insert involves digesting the DNA with SapI as a rare cutter (Abstract, page 3274, column 2, third paragraph when counting from the bottom, and Figure 1). Cited document also discloses that SapI generates small three base pair overhangs connecting the target ORF with expression elements of the vector which allows for generating fusions at both termini of the translated protein. However, it is silent about any concatenation of multiple DNA fragments in the vector.

Document entitled Archive Cloning published in 2014 on www.genome.wisc.edu discloses generation of expression vectors comprising ORF by using restriction endonuclease SapI to excise the ORF from archive plasmid and transfer it into an expression vector without a tag for expression of a full-length native proteins and with a 6xHIS tag at the amino terminus for expression of the fusion polypeptide therefrom. Nonetheless, it is also silent about generation of concatenated ORFs into a single ORF for the expression of polyepitopic protein.

[0003] The goal of the present invention is to deliver a method of easily obtaining a polyepitopic protein of arbitrary length, a vector useful in the embodiment of the method as well as of obtaining higher order polyepitope structures. The resulting polyepitopic proteins as well as higher order polyepitope structures are useful in a broad selection of uses, and in particular can be used to produce improved vaccines of increased efficacy.

[0004] The subject of the present invention is a method of obtaining polyepitopic protein, characterised in that:

- a) a blunt-ended DNA sequence encoding an epitope is cloned at a site recognised by SmaI into a DNA vector containing a sequence, the amplifying module, encompassing two convergent DNA sequences recognised by SmaI and a DNA sequence found between them containing a site for cloning in an insert recognised by SmaI, wherein the amplifying module contains a nucleotide sequence selected from: SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6, preferably the amplifying module possesses the sequence: GCTCTTCACCCGGGCCAGAGAAGAGC (SEQ ID NO: 11),
- b) the resulting vector is amplified in a bacterial host, isolated and digested with SmaI, and then the fragment containing the DNA sequence encoding the epitope is isolated, which is modified such that it becomes equipped with single-stranded sticky ends facilitating directional ligation of the insert to the concatemer,
- c) the isolated fragment is autoligated,
- d) the autoligation product is cloned at a site recognised by SmaI into said DNA vector as defined in point a),
- e) the resulting vector is used to transform a bacterial host and then the polyepitopic protein is expressed and isolated,

wherein in order to increase the size of the polyepitopic protein, stages from b) to d) are repeated prior to realising stage e), a concatenated DNA sequences encoding the epitopes are present in the resulting vector.

[0005] Preferably, the epitope is the HBV epitope, preferably encoded by the synthetic SEQ ID NO: 9. Preferably, the amplified monomer fragment contains various epitopes, stemming from different proteins or different regions of the same protein, preferably encoded by a synthetic sequence. Preferably, the polyepitopic protein additionally contains a sequence facilitating its purification, particularly a fusion tag containing a sequence of 6 histidine residues, and the isolation conducted in stage e) encompasses the use of metalloaffinity chromatography, preferably on immobilised nickel residues, preferably an additional purification stage selected from a group encompassing: heating the protein sample, fractionation using polyethylenimine, salting out with ammonium sulphate and molecular sieving gel chromatography. Preferably, the stage e) polyepitopic protein is immobilised on a macromolecular carrier, particularly one selected from among: microorganisms, cells, bacteria, bacteriophages, viruses, defective virions, autoaggregating proteins, or nanoparticles, preferably from a Phage Display system in the T7 bacteriophage.

[0006] The subject of the present invention is also a DNA vector for use in the method according to the invention containing the sequence of the amplifying module encompassing two convergent DNA sequences recognised by the SmaI endonuclease and the intervening DNA sequence found between them containing a site for cloning in an insert recognised by the endonuclease SmaI, wherein the amplifying module contains a nucleotide sequence selected from: SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6, preferably the amplifying module possesses the sequence: GCTCTTCACCCGGGCCAGAGAAGAGC (SEQ ID NO: 11).

[0007] Preferably, a DNA vector according to the present invention is a protein expression vector, which additionally contains an origin of replication, preferably p15A, antibiotic resistance gene, preferably chloramphenicol, a transcription promoter, preferably P_R of the lambda bacteriophage, a repressor gene, preferably cI857ts, a translation initiation signal, and possibly a sequence encoding 6 histidine residues, as well as a sequence encoding a translation stop signal. Preferably, it has SEQ ID NO: 7. The following examples contain a detailed description of one possible embodiment variant of the method according to the present invention. An alternative method of cloning in the insert is the use of a SmaI-digested vector with sticky ends filled in using DNA polymerase in the presence of deoxyribonucleotide triphosphates. Following the present invention a person skilled in the art can propose subsequent embodiment variants. Preferably, the epitope is a HBV epitope, particularly that encoded by the synthetic sequence 9 (see also Fig. 3). Preferably, the amplified monomer segment may contain different epitopes, from different proteins or different regions of the same protein, preferably encoded by a synthetic sequence (see schematic in Fig. 1).

[0008] We also disclose a method of constructing as well as using artificial genes that do not occur in nature using genetic engineering methods as well as chemical synthesis, containing multiple DNA copies encoding repeating segments, containing multiple monomer units of one or more peptides. The amplification of a gene, encoding a peptide (epitope) with a particular biological or chemical function leads to the amplification of the desirable interaction of the resulting (poly)peptide with a specific ligand. In particular such polyepitopic proteins are useful as: (i) artificial antigens - a new generation of vaccines with a magnified potential stimulation of the immune system; (ii) polyproteins containing modules for rare metal chelation for their industrial production or environmental remediation; (iii) a binding module for enzyme cofactors (such as cations, anions, organic molecules) such as proteases acting within a wound in order to stop deleterious activities; (iv) protective multi-epitopic proteins, multiplex modules containing peptides with activators or inhibitors of biological functions for the treatment of molecular, viral and bacterial diseases; (v) multi-epitopic proteins containing multimers of peptide hormones or biologically active fragments of signalling proteins and those that stimulate tissue regeneration. Such proteins, placed in a wound, would gradually release biologically active peptides under the influence of proteinases, stimulating the regeneration of tissue; (vi) the polyepitopic protein is immobilized on macromolecular carriers, such as microorganisms, cells, bacteria, bacteriophages, viruses, defective virions, autoaggregating

proteins, or nanoparticles. The immobilization may be performed using genetic or chemical means. Immobilized poly-epitopic proteins, may magnify the effect of the envisaged uses (i) - (vi). In particular, we designed a vector-enzymatic system for the amplification of a DNA segment. The amplified DNA segment may be natural origin or the result of a chemical synthesis.

Example 1. General schematic of the embodiment of the method according to the present invention.

[0009] Fig. 1 illustrates a general schematic of the method according to the present invention described using the example of the amplification of the surface antigen of HBV. This represents a schematic of the amplification vector as well as amplification reaction of the epitope (in the figure: synthetic HBV epitope). The amplifying vector contains 2 convergent DNA sequences recognised by the sub-type IIS restriction endonuclease, that preferentially recognises a relatively long DNA sequence, which cuts DNA and generates 3-nt (or multiples of 3 nt) sticky ends. We used the Sapl endonuclease, whose particular characteristic is that it recognises a relatively long sequence of 7 base pairs (unique in the vector and amplified DNA segment) which cuts DNA at a distance of 1 nt in the upper chain and 4 nt in the lower chain, thereby generating 3-nt sticky ends, or the equivalent of a single codon. The Sapl sites are adjacent in the vector to the sequence of the classic Type II endonuclease, which is designed for cloning in the inserted DNA. We used the SmaI endonuclease, which cuts DNA within the recognition sequence, generating the so-called "blunt" ends. A vector cut with SmaI may be cloned with any arbitrary DNA segment, synthetic or natural, which is then to be amplified. In a preferable embodiment, the amplified DNA segment encodes an antigen or amino-acid sequence encompassing several identical or differing antigens. The only limit is the length of the amplified fragment, as dictated by the length of the insert DNA accepted by a given class of DNA vector. The amplifying module may be transferred to different classes of vectors using cloning.

Example 2. A series of 6 designed pAMP1 vectors.

[0010] In the example embodiment shown in Fig. 2 we designed the 1st series of pAMP vectors, based on the vector skeleton, containing a p15A origin of replication. From among the designed vectors, we built pAMP1-A (not containing the 6 histidine residue fusion tag) as well as pAMP1-HisA (containing the 6 histidine residue tag). Additionally, the pAMP vector series was built with the option of the high expression of proteins under the control of the lambda phage P_R as well as the option of fusing with the His6 tag, which enables the isolation of the polyepitopic protein using an efficient metalloaffinity chromatography protocol.

[0011] Vectors contain the origin of replication p15A, an antibiotic resistance gene against chloramphenicol, the strong transcription promoter P_R from the lambda bacteriophage, a repressor gene, cl857ts, translation initiation signals, a sequence of 6 histidine residues with an affinity for nickel ions, a restriction site system for fusing to the translation start codon ATG as well as a module for the directional amplification of a DNA fragment maintaining the ORF, containing convergent restriction sites for a IIS subtype endonuclease, preferably Sapl, separated by a short DNA segment, which can contain ancillary restriction sites for cloning in the insert DNA, preferably SmaI. The variants differ in terms of the possibility of manipulating three reading frames (which may be significant when amplifying natural, non-synthetic DNA sequences) as well as the presence or absence of a His6 tag (excellent for easing the subsequent isolation of the expressed polyepitopic protein, regardless of its charge, solubility and other biochemical parameters). Variant 4 was used for the following example of the amplification of the epitope from the surface antigen of HBV. The amplifying module may be introduced by way of cloning in various classes of vector, containing, for example, alternative origins of replication, antibiotic resistance genes, transcriptional promoters and translation signals. For example, we transferred the amplifying module to the vector pBAD/Myc-HisA as well as pET21d21d(+), possessing ampicillin resistance, a colE1 origin of replication as well as araBAD or T7 transcription promoters, respectively. Synthetic modules with sticky ends for the enzymes NcoI and SacI, in versions containing and not containing the His6 residue affinity tag, were cloned into a vector cut with these enzymes, thereby enabling the expression of the polyepitopic proteins using the araBAD or T7 promoters (respectively). We obtained the vectors: pBADAMP1-A, pBADAMP1-HisA, pETAMP1-A, pETAMP1-HisA possessing the inserted amplifying module at the standard MCS (multiple cloning site).

[0012] The full sequence of the pAMP1-HisA vector used in example 3 is shown as sequence 7. Furthermore, in Fig. 8 we show a restriction map of the pAMP1-HisA vector, and in Fig. 9 a restriction map of the pAMP1-HisA vector with a cloned-in 13-mer antigenic peptide.

[0013] The sequences of the synthetic oligodeoxyribonucleotides, encoding the amplifying module transferred into pBAD and pET vectors, used in example 2 are shown as sequences 12, 13, 14 and 15.

Example 3. Production of a polyepitopic protein containing a model 7 amino-acid epitope of the HBV surface antigen.

[0014] A model 7 amino-acid epitope of the HBV surface antigen subjected to the amplification reaction is shown in

Fig. 3.

[0015] The synthetic DNA fragment encoding the epitope of the HBV surface antigen was subjected to a pilot amplification experiment in the vector pAMP1-HisA. We obtained >60 copies of the epitope in the DNA concatemer *in vitro* as well as 13 copies IN THE HYBRID polyepitopic protein cloned *in vivo*.

[0016] Fig. 4 shows the results of the amplification reaction *in vitro* of a 7 amino-acid epitope from the model HBV surface protein using the pAMP1-HisA vector, the Sapl endonuclease as well as DNA ligase.

[0017] We analysed the amplification reaction using PAGE. We cloned a synthetic DNA fragment of 21 bp into the amplifying vector pAMP1-HisA, encoding the 7 amino-acid epitope of HBV. A plasmid containing the monomer HBV epitope digested with the Sapl endonuclease, excising the modified epitope gene from the plasmid construct. The modification consisted of adding to it 3-nt, single-stranded 5' sticky ends. Aside from the amplification function, in the final polymeric hybrid protein, these ends are responsible for the addition of a proline residue, the so-called "*helical breaker*", which separate the epitope monomers and facilitate the independent folding of the epitope into tertiary structures, thereby help to maintain their natural spatial structure. The number of added "*helical breakers*" can be regulated arbitrarily by incorporating amino-acids encoding them to the end of the synthetic epitope (at the level of its encoding DNA). The excised modified DNA encoding the epitope was subjected to autoligation *in vitro*. Lanes from 5 to 160 minutes show the autoligation kinetics. Reaction products were analysed electrophoretically, yielding a series of DNA segments of increasing length, that are directional concatemers (polymers) of the epitope gene in relation to the control reaction without the DNA ligase (K). The resulting *in vitro* concatemers were recloned into pAMP1-HisA, where they could be subjected to another amplification cycle or expression of the encoded multimeric protein.

[0018] Fig. 5 shows the results of hybrid clone analysis of genes encoding the polyepitope HBV protein obtained during the first amplification round.

[0019] The mixture of *in vitro* polymerised synthetic HBV epitope genes (Fig. 4, lane: MIX) using the following system: the amplification vector pAMP1-HisA/Sapl endonuclease/DNA ligase was recloned into the amplification vector pAMP1-HisA, so as to fix the variants of poly-HBV epitope genes, express the encoded polyepitopic proteins as well as to reiterate the amplification cycle. In the first round of the amplification reaction, we obtained a series of clones, containing from 2 to 13 copies of the epitope (lanes 1-13, PCR analysis, DNA fragments migrate slower and slower due to their increasing length, a direct sign of the number of attached epitope copies). An analysis of the DNA sequence showed the continuity of the ORF in the resulting constructs, thus each such synthetic poly-gene encodes different polyepitopic HBV protein, which is significant in terms of the target role of these synthetic proteins. This is due to the fact that their different variants will induce differing intensities of immune response as well as differing in solubility. After re-excising with Sapl, each of the concatemer gene variants may be in completely subjected to another amplification reaction, leading to subsequent poly-genes composed of hundreds of copies of the HBV epitope gene (separated by proline residues), within the hybrid construct, maintaining ORF continuity in an *Escherichia coli* bacterial hyper-expression system. Thereby, each consecutive amplification round increases the number of monomer copies in the concatemer at a geometric rate, leading in a short time to obtaining the planned number of copies in the target plasmid construct variants.

[0020] Fig. 6 shows the results of the second round of amplification of a pentamer concatemer of the HBV epitope.

[0021] The DNA fragment was excised using the Sapl endonuclease from a pAMP1-HisA construct containing a concatemer of 5 epitope copies, obtained during the 1st round of amplification and subjected to amplification again. The largest concatemer, visible at the edge of agarose gel resolution, contains 12 copies of the pentamer, constituting a 60-fold directionally polymerised HBV epitope. Larger concatemers are evidently visible, although not separated into distinct bands. The resulting 2nd round products were recloned into pAMP-HisA and may be subjected to a third round of round of amplification, leading to the production of hundreds or thousands of HBV epitope copies, set out in a single recombinant polypeptide (protein) with a continuous ORF. These clones were also subjected to analytical expression in order to obtain variants of epitope multiplication within the polyepitopic protein.

[0022] Fig. 7 shows an example expression analysis of the hybrid gene encoding the polyepitopic HBV protein, obtained during the first, model, round of amplification (13-mer) in the pAMP1-HisA vector (Seq. Id. No. 7, Fig. 8). We performed an electrophoretic analysis of the expressed polyHBV epitope 13-mer in a polyacrylamide gel under denaturing conditions and Coomassie Blue staining. Lane M, protein mass marker (GE LMW Calibration Kit); lanes K0-3: recombinant *Escherichia coli* culture, containing a 13-mer HBV construct prior to induction, and sampled at 0, 1, 2 and 3 hours. lanes 1h-16h: recombinant *Escherichia coli* cultures, containing the 13-mer HBV construct after thermal induction, sampled at 1, 2, 3, 4 and 16 hours. The red arrow indicates the growing HBV 13-mer band, stained purple-red. The unusual purple-red staining of the 13-mer band, whose concentration increases over the duration of the expression induction in bacteria (consecutive lanes), stems from the ordered arrangement of the stain on repeating sequence amino-acid segments and may thus serve as an additional tests for the presence of polyepitopic proteins during detection and isolation. We used expression analysis (protein synthesis *in vivo*) on a clone containing a variant of the 13-copy concatemer genu resulting from the amplification of a 7-amino-acid epitope of HBV, where each epitope is separated by a proline residue (expression plasmid see: Seq. Id. No. 8, Fig. 9). This technology maintains ORF continuity. The expression was conducted in the system: *Escherichia coli* / P_R lambda bacteriophage promoter, yielding the over-expression of the artificial poly-

epitopic HBV protein.

[0023] Fig. 10 ABC represents the isolation procedure of the expression of the polyepitopic protein example variant: 13-copy concatemer gene, as well as Western blot detection. From a preparative (10 L) recombinant bacterial culture (bearing the amplification vector, expressing the 13-mer HBV epitope) and after induction and expression, w centrifuged off the bacterial biomass and subjected it to the following procedures: (i) ultrasonic disintegration and centrifugation of cellular debris; (ii) heating the lysate to 65°C to denature the host proteins and centrifugation of the precipitated proteins; (iii) treatment of the preparation with buffered polyethylenimine in order to remove nucleic acids as well as acidic host proteins and centrifugation of the precipitate; (iv) fractioning by salting out with ammonium sulphate, centrifuging the precipitate containing the 13-mer HBV epitope and its dissolution in a buffer for use in a subsequent isolation stage; (v) AKTA high-performance affinity chromatography in a gel for chelating nickel ions - HiTrap IMAC HP, which binds the 6 histidine residue tag at the N-end of the 13-mer; (vi) AKTA high-performance molecular sieving chromatography in Superdex 200 pg. The purified preparation of the electrophoretically homogenous 13-mer polyepitopic protein was subjected to Western blotting using anti-6 histidine residue tag monoclonal antibodies, where the chromogenic reaction was performed using conjugated HRP as well as 3,3',4,4'-tetraaminobiphenyl tetrahydrochloride. Fig. 10 shows: panels A-B: denaturing PAGE of samples from consecutive stages of isolating the 13-mer HBV epitope. Lane: M - molecular mass marker (GE LMW Calibration Kit); lane 1 - recombinant bacteria cells expressing the 13-mer; lane 2 - cell extract resulting from ultrasonication; lane 3 - denatured protein precipitate after heating; lane 4 - supernatant containing the 13-mer after treatment with polyethylenimine; lane 5 and 6 - supernatant containing 13-mer after fractionation with ammonium sulphate; lane 7 - HBV 13-mer after purification using metalloaffinity chromatography; lane 8 - homogenous 13-mer after molecular sieving chromatography; panel C: Western blot detection. Lane 1 - molecular mass marker (GE LMW Calibration Kit); lane 2 - purified HBV 13-mer preparation; lane 3 - Western blot detection with anti-6 histidine residue tag (Merck) antibodies.

[0024] Because none of the recombinant host's own proteins contain the sequence of 6 histidines, the positive reaction indisputably confirms that the isolated protein is the polyepitope 13-mer of HBV. An additional confirmation is the expected size of the isolated protein in comparison to mass markers as well as specific binding to the HiTrap IMAC HP gel. The procedure is universal, successfully confirmed in the isolation of other variants of polyepitopic proteins, fused with with 6 histidine residue tags, and contains varying amounts of polymerised HBV epitope.

Example 4. Production of higher order multimeric structures containing several hundred immobilised copies of the polyepitopic protein, containing the model 7 amino-acid epitope of the HBV surface antigen on the capsid surface of T7 bacteriophage.

[0025] The construction of polyepitopic proteins facilitates a multiple increase in the concentration of the epitope in a single protein molecule, and a subsequent amplification stage is based on the combination of many polyepitopic protein molecules in one higher order structure. This may be achieved: (i) biologically, through the fusion of the gene encoding the polyepitopic protein with the genetic material of a higher order structural carrier such as a chromosome, nucleoid, microorgamal plasmid, a bacterium, bacteriophage or virus or (ii) by chemical means, through the use of factors that conjoin macromolecules.

[0026] Fig. 11 represents the sequence of a poly-genic polyepitopic protein of a pentamer epitope with indicated primers that serve to amplify the HBV epitope polygene and to introduce the restriction sites for EcoRI and HindIII, which are used for genetically fusing it with a bacteriophage vector. This sequence was used to design an example biological immobilisation procedure for the polyepitope HBV pentamer epitope protein on the surface of the T7 bacteriophage capsid which infects *Escherichia coli*. We used the commercial Novagen T7Select® Phage Display System in the version using the phage vector T7Select415-1b. The system precisely immobilised 415 polypeptide copies with a length not exceeding ca. 50 amino-acid residues, thus the complete macromolecular structure, where we preferably used a pentamer HBV epitope, contains 2075 copies HBV epitope. The immobilisation procedure for the polyepitopic protein on the bacteriophage capsid surface consists of the following stages: (i) PCR amplification of the fragment encoding the pentamer HBV epitope, introducing restriction sites for the endonucleases HindIII and EcoRI, dedicated for the vector T7Select415-1b for cloning DNA fragments, while maintaining ORF continuity for its translation with a protein forming the T7 capsid; (ii) digestion of the PCR-amplified fragment with the restriction endonucleases HindIII and EcoRI; (iii) directional ligation of the digested PCR fragment with pre-digested HindIII and EcoRI left and right arms of the T7 bacteriophage; (iv) packing *in vitro* of the recombinant T7 bacteriophage and infection of the host *Escherichia coli*; (v) isolation of individual clones of the recombinant T7 bacteriophage, amplification of the clones as well as validation of the DNA sequence produced with the designed sequence.

[0027] The sequence of the PCR product being the substrate for producing the insert for the Phage Display System (encoding a variant of the pentamer HBV epitope) used in Example 4 is shown as sequence 16, and the sequences of the PCR primers used to amplify the variant of 5 HBV epitopes are shown as sequences 17 and 18.

Example 5. Immunological activity of the polyepitope HBV protein.

[0028] The 25-mer polyepitope HBV protein of Example 2 was isolated from recombinant bacteria using the procedure shown in Example 4. Representative groups of mice (6 individuals each) were inoculated with 20 ug/mouse with a purified 25-mer in PBS buffer mixed with Freund's incomplete adjuvant. In parallel, we inoculated a control group with PBS and Freund's incomplete adjuvant. The vaccination cycle encompassed 3 injections at 3-week intervals.

[0029] Fig. 12 shows a Western blot of a purified 25-mer used in example 3 as well as a recombinant T7 bacteriophage displaying several hundred copies of the pentamer HBV epitope used in example 4. The samples, separated using denaturing PAGE, were Western blotted, using sera from mice inoculated with the 25-mer polyepitope HBV protein mixed with incomplete Freund's adjuvant as well as the sera of control mice inoculated with PBS and incomplete Freund's adjuvant. Lane M, Page Ruler marker; lane 1, purified recombinant bacteriophage containing the pentamer HBV epitope. Panel A left-hand arrow denotes the position of the fusion protein of the bacteriophage T7 capsid and the pentamer HBV epitope. The right-hand arrows in Panel A show multimeric forms of the purified 25-mer protein; Lane 2: 25-mer polyepitope HBV protein according to Fig. 10 purified to homogeneity. The results show a specific humoral immune response after 3 initial immunisation cycles, directed against multimers of the 7-amino-acid HBV epitope, present both in the polyepitopic protein as well as polyepitopic proteins displayed on the capsid of the recombinant bacteriophage T7. At the same time, a positive Western blot result, in addition to the DNA sequence of the recombinant bacteriophage T7, validates the correctness of the Example 4 cloning procedure and the variants of recombinant bacteriophages displaying the pentamer HBV epitope, as well as indicating the possibility of using the described technology in variants of new generation vaccines. The resulting polyepitopic HBV proteins as well as the recombinant T7 bacteriophage, endowed with multiple copies of the polyepitopic HBV protein constitute a prototype anti-HBV vaccine.

SEQUENCE LISTING

[0030]

<110> BioVentures Institute Sp. z o.o.

<120> A method of obtaining a polyepitope protein as well as a DNA vector for embodying the method<130> PK/2483/RW

<160> 18

<170> PatentIn version 3.5

<210> 1

<211> 41

<212> DNA

<213> artificial

<220>

<223> pAMP1-A: amplifying module

<400> 1

atgcgctctt caccggggcc cagaagagct aagtaagtaa g 41

<210> 2

<211> 40

<212> DNA

<213> artificial

<220>

<223> pAMP1-B: amplifying module

<400> 2

atggctcttc acccggggccc agaagagcta agtaagtaag 40

<210> 3

<211> 42

<212> DNA

<213> artificial

<220>

<223> pAMP1-C: amplifying module

<400> 3

atgccgctct tcaccggggc ccagaagagc taagtaagta ag 42

<210> 4

<211> 59

<212> DNA

EP 3 134 426 B1

<213> artificial
 <220>
 <223> pAMP1-HisA: amplifying module
 <400> 4
 5 atgcaccacc accaccacca cagctcttca cccgggcccc gaagagctaa gtaagtaag 59
 <210> 5
 <211> 58
 <212> DNA
 <213> artificial
 10 <220>
 <223> pAMP1-HisB: amplifying module
 <400> 5
 atgcaccacc accaccacca cgctcttcac cggggccag aagagctaag taagtaag 58
 <210> 6
 15 <211> 60
 <212> DNA
 <213> artificial
 <220>
 <223> pAMP1-HisC: amplifying module
 20 <400> 6
 atgcaccacc accaccacca cccgctcttc accggggccc agaagagcta agtaagtaag 60
 <210> 7
 <211> 3003
 <212> DNA
 25 <213> artificial
 <220>
 <223> vector pAMP1-HisA
 <400> 7
 30 catgcaccac caccaccacc acagctcttc acccgggccc agaagagcta agtaagtaag 60
 agctcgaccc ggtcgaaattt gctttcgaat ttctgccatt catccgctta ttatcactta 120
 ttcaggcgta gcaccaggcg ttttaaggga ccaataactg ccttaaaaaa attacgcccc 180
 gccctgccac tcatcgcagt actggtgtaa ttcattaagc attctgccga catggaagcc 240
 atcacagacg gcatgatgaa cctgaatcgc cagcggcatc agcaccttgt cgccttgctg 300
 35 ataatatattg cccatgggtga aaacgggggc gaagaagtgt tccatatttg ccacgtttta 360
 atcaaaaactg gtgaaaactca cccagggtatt ggctgagacg aaaaacatat tctcaataaa 420
 cccttttaggg aaataggcca ggtttttcacc gtaacacgcc acatcttgcg aatatatgtg 480
 tagaaactgc cggaaatcgt cgtgggtattc actccagagc gatgaaaacg tttcagtttg 540
 ctcatggaaa acggtgtaac aagggtgaac actatcccat atcaccagct caccgtcttt 600
 cattgccata cggaattccg gatgagcatt catcaggcgg gcaagaatgt gaataaaggc 660
 40 cggataaaaac ttgtgcttat ttttctttac ggtcttttaa aaggccgtaa tatccagctg 720
 aacgggtctgg ttatagggtac attgagcaac tgactgaaat gcctcaaaaat gttcttttacg 780
 atgccattgg gatatatcaa cgggtgtata tccagtgatt tttttctcca ttttagcttc 840
 cttagctcct gaaaatctcg ataactcaaa aaatacgccc ggtagtgatc ttatttcatt 900
 atggtgaaaag ttggaacctc ttacgtgccg atcaacgtct cattttcgcc aaaagttggc 960
 45 ccagggtcttc ccggtatcaa cagggaacac aggatttatt tattctgcga agtgatcttc 1020
 cgtcacagggt atttattcgg cgcaaagtgc gtcgggtgat gctgccaact tactgattta 1080
 gtgtatgatg gtgtttttga ggtgctccag tggcttctgt ttctatcagc tgtccctcct 1140
 gttcagctac tgacgggggtg gtgcgtaacg gcaaaagcac cgccggacat cagcgtagc 1200
 ggagtgtata ctggcttact atgttggcac tgatgagggt gtcagtgaag tgcttcattg 1260
 ggcaggagaa aaaaggctgc accggtgcgt cagcagaata tgtgatacag gatataattcc 1320
 50 gcttcctcgc tcaactgactc gctacgctcg gtcgttcgac tgccggcgagc ggaaatggct 1380
 tacgaacggg gcggagattt cctggaagat gccaggaaga tacttaacag ggaagtgaga 1440
 gggccgcggc aaagcogttt ttccataggc tccgcccccc tgacaagcat caggaatct 1500
 gacgctcaaa tcagtgggtg cgaaaccgca caggactata aagataccag gcgtttcccc 1560
 55

EP 3 134 426 B1

	ctggcggtc	cctcgtgcgc	tctcctgttc	ctgcctttcg	gtttaccggt	gtcattccgc	1620
	tgttatggcc	gcgtttgtct	cattccacgc	ctgacactca	gttcggggt	ggcagttcgc	1680
	tccaagctgg	actgtatgca	cgaaccccc	gttcagtcgc	accgctgcgc	cttatccggt	1740
5	aactatcgtc	ttgagtccaa	cccgaaaga	catgcaaaag	caccactggc	agcagccact	1800
	ggtaattgat	ttagaggagt	tagtcttgaa	gtcatgcgcc	ggttaaggct	aaactgaaag	1860
	gacaagtttt	ggtagactgcg	ctcctccaag	ccagttacct	cggttcaaag	agttggtagc	1920
	tcagagaacc	ttcgaaaaac	cgccctgcaa	ggcggttttt	tcgttttcag	agcaagagat	1980
	tacgcgcaga	ccaaaacgat	ctcaagaaga	tcattcttatt	aatcagataa	aatattttcta	2040
	gattttcagt	caatttatct	cttcaaagt	agcacctgaa	gtcagcccca	tacgatataa	2100
10	gttgtaattc	tcatgtttga	cagcatgcc	tcgatcgcga	tcttgctcaa	ttgttatcag	2160
	ctatgocccg	accagaacac	cttgccgatc	agccaaacgt	ctcttcaggc	cactgactag	2220
	cgataacttt	ccccacaacg	gaacaactct	cattgcatgg	gatcattggg	tactgtgggt	2280
	ttagtgggtg	taaaaacacc	tgaccgctat	ccctgatcag	tttcttgaag	gtaaactcat	2340
	cacccccaa	tctggctatg	cagaaatcac	ctggctcaac	agcctgctca	gggtcaacga	2400
15	gaattaacat	tccgtcagga	aagcttggct	tggagcctgt	tggtgcggtc	atggaattac	2460
	cttcaacctc	aagccagaat	gcagaatcac	tggctttttt	ggttgtgctt	acccatctct	2520
	ccgcatcacc	tttggttaaag	gttctaagct	caggtgagaa	catccctgcc	tgaacatgag	2580
	aaaaaacagg	gtactcatac	tcacttctaa	gtgacggctg	cataactaac	gcttcataca	2640
	tctcgtagat	ttctctggcg	attgaagggc	taaattcttc	aacgctaact	ttgagaattt	2700
	ttgcaagcaa	tgcggcggtt	taagcattta	atgcattgat	gccattaaat	aaagcaccaa	2760
20	cgctgactg	ccccatcccc	atcttgtctg	cgacagattc	ctgggataag	ccaagttcat	2820
	ttttcttttt	ttcataaatt	gctttaaggc	gacgtgcgtc	ctcaagctgc	tcttgtgtta	2880
	atggtttctt	ttttgtgctc	atacgttaaa	tctatcaccg	caagggataa	atatctaaca	2940
	ccgtgcgtgt	tgactatttt	acctctggcg	gtgataatgg	ttgcatgtac	taaggagggt	3000
	gtt						3003
25	<210> 8						
	<211> 3348						
	<212> DNA						
	<213> artificial						
30	<220>						
	<223> vector pAMP1-HisA-13-epitope						
	<400> 8						
35							
40							
45							
50							
55							

EP 3 134 426 B1

	catgcaccac	caccaccacc	acagctcttc	acccaccaaa	ccgaccgacg	gtaacggggc	60
	caccaaaccg	accgacggta	acgggcccac	caaaccgacc	gacggtaacg	ggcccaccaa	120
	accgaccgac	ggtaacgggc	ccaccaaacc	gaccgacggt	aacggggcca	ccaaaccgac	180
	cgacggtaac	ggggccacca	aaccgaccga	cggtaacggg	cccaccaaac	cgaccgacgg	240
5	taacggggcc	accaaaccga	ccgacggtaa	cgggcccacc	aaaccgaccg	acggtaacgg	300
	gcccaccaa	ccgaccgacg	gtaacggggc	caccaaaccg	accgacggta	acggggcccac	360
	caaaccgacc	gacggtaacg	ggcccagaag	agctaagtaa	gtaagagctc	gaccgggtcg	420
	aatttgcttt	cgaattttctg	ccattcatcc	gcttattatc	acttattcag	gcgtagcacc	480
	aggcggttaa	gggcaaccaat	aactgcctta	aaaaaattac	gccccgccct	gccactcatc	540
10	gcagtactgt	tgtaattcat	taagcattct	gccgacatgg	aagccatcac	agacggcatg	600
	atgaacctga	atcgccagcg	gcacgacac	cttgctgcct	tgcgataaat	atttgcccat	660
	ggtgaaaacg	ggggcgaaga	agttgtccat	attggccaacg	tttaaataca	aactggtgaa	720
	actcaccacg	ggattggctg	agacgaaaaa	catattctca	ataaacccctt	tagggaaata	780
	ggccagggtt	tcaccgtaac	acgccacatc	ttgcgaatat	atgtgtagaa	actgccggaa	840
	atcgctcggt	tattcactcc	agagcgatga	aaacgtttca	gtttgctcat	ggaaaacggg	900
15	gtaacaaggg	tgaacactat	cccataatcc	cagctcaccg	tctttcattg	ccatacggaa	960
	ttccggatga	gcattcatca	ggcggggaag	aatgtgaata	aaggccggat	aaaactttgtg	1020
	cttatttttt	tttacgggtc	ttaaaaaggc	cgtaatatcc	agctgaacgg	tctggttata	1080
	ggtacattga	gcaactgact	gaaatgcctc	aaaatgttct	ttacgatgcc	attgggatat	1140
	atcaacgggtg	gtatatccag	tgattttttt	ctccatttta	gcttccttag	ctcctgaaaa	1200
20	tctcgataac	tcaaaaaata	cgcccggtag	tgatcttatt	tcattatggt	gaaagtggga	1260
	acctcttacg	tgccgatcaa	cgtctcattt	tcgccaataa	ttggcccagg	gcttcccggg	1320
	atcaacaggg	acaccaggat	ttattttatt	tgcgaaagtga	tcttccgtca	caggatattta	1380
	ttcgggcgaa	agtgcgtcgg	gtgatgctgc	caacttactg	athtagtgta	tgatggtggt	1440
	tttgagggtg	tccagtggct	tctgtttcta	tcagctgtcc	ctcctgttca	gctactgacg	1500
	gggtggtgcg	taacggcaaa	agcaccggcg	gacatcagcg	ctagcggagt	gtatactggc	1560
25	ttactatggt	ggcactgatg	aggggtgcag	tgaagtgtct	catgtggcag	gagaaaaaag	1620
	gctgcaccgg	tgcgctcagca	gaatatgtga	tacagtatat	attccgcttc	ctcgctcact	1680
	gactcgctac	gctcgggtcgt	tcgactgcgg	cgagcggaaa	tggtttacga	acggggcgga	1740
30							
	gatttcctgg	aagatgccag	gaagatactt	aacagggaag	tgagagggcc	gcggaacagc	1800
	cgtttttcca	taggctccgc	ccccctgaca	agcatcacga	aatctgacgc	tcaaatcagt	1860
	ggtggcgaaa	cccgacagga	ctataaagat	accaggcggt	tccccctggc	ggctccctcg	1920
	tgcgctctcc	tggtcctgcc	tttcgggtta	ccggtgtcat	tccgctgtta	tgcccgcggt	1980
35	tgtctcattc	cacgcctgac	actcagttcc	gggtaggcag	ttcgctccaa	gctggactgt	2040
	atgcacgaac	cccccggtta	gtccgaccgc	tgcgcttat	ccggtaacta	tcgtcttgag	2100
	tccaacccgg	aaagacatgc	aaaagcacca	ctggcagcag	ccactggtaa	ttgatttaga	2160
	ggagttagtc	ttgaagtcat	gcgcgggtta	aggctaaact	gaaaggacaa	gttttggtga	2220
	ctcgctcctc	ccaagccagt	tacctcggtt	caaagagtgt	gtagctcaga	gaacctcga	2280
	aaaaccgccc	tgaaggcg	ttttttcggt	ttcagagcaa	gagattacgc	gcagaccaa	2340
40	acgatctcaa	gaagatcatc	ttattaatca	gataaaatat	ttctagattt	cagtgcattt	2400
	tatctcttca	aatgtagcac	ctgaagtcag	ccccatacga	tataagttgt	aattctcatg	2460
	tttgacagca	tgccatcgat	cgcatcttg	ctcaattggt	atcagctatg	cgccgaccag	2520
	aacaccttgc	cgatcagcca	aacgtctctt	caggccactg	actagcgata	actttcccca	2580
	caacggaaca	actctcattg	catgggatca	ttgggtactg	tggttttagt	ggttgtaaaa	2640
45	acacctgacc	gctatccctg	atcagtttct	tgaaggtaaa	ctcatcacc	ccaagtctgg	2700
	ctatgcagaa	atcacctggc	tcaacagcct	gctcagggtc	aacgagaatt	aacattccgt	2760
	caggaaagct	tggttgagg	cctgttggtg	cggtcatgga	attaccttca	acctcaagcc	2820
	agaatgcaga	atcactggct	tttttggttg	tgcttaccga	tctctccgca	tcacctttgg	2880
	taaaggttct	aagctcaggt	gagaacatcc	ctgcctgaac	atgagaaaaa	acagggtact	2940
50	catactcact	tctaagtgac	ggctgcatac	taaccgcttc	atacatctcg	tagattttct	3000
	tggtcgattga	agggtcaaat	tcttcaacgc	taactttgag	aatttttgca	agcaatgcgg	3060
	cgttataaag	atttaaatgca	ttgatgccat	ttaaataaag	accaacgcct	gactgccccca	3120
	tccccatctt	gtctgcgaca	gattcctggg	ataagccaag	ttcatttttc	tttttttcat	3180
	aaattgcttt	aaggcgacgt	gcgtcctcaa	gctgctcttg	tgtaaatggt	ttcttttttg	3240
	tgctcatacg	ttaaatctat	caccgcaagg	gataaatatc	taacaccggtg	cgtgttgact	3300
55	attttacctc	tggcggtgat	aatggttgca	tgtactaagg	agggttggt		3348

<210> 9

<211> 21

EP 3 134 426 B1

<212> DNA
 <213> artificial
 <220>
 <223> sequence encoding the HBV epitope
 5 <220>
 <221> CDS
 <222> (1)..(21)
 <400> 9

 10 acc aaa ccg acc gac ggt aac 21
 Thr Lys Pro Thr Asp Gly Asn
 1 5

 <210> 10
 15 <211> 7
 <212> PRT
 <213> artificial
 <220>
 <223> Synthetic Construct
 20 <400> 10

 Thr Lys Pro Thr Asp Gly Asn
 1 5

 25 <210> 11
 <211> 25
 <212> DNA
 <213> artificial
 <220>
 30 <223> Amplifying module
 <400> 11
 gctcttcacc cgggccaga agagc 25
 <210> 12
 <211> 42
 35 <212> DNA
 <213> artificial
 <220>
 <223> pAMP1-A upper oligo
 <400> 12
 40 catgcgctct tcaccgggc ccagaagagc taagtaagta ag 42
 <210> 13
 <211> 42
 <212> DNA
 <213> artificial
 45 <220>
 <223> pAMP1-A lower oligo
 <400> 13
 agctcttact tacttagctc ttctgggccc gggtaagag cg 42
 <210> 14
 50 <211> 60
 <212> DNA
 <213> artificial
 <220>
 <223> pAMP1-HisA upper oligo
 55 <400> 14
 catgcaccac caccaccacc acagctcttc acccgggccc agaagagcta agtaagtaag 60
 <210> 15
 <211> 60

EP 3 134 426 B1

<212> DNA
 <213> artificial
 <220>
 <223> pAMP1-HisA lower oligo
 5 <400> 15
 agctcttact tacttagctc ttctgggccc ggggaagag ctgtggtggt ggtggtggtg 60
 <210> 16
 <211> 182
 <212> DNA
 10 <213> artificial
 <220>
 <223> PCR sequence containing the insert (variant with 5 HBV antigen repeats) for the Phage Display System
 <400> 16
 15 caccaccacc acaggaattc acccaccaaa ccgaccgacg gtaacggggcc caccaaaccg 60
 accgacggta acggggccac caaacccgacc gacggtaacg ggcccaccaa accgaccgac 120
 ggtaacgggc ccaccaaacc gaccgacggt aactaagctt gaagagctaa gtaagtaaga 180
 gc 182
 20 <210> 17
 <211> 28
 <212> DNA
 <213> artificial
 <220>
 25 <223> Top primer for the Phage Display System
 <400> 17
 caccaccacc acaggaattc acccacca 28
 <210> 18
 <211> 39
 30 <212> DNA
 <213> artificial
 <220>
 <223> Bottom primer for the Phage Display System
 <400> 18
 35 gctcttactt acttagctct tcaagcttag ttaccgtcg 39

SEQUENCE LISTING

[0031]

40 <110> BioVentures Institute Sp. z o.o.
 <120> Sposób uzyskiwania białka polipeptopowego oraz wektor DNA do realizacji tego sposobu
 45 <130> PK/2483/RW
 <160> 19
 <170> PatentIn version 3.5
 50 <210> 1
 <211> 41
 <212> DNA
 <213> artificial
 55 <220>
 <223> pAMP1-A: modul powielający

EP 3 134 426 B1

<400> 1
 atgcgctctt caccggggcc cagaagagct aagtaagtaa g 41

 <210> 2
 <211> 40
 <212> DNA
 <213> artificial

 <220>
 <223> pAMP1-B: modul powielajŝcy

 <400> 2
 atggctcttc acccggggccc agaagagcta agtaagtaag 40

 <210> 3
 <211> 42
 <212> DNA
 <213> artificial

 <220>
 <223> pAMP1-C: modul powielajŝcy

 <400> 3
 atgccgctct tcacccgggc ccagaagagc taagtaagta ag 42

 <210> 4
 <211> 59
 <212> DNA
 <213> artificial

 <220>
 <223> pAMP1-HisA: modul powielajŝcy

 <400> 4
 atgcaccacc accaccacca cagctcttca cccggggccc gaagagctaa gtaagtaag 59

 <210> 5
 <211> 58
 <212> DNA
 <213> artificial

 <220>
 <223> pAMP1-HisB: modul powielajŝcy

 <400> 5
 atgcaccacc accaccacca cgctcttcac ccggggcccag aagagctaag taagtaag 58

 <210> 6
 <211> 60
 <212> DNA
 <213> artificial

 <220>
 <223> pAMP1-HisC: modul powielajŝcy

 <400> 6
 atgcaccacc accaccacca cccgctcttc acccggggccc agaagagcta agtaagtaag 60

EP 3 134 426 B1

<210> 7
 <211> 3003
 <212> DNA
 <213> artificial

<220>
 <223> wektor pAMP1-HisA

<400> 7

catgcaccac	caccaccacc	acagctcttc	acccgggccc	agaagagcta	agtaagtaag	60
agctcgaccc	ggtcgaat	gctttcgaat	ttctgccatt	catccgctta	ttatcactta	120
ttcaggcgta	gcaccaggcg	tttaagggca	ccaataactg	ccttaaaaaa	attacgcccc	180
gccctgccac	tcatcgcagt	actgttgtaa	ttcattaagc	attctgccga	catggaagcc	240
atcacagacg	gcatgatgaa	cctgaatcgc	cagcggcatc	agcaccttgt	cgccttgctg	300
ataatatattg	cccatggtga	aaacgggggc	gaagaagttg	tccatattgg	ccacgtttaa	360
atcaaaaactg	gtgaaaactca	cccagggtatt	ggctgagacg	aaaaacatat	tctcaataaa	420
cccttttaggg	aaataggcca	ggttttcacc	gtaacacgcc	acatcttgcg	aatatatgtg	480
tagaaaactgc	cggaaatcgt	cgtggtattc	actccagagc	gatgaaaacg	tttcagtttg	540
ctcatggaaa	acggtgtaac	aagggtgaac	actatcccat	atcaccagct	caccgtcttt	600
cattgccata	cgggaattccg	gatgagcatt	catcaggcgg	gcaagaatgt	gaataaaggc	660
cggataaaac	ttgtgcttat	ttttctttac	ggtcttttaa	aaggccgtaa	tatccagctg	720
aacgggtctgg	ttataggtac	attgagcaac	tgactgaaat	gcctcaaaat	gttcttttacg	780
atgccattgg	gatatatcaa	cggtggtata	tccagtgatt	tttttctcca	ttttagcttc	840

EP 3 134 426 B1

	cttagctcct gaaaatctcg ataactcaaa aaatacgccc ggtagtgatc ttatttcatt	900
	atgggtgaaag ttggaacctc ttacgtgccg atcaacgtct cattttogcc aaaagttggc	960
5	ccagggcttc ccggtatcaa cagggacacc aggatttatt tattctgcga agtgatcttc	1020
	cgtcacagggt atttattcgg cgcaaagtgc gtcgggtgat gctgccaaact tactgattta	1080
	gtgtatgatg gtgtttttga ggtgctccag tggcttctgt ttctatcagc tgtccctcct	1140
10	gttcagctac tgacggggtg gtgcgtaacg gcaaaagcac cgccggacat cagcgctagc	1200
	ggagtgtata ctggcttact atgttggcac tgatgagggt gtcagtgaag tgcttcatgt	1260
	ggcaggagaa aaaaggctgc accggtgcgt cagcagaata tgtgatacag gatataattcc	1320
15	gcttcctcgc tcaactgactc gctacgctcg gtcgttcgac tgcggcgagc ggaaatggct	1380
	tacgaacggg gcggagattt cctggaagat gccaggaaga tacttaacag ggaagtgaga	1440
	gggccgcggc aaagccgttt ttccataggc tccgcccccc tgacaagcat cacgaaatct	1500
20	gacgctcaaa tcagtgggtg cgaaaccgca caggactata aagataccag gcgtttcccc	1560
	ctggcggtc cctcgtgcgc tctcctgttc ctgcctttcg gtttacgggt gtcattccgc	1620
25	tgttatggcc gcgtttgtct cattccacgc ctgacactca gttccgggta ggcagttcgc	1680
	tccaagctgg actgtatgca cgaaccccc gttcagtcgg accgctgogc cttatccgggt	1740
	aactatcgtc ttgagtccaa cccggaaaga catgcaaaag caccactggc agcagccact	1800
30	ggtaattgat ttagaggagt tagtcttgaa gtcatgcgcc gggttaaggct aaactgaaag	1860
	gacaagtttt ggtgactgcg ctctccaag ccagttacct cggttcaaag agttggtagc	1920
	tcagagaacc ttcgaaaaac cgccctgcaa ggcggttttt tcgttttcag agcaagagat	1980
35	tacgcgcaga ccaaaacgat ctcaagaaga tcatcttatt aatcagataa aatattttcta	2040
	gatttcagtg caatttatct cttcaaatgt agcacctgaa gtcagcccca tacgatataa	2100
40	gttgtaattc tcatgtttga cagcatgcc a tcgatcgca tcttgctcaa ttgttatcag	2160
	ctatgcgccg accagaacac cttgccgatc agccaaacgt ctcttcaggc cactgactag	2220
	cgataacttt cccacaacg gaacaactct cattgcatgg gatcattggg tactgtgggt	2280
45	ttagtggttg taaaaacacc tgaccgctat ccctgatcag tttcttgaag gtaaactcat	2340
	cacccccaa g tctggctatg cagaaatcac ctggctcaac agcctgctca ggggtcaacga	2400
	gaattaacat tccgtcagga aagcttggct tggagcctgt tggtgoggtc atggaattac	2460
50	cttcaacctc aagccagaat gcagaatcac tggtttttt gggttggtt acccatctct	2520
	ccgcatcacc tttggtaaag gttctaagct cagggtgagaa catccctgcc tgaacatgag	2580
55	aaaaaacagg gtactcatac tcacttctaa gtgacggctg cataactaacc gcttcataca	2640
	tctcgtagat ttctctggcg attgaagggc taaattcttc aacgctaact ttgagaattt	2700

EP 3 134 426 B1

	ttgcaagcaa tgcggcggtta taagcattta atgcattgat gccattaaat aaagcaccaa	2760
	cgcctgactg ccccatcccc atcttgtctg cgacagattc ctgggataag ccaagttcat	2820
5	ttttctttttt ttcataaaatt gctttaaggc gacgtgcgtc ctcaagctgc tcttgtgtta	2880
	atgggtttctt ttttgtgctc atacgttaaa tctatcacccg caagggataa atatctaaca	2940
10	cogtgogtgt tgactatttt acctctggcg gtgataatgg ttgcatgtac taaggaggtt	3000
	gtt	3003
	<210> 8	
	<211> 3348	
15	<212> DNA	
	<213> artificial	
	<220>	
20	<223> wektor pAMP1-HisA-13-epitop	
	<400> 8	
	catgcaccac caccaccacc acagctcttc acccaccaaa ccgaccgacg gtaacggggcc	60
25	caccaaaccg accgacggta acgggcccac caaaccgacc gacggtaacg ggcccaccaa	120
	accgaccgac ggtaacgggc ccaccaaacc gaccgacggt aacggggcca ccaaaccgac	180
30	cgacggtaac gggcccacca aaccgaccga cggtaacggg cccaccaaac cgaccgacgg	240
	taacggggccc accaaaccga ccgacggtaa cggggcccacc aaaccgaccg acggtaacgg	300
	gcccaccaa cccgaccgac gtaacggggc caccaaaccg accgacggta acggggcccac	360
35	caaaccgacc gacggtaacg ggcccagaag agctaagtaa gtaagagctc gacccggtcg	420
	aatttgcttt cgaattttctg ccattcatcc gcttattatc acttattcag gcgtagcacc	480
40	aggcgtttaa gggcaccaat aactgcctta aaaaaattac gccccgccct gccactcatc	540
	gcagtactgt tgtaattcat taagcattct gccgacatgg aagccatcac agacggcatg	600
	atgaacctga atcgccagcg gcatcagcac cttgtcgcct tgcgtataat atttgcccat	660
45	ggtgaaaacg ggggcgaaga agttgtccat attggccacg tttaaataca aactggtgaa	720
	actcaccacg ggattggctg agacgaaaaa catattctca ataaaccctt tagggaaata	780
	ggccaggttt tcaccgtaac acgccacatc ttgcgaatat atgtgtagaa actgccggaa	840
50	atcgtcgtgg tattcactcc agagcgatga aaacgtttca gtttgctcat ggaaaacggc	900
	gtaacaaggg tgaacactat cccatatcac cagctcacccg tctttcattg ccatacgga	960
55	ttccggatga gcattcatca ggcgggcaag aatgtgaata aaggccggat aaaacttgtg	1020
	cttattttttc ttacgggtct ttaaaaaggc cgtaatatcc agctgaacgg tctgggtata	1080
	ggtacattga gcaactgact gaaatgcctc aaaatgttct ttacgatgcc attgggatat	1140

EP 3 134 426 B1

	atcaacggtg	gtatatccag	tgatTTTTTT	ctccattTTT	gotttccttag	ctcctgaaaa	1200
	tctcgataac	tcaaaaaata	cgcccggtag	tgatcttatt	tcattatggt	gaaagttgga	1260
5	acctcttacg	tgccgatcaa	cgtctcattt	tcgccaaaag	ttggcccagg	gcttcccggg	1320
	atcaacaggg	acaccaggat	ttattttatt	tgCGAAGTGA	tcttccgtca	caggtattta	1380
	ttcggcgcaa	agtgcgtcgg	gtgatgctgc	caacttactg	atttagtgta	tgatggtggt	1440
10	tttgagggtgc	tccagtggct	tctgtttcta	tcagctgtcc	ctcctgttca	gctactgacg	1500
	gggtggtgcg	taacggcaaa	agcaccgccc	gacatcagcg	ctagcggagt	gtatactggc	1560
	ttactatggt	ggcactgatg	aggggtgtcag	tgaagtgtt	catgtggcag	gagaaaaaag	1620
15	gctgcaccgg	tgCGTCAGCA	gaatatgtga	tacaggatat	attccgcttc	ctcgtcact	1680
	gactcgctac	gctcggctgt	tcgactgcgg	cgagcggaaa	tggtttacga	acggggcgga	1740
20	gatttcctgg	aagatgccag	gaagatactt	aacagggaag	tgagagggcc	gcggcaaagc	1800
	cgTTTTTCCA	taggtccgc	ccccctgaca	agcatcacga	aatctgacgc	tcaaatcagt	1860
	ggtggcgaaa	cccgacagga	ctataaagat	accaggcggt	tccccctggc	ggctccctcg	1920
25	tgCGCTCTCC	tgTTCCTGCC	tttcggTTTA	ccggtgtcat	tccgctgtta	tgGCCGCGTT	1980
	tgtctcattc	cacgcctgac	actcagttcc	gggtaggcag	ttcGCTCCAA	gctggactgt	2040
	atgcacgaac	cccccgTTCA	gtccgaccgc	tgCGCCTTAT	ccggtaacta	tcgtcttgag	2100
30	tccaacccgg	aaagacatgc	aaaagcacca	ctggcagcag	ccactggtaa	ttgatttaga	2160
	ggagttagtc	ttgaagtcat	gcgCCGGTTA	aggctaaact	gaaaggacaa	gttttggtga	2220
	ctgcgctcct	ccaagccagt	tacctcggtt	caaagagttg	gtagctcaga	gaaccttcga	2280
35	aaaaccgccc	tgcaaggcgg	TTTTTTCGTT	ttcagagcaa	gagattacgc	gcagacccaa	2340
	acgatctcaa	gaagatcatc	ttattaatca	gataaaatat	ttctagattt	cagtgcatt	2400
40	tatctcttca	aatgtagcac	ctgaagtcag	ccccatacga	tataagttgt	aattctcatg	2460
	tttgacagca	tgccatcgat	cgcgatcttg	ctcaattggt	atcagctatg	cgccgaccag	2520
	aacaccttgc	cgatcagcca	aacgtctctt	caggccactg	actagcgata	actttcccca	2580
45	caacggaaca	actctcattg	catgggatca	ttgggtactg	tggttttagt	ggttgtaaaa	2640
	acacctgacc	gctatccctg	atcagtttct	tgaaggtaaa	ctcatcaccc	ccaagtctgg	2700
	ctatgcagaa	atcacctggc	tcaacagcct	gctcagggtc	aacgagaatt	aacattccgt	2760
50	caggaaagct	tggtttggag	cctgttggtg	cggtcatgga	attaccttca	acctcaagcc	2820
	agaatgcaga	atcactggct	TTTTTGGTTG	tgcttaccga	tctctccgca	tcacctttgg	2880
	taaaggttct	aagctcaggt	gagaacatcc	ctgcctgaac	atgagaaaaa	acagggtact	2940
55	catactcact	tctaagtgac	ggctgcatac	taaccgcttc	atacatctcg	tagattttct	3000

EP 3 134 426 B1

	tggcgattga agggctaaat tcttcaacgc taactttgag aatttttgca agcaatgcgg	3060
	cgttataagc atttaatgca ttgatgccat taaataaagc accaacgcct gactgccccca	3120
5	tccccatctt gtctgcgaca gattcctggg ataagccaag ttcatttttc tttttttcat	3180
	aaattgcttt aaggcgacgt gcgtcctcaa gctgctcttg tgттаатggт ttcttttttg	3240
10	tgctcatacg ttaaatctat caccgcaagg gataaatatc taacaccgtg cgtgttgact	3300
	attttacctc tggcggtgat aatggttgca tgtactaagg aggttggt	3348

<210> 9
 <211> 21
 <212> DNA
 <213> artificial

<220>
 <223> sekwecja kodujсca epitop HBV
 <220>
 <221> CDS
 <222> (1)..(21)

<400> 9

	acc aaa ccg acc gac ggt aac	21
	Thr Lys Pro Thr Asp Gly Asn	
	1 5	

<210> 10
 <211> 7
 <212> PRT
 <213> artificial

<220>
 <223> Synthetic Construct

<400> 10

	Thr Lys Pro Thr Asp Gly Asn
	1 5

<210> 11
 <211> 25
 <212> DNA
 <213> artificial

<220>
 <223> Modul powielajсcy

<400> 11
 gctcttcacc cgggcccgaga agagc 25

<210> 12
 <211> 42
 <212> DNA

EP 3 134 426 B1

<213> artificial

<220>
<223> pAMP1-A oligo górny

5

<400> 12
catgcgctct tcacccgggc ccagaagagc taagtaagta ag 42

10

<210> 13
<211> 42
<212> DNA
<213> artificial

15

<220>
<223> pAMP1-A oligo dolny

<400> 13
agctcttact tacttagctc ttctgggccc ggggaagag cg 42

20

<210> 14
<211> 60
<212> DNA
<213> artificial

25

<220>
<223> pAMP1-HisA oligo górny

<400> 14
catgcaccac caccaccac acagctcttc acccgggccc agaagagcta agtaagtaag 60

30

<210> 15
<211> 60
<212> DNA
<213> artificial

35

<220>
<223> pAMP1-HisA oligo dolny

<400> 15
agctcttact tacttagctc ttctgggccc ggggaagag ctgtggtggt ggtggtggtg 60

40

<210> 16
<211> 182
<212> DNA
<213> artificial

45

<220>
<223> Sekwencja PCR zawierająca insert (wariant 5 powtórzeń antygenu HBV) do Phage Display System

50

<400> 16

caccaccacc acaggaattc acccacaaa ccgaccgacg gtaacggggc caccaaaccg 60

55

accgacggta acggggccac caaaccgacc gacggtaacg ggcccaccaa accgaccgac 120

ggtaacgggc ccaccaaacc gaccgacggt aactaagctt gaagagctaa gtaagtaaga 180

gc

182

5 <210> 17
 <211> 28
 <212> DNA
 <213> artificial

10 <220>
 <223> Starter górny do Phage Display System

<400> 17
 caccaccacc acaggaattc acccacca 28

15 <210> 18
 <211> 39
 <212> DNA
 <213> artificial

20 <220>
 <223> Starter dolny do Phage Display System

<400> 18
 gctcttactt acttagctct tcaagcttag ttaccgtcg 39

25 <210> 19
 <211> 220
 <212> DNA
 <213> artificial

30 <220>
 <223> sekwencja z Fig. 11

<400> 19

35
 aggaggttgt tcatgcacca ccaccaccac cacagctctt caccaccaa accgaccgac 60
 ggtaacggggc ccaccaaacc gaccgacggt aacggggccca ccaaaccgac cgacggtaac 120
 40 gggcccacca aaccgaccga cggtaacggg cccaccaaac cgaccgacgg taacggggccc 180
 agaagagcta agtaagtaag agctcgaccc ggtcgaattt 220

45 **Claims**1. A method of obtaining polypeptopic protein, **characterised in that:**

- 50 a) a blunt-ended DNA sequence encoding an epitope is cloned at a site recognised by SmaI into a DNA vector containing a sequence, the amplifying module, encompassing two convergent DNA sequences recognised by SapI and a DNA sequence found between them containing a site for cloning in an insert recognised by SmaI, wherein the amplifying module contains a nucleotide sequence selected from: SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6, preferably the amplifying module possesses the sequence:
- 55 GCTCTTCACCCGGGCCCCAGAAGAGC (SEQ ID NO: 11),
- b) the resulting vector is amplified in a bacterial host, isolated and digested with SapI, and then the fragment containing the DNA sequence encoding the epitope is isolated, which is modified such that it becomes equipped with single-stranded sticky ends facilitating directional ligation of the insert to the concatemer,

- c) the isolated fragment is autoligated,
- d) the autoligation product is cloned at a site recognised by SapI into said DNA vector as defined in point a),
- e) the resulting vector is used to transform a bacterial host and then the polyepitopic protein is expressed and isolated,

wherein in order to increase the size of the polyepitopic protein, stages from b) to d) are repeated prior to realising stage e), a concatenated DNA sequences encoding the epitopes are present in the resulting vector.

2. A method according to Claim 1, **characterised in that** epitope is the HBV epitope, preferably encoded by the synthetic SEQ ID NO: 9.
3. A method according to Claim 1, **characterised in that** the amplified monomer fragment contains various epitopes, stemming from different proteins or different regions of the same protein, preferably encoded by a synthetic sequence.
4. A method according to Claim 1, **characterised in that** the polyepitopic protein additionally contains a sequence facilitating its purification, particularly a fusion tag containing a sequence of 6 histidine residues, and the isolation conducted in stage e) encompasses the use of metalloaffinity chromatography, preferably on immobilised nickel residues, preferably an additional purification stage selected from a group encompassing: heating the protein sample, fractionation using polyethylenimine, salting out with ammonium sulphate and molecular sieving gel chromatography.
5. A method according to Claim 1, **characterised in that** the stage e) polyepitopic protein is immobilised on a macromolecular carrier, particularly one selected from among: microorganisms, cells, bacteria, bacteriophages, viruses, defective virions, autoaggregating proteins, or nanoparticles, preferably from a Phage Display system in the T7 bacteriophage.
6. A DNA vector for use in the method of claims 1-4 containing a sequence, the amplifying module, encompassing two convergent DNA sequences recognised by SapI and a DNA sequence found between them containing a site for cloning in an insert recognised by SmaI, wherein the amplifying module contains a nucleotide sequence selected from: SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6, preferably the amplifying module possesses the sequence:
GCTCTTCACCCGGGCCAGAGAAGC (SEQ ID NO: 11).
7. The DNA vector according to Claim 6, **characterised in that** it is a protein expression vector, and additionally contains an origin of replication, preferably p15A, an antibiotic resistance gene, preferably chloramphenicol, a transcription promoter, preferably P_R of the lambda bacteriophage, a repressor gene, preferably cI857ts, a translation initiation signal, and possibly a sequence encoding 6 histidine residues, as well as a sequence encoding a translation stop signal.
8. The DNA vector according to Claim 6, **characterised in that** it has SEQ ID NO: 7.

Patentansprüche

1. Verfahren zum Gewinnen von polyepitopischem Protein, **dadurch gekennzeichnet, dass:**

- a) eine DNA Sequenz mit stumpfem Ende, die für ein Epitop kodiert, an einer von SmaI erkannten Stelle in einen DNA Vektor kloniert wird, enthaltend eine Sequenz, das Amplifikationsmodul, umfassend zwei von SapI erkannte konvergierende DNA Sequenzen und eine DNA Sequenz, die zwischen ihnen gefunden wurde, die eine von SmaI erkannte Klonierungsstelle für ein Insert enthält, wobei das Amplifikationsmodul eine Nukleotidsequenz enthält, die ausgewählt ist aus: SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 oder SEQ ID NO: 6, vorzugsweise besitzt das Amplifikationsmodul die Sequenz:
GCTCTTCACCCGGGCCAGAGAAGC (SEQ ID NO: 11),
- b) der resultierende Vektor wird in einem bakteriellen Wirt amplifiziert, isoliert und mit SapI gespalten, und dann wird das Fragment, das die DNA Sequenz enthält, die das Epitop kodiert, isoliert, das so modifiziert ist, dass es mit einzelsträngigen klebrigen Enden ausgestattet ist, welche die gerichtete Ligation des Inserts mit dem Concatemer erleichtern,
- c) das isolierte Fragment wird autoligiert,
- d) das Autoligationsprodukt wird in eine von SapI erkannten Stelle in den DNA Vektor, wie unter Punkt a)

definiert, kloniert,

e) der resultierende Vektor wird verwendet, um einen bakteriellen Wirt zu transformieren, und dann wird das polypeptidisches Protein exprimiert und isoliert, wobei zur Erhöhung der Größe des polypeptidischen Proteins die Schritte von b) bis d) vor der Durchführung von Schritt e) wiederholt werden, eine verkettete DNA Sequenz, die für die Epitope kodieren, sind in dem resultierenden Vektor vorhanden.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** das Epitop das HBV Epitop ist, vorzugsweise durch die synthetische SEQ ID NO: 9 kodiert.

3. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** das amplifizierte Monomerfragment verschiedene Epitope enthält, die von verschiedenen Proteinen oder verschiedenen Regionen desselben Proteins stammen, vorzugsweise durch eine synthetische Sequenz kodiert.

4. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** das polypeptidische Protein zusätzlich eine Sequenz enthält, die seine Reinigung erleichtert, insbesondere eine Fusionsmarkierung, die eine Sequenz von 6 Histidinresten enthält, und die in Schritt e) durchgeführte Isolierung umfasst die Verwendung der Metallaffinitätschromatographie, vorzugsweise an immobilisierten Nickelresten, vorzugsweise eine zusätzliche Reinigungsstufe, ausgewählt aus einer Gruppe, umfassend: Erhitzen der Proteinprobe, Fraktionierung unter Verwendung von Poly-ethylenimin, Aus-salzen mit Ammoniumsulfat und Molekularsieb-Gel-Chromato-graphie.

5. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** das polypeptidische Protein aus Schritt e) auf einem makromolekularen Träger immobilisiert ist, insbesondere einem, der ausgewählt ist aus der Gruppe: Mikroorganis-men, Zellen, Bakterien, Bakteriophagen, Viren, defekte Virionen, autoaggregierende Proteine oder Nanopartikel, vorzugsweise aus einem Phage-Display-System in dem T7-Bakteriophagen.

6. DNA Vektor zur Verwendung in dem Verfahren nach den Ansprüchen 1-4, enthaltend eine Sequenz, das Amplifi-kationsmodul, umfassend zwei von SapI erkannte konvergierende DNA Sequenzen und eine DNA Sequenz, die zwischen ihnen gefunden wurde, die eine von SmaI erkannte Klonierungsstelle für ein Insert enthält, wobei das Amplifikationsmodul eine Nukleotidsequenz enthält, die ausgewählt ist aus: SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 oder SEQ ID NO: 6, vorzugsweise besitzt das Amplifikationsmodul die Sequenz: GCTCTTACCCGGGCCAGAGC (SEQ ID NO: 11).

7. DNA Vektor nach Anspruch 6, **dadurch gekennzeichnet, dass** er ein Proteinexpressionsvektor ist, und zusätzlich einen Replikationsursprung, vorzugsweise p15A, ein Antibiotikumresistenzgen, vorzugsweise Chloramphenicol, einen Transkriptionspromotor, vorzugsweise Pr des Lambda-Bakteriophagen, ein Repressorgen, vorzugsweise cl85/7ts, ein Translations-Initiations-Signal, und möglicherweise eine Sequenz, die 6 Histidinreste kodiert, sowie eine Sequenz, die ein Translations-Stopp-Signal kodiert, enthält.

8. DNA Vektor nach Anspruch 6, **dadurch gekennzeichnet, dass** er die SEQ ID NO: 7 aufweist.

Revendications

1. Méthode d'obtention d'une protéine polyépitopique, **caractérisée en ce que** :

a) une séquence d'ADN à extrémités franches codant pour un épitope est clonée au niveau d'un site reconnu par SmaI dans un vecteur d'ADN contenant une séquence, le module d'amplification, englobant deux séquences d'ADN convergentes reconnues par SapI et une séquence d'ADN trouvée entre celles-ci, contenant un site pour un clonage dans un insert reconnu par SmaI, où le module d'amplification contient une séquence nucléotidique choisie parmi SEQ ID n° 1, SEQ ID n° 2, SEQ ID n° 3, SEQ ID n° 4, SEQ ID n° 5 ou SEQ ID n° 6, préférentiellement le module d'amplification possède la séquence :

GCTCTTACCCGGGCCAGAGC (SEQ ID n° 11),

b) le vecteur résultant est amplifié dans un hôte bactérien, isolé et digéré par SapI, puis le fragment contenant la séquence d'ADN codant pour l'épitope est isolée, laquelle est modifiée de telle façon qu'elle soit pourvue d'extrémités cohésives à brin simple facilitant la ligation directionnelle de l'insert au concatémère,

c) le fragment isolé est auto-ligaturé,

d) le produit d'auto-ligature est cloné au niveau d'un site reconnu par SapI dans ledit vecteur d'ADN tel que défini au point a),

e) le vecteur résultant est utilisé pour transformer un hôte bactérien, puis la protéine polyépitopique est exprimée et isolée,

dans laquelle, afin d'augmenter la taille de la protéine polyépitopique, les étapes b) à d) sont répétées préalablement à la mise en œuvre de l'étape e), des séquences d'ADN concaténées codant pour les épitopes sont présentes dans le vecteur résultant.

2. Méthode selon la revendication 1, **caractérisée en ce que** l'épitope est l'épitope du VHB, préférablement codé par la SEQ ID n° 9 synthétique.

3. Méthode selon la revendication 1, **caractérisée en ce que** le fragment de monomère amplifié contient divers épitopes, résultant de différentes protéines ou de différentes régions de la même protéine, préférablement codés par une séquence synthétique.

4. Méthode selon la revendication 1, **caractérisée en ce que** la protéine polyépitopique contient de plus une séquence facilitant sa purification, en particulier une étiquette de fusion contenant une séquence de 6 résidus histidine, et l'isolement mis en œuvre dans l'étape e) englobe l'utilisation d'une chromatographie par affinité sur métaux, préférablement sur des restes de nickel immobilisés, préférablement une étape de purification supplémentaire choisie dans le groupe constitué par : le chauffage de l'échantillon protéique, le fractionnement à l'aide de polyéthylèneimine, le relargage à l'aide de sulfate d'ammonium et la chromatographie de perméation moléculaire sur gel.

5. Méthode selon la revendication 1, **caractérisée en ce que** la protéine polyépitopique de l'étape e) est immobilisée sur un support macromoléculaire, en particulier un support choisi parmi : les microorganismes, les cellules, les bactéries, les bactériophages, les virus, les virions défectueux, les protéines à auto-agrégation, ou les nanoparticules, préférablement parmi un Système d'Exposition sur Phage chez le bactériophage T7.

6. Vecteur d'ADN destiné à une utilisation dans la méthode selon les revendications 1-4, contenant une séquence, le module d'amplification, englobant deux séquences d'ADN convergentes reconnues par Sapl et une séquence d'ADN trouvée entre celles-ci, contenant un site pour un clonage dans un insert reconnu par SmaI, où le module d'amplification contient une séquence nucléotidique choisie parmi SEQ ID n° 1, SEQ ID n° 2, SEQ ID n° 3, SEQ ID n° 4, SEQ ID n° 5 ou SEQ ID n° 6, préférablement le module d'amplification possède la séquence : GCTCTTCACCCGGCCCAAGAGC (SEQ ID n° 11).

7. Vecteur d'ADN selon la revendication 6, **caractérisé en ce qu'il** s'agit d'un vecteur d'expression protéique, et contient de plus une origine de répllication, préférablement p15A, un gène de résistance à un antibiotique, préférablement le chloramphénicol, un promoteur de transcription, préférablement P_R du bactériophage lambda, un gène répresseur, préférablement cl857ts, un signal d'initiation de traduction, et éventuellement une séquence codant pour 6 résidus histidine, ainsi qu'une séquence codant pour un signal d'arrêt de la traduction.

8. Vecteur d'ADN selon la revendication 6, **caractérisé en ce qu'il** contient SEQ ID n° 7.

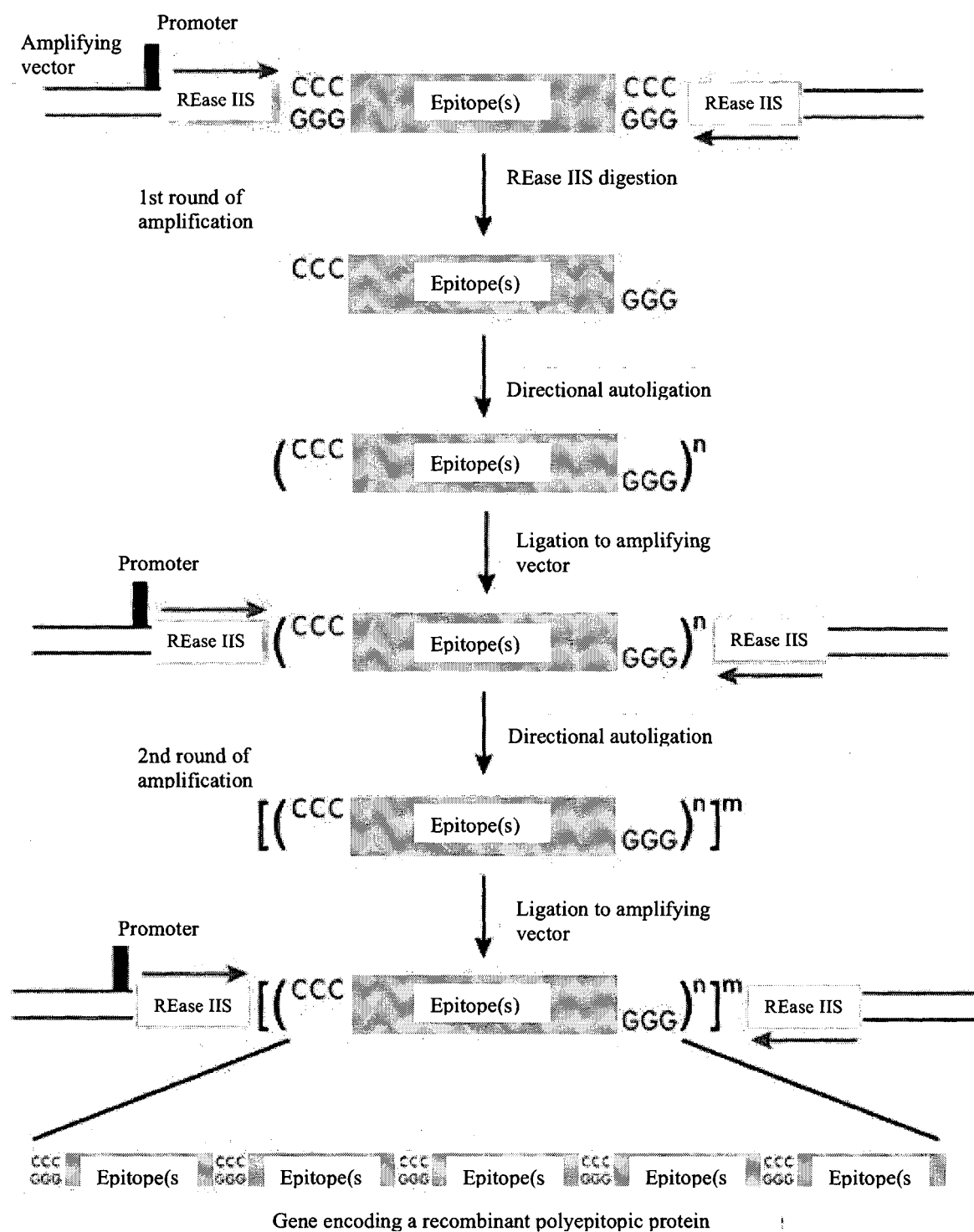


Fig. 1

Vector 1: pAMP1-A: ORF (during reading of SmaI digestion site)

ATG C ^{SapI} GCTCTTC ^{SmaI} ACCC | GGG CCC AGAAGAGC ^{SapI} TAAGTAAGTAAG
Met Arg Ser Ser Pro

Vector 2: pAMP1-B: ORF B (phase -1 bp)

ATG ^{SapI} GCTCTTC ^{SmaI} ACCC | GGGCCC AGAAGAGC ^{SapI} TAAGTAAGTAAG
Met Ala Leu His

Vector 3: pAMP1-C: ORF C (phase +1 bp)

ATG CC ^{SapI} GCTCTTC ^{SmaI} ACCC | GGG CCC AGAAGAGC ^{SapI} TAAGTAAGTAAG
Met Pro Leu Phe Thr

Vector 4: pAMP1-HisA: ORF/N-His tag (during reading of SmaI digestion site)

ATG CAC CAC CAC CAC CAC CAC A ^{His₆} GCTCTTC ^{SapI} ACCC | GGG CCC AGAAGAGC ^{SmaI} TAAGTAAGTAAG ^{SapI} STOP (3 x ORF)
Met His His His His His His Ser Ser Ser Pro

Vector 5: pAMP1-HisB: ORF B (phase -1 bp)/N-His tag

ATG CAC CAC CAC CAC CAC CAC ^{His₆} GCTCTTC ^{SapI} ACCC | GGGCCC AGAAGAGC ^{SapI} TAAGTAAGTAAG ^{SapI} STOP (3 x ORF)
Met His His His His His His Ala Leu His

Vector 6: pAMP1-HisC: ORF C (phase +1 bp)/N-His tag

ATG CAC CAC CAC CAC CAC CAC ^{His₆} CC GCTCTTC ^{SapI} ACCC | GGGCCC AGAAGAGC ^{SapI} TAAGTAAGTAAG ^{SapI} STOP (3 x ORF)
Met His His His His His His Pro Leu Phe Thr

Fig. 2

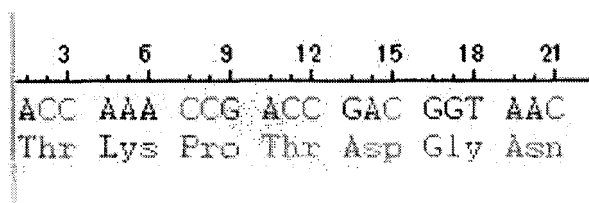


Fig. 3

HBV epitope monomer (21 bp + 3-nt sticky end DNA) pz = bp

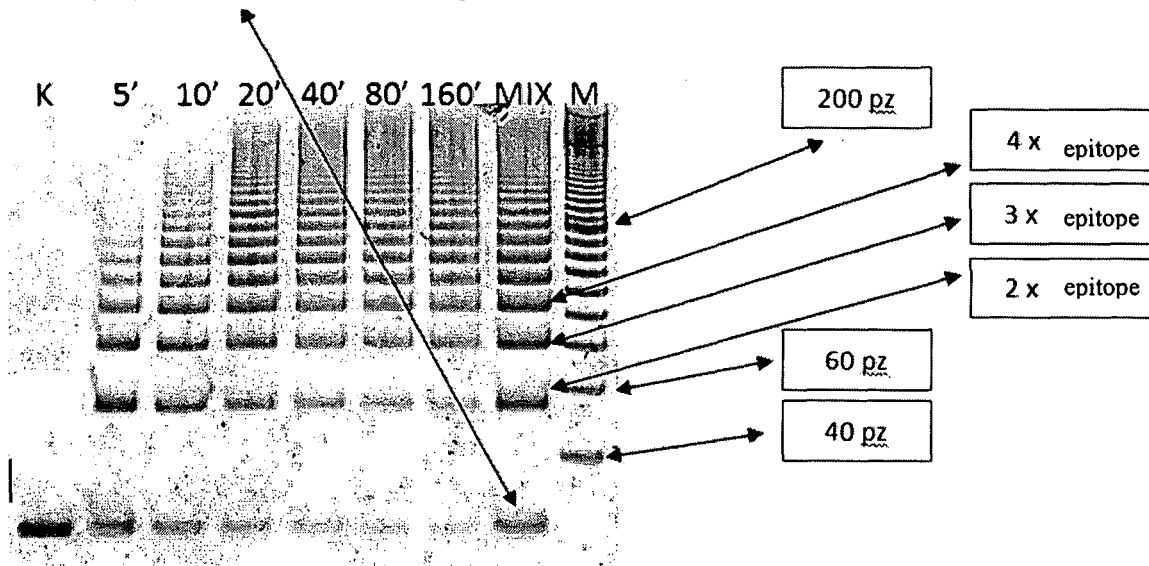


Fig. 4

HBV epitope monomer (amplified monomer + adjacent vector sequences, PCR analysis) pz = bp

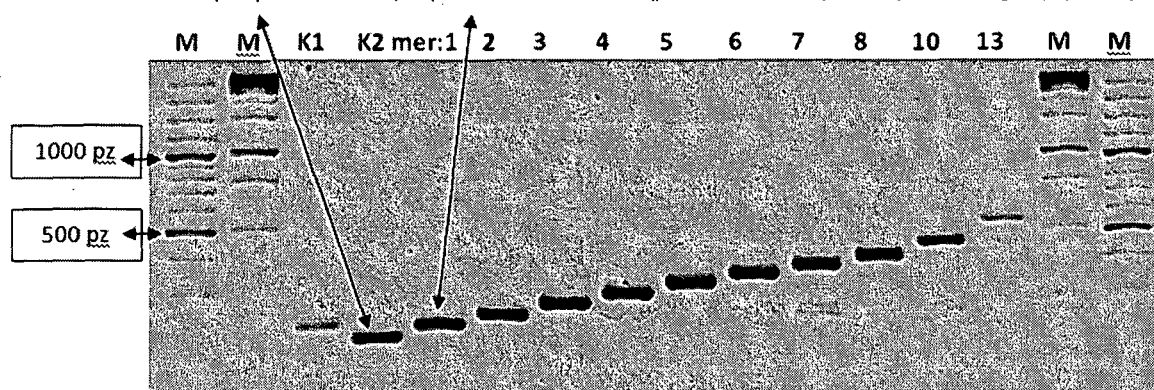


Fig. 5

12-mer of a 5-mer concatemer from the 1st round of HBV epitope amplification (+ 3-nt sticky end DNA_

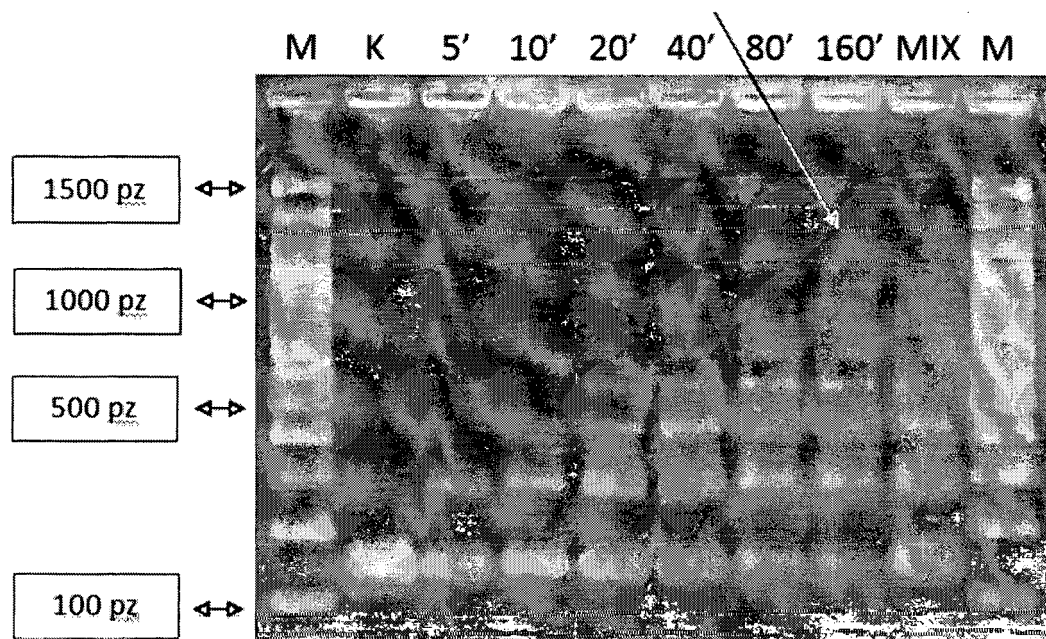


Fig. 6

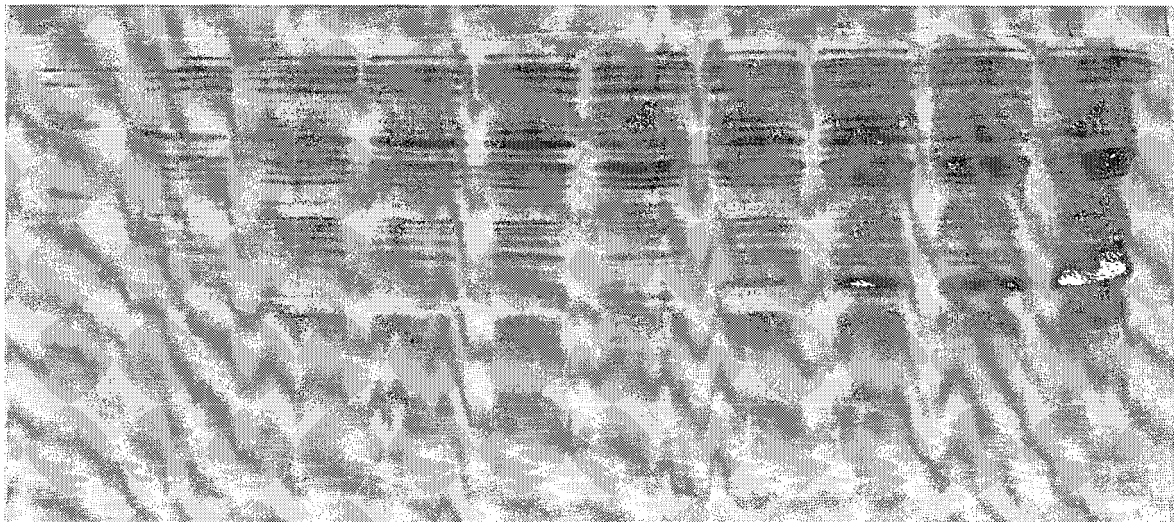


Fig. 7

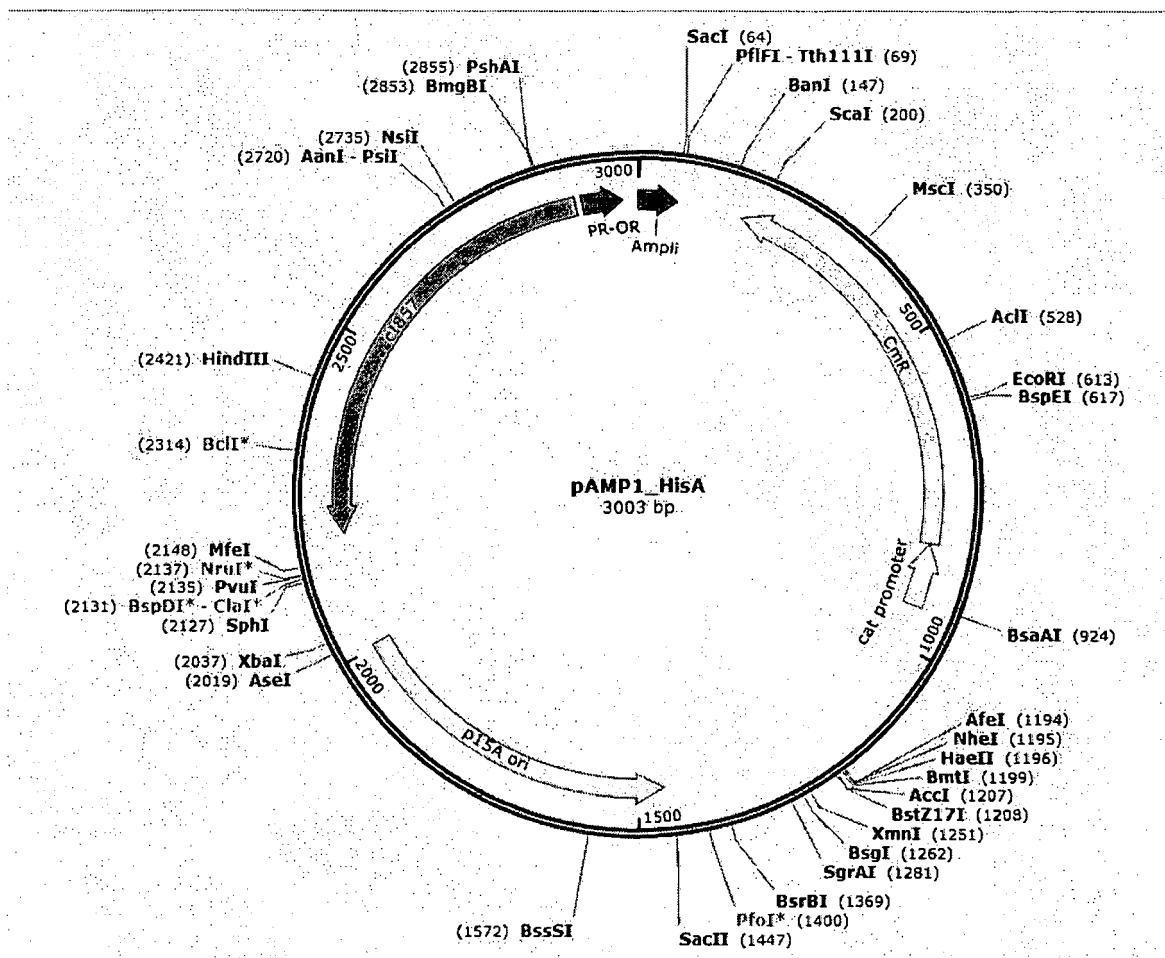


Fig. 8

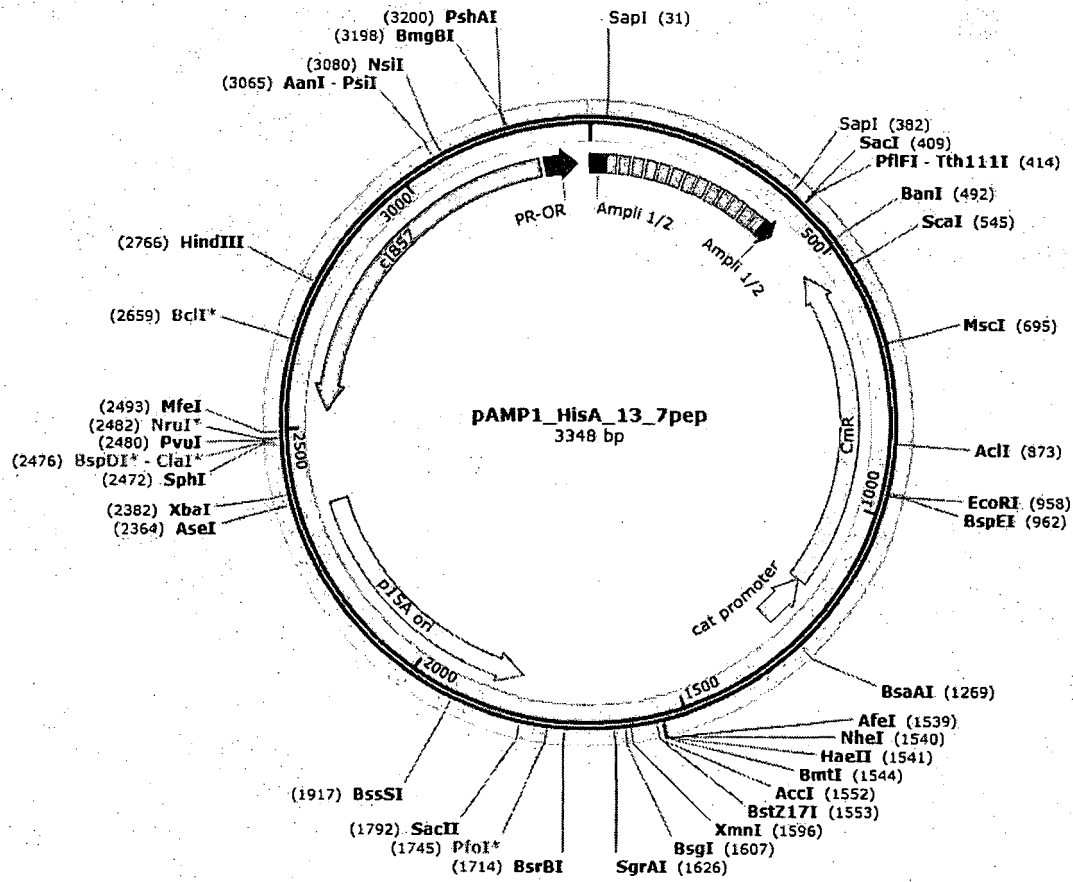


Fig. 9

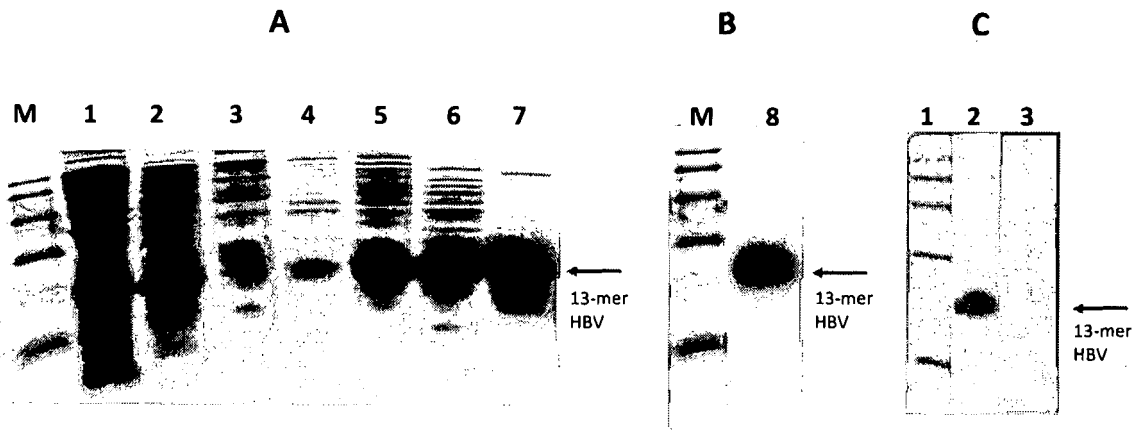


Fig. 10

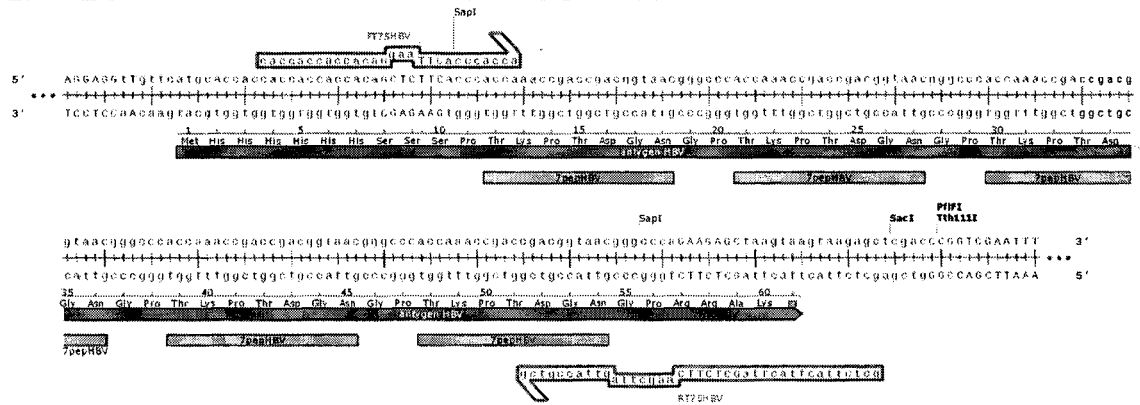


Fig. 11

A

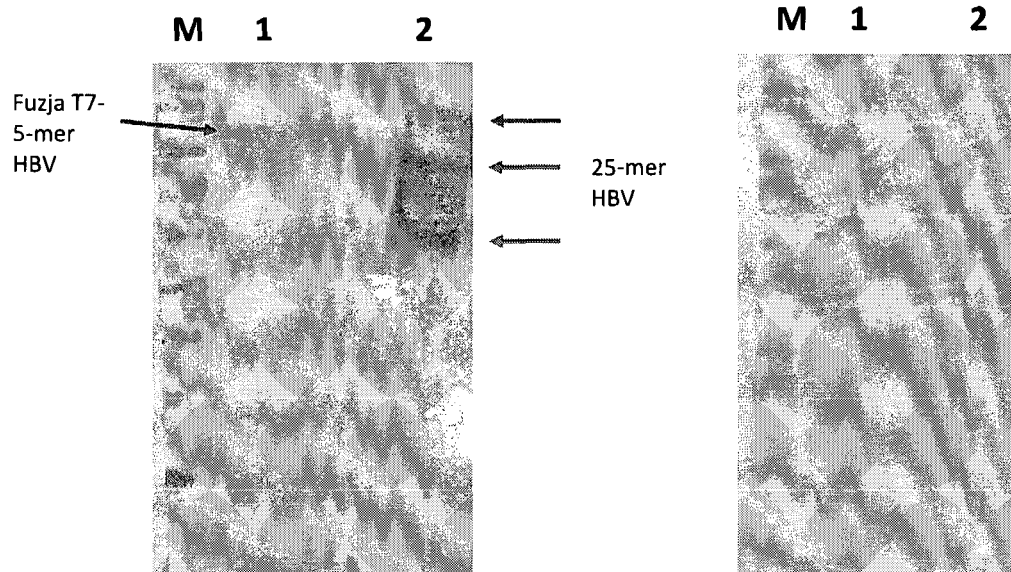
B

Fig. 12

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- **SHEN, S.-H** *P.N.A.Sci USA*, 1984, vol. 81, 4627-4631 [0002]
- **LENNICK et al.** *Gene*, 1987, vol. 61, 103-112 [0002]
- **KIM, S. C. ; SZYBALSKI, W.** *Gene*, 1988, vol. 71, 1-8 [0002]
- **LEE, J. H. et al.** *Genetic Analysis: Biomolecular Engineering*, 1996, vol. 13, 139-145 [0002]
- **LEE, J. H. et al.** *Protein Expression and purification*, 1998, vol. 12, 53-60 [0002]
- **WANG, Y.-Q ; CAI, J.Y.** *Appl Biochem Biotechnol.*, 2007, vol. 141, 203-13 [0002]
- **LEE et al.** *Genetic Analysis: Biomolecular Engineering*, 1996, vol. 13, 139-145 [0002]
- **GREETSMA ; DUTZLER.** *Biochemistry*, 2011, 3272-3278 [0002]
- *Archive Cloning*, 2014, www.genome.wisc.edu [0002]